Energetic, Structural and Dynamic Evaluation of HIV-1 Proteases

by

Previn Naicker, 587420

August 2014

A thesis submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Doctor of Philosophy.

Supervisor: Prof. Yasien Sayed

Co-supervisor: Dr. Ikechukwu Achilonu

Declaration

I, Previn Naicker (587420), am a student registered for the degree of Doctor of Philosophy (PhD) in the academic year 2014.

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 I have explicitly indicated otherwise.

- 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: 21 August 2014

Abstract

Human immunodeficiency virus (HIV), the causative agent of the acquired immunodeficiency syndrome (AIDS), remains a topic of global concern even though great strides have been made to combat the virus. The high replicative rate of the virus and recombination of the variety of viral strains complicate the treatment of AIDS. There has been an increasing prevalence of African HIV strains in the Americas and Europe. The viral protease (PR) is vital for the propagation of the virus; and thus, is a major target in antiviral therapy. The HIV-1 PR enzyme from the subtype C strain; which predominates in sub-Saharan Africa, has been greatly under-investigated in comparison to the protease from the subtype B strain which predominates in North America and Europe. Enzyme activity data which were part of this work suggested that the South African HIV-1 subtype C protease (C-SA PR) displays improved substrate turnover in comparison to the subtype B PR. Thermodynamics and inhibition kinetics of drug binding showed that the C-SA PR is less susceptible to certain clinically-used protease inhibitors when compared to the subtype B PR. A crystal structure of the C-SA PR was solved and showed no difference to the global structure of the subtype B PR. Molecular dynamics simulations showed that the C-SA PR exhibits a wider range of open conformations. Hydrogen/deuterium exchange-mass spectrometry (HDX-MS) was performed to elucidate the mechanism of reduced drug susceptibility displayed by the C-SA PR. HDX-MS data provided insights into the basis of the increased preference for open conformers displayed by the C-SA PR and the stability of the terminal dimer interface which is a target in protease inhibition.

Research outputs

Publications forming part of PhD thesis

Research articles:

1. <u>Naicker, P.</u>, Achilonu, I., Fanucchi, S., Fernandes, M., Ibrahim, M. A. A., Dirr, H. W., Soliman, M. E. S., and 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.

2. <u>Naicker, P.</u>, 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.*, (In revision).

Review articles:

3. <u>Naicker, P.</u>, and Sayed, Y. (2014) Non-B HIV-1 subtypes in sub-Saharan Africa: impact of subtype on protease inhibitor efficacy, *Biol. Chem.*, In Press.

Related publications:

4. <u>Naicker, P.</u>, Seele, P., Dirr, H. W., and Sayed, Y. (2013) F99 is critical for dimerization and activation of South African HIV-1 subtype C protease, *Protein J.* **32**, 560-567.

5. Ahmed, S. M., Kruger, H. G., Govender, T., Maguire, G. E. M., Sayed, Y., Ibrahim, M. A. A., <u>Naicker, P.</u>, and 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.

Conference outputs

1. Molecular Biosciences Research Thrust Research day, Johannesburg 2013. *Oral presentation*. " Subtype analysis of Protease inhibition in HIV-1: mechanism of reduced drug susceptibility". Previn Naicker, Stoyan Stoychev, Heini W. Dirr and Yasien Sayed.

2. 8th General Meeting of the International Proteolysis Society, Stellenbosch 2013. *Poster presentation*. " Amide hydrogen exchange of the South African HIV-1 subtype C Protease: Apparent mechanism of reduced drug susceptibility". Previn Naicker, Stoyan Stoychev, Heini W. Dirr and Yasien Sayed.

3. Wits Cross Faculty Postgraduate Symposium, Johannesburg 2013. *Oral presentation*. " Protease dynamics reveals mechanism of HIV drug resistance". Previn Naicker, Stoyan Stoychev, Heini W. Dirr and Yasien Sayed.

4. Wits Research Day, Johannesburg 2012. *Oral presentation*. " Structural insights into the South African HIV-1 subtype C protease: Importance of hinge region dynamics and flap flexibility in drug resistance". Previn Naicker, Mahmoud Ibrahim, Mahmoud Soliman, Heini W. Dirr and Yasien Sayed.

5. South African Society for Biochemistry and Molecular Biology conference, Drakensberg 2012. *Poster presentation*. "Structural characterisation of the South African wild-type HIV-1 Subtype C Protease". Previn Naicker, Ikechukwu Achilonu, Sylvia Fanucchi, Manuel Fernandes, Heini W. Dirr and Yasien Sayed.

Acknowledgements

I would like to thank my supervisor, Prof. Yasien Sayed, and co-supervisor, Dr. Ikechukwu Achilonu, for their exceptional supervision and mentorship during my studies.

Dr. Stoyan Stoychev for his assistance with mass spectrometry-based experiments.

Prof. Heini Dirr for his guidance and other members of the Protein Structure-Function Research Unit for sharing in this journey.

National Research Foundation (South Africa) and the University of the Witwatersrand for funding.

A special thanks to my family and Dr. Bianca Dias for their endless support and love.

Contents

List of figures

List of abbreviations

AIDS   Acquired immunodeficiency syndrome

CRF   Circulating recombinant form

C-SA   South African subtype C

DNA   Deoxyribonucleic acid

EPR   Electron paramagnetic resonance

FDA   Food and drug administration (USA)

HDX-MS   Hydrogen/deuterium exchange-mass spectrometry

HIV   Human immunodeficiency virus

ITC   Isothermal titration calorimetry

MD   Molecular dynamics

NMR   Nuclear magnetic resonance

PDB   Protein Data Bank

PI   Protease inhibitor

PR   Protease

RNA   Ribonucleic acid

RT   Reverse transcriptase

URF   Unique recombinant form

CHAPTER 1

Non-B HIV-1 Subtypes in sub-Saharan Africa: Impact of Subtype on Protease Inhibitor Efficacy

Previn Naicker and Yasien Sayed.

In this publication, a detailed introduction into HIV is given. Information on the genetic diversity of HIV type 1 and characteristics of the HIV-1 subtype C is presented. Comparisons between the subtype B and C proteases are made and the secondary resistance mutations inherent to the C-SA PR are discussed. An account of some novel anti-HIV drugs is also provided.

Review

Previn Naicker and Yasien Sayed*

Non-B HIV-1 subtypes in sub-Saharan Africa: impact of subtype on protease inhibitor efficacy

Abstract: In 2012, 25 million people [71% of global human immunodeficiency virus (HIV) infection] were estimated to be living with HIV in sub-Saharan Africa. Of these, approximately 1.6 million were new infections and 1.2 million deaths occurred. South Africa alone accounted for 31% of HIV/acquired immunodeficiency syndrome (AIDS) deaths in sub-Saharan Africa. This disturbing statistic indicates that South Africa remains the epicenter of the HIV/AIDS pandemic, compounded by the fact that only 36% of HIV-positive patients in South Africa have access to antiretroviral (ARV) treatment. Drug resistance mutations have emerged, and current ARVs show reduced efficacy against non-B subtypes. In addition, several recent studies have shown an increased prevalence of non-B African HIV strains in the Americas and Europe. Therefore, the use of ARVs in a non-B HIV-1 subtype context requires further investigation. HIV-1 subtype C protease, found largely in sub-Saharan Africa, has been under-investigated when compared with the subtype B protease, which predominates in North America and Europe. This review, therefore, focuses on HIV-1 proteases from B and C subtypes.

Keywords: African subtypes; AIDS; drug resistance; novel protease inhibitors.

Introduction

Acquired immunodeficiency syndrome (AIDS) has caused major global health and socioeconomic challenges over

*Corresponding author: Yasien Sayed, Protein Structure-Function Research Unit, School of Molecular and Cell Biology, University of Witwatersrand, Johannesburg 2050, South Africa,
e-mail: yasien.sayed@wits.ac.za
Previn Naicker: Protein Structure-Function Research Unit, School of Molecular and Cell Biology, University of Witwatersrand, Johannesburg 2050, South Africa

the past three decades. Human immunodeficiency virus (HIV), the causative agent of AIDS, has been extensively investigated in research laboratories throughout the world. Genotypic characterization of the virus and antiretroviral (ARV) drug development efforts have been exhaustive and are ongoing. Understanding the virus at the molecular level is a basic requirement, and the knowledge gained from these studies would help drive drug development. Globally, there are roughly 35 million people living with HIV and 71% reside in sub-Saharan Africa (The Joint Nations Program on HIV/AIDS, 2013). Statistics on ARV efficacy during long-term treatment are generally poor. Statistics from sub-Saharan countries are especially poor due to the cost of HIV monitoring and various socioeconomic challenges. This is an unfortunate reality because reports of drug resistance as well as the genetic variation of HIV are on the rise (Rousseau et al., 2007; Shafer and Schapiro, 2008; Hemelaar et al., 2011).

HIV is a retrovirus. The survival of the virus in a human host requires the viral genetic material (RNA) to be reverse transcribed to DNA by the viral reverse transcriptase (RT). Owing to the RT enzyme lacking proofreading ability and the high replication rate of HIV, mutations occur frequently in the viral proteins that are targeted by antiviral drugs, thereby providing a challenge for the production of effective antiviral compounds. Currently, there are six classes of anti-HIV drugs (Table 1): nucleoside RT inhibitors (NRTIs), non-nucleoside RT inhibitors (NNRTIs), protease inhibitors (PIs), integrase inhibitors, fusion inhibitors, and entry inhibitors. RT and PIs are recommended in treatment regimens globally, and extensive data are available on the drug-resistance profiles of these therapeutics (World Health Organization, 2013). Fewer information is available on the newer classes of inhibitors and reduced efficacies due to drug-resistance mutations (DRMs) are inevitable. Efforts to produce an effective HIV vaccine are ongoing (Moore et al., 2012), and the control of HIV epidemiology in the interim is reliant on RT and PIs. Importantly, more mutations are selected in response to PIs than any other class of ARV (Shafer and Schapiro, 2008). In this review,

Table 1 FDA-approved ARV drugs for treatment of HIV infection.

Generic name	Manufacturer	Approval date
NRTIs		
Zidovudine, AZT	GlaxoSmithKline	March 19, 1987
Didanosine, dideoxyinosine (ddI)	Bristol Myers-Squibb	October 9, 1991
Zalcitabine, dideoxycytidine (ddC) (no longer marketed)	Hoffmann-La Roche	June 19, 1992
Stavudine (d4T)	Bristol Myers-Squibb	June 24, 1994
Lamivudine (3TC)	GlaxoSmithKline	November 17, 1995
Abacavir sulfate	GlaxoSmithKline	December 17, 1998
Tenofovir disoproxil fumarate (TDF)	Gilead Sciences	October 26, 2001
Emtricitabine (FTC)	Gilead Sciences	July 02, 2003
NNRTIs		
Nevirapine (NVP)	Boehringer Ingelheim	June 21, 1996
Delavirdine (DLV)	Pfizer	April 4, 1997
Efavirenz (EFV)	Bristol Myers-Squibb	September 17, 1998
Etravirine	Tibotec Therapeutics	January 18, 2008
Rilpivirine	Tibotec Therapeutics	May 20, 2011
Pis		
Saquinavir mesylate (SQV)	Hoffmann-La Roche	December 6, 1995
Ritonavir (RTV)	Abbott Laboratories	March 1, 1996
Indinavir (IDV)	Merck	March 13, 1996
Nelfinavir mesylate (NFV)	Agouron Pharmaceuticals	March 14, 1997
Amprenavir (APV) (no longer marketed)	GlaxoSmithKline	April 15, 1999
Lopinavir (LPV)	Abbott Laboratories	September 15, 2000
Atazanavir sulfate (ATV)	Bristol-Myers Squibb	June 20, 2003
Fosamprenavir calcium (FOS-APV)	GlaxoSmithKline	October 20, 2003
Tipranavir (TPV)	Boehringer Ingelheim	June 22, 2005
Darunavir (DRV)	Tibotec	June 23, 2006
Fusion inhibitors		
Enfuvirtide (T-20)	Hoffmann-La Roche and Trimeris	March 13, 2003
Entry inhibitors (CCR5 co-receptor antagonist)		
Maraviroc	Pfizer	August 06, 2007
HIV integrase strand transfer inhibitors		
Raltegravir	Merck	October 12, 2007
Dolutegravir	GlaxoSmithKline	Auguste 13, 2013

List excludes approved inhibitors that are combined or modified.
Adapted from the U.S. Food and Drug Administration, http://www.fda.gov/ForConsumers/byAudience/For-Patient Advocates/HIVandAIDS-Activities/ucm118915.htm.

we focus on the subtype C protease (PR) prevalent in sub-Saharan Africa with special attention to the South African HIV-1 subtype C PR. Approximately 12% of the South African population are HIV positive (6.1 million out of a total population of 52 million people) (The Joint Nations Program on HIV/AIDS, 2013). This alarming statistic reflects that South Africa is the epicenter of the HIV/AIDS pandemic. Although there are excellent reviews on HIV-1 subtype B and non-B subtypes (Kantor and Katzenstein, 2003; Martinez-Cajas et al., 2009; Wainberg and Brenner, 2010), there are very few reviews on PRs from sub-Saharan Africa (Bessong, 2008). Therefore, a clear understanding of the role of genetic diversity and subtype variation and their impact on viral fitness and drug efficacy must be addressed.

Genetic diversity of HIV

Two types of HIV have been identified: HIV-1 and HIV-2. HIV-1 is the predominant type and is separated into groups M, N, O, and P (Taylor et al., 2008; Plantier et al., 2009). Group M (comprising 95% of the complete genome sequences of HIV-1) is further separated into nine subtypes (A, B, C, D, F, G, H, J, and K) and a growing number of circulating recombinant forms (CRFs) as shown in Figure 1 (Taylor et al., 2008). A recombinant strain is defined as a CRF if it has been fully sequenced and isolated in three or more epidemiologically distinct individuals (Robertson et al., 2000). Currently, recombinants are estimated to contribute to a minimum of 20% of the global HIV-1 infections (Hemelaar et al., 2011). Unique recombinant forms (URFs)

against viral proteins of non-B HIV-1 subtypes (Velazquez-Campoy et al., 2003; Mosebi et al., 2008; Ahmed et al., 2013). Interestingly, the proportion of non-B subtypes in the Americas and Europe is on the rise, especially those subtypes that predominate in Africa (Descamps et al., 2005; Holguin et al., 2008; Kanizsai et al., 2010; Foster et al., 2014; Mendoza et al., 2014). Therefore, the sustainability and future efficacy of ARVs in clinical use today are a concern.

HIV-1 subtype C

Subtype C viruses account for approximately 50% of HIV-1 infections worldwide (Hemelaar et al., 2011). The increased prevalence of this subtype in the Western world and its dominance in sub-Saharan Africa and India pose vital questions on the approach of current drug development strategies and the efficacy of long-term use of ARVs. Irrespective of the subtype, efficient processing of Gag and Gag-Pol polyproteins is required for viral propagation (Kaplan et al., 1993). In total, there are 12 HIV-1 cleavage sites in the Gag, Gag-Pol, and nef precursor proteins (Figure 2) (de Oliveira et al., 2003). The majority of these sites are significantly more diverse in subtype C viruses when compared with subtype B viruses (de Oliveira et al., 2003). Natural variation at cleavage sites in subtype C viruses may be involved in disease progression, response to therapy, and regulation of the viral cycle (de Oliveira et al., 2003). The disproportionate increase in the prevalence of the subtype C strain in comparison to other HIV-1 viruses implies that subtype C HIV-1 may have a greater level of viral fitness at the population level or may be more easily transmitted (Bessong, 2008). These intriguing details warrant further investigation into the structural and functional features of the subtype C PR and a comparison to the well-investigated subtype B PR.
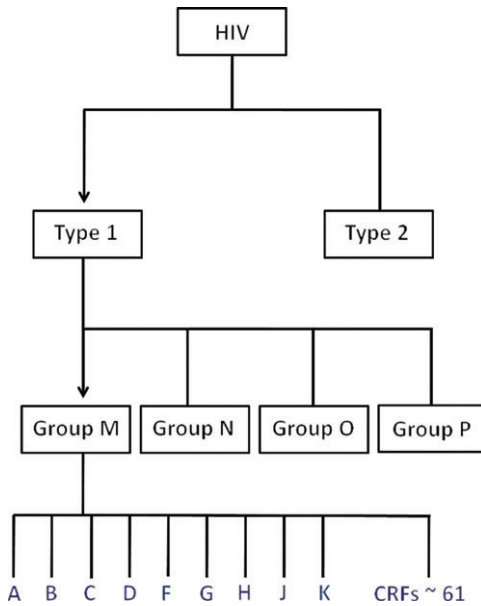
Figure 1 Groups of HIV-1.
The nine subtypes of the major group are displayed. There have been reports of 61 CRFs in HIV-1 group M to date (Los Alamos HIV sequence database, http://www.hiv.lanl.gov/content/sequences/HIV/CRFs/CRFs.html).

refer to strains that do not satisfy the aforementioned criteria (Robertson et al., 2000). The extension designated to a CRF indicates the subtype from which the CRF is derived (Hemelaar, 2012). If the CRF consists of contributions from more than two subtypes, it bears a 'cpx' (complex) extension (Hemelaar, 2012). Recombination events have been reported between different strains from different HIV-1 groups (M and O) as well as within and between group M subtypes (Rousseau et al., 2007). Importantly, intra-subtype recombination has been established as being widespread within group M, subtype C HIV-1 (Rousseau et al., 2007). Recombination events are ongoing in many countries worldwide, where different CRFs and subtypes co-circulate (Hemelaar, 2012). A high prevalence of URFs is observed in West Africa, and the CRFs involved include CRF02, CRF06_cpx, and CRF09_cpx (Delgado et al., 2008). Some CRFs have recombined even further with other CRFs or subtypes, resulting in second-generation recombinants (SGRs) (Hemelaar, 2012).

HIV-1 subtype C is the most common subtype worldwide and occurs mainly in sub-Saharan Africa, India, Brazil, and China, and these regions are most heavily afflicted by HIV and AIDS (Hemelaar et al., 2011). Studies have been performed on the interaction of PIs with PRs of subtype B HIV origin, which occurs mainly in regions of North America and Europe (Hemelaar et al., 2011). More recently, studies have emerged on the efficacy of ARVs

Comparison of viral fitness between subtype B and C proteases

Mature dimeric PR is released from the Gag-Pol polyprotein during precursor processing (Kaplan et al., 1993). Production of mature PR appears to be catalyzed by the Gag-Pol precursor itself (Pettit et al., 2005). It is still unclear whether the initial cleavage is accomplished intra- or inter-molecularly (Wiegers et al., 1998; Pettit et al., 2005). The active site of the PR is formed at the dimer
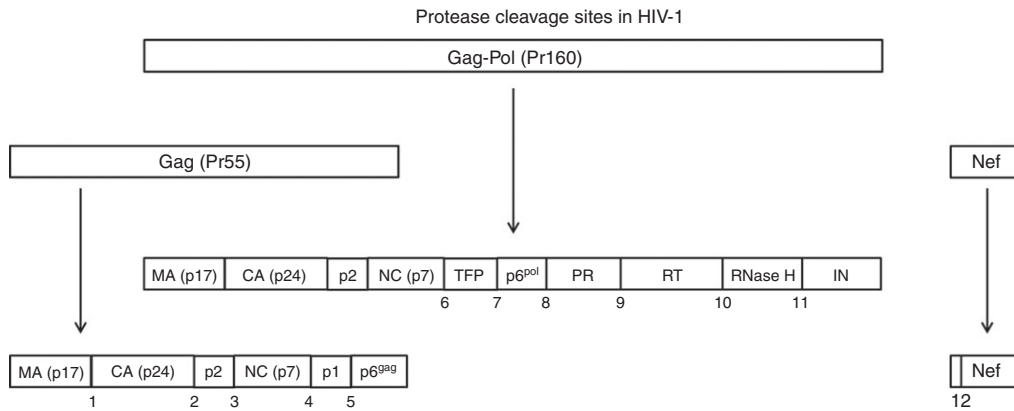
Protease cleavage sites in HIV-1

Figure 2 HIV-1 PR cleavage sites within Gag, Gag-Pol, and nef proteins.
Proteins released following cleavage of Gag and Gag-Pol are matrix (MA), capsid (CA), nucleocapsid (NC), p2, p1, transframe protein (TFP), p6^{gag}, p6^{pol}, PR, RT, RNase H, and integrase (IN) (de Oliveira et al., 2003).

interface by the enzymatic residues (D25/D25′) shown in Figure 3. Aberrant virions that are significantly less infective may appear due to defects in the PR activity (Kaplan et al., 1993; Wiegers et al., 1998; Huang and Chen, 2010). Such defects include mutations that modify the order of cleavage or the rate of processing or that result in incorrect cleavage at any site (Kaplan et al., 1993; Wiegers et al., 1998; Huang and Chen, 2010). Therefore, the PR remains a major target of anti-HIV drug development and one of the most well-studied enzymes by researchers in academia and industry. This is evidenced by the development of 10

U.S. Food and Drug Administration (FDA)-approved PIs from December 1995 to June 2006 (Table 1).

The PDB is replete with HIV PR structures, many of which, however, are redundant. Thus far, the majority of PR crystal structures in the PDB are representative of HIV-1 subtype B, and only four crystal structures of HIV-1 subtype C PR have been solved. We recently published the first crystal structure of the South African HIV-1 subtype C PR (Naicker et al., 2013a,b). Interestingly, even within HIV subtypes, genetic diversity occurs due to the error-prone nature of the virus as well as mutations resulting from
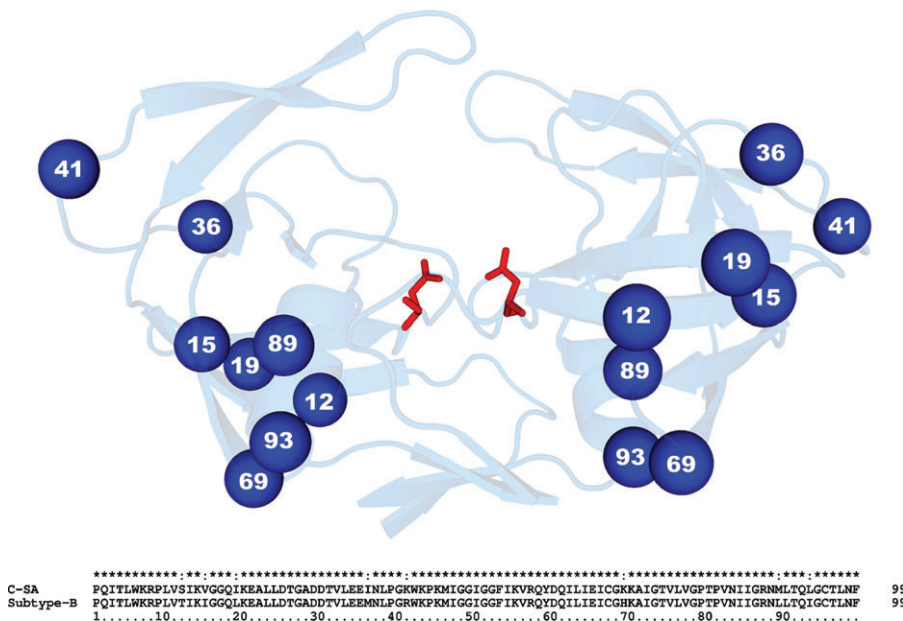
Figure 3 Cartoon representation of the HIV-1 PR.
Spheres represent positions of the polymorphic sites that exist between subtype B and C-SA PRs. Active site residues (D25) shown as sticks. Sequence alignment below shows differences between the PRs at the level of primary structure. PDB ID: 3U71 (Naicker et al., 2013a,b).

environmental factors and drug pressure (Taylor et al., 2008; Hemelaar et al., 2011). The consensus sequence of the HIV-1 subtype C PR occurring in South Africa (C-SA PR) shares the same sequence identity as the consensus subtype C PR in group M (Los Alamos HIV sequence database, http://www.hiv.lanl.gov/). It is anticipated that approximately 15%–20% of adults in South Africa will experience virologic failure within the first 2 years of treatment with the first-line ARV regimen (Coetzee et al., 2004). Therefore, there is still a dependency on the second-line regimen comprising RT inhibitors and PIs in South Africa. The C-SA PR differs from the consensus subtype B PR at eight positions in each monomer (Los Alamos HIV sequence database, http://www.hiv.lanl.gov/) shown in Figure 3. The polymorphisms occur distal to the active site and mutations at these sites do not affect drug binding/substrate cleavage directly.

From among the plethora of biochemical techniques available, isothermal titration calorimetry represents the most powerful and versatile method to evaluate the thermodynamics and binding affinity of a drug toward its target in a single experiment. Thermodynamic analyses of interactions between the PR and FDA-approved PIs indicate that the C-SA PR, generally, displays reduced susceptibility to PIs (Velazquez-Campoy et al., 2003; Mosebi et al., 2008) and that binding affinities of PIs were 3- to 24-fold weaker for the C-SA PR than for the subtype B PR (Velazquez-Campoy et al., 2003). Amprenavir, indinavir, nelfinavir, and saquinavir displayed 3- to 6-fold weaker binding, whereas lopinavir and ritonavir displayed significant reductions in affinity with 8- to 24-fold weaker binding (Velazquez-Campoy et al., 2003). The calculated binding affinity of PIs measured during a 2-ns molecular dynamics (MD) simulation showed a general reduction of 0.5–2 kcal/mol in binding energy for the C-SA PR relative to the subtype B PR (Ahmed et al., 2013). Apart from the genetic variability observed between B and C-PRs, phenotypic assays revealed that IC_{90} values for subtype C HIV-1 strains were 0.16- to 3-fold greater than the control subtype B strain in the presence of indinavir, nelfinavir, ritonavir, and saquinavir (Shafer et al., 1999). These results suggest that subtype C HIV-1 may have improved viral fitness and marginally reduced drug susceptibility to most FDA-approved PIs in comparison to the subtype B HIV-1.

Clinical data from ARV treatment programs in sub-Saharan Africa reveal that drug-resistance patterns are similar between B and non-B HIV-1 subtypes (Barth et al., 2010). Interestingly, naturally occurring polymorphisms in strains such as subtype C HIV-1 have an additive effect on resistance mutations (Velazquez-Campoy et al., 2002). Only few studies have directly compared the clinical outcome of ARV treatments among patients harboring different subtypes (Del Amo et al., 1998; Alaeus et al., 1999; Alexander et al., 2002; Pillay et al., 2002). Owing to the large diversity of HIV strains globally, clinical studies compare disease progression in individuals infected with varying HIV subtypes; however, these studies have not directly compared individuals infected with either subtype B or subtype C HIV-1 alone. However, in drug-experienced HIV-infected individuals, there is a strong correlation between unfavorable virologic response to ritonavir/saquinavir treatments and mutations at protease positions 10, 36, and 93 (Harrigan et al., 1999; Zolopa et al., 1999). Positions 36 and 93 are natural polymorphic sites in the C-SA PR. In contrast, studies of patients infected with B or non-B HIV-1 strains in Canada (Alexander et al., 2002) and Europe (Del Amo et al., 1998; Alaeus et al., 1999; Pillay et al., 2002) revealed no association between HIV-1 subtype and disease progression. These findings may indeed be complicated by patients infected with multiple HIV-1 strains and disparity in treatment regimens.

A recent structural and computational study indicated that the differences drug efficacy may be explained by key differences at the hinge region of the PR comprising amino acids at residue positions 35–42 and 57–61 (Naicker et al., 2013a,b). The stability of the hinge region of the PR influences the movement of the closely associated flaps comprising amino acids at residue positions 46–54. The flaps shield the PR-active site (seen above the active site in Figure 3), and therefore, flap opening is required for substrate/inhibitor entry into the active site. Increased flap flexibility is associated with weaker interactions with inhibitors. MD simulation results showed that although the extent of flap opening of both the C-SA and subtype B PRs are similar, the inter-flap distance measured showed greater fluctuation for the C-SA PR for the majority of the 10-ns simulation (Naicker et al., 2013a,b). Thus, C-SA PR flaps sample a larger range of conformations, which implies greater flexibility around the flap tips. Because the flap regions of both PRs are identical, the mechanism of increased flexibility may not be straightforward. Detailed structural analysis of the apo-forms of the C-SA and subtype B PRs showed that increased flexibility of the flap region may be due to subtle differences in the hinge region (Naicker et al., 2013a,b). The difference is due to the absence of the E35-R57 salt bridge in the C-SA PR. In the subtype B PR, M36 most likely facilitates formation of the salt bridge, whereas I36 in the C-SA PR does not facilitate the formation of the salt bridge. The E35-R57 ionic interaction is the only salt bridge in the flap-hinge region of the PR and may be a key requirement for maintaining the rigidity of the hinge region. Therefore, polymorphisms at

position 36 are likely to increase flexibility of the flap tips, which is clearly evident in the C-SA PR crystal structure and MD simulations (Naicker et al., 2013a, b).

Secondary resistance mutations in HIV-1 subtype C PR

Secondary resistance mutations occur distal to the active site and are not implicated by themselves in high-level drug resistance (Rose et al., 1996; Nijhuis et al., 1999). They either contribute to drug resistance in the presence of selected primary resistance mutations or upregulate PR activity to compensate for the reduction in catalytic efficiency caused by the selected primary PR resistance mutations (Rose et al., 1996; Nijhuis et al., 1999). Certain secondary resistance mutations (e.g., M36I) may also increase the genetic barrier, thereby enhancing the emergence of primary resistance mutations (Perno et al., 2004). Both D30N and L90M primary DRMs occur in non-B viruses during nelfinavir therapy. However, D30N occurs more commonly in subtype B viruses, whereas L90M occurs more frequently in subtype C, F, G, and CRF 01_AE viruses (Cane et al., 2001; Abecasis et al., 2005). Importantly, the increased preference for certain subtypes to develop L90M may be associated with the occurrence of residues other than leucine (subtype B consensus) at position 89 (e.g., methionine in subtype C) (Cane et al., 2001; Abecasis et al., 2005; Shafer and Schapiro, 2008). The residue at position 90 is situated on the α-helix of the terminal domain, which interacts with the surface of the hydrophobic core domain and L90M may disrupt the packing of the two domains (Rose et al., 1998). The same mutation often occurs in response to inhibitors of entirely different chemical scaffolds. Therefore, these drug-resistant mutations are conserved and suggest a mechanism of drug resistance independent of the inhibitor used (Rose et al., 1998). T47S is another polymorphism that is linked to reduced susceptibility toward nelfinavir and occurs in approximately 8% of subtype C viruses, but rarely in other subtypes (Rhee et al., 2006). These findings suggest that subtype C viruses are more prone to develop drug resistance to PIs than subtype B viruses.

Specific mutations (e.g., V82I) within the PR have been found to play a greater role in drug resistance. The occurrence of such mutations appears to differ between subtypes (Kantor and Katzenstein, 2003). The V82I mutation occurs in approximately 1% of untreated HIV-1 subtype B-infected individuals. Interestingly, this mutation has been observed in 6% and 9% of subtype C- and subtype F-infected individuals, respectively, and is the consensus (>50%) in untreated subtype G-infected individuals (Kantor and Katzenstein, 2003). In South Africa, as in other European countries (Pereira-Vaz et al., 2009), amino acid insertions are now being detected more frequently in the PR hinge region in treatment-naïve HIV-1 subtype C patients (Bessong et al., 2006; Pereira-Vaz et al., 2009). With improving HIV surveillance in Africa and India, it is expected that more of these HIV-1 variants will be detected. The exact role of insertions have not been elucidated, however, they may be implicated in improving viral fitness by increasing the catalytic efficiency of the PR (Kozisek et al., 2008). Inhibition data suggest that insertions in the hinge region contribute to PI resistance only in combination with other mutations in the PR or Gag and Gag-Pol (Kozisek et al., 2008).

Novel strategies for overcoming drug resistance and future directions

All currently available PIs are active site inhibitors. Efforts are currently underway to design more adaptable PIs that will adjust to slight structural changes due to resistance mutations. New alternative classes of PIs may provide a higher barrier to resistance and improve HIV PR selectivity. Improved drug selectivity may help reduce side effects associated with antiviral therapy (e.g., lypodystrophy syndrome, hyperlipidemia, diarrhea). One such alternative is dimerization inhibitors. HIV PR is an obligate homodimer, and disruption of dimer formation leads to inactivation of the PR (Hansen et al., 1988). The terminal dimer interface (residue positions 1–5 and 95–99) is conserved across all HIV subtypes, and resistance mutations at these sites are not likely to develop. Disruption of a few key interactions at the terminal dimer interface has been shown to eliminate the formation of a PR dimer (Choudhury et al., 2003; Pettit et al., 2003; Naicker et al., 2013a, b). Mutation of the ultimate amino acid, F99, to an alanine resulted in the formation of two monomers with the concomitant loss in enzyme activity (Naicker et al., 2013a, b). Experimental dimerization inhibitors showed reduced potency relative to their clinically available active site inhibitor counterparts (Babe et al., 1992). Slight improvements of dimerization inhibitors are required to make them clinically effective, especially when considering they have a high barrier to resistance and may display reduced side effects.

Improved pharmacological features such as drug absorption and half-life are always important considerations during drug development. Recent investigations with our collaborators has led to the discovery and synthesis of novel transition-state analogues, pentacycloundecane (PCU) cage-incorporated peptides, with anti-HIV PR activity (Makatini et al., 2011; Karpoormath et al., 2013; Makatini et al., 2013). The lipophilic nature of the cage compounds confers excellent chemical properties required for drug delivery. Cage compounds covalently linked to pharmaceutically active molecules have been shown to reduce drug biodegradation, thereby increasing the half-life of the drug (Ito et al., 2007). The bulky hydrocarbon scaffold aids in the transport of the drugs across cellular membranes (Ito et al., 2007). The PCU-incorporated compound (PCU-lactam-EAIS) with the greatest potency (IC_{50}=0.078 μM) serves as a template for compounds with higher efficacy (Makatini et al., 2011). The coupling of the PCU-lactam (Figure 4A) to a variety of peptide side chains showed minimal effect on the inhibitory potential (Makatini et al., 2011). Removal of the cage resulted in a complete loss of inhibitory activity (Makatini et al., 2011). Therefore, the cage lactam is the key component for inhibition. PCU-peptides showed between 6000 and 8000 times less toxicity to human MT-4 cells than lopinavir (Makatini et al., 2011), further enhancing the potential of these novel PIs.

