6-12-2006

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

Yunfeng Tie

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

by

YUNFENG TIE

Under the Direction of Irene T. Weber

ABSTRACT

HIV-1 protease is the most effective target for drugs to treat AIDS, however, the long-term therapeutic efficiency is restricted by the rapid development of drug resistant variants. To better understand the molecular basis of drug resistance, crystallographic and kinetic studies were applied to wild-type HIV-1 protease (PR) and drug-resistant mutants, PR_{V82A}, and PR_{I84V}, in complex with substrate analogs, the current drug saquinavir and the new inhibitor UIC-94017 (TMC-114). UIC-94017 was also studied with mutants PR_{D30N} and PR_{I50V}. The drug-resistant mutations V82A, I84V, D30N and I50V participate in substrate binding. Eighteen crystal structures were refined at resolutions of 0.97-1.60Å. The high accuracy of the atomic resolution crystal structures helps understand the reaction mechanism of HIV-1 PR.

Different binding modes are observed for different types of inhibitors. The substrate analogs have more extended interactions with PR subsites up to S5-S5', while the clinical inhibitors maximize the contacts within S2-S2'. Hydrophobic interactions are the major force for saquinavir binding since it was designed with enhanced hydrophobic groups based on substrate side-chains. In contrast, the new clinical inhibitor UIC-94017

was designed to mimic the hydrogen bonds between substrates and PR. UIC-94017 forms polar interactions with the PR main-chain atoms of Asp29/30, which have been proposed to be critical for its potency against resistant HIV.

The mutants showed different structural and kinetic effects, depending on the inhibitor and location of the mutations. The observed structural changes were consistent with the relative inhibition data. Both PR_{I84V} and PR_{I50V} lost favorable hydrophobic interactions with inhibitor compared with PR. Similarly, in PR_{D30N} the UIC-94017 had a water-mediated interaction with the side-chain of Asn30 rather than the direct interaction observed in PR. However, PR_{V82A} compensated for the mutation by shifts of the backbone of Ala82. Furthermore, the complexes of PR_{V82A} showed smaller shifts relative to PR, but more movement of the peptide analog, compared to complexes with clinical inhibitors. The structures suggest that substrate analogs have more flexibility than the drugs to accommodate the structural changes caused by mutation, which may explain how HIV can develop drug resistance while retaining the ability of PR to hydrolyze natural substrates.

INDEX WORDS: HIV-1 protease, drug resistance, V82A, I84V, D30N, I50V, substrate analog, UIC-94017, TMC-114, saquinavir, crystal structure, kinetics.

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

by

YUNFENG TIE

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

Doctor of Philosophy

College of Arts and Sciences

Georgia State University

2006

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

by

YUNFENG TIE

Major Professor: Irene T. Weber
Committee: Stuart A. Allison
Giovanni Gadda
Robert W. Harrison

Electronic Version Approved:

Office of Graduate Studies
College of Arts and Sciences
Georgia State University
May 2006

ACKNOWLEDGEMENTS

I would like to first thank my great advisor Dr. Irene T. Weber for all of her patience, encouragement, and guidance. You have been so supportive throughout my degree period. You spent much time not only correcting my writing, discussing the science but also giving me suggestions and directions for my career. I have learned so much from you, which can be of benefit for my career as well as for my whole life. Thank you, Dr. Harrison! You are so open-minded and have encouraged me so much with your enthusiasm to science and brilliant research ideas. I am also very thankful to my other two committee members, Dr. Gadda and Dr. Allison, for your interest, advice and critical review of my dissertation.

Thank you, all my labmates, Peter, Andrey, Patra, Yuanfang, John, Bhuv, Augustin, Ping, Lynn, Bin, Jeff, Anna, Johnson and Sasha. All of you make the lab such an exciting place and I have had a great time with all of you. Patra, you always listen to me, share with me all the happiness and sadness and have helped me so much with my English. Yuanfang taught me all the basics of solving structures and kept all computers working properly. It really helped me saving a lot of time. Great thanks to Peter, he taught me everything from scratch, trained me, a chemistry student to enjoy protein studies. He had also given me lots of help on kinetics studies and advised on almost everything I wrote. John, Bhuv and Augustin, thank you for teaching me all the basic lab techniques for crystallography, from setting up trays, to seeding and crystal mounting. Being a new member of our lab, Andrey has brought brand new ideas from small molecule crystallography to improve my understanding of protein crystallography. Lynn,

Ping and Bin, my Chinese friends, thank you so much for all the support during these years.

The support from my family, my parents, sister, and husband are greatly appreciated. Hao, my husband, has helped me so much on everything, particularly for the programming and debugging. I am so thankful to you for not only being the greatest life-partner, but also for being the best friend of mine.

Finally, I would also like to thank Dr. Baumstark and Dr. Dixon for admitting me into Department of Chemistry program and providing me this precious opportunity to complete this degree, especially for giving me the freedom to choose the research area I am interested in.

Best wishes to all the people! I just hope one day I will be able to help you as much as you all have helped me!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Å	angstrom
ALA	alanine
AMMP	another molecular mechanics/modeling program
ARG	arginine
AS	ammonium sulfate
ASP	aspartic acid
ASN	asparagine
C	carbon
CYS	cysteine
DMSO	dimethylsulfoxide
GEL	1-O-octyl-2-heptylphosphonyl-SN-glycero-3-phosphoethanolamine
GLU	glutamic acid
GLN	glutamine
GLY	glycine
HIS	histidine
HIV-1	human immunodeficiency virus 1
ILE	isoleucine
IPTG	isopropyl-D-thio-galactopyranoside
LB	Luria-Bertani
LEU	leucine
LYS	lysine

M	molar
MET	methionine
MR	molecular replacement
Na^+	sodium ion
PDB	protein data bank
PEG	polyethylene glycol
PHE	phenylalanine
PO_4^{3-}	phosphate ion
PR	wild type HIV-1 protease
PR_{D30N}	PR with D30N mutation
PR_{I50V}	PR with I50V mutation
PR_{I84V}	PR with I84V mutation
PR_{V82A}	PR with V82A mutation
PRO	proline
RMS	root mean square
SER	serine
SO_4^{2-}	sulfate ion
SAQUINAVIR	cis-N-tert-butyl-decahydro-2-[2(R)-hydroxy-4-phenyl-3(S)-[[N-2-quinolylcarbonyl-L-asparaginyl]amino]butyl]-(4AS)-isoquinoline-3(S)-carboxamide
THR	threonine
TRP	tryptophan

TYR	tyrosine
UIC-94017	(3R,3AS,6AR)-hexahydrofuro[2,3-B]furan-3-yl (1S,2R)-3-[[(4-aminophenyl)sulfonyl](isobutyl)amino]-1-benzyl-2-hydroxypropylcarbamate
µl	microliter
VAL	valine

Crystallographic analysis and kinetic studies of HIV-1 protease and drug-resistant mutants

1. Introduction

1.1. General introduction

Current general situation of AIDS

Two decades ago the first case of the acquired immunodeficiency syndrome (AIDS) was recognized in the United States. By 2005, more than 40 million people are known to be infected with AIDS and more than 25 million people have died of AIDS (AVERT). AIDS is now the fourth-biggest killer globally. Currently, no cure or effective vaccine has yet been developed, but modern clinical treatment, Highly Active Anti-retroviral Therapy or HAART, is available to extend and improve the lives of patients afflicted with AIDS. HAART reduces viral load and prevents many opportunistic infections associated with AIDS. However, AIDS is still a fatal disease and even HAART is not widely available outside high-income developed countries because of its high cost. Thus, it is critically important to improve current clinical accessible treatments, and more efforts should be put into the research to continue extending such advances. More effective and low-cost drugs need to be designed, tested and supplied to many developing countries.

The HIV-1 life cycle

AIDS is caused by human immunodeficiency virus type 1 (HIV-1), a member of the family of retroviruses that possess complex genomes and exhibit cone-shaped capsid core particles (Barre-Sinoussi et al. 1983; Gallo et al. 1984; Levy et al. 1984). Other examples of retroviruses include the simian immunodeficiency virus (SIV), bovine leukemia virus (BLV), human T-cell leukemia virus type-1 (HTLV-1), and equine infectious anemia virus (EIAV). As a characteristic of all retroviruses, HIV's genome is encoded by RNA. The viral reverse transcriptase (RT) will reverse-transcribe RNA to viral DNA when virus enters a new host cell. The general features of the mature HIV virion and ribbon representations of the structurally characterized viral proteins are shown in Figure 1.1 (Turner et al. 1999). The surface of virus is a lipid bilayer embedded by envelope proteins (SU) which is derived from the membrane of its host. The middle part is a matrix shell composed by matrix protein (MA) and a capsid core particle comprising of the capsid protein (CA). The inner of the capsid particle contains two copies of the unspliced viral genome and three most essential viral enzymes: protease (PR), reverse transcriptase (RT) and integrase (IN).

Figure 1.2 (Turner et al. 1999) presents the general scheme of the HIV replication cycle. It contains six steps. The replication cycle starts by the binding of the virus envelope protein to CD4 on the cell surface, a protein that normally functions in immune recognition. The viral protein then fuses into the cell membrane (Doranz et al. 1996; Feng et al. 1996; Clapham et al. 1997; Moore 1997). After uncoating of the virus, an intracellular reverse transcriptase (RT) catalyzes the reverse transcription of the viral

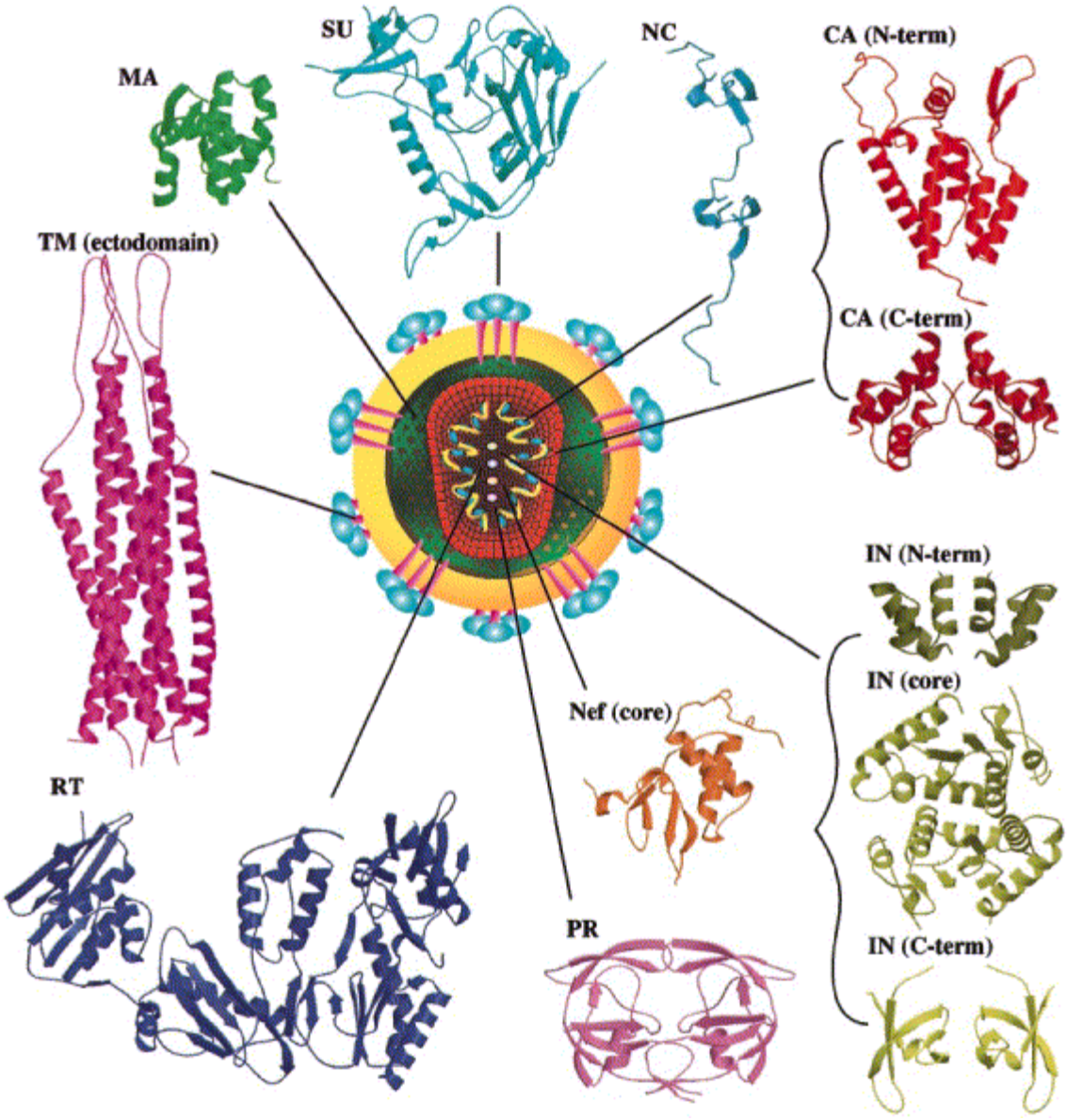


Figure 1.1 Representation of the mature HIV-1 virion surrounded by the known structures of the viral proteins and protein fragments. (Turner et al. 1999)

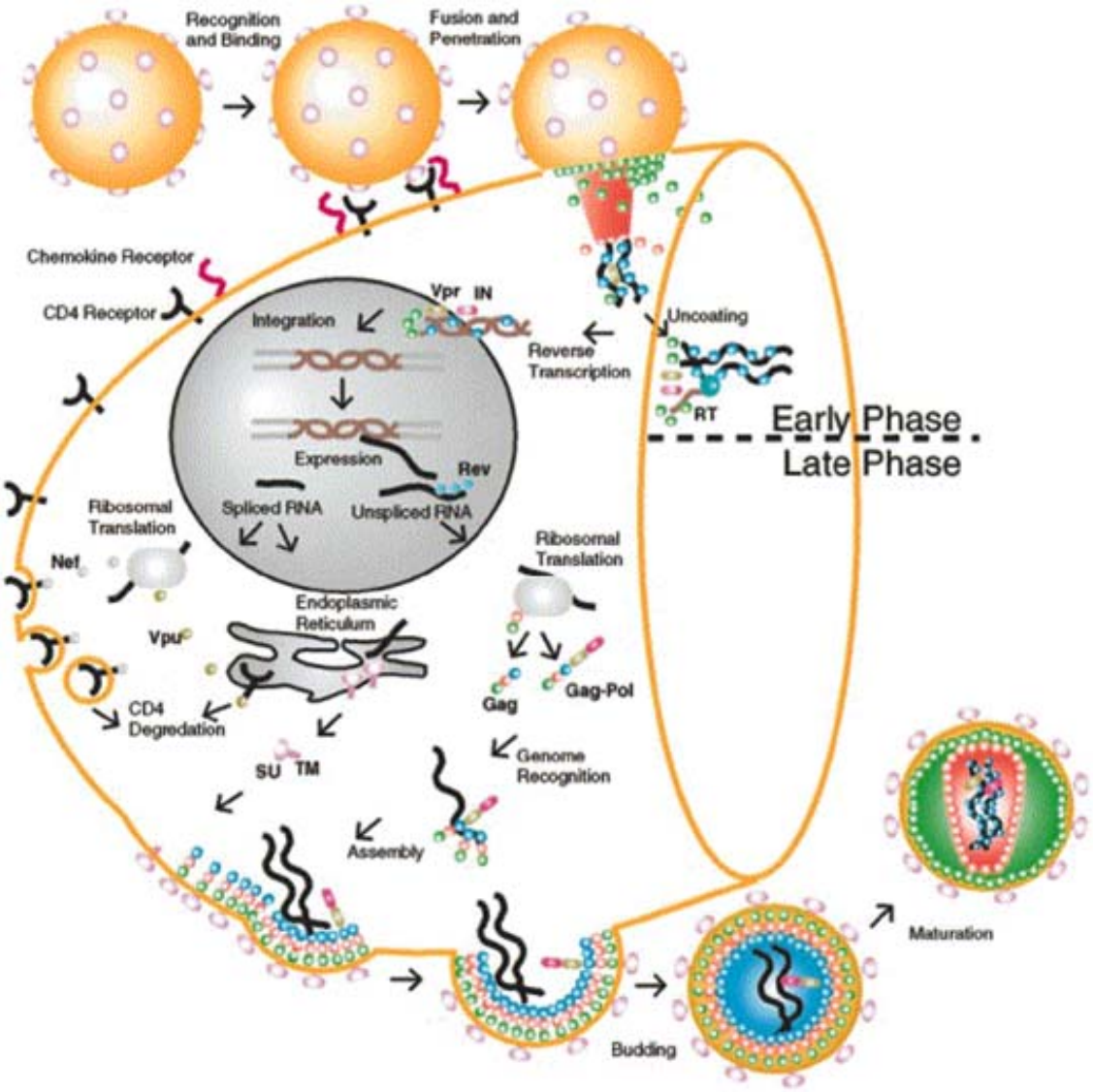


Figure 1.2 General features of the HIV-1 life cycle. (Turner et al. 1999)

RNA in the cytosol. The mechanism of this RT-dependent DNA synthesis has been well studied *in vitro* and *in vivo* (Goff 1990; Whitcomb et al. 1992; Skalka et al. 1993; Katz et al. 1994; Telesnitsky et al. 1997). Once synthesized, the viral cDNA is translocated to the nucleus with matrix protein, IN and other viral proteins as a preintegration complex (Miller et al. 1997). The nuclear localization of the preintegration complex is directed by the accessory protein Vpr (Freed et al. 1995; Fouchier et al. 1997; Nie et al. 1998). The viral cDNA is then integrated into the host genome by the catalytic action of IN. The integration process also requires a host chromosome protein. Multiple unspliced and spliced mRNA are synthesized and transported out of the nucleus for translation. The translation is a very complicated process initialized by the synthesis of regulatory proteins Tat, Rev and Nef, and then the polypeptides: Gag and Gag-pol and protease are formed. Viral RNA and proteins assemble at the host cell surface and buds from the cell surface to form an immature particle. Finally, as protease cleaves polyproteins into structural proteins, such as the MA, CA and NC and functional enzymes (including itself), the structural proteins rearrange and form the mature infectious virus particle. The whole life cycle is usually divided into two phases: the early phase ends with the integration of the viral genetic information into the chromosome of the host cell, while the late phase includes viral protein expression and virus maturation (Turner et al. 1999; Freed 2001).

The purpose of studying HIV-1 protease

Obviously, through the HIV replication cycle, the retroviral enzymes: reverse transcriptase (RT), integrase (IN), and protease (PR) are three important targets for drug discovery and PR has proved to be the most effective one. PR is required for proteolytic cleavage of viral polyprotein precursors Gag and Gag-Pol into individual structural and functional proteins during viral maturation (Darke et al. 1988; Oroszlan et al. 1990). Two steps are critical in this process: one is the formation of an active PR dimer from two Gag-Pol precursors, which is crucial for the release of mature PR; the other is the cleavage at the N terminus of the PR which is necessary for the liberation of catalytic activity (Louis et al. 1994) and optimal ordered processing of the Gag and viral infectivity (Tessmer et al. 1998).

In HIV-1, the Gag polypeptide is composed of structural proteins in the arrangement of MA-CA-p2-NC-p1-p6 (Genbank: HIVHXB2CG) (Henderson et al. 1992; Wondrak et al. 1993), while Gag-Pol of MA-CA-p2-NC-TFR-PR-RT-IN (Louis et al. 2000), as shown in Figure 1.3. A protease cleavage site separates each of the subdomains on the polyprotein and the numbers (Figure 1.3) represent the order of cleavage. The ordered process of polypeptides is controlled by different intrinsic proteolysis rates at the different cleavage sites (Wiegers et al. 1998). The catalytic activity of the mature PR and sequential processing of the polyproteins have been shown to be critical for the liberation of infective progeny virus (Oroszlan et al. 1990; Debouck 1992; Rose et al. 1998).

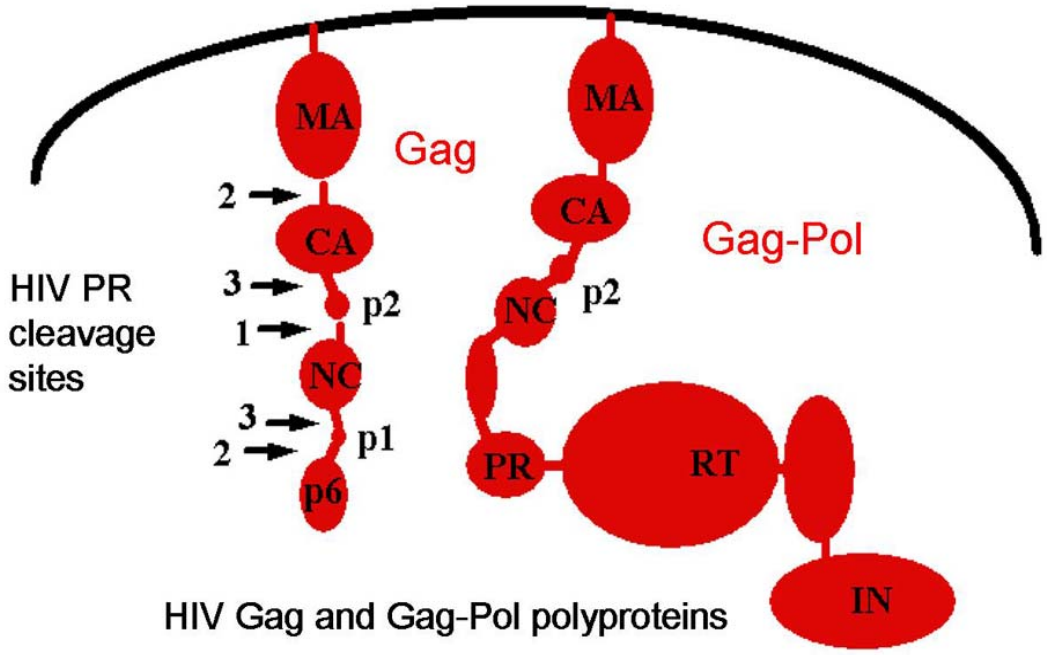
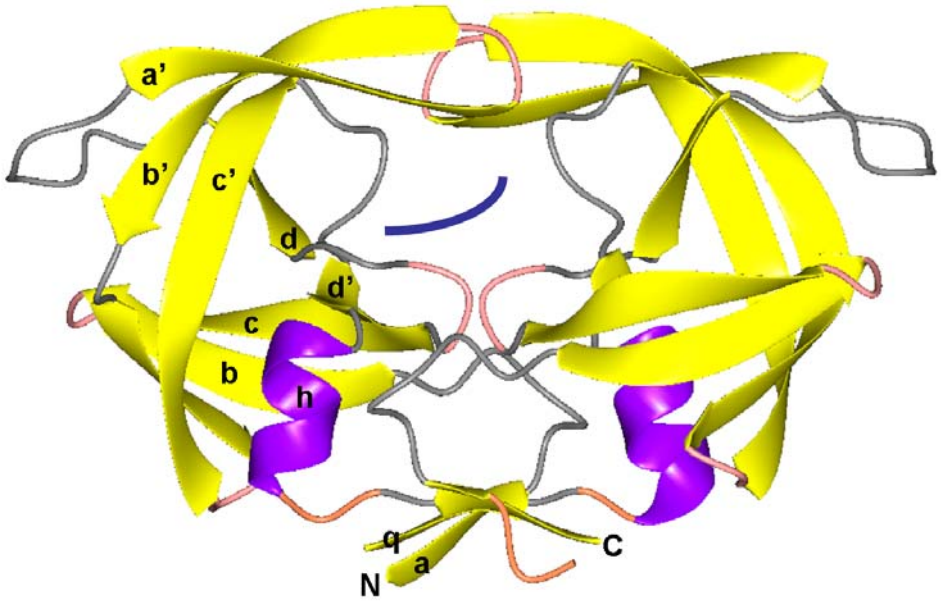


Figure 1.3 Composition of the Gag and Gag-Pol polyproteins and HIV-1 PR cleavage sites. The number represents the order of cleavage of the polyproteins with 1 for the first and 3 for the last.

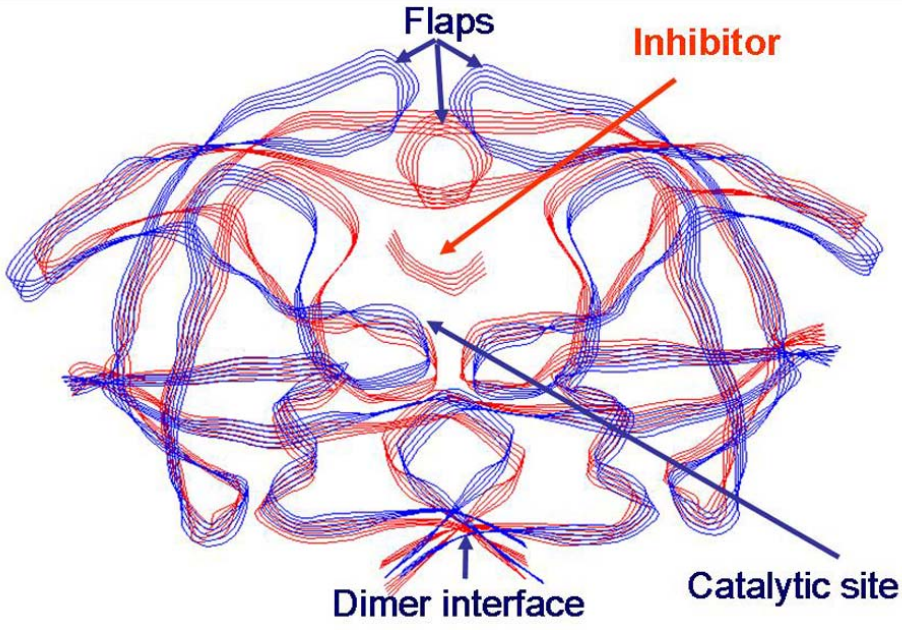
Structural properties of HIV-1 protease

HIV-1 PR has 99 amino acid residues and is enzymatically active as a homo-dimer under acidic conditions. It is a member of the aspartic protease family. Crystal structures have been determined for HIV PR in the absence and presence of inhibitor, as reviewed in (Wlodawer et al. 1993) (Figure 1.4B). The two chemically identical subunits are arranged in a nearly symmetric manner. Each subunit folds into a compact structure of β-strands with a short α-helix near the C terminus. As shown in Figure 1.4A, the N-terminal residues 1–4 forms β-strand *a,* which is the outer part of the interface β-sheet. Residues 9–15 forms β-strand *b* and through a turn, continue into β-strand *c*, which ends at the active-site triplet (Asp25-Thr26-Gly27). Following the active-site loop is β-strand *d* consisting of residues 30–35. A broad loop (residues 36–42) links to β-strand *a'* formed by residues 43–49. Residues 52–66 forms a long β-chain *b'* and the β-chain *c'* comprises residues 69–78. After a loop by residues 79–82, short strand *d'* (residues 83–85) leads directly to the only α-helix *h* (residues 86–94). Helix h is followed by a straight C-terminal strand *q* (residues 95–99), which forms the inner part of the dimer interface. Four of the β-strands in the molecular core are organized into a ψ-shaped sheet characteristic of all aspartic proteases.

