

Direktor:
Prof. Dr. rer. nat. Rolf Hilgenfeld

**Crystallographic Analysis of
the Molecular Evolution of HIV-1 Proteinase:
Structure Correlates With Viral Fitness**

Inauguraldissertation
zur
Erlangung der Doktorwürde
der Universität zu Lübeck
- Aus der Technisch-Naturwissenschaftlichen Fakultät -

Vorgelegt von

Aida Baharuddin

aus Malaysia

Lübeck, 2007

1. Berichterstatter : Prof. Dr. rer. nat. R. Hilgenfeld
2. Berichterstatter : HD Dr. Norbert Tautz
Tag der mündlichen Prüfung : 19.09.2007
zum Druck genehmigt. Lübeck, den : 19.09.2007

gez. Prof. Dr. rer. nat. Enno Hartmann

- Dekan der Technisch-Naturwissenschaftlichen Fakultät –

ABSTRACT

Ritonavir, an HIV-1 proteinase inhibitor, is widely used in AIDS therapy. Its effectiveness however has been hindered by the emergence of resistant mutations, a common problem for drugs that target HIV viral enzymes. Several proteinase variants were studied, that evolved during 115 days of ritonavir monotherapy. In total, eleven X-ray structures of six variants have been determined, six of which are inhibitor-free, three are in complex with saquinavir and two are in complex with ritonavir. The variants were isolated on day 0 (the patient's "wild type" strain), day 28a (Val82Thr), day 28b (Met36Ile-Ile54Val), day 56 (Met36Ile-Ile54Val-Val82Thr), day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr) and day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr). X-ray diffraction datasets were collected at the Deutsche-Elektronen Synchrotron beamline X13. The structures were refined to resolutions ranging from 2.65 to 1.77 Å. Comparative analysis of the different proteinase structures, representing both early and late mutations, was used in order to dissect how the amino acid substitutions can have profound effects on the overall structure and fitness. Especially, the unique series of the six unliganded (i.e. inhibitor free) variant structures, covering 115 days of ritonavir monotherapy, provided insights into the emergence of early and late mutations and the observed changes in overall viral fitness. Most notably, late variants (day 82 and 115) both crystallized with the so-called flaps in a closed conformation (i.e. flaps closing-in over the active-site after substrate or drug binding) even though no inhibitor was present.

Zusammenfassung

Ritonavir ist ein Inhibitor der HIV-1 Proteinase, der zur Therapie von AIDS eingesetzt wird. Seine Effektivität wird jedoch durch Resistenzmutationen beeinträchtigt. Das ist ein allgemein bekanntes Problem für Wirkstoffe, die gegen Viren gerichtet sind. In der vorliegenden Arbeit wurde eine Anzahl von Proteinase-Varianten untersucht, die während der 115-tägigen Monotherapie mit Ritonavir entstanden sind. Elf Kristallstrukturen von sechs Varianten wurden gelöst, sechs davon sind inhibitor-frei, drei sind im Komplex mit Saquinavir und zwei mit Ritonavir. Die Varianten wurden am Tag 0 (vor Beginn der Therapie), 28 (Val82Thr bzw. Met36Ile-Ile54Val), 56 (Met36Ile-Ile54Val-Val82Thr), 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr) und am Tag 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr) isoliert. Die Strukturdaten der kristallisierten Proteine wurden am DESY gesammelt. Die Kristallstrukturen wurden bis zu einer Auflösung von 2,65 bis 1,77 Å verfeinert. Mit Hilfe der vergleichenden Analyse verschiedener Proteinase-Strukturen, die sowohl frühere als auch spätere Mutationen darstellen, wurde der Einfluss des jeweiligen Aminosäuren-Austausches auf die Gesamtstruktur untersucht. Die einzigartige Reihe von den sechs inhibitor-freien Strukturvarianten, die 115 Tage einer Ritonavir-Monotherapie repräsentierten, ermöglichte einen tiefen Einblick in die zeitliche Abfolge des Mutationsverlaufs und in die Überlebensfähigkeit der Viren. Auffällig war, dass die beiden späten Varianten der Proteinase (Tag 82 und 115) in der geschlossenen Konformation kristallisiert haben, obwohl kein Inhibitor anwesend war.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Prof. Rolf Hilgenfeld, Director of the Institute of Biochemistry, University of Lübeck, for giving me the opportunity to learn Structural Biochemistry during my Ph.D. work. His knowledge and guidance have been of great value to me. Dr Jeroen R. Mesters for his constructive comments and support throughout this work.

This work has been made possible with the financial support of the Malaysian Government and the University of Lübeck.

I would like to express my appreciation to my fellow lab friends and staff for their sharing, encouragement and support: Dr. Tanis Hogg, Dr. Kanchan Anand, Dr. Jörg Deiwick, Dr. Christian Schmidt, Dr. Ralf Moll, Prof. Dr. Stefan Anemüller, Dr. Koen Verschueren, Dr. Jinzhi Tan, Dr. Ksenia Pumpor, Sabine Kahle, Doris Mutschall, Angelika, Hans-Joachim Kraus, Silke Schmidtke, Walter Verheyen, Rajesh, Krishna, Raspudin, Helgo, Mirko, Yvonne, Nele M, Nele P, Robert and all others that I forgot to mention.

I thank my parents who really understood what I want in my life, and my sisters and brothers for their support and motivation. Last but not least, I thank Associate Dr. Mohd Razip Samian and Associate Dr. Mohd Nazalan Najimudin.

CONTENTS

ZUSAMMENFASSUNG	iii
ABSTRACT	iv
ACKNOWLEDGEMENTS	v
ABBREVIATIONS	ix
1 INTRODUCTION	1
2 LITERATURE REVIEW	2
2.1 HIV structure and lifecycle	2
2.2 HIV genome and structure components	7
2.3 HIV-1 proteinase cleavage sites	7
2.4 HIV-1 proteinase (PR)	10
2.4.1 Characterization of HIV-1 PR	10
2.4.2 The active-site of aspartyl proteases	12
2.4.3 The flaps	17
2.5 Anti-HIV drug design	18
2.5.1 HIV-1 PR inhibitors	19
2.5.1.1 Peptide-based inhibitors	21
2.5.1.1.1 Peptidomimetics	22
2.5.1.1.2 C ₂ -symmetric and pseudo-symmetric inhibitors	23
2.5.1.2 Non-peptidic inhibitors	23
2.5.2 The Schechter and Berger nomenclature for the description of protease subsites	25
2.6 Resistance to proteinase inhibitors	27
2.6.1 Cross-resistance and sequential therapy	28
2.7 Motivation and scope of this thesis	29
3 MATERIALS AND METHODS	31
3.1 Chemicals	31
3.2 Crystallization buffers	31
3.3 Crystallization of variants of HIV-1 proteinase	31
3.4 Crystal growth in the meta-stable zone	35

3.5	Seeding: streak seeding and macro seeding	37
3.6	Crystallographic analysis	37
3.6.1	Data collection, processing, and reduction	40
3.6.2	Phasing – molecular replacement	41
3.7	Model building and refinement strategy	46
3.7.1	Model refinement using REFMAC5	46
3.7.2	Inhibitor parameterization and ligand minimization for REFMAC	47
3.7.3	Electron density fitting	48
3.7.4	Modelling and refinement of solvent molecules	50
3.7.5	Structure validation	51
3.8	Structural comparisons	52
4	RESULTS	55
4.1	Six HIV-1 proteinase variants	55
4.2	Crystallographic analysis	58
4.3	Crystallization	60
4.3.1	Crystals of inhibitor-free HIV-1 proteinase, day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)	60
4.3.2	Crystals of inhibitor-free HIV-1 proteinase, day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)	61
4.3.3	Crystals of HIV-1 proteinase day 82 (Met36Ile-Ile54Val- Ala71Val-Val82Thr) in complex with ritonavir	61
4.3.4	Crystals of HIV-1 proteinase day 115 (Lys20Arg-Met36Ile- Ile54Val-Ala71Val-Val82Thr) in complex with ritonavir	62
4.4	Structural comparison between the HIV-1 proteinase variants	62
4.5	Crystallographic analysis – Ritonavir and saquinavir complexes	64
4.5.1	Variants of HIV-1 proteinase: day 28 ^b (Met36Ile-Ile54Val), day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr) and day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr) in complex with saquinavir	64
4.5.2	Proteinase-saquinavir interactions	66

4.5.3	Inter- and intra-monomer interactions in variants day 28 ^b , day 82, and day 115 in complex with saquinavir	80
4.5.4	Complexes with ritonavir of day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr) and day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)	86
4.5.5	Proteinase-ritonavir interactions	88
4.5.6	Inter- and intra-monomer interactions in variants day 82 and day 115 in complex with ritonavir	97
4.6	Structural analysis of inhibitor-free HIV-1 proteinase of variant day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr) and day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)	103
4.6.1	Description of the structures: inhibitor-free variant day 82 and day 115 in the closed form	104
5	DISCUSSION	110
5.1	Saquinavir (day28 ^b , day 82, day 115) and ritonavir (day 82 and day 115) complexes	110
5.2	Unliganded variants day 0, day 28 ^a , day 28 ^b , day 56, day 82 and day 115	111
6	CONCLUSIONS	116
	BIBLIOGRAPHY	117

ABBREVIATIONS

AIDS	Acquired Immune Deficiency Syndrome
CA	Capsid Protein
CCP4	Collaborative Computational Project Number 4
CCR5	Chemokine (C-C motif) Receptor 5
CD4	Cluster of Differentiation 4
CXCR4	Chemokine (C-X-C motif) Receptor 4
DNA	Deoxyribonucleic Acid
DTT	Dithiothreitol
EDTA	Ethylene Diamine Tetraacetic Acid
ENV	Envelope Protein
FDA	Food and Drug Administration
gp120	Glycoprotein 120
gp41	Glycoprotein 41
HAART	Highly Active Anti-Retroviral Therapy
HIV	Human Immunodeficiency Virus
IC ₉₀	Inhibitor Concentration for 90% Inhibition
IC ₅₀	Inhibitor Concentration for 50% Inhibition
IN	Integrase
k_{cat}	The Turnover number of the Enzyme
K_d	Dissociation Constant
K_i	Inhibitor Dissociation Constant
K_m	Substrate Binding Affinity
LBHB	Low-Barrier Hydrogen Bond
LTR	Long Terminal Repeat
MA	Matrix Protein
MES	2-(N-Morpholino)ethanesulfonic Acid
MR	Molecular Replacement
mRNA	messenger Ribonucleic Acid
NC	Nucleocapsid
NEF	Negative Factor
PDB	Protein Data Bank

PEG	Polyethylene Glycol
PI	Protein Inhibitor
PR	Proteinase
RC	Replication Capacity
REF	Anti-Repression Transactivator Protein
RNA	Ribonucleic Acid
RT	Reverse Transcriptase
RTp55	Reverse Transcriptase, Polymerase
RTp66	Reverse Transcriptase-RNaseH, Polymerase plus RNase H
TAT	Transactivating Regulatory Protein
TFP	Trans-Frame-Protein
VDW	Van der Waal
Vif	Virion Infectivity Factor
Vpr	Viral Protein R
Vpu	Viral Protein U

1 INTRODUCTION

AIDS stands for Acquired Immune Deficiency Syndrome. People with AIDS are infected with the human immunodeficiency virus (HIV), a retrovirus that damages the immune system. Since the immune system protects the body from illness, people with AIDS are more susceptible to all sorts of other health problems. Despite the enormous amount of research performed since the discovery in 1984 of HIV as the causative agent of AIDS, to date there is still no vaccine that can successfully prevent HIV infection, nor is there an antiviral strategy that can entirely remove the virus from infected people.

The present thesis is concerned with the crystallographic analysis of the molecular evolution of HIV-1 proteinase occurring in a patient under ritonavir-monotherapy. Nijhuis *et al.*, (1999) have reported that an increase in K_i values for ritonavir together with different fitness profiles are observed in the emerging, resistant variants. Interestingly, after merely 115 days of ritonavir therapy, two resistant proteinase variants have evolved with a slightly higher fitness level than that of the wild-type enzyme. To gain an understanding of how amino-acid substitutions modulate inhibitor binding and viral fitness, studies at the structural level have been conducted. It is electrifying to see the result of the cumulative substitutions on the X-ray structures of the six, inhibitor-free, HIV-1 PR variants under investigation, and on the structures of two complexes with ritonavir and three complexes with saquinavir.

In the following, a brief introduction to the HIV lifecycle, HIV-1 proteinase function, and current progress in drug development to combat HIV is given. Finally, the motivation for and scope of the current work is presented.

2 LITERATURE REVIEW

2.1 HIV structure and lifecycle

How does HIV evade the immune system? Why are so many variants of the virus found in a single patient? Understanding HIV structure and lifecycle is the key to answering these questions and essential for the design of effective treatments. HIV has several major genes coding for structural proteins (p24, p6, p7, p17, gp120 and gp41), nonstructural proteins (reverse transcriptase, integrase and protease) that are found in all retroviruses, and several accessory genes (*vif*, *nef*, *tat*, etc.) that are unique to HIV (Table 2.1). HIV-1 is composed of two copies of (+)-sense single-stranded RNA (diploid genome) enclosed by a conical capsid comprising the viral protein p24 (Figure 2.1). By associating with the genomic RNA, the nucleocapsid protein p7 protects the RNA from digestion by nucleases.

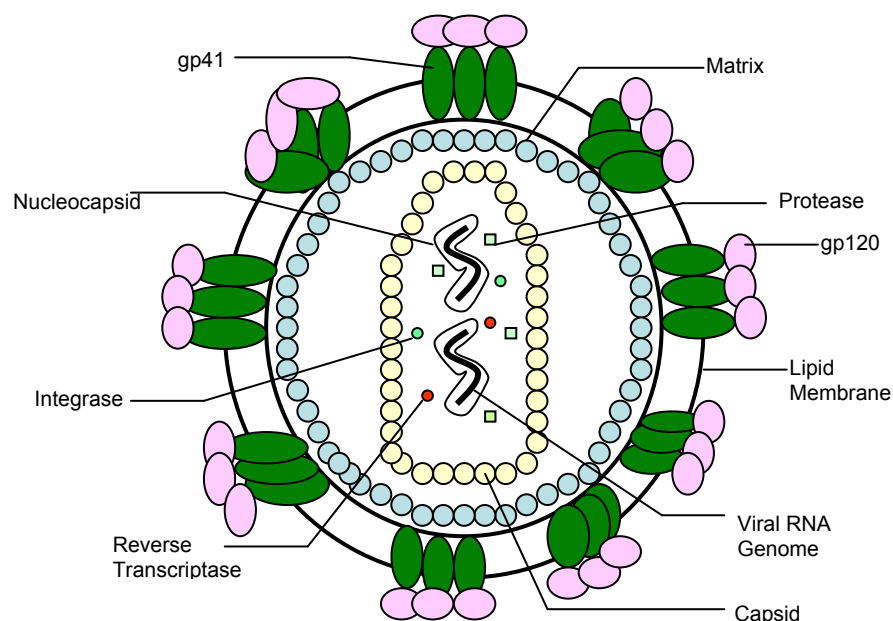


Figure 2.1: The HIV virion particle.

The integrity of the virion capsid is ensured by the surrounding matrix composed of the viral protein p17. Vif, Vpr, Nef, p7 and viral proteinase are also enclosed within the virion particle. When the HIV-1 virion buds from the host-cell surface, a lipid envelope forms by taking some host-cell membrane with it. The envelope contains glycoproteins gp41 bound non-covalently to gp120 on the lipid surface. These two glycoproteins are synthesized as a single 160 kDa precursor that is then cleaved by the cellular host-cell's prohormone convertases PC1 and furin

(Decroly *et al.*, 1997). The products of this cleavage reaction remain associated until the process of viral entry into the cell begins.

Table 2.1: Nonstructural and structural HIV proteins.

Gene	Protein	Function
<i>pol</i>	<i>Nonstructural</i>	
	Reverse transcriptase	This enzyme transcribes the viral RNA into double-stranded DNA.
	Integrase	Enzyme that integrates the provirus DNA produced by RT into the host's genome.
	Proteinase	Enzyme that cleaves the polyproteins derived from <i>gag</i> and <i>pol</i> into functional proteins.
	<i>Regulatory proteins</i>	
<i>tat</i>	Tat	Binds to HIV-LTR TAR-sequence and activates transcription.
<i>rev</i>	Rev	Binds to the rev-responsive-element (RRE) of viral mRNAs and regulates the expression of HIV proteins by controlling the export of mRNAs from nucleus to cytoplasm.
<i>nef</i>	Nef	The expression of Nef early in the viral lifecycle ensures T-cell activation and establishment of a persistent state of infection, two basic attributes of HIV infection. Nef down-regulating the CD4 molecules on the cell surface.
<i>vif</i>	Vif	Affects virion assembly and it is an antagonist of innate cellular protein APOBEC3G.
<i>vpr</i>	Vpr	Plays an important role in regulating nuclear import of the HIV-1 pre-integration complex, and is required for virus replication in non-dividing cells.
<i>vpu</i>	Vpu	Involved in viral budding, enhancing virion release from the cells. In HIV-2, the <i>vpu</i> gene is called <i>vpx</i> .
<i>tev</i>	Tev	Present in a few HIV-1 isolates. It is a fusion of parts of the <i>tat</i> and <i>rev</i> genes and codes for a protein displaying some of the properties of Tat but little or none of the properties of Rev.

	<i>Structural</i>	
<i>gag</i>	p24	Viral capsid protein.
	p6	Mediates the interactions between p55 and Vpr, leading to the incorporation of Vpr into assembling virions.
	p7	Nucleocapsid protein.
	p17	Matrix protein.
<i>env</i>	gp120	Envelope surface protein that binds to the CD4 receptor on any target cell that has such a receptor, particularly helper T-cell, monocytes and macrophages.
	gp41	The interaction of gp120 with CD4 changes gp120's shape. This in turn changes the shape of gp41 in such a way that this protein helps initiate the process of membrane fusion between virus and cell.

Although helper T-cells seem to be the main target of HIV, other cells can become infected as well. These include monocytes and macrophages, which can accommodate large numbers of virus particles. Entry of HIV into the host-cell requires the binding of one or more gp120 surface molecules on the virus to CD4 receptors on the host-cell's surface. Binding to a second receptor is also required. There are two different co-receptors, chemokine receptors CCR5 and CXCR4, which act at an early or late state of infection, respectively (Coffin *et al.*, 1997).

HIV belongs to the group of retroviruses (Baltimore's class VI), all of which share a unique lifecycle (Figure 2.2). Once HIV binds to a host-cell, the viral envelope fuses with the cell membrane, and the viral capsid with RNA enters the cytoplasm. HIV, like other retroviruses, contains an enzyme called reverse transcriptase. This allows the (+)-sense single-stranded RNA of the virus to be reverse-transcribed and the double-stranded DNA (called provirus) to be generated. The enzyme integrase then facilitates the integration of this viral DNA into the cellular genome. The provirus is replicated along with the genome when the cell divides. The integration of the provirus into the host's DNA provides the latency that enables the virus to evade the host immune responses so effectively (Wyatt and Sodroski, 1998).

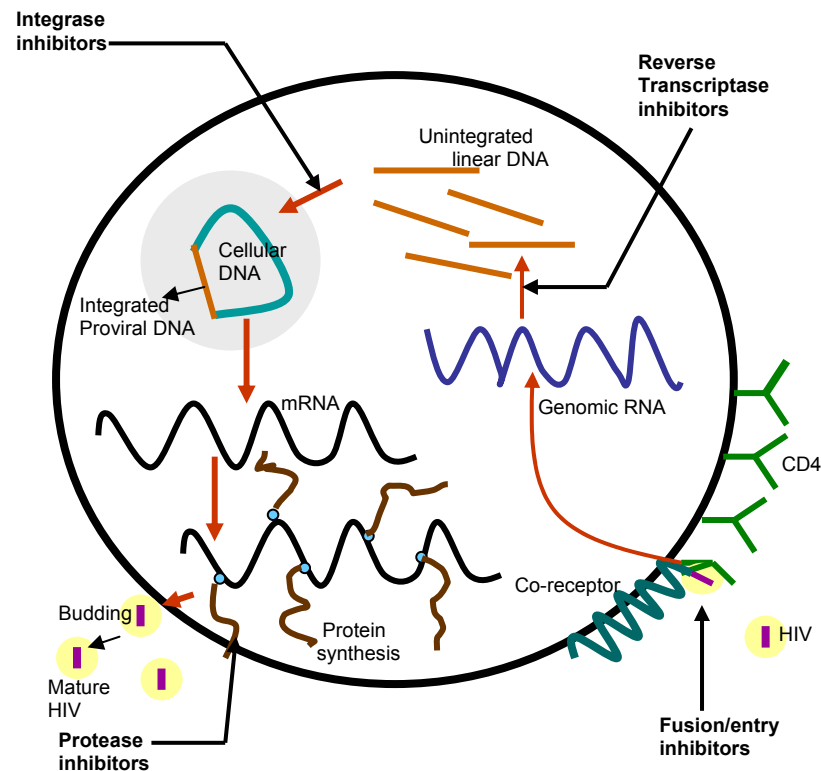


Figure 2.2: HIV lifecycle, including points of potential therapeutic intervention.

The integrated viral DNA can be transcribed into mRNA (messenger RNA) and after being transported to the cytoplasm, the mRNA is translated into large polyproteins. The components of the virus assemble near the cell membrane and a new virus particle forms by a 'pinching' action on the host's lipid membrane. The new virus buds off from the host-cell and as it buds, maturation takes place. During maturation, HIV proteinase cleaves the polyproteins into the individual functional HIV proteins.

The RNA of retroviruses is replicated through a DNA intermediate, the product of the virus-encoded reverse transcriptase (RT), which is an error-prone DNA-polymerase that lacks a proofreading step. RT makes approximately five to ten errors per round of genome replication (Preston *et al.*, 1988). Selective pressures affect replication, cell tropism (the ability of a virus to enter particular cell types) and escape from host immunity, and thus contribute to the genetic differences in HIV-1 isolates between individuals (Barrie *et al.*, 1996). There is no 'wild-type' HIV-1 in an infected individual, but rather a complex mixture of related sequences also referred to as quasispecies (Eigen, 1993).

Combating viruses is always difficult because they make use of the translational machinery of the host-cell. Most inhibitors that target the virus also damage the host. Compounds that inhibit enzymes specific to the virus are therefore less likely to cause side effects in the host. The current anti-HIV drugs block key steps in viral reproduction and uptake (Figure 2.2). Several anti-retroviral drugs work by binding to RT, the key enzyme of retroviruses (Erickson and Burt, 1996). Examples of this group of drugs are the nucleoside analogues (pro-drugs without phosphate) like abacavir, didanosine, emtricitabine, lamivudine, stavudine, zalcitabine, zidovudine and others; these compounds only work in cells that have the machinery to activate the pro-drug through phosphorylation. Further, nucleotide analogues like tenofovir [(R)-9-(2-phosphonylmethoxypropyl)adenine] require no supplementary phosphorylation. There are also non-nucleoside RT inhibitors, for example delavirdine, efavirenz and nevirapine. Proteinase inhibitors, another major class of drugs, act later in the lifecycle of the virus. These drugs interfere with the cleavage of the viral polyproteins into functional viral proteins (Wlodawer and Vondrasek, 1998). The integrase is necessary for integration of the provirus into the host's DNA and specific inhibitors can block the enzyme. Recently, GS 9137 has been discovered by Japan Tobacco, Inc., as a novel HIV integrase inhibitor that is currently in phase-II clinical trials (<http://clinicaltrials.gov/show/NCT00298350>). Another class of drugs are the fusion inhibitors. Enfuvirtide and T-20 act by binding to the leucine zipper-region of the HIV envelope-glycoprotein gp41, thereby preventing viral fusion with the target-cell membrane.

Despite the progress in developing new anti-viral regimes, a fully effective therapy for AIDS is not yet available. Unfortunately, emerging resistance and serious drug side effects pose strong limitations onto the treatment efficacy. In response to unsuccessful mono-drug therapies, clinicians developed drug regimes called HAART (highly active anti-retroviral therapy) and mega-HAART, in which a combination of three or more drugs is applied simultaneously. These cocktails of drugs can lower the virus titers to undetectable levels and extend patient life. Unfortunately, HAART has several long-term side effects including kidney, liver and pancreatic problems and changes in fat metabolism, which result in elevated cholesterol and triglyceride levels and an increased risk of strokes and heart attacks. In addition, some HIV mutants have gained resistance to HAART. Given these side effects, HAART treatment is intentionally delayed until HIV-positive patients exhibit clear signs of AIDS (Spire *et*

al., 2002).

2.2 HIV genome and structure components

There are two causative agents of AIDS, HIV-1 and HIV-2. Epidemiologic, clinical and virologic data suggest that HIV-2 is less virulent than HIV-1. The reduced virulence of HIV-2 is marked by a longer asymptomatic period and a reduced transmission rate (Pereira *et al.*, 2005). The HIV genome contains three major genes, *gag*, *pol* and *env* as listed in Table 2.1. HIV genes are organized as shown in Figure 2.3. The HIV genome also harbours two identical 'long terminal repeats' (LTRs), one at each end of its genome. The LTRs fulfil two important functions. First, they are 'sticky ends' that the integrase uses for inserting the provirus into the host-cell DNA. Second, when integrated into the host genome, LTR acts as promoter or enhancer elements, influencing the cell machinery that transcribes the proviral DNA.

2.3 HIV-1 proteinase cleavage sites

The cleavage reactions catalyzed by HIV-1 proteinase (PR) occur late in the viral life-cycle, during virion assembly and maturation at the cell surface. The process is highly specific, temporally regulated and essential for the production of infectious virus particles (Kräusslich *et al.*, 1989; Swanstrom and Wills, 1997). In total, 11 proteolytic reactions (Figure 2.5) are required to generate a viable virion (Billich *et al.*, 1988). Translation of the retroviral *gag-pol* mRNA produces, in most cases, a Gag polyprotein of 55 kDa, ending before the PR gene (Kräusslich *et al.*, 1989). In about 5% of the *gag-pol* transcripts, a translational frameshift occurs upstream of the PR gene. As a consequence, the stop codon of the *gag* locus is no longer in frame, thus resulting in the production of a Gag-Pol fusion protein (Figure 2.4).

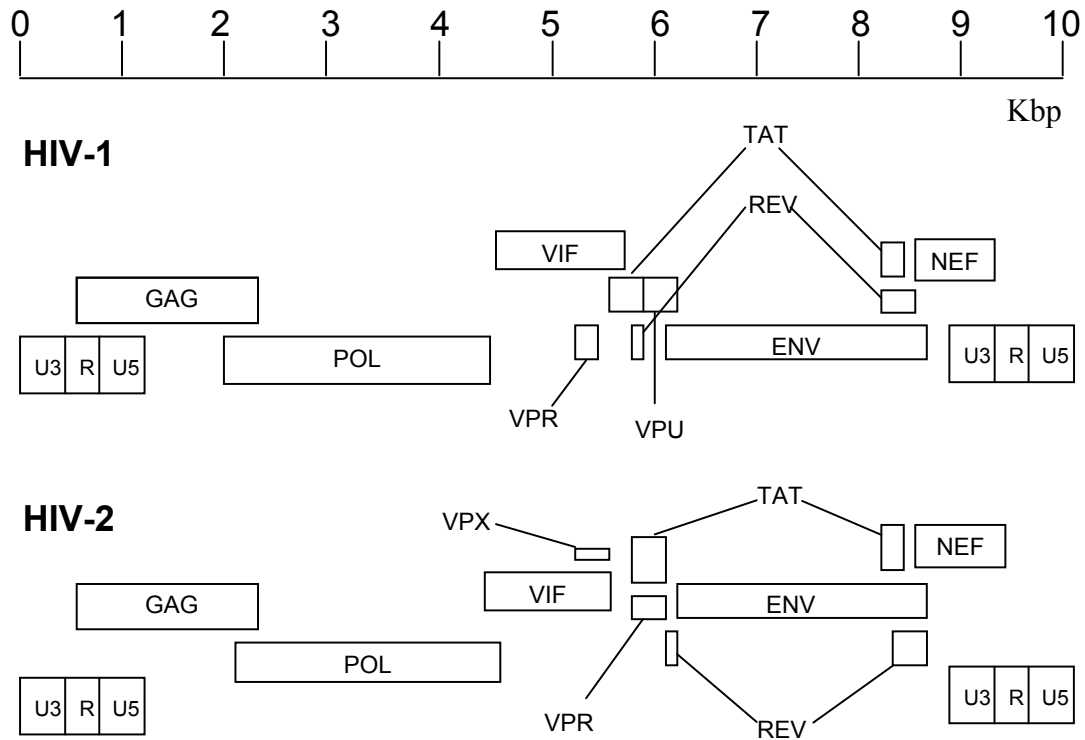


Figure 2.3: Genome map of HIV-1 and HIV-2. Long terminal repeat (LTR) comprises U3, R and U5 regions. Descriptions of the proteins are given in Table 2.1.

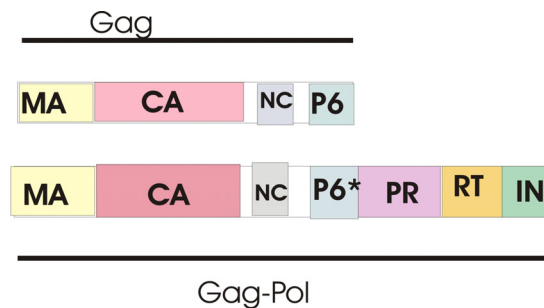


Figure 2.4: Translational products from the retroviral *gag-pol* mRNA. A translational frameshift within the P6 region allows translation beyond the p6 *gag* gene, resulting in a Gag-Pol fusion protein. The Gag-Pol fusion protein contains a p6* protein, the sequence of which differs from the p6 protein as a result of the frameshift. MA, p17 matrix protein; CA, p24 capsid protein; NC, p7 nucleocapsid protein; PR, protease; RT, reverse transcriptase; IN, integrase.

The rates of cleavage can differ by up to 400-fold between sites (Erickson-Viitanen *et al.*, 1989). These differences may be related to different steps in the assembly of virions. As shown in Figure 2.5, the Pr160Gag-Pol polyprotein is cleaved into matrix protein (MA), capsid (CA), p2, nucleocapsid (NC), trans-frame-protein

(TFP), p6pol, protease (PR) and integrase (IN). In addition, either RTp66 (RT-RNaseH) or the shorter RTp51 (RT) is generated. Pr55Gag is proteolytically cleaved during maturation to yield matrix protein (MA), capsid (CA) and nucleocapsid (NC), in addition to the low-molecular mass cleavage products, p1, p2 and p6gag. HIV p1 and p2 proteins have no known function.

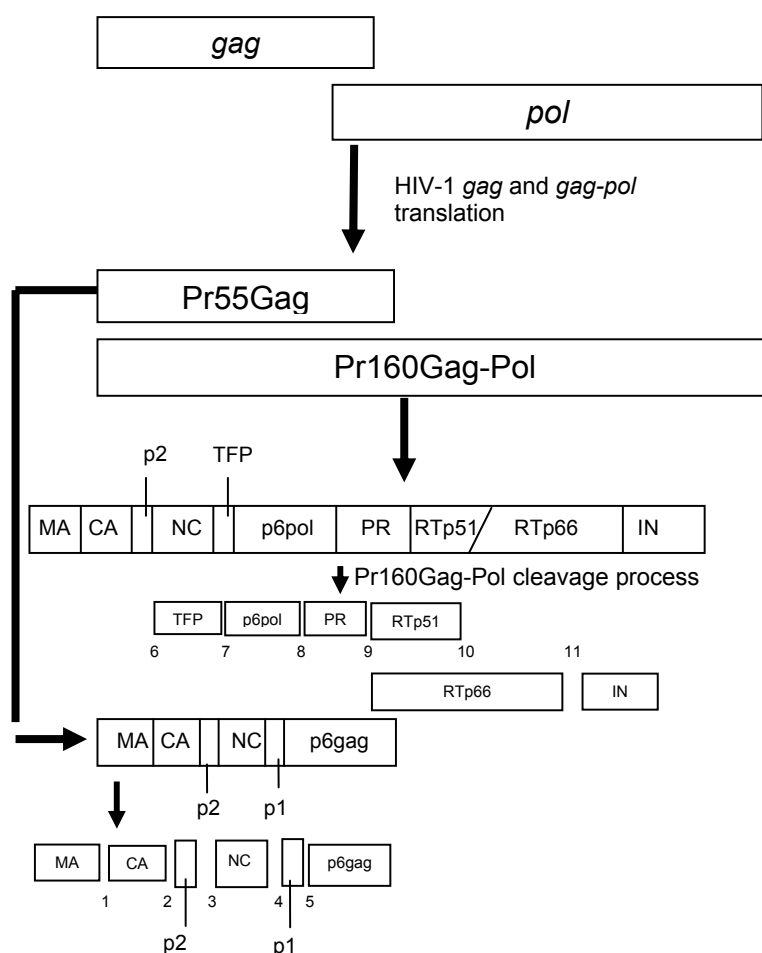


Figure 2.5: Proteins formed by the 11 proteolytic reactions necessary to generate a viable virion. Abbreviations: MA, p17 matrix protein; CA, p24 capsid protein; NC, p7 nucleocapsid protein; p2; p1; p6gag; TFP, trans-frame-protein; p6pol; PR, proteinase; RTp51, reverse transcriptase; RTp66, RT-RNase H; IN, integrase. Descriptions of the proteins are listed in Table 2.1.

Three of the cleavage sites are either tyrosine-proline or phenylalanine-proline bonds, and are unique to retroviral proteinases. These cleavage sites would be unusual for mammalian proteases (Flexner, 1998): therefore, Phe-Pro was used as the basis for selectivity for most HIV proteinase inhibitors (Kiso, 1996; Erickson and Burt, 1996; Flexner, 1998).

2.4 HIV-1 proteinase (PR)

2.4.1 Characterization of HIV-1 PR

Navia *et al.* (1989), were the first to obtain a crystal structure of the HIV-1 proteinase (HIV-1 PR) and since then, the PR structure has been extensively characterized in terms of function, substrate specificity and inhibitor binding. In the same year, Wlodawer and colleagues at the Frederick Cancer Research Facility in Maryland refined and corrected the PR structure since the one solved by Navia and colleagues was partly wrong (Wlodawer *et al.*, 1989). HIV-1 PR is an aspartyl proteinase belonging to the peptidase family A2, retropeptidases (Davies, 1990). It functions as a homodimer with one active-site. Each monomer consists of 99 amino-acid residues (Figure 2.6). The HIV PR structure can be divided into three parts: the flaps, the dimerization sheet and the core (Figure 2.6). It is known that the HIV PR nucleotide sequence slightly changes from generation to generation. This is the result of a continuous selection for function (e.g. drug resistance) and hence, PRs purified from HIV infected people will differ in amino-acid sequence (quasispecies). However, residues at the catalytic dyad (Asp25 of each monomer) are conserved throughout generations.

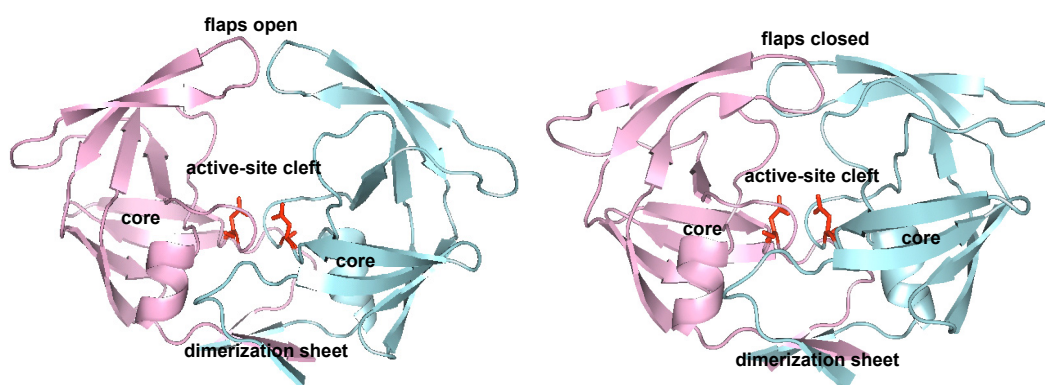


Figure 2.6: The Ca traces of the open and closed structures of the HIV-1 proteinase. Chain A is coloured in pink and chain B in cyan. The active-site residues are shown as red sticks.

The dimerization sheet consists of four β -strands comprising residues 1-4 and 95-99 from each monomer. This part is crucial to dimer formation and stabilizes the active PR. The N-terminal β -strand (Figure 2.6) forms the outer part of the interface β -sheet. The flaps, composed of an antiparallel β -sheet with a β -turn, are known to be

very flexible. They can exist in two conformations; the open conformation allows the polyprotein to access the active-site and the closed conformation embraces the substrate for efficient cleavage (Ishima *et al.*, 1999). There are also two cores, one per monomer, composed of four β -strands plus one α -helix.

As shown in Figure 2.7, the PR chain starts at β -strand 'a' (residues 1-4), continues through a loop into β -strand 'b' (residues 10-15) followed by a turn into β -strand 'c' (residues 18-23) that ends at the active-site (Asp25). The conserved Asp25-Thr26-Gly27 sequence is situated at the interface between cores from the two monomers. Following the active-site loop is β -strand 'd' (residues 30-35), followed by an extended loop that is connected to the mobile flap (residues 36-61), made up of β -strands 'e' and 'f'. Following are β -strands 'g' (residues 62-66) and 'h' (residues 69-78) and a long loop that is connected to the well-defined helix 'i' (residues 87-93). The hydrogen-bonding pattern within this helix corresponds to either α - or 3_{10} -helix for the N- or C-terminal part, respectively. Helix 'i' is followed by the C-terminal β -strand 'j' (residues 96-99) that forms the inner part of the dimer interface. The N- and C-termini of each monomer interdigitate to form a compact four-stranded sheet at the dimer interface.

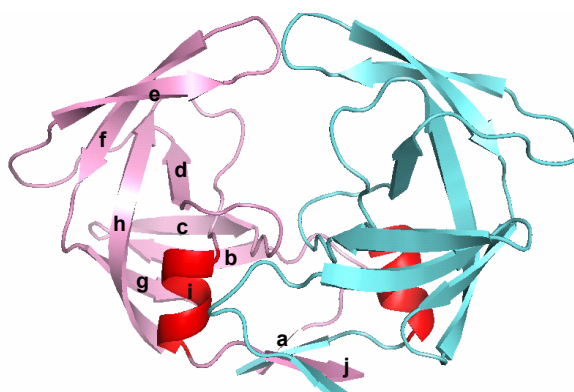


Figure 2.7: Topology of the HIV-1 PR structure. Chain A is coloured in pink and chain B in cyan: β strand a, residues 1-4; b, residues 10-15; c, residues 18-23; d, residues 30-35; e, residues 43-49; f, residues 53-59; g, residues 62-66; h, residues 69-78; j, residues 96-99. Helix i (coloured in red), residues 87-93.

The HIV-1 PR, upon maturation of the Gag-Pol precursor, forms a stable homodimer. The following factors contribute to the observed dimer stability (Ingr *et al.*, 2003):

- (1) The compact four-stranded β -sheet of N- and C-termini at the homodimer

interface.

(2) The conserved, complex scaffold of hydrogen bonds (termed fireman's grip) supporting the active-site residues.

Ingr *et al.* (2003) proposed that the fireman's grip might be important in the initial stage of dimer formation in order to help properly orientate the two subunits of a retroviral proteinase. The fireman's grip is formed by the two Asp25-Thr26-Gly27 sequences of the homodimer: the hydroxyl group of Thr26 of one protein chain forms two hydrogen bonds to the other chain of the dimer: namely, the carbonyl group of Leu24 and the amide group of Thr26 (Figure 2.8).

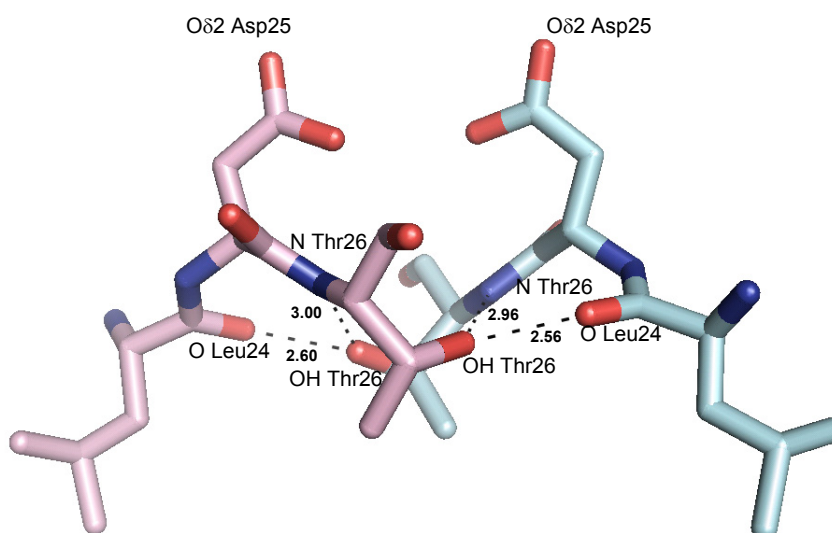


Figure 2.8: Details of the active-site of the HIV-1 proteinase with the fireman's grip. Hydrogen bonds are shown as dashed lines. Chain A is coloured pink and chain B is coloured cyan. Distances are in Å.

2.4.2 The active-site of aspartyl proteases

Aspartyl proteases, also known as acidic proteases, constitute a widely distributed family of proteolytic enzymes known to exist in vertebrates, plants, plant viruses, fungi, and protozoa as well as in retroviruses. This family of enzymes features two highly conserved Asp-Thr/Ser-Gly triplets (Pearl and Taylor, 1987). Aspartate proteases in general, with the exception of retroviral proteinases that are dimers of two identical subunits (with one Asp-Thr/Ser-Gly triplet per chain), are monomeric enzymes. HIV-1 PR harbours, in addition to its family-classifying Asp-

Thy-Gly sequence, a second highly conserved sequence of Gly-Arg-Asn (Asp in retroviruses other than HIV), namely amino-acid residues 86-88 (Pearl and Taylor, 1987). This is unique to the viral enzyme and is not present in cellular aspartic proteases. It is believed that the highly conserved Arg87 of HIV-1 PR is involved in ion pairing with the similarly conserved Asp29 to form a specific structure for substrate binding (Blundell and Pearl, 1989). The Asp25 of monomers A and B are essential to PRs, both catalytically and structurally. The β -carboxyl groups of the two Asp25 residues are nearly coplanar, with close contact between the O δ 1 atoms. On the basis of experimental work, there have been several proposals for the reaction mechanism (Hyland *et al.*, 1991a,b; Rodriguez *et al.*, 1993; Chatfield and Brooks, 1995; Lee *et al.*, 1996; Silva *et al.*, 1996; Chatfield *et al.*, 1998; Park *et al.*, 2000; Northrop, 2001). Initially, a nucleophilic mechanism for the aspartic proteinases was proposed (Hsu *et al.*, 1977). This mechanism implies formation of a covalently bound PR-substrate complex during the course of reaction. There is as yet no direct experimental support for this mechanism, although recent state-of-the-art computational studies do indicate that a direct nucleophilic reaction is feasible (Park *et al.*, 2000; Trylska *et al.*, 2004). In contrast, the general acid/base mechanism as originally proposed by Fruton (1987) is supported by numerous studies. According to the general acid/base mechanism, a lytic water molecule takes part in the reaction. Notable, a water molecule hydrogen bonded to the O δ 1 and O δ 2 atoms of Asp25 from each monomer is usually observed in the unliganded HIV-1 PR crystal structures (Figure 2.9). Common to both these mechanisms is the requirement that one of the catalytic aspartates be charged while the other remains neutral. Moreover, the stability and activity of HIV-1 PR as a function of pH is roughly between pH 3.5 to 6.5 with the maximum stability and enzymatic activity occurring at pH 5.0-6.0 (Billich *et al.*, 1990; Hyland *et al.*, 1991a; Ido *et al.*, 1991; Polgar *et al.*, 1994).

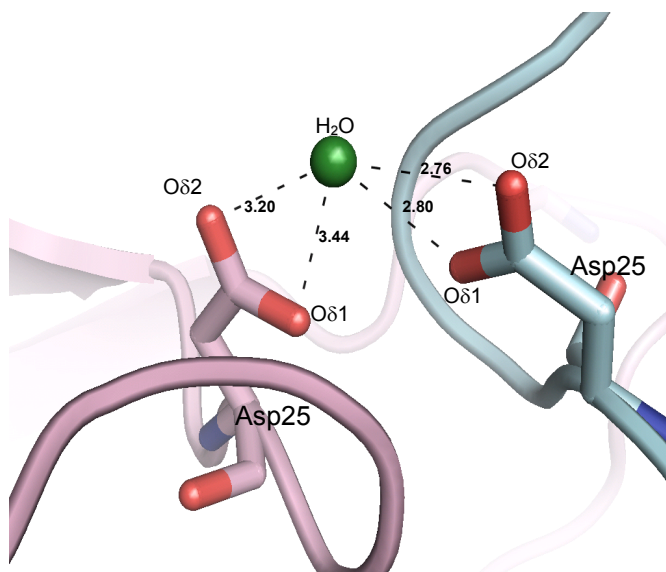


Figure 2.9: The active-site of the HIV proteinase with a water molecule (green ball) making hydrogen-bonding contacts to O δ 1 and O δ 2 of Asp25 (shown as sticks) of chains A (pink) and B (cyan). Distances are in Å.