Other novel PIs currently under development and that are being tested against the C-SA PR include linear and cyclic glycopeptides incorporating a hydroxy propyl amine as a transition-state analogue (Pawar et al., 2013). Nuclear magnetic resonance studies have shown that the

sugar moiety induces a turn in the linear peptides and shows a good correlation with the corresponding cyclic peptides (Pawar et al., 2013). The hydroxy propyl amine forms hydrogen bonds with the enzymatic residues (D25/D25′) of the PR, showing great potential as a transition-state analogue. Another exciting development is the formulation of dual-action HIV-1 PR/RT inhibitors (Figure 4B). Coumarin moieties conjugated to the RT inhibitor azidothymidine (AZT) have shown promise as dual-action inhibitors (Olomola et al., 2013). The AZT moiety showed complete inhibitory potency, whereas the coumarin derivative exhibited moderate inhibition of the PR (two to four times weaker inhibition than ritonavir) (Olomola et al., 2013). The N-benzylated coumarin-AZT analogues displayed great promise, with the N-benzyl group enhancing the RT inhibition activity. These compounds were isolated in high yields, which ranged between 70% and 80% (Olomola et al., 2013), further validating the potential of these compounds as novel dual-action inhibitors.

ARV interventions that are effective against a variety of HIV subtypes and drug-resistant strains are focus areas in HIV/AIDS research at present. Overcoming drug resistance is not limited to the development of novel PIs. Additionally, the production of an effective HIV vaccine is the aspiration of many researchers, particularly a vaccine that confers immunity to HIV-1 strains from sub-Saharan Africa. Efforts to stimulate production of broadly cross-neutralizing (BCN) antibodies through vaccination have been largely unsuccessful (Haynes and Montefiori, 2006; Hu and Stamatatos, 2007). A minority of individuals develop antibodies naturally after several years post infection (Stamatatos et al., 2009). Little is known about

Q6: Please confirm expansion for NMR

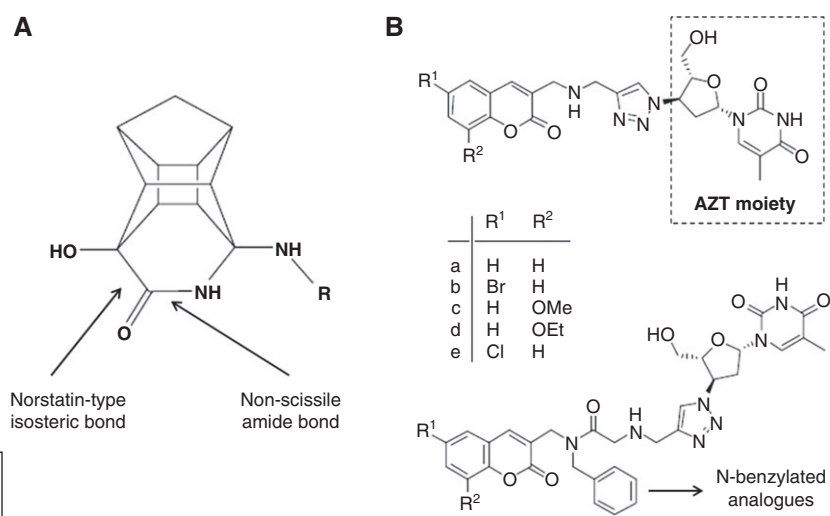Q7: Please supply main caption for Figure 4

Figure 4
(A) Scaffold of the non-cleavable PCU-lactam peptides (R=peptide) (Makatini et al., 2013). (B) Scaffold of the coumarin-AZT conjugates. A series of both conjugates were produced varying from compounds a to e (Olomola et al., 2013).

how these specific neutralizing antibodies arise. However, recent encouraging strides have been made that enhance our understanding of BCN antibodies (Moore et al., 2012). Glycans are known to play an important role in shielding neutralizing epitopes on the HIV envelope and thus viral escape. BCN responses have been shown to directly target these glycans including one at position 332 in the immunogenic C3 region of the gp120 subunit of the HIV-1 envelope protein (Walker et al., 2011; Moore et al., 2012).

Importantly, in HIV-1 subtype C viruses, the absence of the 332 glycan is favored at transmission (Moore et al., 2012). Absence of the 332 glycan confers resistance to Asn332-dependent BCN monoclonal antibodies (Walker et al., 2011). Shortly after transmission, immune pressure by strain-specific neutralizing antibodies triggers the evolution of the 332 glycan (Moore et al., 2012). The evolution of such epitopes provide vital targets for neutralization of subtype C viruses. Two patients from a cohort of 79 subtype C HIV-1-infected women developed Asn332-dependent BCN antibodies (Moore et al., 2012). The glycan that formed the core of the BCN epitope was absent on the infecting virus for both patients but evolved shortly thereafter (Moore et al., 2012). Using clones of the acute virus and a 6-month representative virus of both patients, novel BCN epitopes were identified. The transfer of a glycan from position 334 to 332 allowed the virus to escape autologous neutralizing antibodies; however, this created a new BCN epitope that provided the antigenic stimulus to produce BCN antibodies targeting the 332 glycan (Moore et al., 2012). A glycan at position 160, which was initially absent was also shown to emerge in patients after 3–6 months of transmission with HIV-1 (Moore et al., 2012). Interestingly, BCN antibodies develop more frequently in infected individuals for those targeting epitopes directed toward the 160 or 332 glycans rather than the membrane-proximal region or CD4 binding site (Walker et al., 2010; Tomaras et al., 2011). These useful insights are of great value for the production of effective HIV vaccines, especially those conferring immunity to subtype C HIV-1.

Great advances are being made to elucidate the mechanisms behind the reduced drug efficacy associated with subtype variation. A number of drug development strategies to improve PIs are being explored. The C-SA PR remains an epidemiologically relevant protein to investigate, and further structural and dynamic analyses of this enzyme may reveal key insights for future drug development.

References

Abecasis, A.B., Deforche, K., Snoeck, J., Bacheler, L.T., McKenna, P., Carvalho, A.P., Gomes, P., Camacho, R.J., and Vandamme, A.M. (2005). Protease mutation M89I/V is linked to therapy failure in patients infected with the HIV-1 non-B subtypes C, F or G. AIDS *19*, 1799–1806.

Ahmed, S.M., Kruger, H.G., Govender, T., Maguire, G.E.M., Sayed, Y., Ibrahim, M.A.A., Naicker, P., and 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.

Alaeus, A., Lidman, K., Bjorkman, A., Giesecke, J., and Albert, J. (1999). Similar rate of disease progression among individuals infected with HIV-1 genetic subtypes A–D. AIDS *13*, 901–907.

Alexander, C.S., Montessori, V., Wynhoven, B., Dong, W., Chan, K., O'Shaughnessy, M.V., Mo, T., Piasecny, M., Montaner, J.S., and Harrigan, P.R. (2002). Prevalence and response to antiretroviral therapy of non-B subtypes of HIV in antiretroviral-naive individuals in British Columbia. Antivir. Ther. *7*, 31–35.

Babe, L.M., Rose, J., and Craik, C.S. (1992). Synthetic "interface" peptides alter dimeric assembly of the HIV 1 and 2 proteases. Protein Sci. *1*, 1244–1253.

Barth, R.E., van der Loeff, M.F., Schuurman, R., Hoepelman, A.I., and Wensing, A.M. (2010). Virological follow-up of adult patients in antiretroviral treatment programmes in sub-Saharan Africa: a systematic review. Lancet Infect. Dis. *10*, 155–166.

Bessong, P.O. (2008). Polymorphisms in HIV-1 subtype C proteases and the potential impact on protease inhibitors. Trop. Med. Int. Health *13*, 144–151.

Bessong, P.O., Mphahlele, J., Choge, I.A., Obi, L.C., Morris, L., Hammarskjold, M.L., and Rekosh, D.M. (2006). Resistance mutational analysis of HIV type 1 subtype C among rural South African drug-naive patients prior to large-scale availability of antiretrovirals. AIDS Res. Hum. Retroviruses *22*, 1306–1312.

Cane, P.A., de Ruiter, A., Rice, P., Wiselka, M., Fox, R., and Pillay, D. (2001). Resistance-associated mutations in the human immunodeficiency virus type 1 subtype c protease gene from treated and untreated patients in the United Kingdom. J. Clin. Microbiol. *39*, 2652–2654.

Choudhury, S., Everitt, L., Pettit, S.C., and Kaplan, A.H. (2003). Mutagenesis of the dimer interface residues of tethered and untethered HIV-1 protease result in differential activity and suggest multiple mechanisms of compensation. Virology *307*, 204–212.

Coetzee, D., Hildebrand, K., Boulle, A., Maartens, G., Louis, F., Labatala, V., Reuter, H., Ntwana, N., and Goemaere, E. (2004). Outcomes after two years of providing antiretroviral treatment in Khayelitsha, South Africa. AIDS *18*, 887–895.

de Oliveira, T., Engelbrecht, S., Janse van Rensburg, E., Gordon, M., Bishop, K., zur Megede, J., Barnett, S.W., and Cassol, S. (2003).

Variability at human immunodeficiency virus type 1 subtype C protease cleavage sites: an indication of viral fitness? J. Virol. *77*, 9422–9430.

Del Amo, J., Petruckevitch, A., Phillips, A., Johnson, A.M., Stephenson, J., Desmond, N., Hanscheid, T., Low, N., Newell, A., Obasi, A., et al. (1998). Disease progression and survival in HIV-1-infected Africans in London. AIDS *12*, 1203–1209.

Delgado, E., Ampofo, W.K., Sierra, M., Torpey, K., Perez-Alvarez, L., Bonney, E.Y., Mukadi, Y.D., Lartey, M., Nyarko, C., Amenyah, R.N., et al. (2008). High prevalence of unique recombinant forms of HIV-1 in Ghana: molecular epidemiology from an antiretroviral resistance study. J. Acquir. Immune. Defic. Syndr. *48*, 599–606.

Descamps, D., Chaix, M.L., Andre, P., Brodard, V., Cottalorda, J., Deveau, C., Harzic, M., Ingrand, D., Izopet, J., Kohli, E., et al. (2005). French national sentinel survey of antiretroviral drug resistance in patients with HIV-1 primary infection and in antiretroviral-naive chronically infected patients in 2001–2002. J. Acquir. Immune. Defic. Syndr. *38*, 545–552.

Foster, G.M., Ambrose, J.C., Hue, S., Delpech, V.C., Fearnhill, E., Abecasis, A.B., Leigh Brown, A.J., and Geretti, A.M. (2014). Novel HIV-1 recombinants spreading across multiple risk groups in the United Kingdom: the identification and phylogeography of circulating recombinant form (CRF) 50_A1D. PLoS One *9*, e83337.

Hansen, J., Billich, S., Schulze, T., Sukrow, S., and Moelling, K. (1988). Partial purification and substrate analysis of bacterially expressed HIV protease by means of monoclonal antibody. EMBO J. *7*, 1785–1791.

Harrigan, P.R., Hertogs, K., Verbiest, W., Pauwels, R., Larder, B., Kemp, S., Bloor, S., Yip, B., Hogg, R., Alexander, C., et al. (1999). Baseline HIV drug resistance profile predicts response to ritonavir-saquinavir protease inhibitor therapy in a community setting. AIDS *13*, 1863–1871.

Haynes, B.F. and Montefiori, D.C. (2006). Aiming to induce broadly reactive neutralizing antibody responses with HIV-1 vaccine candidates. Expert Rev. Vaccines *5*, 347–363.

Hemelaar, J. (2012). The origin and diversity of the HIV-1 pandemic. Trends. Mol. Med. *18*, 182–192.

Hemelaar, J., Gouws, E., Ghys, P.D., and Osmanov, S. (2011). Global trends in molecular epidemiology of HIV-1 during 2000–2007. AIDS *25*, 679–689.

Holguin, A., de Mulder, M., Yebra, G., Lopez, M., and Soriano, V. (2008). Increase of non-B subtypes and recombinants among newly diagnosed HIV-1 native Spaniards and immigrants in Spain. Curr. HIV Res. *6*, 327–334.

Hu, S.L. and Stamatatos, L. (2007). Prospects of HIV Env modification as an approach to HIV vaccine design. Curr. HIV Res. *5*, 507–513.

Huang, L. and Chen, C. (2010). Autoprocessing of human immunodeficiency virus type 1 protease miniprecursor fusions in mammalian cells. AIDS Res. Ther. *7*, 27.

Ito, F.M., Petroni, J.M., de Lima, D.P., Beatriz, A., Marques, M.R., de Moraes, M.O., Costa-Lotufo, L.V., Montenegro, R.C., Magalhaes, H.I., and Pessoa Cdo, O. (2007). Synthesis and biological evaluation of rigid polycyclic derivatives of the diels-alder adduct tricyclo[6.2.1.02,7]undeca-4,9-dien-3,6-dione. Molecules *12*, 271–282.

Kanizsai, S., Ghidan, A., Ujhelyi, E., Banhegyi, D., and Nagy, K. (2010). Monitoring of drug resistance in therapy-naive HIV infected patients and detection of African HIV subtypes in Hungary. Acta Microbiol. Immunol. Hung. *57*, 55–68.

Kantor, R. and Katzenstein, D. (2003). Polymorphism in HIV-1 non-subtype B protease and reverse transcriptase and its potential impact on drug susceptibility and drug resistance evolution. AIDS Rev. *5*, 25–35.

Kaplan, A.H., Zack, J.A., Knigge, M., Paul, D.A., Kempf, D.J., Norbeck, D.W., and Swanstrom, R. (1993). Partial inhibition of the human immunodeficiency virus type 1 protease results in aberrant virus assembly and the formation of noninfectious particles. J. Virol. *67*, 4050–4055.

Karpoormath, R., Sayed, Y., Govender, T., Kruger, H.G., Soliman, M.E.S., and Maguire, G.E.M. (2013). Novel PCU cage diol peptides as potential targets against wild-type CSA HIV-1 protease: synthesis, biological screening and molecular modelling studies. Med. Chem. Res. *22*, 3918–3933.

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

Makatini, M.M., Petzold, K., Sriharsha, S.N., Ndlovu, N., Soliman, M.E., Honarparvar, B., Parboosing, R., Naidoo, A., Arvidsson, P.I., Sayed, Y., et al. (2011). Synthesis and structural studies of pentacycloundecane-based HIV-1 PR inhibitors: a hybrid 2D NMR and docking/QM/MM/MD approach. Eur. J. Med. Chem. *46*, 3976–3985.

Makatini, M.M., Petzold, K., Alves, C.N., Arvidsson, P.I., Honarparvar, B., Govender, P., Govender, T., Kruger, H.G., Sayed, Y., Lameira, J., et al. (2013). Synthesis, 2D-NMR and molecular modelling studies of pentacycloundecane lactam-peptides and peptoids as potential HIV-1 wild type C-SA protease inhibitors. J. Enzyme Inhib. Med. Chem. *28*, 78–88.

Martinez-Cajas, J.L., Pai, N.P., Klein, M.B., and Wainberg, M.A. (2009). Differences in resistance mutations among HIV-1 non-subtype B infections: a systematic review of evidence (1996–2008). J. Int. AIDS Soc. *12*, 11.

Mendoza, Y., Bello, G., Castillo Mewa, J., Martinez, A.A., Gonzalez, C., Garcia-Morales, C., Avila-Rios, S., Reyes-Teran, G., and Pascale, J.M. (2014). Molecular epidemiology of HIV-1 in Panama: origin of non-B subtypes in samples collected from 2007 to 2013. PLoS One *9*, e85153.

Moore, P.L., Gray, E.S., Wibmer, C.K., Bhiman, J.N., Nonyane, M., Sheward, D.J., Hermanus, T., Bajimaya, S., Tumba, N.L., Abrahams, M.R., et al. (2012). Evolution of an HIV glycan-dependent broadly neutralizing antibody epitope through immune escape. Nat. Med. *18*, 1688–1692.

Mosebi, S., Morris, L., Dirr, H.W., and 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.

Naicker, P., Achilonu, I., Fanucchi, S., Fernandes, M., Ibrahim, M.A.A., Dirr, H.W., Soliman, M.E.S., and 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.

Naicker, P., Seele, P., Dirr, H.W., and Sayed, Y. (2013b). F99 is critical for dimerization and activation of South African HIV-1 subtype C protease. Protein J. *32*, 560–567.

Nijhuis, M., Schuurman, R., de Jong, D., Erickson, J., Gustchina, E., Albert, J., Schipper, P., Gulnik, S., and 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.

Olomola, T.O., Klein, R., Mautsa, N., Sayed, Y., and Kaye, P.T. (2013). Synthesis and evaluation of coumarin derivatives as potential dual-action HIV-1 protease and reverse transcriptase inhibitors. Bioorg. Med. Chem. 21, 1964–1971.

Pawar, S.A., Jabgunde, A.M., Maguire, G.E., Kruger, H.G., Sayed, Y., Soliman, M.E., Dhavale, D.D., and Govender, T. (2013). Linear and cyclic glycopeptide as HIV protease inhibitors. Eur. J. Med. Chem. 60, 144–154.

Pereira-Vaz, J., Duque, V., Trindade, L., Saraiva-da-Cunha, J., and Melico-Silvestre, A. (2009). Detection of the protease codon 35 amino acid insertion in sequences from treatment-naive HIV-1 subtype C infected individuals in the Central Region of Portugal. J. Clin. Virol. 46, 169–172.

Perno, C.F., Cozzi-Lepri, A., Forbici, F., Bertoli, A., Violin, M., Stella Mura, M., Cadeo, G., Orani, A., Chirianni, A., De Stefano, C., et al. (2004). Minor mutations in HIV protease at baseline and appearance of primary mutation 90M in patients for whom their first protease-inhibitor antiretroviral regimens failed. J. Infect. Dis. 189, 1983–1987.

Pettit, S.C., Gulnik, S., Everitt, L., and Kaplan, A.H. (2003). The dimer interfaces of protease and extra-protease domains influence the activation of protease and the specificity of GagPol cleavage. J. Virol. 77, 366–374.

Pettit, S.C., Clemente, J.C., Jeung, J.A., Dunn, B.M., and Kaplan, A.H. (2005). Ordered processing of the human immunodeficiency virus type 1 GagPol precursor is influenced by the context of the embedded viral protease. J. Virol. 79, 10601–10607.

Pillay, D., Walker, A.S., Gibb, D.M., de Rossi, A., Kaye, S., Ait-Khaled, M., Munoz-Fernandez, M., and Babiker, A. (2002). Impact of human immunodeficiency virus type 1 subtypes on virologic response and emergence of drug resistance among children in the Paediatric European Network for Treatment of AIDS (PENTA) 5 trial. J Infect. Dis. 186, 617–625.

Plantier, J.C., Leoz, M., Dickerson, J.E., De Oliveira, F., Cordonnier, F., Lemee, V., Damond, F., Robertson, D.L., and Simon, F. (2009). A new human immunodeficiency virus derived from gorillas. Nat. Med. 15, 871–872.

Rhee, S.Y., Taylor, J., Wadhera, G., Ben-Hur, A., Brutlag, D.L., and Shafer, R.W. (2006). Genotypic predictors of human immunodeficiency virus type 1 drug resistance. Proc. Natl. Acad. Sci. USA 103, 17355–17360.

Robertson, D.L., Anderson, J.P., Bradac, J.A., Carr, J.K., Foley, B., Funkhouser, R.K., Gao, F., Hahn, B.H., Kalish, M.L., Kuiken, C., et al. (2000). HIV-1 nomenclature proposal. Science 288, 55–56.

Rose, R.E., Gong, Y.F., Greytok, J.A., Bechtold, C.M., Terry, B.J., Robinson, B.S., Alam, M., Colonno, R.J., and Lin, P.F. (1996). Human immunodeficiency virus type 1 viral background plays a major role in development of resistance to protease inhibitors. Proc. Natl. Acad. Sci. USA 93, 1648–1653.

Rose, R.B., Craik, C.S., and Stroud, R.M. (1998). Domain flexibility in retroviral proteases: structural implications for drug resistant mutations. Biochemistry 37, 2607–2621.

Rousseau, C.M., Learn, G.H., Bhattacharya, T., Nickle, D.C., Heckerman, D., Chetty, S., Brander, C., Goulder, P.J., Walker, B.D., Kiepiela, P., et al. (2007). Extensive intrasubtype recombination in South African human immunodeficiency virus type 1 subtype C infections. J. Virol. 81, 4492–4500.

Shafer, R.W. and Schapiro, J.M. (2008). HIV-1 drug resistance mutations: an updated framework for the second decade of HAART. AIDS Rev. 10, 67–84.

Shafer, R.W., Chuang, T.K., Hsu, P., White, C.B., and 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.

Stamatatos, L., Morris, L., Burton, D.R., and Mascola, J.R. (2009). Neutralizing antibodies generated during natural HIV-1 infection: good news for an HIV-1 vaccine? Nat. Med. 15, 866–870.

Taylor, B.S., Sobieszczyk, M.E., McCutchan, F.E., and Hammer, S.M. (2008). The challenge of HIV-1 subtype diversity. N. Engl. J. Med. 358, 1590–1602.

The Joint Nations Program on HIV/AIDS (UNAIDS) (2013). Global report: UNAIDS report on the global AIDS epidemic 2013 (Geneva, Switzerland: UNAIDS).

Tomaras, G.D., Binley, J.M., Gray, E.S., Crooks, E.T., Osawa, K., Moore, P.L., Tumba, N., Tong, T., Shen, X., Yates, N.L., et al. (2011). Polyclonal B cell responses to conserved neutralization epitopes in a subset of HIV-1-infected individuals. J. Virol. 85, 11502–11519.

Velazquez-Campoy, A., Vega, S., and Freire, E. (2002). Amplification of the effects of drug resistance mutations by background polymorphisms in HIV-1 protease from African subtypes. Biochemistry 41, 8613–8619.

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

Wainberg, M.A. and Brenner, B.G. (2010). Role of HIV subtype diversity in the development of resistance to antiviral drugs. Viruses 2, 2493–2508.

Walker, L.M., Simek, M.D., Priddy, F., Gach, J.S., Wagner, D., Zwick, M.B., Phogat, S.K., Poignard, P., and Burton, D.R. (2010). A limited number of antibody specificities mediate broad and potent serum neutralization in selected HIV-1 infected individuals. PLoS Pathog. 6, e1001028.

Walker, L.M., Huber, M., Doores, K.J., Falkowska, E., Pejchal, R., Julien, J.P., Wang, S.K., Ramos, A., Chan-Hui, P.Y., Moyle, M., et al. (2011). Broad neutralization coverage of HIV by multiple highly potent antibodies. Nature 477, 466–470.

Wiegers, K., Rutter, G., Kottler, H., Tessmer, U., Hohenberg, H., and Krausslich, H.G. (1998). Sequential steps in human immunodeficiency virus particle maturation revealed by alterations of individual Gag polyprotein cleavage sites. J. Virol. 72, 2846–2854.

World Health Organization (WHO) (2013). The use of antiretroviral drugs for treating and preventing HIV infection (Geneva, Switzerland: WHO Press).

Zolopa, A.R., Shafer, R.W., Warford, A., Montoya, J.G., Hsu, P., Katzenstein, D., Merigan, T.C., and Efron, B. (1999). HIV-1 genotypic resistance patterns predict response to saquinavir-ritonavir therapy in patients in whom previous protease inhibitor therapy had failed. Ann. Intern. Med. 131, 813–821.

Previn Naicker received his BSc and BMedSci (Hons) from the University of KwaZulu-Natal, Durban, South Africa. He is currently a PhD candidate under the supervision of Professor Yasien Sayed. His research interests include protein dynamics and biochemistry.

Yasien Sayed obtained his BSc (Hons) and PhD (2001) in Biochemistry from the University of the Witwatersrand, Johannesburg, South Africa. He is an associate professor and research leader of the HIV/AIDS Proteins Research Group located in the Protein Structure-Function Research Unit, School of Molecular and Cell Biology. His research interests are in the areas of HIV protein structure, stability, enzyme kinetics, biological thermodynamics and x-ray crystallography.

CHAPTER 2

General introduction

2.1 Overview

Human immunodeficiency virus (HIV) continues to be one of the most problematic pathogens. HIV infection progresses to Acquired Immunodeficiency Syndrome (AIDS) resulting in severe suppression of the immune system in infected individuals. The latest statistics on world-wide HIV infection shows approximately 35 million people were living with HIV in 2012[1]. Therefore, HIV remains a serious pandemic that occurs predominantly in sub-Saharan Africa (approximately 25 million living with HIV); with 1.6 million new infections occurring in 2012[1]. Producing effective drugs against HIV has been an uphill battle since it was first isolated in 1983 by Barre-Sinoussi and co-workers[2]. This difficulty is due mainly to the viral reverse transcriptase (RT), which is prone to introducing mutations into the genes of the virus[3]. RT is responsible for transcribing the viral RNA to DNA that integrates into the host cell's genome[3]. Owing to the mutation-prone nature of RT and the high replication rate of HIV, alterations and mutations are incorporated rapidly in viral proteins which antiviral drugs target; thereby, providing a challenge for the production of effective antivirals[3]. One such protein is the HIV protease (PR) which cleaves immature viral polyproteins into mature and functional proteins for the production of infective virions[4]. HIV PR targets a total of 12 cleavage sites in the Gag and Gag-pol polyproteins and the HIV nef protein[5]. Thus, PR has been identified as a major target in HIV therapy[4; 6]. The life cycle of HIV is depicted in Figure 1.
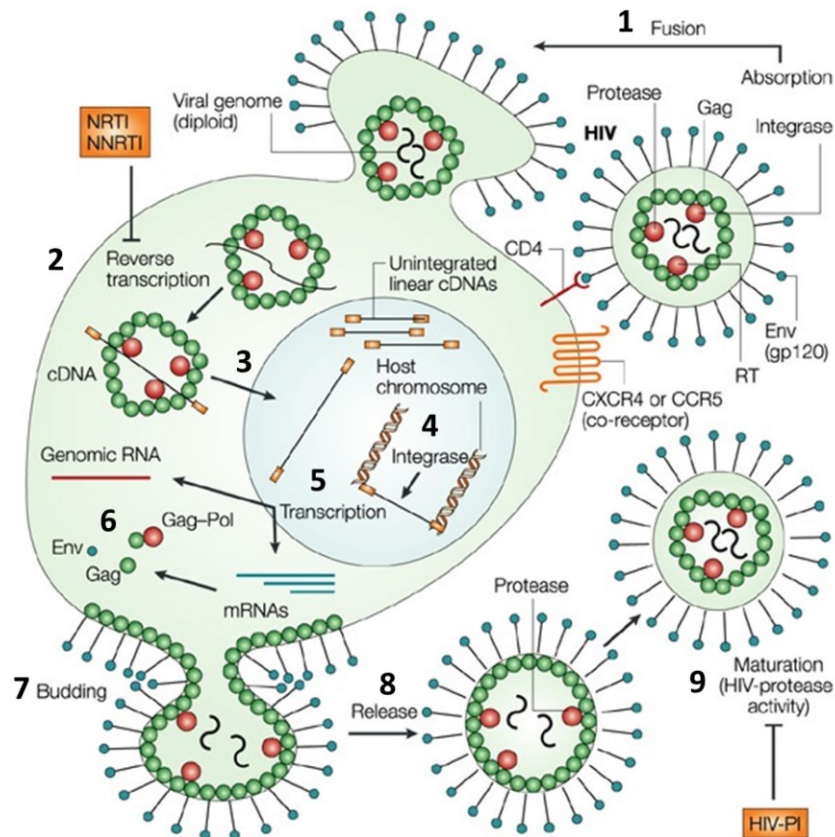
Figure 1: HIV life cycle. Stages of the life cycle are numbered. 1: Binding of HIV to receptors on the surface of the CD4$^+$ T-lymphocyte and fusion with the host cell. 2: Reverse transcription of single-stranded viral RNA to double-stranded viral DNA. 3: Entry of viral DNA into the nucleus of the host cell. 4: Integration of the viral DNA into the host cell's DNA. 5: Transcription of viral mRNA from the HIV genomic material. 6: Assembly of newly translated viral enzymes, structural proteins, and RNA. 7: Budding of the assembled virion from the host cell. 8: Release of the HIV virion into the surrounding environment. 9: Maturation of HIV into its infective form. The HIV protease cleaves viral proteins into their functional forms. Orange boxes depict stages of the life cycle that are major targets for drug therapy. Nucleoside reverse transcriptase inhibitors (NRTIs) and non-nucleoside reverse transcriptase inhibitors (NNRTIs) target the RT enzyme. Protease inhibitors (HIV-PIs) target the protease enzyme. Fusion and entry inhibitors and integrase inhibitors are newer targets for drug therapy. Image adapted from[7].

2.2 HIV genetic diversity

There are the two types of HIV that have been identified; namely, HIV-1 and HIV-2[8; 9]. The predominant type (HIV-1) is separated into groups M, N, O and P[9; 10]. Group M (comprising 95% of the complete genome sequences of HIV-1) is further separated into nine subtypes being A, B, C, D, F, G, H, J and K[8] and a growing number of circulating recombinant forms (CRFs)[9] as seen in Figure 2. Subtype C which is the most common subtype world-wide accounts for roughly 50% of HIV-1 infections and occurs mainly in sub-Saharan Africa, India, Brazil and China, which are the regions of the world that are most heavily affected by HIV and AIDS[8; 11].

A vast number of studies have been done on the interaction of protease inhibitors (PIs) with PRs of subtype B HIV origin[12-15], which occurs mainly in regions of North America and Europe. However, fewer studies have been done on the inhibition of non-B subtypes, in particular, the subtype C PR[16]. Variation in PI sensitivity between HIV-1 subtypes may be significant and contribute to the high prevalence of HIV in sub-Saharan Africa. Genetic variation between subtypes usually ranges between 25-35% at the nucleotide level and variation within subtypes can range from 15-20%[9; 17]. Thus, both the high mutation rate of HIV and the genetic diversity of HIV make it seemingly improbable to create a single PI that will maintain potency against all the subtypes of HIV proteases found within different individuals[18].
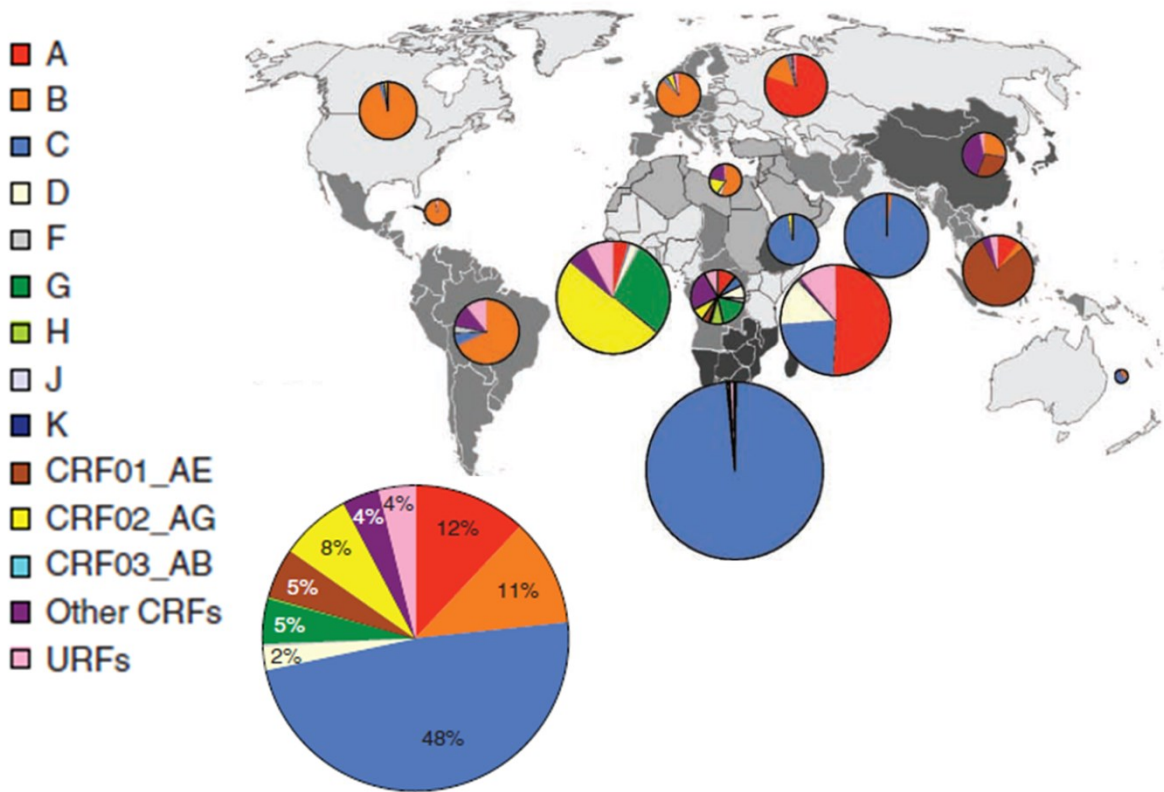
Figure 2: Global HIV-1 subtype distribution. Nine subtypes of the major group are displayed. As of 16 June 2014, there have been reports of 61 CRFs in HIV-1 group M (Los Alamos HIV sequence database, http://www.hiv.lanl.gov/content/sequences/HIV/CRFs/CRFs.html). URFs refer to unique recombinant forms. Image adapted from[11].

.