Residues 44-57 in each subunit create a flexible "flap" composed of a pair of anti-parallel β-strands (*a'* and part of *b'*). Inhibitors or substrates bind between two subunits in the dimer by hydrogen bonds and van der Waals interactions with the active site, which consists of two characteristic triplets and two flexible flaps. The flaps fold down over the substrate or inhibitor and act during catalysis both to bind substrate and exclude



A.



B.

Figure 1.4 Ribbon drawing of HIV-1 PR structure.
A. Representation of secondary structure for PR with inhibitor (or substrate).
B. PR structures with / without inhibitor (or substrate). Red one presents PR with binding of inhibitor (or substrate) and blue for PR without inhibitor (or substrate).

water molecule from the active site (Figure 1.4B) (Wlodawer et al. 1993; Louis et al. 2000). There is a conserved buried water molecule that bridges the inhibitor and Ile50/50' NH groups of the flaps. This water has proved to be essential for the PR hydrolysis and could be displaced by a carbonyl group in an appropriately designed inhibitor (Lam et al. 1994). NMR studies confirmed the displacement of the water molecule (Wang et al. 1996) and also verified the presence of other structurally important water molecules (Grzesiek et al. 1994). NMR studies also confirmed that the PR flaps are indeed flexible, and suggest that this flexibility may be important for enzyme activity (Nicholson et al. 1995; Yamazaki et al. 1996). Mutation of flap residues can cause dramatic reduction in enzyme activity (Tozser et al. 1997). Besides the catalytic triad, Gly86-Arg87-Asn88 in HIV-1 PR is another highly conserved region in retroviral proteases which may contribute to dimer formation (Louis et al. 1989). It has been reported that either mutating catalytic residues at the active site (Kohl *et al.* 1991) or disrupting the dimerization of active HIV-1 protease monomers (Zhang *et al.* 1999) will eliminate the catalytic activity of the PR and thus block the infectivity of the virus.

Although the native protein is a symmetric dimer, inhibitors bind in an asymmetric manner, as must the natural protein substrates. The substrate binding site consists of subsites S3 to S4' and is mainly formed by residues 8, 23-30, 32, 45-50, 53, 56, 76, 80-82 and 84. PR residues 1-3, 5-9, 23-27, 29, 47-52, 54, 67, 81, 87 and 90-99 contribute to the dimer interface contacts.

Catalytic mechanism of HIV-1 PR

The catalytic mechanism of HIV-1 PR is still controversial. HIV PR is a member of the aspartic protease family. The reaction involves two catalytic Asps and one or two flap regions are important for substrate binding and catalysis. The retroviral PRs, like HIV PR, have differences in their tertiary and quaternary structures compared to the cellular or pepsin-like proteases. HIV PR is catalytically active as a dimer with one catalytic aspartic acid and one flap from each subunit, while pepsin-like proteases act as monomers with one catalytic Asp from each of two domains and generally a single flap region. Kinetic studies and molecular dynamic simulations have been used to explore the detailed mechanism but it is hard to demonstrate the feasibility of a particular mechanism. Overall, the reaction is a general-acid/general-base catalysis and involves both physical changes (such as the folding down of the flaps over substrate) and chemical steps (formation and dissociation of a tetrahedral intermediate). Unlike the serine and cysteine proteases, no covalent acyl-enzyme intermediate is formed during this process. The efficiency of the cleavage depends on the amino acid sequence of the substrate (Polgar et al. 1994).

The general catalytic mechanism of HIV-1 PR was proposed by Hyland (Hyland et al. 1991) by kinetic measurements, and is very similar to the "push-pull" mechanism of cellular aspartic proteases (Polgar 1987). The catalytic Asp25 residues from both subunits interact closely at the subunit interface and dispose on opposite faces of the peptide bond to be cleaved. One Asp (Asp-COO-) which is un-protonated acts as the general base to activate the attacking H-OH, and the second Asp (Asp-COOH) which is

protonated acts as the general acid to protonate the departing amine product. After the formation of a tetrahedral intermediate, the two aspartic acids act in a complementary fashion for breakdown of the tetrahedral adduct (Figure 1.5A). Moreover, Hyland suggested a possible change in rate-limiting step from a chemical one coupled with proton transfer at high pH to a physical one in which a product release step or an enzyme isomerization step becomes partly rate-limiting at low pH. Kinetic experiments of different substrates by Polar indicated that the rate-limiting step can be either a chemical step or a physical step depending on the charge of the residue in P2'. A charged residue at P2' can alter the nature of the enzyme-substrate complex and result in different mechanisms for the different substrates (Polgar et al. 1994). Later studies by Szeltner and Polgar showed that, independent of the salt concentration, the rate-limiting step of the catalysis is a chemical process rather than a physical one (Polgar et al. 1994). Besides the argument of the rate-limiting step, other debate regarding the catalytic process is which Asp is protonated and where the H_2O molecule is from. One possibility is the conserved water molecule between the PR flaps and the inhibitor observed in most crystal structures. However, this water molecule is located at the opposite side of the cleaved peptide bond and far away from the catalytic Asps; so to be able to attack the catalytic Asp, energy would be needed to transfer it to a proper position between the catalytic Asps and substrate. Studies by Baca have shown that the hydrogen bonds between this water molecule and the PR flaps do not contribute to the binding of substrate, but are essential for normal turnover of substrate under steady-state conditions by chemically modified PR (Baca et al. 1993). Molecular dynamics simulation studies by

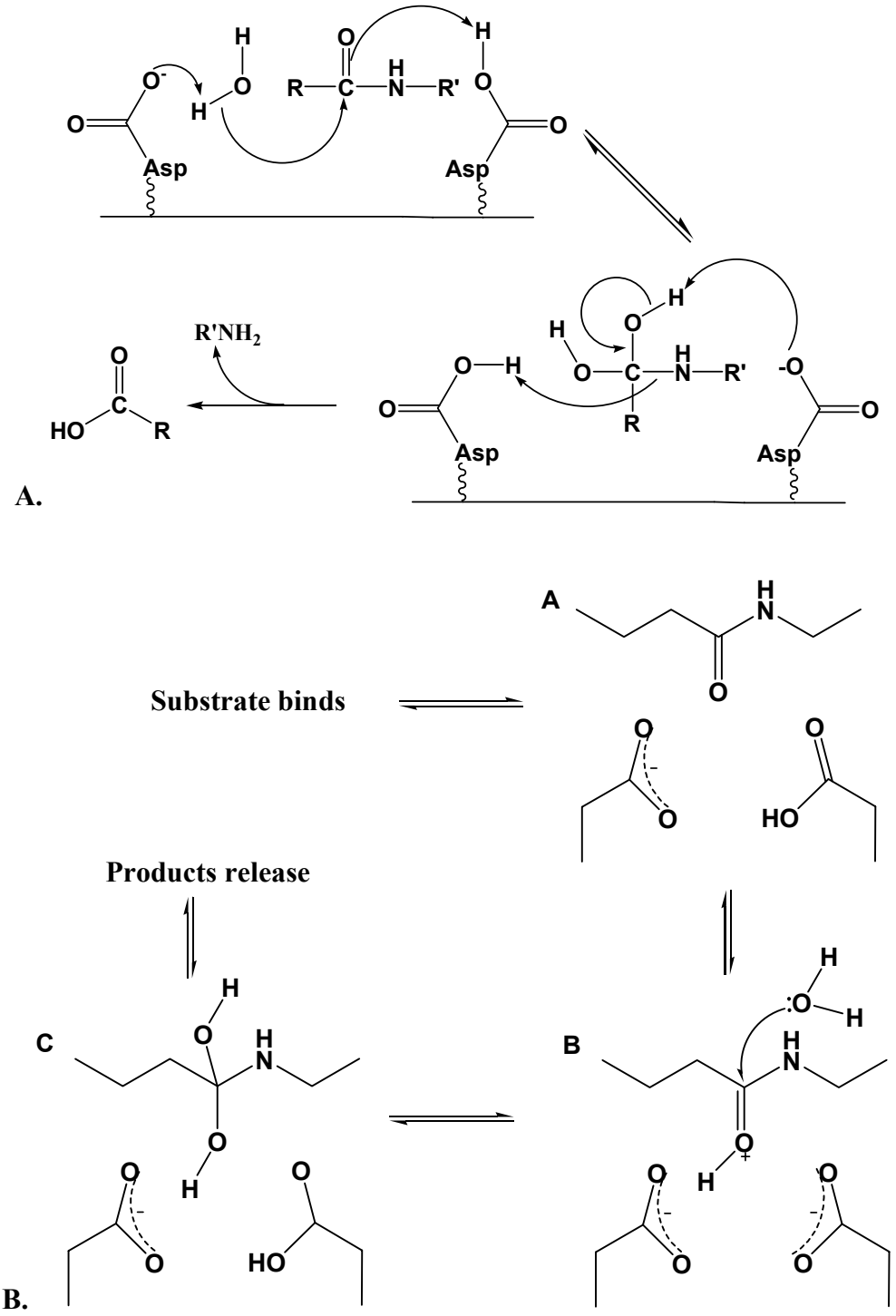


Figure 1.5 Reaction mechanism of HIV-1 PR.
A. One reaction mechanism proposed by Hyland (Hyland et al. 1991).
B. Another mechanism for the formation of reaction intermediate proposed by Harrison (Harrison et al. 1994).

Chatfield further suggested that this H_2O does not participate directly in the reaction but stabilizes productive conformations of reactants and intermediates, and in addition, there would be a lytic H_2O near the active site with very little mobility to act as nucleophilic water (Chatfield et al. 1995). The latter assumption can hardly agree with the crystal structures of HIV-1 PR with transition-state analog inhibitors because there is no space for such a water molecule near the catalytic aspartic residues. Some other studies of pepsin-like proteases have suggested that there was a water/solvent molecule between the two catalytic Asps and at a hydrogen bonding distance from all four oxygen atoms in the crystal structures without bound inhibitor and this water/solvent might act as the nucleophile (Pearl et al. 1984; Suguna et al. 1987). However, the mechanism of HIV-1 PR could be different from that of pepsin-like aspartic proteases. Crystal structures suggested the presence of a uranyl ion at similar site of Rous sarcoma virus PR (Miller et al. 1989) and HIV-1 PR structure in the absence of inhibitor (Wlodawer et al. 1989). Thus, Harrison had analyzed another possibility by molecular dynamics simulations (Harrison et al. 1994). As shown in Figure 1.5B, after substrate binds (A), the proton moves from the protonated Asp to the carbonyl oxygen on the substrate and the positive charge is stabilized by the Asps (B). The protonated carbonyl carbon is then attacked by a nucleophile, such as water, to form the transition-state intermediate (B, C). The gem-diol groups are unstable and will rapidly break down to release two product peptides. The proton is restored. In this case, the water molecule is attacking from the opposite side of the scissile bond so the flap water molecule mentioned above becomes an ideal nucleophile in processing the reaction. Moreover, once the reaction completes, this H_2O

will no longer stabilize the protease/substrate binding and facilitate the dissociation of the peptide products. A recent quantum-classical molecular dynamics simulations by Trylska has proposed another direct nucleophilic mechanism in which the un-protonated Asp25 acts as a nucleophile and attacks the peptide carbon of the substrate, and tried to test the possibility of this direct mechanism and general-acid/general-base mechanism, but no conclusions have been made at this stage (Trylska et al. 2004).

Substrate preference of HIV-1 PR

The naturally occurring cleavage site sequences on Gag and Gag-Pol polyprotein are listed in Table 1.1. The sequences are quite diverse, but still show certain preferences. At P1 and P1', on either side of the cleavage bond, hydrophobic residues dominate, but Val and Ile are absent. Pro appears frequently at P1' but never at P1, instead, Phe is the most frequent amino acid at P1. Small hydrophobic or polar residues occupy P2 and P2' and more diversity is found at the outer positions. The subsite preference has also been studied by enzyme kinetics using oligopeptide substrates. The PR can not hydrolyze peptides with β–branched amino acids, such as Val, substituted in the P1 position, as well as Pro, Ser or Gly (Phylip et al. 1990; Tozser et al. 1992). Charged amino acids, such as Arg and Glu, at P1 or P1' positions are also not preferred in substrates (Konvalinka et al. 1990; Cameron et al. 1993). In agreement with the natural cleavage site sequences, HIV PR prefers small hydrophobic residues at P2 and can accommodate various residues at P3

Table 1.1 Protease cleavage sites in HIV-1 Gag and Gag-Pol polyproteins (Louis et al. 2000).

Location of the cleavage site	**P4**	**P3**	**P2**	**P1**	**P1'**	**P2'**	**P3'**
In Gag							
MA-CA	Ser	Gln	Asn	Tyr	Pro	Ile	Val
CA-p2	Ala	Arg	Val	Leu	Ala	Glu	Ala
p2-NC	Ala	Thr	Ile	Met	Met	Gln	Arg
NC-p1	Arg	Gln	Ala	Asn	Phe	Leu	Gly
p1-p6	Pro	Gln	Asn	Phe	Leu	Gln	Ser
in p6	Lys	Glu	Leu	Tyr	Pro	Leu	Thr
In Pol							
TFP-$p6^{pol}$	Asp	Leu	Ala	Phe	Leu	Gln	Gly
$p6^{pol}$-PR	Ser	Phe	Asn	Phe	Pro	Gln	Ile
PR-RT	Thr	Leu	Asn	Phe	Pro	Ile	Ser
p66-p51	Ala	Glu	Thr	Phe	Tyr	Val	Asp
RT-IN	Arg	Lys	Ile	Leu	Phe	Leu	Asp

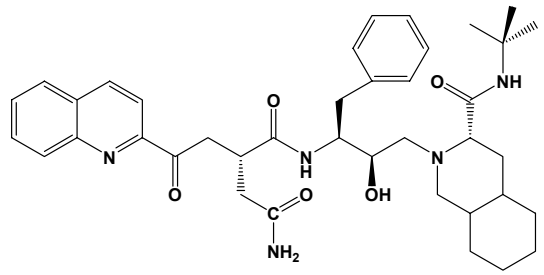
and P3' while the P2' preference depends on the amino acid of P1 and P1' (Konvalinka et al. 1990; Tozser et al. 1991a; Tozser et al. 1991b; Tozser et al. 1992; Cameron et al. 1993).

Drug-resistance

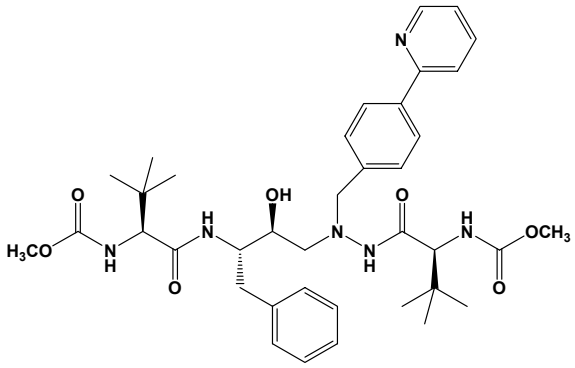
Since the reverse transcriptase inhibitor, azidothymidine (AZT), previously known as a potential anticancer agent, was first approved by the U.S. Food and Drug Administration (FDA) for treating AIDS in 1987, seven FDA-approved RT inhibitors and eight PR inhibitors (PIs) have become commercially available. The protease inhibitors are saquinavir, amprenavir (and fosamprenavir), indinavir, nelfinavir, ritonavir, atazanavir, lopinavir and tipranavir (chemical structures shown in Figure 1.6). The combination therapy using protease inhibitors as well as inhibitors of reverse transcriptase has had a great impact in extending the life of AIDS patients (Gulick et al. 1997; Palella et al. 1998). However, the long-term therapeutic efficiency is restricted by the rapidly development of drug resistant variants (Tamalet et al. 2000).

Similar to other retroviral polymerases, HIV RT does not have a proofreading function (Roberts et al. 1988; Ji et al. 1992) and thus increases the mutational rates of the virus. Divergent viral populations are present during infection and the proteases from these different strains differ in their sequences. Besides, HIV virus has high replication capability. In a fully infected patient about 10^9 virus particles are produced every day (Ho et al. 1995; Wei et al. 1995). It has been estimated that at least 109 new cells are infected daily in latent HIV-infected patients, and the point mutation occurs at the rate of 105

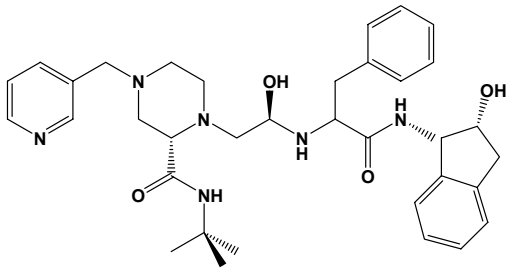
Saquinavir



Ritonavir



Indinavir



Nelfinavir

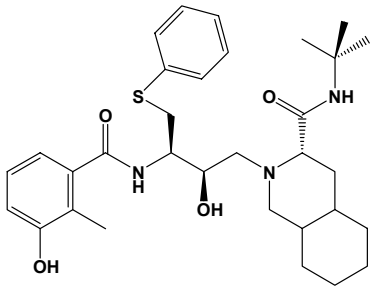
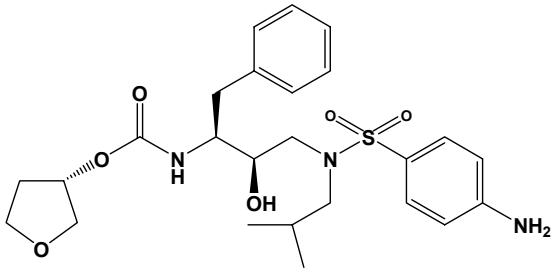
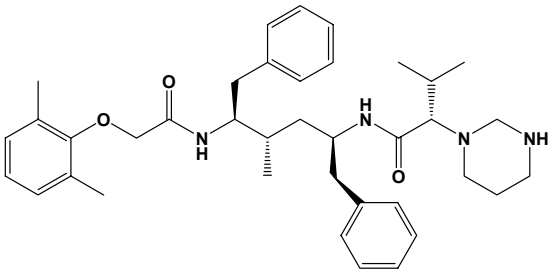


Figure 1.6 Chemical structures of currently eight FDA-approved HIV protease inhibitors (Part 1).

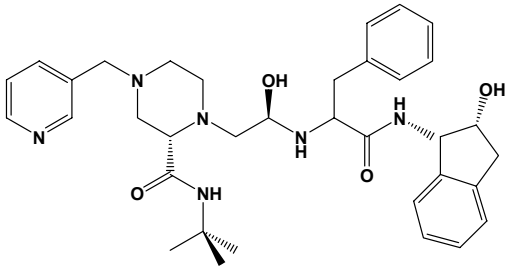
Amprenavir



Lopinavir



Atazanavir



Tipranavir

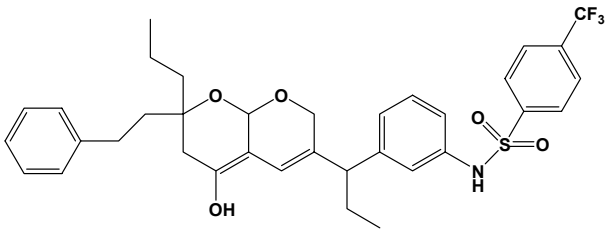


Figure 1.6 Chemical structures of currently eight FDA-approved HIV protease inhibitors (Part 2).

times per day along the entire length of the genome (Coffin 1995). Many of these mutations do not affect the activity of the enzyme, and the structures of fully active proteases that differ in sequences have been reported. However, the virus with certain drug-resistant mutations apparently can be retained in patients' body and is not susceptible to current drugs (Chun et al. 1997; Finzi et al. 1997; Wong et al. 1997). High rates of viral turnover rate may eventually produce high populations of resistant virus during antiviral therapy due to drug selection of viral strains (Boucher 1996). It has been reviewed that 62 mutations of 43 different residues in HIV-1 PR were associated with drug resistance (Hammond et al.; Johnson et al. 2005). The drug-resistant level of some of these mutants versus seven current PIs is shown in Figure 1.7 and their positions in PR structure are mapped in Figure 1.8. Furthermore, cross-resistance has been reported in AIDS patients on combination therapy (Shafer et al. 1998; Hertogs et al. 2000).

The mechanism for the development of resistance to PIs is quite complicated. As shown in Figure 1.8, the locations of PI-resistant mutations vary from substrate/inhibitor binding site, dimer interface, and flap region to surface of PR. One explanation is that the residues near the active site can directly interfere with the binding of the inhibitor to the enzyme by modifying its specificity and/or affinity. Another cause for those mutations which do not directly change the shape or property of the binding cavity is that they may affect the efficiency of catalysis and the stability of the PR (Gulnik et al. 1995; Mahalingam et al. 1999), or influence enzyme active site via long-range structural perturbations. Still, drug resistance can be caused by various reasons or combination of different subtle changes.

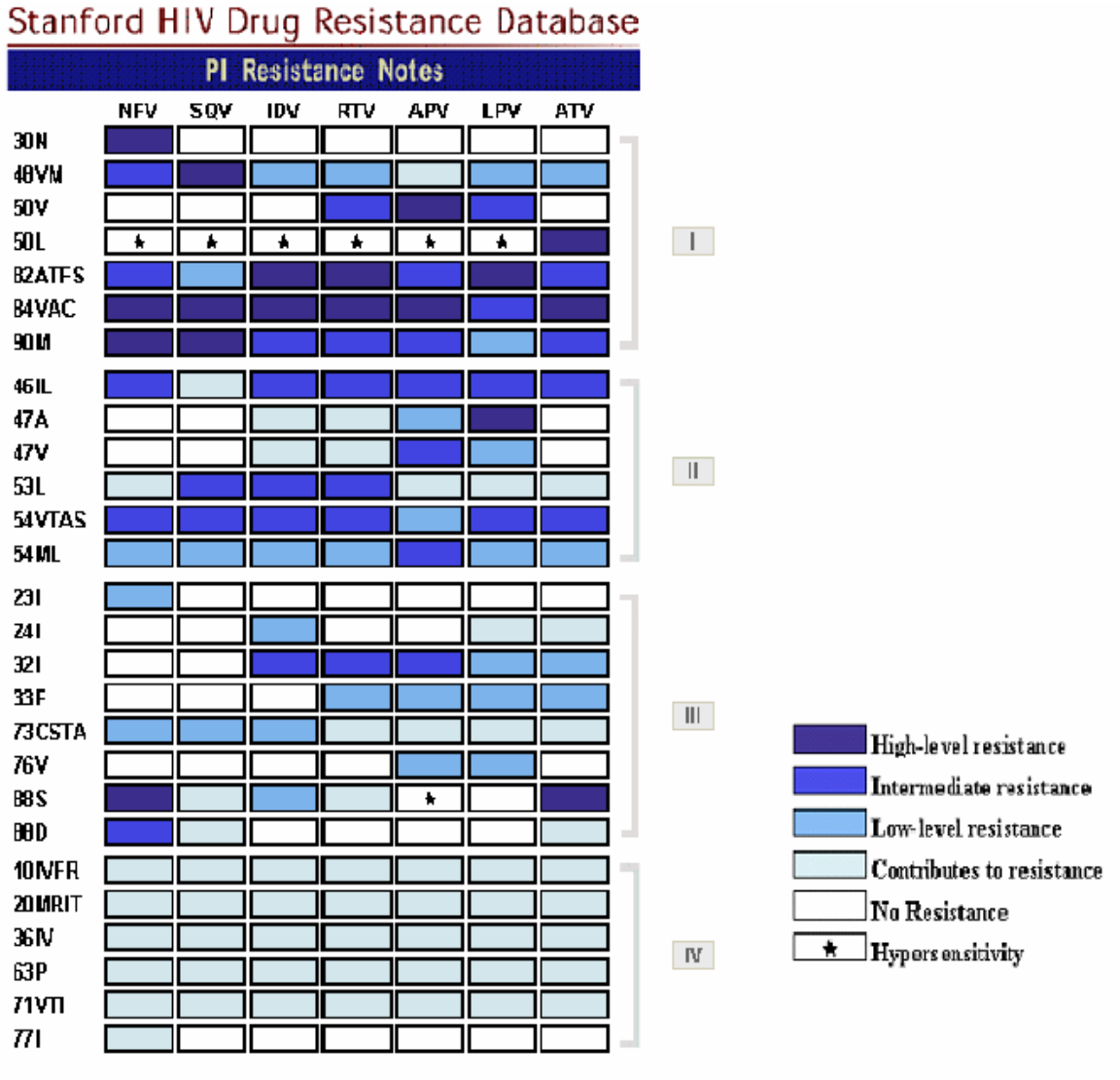


Figure 1.7 Drug resistant level of different mutants against seven FDA approved PIs. (http://hivdb.stanford.edu/cgi-bin/PIResiNote.cgi)
I) Major mutations. With the exception of L90M, these mutations are in the substrate cleft site. Each by itself is capable of reducing susceptibility to one or more PIs. With the exception of V82I, which does not cause drug resistance, mutations at these positions are nonpolymorphic in that they do not occur in PI-naïve persons. II) Flap mutations. These mutations are in the protease flap and are second in importance only to the major mutations. Some may cause resistance by themselves, but these mutations are more often accessory. They are also nonpolymorphic. III) Other nonpolymorphic PI-resistance mutations. These mutations indicate past PI exposure. Several cause resistance by themselves but more often they are accessory. IV) With the exception of L10F, these mutations are polymorphic. They contribute to resistance only when present in combination with one or more category I-III mutations.