Most studies indicate a general acid/base mechanism for peptide bond cleavage, in which the hybridization of the peptide carbon is changed from sp^2 to sp^3 through a nucleophilic attack by a lytic water molecule (Figure 2.10). One of the aspartates helps in ionizing the water molecule, taking away its proton. The resulting hydroxy- anion attacks and binds covalently to the peptide carbon, initiating C-N bond breakage.

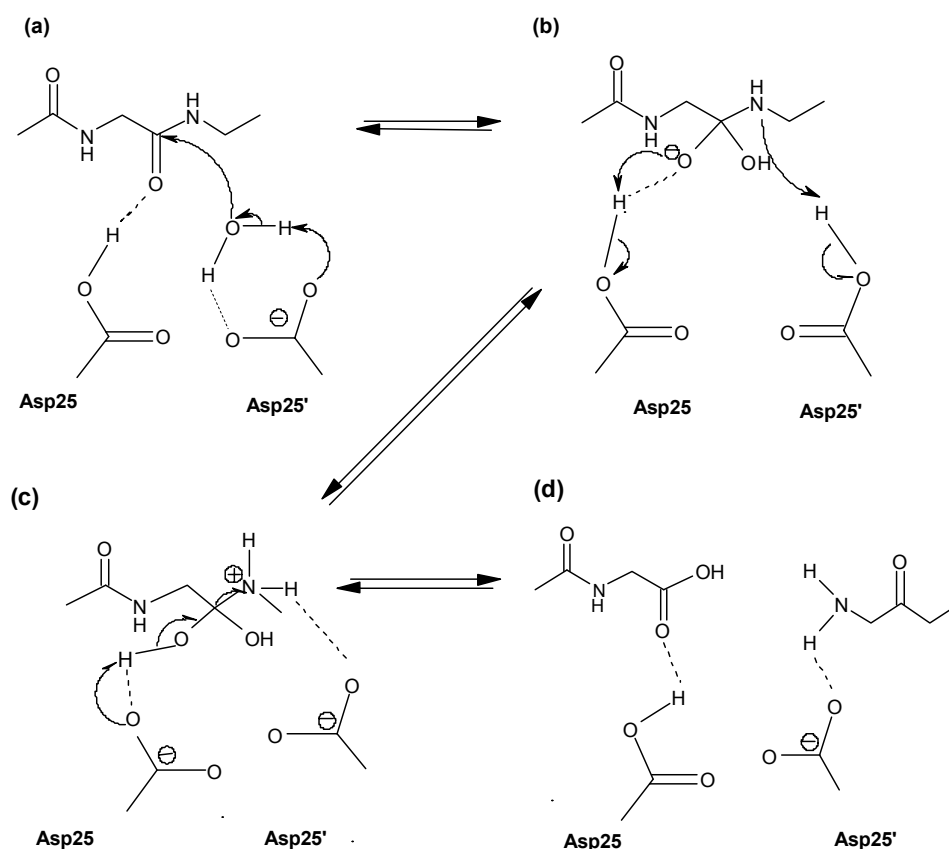


Figure 2.10: HIV proteinase general acid/base mechanism with the initial proton positioned on the outer oxygen of Asp25, (a) Following the dissociation of H₂O, OH⁻ acts as a nucleophile on the peptide carbon; (b) two proton transfer, from Asp25 onto carbonyl O and from Asp25' onto peptide N; (c) C-N bond breakage and return of proton to Asp25; (d) final state.

Recently, Tyrłska *et al.* (2004) used quantum-classical molecular dynamics simulations to study the hydrogen-bonding network in various stages of the enzymatic reaction catalyzed by HIV-1 PR. In a general acid/base mechanism with Asp25 protonated at the inner oxygen atom (Oδ1), they observed a short hydrogen bond between the inner oxygens of the catalytic dyad (Figure 2.11). However, the hydrogen bond was not a single-well type and asymmetry was observed, that is, the proton remained bonded to Asp25 in most cases or it transferred and bonded covalently to Asp25'. This short distance and the possibility of proton transfer on a small time scale, indicate that this bond is a low-barrier hydrogen bond (LBHB), first suggested by Gerlt and Gassman (1993), Cleland and Krevoy (1994), and Frey *et al.*, (1994). Formation of a LBHB between the inner oxygens of aspartates in the HIV-1 PR active-site was confirmed by Northrop (2001) on the basis of isotope effects and by Piana and Carloni (2000) on the basis of *ab-initio* molecular dynamics. The general

acid/base mechanism with Asp25 protonated at the inner oxygen atom is shown in Figure 2.11 (Trylska *et al.*, 2004). Asp25/25' and the catalytic water induce simultaneously general acid/base protein hydrolysis. The initial proton on O δ 1 of Asp25 interacts with O δ 1 of Asp25' as indicated by the dashed line. The Asp25 closest to the catalytic water molecule acts as a base to accept a proton from the water molecule (a). The resulting hydroxide ion is a better nucleophile than aspartate and attacks the carbonyl of the substrate to generate an oxyanion tetrahedral transition-state. Asp25' act as an acid and donates a proton to the peptide nitrogen (b) and the subsequent rearrangement results in the breakdown of the tetrahedral intermediate (c) to yield the hydrolysis products (d).

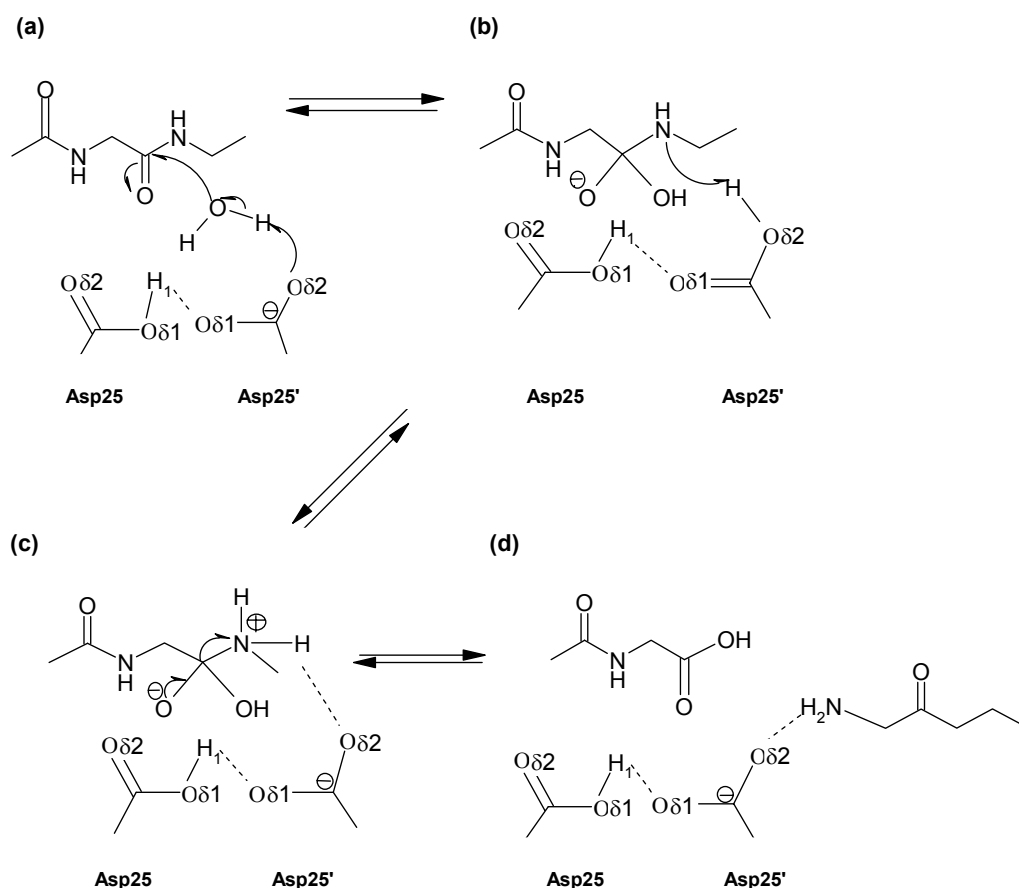


Figure 2.11: HIV proteinase catalytic mechanism with Asp25 protonated at the inner oxygen, O δ 1. (a) Following the dissociation of H₂O, OH⁻ acts as a nucleophile on the peptide carbon; (b) proton transfer from Asp25' onto peptide N; (c) peptide bond breakage; (d) final state.

2.4.3 The flaps

The flaps consist of residues 46-56 of each monomer and can be characterized as follows (Figure 2.12):

- (1) Side-chains that point outwards (Met46, Phe53, and Lys55)
- (2) Hydrophobic side-chains that point inwards (Ile47, Ile50, and Val54)
- (3) A glycine-rich tip (residues 48-52)

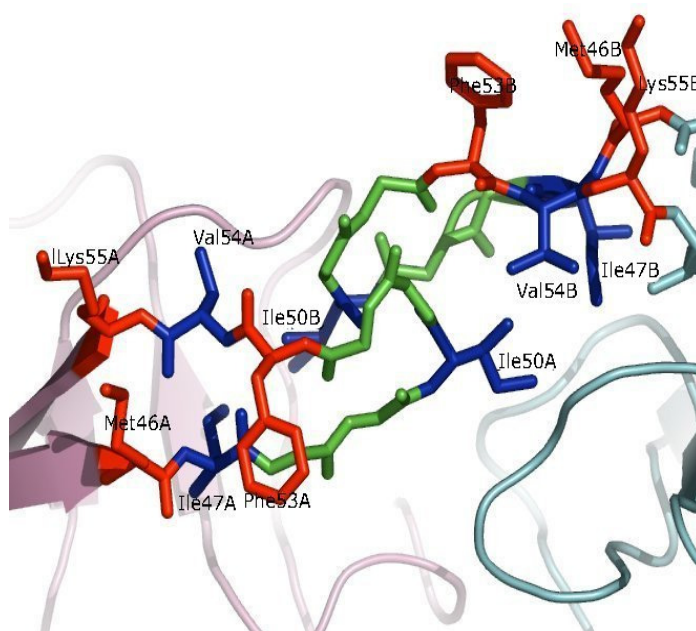


Figure 2.12: Flap region. Met46, Lys55, and Phe53 (red, pointing outwards); Ile47, Val54 and Ile50 (blue, pointing inwards); Gly48, Gly49, Gly51, and Gly52 (green, glycine-rich tip). Chain A is coloured pink and chain B is coloured cyan.

The high flexibility and mobility of the flaps are thought to be necessary for substrate binding and product release. It has been observed that in most of the crystal structures of HIV PR bound to different inhibitors, a unique water molecule is present between the flaps and the inhibitors (Figure 2.13). This water molecule is known as the flap water and makes hydrogen bonds with the carbonyl group of P1 and P1' of the inhibitor and the amide group of Ile50 and Ile50' (Birk and Wong, 2003).

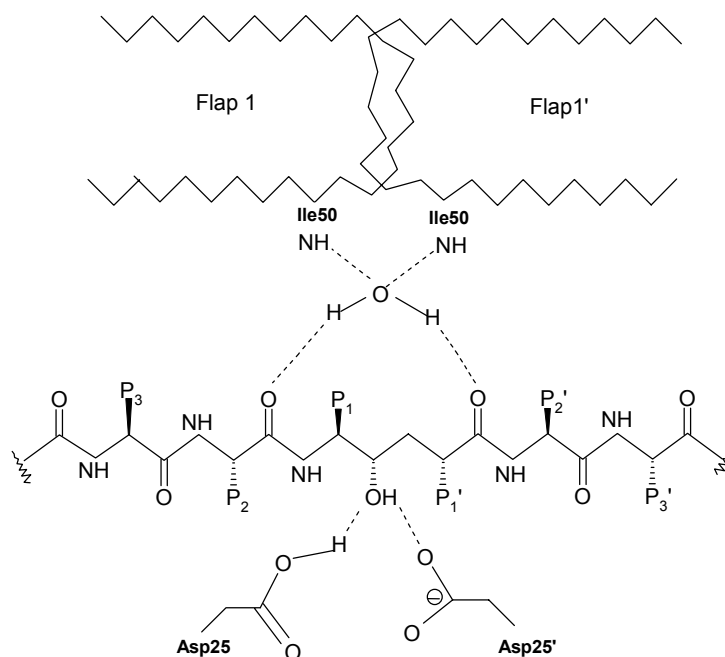


Figure 2.13: Schematic representation of the two flaps of HIV proteinase and their hydrogen bonds with the flap water. This water molecule makes two additional hydrogen bonds with the peptidic inhibitor. Also shown is the interaction of the central hydroxyl group of the inhibitor with the catalytic aspartates of the enzyme.

2.5 Anti-HIV drug design

Chemotherapy has focused on the discovery and design of inhibitors for two main enzymes of HIV-1, reverse transcriptase (RT) and proteinase (PR). Despite several classes of promising new anti-HIV agents that have been developed, the emergence of drug-resistant variants of HIV has limited the long-term effectiveness of current drugs. The crucial role of the PR in the viral lifecycle has made it an important therapeutic target. All the available proteinase inhibitors (PIs) were designed after the first three-dimensional structure of HIV-1 PR was solved (Navia *et al.*, 1989), indicating the importance of structure-based drug design (Hillisch and Hilgenfeld, 2002). However, efficacy of these drugs is limited due to the natural selection of PR variants that are still catalytically competent while displaying a lower affinity for the PIs than the wild-type enzyme (Abdel-Rahman *et al.*, 2002).

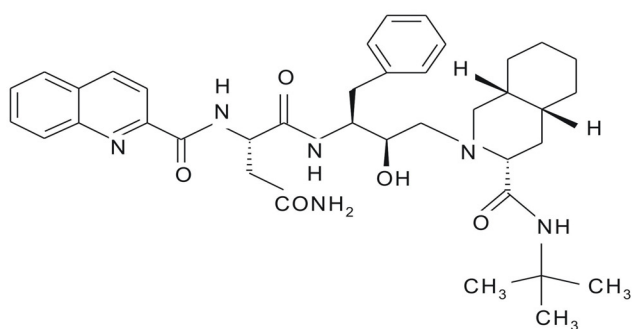
2.5.1 HIV-1 PR inhibitors

Knowledge of the active-site and the mechanism of catalysis have been widely exploited in the design of PIs. They interrupt virus replication at a late step in the HIV lifecycle. PIs can reduce the amount of virus in the blood and increase CD4 cell counts, helping preserve or even restore immune system functions. The US Food and Drug Administration (FDA) approved the use of several HIV PR inhibitors such as amprenavir (VX-478), indinavir (MK-639), nelfinavir (AG-1343), saquinavir (RO 31-8959), ritonavir (ABT-538), lopinavir (ABT-378), atazanavir (BMS-232632), and fosamprenavir (VX-175). A few first-generation HIV-1 proteinase inhibitors are shown in Figure 2.14. Despite the considerable success of rational drug design, the need for more effective PIs is still urgent. Most of the current PIs are pseudopeptide compounds with limited bio-availability and -stability.

The use of proteinase inhibitors in HIV patients over a longer period results in resistant variants of the HIV PR. All of the current drugs are reversible binders, even though some of them bind very tightly to the active-site (Appelt *et al.*, 1991). Another promising approach is to design irreversible inhibitors able to bind covalently to the catalytic aspartates of the enzyme. One example of an irreversible inhibitor is a symmetric epoxide-containing compound (Turner, 2002). If properly oriented within the active-site, an epoxide functional group could become protonated by one of the catalytic aspartates acting as a nucleophile to open the epoxide ring. The resulting covalent modification would destroy the catalytic activity of the enzyme. However, such a design is quite challenging, because it is necessary to maintain the correct orientation of the epoxide ring so that it can make proper contact with the two active-site aspartates. Several inhibitors based on these concepts were developed and found to be potent compounds against HIV PR. Although some of these inhibitors were clinically tested, none of them has yet reached the market (Brik and Wong, 2003).

Name
Synonyms

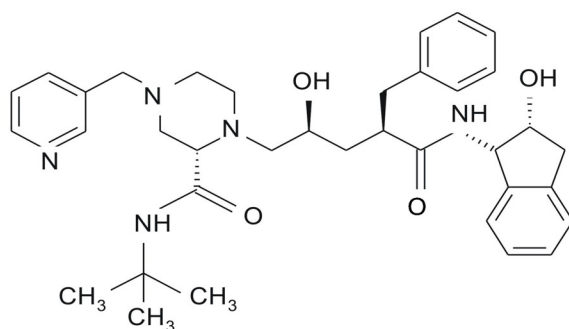
Saquinavir
Invirase
Fortovase
Ro3-8959



Launch

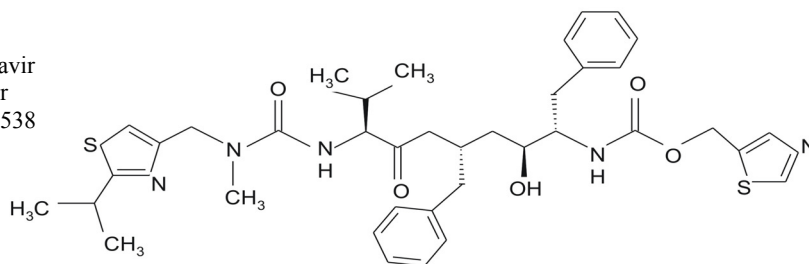
12/95

Indinavir
Crixivan
L735, 534
MK 639



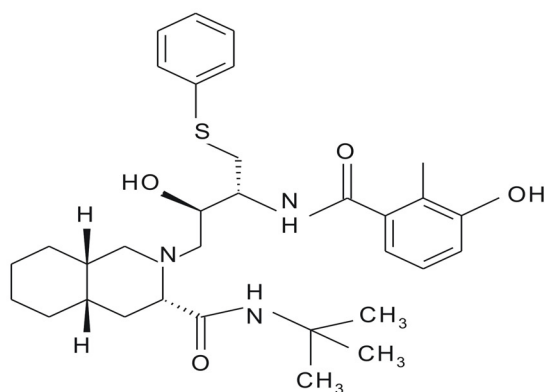
3/96

Ritonavir
Norvir
ABT-538



3/96

Nelfinavir
Viracept
AG 1343



3/97

Figure 2.14: First generation HIV-1 proteinase inhibitors.

The conformational stability of the PR and its catalytic activity are both enhanced at high salt concentrations (1 M) as compared to low salt conditions (0.1 M). Low-molecular-weight ligands with an affinity for the dimer interface could prevent the dimerization. Short peptides from the termini of HIV PR were confirmed to be weak inhibitors of the enzyme but showed a better inhibition in *in-vitro* studies than in cell cultures (Schramm *et al.*, 1993). The reported K_d values for HIV-1 PR dimer dissociation range from a wild-type value of 39 pM at pH 6.0, 25°C and 0.25 M ionic strength, as followed by spectroscopic monitoring of unfolding (Grant *et al.*, 1992), to 5.8 μ M for an autolysis-resistant mutant (containing substitutions Gln7Lys, Leu33Ile, Leu63Ile) assayed at pH 7.0 and 4°C by sedimentation equilibrium analysis (Xie *et al.*, 1999).

Recently, Cígler *et al.* (2005) reported a group of inorganic compounds, icosahedral metallocarboranes, as candidates for a novel class of competitive non-peptidic PIs. These compounds block proper flap closure in addition to filling the corresponding binding pockets as conventional PIs do.

The currently available HIV-1 PIs can be divided into two classes:

- (1) Peptide-based inhibitors (peptidomimetics and pseudo-symmetric inhibitors)
- (2) Non-peptide based inhibitors such as cyclic ureas and sulfonamides.

2.5.1.1 Peptide-based inhibitors

Most peptide-based drugs are designed so that they imitate substrate binding in terms of the hydrogen-bonding pattern. Hydrogen-bonding between Ile50 and Ile50' at the tips of the protein flaps and the inhibitor is a common feature in PR-inhibitor complexes. Most of the peptide inhibitors also form hydrogen bonds with the catalytic dyad Asp25 and Asp25' residues (Tie *et al.*, 2004). Further, many of the inhibitors utilize the flap water molecule found in the cleft of active-site as a link to the flaps (Figure 2.15).

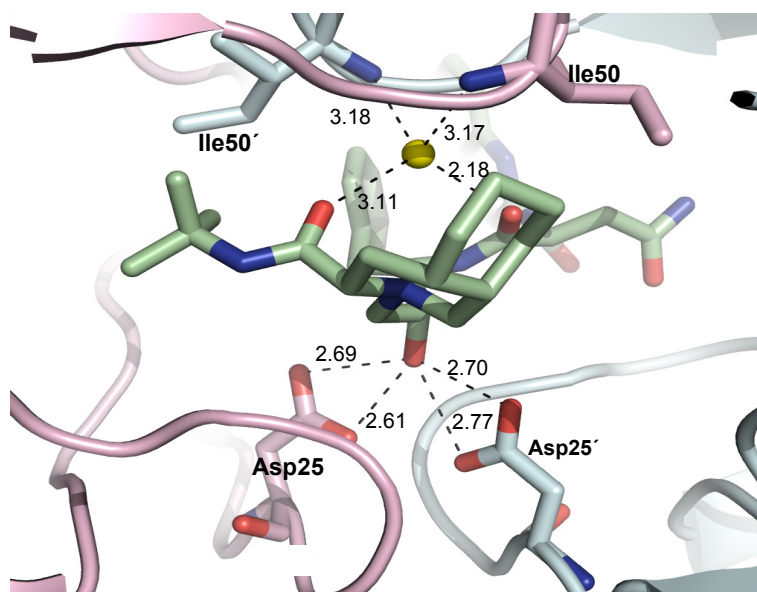


Figure 2.15: Hydrogen-bonding interactions of proteinase flaps with saquinavir (coloured green) mediated by a flap water molecule (coloured yellow) found in the cleft of the active-site. The main-chain amides of Ile50/50' form hydrogen bonds with this water molecule. Chain A is coloured pink and chain B is coloured cyan. Dashed lines indicate hydrogen bonds and distances are in Å.

2.5.1.1.1 Peptidomimetics

A peptidomimetic is a small protein-like chain designed to mimic an oligopeptide. They compete with the natural substrate for binding to a receptor or enzyme and typically arise from modification of an existing peptide. These modifications involve changes to the peptide through modification of the backbone by amide-bond isosteres, cyclizations, and attachment of pharmacophores to a template or scaffold. The advantage of peptidomimetics over native peptides has been demonstrated by their increased potency and selectivity, decreased side effects, improved oral bioavailability and half-life of activity through minimizing enzymatic degradation (Goodman and Ro, 1995). HIV PR inhibitors are designed to mimic the tetrahedral transition-state intermediate formed during the catalysis event. The tight binding of the transition state is well exploited in the design of PIs like saquinavir and indinavir, which do no longer have classical peptide characteristics such as scissile peptide bonds.

2.5.1.1.2 C₂-symmetric and pseudo-symmetric inhibitors

The initial observations that retroviral proteases contain the amino-acid triplet, aspartic acid-threonine/serine-glycine, and that their proteolytic activity could be inhibited by pepstatin, led to the early proposal that these enzymes were related mechanistically to the aspartic protease family of enzymes (Toh *et al.*, 1985; Katoh *et al.*, 1987). In contrast to other aspartic proteases such as pepsin and renin, the HIV-1 protease crystal structures (Navia *et al.*, 1989; Wlodawer *et al.*, 1989) revealed that the viral enzyme forms a C₂-symmetric homodimer, with a twofold axis transversing the active-site. The C₂-symmetry of the HIV-1 protease has prompted the synthesis of C₂-symmetric inhibitors (Budt *et al.*, 1993; Kempf *et al.*, 1991; Kempf *et al.*, 1993; Peyman *et al.*, 1993). HOE/BAY 793 is one of the most active compounds of this series that features a central vicinal diol serving as a transition state analogue in a C₂-symmetric peptidic environment (Budt *et al.*, 1993; 1995). However, the crystal structure of HOE/BAY 793 complexed to HIV-1 (Lange-Savage *et al.*, 1997) unexpectedly showed that the inhibitor binds in an asymmetric fashion with only one hydroxyl group interacting with both catalytic aspartates. This suggests that upon binding of a non-symmetric peptide substrate an asymmetry is introduced into the system and that the natural protease-substrate complex was not designed to be symmetric in evolution. Obviously, the HIV protease itself is only pseudo C₂-symmetric. An asymmetric binding mode is also manifest in the preference for nonequivalent side-chains in P1 and P1' as found in structure-activity investigations (Thanki *et al.*, 1992; Dreyer *et al.*, 1993).

2.5.1.2 Non-peptidic inhibitors

One example of non-peptidic compounds are cyclic-urea inhibitors. DMP 323, DMP 450 and DMP 850 are examples of this group of PIs (Turner, 2002). These compounds contain a seven-membered cyclic urea template (Figure 2.16) and were developed at DuPont Merck using *de-novo* computational design approaches (Lam *et al.*, 1994; Eyermann *et al.*, 1997). A seven-membered ring system was selected to allow the use of a diol, for optimal interaction with the catalytic aspartate residues.

The endocyclic urea functionality was chosen for enhancement of hydrogen-bonding to the flap region of the enzyme and for displacement of the flap water.

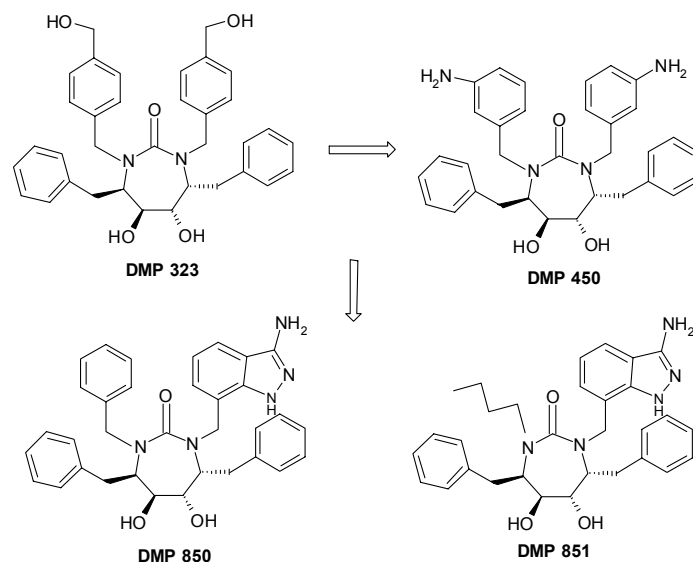


Figure 2.16: Cyclic-urea based HIV protease inhibitors: DMP 323, DMP 450, DMP 850, and DMP 851.

For the first candidate, DMP 323, the clinical trial was terminated because of highly variable oral bioavailability resulting from low aqueous solubility (6 $\mu\text{g/ml}$) and metabolic instability of the benzylic hydroxymethyl groups. DMP 323 had a K_i value of 0.34 nM and a cellular IC_{90} value of 57 nM (Lam *et al.*, 1996). The IC_{90} is the concentration of drug needed to inhibit 90% of viral reproduction in a MT-2 cell-based assay (Bachelier *et al.*, 1994). The DMP 450 compound was a second-generation clinical candidate, obtained after adjustment of the lipophilicity of the first generation clinical candidate, DMP 323. DMP 450 had good aqueous solubility (130-170 mg/ml) and showed good oral bioavailability in humans (Hodge *et al.*, 1996). Furthermore, DMP 450 inhibited HIV PR with a K_i value of 0.28 nM, and was effective in cell culture with an IC_{90} value of 130 nM (Lam *et al.*, 1996). Resistant variants to DMP 450 feature five substitutions in the HIV PR, Lys45Ile, Met46Leu, Val82Ile, Ile84Val and Leu90Met, as well as a 45-fold increase in IC_{90} levels (Ala *et al.*, 1997). The key residues associated with resistance among the five mentioned are Val82Ile and Ile84Val, leading to a 1000-fold increase in the K_i value compared to wild-type. Further improvement of DMP 450 was achieved with DMP 850 and DMP 851 (Rodgers *et al.*, 1998), which maintained an excellent inhibition constant (K_i

values of 31 and 21 pM, respectively) and antiviral activity (IC₉₀ values of 62, and 56 nM, respectively). These compounds also showed very good pharmacokinetic behaviour compared to DMP 450. Unlike DMP 323 and DMP 450, resistance to DMP 850 and 851 develops more slowly. Importantly, the Ile84Val mutation features prominently in viruses in the presence of DMP 850 and 851, showing phenotypic resistance as measured by an increase in IC₉₀ value (Rodgers *et al.*, 1998) relative to wild type. The largest losses in potency of DMP 850 and DMP 851 occur against multiply mutant viruses that carry mutations in both amino-acid residues 82 and 84. DMP 323 had resulted in viruses carrying the single mutation of Ile84Val, the single mutation of Val82Ile or the double mutation Val82Ile/Ile84Val (King *et al.*, 1995). The single mutation of Ile84Val has also been observed as a predominant mutation in *in-vitro* selection experiments with DMP 450 (Hodge *et al.*, 1996). In phase-II clinical trials, DMP 450 was found to have only modest potency in patients (Davis *et al.*, 2003). The compound is therefore no longer under development (Rusconi and Viganò, 2006).

2.5.2 The Schechter and Berger nomenclature for the description of protease subsites

The nomenclature to describe the interaction of a substrate with a protease has been introduced in 1967 by Schechter and Berger. In this classification, it is considered that the amino-acid residues of the polypeptide substrate bind to enzyme subsites around the active-site. By convention, these subsites on the protease are called S and the substrate or peptide amino-acid residues are called P. The amino-acid residues at the N-terminal side of the scissile bond are numbered Pn.....P3, P2, P1 and those at the C-terminal side are numbered P1', P2', P3'....Pn'. The two substrate residues that flank the scissile bond are P1 and P1'. The subsites on the protease that accommodate Pn...P1 and P1'....Pn' are numbered Sn.... S1 and S1'Sn', respectively. A number of apparent subsites, for holding the side-chains of the inhibitors (or substrates), can be identified in HIV-1 PR (Figure 2.17).

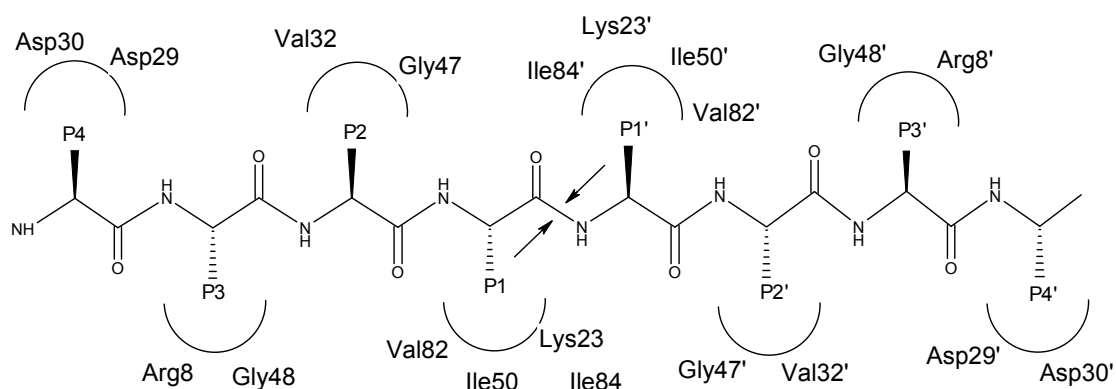


Figure 2.17: HIV-1 proteinase subsites and a modelled substrate. The nomenclature for the subsites is according to Schechter and Berger (1967). The arrows indicate the scissile bond.

Except for the active-site aspartates, the PR side-chains comprising the pockets S1 and S1' are mostly hydrophobic. The residues which contact the P1 or P1' side-chain of the inhibitors (or substrates) in HIV-1 PR include Leu23/Leu23', Asp25/Asp25', Ile50/Ile50', Pro81/Pro81', Val82/Val82' and Ile84/Ile84'. Almost all known inhibitors have hydrophobic moieties at P1 and P1' (Wlodawer and Gustchina, 2000). Subsites S2 and S2' are interior pockets in all retroviral PRs and are smaller than the S1/S1' or S3/S3' sites. In HIV-1 PR, S2 and S2' are formed by Ala28/28', Asp29/29', Asp30/30', Val32/32', Ile47/47', Gly48/48' and Gly49/49'. Even though the S2 and S2' pockets are hydrophobic, both hydrophilic or hydrophobic amino-acid residues of substrates/inhibitors can occupy the individual sites (Dunn *et al.*, 1994). Subsites S3 and S3' in retroviral proteases are known to have a rather broad specificity, able to accept residues of different types and sizes (Tözsér *et al.*, 1992; Wlodawer and Erickson, 1993). These subsites are quite large and exposed to the solvent. They are formed by residues Arg8/Arg8', Gly48/Gly48' and Phe53/Phe53'. The S3/S3' subsites can accept a variety of side-chains, some of which are very large, for example, naphthylalanine in inhibitors such as LP-130 (Kervinen *et al.*, 1998), LP-149 (Wlodawer *et al.*, 1995) and HOE/BAY 793 (Lange-Savage *et al.*, 1997).

2.6 Resistance to proteinase inhibitors

The X-ray crystal structures of numerous HIV PR inhibitor complexes have been solved to aid the process of inhibitor design (Wlodawer and Vondrasek, 1998). These structures allow us to understand many detailed features of drug binding as well as to begin examine resistance mechanisms at an atomic level. According to Stanford's HIV drug resistance database (<http://hivdb.stanford.edu/cgi-bin/PIResiNote.cgi>), proteinase resistance pinpoints modifications in 55 residues out of a total of 99 (Figure 2.18).

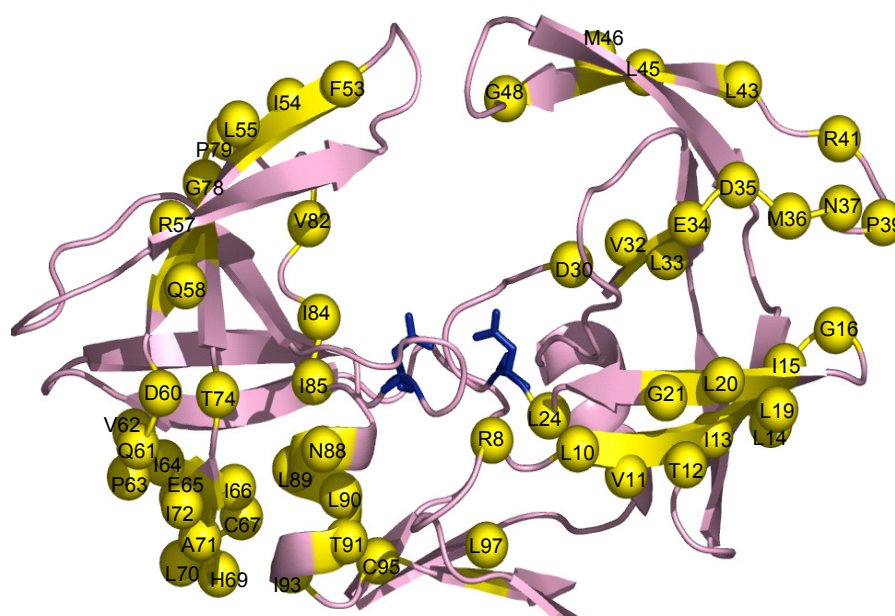


Figure 2.18: Backbone representation of HIV proteinase showing the location of possible mutations (coloured yellow). The active-site residues, Asp25/25', are shown as blue sticks.

Resistance to HIV PR inhibitors is caused by changes of specificity-determining residues (subsites mutations) that can directly interfere with the binding of the inhibitor to the enzyme. Mutations outside the subsites can indirectly affect inhibitor binding through long-range perturbations or changes in the efficiency of catalysis, enzyme stability and flexibility (Stefano *et al.*, 2002). Mutations outside the PR coding region that alter the amino-acid sequences in the processing sites of the Gag-Pol polyprotein increase the capability of the mutated PR to cleave at these sites (Wei *et al.*, 1995; Coffin, 1995).

In 1995, a characterization of resistant variants isolated from patients undergoing therapy with the proteinase inhibitor MK-639 (indinavir) was published

(Condra *et al.*, 1995). Some of these variants showed cross-resistance to six structurally diverse proteinase inhibitors (saquinavir, ritonavir, nelfinavir, amprenavir and lopinavir). This study provided the first evidence that inhibition of the HIV-1 PR can lead to the emergence of drug-resistance *in vivo* and that a combination therapy with multiple PIs may not prevent the loss of antiviral activity resulting from resistance selection. The effect of the Arg8Gln mutation on the binding of A77003 (norvir) was shown to be due to the loss of a charge-induced dipole interaction between the guanidinium side-chain of the Arg8 and the pyridine ring of the inhibitor (Gulnik *et al.*, 1995). While this mutation resulted in a dramatic loss of inhibition, it also produced a virus with severely impaired growth kinetics. A possible explanation for this defect is the loss of the inter-subunit salt bridge between Arg8 and Asp29', which may decrease the stability of the enzyme. While the precise structural mechanism of drug resistance can often be pinpointed for active-site mutations that directly affect inhibitor binding, the evaluation of non-active-site mutations is more challenging and there may be several different mechanisms at work. Some mutations may act in concert with active-site mutations by compensating for a functional deficit caused by the latter (Gulnik *et al.*, 1995). For example, the Val82Thr mutation is found almost exclusively in combination with one or more mutations outside the active-site region, such as Met36Ile, Ile54Val and Ala71Val (Nijhuis *et al.*, 1999).

2.6.1 Cross-resistance and sequential therapy

Most of the inhibitors are structurally derived from pseudopeptide substrate analogues and, as few as four to five amino-acid exchanges lead to cross-resistance (Condra *et al.*, 1995). Among the mutations observed in cross-resistant PR variants, some exchanges appear to be particularly important, *i.e.* Val32Ile, Met46Ile/Leu/Phe, Ile47Val, Gly48Val, Val82Ala/Ile/Thr, and Leu90Met (Erickson and Burt, 1996). The virus's responses to individual PIs differ significantly. Resistance towards saquinavir develops in 50% of the treated patients within the first year of therapy (Rose *et al.*, 1996) and is mostly associated with mutations Gly48Val and Leu90Met (Jacobsen *et al.*, 1996). Treatment with ritonavir usually leads to multiple mutations (3-11 substitutions, mostly in positions 30, 32, 46, 63, 71, 82, 84, 88 and 90) (Molla *et al.*, 1996). Indinavir typically selects resistant mutations relatively rapidly (12-24 weeks),

usually Val82Ala and Met46Ile, accompanied by mutations in positions 10, 20, 24, 63, 71 and 90 (Zhang *et al.*, 1997; Rose *et al.*, 1996). Nelfinavir selects the unusual mutation Asp30Asn (observed in 50% of the cases as early as 13 weeks after the start of the treatment). Mutations at positions 35, 36, 46 and 88 (Patick *et al.*, 1998) have also been observed. Amprenavir treatment selects the mutation Ile50Val, accompanied by further exchanges at positions 10, 46, 47 and 84 (Markland *et al.*, 2000).

The amino-acid residues of the enzyme forming the subsites S3 to S3' are thus obvious candidates for mutations that could directly influence inhibitor binding (Condra *et al.*, 1995; Molla *et al.*, 1996). Such mutations induce only minor structural changes (often gain or loss of a methyl group) at positions 8, 30, 32, 47, 48, 50, 82 and 84 (Weber *et al.*, 2002). In addition to mutations directly involving the substrate-binding site, the virus responds to the selection pressure of the PI by the concerted accumulation of mutations, often as many as 8-12 (Wlodawer and Vondrasek, 1998), several of which occurs outside the substrate binding cleft of the enzyme. Several of these mutations improve the catalytic efficiency of the enzyme rather than decrease inhibitor binding ("compensatory mutations" Arg20Lys, Met46Ile, Leu63Pro, Ala71Val) (Markowitz *et al.*, 1998; Rose *et al.*, 1996).

2.7 Motivation and scope of this thesis

Despite the numerous studies that have been conducted on HIV-1 proteinase-inhibitor complexes over the last two decades, an understanding of the (structural) evolution of the virus (*i.e.*, viral enzymes) during therapy is still lacking. In light of the importance of understanding and predicting viral evolution in the presence of potential drugs, a series of six HIV-1 proteinase variants were selected for our structural research reported herein. These variants were selected on the basis of PI resistance and the occurrence of secondary mutations, that cover in total 115 days of *in-vivo* viral evolution in an AIDS patient under suboptimal ritonavir monotherapy. The variants under investigation were day 0 (wild type with a valine at positions 62 and 77), day 28^a (Val82Thr), day 28^b (Met36Ile-Ile54Val), day 56 (Met36Ile-Ile54Val-Val82Thr), day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr, this variant was also observed on day 115) and day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-

Val82Thr). The variants evolved from having a quite normal replication capacity (day 0 and day 28^a) followed by variants with a lower replication capacity (day 28^b and day 56) and interestingly, the late variants (day 82 and day 115) exhibit a slightly higher replication capacity than variant day 0.

The aim of this thesis is to explain the effects of these mutations in terms of conformational changes of the proteinase and to understand the flexibility of the enzyme in relation to viral fitness during drug treatment in HIV infected people.

3 MATERIALS AND METHODS

3.1 Chemicals

Crystallization-grade chemicals were purchased from Fluka (Taufkirchen, Germany), ICN (Eschwege, Germany), Sigma (Deisenhofen, Germany), and Roth (Karlsruhe, Germany).

3.2 Crystallization buffers

Buffers for crystallization were prepared using deionised water (SimplicityTM, Millipore). Stock solutions of buffers were filtered and sterilized using 0.20 µm non-pyrogenic filters (Millipore), and kept at -20°C. High-concentration PEG stock solutions were not filtered because of the high viscosity.

3.3 Crystallization of variants of HIV-1 proteinase

The following variants of HIV-1 proteinase (Table 3.1) were prepared by AssoProt Biotech Ltd., Czech Republic, and were stored at -70°C in 20 mM sodium acetate, pH 5.0, 200 mM sodium chloride, 0.5 mM DTT, 10% (v/v) glycerol, 1 mM EDTA. Saquinavir and ritonavir were kindly provided by Dr. Jan Konvalinka (Prague, Czech Republic). The inhibitors were dissolved in 100% dimethylsulfoxide to a final concentration of 20 mM and stored at -20°C.

Table 3.1: Variants of HIV-1 proteinase.

Variants of HIV-1 proteinase	Concentration (mg/ml)
Day 0 (Val62-Val77)	0.45
Day 28 ^a (Val82Thr)	1.00
Day 28 ^b (Met36Ile-Ile54Val)	0.68
Day 56 (Met36I-Ile54Val-Val82Thr)	0.40
Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)*	0.45
Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)	0.84

*Still present after 115 days of therapy

All variant PRs were dialysed before crystallization for 24 hrs at 4°C. The diluted protein was concentrated to 2 mg/ml using a Microcon YM-10 (Millipore) with a molecular weight cut-off of 10 kDa at 5,000 x g. The protein concentration was estimated according to the formula

$$M2 = M1 \cdot V1 / V2$$

where M1 = initial molar concentration, V1 = initial volume, M2 = final molar concentration and V2 = final volume. For example, to achieve a final concentration of 2 mg/ml from 300 µl of 0.68 mg/ml, the protein solution was concentrated to 100 µl. For unliganded HIV-1 PR variant day 115, the initial crystals were grown from 25% (NH₄)₂SO₄, 0.1 M sodium acetate pH 5.0 at room temperature by the hanging drop vapour diffusion technique, producing bunches of tiny needles. Crystals of HIV-1 PR-inhibitor complexes were grown by co-crystallization with a protein-to-ligand molar ratio of 1:10. Similar to the unliganded proteinase, the initial crystals of complexes with ritonavir, variants day 82 and 115, grown from 1.25 M KCl, 0.1 M sodium acetate pH 5.5 at room temperature, were clusters of tiny needles. Both unliganded and liganded needles grew after about one week. The hanging-drop vapour diffusion method in combination with a matrix screen (Table 3.2) was used in an effort to obtain single, big crystal at room temperature (approximately 25°C). After two weeks, crystals of hexagonal or orthorhombic morphology appeared. Out of the six variants, only three (Met36Ile-Ile54Val; Met36Ile-Ile54Val-Ala71Val-Val82Thr; Lys20Arg-

Met36Ile-Ile54Val-Ala71Val-Val82Thr) in complex with saquinavir and two in complex with ritonavir (Met36Ile-Ile54Val-Ala71Val-V82T; Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr) were successfully crystallized and these diffracted X-rays to a resolution of up to 1.76-1.95 Å, while no crystals were obtained for the other variants. All six variants were successfully crystallized in the unliganded form and these diffracted X-rays to a resolution between 2.18-2.65 Å. A summary of the different crystallization conditions, which produced diffracting crystals, is presented in Tables 3.3, 3.4 and 3.5.

Table 3.2: A matrix screen, pH versus salt.

pH versus (NH₄)₂SO₄ (%)

10 0.1 M sodium acetate pH 5.5	15	20	25
10 0.1 M Mes pH 6.5	15	20	25
10 50 mM sodium cacodylate pH 7.0	15	20	25
10 0.1 M Tris.HCl pH 7.5	15	20	25

pH versus KCl (M)

0.50 0.1 M sodium acetate pH 5.5	0.75	1.00	1.25
0.50 0.1 M Mes pH 6.5	0.75	1.00	1.25
0.50 50 mM sodium cacodylate pH 7.0	0.75	1.00	1.25
0.50 0.1 M Tris.HCl pH 7.5	0.75	1.00	1.25

pH versus NaCl (M)

0.25 0.1 M sodium acetate pH 5.5	0.50	0.75	1.00
0.25 0.1 M Mes pH 6.5	0.50	0.75	1.00
0.25 50 mM sodium cacodylate pH 7.0	0.50	0.75	1.00
0.25 0.1 M Tris.HCl pH 7.5	0.50	0.75	1.00

Table 3.3: Crystallization conditions for HIV-1 proteinase in complex with saquinavir (Taken from the Bachelor thesis of Marcus Fischer, University of Lübeck, 2004).