2.3 HIV protease and mutations

HIV PR is a 22 kDa homodimer, with each monomer containing 99 amino acid residues[19; 20]. Figure 3A, is an illustration of the HIV PR. It belongs to the class of aspartic proteases which contain the signature active site amino acid triplet (aspartic acid-threonine-glycine)[21]. Amino acid residues comprising the active site of the PR are conserved because mutations in this region will be detrimental to its ability to cleave natural substrates[21]. However, mutations near/at the active site (primary resistance mutations) and distal (secondary resistance mutations) from the active site can impact negatively on the PIs ability to bind the PR[22; 23]. Secondary resistance mutations either contribute to drug resistance in the presence of selected primary resistance mutations or they may compensate for the reduction in catalytic efficiency caused by the selected primary PR resistance mutations[24; 25]. A combination of secondary resistance mutations are required to affect the susceptibility of the PR to PIs[12].

2.4 Protease inhibitors

PIs are designed to target the active site of the PR; thereby, inhibiting its ability to bind and cleave natural substrates. The first-line antiretroviral therapy (ART) guidelines for adults and adolescents include the use of two nucleoside reverse transcriptase inhibitors in combination with one non-nucleoside reverse transcriptase inhibitor[26]. Second-line ART may proceed upon failure of the first-line ART regimen in an individual. Currently, there are 10 FDA approved PIs; namely, saquinavir, indinavir, amprenavir, ritonavir, nelfinavir, lopinavir, atazanavir, fosamprenavir, tipranavir and darunavir. Protease inhibitors (PIs) preferred in second-line ART for adults and adolescents are either atazanavir, darunavir or lopinavir in combination with the PI booster, ritonavir[26].
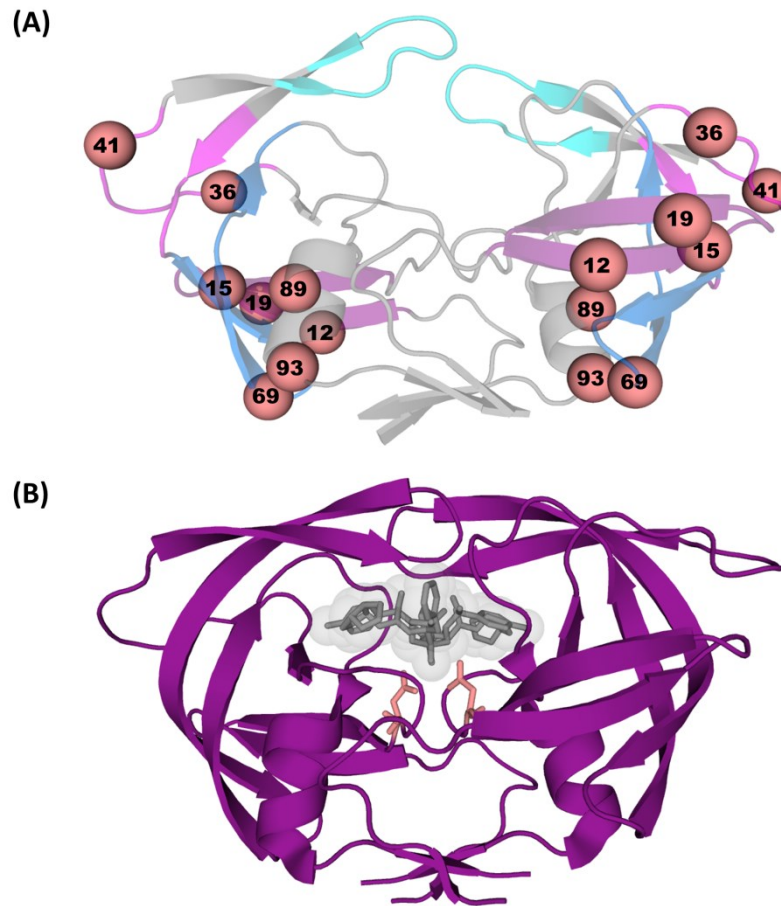
Figure 3: Polymorphic sites in subtype B and C-SA PRs. (A) The flexible flaps of PR (residues at positions 46–54) are coloured cyan. The fulcrum coloured purple (residues at positions 10–23), hinge coloured pink (residues at positions 35–42 and 57–61) and cantilever coloured blue (residues at positions 62–78) are regions implicated in flap opening. Locations of the eight polymorphic amino acids in each monomer are shown as spheres. PDB ID: 3U71[27]. Spheres indicate the position of polymorphisms in the structure in comparison to the consensus subtype B PR and numbers correspond to the amino acid position in the primary sequence. **(B)** Representation of darunavir (grey sticks surrounded by spheres) bound to the HIV-1 subtype B PR. PDB ID: 4LL3[28]. The spheres, showing the area filled by darunavir upon binding, encompass the active site region of the PR. The catalytic aspartic acid (position 25) of each monomer of the PR is represented by sticks. Figures generated using PyMOL (Schrödinger LLC., Portland, USA; http://www.schrodinger.com).

All of the clinically used PIs are competitive active site inhibitors and, except for tipranavir, all are peptidomimetics[29]. Peptidomimetics are non-cleavable chemical structures designed to convey characteristic information contained in peptides into small non-peptide structures, thereby mimicking peptides[30]. Figure 3B is an illustration of the HIV-1 subtype B PR bound to the peptidomimetic PI, darunavir[28].

The first generation PI, ritonavir, is no longer used as an inhibitor alone, rather as a boosting agent in highly active antiretroviral therapy (HAART) with other PIs, due to its ability to inhibit the CYP-450 3A4 isoform[31]. Atazanavir is a second generation PI which is the bulkiest of the currently available FDA-approved PIs[29]. The second generation inhibitor, darunavir, is the latest FDA-approved PI and has been described as the most potent PI currently available. Darunavir is also described as being highly active against multi-drug resistant PRs which is attributed to its ability to fit within the "substrate envelope"[32; 33] of the active site during binding[34]. Darunavir differs only by the addition of a *bis*-tetrahydrofuranyl (*bis*-THF) moiety to the structure of amprenavir, which form vital hydrogen bonding interactions with the main chain of Asp29 and Asp30 in the PR[29; 34], allowing darunavir to mimic conserved hydrogen bonds made by Gag and Gag-Pol substrates[35].

2.5 Catalytic mechanism of HIV protease

The mature HIV PR is released from the Gag-Pol polyprotein precursor via a two-step mechanism[36-40]. The first step appears to occur via intramolecular cleavage of the p6^{Pol}–PR junction (Chapter 1, Figure 2)[39]. This is the faster step of PR maturation. The flanking C-terminal sequence does not appear to hinder catalytic activity. The activity of PR attached to flanking RT residues is comparable to mature PR[39; 40]. The slower cleavage of the PR-RT junction is achieved by an intermolecular catalytic mechanism[39; 40]. Therefore, processing of other cleavage sites in Gag, Gag-Pol and Nef may be performed by either PR attached to flanking RT residues or mature PR. Previous ^{18}O-exchange mass spectrometry studies involving HIV PR and its substrate analogues show that peptide hydrolysis proceeds through the formation of a reversible and metastable gem-diol intermediate[41]. Near-atomic resolution crystal structures of HIV PR have captured important components of the reaction mechanism[42]. A substrate is initially recognised and binds to the PR enzyme (ES). Thereafter, it is converted to a gem-diol intermediate (ES*) through nucleophilic attack by an activated water molecule (Figure 4). Covalent linkages in the substrate are broken resulting in the formation of two products (P$_1$ and P$_2$). The products are then released sequentially and the PR may hydrolyse another substrate. The two catalytic residues (D25 and D25′) perform a vital general acid-base role to activate the nearby water molecule which acts as a nucleophile and attacks the carbonyl carbon of the scissile bound. Previous studies on the pH dependence of the hydrolysis reaction catalysed by HIV PR reveal that substrates and inhibitors only bind to the PR when one of the two catalytic aspartic acid side-chains are protonated[43-45]. The pKa values of the catalytic aspartic acid side-chains are highly dependent on their surrounding environment. These side-chains have different pKa values of 3.1 and 5.2[43] and in the presence of a pepstatin inhibitor, one aspartic acid is protonated in the pH range of 2.5–6.5[44].
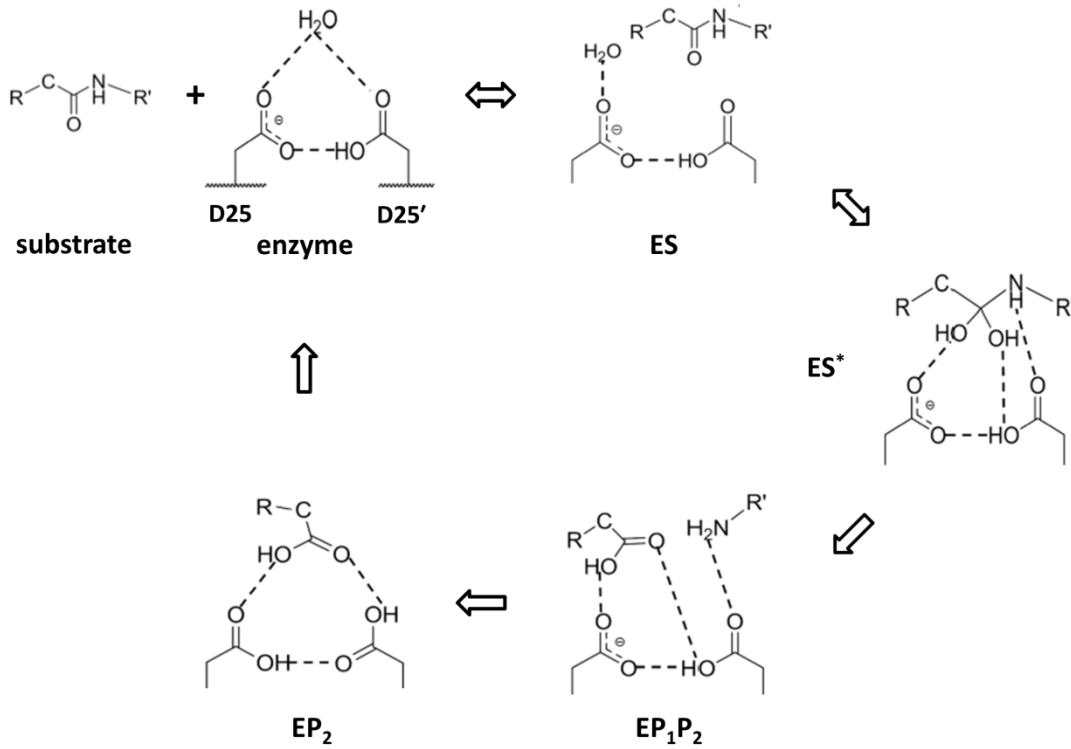
Figure 4: HIV protease reaction mechanism. Generalised reaction mechanism for all HIV PR substrates is displayed. A water molecule is located at the catalytic dyad (D25 and D25′) in the unbound PR and is activated by the unprotonated aspartate group. Image adapted from[42].

2.6 South African HIV-1 subtype C protease (C-SA PR)

Currently, over 100 000 macromolecular structures have been deposited in the Protein Data Bank (PDB). Roughly 600 of these are of HIV PRs. Thus far, the majority of PR crystal structures in the PDB are representative of HIV-1 subtype B and only three crystal structures of HIV-1 subtype C PRs had been solved prior to the current study[46; 47]. The previous structures of subtype C PRs represent those with a sequence matching a patient from India (comprising the N37A and K41R polymorphisms in comparison to the C-SA PR), including the 1.2 Å high-resolution structure of the unbound subtype C PR which provided important basic information about the subtype C PR and a basis for expansion in the current study[47]. Interestingly, genetic diversity occurs within subtypes of HIV owing to the ever-mutating nature of the virus as well as mutations resulting from selective pressure from antiviral drugs[9; 17]. The consensus sequence of the HIV subtype C PR occurring in South Africa (C-SA PR) shares the same sequence identity as the consensus subtype C PR in group M (Los Alamos HIV sequence database, http://www.hiv.lanl.gov/). The C-SA PR differs at eight amino acid residues in each monomer (Figure 3A) to the consensus subtype B PR (Los Alamos HIV sequence database, http://www.hiv.lanl.gov/). The consensus C-SA PR (wild-type) sequence data were obtained from Prof Lynn Morris (AIDS Virus Research Unit, National Institute of Communicable Diseases, South Africa). The M36I, L89M and I93L polymorphisms inherent to the C-SA PR are associated with drug resistance to multiple PIs in subtype B HIV-1. Therefore, these polymorphisms are referred to as secondary resistance mutations. Amprenavir, indinavir, nelfinavir and saquinavir displayed 3–6 fold weaker binding, whereas, lopinavir and ritonavir displayed 8–24 fold weaker binding to the C-SA PR in comparison to the subtype B PR[16]. Importantly, naturally occurring polymorphisms in

strains such as subtype C HIV-1 have an additive effect to primary resistance mutations and amplification of drug resistance mutations[48; 49].

2.7 Protease dynamics and flap conformers

Crystal structures of HIV-1 PRs when combined with activity data provide a good idea of structure-function relations. However, proteins are naturally highly dynamic and crystal structures only provide a static representation of proteins. The vast majority of studies measuring PR dynamics are investigated using molecular dynamics (MD) simulations. MD simulation results require careful attention for correct interpretation and are useful when complemented with experimental data. Other analyses of PR dynamics have been investigated using nuclear magnetic resonance (NMR)[50-52], which is limited by the modest sensitivity of the technique, and spin-labelled pulsed electron paramagnetic resonance (EPR) spectroscopy[53; 54], which was used to exclusively measure distances between flap residues.

HIV PR may exhibit a number of flap conformations in solution. The majority of PR molecules prefer a semi-open conformation in solution (Figure 5B)[54]. At baseline, an ensemble of semi-open conformers may exist. The semi-open conformers may directly interchange with a closed conformer (Figure 5A). Structural and dynamics studies of HIV-1 PR indicate that a curled flap conformation (top view of Figure 5C) is an intermediate state which allows for the formation of the fully-open conformer (Figure 5C)[50; 55; 56]. No x-ray diffracted structural data is available for both the curled and fully-open conformer. The curled conformer is described by a pronounced curling of the flap tips[56]. The fully-open conformer also displays inward curling of the flap tips (residue positions 48–52) accompanied by upward and outward motion of the flap-hinge region from the active site[57]. Entry of a

substrate or inhibitor to the active site of the PR requires substantial movement of the flaps (~ 15 Å from their position in the closed conformer)[58]. Therefore, it is accepted that only the fully-open conformation allows for substrate/inhibitor binding.

Complete flap opening occurs through concerted downward movement of the hinge (residue positions 35–42 and 57–61 of each monomer), cantilever (residue positions 62–78) and fulcrum (residue positions 10–23) regions (Figure 3)) which results in the upward and outward motion of the flaps (Figure 5C)[57]. Distance measurements using spin-labelled pulsed EPR spectroscopy, displayed an increased proportion of the curled and fully-open conformers (Figure 5C) in the subtype C PR population relative to other HIV-1 PRs[54]. The majority of C-SA PR molecules display semi-open flaps; though, there is a definite shift in the equilibrium between semi-open, curled and fully-open conformers. Combined EPR spectroscopy and NMR experiments have identified and verified the population of flap conformers in the apo-state and during inhibitor binding for the subtype B and C HIV-1 PRs[59]. These experiments did not fully elucidate the mechanism for the reduced drug susceptibility displayed by the subtype C HIV-1 PR; however, differences in the population of flap conformers may be vital for the interpretation of other dynamic comparisons of the subtype B and C-SA PRs.
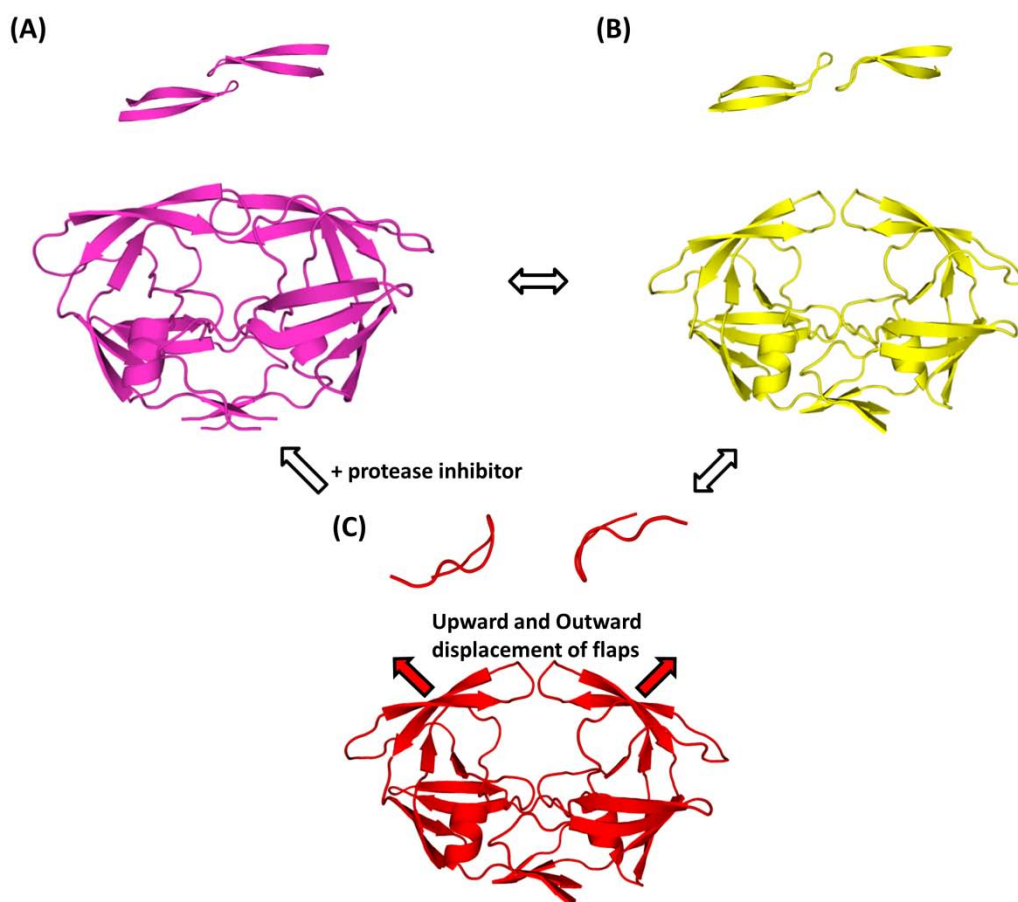
Figure 5: Overview of flap conformers displayed by HIV-1 PR. The closed **(A)**, semi-open **(B)** fully-open **(C)** conformers are in dynamic equilibrium. Above the respective three-dimensional structures are the top views of the flaps. The semi-open flap conformation is the most prevalent conformer in the absence of inhibitor. The presence of protease inhibitors promote the closed form of the PR. Crystal structure data is unavailable for a fully-open conformer which is characterised by upward and outward displacement of the flap-hinge region. Representation of the flap positioning of the fully-open conformer is based on MD simulation models[57]. The fully-open conformer may also be referred to as a wide-open conformer; however, fully-open is preferred in the current study due to previous misuse of the term "wide-open"[60]. PDB ID: 1HXW (A)[61], 1HHP (B and C)[62] and 3UHL (C)[63].

Hydrogen/deuterium exchange-mass spectrometry (HDX-MS) experiments may provide a straight forward comparison of protein dynamics between the C-SA PR and subtype B PR. Measuring amide hydrogen/deuterium exchange (HDX) via mass spectrometry over several time points allows for the determination of HDX rates in detected peptides. HDX-MS has been employed to elucidate the solvent protection patterns and protein dynamics of reverse transcriptase monomers[64] and the mature and immature HIV-1 capsid proteins[65-67]. There has been no report on the investigation of HIV PR dynamics using HDX-MS. Thus, the current study provides novel results and insights into the dynamics at different regions of the HIV-1 PR and the mechanism of reduced drug susceptibility displayed by the C-SA PR.

2.8 Aim and objectives

The HIV-1 subtype C PR may be described as having increased virulence in comparison to other HIV-1 subtypes. This implies that HIV-1 subtype C strains may display improved viral fitness. The recent increased prevalence of this subtype in the Americas and Europe, where it had previously been detected at low levels in these populations, supports this statement[68-72]. The aim of the current research was to identify functional differences between the subtype B and C-SA PRs and explain any differences through structural and dynamic characterisation.

The specific objectives of this research were to:

1. Express and purify the subtype B and C-SA PRs at high yields to allow for energetic, structural and dynamic evaluation.

2. Investigate the steady-state kinetics for substrate processing for both PRs using fluorescence-based activity assays.

3. Determine inhibition constants of atazanavir, darunavir and ritonavir for both PRs by performing fluorescence-based inhibition assays.

4. Evaluate the energetics and mode of binding of the aforementioned inhibitors to both PRs using isothermal titration calorimetry. Binding energetics of atazanavir and darunavir for the subtype B PR have been determined previously under the conditions used in the current research[13; 14].

5. Resolve the three-dimensional crystal structure of the apo-C-SA PR using x-ray crystallography. The atomic structure of the consensus subtype B PR has been solved previously[55].

6. Identify differences in flap dynamics between the subtype B and C-SA PRs by performing molecular dynamics simulations.

7. Investigate the dynamics of HIV-1 PR at all regions and determine differences in protein dynamics between the subtype B and C-SA PRs via hydrogen/deuterium exchange-mass spectrometry experiments.

CHAPTER 3

Structural Insights into the South African HIV-1 Subtype C Protease: Impact of hinge region dynamics and flap flexibility in drug resistance

Previn Naicker, Ikechukwu Achilonu, Sylvia Fanucchi, Manuel Fernandes, Mahmoud A.A. Ibrahim, Heini W. Dirr, Mahmoud E.S. Soliman and Yasien Sayed.

In this publication, the crystal structure of the apo-C-SA PR was determined and compared to that of the apo-subtype B PR. Altered dynamics at the flap tips of the C-SA PR was identified.

Author contributions: Previn Naicker analysed the data, wrote the manuscript and performed all experimental work excluding the molecular dynamics simulation,. Ikechukwu Achilonu assisted with crystal growth and revision of the manuscript. Sylvia Fanucchi assisted with crystal structure refinement. Manuel Fernandes assisted with loading crystals on the x-ray diffractometer and initial data processing. Mahmoud A.A. Ibrahim and Mahmoud E.S. Soliman performed molecular dynamics simulation. Heini W. Dirr assisted in manuscript revision. Yasien Sayed supervised the project and assisted in data analysis and interpretation.

Structural insights into the South African HIV-1 subtype C protease: impact of hinge region dynamics and flap flexibility in drug resistance

Previn Naicker[a], Ikechukwu Achilonu[a], Sylvia Fanucchi[a], Manuel Fernandes[b], Mahmoud A.A. Ibrahim[cd], Heini W. Dirr[a], Mahmoud E.S. Soliman[ef] & Yasien Sayed[a]

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

[b] School of Chemistry, University of Witwatersrand, Johannesburg, 2050, South Africa.

[c] School of Chemistry, University of Manchester, Oxford Road, Manchester, M13 9PL, UK.

[d] Faculty of Science, Department of Chemistry, Minia University, Minia, 61519, Egypt.

[e] School of Health Sciences, University of KwaZulu-Natal, Durban, 4001, South Africa.

[f] Faculty of Pharmacy, Department of Pharmaceutical Organic Chemistry, Zagazig University, Zagazig, 44519, Egypt.
Published online: 12 Nov 2012.

PLEASE SCROLL DOWN FOR ARTICLE

31

Taylor & Francis
Taylor & Francis Group

Structural insights into the South African HIV-1 subtype C protease: impact of hinge region dynamics and flap flexibility in drug resistance

Previn Naicker[a], Ikechukwu Achilonu[a], Sylvia Fanucchi[a], Manuel Fernandes[b], Mahmoud A.A. Ibrahim[c,d], Heini W. Dirr[a], Mahmoud E.S. Soliman[e,f] and Yasien Sayed[a]*

[a]*Protein Structure-Function Research Unit, School of Molecular and Cell Biology, University of Witwatersrand, Johannesburg 2050, South Africa;* [b]*School of Chemistry, University of Witwatersrand, Johannesburg 2050, South Africa;* [c]*School of Chemistry, University of Manchester, Manchester M13 9PL, UK;* [d]*Faculty of Science, Department of Chemistry, Minia University, Minia 61519, Egypt;* [e]*School of Health Sciences, University of KwaZulu-Natal, Durban 4001, South Africa;* [f]*Faculty of Pharmacy, Department of Pharmaceutical Organic Chemistry, Zagazig University, Zagazig 44519, Egypt*

Communicated by Ramaswamy H. Sarma

The HIV protease plays a major role in the life cycle of the virus and has long been a target in antiviral therapy. Resistance of HIV protease to protease inhibitors (PIs) is problematic for the effective treatment of HIV infection. The South African HIV-1 subtype C protease (C-SA PR), which contains eight polymorphisms relative to the consensus HIV-1 subtype B protease, was expressed in *Escherichia coli*, purified, and crystallized. The crystal structure of the C-SA PR was resolved at 2.7 Å, which is the first crystal structure of a HIV-1 subtype C protease that predominates in Africa. Structural analyses of the C-SA PR in comparison to HIV-1 subtype B proteases indicated that polymorphisms at position 36 of the homodimeric HIV-1 protease may impact on the stability of the hinge region of the protease, and hence the dynamics of the flap region. Molecular dynamics simulations showed that the flap region of the C-SA PR displays a wider range of movements over time as compared to the subtype B proteases. Reduced stability in the hinge region resulting from the absent E35-R57 salt bridge in the C-SA PR, most likely contributes to the increased flexibility of the flaps which may be associated with reduced susceptibility to PIs.

An animated interactive 3D complement (I3DC) is available in Proteopedia at http://proteopedia.org/w/Journal:JBSD:36

Introduction

Human immunodeficiency virus (HIV) continues to be one of the most problematic pathogens known to man. HIV infection progresses to acquired immunodeficiency syndrome (AIDS) resulting in extreme suppression of the immune system in infected individuals. The production of effective drugs against HIV has been an uphill battle which is attributed to its reverse transcriptase (RT), which is prone to introducing mutations into the viral genome (Bebenek, Abbotts, Roberts, Wilson, & Kunkel, 1989). Due to the mutation-prone nature of RT and the high replication rate of HIV, alterations and mutations occur often in viral proteins for which antiviral drugs have been designed to target; thereby, providing a challenge for the production of effective antivirals (Bebenek et al., 1989). One such protein is the HIV protease (PR), which cleaves

immature viral polyproteins into mature and functional proteins, for the production of infective virions (Tomasselli & Heinrikson, 2000). HIV PR targets 12 cleavage sites in Gag and Gag-Pol polyproteins (De Oliveira et al., 2003). Thus, the PR has been identified as a major target in HIV therapy (Tomasselli & Heinrikson, 2000).

Two types of HIV have been identified: HIV-1 and HIV-2 (McCutchan, 2006; Taylor, Sobieszczyk, McCutchan, & Hammer, 2008). The predominant type (HIV-1) is separated into groups M, N, O, and P (Plantier et al., 2009; Taylor et al., 2008). Group M (Major group) is further separated into nine subtypes (A, B, C, D, F, G, H, J, and K) and a growing number of circulating recombinant forms (McCutchan, 2006). HIV-1 subtype C is the most common subtype worldwide, occurring mainly in sub-Saharan Africa, India, Brazil,

*Corresponding author. Email: yasien.sayed@wits.ac.za

and China (McCutchan, 2006), and is responsible for 50% of global HIV-1 group M infections (Buonaguro, Tornesello, & Buonaguro, 2007). Genetic variation between subtypes usually ranges between 25 and 35% at the nucleotide level and variation within subtypes can range from 15 to 20% (Hemelaar, Gouws, Ghys, & Osmanov, 2006; McCutchan, 2006). Thus, both the high mutation rate within the HIV PR and the genetic diversity of HIV increase the challenge of producing a protease inhibitor (PI) that is capable of inhibiting the PRs of various HIV subtypes found within different individuals (Kantor & Katzenstein, 2003).

The homodimeric HIV-1 PR, consisting of 99 amino acids per monomer (Figure 1(A)), belongs to a class of aspartyl PRs which contain the signature active site amino acid triplet (aspartic acid-threonine-glycine) (Wlodawer & Vondrasek, 1998). The hinge region of the PR is closely associated with the stability and movement of the flap region (position 46–54); Figure 1(B). We define the hinge region as comprising residues from positions 35–42 and 57–61. The PR flaps undergo substantial movement allowing for substrate/inhibitor entry (open conformation) and form key interactions during the binding of substrate/inhibitor (closed conformation) (Gustchina & Weber, 1990). Therefore, the flaps are required to display flexibility; however, increased flexibility may result in reduced substrate processing and PI resistance (Meiselbach, Horn, Harrer, & Sticht, 2007). Increased flexibility of the flap-hinge region is likely to result in increased flexibility of the flaps, as seen in PRs

(B)

```
C-SA 3U71      PQITLWKRPLVSIKVGGQIKEALLDTGADDTVLEEINLPGKWKPKMIGGIGGFI 54
Subtype B      PQITLWKRPLVTIKIGGQLKEALLDTGADDTVLEEMNLPGRWKPKMIGGIGGFI 54
Subtype B-MDR  PQITLWQRPIVTIKIGGQLKEALLNTGADDTVLEEVNLPGRWKPKLIGGIGGFV 54
Subtype C-2R8N PQITLWKRPLVSIKVGGQIKEALLDTGADDTVIEEIALPGRWKPKMIGGIGGFI 54
               ****** **:.*:.**:***:***** :****:**. ***:****:*******:
```

```
C-SA 3U71      KVRQYDQILIEICGKKAIGTVLVGPTPVNIIGRNMLTQLGCTLNF 99
Subtype B      KVRQYDQILIEICGHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNF 99
Subtype B-MDR  KVRQYDQVPIEICGHKVIGTVLVGPTPANVIGRNLMTQIGCTLNF 99
Subtype C-2R8N KVRQYDQIIIEICGKKAIGTVLVGPTPVNIIGRNMLTQLGCTLNF 99
               ****** : *****:*.*********.*:****::**:*:******
```

Figure 1. (A) Structure of unbound HIV-1 PR with the active site triplet shown as sticks, hinge region in magenta, and flaps in blue (PDB ID: 2PC0). (B) Multiple sequence alignment of the C-SA PR, consensus subtype B PR, subtype B-MDR PR, and subtype C-2R8N PR. Active site triad shown in bold, hinge region in magenta, and flaps in blue. Both the C-SA PR and the consensus subtype B PR under investigation exhibit the Q7K mutation which reduces protease autocatalysis (Mildner et al., 1994).

that are resistant to the PI, amprenavir (Meiselbach et al., 2007).

The HIV-1 subtype C PR that predominates in South Africa (C-SA PR) is defined as the consensus subtype C PR (Los Alamos HIV sequence database, http://www.hiv.lanl.gov/). In this study, the C-SA PR was recombinantly expressed in *E. coli*, purified from inclusion bodies, crystallized, and its detailed three-dimensional structure solved. Structural analyses of the C-SA PR, as well as the consensus subtype B PR (Los Alamos HIV sequence database, http://www.hiv.lanl.gov/), and a multidrug resistant subtype B PR (subtype B-MDR PR) were performed. Molecular dynamics (MD) simulations were performed to determine the differences in flap movement and residue-specific fluctuation amongst the PRs. This study highlights the variability found in the flap-hinge region amongst subtype B PRs and the C-SA PR, as well as its impact on flap movement and resistance to clinically available PIs.

Materials and methods

Expression and purification

The C-SA PR sequence data were obtained from Prof Lynn Morris (AIDS Virus Research Unit, NICD of Johannesburg, South Africa). The recombinant pET-11b plasmid encoding a subtype B PR gene was a kind gift from Dr. J. Tang (University of Oklahoma Health Sciences Center, Oklahoma City). The mutations (T12S, I15V, L19I, M36I, R41K, H69K, L89M, and I93L) constituting the baseline sequence of the C-SA PR were generated previously in our lab (Mosebi, Morris, Dirr, & Sayed, 2008), by site-directed mutagenesis using the QuikChange method (Stratagene, La Jolla, CA, USA). A Q7K point mutation was introduced to reduce PR autocatalysis (Mildner et al., 1994). DNA sequencing confirmed the coding region of the C-SA PR.

Escherichia coli BL21 (DE3) pLysS cells transformed with the plasmid encoding the C-SA PR gene were induced to express the C-SA PR as inclusion bodies by the addition of isopropyl-β-D-thiogalactopyranoside (Ido, Han, Kezdy, & Tang, 1991). Cells were disrupted by sonication and centrifuged after overexpression of the PR. PR was recovered from the inclusion bodies using 8 M urea, 10 mM Tris-HCl, and 2 mM dithiothreitol (DTT) (pH 8.0) buffer. The positively charged PR was allowed to flow through a diethylaminoethyl ion-exchange matrix at pH 8.0. The PR was dialyzed against formic acid and was refolded by dialysis against 10 mM sodium acetate, 2 mM NaCl, 1 mM DTT, and 5% glycerol (pH 5.0) buffer. The PR was further purified using CM-Sepharose ion-exchange column chromatography, with a 0–1 M NaCl gradient elution. The PR was finally dialyzed against 10 mM sodium acetate, 2 mM NaCl, and
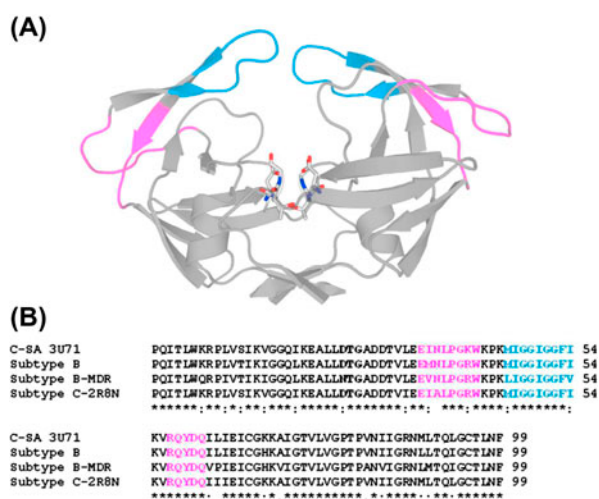
1 mM DTT (pH 5.0) buffer. Purity of the PR was evaluated by tricine-SDS-PAGE (Laemmli, 1970; Schägger, 2006). The concentration of the PR was determined using a molar extinction coefficient of $25,480\,M^{-1}\,cm^{-1}$ from absorbance spectra obtained on a Jasco V-630 spectrophotometer.