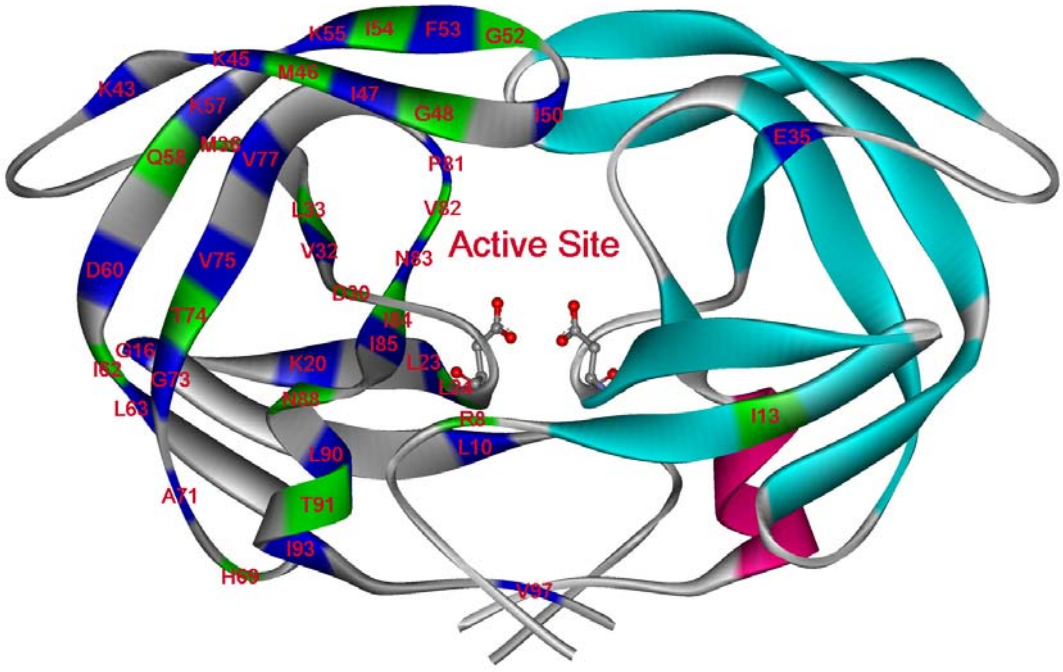


Figure 1.8 Sites of drug resistant mutations on PR dimer.

1.2 Rationale and Specific Aims

Up to today, more than 240 crystal structures of HIV-1 PR and PR-inhibitor complexes were determined (PDB), most of which were used to guide the development of protease inhibitors. Indeed, these efforts have illustrated enormous potential and utility of the rational structure-based drug design strategy. However, the molecular basis for the drug resistance is still not fully understood. The long-term effectiveness of antiviral treatment is severely limited by the drug resistance and narrow choices of available drugs. To address this question and design more effective PIs to combat drug resistance, crystallographic and kinetic studies were applied to wild-type HIV-1 protease and drug-resistant mutants, PR_{V82A}, and PR_{I84V}, in complex with substrate analogs, current drug saquinavir and latest design UIC-94017 (also called TMC-114). UIC-94017 was also studied with PR_{D30N} and PR_{I50V}. Crystal structures of some resistant mutants have provided the molecular basis for the altered enzymatic properties and the resistant phenotype (Hong et al. 2000; Mahalingam et al. 2001; Mahalingam et al. 2002; Prabu-Jeyabalan et al. 2003). These studies of detailed structural information and kinetic properties suggested that individual mutations, double or multiple mutations, had a range of effects, such as changed binding affinity, catalytic activity and dimer stability, which depend on the specific combination of mutant with substrate or inhibitor. Therefore, it is important to obtain high-resolution crystal structures in order to observe the small structural changes associated with the mutations.

The purpose of choosing PR and mutants PR_{V82A}, PR_{I84V}, PR_{D30N} and PR_{I50V}

Mutations in the substrate binding site can cause resistance by reducing the PR binding affinity by 2~5-fold for inhibitors (Erickson et al. 1999). Resistant mutations are commonly observed at D30, M46, I50, V82 and I84 (Shafer et al. 1998; Shafer 2002). Mutations of residue 82 show decreased susceptibility to indinavir, ritonavir, and lopinavir in vitro; PR_{V82A} is the most common mutation at this position. Mutation V82A is observed predominantly in HIV-1 isolates from patients receiving treatment with indinavir and ritonavir (Condra et al. 1996; Molla et al. 1996). It also occurs in isolates from patients receiving prolonged therapy with saquinavir after the development of the mutation G48V (Winters et al. 1998). Mutation I84V has been reported in patients receiving indinavir, ritonavir, saquinavir, and amprenavir (Hertogs et al. 2000; Maguire et al. 2002). It causes phenotypic and clinical resistance to each PR inhibitor (Kempf et al. 2001; Prado et al. 2002; Colonno et al. 2003). I84V mutation tends to develop in isolates that already have the mutation L90M and is rarely the first major mutation to develop in patients receiving a PR inhibitor (Wu et al. 2003). Mutations D30N and I50V were especially chosen to study with UIC-94017. PR_{D30N} was selected because Asn30 was expected to affect the hydrogen bond interaction between Asp30 side chain and the aniline group of UIC-94017. PR_{D30N} occurs solely in patients receiving nelfinavir and confers to reduce nelfinavir susceptibility by 5-20 fold (Jarvis et al. 1998; Markowitz et al. 1998; Patick et al. 1998). The Asp-to-Asn mutation neutralized negative charge on aspartic acid without significantly changing the geometry of the side chain. Previous study in our group suggested that PR_{D30N} had varied activity for the hydrolysis of

peptides resembling different Gag and Gag-Pol cleavage sites while retaining its stability compared to the wild type PR (Mahalingam et al. 1999). Crystal structures of PR_{D30N} in complexes with substrate analogs corresponding to CA-p2 and p2-NC cleavage sites revealed distinct changes in the position of the Asn30/30' side chains when they were compared to each other or to the wild type structures, in agreement with the observed different specificity (Mahalingam et al. 2001). Mutation I50V was selected for study since HIV with this mutation shows intermediate-to-high level resistance with amprenavir (Maguire et al. 2002) and low resistance to ritonavir and lopinavir (Partaledis et al. 1995; Tisdale et al. 2000; Prado et al. 2002). Since UIC-94017 was designed based on the chemical structure of amprenavir, the analysis of PR_{I50V}/UIC-94017 complex will help evaluate the effectiveness of potent compound UIC-94017. The mutation from Ile to Val reduces the size of the hydrophobic side chain by one methyl group and is expected to increase the size of the binding cavity and reduce the affinity for inhibitor.

The specific aims of my investigations are as follows:

Aim 1: To determine & analysis crystal structures of PR and mutants PR_{V82A} and PR_{I84V} with peptide analogs.

HIV PR hydrolyzes several different cleavage sites in the natural polyprotein substrates. The cleavage sites share little sequence homology (as shown in Table 1.1) so the specificity of the PR is not fully defined. Furthermore, the mechanisms for how the mutants maintain sufficient enzymatic activity for viral replication can be better understood by studying the structures of PR with natural cleavage sites. Two alternative

strategies have been applied to overcome the difficulty of crystallizing catalytically active enzyme with peptide substrates. Our strategy has been to analyze structures of active PR with substrate analogs, while other groups have used an alternative strategy of crystallizing an inactive enzyme with peptide substrates. Crystal structures have been reported of the inactive PR variant (PR_{D25N}) in complex with peptides representing the MA-CA, CA-p2, p2-NC, NC-p1, p1-p6, p2-NC, RT-RH and RH-IN cleavage sites, and inactive mutant $PR_{V82A/D25N}$ with MA-CA, CA-p2, NC-p1 and p1-p6 peptides (Prabu-Jeyabalan et al. 2002; Prabu-Jeyabalan et al. 2003; Prabu-Jeyabalan et al. 2004). These inactive complexes were refined at about 1.9 Å resolution or lower. We have reported crystal structures of PR, single mutants in complex with substrate analogs CA-p2 and p2-NC (PR_{D30N}, PR_{K45I}, PR_{N88D} and PR_{L90M}), and double mutants with CA-p2 ($PR_{K45I/V82S}$, $PR_{K45I/L90M}$, $PR_{D30N/V82S}$ and $PR_{N88D/L90M}$) at resolutions ranging from 2.2 to 1.2 Å (Weber et al. 1997; Mahalingam et al. 2001; Mahalingam et al. 2002). Reduced peptide analogs mimic the transition state of the hydrolytic reaction since they contain an amine and tetrahedral carbon at the non-hydrolyzable peptide bond. A complete set of accurate high resolution structures of active PR with different substrate analogs will more accurately reflect the structures of reaction intermediates. Therefore, high resolution crystal structures of PR and the common drug resistant variants PR_{V82A} and PR_{I84V} were determined with reduced peptide analogs that represent the CA-p2, p2-NC, $p6^{pol}$-PR and p1-p6 polyprotein cleavage sites. These structures and kinetic data provide details of the PR interaction with reaction intermediates and better understanding of the substrate specificity. Comparison of protease complexes with clinical inhibitors has improved our

knowledge of the molecular basis of drug resistance and will assist in the structure-based design of more potent anti-viral inhibitors.

Aim 2: To determine & analysis crystal structures of PR and mutants PR_{V82A}, PR_{I84V}, PR_{D30N} and PR_{I50V} with UIC-94017

Most current PIs were designed based on substrate peptides, so they generally are targeted on wild type PR and have enhanced hydrophobic groups on the side chains and high molecular weight. Novel inhibitors that target critical conserved regions of the PR different from those of the current drugs are expected to be more active against resistant variants of HIV-1 PR. A novel non-peptide inhibitor UIC-94017 (also called TMC-114) has been developed that is extremely potent *in vivo* against a wide spectrum of HIV strains including multi-drug resistant clinical strains, with the IC50 values of 3-29 nM for the inhibition of the PR mutants selected for resistance to saquinavir, indinavir, nelfinavir, or ritonavir (Koh et al. 2003; Surleraux et al. 2005). This promising non-peptidic inhibitor is in Phase III clinical trials and shows superiority to the approved clinical PIs (Shurtleff 2004). This compound is chemically related to the clinical inhibitor amprenavir, and was designed with a novel *bis*-tetrahydrofuranylurethane (*bis*-THF) group to incorporate additional polar interactions with main chain atoms of the PR dimer. UIC-94017 was observed to bind to the PR with favorable change in enthalpy with ΔH of 12.1 kcal/mol (King et al. 2004), whereas other PIs, such as indinavir, nelfinavir, and saquinavir, their binding is an enthalpically unfavorable process with ΔH small or even negative (Velazquez-Campoy et al. 2000; Velazquez-Campoy et al. 2001). Alternatively,

the favorable entropy gain by the desolvation of big hydrophobic side chains upon binding for most of the PIs in clinical use compensates for the lack of the favorable enthalpy change. Therefore, it suggests that most PIs have very few attractive interactions when bound to the PR, and changes in the geometry of the cavity due to mutations in the binding site may result in considerable loss of the PI's affinity to the PR. Crystal structures of complexes of the UIC-94017 with PR and four common drug resistant mutants PR_{V82A}, PR_{I84V}, PR_{D30N} and PR_{I50V} have been determined at high resolution to understand the molecular basis for the potency of this compound. Such structures are also beneficial in predicting which active site mutations may generate resistant HIV. Furthermore, a complete set of kinetic data of UIC-94017 with comparison to indinavir under same conditions will assist the evaluation of these structure-based designed compounds and help to understand the role of drug resistance.

Aim 3: To determine & analysis crystal structures of PR and mutants PR_{V82A} and PR_{I84V} with saquinavir.

Saquinavir (Figure 1.6), one of the eight protease inhibitors proved by FDA, is a peptide-like substrate analog that binds to the protease active site and inhibits the activity of the enzyme. This inhibition prevents the cleavage of the viral poly-proteins and results in the formation of immature noninfectious virus particles. Saquinavir was first introduced in 1995 (FDA 2005). Saquinavir has been shown to inhibit HIV activity in both acutely and chronically infected cells with EC50 values in the range of 1 to 30 nM (Craig et al. 1991). Combinations of saquinavir with other protease inhibitors, such as

amprenavir, atazanavir, or lopinavir resulted in synergistic antiviral activity. In cell culture, saquinavir also exhibited additive to synergistic effects against HIV-1 when using with reverse transcriptase inhibitors without enhanced cytotoxicity (Deminie et al. 1996). However, similar to other PIs, the long-term therapeutic efficiency is limited by the rapid development of drug resistant mutants with reduced susceptibility to saquinavir (Ji and Leob, 1992; Ho *et al*., 1995; Wei *et al.,* 1995). Genotypic analysis of HIV-1 isolates with reduced susceptibility to saquinavir identified resistance conferring primary mutations G48V and L90M in the protease gene, and secondary mutations L10I/R/V, I54V/L, A71L/T, G73S, V77I, V82A and I84V that contributed additional resistance to saquinavir (Jacobsen et al. 1995; Boucher 1996; Jacobsen et al. 1996). The degree of reduction in susceptibility to saquinavir of clinical isolates bearing substitutions G48V and L90M depends on the number of secondary mutations present. In general, higher levels of resistance are associated with greater number of mutations only in association with either or both of the primary mutations G48V and L90M (Boucher 1996). Nevertheless, little structural study has been done on this inhibitor, only a few structures (PR (Krohn et al. 1991), $PR_{D25N/V82A}$ (Prabu-Jeyabalan et al. 2003), $PR_{G48V/L90M}$ (Hong et al. 2000), 9X mutant structure (Munshi et al. 2000)) are available and they are all around 2.5 Å resolution. Here, the kinetic data and atomic resolution structures of wild-type, PR_{V82A} and PR_{I84V} in complex with saquinavir were studied. PR_{V82A}/SQV structure was refined at 0.97 Å resolution, which is by far the highest resolution structure for this complex. As mentioned before, V82A and I84V are two particularly important mutants which are observed to have cross-resistance. They are associated with high level drug

resistance for almost all seven available PIs (Figure 1.7). Thus the structures presented here will give the detail molecular based information to help understand the drug resistance to saquinavir and provide valuable reference to compare with other PI drugs. The comparison will help understand the structural advantage and disadvantage of compounds and further help the design of more efficient next generation inhibitors.

Aim 4: To develop a small database for organizing and automating the analysis of protease-inhibitor structures.

Although great efforts have been made to study of HIV-1 protease, only a small number of databases have been developed to summarize these valuable data about sequence, resistance, immunology, and vaccine trials information. These databases are available online, such as Los Alamos HIV PR databases at http://hiv-web.lanl.gov/content/index, Stanford HIV PR sequence and resistance databases at http://hivdb.stanford.edu and the National Institute of Allergy and Infectious Disease site of anti-HIV compounds at http://www.niaid.nih.gov/daids/dtpdb. A HIV PR structural database containing about 200 structures is also available at National Institute of Standards and Technology (NIST) site at http://srdata.nist.gov/hivdb. This database not only provides certain descriptive information, for example, cell parameters, space group and citations, but also puts great emphasis on the study of the inhibitors. The best aspect is that it provides a tree of fragments within the selected inhibitor and allows a further tool to perform fragment searches on the inhibitor. Using this tool a user may rapidly search for multiple fragments and thus perform homology searches. However, there are

some limitations. One is that it is a website based program, so, although it offers broad information it is not designed for local lab use and is not secured for those unpublished data. Besides, it focuses less on the structural study, for example, no tools are available to analyze the detailed interactions between atoms or superimpose different complexes. For this instance, and as more and more structures of different HIV-1 protease complexes (>50) will soon be completed in our lab and overall hundreds of HIV protease structures have been published, a database which can be incorporated with some analytical programs as well as present more efficient management would be beneficial. There are at least two reasons for creating this database. First, as only a few inhibitors have reached approval by the FDA as anti-AIDS drugs, we are working on a lot of those compounds still being synthesized or under development by Dr. Arun Ghosh's lab. The information about these compounds is valuable and confidential. Some of them may not have direct clinical significance but still be important for the purpose of design. Second, the database would provide an efficient way to analysis and compare ligand-enzyme interactions in the structures and provide useful information that could help exploring new drug-design principles and synthetic strategies. Its shell will contain structural files that have been previously deposited in the Protein Data Bank (PDB) or can be directly uploaded into the HIV PR database. After filtering by protease or inhibitor name (decided by the user), the files of complexes are placed in a list. A branched link can display the pdb file and the 3D structure can be visualized by Rasmol (Kohl et al. 1988). At this stage, the analytical part of the database consists of two basic functions: distance calculation and superimposing program FUD. For distance calculation, the residues (or atoms) and the

range of distance to be calculated can be defined by the user and be applied to multiple structural files. Whilst for superimposing structures, a maximum of eight structures can be superimposed on the same reference file at the same time. A plot of the RMS deviations on Cα against each residue will be shown, with the global RMS deviation for each pair of comparison presented at the top. Later on, more calculations or analytical programs can be incorporated easily for further development.

2. Methods

2.1 Protein expression and purification

The HIV-1 PR has been optimized for structural and kinetic studies, with 5 mutations: Q7K, L33I, L63I to minimize the autoproteolysis of the PR, and C67A and C95A to prevent cysteine-thiol oxidation (Wondrak et al. 1996). The construction and expression of HIV-1 PR, PR_{D30N}, PR_{I50V}, PR_{V82A} and PR_{I84V} was carried out as described (Louis et al. 1989; Wlodawer et al. 1993). Mutations were confirmed by protein mass spectrometry. The purification and refolding procedures were similar to those used by (Louis et al. 1989; Wondrak et al. 1996).

Escherichia coli cells with PR constructs were incubated in 2 L LB (Luria Bertani) broth containing 100 μg/ml carbenicillin and 1% glucose at 37 ˚C. After A_{600} reached 0.5-0.6, the cells were induced by IPTG (2mM) for around 4 hours. After pelleted by centrifuge (4,000 rpm, 20 min, 4 ˚C), the cells were suspended in 20 volumes of TE buffer (50 mM Tris-HCl, 10 mM EDTA, pH 8.2) and add to it 0.1 mg/ml lysozyme. After sitting on ice for 10 min, the cells were lysed by sonication at 4 ˚C for 10 min with 1 min pulses at 1 min intervals. The inclusion bodies (IB) were clarified by centrifugation at 12,000 rpm for 20 min at 4 ˚C. The IB pellet was then suspended in TE buffer with 1% triton X-100 and 2 M urea. After centrifuge at 12,000 rpm for 20 min at 4 ˚C, the pellet was washed by TE buffer and dissolved in 3 volumes of TE buffer containing 8 M urea and 10 mM DTT. The supernatant was subjected to gel filtration chromatography on Superdex 75 (HiLoad 26/60, Pharmacia Biotech, Uppsala, Sweden) with elution buffer (50 mM Tris-HCl (pH 8.2), 4 M guanidine hydrochloride and 5 mM

EDTA) at a flow rate of 1 ml/min. The proteins in the peak fractions around 135 ml were collected and further purified by HPLC with RPC-15 column (reverse phase) (ST 4.6/100, Amersham Biosciences) using a gradient of 0 to 100% acetonitrile with 0.05% trifluoroacetic acid at a flow rate of 2 ml/min. The major peak corresponding to HIV-1 protease eluted at 40% acetonitrile. The presence of protein was confirmed by SDS-PAGE. The desired HPLC fractions were dialyzed against 200 volumes of first, 25 mM formic acid and then, 50 mM sodium acetate (pH 5.0) at 4 ˚C to obtain folded active protease.

2.2 Substrates, peptide analogs and small organic inhibitors

The chromogenic substrate L6525 was purchased from Sigma. Two series of inhibitors, substrate analogs and small organic molecules, were used to grow crystals with wild-type protease and mutants (details shown in Table 2.1 and Figure 2.1). CA-p2 and p2-NC reduced peptide analogs were purchased from BACHEM. The NC-p1, p1-p6 and $p6^{pol}$-PR reduced peptides and the substrate for fluorescence essay were synthesized by Dr. Ivo Blaha (Ferring Leciva, Prague). The substrate analog inhibitors were dissolved in deionized water by vortexing for several minutes, and then centrifuged briefly to remove any insoluble material. The small organic inhibitors were dissolved in dimethyl sulfoxide (DMSO). UIC-94017 was designed and synthesized by our collaborator Dr. Arun Ghosh's group.

Table 2.1 Substrate analogs.

Cleavage site	Number	Sequence
CA-p2	N1270	R-V-L-r-F-E-A-Nle
p2-NC	N1465	Ac-T-I-Nle-r-Nle-Q-R
NC-p1	IB 308	E-R-Q-A-N-r-F-L-G-K-I
p1-p6	IB 307	R-P-G-N-F-r-L-Q-S-P
p6-PR	IB 306	V-S-F-N-F-r-P-Q-I-T-K-K

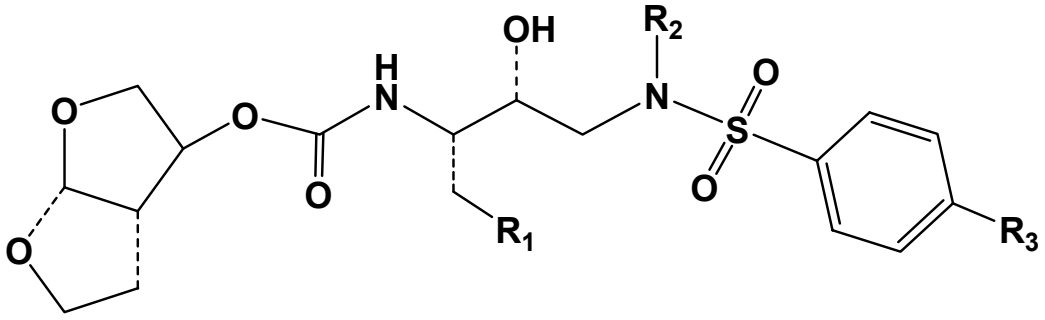


Figure 2.1 Backbone structure of small organic compounds.
(A series of compounds designed and synthesized by Dr. Arun Ghosh)

2.3 Crystallization and data collection

The refolded PR or mutant was concentrated to 3 to 5 mg/ml, and mixed with the substrate analog inhibitor at a 20-fold molar excess (2- fold for synthesized small inhibitions). The mixture was incubated at 4 ˚C for an hour and then centrifuged. Crystallization was achieved by the hanging-drop vapor-diffusion method at 297 K using 24-well VDX plates (Hampton Research). Equal volumes of enzyme-inhibitor stock and reservoir solution were used. The screening was performed with combinations of the following solutions: 0.1 M sodium acetate buffer, pH from 4.2 to 5.0, 0.1 M citrate phosphate buffer, pH from 5.0 to 6.4, 5% DMSO, 0 to 5% dioxane, 0.4 M to 1.2 M sodium chloride and 15% to 40% saturated ammonium sulfate (crystal pictures are shown in Figure 2.2). Generally, the kind of salt used in crystallization will not affect the space group of the crystal. Except for PR_{V82A}/SQV and PR_{I84V}/SQV complex, crystals in space group $P2_12_12_1$ were obtained with ammonium sulfate as precipitant, while crystals with space group $P2_12_12$ grew from sodium chloride.

The crystals were frozen in liquid nitrogen using glycerol as a cryoprotectant, which was added to the reservoir solution to a final concentration of 30% (v/v). X-ray diffraction data were collected at National Synchrotron Light Source, beamline X26C or Advanced Photon Source, beamline SER-CAT 22. An example of diffraction pattern of HIV-1 protease is shown in Figure 2.3.

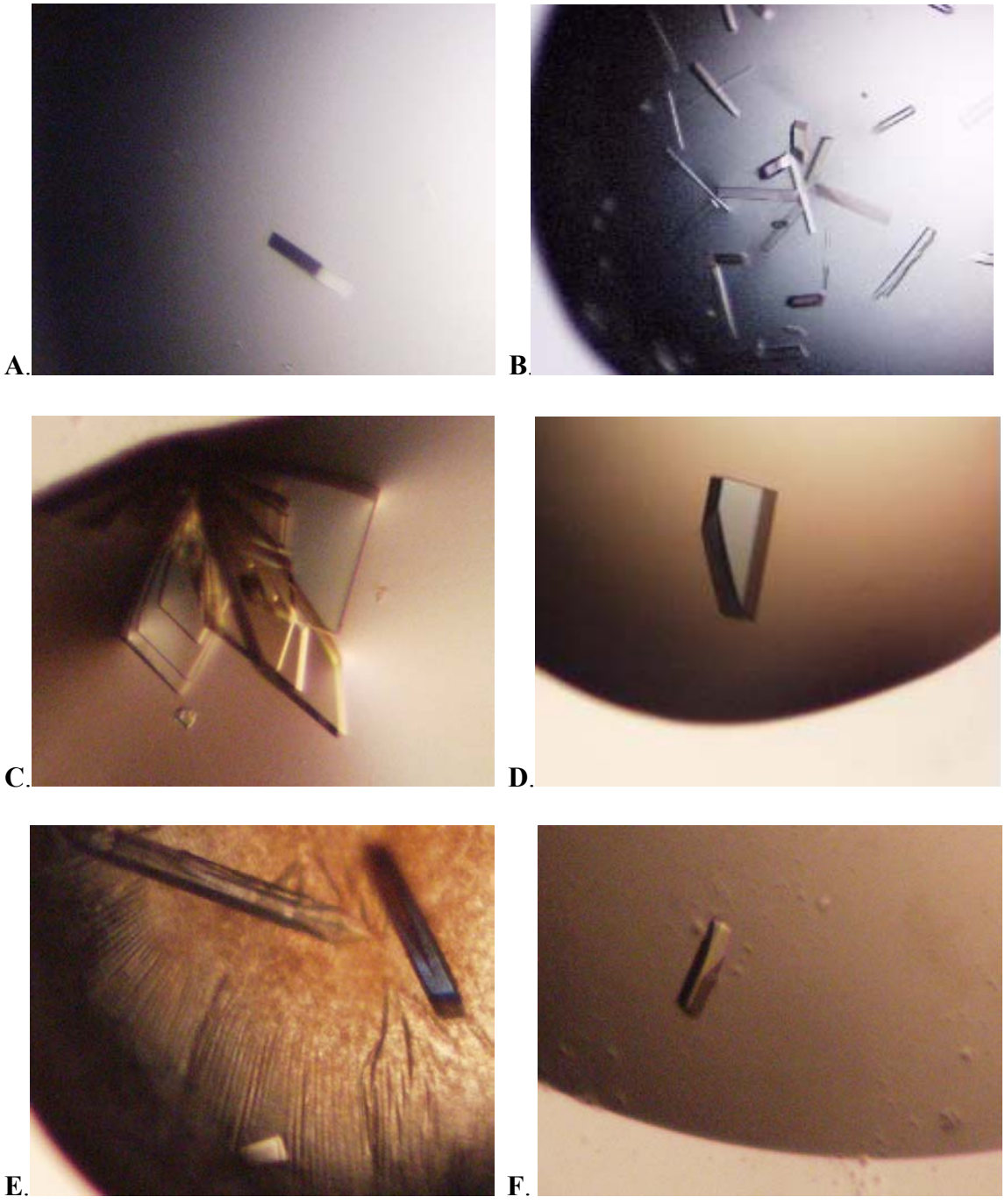


Figure 2.2 Crystal pictures of some complexes.
A. B. C. crystals for substrate analogs; D. crystal for PR_{V82A}/UIC-94017; E. crystals for PR_{V82A}/Saquniavir (space group $P2_12_12_1$); F. crystal for PR_{V82A}/Saquniavir (space group $P2_12_12$)

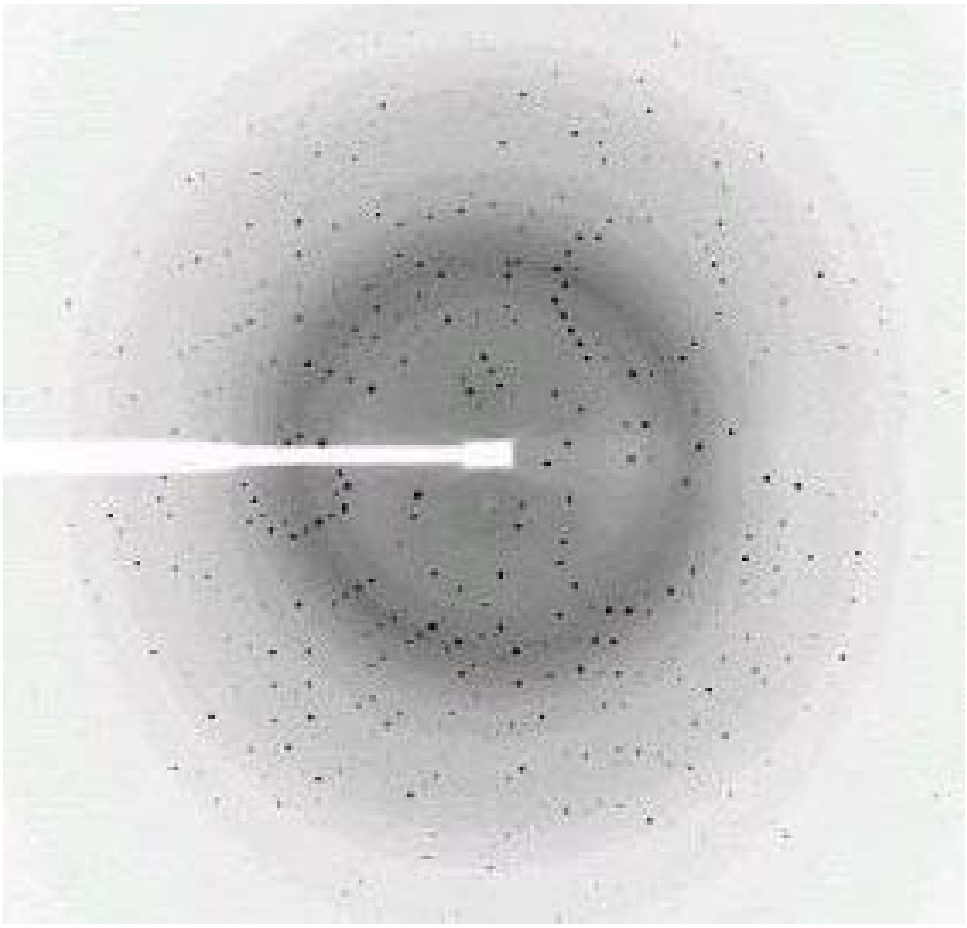


Figure 2.3 Example of a diffraction pattern of HIV-1 protease.
The dark spots are called reflections. The reflection spots are those X-rays that are diffracted after traveling through a crystal. The white shadow at the center of the image is produced by a beamstop which is used to block high energy X-rays from directly hitting the image plate detector to protect the detector.