HIV-1 PR-saquinavir complexes (Protein buffer: 20 mM MES pH 6.0, 10% (v/v) glycerol, 1 mM EDTA, 1 mM DTT)	Conditions
Day 28 ^b (Met36Ile-Ile54Val)	1) 1.00 M KCl, 0.1 M MES pH 6.0, 3 mM sodium azide, 10 mM DTT 2) 16% (NH ₄) ₂ SO ₄ pH 5.8, 200 mM sodium citrate tribasic dihydrate
Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)	1) 1.00 M KCl, 0.1 M MES pH 6.0, 3 mM sodium azide, 10 mM DTT 2) 16% (NH ₄) ₂ SO ₄ pH 5.8, 200 mM sodium citrate tribasic dihydrate
Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)	1) 1.00 M KCl, 0.1 M MES pH 6.0, 3 mM sodium azide, 10 mM DTT 2) 16% (NH ₄) ₂ SO ₄ pH 5.8, 200 mM sodium citrate tribasic dihydrate

Table 3.4: Crystallization conditions for HIV-1 proteinase in complex with ritonavir. Conditions that produced well-diffracting crystals are printed in bold.

HIV-1 PR-ritonavir complexes (Protein buffer: 20 mM sodium acetate pH 5.0, 5 mM DTT, 1 mM EDTA, 20 mM NaCl)	Conditions
Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)	1) 25% (NH ₄) ₂ SO ₄ , 100 mM MES pH 6.5 2) 25% (NH ₄) ₂ SO ₄ , 100 mM Hepes pH 6.5 3) 25% (NH₄)₂SO₄, 100 mM NaH₂PO₄ pH 6.5
Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)	1) 20% (NH₄)₂SO₄, 100 mM MES pH 6.0 2) 25% (NH ₄) ₂ SO ₄ , 100 mM NaH ₂ PO ₄ pH 6.0 3) 1 M KCl, 100 mM Hepes pH 6.5, 10 mM DTT

Table 3.5: Crystallization conditions that produced crystals for inhibitor-free HIV-1 proteinase. Conditions that produced well-diffracting crystals are printed in bold. Crystallization conditions of day 0, day 28^a (Val82Thr), day 28^b (Met36Ile-Ile54Val) and day 56 (Met36Ile-Ile54Val-Val82Thr) are taken from the Bachelor thesis of Marcus Fischer, University of Lübeck, 2004.

Inhibitor-free HIV-1 PR	Conditions
Day 0 (in 25 mM sodium cacodylate pH 6.0, 50 mM sodium chloride)	1) 10% (NH ₄) ₂ SO ₄ 2) 20 % (NH₄)₂SO₄
Day 28 ^a (Val82Thr) (in 50 mM sodium cacodylate pH 7.0, 100 mM sodium chloride)	1) 20% (NH ₄) ₂ SO ₄
Day 28 ^a (Val82Thr) (in 25 mM sodium cacodylate pH 6.0, 50 mM sodium chloride)	2) 10% (NH₄)₂SO₄
Day 28 ^b (Met36Ile-Ile54Val) (in 50 mM sodium cacodylate pH 7.0, 100 mM sodium chloride)	1) 20% (NH ₄) ₂ SO ₄
Day 28 ^b (Met36Ile-Ile54Val) (in 25 mM sodium cacodylate pH 6.0, 50 mM sodium chloride)	2) 20 % (NH₄)₂SO₄
Day 56 (Met36Ile-Ile54Val) (in 50 mM sodium cacodylate pH 7.0, 100 mM sodium chloride)	1) 15% (NH₄)₂SO₄
Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr) (in 20 mM sodium acetate pH 5.0, 5 mM DTT, 1 mM EDTA, 20 mM sodium chloride)	1) 25% (NH ₄) ₂ SO ₄ , 100 mM MES pH 6.5 2) 25% (NH ₄) ₂ SO ₄ , 100 mM Hepes pH 6.5 3) 25% (NH₄)₂SO₄, 100 mM NaH₂PO₄ pH 6.5
Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr) (in 20 mM sodium acetate pH 5.0, 5 mM DTT, 1 mM EDTA, 20 mM sodium chloride)	1) 20% (NH₄)₂SO₄, 100 mM MES pH 6.0 2) 25% (NH ₄) ₂ SO ₄ , 100 mM NaH ₂ PO ₄ pH 6.0 3) 1 M KCl, 100 mM Hepes pH 6.5, 10 mM DTT

3.4. Crystal growth in the meta-stable zone

A schematic phase diagram showing the different zones and types of solid phases for protein solubility is given in Figure 3.1. It has been argued that the meta-stable zone is the optimum zone for crystal growth in contrast to the labile zone, where nucleation occurs. The first task in protein crystallization is therefore to determine the under- and supersaturated zones, which are delineated by the solubility curve. Protein solubility is defined as the concentration of soluble protein in equilibrium with the solid state, at given temperature and pH value, and at given concentrations of solvent compounds (*e.g.* buffer, salts, crystallizing agents, additives, etc.). The advantage of growing crystals in the meta-stable zone is to minimize the

number of crystals and to obtain larger single crystals. However in reality, spontaneous nucleation does not occur in the meta-stable zone, unless it is induced by vibration or a particle that will promote heterogeneous nucleation.

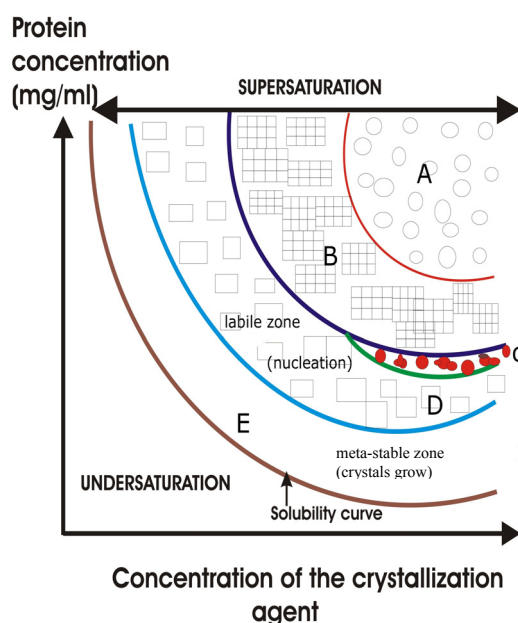


Figure 3.1: The phase diagram showing the different zones and types of solid phases. A, liquid-liquid phase separation (oil droplets); B, precipitation (amorphous); C, spherulites; D, labile zone (nucleation); E, meta-stable zone (clear). Crystals will never grow in the undersaturated zone (Bergfors, 1999).

In order to induce growth of large single crystals in the meta-stable zone, a modified vapour diffusion technique was applied (Saridakis and Chayen, 2000). In this method, hanging drops of HIV PR day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr) (~2 mg/ml) were set up at 25°C using 25% (NH₄)₂SO₄, 0.05 M sodium cacodylate, pH 6.0. After different time intervals, *i.e.* 0, 24, 36, and 48 hrs, the cover slides with the hanging drops were transferred over wells containing 20% (NH₄)₂SO₄, 0.05 M sodium cacodylate, pH 6.0. In this way, nucleation and growth processes can be clearly separated.

3.5 Seeding: streak seeding and macro seeding

The streak-seeding technique was used in order to improve the quality and size of crystals. Cat whiskers were cleaned using 70% (v/v) ethanol, washed with deionised water, and dried. Using the clean cat whiskers, seeds were taken from drops that produced bunches of tiny needles or contained phase separation. Then, the seeds were transferred to freshly prepared hanging drops over reservoirs with lower concentrations of the precipitant (in 2% steps), *e.g.* when bunches of tiny needles were obtained at 24% $(\text{NH}_4)_2\text{SO}_4$, the precipitant concentrations for the subsequent experiments were adjusted to 22%, 20%, 18%, and 16% before streak seeding the newly prepared protein-precipitant drops. These freshly prepared protein-precipitant drops were first pre-equilibrated for 24 hrs.

Macro-seeding was used in cases where no crystals could be obtained. Very small crystals of unliganded PR variant day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr) were used as seeds for variant day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr). The small seeds were washed in a slightly diluted solution, *i.e.* crystals of unliganded PR variant day 115 grown in 1 M KCl, 100 mM Hepes pH 6.5 and 10 mM DTT, were picked using nylon Cryo-Loops (Hampton Research) and transferred to a wash solution of 0.8 M KCl, 100 mM Hepes pH 6.5, 10 mM DTT. From the wash solution, seeds were then transferred to pre-equilibrated droplets (24 hrs over reservoir) containing PR variant day 82. Several rounds of streak-seeding followed in order to avoid cross-contamination of the crystals of variant day 82 with protein of day 115 that was used as the initial seed.

3.6 Crystallographic analysis

Figure 3.2 shows the general strategy used for data collection, processing and integration of HIV-1 PR variants; complexes with saquinavir: day 28^b (Met36Ile-Ile54Val), day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr), and day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr), complexes with ritonavir: day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr), and day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr), and inhibitor-free: day 0 (Val62-Val77), day 28^a (Val82Thr), day 28^b (Met36Ile-Ile54Val), day 56 (Met36Ile-Ile54Val-Val82Thr), day 82

(Met36Ile-Ile54Val-Ala71Val-Val82Thr), and day 115 (Lys20Arg-Met36Ile-I54V-Ala71Val-Val82Thr). Data collection results in a set of images, each representing a part of the rotation of the crystal in the beam. These images are distorted sections of reciprocal space. Each image in a dataset likely has its recorded intensities on a different scale due to absorption effect (e.g. crystal shape), detector quality (e.g. different regions of the detector may have different sensibilities) and variations in intensity of the incident beam, which applies especially to synchrotrons. Data may even be collected from several crystals. In the following steps, the symmetrically-equivalent reflections from different images were put on a common scale and merged using the program SCALA (Evans, 1997) or SCALEPACK (Otwinowski and Minor, 1997). Three programs, namely WILSON, TRUNCATE (French and Wilson, 1978) and UNIQUEIFY were used in the SCALA protocol. The WILSON program reads a file of amplitudes (or intensities) and produces a Wilson plot (Wilson, 1942) of the observed intensities against the resolution in order to determine an absolute scale and temperature factor. UNIQUEIFY creates a list of the unique reflections in an MTZ file, with the appropriate cell parameters, symmetry, and resolution range. A column containing the free-R flag was added to the MTZ file. This was followed by the calculation of structure factor amplitudes from averaged intensities. The SCALA, WILSON, TRUNCATE, UNIQUEIFY and SCALEPACK2MTZ programs are part of the CCP4 program suite (Collaborative Computational Project, Number 4, 1994).

Unlike SCALA, SCALEPACK writes averaged intensities, not structure factor amplitudes. The SCALEPACK2MTZ program converts merged data from SCALEPACK into the MTZ format followed by TRUNCATE to convert the intensities into structure factor amplitudes.

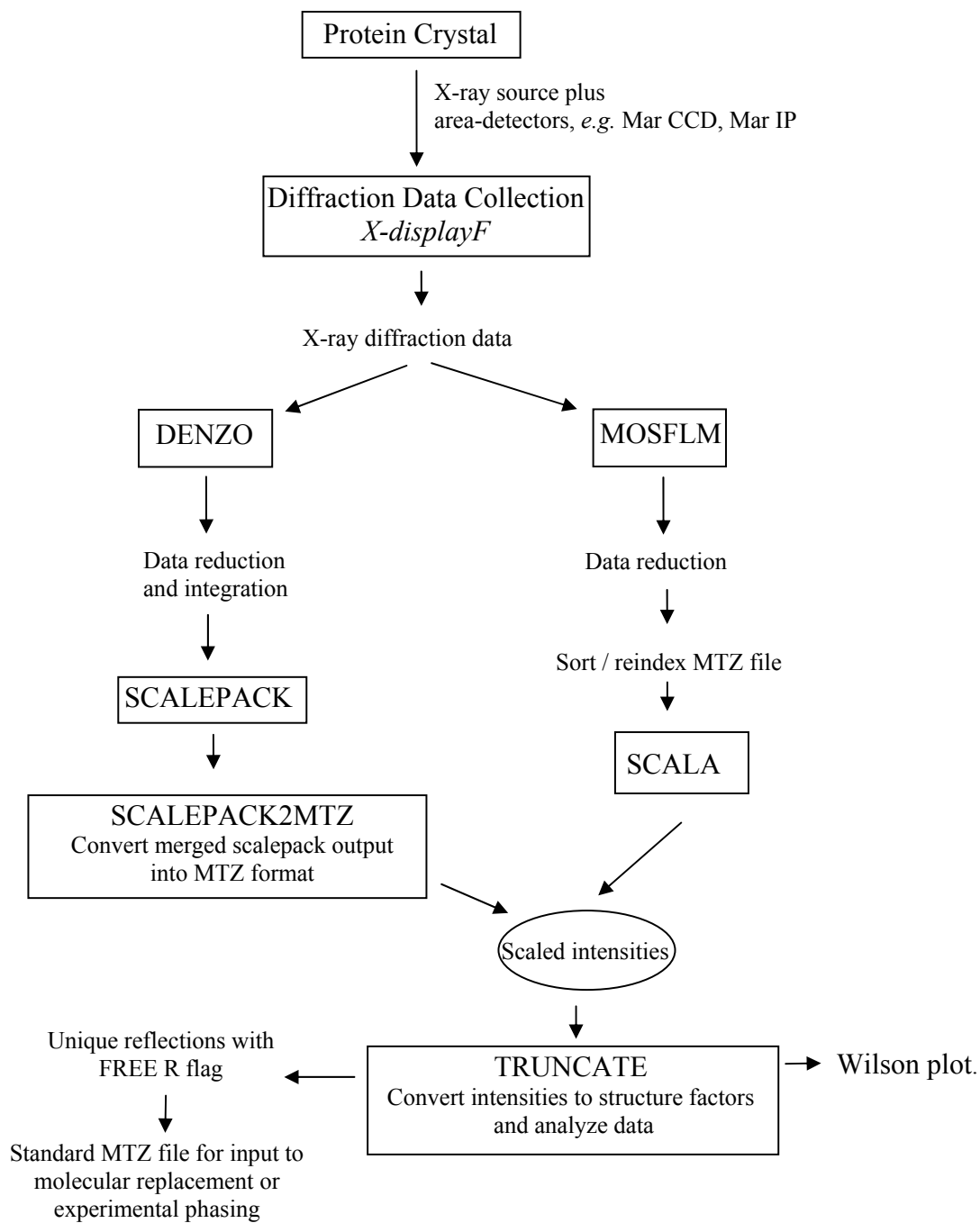


Figure 3.2: Outline of diffraction data collection, processing, and integration used for HIV PR variants.

3.6.1 Data collection, processing, and reduction

The diffraction datasets were collected at the Joint University of Hamburg - IMB Jena - EMBL beamline X13 (wavelength 0.812 Å), Deutsches Elektronen-Synchrotron (DESY, Hamburg) that is equipped with a MarCCD detector (X-ray Research, Hamburg). Each variant of the inhibitor-free HIV-1 PR and of complexes were collected at cryo-temperature (100 K). In order to determine the space group and how many images should be collected, a few images were indexed, reduced, and processed at the beginning of each data collection. Prediction of the space group was also done based on crystal morphology. The ideal data set is 100% complete, with most reflections measured several times (i.e. high redundancy). Naturally, this is not always possible. The fraction of reciprocal space that must be covered to collect all the unique data to a given resolutions depends on the space group. For example, a minimum of 90° of rotation are normally required for a tetragonal (P4) space group, or at least 180° for the triclinic space group. Completeness and redundancy of the data for hexagonal (P6) and orthorhombic (P222) space groups of HIV-1 PR was ensured by collecting no less than 180° of data on a MAR CCD X-ray detector from a single crystal. The crystal to detector distance was usually set to 180.0 mm. Due to the prior knowledge that HIV proteinase is in reality a pseudo C₂-symmetric homodimer, space groups were chosen in which the asymmetric unit contains the homodimer.

Diffraction data were recorded, visualized and reduced using the HKL package (Otwinoski and Minor, 1997) with the DENZO program for indexing and integration, and SCALEPACK for scaling. MOSFLM (Leslie, 1999) and SCALA programs, part of the CCP4 program suite, were also used for integration and scaling, respectively. The data were integrated and optimized for mosaicity, spot ellipticity, and background. The HKLVIEW program, part of the CCP4 program suite, was used to check the correctness of space group assignment by studying possible systematic absences.

Data from inhibitor-free variants of HIV-1 PR and its complexes with either saquinavir or ritonavir were collected to the resolution limit listed in Table 3.6.

Table 3.6: Maximum resolution of diffraction data for inhibitor-free variants of HIV-1 proteinase and its complexes either with saquinavir or ritonavir.

Inhibitor-free HIV-1 proteinase	Resolution limit (Å)	Program for data processing
Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)	2.40	DENZO/SCALEPACK
Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)	2.20	DENZO/SCALEPACK
Variants of HIV-1 proteinase in complex with saquinavir		
Day 28 ^b (Met36Ile-Ile54Val)	1.95	MOSFLM/SCALA
Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)	1.95	MOSFLM/SCALA
Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)	1.76	MOSFLM/SCALA
Variants of HIV-1 proteinase in complex with ritonavir		
Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)	1.77	DENZO/SCALEPACK
Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)	1.99	DENZO/SCALEPACK

Information about the other four inhibitor-free HIV-1 PRs can be found in the Bachelor thesis of Marcus Fischer, University of Lübeck (2004).

3.6.2 Phasing - molecular replacement

The initial phases for the inhibitor-free variants and of complexes with ritonavir or saquinavir were determined by using the MOLREP program (Vagin and Teplyakov, 1997) that is part of the CCP4 program suite. As mentioned earlier, the homodimer was chosen as the unit to refine. The collected dataset for each inhibitor-free HIV-1 PR variant and of complexes revealed that the crystals belonged to either space group P6₁ or P2₁2₁2₁.

The equation for calculating electron density is as follows

$$\rho(xyz) = 1/V \sum_h \sum_k \sum_l |F_{hkl}| e^{-2\pi i(hx + ky + lz) + i\alpha(hkl)}$$

where $\rho(xyz)$ stands for the electron-density distribution in the unit cell, V is the volume of the unit cell, $\sum_h \sum_k \sum_l$ is the sum over all the reflections, $|F_{hkl}|$ is the observed structure factor amplitude of a given reflection and α is the phase of the structure factor. The calculated electron-density distribution $\rho(xyz)$ is further improved by several rounds of model adjustment and refinement.

After each round of refinement, the corresponding Fourier syntheses were calculated to generate the electron density maps (Drenth, 1994):

- (1) a σ^A -weighted electron-density map, $2F_o-F_c$ (Read, 1986)
- (2) a F_o-F_c difference electron-density map

Table 3.7: The starting model for each inhibitor-free HIV-1 proteinase and of complexes with saquinavir or ritonavir.

Inhibitor-free HIV-1 proteinase	Space group	PDB code of starting model and its space group
Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)	P6 ₁	1g6l.pdb (Pillai <i>et al.</i> , 2000), P6 ₁
Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)	P6 ₁	1g6l.pdb (Pillai <i>et al.</i> , 2000), P6 ₁
HIV-1 PR in complex with saquinavir		
Day 28 ^b (Met36Ile-Ile54Val)	P2 ₁ 2 ₁ 2 ₁	1nh0.pdb (Brynda <i>et al.</i> , 2004), P2 ₁ 2 ₁ 2 ₁
Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)	P6 ₁	1pro.pdb (Sham <i>et al.</i> , 1996), P6 ₁
Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)	P6 ₁	1pro.pdb (Sham <i>et al.</i> , 1996), P6 ₁
HIV-1 PR in complex with ritonavir		
Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)	P6 ₁	1pro (Sham <i>et al.</i> , 1996), P6 ₁
Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)	P6 ₁	1pro (Sham <i>et al.</i> , 1996), P6 ₁

Traditional molecular replacement (MR) methods make use of the properties of the Patterson function because it can be computed without phase information, *i.e.* all phases are set to zero. MOLREP is an automated program for MR that utilizes the properties of the Patterson function (Vagin and Teplyakov, 1997). Meanwhile, non-traditional MR methods, based on (maximum) likelihood, have been developed that are more sensitive and that have been shown to be very powerful in difficult cases of phase determination by MR (Storoni *et al.*, 2004). The likelihood-based method has been implemented, for example, in the program Phaser (McCoy *et al.*, 2005). The Patterson function will give a distribution of the interatomic vectors in the unit cell in terms of a Patterson density

$$P(uvw) = 1/V \sum_h \sum_k \sum_l |F_{hkl}|^2 \cos [2\pi(hx + ky + lz)]$$

or

$$P(uvw) = \int_V \rho(x_1, y_1, z_1) \cdot \rho(x_2, y_2, z_2) dV$$

The real space (*i.e.*, of the crystal) and Patterson space unit-cells are identical.

If the unit cell contains N atoms, the Patterson map will show N^2 peaks, corresponding to the N vectors that exist for each atom. One vector of each atom points to itself (*i.e.*, vector length = 0), so that the N atoms within the unit cell produce a large origin peak at the origin of the map. The unit cell of the Patterson will therefore contain $N^2 - N$ peaks outside the origin, *e.g.* if a simple structure contains $N = 3$ atoms (Figure 3.4; for space group P1), $N^2 - N = 6$ peaks are found outside of the origin. For HIV-1 PR, with $N = 1520$ atoms, the Patterson map will be very crowded with the majority of peaks overlapping one another.

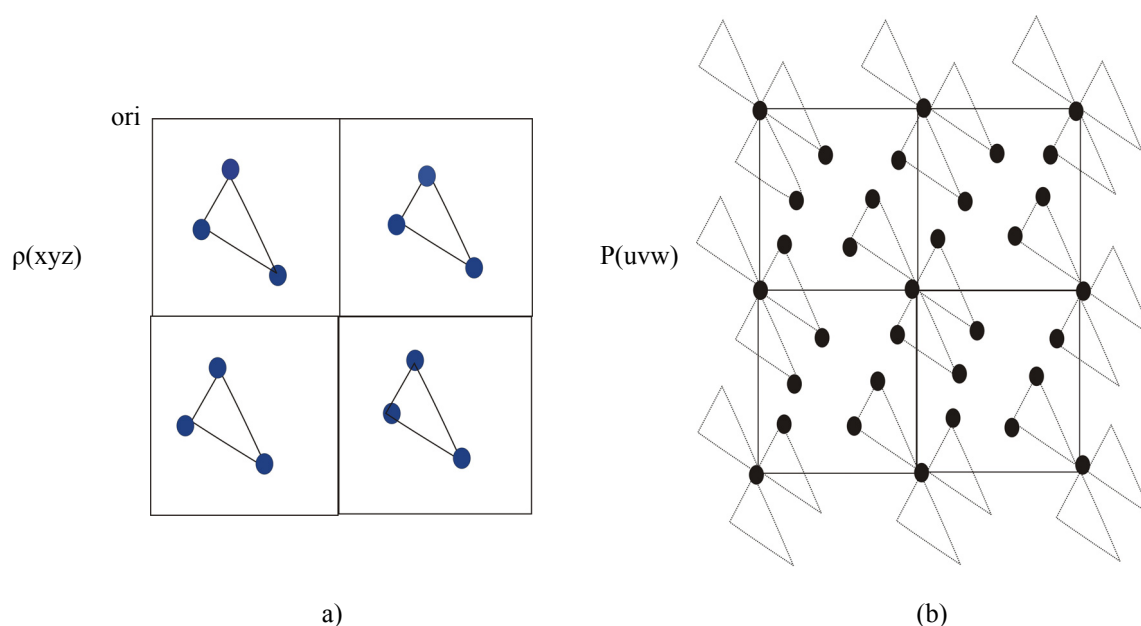


Figure 3.4: (a) Example showing four unit cells, space group P1, each containing a 3-atom structure. (b) Four unit cells of Patterson space with 6 peaks per cell outside the origin (black).

In the molecular replacement method, the steps performed between the two rigid bodies are as shown in Figure 3.5. The relationships are described by the following operators: the rotation matrix \mathbf{R} expressed in the three Eulerian angles α , β , γ or θ_1 , θ_2 , θ_3 and the translation vector \mathbf{t} , expressed as t_x , t_y and t_z .

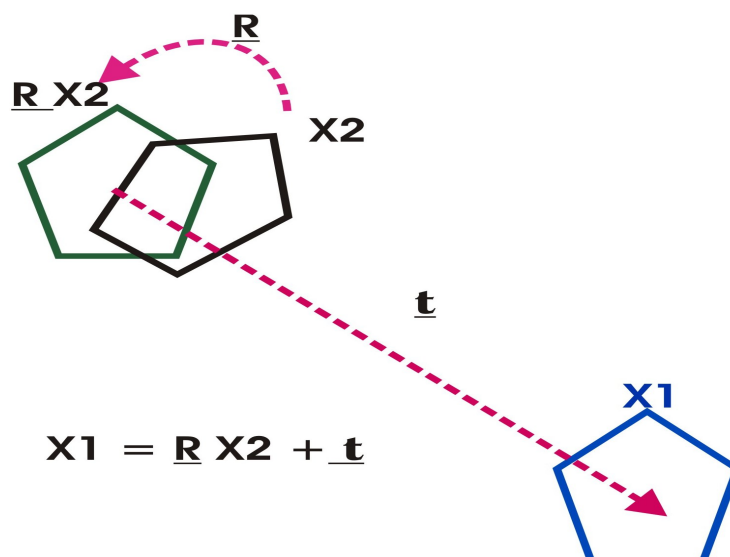


Figure 3.5: The relationship between two identical rigid bodies, X1-unknown, observed or target, and X2-known, starting or search model, is expressed by six parameters, three rotational and three translational.

Coordinates for water molecules, inhibitors and other small molecules/ions (*e.g.* zinc, magnesium, DTT, calcium, etc.) were removed from the starting/search model before using it as an input model for the MOLREP program. For the inhibitor-free HIV-1 PR variant day 115, residues 45 to 55 were removed from the starting model, PDB code 1g61, in order to reduce bias. One homodimer was assumed in the asymmetric unit. The experimental data to a resolution of 3.5 Å were used for molecular replacement. Results from the molecular replacement analysis (Table 3.8) supported that the space groups were correctly determined.

The initial diffraction analysis indicated that crystals of the inhibitor-free variants day 82 and day 115 in complexes with saquinavir or ritonavir are of a hexagonal space group. On the basis of indicators such as R_{sym} , it was not possible to differentiate between space groups $P6_1$ and $P6_122$. Also, an investigation of the systematic absences in the reflections of space groups $P6_1$ and $P6_122$ could not resolve the ambiguity. An SFCHECK (Vaguine *et al.*, 1999) analysis was performed on the different datasets and no problems were identified. Because a homodimer had been chosen as the unit to refine in the asymmetric unit, space group $P6_1$ was finally chosen. The selection of $P6_122$ would have reduced the asymmetric unit content to a single protein chain. Furthermore, refinement of the pseudo C_2 -symmetric homodimer in the lower-symmetry space group resolved the ambiguity in modelling several side-chains into the electron density maps. This finally resulted in an improved R_{factor} and

R_{free} for space group $P6_1$ compared to $P6_122$.

Table 3.8: MOLREP output. Theta, phi and chi represent the rotation matrix expressed in Eulerian angles; translation vector t_x , t_y , t_z

Inhibitor-free HIV-1 proteinase (space group)	MOLREP output							
	theta	phi	chi	t_x	t_y	t_z	R_{factor}	CC
Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr) ($P6_1$)	0.0	0.0	59.58	-0.38	0.00	0.0	0.40	0.65
Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val- Val82Thr) ($P6_1$)	0.0	0.0	0.0	0.0	0.38	0.0	0.43	0.53
HIV-1 proteinase in complex with saquinavir								
Day 28 ^b (Met36Ile-Ile54Val) ($P2_12_12_1$)	156.8	-1.59	22.21	0.46	-0.03	-0.10	0.46	0.52
Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr) ($P6_1$)	0.0	180.0	0.0	0.41	0.40	0.0	0.46	0.53
Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val- Val82Thr) ($P6_1$)	0.0	0.0	0.0	0.0	0.38	0.0	0.45	0.55
HIV-1 proteinase in complex with ritonavir								
Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr) ($P6_1$)	0.0	0.0	0.0	0.0	0.38	0.0	0.43	0.56
Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val- Val82Thr) ($P6_1$)	0.0	0.0	0.0	0.0	0.38	0.0	0.41	0.51

$R_{\text{factor}} = \Sigma (|F_o| - |F_c|) / \Sigma |F_o|$, F_o = observed structure factor, F_c = calculated structure factor.

Correlation coefficient CC = $\{\Sigma_h (|F_o| - \langle F_o \rangle) \times (|F_c| - \langle F_c \rangle)\} / \{\Sigma_h (|F_o| - \langle F_o \rangle)^2 \times (|F_c| - \langle F_c \rangle)^2\}^{1/2}$

In all cases, the rotation was either 0°, 60° (59.58°) or 180° while the translation vector is often zero. This is because the starting and target models have the same space group ($P6_1$) and unit cell. The only exception is variant day 28^b (Met36Ile-Ile54Val) in complex with saquinavir that belongs to space group $P2_12_12_1$.

3.7 Model building and refinement strategy

The initial models obtained as a result of the molecular replacement were subjected to maximum-likelihood refinement using REFMAC5 (Murshudov *et al.*, 1997) as incorporated into the CCP4 program suite (Collaborative Computational Project, Number 4, 1994).

3.7.1 Model refinement using REFMAC5

The first step in the model improvement process was a rigid-body refinement, followed by many rounds of refinement and manual model rebuilding. Model parameters were refined against F_o . Each atom of the model can be described using three positional parameters x , y , z and an isotropic temperature factor (B -factor). At very high resolution, *i.e.* better than 1.0 Å, even the occupancy and anisotropic temperature factors can be refined (Murshudov *et al.*, 1999). According to Brünger *et al.*, (1998), the temperature factor should be refined in addition to the coordinates (x , y , z) only if the data (observations or reflections) to parameter ratio is above 1.0.

During the refinement steps, a proper ‘weighting’ scheme that relates reflection data and geometric restraints (bond lengths, bond angles, etc.) was found to be critical to REFMAC5 refinement, especially for low-resolution data (when the resolution is worse than 2.0 Å). Weighting should be sensible so that the stereochemical parameters of the main-chain bond lengths and bond angles are within limits: a refined model should exhibit r.m.s deviations of no more than 0.02 Å for bond distances and 2.0 degrees for bond angles (Jaskolski *et al.*, 2007). These values are somewhat larger than the ones observed in small molecules (Engh and Huber, 1991). For example, a weighting factor of 0.30 resulted in sensible r.m.s. deviations for HIV-1 PR variant day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr) in complex with ritonavir (1.77 Å resolution data). However, a weighting factor of 0.15 was necessary for the refinement of unliganded variant day 115 (2.4 Å resolution data).

Four parameters, the coordinates x , y , z and the temperature factor, were refined for variants of the HIV-1 proteinase complexes with ritonavir (day 82 and day 115) or saquinavir (day 28^b, day 82, and day 115). However, only three parameters,

coordinates x, y, and z, were refined in unliganded variants day 82 and day 115, and the temperature factor was restricted according to the Wilson plots (Wilson, 1942), 35 Å² for variant day 82 and 37 Å² in variant day 115 because of the limited resolution of 2.20 and 2.40 Å, respectively. In all cases, 5% of the reflections were flagged for the calculation of an R_{free} (Brünger, 1993). The R_{factor} and R_{free} for different stages of the refinement are listed in Table 3.9.

Table 3.9: The R_{factor} values after MOLREP and rigid-body refinement.

	R_{factor} after MOLREP	R_{factor}, R_{free} after the first step of REFMAC5 rigid-body refinement	Resolution range (Å)
Inhibitor-free HIV-1 proteinase			
Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)	40.1	25.9, 37.8	53.76-2.40
Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)	43.3	26.7, 36.2	53.53-2.20
HIV-1 proteinase in complex with saquinavir			
Day 28 ^b (Met36Ile-Ile54Val)	45.8	25.8, 31.6	92.45-1.95
Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)	46.0	30.6, 35.5	83.05-1.76
Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)	44.9	26.8, 32.0	84.51-1.95
HIV-1 proteinase in complex with ritonavir			
Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)	42.9	26.8, 32.1	53.53-1.90
Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)	41.3	28.6, 31.8	53.45-1.77

$R_{\text{factor}} = \sum (|F_o| - k |F_c|) / \sum |F_o|$, where F_o = observed structure factor, F_c = calculated structure factor, k = scaling factor.

R_{free} = as “R-factor” described above, calculated with 5% of the X-ray data selected randomly, and excluded from the refinement (Brünger, 1993).

3.7.2 Inhibitor parameterization and ligand minimization for REFMAC5

PDB coordinates for saquinavir and ritonavir were taken from the Macromolecular Structure Database, (Berman *et al.*, 2000). The PDB coordinates were used as an input for Sybyl 7.2 (Tripos Associates, Inc., St. Louis, MO, USA) and minimized by the Powell method, with charges according to the Gasteiger-Hückel

method. The first step was to make sure that atom- and bond-types were correctly chosen. Secondly, hydrogen atoms were added. The minimized inhibitor was docked into the different electron density maps and the refinement proceeded.

3.7.3 Electron density fitting

Fitting of all protein atoms, inhibitor molecules, and water molecules into electron density was done with XFIT (McRae, 1999) displaying both the σ^A -weighted $2F_o-F_c$ and the F_o-F_c maps and sequentially checking the model for its quality. The σ^A -weighted $2F_o-F_c$ map is habitually displayed in blue colour while for the F_o-F_c difference map, positive peaks are normally displayed in green and negative peaks in red. The F_o-F_c map was contoured at 3.0σ above the mean and the σ^A -weighted $2F_o-F_c$ map contoured at 1.0σ . During the map fitting and modelling steps, all residues were systematically inspected for side-chain disorder, or misplaced side-chains.

Amino-acid substitutions were located in the electron density and replaced according to the sequence of the HIV-1 PR variant. Strong negative peaks in the F_o-F_c maps were further investigated. When no density was seen for a certain side-chain, mostly for arginine, lysine, aspartate or glutamate, the side-chain was not built until the phases had improved. If the density did not improve at all, the occupancies for the side-chain were fixed to zero, and the B -factors were fixed to 100 \AA^2 , so that atoms would have no contribution in the next round of refinement. The atomic occupancy of unliganded variants and of complexes with ritonavir and saquinavir were kept at 1.0 except for atoms that showed clear disorder, or where electron density was not visible even after a few rounds of refinement. Side-chains exposed to the surface, such as arginine and glutamate, are usually more flexible compared to other amino-acids residues; this flexibility can be expressed by the B -factor and/or the occupancy. The occupancy monitors the presence of an atom at its average coordinate position. It ranges from 0.0 to 1.0 and depends on the time the atom spends in its average position over the time of the data collection. On the other hand, the B -factor is a measure for the positional spread of an atom around its average coordinate position due to thermal motion. Except for high-resolution structures and in cases of multiple side chain conformation, occupancies are nowadays set to 1.0 or zero. This is because the occupancy and B -factor are mutually dependent of one another, *i.e.*, the occupancy

decreases as the B -factor increases and *vice versa*.

The inhibitor configuration was clearly seen in the active-site cleft when the σ^A -weighted $2F_o-F_c$ electron density map was contoured at $1.0\ \sigma$ together with the F_o-F_c maps contoured at $3.0\ \sigma$ above the mean (Figure 3.7). The minimized inhibitor (mention in 3.7.2) was manually docked into the F_o-F_c difference-density maps. As a result of the twofold non-crystallographic symmetry, the electron density map clearly shows two possible orientations for the saquinavir and ritonavir molecule. Two inhibitor orientations, each with occupancy of 0.5, were built into the electron density. Refinement continued for several additional rounds with improvements of stereochemistry, followed by the addition of water molecules to the model.

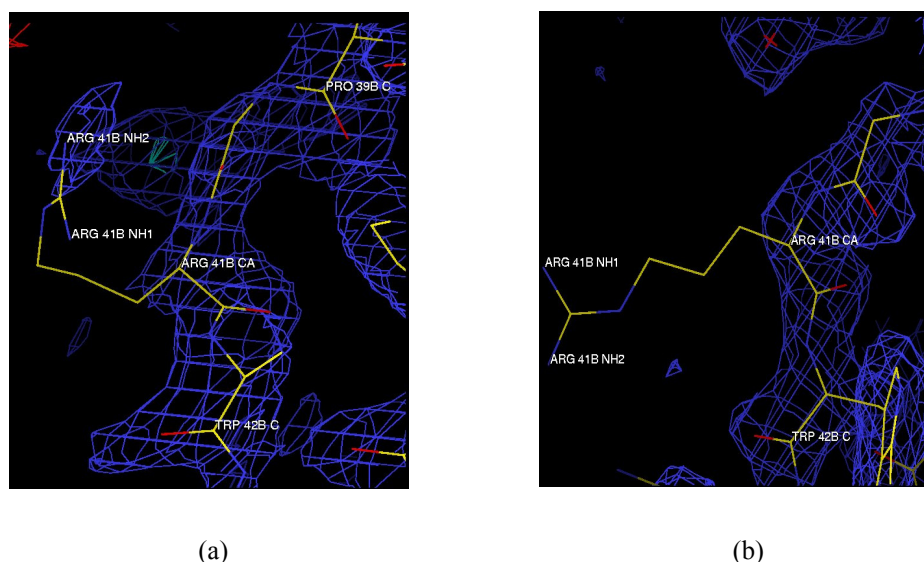
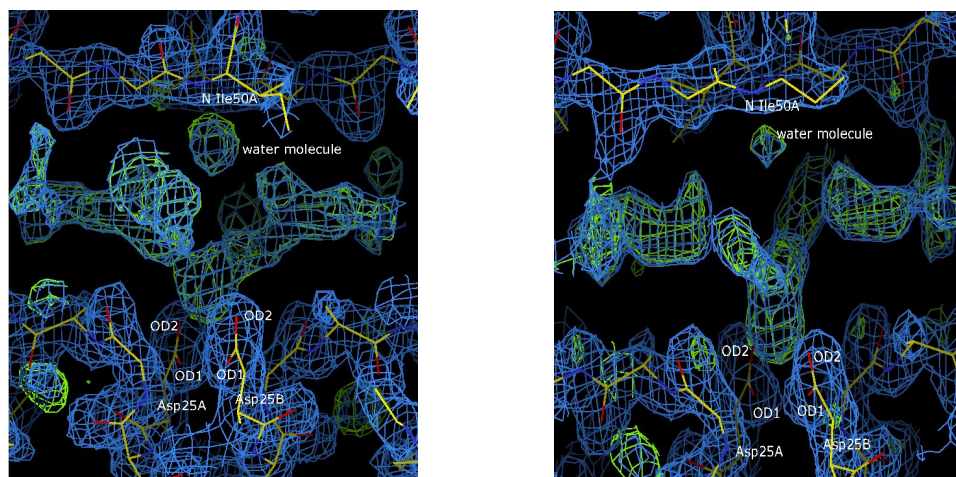


Figure 3.6: Example of a flexible side-chain. Residue Arg41 (chain B) in inhibitor-free variant day 115 at $2.20\ \text{\AA}$ resolution. σ^A -weighted $2F_o-F_c$ map contoured at $1.0\ \sigma$ in blue. (a) After the first round of refinement with full side-chain atom occupancy. (b) After the last round of refinement with occupancies = 0.0 and B -factors = $100.0\ \text{\AA}^2$ for the side-chain atoms of Arg41.



(a)

(b)

Figure 3.7: Difference electron density, F_o-F_c , and σ^A -weighted $2F_o-F_c$ maps of (a) saquinavir bound to variant day 82 at 1.76 Å (b) ritonavir bound to variant day 82 at 1.90 Å after the first steps of refinement. The σ^A weighted $2F_o-F_c$ map contoured at 1.0 σ is coloured blue and the F_o-F_c map contoured at 3.0 σ above the mean is coloured green. The spherically shaped σ^A -weighted $2F_o-F_c$ and F_o-F_c electron density above the inhibitor exposed the presence of the unique flap water molecule.

3.7.4 Modelling and refinement of solvent molecules

The highest remaining peaks in the electron density usually correspond to the oxygen atoms of water molecules even though salt or other molecules, *i.e.* Ca^{2+} , Mg^{2+} , DTT and PEG, may also be bound to the protein. In order to discriminate among these molecules, it is necessary to know the atom coordination geometry, crystallization conditions, purification strategy, and the biological function of the protein. Possible water oxygen atoms were judged by three criteria:

- (1) Peak in the F_o-F_c map is more than 3.0 σ above the mean,
- (3) A spherically shaped maximum in the σ^A -weighted $2F_o-F_c$ map (contoured at 1.0 σ) is observed at the same position, and
- (4) Presence of four hydrogen-bonded partner atoms (2 donors and 2 acceptors) at distances between 2.3 to 3.5 Å, leading to a tetrahedral coordination geometry for the respective water atom or, at least, the presence of one to two hydrogen-bonded partner atoms.

The F_o-F_c map was automatically analyzed to find the highest peaks that had no bad contacts and allowed sensible hydrogen bonding. Water molecules were first added to

the highest peaks in the F_o-F_c electron density map, and, in the later stages of the refinement, water molecules at weaker peaks were modelled (existence of at least one hydrogen-bonding partner) and refined with occupancy = 0.5 where necessary. If the B -value of the refined water molecules increased above 50 \AA^2 , the peaks were considered noise and not modelled as water molecules.



Figure 3.8: Example of water modelling with clear σ^A -weighted $2F_o-F_c$ and F_o-F_c electron density. For this water molecule in variant day 115 in complex with saquinavir, a tetrahedral coordination-geometry is observed.

3.7.5 Structure validation

The stereochemical quality of the structures was checked by using the programmes PROCHECK (Laskowski *et al.*, 1993) and SFCHECK (Vaguine *et al.*, 1999).

The program PROCHECK requires a PDB coordinate file as input and provides information about:

- (1) Ramachandran plot
- (2) χ^1 - χ^2 plots
- (3) Main-chain parameters: a) Ramachandran plot quality, b) peptide bond planarity, c) bad non-bonded interactions, d) $C\alpha$ tetrahedral distortion, e) main-chain hydrogen bond energy and f) overall G-factor (computed properties of torsion angles)
- (4) Side-chain parameters: a) standard deviation of the χ^1 gauche minus torsion angles, b) standard deviation of the χ^1 trans torsion angles, c) standard deviation of

the χ -1 gauche plus torsion angles, d) pooled standard deviation of all χ -1 torsion angles, e) standard deviation of the χ -2 trans torsion angles.

(5) Residue properties - this plot show how the protein's geometrical properties vary along its sequence.

(6) Main-chain bond-length distribution

(7) Main-chain bond-angle distribution

(8) R.m.s. deviation from planarity

(9) Distorted geometry plots

SFCHECK was run for the unliganded variants of HIV-1 proteinase and of complexes with saquinavir and ritonavir against the experimental data. The program thus requires an MTZ file and PDB coordinates and provides information about resolution range, completeness, Wilson plot, pseudo-translation, used reflections, twinning, overall B -factor, R_{factor} , R_{free} , correlation factors and Luzzati plot (Luzzati, 1952; estimation of the errors in the atomic coordinates).

3.8 Structural comparisons

Structures listed in Table 3.10 were used for comparison and analysis. The ALIGN (Cohen, 1997) and LSQKAB (Kabsch, 1976) programmes were used for superimposing Ca atoms. Pictures were prepared using the PyMOL molecular graphics suite (DeLano Scientific LLC, South San Francisco, U.S.A.). Binding affinity can be expressed in terms of free energy, $\Delta G = -RT \ln K$ (K is the equilibrium constant, R is the gas constant and T is the temperature). The estimated total binding energy gives information of the binding constant, *i.e.*, the more negative ΔG , the larger is the binding constant thus indicating stronger binding of the ligand to the proteinase. The estimated total binding energy of saquinavir or ritonavir bound to HIV-1 PR variants was calculated using AutoDock 3.0.5 (Morris *et al.*, 1998) and compared to the wild-type complex with inhibitor. Separately, the surface area of interaction between inhibitor and PR molecule was calculated in order to determine whether the binding pocket of the HIV-1 PR variant had increased or decreased when compared to the wild-type complex. The surface area (\AA^2) was calculated using the Molcad module in Sybyl 7.2 (Tripos Associates, Inc., St. Louis, MO, USA). In

Molcad, surfaces were calculated between van der Waals (VDW) surfaces of all atoms and were defined by a set of points halfway between the shortest distance separating the inhibitor and the enzyme. The grid width was set to 1.0 Å and the virtual sphere radius was set to 10.0 Å. Structural alignments of the backbone atoms were done using the Biopolymer module in Sybyl 7.1. The volume (Å³) of the binding site cavity was calculated, using SURFNET (Laskowski, 1995), in order to determine whether there is an expansion or reduction of cavity volume between variant and wild-type complex. The hydrophobic and VDW contacts of the ligand with the active-site residues were computed using CONTACT as incorporated into the CC4 program suite (Collaborative Computational Project 4, 1994).

Table 3.10: Structures used for alignment and analysis.