Enzyme activity determination

The specific activity of the C-SA PR was determined following hydrolysis of the HIV-1 PR fluorogenic substrate (Abz-Arg-Val-Nle-Phe(NO_2)-Glu-Ala-Nle-NH_2) which mimics the capsid/P_2 cleavage site in the HIV-1 Gag polyprotein. An increase in fluorescence emission from the 2-aminobenzoyl group was detected at an excitation wavelength of 337 nm and emission wavelength of 425 nm, resulting from cleavage of the Nle-Phe(NO_2) peptide bond (Carmel & Yaron, 1978; Szeltner & Polgar, 1996). A substrate concentration of 50 μM and an enzyme concentration of 10–50 nM, with an excitation bandwidth of 2.5 nm and emission bandwidth of 5 nm were used for the 1 min measurements during steady state. The intensity signal associated with complete cleavage of 1 nmol of substrate was measured and used to convert intensity to activity.

Active site titration

The percentage of active enzyme in the protein preparation was determined using isothermal titration calorimetry (ITC). Briefly, 192.5 μM of acetyl-pepstatin (an aspartyl PR inhibitor) was titrated against 12.75 μM of C-SA PR in 10 μl injections at 293.8 K using a VP-ITC microcalorimeter. The heat due to dilution of acetyl-pepstatin was subtracted from the data-set and any baseline errors were corrected. The changes in heats were integrated and fitted using one set of binding sites (Origin 5.0 software package). The percentage of active sites was determined from the calculated stoichiometry value; a value of 1 indicates a 100% active enzyme preparation.

Crystallization and data collection

C-SA PR crystals were grown at 293 K using the hanging-drop vapor diffusion method in a 24-well microplate. Conditions allowing for optimal crystal growth were screened using the Hampton Research Index HR2-144 (Hampton Research). Using a reservoir buffer comprising 0.2 M sodium citrate tribasic dihydrate in 20% polyethelene glycol 3 350 resulted in the formation of diffraction quality PR crystals. The stock protein concentration used for crystallization was 0.5 mg/ml in 10 mM sodium acetate, 2 mM NaCl, and 1 mM DTT (pH 5.0). Each hanging drop (4 μl or 8 μl) comprised equal volumes of reservoir buffer and stock protein. A paraffin–silicone oil mixture (1:1 ratio) covered the reservoir buffer in each well (Chayen & Stewart, 1992). Single crystals were mounted on a cryoloop, briefly soaked in reservoir solution containing a higher concentration of cryoprotectant (polyethelene glycol), and frozen in liquid nitrogen. X-ray diffraction data were collected in-house on a Bruker X8 Proteum system with a Microstar copper rotating-anode generator with Montel 200 optics, a PLATINUM 135 CCD detector, and an Oxford Cryostream Plus system. During data collection, the crystal was maintained at 113 K in a stream of nitrogen gas. Images were collected covering an oscillation angle of 0.5° per image. The data-set was processed using APEX and SAINT software (Bruker AXS Inc., Madison, WI, USA).

Model building and structure refinement

The phases of the structure were solved by the molecular replacement method using MOLREP (Vagin & Teplyakov, 2000); a component of the *CCP*4i suite of programs (Potterton, Briggs, Turkenburg, & Dodson, 2003). A subtype C isoform with the following variations as compared to the C-SA PR: L33I, N37A, K41R, and L63I was used as a search probe for molecular replacement (PDB ID: 2R8N; Coman, Robbins, Goodenow, McKenna, & Dunn, 2007). Model building was performed using *Coot* (Emsley & Cowtan, 2004); thereafter, cycles of global reciprocal space refinement using *REFMAC*5 (Murshudov, Vagin, & Dodson, 1997) and local real space refinement using *Coot* (Emsley & Cowtan, 2004) were performed. Stereochemical validation of the target model was performed using *MolProbity* (Chen et al., 2010) and *PROCHECK* (Laskowski, 1993). The PyMOL Molecular Graphics System (Schrödinger LLC., Portland, USA; http://www.schrodinger.com) was used to generate images of the structure. Sequence alignment was performed using Clustal X 2.0 (Higgins & Sharp, 1988; Larkin et al., 2007).

MD simulations

The simulations of the three enzyme systems: C-SA PR (PDB ID: 3U71), consensus subtype B PR (PDB ID: 2PC0), and subtype B-MDR PR (PDB ID: 1RPI) were performed under physiological pH conditions. The aspartic acid residue (D25) of the catalytic site exhibits an increased pK_a value of 5.2 in the inhibitor bound PR (Shen, Wang, Kovalevsky, Harrison, & Weber, 2010) while no increased pK_a was reported for the free form of the PR ($pK_a = 4.5$) (Smith, Brereton, Chai, & Kent, 1996). Therefore, an unprotonated active site was used throughout the MD studies, as this is the prevalent form at physiological pH. Hydrogen atoms of the proteins were added using the Leap module in Amber10 (Case et al., 2005). The standard AMBER force field for bioorganic systems (ff03) (Duan et al., 2003) was used to define the HIV-1 PR enzyme parameters. Counter ions were added to neutralize the enzyme. The system was enveloped in a box of equilibrated TIP3P (Jorgensen,

Chandrasekhar, Madura, Impey, & Klein, 1983) water molecules with 8 Å distance around the enzyme.

The MD package Amber10 (Case et al., 2005) was used for the minimization and equilibration protocols. Cubic periodic boundary conditions were imposed and the long-range electrostatic interactions were treated with the particle-mesh Ewald method (Essmann et al., 1995) implemented in Amber10 with a nonbonding cut-off distance of 10 Å. The energy minimization was performed using the steepest descent method in Amber10 for 1000 iterations and switched to conjugate gradient for 2000 steps, with a restraint potential of $500\,kcal/mol\,Å^2$ applied to the solute. The entire system was then freely minimized for 1000 iterations. For the equilibration and subsequent production run, the SHAKE algorithm (Ryckaert, Ciccotti, & Berendsen, 1977) was employed to constrain all bonds involving hydrogen atoms, allowing for an integration time step of 2 fs. Harmonic restraints with force constants $2\,kcal/mol\,Å^2$ were applied to all solute atoms. A canonical ensemble (NVT) MD was carried out for 50 ps, during which the system was gradually annealed from 0 to 300 K using a Langevin thermostat with a coupling coefficient of 1/ps. Harmonic restraints with force constraints $10\,kcal/mol\,Å^2$ were applied to all solute atoms during the heating stage. Subsequently, the system was equilibrated at 300 K with a 2 fs time step for 500 ps while maintaining the force constants on the restrained solute. With no restraints imposed, a production run was performed for 10 ns in an isothermal isobaric (NPT) ensemble using a Berendsen barostat (Berendsen, Postma, van Gunsteren, DiNola, & Haak, 1984) with a target pressure of 1 bar and a pressure coupling constant of 2 ps. The coordinate file was saved every 1 ps and the trajectory was analyzed every 10 ps using the Ptraj module implemented in Amber10.

Results and discussion

C-SA PR preparation

The recombinant expression of the C-SA PR yielded 0.2 mg/l of culture. The purity of the C-SA PR was estimated to be > 99%, which is required for the ITC and crystallography studies. The C-SA PR was shown to be enzymatically active following hydrolysis of the HIV-1 PR fluorogenic substrate (Abz-Arg-Val-Nle-Phe(NO₂)-Glu-Ala-Nle-NH₂) (Figure 2). The specific activity of the C-SA PR derived from the slope of the plot in Figure 2 is 79 μmoles min⁻¹ mg⁻¹. Figure 3 shows a binding thermogram of the C-SA PR titrated with the aspartyl PR inhibitor, acetyl-pepstatin. Acetyl-pepstatin binds weakly to the C-SA PR showing a calculated K_d of 192 nM which is employed in displacement titration reactions for high affinity inhibitors (Velázquez-Campoy, Kiso, & Freire, 2001). This titration allows for the calculation of the
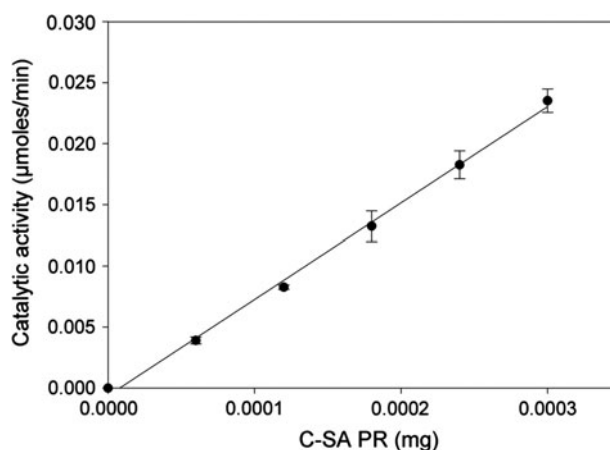
Figure 2. Specific activity of the C-SA PR. Determined following hydrolysis of the HIV-1 protease substrate (Abz-Arg-Val-Nle-Phe(NO₂)-Glu-Ala-Nle-NH₂) in 50 mM sodium acetate and 1 M NaCl (pH 5.0) at 293 K.

percentage of active PR in solution (active site concentration), which was calculated to be 82%, derived from a stoichiometry value of 0.82. These findings indicate that the solved C-SA PR is enzymatically active.

Crystal structure of the C-SA PR

The current study is the first to report on the detailed structure of the C-SA PR. The only other subtype C PR structures to be solved to date are of a PR which predominates in India (subtype C-2R8 N PR) (Coman et al., 2007, 2008). The C-SA PR being the consensus HIV-1 subtype C PR and subtype C being the most common subtype worldwide (McCutchan, 2006), emphasize the importance of this study. Structural homology analysis of the C-SA PR and the subtype C-2R8 N PR revealed no significant RMSD in respect to the hinge and flap regions. The coordinates for the 2.7 Å resolution C-SA PR structure were deposited in the Protein Data Bank (PDB ID: 3U71). Stereochemical validation of the C-SA PR structure revealed no disallowed bond angles or rotamers for all of the constituent residues (Table 1). All water molecules were removed from the structure prior to being deposited in the PDB. The removal of water molecules was implemented because all solvent molecules could not be resolved accurately at 2.7 Å. However, this had no impact on the overall structure of the PR. The space group and unit cell parameters match previously described unbound HIV-1 PR crystals (Heaslet et al., 2007; Yedidi et al., 2011). A Matthews coefficient (V_M) of $2.45\,Å^3\,Da^{-1}$ and a solvent content of 49.75% indicate a PR monomer as the asymmetric unit of each unit cell of the C-SA PR crystal, which is a common finding for crystals of unbound HIV PR (Heaslet et al., 2007; Yedidi et al., 2011).

Figure 3.   Active site titration of the C-SA PR with acetyl-pepstatin. The raw calorimetric data is indicated in the upper panel and the integrated heats for the above peaks plotted against the molar ratio of acetyl-pepstatin to protease dimer in the lower panel. The peak contributing to the heat signal below baseline was corrected for in the fitting procedure. Acetyl-pepstatin was titrated until binding sites on the PR were saturated in a buffer comprising 10 mM sodium acetate (pH 5.0) at 293 K.
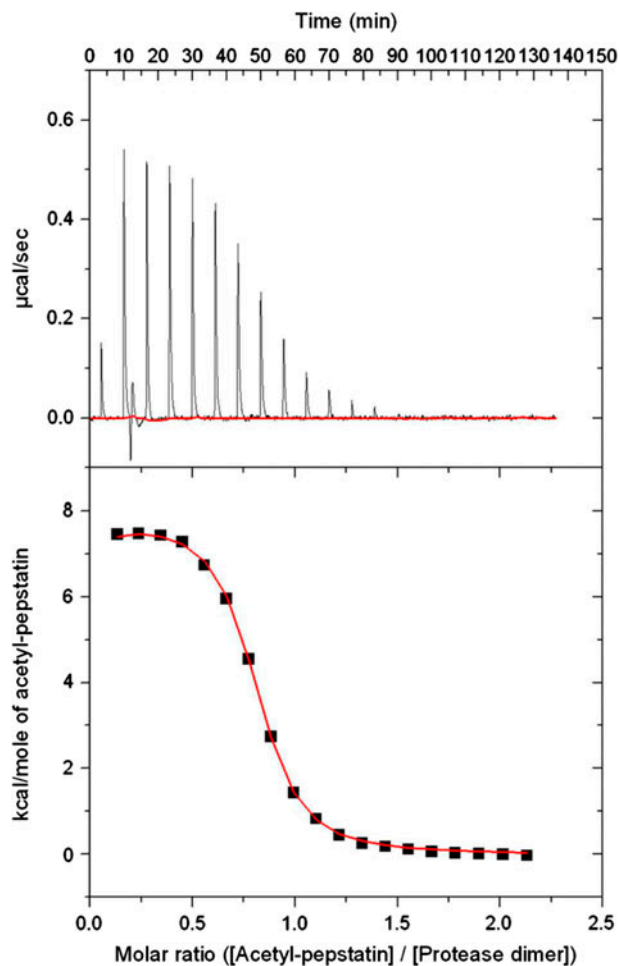
Structural analyses of the C-SA PR

Figure 1(B) shows a sequence alignment of the C-SA PR with the consensus subtype B PR (Heaslet et al., 2007), subtype B-MDR PR (Logsdon et al., 2004), and subtype C-2R8N PR (Coman et al., 2008). The subtype C-2R8N PR contains the following mutations relative to the C-SA PR: L33I, N37A, K41R, and L63I. The C-SA PR displays high structural homology with the subtype C-2R8N PR; however, notable differences occur at positions 37 and 41 which are located in the hinge region of the PR. In subtype B PRs, M36 was shown to form contacts with residues in the 10s loop (Clemente et al., 2004; Martin et al., 2005). These contacts are absent in the subtype C-2R8N due to the M36I and I15V polymorphisms (Coman et al., 2008). These contacts as

Table 1.   Data collection and refinement statistics for the C-SA PR structure.

Wavelength (Å)	1.5418
Space group	P 41 21 2
Unit-cell parameters	
a, b, c (Å)	45.928, 45.928, 99.960
α, β, γ (°)	90.00, 90.00, 90.00
Resolution range (Å)	2.72–41.73 (2.72–2.79)
No. of observed reflections	43,511
No. of unique reflections	3193
Completeness (%)	98.8
$I/\sigma(I)$	5.18 (3.00)
R_{merge}[a]	0.200 (0.413)
Final overall R factor	0.220
R_{work}[b]	0.217
R_{free}[b]	0.283
No. of protein atoms	755
No. of ligand atoms	0
Average B value (Å^2)	25.79
RMSD in bond length (Å)	0.017
RMSD in bond angles (°)	1.776
Ramachandran statistics	
Outliers (%)	0
Favored (%)	96.91
V_M (Å^3 Da^{-1})	2.45
Solvent content (%)	49.75
Asymmetric unit content	monomer
PDB ID	3U71

[a]$R_{merge} = \sum_{hkl} \sum_i |I_i(hkl) - (I(hkl))| / \sum_{hkl} \sum_i |I_i(hkl)|$, where $I(hkl)$ is the intensity of reflection *hkl*, $\sum_{hkl}$ is the sum over all reflections and   is the sum over *i* measurements of reflection *hkl*.
[b]R_{free} is calculated for a randomly chosen 5% of reflections which were not used for refinement of the structure and R_{work} is calculated for the remaining reflections. Data in brackets refer to the highest resolution shell.

expected are also absent in the C-SA PR, and probed further investigation of the contacts occurring in the hinge region of the C-SA PR which are closely related to flap dynamics (Coman et al., 2008). The subtype B-MDR PR exhibits the following 11 mutations relative to the consensus subtype B PR: L10I, D25N, M36V, M46L, I54V, I62V, L63P, A71V, V82A, I84V, and L90M. The subtype B-MDR PR displays high-level drug resistance and cross-resistance to PIs (Palmer, Shafer, & Merigan, 1999; Logsdon et al., 2004). The V82A and I84V active site mutations allow for expansion of the active site which is a mechanism implicated in drug resistance (Logsdon et al., 2004). The implication of the M36V mutation in drug resistance is an important consideration, as the M36I mutation in HIV-1 PRs is thought to be a secondary drug resistance mutation (Patick et al., 1998). The C-SA PR exhibits the following eight mutations relative to the consensus subtype B: T12S, I15V, L19I, M36I, R41K, H69K, L89M, and I93L, which occur distal to the active site. The M36I mutation being considered to be a secondary drug resistance mutation (Patick et al., 1998), which occurs in
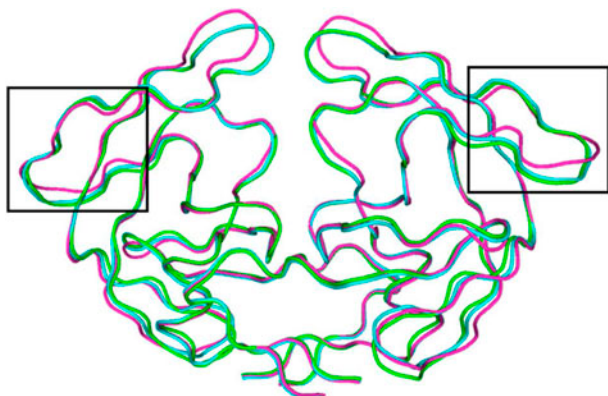
Figure 4. Structural alignment of ribbon representations of the C-SA PR (green, PDB ID: 3U71), consensus subtype B PR (cyan, PDB ID: 2PC0), and subtype B-MDR PR (magenta, PDB ID: 1RP1). The hinge region of each monomer is indicated in the frames.

the hinge region of the PR, probed further investigation into the interactions occurring in this region. The overall tertiary folds of the PRs under investigation do not differ significantly (Figure 4). The subtype B-MDR PR (magenta) is shown to crystallize with its flaps in a wider open conformation. More detailed analyses of the structures are required in order to determine the effects of polymorphisms between the subtypes on the dynamics of these enzymes.

E35 in HIV PR maintains long-range interactions within the PR polypeptide chain (Swairjo, Towler, Debouck, & Abdel-Meguid, 1998). The absence of the E35-R57 salt bridge (ion pair) in both monomers of the C-SA PR is indicated in Figure 5(A). R57 in the C-SA PR adopts a different rotamer to that of the R57 in the subtype B PRs, resulting in an inability to form a salt bridge with E35. The backbone of R57 only forms backbone hydrogen bonds with V77 that form part of a β-sheet located in a region interior to the flaps which are maintained in all the structures under investigation. I36 (green) displays no contacts with other residues in the hinge region. The absence of the E35-R57 (side chain-side chain) salt bridge in PRs results in an apparent outward movement of the flaps (Swairjo et al., 1998). Figure 5(B) shows that the consensus subtype B PR exhibits the E35-R57 salt bridge in one of its monomers, as well as a hydrogen bond between the side chain of R57 and backbone of M36 (cyan). The other monomer shows the side chain of R57 also interacts with the backbone of M36 which may assist in maintaining the compact conformation of the hinge region, as seen by the close proximity of E35 and R57. Figure 5(C) shows the subtype B-MDR PR exhibits the E35-R57 salt bridge in both monomers. The backbone of V36 (magenta) exhibits hydrogen bonds with the side chain of E35 and not R57.

The E35-R57 salt bridge is the only possible ion pair present in the hinge region of HIV-1 PR. Thermodynamic studies using ITC revealed that the CRF01_AE strain of PR containing D35 and I36, which also lacks a salt bridge between the residues at position 35 and 57, displays reduced binding affinity to inhibitors nelfinavir and darunavir when compared to the E35 and M36 subtype B PR (Bandaranayake et al., 2010). Calculated binding free energies reveal reduced binding affinity to amprenavir for a E35D HIV-1 PR mutant (Meiselbach et al., 2007). The mutant showed a stable salt bridge over short periods (0–1 and 8.3–9.3 ns), whereas a stable salt bridge is established over the entire 10 ns simulation in the E35 HIV-1 PR (Meiselbach et al., 2007). These findings suggest that an absence of the E35-R57 salt bridge highlighted in this study may be implicated in drug resistance mechanisms.

The displacement of the flap tips, residues 49–52 forming a hairpin loop, in the absence of a hinge region salt bridge occurs in the plane perpendicular to the plane of the hairpin loop (Swairjo et al., 1998). In some cases, minor displacement (<1.2 Å) results in the backbone atoms of the flap tips being out of hydrogen bonding range (Swairjo et al., 1998). A resultant loss in a hydrogen bond due to flap displacement may also be seen between peptide backbone G51 and carbonyl oxygen of I50′, which plays a role in stabilizing inhibitor binding (Swain et al., 1990). Polymorphic variations at position 36, which are seemingly closely related to E35-R57 salt bridge formation as seen in the C-SA PR, consensus subtype B PR, and subtype B-MDR PR, may be implicated in drug resistance via effects on flap dynamics.

MD simulations

MD simulations were performed to explore the dynamics of the flap region as well as the overall flexibility of the PRs. The root-mean-square fluctuation (RMSF) for each of the 198 residues of the homodimeric HIV PRs was calculated over the 10 ns MD trajectory (Figure 6(A)). No significant differences in main chain B-factors between the C-SA PR and the subtype B PR structures were observed. Seibold and Cukier (2007) reported that the D25 N mutant exhibited the same dynamic behavior as the wild-type PR and has no effect on the overall conformational behavior of the flap regions (Seibold & Cukier, 2007). Moreover, as evident from the RMSF calculations (Figure 6(A)), no noticeable difference in the fluctuation is observed at positions 25/25′. Therefore, the D25N mutation does not affect the overall flap flexibility.

The RMSF for the 99 residues of each monomer of all three PRs shows close similarity over the 10 ns simulation, as is expected for a homodimeric protein. The residues of the hinge region show high RMSF (residues 35–42 and 57–61) when compared to the rest of the protein, as can be seen by the peaks in Figure 6(A).

1376   *P. Naicker* et al.
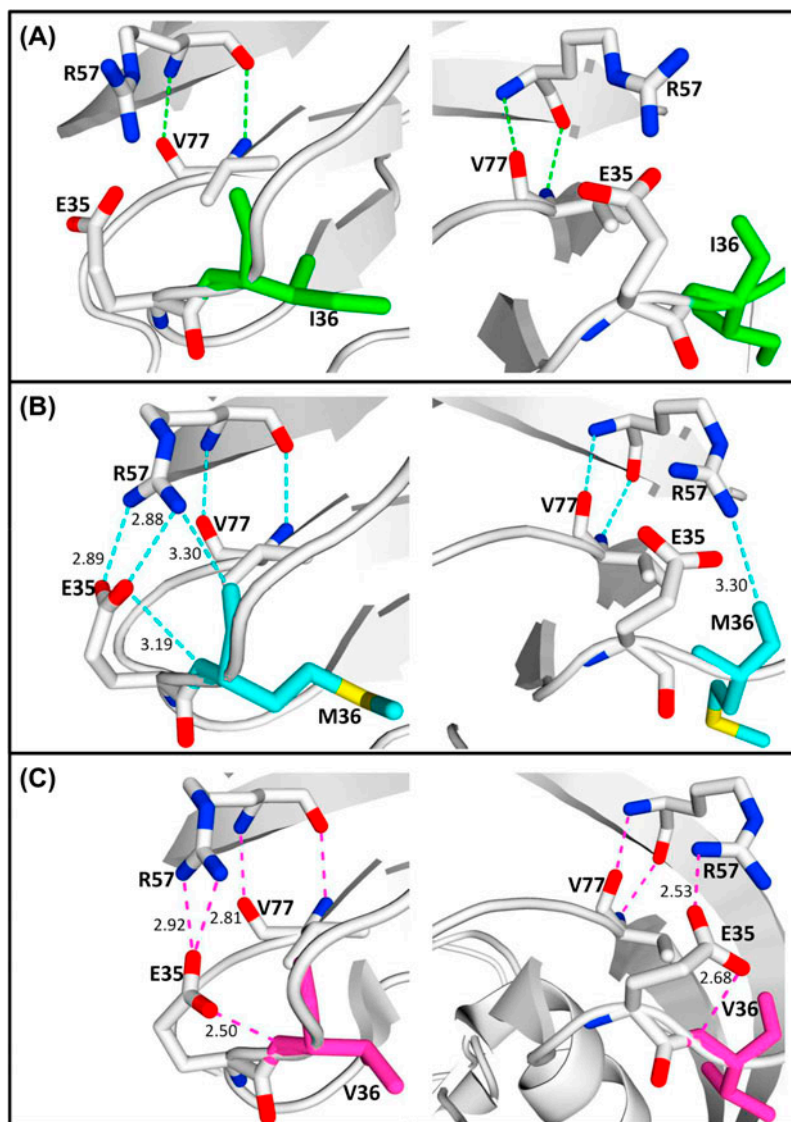
Figure 5.   (A), (B), and (C) show images of the hinge regions of both monomers (left – subunit 1, right – subunit 2) for the C-SA PR, consensus subtype B PR, and subtype B-MDR PR, respectively. The differing hinge-located residue of each PR is highlighted (green: isoleucine, cyan: methionine, and magenta: valine); hydrogen bonds are depicted by dashed lines and interatomic distances in Å are shown.

These residues are solvent exposed and are expected to display high RMSF compared to buried residues such as the catalytic aspartic acids at position 25 of each monomer. Regions of the C-SA PR (red) and consensus subtype B PR (black) that differ significantly in RMSF per residue are the flap tip (residues 46–54) of the first monomer and residues 55–75 of the second monomer. These regions are identical in primary sequence between the two PRs with the exception of the polymorphism at position 69, suggesting that the polymorphisms occurring in the C-SA PR impart local stability changes as well as changes distal to the sites of polymorphisms.

The open conformation of the HIV-1 PR has been inferred from NMR (Ishima, Freedberg, Wang, Louis, &

Torchia, 1999) and MD simulation experiments (Scott & Schiffer, 2000). Both the C-SA PR and consensus subtype B PR displayed open flap conformations (12–14 Å), and relaxed to semi-open conformations ($\sim$ 7 Å) for most of the duration of the simulation, Figure 6(B) showing the inter-flap distance (C_α I50–C_α I50′) of the three PRs. Crystal structures of inhibitor bound subtype B and C PRs display inter-flap distances of 5.86–6.02 Å (Coman et al., 2007; Kempf et al., 1995), representative of a closed conformation. The semi-open conformations of the C-SA PR and consensus subtype B PR overlap with closed inter-flap distances. The subtype B-MDR PR opens wider than the other PRs for most of the simulation, fluctuating with an inter-flap distance of $\sim$ 9 Å in

Figure 6. (A) RMSF of each residue and (B) distance between the flap tips (C_α I50 and C_α I50′) over the 10 ns simulation for the structures under investigation; red (C-SA PR, PDB ID: 3U71), black (consensus subtype B PR, PDB ID: 2PC0), and blue (subtype B-MDR PR, PDB ID: 1RPI). Magenta and blue bars in (A) correspond to the hinge and flap region, respectively.

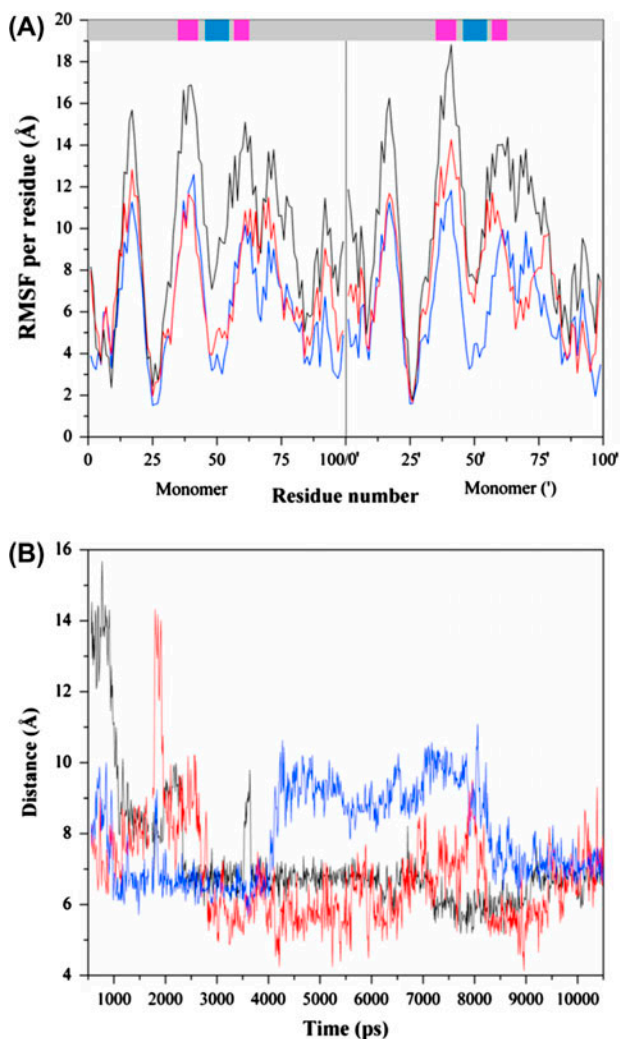its semi-open conformation. Although the inter-flap distances between the C-SA PR and the consensus subtype B PR do not differ greatly for the duration of the simulation, it is important to note that the C-SA PR shows greater fluctuation in inter-flap distance in its semi-open conformation as compared to both subtype B PRs. At, approximately, all conformations of the C-SA PR, its flap tips fluctuate within 3 Å whereas the flap tips of the consensus subtype B PR generally fluctuate within 1 Å of each other. This difference is highlighted in the final ns of the simulation where the inter-flap distances of all three structures coincide and it can be clearly seen that the inter-flap distance of the C-SA PR fluctuates between a wider range than both subtype B PRs. This large fluctuation implies increased flexibility of the flaps

of the C-SA PR in comparison to the subtype B PRs. Although the subtype B-MDR PR exhibits the E35-R57 salt bridge in both its monomers, it still displays a fair degree of fluctuation in inter-flap distance and a wider semi-open conformation for most of the simulation. This discrepancy is likely due to the M46L and I54V mutations which reduces the bulk around the flap tips and has been shown to destabilize the flaps and decrease PI binding (Clemente et al., 2004). As PIs are rigid, they preferentially bind to more closed flap conformations (Clemente et al., 2004). Once again, this emphasizes the importance of the E35-R57 salt bridge which we postulate stabilizes the hinge region of the PR resulting in restrained movement of the flap-hinge region of the PR allowing for more closed flap conformations. In the C-SA PR, the absence of the salt bridge may allow for less restricted movement of the flap-hinge region resulting in the large fluctuations in the inter-flap distances seen in the simulation.

The flexible flaps which are found in all aspartyl PRs are highly conserved in different isolates of HIV-1 and HIV-2 PRs (Gustchina & Weber, 1991). All isolates of HIV-1 and HIV-2 PR have a conserved flap tip (Gustchina & Weber, 1991), 47-IGGIGGFI-54, highlighting its importance in substrate entry and binding. The central isoleucine surrounded by two glycine residues on either side confers flexibility to the flaps, in order for them to adapt to the asymmetry in the substrate sequence (Gustchina & Weber, 1991; Prabu-Jeyabalan, Nalivaika, & Schiffer, 2000). Importantly, I50 and I50′ have been shown to be involved in water-mediated hydrogen bonds with peptide substrate (Prabu-Jeyabalan et al., 2000). Although increased flexibility of the flaps could result in increased substrate entry, interaction of substrate/inhibitor with PR is likely to be of lower affinity. This reduced affinity may be seen as the flaps which are involved in vital hydrogen bonds with substrate/inhibitor, may be displaced to a level that weakens or causes loss of these hydrogen bonds. Previously reported inhibition data shows that the C-SA PR displays reduced affinity for FDA-approved PR inhibitors compared to the subtype B PR, with the C-SA PR displaying the greatest reduction for the inhibitor ritonavir (Mosebi et al., 2008; Velázquez-Campoy et al., 2003). The reduced affinity for PR inhibitors may be a result of increased flexibility of the flaps due to the absence of the E35-R57 salt bridge in the hinge region, as suggested in this study.

Conclusion

The absence of the E35-R57 salt bridge in the C-SA PR results in increased flexibility of the flaps of the enzyme, suggested by the large fluctuations in inter-flap distances during simulation studies. We propose that this increased flexibility is a result of less restrained movement of the

39

flap-hinge region. The increased flap flexibility likely contributes to the increased drug resistance seen in the C-SA PR (Mosebi et al., 2008; Velázquez-Campoy et al., 2003) as suitable conformations of the flaps are vital for effective drug binding.

The polymorphisms between HIV-1 subtype B PR and the C-SA PR do not impact on the overall structure. However, we have identified dynamic differences between the C-SA PR and subtype B PRs. This is the first study to report on the crystal structure of the C-SA PR which is an epidemiologically relevant PR. The crystal structure of the C-SA PR will serve as a foundation to improve the rational design of PIs which will have a greater impact on antiretroviral chemotherapy in sub-Saharan Africa.

Abbreviations

PR	Protease
PI	PR inhibitor
C-SA PR	South African subtype C PR
PDB	Protein data bank
MD	Molecular dynamics
RMSF	Root-mean-square fluctuation

Acknowledgments

References

Bandaranayake, R. M., Kolli, M., King, N. M., Nalivaika, E. A., Heroux, A., Kakizawa, J., … Schiffer, C. A. (2010). The effect of clade-specific sequence polymorphisms on HIV-1 protease activity and inhibitor resistance pathways. *Journal of Virology, 84*, 9995–10003.