2.4 Data processing and refinement

The datasets collected at the National Synchrotron Light Source were processed using the HKL suite 1.96, and the other datasets collected at the Advanced Photon Source were processed by HKL 2000 package (Otwinowski et al. 1997). Molecule replacement was done using AMoRe (Navaza 1994). The starting model for molecular replacement was chosen from the highest resolution structure available at the moment. The structures were refined using the program SHELXL (Sheldrick et al. 1997), and map display and refitting used the molecular graphics program O (Jones et al. 1991). Structure solution in the $P2_1$ space group was tested when the structures showed two alternate conformations of inhibitor in $P2_12_12_1$. Alternate conformations for PR residues, water and other solvent molecules were modeled when observed. The type of ion and other solvent molecules were identified by the shape of the 2Fo-Fc electron density map, the potential for hydrogen bonding, the coordination state, and inter-atomic distances, for molecules present in the crystallization conditions. Anisotropic B factors were applied. Hydrogen atoms were calculated in the last round of refinement by SHELXL (except for $p6^{pol}$-PR complexes which had low resolution of 1.42 and 1.6 Å).

Structures were compared by superimposing their main chain atoms using the algorithm described in (Mahalingam et al. 2004). Distances between atoms were calculated. Two atoms with separation of about 4 Å (3.85 Å is the average) are considered to form van der Waals interactions, while those of 2.6 to 3.2 Å (2.85 Å is the average) in distance and between a hydrogen bond donor atom and a hydrogen bond acceptor atom are hydrogen bonds. To analysis more accurately and efficiently, a small

database incorporated with basic structural analysis functions was designed to automate this process. The database interface was programmed by JAVA language (Sun Microsystems, Inc.). Structural figures were made using Molscript (Kraulis 1991), Bobscript (Esnouf 1997; Esnouf 1999) and Weblab viewer (Molecular Simulations Inc.). All structures and diffraction data were deposited to the PDB (Berman et al. 2000). The PDB codes are listed in Appendix 1.

2.5 Kinetic measurements

The kinetic parameters, inhibition constant K_i, for peptide analogs of the CA-p2, p2-NC, $p6^{pol}$-PR and p1-p6 natural PR cleavage sites, and the catalytic efficiency k_{cat}/Km were measured by UV spectroscopic assay with the chromogenic substrate (Lys-Ala-Arg-Val-Nle-p-nitroPhe-Glu-Ala-Nle-amide, L6525, Sigma), which is an analog of the CA-p2 cleavage site. The assay solution contained 50mM sodium acetate, pH 5.0, 0.1 M NaCl, 1mM EDTA. The reaction concentrations of enzyme and the substrate were 70 to 120 nM and 300 μM, respectively. The PR concentrations were determined by active site titrations with indinavir and the substrate concentration was determined by converting absorbance of the substrate to concentration via a calibration curve. The decrease in absorbance at 310 nm of the reaction mixture was measured on a Hitachi U-2000 spectrophotometer. Michaelis-Menten curves over 25-400 μM substrate and the inhibition curves were fit by SigmaPlot 8.0.2 (SPSS Inc.). Inhibition constants of each analog inhibitor were obtained from the IC_{50} values estimated from a dose-response curve

using the equation $Ki = (IC_{50} - 0.5*[E])/(1+[S]/Km)$, where [E] and [S] are the PR and substrate concentrations, respectively.

Kinetic parameters for small chemical inhibitors were determined by a fluorescence assay as described (Bagossi et al. 2004). A synthetic peptide, RE(Edans)SQNY-PIVRK(Dabcyl)R (Matayoshi et al. 1990; Wang et al. 1990) based on the HIV-1 MA-CA cleavage site in the Gag-Pol polyprotein was used. Protease 10 µl (final concentration of 7-12 nM) was mixed with 100 µl PNF buffer (250mM phosphate buffer, pH = 5.6 containing 500 mM NaCl, 1mM EDTA, 5 mM DTT and 5% glycerol) and pre-incubated at 37 ^{o}C for 5 min. The reaction was initiated by adding 90 µl substrate (final concentration of 2-40 µM). The mixture was assayed over 5 min for the increase in fluorescence, which was detected at 460 nm, using 355 nm excitation wavelength in a POLARstar OPTIMA fluorimeter-luminometer (BMG Labtechnologies Inc.). Inhibition assays were performed in the same way, but the reaction mixture contained 2 µl DMSO or inhibitor dissolved in DMSO. Inner filter effect was determined by measuring the fluorescence as a function of the concentration of RE (Edans), at the substrate concentration range used for the kinetic measurements. Data analysis was performed same way as discussed for the UV spectroscopic assay.

3. Results and discussion

3.1 Structural analysis of complexes with substrate analogs

Specific aim: To determine & analysis crystal structures of PR and mutants PR_{V82A} and PR_{I84V} with peptide analogs.

High resolution (1.1 – 1.6 Å) crystal structures of PR and the common drug resistant variants PR_{V82A} and PR_{I84V} were determined with reduced peptide analogs that represent the CA-p2, p2-NC, $p6^{pol}$-PR and p1-p6 polyprotein cleavage sites. These structures and kinetic data provide details of the PR interaction with reaction intermediates and better understanding of the substrate specificity. Furthermore, comparison of substrate analog complexes with clinical inhibitors has improved our knowledge of the molecular basis of drug resistance and will assist in the structure-based design of more potent anti-viral inhibitors.

Inhibition of PR, PR_{V82A} and PR_{I84V}

The reduced peptide analogs represented five different HIV-1 cleavage sites (Table 3.1.1). The p2-NC site is the first and CA-p2 the last in sequential processing of Gag precursor (Pettit et al. 1994). Cleavage of $p6^{pol}$-PR is essential for release of mature active protease (Louis et al. 1994). Mutations in the NC-p1 and p1-p6 sites contribute to drug resistance both *in vitro* and *in vivo* (Croteau et al. 1997; Zhang et al. 1997; Robinson et al. 2000). These two cleavage sites show significant sequence polymorphism (Barrie et al. 1996; Bally et al. 2000), and the specificity of cleavage has been studied with PR and several mutants (Feher et al. 2002). The peptide analogs differed in length. CA-p2

Table 3.1.1 Sequence of the substrate analog inhibitors and inhibition constants.

Values are listed for Ki (μM) and (Ki relative to PR).

Cleavage Site	Peptide Sequence	PR	PR_{V82A}	PR_{I84V}
CA-p2	R-V-L-r-F-E-A-Nle	**0.075 ± 0.009**	**0.024 ± 0.004 (0.3)**	**0.275 ± 0.031 (3.7)**
p2-NC	Ace-T-I-Nle-r-Nle-Q-R	**2.17 ± 0.28**	**0.53 ± 0.078 (0.24)**	**13.0 ± 1.5 (6.0)**
$p6^{pol}$-PR	V-S-F-N-F-r-P-Q-I-T-K-K	**22.1 ± 2.8**	**36.3 ± 5.4 (1.6)**	**46.6 ± 5.3 (2.1)**
p1-p6	R-P-G-N-F-r-L-Q-S-R-P	**96.7 ± 12.3**	**28.2 ± 4.2 (0.3)**	**> 500 (>5)**
NC-p1	E-R-Q-A-N-r-F-L-G-K-I	**> 500**	**> 500**	**> 500**

and p2-NC were the two shortest peptides, extending from P3 to P4' and P3 to P3', respectively. The analogs NC-p1 and p1-p6 extended from P5 to P5', while $p6^{pol}$-PR extended from P5 to P6' since lysine was added to provide greater solubility.

The catalytic activities of PR and the mutants PR_{V82A} and PR_{I84V} were competitively inhibited by the five substrate analogs (Table 3.1.1). Previous studies demonstrated that reduced peptide bond-containing inhibitors act in a competitive manner on HIV-1 PR, and the same type of inhibition was assumed in this study for the mutants (Tozser et al. 1991a). The inhibition constants for PR were in the order CA-p2 < p2-NC < $p6^{pol}$-PR < p1-p6 < NC-p1. The CA-p2 analog was the best inhibitor of PR and the mutants, while the NC-p1 analog was the weakest inhibitor for all enzymes with Ki > 500 μM. PR_{V82A} was better inhibited than PR (about 3-fold) by all the analogs, except for $p6^{pol}$-PR. PR_{I84V} was poorly inhibited relative to PR for all analogs with 2 to 6-fold higher Ki values. This variation in Ki was smaller than observed for the clinical inhibitors that showed 2 to 11-fold relative inhibition of the mutants PR_{V82A} or PR_{I84V} compared to wild type PR (Klabe et al. 1998).

Description of the high resolution crystal structures

Eight crystal structures have been determined of PR and drug resistant mutants PR_{V82A} and PR_{I84V} in the complexes with four different substrate analogs. Crystallographic statistics are summarized in Table 3.1.2. Seven of these are new structures, while PR/CA-p2 complex was determined at the higher resolution of 1.4 Å compared to 1.9 Å for the previously reported structure (Mahalingam et al. 2001).

Table 3.1.2 Crystallographic data statistics.

Protease		***PR$_{V82A}$**	***PR**	***PR$_{V82A}$**	**PR$_{I84V}$**	**$^{\#}$PR**	**$^{\#}$PR$_{V82A}$**	***PR**	**PR$_{V82A}$**
Inhibitor		CA-p2	p2-NC	p2-NC	p2-NC	p6^{pol}-PR	p6^{pol}-PR	p1-p6	p1-p6
Space group		$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$
Unit cell dimensions (Å)	**a**	57.89	58.00	58.02	57.80	59.45	58.88	58.91	58.46
	b	85.96	85.78	85.89	85.59	87.00	86.27	86.07	85.85
	c	46.19	46.53	46.61	46.46	46.32	46.40	46.54	46.39
Resolution range (Å)		50-1.54	50-1.40	50-1.10	50-1.30	50-1.60	50-1.42	50-1.38	50-1.32
Unique reflections		34544	44291	95318	55009	32005	40847	47418	55317
R_{merge} (%) **Overall (final shell)**		6.5 (37.1)	10.1 (45.6)	10.2 (37.7)	7.9 (33.0)	9.1 (41.3)	9.8 (70.8)	10.4 (28.6)	8.4 (57.7)
<I/sigma> **Overall (final shell)**		13.4 (5.8)	8.9 (2.1)	10.4 (2.1)	14.4 (3.2)	13.2 (3.7)	17.1 (2.6)	14.1 (13.8)	10.4 (2.9)
Data range for refinement (Å)		10-1.54	10-1.40	10-1.10	10-1.30	10-1.60	10-1.42	10-1.38	10-1.32
R_{work} (%)		0.12	0.15	0.13	0.12	0.15	0.18	0.16	0.13
R_{free} (%)		0.19	0.19	0.17	0.16	0.23	0.23	0.21	0.18
No. of waters (total occupancies)		194	188	206	223.5	155	199	190	297
Completeness (%) **Overall (final shell)**		99.6 (100)	95.6 (66.6)	94.8 (68.4)	96.0 (72.4)	99.2 (93.1)	90.3 (100)	89.2 (96.2)	99.9 (100)
RMS deviation from ideality									
Bonds (Å)		0.010	0.011	0.015	0.013	0.009	0.011	0.010	0.012
Angle distance (Å)		0.029	0.029	0.033	0.030	0.029	0.031	0.030	0.034
Average B-factors (Å^2)									
Main chain		15.9	8.0	9.0	10.6	22.4	16.7	14.2	12.0
Side chain		21.9	12.2	13.7	15.4	27.6	21.1	18.4	17.0
Inhibitor		27.5	14.0	10.6	14.7	29.9	19.1	18.3	24.0
Solvent		31.5	23.7	24.4	25.8	37.9	30.9	27.3	26.7

*: Diffraction data collected at Advanced Photon Source, beamline SER-CAT 22. All other data were collected at National Synchrotron Light Source, beamline X26C. $^{\#}$: Structures in which hydrogen atoms were not added.

Crystals and diffraction data were obtained for PR with the NC-p1 analog; however, the electron density was disordered and not interpretable for the analog probably due to weak binding and consistent with the high Ki values of >500 μM. The asymmetric unit of the crystals contained a PR dimer with the residues in the two subunits numbered 1-99 and 1'-99'. All structures are in space group $P2_12_12$ and were refined to R-factors of 0.12-0.18 including solvent molecules, and anisotropic B-factors. The resolution ranges from 1.10-1.60 Å. The complex PR_{V82A}/p2-NC was determined at 1.1 Å resolution, which is the highest resolution to date for a substrate analog complex. The quality of the electron density map is shown in Figure 3.1.1. These crystal structures showed clear electron density for all the protease atoms, P4-P4' residues in the peptide analog, and solvent molecules. All the peptide analogs, except for p2-NC, showed two pseudosymmetric conformations bound to both subunits of the PR dimer. The $p6^{pol}$-PR and p1-p6 analogs have 11 and 10 residues, respectively, compared to 6-7 for the CA-p2 and p2-NC analogs. The longer analogs extended out of the PR binding pocket, and showed poor electron density at both termini. The average B-factors ranged from 8.0 to 22.4 $Å^2$ for protein main chain atoms and 10.6 to 29.9 $Å^2$ for protein side chain and inhibitor atoms. The lower average B-factors were observed for the higher resolution structures.

Alternate conformations were modeled for the side chain atoms of about 30 residues in all the crystal structures based on the shape of the electron density (Figure 3.1.2). Only Lys 7 had alternate conformations in both subunits of all structures, while Met 46 had alternate conformations in all but one subunit. Most of these alternate conformations were observed for residues with longer and flexible side chains, such as

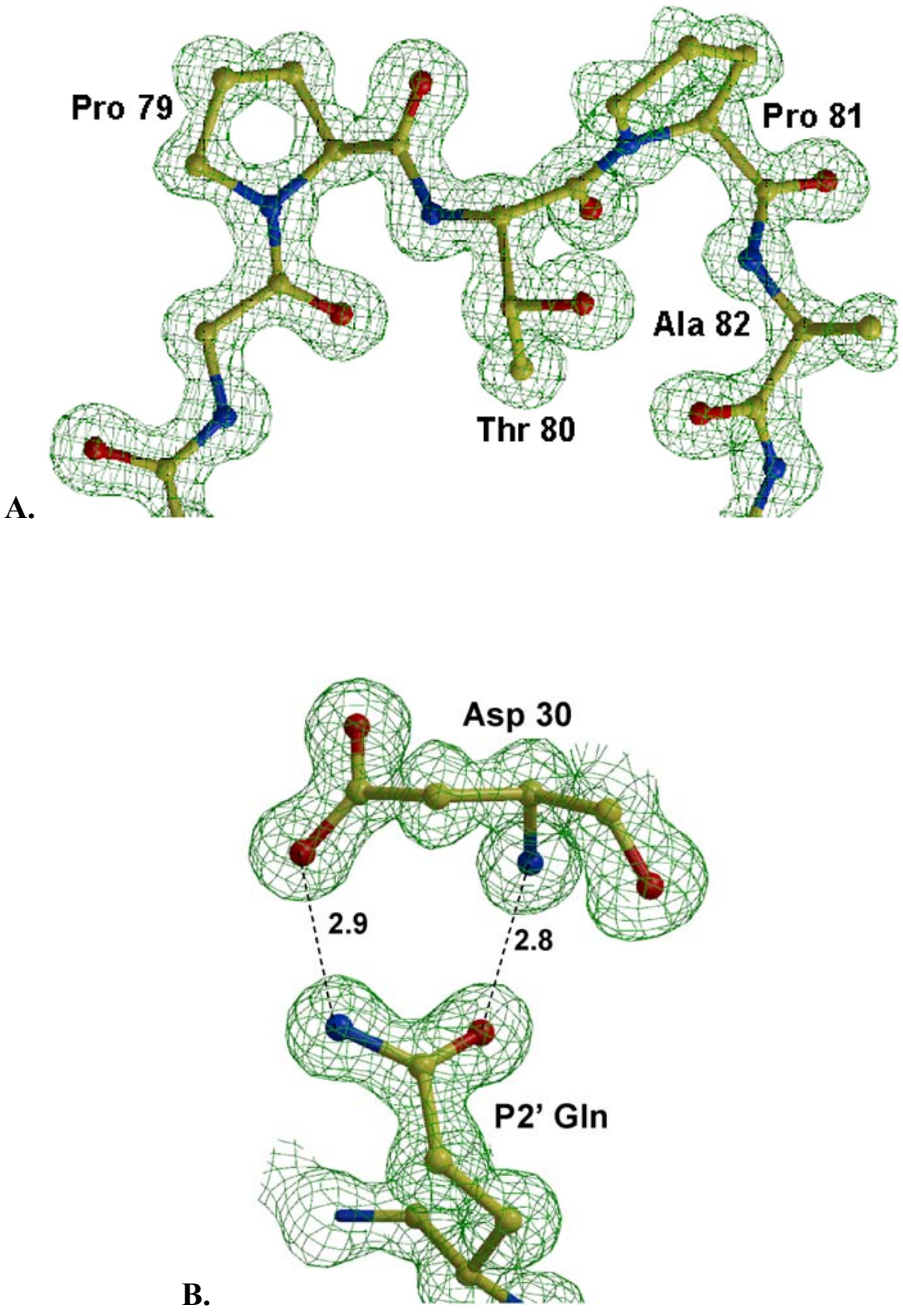


Figure 3.1.1 Electron density map for PR_{V82A}/p2-NC crystal structure.
The 2Fo−Fc map was contoured at a level of 2.2σ.
A. Residues 78-82.
B. Asp 30 interacting with P2' Gln.

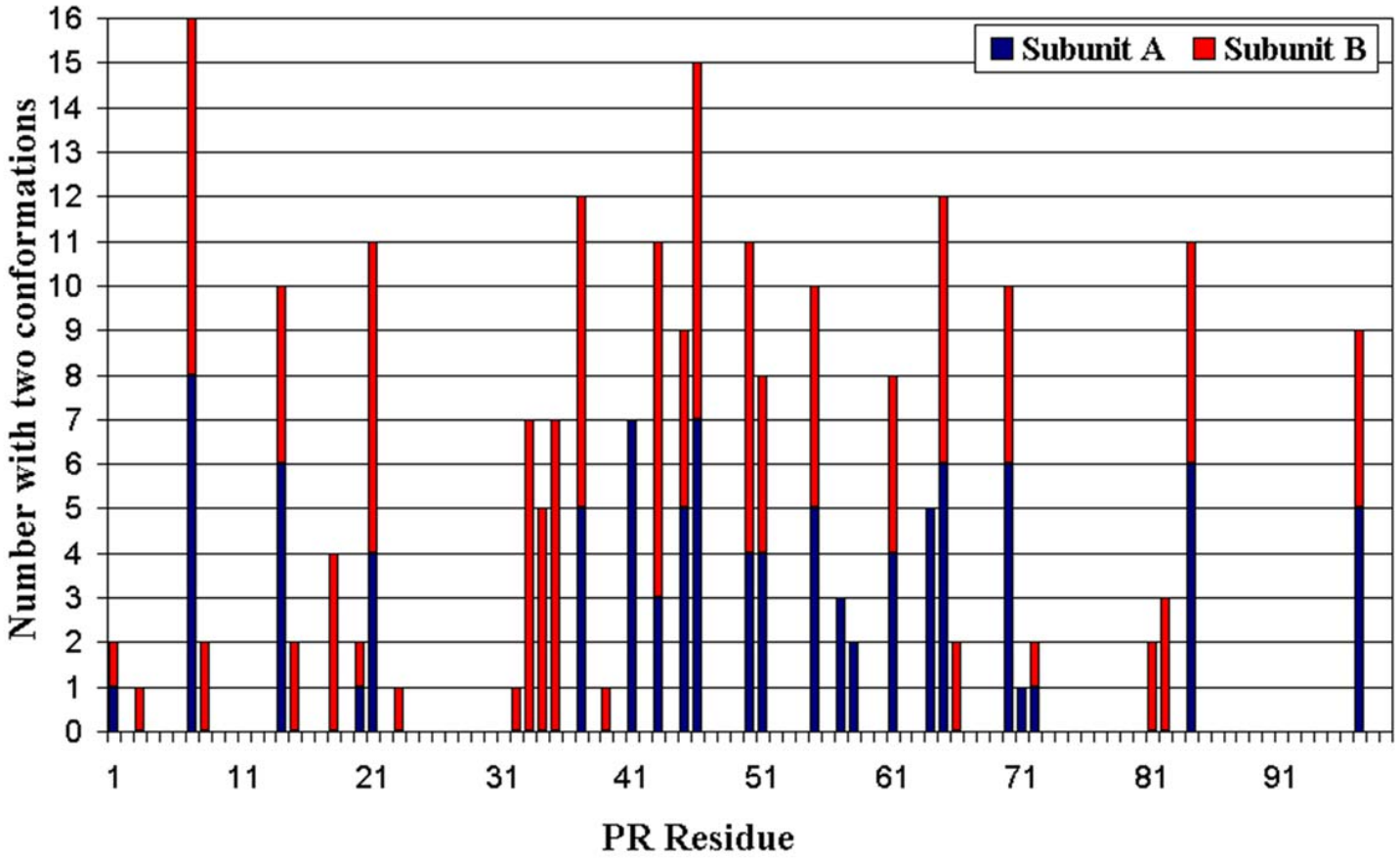


Figure 3.1.2 Statistics of the residues with alternate conformations.

Lys and Glu, while alternate conformations were observed also for hydrophobic residues including Leu, Ile and Val. The number of alternate conformations was in the order of Lys (68 with alternate conformations), Ile (41), Glu (35), Met (15), followed by Gln, Arg, Ser and Leu at about 10 each. Some residues including 33', 34' and 35' were observed to have two conformations only in one subunit of the dimer. These residues were located on the surface of structure so they are either very flexible or involved in interactions with symmetry related molecules. The presence of alternate conformations of side chains for Leu 23, Lys 45/45', Met 46/46', Ile 50/50', Val 82/82' and Ile 84/84' in the inhibitor binding site was consistent with previous descriptions (Mahalingam et al. 2001; Tie et al. 2004). Alternate positions with a 180 degree flip of the main chain atoms of Ile 50 and 50' were observed in complexes PR/p1-p6, PR_{V82A}/p1-p6, PR/$p6^{pol}$-PR and PR_{V82A}/$p6^{pol}$-PR, as described for the PR structures with the potent antiviral inhibitor UIC-94017 (Tie et al. 2004).

Overall comparison of the crystal structures

All the PR dimers shared very similar backbone structures with RMS (root-mean-square) deviation for all Cα atoms ranging from 0.09 to 0.26 Å compared to the PR/p2-NC complex. The least variation was observed for the complexes with p2-NC. The deviations for residues in subunit A (within 1 Å) were larger than those of subunit B (within 0.8 Å). Larger deviations for all the structures were located at external loop residues 38 to 41 and residues 79-84 near the mutations. The biggest difference from PR/p2-NC was observed for the complexes with $p6^{pol}$-PR analog, and occurred at

residues 47 to 53 in the hairpin loop that links two β-strands in the flap in both subunits. Residues 25 to 28 at the active site had the least deviation in both subunits for all structures.

Protease interactions with substrate analogs

This series of high resolution crystal structures allowed more precise description of the PR interactions with transition state mimics. Important conserved features of the PR interaction with substrate analogs of size P4-P4' have been identified from crystal structures (Gustchina et al. 1994; Mahalingam et al. 2001). These new structures can extend the description of the interactions to more distal residues with greater accuracy. Peptide analogs are bound with each side chain from P5 to P5' lying within a set of subsites S5 to S5' formed by PR residues. Figure 3.1.3 shows the superimposed substrate analogs in the PR_{V82A} complexes. The P2-P2' side chains were in similar positions while the more distal residues had greater conformational variation. PR recognizes substrates by means of a series of hydrogen bond interactions with the main chain atoms of the peptide (Figure 3.1.4). All the structures showed similar hydrogen bond interactions between PR and inhibitor in P3-P3' positions, while there was more variation at the distal ends. Two water molecules are conserved in all eight structures and mediate the interactions between PR and inhibitor. One conserved water lies between the flap region (Ile 50 and 50') and P2 and P1' of inhibitor, which has been proved to be important for catalysis (Grzesiek et al. 1994; Wang et al. 1996), the other mediates interactions of P2' with Gly 27' and Asp 29'.