Structures	Space group	Reference
Day 28 ^b (Met36Ile-Ile54Val) with saquinavir	P6 ₁	This study
Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr) with saquinavi	P6 ₁	This study
Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr) with saquinavir	P6 ₁	This study
Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr) with ritonavir	P6 ₁	This study
Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr) with ritonavir	P6 ₁	This study
Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr) inhibitor-free	P6 ₁	This study
Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr) inhibitor-free	P6 ₁	This study
Day 0 (Val62-Val77)	P4 ₁	Fischer, 2004
Inhibitor-free day 28 ^a (Val82Thr)	P4 ₁	Fischer, 2004
Inhibitor-free day 28 ^b (Met36Ile-Ile54Val)	P2 ₁ 2 ₁ 2 ₁	Fischer, 2004
Inhibitor-free day 56 (Met36Ile-Ile54Val-Val82Thr)	P4 ₁	Fischer, 2004

Structures	Space group	Reference
3phv.pdb (wild-type)	P4 ₁ 2 ₁ 2	Lapatto <i>et al.</i> , 1989
1fb7.pdb (Gly48Val-Leu90Met double mutant complex with saquinavir)	I2 ₁ 3	Hong <i>et al.</i> , 2000
1hxb.pdb (wild-type complex with saquinavir)	P6 ₁	Krohn <i>et al.</i> , 1991
1sh9.pdb (Lys20Arg-Val32Ile-Leu33Phe-Met36Ile-Ile54Val-Leu63Pro-Ala71Val-Val82Ala-Ile84Val-Leu90Met, mutant complex with ritonavir)	P6 ₁	Clemente <i>et al.</i> , 2004
1hxx.pdb (wild-type with ritonavir)	P2 ₁ 2 ₁ 2	Kempf <i>et al.</i> , 1995
1izi.pdb (Ala71Val-Val82Thr-Ile84Val, mutant complex with Q50, {Q50=C38 H47 N5 O7})	P6 ₁	Weber <i>et al.</i> , 2002
1g6l.pdb (Cys95Met, inhibitor-free HIV-1 proteinase)	P6 ₁	Pillai <i>et al.</i> , 2000
1rpi.pdb (MDR769, mutation at position Leu10Ile-Asp25Asn-Met36Val-Met46Leu-Ile54Val-Ile62Val-Leu63Pro-Ala71Val-Val82Ala-Ile84Val-Leu90Met, inhibitor-free HIV-1 proteinase)	P4 ₁	Logsdon <i>et al.</i> 2004

4 RESULTS

4.1 Six HIV-1 proteinase variants

A set of six proteinase variants (Table 4.1) with different replication capacities (RC), derived from a HIV-1 infected patient during ritonavir monotherapy, was obtained (Nijhuis *et al.*, 1999). In total, eleven X-ray structures of the six different variants were examined in this study,

1. six of which are inhibitor-free: day 0 (Ile62Val-Ile77Val, the patient's 'wild type' strain), day 28^a (Val82Thr), day 28^b (Met36Ile-Ile54Val), day 56 (Met36Ile-Ile54Val-Val82Thr), day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr), and day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr), [Inhibitor-free: day 0, day 28^a, day 28^b and day 56 were crystallized and refined by Fischer (2004)]
2. three are in complex with the inhibitor saquinavir, day 28^a (Val82Thr), day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr), and day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr), [Crystals of saquinavir complexes were grown by Fischer (2004)]
3. two are in complex with ritonavir, day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr) and day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr).

Table 4.1: Effect of the substitutions of amino-acid residues during ritonavir monotherapy on viral fitness of the variants examined in this study according to Nijhuis *et al.*, (1999).

Variant	Mutations	K_i (nM) (n-fold increased)	IC ₅₀ (μM)	k_{cat}/K_m	RC value
Day 0	Ile62Val-Ile77Val	4.7	0.02	20 / 12 = 1.7	1.0
Day 28 ^a	Val82Thr	80 (17x)	0.16	n.d. / 11	0.9
Day 28 ^b	Met36Ile-Ile54Val	41 (9x)	0.21	n.d. / 34	< 0.6
Day 56	Met36Ile-Ile54Val-Val82Thr	1930 (410x)	0.96	20 / 36 = 0.5	< 0.6
Day 82*	Met36Ile-Ile54Val-Ala71Val-Val82Thr	1870 (398x)	0.75	67 / 13 = 5.3	1.2
Day 115	Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr	1820 (387x)	0.65	n.d. / 15	1.1

*This variant was also observed on day 115.

Abbreviation: n.d., not determined.

K_i : dissociation constant for inhibitor binding

IC₅₀: concentration of drug needed to reduce HIV replication in cell culture by 50%

k_{cat} : catalytic constant, which is the maximum number of substrate molecules converted to product per active-site of the enzyme per unit time.

K_m : the substrate concentration that produces half-maximal velocity, in the presence of a competitive inhibitor.

k_{cat}/K_m : measure of the overall efficiency of the enzyme catalysis and is known as the specificity constant.

On day 0, prior to therapy start, two substitutions in HIV-1 proteinase are observed that are the result of sequence polymorphism (so-called quasi species). The two amino-acid substitutions, Ile62Val and Ile77Val, are clearly visible in the electron density maps. According to Nijhuis *et al.* (1999), one variant observed on day 28 carries the Val82Thr substitution (day 28^a, see Table 4.1). Variant day 28^a shows a similar replication capacity as wild type along with a 17-fold increase in K_i value. This variant, day 28^a, exists alongside a variant that carries two substitutions, namely Met36Ile and Ile54Val. The latter variant, day 28^b, shows a 9-fold increase in K_i value and a clear reduction in viral replication capacity.

The unification of Val82Thr and Met36Ile-Ile54Val substitutions leads to a new variant, day 56 (Met36Ile-Ile54Val-Val82Thr), which displays a 411-fold increase in K_i value when compared to day 0. The viral replication capacity is the same as that seen for day 28^b (Met36Ile-Ile54Val). Amino-acid residue 82 is located in the active-site of the proteinase and directly influences substrate / inhibitor binding (Figure 4.1). Met36Ile and Ile54Val are distal to the substrate binding-site. Met36Ile is located at the 30's loop while Ile54Val is located at a β -strand in the flap region. This triple mutant, Met36Ile-Ile54Val-Val82Thr, shows the highest K_i value when compared to the other variants alongside the lowest replication efficiency.

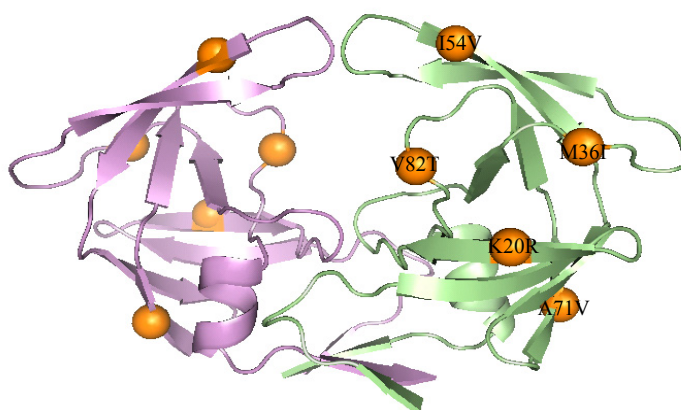


Figure 4.1: The HIV-1 proteinase with mutations examined in this study. Orange spheres mark residue changes acquired during ritonavir monotherapy.

Continuous viral replication under antiretroviral drug pressure selects for proteinase variants that acquire the Ala71Val and Ala71Val-Lys20Arg substitutions in addition to Met36Ile-Ile54Val-Val82Thr. The variant harboring a total of four substitutions, day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr), is observed not only

on day 82 but also on day 115. Finally, a variant with five substitutions, Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr, emerges on day 115. Lys20Arg is located at the β -strand that connects the N-terminus and the catalytic region. Ala71Val is located at the β -strand that connects the flap and 80's loop (Figure 4.1). Combining either Ala71Val or Ala71Val-Lys20Arg with Met36Ile-Ile54Val-Val82Thr doesn't further increase the K_i value but drastically improves the viral replication efficacy to better than wild type (Table 4.1).

Cross-resistance is a serious problem for long-term antiretroviral therapy. In order to investigate the molecular basis of HIV proteinase drug resistance, the variants were co-crystallized with saquinavir and ritonavir. However, only three proteinase-inhibitor complexes (day 28^b, day 82, and day 115) with saquinavir and two complexes (day 82 and day 115) with ritonavir produced crystals and these diffracted X-rays to a good resolution. Structural comparison of the complexes was performed to observe the effects of the amino-acid substitutions on the ligand-proteinase binding interactions, *e.g.* van der Waals (VDW), hydrophobic, and hydrogen-bonding interactions.

4.2 Crystallographic analysis

Crystallographic analysis for each of the inhibitor-free and complexes of HIV-1 proteinase variants are as listed in Table 4.2.

Table 4.2: Analysis of the eleven X-ray structures. Crystallographic analysis of variants days 0, 28^a, 28^b and 56 were taken from Fischer (2004).

	Inhibitor-free HIV-1 proteinase					
	Day 0 (Ile62Val- I77V)	Day 28 ^a (Val82Thr)	Day 28 ^b (Met36Ile- Ile54Val)	Day 56 (Met36Ile- Ile54Val- Val82Thr)	Day 82 (Met36Ile- Ile54Val- Ala71Val- Val82Thr)	Day 115 (Lys20Arg- Met36Ile- Ile54Val- Ala71Val- Val82Thr)
Resolution (Å)	50.00-2.18	50.00-2.65	64.55-2.40	46.63-2.35	53.76-2.40	53.53-2.20
Space group	P4 ₁	P4 ₁	P2 ₁ 2 ₁ 2 ₁	P4 ₁	P6 ₁	P6 ₁
Unit cell dimensions (Å)						
a	49.889	49.780	30.828	46.590	62.11	61.85
b	49.889	49.780	49.176	46.590	62.11	61.85
c	108.130	107.959	129.643	101.996	83.58	83.28
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 120	90, 90, 120
<I/>>/<σ (I)> in the final shell	132.3/60.3 = 2.19	142.1/64.0 = 2.22	275.7/75.6 = 3.65	414.9/189.0 = 2.19	370.1/118.3 = 3.13	333.8/100.1 = 3.33
Overall R _{sym} (%) ^a , all data final shell	4.4 38.4	5.9 31.7	12.7 25.1	8.8 26.7	7.6 39.2	7.5 47.4
Number of reflections	60,781	25,981	51,078	24,169	23,660	31,958
Number of unique reflections	13,348	7,524	7,938	8416	7165	9209
Redundancy	4.5	3.4	6.4	2.9	3.3	3.5
R _{factor} (%) ^b	20.28	17.99	19.70	22.06	20.76	21.29
R _{free} (%) ^c	27.12	25.15	28.99	26.66	30.86	29.58
No. of water molecules	22	14	20	11	45	49
Completeness (%)	96.4	97.9	96.0	90.7	99.6	81.9
R.m.s. deviations Bonds (Å) Angles (°)	0.023 1.977	0.025 1.970	0.021 1.968	0.021 1.934	0.021 2.071	0.019 1.969
Ramachandran plot Favorable (%) Additional (%) Generous (%) Forbidden (%)	96.8 3.2 0.0 0.0	93.6 5.8 1. (Ile50) 0.0	94.2 5.1 0.0 0.6 (I47')	89.7 7.1 (Arg41, Asp35', Leu38') 1.3 (Leu38, Arg41')	91.0 8.3 0.6 (Cys67) 0.0	93.2 5.4 1.4 (Cys67) 0.0

^aR_{sym} = $\sum |I_i - I_m| / \sum I_m \times 100$, where I_i and I_m are the observed intensity and mean intensity of related reflections h, k, l , respectively.

^bR_{factor} = $\sum |F_{obs} - F_{calc}| / \sum F_{obs}$ for the 95% of the data included in the refinement.

^cR_{free} = $\sum |F_{obs} - F_{calc}| / \sum F_{obs}$ for the 5% of the data randomly selected and excluded from refinement.

I / σ (I) = Integrated intensity/Standard deviation of intensity.

Complexes of HIV-1 proteinase variants with saquinavir or ritonavir

	Complexes with saquinavir			Complexes with ritonavir	
	Day 28^b (Met36Ile-Ile54Val)	Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)	Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)	Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)	Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)
Resolution (Å)	46.25-1.95	29.01-1.76	45.18-1.95	30.89-1.90	26.74-1.77
Space group	P2 ₁ 2 ₁ 2 ₁	P6 ₁	P6 ₁	P6 ₁	P6 ₁
Unit cell dimensions (Å)					
a	28.907	61.912	62.057	61.740	61.775
b	66.492	61.912	62.057	61.740	61.775
c	92.503	83.099	83.293	83.817	83.601
α, β, γ (°)	90, 90, 90	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120
<I>/<σ (I)> in the final shell	1169.1/249.5 = 4.68	324.8/69.7 = 4.65	176.6/33.2 = 5.32	620.4/124.2 = 5.0	397.3/98.9 = 4.0
Overall R _{sym} (%) ^a , all data final shell	8.8 26.7	7.6 39.2	7.5 47.4	6.5 29.6	5.2 46.5
Number of reflections	99,758	90,685	112,638	95,107	115,958
Number of unique reflections	11,440	16,901	12,683	13,559	16,773
Redundancy	8.7	5.4	8.9	7.0	6.9
Completeness (%)	88.21	98.82	99.72	99.90	99.82
R _{factor} (%) ^b	17.06	21.03	19.80	18.74	18.32
R _{free} (%) ^c	24.23	27.44	26.80	25.75	23.68
No. of water molecules	124	131	121	141	177
R.m.s. deviations					
Bonds (Å)	0.021	0.019	0.019	0.019	0.020
Angles (°)	1.910	2.039	2.073	2.040	2.047
Average refined B-factors (Å ²)					
Protein atoms	16.86	27.85	25.12	24.76	21.90
Inhibitor	23.52	19.60	15.35	24.63	27.98
Solvent	20.02	32.71	26.32	31.78	32.43
Average B-factors (Å ²) from the Wilson plot	16.62	27.51	24.33	24.07	22.03
Ramachandran plot					
Favorable (%)	94.2	93.5	94.9	94.2	94.9
Additional (%)	5.8	6.5	4.5	5.8	5.1
Generous (%)	0.0	0.0	0.6 (C67)	0.0	0.0
Forbidden (%)	0.0	0.0	0.0	0.0	0.0

^aR_{sym} = Σ|I_i - I_m|/Σ I_m, where I_i and I_m are the observed intensity and mean intensity of related reflections, respectively.

^bR_{factor} = Σ|F_{obs} - F_{calc}|/Σ F_{obs} for the 95% of the data included in the refinement.

^cR_{free} = Σ|F_{obs} - F_{calc}|/Σ F_{obs} for the 5% of the data randomly selected and excluded from refinement.

I / σ (I) = Integrated intensity/Standard deviation of intensity.

Crystals of saquinavir complexes were grown by Fischer, 2004

4.3 Crystallization

4.3.1 Crystals of inhibitor-free HIV-1 proteinase, day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)



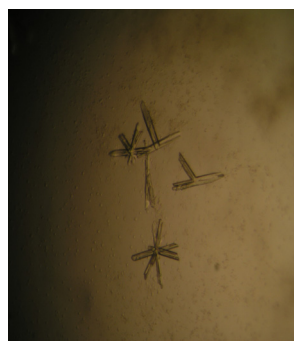
A

Precipitant

A. 20% $(\text{NH}_4)_2\text{SO}_4$
100 mM Na cacodylate pH 6.5

protein buffer

20 mM Na acetate pH 5.0
5 mM DTT, 1 mM EDTA
20 mM NaCl



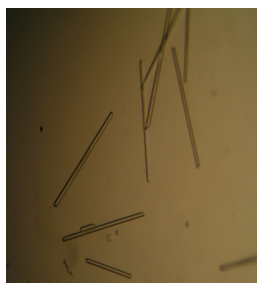
B

Precipitant

B. 25% $(\text{NH}_4)_2\text{SO}_4$
100 mM Mes pH 6.5

protein buffer

20 mM Na acetate pH 5.0
5 mM DTT, 1 mM EDTA
20 mM NaCl



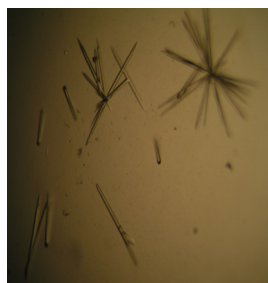
C

Precipitant

C. 1.0 M NaCl
100 mM Mes pH 6.5
10 mM DTT

protein buffer

20 mM Na acetate pH 5.0
5 mM DTT, 1 mM EDTA
20 mM NaCl



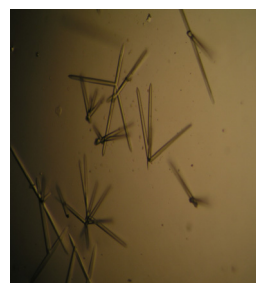
D

Precipitant

D. 1.25 M NaCl
100 mM Mes pH 6.5
10 mM DTT

protein buffer

20 mM Na acetate pH 5.0
5 mM DTT, 1 mM EDTA
20 mM NaCl



E

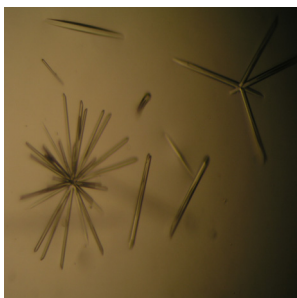
Precipitant

E. 1.50 M NaCl
100 mM Mes pH 6.5
10 mM DTT

protein buffer

20 mM Na acetate pH 5.0
5 mM DTT, 1 mM EDTA
20 mM NaCl

4.3.2 Crystals of inhibitor-free HIV-1 proteinase, day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)



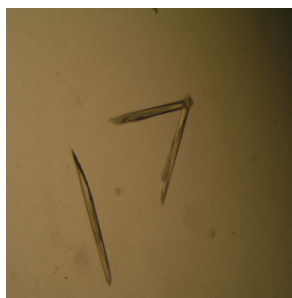
F

Precipitant

F. 20% $(\text{NH}_4)_2\text{SO}_4$
100 mM Hepes pH 6.5

protein buffer:

20 mM Mes pH 6.0
5 mM DTT, 1 mM EDTA
20 mM NaCl



G

Precipitant

G. 20% $(\text{NH}_4)_2\text{SO}_4$
50 mM Na cacodylate
pH 6.0

protein buffer:

20 mM Mes pH 6.0
5 mM DTT, 1 mM EDTA
20 mM NaCl



H

Precipitant

H. 1.0 M KCl
100 mM Mes pH 6.0

protein buffer:

20 mM Mes pH 6.0
5 mM DTT, 1 mM EDTA
20 mM NaCl

4.3.3 Crystals of HIV-1 proteinase day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr) in complex with ritonavir



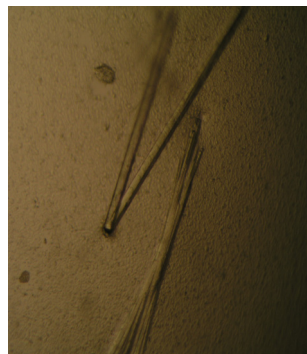
I

Precipitant

I. 25% $(\text{NH}_4)_2\text{SO}_4$
100 M Mes pH 6.5

protein buffer:

20 mM Na acetate pH 5.0
5 mM DTT, 1 mM EDTA
20 mM NaCl



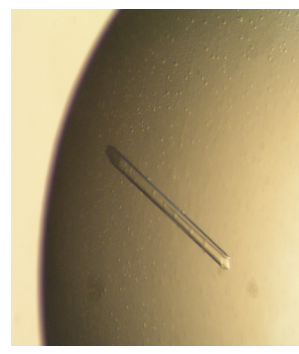
J

Precipitant

J. 25% $(\text{NH}_4)_2\text{SO}_4$
100 mM Hepes pH 6.5

protein buffer:

20 mM Na acetate pH 5.0
5 mM DTT, 1 mM EDTA
20 mM NaCl



K

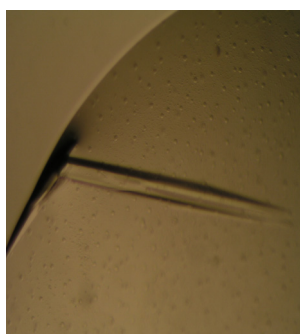
Precipitant

K. 25% $(\text{NH}_4)_2\text{SO}_4$
200 mM NaH_2PO_4 pH 6.5

protein buffer:

20 mM Na acetate pH 5.0
5 mM DTT, 1 mM EDTA
20 mM NaCl

4.3.4 Crystals of HIV-1 proteinase day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr) in complex with ritonavir



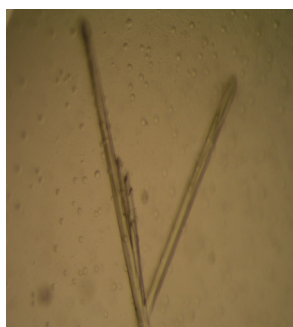
L

Precipitant

L: 20% (NH₄)₂SO₄
100 mM Mes pH 6.0

protein buffer:

20 mM Na acetate pH 5.0
5 mM DTT, 1 mM EDTA
20 mM NaCl



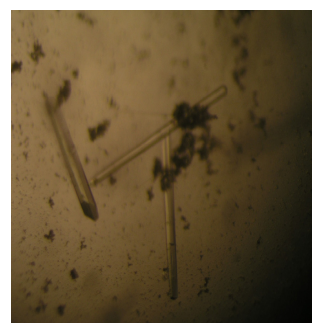
M

Precipitant

M: 25% (NH₄)₂SO₄
100 mM Mes pH 6.0

protein buffer:

20 mM Na acetate pH 5.0
5 mM DTT, 1 mM EDTA
20 mM NaCl



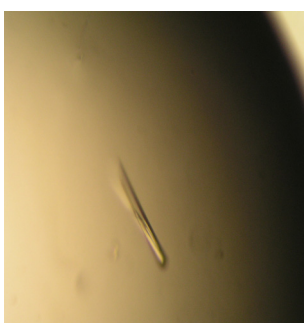
N

Precipitant

N: 25% (NH₄)₂SO₄
200 mM NaH₂PO₄ pH 6.0

protein buffer:

20 mM Na acetate pH 5.0
5 mM DTT, 1 mM EDTA
20 mM NaCl



20 mM NaCl
O

Precipitant

O: 1.0 M KCl
100 mM MES pH 6.0
3 mM NaN₃
10 mM DTT

protein buffer:

20 mM MES pH 6.0
1 mM DTT, 1 mM EDTA

Co-crystallization of variants day 0, day 28^a, day 28^b and day 56 (Met36Ile-Ile54Val-Val82Thr) with ritonavir did not produce any crystals. Also, no crystals were produced from co-crystallization of variants day 0, day 28^a and day 56 (Met36Ile-Ile54Val-Val82Thr) with saquinavir. Of all the complexes, the late variants day 82 and day 115 in complex with ritonavir crystallized promptly.

4.4 Structural comparison between the HIV-1 proteinase variants

The root-mean-square deviations (r.m.s. deviations) for C α atoms (of 198 amino-acid residues) of the variants, in comparison to variant day 0 (inhibitor-free),

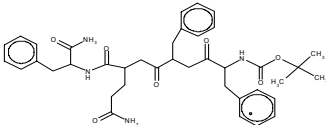
are presented in Tables 4.3 and 4.4.

Table 4.3: R.m.s. deviation values (Å) for Ca atoms (of 198 amino-acid residues) between inhibitor-free HIV-1 proteinase and its complexes with ritonavir or saquinavir.

	Inhibitor-free HIV-1 proteinase					
	Day 0	Day 28 ^a (Val82Thr)	Day 28 ^b (Met36Ile-Ile54Val)	Day 56 (Met36Ile-Val54Ile-Val82Thr)	Day 82 (Met36Ile-Val54Ile-Ala71Val-Val82Thr)	Day 115 (Lys20Arg-Met36Ile-Val54Ile-Ala71Val-Val82Thr)
Day - 0 average r.m.s. deviation	-	0.14	0.53	0.69	1.45	1.43
Saquinavir - day 28^b average r.m.s. deviation	1.20	1.14	0.96	1.17	0.66	0.64
Saquinavir - day 82 average r.m.s. deviation	1.39	1.37	1.04	1.35	0.35	0.35
Saquinavir - day 115 average r.m.s. deviation	1.38	1.35	0.99	1.44	0.30	0.29
Ritonavir - day 82 average r.m.s. deviation	1.41	1.42	1.03	1.29	0.31	0.27
Ritonavir - day 115 average r.m.s. deviation	1.40	1.39	1.04	1.37	0.27	0.24

Table 4.4: R.m.s. deviation values (Å) for Ca atoms (of 198 amino-acid residues) for complexes with ritonavir or saquinavir.

	Complexes with saquinavir			Complexes with ritonavir	
Reference	Day 28 ^b (Met36Ile-Ile54Val)	Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)	Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)	Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)	Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)
^a 1izi.pdb average r.m.s. deviation	0.46	0.30	0.29	0.32	0.32
^b 1hxb.pdb average r.m.s. deviation	0.43	0.31	0.32	0.39	0.40
^c 1hwx.pdb average r.m.s. deviation	0.61	0.54	0.53	0.50	0.49
^d 1sh9.pdb average r.m.s. deviation	0.36	0.38	0.37	0.34	0.34
Day 28^b - saquinavir average r.m.s. deviation	-	0.47	0.48	0.50	0.52
Day 82 - saquinavir average r.m.s. deviation	-	-	0.15	0.19	0.20
Day 115 - saquinavir average r.m.s. deviation	-	-	-	0.18	0.16

^a1izi.pdb = complex of Q50 (C₃₈H₄₇N₅O₇, ) with variant HIV-1 proteinase (Ala71Val-Val82Thr-Ile84Val)

^b1hxb.pdb = complex of saquinavir with wild-type HIV-1 proteinase

^c1hwx.pdb = complex of ritonavir with wild-type HIV-1 proteinase

^d1sh9.pdb = complex of ritonavir with variant HIV-1 proteinase (Lys20Arg-Val32Ile-Leu33Phe-Met36Ile-Leu63Pro-Ala71Val-Val82Ala-Ile84Val-Leu90Met)

4.5 Crystallographic analysis – Ritonavir and saquinavir complexes

Crystals of day 28^b, day 82 and day 115 complex with saquinavir were grown by Fischer (2004). In order to elucidate the structural basis for the increasing resistance towards ritonavir and the overall viral fitness, the following structures (inhibitor-free and complexes with ritonavir or saquinavir) are dissected. Chemical structures of saquinavir and ritonavir are as shown in Figure 4.2.

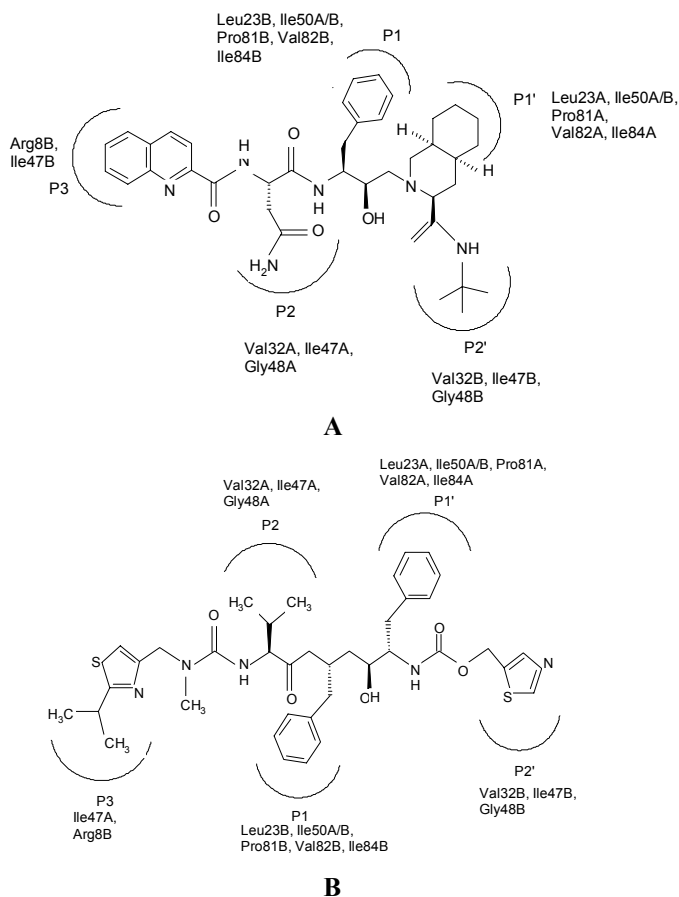


Figure 4.2: Chemical structures of the HIV-1 proteinase inhibitors used in the present study. Ligand residues recognized by the proteinase are labelled as P3, P2, P1, P1' and P2'. (A) Saquinavir: P1; Phenylalanine, P2; Asparagine, P3; 2-Carbonylquinoline, P1'; 2-Methyl-decahydro-isoquinoline-3-carboxylic acid, P2'; Tertiary-butylamine. (B) Ritonavir: P1; Phenylalanine, P2; Valine, P3; 2-Isopropyl-4-thiazoyl, P1'; Phenylalanine, P2'; 5-Thiozoyl.

4.5.1 Variants of HIV-1 proteinase: day 28^b (Met36Ile-Ile54Val), day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr) and day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr) in complex with saquinavir

Inhibitor complexes of day 82 and day 115 crystallized in space group P6₁ whereas day 28^b crystallized in space group P2₁2₁2₁ (Figure 4.3). In each structure, the inhibitor was observed to bind to the active-site cavity in two orientations, which are related to one another by the twofold symmetry axis of the homodimeric enzyme (Figure 4.4). Krohn *et al.* (1991) reported a similar binding mode for wild-type proteinase in complex with saquinavir.

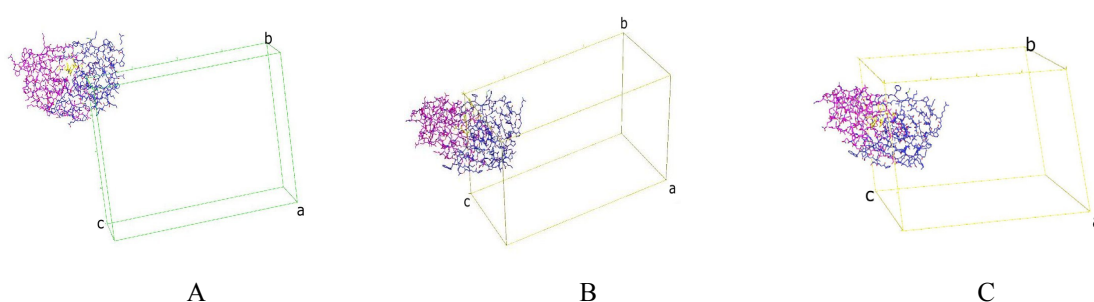


Figure 4.3: Unit cells of inhibitor complexes of (A) day 28, unit cell dimensions 28.4 Å x 66.5 Å x 92.5 Å, (B) day 115, 62.1 Å x 62.1 Å x 83.3 Å, and (C) day 82, 61.9 Å x 61.9 Å x 83.1 Å. Monomer A, pink; monomer B, blue; saquinavir, yellow.

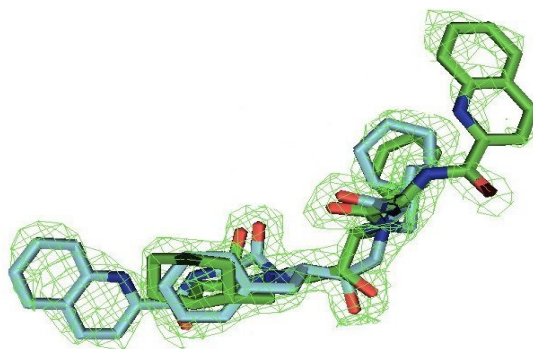


Figure 4.4: Saquinavir binds in two modes related by the twofold symmetry of the homodimer in variant day 28^b. The 2F_o-F_c electron density is contoured at 1.0 σ above the mean. Complexes of day 115 and day 82 with saquinavir also exhibit two binding modes of the saquinavir molecule at the active-site cavity.

All HIV-1 proteinase variant complexes with saquinavir had good electron density. Only few residues lack side-chain electron density. The side-chains of arginine 41A and 41B are disordered in all structures and their occupancies were set

to zero whilst the B -factors were set to 100 \AA^2 . This flexible residue is situated at the surface of the protein. The region containing the Met36Ile mutation (residues Glu34 to Leu38) from chains A and B show well-defined electron density except for the side-chain of Glu34 in chain B.

4.5.2 Proteinase-saquinavir interactions

The overall B -factor per structure decreases in the order, day 82 (green) > day 115 (blue) > 1hxb.pdb (wild-type complex with saquinavir, black) > day 28^b (red). The distribution of the averaged B -factor per residue of 1hxb.pdb, day 28^b, day 82 and day 115 exhibits similar patterns as indicated in Figure 4.5. The flap regions of variants day 82 and day 115 display slightly increased B -factors (overall 34 \AA^2) when compared to 1hxb.pdb (overall 26 \AA^2 ; Table 4.5). The B -factors of variant day 28^b are significantly lower: This could result from differences in crystal packing between the two space groups, $P2_1,2_1,2_1$ (day 28^b) and $P6_1$ (day 82, day 115, and 1hxb.pdb).

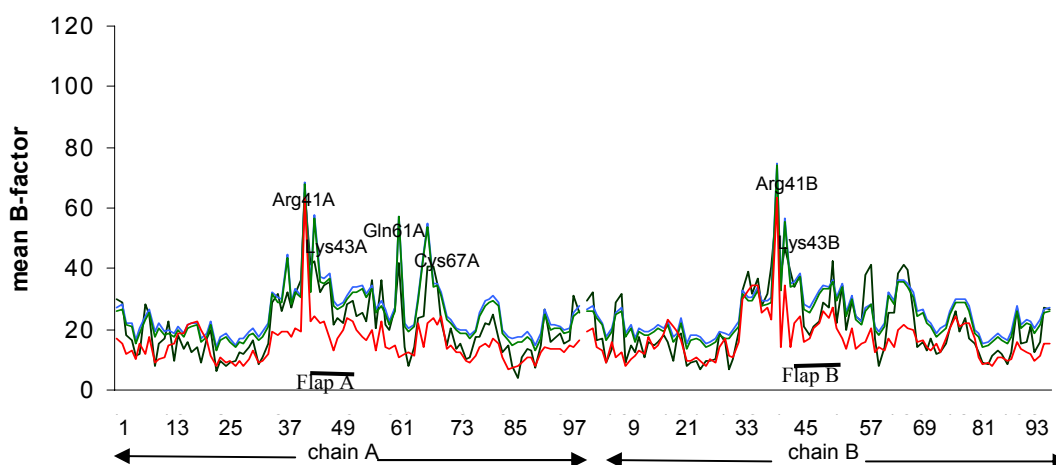


Figure 4.5: Averaged B -values per residue for the wild-type HIV-1 proteinase (1hxb.pdb, black), day 28^b (red), day 82 (green) and day 115 (blue), in complex with saquinavir. The residues in the two monomers are numbered 1-99 (chain A) and 101-199 (chain B).

Table 4.5: Comparison of the averaged B -values (\AA^2) per residue of the flap regions for 1hxb.pdb, day 28^b, day 82, and day 115.

Residues	1hxb.pdb		Day 28 ^b (Met36Ile-Ile54Val)		Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)		Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)	
	Chain A	Chain B	Chain A	Chain B	Chain A	Chain B	Chain A	Chain B
Lys45	34.6	33.9	23.2	22.8	35.3	34.5	36.8	35.8
Met46	35.8	38.2	17.3	24.3	36.7	37.2	38.4	38.3
Ile47	21.6	20.9	14.1	16.4	27.9	26.6	29.3	28.3
Gly48	23.8	18.3	17.5	17.5	26.6	25.6	27.7	27.2
Gly49	21.9	20.8	20.8	21.1	27.8	27.7	29.0	29.4
Ile50	28.4	22.9	24.9	23.2	29.8	31.5	31.1	33.1
Gly51	29.4	28.6	23.8	26.9	32.2	33.4	33.8	34.8
Gly52	24.2	27.4	20.3	24.5	32.4	33.2	34.1	34.2
Phe53	25.4	42.7	18.4	27.5	33.2	35.6	34.7	36.5
Ile54	22.6	21.9	16.9	20.8	30.4	29.7	31.9	30.6
Lys55	36.2	34.4	20.6	17.4	33.3	33.9	34.4	34.9

The **S1** subsite (Schechter & Berger, 1967) is lined by residues Leu23B, Ile50A/50B, Pro81B, Val82B, Ile84B, **S2** by residues Val32A, Ile47A, Gly48A, **S3** by Ile47A, Arg8B, **S1'** by Leu23A, Ile50A/50B, Pro81A, Val82A, Ile84A, **S2'** by Val32B, Ile47B, Gly48B and finally, **S3'** by Arg8A, Val82A, Ile47B (Figure 4.6).

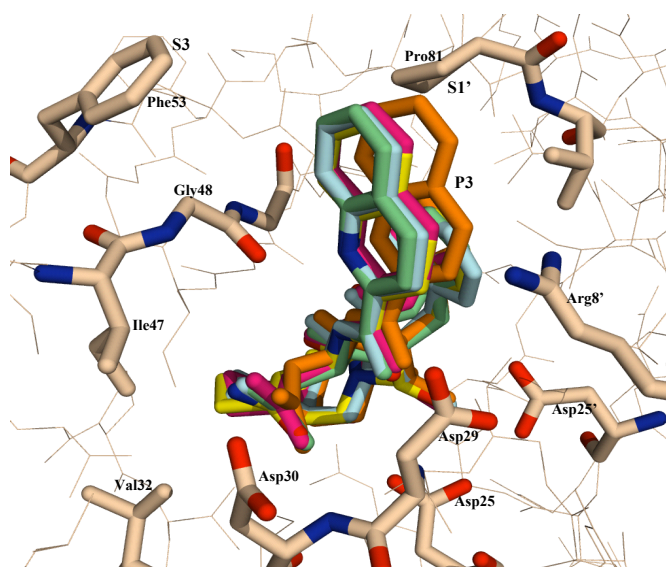


Figure 4.6: Residues that interact directly with saquinavir are labelled. Pink - day 28^b; cyan - day 82; green - day 115; yellow - 1hxb.pdb; orange - 1fb7.pdb (mutations Gly48Val and Leu90Met).

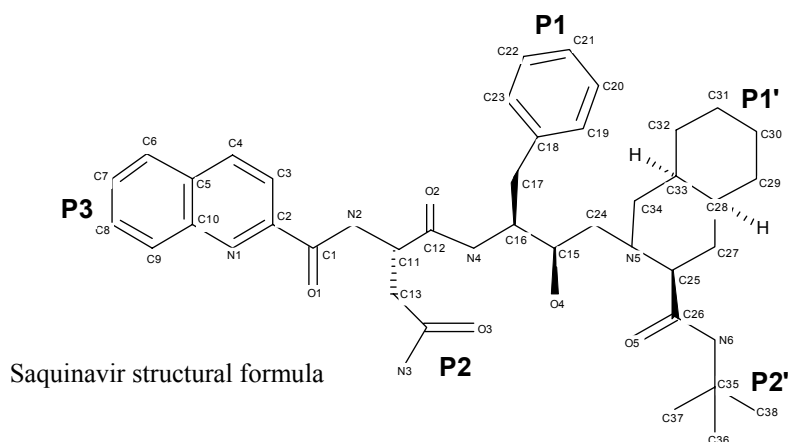
In one reported structure (PDB code 1fb7), the quinoline ring (Figure 4.6, P3 residue) of saquinavir has moved slightly away from the flaps to compensate for the Gly48Val substitution (Hong *et al.*, 2000) when compared to wild type (PDB code 1hxb). However, no significant movement of the P3 was observed in the complexes of

day 28^b, day 82 and day 115. Variant day 28^b does not contain any mutations at the active-site residues Val32, Val82 or Ile84, whereas variant day 82 and day 115 contain one substitution at the active-site, Val82Thr.

Using CONTACT, the hydrophobic and VDW interactions of saquinavir with amino-acid residues at the substrate-binding site of variants day 28^b, day 82, day 115, and 1hxb.pdb were computed. The hydrophobic proteinase-saquinavir interactions are shown in Table 4.6 (cut-off value of 3.7 Å). The hydrogen-bonded atoms, with a cut-off value of 3.5 Å, are listed in Table 4.7 (two orientations of inhibitor binding are included in the analysis).

Tighter hydrophobic contacts are observed for variants day 82 and day 115 in complex with saquinavir when compared to 1hxb.pdb (Table 4.6). However, chain X of bound saquinavir of day 28^b exhibits fewer contacts (31 contacts) when compared to 1hxb.pdb (34 contacts), variant day 82 (35 contacts) and day 115 (39 contacts). This was due to the loss of some hydrophobic contacts at the P1' residue of chain X. Another important feature was the loss of a critical hydrophobic interaction between Arg8B and 2-carbonylquinoline (P3) in variants day 82 and day 115 when compared to 1hxb.pdb and day 28^b. The loss of hydrophobic contacts at the P3 residue in variants day 82 and day 115 could be compensated for by the increase of hydrophobic contacts at residues P1, P1' and P2'. In variant day 115, tighter hydrophobic interactions were observed between the main-chain atoms of the inhibitor chain X and residues Ala28A, Asp25A, Gly27B of the proteinase. For the P1 residue of chain Y, tighter hydrophobic interactions were observed with residues Ile84A and Thr82A. The total number of hydrophobic interactions for inhibitor chains X and Y differs, which might be the result of a slightly incorrect occupancy refinement (fixed at 0.5 for each orientation).

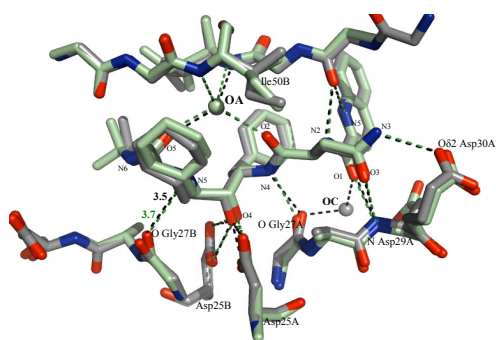
Table 4.6: Hydrophobic interactions between HIV-1 proteinase and saquinavir. The number of hydrophobic contacts is shown in parentheses (cut-off value of 3.7 Å). Saquinavir was observed to bind in two different orientations; chain A of saquinavir in 1hxb.pdb is equivalent to chain X and chain B of saquinavir in 1hxb.pdb is equivalent to chain Y in variants day 28^b, day 82 and day 115.



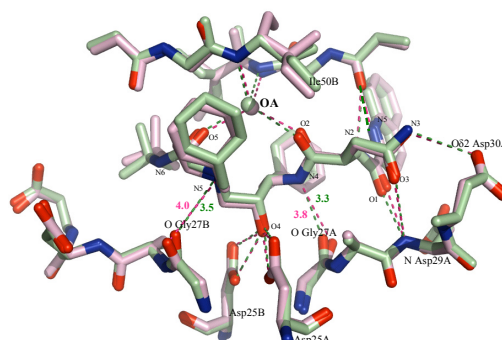
Saquinavir	1hxb.pdb (wild-type complex)	Day 28 ^b (Met36Ile-Ile54Val)	Day 82 (Met36Ile-Ile54Val- Ala71Val-Val82Thr)	Day 115 (Lys20Arg-Met36Ile- Ile54Val-Ala71Val- Val82Thr)
	Chain A	Chain X	Chain X	Chain X
P3 - 2- carbonylquinoline	Arg8B (2) Pro81B (2) Asp29A (1) Gly48A (2) 7 contacts	Arg8B (2) Pro81B (2) Gly48A (2) 6 contacts	Pro81B (1) Gly48A (2) 3 contacts	Gly48A (2) Gly49A (2) 4 contacts
P2 - Asparagine	-	-	-	-
P1 - Phenylalanine	Leu23B (1) Val82B (3) Gly49A (1) Ile50A (1) 6 contacts	Leu23B (1) Val82B (2) Ile84B (2) Gly49A (1) Ile50A (1) 7 contacts	Leu23B (1) Thr82B (2) Ile84B (3) 6 contacts	Leu23B (1) Thr82B (2) Ile84B (3) Gly49A (2) 8 contacts
P1' - 2-methyl-decahydro- isoquinoline-3-carboxyl	Gly49B (2) Ile50B (3) Pro81A (3) Val82A (3) Ile84A (3) 14 contacts	Gly27B (1) Gly49B (3) Ile50B (2) Thr80A (1) Pro81A (2) Ile84A (2) 11 contacts	Gly49B (3) Ile50B (4) Pro81A (4) Thr82A (1) Ile84A (3) 15 contacts	Gly49B (1) Ile50B (1) Pro81A (4) Thr82A (4) Ile84A (4) 14 contacts
P2' - Tertiary-butylamino	Ala28B (1) Asp30B (1) Val32B (1) Ile47B (1) Gly48B (1) Ile84B (1) 6 contacts	Val32B (1) Ile47B (1) Ile84B (1) Ile50A (3) 6 contacts	Ala28B (1) Gly48B (1) Ile47B (1) Ile84B (1) Ile50A (5) 9 contacts	Ala28B (3) Asp30B (1) Val32B (1) Ile47B (1) Ile84B (1) Ile50A (2) 9 contacts
C16 atom	Asp25A (1)			
C11 atom				Ala28A (2)
C15 atom			Asp25A (1)	Asp25A (1)
C24 atom		Gly27B (1)	Gly27B (1)	Gly27B (1)
Total	34	31	35	39

Saquinavir	1hxb.pdb (wild-type complex)	Day 28^b (Met36Ile-Ile54Val)	Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)	Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)
	Chain B	Chain Y	Chain Y	Chain Y
P3 - 2- carbonylquinoline	Gly48B (3) Gly49B (1) Arg8A (2) Pro81A (3) 9 contacts	Gly48B (2) Arg8A (2) Pro81A (3) 7 contacts	Asp29B (1) Gly48B (2) Pro81A (1) 4 contacts	Asp29B (1) Gly48B (2) Pro81A (3) 6 contacts
P2 - Asparagine	-	-	-	-
P1-Phenylalanine	Gly49B (2) Ile50B (1) Val82A (2) Ile84A (1) 6 contacts	Val82A (3) Gly49B (4) Ile50B (1) 8 contacts	Asp25B (1) Leu23A (2) Ile84A (2) Thr82A (2) 7 contacts	Asp25B (2) Gly49B (2) Leu23A (1) Ile84A (4) Thr82A (3) 12 contacts
P1' - 2-methyl-decahydro- isoquinoline-3-carboxyl	Pro81B (4) Val82B (3) Ile84B (1) Ile50A (2) 10 contacts	Pro81B (4) Val82B (1) Ile84B (3) Ile50A (3) Gly49A (2) 13 contacts	Pro81B (4) Thr82B (1) Ile84B (5) Gly49A (2) Ile50A (3) 15 contacts	Pro81B (4) Thr82B (1) Ile84B (3) Gly49A (2) Ile50A (1) 11 contacts
P2' - Tertiary-butylamino	Ile50B (3) Asp30A (1) Val32A (1) Ile47A (1) Gly48A (1) Ile84A (1) 8 contacts	Ile50B (2) Ala28A (1) Val32A (1) Ile47A (1) Gly48A (1) Ile84A (3) 9 contacts	Ile50B (2) Ala28A (2) Asp30A (1) Ile47A (1) Ile84A (1) 7 contacts	Ile50B (4) Ala28A (2) Asp30A (1) Gly49A (1) Val32A (1) 9 contacts
C11 atom	Ala28A (1)		Ala28B (2)	
C24 atom	Gly27A (1)	Gly27A (1)	Gly27A (1)	Gly27A (1)
Total	35	38	36	39

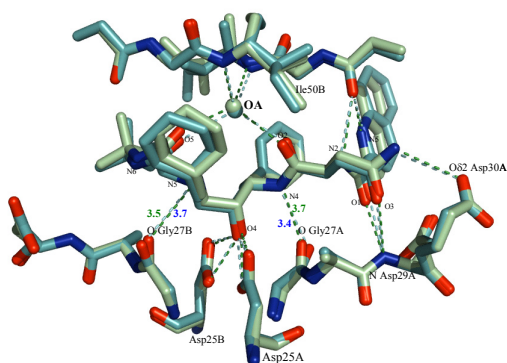
The proteinase-saquinavir hydrogen-bonding interactions for variants day 28^b, day 82 and day 115 compared to wild type, 1hxb.pdb, are shown in Figure 4.7.