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. *Journal of Biological Chemistry, 264*, 16948–16956.

Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *Journal of Chemical Physics, 81*, 3684–3690.

Buonaguro, L., Tornesello, M. L., & Buonaguro, F. M. (2007). Human immunodeficiency virus type 1 subtype distribution in the worldwide epidemic: Pathogenetic and therapeutic implications. *Journal of Virology, 81*, 10209–10219.

Carmel, A., & Yaron, Y. (1978). An intramolecularly quenched fluorescent tripeptide as a fluorogenic substrate of angiotensin-i-converting enzyme and of bacterial dipeptidyl carboxypeptidase. *European Journal of Biochemistry, 87*, 265–273.

Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., … Woods, R. J. (2005). The amber biomolecular simulation programs. *Journal of Computational Chemistry, 26*, 1668–1688.

Chayen, N. E., & Stewart, P. D. S. (1992). Microbatch crystallization under oil – a new technique allowing many small-volume crystallization trials. *Journal of Crystal Growth, 122*, 176–180.

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., … Richardson, D. C. (2010). *MolProbity*: All-atom structure validation for macromolecular crystallography. *Acta Crystallographica, Section D Biological Crystallography, 66*, 12–21.

Clemente, J. C., Moose, R. E., Hemrajani, R., Whitford, L. R. S., Govindasamy, L., Reutzel, R., … Dunn, B. M. (2004). Comparing the accumulation of active- and nonactive-site mutations in the HIV-1 protease. *Biochemistry, 43*, 12141–12151.

Coman, R. M., Robbins, A. H., Goodenow, M. M., Dunn, B. M., & McKenna, R. (2008). High-resolution structure of unbound human immunodeficiency virus 1 subtype C protease: Implications of flap dynamics and drug resistance. *Acta Crystallographica, Section D Biological Crystallography, 64*, 754–763.

Coman, R. M., Robbins, A., Goodenow, M. M., McKenna, R., & Dunn, B. M. (2007). Expression, purification and preliminary X-ray crystallographic studies of the human immunodeficiency virus 1 subtype C protease. *Acta Crystallographica, Section F Structural Biology and Crystallization Communications, 63*, 320–323.

De Oliveira, T., Engelbrecht, S., Rensburg, E. J. V., Gordon, M., Bishop, K., Zur Megede, J., … Cassol, S. (2003). Variability at human immunodeficiency virus type 1 subtype C protease cleavage sites: An indication of viral fitness? *Journal of Virology, 77*, 9422–9430.

Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G. M., Zhang, W., … Kollman, P. A. (2003). Point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *Journal of Computational Chemistry, 24*, 1999–2012.

Emsley, P., & Cowtan, K. (2004). *Coot*: Model-building tools for molecular graphics. *Acta Crystallographica, Section D Biological Crystallography, 60*, 2126–2132.

Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., & Pedersen, L. G. (1995). A smooth particle mesh Ewald method. *Journal of Chemical Physics, 103*, 8577–8593.

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 Letters, 269*, 269–272.

Gustchina, A., & Weber, I. T. (1991). Comparative analysis of the sequences and structures of HIV-1 and HIV-2 proteases. *Proteins, 10*, 325–339.

Heaslet, H., Rosenfeld, R., Giffin, M., Lin, Y. C., Tam, K., Torbett, B. E., … Stout, C. D. (2007). Conformational flexibility in the flap domains of ligand-free HIV protease. *Acta Crystallographica, Section D Biological Crystallography, 63*, 866–875.

Hemelaar, J., Gouws, E., Ghys, P. D., & Osmanov, S. (2006). Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. *AIDS, 20*, 13–23.

Higgins, D. G., & Sharp, P. M. (1988). CLUSTAL: A package for performing multiple sequence alignment on a microcomputer. *Gene, 73*, 237–244.

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

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

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics, 79*, 926–935.

Kantor, R., & Katzenstein, D. (2003). Polymorphism in HIV-1 non-subtype B protease and reverse transcriptase and its potential impact on drug susceptibility and drug resistance evolution. *AIDS Reviews, 19*, 25–35.

Kempf, D. J., Marsh, K. C., Denissen, J. F., McDonald, E., Vasavanonda, S., Flentge, C. A., … Norbeck, D. W. (1995). ABT-538 is a potent inhibitor of human immunodeficiency virus protease and has high oral bioavailability in humans. *Proceedings of the National Academy of Sciences USA, 92*, 2484–2488.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature, 227*, 680–685.

Larkin, M. A., Blackshields, G., Brown, N .P., Chenna, R., McGettigan, P. A., McWilliam, H., … Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics, 23*, 2947–2948.

Laskowski, R. A. (1993). *PROCHECK*: A program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography, 26*, 283–291.

Logsdon, B. C., Vickrey, J. F., Martin, P., Proteasa, G., Koepke, J. I., Terlecky, S. R., … Kovari, L. C. (2004). Crystal structures of a multidrug-resistant human immunodeficiency virus type 1 protease reveal an expanded active-site cavity. *Journal of Virology, 78*, 3123–3132.

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

McCutchan, F. E. (2006). Global epidemiology of HIV. *Journal of Medical Virology, 78*, 7–12.

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.

Mildner, A. M., Rothrock, D. J., Leone, J. W., Bannow, C. A., Lull, J. M., Reardon, I. M., … Tomasselli, A. G. (1994). The HIV-1 protease as enzyme and substrate: Mutagenesis of autolysis sites and generation of a stable mutant with retained kinetic properties. *Biochemistry, 33*, 9405–9413.

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. *Journal of Virology, 82*, 11476–11479.

Murshudov, G. N., Vagin, A. A., & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallographica, Section D Biological Crystallography, 53*, 240–255.

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

Patick, K. A., Duran, M., Cao, Y., Shugarts, D., Keller, M. R., Mazabel, E., … Markowitz, M. (1998). Genotypic and phenotypic characterization of human immunodeficiency virus type 1 variants isolated from patients treated with the protease inhibitor nelfinavir. *Antimicrobial Agents and Chemotherapy, 42*, 2637–2644.

Plantier, J. C., Leoz, M., Dickerson, J. E., De Oliveira, F., Cordonnier, F., Lemee, V., … Simon, F. (2009). A new human immunodeficiency virus derived from gorillas. *Nature Medicine, 15*, 871–872.

Potterton, E., Briggs, P., Turkenburg, M., & Dodson, E. J. (2003). A graphical user interface to the *CCP*4 program suite. *Acta Crystallographica, Section D Biological Crystallography, 59*, 1131–1137.

Prabu-Jeyabalan, M., Nalivaika, E., & Schiffer, C. A. (2000). How does a symmetric dimer recognize an asymmetric substrate? A substrate complex of HIV-1 protease. *Journal of Molecular Biology, 301*, 1207–1220.

Ryckaert, J. P., Ciccotti, G., & Berendsen, H. J. C. (1977). Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics ofn-alkanes. *Journal of Computational Physics, 23*, 327–341.

Schägger, H. (2006). Tricine-SDS-PAGE. *Nature Protocols, 1*, 16–23.

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.

Seibold, S. A., & Cukier, R. I. (2007). A molecular dynamics study comparing a wild-type with a multiple drug resistant HIV protease: Differences in flap and aspartate 25 cavity dimensions. *Proteins, 69*, 551–565.

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 Journals, 277*, 3699–3714.

Smith, R., Brereton, I. M., Chai, R. Y., & Kent, S. B. H. (1996). Ionization states of the catalytic residues in HIV-1 protease. *Nature Structural Biology, 3*, 946–950.

Swain, A. L., Miller, M. M., Green, J., Rich, D. H., Schneider, J., Kent, S. B. H., & 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. *Proceedings of the National Acedemy of Sciences of the USA, 87*, 8805–8809.

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.

Szeltner, Z., & Polgar, L. (1996). Rate-determining steps in HIV-1 protease catalysis. *Journal of Biological Chemistry, 271*, 32180–32184.

Taylor, B., Sobieszczyk, M. E., McCutchan, F. E., & Hammer, S. M. (2008). The challenge of HIV1 subtype diversity. *New England Journal of Medicine, 358*, 1590–1602.

41

Tomasselli, A. G., & Heinrikson, R. L. (2000). Targeting the HIV-protease in AIDS therapy: A current clinical. *Science, 1477*, 189–214.

Vagin, A., & Teplyakov, A. (2000). An approach to multi-copy search in molecular replacement. *Acta Crystallographica, Section D Biological Crystallography, 56*, 1622–1624.

Velázquez-Campoy, A., Vega, S., Fleming, E., Bacha, U., Sayed, Y., & Dirr, H. W. (2003). Protease inhibition in African subtypes of HIV-1. *AIDS Reviews, 410*, 165–171.

Velázquez-Campoy, A., Kiso, Y., & Freire, E. (2001). The binding energetics of first- and second-generation HIV-1 protease inhibitors: Implications for drug design. *Archives of Biochemistry and Biophysics, 390*, 169–175.

Wlodawer, A., & Vondrasek, J. (1998). Inhibitors of HIV-1 Protease: A major success of structure-assisted drug design. *Annual Review of Biophysics & Biomolecular Structure, 27*, 249–284

Yedidi, R. S., Proteasa, G., Martinez, J. L., Vickrey, J. F., Martin, P. D., Wawrzak, Z., … 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 Crystallographica, Section D Biological Crystallography, 67*, 524–532.

CHAPTER 4

Amide Hydrogen Exchange in HIV-1 subtype B and C Proteases: Insights into reduced drug susceptibility and dimer stability

Previn Naicker, Stoyan Stoychev, Heini W. Dirr and Yasien Sayed.

In this article, differences in substrate processing and inhibitor susceptibility between the subtype B and C-SA PRs were identified. Dynamics at all regions of the PRs were determined and the results were used to explain the aforementioned differences.

Author contributions: Previn Naicker performed all experimental work, analysed the data and wrote the manuscript. Stoyan Stoychev assisted with the HDX-MS experiments. Heini W. Dirr assisted in experimental design and revision of manuscript. Yasien Sayed supervised the project and assisted in data analysis and interpretation.

Amide hydrogen exchange in HIV-1 subtype B and C proteases - insights into reduced drug susceptibility and dimer stability

Previn Naicker [1], Stoyan Stoychev [2], Heini W. Dirr [1], Yasien Sayed [1,*]

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

[2] Council for Scientific and Industrial Research, Biosciences, Pretoria 0001, South Africa

* Corresponding author. Tel.: +2711 7176350; Fax: +2711 7176351.

E-mail address: yasien.sayed@wits.ac.za

Running title: Amide hydrogen exchange in HIV-1 subtype B and C proteases

Abbreviations: FDA, United States food and drug administration; HDX, hydrogen/deuterium exchange; MS, mass spectrometry; PR, protease; PI, protease inhibitor; PDB, Protein Data Bank; C-SA, South African subtype C.

Enzyme commission number: **EC 3.4.23.16** (HIV-1 retropepsin)

Key words: Protein dynamics; hydrogen/deuterium exchange; drug binding; flap conformations; dimer interface.

Abstract

Since its identification, HIV continues to have a detrimental impact on the lives of millions of people throughout the world. The protease (PR) of HIV is a major target in antiviral treatment. The South African HIV-1 subtype C protease (C-SA PR) displays weaker binding affinity for some clinically-approved protease inhibitors in comparison to the HIV-1 subtype B protease. The heavy HIV burden in sub-Saharan Africa, where subtype C HIV-1 predominates, makes this disparity a topic of great interest. In light of this, the enzyme activity and affinity of protease inhibitors for the subtype B and C-SA PRs were determined. The relative vitality, indicating the selective advantage of polymorphisms, of the C-SA PR relative to the subtype B PR in the presence of ritonavir and darunavir was 4-fold and 10-fold greater, respectively. Dynamic differences that contribute to the reduced drug susceptibility of the C-SA PR were investigated by performing hydrogen/deuterium exchange-mass spectrometry (HDX-MS) on the unbound subtype B and C-SA PRs. The reduced propensity to form the E35-R57 salt bridge and alterations in the hydrophobic core of the C-SA PR is proposed to affect the anchoring of the flexible flaps resulting in an increased proportion of the fully-open flap conformation. HDX-MS data showed that the N-terminus of both PRs appears to be less stable than the C-terminus of the PRs, thus rationalising the increased efficacy of dimerisation inhibitors targeted toward the C-terminus of HIV PRs. This is the first known report on HIV protease dynamics using HDX-MS.

Introduction

Human immunodeficiency virus (HIV) infections are a global health and socio-economic challenge with approximately 35 million people living with the virus in 2012 [1]. Disturbingly, ~ 70% of these people reside in sub-Saharan Africa and HIV-1 subtype C is the most prevalent subtype in this region and globally [1, 2]. In South Africa, an estimated 6.1 million people (~ 12% of the total population) are living with HIV [1]. The homodimeric HIV protease (PR) is a major drug candidate because it has a central role in the life cycle of the virus. The first-line antiretroviral therapy (ART) guidelines for adults and adolescents include the use of two nucleoside reverse transcriptase inhibitors in combination with one non-nucleoside reverse transcriptase inhibitor [3]. Second-line ART may proceed upon failure of the first-line ART regimen in an individual. The preferred protease inhibitors (PIs) in second-line ART for adults and adolescents are atazanavir, darunavir or lopinavir in combination with the PI booster, ritonavir [3].

The present study investigates the differences in binding affinity to clinically-used PIs and the dynamic differences of the consensus HIV-1 subtype B protease (subtype B PR) and South African HIV-1 subtype C protease (C-SA PR) which is effectively the consensus HIV-1 subtype C protease (Los Alamos HIV sequence database, http://www.hiv.lanl.gov/). Importantly, the C-SA PR differs by eight amino acids per PR monomer to the subtype B PR (Fig. 1). FDA-approved PIs have been designed against the HIV-1 subtype B PR which predominates in North America and Europe and is less prevalent globally [2]. When comparing the C-SA PR to the subtype B PR, the following polymorphisms are observed: 3 situated in the fulcrum region

(T12S, I15V and L19I), 2 in the hinge region (M36I and R41K), 1 in the 60's loop (H69K), 1 in the 80's loop (L89M) and 1 at the end of the single α-helix of each monomer (I93L) [4]. These polymorphisms represent non-active site mutations and therefore, do not directly affect inhibitor binding [4].

Previous comparison of PI binding to the C-SA PR showed a ~ 8-fold weaker affinity for lopinavir than the subtype B PR [5]. Only recently has the crystal structure of the C-SA PR been elucidated [6], whereas, the structure of the subtype B PR has been well studied. A structural comparison of the PRs does not definitively provide a clear basis for the differences in drug binding. Molecular dynamics (MD) simulations showed that the inter-flap distance (distance between C_α of I50 on adjacent monomers) of the subtype B PR appeared to fluctuate within a 1 Å distance, whereas, that of the C-SA PR fluctuated within 3 Å, implying altered flexibility of the flap tips (residues at positions 46–54) of the C-SA PR [6]. Increased flap flexibility may result in reduced binding affinity to PIs as the flap tips are involved in water-mediated interactions with substrates/inhibitors [7]. Distance measurements using spin-labelled pulsed electron paramagnetic resonance (EPR) spectroscopy, displayed an increased proportion of the curled and fully-open conformations (Fig. 2C) of the flaps in the subtype C PR population relative to other HIV-1 PRs [8]. No x-ray diffracted structural data is available for the curled and fully-open conformers. The curled conformer is described by pronounced curling of the flap tips (residue positions 48–52) [9]. The fully-open conformer also displays flap tip curling, accompanied by upward and outward displacement of the flap-hinge region from the active site [10]. The majority of C-SA PR molecules exist in the semi-open conformation in solution (Fig. 2B) [8, 10-12]; however, there is a definite shift in the equilibrium between semi-open and other

flap conformers. At baseline, an ensemble of semi-open conformers may exist. These semi-open conformers may directly interchange with closed conformers (Fig. 2A). Structural and PR dynamics studies indicate that the curled conformer may be an intermediate state which allows for the formation of the fully-open conformer [9, 11, 13]. Entry of a substrate or inhibitor to the active site of the PR requires substantial movement of the flaps (~ 15 Å from their position in the closed conformer) [14]. Structure-based calculations reveal that semi-open conformers do not permit the entry of substrate/inhibitor to the active site [15]. Therefore, it is accepted that only the fully-open conformation allows for substrate/inhibitor binding. Complete flap opening occurs through concerted downward movement of the hinge (residue positions 35–42 and 57–61), cantilever (residue positions 62–78) and fulcrum (residue positions 10–23) regions (Fig. 1) which results in the upward and outward motion of the flaps (Fig. 2C) [10]. The vast majority of studies measuring PR dynamics are investigated using MD simulations. MD simulation results require careful attention for correct interpretation; however, are useful when complemented with experimental data. Other analyses of PR dynamics were investigated using nuclear magnetic resonance (NMR) [11, 12, 16], which is limited by the modest sensitivity of the technique, and spin-labelled pulsed EPR spectroscopy [8, 17], which was used to exclusively measure distances between flap residues. Combined EPR spectroscopy and NMR experiments have identified and verified the population of flap conformers in the apo-state and during inhibitor binding for the subtype B and C HIV-1 PRs [18]. However, these experiments did not fully elucidate the mechanism for the reduced drug susceptibility displayed by the subtype C HIV-1 PR. Here we probed the mechanism for the altered flap flexibility and reduced susceptibility of the C-SA PR to PIs by performing hydrogen/deuterium exchange-mass spectrometry (HDX-MS) on apo-

subtype B and apo-C-SA PRs. To our knowledge, this is the first HDX-MS study performed on HIV-1 proteases.

Results

C-SA PR viral fitness

Following expression and purification, the subtype B and C-SA PRs were > 99% pure. The HIV-1 PR fluorogenic substrate (Abz-Arg-Val-Nle-Phe(NO$_2$)-Glu-Ala-Nle-NH$_2$), which mimics the capsid/p2 cleavage site in the HIV-1 Gag polyprotein, was used to determine the kinetic parameters of the PR-substrate reaction (Table 1). Both PRs display comparable kinetic parameters; however, the C-SA PR shows a two-fold increase in the turnover of substrate, as seen by a higher k_{cat} value (Table 1).

Inhibition constants (K_i) of PIs for the subtype B and C-SA PRs were determined from their respective IC_{50} determinations (Table 2). The thermodynamic parameters (ΔG, ΔH, $-T\Delta S$, K_d and n), determined by ITC, describe the overall energetic contribution to the interaction. For all ITC experiments, a stoichiometry value (n) of > 0.95 was achieved. Binding thermodynamics of acetyl-pepstatin, a substrate analogue, is similar for both PRs. Displacement titration ITC was employed to improve the accuracy of the determined K_d values, the affinity of the FDA-approved PIs for both the subtype B [19, 20] and C-SA PRs were in the nM–pM range (Table 2). The binding affinities of the PIs for both PRs were also compared using the K_i and relative vitality values. Atazanavir displays a similar affinity for the subtype B and C-SA PRs. Darunavir shows

a 7-fold reduced affinity for the C-SA PR compared to the subtype B PR. The binding of darunavir to both PRs is enthalpically driven; however, binding to the C-SA PR displayed an unfavourable entropy change ($-T\Delta S > 0$) which contributes to the reduced affinity. The C-SA PR displays a relative vitality of ~ 4 and ~ 10 in the presence of ritonavir and darunavir, respectively. This suggests that subtype C HIV-1 may have a greater capacity to replicate in the presence of ritonavir and darunavir when compared to subtype B HIV-1. Therefore, the C-SA PR displays reduced drug susceptibility to ritonavir and darunavir in comparison to the subtype B PR.

Amide hydrogen-deuterium exchange

Two aspartyl proteases, porcine pepsin and protease XIII from *Aspergillus saitoi*, were tested for optimum cleavage of the HIV-1 PRs following deuterium labelling. Due to the pH range of their optimal activity, aspartyl proteases are the only enzymes used for fragmentation during a HDX-MS experiment. The competitive inhibition of pepsin by FDA-approved HIV-1 PIs has been described previously [21]. K_i values of 0.6 µM and 2.1 µM were obtained previously for binding of porcine pepsin to ritonavir and darunavir, respectively [21]. In the present study, attempts to determine HDX in the presence of PIs were unsuccessful. After optimisation, cleavage of the HIV-1 PRs in the presence of HIV-1 PIs by both pepsin and protease XIII was poor, producing large amounts of undigested PR and only 5 and 2 detected peptide fragments, respectively. Therefore, HDX-MS analysis was limited to the unbound form of the PRs.

The conformational dynamics of the apo-PRs were measured by detecting the exchange of backbone amide hydrogen atoms with deuterium atoms in solution. In this study, amide hydrogen/deuterium exchange (HDX) was detected via mass spectrometry. HDX was measured at pH 5.0 in accordance with the activity and inhibition assays. The catalytic dyad (D25 and D25′) of HIV-1 PR is monoprotonated at pH 5–6 and the PR displays maximal activity and stability in this pH range [22]. At pH 7.0, the catalytic dyad loses its proton and becomes more unstable [23]. Measuring HDX at pH 5.0 also enables the detection of amide protons that may exchange too fast at pH 7.0. At pH > 4.0, HDX is catalysed by direct attack and H-bonding of OD^- to the exchangeable proton, thus, HDX at pH 5.0 reports on PR dynamics [24-26]. The HDX reaction was quenched at several time points (10 s–1 hr) to map fast as well as intermediate exchanging regions of the PRs. During analysis, 35 peptides derived from the subtype B PR and 41 peptides from the C-SA PR were used to provide complete coverage of both PRs. Thirty-four of these peptides were in common to both PRs. The first and second amide protons of each peptide are unable to retain deuterium due to cleavage of the PRs by pepsin and binding of the resultant peptides to the C18 reversed-phase column [27]. Proline residues do not possess an amide hydrogen and HDX at these residues along with the first and second residues of each peptide are not measured. The heat maps in Figure 3 display the percentage of deuteration across all regions of the PRs barring the residues at positions 1–2. Proteins exhibit highly mobile behaviour in solution. Exchange in a folded protein is believed to occur through low amplitude atomic motions, of ~ 1 Å, which are sufficient to allow for diffusion of D_2O and OD^- to backbone amide linkages [28, 29]. Thus, there is a steady increase in the percentage of deuteration at various regions of the PRs with increasing incubation time (Fig. 3).

The amplitude and rates of HDX in peptides covering the regions near polymorphic sites were measured (Table 3). During analysis, fast and intermediate exchange rates were determined and are similar in the respective peptides of both PRs. Engen et al. have proposed that amide protons exchanging at very fast rates (i.e. unprotected and solvent exposed) must exchange from the folded form of the protein [30]. Amide protons exchanging at fast to slow rates require local structural fluctuations prior to exchange. In peptides 13–23, 34–53 and 64–90 the number of amide protons with very fast and/or fast exchange rates is less in the C-SA PR than in the subtype B PR. The number of amide protons in these peptides exhibiting intermediate exchange rates is greater in the C-SA PR. The total number of amide protons exchanged is similar in both PRs after 1 hour. Therefore, the number of slow exchanging amide protons in these peptides is similar and differences are restricted to the number of fast and intermediate exchanging amide protons. The regions of HIV-1 PR covered by these peptides are displayed in Figure 4. Deuterium incorporation plots show the number of deuterons incorporated into the relevant peptides with respect to the incubation time. Differences in deuterium incorporation at the N- and C- termini are highlighted in Table 4. Exchange at the N-terminus of both PRs is similar. The C-SA PR shows increased dynamics at C95 and T96 in comparison to the subtype B PR.

Discussion

Dynamics of the HIV-1 PR

Heat maps (Fig. 3) show results as percentage deuteration, thereby, normalising the deuterium incorporation in the measured peptides and allowing for a direct comparison of different regions

of the HIV-1 PR. Increased spatial resolution in the heat maps is allowed for by peptide overlap. Improved spatial resolution for the C-SA PR is evident in parts of the heat maps due to increased peptide detection and overlap. Both PRs display similar exchange profiles across different regions. Residues around the active site triplet (D25, T26 and G27) display slow deuterium incorporation in comparison to the other regions of the PRs. These residues are positioned deep within the active site cavity. Slow exchange is expected here due to low solvent accessibility and a complex network of interactions around the active site core. The flap tips (residues at positions 49–53) are the regions of the PRs that display the fastest deuterium incorporation. The flaps are completely solvent exposed and highly mobile regions of the PRs. The reported HDX-MS profile of HIV-1 PR is in agreement with previous NMR data [11, 12, 16].

HDX in macromolecules resulting from local unfolding events may either occur via a correlated (EX1) or uncorrelated (EX2) mechanism [26, 31, 32]. HDX in the subtype B and C-SA PRs displays EX2 kinetics indicated by the single binomial isotopic distribution of the mass spectra in the current study [33]. This uncorrelated exchange mechanism may be well-illustrated by HDX at α-helices [34]. As seen in Figure 3, exchange at residues in the single α-helix of the HIV-1 PR (residues at positions 87–93) occurs independently of each other. HDX via the EX2 mechanism suggests that segments of the PRs must unfold and refold many times before exchange within it is complete [26, 31, 32].

Comparison of conformational stability between the subtype B and C-SA proteases

Previous differential scanning calorimetry studies showed that the overall conformational stability of the C-SA PR was slightly reduced in comparison to the subtype B PR [5]. In the current study, subtle differences in stability at different regions of the PRs were observed. In most peptides, the number of amide protons exchanging at fast rates increased by ~ 1 amide proton in the subtype B PR relative to the C-SA PR. This may imply that the C-SA PR is more stable than the subtype B PR. However, residues at the C-terminus of the C-SA PR appear to be more dynamic (Table 4). Thermodynamic analyses showed that the N- and C-terminal antiparallel β-sheet contributes 75% to the total Gibbs energy [35]. A more dynamic C-terminus will affect the stability of the terminal β-sheet and greatly impact the overall conformational stability. Salt bridges also affect the conformational stability of a macromolecule [36, 37]. Crystal structure analysis shows that the C-SA PR lacks the K20-E34 and E35-R57 salt bridges relative to the subtype B PR in their crystalline forms. Therefore, the reduced number of ionic interactions and a less stable terminal β-sheet evident in the C-SA PR are major determinants for the apparent overall reduced conformational stability of the C-SA PR.

Dynamics around the polymorphic sites

Figure 1 shows the position of the eight polymorphisms in the subtype B and C-SA PRs. As mentioned previously, 3 polymorphic sites are situated in the fulcrum region (T12S, I15V and L19I), 2 in the hinge region (M36I and R41K), 1 in the 60's loop (H69K), 1 in the 80's loop (L89M) and 1 at the end of the single α-helix of each monomer (I93L). For peptides 13–23, 34–

53 and 64–90 (Fig. 4) subtle differences in deuterium incorporation between the subtypes are evident. In all of the aforementioned peptides covering the 60's and 80's loops, the fulcrum region and the flap-hinge region, there is a reduction in the number of very fast and/or fast exchanging amide protons ($12 > k > 0.1$ min^{-1}) and a concomitant increase in the number of intermediate exchanging amide protons ($0.1 > k > 0.01$ min^{-1}) in the C-SA PR (Table 3). These results indicate that these regions of the C-SA PR are more stable than that of the subtype B PR.

The differences observed in peptides 13–23 (covering the fulcrum region), 34–53 (covering the hinge region) and 64–90 are explained in part by the varying preference of flap conformations between the PRs. Distance measurements during pulsed EPR experiments showed that the total population of the fully-open conformer increased by 23% with a concomitant decrease by 15% in the total population of semi-open conformers in the subtype C PR relative to the subtype B PR. Complete flap opening occurs through concerted downward movement of the hinge (residue positions 35–42 and 57–61), cantilever (residue positions 62–78) and fulcrum (residue positions 10–23) regions which results in the upward and outward motion of the flaps (Fig. 2C) [10]. The cantilever region is covered by peptide 64–76 in the current study (Table 3). Slightly increased dynamics is also displayed in the cantilever region of the subtype B PR. In the fully-open conformer, the parts of the hinge and cantilever regions are anchored in the hydrophobic core of the PR [38]. Therefore, parts of the hinge, cantilever and fulcrum regions are likely stabilised in the fully-open conformer. This detail may explain in part the reduced dynamics at peptides 13–23, 34–53 and 64–90 in the C-SA PR.

Dynamics at the flap tips

A recent simulation study indicated altered flap tip flexibility between the apo-form of the C-SA PR and subtype B PR on a picosecond to nanosecond time scale [6]. The C-SA PR displayed a wider range of open conformers than the subtype B PR. These findings hint at an altered stability around the flap tips of the C-SA PR which allow for the variation in open conformers. Similar to the thermodynamic data reported here (Table 2), HIV-1 PRs with flap tip mutations displayed reduced binding entropy in comparison to wild-type PR, which is compensated for by enhanced binding enthalpy during atazanavir and darunavir binding [39]. Only subtle packing rearrangements around the flap region and inhibitors were identified in the inhibitor-bound structures of these mutants [39]. Observed entropic effects during inhibitor binding to PRs with flap tip mutations as well as the C-SA PR indicate changes in flap flexibility and/or solvation in comparison to the wild-type subtype B PR. Because the C-SA PR does not possess any flap tip mutations, increased flexibility of the flap tips is unexpected. As mentioned, the C-SA PR has a greater preference for the fully-open conformation than other HIV-1 subtypes including multi-drug resistant variants [8]. An increased proportion of fully-open conformers are also evident in C-SA PR in comparison to the subtype B PR when saturated with atazanavir and darunavir [18]. Fully-open conformers allow access of more ordered solvent to the active site than semi-open conformers. These findings may explain the less favourable binding entropy displayed by the C-SA PR during binding to atazanavir and darunavir (Table 2).

Data for peptide 47–53 measures deuterium incorporation at backbone amides for residues at positions 49–53 (flap tips). Importantly, the flap tips are solvent exposed in all flap conformers.

Therefore, HDX at the flap tips are mostly dependent on local atomic fluctuations which reduce the number of protected amide protons (amide protons involved in interactions) rather than an increase in solvent accessibility to unprotected amide protons. The C-SA PR displays a reduction in the number of very fast exchanging amide protons (Table 3) which reflect the amides that are available for exchange at baseline. The C-SA PR also displays a slight reduction in the number of fast exchanging amide protons. This difference is explained in part by the increase in the proportion of curled conformers in the C-SA PR population in comparison to that of the subtype B PR. Distance measurements during pulsed EPR experiments showed that the total population of the curled conformer increased by 8% in the subtype C PR relative to the subtype B PR [8]. Importantly, during flap tip curling I50 of each monomer is buried and stabilised by hydrophobic contacts with I47, I54, P79 and P81 [9]. This affects HDX-MS data, as more interactions need to be broken in a curled conformation before amide proton exchange proceeds relative to the semi-open conformation. Moderate fluctuations of > 3 Å between interacting groups are required for OD$^-$ catalyst attack when exchangeable amide protons are involved in interactions [40]. In a similar manner, MD simulation data suggest that the fully-open conformer displays curling of the flap tips [10]. As mentioned, the C-SA displays a greater population of fully-open conformers in comparison to the subtype B PR. Therefore, the increased proportion of fully-open conformers in the C-SA PR population may also account for the reduction in fast exchanging amide protons (Table 3) in the flap tips of the C-SA PR. When dealing with multiple conformers, it is important that data has been standardised. Importantly, all HDX-MS data was derived from subtype B and C-SA PR solutions with an active site concentration of greater than 95% (see methods; active site determination). This active concentration illustrates the population of PRs in solution that

display an active conformation. Therefore, a fair comparison of PR dynamics may be made based on differences in flap conformers between the subtype B and C-SA PRs.

Effect of secondary resistance mutations on flap movement

Globally, the structure of the subtype B and C-SA PRs do not differ. Detailed comparison of the structures revealed no significant differences at any of the polymorphic sites with the exception of the residue at position 36. Subtle differences are evident in the hinge region (residue positions 35–42 and 57–61) of the PRs (Fig. 5) [6]. The hinge region is implicated in the control of the flexible flaps of the PR and the flap tips are involved in water-mediated contacts with substrates/inhibitors. Importantly, M36 in the subtype B PR shows no exchange over all time points, whereas, I36 in the C-SA PR has exchanged its amide hydrogen at 10 seconds (Fig. 3). This supports the crystallographic evidence which shows the amide group of M36 in the subtype B PR interacting with the E35 side chain, thereby, facilitating the E35-R57 salt bridge (Fig. 5A) which is the only salt bridge in the flap-hinge region. I36 of the C-SA PR is not involved in a similar interaction (allowing for exchange of its amide hydrogen). Therefore, the C-SA PR has a reduced propensity to form the E35-R57 salt bridge likely contributing to altered flap dynamics.

The M36I mutation is considered to be a secondary drug resistance mutation. The M36I mutation is the only polymorphic mutation in the C-SA PR which is associated with drug resistance to most of the FDA-approved PIs in combination with ritonavir (Stanford University HIV drug resistance database [41]). This mutation is associated with virologic failure or reduced virologic response to atazanavir, darunavir, fosamprenavir, indinavir, lopinavir, saquinavir and tipranavir

in combination with ritonavir [42-49]. Mutations at residue position 89 are associated with virologic failure to darunavir and tipranavir in combination with ritonavir [44, 50] and mutations at residue position 93 are associated with virologic failure to saquinavir in combination of ritonavir [42, 49]. Although the other polymorphisms inherent in the C-SA PR have not been associated with virologic failure in clinical studies, they may still contribute to the observed reduced drug susceptibility of the C-SA PR.