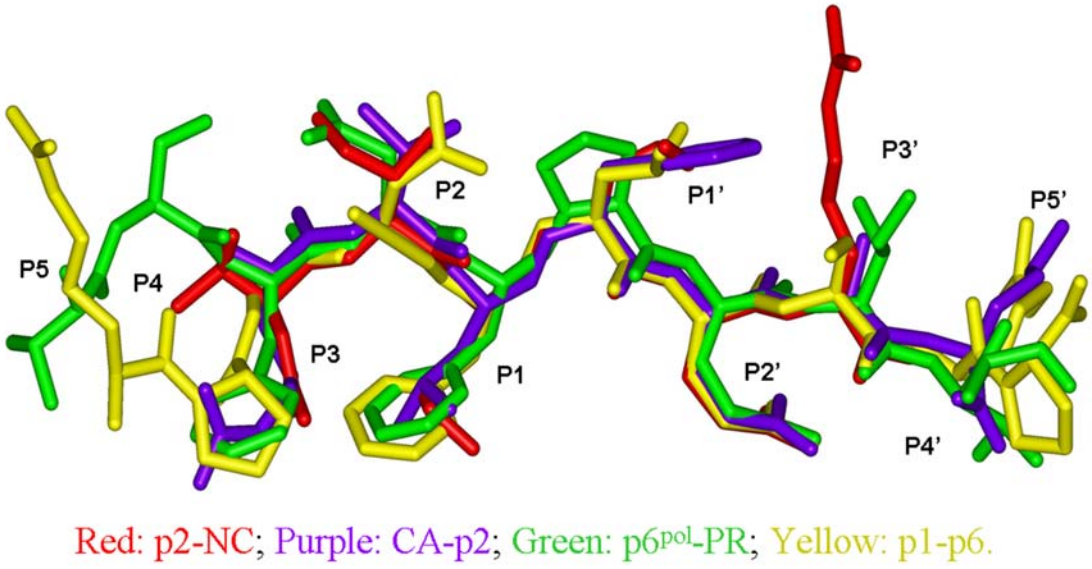
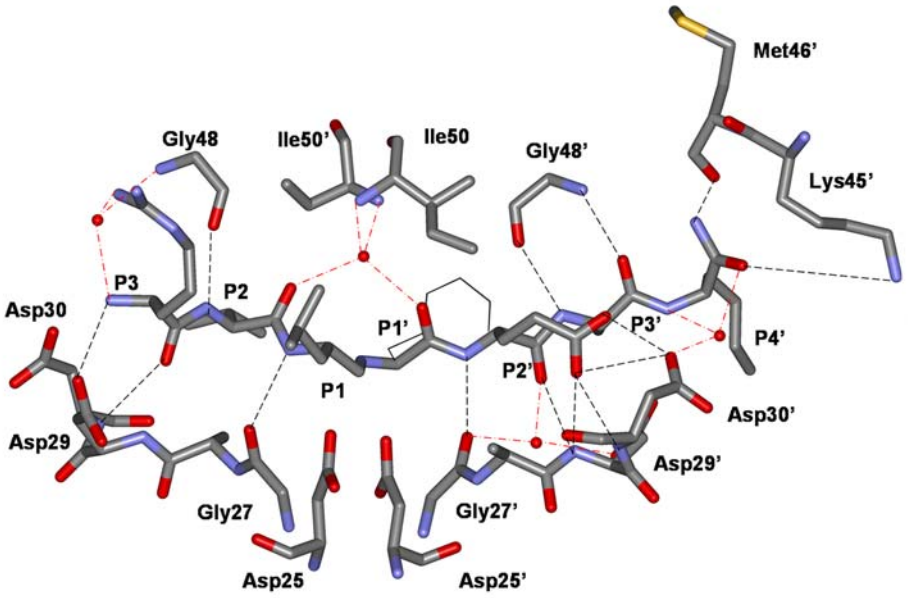
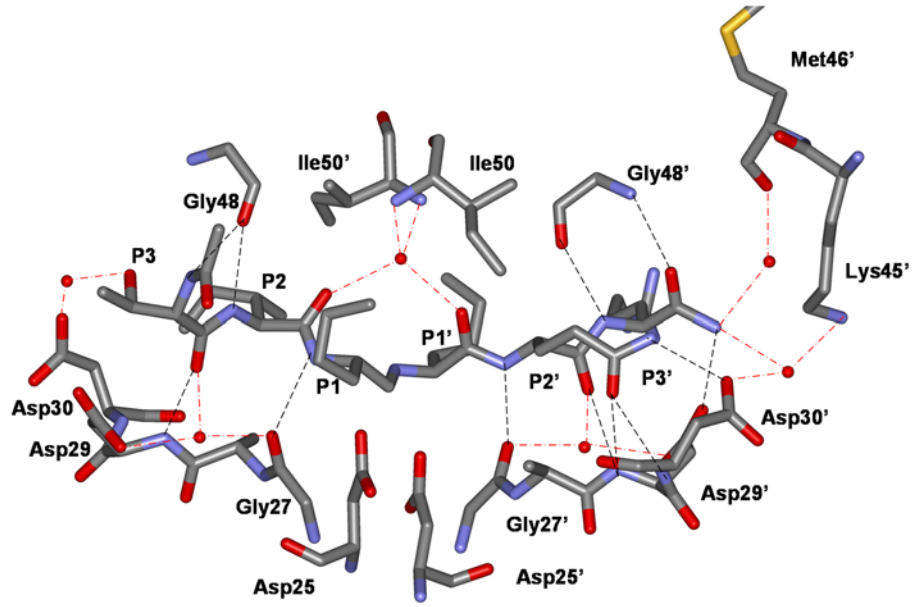


Figure 3.1.3 Superposition of four PR_{V82A} complexes with CA-p2, p2-NC, $p6^{pol}$-PR and p1-p6.

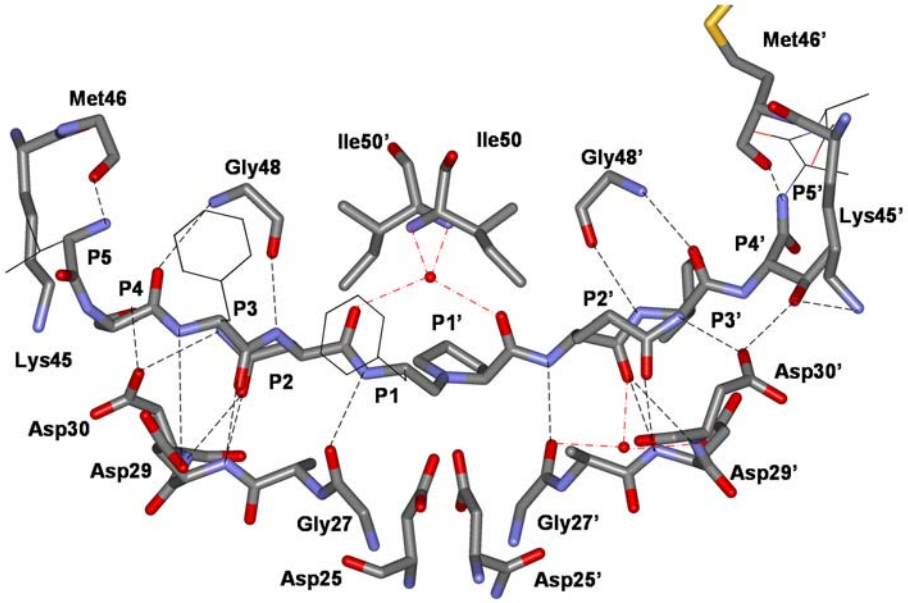


A.

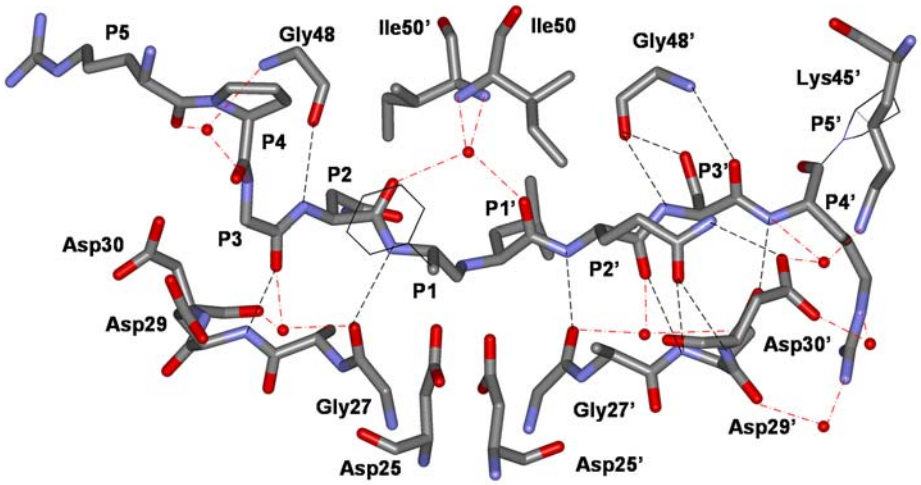


B.

Figure 3.1.4 Hydrogen bond interactions between protein and inhibitor.
Hydrogen bond interactions are considered for interatomic distances of 2.5-3.3 Å. Water molecules are indicated by red spheres. Water-mediated hydrogen bonds are shown in red dashed lines, while direct interactions between protease and inhibitor are in black.
A. Hydrogen bond interactions between PR_{V82A} and CA-p2. (One water-mediated interaction between P3Arg and Pro 81' is not shown.)
B. Hydrogen bond interactions between PR and p2-NC. (Water-mediated interactions of both termini of inhibitor with Arg 8 and 8' are not shown.)



C.



D.

Figure 3.1.4 Hydrogen bond interactions between protein and inhibitor.
C. Hydrogen bond interactions between PR and p6^{pol}-PR. (Water-mediated interactions of C termini of p6^{pol}-PR with Asp 60 and Gln 61 are not shown.)
D. Hydrogen bond interactions between PR$_{V82A}$ and p1-p6. (Water-mediated interactions of C termini of p1-p6 with Trp 6, Arg 8 and Arg 87' are not shown.)

Complex with CA-p2 analog

The CA-p2 analog was bound to PR_{V82A} in two orientations with relative occupancy of 0.65/0.35. The complex structure was similar to the structure of PR/p2-NC with RMS deviation on alpha carbon atoms of 0.20 Å. Residues P3-P4' of CA-p2 interacted with PR (Figure 3.1.4A). Compared to p2-NC, the CA-p2 analog lacked an acetyl group at P4 and cannot form the same van der Waals interactions with PR. However, the CA-p2 analog with P2' Glu had two proton-mediated hydrogen bond interactions with the Asp 30 carboxylate side chain, instead of the single hydrogen bond of P2' Gln in p2-NC (Figure 3.1.1B), as observed previously (Mahalingam et al. 2001). The Nle at P4' position in CA-p2, instead of the NH_2 in p2-NC, allowed formation of hydrogen bonds with the carbonyl oxygen of Met 46' and the side chain of Lys 45'. Furthermore, P3 is Arg in CA-p2 and Thr in p2-NC. As a result, the carbonyl oxygen of P3 Arg interacted with the amide of Asp 29 instead of the interaction of the amide of P3 Thr with carbonyl oxygen of Gly 48 in the flap. Also, the longer side chain of Arg provided more van der Waals interactions with the PR. These differences agreed with the 25-fold stronger inhibition observed for the analog CA-p2 compared to p2-NC.

Structural comparison of the complexes with p2-NC analog

Crystal structures were determined of complexes of the substrate analog p2-NC with PR and the two mutants PR_{V82A} and PR_{I84V}. The p2-NC showed one conformation in all three protease complexes. The interactions between PR and analog extended to P4-P4' with the P4 acetyl group and there was clear density for all the residues. All three

structures were very similar with RMS deviation of 0.09 and 0.11 Å for PR_{I84V} and PR_{V82A}, respectively. There were conserved hydrogen bond interactions with main chain amide and oxygen atoms extending from P3 O to P4' N, and the P2' Gln side chain formed hydrogen bond interactions with Asp 29' and Asp 30' in all three structures (Figure 3.1.4B). Multiple conformations were modeled for the side chain of P1' Nle depending on the shape of the electron density in the complexes with mutants. The main chain oxygen of P3 Thr had water mediated interactions with Gly 27 and Asp 29. Moreover, the hydroxyl oxygen of P3 Thr interacted with Asp 30 via a water molecule. These two waters may stabilize the protease inhibitor interactions, as suggested previously (Mahalingam et al. 2001).

There were small compensating changes in the interactions with p2-NC in the complexes with PR_{V82A} and PR. P3' in PR_{V82A}/p2-NC showed more interactions with water molecules than observed in PR complex. Alternatively, it is possible that more water molecules were identified due to the higher resolution of the PR_{V82A}/p2-NC complex. In mutant PR_{V82A}, Ala 82 had lost the van der Waals interaction with P4 Ace, and had weaker interactions with P1 (interatomic distances were longer than 4 Å). However, the interactions at P1'and P3' were enhanced (Figure 3.1.5A) mainly by the shifts of the side chains of P1 Nle and P3' Arg and partially by the small (0.3 Å) shift of the CA atom of Ala 82/82'. The side chain of P1' Nle showed three orientations which formed closer contacts with the CB atom of Ala 82 in PR_{V82A} than observed with Val 82 in PR. The CE atom of P3' Arg moved about 1.2 Å and formed closer interactions with CB atom of Ala 82. All these observed structural changes and closer van der Waals

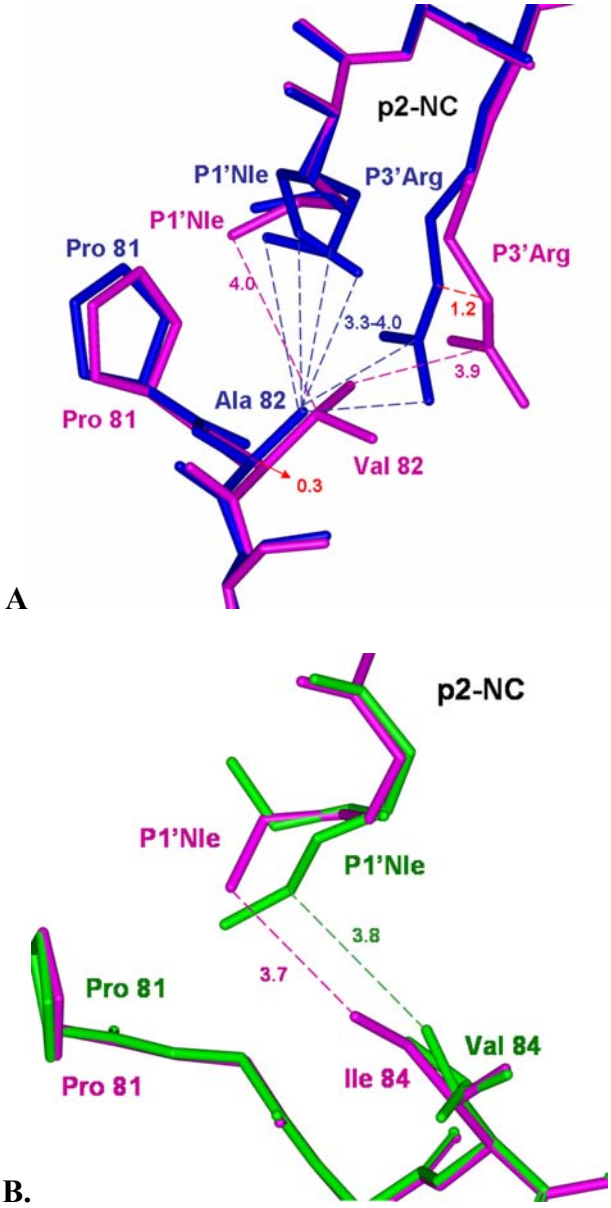


Figure 3.1.5 Structural variation around residues 81–84 in p2-NC, p6^{pol}-PR, p1-p6, UIC-94017 and indinavir complexes.
The PR structure is in purple, PR_{I84V} in green and PR_{V82A} in blue bonds. Interatomic distances (Å) are indicated by dashed lines.

A. PR_{V82A}/p2-NC superimposed on PR/p2-NC.
B. PR_{I84V}/p2-NC superimposed on PR/p2-NC.

interactions with p2-NC were consistent with 4-fold better inhibition for PR_{V82A} than PR (Table 3.1.1). Small changes in the backbone atoms of residue 82 in mutant PR_{V82A} in complex with non-peptidic inhibitors, have been described (Baldwin et al. 1995).

PR_{I84V}/p2-NC had lost some interactions with p2-NC due to the mutation. Val 84 in PR_{I84V}/p2-NC had fewer interactions with P2, compared to those of Ile 84 in the PR/p2-NC structure. Similar to the PR_{V82A} complex, the flexibility of P1'Nle side chain had compensated partially for the loss of van der Waals interactions caused by shortening the side chain from Ile to Val (Figure 3.1.5B). These changes also agreed with the 6-fold worse inhibition of p2-NC for PR_{I84V} than PR. These observations suggested that analog p2-NC modulated the binding affinity for mutants through small conformational changes of the side chain of P1' Nle. Similarly, the conformational flexibility of the Met side chain in the natural substrate is expected to compensate for the drug resistant mutations like V82A and I84V.

Structural comparison of the complexes with $p6^{pol}$-PR analog

The two structures of PR and PR_{V82A} with $p6^{pol}$-PR analog were very similar with RMS deviation on alpha carbon atoms of 0.17 Å, and were similar to PR/p2-NC structure with RMS deviation values of about 0.25 Å. The $p6^{pol}$-PR showed alternative conformations with relative occupancies of 0.6/0.4 and 0.8/0.2 for PR and PR_{V82A}, respectively. Since the $p6^{pol}$-PR analog had eleven residues and extended out of the PR binding pocket, both N and C terminal residues were quite flexible with poor electron density. Due to the absence of electron density for the side chain at the C-terminus Ala

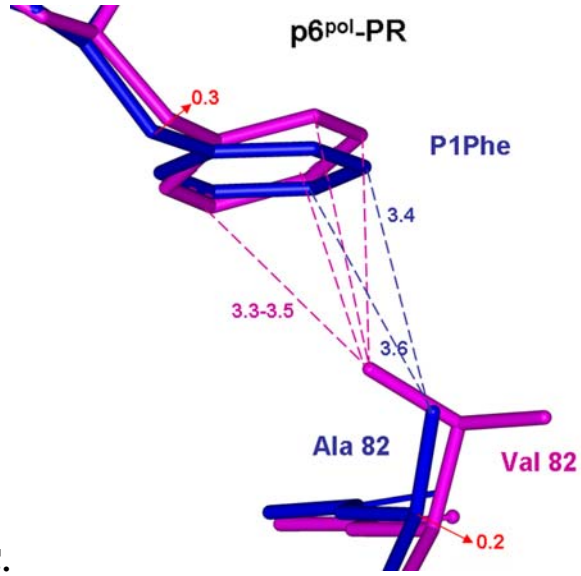
was modeled instead of P6' Lys. On the other hand, the longer peptide provided interactions extending from P5 to P5'as illustrated in Figure 3.1.4C. P2 is the polar Asn, unlike the hydrophobic Val and Ile in CA-p2 and p2-NC. Hence, the side chain of P2 Asn formed hydrogen bonds with Asp 29 and Asp 30, which cannot occur in CA-p2 or p2-NC. The $p6^{pol}$-PR had the larger hydrophobic Phe at P3 compared to Arg, Thr or Gly in the other analogs. Therefore, P3 Phe occupied more space in the binding pocket, and fewer water-intermediated interactions were observed. The smaller amino acid Pro at P1' resulted in fewer van der Waals contacts with PR than for other analogs with Phe, Nle or Leu at P1'. The two complexes showed the largest deviation from PR/p2-NC for the residues 48-52 in the flap region of both subunits. This structural change in the flaps and the differences in interactions with the substrate side chains were consistent with the poorer inhibition of PR by $p6^{pol}$-PR compared to CA-p2 and p2-NC analogs.

PR_{V82A} showed almost identical hydrogen bond and van der Waals interactions with $p6^{pol}$-PR compared to PR, except for interactions with the terminal P5Val and P4'Thr. As noted previously, the N-terminus was very flexible and had two different orientations when bound to PR or $PR_{V82A.}$ Thus, the N-terminal residues had van der Waals interactions with totally different residues in the two complexes. In the case of PR_{V82A}, the N-terminus had lost the hydrogen bond at P5 position and, instead, had a water-mediated interaction of P4 with Met 46. PR/$p6^{pol}$-PR showed interactions of the C-terminus of $p6^{pol}$-PR with Asp 60 and Gln 61 through a water molecule, while PR_{V82A}/$p6^{pol}$-PR did not have those interactions. At the mutation site 82, however, most interactions between protease and analog were conserved. Residue 82 interacted with P1

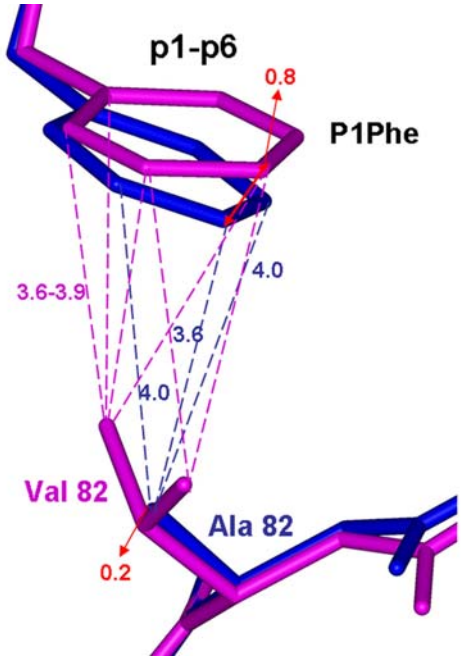
and P1' of $p6^{pol}$-PR and small shifts were observed for both Ala 82 and P1Phe in $PR_{V82A}/p6^{pol}$-PR compared to the positions in the PR complex (Figure 3.1.5C). These structural changes resulted in good van der Waals interactions of Ala 82/82' CB atoms with P1'Pro/P1Phe and compensated for the loss of the methyl groups of Val 82 in PR. The structural adjustment of the PR_{V82A} mutant to accommodate inhibitor binding was consistent with the similar inhibition constants observed for PR_{V82A} and PR with $p6^{pol}$-PR (36 and 22 μM, respectively).

Structural comparison of the complexes with p1-p6 analog

The two complexes of PR/p1-p6 and PR_{V82A}/p1-p6 were very similar with only small changes in the interactions with p1-p6. The RMS deviation values on alpha carbon atoms were 0.17-0.18 Å between these two complexes and PR/p2-NC. Analog p1-p6 was bound to PR and PR_{V82A} in two orientations with relative occupancy of 0.6/0.4. Residues P5-P5' of p1-p6 interacted with PR and PR_{V82A} (Figure 3.1.4D). The N terminal P4 and P3 of p1-p6 showed similar hydrogen bond and van der Waals interactions with protease as in the p2-NC complexes, however, these differed from the interactions of $p6^{pol}$-PR. The long side chain of P4' Arg at the C-terminus formed extra water-mediated interactions with PR residues Trp 6, Arg 8, Asp 29', Asp 30' and Arg 87'. The major difference from other substrate analogs was the presence of the small Gly at the P3 position in p1-p6. The P3 Gly had fewer van der Waals interactions with PR, and p1-p6 had more space to move around the binding pocket. As a result, although both p1-p6 and $p6^{pol}$-PR had Asn at P2, it showed different hydrogen bonds with PR. In $p6^{pol}$-PR, the big



C.



D.

Figure 3.1.5 Structural variation around residues 81 84 in p2-NC, p6^{pol}-PR, p1-p6, UIC-94017 and indinavir complexes.
The PR structure is in purple, PR_{I84V} in green and PR_{V82A} in blue bonds. Interatomic distances (Å) are indicated by dashed lines.

C. PR_{V82A}/p6^{pol}-PR superimposed on PR/ p6^{pol}-PR.
D. PR_{V82A}/p1-p6 superimposed on PR/p1-p6.

ring of P3Phe restricted movement in the binding site and pushed P2Asn more towards the active site, which enabled P2Asn to form hydrogen bonds with Asp 29 and Asp 30. While, with Gly at P3 the backbone of p1-p6 had moved in the binding site and provided more flexibility for P2Asn. The side chain of P2Asn in p1-p6 adopted two conformations differing by a rotation of about 90 degrees. One conformation of P2Asn maintained weaker hydrogen bonds with Asp 29 and 30, while the other conformation was surrounded by the hydrophobic side chains of Ile 50', Ile 84, and P1'Leu. Also, there were more water-intermediated interactions of PR with p1-p6. The loose binding of PR and p1-p6 primarily due to P3Gly was consistent with its more than 50 times worse inhibition than that of CA-p2 and p2-NC.

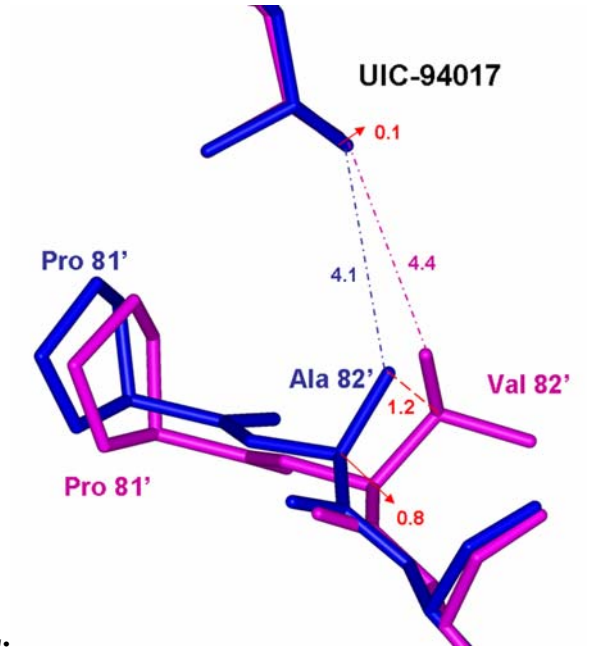
PR_{V82A} showed similar movements of Ala 82 and p1-p6 compared to PR/p1-p6 (Figure 3.1.5D) as described for the complexes with p2-NC and $p6^{pol}$-PR. Again, these subtle structural changes allowed formation of improved van der Waals interactions between PR_{V82A} and P1' and P1 of the substrate analog. The improved interactions with substrate analog observed in PR_{V82A} were compatible with its three times better inhibition than PR, and with the higher relative k_{cat}/Km for hydrolysis of the p1/p6 substrate (Feher et al. 2002).