A. Day 28^b (Met36Ile-Ile54Val)



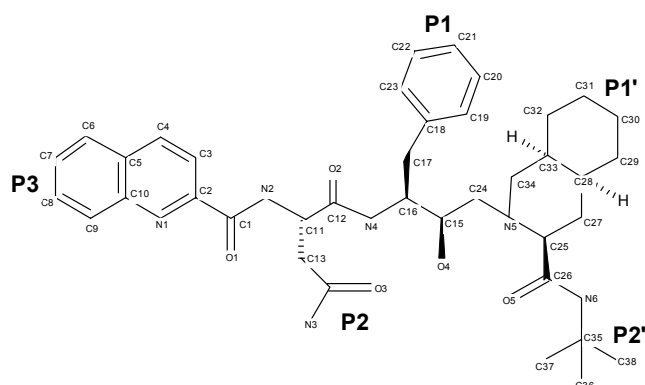
B. Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)



C. Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)

Figure 4.7: Hydrogen-bonding pattern of the proteinase-saquinavir complex (cut-off at 3.5 Å). (A) Variant day 28^b is coloured in black and wild type, 1hxb.pdb, is coloured in green; (B) Variant day 82 is coloured in pink; (C) Variant day 115 coloured in cyan.

Table 4.7: Proteinase-saquinavir hydrogen-bonding interactions (cut-off value of 3.5 Å) in 1hxb.pdb, day 28^b, day 82 and day 115. Only hydrogen-bonding interactions of the proteinase with chain X of saquinavir are indicated since similar interactions are observed for chain Y.



Saquinavir structural formula

Atoms			Distance (Å)			
Chain X of saquinavir in variants day 28 ^b , day 82, day 115	Water molecule	Proteinase	1hxb.pdb (wild-type complex)	Day 28 ^b (Met36Ile-Ile54Val)	Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)	Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)
Direct interactions						
O1		N Asp29A	2.9	2.9	3.1	3.1
O3		N Asp30A	3.2	2.9	3.1	3.4
O4 (OH)		Oδ1 Asp25A	2.8	2.4	2.8	2.8
		Oδ2 Asp25A	2.8	3.1	3.1	3.0
		Oδ1 Asp25B	3.0	2.8	2.7	2.7
		Oδ2 Asp25B	3.0	2.6	2.7	2.7
N1		O Gly48A	3.3	3.0	3.0	3.0
N2		O Gly48A	3.3	3.2	3.1	3.2
N3 (NH ₂)		Oδ2 Asp30A	3.4	3.1	3.4	3.5
N4		O Gly27A	3.3	3.3	(3.8)	(3.7)
N5		O Gly27B	3.5	(3.7)	(4.0)	(3.7)
Water-mediated interactions						
O2	OA (flap water)		2.9	2.8	2.6	3.1
	OA	N Ile50A	3.2	3.1	3.2	3.2
	OA	N Ile50B	2.9	3.0	3.0	3.2
O5	OA		2.8	2.8	2.6	2.8
	OA	N Ile50A	3.2	3.1	3.2	3.2
	OA	N Ile50B	2.9	3.0	3.0	3.2
O1	OC (water #56)			3.10		
	OC	O Gly27A		2.80		

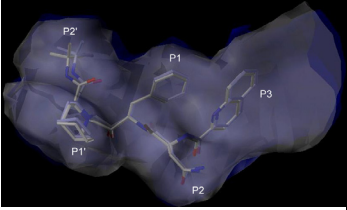
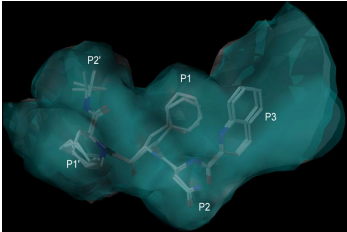
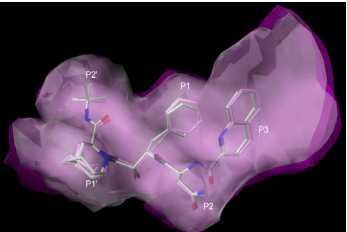
*Interactions greater than 3.5 Å (between brackets) are considered too long for hydrogen-bonding.

When compared to wild-type complex (PDB code 1hxb), one hydrogen bond is lost in variant day 28^b, while two hydrogen bonds are lost in variants day 82 and day 115 (Figure 4.7 and Table 4.7). Saquinavir establishes eleven (wild-type complex), ten (28^b) and nine (day 82 and day 115) direct contacts in addition to one conserved water-mediated hydrogen bond with the active-site residues. The conserved flap water molecule, Wat1 (wild-type complex), Wat19 (day 28^b) or Wat3 (day 82 and day 115), mediates contacts between the carbonyl O2 and O5 atoms of saquinavir on the one hand and the amide groups of Ile50A and Ile50B on the other. Further, quite similar hydrogen-bonding arrangements were observed in all saquinavir containing complexes between the O4 hydroxyl group of saquinavir and the, in total, four carboxylate oxygens of Asp25A and B with distances ranging between 2.4 and 3.1 Å.

Hydrogen-bonding interactions between the O1 atom of saquinavir with the amide of Asp29A, the O3 atom with the amide of Asp30A and, the N1 atom with the carbonyl of Gly48A exist in all structures. In addition, the O1 atom of saquinavir in variant day 28^b makes a hydrogen-bond to a nearby water molecule, Wat56. When compared to wild-type complex, an increase in the distance between the N5 atom of the inhibitor and the carbonyl oxygen of Gly27B is detected in all variant structures, ranging from 3.5 Å (wild-type complex) to 3.7 Å (day 28 and day 115) and 4.0 Å (day 82). Alternatively, the carbonyl oxygen of Gly27B is observed to make hydrogen contact to a water molecule, Wat15 at a distance of 2.65 Å in variant day 28^b, Wat22 (2.58 Å) in day 82 and Wat28 (2.69 Å) in day 115.

The buried protein-surface area (Å²) upon ligand binding, which is indicative of the extensiveness of the ligand-protein interface, was computed using Molcad (Sybyl[®] 7.2). The volume of the binding site cavity and the gap volume between inhibitor and proteinase were calculated using SURFNET (Laskowski, 1995). The gap regions between two molecules can give an indication of how closely molecules pack against one another while the active-site cavity volume indicates whether there is an expansion or reduction of the binding cavity. Table 4.8 summarizes the computed volumes (Å³) and surface areas (Å²).

Table 4.8: Buried protein surface area upon ligand binding (\AA^2), gap volume between inhibitor and proteinase (\AA^3), and volume of the active-site cavity (\AA^3) for variants day 28^b, day 82, day 115 and 1hxb.pdb.

Gap volume (\AA^3) between inhibitor and proteinase		Volume of the active-site cavity (\AA^3)		Superimposition of surface areas		
1hxb	Day 28 ^b	1hxb	Day 28 ^b	1hxb	Day 28 ^b	Superimposition of day 28 ^b (purple) and 1hxb.pdb (white).
301	286	1270	1281	778	782	
1hxb	Day 82	1hxb	Day 82	1hxb	Day 82	Superimposition of day 82 (green) and 1hxb.pdb (white).
301	295	1270	1290	778	799	
1hxb	Day 115	1hxb	Day 115	1hxb	Day 115	Superimposition of day 115 (pink) and 1hxb.pdb (white).
301	289	1270	1285	778	793	

The buried surface areas (\AA^2) of the different complexes show almost no increase or decrease. Also, the volume of the active-site cavity of the variants did not change upon inhibitor binding: day 28^b (1281 \AA), day 82 (1290 \AA), day 115 (1285 \AA), and wild-type complex (1270 \AA). Similarly, the gap volume (between the inhibitor and proteinase) was unaltered.

In order to detect any influence of the amino-acid substitutions on the VDW ligand-proteinase interaction, complementation of shape between the VDW surfaces of ligand and proteinase was investigated. Additionally, possible effects on inhibitor binding were further analysed by computing ligand-proteinase total binding

interaction energies (Table 4.9).

Table 4.9: The VDW interactions and ligand-proteinase estimated total binding energies. The binding mode of chain A of saquinavir in 1hxb.pdb is similar to chain X while chain B in 1hxb.pdb is similar to chain Y in the variant complexes (day 28^b, day 82 and day 115). The cut-off distance used for the VDW contacts is 3.7 Å. The estimated total binding energy was calculated by Dr. Jinzhi Tan, Institute of Biochemistry, University of Lübeck.

	Complexes with saquinavir			
Interactions for two saquinavir binding modes are shown	1hxb.pdb (wild-type complex)	Day 28 ^b (Met36Ile-Ile54Val)	Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)	Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)
Number of VDW interactions between residue 82 and inhibitor	6^a (Cγ1 Val82A-P1', 3 contacts; Cγ2 Val82B-P1, 3 contacts)	2^d (Cγ1 Val82B-P1, 2 contacts)	5^d (Cγ2, Thr82A-P1', 1 contact; Cγ2, Oγ1 Thr82B-P1, 4 contacts)	7^d (Cγ2 Thr82A-P1', 4 contacts; Cγ2, Oγ1 Thr82B-P1, 3 contacts)
	5^b (Cγ1 Val82A-P1, 2 contacts; Cγ1, Cγ2 Val82B-P1', 3 contacts)	4^e (Cγ2, Cβ Val82A-P1, 3 contacts; Cγ1 Val82B-P1', 1 contact)	4^e (Cγ2, Oγ1 Thr82A-P1, 3 contacts; Cγ2 Thr82B-P1', 1 contact)	4^e (Cγ2, Thr82A-P1, 3 contacts; Cγ2 Thr82B-P1', 1 contact)
Total numbers of ligand-proteinase VDW interactions	107 ^a	109 ^d	111 ^d	114 ^d
	108 ^b	111 ^e	109 ^e	111 ^e
^c Estimated total binding energy (kcal/mol)	-20.26	-20.62	-19.91	-19.51

^aVDW contacts between the proteinase and chain A of saquinavir in 1hxb.pdb

^bVDW contacts between the proteinase and chain B of saquinavir in 1hxb.pdb

^c $\Delta G = -RT \ln K$ where ΔG is the change in Gibbs free energy, R is the gas constant (8.314 J mol⁻¹ K⁻¹), T is the temperature (298.2 Kelvin), and K is the equilibrium constant

^dVDW contacts between the variants and chain X of saquinavir

^eVDW contacts between the variants and chain Y of saquinavir

In particular, VDW packing in the wild-type complex is achieved by the Cγ1/Cγ2 side-chain atom of Val82A and Val82B contacting the phenylalanine group at P1' and P1 residues, respectively. According to the Stanford's HIV database, mutations at positions 20, 36 and 71 contribute to saquinavir resistance only when present in combination with mutations at positions 46, 48, 54, 82, 84 and 90. Single mutations at position 82 confer little resistance to saquinavir. The HIV-1 proteinase variants in complex with saquinavir display a similar binding mode as the wild-type complex except for the interaction between residue 82A and the P1' group. Variants day 28^b and day 82 display a different side-chain orientation for Val82A/Thr82A (Figure 4.8). The Val82A side-chain now faces the proteinase instead of the 2-methyl-

decahydro-isoquinoline-3-carbonyl group (P1' position), thus failing to make VDW contacts as observed in wild type and day 115. The different orientation of the Thr82A side-chain in variant day 82 caused the loss of extensive VDW contacts between the C γ 2 atom of Thr82A and the P1' residue of chain X of saquinavir (one contact) when compared to wild type (three contacts) and the Thr82A C γ 2 atom of day 115 (four contacts; see Table 4.9).

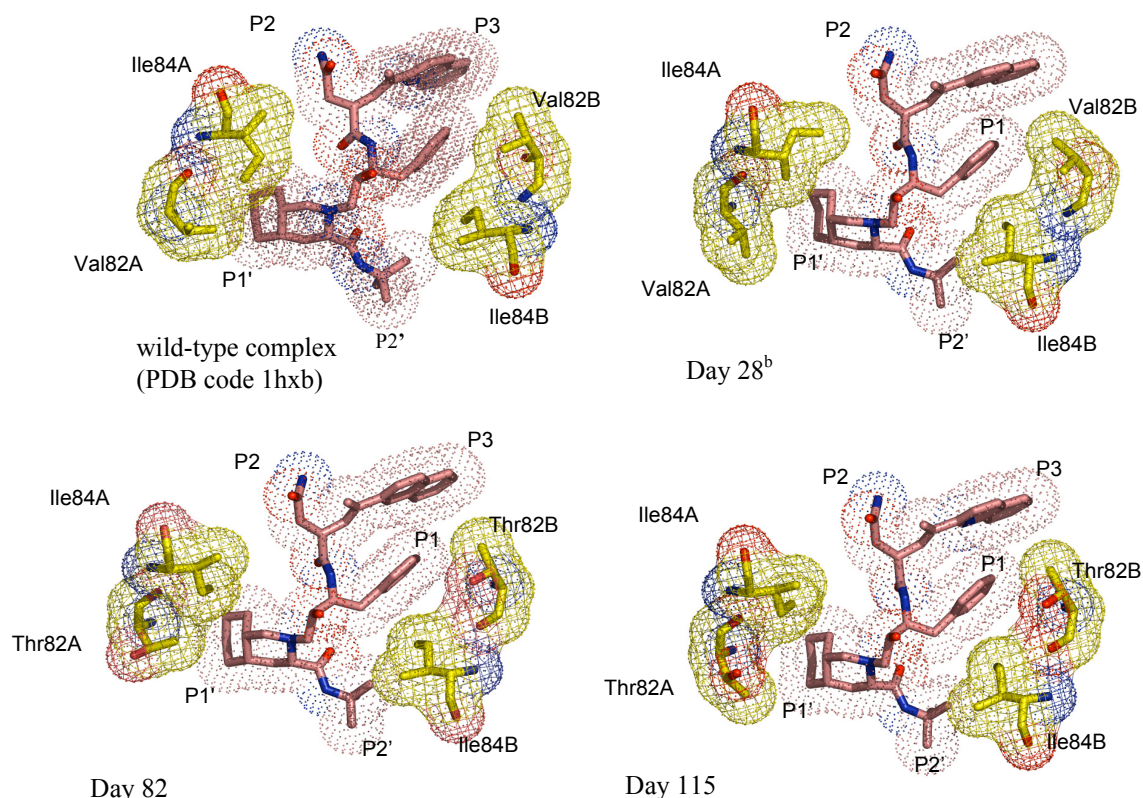
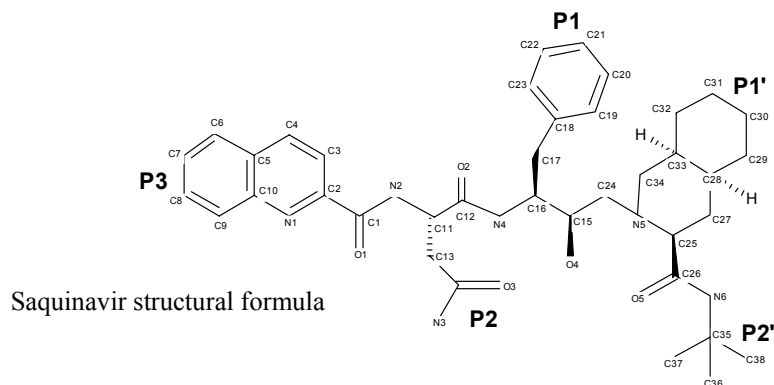


Figure 4.8: Proteinase-saquinavir VDW interactions between residues 82, 84, P1 and P1'. Only one binding mode of saquinavir is depicted. Interactions are between chain A of saquinavir in 1hxb.pdb and chain X of saquinavir in HIV-1 PR variants. Saquinavir is coloured in pink. P1, Phenylalanine; P2, Asparagine; P3, 2-Carbonylquinoline; P1', 2-Methyl-decahydro-isoquinoline-3-carboxylic acid, P2', Tertiary-butylamine.

It is unclear how to explain or predict the saquinavir cross-resistance in terms of the molecular interactions for the variants, since the overall VDW contacts between residues 82A/B and P1/P1' of chain B/Y of saquinavir are almost identical (Table 4.9). Similarly, VDW contacts between residues 82A/B and P1'/P1 of chain A/X of saquinavir are observed, with six contacts in the wild-type complex, five contacts in day 82, seven contacts in day 115 and finally two for day 28^b. Actually, the overall number of ligand-proteinase VDW interactions for variants day 28^b, day 82 and day 115 increases when compared to wild type. Table 4.10 shows the details

of the VDW interactions.

Table 4.10: Details of VDW interactions (up to 3.7 Å) between proteinase saquinavir. Two binding modes of saquinavir, namely chains A and B, are observed in the wild-type complex. The binding mode of chain A of saquinavir in the wild-type complex is similar to chain X while chain B is similar to chain Y in the variant complexes (day 28^b, day 82 and day 115). The cut-off distance used for the VDW contacts is 3.7 Å.



Saquinavir	1hxb.pdb (chain A)	Day 28 ^b (chain X)	Day 82 (chain X)	Day 115 (chain X)	1hxb.pdb (chain B)	Day 28 ^b (chain Y)	Day 82 (chain Y)	Day 115 (chain Y)
P3 (C2, C3, C4, C5, C6, C7, C8, C9, C10, N1)	18	15	12	13	19	14	10	16
P2 (C14, N3, O3)	14	17	13	14	19	15	15	12
P1 (C18, C19, C21, C23, C22, C20)	7	11	9	11	9	10	8	12
P1' (N5, C34, C33, C32, C31, C30, C29, C28, C27, C25)	21	17	23	20	15	23	23	17
P2' (N6, C35, C36, C37, C38)	11	9	12	14	11	17	13	12
Saquinavir backbone								
C26	1		1		4		6	6
C24	4	3	6	6	3	4	5	6
C15	4	4	5	5		1	1	1
C16	1	1		1	1	1	1	1
N4		2	1	1				
C12								
C11	1	2	1	2	2	1	2	1
N2		1	1	1	1	1	1	1
C1		2	1	1			1	1
O5	3	3	2	1	2	2		1
O4	9	10	11	10	9	10	11	10
O2	1	1	2	2	2	1		2
O1	9	9	9	9	8	9	9	9
C17	2	1	2	2	1	1	2	2
C13	1	1	1	1	1	1	1	1
Total	107	109	111	114	108	111	109	111

The average r.m.s. deviation for C α atoms of the proteinase monomers between variants day 82, day 115 and wild type (chain A and B) are about 0.30 Å,

suggesting that the overall structure of these saquinavir-complexes is very similar (Table 4.12). The slight increase in r.m.s. deviations (0.43 and 0.40 Å for chains A and B, respectively) in the proteinase variant day 28^b could be due to the different space groups. The C α atoms of all superimposed catalytic dyad residues show little deviation between the structures with average r.m.s. values of 0.14 to 0.16 Å. In contrast to the flaps of day 82 and day 115 (r.m.s. deviation 0.25 - 0.31 Å), the chain-A flap of day 28^b shows a significant shift when compared to the wild-type complex, with an averaged r.m.s. deviation of 0.46 Å (0.30 Å for chain B). Also, for certain residues of the 80's loop, somewhat larger r.m.s. deviations are observed (see Table 4.11).

Table 4.11: The average r.m.s. deviations (Å) for C α atoms of the monomer (residues 1-99), flap (residues 41-57), catalytic dyad (residues 24-28) and the 80's loop (residues 75-86) in complexes with saquinavir of day 28^b, day 82, and day 115 (compared to wild type).

	Day 28^b (Met36Ile-Ile54Val)		Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)		Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)	
	chain A	chain B	chain A	chain B	chain A	chain B
1hxb.pdb C α atoms of the 99 amino- acid residues	0.43	0.40	0.31	0.30	0.32	0.33
Catalytic dyad (residues 24 - 28)	0.14 (0.23 - Gly27A)	0.15 (0.28 - Gly27B)	0.13 (0.26 - Gly27A)	0.16 (0.31 - Gly27B)	0.15 (0.28 - Gly27A)	0.16 (0.25 - Gly27B)
Flap (residues 41 - 57)	0.46 (0.86 - Trp42A)	0.30 (0.62 - Arg41A)	0.27 (0.53 - Pro44A, Lys45A)	0.25 (0.49 - Arg41B)	0.31 (0.53 - Met46A)	0.26 (0.56 - Arg41B)
80's loop (residues 75 - 86)	0.27 (0.61 - Gly78A)	0.28 (0.50 - Thr80B)	0.34 (0.74 - Pro79A)	0.29 (0.69 - Pro81A)	0.29 (0.70 - Pro79A)	0.25 (0.63 - Pro81B)

Numbers in parentheses indicate the maximum C α deviation when compared to the wild-type complex (PDB code 1hxb).

The distances (Å) between C α atoms of certain amino-acid residues located at the binding cavity (including the catalytic dyad) and the flaps were measured in order to determine whether any changes take place that could account for cross resistance (Table 4.12, Figures 4.9 and 4.10).

Table 4.12: Distances (Å) between residues located at the binding cavity, between the flaps and between flaps and catalytic dyad.

Distance (Å) between residues in the binding cavity	Wild-type complex (PDB code 1hxb)	Day 28 ^b (Met36Ile-Ile54Val)	Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)	Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)
Thr80A-Thr80B	17.67	18.11 (+ 0.44)	17.22 (- 0.45)	17.14 (- 0.53)
Pro81A-Pro81B	19.50	20.15 (+ 0.65)	18.37 (- 1.13)	18.54 (- 0.96)
Val82A-Val82B	20.26	20.39 (+ 0.13)	19.56 (- 0.70)	19.81 (- 0.44)
Thr82A-Thr82B				
Asn83A-Asn83B	21.84	21.81 (- 0.03)	21.25 (- 0.59)	21.44 (- 0.40)
Ile84A-Ile84B	16.17	16.19 (+ 0.02)	15.99 (- 0.18)	16.11 (- 0.06)
Ile85A-Ile85B	17.79	17.62 (- 0.15)	17.57 (- 0.22)	17.65 (- 0.14)
Distance between flaps residues				
Gly52A-Gly51B	4.23	4.41 (+ 0.18)	3.76 (- 0.57)	3.94 (- 0.29)
Gly51A-Gly52B	3.96	4.25 (+ 0.29)	4.17 (+ 0.21)	4.19 (+ 0.23)
Ile50A-Gly49B	5.46	5.30 (- 0.16)	4.97 (- 0.49)	5.13 (- 0.33)
Gly49A-Ile50B	4.85	4.83 (- 0.02)	4.93 (+ 0.08)	4.98 (+ 0.13)
Distance between flaps and catalytic dyad residues				
Ile50A-Asp25B	13.06	12.49 (- 0.77)	12.93 (- 0.13)	12.73 (- 0.33)
Ile50B-Asp25A	12.92	12.70 (- 0.22)	12.92 (0.0)	12.91 (- 0.01)

For the values in parentheses, (-) indicates how much the distance has decreased and (+) indicates the distance has increased in comparison to the wild-type complex (PDB code 1hxb).

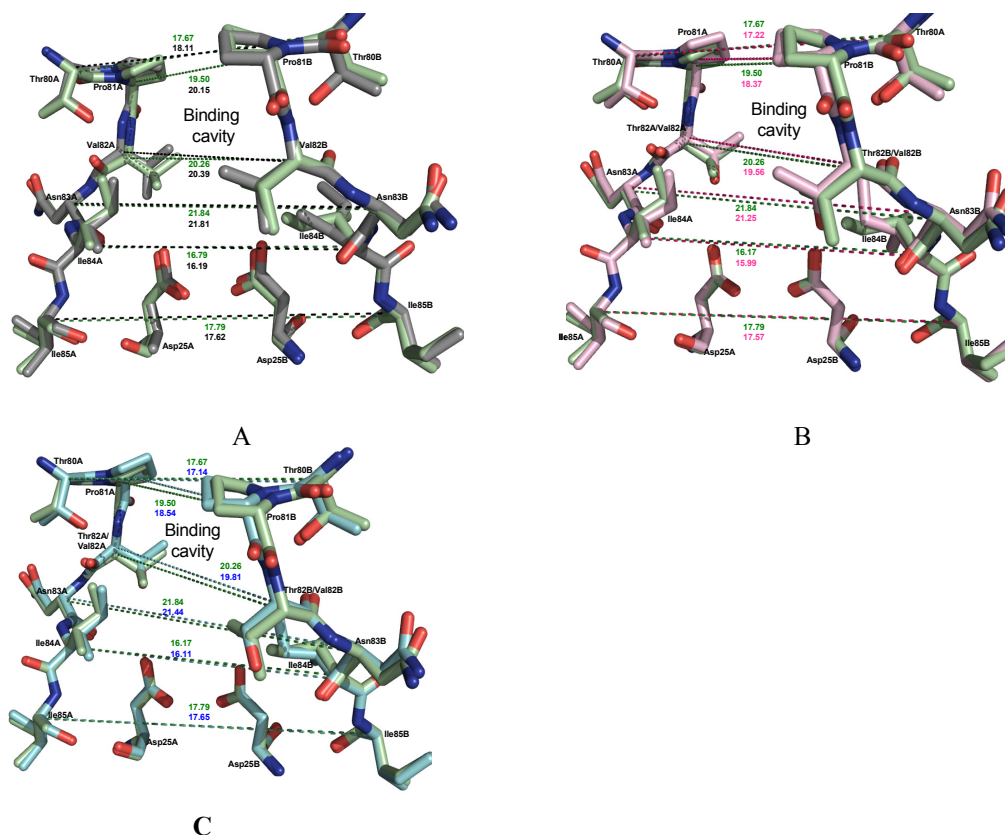


Figure 4.9: Schematic showing the distances (Å) between residues located at the binding cavity. The wild-type complex (PDB code 1hxb) is coloured in green (as is the distance). Panel A, variant day 28^b coloured in black, B, day 82 is coloured in pink, C, day 115 is coloured in cyan.

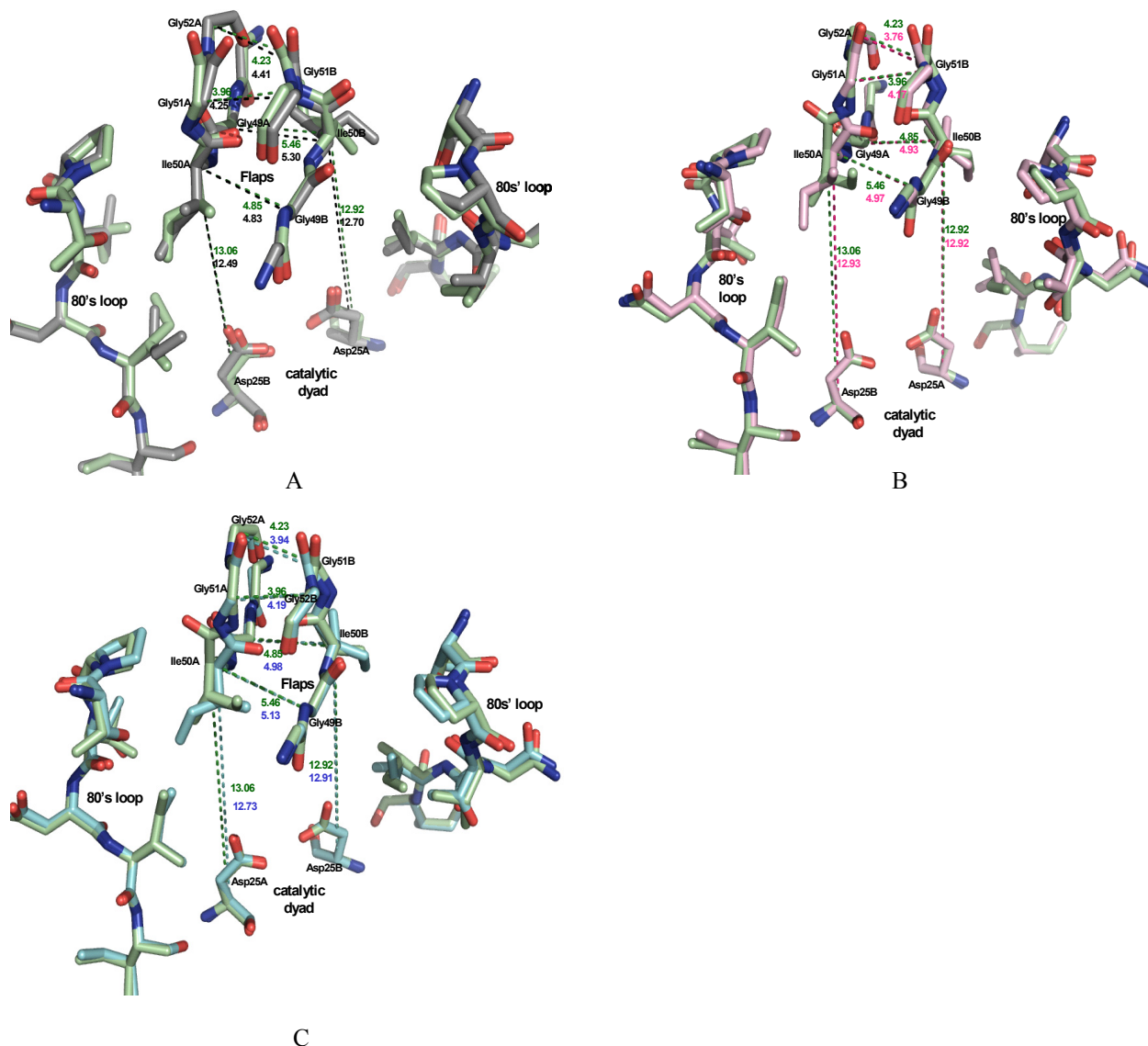


Figure 4.10: Schematic showing distances (Å) between residues located at the flaps and catalytic dyad. The wild-type complex (PDB code 1hxb) is coloured in green (as is the distance). Panel A, variant day 28^b coloured in black, B, day 82 is coloured in pink, C, day 115 is coloured in cyan.

Variant day 82 and day 115 show slight decreases in the $\text{Ca} - \text{Ca}$ distances of Thr80A-Thr80B, Pro81A-B, Val82A-B, Asn83A-B when compared to wild type and day 28^b. The $\text{Ca} - \text{Ca}$ distance between residues Ile50, located at the tip of the flap, and Asp25 (of the catalytic dyad) is slightly decreased in all variant complexes with saquinavir when compared to wild type. These small changes have, however, no definite effect on the volume of the active-site cavity (less than 2% increase for the variants when compared to wild type; see Table 4.8).

4.5.3 Inter- and intra-monomer interactions in variants day 28^b, day 82, and day 115 in complex with saquinavir

Only Thr82 (variant day 82 and 115) may directly influence inhibitor binding. In order to determine whether substitutions outside the active site (Met36Ile, Ile54Val or Ala71Val) have an effect on residues that line the binding cavity (for example Arg8, Val32, Gly49 and Ile84), several inter/intra-monomer interactions were investigated.

One of the well-known features of wild-type HIV-1 proteinase-inhibitor complexes is the inter-monomer salt-bridge between Arg8 and Asp29. The Arg8A/B and Asp29A/B residues in all three variants are clearly visible in their respective 2F_o-F_c electron-density map (resolutions ranging from 1.76 to 1.95 Å). Except for the salt-bridge between Arg8A and Asp29B in variant day 82, this salt-bridge is conserved throughout (Table 4.13, Figure 4.15).

Table 4.13: Inter- and intra-monomer salt-bridges (cut-off distance of 3.50 Å).

Salt-bridge interactions (cut-off distance of 3.50 Å)	1hxb.pdb	Day 28 ^b (Met36Ile-Ile54Val)	Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)	Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)
Nη2 Arg8B-Oδ2 Asp29A	3.35	2.61	2.69	2.67
Nη2 Arg87A-Oδ1 Asp29A	3.14	2.79	2.87	2.95
Nη2 Arg8A-Oδ2 Asp29B	3.42	2.58	-	2.75
Nη2 Arg87B-Oδ1 Asp29B	3.18	2.99	3.15	2.94
Nη2 Arg57B-Oε2 Glu35B	3.30	-	-	-
Nη2 Arg57A-Oε2 Glu35A	3.30	-	-	-

Distances are in Å.

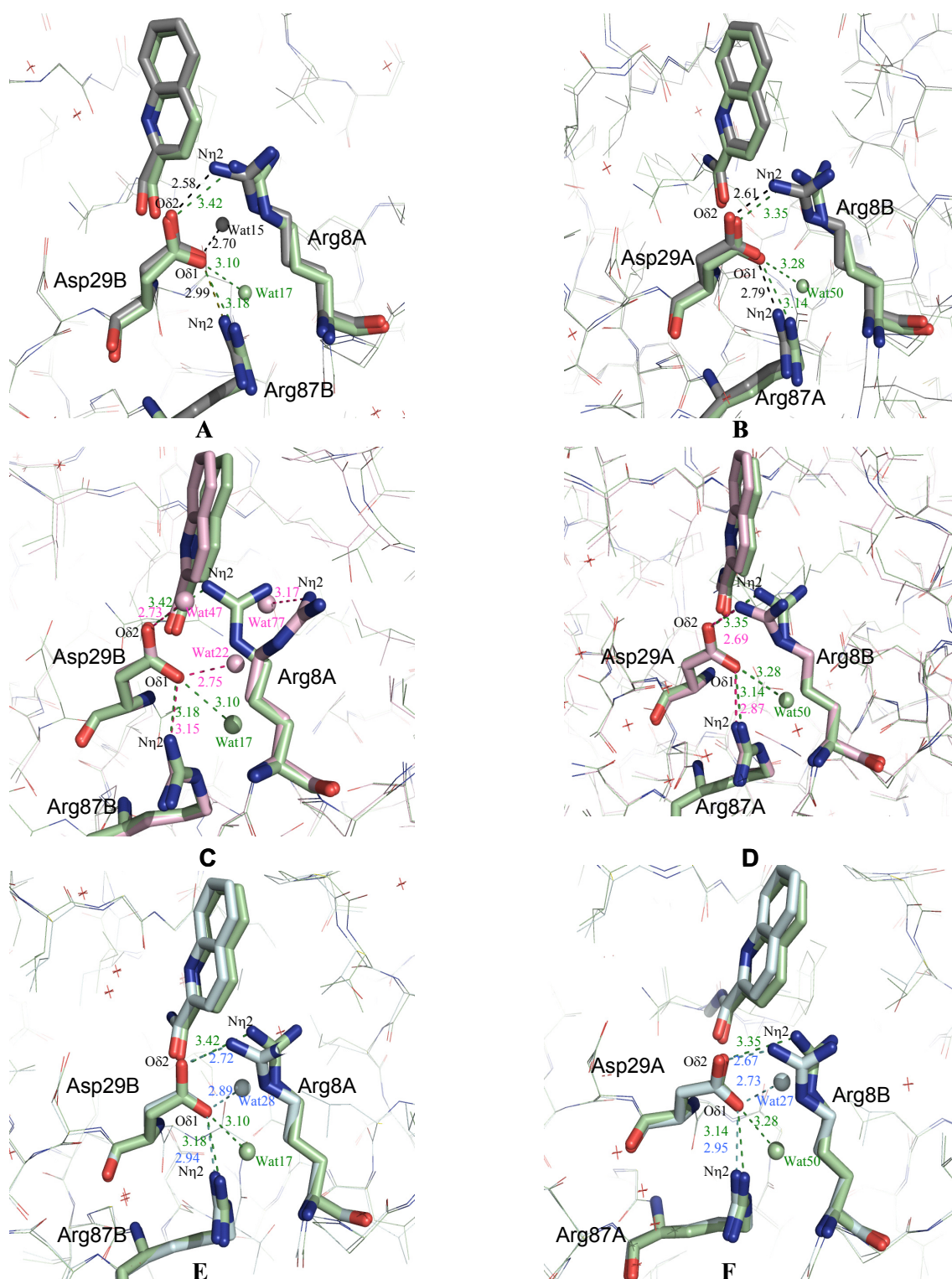
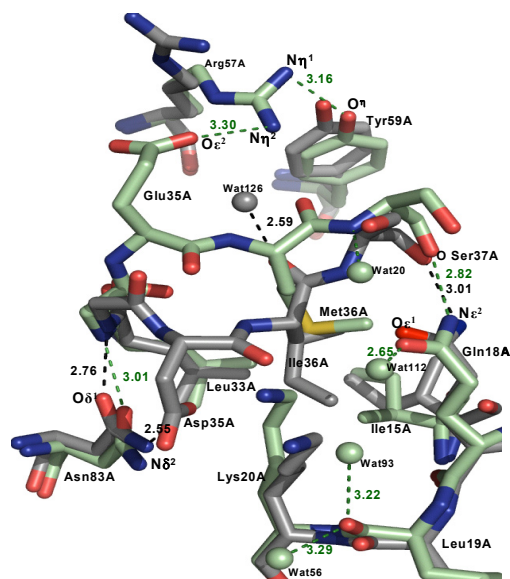


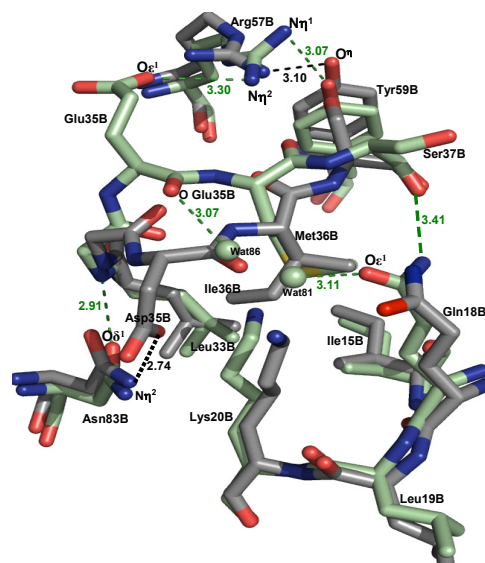
Figure 4.11: The inter-monomer salt-bridge interaction between Arg8A/B and Asp29B/A in day 28^b (coloured in black, panels A and B), day 82 (coloured in pink, C and D) and day 115 is (coloured in cyan, E and F), superimposed onto the wild-type complex (coloured in green, PDB code 1hxb).

The O δ 1 atom of Asp29B in the wild-type complex forms an intra-monomer salt-bridge interaction with the N η 2 atom of Arg87B (distance of 3.18 Å) and a hydrogen bond with a water molecule, Wat17, at a distance of 3.10 Å. The O δ 1 atom of Asp29A makes an intra-molecular salt-bridge with N η 2 atom of Arg87A (3.14 Å) and a hydrogen bond (3.28 Å) with water molecule Wat50 (Figure 4.11, B). The N η 2 atom of the Arg8A side-chain of variant day 82 makes a hydrogen bond with a water molecule, Wat77 (3.17 Å) instead of forming a salt-bridge with the O δ 2 atom of Asp29B (Figure 4.11, C). Besides, the O δ 2 atom of Asp29B makes a hydrogen bond with a water molecule, Wat47 (2.73 Å), suggesting that the water molecule structurally compensated for the removal of the salt-bridge. The O δ 1 atom of Asp29B of variant day 28^b, day 82 and day 115 makes an intra-subunit salt-bridge with N η 2 atom of Arg87B and form a hydrogen bond with a water molecule, Wat15 (2.70 Å), Wat22 (2.75 Å), Wat28 (2.89 Å), respectively. Further, in contrast to day 115 and wild type (PDB code 1hxb), no water molecules are hydrogen bonded to the O δ 1 atom of Asp29A of variant day 28^b and day 82.

The intra-molecular salt-bridge between Arg57A-Asp35A is lost in variant day 28^b, day 82 and day 115 when compared to wild type (Figure 4.12): The N η 1 atom of the Arg57A of variant day 28^b, day 82 and day 115 makes a hydrogen bond with the OH atom of Tyr59 instead of with the O ϵ 2 atom of the Asp35. Actually, in the wild-type complex (PDB code 1hxb) there is an Asp35Glu substitution. The loss of this intra-molecular salt-bridge is a result of the Met to Ile amino-acid substitution at neighboring residue 36 in variant day 28^b, day 82 and day 115 (Figure 4.12). The backbone of residues Asp35, Ile36 and Ser37 is pushed towards the backbone of residues Glu21 and Lys20/Arg20 in order to accommodate the Ile36 side-chain in day 28^b, day 82 and day 115. Leu33, Ile15 and the main-chain of Lys/Arg20 form a hydrophobic pocket, into which the side-chain of Ile36 fits. Asp35 now faces residue Lys20/Arg20 instead of Arg57. In addition, the O δ 2 atom of Asp35 forms a hydrogen bond with the N δ 2 atom of Asn83 of variant day 28^b, day 82 and day 115.

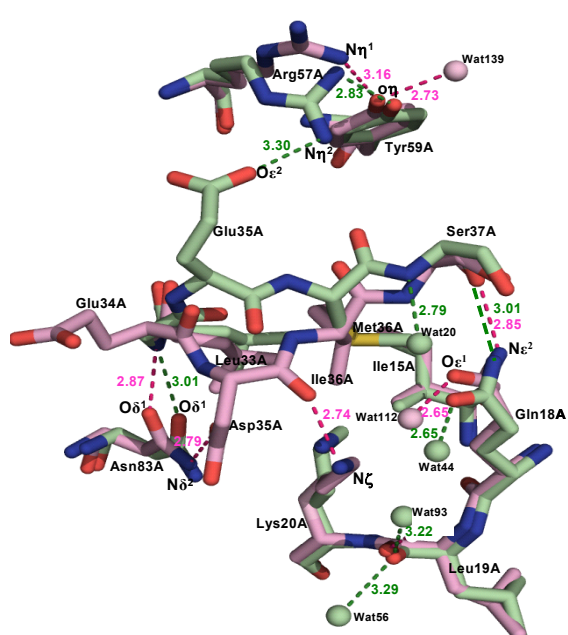


A

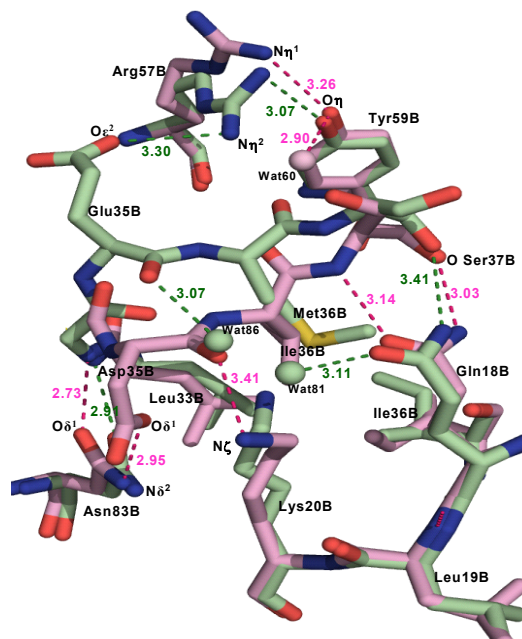


B

Interactions at the 10's (residues 15-25) and 30's (residues 30-40) loops in variant day 28^b (grey), superimposed onto wild type (green), 1hxb.pdb. A; chain A of the proteinase, B; chain B of the proteinase. Spheres indicate water molecules.