The importance of other polymorphisms inherent to the C-SA PR in relation to reduced drug susceptibility may be explained through the process of hydrophobic sliding [38]. A molecular dynamics simulation study identified hydrophobic residues whose side chains were buried for the majority of the simulation [38]. These 19 residues per monomer are referred to as the hydrophobic core of the PR (Fig. 6) [38]. Although the residue at position 19 was not identified as part of the hydrophobic core, the L19I mutation may contribute to rearrangement of interactions in the hydrophobic core. Therefore, L/I19 was included in the hydrophobic core in the current study. In the unliganded PR, rearrangement of interactions in the hydrophobic core facilitate flap movement upward and outward from the active site [38]. Seven isoleucines per monomer are present in the core and may provide increased flexibility because they can exist in more possible conformations than other hydrophobic residues [38]. In the crystal structure of the subtype B PR, I15 makes van der Waal contacts with L33, M36 and L38 which form part of a loop in the hinge region. I15 also makes van der Waals contacts with I62, I64 and V75 in the cantilever region comprising residues at positions 62–78 which is closely associated with the hinge region. During flap movement, the loops (hinge and cantilever regions) slide over I15 and the contacts that are lost by I15 are gained by I13 (both I13 and I15 forming part of the fulcrum

region) [38]. Interaction of these loops with the hydrophobic core allow for anchoring of the flaps. When the loops slide over residues of the hydrophobic core, wider opening of the flaps occur. Cross-linking studies have confirmed that altered flexibility in the hydrophobic core modulates PR activity and movement of the core is required for function [51]. Rearrangement of the hydrophobic core residues and exchange of one set of hydrophobic contacts for another likely result in modest energetic penalties [38].

In the C-SA PR, I15V, L19I, M36I, L89M and I93L polymorphisms occur in the hydrophobic core. These polymorphisms may alter the network of hydrophobic contacts in the core. In the crystal structure of the C-SA PR, V15 does not interact with V75. However, this interaction may occur during flap movement. It is evident from the crystal structures that these polymorphisms bring about slight changes in the complex network of hydrophobic interactions. Due to its high incidence in drug-resistant isolates, the M36I mutation likely contributes the greatest to differences in flap dynamics between the PRs. Rotation of this I36 through different rotamers may assist in sliding over other hydrophobic residues. Only subtle differences in the hydrophobic core are evident between the crystal structures of the subtype B and C-SA PRs. Similarly, the HIV-1 A_E variant which has polymorphisms in the hydrophobic core does not display major structural differences in this region relative to the subtype B PR [52]. Analysis of the dynamics of the hydrophobic core is limited because residues in the core are disconnected in the primary structure of the HIV-1 PR. Residue-level resolution of hydrophobic core residues was not obtained in this study. However, the dynamic nature of the peptides covering the hinge loop and cantilever loop which anchor the flaps were measured. As mentioned previously, the hinge region of the C-SA PR undergoes reduced rates of amide hydrogen exchange in

comparison to the subtype B PR. For peptide 64–76 (covering part of the cantilever region), a slight reduction in fast exchanging amide protons (meaning more protected amide protons) is observed in the C-SA PR (Table 3). These results indicate that the network of interactions in the hydrophobic core of the C-SA PR may be altered. The hinge, cantilever and fulcrum regions which are important for flap opening display increased stability in the C-SA PR and may stabilise the wide range of fully-open conformers exhibited by the C-SA PR.

Preferential targeting of the C-terminus to prevent dimerisation

The N-terminus (positions 1–5) and C-terminus (positions 95–99) of HIV-1 PRs are pivotal in dimer stability; structure-based thermodynamic analyses showed that the N- and C-termini antiparallel β-sheet contributes ~ 75% to the total Gibbs energy [35]. Thermodynamic and structural analyses of recombinant HIV-1 PRs with a short segment of the transframe region tethered to the N-terminus of the PR, displayed significantly reduced thermal stability compared to wild-type HIV-1 PR and crystallographic data showed evidence of random coils at the N-terminus [53]. Recombinant HIV-1 PR with a short segment of reverse transcriptase tethered to the C-terminus showed no significant reduction in thermal stability, implying increased stability of the C-terminus relative to the N-terminus [53]. A similar result is seen in the HDX-MS experiment for both the subtype B and C-SA PRs (Table 4).

Due to cleavage of the PRs by pepsin and binding of the resultant peptides to the C18 reversed-phase column [27], the HDX data for peptide 1–5 reports on the peptide-amide linkages at residue positions 3–5 and data for peptide 95–99 reports on the peptide-amide linkages at residue

positions 97–99. The amide groups of L97 and F99 participate in hydrogen bonds with the N-terminus of the adjacent subunit; meaning, interactions need to be broken for HDX to proceed. After 1 minute of incubation in deuterium oxide, both PRs display low levels of deuterium incorporation at the N- and C-termini (< 0.5 deuterons). After 1 hour of incubation, the differences are more pronounced; more than 1.6 amide protons are exchanged in peptide 1–5 (reporting on I3, T4 and L5) and less than 1 proton is exchanged in peptide 95–99 (reporting on L97, N98 and F99) in both PRs. Therefore, only a fraction of the PRs in solution exchange an amide proton after 1 hour of incubation at their final 3 residues (Table 4). Rates of HDX in this region reflect the internal motions of the β-sheet formed by the N- and the C-termini. Therefore, the N-terminus of both PRs is seemingly more dynamic than the C-terminus. Targeting the residues at positions 1–5 and 95–99 in HIV-1 and HIV-2 PRs with complementary synthetic peptides has shown inhibitory potential [54]. Inhibitors targeted to the C-terminus alone displayed greater potency than inhibitors targeted to the N-terminus or both ends of the PR [54]. Drug resistance to currently available PIs are attributed to active site and flap-hinge region mutations. Dimerisation inhibitors attempt to solve this complication by targeting an alternative site.

Using overlapping peptides, deuterium incorporation of other C-terminal amide protons (C95 and T96) may be determined. Calculating the difference in deuterium incorporation between peptides 77–99 and 77–94 allows for HDX measurement at residue positions 95–99. Subtracting the known incorporation at residue positions 97–99 from the calculated 95–99 allows for resolution of C95 and T96 (Table 4). Differences in deuterium incorporation at peptide 77–94 between the PRs is evident from 5 minutes of incubation and is most pronounced after 1 hour of

incubation. Therefore, detailed analysis of deuterium incorporation was performed using data resulting from 1 hour of incubation. Analysis of exchange at C95 and T96 shows that ~ 0.6 amide protons have exchanged in the subtype B PR, whereas, these amide protons are fully exchanged (2 amide protons) in the C-SA PR. This suggests that the C-terminal residues, C95 and T96, are more dynamic in the C-SA PR. Therefore, these residues may contribute less to the stability of the N- and C-termini β-sheet than other C-terminal residues. Mutation of the ultimate amino acid in the HIV-1 PR sequence, phenylalanine, to an alanine was shown to reduce the β-sheet content of the PR and disrupt dimerisation which results in complete loss of activity [55]. Contacts made by F99 with neighbouring residues are shown in Figure 7. F99 is central to the large network of interactions present in the 'lock-and-key' motif (Fig. 7). Therefore, the importance of the C-terminus for correct folding of the PR and its increased stability in comparison to the N-terminus, make the C-terminus, particularly F99, a preferred target for antiviral therapy.

This study represents the first known report of amide hydrogen/deuterium exchange-mass spectrometry performed on HIV PR. The C-SA PR displays weaker binding to clinically used PIs than the subtype B PR. The C-SA PR has an increased propensity to exist in a fully-open conformation (Fig. 2C) [8] and displays a wider range of open conformers than the subtype B PR [56]. Altered interactions in the hydrophobic core of the C-SA PR likely contribute to the difference in flap ensembles. Polymorphisms in the C-SA PR alter the complex network of interactions in the hydrophobic core thus favouring the formation and increasing the stability of the fully-open conformer in comparison to the subtype B PR. Importantly, polymorphisms in the C-SA PR conserve hydrophobic residues but may alter the anchorage of flaps through altered

hydrophobic sliding. The reduced propensity to form the E35-R57 salt bridge in the C-SA PR in combination with altered dynamics at the hinge, cantilever and fulcrum regions contribute to the altered anchorage of the flaps and a shift in the equilibrium between semi-open (Fig. 2B) and fully-open conformers (Fig. 2C). Because the C-SA PR displays reduced drug susceptibility and similar catalytic efficiency in comparison to the subtype B PR, the increased stability of the fulcrum, hinge and cantilever regions leading to the shift in flap conformers preferentially affects inhibitor binding over substrate binding. The increase in substrate turnover (Table 1) for the C-SA PR may be explained by its increased proportion of the fully-open conformer when compared to the subtype B PR. Substrate processing is only allowed following full opening of the PR flaps. This highlights the need to further characterise the fully-open conformer which is a viable target for drug design because this conformer is a prerequisite for PR function. Furthermore, the N-terminus of both PRs appears to be more dynamic than the C-terminus of the PRs, thereby, substantiating the targeting of the C-terminus, particularly F99, for the development of dimerisation inhibitors.

Materials and Methods

Expression and purification

Escherichia coli BL21 (DE3) pLysS cells were transformed with the pET-11b plasmid encoding either the subtype B or C-SA PR insert. The consensus C-SA PR (wild-type) sequence data were obtained from Prof Lynn Morris (AIDS Virus Research Unit, National Institute of Communicable Diseases, South Africa). A Q7K point mutation was introduced in the PR

sequences by site-directed mutagenesis using the QuikChange® method (Stratagene, La Jolla, CA, USA) to reduce PR autocatalysis [57].

Expression and purification of the PRs were performed as previously described [6]. Briefly, transformed cells were induced to over-express the PRs as inclusion bodies using isopropyl-β-D-thiogalactopyranoside [58]. Inclusion bodies were solubilised using 8 M urea and refolded by dialysis into a buffer containing 10 mM sodium acetate, 2 mM dithiothreitol and 0.02% sodium azide (pH 5.0). The PRs were purified using CM-Sepharose cation exchange resin, with a 0–1 M NaCl gradient elution. The PRs were dialysed into 10 mM sodium acetate, 2 mM dithiothreitol and 0.02% sodium azide (pH 5.0) for storage and were resolved on a 10% tricine-SDS polyacrylamide gel to evaluate the purity of the proteins [59, 60].

Active site determination

The percentage of active enzyme in the PR preparations was investigated using isothermal titration calorimetry (ITC). Active site determinations were performed at 20 °C using a VP-ITC microcalorimeter (MicroCal Inc.). Acetyl-pepstatin (Bachem), a well-known aspartyl protease inhibitor, was titrated into the sample cell containing the PR. Saturation of the PRs was achieved during titrations using 200 µM acetyl-pepstatin, 10.59 µM subtype B PR and 15.50 µM C-SA PR. Calorimetric data were fit using the Origin 7.0 software package. The percentage of active sites was determined from the stoichiometry value.

Steady-state and inhibition kinetics

Kinetic parameters, K_M, k_{cat} and k_{cat}/K_M, were determined in separate experiments following hydrolysis of the HIV-1 PR fluorogenic substrate (Abz-Arg-Val-Nle-Phe(NO$_2$)-Glu-Ala-Nle-NH$_2$). An active enzyme concentration of 30–50 nM and a substrate concentration ranging 5–200 µM and 3–10 µM were used to determine K_M and k_{cat}/K_M, respectively. Varying amounts of enzyme (1–10 pmole) were used with a constant substrate concentration (50 µM) for k_{cat} determination.

The FDA-approved drugs competitively inhibit HIV PR with dissociation constants within the nM–pM range. Inhibition constants (K_i) of atazanavir, darunavir and ritonavir (NIH AIDS Reagent Program) were determined using the following equation for tight-binding inhibitors [61]:

$$K_i = \left(IC_{50} - \frac{[E]}{2} \right) \Big/ \left(\frac{[S]}{K_M} + 1 \right) \qquad (1)$$

E is the active enzyme concentration (50 nM), [S] is the substrate concentration (50 µM) and IC_{50} is the concentration of inhibitor which results in half-maximal activity of the PR. IC_{50} values were determined using inhibitor concentrations ranging 0–200 µM. For all kinetic measurements, an excitation wavelength of 337 nm and emission wavelength of 425 nm were used for the 30 second measurements during steady-state. Activity assays were performed in 50 mM sodium

acetate and 1 M sodium chloride (pH 5.0) at 20 °C. For k_{cat}/K_M determinations, 0.1 M sodium

chloride was used to obtain measurable rates of substrate cleavage. A final dimethyl sulfoxide

concentration of 2% was used during IC_{50} determinations to ensure inhibitor solubility. All

kinetic experiments were performed in triplicate using a Jasco FP-6300 spectrofluorometer and

the data were fit using SigmaPlot (version 11.0). Kinetic parameters were used to calculate the

relative vitality of the C-SA PR in the presence of PIs and compared to the subtype B PR

according to the following equation [62]:

$$Relative\ vitality = \left(K_i * \frac{k_{cat}}{K_M}\right) C\text{–}SA \Bigg/ \left(K_i * \frac{k_{cat}}{K_M}\right) subtype\ B \qquad (2)$$

Displacement titration ITC

Binding thermodynamics of atazanavir, darunavir and ritonavir binding to the C-SA PR were

determined using a displacement titration method [63]. PIs were titrated into the sample cell

containing the C-SA PR pre-bound to acetyl-pepstatin (200 µM final concentration).

Thermodynamic parameters of C-SA PR binding were measured following titration of 150 µM

atazanavir into 13.89 µM PR, 100 µM darunavir into 13.78 µM PR and 100 µM ritonavir into

12.66 µM PR. For subtype B PR binding, 100 µM of ritonavir was titrated into 11.86 µM PR.

Displacement titrations were performed in 10 mM sodium acetate and 2% dimethyl sulfoxide

(pH 5.0) at 20 °C using a VP-ITC microcalorimeter (MicroCal Inc.).

Hydrogen/deuterium exchange-mass spectrometry (HDX-MS)

The dynamics of the subtype B and C-SA PRs were investigated by HDX-MS as previously described [64-66]. For exchange experiments, 10 µl of PR stock solutions (1.3 mg/ml) were diluted 4-fold in 100% D_2O at 20 °C. At appropriate time intervals, hydrogen-deuterium exchange (HDX) was minimised with a 1:1 dilution in 4 M guanidine hydrochloride and 1% formic acid (pH 2.3) and incubated on ice for 30 seconds. The PRs were fragmented with the addition of 13 µl of pepsin (1 µg/µl) and kept on ice for 5 minutes, resulting in a 1:1 (pepsin:PR) ratio. Following fragmentation, 50 µl of the PR solution was injected onto an Aeris PEPTIDE 3.6 µm XB-C18 reversed-phase column (Phenomenex); submerged in ice; coupled to an AB SCIEX QSTAR® Elite mass spectrometer via a 6-port switching valve. Peptides were eluted at 300 µl/min from the reversed-phase column with a 5–95% B gradient in 10 minutes (A: 0.1 % formic acid; B: acetonitrile/0.1 % formic acid). Initial peptide identification was carried out using Collision Induced Dissociation (CID) in Information Dependent Acquisition (IDA) mode. All subsequent samples, including the non-deuterated and fully-deuterated, were analysed in MS mode. PR stock solutions were maintained in the standard storage buffer comprising 10 mM sodium acetate, 2 mM dithiothreitol and 0.02% sodium azide (pH 5.0). All experiments were performed in duplicate and data collection for each PR subtype was completed within 2 days.

Undeuterated controls were analysed using PEAKS 6 (Bioinformatics Solutions Inc., Ontario, Canada; http://www.bioinfor.com) to determine the peptide coverage and peptide pool for HDX analysis. The percentage back-exchange that occurs once the HDX reaction is quenched (during fragmentation, peptide separation and mass spectrometry analysis) was determined using a fully-

deuterated control. On average, the percentage back-exchange for subtype B PR experiments was 20.75% and 22.30% for C-SA PR experiments. HDX analysis was performed using HDExaminer 1.2 (Sierra Analytics Inc., California, USA; http://www.masspec.com). Manual quality control was performed on analysed peptides (high and medium confidence peptides only) to ensure that theoretical and experimental isotopic profiles matched. Deuterium incorporation versus exchange time was fit using SigmaPlot (version 11.0) according to the following equation [65]:

$$D = N - \sum_i^N \exp(-k_i t) \qquad (3)$$

were D is the deuterium content of a peptide, N is the number of peptide amide protons, k_i is the exchange rate constant for each peptide amide proton and t is the time allowed for isotopic exchange. The total number of exponential terms (either one or two in this study) was selected based on the goodness of each fit. Structural images were constructed using the PyMOL Molecular Graphics System (Schrödinger LLC., Portland, USA; http://www.schrodinger.com).

Acknowledgements

Author contributions

PN performed all experimental work, analysed the data and wrote the manuscript. SS assisted with the HDX-MS experiments. HWD assisted in experimental design and revision of manuscript. YS supervised the project and assisted in data analysis and interpretation.

References

1. The Joint Nations Program on HIV/AIDS (UNAIDS). (2013) *Global report: UNAIDS report on the global AIDS epidemic 2013*, UNAIDS, Geneva, Switzerland.

2. Walker, PR, Pybus, OG, Rambaut, A & Holmes, EC (2005) Comparative population dynamics of HIV-1 subtypes B and C: subtype-specific differences in patterns of epidemic growth, *Infect Genet Evol.* **5**, 199-208.

3. World Health Organisation (WHO). (2013) *The use of antiretroviral drugs for treating and preventing HIV infection*, Geneva, Switzerland: WHO Press.

4. Naicker, P & Sayed, Y (2014) Non-B HIV-1 subtypes in sub-Saharan Africa: impact of subtype on protease inhibitor efficacy *Biol Chem.*, In Press.

5. Velazquez-Campoy, A, Vega, S, Fleming, E, Bacha, U, Sayed, Y & Dirr, HW (2003) Protease Inhibition in African Subtypes of HIV-1, *AIDS Reviews.* **5**, 165-171.

6. Ahmed, SM, Kruger, HG, Govender, T, Maguire, GEM, Sayed, Y, Ibrahim, MAA, Naicker, P & Soliman, MES (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.

7. Prabu-Jeyabalan, M, Nalivaika, E & Schiffer, CA (2000) How does a symmetric dimer recognize an asymmetric substrate? A substrate complex of HIV-1 protease, *J Mol Biol.* **301**, 1207-1220.

8. Kear, JL, Blackburn, ME, Veloro, AM, Dunn, BM & Fanucci, GE (2009) Subtype polymorphisms among HIV-1 protease variants confer altered flap conformations and flexibility, *J Am Chem Soc.* **131**, 14650-1.

9. Scott, WR & Schiffer, CA (2000) Curling of flap tips in HIV-1 protease as a mechanism for substrate entry and tolerance of drug resistance, *Structure.* **8**, 1259-65.

10. Hornak, V, Okur, A, Rizzo, RC & Simmerling, C (2006) HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations, *Proc Natl Acad Sci USA.* **103**, 915-20.

11. Freedberg, DI, Ishima, R, Jacob, J, Wang, YX, Kustanovich, I, Louis, JM & Torchia, DA (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**, 221-32.

12. Ishima, R, Freedberg, DI, Wang, YX, Louis, JM & Torchia, DA (1999) Flap opening and dimer-interface flexibility in the free and inhibitor-bound HIV protease, and their implications for function, *Structure.* **7**, 1047-55.

13. Heaslet, H, Rosenfeld, R, Giffin, M, Lin, YC, Tam, K, Torbett, BE, Elder, JH, McRee, DE & Stout, CD (2007) Conformational flexibility in the flap domains of ligand-free HIV protease, *Acta Crystallogr D.* **63**, 866-875.

14. Gustchina, A & Weber, IT (1990) Comparison of inhibitor binding in HIV-1 protease and in non-viral aspartic proteases: the role of the flap, *FEBS letters.* **269**, 269-72.

15. Rick, SW, Erickson, JW & Burt, SK (1998) Reaction path and free energy calculations of the transition between alternate conformations of HIV-1 protease, *Proteins.* **32**, 7-16.

16. Nicholson, LK, Yamazaki, T, Torchia, DA, Grzesiek, S, Bax, A, Stahl, SJ, Kaufman, JD, Wingfield, PT, Lam, PY, Jadhav, PK & et al. (1995) Flexibility and function in HIV-1 protease, *Nat Struct Biol.* **2**, 274-80.

17. Galiano, L, Bonora, M & Fanucci, GE (2007) Interflap distances in HIV-1 protease determined by pulsed EPR measurements, *J Am Chem Soc.* **129**, 11004-5.

18. Huang, X, de Vera, IM, Veloro, AM, Blackburn, ME, Kear, JL, Carter, JD, Rocca, JR, Simmerling, C, Dunn, BM & Fanucci, GE (2012) Inhibitor-induced conformational shifts and ligand-exchange dynamics for HIV-1 protease measured by pulsed EPR and NMR spectroscopy, *J Phys Chem B.* **116**, 14235-44.

19. King, NM, Prabu-Jeyabalan, M, Bandaranayake, RM, Nalam, MNL, Nalivaika, EA, Ozen, A, Haliloglu, T, Yilmaz, NK & Schiffer, CA (2012) Extreme Entropy-Enthalpy Compensation in a Drug-Resistant Variant of HIV-1 Protease, *ACS Chem Biol.* **7**, 1536-1546.

20. Muzammil, S, Armstrong, AA, Kang, LW, Jakalian, A, Bonneau, PR, Schmelmer, V, Amzel, LM & Freire, E (2007) Unique thermodynamic response of tipranavir to human immunodeficiency virus type 1 protease drug resistance mutations, *J Virol.* **81**, 5144-54.

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

22. Szeltner, Z & Polgar, L (1996) Rate-determining steps in HIV-1 protease catalysis. The hydrolysis of the most specific substrate, *J Biol Chem.* **271**, 32180-4.

23. Cheng, YS, Yin, FH, Foundling, S, Blomstrom, D & Kettner, CA (1990) Stability and activity of human immunodeficiency virus protease: comparison of the natural dimer with a homologous, single-chain tethered dimer, *Proc Natl Acad Sci USA*. **87**, 9660-4.

24. Berger, A & Linderstrom-Lang, K (1957) Deuterium exchange of poly-DL-alanine in aqueous solution, *Arch Biochem Biophys*. **69**, 106-18.

25. Eigen, M (1964) Proton transfer, acid-base catalysis, and enzymatic hydrolysis, *Angew Chem Int Ed Eng*. **3**, 1-19.

26. Englander, SW & Kallenbach, NR (1983) Hydrogen exchange and structural dynamics of proteins and nucleic acids, *Q Rev Biophys*. **16**, 521-655.

27. Bai, YW, Milne, JS, Mayne, L & Englander, SW (1993) Primary structure effects on peptide group hydrogen exchange, *Proteins*. **17**, 75-86.

28. Kim, KS, Fuchs, JA & Woodward, CK (1993) Hydrogen exchange identifies native-state motional domains important in protein folding, *Biochemistry*. **32**, 9600-9608.

29. Elber, R & Karplus, M (1987) Multiple conformational states of proteins: a molecular dynamics analysis of myoglobin, *Science*. **235**, 318-21.

30. Engen, JR, Gmeiner, WH, Smithgall, TE & Smith, DL (1999) Hydrogen exchange shows peptide binding stabilizes motions in Hck SH2, *Biochemistry*. **38**, 8926-35.

31. Roder, H, Wagner, G & Wuthrich, K (1985) Individual amide proton exchange rates in thermally unfolded basic pancreatic trypsin inhibitor, *Biochemistry*. **24**, 7407-11.

32. Woodward, C, Simon, I & Tuchsen, E (1982) Hydrogen exchange and the dynamic structure of proteins, *Mol Cell Biochem*. **48**, 135-60.

33. Miranker, A, Robinson, CV, Radford, SE, Aplin, RT & Dobson, CM (1993) Detection of transient protein folding populations by mass spectrometry, *Science*. **262**, 896-900.

34. Maity, H, Lim, WK, Rumbley, JN & Englander, SW (2003) Protein hydrogen exchange mechanism: local fluctuations, *Protein Sci.* **12**, 153-60.

35. Todd, MJ, Semo, N & Freire, E (1998) The structural stability of the HIV-1 protease, *J Mol Biol.* **283**, 475-488.

36. Fersht, AR (1972) Conformational equilibria in -and -chymotrypsin. The energetics and importance of the salt bridge, *J Mol Biol.* **64**, 497-509.

37. Perutz, MF (1978) Electrostatic effects in proteins, *Science.* **201**, 1187-91.

38. Foulkes-Murzycki, JE, Scott, WR & Schiffer, CA (2007) Hydrophobic sliding: a possible mechanism for drug resistance in human immunodeficiency virus type 1 protease, *Structure.* **15**, 225-33.

39. Mittal, S, Bandaranayake, RM, King, NM, Prabu-Jeyabalan, M, Nalam, MN, Nalivaika, EA, Yilmaz, NK & Schiffer, CA (2013) Structural and thermodynamic basis of amprenavir/darunavir and atazanavir resistance in HIV-1 protease with mutations at residue 50, *J Virol.* **87**, 4176-84.

40. Milne, JS, Mayne, L, Roder, H, Wand, AJ & Englander, SW (1998) Determinants of protein hydrogen exchange studied in equine cytochrome c, *Protein Sci.* **7**, 739-45.

41. Rhee, SY, Gonzales, MJ, Kantor, R, Betts, BJ, Ravela, J & Shafer, RW (2003) Human immunodeficiency virus reverse transcriptase and protease sequence database, *Nucleic Acids Res.* **31**, 298-303.

42. Harrigan, PR, Hertogs, K, Verbiest, W, Pauwels, R, Larder, B, Kemp, S, Bloor, S, Yip, B, Hogg, R, Alexander, C & Montaner, JS (1999) Baseline HIV drug resistance profile predicts response to ritonavir-saquinavir protease inhibitor therapy in a community setting, *AIDS.* **13**, 1863-71.

43. King, MS, Rode, R, Cohen-Codar, I, Calvez, V, Marcelin, AG, Hanna, GJ & Kempf, DJ (2007) Predictive genotypic algorithm for virologic response to lopinavir-ritonavir in protease inhibitor-experienced patients, *Antimicrob Agents Chemother.* **51**, 3067-74.

44. Marcelin, AG, Masquelier, B, Descamps, D, Izopet, J, Charpentier, C, Alloui, C, Bouvier-Alias, M, Signori-Schmuck, A, Montes, B, Chaix, ML, Amiel, C, Santos, GD, Ruffault, A, Barin, F, Peytavin, G, Lavignon, M, Flandre, P & Calvez, V (2008) Tipranavir-ritonavir genotypic resistance score in protease inhibitor-experienced patients, *Antimicrob Agents Chemother.* **52**, 3237-43.

45. Naeger, LK & Struble, KA (2006) Effect of baseline protease genotype and phenotype on HIV response to atazanavir/ritonavir in treatment-experienced patients, *AIDS.* **20**, 847-53.

46. Pellegrin, I, Breilh, D, Coureau, G, Boucher, S, Neau, D, Merel, P, Lacoste, D, Fleury, H, Saux, MC, Pellegrin, JL, Lazaro, E, Dabis, F & Thiebaut, R (2007) Interpretation of genotype and pharmacokinetics for resistance to fosamprenavir-ritonavir-based regimens in antiretroviral-experienced patients, *Antimicrob Agents Chemother.* **51**, 1473-80.

47. Pellegrin, I, Wittkop, L, Joubert, LM, Neau, D, Bollens, D, Bonarek, M, Girard, PM, Fleury, H, Winters, B, Saux, MC, Pellegrin, JL, Thiebaut, R & Breilh, D (2008) Virological response to darunavir/ritonavir-based regimens in antiretroviral-experienced patients (PREDIZISTA study), *Antivir Ther.* **13**, 271-9.

48. Shulman, N, Zolopa, A, Havlir, D, Hsu, A, Renz, C, Boller, S, Jiang, P, Rode, R, Gallant, J, Race, E, Kempf, DJ & Sun, E (2002) Virtual inhibitory quotient predicts response to ritonavir boosting of indinavir-based therapy in human immunodeficiency virus-infected patients with ongoing viremia, *Antimicrob Agents Chemother.* **46**, 3907-16.

49. Zolopa, AR, Shafer, RW, Warford, A, Montoya, JG, Hsu, P, Katzenstein, D, Merigan, TC & Efron, B (1999) HIV-1 genotypic resistance patterns predict response to saquinavir-ritonavir therapy in patients in whom previous protease inhibitor therapy had failed, *Ann Intern Med.* **131**, 813-21.

50. de Meyer, S, Vangeneugden, T, van Baelen, B, de Paepe, E, van Marck, H, Picchio, G, Lefebvre, E & de Bethune, MP (2008) Resistance profile of darunavir: combined 24-week results from the POWER trials, *AIDS Res Hum Retroviruses.* **24**, 379-88.

51. Mittal, S, Cai, Y, Nalam, MN, Bolon, DN & Schiffer, CA (2012) Hydrophobic core flexibility modulates enzyme activity in HIV-1 protease, *J Am Chem Soc.* **134**, 4163-8.

52. Bandaranayake, RM, Kolli, M, King, NM, Nalivaika, EA, Heroux, A, Kakizawa, J, Sugiura, W & Schiffer, CA (2010) The effect of clade-specific sequence polymorphisms on HIV-1 protease activity and inhibitor resistance pathways, *J Virol.* **84**, 9995-10003.

53. Agniswamy, J, Sayer, JM, Weber, IT & Louis, JM (2012) Terminal Interface Conformations Modulate Dimer Stability Prior to Amino Terminal Autoprocessing of HIV-1 Protease, *Biochemistry.* **51**, 1041-1050.

54. Babe, LM, Rose, J & Craik, CS (1992) Sythetic interface peptides alter dimeric assembly of the HIV-1 and HIV-2 proteases, *Protein Sci.* **1**, 1244-1253.

55. Naicker, P, Seele, P, Dirr, HW & Sayed, Y (2013) F99 is critical for dimerization and activation of South African HIV-1 subtype C protease, *Protein J.* **32**, 560-567.

56. Naicker, P, Achilonu, I, Fanucchi, S, Fernandes, M, Ibrahim, MAA, Dirr, HW, Soliman, MES & 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.

57. Mildner, AM, Rothrock, DJ, Leone, JW, Bannow, CA, Lull, JM, Reardon, IM, Sarcich, JL, Howe, WJ, Tomich, CC, Smith, CW, Heinrikson, RL & Tomasselli, AG (1994) The HIV-1 Protease as Enzyme and Substrate: Mutagenesis of Autolysis Sites and Generation of a Stable Mutant with Retained Kinetic Properties, *Biochemistry.* **33**, 9405-9413.

58. Ido, E, Han, HP, Kezdy, FJ & 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.

59. Laemmli, UK (1970) Cleavage of structural proteins during assembly of head of bacteriophage-T4, *Nature.* **227**, 680-685.

60. Schagger, H (2006) Tricine-SDS-PAGE, *Nat Protoc.* **1**, 16-22.

61. Copeland, RA, Lombardo, D, Giannaras, J & Decicco, CP (1995) Estimating K_I values for tight-binding inhibitors from dose-response plots, *Bioorg Med Chem Lett.* **5**, 1947-1952.

62. Gulnik, SV, Suvorov, LI, Liu, BS, Yu, B, Anderson, B, Mitsuya, H & Erickson, JW (1995) Kinetic characterization and cross-resistance patterns of HIV-1 protease mutants selected under drug pressure, *Biochemistry.* **34**, 9282-9287.

63. 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.

64. Stoychev, SH, Nathaniel, C, Fanucchi, S, Brock, M, Li, S, Asmus, K, Woods, VL & Dirr, HW (2009) Structural Dynamics of Soluble Chloride Intracellular Channel Protein CLIC1 Examined by Amide Hydrogen-Deuterium Exchange Mass Spectrometry, *Biochemistry.* **48**, 8413-8421.

65. Zhang, ZQ & Smith, DL (1993) Determination of amide hydrogen exchange by mass spectrometry: A new tool for protein structure elucidation, *Protein Sci.* **2**, 522-531.

66. Balchin, D, Stoychev, SH & Dirr, HW (2013) S-Nitrosation destabilizes glutathione transferase P1-1, *Biochemistry.* **52**, 9394-402.

67. Higgins, DG & Sharp, PM (1988) CLUSTAL: a package for performing multiple sequence alignment on a microcomputer, *Gene.* **73**, 237-244.

68. Larkin, MA, Blackshields, G, Brown, NP, Chenna, R, McGettigan, PA, McWilliam, H, Valentin, F, Wallace, IM, Wilm, A, Lopez, R, Thompson, JD, Gibson, TJ & Higgins, DG (2007) Clustal W and clustal X version 2.0, *Bioinformatics.* **23**, 2947-2948.

Table 1: Enzymatic parameters of the subtype B and C-SA PRs.

HIV-1 Subtype	K_M (µM)	k_{cat} (s^{-1})	k_{cat}/K_M (M^{-1}s^{-1})
B	38.7±6.0	9.0±0.2	$3.2 \times 10^{-7} \pm 0.1 \times 10^{-7}$
C-SA	29.4±4.4	20.6±0.8	$4.6 \times 10^{-7} \pm 0.5 \times 10^{-7}$

Table 2: Comparison of inhibitor binding between the subtype B and C-SA PRs.

HIV-1 Subtype	ΔH (kcal/mol)	$-T\Delta S$ (kcal/mol)	ΔG (kcal/mol)	K_d (nM)	K_i (nM)	$K_i/K_{i(B)}$	Relative vitality
			Acetyl-pepstatin				
B	11.5	-20.2	-8.7	300			
C-SA	9.1	-17.9	-8.8	252			
			Ritonavir				
B	3.0	-16.8	-13.8	0.049	6.47±2.17	1	1
C-SA	2.0	-15.5	-13.5	0.083	18.25±1.72	2.82	4.05
			Atazanavir				
B	-4.2 [a]	-10.1 [a]	-14.3 [a]	0.035 [a]	10.96±4.29	1	1
C-SA	-4.8	-9.2	-14.0	0.036	5.53±1.88	0.50	0.72
			Darunavir				
B	-12.1 [b]	-3.1 [b]	-15.0 [b]	0.005 [b]	0.49±0.20	1	1
C-SA	-19.1	4.5	-14.6	0.012	3.30±1.42	6.73	9.67

[a] Values cited from reference [20].
[b] Values cited from reference [19].