PR interactions with substrate analogs compared to those with clinical inhibitors

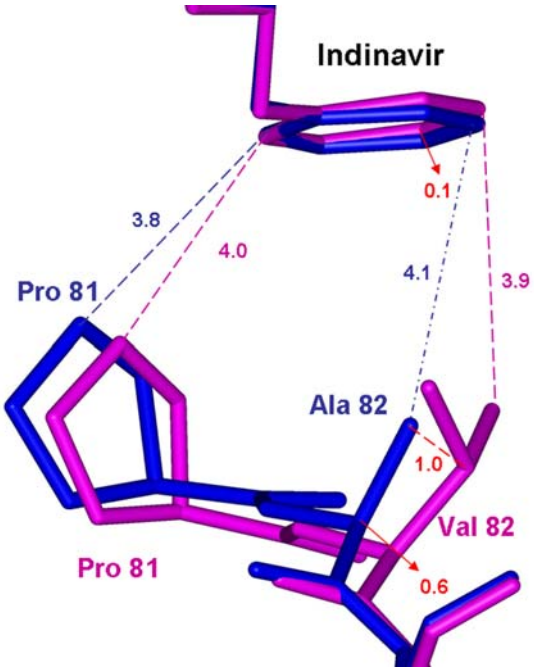
Substrate analogs showed more flexibility than clinical inhibitors in binding to the mutant PRs. The high resolution crystal structures of PR, PR_{V82A} and PR_{I84V} complexes indicated that the binding affinity for mutants was modulated by the conformational

flexibility of P1 and P1' side chains in the substrate analogs (Figure 3.1.5). It is instructive to compare the PR and mutant complexes with the clinical inhibitors. The crystal structures of PR, PR_{V82A} and PR_{I84V} with UIC-94017, an inhibitor in phase IIB clinical trials, and PR, PR_{V82A} and PR_{L90M} with the drug indinavir, were determined at resolutions of 1.1 – 1.6 Å (Mahalingam et al. 2004; Tie et al. 2004). All these structures were superimposed on PR/UIC-94017 with RMS deviations on alpha carbon atoms of 0.15 to 0.25 Å. The clinical inhibitors maximize the interactions within PR subsites S2 to S2', while the longer substrate analogs have more extended interactions within S4 to S4'. The UIC-94017 is smaller than the substrate analogs but formed similar hydrogen bonds to PR main chain atoms. Compared to indinavir and other clinical inhibitors, UIC-94017 formed additional polar interactions with the main chain atoms of Asp 29 and Asp 30 (Mahalingam et al. 2004; Tie et al. 2004). These interactions resembled those of the P2' Gln or Glu side chain of peptide analogs (Figures 3.1.1B and 3.1.5).

Similar rearrangements of residue 82/82' and P1/P1' were observed in PR_{V82A} complexes compared to PR complexes (Figure 3.1.5). These shifts allowed closer contacts of Ala 82 and 82' with the inhibitor, and partially compensated for the smaller side chain of Ala compared to wild type Val. However, Ala 82/82' showed smaller shifts (0.1–0.4 Å of Cα) with substrate analogs and larger changes (0.5-0.8 Å) with clinical inhibitors. These changes were coupled with larger movements or multiple conformations of P1/P1' side chains in substrate analogs (Figures 3.1.5A-D) than observed for the inhibitors UIC-94017 or indinavir (Figures 3.1.5E and F). The flexible side chain of P1'Nle in p2-NC adopted different conformations to maintain the interactions with



E.



F.

Figure 3.1.5 Structural variation around residues 81 84 in p2-NC, p6^{pol}-PR, p1-p6, UIC-94017 and indinavir complexes.
The PR structure is in purple, PR_{I84V} in green and PR_{V82A} in blue bonds. Interatomic distances (Å) are indicated by dashed lines.

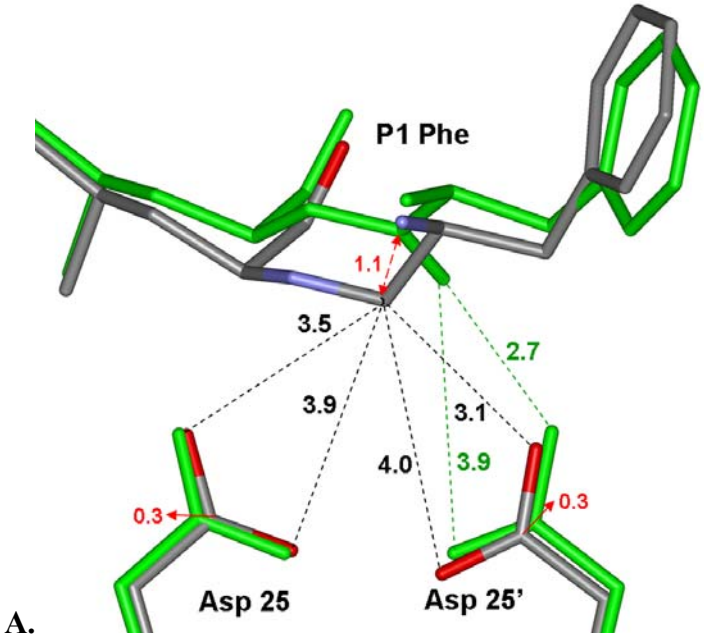
E. PR_{V82A}/UIC-94017 superimposed on PR/ UIC-94017.
F. PR_{V82A}/indinavir superimposed on PR/indinavir.

PR_{V82A}. P1Phe in $p6^{pol}$-PR and p1-p6 showed adjustments to interact better with Ala 82. In contrast, Val 84 in PR_{I84V} was less flexible than Ala 82 and adaptation in the PR_{I84V}/p2-NC complex was due to the alternate conformations of P1'Nle (Figure 3.1.5B). Consequently, similar K_i values for PR and mutant PR_{V82A} were observed for both substrate analogs and UIC-94017 (0.3 to 1.6-fold) and increased by 3-fold for indinavir (Wang et al. 1996), while the Ki increased from 2 to 6-fold for PR_{I84V} (Tie et al. 2004). Similar structural changes were reported for the inactive double mutant V82A/D25N compared to D25N mutant in complexes with peptides or ritonavir (Prabu-Jeyabalan et al. 2004). These observations suggested that the substrate analogs have more flexibility to accommodate the structural changes caused by mutation of PR. Hence, the comparison of PR complexes with substrate analogs or drugs helps to explain how the virus can develop drug resistance while retaining the ability to catalyze the hydrolysis of natural substrates.

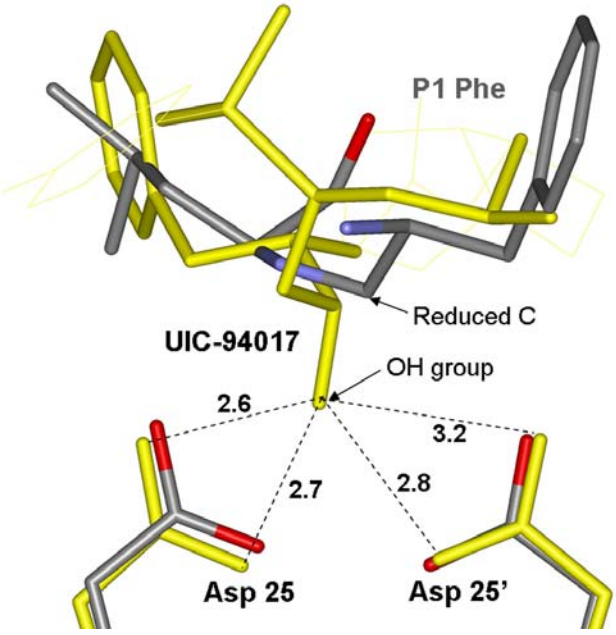
Structure of active site and implications for reaction mechanism

These crystal structures of PR with reduced peptide analogs represent a transition state in the reaction. Ideally, the reaction mechanism would be analyzed using a series of crystal structures of active PR with peptide substrates and transition state analogs representing different steps in the reaction. However, it is difficult to obtain crystal structures of active PR with peptide substrates. Two strategies have been used to analyze the structures of the transition state(s). We have analyzed structures of active PR with reduced peptide analogs that mimic the transition state of the hydrolytic reaction since

they contain an amine and tetrahedral carbon at the non-hydrolysable peptide bond. Other groups have used an alternative strategy of crystallizing an inactive enzyme with the D25N mutation in complex with peptide substrates (Prabu-Jeyabalan et al. 2002; Prabu-Jeyabalan et al. 2003; Prabu-Jeyabalan et al. 2004). Comparison of the crystal structures has helped to understand two different steps in the reaction. There were several differences between our crystal structures of PR with peptide analogs and those of the D25N inactive enzyme with peptides. The PR sequence differed in six amino acids, as well as the D25/N25 difference. Also, most of the peptides had differences in sequence. The two structures of D25N/p1-p6 (1KJF) and PR/p1-p6 that share similar peptide sequences were compared. Overall, the RMS differences were 0.6 Å for main chain atoms, as usually observed for PR crystal structures in different space groups. The most striking difference was in the conformation of the peptide or reduced peptide backbone atoms between P1 and P1' (Figure 3.1.6A). These differences arise from the presence of the planar peptide bond (CO-NH) in the peptide instead of the tetrahedral carbon in the reduced peptide (CH_2-NH). The tetrahedral carbon in the reduced peptide was much closer to the Asp 25 and 25' side chains than was the carbonyl carbon in the peptide bond (the two carbon atoms were separated by 1.1 Å). The tetrahedral carbon atom of the reduced peptide interacted with the four carboxylate oxygen atoms of Asp 25 and 25' at distances of 3.1-4.0 Å. In contrast, the peptide carbonyl oxygen of D25N/p1-p6 showed one hydrogen bond interaction and one van der Waals interaction with the carboxylate oxygens of Asp 25'. Furthermore, the tetrahedral carbon in the reduced peptide was in a similar position to the tetrahedral carbon of CH-OH in the UIC-94017 inhibitor, which



A.



B.

Figure 3.1.6 Structural variation around active site.

A. PR/p1-p6 is shown (colored by atom type) superimposed on D25N/p1-p6 (1KJF) in green bonds. Distances within 4.0 Å are shown.
B. PR/UIC-94017 is shown in yellow bonds superimposed on PR/p1-p6 complex (colored by atom).

mimics the transition state and showed interactions of the hydroxyl group with all four Asp 25/25' carboxylate oxygen atoms (Figure 3.1.6B). Therefore, the PR complexes with reduced peptide analogs more closely represented the tetrahedral transition state of the reaction, while the D25N/peptide structures are likely to represent the initial step of substrate binding to the PR.

Special feature in electron density map around the active site

The atomic resolution structure of PR_{V82A}/p2-NC showed unusual Fo-Fc difference density at the catalytic site that may relate to the reaction mechanism. The other crystal structures showed little or no difference density around the catalytic site. In these substrate analogs, the carbonyl group of P1 has been reduced to a methylene group to prevent hydrolysis. However, significant Fo – Fc positive difference density was observed close (about 1.4 Å) to the reduced carbon atom on P1Nle (norleucine) (Figure 3.1.7). Previous crystallographic studies of HIV-1 PR in complex with a pseudo-C2 symmetric inhibitor and molecular dynamic calculations suggested that the difluoroketone core was hydrated and that the hydration of the carbonyl group is the initial step for HIV-1 PR catalysis (Slee et al. 1995; Silva et al. 1996). Therefore, a hydroxyl was tested in the positive density. No reduction in the difference density was observed in tests with various other atoms (H, Na or O). The positive difference density was decreased, but not eliminated, only when a hydroxyl group was added to the reduced carbon atom. The refinement used a standard Nle and hydroxyl-Nle with relative occupancy of 0.7/0.3. Mass spectroscopic studies of crystals and separated peptide analog

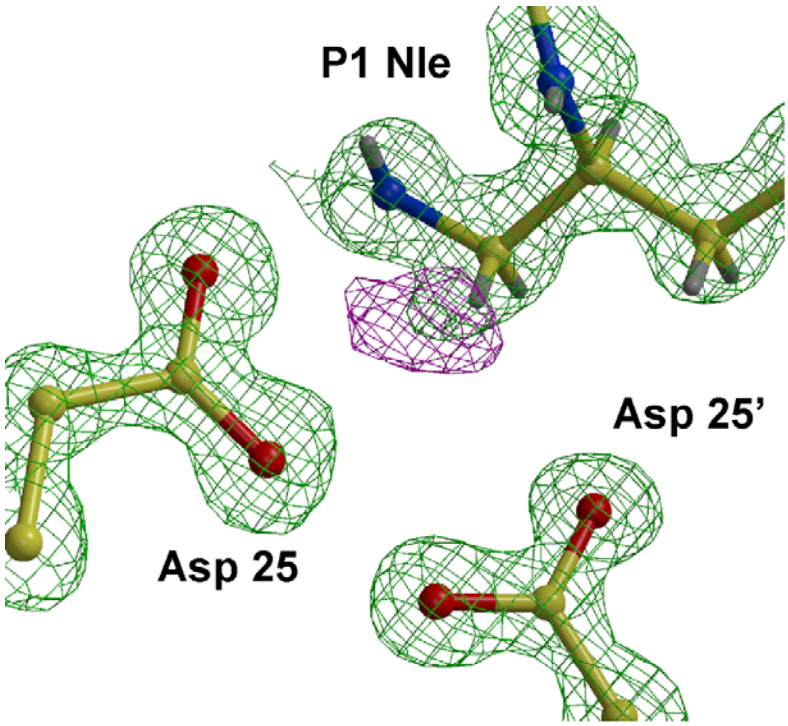


Figure 3.1.7 Electron density maps at the active site of PR_{V82A}/p2-NC complex. The 2Fo−Fc map is green and was contoured at a level of 2.2σ, whereas the Fo−Fc map is contoured at 3.2σ and colored purple for positive.

did not show any significant change in molecular mass of either PR or inhibitor. Therefore, any modification of the p2-NC analog must be transient at best and occurred only in the crystal structure. Moreover, hydration of the reduced carbon is an energetically unfavorable event. Thus, it is not clear whether the hydroxyl-Nle exists. Further analysis of the data by charge density analysis or quantum calculations will be necessary to understand the observed difference density at the active site, and may help to understand the catalytic mechanism.

These high resolution crystal structures of HIV PR with natural cleavage substrate analogs provide new molecular details to understand the specificity of substrate recognition and a basic framework for the design of new inhibitors that are more effective against resistant HIV. The results have been published in (Tie et al. 2005).

3.2 Structural analysis of protease complexes with UIC-94017 (TMC-114)

Specific aim: To determine & analysis crystal structures of PR and mutants PR_{V82A}, PR_{I84V}, PR_{D30N} and PR_{I50V} with UIC-94017

Comparison of these atomic resolution crystal structures will help to understand the effects of mutation in protease and evaluate the structural modification of protease-inhibitor complexes. Crystal structures of wild-type PR, PR_{V82A}, PR_{I84V}, PR_{D30N} and PR_{I50V} in complex with UIC-94017 have been determined at 1.10-1.53 Å resolution and similar hydrogen bonding interactions are shown between protease and inhibitor. The hydrogen bonds between UIC-94017 with the main chain atoms of Asp 29 and 30 in the protease are demonstrated, which is critical for the potency and activity of this compound against multiple PI-resistant HIV.

Chemical synthesis of UIC-94017

The synthesis of UIC-94017 (**1**) was done by Dr. Arun Ghosh's lab and is outlined in Figure 3.2.1. Commercially available optically active epoxide **3** (Ghosh et al. 1998a) was reacted with isobutylamine in 2-propanol at 84°C for 6 h. The resulting amino alcohol was reacted with 4-nitrobenzenesulfonyl chloride to provide the sulfonamide derivative **3** (Ghosh et al. 1998b). Reduction of nitro group of **4** by catalytic hydrogenation over 10% Pd-C in ethyl acetate provided the corresponding aromatic amine. Trifluoroacetic acid promoted removal of BOC-group followed by reaction of the resulting diamine with mixed carbonate **5** (Ghosh et al. 1995) furnished inhibitor **1** (Ghosh et al. 1996) selectively. Synthesis of UIC-94017 was confirmed by ^{1}H-NMR.

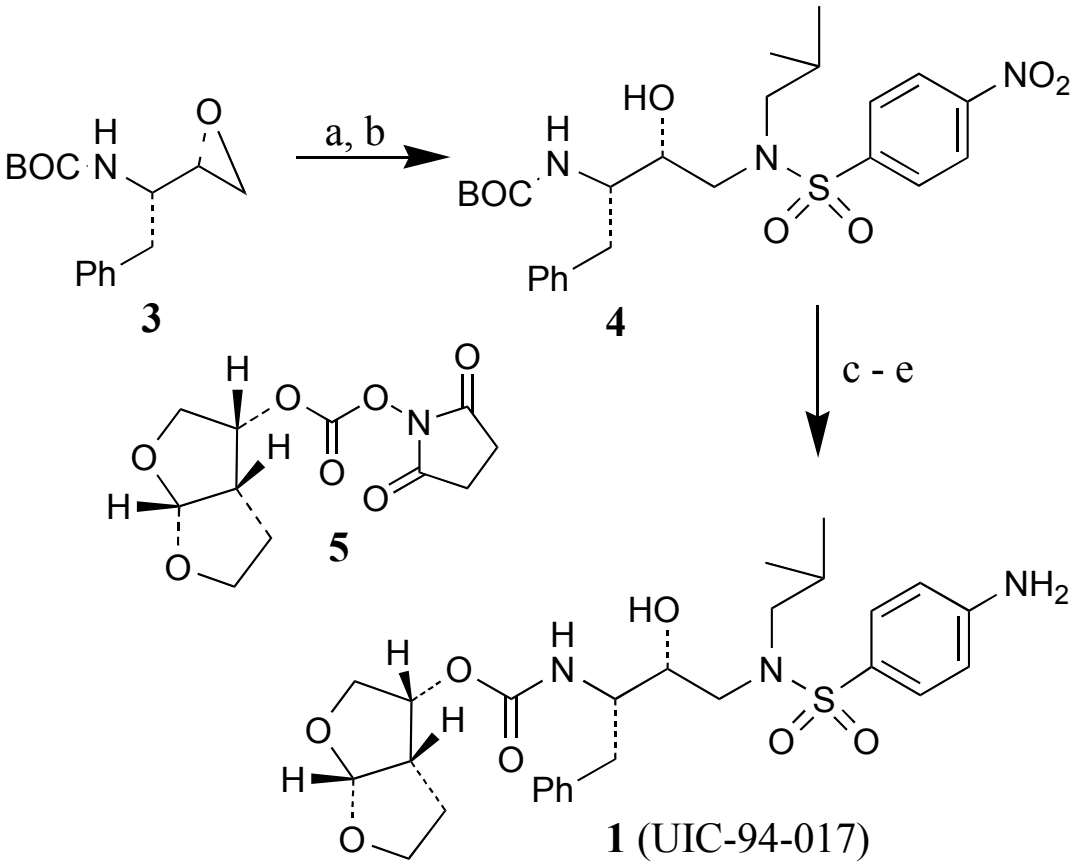


Figure 3.2.1 Synthesis of UIC-94017.
Reagents and conditions: (a) $Me_2CHCH_2NH_2$, 2-propanol, 84°C, 6 h; (b) *p*-NO_2-$C_6H_4SO_2Cl$, aq. $NaHCO_3$, CH_2Cl_2; (c) H_2, 10% Pd-C, EtOAc; (d) CF_3CO_2H, CH_2Cl_2; (e) **5**, Et_3N then diamine.

Enzyme kinetics and relative inhibition of PR_{V82A}, PR_{I84V}, PR_{D30N} and PR_{I50V}

PR and the set of drug-resistant mutant proteases (PR_{V82A}, $PR_{I84V,}$ PR_{D30N}, PR_{I50V}, and PR_{L90M}) catalyze the hydrolysis of the fluorescence substrate, an analog of the MA-CA cleavage site in the Gag-Pol polyprotein, and their catalytic activities are competitively inhibited by UIC-94017. Saturation kinetics were observed for PR, PR_{V82A}, $PR_{I84V,}$ PR_{D30N}, PR_{I50V} and PR_{L90M} - catalyzed hydrolysis of substrate. The kinetic parameters are summarized in Table 3.2.1. Mutants varied in their catalytic activity for the fluorescence substrate compared to PR. Except for the dramatically lower k_{cat} for PR_{D30N} (0.2 min^{-1}) and >2 times higher k_{cat} for PR_{L90M} (7.6 min^{-1}), the k_{cat} value for the others was in the range of 1.7-2.7 min^{-1}. The Km value had less change (26-57 μM). Overall, similar k_{cat}/Km values were observed for PR, PR_{V82A} and PR_{I84V}, while PR_{D30N} and PR_{I50V} showed only 0.09 and 0.38 fold of the PR value, respectively, and that of PR_{L90M} was 2.6 fold of PR value. PR_{V82A} and PR_{I84V} showed little increase in K_i for UIC-94017, with value of 3.6 and 5 fold in relative to that of PR, respectively. PR_{I50V} and PR_{D30N} had larger change of 9 and 30 fold compared to the Ki of PR. Surprisingly, PR_{L90M} showed only 0.14 fold of the Ki value for PR, which suggested that it was 7 times better inhibited by UIC-94017 than PR. The order of relative inhibition for UIC-94017 was PR_{L90M} >> PR > PR_{V82A} > PR_{I84V} > PR_{I50V} > PR_{D30N}. A slightly different order of inhibition was shown for indinavir (IDV), which was PR ≈ PR_{L90M} > PR_{V82A} > PR_{I84V} > PR_{D30N} > PR_{I50V}. Mutants PR_{V82A} and PR_{I84V} had similar Ki relative to PR value for UIC-94017 and indinavir. Most of the other clinical inhibitors showed significantly poorer relative inhibition of the protease with V82A or I84V mutations, except for amprenavir,

Table 3.2.1 Kinetic data.
Kinetic parameters for protease-catalyzed hydrolysis of fluorescence substrate [Ac-RE(Edans)SQNY*PIVRK(Dabcyl)R-CO-NH2, where * indicates the cleavage site], and inhibition constants for the hydrolytic reaction with the inhibitor UIC-94017 (or IDV)

	K_M (μM)	k_{cat} (min^{-1})	k_{cat}/K_M ($mM^{-1}min^{-1}$)	Ki (UIC-94017) (nM)	relative Ki (UIC-94017)	Ki (IDV) (nM)	relative Ki (IDV)
PR	36 ± 8	2.7 ± 0.5	77 ± 20	0.22 ± 0.05	1	0.6 ± 0.1	1
PR_{V82A}	26 ± 2	2.4 ± 0.1	90 ± 8	0.8 ± 0.06	3.6	1.3 ± 0.1	2.2
PR_{I84V}	39 ± 7	2.0 ± 0.2	50 ± 10	1.1 ± 0.2	5	2.5 ± 0.3	4.2
PR_{I50V}	57 ± 11	1.7 ± 0.2	29 ± 7	2.0 ± 0.4	9	10.4 ± 1.0	17
PR_{D30N}	31 ± 5	0.21 ± 0.02	7 ± 1	6.6 ± 1.0	30	7.0 ± 1.0	12
PR_{L90M}	38 ± 5	7.6 ± 0.7	200 ± 34	0.03 ± 0.004	0.14	0.8 ± 1.0	1.3

which had only 2- and 3-fold increased K_i for PR_{V82A} and PR_{I84V} mutants, respectively (Gulnik et al. 1995; Klabe et al. 1998). UIC-94017 was 3 fold worse than indinavir for the inhibition of PR_{D30N}, a mutant most commonly found on exposure to nelfinavir. However, UIC-94017 had 2 to 10 fold lower relative Ki (better inhibition) than indinavir for PR_{I50V} and PR_{L90M}. This is consistent with the potency of UIC-94017 against HIV-1 drug- resistance strains (Koh et al. 2003). Furthermore, UIC-94017 is more effective on PR_{I50V} than is indinavir, even though PR_{I50V} commonly occurs with the treatment of amprenavir, which is chemically most similar to UIC-94017.

Description of the high resolution crystal structures

The crystal structures of PR and drug resistant mutants (PR_{V82A}, PR_{D30N}, PR_{I50V} and PR_{I84V}) have been determined in complexes with UIC-94017. Crystallographic statistics are summarized in Table 3.2.2. The crystallographic asymmetric units contain a PR dimer with the residues in the two subunits numbered 1-99 and 1'-99'. The three structures were determined in two different space groups ($P2_12_12$ and $P2_12_12_1$) and were refined to resolutions of 1.10-1.53 Å and R-factors of 0.12-0.15 including solvent molecules, anisotropic B-factors and hydrogen atoms. These high resolution crystal structures showed excellent electron density for all the protease atoms, the inhibitor, and solvent molecules. The electron density map for Tyr 59' and interacting atoms in the PR structure is shown in Figure 3.2.2. The average B-factors were low for protein and inhibitor atoms. The highest resolution structure of PR_{V82A} had the lowest R-factor. PR_{V82A} showed the lowest average B-factors of 8.0 $Å^2$ for protein main chain atoms

Table 3.2.2 Crystallographic data statistics.

		PR	PR_{V82A}	PR_{I84V}	PR_{D30N}	PR_{I50V}
Space group		$P2_12_12$	$P2_12_12_1$	$P2_12_12$	$P2_12_12_1$	$P2_12_12$
Unit cell dimensions (Å)	**a**	58.26	50.79	50.79	58.49	58.70
	b	85.91	57.87	57.87	86.13	85.79
	c	46.05	61.94	61.94	45.94	46.15
Resolution range (Å)		50-1.30	50-1.10	50-1.53	50-1.45	50-1.22
Unique reflections		56890	74707	35966	40644	64113
R_{merge} (%)		6.5	5.2	8.2	8.9	9.0
Overall (final shell)		(26.8)	(19.6)	(31.2)	(57.5)	(40.2)
Data range for refinement (Å)		10-1.30	10-1.10	10-1.53	10-1.45	10-1.22
R_{work} (%)		13.42	11.74	12.98	15.2	13.0
R_{free} (%)		17.34	14.53	19.65	21.8	18.8
No. of solvent (total occupancies)		210	221.6	193	178.0	171.1
Completeness (%)		98.6	99.7	99.3	98.0	91.9
Overall (final shell)		(90.2)	(97.6)	(97.3)	(86.0)	(53.3)
RMS deviation from ideality						
Bonds (Å)		0.014	0.016	0.012	0.011	0.014
Angle distance (Å)		0.038	0.037	0.035	0.034	0.037
Average B-factors ($Å^2$)						
Main chain		14.9	8.0	15.9	16.2	15.4
Side chain		19.2	13.1	21.9	22.7	19.9
Inhibitor		10.8	7.4	14.1	12.9	12.9
Solvent		31.1	24.1	32.1	33.4	28.8
Relative occupancies of alternate conformations of UIC-94017		0.6/0.4	1*	0.7/0.3	0.5/0.5	0.6/0.4

*Only the aniline ring had two orientations with occupancy distribution of 0.6/0.4.

compared to values of about 15 $Å^2$ for the other structures. The inhibitor atoms also had the lowest average B-factors for the complex with PR_{V82A}: 7.4 $Å^2$ compared to 10.8, 14.1, 12.9 and 12.9 $Å^2$ for PR, PR_{I84V}, PR_{D30N} and PR_{I50V}, respectively.