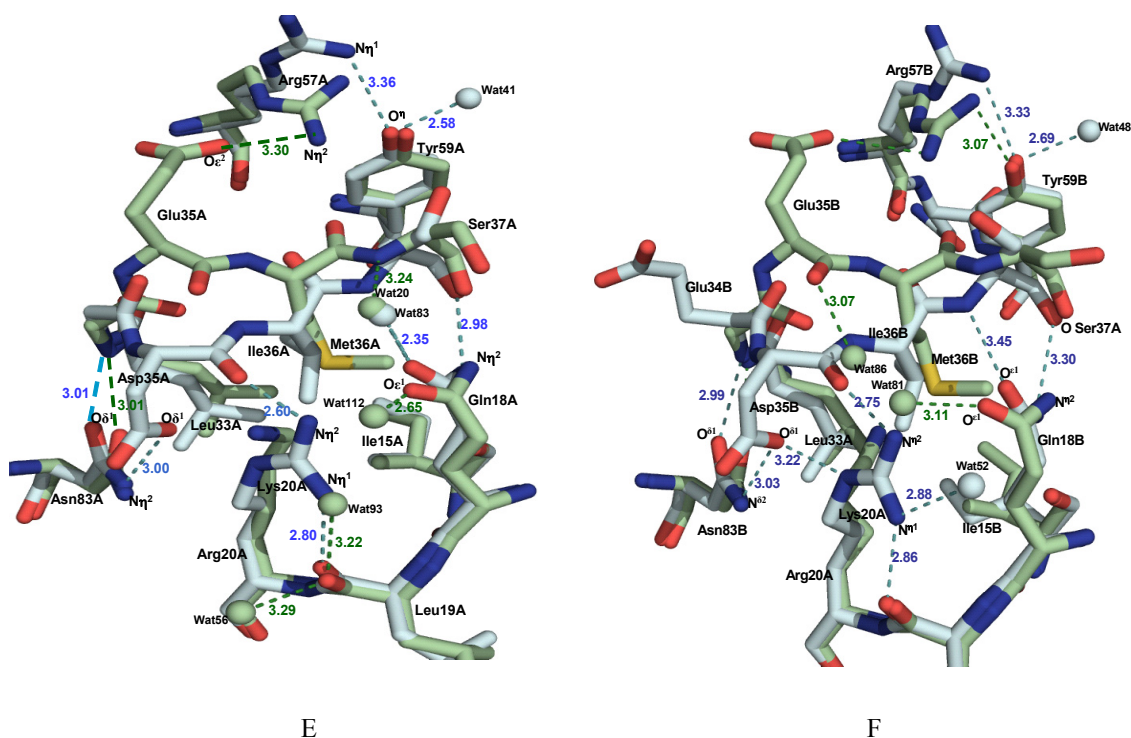


C



D

Interactions at the 10's (residues 15-25) and 30's (residues 30-40) loops in variant day 82 (pink) complex with saquinavir, superimposed onto the wild type (green), 1hxb.pdb. C; chain A of the proteinase, D; chain B of the proteinase.



Interactions at the 10's (residues 15-25) and 30's (residues 30-40) loops in variant day 115 (cyan) complex with saquinavir, superimposed onto the wild type (green), 1hxb.pdb. E; chain A of the proteinase, F; chain B of the proteinase.

Figure 4.12: The hydrogen bond interactions (Å) at the 10's (residues 15-25) and 30's (residues 30-40) loops in variant day 28^b (Met36Ile-Ile54Val), day 82 (Met36Ile-Ile54Val-Ala71Val) and day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr) compared to the wild-type complex, 1hxb.pdb.

The Nε2 atom of Gln18 of variant day 82, day 115 and wild-type complex makes a hydrogen bond with the carbonyl O atom of Ser37 except in chain B of variant day 28^b (Figure 4.12). The Oε1 atom of Gln18A also interacts with a water molecule, Wat44 (2.65 Å), Wat83 (2.35 Å) and Wat112 (2.65 Å) in variant day 82, day 115 and the wild-type complex, 1hxb.pdb, respectively (Figure 4.12, C and D). On the other hand, in chain B of variant day 82 and day 115, the Oε1 atom makes a hydrogen bond to the N atom of Ser37. The Oε1 atom of Gln18 of 1hxb.pdb interacts with a water molecule, Wat44 for chain A (at 2.65 Å) and Wat81, at a distance of 3.11 Å (chain B). The Oε1 atom of Gln18B of variant day 28^b makes no interactions, neither to a water molecule nor to the N atom of Ser37. The residue Lys20A/B of the wild type, 1hxb.pdb, and variant day 28^b makes no hydrogen contact to neighboring residues or water molecule (Figure 4.12, A and B). In variant day 82, the Nζ atom of Lys20A/B make a hydrogen bond to the carbonyl O atom of the Asp35A with a

distance of 2.74 and 3.41 Å, respectively. The Nη1 atom of Arg20 in the variant day 115 makes a hydrogen bond to the carbonyl O atom of Leu19 whereas the Nη2 atom of Arg20 makes a hydrogen bond to the carbonyl O atom of Asp35 (Figure 4.12, E and F).

One significant difference in variant day 28^b is the flip of the Phe53B side-chain towards the 2-carbonylquinoline ring (P3 position) of saquinavir instead of being exposed at the proteinase surface as in 1hxb.pdb, day 82 and day 115 (Figure 4.13). The reason for the different orientation of Phe53B in day 28^b is not so clear; the distance between the Phe53B and the quinoline ring of the saquinavir is more than 4.5 Å. One possible reason for the different side-chain orientations could be that variant day 28 crystallized in the P2₁2₁2₁ space group while 1hxb.pdb, day 82 and day 115 in complex with saquinavir crystallized in space group P6₁.

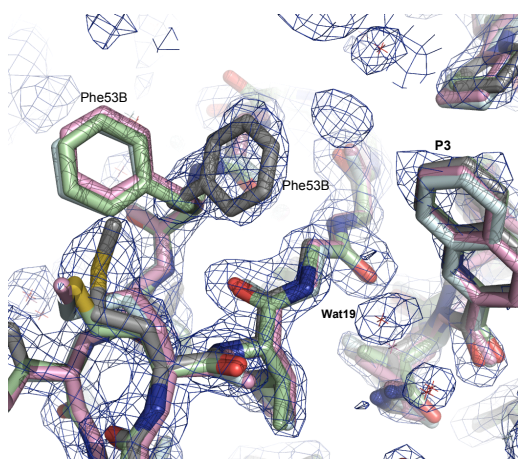


Figure 4.13: The 2F_o-F_c electron-density map of the variant day 28^b in complex with saquinavir, contoured at 1.0 σ above the mean, showing the Phe53B side-chain facing towards the P3 of saquinavir. Variant day 28^b (black), day 82 (pink), day 115 (cyan) and the wild type, 1hxb.pdb (green).

4.5.4 Complexes with ritonavir of day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr) and day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)

The variant day 82 and day 115 both crystallized in space group $P6_1$. Like the saquinavir-complexes, the crystallographic asymmetric unit contains a dimer of HIV-1 proteinase. In each structure, the inhibitor is bound in two orientations (Figure 4.14), which has also been observed in 1rl8.pdb, a complex with ritonavir (Rezacova *et al.*, 2005).

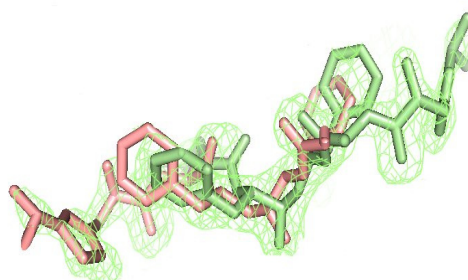


Figure 4.14: The $2F_o - F_c$ electron-density map contoured at 1.0σ above the mean for ritonavir with two binding modes, namely RIT 1X and RIT 2Y of variants day 82 (similar in day 115).

The overall quality of electron density for the ritonavir complexes is similar to those of the saquinavir-complexes, displaying few disordered side-chains (Table 4.14).

Table 4.14: Disordered amino-acid side-chains in variant complexes with ritonavir.

Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)		Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)	
A	B	A	B
Arg41	Arg41	Arg41	Arg41
Lys43	Lys43	Lys43	Lys43
Gln61		Lys55	Lys55
Arg41	Arg41	Gln61	

For comparison studies, both wild-type HIV-1 proteinase (PDB code 1hwx; Kempf *et al.*, 1995) and a variant with nine mutations (PDB code 1sh9; Clemente *et al.*, 2004) in complex with ritonavir have been chosen. Table 4.15 shows a list amino-acid substitutions.

Table 4.15: Variants and residue changes

Variants	Mutations	Space group
1hwx.pdb P2 ₁ 2 ₁ 2	Ser37Asn (a mutation resulting from sequence polymorphism)	
1sh9.pdb	Lys20Arg, Val32Ile, Leu33Phe, Met36Ile, Ile54Val, Ala71Val, Val82Ala, Ile84Val, Leu90Met	P6 ₁
day 82	Met36Ile-Ile54Val-Ala71Val-Val82Thr	P6 ₁
day 115	Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr	P6 ₁

Superimposition of all the ritonavir complexes mentioned above indicates that the inhibitor conformations are almost similar, as shown in Figure 4.15, except for small displacements at P1' and P2'. In 1sh9.pdb, the P2' position (Figure 4.15, pink) deviates from wild type (1hwx.pdb) and the variants day 82 and day 115, which could be the result of the V32I substitution.

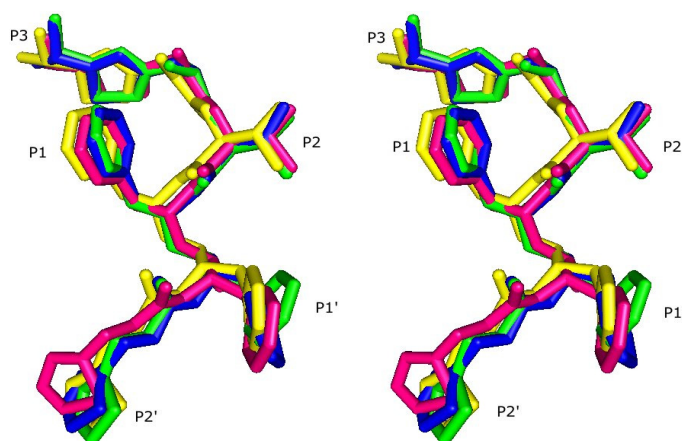


Figure 4.15: Superimposition of ritonavir in 1hwx.pdb (yellow), 1sh9.pdb (pink), day 82 (green), and day 115 (blue). The ritonavir side chains are labelled P3, P2, P1, P1', P2'.

The side chains of residues Arg41, Lys55 and Gln61 in variant day 82 and day 115 display weak electron density as evident from a higher overall residue *B*-factor (Figure 4.16 and Table 4.16). Obviously, some differences originate from different crystal packing, i.e., different space groups (compare 1hwx with other structures in Figure 4.16 and Table 4.16).

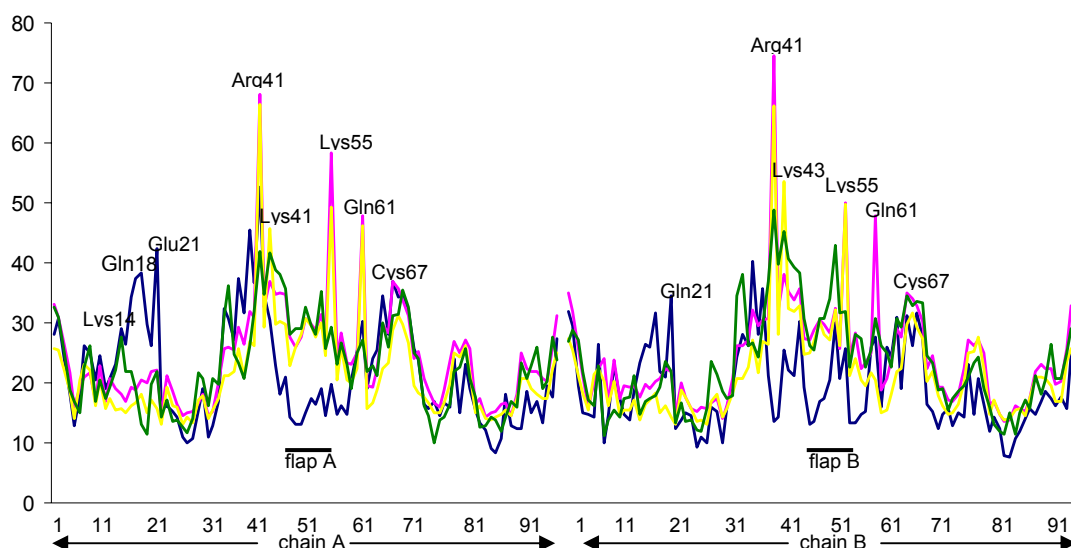


Figure 4.16: Mean *B*-values per residue of wild type (PDB code 1hwx.pdb, blue), variant 1sh9.pdb (green), day 82 (pink) and day 115 (yellow) in complex with ritonavir. The residues of the two monomers are numbered 1-99.

Table 4.16: Comparison of the mean *B*-values (\AA^2) per residue at the flaps of the proteinase for wild type, 1sh9.pdb, day 82 and day 115 complex in complex with ritonavir.

	1hwx.pdb		1sh9.pdb		Day 82		Day 115	
	chain A	chain B	chain A	chain B	chain A	chain B	chain A	chain B
Lys45	18.1	21.2	38.1	39.3	34.9	33.7	30.2	31.9
Met46	20.9	30.3	35.7	38.4	34.7	35.7	29.8	32.9
Ile47	14.3	19.2	28.0	30.8	27.6	27.5	22.9	24.7
Gly48	13.2	13.0	29.1	26.3	25.9	27.1	25.9	24.9
Gly49	13.1	13.6	29.1	25.4	27.7	28.7	28.1	26.8
Ile50	15.1	17.0	32.6	30.7	30.6	30.5	30.8	30.3
Gly51	17.3	17.5	29.5	30.8	29.8	29.7	29.5	28.2
Gly52	16.5	20.4	28.0	34.0	28.8	28.4	27.7	27.2
Phe53	19.0	30.1	35.3	42.9	31.4	32.5	30.0	32.1
Ile54	14.6	20.7	25.7	31.7	26.2	25.6	24.6	26.2
Lys55	19.8	25.7	29.4	31.8	58.4	50.1	49.2	49.7

4.5.5 Proteinase-ritonavir interactions

The interactions of ritonavir with the active-site residues of variant day 82 and day 115 were compared to the wild-type complex, 1hwx.pdb. Analysis of the hydrogen-bonding pattern of the variant complexes shows similar interactions as seen in the wild type, except for the loss of one water-mediated hydrogen bond in the variant complexes. In the wild-type complex, 1hwx.pdb, nine direct hydrogen bonds and two water mediated-interactions are present while in variant day 82 and day 115

nine hydrogen bonds and only one water-mediated interaction are observed (Table 4.17 and Figure 4.17).

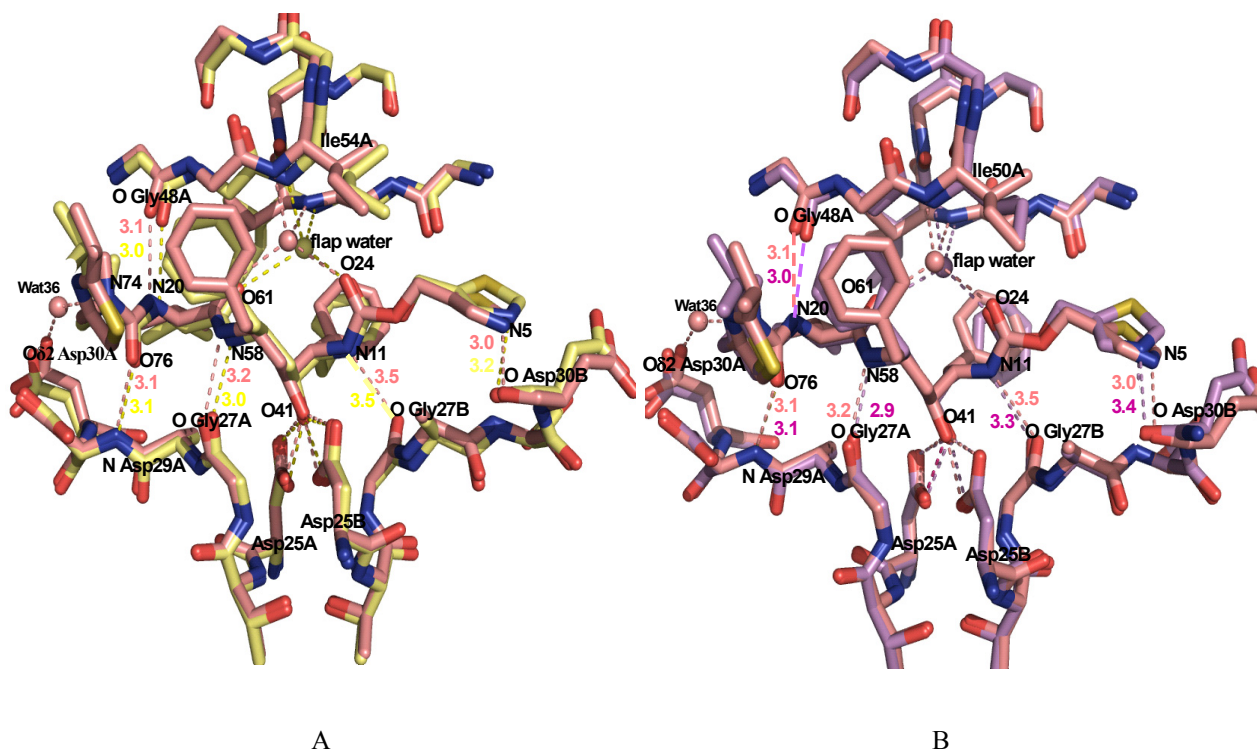
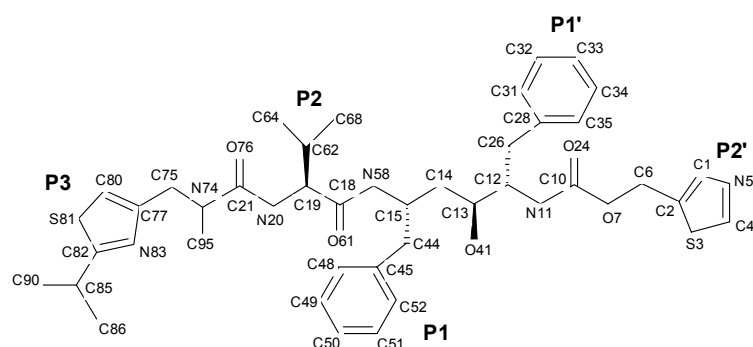


Figure 4.17: Hydrogen-bonding interactions in complexes with ritonavir. A. Variant day 82 coloured in yellow superimposed onto the wild-type complex, 1hxx.pdb, coloured in pink; B. Variant day 115 coloured in purple, superimposed onto wild-type, 1hxx.pdb.

Table 4.17: Hydrogen-bonding interactions between the inhibitor and residues of variants day 82 and day 115 (distances in Å).

Ritonavir structural formula

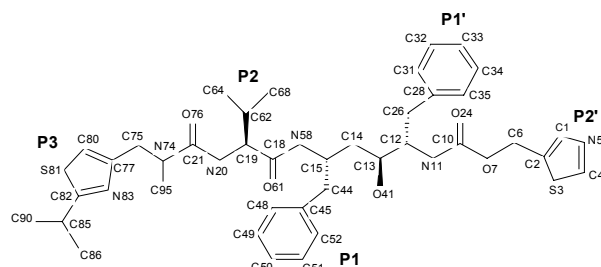


Ritonavir	Water	Proteinase	1hxb.pdb (wild-type complex)	Day 82 (Met36Ile-Ile54Val- Ala71Val- Val82Thr)	Day 115 (Lys20Arg-Met36Ile- Ile54Val-Ala71Val- Val82Thr)
Direct interactions					
Rit O41		Oδ1 Asp25A	2.7	2.5	2.6
		Oδ2 Asp25A	3.1	2.7	2.7
		Oδ1 Asp25B	2.7	2.9	2.9
		Oδ2 Asp25B	2.7	3.1	3.1
Rit N11		O Gly 27B	3.5	3.5	3.3
Rit N5		N Asp 30B	3.0	3.2	3.4
Rit N58		O Gly 27A	3.2	3.0	2.9
Rit O76		N Asp29A	3.1	3.1	3.1
Rit N20		O Gly 48A	3.1	3.0	3.0
Water-mediated interactions					
Rit N74	OB		3.4 (Wat36)	-	-
		Oδ2 Asp30A	2.7	-	-
Rit O61	OC (flap water)		2.7	3.1	3.0
		N Ile50A	2.9	3.0	3.0
Rit O24	OC (flap water)		2.7	2.9	2.5
		N Ile50B	3.1	3.0	3.0

Despite similar hydrogen-bonding interactions, significant increases in the hydrophobic interactions are observed especially for variant day 115 when compared to wild type, 1hxb.pdb (see Table 4.18).

Table 4.18: Hydrophobic interactions between the proteinase and ritonavir. The total number of hydrophobic contacts is shown in parentheses (cut-off value of 3.7 Å). In the wild type, only one orientation of ritonavir, namely RIT 301, is reported, while two orientations for ritonavir, namely RIT 1X and RIT 2Y, are observed for variant day 82 and day 115. The RIT 301 binding mode is similar to RIT 2Y binding mode.

Ritonavir structural formula



Ritonavir	1hxx.pdb (wild-type)	Day 82 (Met36Ile-Ile54Val- Ala71Val- Val82Thr)	Day 115 (Lys20Arg- Met36Ile-Ile54Val- Ala71Val- Val82Thr)	Day 82 (Met36Ile-Ile54Val- Ala71Val- Val82Thr)	Day 115 (Lys20Arg- Met36Ile-Ile54Val- Ala71Val- Val82Thr)
	RIT 301	RIT 2Y	RIT 2Y	RIT 1X	RIT 1X
P3, 2-isopropyl-4- thiozoyl	Arg8B (3) 3 contacts	Arg8B (2) Thr82B (2) 4 contacts	Arg8B (3) Thr82B (3) 6 contacts	Arg8A (3) Thr82A (3) 6 contacts	Arg8A (3) Pro81A (2) Thr82A (4) 9 contacts
P2, Valine	Ile84A (1) 1 contact	Ala28A (1) Ile50A (1) 2 contacts	Ile50B (1) Ile47A (2) Val32A (1) Ala28A (1) 5 contacts	Ile84B (1) Ala28B (1) Ile50A (1) 3 contacts	Ile84B (1) Val32B (1) Ala28B (1) Ile50A (1) 4 contacts
P1, Phenylalanine	Ile50B (1) Gly49B (2) Ile84A (3) Val82A (2) Pro81A (2) 10 contacts	Ile50B (3) Ile84A (4) Thr82A (2) Pro81A (2) 11 contacts	Gly49B (3) Thr82A (2) Ile84A (2) Leu23A (1) 8 contacts	Pro81B (2) Ile84B (3) Ile50A (2) 7 contacts	Ile84B (5) Ile50A (2) Thr82B (2) 9 contacts
P1', Phenylalanine	Val82B (4) Pro81B (3) Ile50B (2) Gly49A (5) Ile50A (1) 15 contacts	Pro81B (5) Gly49A (5) 10 contacts	Ile84B (1) Thr82B (1) Pro81B (5) Gly49A (5) 12 contacts	Gly48B (1) Gly49B (6) Ile84A (1) Thr82A (4) Pro81A (5) 17 contacts	Gly49B (3) Thr82A (4) Pro81A (5) 12 contacts
P2', Thiozoyl	Ala28B (1) Asp 30B (1) Val32B (1) 3 contacts	Asp30B (3) Val32B (1) 4 contacts	Ile47B (1) Val32B (1) Asp30B (3) Ala28B (1) 6 contacts	Ala28A (1) Asp30A (2) Val32A (1) 4 contacts	Ile84A (1) Asp30A (3) Val32A (2) Ala28A (2) Ile50B (1) 9 contacts
RIT C13		Asp25A (1)	Asp25A (1)		Asp25B (1)
RIT C14		Asp25B (1)	Gly27A (1) Asp25B (1)		
RIT C75	Asp29A (1)	Asp29A (1)	Asp29A (1)	Asp29B (1)	Asp29B (1)
RIT C6				Ile50B (1)	
RIT C95		Gly48A (1)	Gly48A (1)	Gly48B (1)	Gly48B (1)
Total	33	35	42	41	48

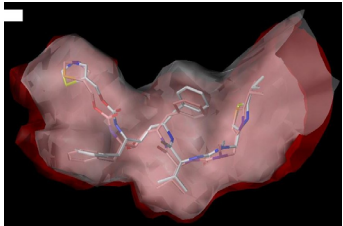
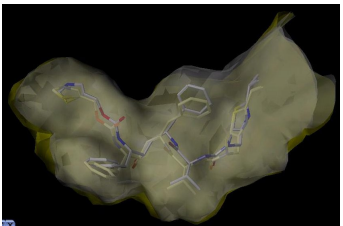
In total, 33 contacts (RIT 301) in 1hxb.pdb, 35 contacts (RIT 2Y) in day 82 and 42 contacts (RIT 2Y) in day 115 are observed (Table 4.18). At the P1' residue of ritonavir, a reduced number of hydrophobic interactions are seen in day 82 (ten contacts) and day 115 (twelve contacts) when compared to the wild type, 1hxb.pdb (fifteen contacts). An increase in the number of hydrophobic interactions is observed at (see Table 4.18):

1. residue P2 (valine), closer contacts are made with Ala28A, Val32A, Ile47A and Ile50B (one contact only between Ile84A and P2 in 1hxb.pdb);
2. residue P3 (2-isopropyl-4-thiozoyl), closer contacts are made with Thr82A and Arg8B (one contact only between Arg8B and P2 in 1hxb.pdb);
3. residue P2' (thiozoyl), the three contacts in 1hxb.pdb are expanded to four and six contacts in variant day 82 and day 115, respectively.

Thr82B in variant day 82 and day 115 moves somewhat away from P1', *i.e.*, makes closer contacts with residue P3 (2-isopropyl-4-thiozoyl) when compared to the wild type, 1hxb.pdb (Table 4.18). In both day 82 and day 115, closer hydrophobic interactions between the ritonavir backbone and the proteinase are observed when compared to wild type. Apparently, the active-site cavity in day 82 and day 115 is slightly decreased, which could result in unfavourable hydrophobic interactions between the inhibitor and the proteinase. The unfavourable hydrophobic interactions could explain the high K_i value, 1870 and 1820 nM, for day 82 and day 115, respectively.

No significant increase or decrease of the active-site cavity volume or gap volume is observed when comparing day 82 or day 115 with the wild-type complex, 1hxb.pdb (Table 4.19). The number of van der Waals interactions between proteinase and ritonavir increases slightly in variant day 82 and day 115 when compared to the wild type (Table 4.20). Also, the calculated estimated binding energy of variant day 82 (-18.92 kcal/mol) and day 115 (-18.34 kcal/mol) deviate only a bit when compared to wild type (-19.84 kcal/mol) and as such, these values do not correlate with the K_i values of 1870 and 1820 nM for day 82 and day 115, respectively.

Table 4.19: The buried surface area (\AA^2) and active-site cavity volume (\AA^3).

Gap volume (\AA^3) between the ligand and the proteinase		Volume of the active-site cavity (\AA^3)		Surface area (\AA^2) of the ligand-proteinase interface		Superimposition of the surface area of ritonavir-proteinase between the variants and 1hwx.pdb
1hwx.pdb (wild-type)	Day 82	1hwx.pdb	Day 82	1hwx.pdb	Day 82	surface area of ritonavir-proteinase of day 82 (red) superimposed onto the surface area of the ritonavir-proteinase of 1hwx.pdb (white)
205	190	1315	1290	860	840	
1hwx.pdb (wild-type)	Day 115	1hwx.pdb	Day 115	1hwx.pdb	Day 115	surface area of ritonavir-proteinase of day 115 (yellow) superimposed onto the surface area of ritonavir-proteinase of 1hwx.pdb (white)
205	196	1315	1298	860	845	

According to the Stanford database (<http://hivdb.stanford.edu/cgi-bin/PIResiNote.cgi>), the amino-acid substitution Val82Ala/Thr/Phe confers high-level resistance to indinavir, ritonavir, and lopinavir, intermediate resistance to nelfinavir, amprenavir, and atazanavir, and low-level resistance to saquinavir. The amino-acid substitution Ile84Val/Ala/Cys confers resistance to all seven drugs. Most probably, a divergence of complementation of shape between the VDW surface of residues 82, 84 and the VDW surface of the ligand could account for the ritonavir resistance.

Table 4.20: VDW interactions (up to a 3.7 Å) and ligand-proteinase estimated binding energy. Only one orientation for ritonavir is reported in the wild-type complex, 1hwx.pdb, namely RIT 301. Two orientations are observed in the variants day 82 and day 115, namely RIT 2X and RIT 2Y. The RIT 301 binding mode is similar to RIT 2Y. VDW energy for inhibitor and estimated binding energy were calculated by Dr. Jinzhi Tan, Institute of Biochemistry, University of Lübeck.

	1hwx.pdb	Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)	Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)	Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)	Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)
	RIT 301	RIT 2Y (chain Y of the ritonavir)	RIT 2Y (chain Y of the ritonavir)	RIT 1X (chain X of the ritonavir)	RIT 1X (chain X of the ritonavir)
Number of VDW interactions (atom) between residue 82 and the ligand. (up to a 3.7 Å)	7 (Cγ1 Val82A - P1', 2 contacts; Cγ1, Cγ2 Val82B - P1, P3, 5 contacts)	7 (Cγ2 Thr82A - P1', 2 contacts; Oγ1, Cγ2 Thr82B - P1, P3, 5 contacts)	8 (Cγ2 Thr82A - P1', 2 contacts; Oγ1, Cγ2 Thr82B - P1, P3, 6 contacts)	8 (Oγ1, Cγ2 Thr82A - P1, P3, 6 contacts; Oγ1, Thr82B - P1' 2 contacts)	9 (Oγ1, Cγ2 Thr82A - P1, P3, 7 contacts; Cγ2 Thr82B-P1' 2 contacts)
Number of VDW ligand-proteinase interactions (atoms). (up to a 3.7 Å)	109	114	120	112	115
^a VDW energy for inhibitor conformation (kcal/mol) – SYBYL 7.2	55.55	79.05	166.53	n.d	n.d
^b Estimated binding energy (kcal/mol) – (AutoDock, Version 3.0.5)	-19.84	-18.92	-18.34	n.d	n.d

^aVDW energy for inhibitor conformation

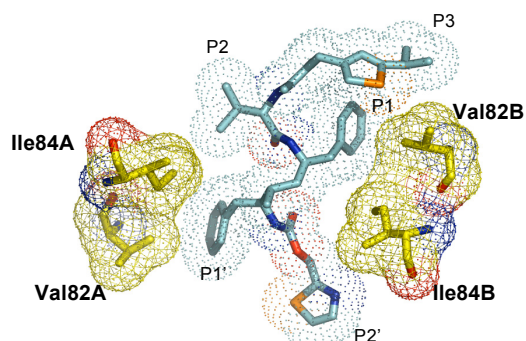
^bThe estimated binding energy evaluates the total energy of the binding interaction between the proteinase and the ligand where, $\Delta G = -RT \times \ln K$, ΔG is the change in Gibbs free energy, R is the gas constant ($8.314 \text{ J mol}^{-1}\text{K}^{-1}$), T is the temperature (K) = 298.2 K, and K is the equilibrium constant.

n.d – not determined

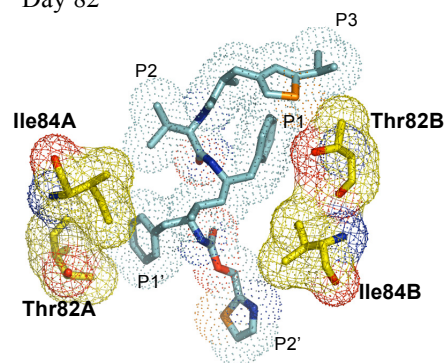
In particular, VDW packing in the wild-type complex (Table 4.20) is achieved by the side-chain atom Cγ1/Cγ2 of Val82A/B, contacting the phenylalanine group at P1/P1' and the 2-isopropyl-4-thiozoyl at P3. Since the phenylalanine group at the P1/P1' positions have been designed to target the HIV-1 proteinase pockets S1/S1' lined with Val82, a movement of the side-chain, or a shift of the Cα of Val82 will affect the VDW interactions. The Thr82B (but also Thr82A) side-chain of day 82 and day 115 exhibits a different conformation when compared to the wild type. In the variants, the Oγ1 atom is facing towards P1 while Cγ2 points towards P3, thus increasing the VDW interactions at P3 (Figure 4.18, B and C). In wild type, Val82B Cγ1 is facing towards the hydrophobic environment created by the side-chains of Leu10B and Leu23B, making only one contact with P3 while Cγ2 faces P1, making

four contacts.

Wild-type complex (1hwx.pdb)



Day 82



Day 115

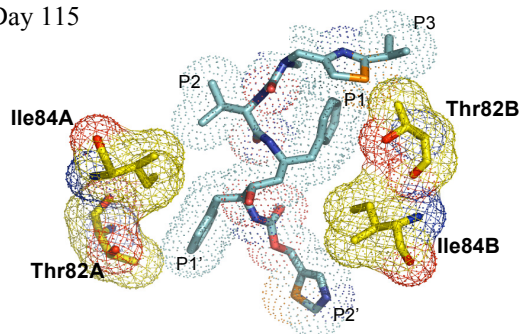
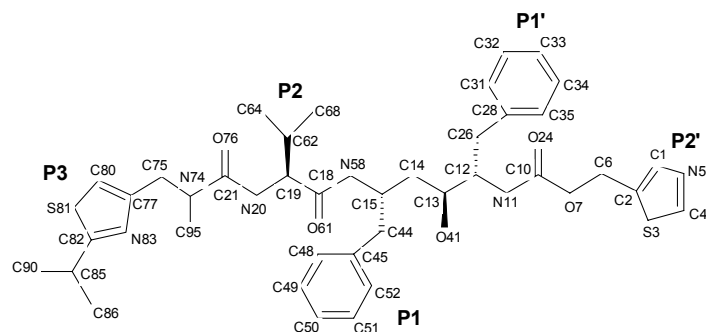


Figure 4.18: VDW interactions between ritonavir and residues Val/Thr82, Ile84 of the proteinase.

Closer VDW interactions in the binding cavity, as observed between for P2', P1 and P3 (Table 4.21), can finally result in repulsion, *i.e.*, creating unfavourable interactions, explaining the 400-fold increase in K_i values.

Table 4.21: Details of VDW interactions between the residues of ritonavir, P3, P2, P1, P1', P2' and the proteinase. The RIT 301 binding mode is similar to RIT 2Y. Listed are the number of VDW contacts up to 3.7 Å.

Ritonavir structural formula



Ritonavir side chains	1hwx.pdb (RIT 301)	Day 82 (chain Y)	Day 115 (chain Y)	Day 82 (chain X)	Day 115 (chain X)
P3 (C77, C80,S81, C82, N83, C85, C15, C14, C13, C12, N11, C10, O7, C6)	12	14	18	17	22
P2 (C62, C64, C68)	2	2	5	3	4
P1 (C45, C48, C49, C50, C51, C52)	15	18	18	20	14
P1'(C28, C31, C32, C33, C34, C35)	14	13	12	13	14
P2' (C1, C2, S3, C4, N5)	16	17	18	16	19
Ritonavir backbone					
C75	3	4	4	4	4
N74	1	1	1	1	1
N20	1	1		1	1
C19	2	1	1	1	1
C18			2		
N58	1	2	2	1	1
C15	1	1	1	1	1
C14	3	3	4	3	4
C13	5	5	5	4	6
C12	1	1	1	2	2
N11	2	1	1	1	2
C10					
O7	2				1
C6	1	2	1	2	1
C95	1	2	2	2	2
O76	6	7	7	7	6
O61	4	2	2	2	3
O41	10	12	9	9	9
O24	2	1	2	1	
C44	2	1	2	2	1
C26	2	3	2	2	1
Total	109	114	120	115	120

4.5.6 Inter- and intra-monomer interactions in variants day 82 and day 115 in complex with ritonavir.

The average r.m.s. deviations of the C α atoms of the 99 amino-acid residues for day 82 and day 115 is approximately 0.50 Å when compared to 1hwx.pdb (Table 4.22), suggesting only small changes. The catalytic dyad region of day 82 and day 115 remains unaltered with an average r.m.s. deviation of 0.12 and 0.10 Å, respectively. The average r.m.s. deviation of the flaps is about 0.40 Å for both variants, with residue Gly49A/B showing the largest deviation of about 1.0 Å. Significant conformational changes at the 80's loop with an average r.m.s. deviations of 0.42 to 0.47 Å are observed, with residues Pro79A and Pro81B showing largest distance from the wild type.

Table 4.22: The average r.m.s. deviations of C α atoms (Å) for the 99 amino-acid residues, the flaps (residues 41-57), catalytic dyad (residues 24-28) and the 80's loop (residues 75-86).

	Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)		Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)	
	chain A	chain B	chain A	chain B
1hwx.pdb C α atoms of the 99 amino-acid residues	0.46	0.44	0.43	0.45
Catalytic dyad (residues 24-28)	0.10 (Gly27A, Ala28A: 0.14)	0.12 (Asp25B: 0.21)	0.10 (Gly27A: 0.18)	0.10 (Asp25B: 0.19)
Flap (residues 41-57)	0.41 (Gly49A: 0.98)	0.37 (Gly49B: 0.73)	0.43 (Gly49A: 1.01)	0.40 (Gly49B: 0.94)
80's loop (residues 75-86)	0.47 (Pro79A: Thr80A:0.81)	0.46 (Pro81B: 1.51)	0.42 (Pro79A: 0.86)	0.47 (Pro81B: 1.50)

Numbers in parentheses indicate the maximum deviation from wild type, 1hwx.pdb.

C α - C α distances (Å) between amino acids located at the binding cavity (flaps and catalytic dyad) are presented in Table 4.23 and Figure 4.19.

Table 4.23: C α - C α distances (Å) between amino acids located at the binding cavity.

Cα - Cα distances (Å)	1hwx.pdb (wild-type complex)	Day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr)	Day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)
Thr80A-Thr80B	18.11	17.24 (- 0.87)	17.18 (- 0.93)
Pro81A-Pro81B	19.70	18.72 (- 0.98)	18.50 (- 1.20)
Val82A-Val82B	20.06	-	-
Thr82A-Thr82B	-	19.37 (- 0.69)	19.29 (- 0.77)
Asn83A-Asn83B	21.91	20.95 (- 0.96)	21.15 (- 0.76)
Ile84A-Ile84B	16.19	15.85 (- 0.34)	16.08 (- 0.11)

Ile85A-Ile85B	17.77	17.66 (- 0.11)	17.73 (- 0.04)
The gap between the flaps			
Gly51B-Gly52A	4.29	4.25 (- 0.04)	4.32 (+ 0.03)
Gly51A-Gly52B	4.21	4.25(+ 0.04)	4.39 (+ 0.18)
Gly49B-Ile50A	4.60	4.89 (+ 0.29)	5.08 (+ 0.48)
Gly49A-Ile50B	5.62	4.92 (- 0.70)	4.90 (- 0.72)
Distance between flap and catalytic dyad			
Ile50A-Asp25B	13.13	12.87 (- 0.26)	12.76 (- 0.37)
Ile50B-Asp25A	12.42	12.88 (+ 0.46)	12.82 (+ 0.40)

Values in parentheses are deviations from wild type values; (-) indicates the distance has decreased and (+) indicates the distance has increased.

Especially the distances between the $C\alpha$ pairs of Thr80A-B, Pro81A-B, Thr82A-B, and Asn83A-B have decreased considerably (Table 4.23), when compared to wild type. Overall, there are no large differences in the positions of the flaps except for the distance between Gly49A and Ile50B.

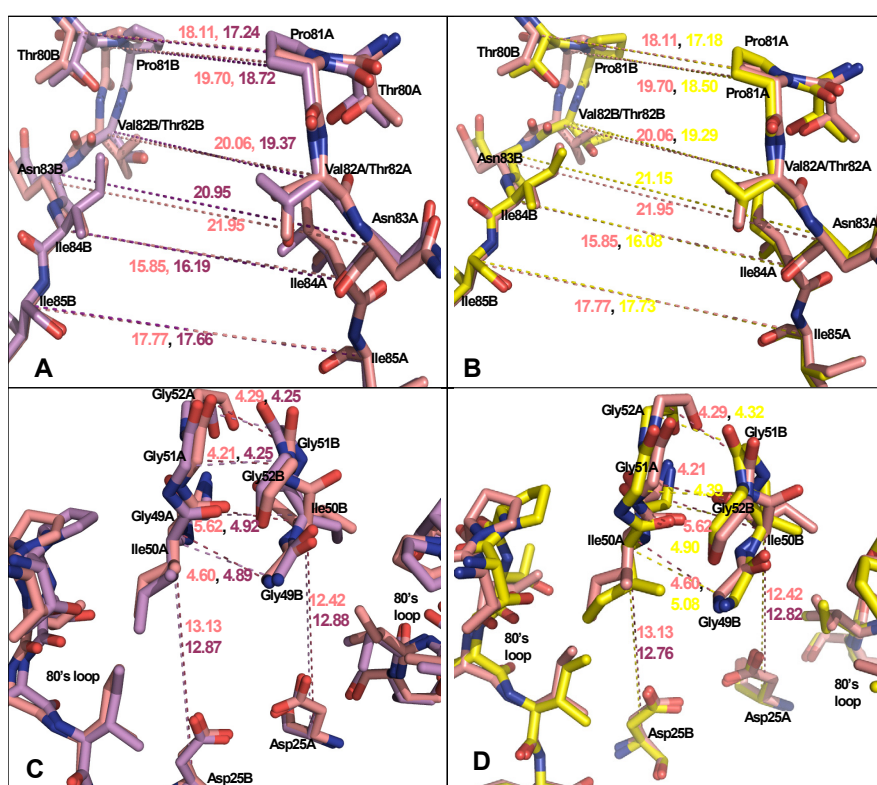


Figure 4.19: $C\alpha$ - $C\alpha$ distances (\AA) between amino acids located at the binding cavity. Distances between the 80's loop (A and B) and between the flaps and catalytic dyad (C and D). Variant day 82 coloured in purple and variant day 115 is coloured in yellow. Wild type, 1hwx.pdb, is coloured in pink.

In order to investigate whether there are any significant movements of the 80's loop to the flap region, distances of the corresponding residues are analysed (Table 4.24 and Figure 4.20).

Table 4.24: Comparison of the $C\alpha$ - $C\alpha$ distances (Å) between the flaps (residues Ile50, Ile54, Lys55, Val56) and the 80's loop (residues Ile84, Asn83, Val82/Thr82, Pro81, Thr80, Pro79, Gly78) in 1hwx.pdb, day 82 and day 115, all in complex with ritonavir.

Distances	1hwx.pdb	Day 82	Day 115	Distances	1hwx.pdb	Day 82	Day 115
Gly78A-Val56A	4.64	4.47 (- 0.17)	4.44 (- 0.20)	Gly78B-Val56B	4.54	4.38 (- 0.16)	4.54 (0.0)
Pro79A-Lys55A	7.87	7.73 (- 0.14)	7.62 (- 0.25)	Pro79B-Lys55B	8.15	7.80 (- 0.35)	7.60 (- 0.55)
Thr80A-Ile54A	8.00	-	-	Thr80B-Ile54B	8.59	-	-
Thr80A-Val54A		8.43 (+ 0.43)	8.27 (+ 0.27)	Thr80B-Val54B		8.30 (- 0.29)	8.21 (- 0.38)
Thr80A-Ile50B	6.00	5.76 (- 0.24)	5.68 (- 0.32)	Thr80B-Ile50A	6.37	5.69 (- 0.68)	5.65 (- 0.72)
Pro81A-Ile50B	6.70	7.06 (+ 0.36)	6.89 (+ 0.19)	Pro81B-Ile50A	7.95	7.02 (- 0.93)	6.91 (- 1.04)
Val82A-Ile50B	8.35	-	-	Val82B-Ile50A	9.29	-	-
Thr82A-Ile50B		8.61 (+ 0.26)	7.42 (- 0.93)	Thr82B-Ile50A		8.63 (- 0.66)	8.49 (- 0.80)
Asn83A-Ile50B	10.36	10.25 (- 0.11)	10.26 (- 0.10)	Asn83B-Ile50A	11.20	10.20 (- 1.0)	10.26 (- 0.94)
Ile84A-Ile50B	9.87	10.21 (+ 0.34)	10.20 (+ 0.33)	Ile84B-Ile50A	10.71	10.22 (- 0.49)	10.20 (- 0.51)

Values in parentheses are deviations from wild type values; (-) indicates the distance has decreased and (+) indicates the distance has increased.

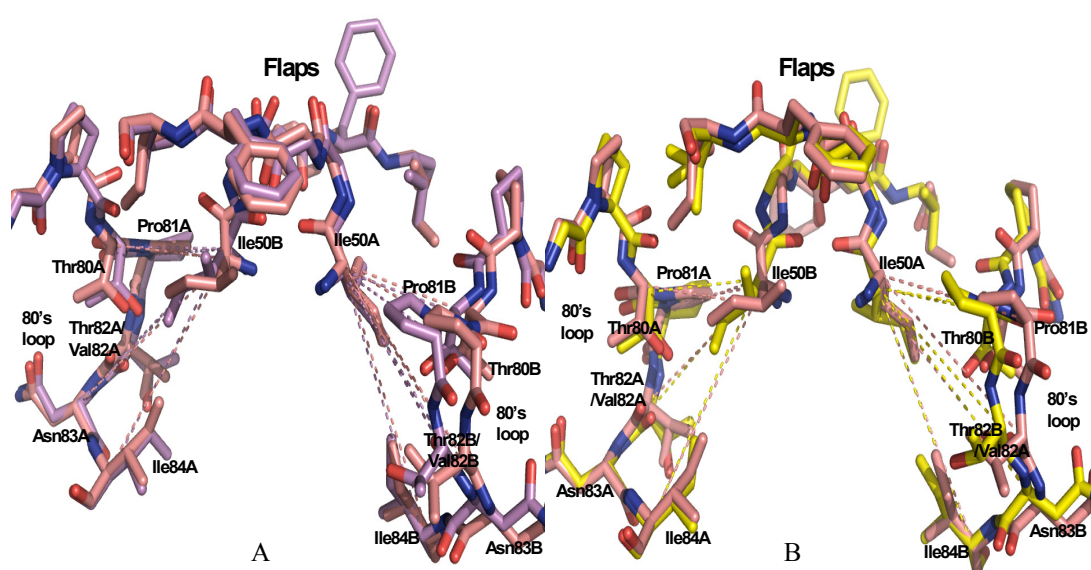


Figure 4.20: Superimposition of variant day 82 (purple) and day 115 (yellow) onto wild type, 1hwx.pdb (pink).