Table 3: Number of very fast, fast, intermediate and slow exchanging amide protons in peptides derived from the subtype B and C-SA PRs.

HIV-1 Subtype	Peptide	Very fast [a] (ND)	Fast $(12>k>0.1$ min$^{-1})$	Intermediate $(0.1>k>0.01$ min$^{-1})$	Slow [b] (ND)
B	13–23	1.7	0.9	2.5	3.9
C-SA		1.4	-	3.4	4.2
B	34–53	3.5	3.8	3.7	5.0
C-SA		1.8	4.0	5.4	4.8
B	47–53	1.3	1.5	0.7	1.5
C-SA		0.8	1.2	1.2	1.8
B	64–76	1.2	0.4	1.7	7.7
C-SA		1.1	-	2.3	7.6
B	64–90	1.2	1.8	3.3	16.7
C-SA		1.2	-	5.2	16.6

[a] Determined by subtracting the number of fast and intermediate exchanging amide protons from the total exchanged at 1 hour of incubation.

[b] Determined by subtracting the total protons exchanged at 1 hour of incubation from the total amide protons available in the peptide.

Rates are displayed in parenthesis. ND = Not determined.

Table 4: Number of amide protons exchanged at the N- and C-termini of the subtype B and C-SA PRs at 1 hour of deuterium oxide incubation.

HIV-1 Subtype	Peptide	Residues analysed [a]	Number of protons exchanged
B	1–5	3–5	1.74±0.04
C-SA			1.60±0.04
B	77–94	79–94	5.84±0.31
C-SA			3.83±0.05
B	77–99	79–99	7.17±0.22
C-SA			7.79±0.05
B	Difference of 77–99 and 77–94	95–99	1.33±0.09
C-SA			2.97±0.08
B	95–99	97–99	0.77±0.05
C-SA			0.91±0.03

[a] Residues analysed exclude the first two amino acids of the peptide.

Figure 1: Polymorphic sites of the subtype B and C-SA PRs. The flexible flaps of PR (residues at positions 46–54) are coloured cyan. The fulcrum (residues at positions 10–23), hinge (residues at positions 35–42 and 57–61) and cantilever (residues at positions 62–78) regions are implicated in flap opening. Locations of the eight polymorphic amino acids in each monomer are shown as spheres. The complete sequence alignment is displayed below the structure. The corresponding secondary structure is shown below the sequence. Arrows represent β-strands and waves represent α-helices. The sequence alignment was performed using Clustal X 2.0 [67, 68]. PDB ID: 3U71 [6].

Figure 2: Overview of flap conformers displayed by the HIV-1 PR. The closed (A), semi-open (B) and fully-open (C) conformers are in dynamic equilibrium. Above the respective three-dimensional structures are the top views of the conformers. The semi-open flap conformation is the most prevalent conformer in the absence of inhibitor. Crystal structure data is unavailable for a fully-open conformer which is characterised by upward and outward displacement of the flaps. Representation of the flap positioning of the fully-open conformer is based on MD simulation models [10].

Figure 3: Heat maps for amide hydrogen/deuterium exchange of the subtype B and C-SA PRs. The heat maps depict the percentage deuteration for all time points measured (10 s, 20 s, 30 s, 1 min, 5 min, 20 min and 1 hr) at different regions of the PRs corresponding to the sequence above. Peptides which were detected and used for data processing are shown above the sequences. Positions of the polymorphisms are shown in red. The heat map shows the averaged deuteration for both monomers in the homodimeric PRs.

Figure 4: Amide hydrogen/deuterium exchange kinetics of different regions of the PRs. (A) The corresponding positions of measured peptides are coloured blue on the structure. Graphs show the deuterium incorporation plots for: peptide 13–23, peptide 34–53 and peptide 64–90. (B) Deuterium incorporation plot of peptide 47–53 covering the flap tips of the PRs.

Figure 5: Interactions localised around residue position 36 of the hinge region. In the subtype B PR (A), M36 seemingly provides anchorage for E35 and stabilises the salt bridge between E35 and R57. I36 in the C-SA PR (B) is implicated in less hinge region interactions and disruption of the E35-R57 salt bridge. PDB ID: 2PC0 (A) [13], 3U71 (B) [6].

Figure 6: Hydrophobic core of HIV-1 PR in stereo view. Residues comprising the hydrophobic core are shown as sticks. Flap tips (residue positions 46–54) are coloured cyan. The hinge regions (residue positions 35–42 and 57–61) coloured pink and the cantilever region (residue positions 62–78) coloured blue anchor the flap tips.

Figure 7: Interactions made by F99. F99 is central to the illustrated 'lock-and-key' motif and is essential for dimerization.

```
           **********:**:***:*****************:****:*****************************:*****************:***:******
C-SA       PQITLWKRPLVSIKVGGQIKEALLDTGADDTVLEEINLPGKWKPKMIGGIGGFIKVRQYDQILIEICGKKAIGTVLVGPTPVNIIGRNMLTQLGCTLNF      99
Subtype-B  PQITLWKRPLVTIKIGGQLKEALLDTGADDTVLEEMNLPGRWKPKMIGGIGGFIKVRQYDQILIEICGHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNF      99
           1.......10........20........30........40........50........60........70........80........90.........
```
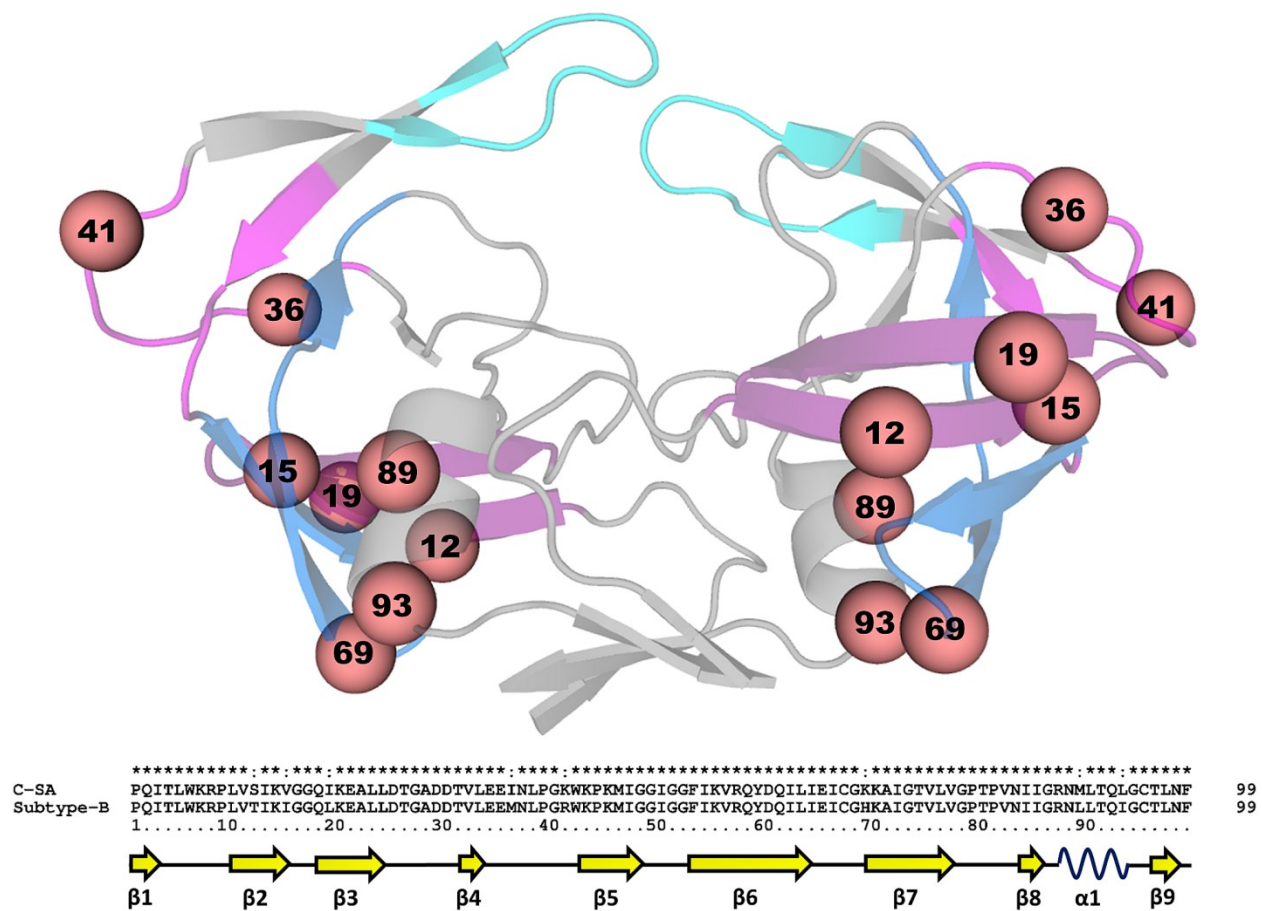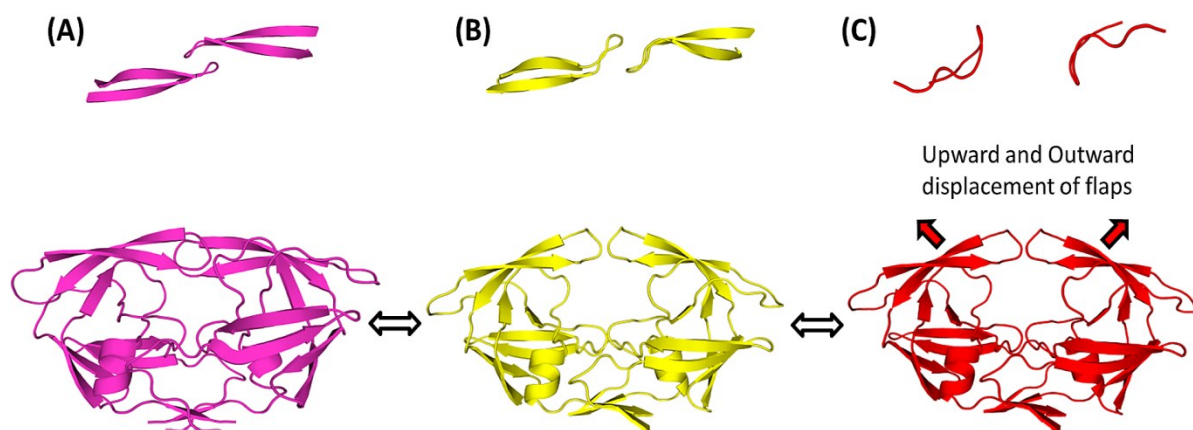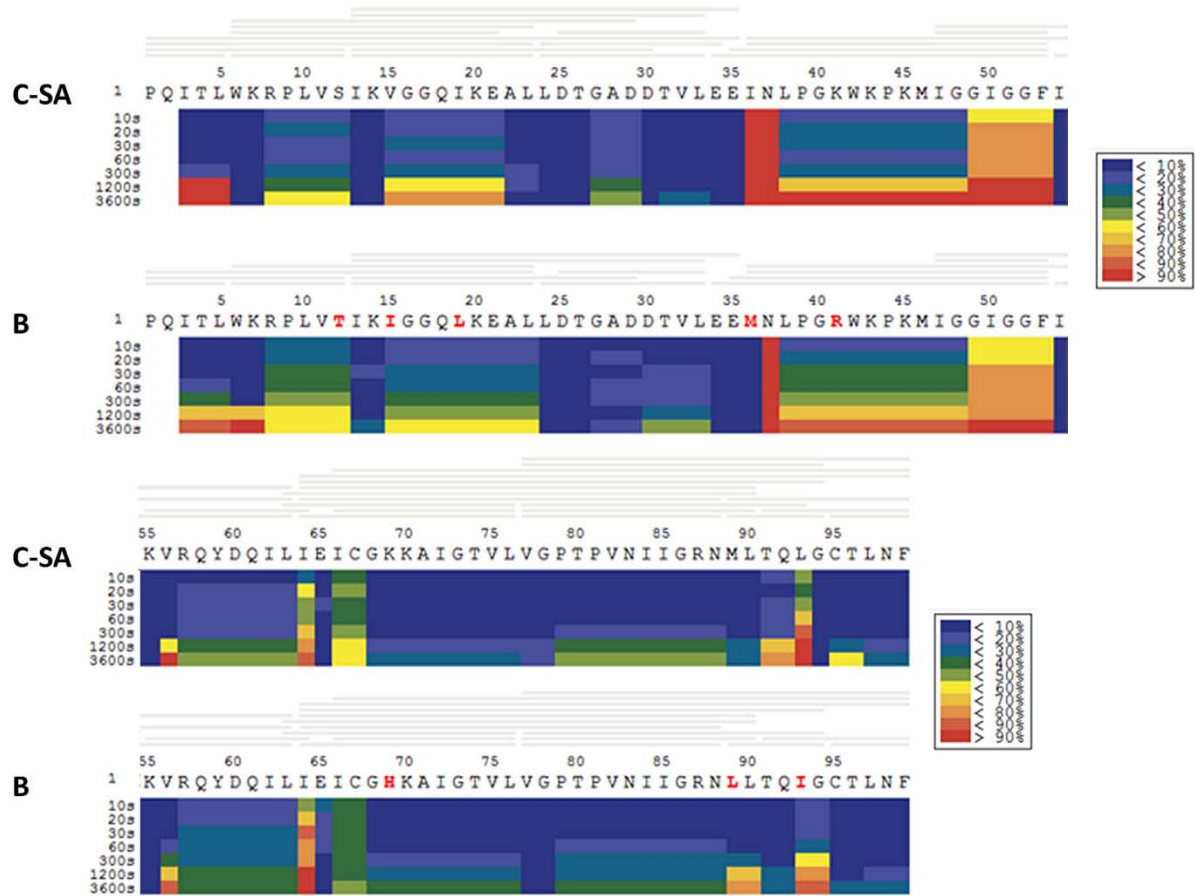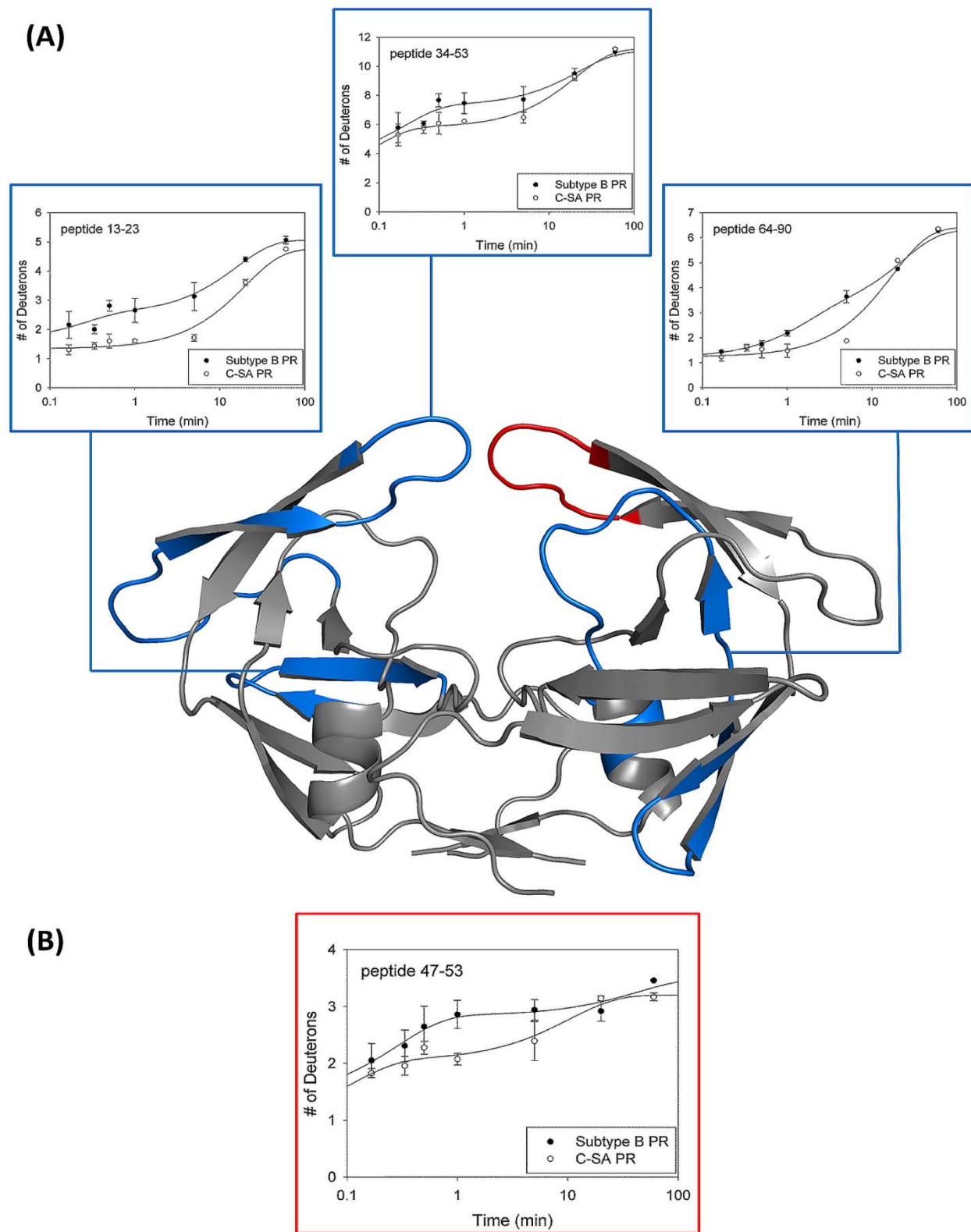
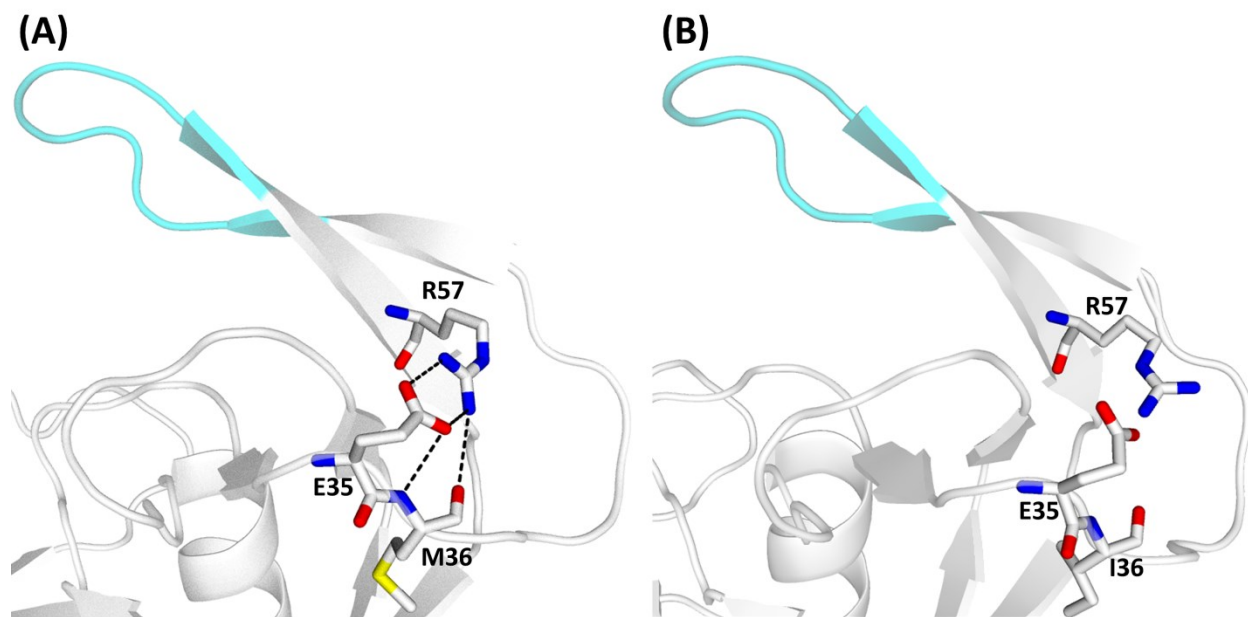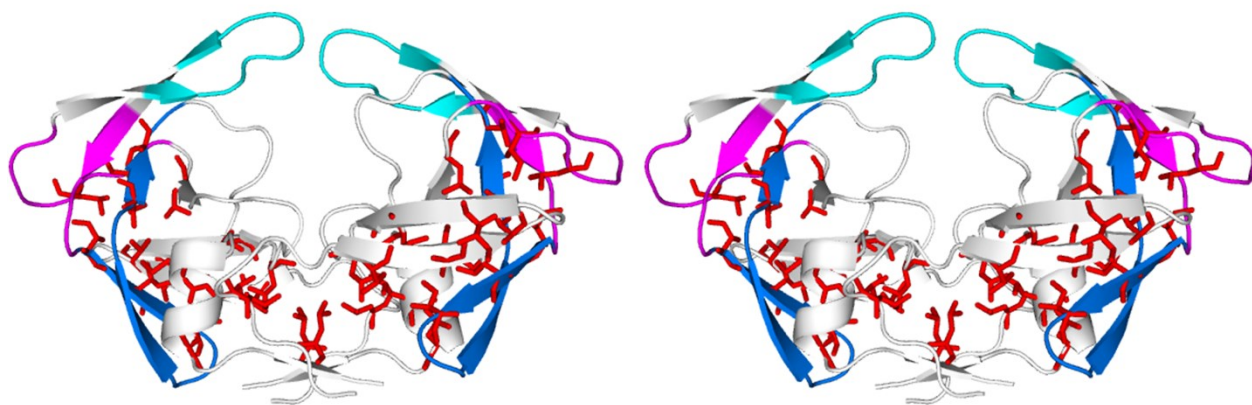Figure 1

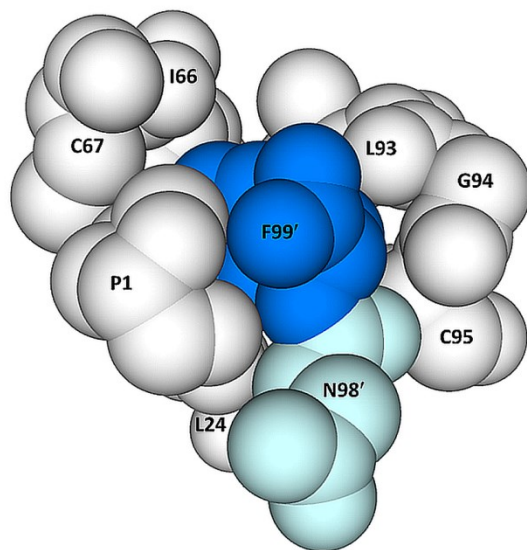Figure 2

Figure 3

Figure 4

Figure 5

Figure 6

Figure 7

Chapter 5

General discussion and conclusions

5.1 C-SA PR displays increased substrate turnover and reduced drug susceptibility in comparison to the subtype B PR

Steady-state kinetic parameters were determined following hydrolysis of the HIV-1 PR fluorogenic substrate (Abz-Arg-Val-Nle-Phe(NO$_2$)-Glu-Ala-Nle-NH$_2$) which mimics the capsid/p2 cleavage site in the Gag and Gag-Pol polyprotein precursors. The K_M and k_{cat}/K_M values displayed by the subtype B and C-SA PRs are comparable. However, the C-SA PR displays a 2-fold increase in substrate turnover per second (higher k_{cat}). This improved rate of substrate processing may stem from a 23% increase in the proportion of fully-open conformers evident in the C-SA PR population relative to that of the subtype B PR[54]. Importantly, structure-based calculations reveal that semi-open conformers do not permit the entry of substrate/inhibitor to the active site[73]. Entry of a substrate or inhibitor to the active site of the PR requires substantial movement of the flaps (~ 15 Å from their position in the closed conformer)[58] and flap flexibility is a requirement for substrate binding and product release[74]. Therefore, only the fully-open conformer may allow a substrate/inhibitor access to the active site. An increased preference for the fully-open conformer may improve the rate substrate entry and/or product release, evidenced by the increase in k_{cat} for the C-SA PR. Binding kinetics of FDA-approved PIs to the HIV PR measured by surface plasmon resonance (SPR) showed that the rate of inhibitor dissociation is far slower than the rate of inhibitor association[75]. The rate of inhibitor dissociation is greater in multi-drug resistant

(MDR) PRs than for the wild-type; however, display similar rates of inhibitor binding[75]. Therefore, polymorphisms which improve the rate of inhibitor dissociation through increased propensity for more open conformers may also improve the rate of substrate release. The increased preference for the fully-open conformer must be a result of some altered dynamics in the C-SA PR. Alterations in the regions which facilitate flap movement are likely to explain the greater presence of fully-open conformers because no polymorphisms occur in the flap region which may directly affect flap flexibility[27; 76].

The C-SA PR displayed reduced drug susceptibility toward ritonavir and darunavir in comparison to the subtype B PR. The relative vitality, indicating the selective advantage of polymorphisms, of the C-SA PR relative to the subtype B PR in the presence of ritonavir and darunavir was 4-fold and 10-fold greater, respectively. The vitality of the C-SA PR in the presence of atazanavir is comparable to that of the subtype B PR. All the aforementioned inhibitors, which are recommended in current antiretroviral therapy guidelines, bound with a reduced entropic contribution to the C-SA PR in comparison to the subtype B PR and an unfavourable entropy change ($-T\Delta S > 0$) was observed during darunavir binding. The reduced entropic contributions are compensated by improved enthalpic contributions with no significant difference in the change in Gibbs free energy during inhibitor binding to either PR. Enthalpy-entropy compensation is thought to be mainly based on the weak interactions of the bulk solvent. The reduced entropic contribution during inhibitor binding to the C-SA PR may be due to increased order of the inhibitor and/or solvent molecules in the binding cavity. Retention of ordered solvent in the binding cavity results in increased hydrogen bonding with the PR and other structural water molecules. These bonds are stronger than the intermolecular forces in the bulk solvent. Therefore, such bonds result in an enhanced enthalpy change.

Altogether, the reduced susceptibility of the C-SA PR toward ritonavir and darunavir also substantiates the classification of the M36I, L89M and I93L mutations, which are inherent to the C-SA PR, as secondary drug resistance mutations.

5.2 Global structures of PRs are similar in their static states

The alignment of the crystal structures of the consensus subtype B and C-SA PRs revealed no major structural differences between the PRs. Protein dynamics data suggest a difference in the equilibrium between flap conformers for the apo-C-SA PR relative to the apo-subtype B PR[54]. Although no structural differences in the flap region of the apo-PRs were identified, the overall dynamics of the PRs need to be investigated to elucidate the mechanism of reduced drug susceptibility displayed by the C-SA PR. The polymorphic sites were investigated in detail and only structural differences at the polymorphic residue at position 36 were identified. The M36I polymorphism may hinder formation of the nearby E35-R57 salt bridge which is the only salt bridge in the flap-hinge region. The reduced propensity of the C-SA PR to form the E35-R57 salt bridge may facilitate altered anchoring of the flap-hinge region.

The C-SA PR crystallised in the semi-open conformation with flaps displaying normal handedness (similar to the closed conformation Figure 5A). The flap tips (residue positions 48–52) are displaced further away and do not overlap in contrast to the closed conformation. Two distinct types of semi-open conformers have been identified experimentally in crystal structures of the apo-HIV-1 PR[55]. The conventional semi-open conformer trapped during protein crystallography exhibits a reversal of flap handedness (Figure 5B). This conventional conformer appears to display slight flap tip curling similar to the curled and fully-open

conformers (Figure 5C). However, both ensembles of the semi-open conformer may exist in solution.

Previous drug-complexed structures of a subtype C PR with a sequence matching a patient from India displayed the same overall fold and no major structural differences to drug-complexed subtype B PR structures[49]. A detailed analysis was possible using the indinavir bound structure of the subtype C PR which showed no significant difference in the number of hydrophobic interactions and hydrogen bonds between the drug and active site in comparison to the indinavir bound structure of the subtype B PR[49]. This subtype C PR did exhibit an increased proportion of the fully-open conformer in the apo-form in comparison to the subtype B PR[54]; however, any structural difference that may allow for the shift in the equilibrium of PR conformers was not identified in the crystal structures of the subtype C PR[49]. The aforementioned subtype C PR exhibits the N37A and K41R polymorphisms in comparison to the C-SA PR. Unfortunately, several attempts to obtain high-resolution drug-complexed structures of the C-SA PR in the current study proved unsuccessful. Obtaining such structures is an objective of upcoming projects in our lab.

5.3 Isotope exchange mechanism of HIV-1 PR

Hydrogen/deuterium exchange (HDX) under conditions where proteins are generally folded may be illustrated using a two-process model (Figure 6)[77-79]. Exchange in a folded protein (Figure 6A) is believed to occur through low amplitude atomic motions, of ~ 1 Å, which are sufficient to allow diffusion of D_2O and OD^- to backbone amide linkages[80; 81]. The motion of a single atom is not large enough to allow penetration of solvent molecules; however, the collective effect of many independent motions occasionally allows for transient penetration of solvent[82]. Simultaneously, short segments as well as the entire protein backbone may exchange through unfolding processes[83].

HDX in macromolecules resulting from local unfolding events (Figure 6B) may either occur via a correlated (EX1) or uncorrelated (EX2) mechanism[82; 84-86]. Correlated HDX (EX1) occurs when the rate of isotopic exchange (k_{int}) far exceeds the rate of refolding (k_{-1}); and thus, all amide hydrogens within a segment exchange for deuterons in solution when the segment is unfolded[82; 84; 85]. HDX in the subtype B and C-SA PRs displays EX2 kinetics indicated by the single binomial isotopic distribution of the mass spectra in the current study[87]. This uncorrelated exchange mechanism may be well-illustrated by HDX at α-helices[88]. Exchange at residues in the single α-helix of the HIV-1 PR (residues at positions 87–93) occurs independently of each other. HDX via the EX2 mechanism occurs when the rate of refolding (k_{-1}) far exceeds the rate of isotopic exchange (k_{int}), suggesting that segments of the PRs must unfold and refold many times before exchange within it is complete[82; 84; 85].
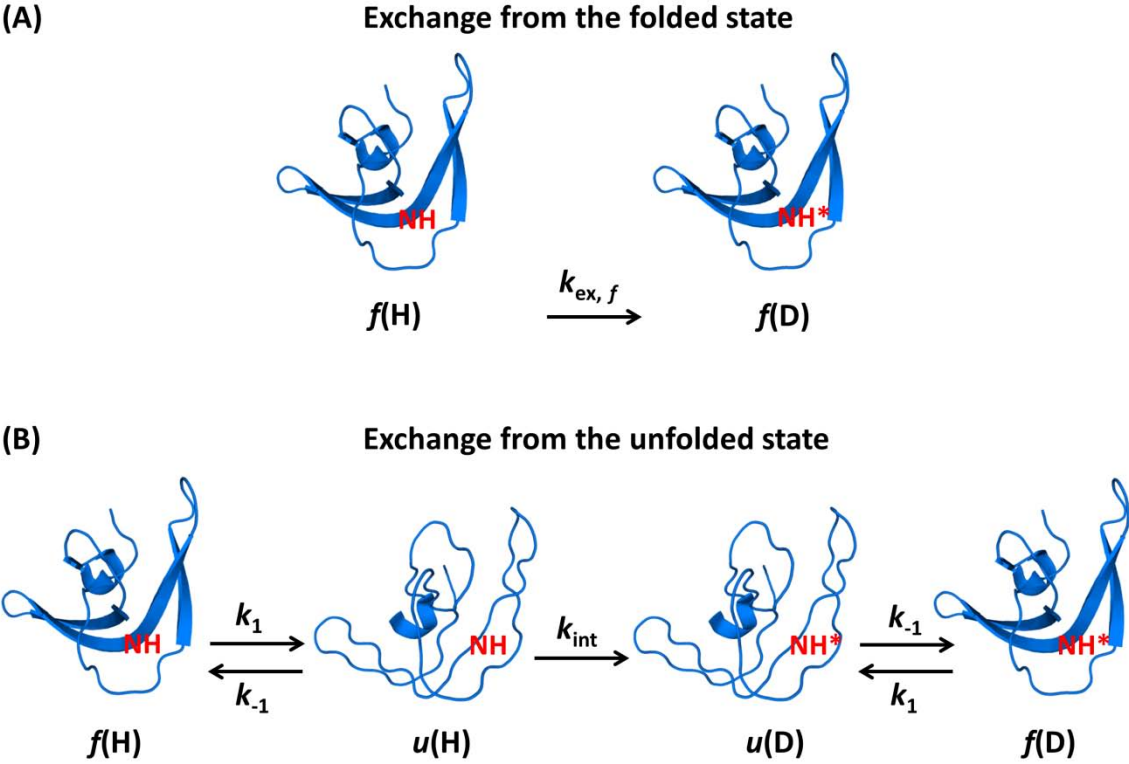
Figure 6: Description of isotope exchange in the folded (A) and unfolded (B) states of a protein. NH* denotes a backbone amide which has exchanged its proton with a deuterium atom in solution. H and D refer to hydrogen and deuterium, respectively, and f and u refer to the folded and unfolded forms. The isotopic exchange rate constant from the folded state is designated as $k_{ex,f}$. The unfolding rate (k_1), refolding rate (k_{-1}) and isotopic exchange rate (k_{int}) constants are also displayed.