Alternate conformations were modeled for residues in all the crystal structures. There were 36 residues with two alternate conformations modeled for PR, 37 for PR_{V82A}, 41 for PR_{I84V} mutant, 18 for PR_{D30N} mutant and 35 for PR_{I50V} mutant (Table 3.2.3). There was clear electron density for the alternate positions of these side chains, as shown by the examples in Figure 3.2.3. The residues with alternate conformations were not the same in both subunits, or in all five structures. However, four of the structures, PR, PR_{I84V}, PR_{D30N} and PR_{I50V} showed a similar pattern, probably because they were in the same space group $P2_12_12$ and both showed two conformations of the inhibitor. Most alternate conformations were observed for the side chains of surface residues, especially residues with longer side chains, such as arginine, lysine, glutamine and glutamic acid. The side chains of Ile 33', Ile 64, Leu 89, and Leu 97 are in internal hydrophobic regions. The side chains of seven amino acids which interact directly with the inhibitor, Arg 8, Asp 30, Val 32, Ile 47, Ile 50, Pro 81, Val 82, and Ile 84, showed alternate conformations in at least one subunit. Disorder was previously reported for Val 32, Val 82 and Ile 84 in ternary complexes of protease bound to substrate analogs (Mahalingam et al. 2001).

The main chain atoms of Ile 50 and 50' in the two subunits had two alternative conformations, with relative occupancies of 0.6 to 0.4, in the four structures with space group $P2_12_12$. The carbonyl groups in these two residues can flip about 180 degree and still form the hydrogen bonds between the carbonyl oxygen atom and amide nitrogen

Table 3.2.3 Residues with alternate conformations.
The side chains have alternate conformations in most cases.
*main chain atoms have alternate positions. #alternate positions for both main chain and side chain atoms. A, B indicate the subunit. Residues in inhibitor binding site are indicated in **bold.**

Residue	PR	PR_{V82A}	PR_{I84V}	PR_{D30N}	PR_{I50V}
Thr 4		A#			
Leu 5		A*			
Trp 6		B			
Lys 7	AB	A	AB	AB	AB
Arg 8	**AB**	**AB**	**AB**		**AB**
Lys 14	AB	AB	AB		AB
Gln 18	B	A	B	B	B
Lys 20		AB			
Glu 21	AB	B	AB	AB	AB
Asp/Asn 30	**AB**	**A**	**AB**	**B#**	**B#**
Val 32	**B**		**B**		**B**
Ile 33	B		B		
Glu 34		B	AB	A	A
Glu 35		B	B		
Met 36		B			
Ser 37	AB	A	A	AB	AB
Leu 38		AB			
Pro 39		B			
Gly 40		B*			
Arg 41	A	AB	A		A
Trp 42			B		
Lys 43	B	AB	AB		B
Lys 45	B		AB		B
Met 46	AB	AB	AB	B	AB
Ile 47		**B**			
Ile/Val 50	**A*B#**		**A*B#**	**A*B***	**A*B#**
Gly 51	**A*B***			**A*B***	**A*B***
Lys 55	A		AB		A
Arg 57		B			A
Gln 61		A	AB		
Ile 63		A			
Ile 64		A#			
Glu 65	AB	AB	AB	A	B
Lys 70	AB	B	AB		AB
Pro 81	**B**				**B**
Val/Ala 82	**AB**		**AB**		**AB**
Ile/Val 84	**AB**	**A**	**B**	**B**	**AB**
Leu 89			A		
Leu 97	AB	A	AB	AB	AB

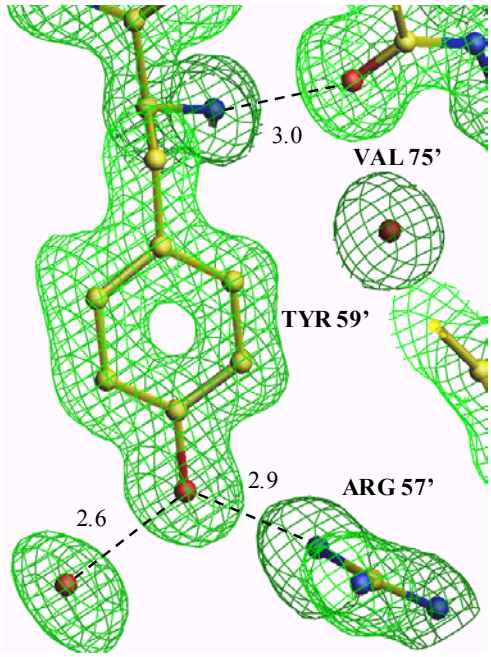


Figure 3.2.2 2Fo-Fc electron density map for Tyr 59' of PR complex.
The contour level is 1.7 σ. Hydrogen bonds are indicated by dashed lines with the interatomic distance in Å.

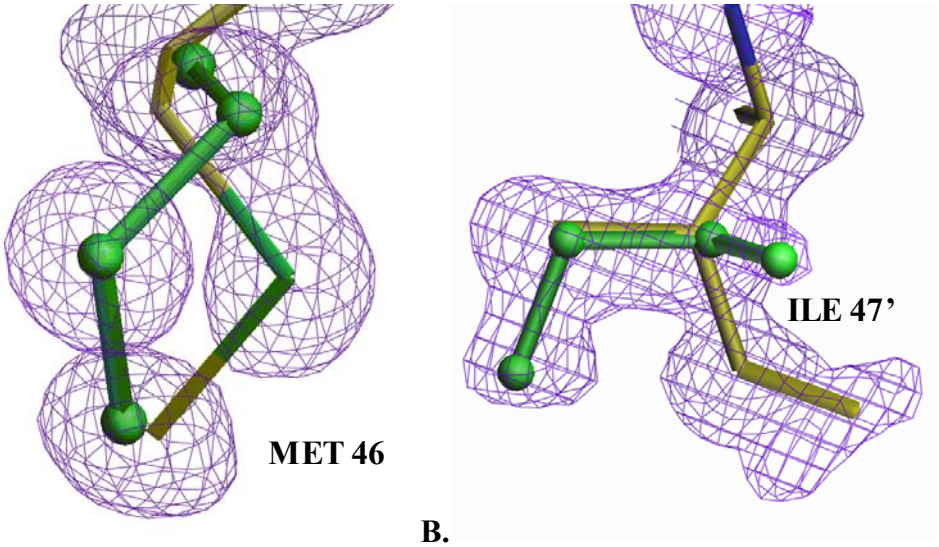
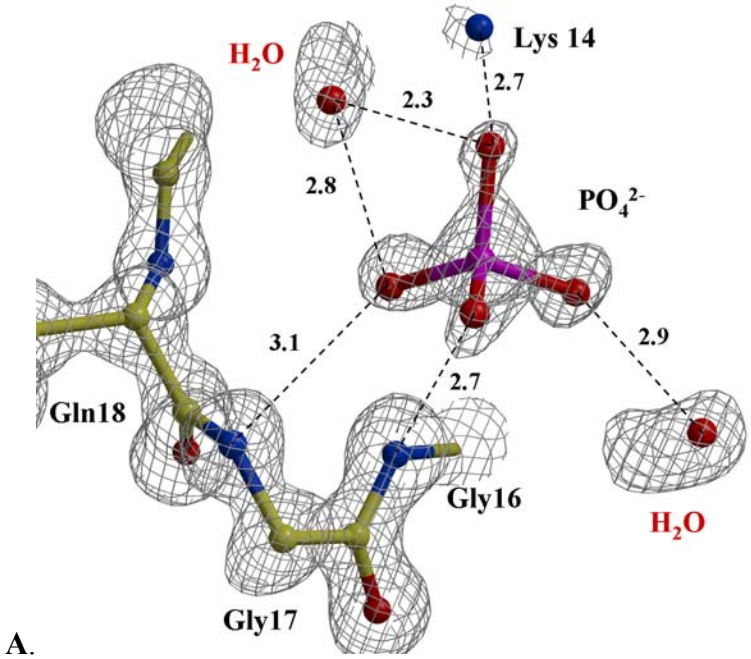


Figure 3.2.3 Alternate side chain positions of residues in PR_{V82A} structure shown in 2Fo-Fc map.
The 2Fo-Fc map was contoured at a level of 1.7 σ. The major conformation is colored by atom type, and the minor conformation is in green.
A) Met 46. **B) Ile 47'**

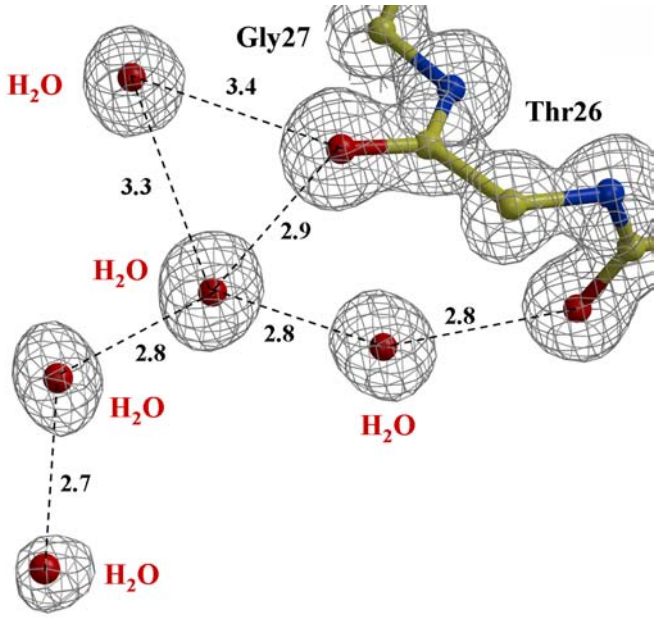
atom of each residue that link the tips of the flaps. The structure of PR_{V82A} in space group $P2_12_12_1$ showed two alternative conformations of the main chain atoms between residues 4 and 5, which are located at the dimer surface. Ile 64 in PR_{V82A} had alternate main chain and side chain conformations in an internal hydrophobic region. The occupancy distribution was 0.7 to 0.3 and the two conformations interacted with the same surrounding residues.

Solvent structure in the crystals

Previous crystal structures of HIV protease-inhibitor complexes were modeled with a single shell of solvent molecules (Mahalingam et al. 1999; Mahalingam et al. 2001). The high quality of diffraction data for the structures described here permitted the modeling of a second shell of solvent (Figure 3.2.4B). The solvent was modeled with more than 200 water molecules, ions and other small molecules from the crystallization solutions, including many with partial occupancy, depending on the shape of the electron density and the interactions with other molecules. The ions modeled were sodium, chloride, acetate, sulfate and phosphate. Glycerol and DMSO were fitted to density in the PR, PR_{V82A}, PR_{D30N} and PR_{I50V}. Examples of the electron density maps for solvent regions are shown in Figure 3.2.4. The average B-factors for the solvent atoms were 24.1 $Å^2$ for PR_{V82A} and 29-33 $Å^2$ for PR, PR_{I84V}, PR_{D30N} and PR_{I50V} complexes.

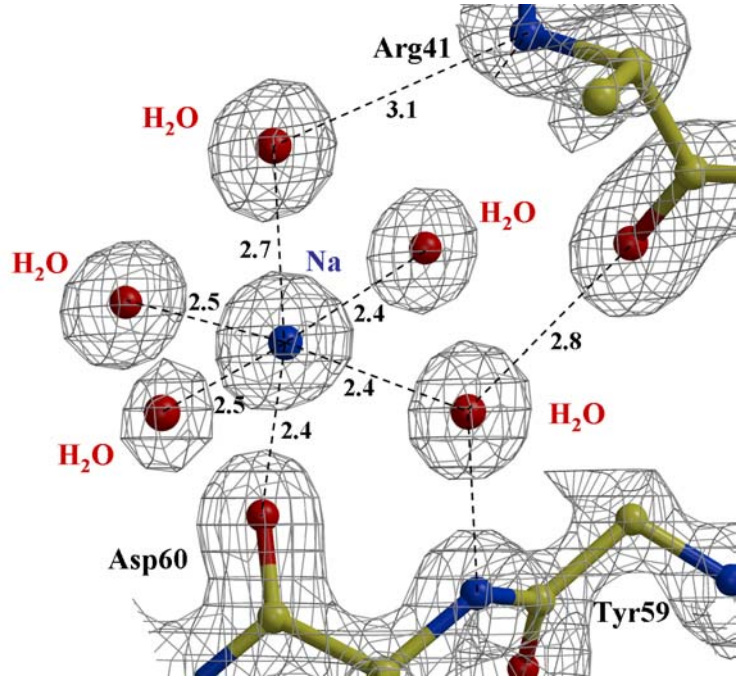


A.

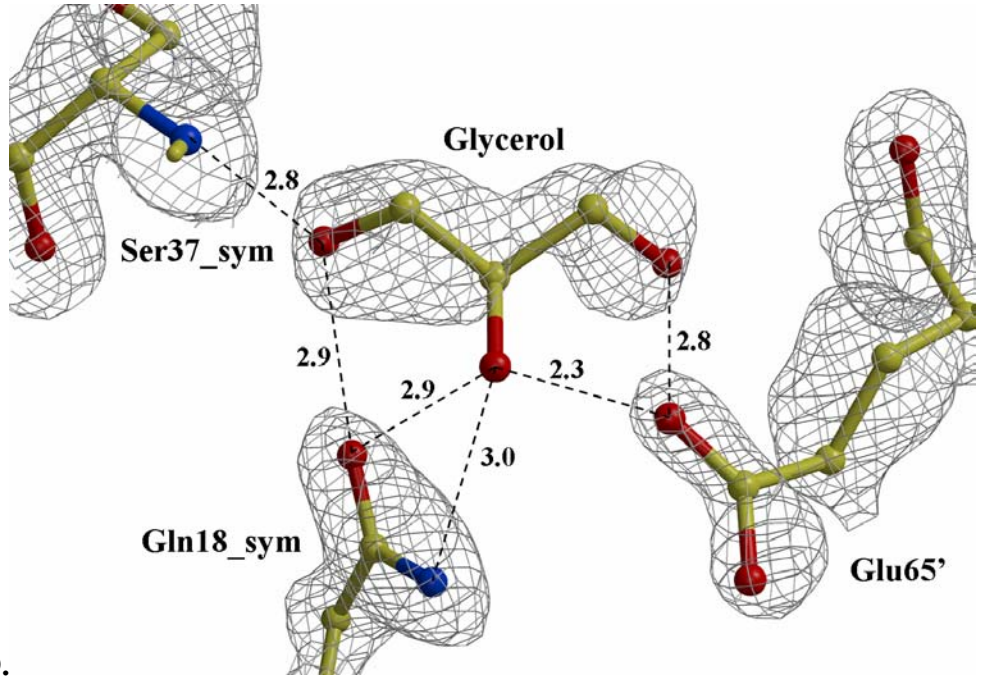


B.

Figure 3.2.4 2Fo-Fc electron density maps for solvent molecules.
The 2Fo-Fc map was contoured at a level of 1.7 σ. Hydrogen bonds are indicated by dashed lines with the interatomic distance in Å.
A) Phosphate ion in PR_{V82A} complex.
B) A set of interacting water molecules in PR_{V82A} complex.



C.

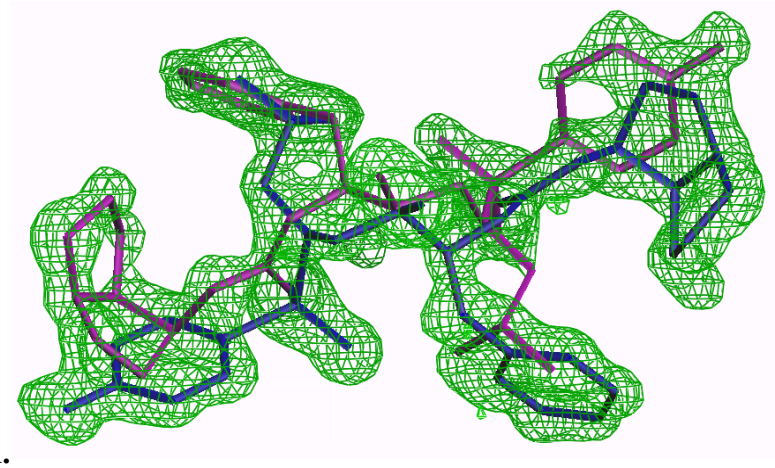


D.

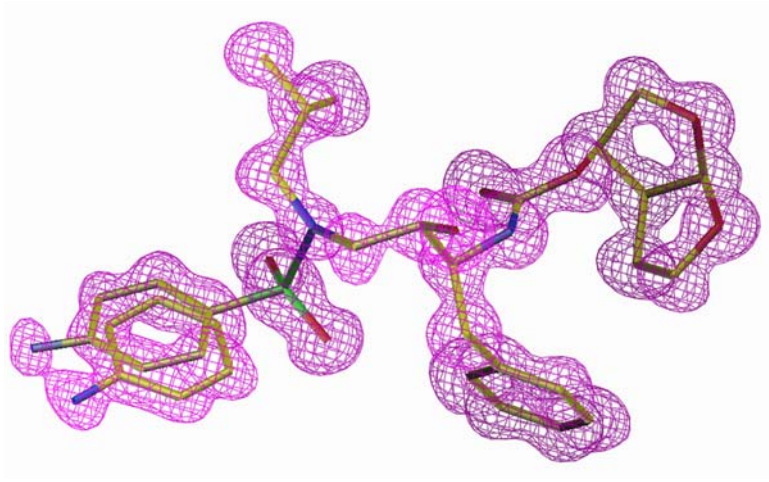
Figure 3.2.4 2Fo-Fc electron density maps for solvent molecules.
The 2Fo-Fc map was contoured at a level of 1.7 σ. Hydrogen bonds are indicated by dashed lines with the interatomic distance in Å.
C) Sodium ion in PR_{I50V} complex.
D) Glycerol molecule in PR complex.

Protease-inhibitor interactions in PR/UIC-94017 complex

The crystal structure of HIV-1 PR showed two pseudosymmetric positions of UIC-94017 with relative occupancy of 0.55/0.45 (Figure 3.2.5A). The side chains of several amino acids in the binding site showed two conformations, as observed previously for hydrophobic side chains near the peptide analogs (Mahalingam et al. 2001). The inhibitor P1 and P1' groups form van der Waals interactions with protease residues Leu 23, Gly 49, Ile 50, Pro 81, Val 82, and Ile 84 from both subunits. The end groups interact with residues Ala 28, Asp 29, Asp 30, Val 32, Ile 47 and Ile 50. The CG atom of Pro 81' in the PR structure interacts closely with the inhibitor and shows alternate positions with the occupancy distribution of 0.55 to 0.45, which is the same as that of the inhibitor. This suggests that the alternate conformations of the amino acid side chains correlate with those of the inhibitor. Theoretically, the greater flexibility of the structure will increase the entropy and in turn reduce the interaction energy between protein and inhibitor. The PR-inhibitor hydrogen bond interactions are shown in Figure 3.2.6A. Equivalent interactions exist for both positions of the inhibitor (Table 3.2.4). The inhibitor hydroxyl group interacted with all four carboxylate oxygens of the catalytic Asp 25 and 25', which mimics the interactions of its tetrahedral transition state of the protease. The carbonyl oxygen and one of the sulfonamide oxygen atoms form hydrogen bonds with the water that interacts with the amides of Ile 50 and 50' on the flaps, as observed in most HIV protease-inhibitor complexes (Gustchina et al. 1994). This water is conserved in all structures of PR with inhibitors/substrates except for those urea-based inhibitors designed explicitly to substitute this H_2O, such as DMP323 and DMP450



A.



B.

Figure 3.2.5 Electron density map of UIC-94017.
The omit map was contoured at a level of 3.2 σ.
A. Electron density map of UIC-94017 in PR complex. UIC-94017 is shown in two pseudosymmetric overlapping positions in red and blue bonds.
B. Electron density map of UIC-94017 in PR_{V82A} complex.
The map shows one position of UIC-94017 with two conformations for the aniline group.

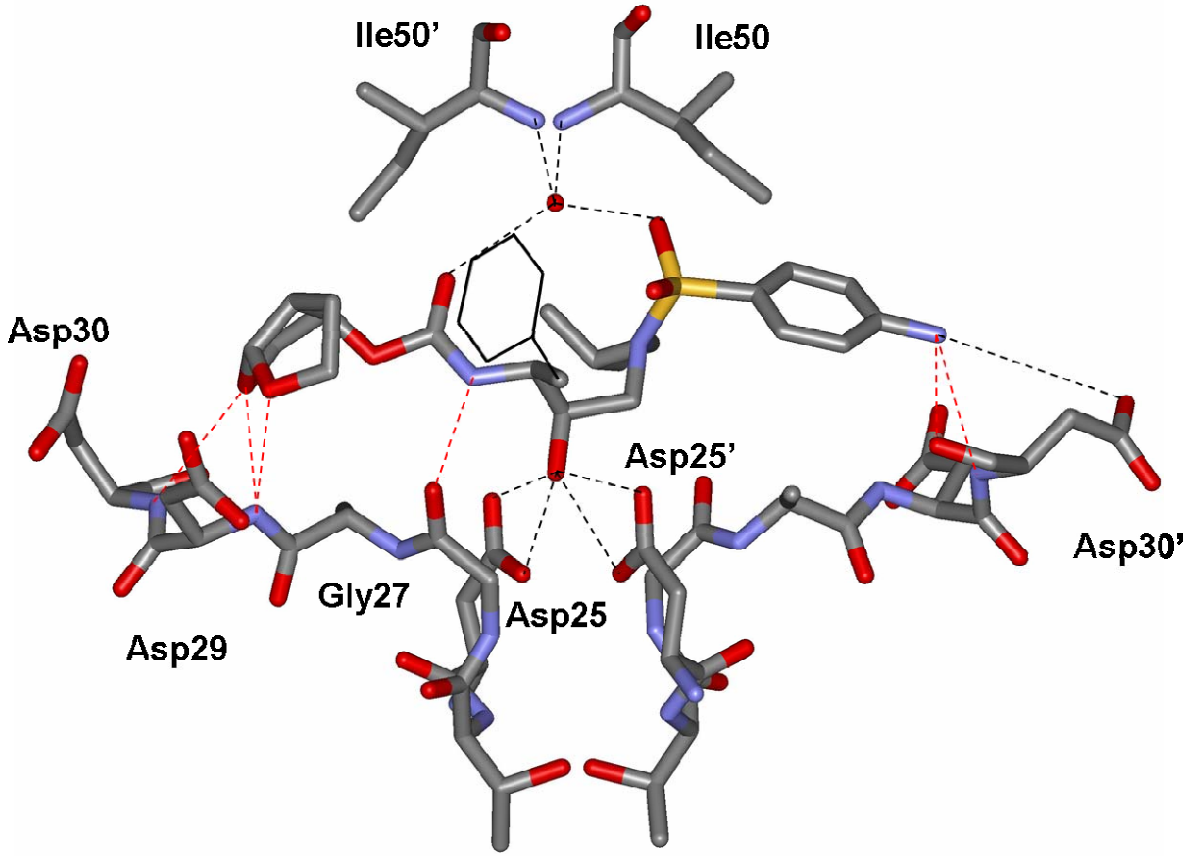
Table 3.2.4 Protease-inhibitor hydrogen bond interactions.
The protease-UIC-94017 interatomic distances are shown for the crystal structures of PR, PR_{V82A}, PR_{I84V}, PR_{D30N} and PR_{I50V} complexes. Except for PR_{V82A}, other complexes have two positions of UIC-94017. The atoms of the major UIC-94017 conformation and interacting PR atoms are indicated in **bold**. The Asp 25 and 25' carboxylate oxygen atoms interacted with OH in both positions of UIC-94017. PR_{V82A} had one UIC-94017 with two conformations for the aniline group that had different interactions. *The minor conformation of the aniline NH_2 interacted with Asp 30 OD2. [m]The major conformation of 30/30'and [#] the minor conformation of 30/30'. Two water molecules are indicated as H2OA and H2OB. Distances of more than 3.4 Å are in parentheses to indicate they are too long for good hydrogen bond interactions.
The interatomic distances are shown for equivalent atoms of amprenavir from 1HPV crystal structure.

Protease atom	UIC-94017 Atom	Distance (Å)					
		PR	PR_{V82A}	PR_{I84V}	PR_{D30N}	PR_{I50V}	WT-AMP
Asp 25 OD1	OH	**2.7**/2.8	3.0	**2.6**/2.9	**2.7**/2.6	**2.6**/2.7	2.6
Asp 25 OD2	OH	**2.6**/3.1	2.5	**2.4**/3.1	**2.5**/2.8	**2.5**/3.1	3.1
Asp 25' OD1	OH	**2.8**/2.6	2.5	**2.7**/2.5	**2.7**/2.5	**3.0**/2.7	2.7
Asp 25' OD2	OH	**3.2**/2.6	3.2	**3.1**/2.5	**3.0**/2.7	**3.3**/2.6	2.8
Gly **27'**/27 O	NH	**3.2**/3.2	3.1	**3.1**/3.2	**3.2**/3.2	**3.2**/3.1	(3.6)
Asp **29'**/29 NH	O26	**3.1**/3.0	3.0	**3.1**/3.1	**3.1**/3.1	**3.1**/3.1	(3.5)
Asp **29'**/29 OD2	O28	**3.2**/3.2	3.3	**3.2**/3.1	**3.3**/3.3	**3.1**/3.3	
Asp **29'**/29 NH	O28	**3.0**/2.9	2.9	**2.9**/3.0	**2.9**/2.9	**3.0**/2.9	
Asp (Asn) **30**/30' O	NH_2	**3.3**/3.2	(3.5)	**3.2**/3.4	**3.2**/3.0[#]	**(4.1)**/3.2[#]2.9[m]	(3.6)
Asp (Asn) **30**/30' OD2(ND2)	NH_2	**2.7**/2.8	**2.8**/2.9*	**2.4**/2.6		**2.5**/2.8	3.1
Asp (Asn) **30**/30' NH	NH_2	**3.2**/3.0	(3.5)*	**3.2**/3.1	**3.2**/3.0	**3.4**/3.0	3.2
Asp (Asn) **30'**/30 NH	O26	**3.3**/3.3	3.1	**3.3**/(3.5)	**3.3**/3.0	**3.4**/3.3	(3.5)
H_2OA	O22	**3.0**/3.0	2.8	**2.9**/3.1	**2.9**/3.1	**3.1**/2.9	3.0
H_2OA	O10=S	**2.4**/2.4	2.8	**2.4**/2.3	**2.5**/2.4	**2.3**/2.6	2.8
H_2OA	O9=S	**3.3**/3.2		(**3.5**)/3.2	**3.4**/3.4	**3.4**/(3.5)	3.4
Ile (Val) 50 N	H_2OA	2.9	3.0	2.9	3.0	3.0	2.7
Ile (Val) 50' N	H_2OA	2.9	2.9	3.0	2.9	2.9	2.9
H_2OB	NH_2		2.4		**3.2**/2.9		

(Lam et al. 1994; Hodge et al. 1996; Lam et al. 1996). No other water-mediated interactions were observed between PR and UIC-94017. The two oxygen atoms of the *bis*-THF group of the inhibitor are positioned to form hydrogen bond interactions with the main chain amides of Asp 29 and Asp 30. The *bis*-THF O28 also formed a close interaction with the OD2 of Asp 29, similar to the proton-mediated interaction observed between Asp 29 and peptide analogs with P2' Glu (Mahalingam et al. 2001). The amino group at the other end of the inhibitor interacts with the main chain carbonyl oxygen and the side chain carboxylate oxygen of Asp 30'.