Distances involving residues Thr80B, Pro81B, Thr82B, Asn83B, Ile84B and Ile50A have decreased significantly (more than 0.5 Å) when compared to 1hwx.pdb. However, this is not the case for distances involving residues Thr80A, Pro81A, Thr82A, Asn83A, Ile84A and Ile50B. Taken together, the asymmetry between chains A and B in the wild-type complex is clearly reduced for variant day 82 and day 115, although no NCS (non crystallographic symmetry) was applied in the refinement.

In variant day 82 in complex with ritonavir, Arg57 Nη2 does not form a salt bridge with Asp35 Oε2 (Table 4.25 and Figure 4.21). The Nη2 Arg57 of variant day 82, and day 115, makes hydrogen bond with the Oη atom of Tyr59. In wild type, 1hwx.pdb, there is a glutamate at position 35. The Glu to Asp substitution is not resulting from ritonavir treatment but reflex sequence polymorphism in HIV-1 proteinase. Other salt-bridge interactions are conserved (Table 4.25).

The loss of the salt-bridge between Arg57 and Asp35 is due to the amino-acid substitution at residue 36 from methionine to isoleucine. Similar to the saquinavir complex, the backbone of Asp35 is forced to come closer to the 10's loop (residues 18 to 25), redirecting the Ile36 side-chain into the hydrophobic environment created by Leu33, Ile15 and the Cα-atoms of Leu19 and Lys/Arg20 (Figure 4.21). As a result, the side chain of Asp35 now points away from Arg57. The Asp35 Oε1 makes a hydrogen bond with the Nδ2 of Asn83 in both chain A and B of variant day 82 (Figure 4.21). In variant day 115, only Asp35 Oε1 of chain B is involved in making a hydrogen bond.

Table 4.25: Salt-bridges (distances in Å) in 1hwx.pdb, day 82 and day 115.

Salt-bridge interactions	1hxb.pdb (wild-type complex)	Day 82 (Met36Ile-Ile54Val- Ala71Val-Val82Thr)	Day 115 (Lys20Arg-Met36Ile- Ile54Val-Ala71Val- Val82Thr)
Nη2 Arg8B-Oδ2 Asp29A	2.78	2.73	2.77
Nη2 Arg8A- Oδ2 Asp29B	3.76	2.69	2.72
Oδ1 Asp29B- Nη2 Arg87B	2.89	3.00	2.86
Oδ1 Asp29A- Nη2 Arg87A	3.23	2.95	2.80
Nη1 Arg57B-Oε1 Glu35B	3.09	-	-
Nη1 Arg57A-Oε1 Glu35A	2.82	-	-

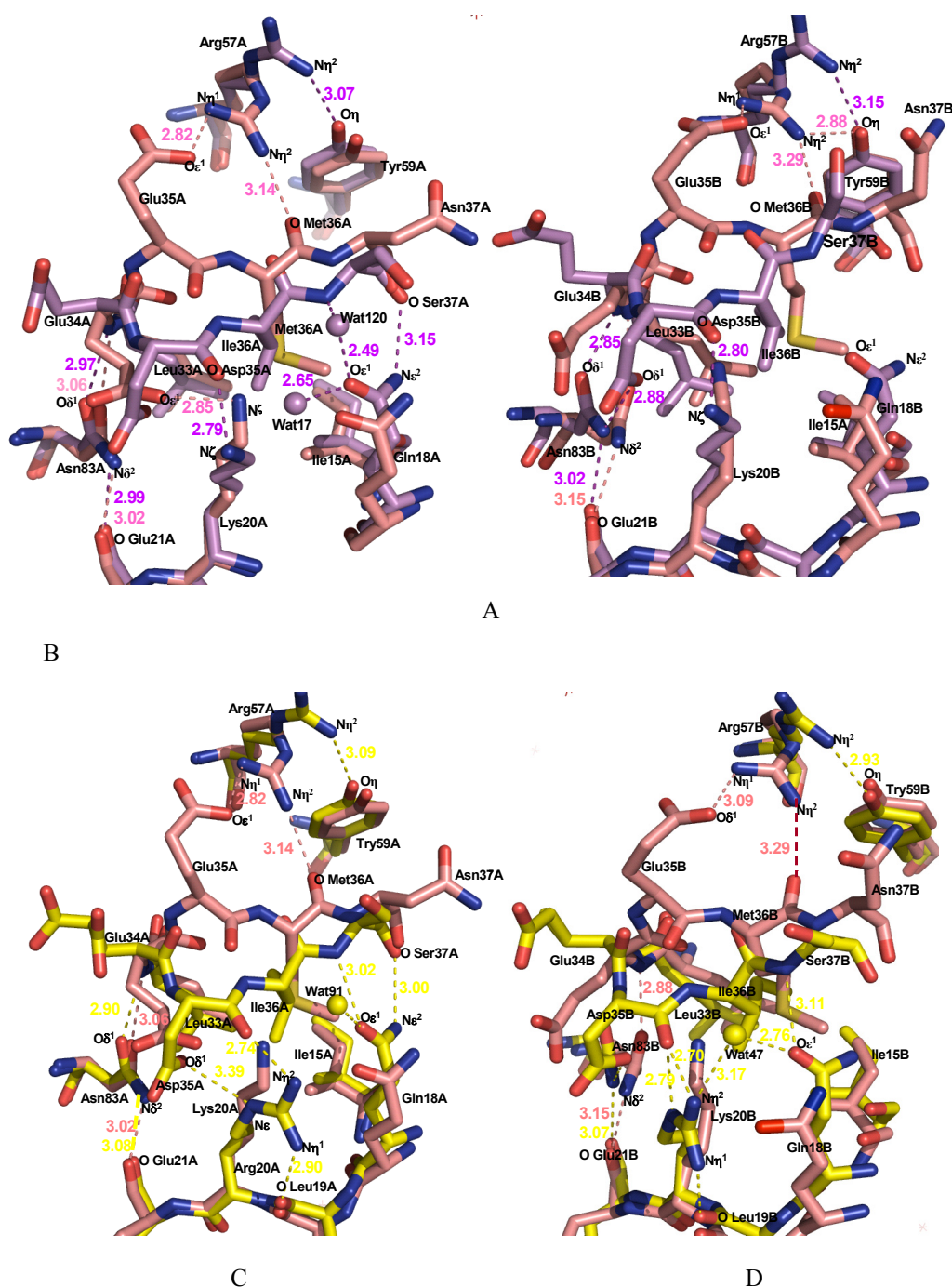


Figure 4.21: The intra-subunit salt-bridges Glu35A and Arg57A, Glu35B and Arg57B; the hydrogen-bonding interactions between the 30's loop (residues Glu34, Asp35, Ser37) and residues Asn83, Lys20/Arg20. Variant day 82 (purple) of chain A (A) and chain B (B) superimposed onto wild type, 1hwx.pdb (pink). Variant day 115 (yellow) of chain A (C) and chain B (D) superimposed onto wild type, 1hwx.pdb (pink).

A hydrogen-bonding network, which is not present in the wild-type complex, is established between the 10's (residues 10 to 20) and the 30's loop in variant day 82 and day 115. This network is also observed in 1sh9.pdb, a variant (Lys20Arg-Val32Ile-Leu33Phe-Met36Ile-Leu63Pro-Ala71Val-Val82Ala-Leu90Met-Met46Ile-Ile84Val) in complex with ritonavir. In day 115, Arg20 establishes hydrogen bonds

with the carbonyl oxygens of Asp35 and Leu19 (Figure 4.21, C and D, Table 4.26). Further, the O ϵ 1 of Gln18 makes a hydrogen bond with the main-chain amide of Ser37. In addition, the N ϵ 2 of Gln18A makes a hydrogen bond with the carbonyl O of Ser37A (Figure 4.21, C and D). In variant day 82, Gln18A O ϵ 1 makes a hydrogen bond with a water molecule, Wat120, while N ϵ 2 of Gln18A makes a hydrogen bond with the carbonyl O of Ser37A. In chain B of day 82, the hydrogen bond between the Gln18 and Ser37 is not observed (Figure 4.21, A and B).

Table 4.26: Hydrogen-bonding interactions (Å) between the 30's loop (residues Glu34, Asp35, Ser37) and residues Asn83, Lys20/Arg20, Leu19, Gln18, in 1hwx.pdb, day 82 and day 115.

Hydrogen-bonding interactions	1hwx.pdb (wild-type complex)	Hydrogen-bonding interactions	Day 82	Hydrogen-bonding interactions	Day 115
N η 2 Arg57A-O Met36A	3.14	N η 2 Arg57A-O Met36A	-	N η 2 Arg57A-O Met36A	-
N η 2 Arg57B-O Met36B	3.29	N η 2 Arg57B-O Met36B	-	N η 2 Arg57B-O Met36B	-
N η 2 Arg57A-O η Tyr59A	-	N η 2 Arg57A-O η Tyr59A	3.07	N η 2 Arg57A-O η Tyr59A	3.09
N η 2 Arg57B-O η Tyr59B	-	N η 2 Arg57B-O η Tyr59B	3.15	N η 2 Arg57B-O η Tyr59B	2.93
N Glu34A-O δ 1 Asn83A	3.06	N Glu34A-O δ 1 Asn83A	2.97	N Glu34A-O δ 1 Asn83A	2.90
N Glu34A-O δ 1 Asn83A	3.07	N Glu34A-O δ 1 Asn83A	2.85	N Glu34A-O δ 1 Asn83A	3.15
N δ 2 Asn83A-O Glu21A	3.02	N δ 2 Asn83A-O Glu21A	2.99	N δ 2 Asn83A-O Glu21A	2.88
N δ 2 Asn83B-O Glu21B	3.42	N δ 2 Asn83B-O Glu21B	3.02	N δ 2 Asn83B-O Glu21B	3.15
O ϵ 1 Glu34A-N ζ Lys20A	2.85	O ϵ 1 Glu34A-N ζ Lys20A	-	-	-
O ϵ 1 Asp35A-N δ 2 Asn83A	-	O ϵ 1 Asp35A-N δ 2 Asn83A	2.92	O ϵ 1 Asp35A-N δ 2 Asn83A	-
O ϵ 1 Asp35B-N δ 2 Asn83B	-	O ϵ 1 Asp35B-N δ 2 Asn83B	2.88	O ϵ 1 Asp35B-N δ 2 Asn83B	2.90
O ϵ 1 Asp35A-N ϵ Arg20A	-	-	-	O ϵ 1 Asp35A-N ϵ Arg20A	3.39
O Asp35B- N ϵ Arg20B	-	-	-	O Asp35B- N ϵ Arg20B	2.79
O Asp35A- N ζ Lys20A	-	O Asp35A- N ζ Lys20A	2.79	O Asp35A-N η 2 Arg20A	2.74
O Asp35B- N ζ Lys20B	-	O Asp35B- N ζ Lys20B	2.80	O Asp35B-N η 2 Arg20B	2.70
Wat120-O ϵ 1 Gln18A	-	Wat120-O ϵ 1 Gln18A	2.49	Wat120-O ϵ 1 Gln18A	-
N Ser37A-Wat120	-	N Ser37A-Wat120	2.51	N Ser37A-O ϵ 1 Gln18A	3.02
N Ser37B-O ϵ 1 Gln18B	-	N Ser37B-O ϵ 1 Gln18B	-	N Ser37B-O ϵ 1 Gln18B	3.11
Wat17-O ϵ 1 Gln18A	-	Wat17-O ϵ 1 Gln18A	2.65	Wat91-O ϵ 1 Gln18A	2.83
Wat47-O ϵ 1 Gln18B	-	Wat47-O ϵ 1 Gln18B	-	Wat47-O ϵ 1 Gln18B	2.76
Wat47-N η 2 Arg20B	-	Wat47-N η 2 Arg20B	-	Wat47-N η 2 Arg20B	3.17
N η 1 Arg20A-O Leu19A	-	N η 1 Arg20A-O Leu19A	-	N η 1 Arg20A-O Leu19A	2.90
N η 1 Arg20B-O Leu19B	-	N η 1 Arg20B-O Leu19B	-	N η 1 Arg20B-O Leu19B	2.89
O Ser37A-N ϵ 2 Gln18A	-	O Ser37A-N ϵ 2 Gln18A	3.15	O Ser37A-N ϵ 2 Gln18A	3.00

4.6 Structural analysis of inhibitor-free HIV-1 proteinase of variant day 82 (Met36Ile-Ile54Val-Ala71Val-Val82Thr) and day 115 (Lys20Arg-Met36Ile-Ile54Val-Ala71Val-Val82Thr)

The crystal structures of the unliganded form are described for day 82 and day 115, diffracting X-rays to a maximum Bragg spacing of 2.40 and 2.20 Å resolution, respectively. Crystallization was difficult compared to obtaining complexes with either saquinavir or ritonavir. Small needle-like crystals were easily grown after two days especially when sodium acetate was used as the reservoir buffer. Several crystallization strategies have been tried as mention in chapter 3 (MATERIALS AND METHODS) in order to obtain better-diffracting crystals. By increasing the proteinase concentration to about two times higher (final concentration of 2.0 mg/ml), bigger crystals were obtained. Both, inhibitor-free, day 82 and day 115 crystallized in the same space group, P6₁, and showed a similar quality of electron density maps.

As mention earlier (Nijhuis *et al.*, 1999), addition of Ala71Val or Lys20Arg-Ala71Val to the triple variant, Met36Ile-Ile54Val-Val82Thr, did not increase resistance but improved the catalytic efficiency to better than wild type values. Mutation of Lys20Arg did not change the electrostatic potential since lysine and arginine are positively charged (substitution of the butyl amine with a guanido group). Interestingly, both variants crystallized in the so-called "flaps closed" form. Four unliganded wild type (Lapatto *et al.*, 1989; Navia *et al.*, 1989; Wlodawer *et al.*, 1989; Spinelli *et al.*, 1991) and two multidrug-resistant HIV-PR structures have been reported with the "flaps open" (Logsdon *et al.*, 2004; Martin *et al.*, 2005). The inhibitor-free multidrug-resistant variants were isolated from a patient who had received long-term antiretroviral therapy (MDR 769, mutations at residues 10-25-36-46-54-63-71-82-84-90). So far, only two variant structures of a tethered dimer, 1g6l.pdb and 1lv1.pdb, of uncomplexed HIV-1 proteinase have been deposited with the Protein Data Bank (see Table 4.27; Berman *et al.*, 2000) displaying a "flaps closed" conformation.

Table 4.27: Crystal structures of unliganded HIV-1 proteinase*

PDB code	Resolution (Å)	R factor (%)	Space group	Comments	Reference
3phv.pdb	2.70	19.1	P ₄ ₁ 2 ₁ 2	open flaps, HXB2 wild-type	Lapatto <i>et al.</i> , 1989
2hvp.pdb	3.00	37.0	P ₄ ₁ 2 ₁ 2	open flaps, NY5 wild-type	Navia <i>et al.</i> , 1989
3hvp.pdb	2.80	18.4	P ₄ ₁ 2 ₁ 2	open flaps, SF2 wild-type, synthetic	Wlodawer <i>et al.</i> , 1989
1hhp.pdb	2.70	19.0	P ₄ ₁ 2 ₁ 2	open flaps, BRU wild-type	Spinelli <i>et al.</i> , 1991
1rpi.pdb	1.86	21.7	P ₄ ₁	wide open flaps	Logsdon <i>et al.</i> , 2004
1tw7.pdb	1.30	14.0	P ₄ ₁	Multidrug-Resistant wide open flaps	
1g6l.pdb	1.90	19.5	P ₆ ₁	Multidrug-Resistant closed flaps, Cys95Met	Martin <i>et al.</i> , 2005
1lv1.pdb	2.10	19.4	P ₆ ₁	closed flaps, Cys95Met/Cys1095Ala	Pillai <i>et al.</i> , 2000
					Kumar <i>et al.</i> , 2002

*1tw7.pdb and 1rpi.pdb are similar structures at different resolution. 1g6l.pdb and 1lv1.pdb are tethered dimers.

4.6.1 Description of the structures: inhibitor-free variant day 82 and day 115 in the closed form

The inhibitor-free variants, day 82 and day 115, exhibit the flaps closed conformation in contrast to the other four inhibitor-free variants, day 0, day 28^a, day 28^b, and day 56 that have their flaps in the open conformation. The variants were crystallized in the absence of inhibitor, and thus the addition of inhibitor or substrate is clearly not a prerequisite for adopting the closed conformation. The flaps are thought to accept two conformations, an open form in order for substrates to bind and a closed form for the catalytic reaction to take place.

Seven and eight water molecules are observed in the substrate binding-cleft of day 82 and day 115, respectively (Figure 4.22, Table 4.28). One water molecule is positioned in between the aspartates of the catalytic dyad, Wat40 in day 82 and Wat49 in day 115. Such a water molecule is also seen in the unliganded crystal structures of day 0, day 28^a and day 56 but it is not visible in unliganded day 28^b. Piana *et al.* (2002) suggested that this particular water molecule, polarized by Asp25A/B, is involved in the proposed reaction mechanism of HIV-1 PR. In day 82, the N amide of Ile50A/B makes a hydrogen contact with water molecule Wat26. However, a similar water molecule is not seen in day 115. Five water molecules were reported in the, flaps closed, active-site cavity of unliganded SIV protease (Wilderspin and Surgue, 1994); Wat1 form hydrogen bonds with Asp25A/B, Wat2 form hydrogen bonds with the N amides of Ile50A/B, and Wat30, Wat28 and Wat29 form hydrogen bonds with flap residues. Only one water molecule, Wat413 was reported by Rose *et al.* (1998) to

make hydrogen bonds with Asp25A/B in the open form of unliganded SIV protease. In the closed formed of inhibitor-free 1lv1.pdb, with amino acid substitutions at position Cys95Met/Cys1095Ala, six water molecules were observed in the active-site cavity. The unique catalytic water molecule (hydrogen bonds with Asp25A/B) and flap water molecule (hydrogen bonded with the inhibitor and amide N atom of Ile50A/B) are also seen in the inhibitor-free variant Cys95Met, 1G6l.pdb.

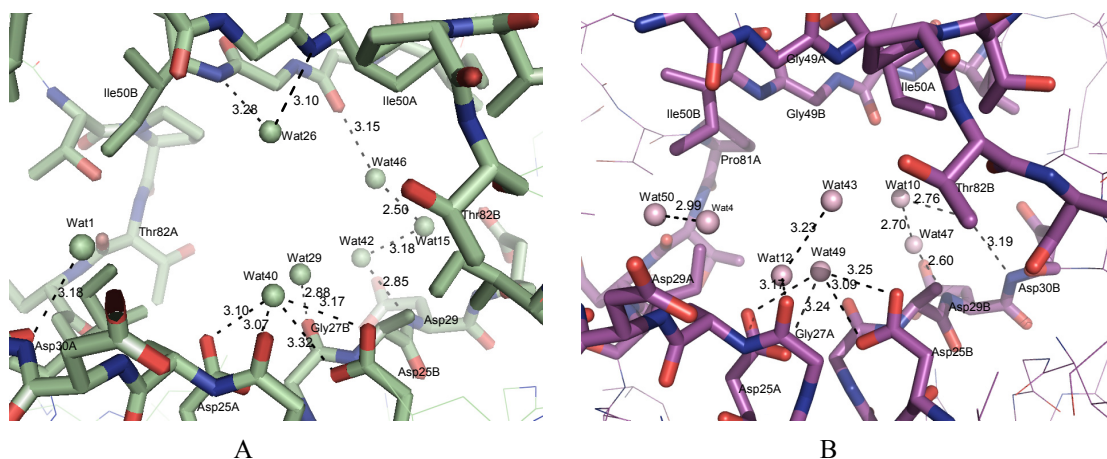


Figure 4.22: Water molecules in the active-site cavity of inhibitor-free variant (A) day 82 and (B) day 115.

Table 4.28: Water molecules in the active site cavity of unliganded variant day 82 and 115.

Day82			Day115		
water molecule	hydrogen-bonding (Å)	residues at the active site, water molecule	water molecule	hydrogen-bonding (Å)	residues at the active site, water molecule
Wat40	3.10 3.07 3.17 3.32	Oδ2 Asp25A Oδ1 Asp25A Oδ2 Asp25B Oδ1 Asp25B	Wat49	3.17 3.24 3.25 3.09	Oδ2 Asp25A Oδ1 Asp25A Oδ2 Asp25B Oδ1 Asp25B
Wat26	3.10 3.28	N Ile50A N Ile50B	Wat12	3.25	N Ala28A
Wat1	3.18	N Asp30A	Wat47	2.60	N Asp29B
Wat42	2.85 3.18	N Asp29B Wat15	Wat41	3.19	N Asp30B
Wat29	2.88	O Gly27B	Wat10	2.70 2.76	Wat47 Wat41
Wat46	3.15	O Gly48B	Wat50	2.99	Wat45
Wat15	2.50	Wat46	Wat43	3.23	Wat12

Both the inhibitor-free and ritonavir-complexes of day 82 and day 115 crystallized in the same space group, P6₁, even though the crystallization conditions were quite different. Unliganded crystals of day 82 grew from 1.0 M NaCl, 100 mM

Mes pH 6.5, while crystals of day 82 in complex with ritonavir were obtained from 25% $(\text{NH}_4)_2\text{SO}_4$, 200 mM NaH_2PO_4 , pH 6.5. For unliganded day 115, crystals were grown from 20% $(\text{NH}_4)_2\text{SO}_4$, 50 mM sodium cacodylate pH 6.0, while crystals of day 115 in complex with ritonavir were grown from 20% $(\text{NH}_4)_2\text{SO}_4$, 100 mM Mes pH 6.0. Intriguingly, acquisition of substitution Ala71Val induced a "flaps closed" conformation in the absence of any inhibitor.

At resolutions of 2.4 Å for day 82 and 2.2 Å for day 115, clear $2F_o - F_c$ electron density was observed for residues Lys/Arg20, Asp35, Ile36, Val54, Val71 and Thr82. An overview of selected hydrogen-bonding interactions is presented in Table 4.29 and Figure 4.23. In day 115, the carbonyl O of Asp35 makes hydrogen bonds with the $\text{N}\eta_2$ and $\text{N}\epsilon$ of Arg20. The $\text{O}\delta_1$ of Asp35 interact with $\text{N}\delta_2$ of Asn83 in day 82 and day 115. Hydrogen bonds are also formed between the carbonyl O of Ser37 and $\text{N}\epsilon_2$ of Gln18 as well as the N of Ser37 and $\text{O}\epsilon_1$ of Gln18 of day 82 and day 115. These hydrogen bonds are not detected in variant day 0, day 28^a, day 28^b, and day 56.

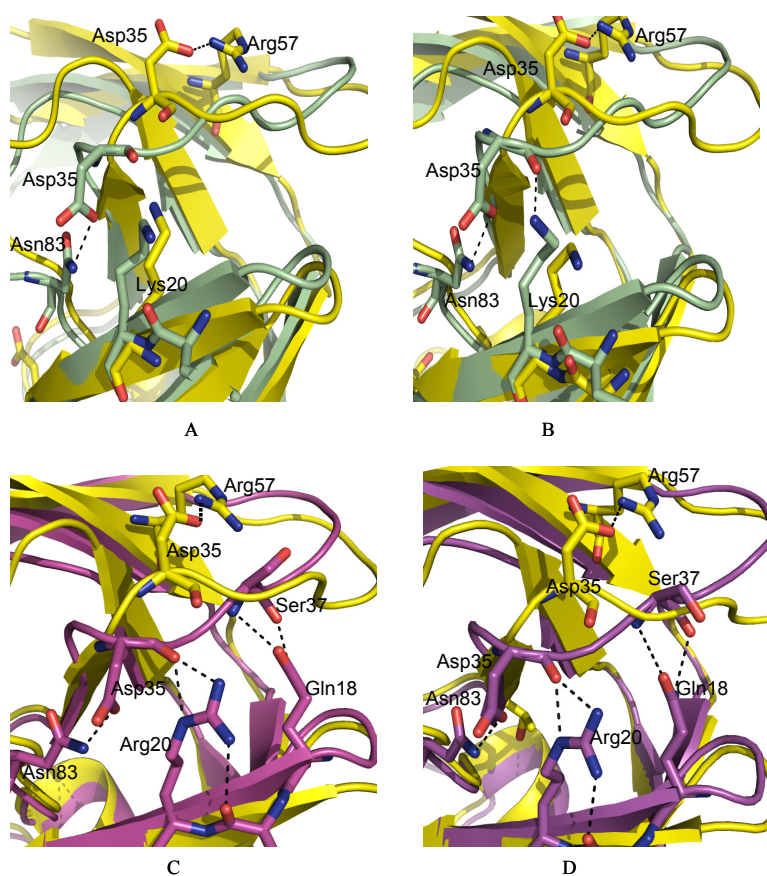


Figure 4.23: Intramolecular salt-bridge between Arg57-Asp35 and hydrogen-bonding between Lys20/Arg20 and neighbouring residues in variant day 82 (Panel A and B) and day 115 (Panel C and D). Day 0 is coloured yellow, day 82 green and day 115 magenta.

Table 4.29: Salt-bridge interactions (cut-off value of 3.5 Å, printed in bold) and hydrogen-bonding contacts (cut-off value 3.5 Å) of the six unliganded variants.

	Day 0		Day28 ^a		Day28 ^b		Day56		Day82		Day115	
Monomer	A	B	A	B	A	B	A	B	A	B	A	B
Nη1 Arg 57-Oδ1 Asp35	2.93	-	2.91	-	-	-	-	-	-	-	-	-
Nη2 Arg57-Oδ2 Asp35	-	2.77	-	2.64	-	-	-	-	-	-	-	-
Nη2 Arg57-O Met36	2.96	2.91	3.00	2.90								
Nη2 Arg57-Wat250	-	-	-	-	2.91	-	-	-	-	-	-	-
O Ile36-Wat250	-	-	-	-	3.01	-	-	-	-	-	-	-
Nη2 Arg57-OH Tyr59	-	-	-	-	-	-	-	-	3.07	3.08	3.18	-
Nζ Lys20-Oε1 Glu34	2.68	-	2.22	-	-	-	-	-	-	-	-	-
Nζ Lys20-Oδ1 Gln18	-	-	-	-	-	3.44	-	-	-	-	-	-
Nζ Lys20-O 34 Glu34	-	-	-	-	3.17	-	3.21	-	-	-	-	-
Nζ Lys20-O Asp 35	-	-	-	-	-	-	2.93	-	-	2.98	-	-
Nη2 Arg20-O Asp35	-	-	-	-	-	-	-	-	-	-	2.42	2.92
Nε Arg20-O Asp35	-	-	-	-	-	-	-	-	-	-	2.76	2.74
Oε1 Glu34-N Thr80	-	-	-	-	-	-	-	3.34			3.39	
Oε2 Glu34-N Thr80	-	-	-	-	-	-	-	-	-	-	-	2.54
Oε1 Glu34-Nη2 Asn83	-	-	-	3.45	-	3.16	-	-	-	-	-	-
Oε2 Glu34-Wat249	-	-	-	-	-	2.69	-	-	-	-	-	-
O Asp35-Wat249	-	-	-	-	-	3.00	-	-	-	-	-	-
N Glu34-Oδ1 Asn 83	2.90	2.92	2.98	2.87	2.87	3.04	3.21	3.20	2.99	2.75	2.87	2.94
Oδ1 Asp35-Nδ2 Asn83	-	-	-	-	-	-	-	-	2.70	3.24	2.75	2.78
O Glu21-Nδ2 Asn83	2.75	2.76	2.76	2.87	2.77	2.70	-	2.91	3.10	3.33	3.11	3.03
Oε1 Gln18-Nη2 Arg20	-	-	-	-	-	-	-	-	-	-	-	3.17
Nε2 Gln18-O Gly16	-	3.22	-	3.02	-	3.23	-	3.23	-	3.33	-	-
Nε2 Gln18-O S37	-	-	-	-	-	-	-	-	3.11	3.25	3.09	3.20
Oε1 Gln18-N S37	-	-	-	-	-	-	-	-	3.45	3.51	2.98	3.42

Residue Ala71Val is buried in a hydrophobic environment adjacent to amino-acid residues Ile66, Ile85, Leu89 and Ile93 (Figure 4.24). The addition of two methyl groups increases the volume of residue 71 considerably. Val71 makes close hydrophobic interactions with helix residues Gln92, Ile93 and Leu89 and β-sheet residue Ile66 (see Table 4.30). In variant day 0, day 28^a, day 28^b, and day 56, no close hydrophobic interactions were observed between Ala71 and the neighboring residues.

The Gln92 side-chain of day 82 and day 115 is pointing towards the C α of Ile72, forming hydrogen bonds with its amide and carbonyl oxygen atom. These hydrogen bonds are not observed in variant day 0, day 28^a, day 28^b and day 56.

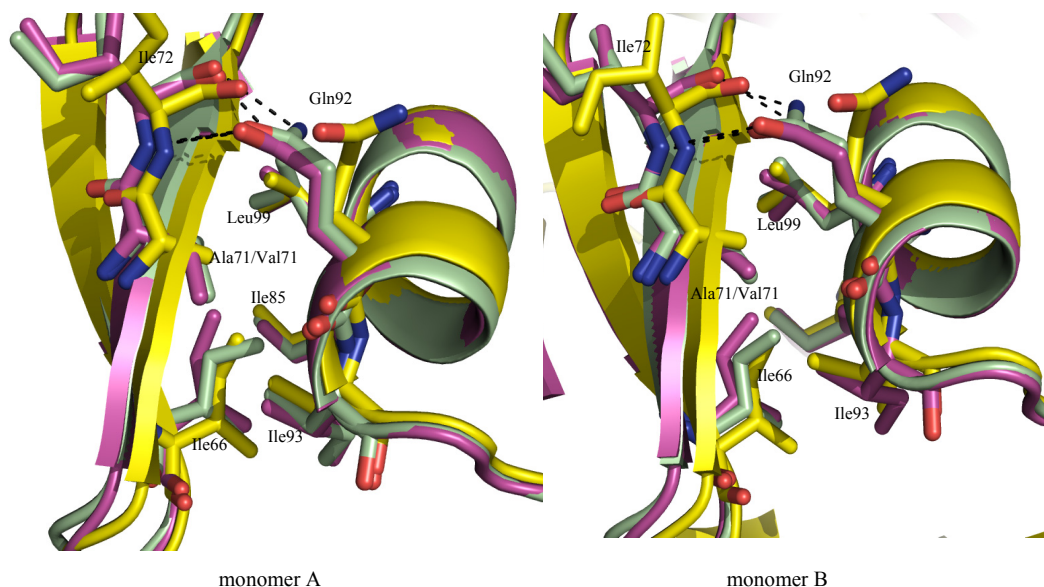


Figure 4.24: Hydrogen-bond interactions between the helix i (Gln92, Ile93, Leu89) and β -strands g and h (Ile62-Ile66 and Gly73-His69). Day 0 is coloured yellow, day 82 green, and day 115 yellow.

Table 4.30: Hydrophobic and hydrogen-bonding interactions (\AA) between helix i and β -strands g and h in variant day 82 and day 115.

Hydrogen-bonding		Monomer A		Monomer B	
day 82					
O Ile72–Nε2 Gln92		3.04		2.70	
N Ile72–Oε1 Gln92		3.26		3.29	
day 115					
O Ile72–Nε2 Gln92		2.46		2.90	
N Ile72–Oε1 Gln92		3.22		3.48	
Hydrophobic interactions					
Day 82			Day 115		
Monomer A		Monomer B	Monomer A		Monomer B
Cγ1 Val71–Cβ Gln92: 3.94		Cγ1 Val71–Cβ Gln92: 3.59	Cγ1 Val71–Cβ Gln92: 3.75		Cγ1 Val71–Cβ Gln92: 3.70
Cγ1 Val71–Cγ2 Ile93: 3.49		Cγ1 Val71–Cγ2 Ile93: 3.62	Cγ1 Val71–Cγ2 Ile93: 3.93		Cγ1 Val71–Cγ2 Ile93: 3.69
Cγ2 Val71–Cδ Ile66: 4.09		Cγ1 Val71–Cδ Ile66: 3.91	Cγ2 Val71–Cδ Ile66: 3.93		Cγ2 Val71–Cδ Ile66: 3.68
Cγ2 Val71–Cδ1 Leu89: 4.00		Cγ2 Val71–Cδ2 Leu89: 3.59	Cγ2 Val71–Cδ1 Leu89: 3.85		Cβ Val71–Cδ2 Leu89: 4.11
Cβ Val71–Cβ Gln92: 4.06		Cβ Val71–Cβ Gln92: 4.03	Cβ Val71–Cβ Gln92: 3.75		
		Cβ Val71–Cδ Gln92: 3.98			

The amino-acid substitution Ala71Val apparently endorses tighter interactions between helix i (residues 87-93) and β -strands g and h (residues 62-66 and 69-73): Ala71Val induces long-range structural changes in the proteinase that mainly stabilize the flaps in the closed form. The overall structural changes are accompanied by a threefold increase in k_{cat}/K_m value in day 82 relative to day 0. The latter is probably also true for variant day 115 (see Table 4.1). The accumulation of small structural changes induced by the different amino-acid substitutions renders the proteinase resistance to ritonavir and, moreover, improves the viral replication capacity to better than wild type values.

The r.m.s. deviation values for C α atoms (of 198 amino-acid residues) of the inhibitor free HIV-1 proteinase, in comparison to variant day 0 (inhibitor-free) are presented in Table 4.31.

Table 4.31: R.m.s. deviation values (Å) for C α atoms (of 198 amino-acids residues) between inhibitor-free HIV-1 proteinase and variant day 0.

	Day 28 ^a	Day 28 ^b	Day 56	Day 82	Day 115
Day 0					
Average r.m.s. deviation	0.14	0.54	0.69	1.45	1.43

5 DISCUSSION

5.1 Saquinavir (day 28^b, day 82, day 115) and ritonavir (day 82 and day 115) complexes

Within 115 days of ritonavir mono-therapy, the virus acquired a total of five mutations that are all located within the coding region for the viral proteinase. Up to day 56, three amino-acid substitutions (resulting from primary mutations) accumulated that reduced both the inhibitor binding and the viral replication capacity; Met36Ile, Ile54Val, and Val82Thr. From day 56 up to day 115, two more amino-acid substitutions (resulting from secondary mutations) are acquired that have little influence on inhibitor binding but greatly improve the viral replication capacity to better than wild type values; Ala71Val and Lys20Arg. A short description of what is known about these amino-acid residues is presented below.

The substitution Ile54 to Val, Thr, Ala, or Ser affords intermediate resistance to nelfinavir, saquinavir, indinavir, ritonavir, lopinavir and atazanavir but only a low level of resistance to amprenavir (<http://hivdb.stanford.edu/>). Actually, the Ile54Val substitution mainly confers resistance in the presence of additional substitutions, *e.g.*, in combination with Met36Ile. An Ile54Met/Leu substitution conveys an intermediate level of resistance to amprenavir. Together with additional substitutions, the Val82Ala exchange makes the proteinase resistant to all clinically used proteinase inhibitors (Johnson *et al.*, 2003; Hoffman *et al.*, 2003). In amprenavir resistance (Xu *et al.*, 2002), Ile54Met and Ile54Leu induce a rotation of the terminal methyl group of Ile50. This re-orientation of the C δ 1 atom results in fewer VDW contacts with amprenavir. Also, according to the HIV Stanford Protease Database (<http://hivdb.stanford.edu/>), exchange of Val82 to Ala, Phe, or Ser renders a high level of resistance to indinavir, ritonavir, and lopinavir, an intermediate level of resistance to nelfinavir, amprenavir, and atazanavir, and a low-level resistance to saquinavir. Mutagenesis and kinetic studies on inhibitor binding of HIV-1 and HIV-2 proteinases demonstrated that residues 31 and 33-37 (*e.g.* Met36Ile) play a role in inhibitor binding/specificity (Towler *et al.*, 1997; Swairjo *et al.*, 1998). The substitution Ala71Val, Met36Ile, or Lys20Arg is not known to confer resistance on its own. Polymorphism at Ala71 (Val or Thr) is observed in 1-2% of the HIV isolates taken from untreated persons but which becomes much more frequent in patient receiving PI treatment (<http://hivdb.stanford.edu/>).

Analogous to the situation in the wild-type complexes (PDB code 1hxb, wild-type complex with saquinavir and 1hwx, wild-type complex with ritonavir), in the variant complexes with saquinavir (day28^b, day 82 and day 115) or ritonavir (day 82 and day 115), the Ile54 to Val exchange did not induce a re-rotation of the C δ 1 atom of Ile50. This was not a surprise since both inhibitors do not make close contacts with amino acid residue 54. Unfortunately, the wild-type complex with ritonavir was crystallized in a different space group (see Table 4.2), a circumstance that made precise comparisons difficult. The different complexes obtained herein have been analyzed and compared exhaustively; hydrophobic interactions, hydrogen bonding patterns, side-chain orientations, intra- and inter-molecular distances, gap volumes (space in between the proteinase and inhibitor molecules), and active-site-cavity volumes. Somewhat unexpected, the broad analysis/comparison revealed more, and closer, contacts between ritonavir and proteinase variant day 82 and day 115, when compared to wild-type complex. Closer contacts in the variant complexes may be indicative of a stronger binding. However, recall that crystals of the complexes were obtained using excess amounts of inhibitor (~1 mM, ~500 \times K_i). More likely, the subtle structural changes in the active-site cavity require the binding of the inhibitor in a slightly distorted conformation (see VDW energy for inhibitor conformation, Table 4.20). More precise binding of a slightly distorted inhibitor (*i.e.*, a loss of entropy) must be paid for by a decrease in apparent affinity (maximization of the Circe effect; Jencks, 1997; Blow, 2000). Indeed, the K_i for ritonavir of variant day 56, day 82 and day 115 is about 400-fold increased compared to day 0 (K_i 4.7 nM; Nijhuis *et al.*, 1999). Also the saquinavir complexes exhibit closer contacts when compared to the wild-type complex. The effect is however not as pronounced as observed for the ritonavir complexes; the ritonavir resistant variants studied herein only display a medium resistance to saquinavir.

5.2 Unliganded variant day 0, day 28^a, day 28^b, day 56, day 82 and day 115

Rapid evolution offers an efficient way to redesign proteins with improved/novel functional characteristics. The selective pressure of ritonavir treatment forces the HIV-1 virus to evolve, escaping drugs by manipulating the inter-

/intra-molecular interactions that affect the structural/conformational stability as well as the proteinase dynamics (*e.g.*, opening and closing of the flaps). Opposite to Ile36, Val54, Val71, and Thr82, Lys/Arg20 is completely exposed at the proteinase surface. Changes in hydrophobicity of buried side-chains have been shown to affect the protein stability (Clemente *et al.*, 2003; Matsumura *et al.*, 1988; Kellis *et al.*, 1988; Shortle *et al.*, 1990).

The X-ray structures of six unliganded HIV-1 PR variants, obtained from a patient treated with ritonavir, demonstrate how viral evolution can traverse complex fitness landscape in order to survive and produce infectious progeny (Figure 5.1). A single substitution, Val82Thr (variant day 28^a), is sufficient to reduce the susceptibility of HIV Pr for ritonavir (<http://hivdb.stanford.edu/>). Variant day 0 and day 28^a exhibit an almost comparable relative replication capacity whereas the K_i value differs 17-fold. The Val82 to Thr exchange, from hydrophobic to polar, can perturb the hydrophobic environment created by the side-chains of Val82, Leu23 and Ile84. Thr82 is in close proximity to the P3 and P1 side-chains of ritonavir and may thus exert a direct effect on inhibitor binding. Further, Thr82 may introduce main-chain movements of the 80's loops that in turn affect the cantilevers (residues 66-62 and 73-69) and flaps of the proteinase dimer. Nevertheless, an alignment of the C α atoms of 198 amino-acid residues of variant day 0 and day 28^a reveals negligible structural changes with an overall r.m.s. deviation of 0.14 Å (see Table 4.31).

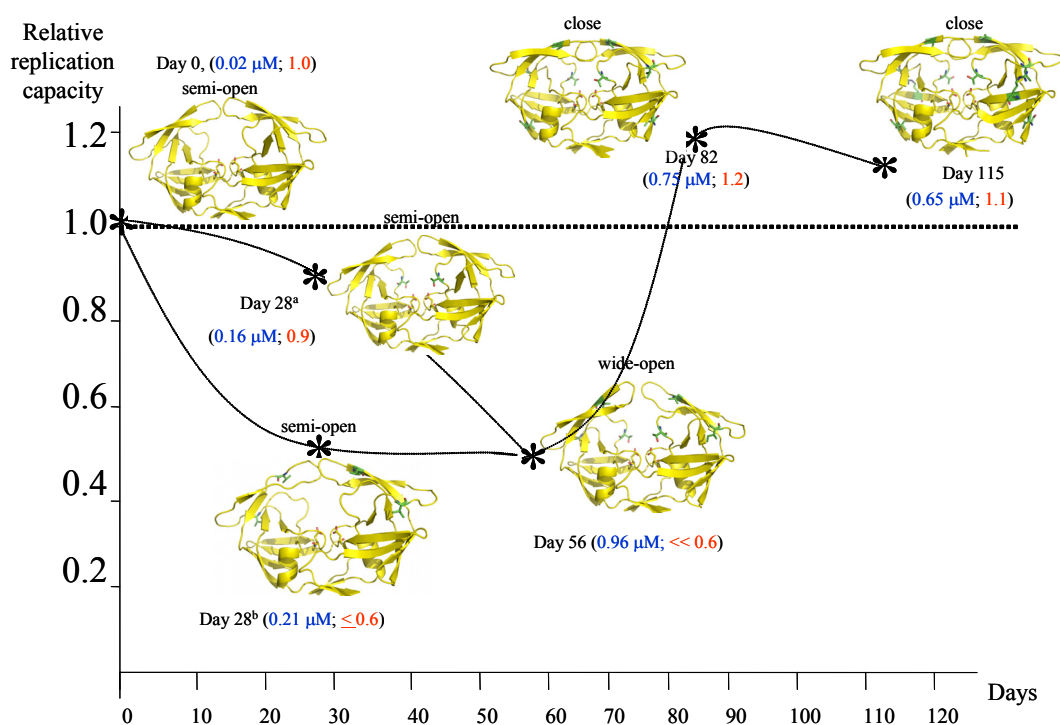


Figure 5.1: Relative replication capacity (RC) levels of variant proteinases. Values in brackets; IC₅₀ values in blue and relative replication capacity in red.

As mentioned above, a substitution at position 54 mainly confers resistance in the presence of an additional amino-acid exchange. Indeed, in day 28^b, the Ile54Val and Met36Ile substitutions occur simultaneously. For variant day 28^b, the viral replication capacity drops to below 0.6 and the K_i value increases 9-fold. The Ca r.m.s. deviation of 0.54 Å suggests somewhat larger differences than observed between wild type and day 28^a. In the flaps closed conformation, the Cδ1 atom of Ile54 (day 0 and day 28^a) is situated in a hydrophobic pocket created by residues Ile47, Ile50, Val56 and Pro79, Thr80, and Pro81. Removal of the Cδ1 atom may slightly destabilize a flaps closed conformation. It has been shown that when large core side-chains are replaced with smaller ones, the loss in stability appears to correlate with both a loss in hydrophobicity and a loss in internal packing efficiency (Ericksson *et al.*, 1993). On the other hand, removal of the Cδ1 atom (Ile54Val) may allow the flaps to embrace the inhibitor more closely, thus affecting the shape and/or volume of one or more substrate binding pockets. In combination with Ile54Val, the Met36Ile substitution significantly affects the 30's loops and, due to main-chain rearrangement, the Asp35-Arg57 and Glu34-Lys20 salt bridges are weakened/disrupted. Possible, these two salt bridges form an important link between the flaps and the bottom of the active-site (Arg57...Asp35-Glu34...Lys20). The flaps

must firmly and correctly embrace the substrate in order for the cleavage reaction to proceed. Occurrence of both Val54 and Ile36, as present in variant day 28^b, may directly influence the volume and/or shape of the active-site cavity, which may well affect both the k_{cat} and K_m values (not determined, Nijhuis *et al.*, 1999).

Interestingly, the two evolutionary separated pathways (*i.e.*, variants day 28^a and 28^b) meet up to create a new variant, day 56. Variant day 28^a (Val82Thr) and day 28^b (Met36Ile-Ile54Val) set the foundation for the virus to escape from ritonavir. With day 56, the K_i value for ritonavir increases 411-fold relative to variant day 0. But, the combination of all three substitutions generates a dramatic effect on the proteinase, severely reducing the replication capacity to far below 0.6. Indeed, the average C α r.m.s. deviation of 0.69 Å (Table 4.31) unveils larger structural changes from the wild-type structure. Not only does Thr82 introduce main-chain movements of the 80's loops that in turn affect the cantilevers and flaps of the proteinase dimer, but also the Ile36-Val54 pair significantly affects the 30's loops that in turn disrupt two important salt bridges (Asp35-Arg57 and Glu34-Lys20; Figure 5.2). As a result, the flaps appear more mobile than in any of the preceding variants; the tips of the flaps switch positions with respect to one another (akin the canonical flaps-closed conformation) and the distance between them increases. The k_{cat} value for variant day 56 is similar to wild type (Nijhuis *et al.*, 1999) whereas the K_m increases threefold. In order to increase the viral replication capacity, the virus will now need to acquire additional mutations while maintaining resistance to ritonavir. Intriguingly, one variant (with a total of 11 amino-acid substitutions), isolated from a patient that received multiple drug-treatments, also crystallizes with the flaps extra wide open (Logsdon *et al.*, 2004; Martin *et al.*, 2005). From an evolutionary point of view, it is unclear if this variant is still undergoing evolution or whether it reached a final stage.