5.4 Conformational stability of HIV-1 PR

Previous differential scanning calorimetry studies showed that the overall conformational stability of the C-SA PR was slightly reduced in comparison to the subtype B PR[16]. Hydrogen/deuterium exchange-mass spectrometry (HDX-MS) data reported in the current study identified subtle differences in stability at different regions of the PRs. In most peptides, the number of amide protons exchanging at fast rates increased by ~ 1 amide proton in the subtype B PR relative to the C-SA PR. This may initially imply that the C-SA PR is more stable than the subtype B PR. However, residues at the C-terminus (C95 and T96) of the C-SA PR appear to be more dynamic. Thermodynamic analyses showed that the N- and C-terminal antiparallel β-sheet contributes 75% to the total Gibbs energy[89]. A more dynamic C-terminus will affect the stability of the terminal β-sheet and greatly impact the overall conformational stability. Salt bridges also affect the conformational stability of macromolecules[90; 91]. Crystal structure analysis shows that the C-SA PR lacks the K20-E34 and E35-R57 salt bridges relative to the subtype B PR. Therefore, the reduced number of ionic interactions and a less stable terminal β-sheet evident in the C-SA PR are major determinants for the apparent overall reduced conformational stability of the C-SA PR.

5.5 Effect of polymorphisms on C-SA PR stability

HDX-MS results indicate that the regions at the polymorphic sites of the C-SA PR are more stable than that of the subtype B PR. In all peptides covering the polymorphic sites (10's, 60's and 80's loops and the hinge region) there is a reduction in the number of very fast

and/or fast exchanging amide protons ($12 > k > 0.1$ min^{-1}) and a concomitant increase in the number of intermediate exchanging amide protons ($0.1 > k > 0.01$ min^{-1}) in the C-SA PR. The differences observed in peptide 13–23 (covering the fulcrum region), 34–53 (covering the hinge region) and 64–76 (covering the cantilever region) may contribute to the varying preference of flap conformations between the PRs. Distance measurements during pulsed EPR experiments showed that the total population of the fully-open conformer increased by 23% with a concomitant decrease by 15% in the total population of semi-open conformers in the subtype C PR relative to the subtype B PR. Complete flap opening occurs through concerted downward movement of the hinge (residue positions 35–42 and 57–61), cantilever (residue positions 62–78) and fulcrum (residue positions 10–23) regions and the exchange of hydrophobic contacts between these regions[92]. In the C-SA PR, I15V, L19I, M36I, L89M and I93L polymorphisms occur in the hydrophobic core. These polymorphisms may alter the network of hydrophobic contacts in the core. It is evident from the crystal structures of the apo-subtype B and C-SA PRs that these polymorphisms bring about slight changes in the complex network of hydrophobic interactions. Due to its high incidence in drug-resistant isolates, the M36I polymorphism likely contributes the greatest to differences in flap dynamics between the PRs. The rotation of I36 through different rotamers may assist in sliding over other hydrophobic residues. In addition, the reduced propensity of the C-SA PR to form the E35-R57 salt bridge may aid in the sliding of hydrophobic residues in the core and not reduce stability in the hinge region as initially predicted[27]. HDX-MS data indicate that the network of interactions in the hydrophobic core of the C-SA PR may be altered. Segments of the hinge, cantilever and fulcrum regions display increased stability in the C-SA PR which may ultimately stabilise the fully-open conformer.

5.6 C-SA PR displays a wider range of fully-open conformers

MD simulations were performed to identify differences in flap tip dynamics between the subtype B and C-SA PRs. The inter-flap distances (distance between C_α of I50 on adjacent monomers) of the PRs were measured to probe flap tip dynamics over 10 nanoseconds. NMR studies have indicated that motion of the flap tips do occur on a time-scale of less than 10 nanoseconds[51]. Simulations were also performed on a PR from a multi-drug resistant clinical isolate (MDR 769)[93]. All PR structures used for simulations displayed normal flap handedness and this simulation study probed movements of the flap tip of this type of semi-open conformer in the PRs. The MDR PR displayed drug resistance to most FDA-approved PIs and reduced susceptibility to the second generation PIs; namely, darunavir, lopinavir and tipranavir[94]. The MDR PR was expected to show a difference in flap dynamics due the presence of flap region mutations (M46L and I54V). A wider displaced semi-open conformation was displayed by the MDR PR for ~ 4 nanoseconds of the 10 nanosecond simulation. However, spin-labelled pulsed EPR spectroscopy studies did not reveal a wider semi-open conformer for the MDR PR. A wider fully-open conformer was rather identified[54; 95]. It is important to note the type of starting structure used during the simulation of the current study. The PRs in the wide semi-open conformation (Figure 7C), displaying normal flap handedness, used in the current study has been likened to a wide-open conformer[47; 55]. Although this type of semi-open conformer does not display a reversal of flap handedness, as seen in conventional semi-open conformers (Figure 7B) and simulations of the fully-open conformer[57], the distance between the flap tips is far greater than that displayed by the conventional semi-open conformer and is likened to the wider opening displayed by the fully-open conformer[55]. Therefore, flap tip dynamics reported here may be indicative of the type of motions which occur when the flaps are displaced further apart, as in the fully-open

conformer, rather than close together as displayed by the conventional semi-open conformer[55]. Large-scale flap motions indicative of full flap opening occur on the microsecond–millisecond time-scale[51]. Sampling fully-open flap conformers via MD simulations on a computationally feasible time-scale has only been possible using a continuum solvent representation with reduced viscosity, allowing for accelerated protein motions[57; 96]. These studies were performed using a closed (Figure 7A) and a conventional semi-open conformer (Figure 7B)[57]. Whether the wide semi-open conformer is physiologically relevant is up for debate. However, it does exhibit some characteristics of the fully-open conformer.
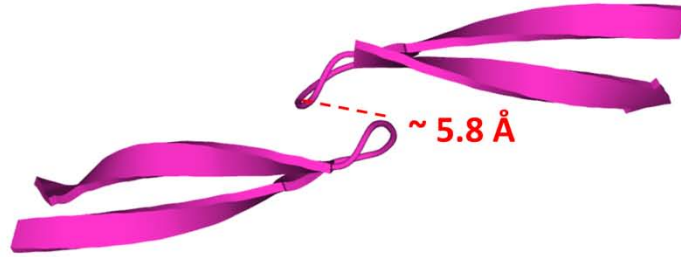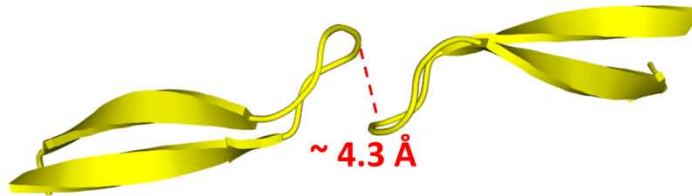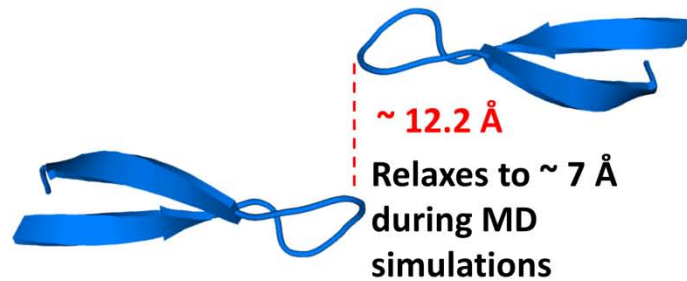
Figure 7: Inter-flap distances of closed and semi-open conformers. Inter-flap distances in red are based on crystal structure data. The wide semi-open conformers of the subtype B and C-SA PRs relax to ~ 7 Å during the simulation experiment in this study. PDB ID: 1HXW (A)[61], 1HHP (B)[62] and 2PC0 (C)[55].

The C-SA PR does not display larger inter-flap distances in comparison to the subtype B PR. Following complete equilibration of all PR starting structures (after ~ 3.8 nanoseconds), the PRs relax to either the closed or the wide semi-open conformer which displays closer inter-flap distances (~ 7 Å) than that of the respective crystal structures. The subtype B and C-SA PRs displayed "closed" inter-flap distances (~ 5.8 Å) for part of the simulation and exhibited inter-flap distances indicative of the wider semi-open conformation at the end of the simulation. However, the range of inter-flap distances following equilibration of both structures is wider for the subtype C PR. These results are consistent with pulsed EPR spectroscopy experiments which showed a wider range of fully-open conformers for the C-SA PR in comparison to the subtype B PR[54]. The range of semi-open conformers were similar[54], consistent with the notion that differences in inter-flap distance measured in the present simulation report on the motion of open flaps. This difference was highlighted by a broader peak width at half maximal intensity for the fully-open population of the C-SA PR during pulsed EPR spectroscopy measurements[54]. These results together with that reported in the current study indicate that the C-SA PR must display altered dynamics to allow for the stabilisation of the wider range of open conformers. It is improbable that the wider range of open conformers is due to altered dynamics at the flap tips because no polymorphisms are located in the flap region. HDX-MS results reported in the current study did not show increased dynamics at the flap tips of the C-SA PR. Our results suggest that altered dynamics around the flaps and the increased propensity for fully-open conformations is a result of altered dynamics in the regions which allow for flap opening and closure.

5.7 Key features of dimerisation in HIV-1 PRs

The N-terminus (positions 1–5) and C-terminus (positions 95–99) of HIV-1 PRs are pivotal in dimer stability; structure-based thermodynamic analyses showed that the N- and C-termini antiparallel β-sheet contributes ~ 75% to the total Gibbs energy[89]. HDX-MS data reveal that the N-terminus of both PRs is more flexible than the C-terminus, supporting previous structural analysis. These findings further substantiate the ordered initial cleavage of the p6^{Pol}-PR junction prior to cleavage of the more stable PR-RT junction[97].

Drug resistance to currently available PIs are attributed to mutations in the active site cavity and flap-hinge region. Dimerisation inhibitors targeted to the conserved C-terminus of HIV PR may have a higher barrier to resistance. The C-terminal residues, C95 and T96, are more dynamic in the C-SA PR in comparison to that of the subtype B PR. Therefore, these residues may contribute less to the stability of the N- and C-termini β-sheet than other C-terminal residues and may not serve as suitable targets for preventing dimerisation. Mutation of the ultimate amino acid in the HIV-1 PR sequence, phenylalanine, to an alanine was shown to reduce the β-sheet content of the PR and disrupt dimerisation which results in complete loss of activity[98]. F99 is central to the large network of interactions present in the 'lock-and-key' motif. Therefore, the importance of the C-terminus for correct folding of the PR and its increased stability in comparison to the N-terminus, make the C-terminus, particularly F99, a preferred target for antiviral therapy.

5.8 Conclusions

Polymorphisms inherent in the C-SA PR represent non-active site and non-flap region mutations. These polymorphisms result in the C-SA PR displaying improved viral fitness in the presence of certain PIs in comparison to the subtype B PR. The improved viral fitness is not explained by the static structure of the C-SA PR which is comparable to that of the subtype B PR. Previous and current investigations of flap dynamics indicate that the C-SA PR has an increased preference for the fully-open conformer and polymorphisms inherent in the C-SA PR allow for a wider range of these fully-open conformers[27; 54]. Increased stability in segments of the hinge, cantilever and fulcrum regions likely allow for altered anchoring of the flaps ultimately increasing the preference of fully-open conformers in the C-SA PR in comparison to the subtype B PR. Because the C-SA PR displays reduced drug susceptibility and similar catalytic efficiency in comparison to the subtype B PR, selection of fully-open conformers preferentially affects inhibitor binding over substrate binding. FDA-approved PIs were designed to target the closed conformation of the PR and are more rigid than substrates of HIV PR. As substrate processing is more dynamic than inhibitor binding, polymorphisms which promote more open conformers do not hinder substrate processing; rather, they may improve substrate turnover as shown by the increase in k_{cat} for the C-SA PR. Inhibitors designed to target the fully-open conformer of the PR, which is a prerequisite for substrate binding, may be preferred for PRs which have a preference for the fully-open conformer such as the epidemiologically relevant C-SA PR.

References

1.   The Joint Nations Program on HIV/AIDS (UNAIDS). (2013). *Global report: UNAIDS report on the global AIDS epidemic 2013*, UNAIDS: Geneva, Switzerland.

2.   Barre-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C., Rozenbaum, W. & Montagnier, L. (1983). Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* **220**, 868-871.

3.   Bebenek, K., Abbotts, J., Roberts, J. D., Wilson, S. H. & Kunkel, T. A. (1989). Specificty and mechanism of error-prone replication by human immonodeficiency virus-1 reverse-transcriptase. *J. Biol. Chem.* **264**, 16948-16956.

4.   Tomasselli, A. G. & Heinrikson, R. L. (2000). Targeting the HIV-protease in AIDS therapy: a current clinical perspective. *Biochim. Biophys. Acta* **1477**, 189-214.

5.   de Oliveira, T., Engelbrecht, S., van Rensburg, E. J., Gordon, M., Bishop, K., zur Megede, J., Barnett, S. W. & Cassol, S. (2003). Variability at human immunodeficiency virus type 1 subtype C protease cleavage sites: an indication of viral fitness? *J. Virol.* **77**, 9422-9430.

6.   Naicker, P. & Sayed, Y. (2014). Non-B HIV-1 subtypes in sub-Saharan Africa: impact of subtype on protease inhibitor efficacy *Biol. Chem.*, In Press.

7.   Monini, P., Sgadari, C., Toschi, E., Barillari, G. & Ensoli, B. (2004). Antitumour effects of antiretroviral therapy. *Nat. Rev. Cancer* **4**, 861-875.

8.   McCutchan, F. E. (2006). Global epidemiology of HIV. *J. Med. Virol.* **78**, 7-12.

9.   Taylor, B. S., Sobieszczyk, M. E., McCutchan, F. E. & Hammer, S. M. (2008). Medical progress: The challenge of HIV-1 subtype diversity. *N. Engl. J. Med.* **358**, 1590-1602.

10.   Plantier, J. C., Leoz, M., Dickerson, J. E., De Oliveira, F., Cordonnier, F., Lemee, V., Damond, F., Robertson, D. L. & Simon, F. (2009). A new human immunodeficiency virus derived from gorillas. *Nat. Med.* **15**, 871-872.

11.   Hemelaar, J., Gouws, E., Ghys, P. D. & Osmanov, S. (2011). Global trends in molecular epidemiology of HIV-1 during 2000-2007. *AIDS* **25**, 679-89.

12.   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.

13.   King, N. M., Prabu-Jeyabalan, M., Bandaranayake, R. M., Nalam, M. N. L., Nalivaika, E. A., Ozen, A., Haliloglu, T., Yilmaz, N. K. & Schiffer, C. A. (2012). Extreme Entropy-Enthalpy Compensation in a Drug-Resistant Variant of HIV-1 Protease. *ACS Chem. Biol.* **7**, 1536-1546.

14.   Muzammil, S., Armstrong, A. A., Kang, L. W., Jakalian, A., Bonneau, P. R., Schmelmer, V., Amzel, L. M. & Freire, E. (2007). Unique thermodynamic response of tipranavir to human immunodeficiency virus type 1 protease drug resistance mutations. *J. Virol.* **81**, 5144-54.

15.   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.

16.   Velazquez-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.

17.   Hemelaar, J., Gouws, E., Ghys, P. D. & Osmanov, S. (2006). Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. *AIDS* **20**, 13-23.

18. Kantor, R. & Katzenstein, D. (2003). Polymorphism in HIV-1 Non-subtype B Protease and Reverse Transcriptase and its Potential Impact on Drug Susceptibility and Drug Resistance Evolution. *AIDS Rev.* **5**, 25-35.

19. Loeb, D. D., Hutchison, C. A., Edgell, M. H., Farmerie, W. G. & Swanstrom, R. (1989). Mutational analysis of human immunodeficiency virus type-1 protease suggests functional homology with aspartic proteinases *J. Virol.* **63**, 111-121.

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

21. Wlodawer, A. & Vondrasek, J. (1998). Inhibitors of HIV-1 protease: A major success of structure-assisted drug design. *Annu. Rev. Biophys. Biomol. Struct.* **27**, 249-284.

22. Flexner, C. (1998). HIV-protease inhibitors. *N. Engl. J. Med.* **338**, 1281-1292.

23. Hirsch, M. S., Brun-Vezinet, F., D'Aquila, R. T., Hammer, S. M., Johnson, V. A., Kuritzkes, D. R., Loveday, C., Mellors, J. W., Clotet, B., Conway, B., Demeter, L. M., Vella, S., Jacobsen, D. M. & Richman, D. D. (2000). Antiretroviral drug resistance testing in adult HIV-1 infection - Recommendations of an International AIDS Society-USA panel. *JAMA* **283**, 2417-2426.

24. 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**, 2349-59.

25. Rose, R. E., Gong, Y. F., Greytok, J. A., Bechtold, C. M., Terry, B. J., Robinson, B. S., Alam, M., Colonno, R. J. & Lin, P. F. (1996). Human immunodeficiency virus

type 1 viral background plays a major role in development of resistance to protease inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 1648-53.

26. World Health Organisation (WHO). (2013). *The use of antiretroviral drugs for treating and preventing HIV infection*, Geneva, Switzerland: WHO Press.

27. 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.

28. Kozisek, M., Lepsik, M., Grantz Saskova, K., Brynda, J., Konvalinka, J. & Rezacova, P. (2014). Thermodynamic and structural analysis of HIV protease resistance to darunavir - analysis of heavily mutated patient-derived HIV-1 proteases. *FEBS J.* **281**, 1834-47.

29. Ali, A., Bandaranayake, R. M., Cai, Y. F., King, N. M., Kolli, M., Mittal, S., Murzycki, J. F., Nalam, M. N. L., Nalivaika, E. A., Ozen, A., Prabu-Jeyabalan, M. M., Thayer, K. & Schiffer, C. A. (2010). Molecular Basis for Drug Resistance in HIV-1 Protease. *Viruses* **2**, 2509-2535.

30. Wiley, R. A. & Rich, D. H. (1993). Peptidomimetics derived from natural-products. *Med. Res. Rev.* **13**, 327-384.

31. Kempf, D. J., Marsh, K. C., Kumar, G., Rodrigues, A. D., Denissen, J. F., McDonald, E., Kukulka, M. J., Hsu, A., Granneman, G. R., Baroldi, P. A., Sun, E., Pizzuti, D., Plattner, J. J., Norbeck, D. W. & Leonard, J. M. (1997). Pharmacokinetic enhancement of inhibitors of the human immunodeficiency virus protease by coadministration with ritonavir. *Antimicrob. Agents Chemother.* **41**, 654-660.

32. Altman, M. D., Ali, A., Reddy, G., Nalam, M. N. L., Anjum, S. G., Cao, H., Chellappan, S., Kairys, V., Fernandes, M. X., Gilson, M. K., Schiffer, C. A., Rana, T.

M. & Tidor, B. (2008). HIV-1 protease inhibitors from inverse design in the substrate envelope exhibit subnanomolar binding to drug-resistant variants. *J. Am. Chem. Soc.* **130**, 6099-6113.

33. 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.

34. Pokorna, J., Machala, L., Rezacova, P. & Konvalinka, J. (2009). Current and Novel Inhibitors of HIV Protease. *Viruses* **1**, 1209-1239.

35. Prabu-Jeyabalan, M., Nalivaika, E. & Schiffer, C. A. (2000). How does a symmetric dimer recognize an asymmetric substrate? A substrate complex of HIV-1 protease. *J. Mol. Biol.* **301**, 1207-1220.

36. Cherry, E., Liang, C., Rong, L., Quan, Y., Inouye, P., Li, X., Morin, N., Kotler, M. & Wainberg, M. A. (1998). Characterization of human immunodeficiency virus type-1 (HIV-1) particles that express protease-reverse transcriptase fusion proteins. *J. Mol. Biol.* **284**, 43-56.

37. Co, E., Koelsch, G., Lin, Y., Ido, E., Hartsuck, J. A. & Tang, J. (1994). Proteolytic processing mechanisms of a miniprecursor of the aspartic protease of human immunodeficiency virus type 1. *Biochemistry* **33**, 1248-54.

38. Louis, J. M., Clore, G. M. & Gronenborn, A. M. (1999). Autoprocessing of HIV-1 protease is tightly coupled to protein folding. *Nat. Struct. Biol.* **6**, 868-75.

39. Louis, J. M., Nashed, N. T., Parris, K. D., Kimmel, A. R. & Jerina, D. M. (1994). Kinetics and mechanism of autoprocessing of human immunodeficiency virus type 1 protease from an analog of the Gag-Pol polyprotein. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 7970-4.

40.   Wondrak, E. M., Nashed, N. T., Haber, M. T., Jerina, D. M. & Louis, J. M. (1996). A transient precursor of the HIV-1 protease. Isolation, characterization, and kinetics of maturation. *J. Biol. Chem.* **271**, 4477-81.

41.   Hyland, L. J., Tomaszek, T. A., Jr., Roberts, G. D., Carr, S. A., Magaard, V. W., Bryan, H. L., Fakhoury, S. A., Moore, M. L., Minnich, M. D., Culp, J. S., DesJarlais, R. L. & Meek, T. D. (1991). Human immunodeficiency virus-1 protease. 1. Initial velocity studies and kinetic characterization of reaction intermediates by 18O isotope exchange. *Biochemistry* **30**, 8441-53.

42.   Shen, C. H., Tie, Y., Yu, X., Wang, Y. F., Kovalevsky, A. Y., Harrison, R. W. & Weber, I. T. (2012). Capturing the reaction pathway in near-atomic-resolution crystal structures of HIV-1 protease. *Biochemistry* **51**, 7726-32.

43.   Hyland, L. J., Tomaszek, T. A., Jr. & Meek, T. D. (1991). Human immunodeficiency virus-1 protease. 2. Use of pH rate studies and solvent kinetic isotope effects to elucidate details of chemical mechanism. *Biochemistry* **30**, 8454-63.

44.   Smith, R., Brereton, I. M., Chai, R. Y. & Kent, S. B. (1996). Ionization states of the catalytic residues in HIV-1 protease. *Nat. Struct. Biol.* **3**, 946-50.

45.   Wang, Y. X., Freedberg, D. I., Yamazaki, T., Wingfield, P. T., Stahl, S. J., Kaufman, J. D., Kiso, Y. & Torchia, D. A. (1996). Solution NMR evidence that the HIV-1 protease catalytic aspartyl groups have different ionization states in the complex formed with the asymmetric drug KNI-272. *Biochemistry* **35**, 9945-50.

46.   Coman, R. M., Robbins, A., Goodenow, M. M., McKenna, R. & Dunn, B. M. (2007). Expression, purification and preliminary X-ray crystallographic studies of the human immunodeficiency virus 1 subtype C protease. *Acta Crystallogr. F* **63**, 320-323.

47.   Coman, R. M., Robbins, A. H., Goodenow, M. M., Dunn, B. M. & McKenna, R. (2008). High-resolution structure of unbound human immunodeficiency virus 1

subtype C protease: implications of flap dynamics and drug resistance. *Acta Crystallogr. D* **64**, 754-763.

48.   Velazquez-Campoy, A., Vega, S. & Freire, E. (2002). Amplification of the effects of drug resistance mutations by background polymorphisms in HIV-1 protease from African subtypes. *Biochemistry* **41**, 8613-9.

49.   Coman, R. M., Robbins, A. H., Fernandez, M. A., Gilliland, C. T., Sochet, A. A., Goodenow, M. M., McKenna, R. & Dunn, B. M. (2008). The contribution of naturally occurring polymorphisms in altering the biochemical and structural characteristics of HIV-1 subtype C protease. *Biochemistry* **47**, 731-43.

50.   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**, 221-32.

51.   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**, 1047-55.

52.   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. & et al. (1995). Flexibility and function in HIV-1 protease. *Nat. Struct. Biol.* **2**, 274-80.

53.   Galiano, L., Bonora, M. & Fanucci, G. E. (2007). Interflap distances in HIV-1 protease determined by pulsed EPR measurements. *J. Am. Chem. Soc.* **129**, 11004-5.

54.   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**, 14650-1.

55.   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* **63**, 866-875.

56.   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**, 1259-65.

57.   Hornak, V., Okur, A., Rizzo, R. C. & Simmerling, C. (2006). HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 915-20.

58.   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**, 269-72.

59.   Huang, X., de Vera, I. M., Veloro, A. M., Blackburn, M. E., Kear, J. L., Carter, J. D., Rocca, J. R., Simmerling, C., Dunn, B. M. & Fanucci, G. E. (2012). Inhibitor-induced conformational shifts and ligand-exchange dynamics for HIV-1 protease measured by pulsed EPR and NMR spectroscopy. *J. Phys. Chem. B* **116**, 14235-44.

60.   Martin, P., Vickrey, J. F., Proteasa, G., Jimenez, Y. L., Wawrzak, Z., Winters, M. A., Merigan, T. C. & Kovari, L. C. (2005). "Wide-open" 1.3 A structure of a multidrug-resistant HIV-1 protease as a drug target. *Structure* **13**, 1887-95.

61.   Kempf, D. J., Marsh, K. C., Denissen, J. F., McDonald, E., Vasavanonda, S., Flentge, C. A., Green, B. E., Fino, L., Park, C. H., Kong, X. P. & et al. (1995). ABT-538 is a potent inhibitor of human immunodeficiency virus protease and has high oral bioavailability in humans. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 2484-8.

62.   Spinelli, S., Liu, Q. Z., Alzari, P. M., Hirel, P. H. & Poljak, R. J. (1991). The three-dimensional structure of the aspartyl protease from the HIV-1 isolate BRU. *Biochimie* **73**, 1391-6.

63.   Agniswamy, J., Shen, C. H., Aniana, A., Sayer, J. M., Louis, J. M. & Weber, I. T. (2012). HIV-1 protease with 20 mutations exhibits extreme resistance to clinical inhibitors through coordinated structural rearrangements. *Biochemistry* **51**, 2819-28.

64.   Braz, V. A., Barkley, M. D., Jockusch, R. A. & Wintrode, P. L. (2010). Efavirenz Binding Site in HIV-1 Reverse Transcriptase Monomers. *Biochemistry* **49**, 10565-10573.

65.   Lanman, J., Lam, T. T., Barnes, S., Sakalian, M., Emmett, M. R., Marshall, A. G. & Prevelige, P. E. (2003). Identification of novel interactions in HIV-1 capsid protein assembly by high-resolution mass spectrometry. *J. Mol. Biol.* **325**, 759-772.

66.   Lanman, J., Lam, T. T., Emmett, M. R., Marshall, A. G., Sakalian, M. & Prevelige, P. E. (2004). Key interactions in HIV-1 maturation identified by hydrogen-deuterium exchange. *Nat. Struct. Mol. Biol.* **11**, 676-677.

67.   Monroe, E. B., Kang, S., Kyere, S. K., Li, R. & Prevelige, P. E. (2010). Hydrogen/Deuterium Exchange Analysis of HIV-1 Capsid Assembly and Maturation. *Structure* **18**, 1483-1491.

68.   Descamps, D., Chaix, M. L., Andre, P., Brodard, V., Cottalorda, J., Deveau, C., Harzic, M., Ingrand, D., Izopet, J., Kohli, E., Masquelier, B., Mouajjah, S., Palmer, P., Pellegrin, I., Plantier, J. C., Poggi, C., Rogez, S., Ruffault, A., Schneider, V., Signori-Schmuck, A., Tamalet, C., Wirden, M., Rouzioux, C., Brun-Vezinet, F., Meyer, L. & Costagliola, D. (2005). French national sentinel survey of antiretroviral drug resistance in patients with HIV-1 primary infection and in antiretroviral-naive chronically infected patients in 2001-2002. *J. Acquir. Immune. Defic. Syndr.* **38**, 545-52.

69.   Foster, G. M., Ambrose, J. C., Hue, S., Delpech, V. C., Fearnhill, E., Abecasis, A. B., Leigh Brown, A. J. & Geretti, A. M. (2014). Novel HIV-1 Recombinants Spreading

across Multiple Risk Groups in the United Kingdom: The Identification and Phylogeography of Circulating Recombinant Form (CRF) 50_A1D. *PLoS One* **9**, e83337.

70. Holguin, A., de Mulder, M., Yebra, G., Lopez, M. & Soriano, V. (2008). Increase of non-B subtypes and recombinants among newly diagnosed HIV-1 native Spaniards and immigrants in Spain. *Curr. HIV Res.* **6**, 327-34.

71. Kanizsai, S., Ghidan, A., Ujhelyi, E., Banhegyi, D. & Nagy, K. (2010). Monitoring of drug resistance in therapy-naive HIV infected patients and detection of African HIV subtypes in Hungary. *Acta Microbiol. Immunol. Hung.* **57**, 55-68.

72. Mendoza, Y., Bello, G., Castillo Mewa, J., Martinez, A. A., Gonzalez, C., Garcia-Morales, C., Avila-Rios, S., Reyes-Teran, G. & Pascale, J. M. (2014). Molecular Epidemiology of HIV-1 in Panama: Origin of Non-B Subtypes in Samples Collected from 2007 to 2013. *PLoS One* **9**, e85153.

73. 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.

74. Tozser, J., Yin, F. H., Cheng, Y. S., Bagossi, P., Weber, I. T., Harrison, R. W. & Oroszlan, S. (1997). Activity of tethered human immunodeficiency virus 1 protease containing mutations in the flap region of one subunit. *Eur. J. Biochem.* **244**, 235-41.

75. Dierynck, I., De Wit, M., Gustin, E., Keuleers, I., Vandersmissen, J., Hallenberger, S. & Hertogs, K. (2007). Binding kinetics of darunavir to human immunodeficiency virus type 1 protease explain the potent antiviral activity and high genetic barrier. *J. Virol.* **81**, 13845-51.

76. Piana, S., Carloni, P. & Rothlisberger, U. (2002). Drug resistance in HIV-1 protease: Flexibility-assisted mechanism of compensatory mutations. *Protein Sci.* **11**, 2393-402.

77.   Bai, Y., Milne, J. S., Mayne, L. & Englander, S. W. (1994). Protein stability parameters measured by hydrogen exchange. *Proteins* **20**, 4-14.

78.   Kim, K. S. & Woodward, C. (1993). Protein internal flexibility and global stability: effect of urea on hydrogen exchange rates of bovine pancreatic trypsin inhibitor. *Biochemistry* **32**, 9609-13.

79.   Li, R. & Woodward, C. (1999). The hydrogen exchange core and protein folding. *Protein Sci.* **8**, 1571-90.

80.   Elber, R. & Karplus, M. (1987). Multiple conformational states of proteins: a molecular dynamics analysis of myoglobin. *Science* **235**, 318-21.

81.   Kim, K. S., Fuchs, J. A. & Woodward, C. K. (1993). Hydrogen exchange identifies native-state motional domains important in protein folding. *Biochemistry* **32**, 9600-9608.

82.   Woodward, C., Simon, I. & Tuchsen, E. (1982). Hydrogen exchange and the dynamic structure of proteins. *Mol. Cell. Biochem.* **48**, 135-60.

83.   Engen, J. R., Gmeiner, W. H., Smithgall, T. E. & Smith, D. L. (1999). Hydrogen exchange shows peptide binding stabilizes motions in Hck SH2. *Biochemistry* **38**, 8926-35.

84.   Englander, S. W. & Kallenbach, N. R. (1983). Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Q. Rev. Biophys.* **16**, 521-655.

85.   Roder, H., Wagner, G. & Wuthrich, K. (1985). Individual amide proton exchange rates in thermally unfolded basic pancreatic trypsin inhibitor. *Biochemistry* **24**, 7407-11.

86.   Englander, S. W., Downer, N. W. & Teitelbaum, H. (1972). Hydrogen exchange. *Annu. Rev. Biochem.* **41**, 903-24.

87.   Miranker, A., Robinson, C. V., Radford, S. E., Aplin, R. T. & Dobson, C. M. (1993). Detection of transient protein folding populations by mass spectrometry. *Science* **262**, 896-900.

88.   Maity, H., Lim, W. K., Rumbley, J. N. & Englander, S. W. (2003). Protein hydrogen exchange mechanism: local fluctuations. *Protein Sci.* **12**, 153-60.

89.   Todd, M. J., Semo, N. & Freire, E. (1998). The structural stability of the HIV-1 protease. *J. Mol. Biol.* **283**, 475-488.

90.   Fersht, A. R. (1972). Conformational equilibria in alpha and delta chymotrypsin. The energetics and importance of the salt bridge. *J. Mol. Biol.* **64**, 497-509.

91.   Perutz, M. F. (1978). Electrostatic effects in proteins. *Science* **201**, 1187-91.

92.   Foulkes-Murzycki, J. E., Scott, W. R. & Schiffer, C. A. (2007). Hydrophobic sliding: a possible mechanism for drug resistance in human immunodeficiency virus type 1 protease. *Structure* **15**, 225-33.

93.   Logsdon, B. C., Vickrey, J. F., Martin, P., Proteasa, G., Koepke, J. I., Terlecky, S. R., Wawrzak, Z., Winters, M. A., Merigan, T. C. & Kovari, L. C. (2004). Crystal structures of a multidrug-resistant human immunodeficiency virus type 1 protease reveal an expanded active-site cavity. *J. Virol.* **78**, 3123-32.

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

95.   de Vera, I. M., Blackburn, M. E. & Fanucci, G. E. (2012). Correlating conformational shift induction with altered inhibitor potency in a multidrug resistant HIV-1 protease variant. *Biochemistry* **51**, 7813-5.

96. Ding, F., Layten, M. & Simmerling, C. (2008). Solution structure of HIV-1 protease flaps probed by comparison of molecular dynamics simulation ensembles and EPR experiments. *J. Am. Chem. Soc.* **130**, 7184-5.

97. Agniswamy, J., Sayer, J. M., Weber, I. T. & Louis, J. M. (2012). Terminal Interface Conformations Modulate Dimer Stability Prior to Amino Terminal Autoprocessing of HIV-1 Protease. *Biochemistry* **51**, 1041-1050.

98. Naicker, P., Seele, P., Dirr, H. W. & Sayed, Y. (2013). F99 is critical for dimerization and activation of South African HIV-1 subtype C protease. *Protein J.* **32**, 560-567.