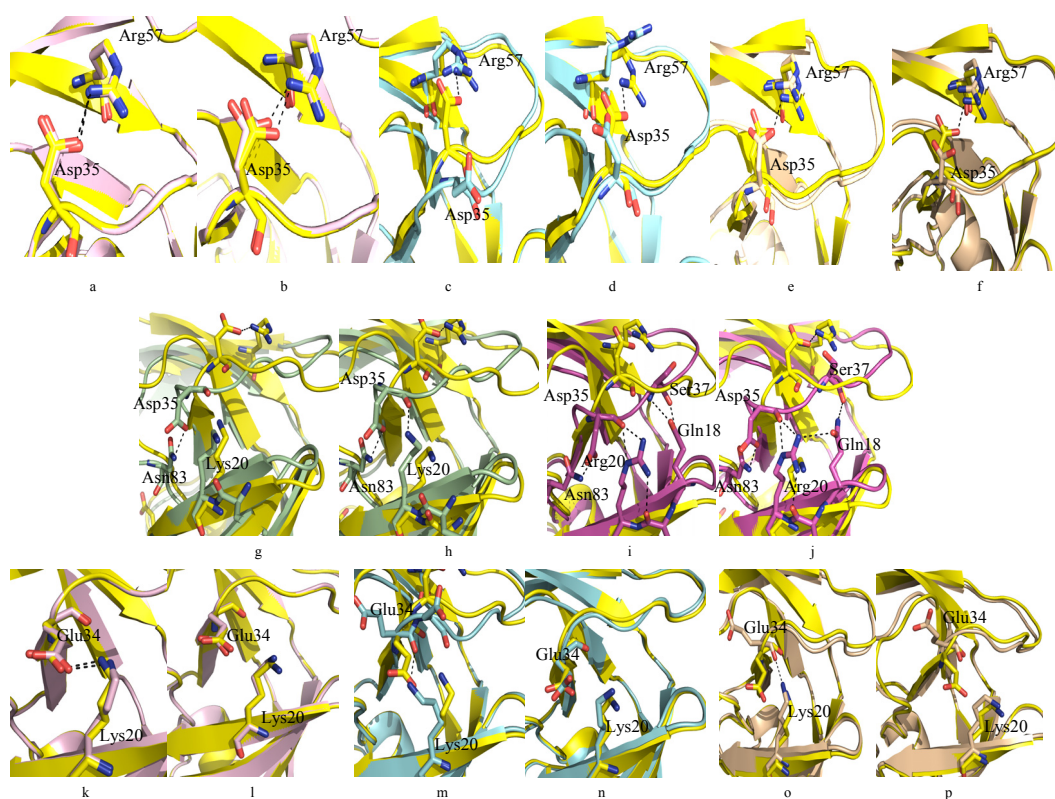


Figure 5.2: Intra-molecular salt bridges between Arg57-Asp35 as well as between Glu34-Lys20 and hydrogen-bonding interactions of day 0 (yellow), day28^a (pink), day 28^b (cyan), day 56 (wheat), day 82 (green) and day 115 (magenta). Panel (a) monomer A, (b) monomer B of day 28^a; (c) monomer A, (d) monomer B of day 28^b; (e) monomer A, (f) monomer B of day 56; (g) monomer A, (h) monomer B of day 82; (i) monomer A, (j) monomer B of day 115; (k) monomer A, (l) monomer B of day 28^a; (m) monomer A, (n) monomer B of day 28^b; (o) monomer A, (p) monomer B of day 56.

Merely one additional substitution, Ala71Val, is able to restore the balance between resistance on the one hand and a good replication capacity on the other (Nijhuis *et al.*, 1999). The K_m value for variant day 82 is similar to wild type whereas the k_{cat} value increases threefold. Thus, the k_{cat}/K_m ratio is three times higher than the wild-type value, which may well explain the better than wild-type relative replication capacity. The substitution Ala71Val is not known to confer resistance on its own (<http://hivdb.stanford.edu/>) but is frequently observed in patients receiving drug treatment. Completely unexpected, the proteinase crystallizes with the flaps closed-in over the active-site as if an inhibitor is present. Hence, the inhibitor-free structure of day 82 is incredibly similar to the structures of the same variant in complex with either ritonavir (overall C α r.m.s. deviation of 0.31 Å) or saquinavir (overall C α r.m.s. deviation of 0.35 Å). The only difference is that the active-site cavity is now filled with several water molecules instead of an inhibitor. For the first time, an amino-acid substitution is acquired that greatly adds to the overall stability of the protein. The

Ala71Val substitution displaces the helix i only slightly but at the same time, tightens the interactions between the helix i (residues 87-93) and β -strands g and h (residues 62-66 and 69-73). Moreover, hydrogen bonds between the 10's and the 30's loops as well as between Asp35 and Asn83 form that are not observed in variant day 0, day28^a, day 28^b and day 56. In this way, long-range structural changes are introduced into the proteinase that stabilize the flaps in the closed conformation. A closed form is also observed for an, inhibitor-free, tethered variant bearing a Cys95Met substitution (Pillai *et al.*, 2000; Kumar *et al.*, 2002). Although it is unclear how a tethering can affect the overall structure, the Cys95Met substitution may exert a similar affect on the helix i as seen for the Ala71Val substitution.

Finally, on day 115, Lys20Arg is acquired in addition to the Met36Ile-Ile54Val-Ala71Val-Val82Thr substitutions. Variants day 82 and day 115 exhibit a similar replication capacity. Once more, the proteinase crystallizes with the flaps closed-in over the active-site as if an inhibitor is present. Consequently, the inhibitor-free structure of day 115 is again similar to the structures of the same variant in complex with either ritonavir (overall C α r.m.s. deviation of 0.24 Å) or saquinavir (overall C α r.m.s. deviation of 0.29 Å). Substitution of lysine to arginine at position 20 further increases the stability of the 30's loop; the carbonyl O of Asp35 makes hydrogen bonds with the Arg20 N ϵ and N η 1 atoms. Also, a strong hydrogen bond is formed between Arg20 N η 2 and the carbonyl O of Leu19.

Antiretroviral therapy may drive the viral population to a low fitness level before arriving at a high fitness level (Nijhuis *et al.*, 1999). This is because the search for both resistance and replication capacity, in sequence space, may require the accumulation of more than one mutation. In the absence of inhibitor, the mutant virus may or may not revert back to wild type. For example, variant day 28^a, day 82 and day 115 do not revert back readily since their RC values are close to wild type (Maarseveen *et al.*, 2006). Interestingly, day 28^b reverts back to wild type (Val54Ile) and so does day 56 (Val54Ile-Thr82Val). Day 56 may also acquire an additional substitution (Ala71Val).

6 CONCLUSIONS

After two decades of intense studies around the world, HIV remains a major therapeutic challenge. Therapy consists of a life-long, laborious treatment because of the ability of HIV to quickly acquire mutations, leading to viruses that are often cross-resistant to various proteinase inhibitors.

The present thesis is concerned with the crystallographic analysis of the molecular evolution of the HIV-1 proteinase (PR) occurring in a patient under ritonavir monotherapy. The results presented herein provide a clear picture of the structural evolution of the proteinase under drug pressure. This picture, of how amino-acid substitutions modulate the overall PR structure, inhibitor binding, and viral fitness, may help to design new, more persistent drugs.

In summary, the structural analysis indicates that:

1. In search of resistance and replication capacity, the HIV PR traverses a complex landscape of possible folds resulting from several consecutive amino-acid substitutions.
2. The substitutions fall into two main groups. First, early substitutions accumulate that increase the K_i value for the drug but which display adverse effect on the structural integrity of the protein, thereby severely reducing the viral replication capacity. In a second stage late substitutions team-up that restore structural integrity and viral replication capacity.
3. Early and late substitutions outside of the active-site cavity manipulate the flaps (e.g., position and mobility) which in turn modulates the subsites.
4. For productive binding to the variant active-site, inhibitors need to adopt a more distorted, *i.e.*, energetically unfavourable, conformation.
5. The HIV proteinase can adopt a flaps-closed conformation in the absence of any ligand. Thus, the central dogma of HIV-1 proteinase needs to be amended.
6. It is not useful to study the effects of late substitutions by themselves since they are specifically selected for by evolution to balance/counteract early mutations. As a result, they will only display their full structural effects in conjunction with preceding substitutions.

Bibliography

Abdel-Rahman, H. M., Al-Karamany, G. S., El-Koussi, N. A., Youssef, A. F. and Kiso, Y. (2002): HIV protease inhibitors: Peptidomimetic drugs and future perspectives. *Curr. Med. Chem.*, **9**: 1905-1922.

Ala, P. J., Huston, E. E. and Klabe, R. M. (1997): Molecular basis of HIV-1 protease drug resistance: structural analysis of mutant proteases complexed with cyclic urea inhibitors. *Biochemistry*, **36**: 1573-1580.

Appelt, K., Bacquet, R. J., Bartlett, C. A., Booth, C. L. and Freer, S. T. (1991): Design of enzyme inhibitors using iterative protein crystallographic analysis. *J. Med. Chem.*, **34**: 1925-1934.

Bacheler, L. T., Paul, M., Otto, M. J., Jadhav, P. K., Stone, B. A. and Miller, J. A. (1994): An assay for HIV RNA in infected cell lysates, and its use for rapid evaluation of antiviral efficacy. *Antiviral Chem. Chemother.*, **5**: 111-121.

Barrie, K. A., Perez, E., Lamers, S. L., Sleasman, J. W., Dunn, B. M. and Goodenow, M. M. (1996): Natural variation in HIV-1 protease, Gag p7 and p6, and protease cleavage sites within Gag/Pol polyproteins: amino acid substitutions in the absence of protease inhibitors in mothers and children infected by human immunodeficiency virus type 1. *Virology*, **219**: 407-416.

Bergfors, T. M. (1999): Chapter 8, Crystallization Strategy. In: Bergfors, T. M. (ed.), *Protein Crystallization: Techniques, Strategies, And Tips. A Laboratory Manual*. International University Line, La Jolla, California.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000): The Protein Data Bank. *Nucl. Acids Res.*, **28**: 235-242.

Birk, A. and Wong, H. C. (2003): HIV-1 protease: mechanism and drug discovery. *Org. Biomol. Chem.*, **1**: 5-14.

Billich, S., Knoop, M. T., Hansen, J., Strop, P., Sedlacek, J., Mertz, R. and Moelling, K. (1988): Synthetic peptides as substrates and inhibitors of human immunodeficiency virus-1 protease. *J. Biol. Chem.*, **263**: 17905-17908.

Billich, A., Hammerschmid, F. and Winkler, G. (1990): Purification, assay and kinetic features of HIV-1 proteinase. *Biol. Chem. Hoppe-Seyler*, **371**: 265-272.

Blow, D. (2000): So do we understand how enzymes work? *Structure*, **8**: 77-81.

Blundell, T. L. and Pearl, L. (1989): A second front against AIDS. *Nature*, **337**: 596-597.

Brünger, A. T. (1993): Assessment of phase accuracy by cross validation. *Acta Crystallogr. D*, **49**: 24-36.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, N., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. and Warren, G. L. (1998): Crystallography and NMR system (CNS): A new software system for macromolecular structure determination, *Acta Crystallogr. D*, **54**: 905-921.

Brynda, J., Rezáčová, P., Fábry, M., Horejsí, M., Stouracová, R., Soucek, M., Hradílek, M., Konvalinka, J. and Sedláček, J. (2004): Inhibitor binding at the protein interface in crystal of a HIV-1 protease complex. *Acta Crystallogr. D*, **60**: 1943-1948.

Budt, K. H., Hansen, J., Knolle, J., Meichsner, C., Ruppert, D., Paessens, A. and Stowasser, B. (1993): HIV protease inhibitor HOE/BAY 793 Part 4: long term treatment in persistently HIV-1 infected cells, *IXth International Conference on AIDS, Berlin*.

Budt, K. H., Peyman, A., Hansen, J., Knolle, J., Meichsner, C., Paessens, A., Ruppert, D. and Stowasser, B. (1995): HIV protease inhibitor HOE/BAY 793, structure-activity relationship in a series of C₂-symmetric diols. *Bioorg. Med. Chem.*, **3**: 559-571.

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

Chatfield, D. C., Eurenus, K. P. and Brooks, B. R. (1998): HIV-1 protease cleavage mechanism: A theoretical investigation based on classical MD simulation and reaction path calculation using a hybrid QM/MM potential. *J. Mol. Struct., (THEOCHEM)* **423**: 79-92.

Cleland, W. W. and Krevoy, M. M. (1994): Low-barrier hydrogen bonds and enzymatic analysis. *Science*, **264**: 1887-1890.

Clemente, J. C., Hemrajani, R., Blum, L. E., Goodenow, M. M. & Dunn, B. (2003). M. Secondary mutations M36I and A71V in the human immunodeficiency virus type 1 protease can provide an advantage for the emergence of the primary mutation D30N. *Biochemistry*, **43**: 15029-15035.

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

Coffin, J. M. (1995): HIV population dynamics *in vivo*: implications for genetics variation, pathogenesis and therapy. *Science*, **267**: 483-489.

Coffin, J. M., Hughes, S. H. and Varmus, H. E. (1997): *Retroviruses*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Cohen, G. E. (1997): ALIGN: a program to superimpose protein coordinates, accounting for insertions and deletions. *J. Appl. Cryst.*, **30**: 1160-1161.

Collaborative Computational Project, Number 4 (1994). *Acta Crystallogr. D.*, **50**: 760-763.

Condra, J. H., Schleif, W. A., Blahy, O. M., Gabryelski, L. J. and Graham, D. J. (1995): *In vivo* emergence of HIV-1 variants resistant to multiple protease inhibitors. *Nature*, **374**: 569-571.

Cígler, P., Millan, K., Rezacova, P., Brynda, J., Otwinowski, Z., Pokorna, J. M., Grüner, B., Lucie, D. M., Mášá, M., Sedláček, J., Bodem, J., Kräusslich, H. G., Král, V. and Konvalinka, J. (2005): From nonpeptide toward noncarbon protease inhibitors: Metallocarboranes as specific and potent inhibitors of HIV protease. *Proc. Natl. Acad. Sci. USA*, **102**: 15394-15399.

Davies, D. R. (1990): The structure and function of the aspartic proteinases. *Annu. Rev. Biophys. Chem.*, **19**: 189-215.

Davis, A. M., Teague, S. J. and Kleywegt, G. J. (2003): Application and limitations of x-ray crystallographic data in structure-based ligand and drug design. *Angew. Chem. Int. Ed.*, **42**: 2718-2736.

Decroly, E., Vandenbranden, M., Ruyschaert, J. M., Cogniaux, J., Jacob, G. S., Howard, S. C., Marshall, G., Kompelli, A., Basak, A., Jean, F., Lazure, C., Benjannet, S., Chrétien, M., Day, R. and Seidah, N. G. (1994): The convertases furin and PC1 can both cleave the human immunodeficiency virus(HIV)-1 envelope glycoprotein gp160 into gp120 (HIV-I SU) and gp41 (HIV-I TM). *J. Biol. Chem.*, **269**: 12240-12247.

Drenth, J. (1994): Chapter 8, Phase Improvement. *Principles of Protein X-ray Crystallography*. Springer, New York.

Dreyer, G. B., Boehm, J. C., Chenera, B., DesJarlais, R. L., Hassell, A. M., Thomas, D., Tomaszek, T. A. and Lewis, M. (1993): A symmetric inhibitor binds HIV-1 protease asymmetrically. *Biochemistry*, **2**: 937-947.

Dunn, B. M., Gustchina, A., Wlodawer, J. and Kay, J. (1994): Subsite preferences of retroviral proteinases. *Methods Enzymol.*, **241**: 254-278.

Eigen, M. (1993): Viral quasispecies. *Sci. Am.*, **269**: 42-49.

Engh, R. A. and Huber, R. (1991): Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr. A*, **47**: 392-400.

Erickson, J. W. and Burt, S. K. (1996): Structural mechanisms of HIV drug resistance. *Annu. Rev. Pharmacol. Toxicol.*, **36**: 545-571.

Erickson-Viitanen, S., Manfredi, J., Viitanen, P., Tribe, D. E., Tritch, R., Hutchison, C. A., Loeb, D. D. and Swanstrom, R. (1989): Cleavage of HIV-1 gag polyprotein synthesized *in vitro* - sequential cleavage by the viral protease. *Aids Res. Hum. Retroviruses*, **5**: 577-591.

Erickson, J. W. (1993): Design and structure of symmetry-based inhibitors of HIV-1 protease. In: Anderson, P. S., Kenyon, G. L. and Marshall, G. R (eds.), *Perspectives in Drug Discovery and Design*. The Netherlands: ESCOM Science.

Evans, P. R. (1997): *Proceedings of the CCP4 Study Weekend. Recent Advances in Phasing*. In: Wilson, K. S., Davies, G., Ashton, A. W. and Bailey, S. (eds.), Warrington: Daresbury Laboratory.

Eyermann, C. J., Jadhav, P. K., Hodge, C. N., Chang, C. H., Rodgers, J. D. and Lam, P. Y. S. (1997): The role of computer-aided and structure-based design techniques in the discovery and optimization of cyclic urea inhibitors of HIV protease. *Adv. Amino Acid Mimetics Peptidomimetics*, **1**: 1-40.

Fischer, M. (2004): Kristallographische Analyse der Evolution von HIV-PR Mutanten während suboptimaler Ritonavir-Monotherapie. *Bachelor Arbeit*, Institut für Biochemie der Universität zu Lübeck.

Flexner, C. (1998): HIV-protease inhibitors. *New Engl. J. Med.*, **338**: 1281-1293.

French, G. S. and Wilson, K. S. (1978): Crystal physics, diffraction, theoretical and general crystallography. *Acta Crystallogr. A*, **34**: 517-525.

Frey, P. A., Whitt, S. A. and Tobin, J. B. (1994). A low-barrier hydrogen bond in the catalytic triad of serine proteases. *Science*, **264**: 1927-1930.

Fruton, J. S. (1987): Aspartyl proteinase. In: Neuberger, A. and Brocklehurst, K. (eds.), *New Comprehensive Biochemistry Vol. 16, Hydrolytic Enzymes*. Elsevier, Amsterdam.

Gerlt, J. A. and Gassman, P. J. (1993): Understanding the rates of certain enzyme-catalyzed reactions: Proton abstraction from carbon acids, acyl-transfer reactions, and displacement reactions of phosphodiesteres. *Biochemistry*, **32**: 11943-11952.

Grant, S. K., Deckman, I. C., Culp, J. S., Minnich, M. D., Brooks, I. S., Hensley, P., Debouck, C. and Meek, T. D. (1992): Use of protein unfolding studies to determine the conformational and dimeric stabilities of HIV-1 and SIV proteases. *Biochemistry*, **31**: 9491-9501

Goodman, M. and Ro, S. (1995): *Burger's Medicinal Chemistry and Drug Discovery, Vol. 1. Principles and Practice*. In: Wolf, M. E. (ed.), John Wiley and Sons, New York, USA.

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

Hillisch, A. and Hilgenfeld, R. (2002): The role of protein 3D-structures in the drug discovery process. In: Hillisch, A. and Hilgenfeld, R. (eds.), *Modern Methods of*

Drug Discovery. Birkhäuser Verlag, Basel, Switzerland.

Hodge, C. N., Aldrich, P. E., Bacheler, L. T., Chang, C. H., Eyermann, C. J., Garber, S., Grubb, M., Jackson, D. A., Jadhav, P. K., Korant, B., Lam P. Y. S., Maurin, M. B., Meek, J. L., Otto, M. J., Rayner, M. M., Reid, C., Sharpe, T. R., Shum, L., Winslow, D. L. and Erickson-Viitanen, S. (1996): Improved cyclic urea inhibitors of the HIV-1 protease: synthesis, potency, resistance, profile, human pharmacokinetics and X-ray crystal structure of DMP 450. *Chem. Biol.*, **3**: 301-314.

Hoffman, N. G., Schiffer, C. A. and Swanstrom, R. (2003): Covariation of amino acid positions in HIV-1 protease inhibitors, *J. Biol. Chem.*, **277**: 5952-5961.

Hong, L., Zhang, X. C., Hartsuck, J. A. and Tang, J. (2000): Crystal structure of an *in vivo* HIV-1 protease mutant in complex with saquinavir: Insights into the mechanisms of drug resistance. *Protein Sci.*, **9**: 1898-1904.

Hsu I-N., Delbaere, L. T. J., James, M. N. G. and Hoffman, T. (1977): Penicillopepsin: 2.8 Å structure, active site conformation and mechanistic implications. In: Tang, J. (ed.), *Acid proteases, structure, function, and biology*. Plenum Press, New York, USA.

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

Hyland, L. J., Tomaszek Jr, T. A., Roberts, G. D., Carr, S. A., Magaard, V. W., Bryan, H. L., Fakhoury, S. A., Moore, M. L. and Minnich, M. D. (1991b): Human immunodeficiency virus-1 protease. 1. Initial velocity studies and kinetic characterization of reaction intermediates by ¹⁸O isotope exchange. *Biochemistry*, **30**: 8441-8453.

Ido, E., Han, H. P., Kedzy, F. J. and Tang, J. (1991): Kinetics studies of human immunodeficiency virus type I protease and its active-site hydrogen bond mutant A28S. *J. Biol. Chem.*, **266**: 24359-24366.

Ingr, M., Uhlikova, T., Strisovsky, K., Majerova, E. and Konvalinka, J. (2003): Kinetics of the dimerization of retroviral proteases: The “fireman's grip” and dimerization. *Protein Sci.*, **12**: 2173-2182.

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

Jacobsen, H., Hanggi, M., Ott, M., Duncan, I. B., Owen, S. and Andreoni, M. (1996): *In vivo* resistance to a human immunodeficiency virus type I proteinase inhibitor: mutations, kinetics, frequencies. *J. Infect. Dis.*, **173**: 1379-1387.

Jaskolski, M., Gilski, M., Dauter, Z. and Wlodawer, A. (2007): Stereochemical restraints revisited: how accurate are refinement targets and how

much should protein structures be allowed to deviate from them? *Acta Crystallogr. D.*, **63**: 611-620.

Jencks, W. P. (1997): From chemistry to biochemistry to catalysis to movement. *Annu. Rev. Biochemistry*, **66**: 1-18.

Johnson, V. A., Brun-Vézinet, F., Clotet, B., Conway, B., D' Aquila, R. T., Demeter, L. M., Kuritzkes, D. R., Pillay, D., Schapiro, J. M., Telenti, A. and Richman, D. D. (2003): Drug resistance mutations in HIV-1. *Top. HIV Med.*, **11**: 215-221.

Kabsch, W. (1976): A solution for the best rotation to relate two sets of vectors. *Acta Crystallogr. A*, **32**: 922-923.

Katoh, I., Yasunaga, T., Ikawa, Y. and Yoshinka, Y. (1987): Inhibition of retroviral protease activity by an aspartyl proteinase inhibitor. *Nature*, **329**: 654-656.

Kellis, J. T., Nyberg, K., Sali, D. & Fersht, A. R. (1988).Contribution of hydrophobic interactions to protein stability. *Nature*, **333**: 784-786.

Kempf, D. J., Marsh, K. C., Paul, D. A., Knigge, M. F. and Norbeck, D. W. (1991): Antiviral and pharmacokinetic properties of C₂-symmetric inhibitors of the human immunodeficiency virus type 1 protease. *Antimicrob. Agents Chemother.*, **35**: 2209-2214.

Kempf, D. J., Codacovi, L., Wang, X. C., Kohlbrenner, W. E., Wideburg, N. E., Saldivar, A., Vasavanonda, S., Marsh, K. C., Brzant, P., Sham, H. L., Green, B. E., Betebenner, D. A., Erickson, J. and Norbeck, D. (1993): Symmetry-based inhibitors of HIV protease. Structure-activity studies of acylated 2,4-diamino-1,5-diphenyl-3-hydroxypentane and 2,5-diamino-1,6-diphenylhexane-3,4-diol. *J. Med. Chem.*, **36**: 320-330.

Kempf, D. J., Marsh, K. C., Denissen, J. F., McDonald, E., Vasavanonda, S., Flentge, C. A., Green, B. E., Fino, L., Park, C. H. and Kong, X. P. (1995): ABT-538 is a potent inhibitor of human immunodeficiency virus protease and has high oral bioavailability in humans. *Proc. Natl. Acad. Sci. USA*, **92**: 2484-2488.

Kervinen, J., Lubskowski, J., Zdanov, A., Bhatt, D., Dunn, B. M., Hui, K. Y., Powell, D. J., Kay, A., Wlodawer, A. and Gustchina, A. (1998): Toward a universal inhibitor of retroviral proteases: comparative analysis of the interactions of LP-130 complexed with proteases from HIV-1, FIV and EIAV. *Protein Sci.*, **7**: 2314-2323.

King, R. W., Garber, S., Winslow, D. L., Reid, C., Bacheler, L. T., Anton, E. and Otto, M. J., (1995): Multiple mutations in the human immunodeficiency virus protease gene are responsible for decreased susceptibility to protease inhibitors. *Antiviral Chem. Chemother.*, **6**: 80-88.

Kiso, Y. (1996): Design and synthesis of substrate-based peptidomimetic human immunodeficiency virus protease inhibitors containing the hydroxymethylcarbonyl isostere. *Biopolymers (Peptide Science)*, **40**: 235-244.

Kräusslich, H. G., Ingraham, F. H., Skoog, M., Wimmer, E., Pallai, P. V. and Carter, C. A. (1989): Activity of purified biosynthetic proteinase of human immunodeficiency virus on natural substrates and synthetic peptides. *Proc. Natl. Acad. Sci. USA*, **86**: 807-811.

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

Kumar, M., Kannan, K. K., Hosur, M. V., Bhavesh, N. S., Chatterjee, A., Mittal, R. And Hosur, R. V. (2002): Effects of remote mutation on the autolysis of HIV-PR: X-ray and NMR investigations. *Biochem. Biophys. Res. Commun.*, **294**: 395-401.

Lam, P. Y. S., Jadhav, P. K., Eyermann, C. J., Hodge, C. N., Ru, Y., Ru, Y., Bacheler, L. T., Meek, J. L., Otto, M. J., Rayner, M. M., Wong, Y. N., Chang, C. H., Weber, P. C., Jackson, D. A., Sharpe, T. R. and Viitanen-Erickson, S. (1994): Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors. *Science*, **263**: 380-384.

Lam, P. Y. S., Ru, Y., Jadhav, P. K., Aldrich, P. E., DeLucca, G. V., Eyermann, C. J., Chang, C. H., Emmet, G., Holler, E. R., Danekar, W. F., Li, L., Confalone, P. N., McHugh, R. J., Han, Q., Li, R., Markwalder, J. A., Seitz, S. P., Sharpe, T. R., Bacheler, L. T., Rayner, M. M., Klabe, R. M., Shum, L., Winslow, D. L., Kornhauser, D. M., Jackson, D. A., Erickson-Viitanen, S. and Hodge, C. N. (1996): Cyclic HIV protease inhibitors: Synthesis, conformational analysis, P2/P2' structure-activity relationship, and molecular recognition of cyclic ureas. *J. Med. Chem.*, **39**: 3514-3535.

Lange-Savage, G., Berchtold, H., Liesum, A., Budt, K. H., Peyman, A., Knolle, J., Sedlacek, J., Fabry, M. and Hilgenfeld, R. (1997): Structure of HOE/BAY 793 complexed to human immunodeficiency virus (HIV-1) protease in two different crystal forms. Structure/function relationship and influence of crystal packing. *Eur. J. Biochemistry*, **248**: 313-322.

Lapatto, R., Blundell, T., Hemmings, A., Overington, J., Wilderspin, A., Wood, S., Merson, J. R., Whittle, P. J. and Geoghegan, K. F. (1989): X-ray analysis of HIV-1 proteinase at 2.7 Å resolution confirms structural homology among retroviral enzymes. *Nature*, **342**: 299-302.

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

Laskowki, R. A., (1995): SURFNET: A program for visualizing molecular surfaces, cavities and intermolecular interactions. *J. Mol. Graph.*, **13**: 323-330.

Lee, H., Darden, T. A. and Pedersen, L. G. (1996): An ab initio quantum mechanical model for the catalytic mechanism of HIV-1 protease. *J. Am. Chem. Soc.*, **118**: 3946-3950.

Leslie, A. G. W. (1999): Mosflm: Integration of macromolecular diffraction data. *Acta Crystallogr. D*, **55**: 1696-1702.

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

Luzzati, P. V. (1952): Traitement Statistique des Erreurs dans la Determination des Structures Cristallines, *Acta Crystallogr.*, **5**: 802-810.

Maarseveen, V. N. M., de Jong, D., Boucher, C. A. & Nijhuis, M. (2206). An increase in viral replicative capacity drives the evolution of protease inhibitor resistant human immunodeficiency virus type 1 in absence of drugs. *J. AIDS*, **42**: 162-168

Markland, W., Rao, B. G., Parsons, J. D., Black, J., Zuchowski, L., Tisdale, M. and Tung, R. (2000): Structural and kinetic analyses of the protease from an amprenavir-resistant to saquinavir and resensitized to amprenavir. *J. Virol.*, **74**: 7636-7641.

Markowitz, M., Conant, M., Hurley, A., Schluger, R., Duran, M. and Peterkin, J. (1998): A preliminary evaluation of nelfinavir mesylate, an inhibitor of human immunodeficiency virus (HIV)-1 protease, to treat HIV infection. *J. Infect. Dis.*, **177**: 1533-1540.

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

Matsumura, M., Becktel, W. J. & Matthews, B. W. (1988). Hydrophobic stabilization in T4 lysozyme determined directly by multiple substitutions of Ile3. *Nature*, **334**: 406-410.

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

McRee, D. E. (1999): XtalView/Xfit-A Versatile Program for Manipulating Atomic Coordinates and Electron Density. *J. Struct. Biol.*, **125**: 156-165.

Molla, A., Korneyeva, M., Gao, Q., Vasavanonda, S., Shipper, P. J. and Mo, H. M. (1996): Ordered accumulation of mutations in HIV protease confers resistance to ritonavir. *Nature Med.*, **2**: 760-766.

Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K. and Olson, A. J. (1998): Automated docking using a lamarckian genetic algorithm and empirical binding free energy function. *J. Computational Chemistry*, **19**: 1639-1662.

Murshudov, G. N., Vagin, A. A. and Dodson, E. J. (1997): Refinement of

macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D*, **53**: 240-255.

Navia, M. A., Fitzgerald, P. M., McKeever, B. M., Leu, C. T. and Heimbach, J. C. (1989): Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. *Nature*, **337**: 615-620.

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

Northrop, D. B. (2001): Follow the protons: A low-barrier hydrogen bond unifies the mechanisms of the aspartic proteinases. *Acc. Chem. Res.*, **34**: 790-797.

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

Park, H., Suh, J. and Lee, S. (2000): Ab initio studies on the catalytic mechanism of aspartic proteinases: Nucleophilic versus general acid/general base mechanism. *J. Am. Chem. Soc.*, **122**: 3901-3908.

Patick, A. K., Duran, M., Cao, Y., Shugarts, D., Keller, M. R. and Mazabel, E. (1998): Genotypic analysis of HIV-1 variants isolated from patients treated with the protease inhibitor, nelfinavir. *Antimicrob. Agents Chemother.*, **42**: 2367-2644.

Pearl, L. H. and Taylor, W. R. (1987): Sequence specificity of retroviral proteases. *Nature*, **328**: 428.

Pereira, J. M. A., Costa, Q. S. and Pereira, J. M. (2005): HIV-2 infection and chemokine receptors usage – clues to reduced virulence of HIV-2. *Curr. HIV Res.*, **3**: 3-16.

Peyman, A., Budt, K. H., Spanig, J. and Ruppert, D. (1993): C₂ symmetrical inhibitors of the HIV-1 protease based on phosphinic acid. *Angew. Chem. Int. Ed. Engl.*, **32**: 1720-1722.

Piana, S. and Carloni, P. (2000). Conformational flexibility of the catalytic aspartate dyad in HIV-1 protease: An ab initio study on the free enzyme. *Proteins Struct. Funct. Genet.*, **39**: 26-36.

Piana, S., Carloni, P. and Parrinello, M. (2002): Role of conformational fluctuations in the enzymatic reaction of HIV-1 protease. *J. Mol. Biol.*, **319**: 567-583.

Pillai, B., Kannan, K. K. and Honsur, R. V. (2000): 1.9 Å x-ray study shows closed flap conformation in crystal of tethered HIV-PR. *Proteins*, **43**: 57-64.

Polgar, L., Szeltner, Z. and Boros, I. (1994): Substrate-dependent mechanisms in the catalysis of human immunodeficiency virus protease. *Biochemistry*, **33**: 9351-9357.

Preston, B. D., Poiesz, B. J. and Loeb, L. A. (1988): Fidelity of HIV-1 reverse transcriptase. *Science*, **242**: 1171-1173.

Read, R. J. (1986): Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr. A*, **42**: 140-149.

Rezacova P., Brynda J., Lescar J., Fabry M., Horejsi M., Sieglöva I., Sedláček J. and Bentley G. A. (2005): Crystal structure of a cross-reaction complex between an anti-HIV-1 protease antibody and an HIV-2 protease peptide. *J. Struct. Biol.*, **149**: 332-337.

Rodriguez, E. J., Angeles, T. S. and Meek, T. D. (1993): Use of nitrogen-15 kinetic isotope effects to elucidate details of the chemical mechanism of human immunodeficiency virus 1 protease. *Biochemistry*, **32**: 12380-12385.

Rodgers, J. D., Lam, P. Y. S., Johnson, B. L., Wang, H., Li, R., Ru, Y., Ko, S. S., Seitz, S. P., Trainor, G. L., Anderson, P. S., Klabe, R. M., Bacheler, L. T., Cordova, B., Garber, S., Reid, C., Wright, M. R., Chang, C. H. and Erickson-Viitanen, S. (1998): Design and selection of DMP 850 and DMP 851: the next generation of cyclic urea HIV protease inhibitors. *Chem. Biol.*, **5**: 597-608.

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.

Rose, R. E., Gong, Y. F., Greytok, J. A., Bechtold, C. M., Terry, B. J. and Robinson, B. S. (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.

Rusconi, S and Viganò, O. (2006): New HIV protease inhibitors for drug-resistant viruses. *Therapy*, **3**: 79-88.

Saridakis, E. and Chayen N. E. (2000): Improving protein crystal quality by decoupling nucleation and growth in vapor diffusion. *Protein Sci.*, **9**: 755-757.

Schechter, I. and Berger, A. (1967): On the size of the active site in proteases, I: papain. *Biochem. Biophys. Res. Commun.*, **27**: 157-162.

Schramm, H. J., Billich, A., Jaeger, E., Rucknagel, K. P. Arnold, G. and Schramm, W. (1993): The inhibition of HIV-1 protease by interface peptides. *Biochem. Biophys. Res. Commun.*, **194**: 595-600.

Shortle, D., Stites, W. E. & Meeker, A. K. (1990). Contributions of the large hydrophobic amino acids to the stability of proteins and an experimental analysis of the evolution of protein stability. *Biochemistry*, **29**: 8033-8041.

Sham, H. L., Zhao, C., Stewart, K. D., Betebenner, D. A., Lin, S., Park, C. H., Kong, X. P., Rosenbrook, W. Jr., Herrin, T., Madigan, D., Vasavanonda, S., Lyons,

N., Molla, A., Saldivar, A., Marsh, K. C., McDonald, E., Wideburg, N. E., Denissen, J. F., Robins, T., Kempf, D. J., Plattner, J. J. and Norbeck, D. W. (1996): A novel picomolar inhibitor of human immunodeficiency virus type 1 protease. *J. Med. Chem.*, **39**: 392-397.

Silva, A. M., Cachau, R. E., Sham, H. L. and Erickson, J. W. (1996): Inhibition and catalytic mechanism of HIV-1 aspartic protease. *J. Mol. Biol.*, **255**: 321-346.

Spinelli, S., Liu, Q. Z., Alzari, P. M., Hirel, P. H. and Poljak, R. J. (1991): The three-dimensional structure of the aspartyl protease from the HIV-1 isolate BRU. *Biochimie*, **73**: 1391-1396.

Spire, B., Duran, S., Souville, M., Leport, C., Raffi, F. and Moatti J., P. (2002): Adherence to highly active antiretroviral therapies (HAART) in HIV-infected patients: from a predictive to a dynamic approach. *Social Science and Medicine*, **54**: 1481-1496.

Stefano, P., Carloni, P. and Rothlisberger, U. (2002): Drug resistance in HIV-1 protease: Flexibility-assisted mechanism of compensatory mutations. *Protein Sci.*, **9**: 1898-1904.

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

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

Swanstrom, R. and Wills, J. W. (1997): Retroviral gene expression. II. Synthesis, processing, and assembly of viral proteins, p. 263-334. In: Coffin, J. M., Hughes, S. H. and Varmus H. E. (eds.), *Retroviruses*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Thanki, N., Mohana Rao, J. K., Foundling, S. I., Howe, W. J., Moon, J. B., Hui, J. O., Tamasseli, A. G., Heinrikson, R. L., Thaisrivongs, S. and Wlodawer, A. (1992): Crystal structure of a complex of HIV-1 protease with a dihydroxyethylene-containing inhibitor: comparisons with molecular modeling. *Protein Sci.*, **1**: 1061-1072.

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

Toh, H., Ono, M. and Miyata, T. (1985): Retroviral gag and DNA endonuclease coding sequence in IgE-binding factor gene. *Nature*, **318**: 245-248.

Towler, E. M., Thompson, S. K., Tomaszek, T. and Debouck, C. (1997): Identification of a loop outside cavity of the human immunodeficiency virus proteases

which confers inhibitor specificity. *Biochemistry*, **36**: 5128-5133.

Tözsér, J., Weber, I. T., Gustchina, A., Blaha, I., Copeland, T. D., Louis, J. M. and Oroszlan, S. (1992): Kinetic and modeling studies of S3-S3' subsites of HIV proteinases. *Biochemistry*, **31**: 4793-4800.

Trylska, J., Grochowski, P. and McCammon, J. A. (2004): The role of hydrogen bonding in the enzymatic reaction catalyzed by HIV-1 protease. *Protein Sci.*, **13**: 513-528.

Turner, S. R. (2002): HIV protease inhibitors – the next generation. *Curr. Med. Chem. - Anti-Infective Agents*, **1**: 141-162.

Vagin, A. and Teplyakov, A. (1997): MOLREP: an automated program for molecular replacement. *J. Appl. Crystallogr.*, **30**: 1022-1025.

Vaguine, A. A., Richelle, J. and Wodak, S. J. (1999): SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallogr. D*, **55**: 191-205.

Weber, J., Mesters, J. R., Lepsik, M., Prejdova, J., Svec, M., Sponarova, J., Mlcochova, P., Skalicka, K., Strisovsky, K., Uhlikova, T., Soucek, M., Machala, L., Stankova, M., Vondrasek, J., Klimkait, T., Kraeusslich, H. G., Hilgenfeld, R. and Konvalinka, J. (2002): Unusual binding mode of an HIV-1 protease inhibitor explains its potency against multi-drug-resistant virus strains. *J. Mol. Biol.*, **324**: 739-754.

Wei, X., Ghosh, S. K., Taylor, M. E., Johnson, V. A., Emini, E. A., Deutsch, P., Lifson, J. D., Bonhoeffer, S. and Nowak, M. A. (1995): Viral dynamics in human immunodeficiency virus type 1 infection. *Nature*, **373**: 117-122.

Wilderspin, A. F. and Sugrue, R. J. (1994): Alternative native flap conformation revealed by 2.3 Å resolution structure of SIV proteinase. *J. Mol. Biol.*, **239**: 97-103.

Wilson, A. J. C. (1942): *Nature*, **150**: 152-155.

Wlodawer, A., Miller, M., Jaskolski, M., Sathyanarayana, B. K. and Baldwin, E. (1989): Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease. *Science*, **245**: 616-621.

Wlodawer, A. and Erickson, J. W. (1993): Structure-based inhibitors of HIV-1 protease. *Annu. Rev. Biochem.*, **62**: 543-585.

Wlodawer, A., Gustchina, A., Reshetnikova, L., Lubkowski, J., Zdanov, A., Hui, K. Y., Angleton, E. L., Farmerie, W. G., Goodenow, M. M., Bhatt, D., Zhang, L. and Dunn, B. M. (1995): Structure of an inhibitor complex of proteinase from feline immunodeficiency virus. *Nat. Struct. Biol.*, **2**: 480-488.

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

27: 249-284.

Wlodawer, A. and Gustchina, A. (2000): Structural and biochemical studies of retroviral proteases. *Biochim. Biophys. Acta*, **1477**: 16-34.

Wyatt, R. and Sodroski, J. (1998): The HIV-1 envelope glycoproteins: fusogens, antigens and immunogens. *Science*, **280**: 1884-1888.

Xie, D., Gulnik, S., Gustchina, E., Yu, B., Shao, W., Qoronfleh, W., Nathan, A. and Erickson, J. W. (1999): Drug resistance mutations can effect dimer stability of HIV-1 protease at neutral pH. *Protein Sci.*, **8**: 1702-1707.

Xu, R., Andrews, W., Spaltenstein, A., Danger, D., Dallas, W., Carter, I., Hanlon, M., Wright, L. and Furfine, E. (2002): Molecular mechanism of I50V HIV-1 protease resistance and cross-resistance to protease inhibitors. 9th Conference on Retroviruses and Opportunistic Infections. Seattle, WA, USA.

Zhang, Y. M., Imamichi, H., Imamichi, T., Lane, H. C., Palloon, J., Vasudevachari, M. B. and Salzman, N. P. (1997): Drug resistance during indinavir therapy is caused by mutations in the protease gene and its Gag substrate cleavage sites. *J. Virol.*, **71**: 6662-6670.

Curriculum vitae

Aida Baharuddin, born 09.10.1974 in Taiping, Malaysia

School

1981-1986	Primary school from Convent Aulong Primary School, Taiping, Malaysia.
1987-1991	Higher secondary school from Darul Ridwan Secondary School, Taiping and MARA Junior Science College, Taiping, Malaysia.

College/University

1992-1993	Matriculation (equivalent to A-Level) from MARA Kulim Matriculation Center, Kedah and University Science Malaysia, Penang Island.
1994-1998	B.Sc (Biotechnology) from University Science Malaysia, Penang Island.
08/1998-12/1998	Student Assistant at School of Biological Sciences, University Science Malaysia, Penang Island.
01/1999-05/2001	M.Sc in Molecular Biotechnology from University Science Malaysia, Penang Island. M.Sc. dissertation project: Cloning and Characterization of polyhydroxyalkanoate synthase gene from <i>Pseudomonas aeruginosa</i>
06/2001-05/2002	Research Assistant at School of Biological Sciences, University Science Malaysia, Penang Island. Project Title: Isolation and Characterization of the Oil Palm (<i>Elaeis guineensis</i>) b-ketothiolase (Acetoacetyl-CoA Acyltransferase) Regulatory Region.
06/ 2002	Started Ph.d work in Structural Biochemistry with Prof. Rolf Hilgenfeld at Institute of Biotechnology, Jena and Institute of Biochemistry, University of Lübeck, Germany.

Scientific conferences and workshops

1. Participated in 1st European Conference on Computational Biology 2002. 6-9 Oct. 2002. Saarbrücken, Germany.
Poster Presentation: Structural Infectomics: Identification and Characterization of Potential Virulence Factors in *Legionella pneumophila*
2. Participated in 6th Heart of Europe Biocrystallography meeting. 25-27 Sept. 2003. Wittemberg, Germany. Oral presentation: Crystallization of MIP protein from *Chlamydiae pneumoniae*
3. First Baltic Sea BioCrystallography Meeting. 2-4 Sept. 2004. Lübeck, Germany.
Oral presentation: Cloning, expression and crystallization of macrophage infectivity potentiator (MIP) protein from *Chlamydiae pneumoniae* and *Chlamydiae trachomatis*.
4. Participated in DGK/DGKK (Germany Crystallography Conference) Pre-Conference Crystallization Course. 11-14 March, 2004. Jena, Germany.
5. Participated in International Conference on SARS one year after the (first) outbreak. 8-11 May, 2004. Lübeck, Germany.
6. Participated in ICCBM10 (International Conference on the Crystallization of Biological Macromolecules). 5-8 June, 2004. Beijing, China.
7. Participated in FEBS (Federation of European Biochemical Society) Advanced Course: Advanced methods in Macromolecules Crystallization. 2-9 Oct. 2004. Nove Hradý, Czech Republic.
8. Participated in CCP4 (Collaborative Computing Project Number 4 in Protein Crystallography) Study weekend 7-8 Jan. 2005. Data Collection and Analysis, University of Reading, UK.
9. Participated in IUCR (XX Congress of the International Union of Crystallography). October 2005. Florence, Italy.
10. Participated in CCP4 (Collaborative Computing Project Number 4 in Protein Crystallography) Study weekend. 5-7 Jan. 2006. Crystallography of Complexes, University of Leeds, UK.
11. International Workshop on Discovery of antiviral compounds (VIZIER). 26-29 April, 2006. Lübeck, Germany.
12. 8th International School on the Crystallography of Biological Macromolecules, 21-25 May, 2006. Como, Italy.