Comparison of the structures of PR/UIC-94017 and PR/amprenavir

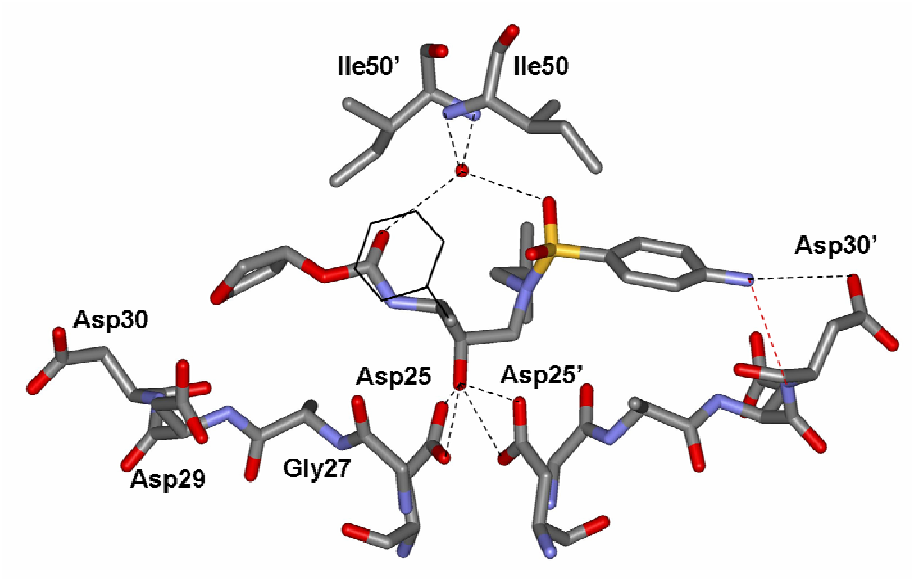
It is instructive to compare the protease interactions of UIC-94017 with those of the structurally related inhibitor, amprenavir. The crystal structure of wild type protease with amprenavir was determined at 1.9 Å resolution (Kim et al. 1995) (PDB: 1HPV). Amprenavir had fewer hydrogen bond interactions with the protease than observed for UIC-94017 (Table 3.2.4; Figure 3.2.6B). The detailed comparison for two distal ends of two inhibitors is shown in Figure 3.2.7. The single THF ring on amprenavir was more than 3.5 Å away from the protease and no hydrogen bond was formed, while the modified *bis*-THF ring on UIC-94017 had tight hydrogen bonds with the main-chain atoms of residues Asp 29' and Asp 30'. The difference in the size of the THF ring further induced slight change in the shape of the inhibitor, so even though amprenavir has the same aniline end as UIC-94017, the NH_2 group of amprenavir formed a weaker hydrogen bond interaction with the carboxylate oxygen of Asp30 (3.1 Å compared to 2.7 Å in



A.

Figure 3.2.6A Hydrogen bond interactions of PR with UIC-94017 (1S6G).
The PR residues forming hydrogen bond interactions with the major position of UIC-94017 are shown. The interactions of the minor position of UIC-94017 are essentially identical (Table 3.2.4). The red sphere indicates a water molecule. Hydrogen bonds are indicated by dashed lines with red lines representing the interactions between the inhibitor to the main chain atoms of PR.

B.



C.

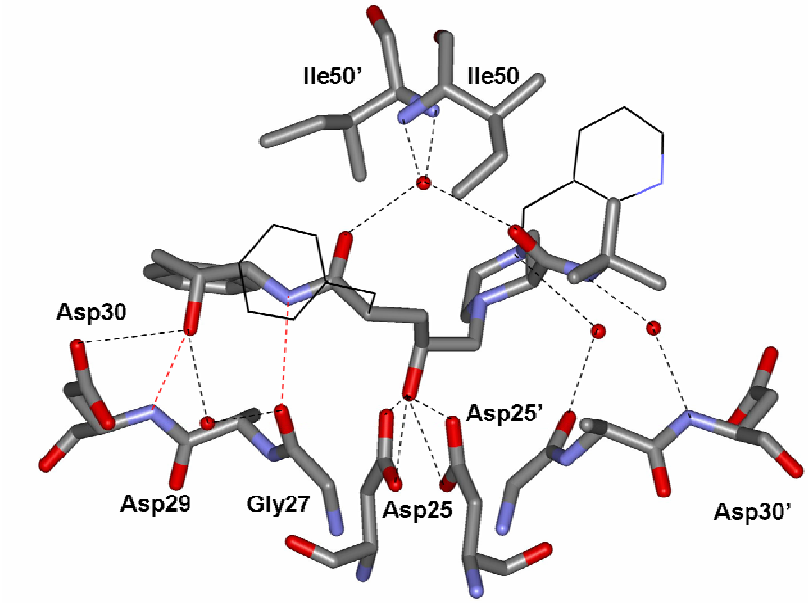
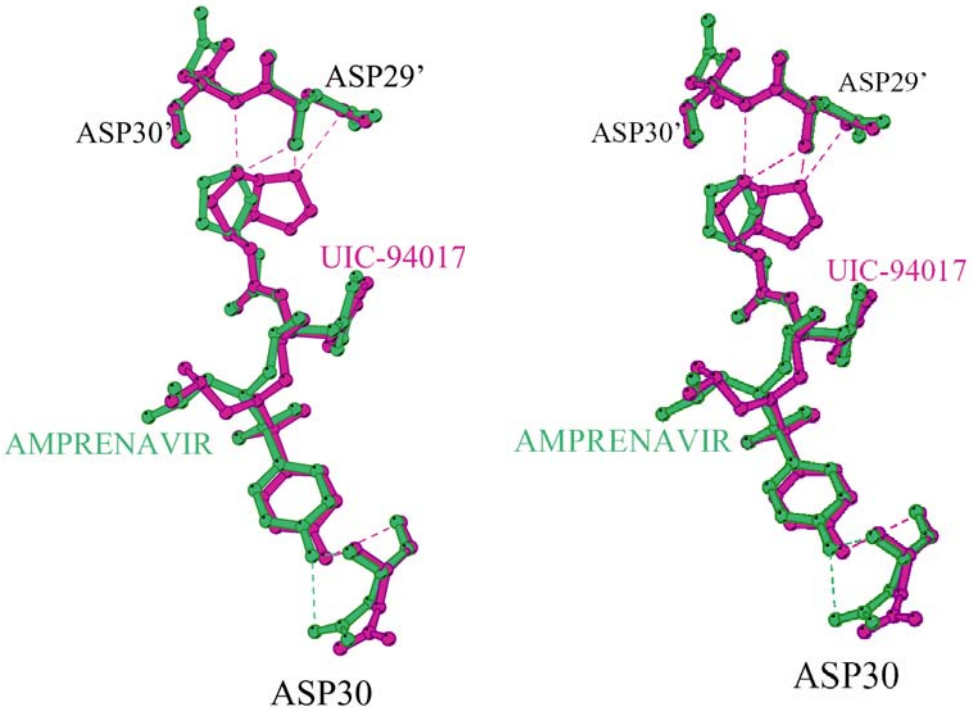
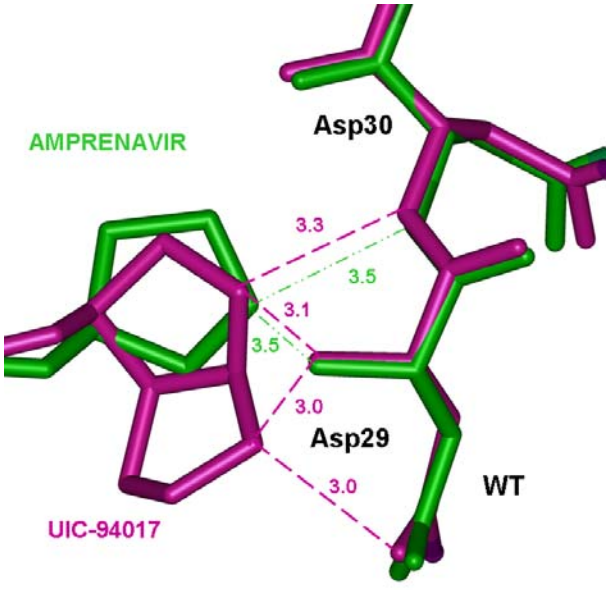


Figure 3.2.6B Hydrogen bond interactions of PR with amprenavir (1HPV).
The PR residues forming hydrogen bond interactions with amprenavir are shown. Hydrogen bonds are indicated by dashed lines with red lines representing the interactions between inhibitor to the main chain atoms of PR.

Figure 3.2.6C Hydrogen bond interactions of PR with indinavir (1SDT).
The PR residues forming hydrogen bond interactions with the major position of indinavir are shown. The interactions of the minor position of indinavir are essentially identical. The red sphere indicates a water molecule. Hydrogen bonds are indicated by dashed lines with red lines representing the interactions between inhibitor to the main chain atoms of PR.



A.



B.

Figure 3.2.7 Major differences in PR interactions with UIC-94017 and amprenavir. The hydrogen bond interactions of PR residues Asp 30, Asp 29' and Asp 30' with UIC-94017 and amprenavir are compared in a stereo view. Hydrogen bond interactions are indicated by dashed lines. The atoms and interactions in the PR/UIC-94017 structure are green, and the PR/amprenavir atoms and interactions are in red.

A. Overall difference

B. Detail difference at *bis*-THF ring

PR/UIC-94017 complex) and lost the one with the main-chain oxygen atom of Asp30, as seen in PR/UIC-94017 complex. Amprenavir and UIC-94017 have similar hydrophobic groups with the exception of the *bis*-THF at one end of UIC-94017. Therefore, many of the hydrophobic interactions with PR are similar for both inhibitors. The *bis*-THF group of UIC-94017 introduced one very short van der Waals contact of 2.9-3.0 Å with the main chain carbonyl oxygen of Gly 48, which cannot occur in amprenavir. However, both UIC-94017 and amprenavir showed little change in K_i values for PR and mutants PR_{V82A} and PR_{I84V} (Table 3.2.1) (Klabe et al. 1998). And surprisingly, an improvement of 9 times lower relative Ki to wild-type PR was obtained for PR_{I50V} with UIC-94017 (relative Ki of 9) compared to that of amprenavir (relative Ki of 83) (Pazhanisamy et al. 1996), although different assay conditions may have been used.

Comparison of protease interactions in UIC-94017 and indinavir complexes

Since crystal structures with another clinical drug indinavir (IDV) have been obtained for the wild-type PR and mutant proteases PR_{I50V}, PR_{V82A} and PR_{L90M} in our lab (Mahalingam et al. 2004; Liu et al. 2005) and a complete set of kinetic data had been measured (Table 3.2.1), it will be useful to compare the protease interactions of UIC-94017 with those of indinavir. The PR_{I50V}/IDV complex had the highest resolution of 1.10Å, while the others were of resolution of 1.25-1.40 Å, which are comparable to the values for the protease/UIC-94017 crystals. The indinavir structures showed mostly one orientation of the inhibitor, except for PR_{I50V}/IDV, which had two orientations of IDV similar to the complexes with UIC-94017. Low RMS deviation values (less than 0.3 Å)

were calculated for the structures in the same space group, while higher values of about 0.6 Å were obtained for those complexes in different space groups. The same trend has been observed for all other protease-inhibitor/substrate structures.

Hydrogen bonds are the strongest interactions that play a major role in stabilizing inhibitor/substrate binding in the protease cleavage site. To minimize the effect caused by mutation at side-chain, UIC-94017 inhibitor was designed to increase the number of strong favorable interactions with the protease, especially with the main-chain atoms. Both indinavir and UIC-94017 have the hydroxyethyl isostere as the center of the molecule, which mimics the transition state of the natural substrate during proteolysis reaction. This hydroxyl group forms strong hydrogen bonds with all four carboxylate oxygens of catalytic Asp25 and Asp25'. Other than these four hydrogen bonds, in the protease/IDV complex, only three direct interactions were made between one end of the inhibitor and protease residues Gly27, Asp29 and Asp30, two of which were with the main chain atoms (Figure 3.2.6C). In PR_{I50V}/IDV and PR_{L90M}/IDV complexes, additional hydrogen bonds were observed between IDV and the side chain of Arg8. Three water molecules (besides the one between flap and inhibitor) mediated extra interactions between PR and indinavir. Compared to IDV, in protease/UIC-94017 structures, the *bis*-THF group formed three extra direct hydrogen bonds to the main-chain NH groups of Asp29 and Asp/Asn30 and lost the one with the side chain of Asp/Asn30; and the presence of NH_2 group at the other end of the inhibitor allowed for three additional direct hydrogen interactions to Asp29' and Asp/Asn30' (Figure 3.2.6A). Except for the one H_2O between flap and inhibitor, no other water-mediated interaction between protease

and inhibitor was observed in PR/UIC-94017 complexes. However, direct hydrogen bonding is more favorable and conserved than water-mediated interactions. As expected, the water-mediated hydrogen bonds were not conserved in all mutant/IDV structures. So apparently, the inhibitor-main chain contacts are more conserved and less susceptible to mutations than the interactions with the side chains or intermediated by water. The significant difference in the hydrogen bond network of UIC-94017 and indinavir, especially the capability of making more hydrogen bonds to the main-chain atoms of protease by UIC-94017 provides the basis of the superiority of new inhibitor UIC-94017 over other known inhibitors.

Hydrophobic interactions are another important component for an inhibitor binding. Yet, the relationship between the amount of hydrophobic contacts within the protease and an inhibitor and the inhibitor's ability to suppress the activity of the protease mutants is not clear. A similar number of protease-inhibitor hydrophobic interactions was found in PR_{V82A}/IDV and PR/IDV complexes, while the K_i value of mutant V82A increased to two fold of that of PR. In contrast, the K_i value for UIC-94017 was three times higher for PR_{V82A} than for PR, while about 12 interactions were lost in PR_{V82A}/UIC-94017 structure compared to PR/UIC-94017.

Comparison of inhibitor interactions in PR and mutants

Four protease/inhibitor structures in the current series were obtained in same space group, $P2_12_12$, while the PR_{V82A} was the only one crystallized with UIC-94017 in the space group $P2_12_12_1$. The mutant structures in the same space group were very

similar, with the RMS deviations from PR of around 0.1 Å for the Cα atoms. The PR_{V82A} complex had a larger RMS deviation of 0.35 Å compared to PR, due to the differences in lattice packing in the two space groups. Big deviations of more than 1 Å were observed for regions of residues 37-40, 37'-48' and 54'-56', which were mostly located on the surface of the protein. The catalytic residues 25-27 showed very low RMS deviations of 0.02-0.12 Å for comparison of main chain atoms, consistent with the uniquely conserved active site. UIC-94017 had very similar conformation and interactions in all five structures, with the exception of the aniline group, which had two conformations in the PR_{V82A} structure. The occupancy distributions of two conformations of inhibitor are listed in Table 3.2.2. The hydrogen bond network and hydrophobic interactions were comparable for PR and all four mutants with some small variations caused by the point mutations.

PR_{V82A}/UIC-94017 showed ordered density for the inhibitor and two conformations for the aniline group as shown in Figure 3.2.5B. The hydrogen bond interactions with UIC-94017 were very similar to those of wild type enzyme except for those of the amino group (Table 3.2.4). The two conformations of the aniline group had relative occupancy of 0.56/0.44. Both conformations of the NH_2 group formed hydrogen bond interactions with the carboxyl oxygen of Asp30, as observed for PR. One conformation also had a weak (3.5 Å) hydrogen bond with the carbonyl oxygen of Asp30, similar to that observed for wild type enzyme. However, the NH_2 group in the other conformation interacted with a water molecule instead. These small differences suggested that PR_{V82A} had weaker polar interactions with UIC-94017 compared to those

of PR. Almost identical numbers of van der Waals interactions between protease and inhibitor were observed in PR and PR_{V82A} (91 contacts of less than 4.0 Å for PR_{V82A} compared to 91 for PR). However, PR_{V82A} showed a shift in the position of the main chain atoms compared to those observed for PR (Figure 3.2.8). The shift of 0.5 Å on the α-carbon atom of residue 82, together with the small shift of the benzene ring of UIC-94017, placed the β-carbon atom of Ala82 at 3.8-4.0 Å from the closest inhibitor atom, compared to 4.1 Å for β-carbon atom and 3.5-3.6 Å for γ-carbon atom of Val82 (Figure 3.2.8A). The other subunit showed a larger shift of 0.8 Å of Ala82 Cα that positioned the β-carbon atom of Ala82' at 4.1 Å from inhibitor, compared to 3.9 Å for Val82' (Figure 3.2.8B). These small shifts result in good van der Waals interactions of β-carbon atom of Ala82 with inhibitor and partially offset the loss of the methyl groups of Val 82. The structural adaptation of the PR_{V82A} mutant to accommodate the UIC-94017 inhibitor is consistent with the similar inhibition values observed for PR_{V82A} (relative K_i of 3.6) and PR.

The PR_{I84V}-inhibitor hydrogen bond interactions were essentially identical to those of PR, except that the lower occupancy conformation of inhibitor showed a weak (3.5 Å) interaction with the amide of Asp 30 (Table 3.2.4). The PR and PR_{I84V} had almost same hydrophobic contacts with inhibitor. There were 93 and 90 contacts of less than 4.0 Å calculated for the major conformation of UIC-94017 in PR and PR_{I84V}, respectively. PR_{I84V} mutant showed very similar main chain structure around the mutated residue compared to the wild type structure, as shown in Figure 3.2.9. Residue Ile84 in PR had two conformations of its side chain and both showed good van der Waals interactions

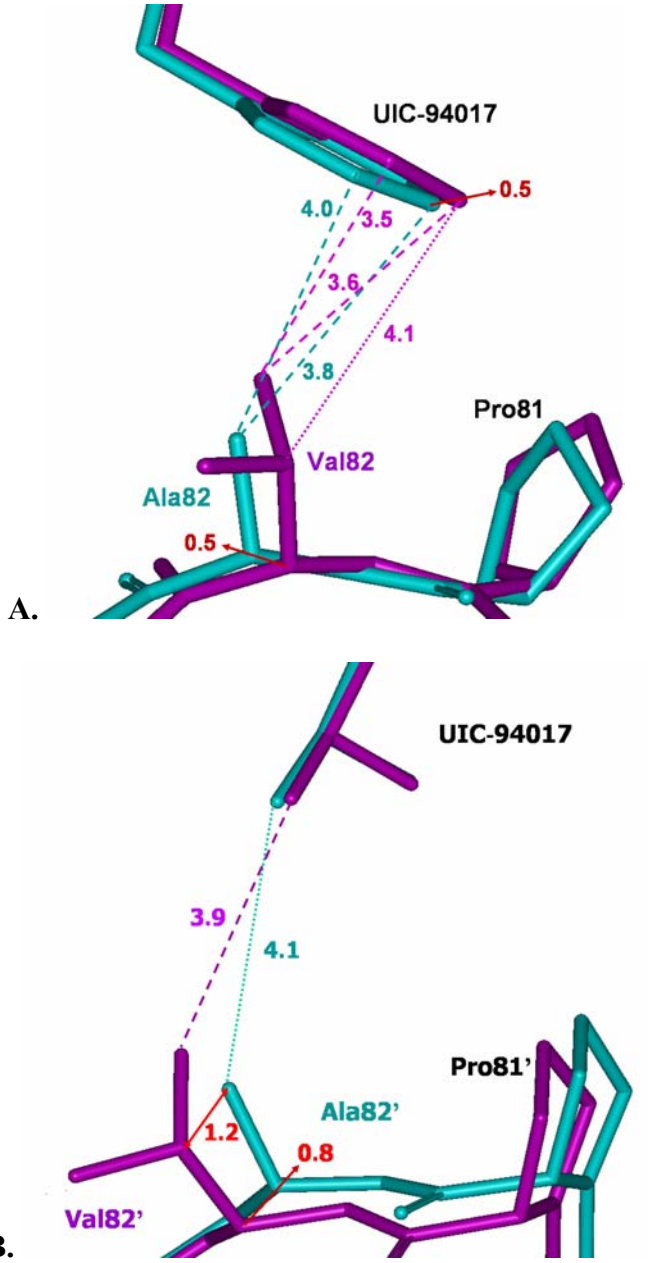


Figure 3.2.8 Interactions of mutated residue with UIC-94017 shown in superimposed PR_{V82A} and PR structures.
Only the P1 and P1' groups of UIC-94017 are shown with the interacting protease residues. Interatomic distances are indicated by dashed lines with the separation in Å. PR_{V82A} is shown in blue bonds and PR in purple. The major conformation is shown for the side chains of Val 82 and 82' in PR for clarity, whereas the minor orientation had the similar interactions compared to major one.
A) PR compared to PR_{V82A} near residue 82.
B) PR compared to PR_{V82A} near residue 82'.

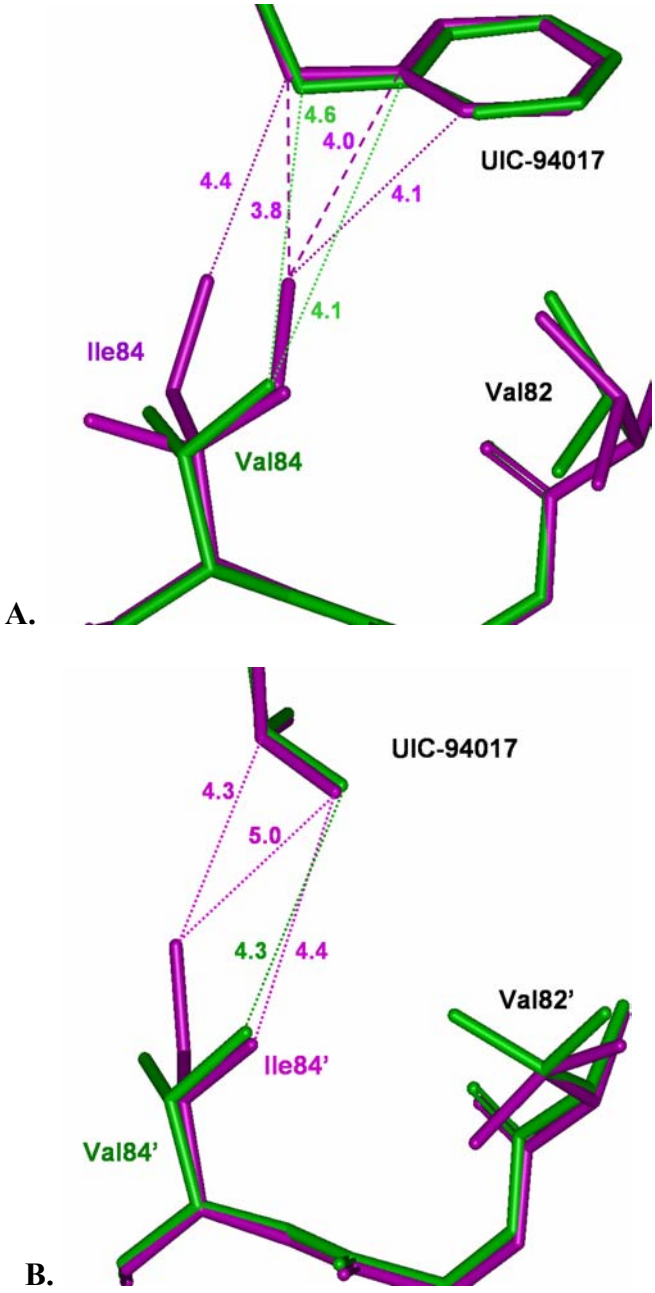


Figure 3.2.9 Interactions of mutated residue with UIC-94017 shown in superimposed PR_{I84V} and PR structures.
Only the P1 and P1' groups of UIC-94017 are shown with the interacting protease residues. Interatomic distances are indicated by dashed lines with the separation in Å. PR_{I84V} is shown in green bonds and PR in purple. The two alternate conformations with similar occupancy are shown for the Ile84 side chain in PR.
A) PR compared to PR_{I84V} near residue 84.
B) PR compared to PR_{I84V} near residue 84'.

with the inhibitor with interatomic distances of 3.8 to 4.1 Å (Figure 3.2.9A), whilst Ile84'in other subunit had only one position of its side chain and was at least 4.3 Å away from the inhibitor (Figure 3.2.9B). These close van der Waals interactions with inhibitor of Ile84 were lost in PR_{I84V}, consistent with the observed 5-fold lower inhibition of PR_{I84V} compared to PR.

As mentioned above, the two ends of the UIC-94017 are the good hydrogen-bond donor (aniline group) and acceptor (*bis*-THF group). Due to the presence of two inhibitor conformations these groups occupied almost the same sites and formed a tight hydrogen bond network with both Asp29/30, and Asp29'/30'. The *p*-NH_2 substituent of the aniline moiety interacted closely with residue 30 or 30', including contacts with both the main-chain atoms: amide nitrogen and carbonyl oxygen, and the side-chain carboxyl group. These hydrogen bonds were similar in the wild type and the mutant structures with differences observed in PR_{D30N} and PR_{I50V} complexes. In PR structure, Asp30 had alternate conformations with occupancies of about 50/50 % in both subunit, probably related to the two conformations of inhibitor, and same direct hydrogen bonds were established, as shown in Figure 3.2.10A. However, in PR_{D30N} structure, clear electron density suggested a single position for Asn30 (Figure 3.2.10B) and two possible positions for Asn30' (Figure 3.2.10C), and both of them had a water linking the side-chain carboxylate of residue 30 to the *p*-amino of inhibitor instead of the direct contact observed in PR, while the interactions of the main-chain atoms of Asn30/30' were maintained. Such water-mediated interactions with both residues 30 and 30' were not present in any other structure. Yet, in the PR_{D30N} structure, the *bis*-THF oxygens had

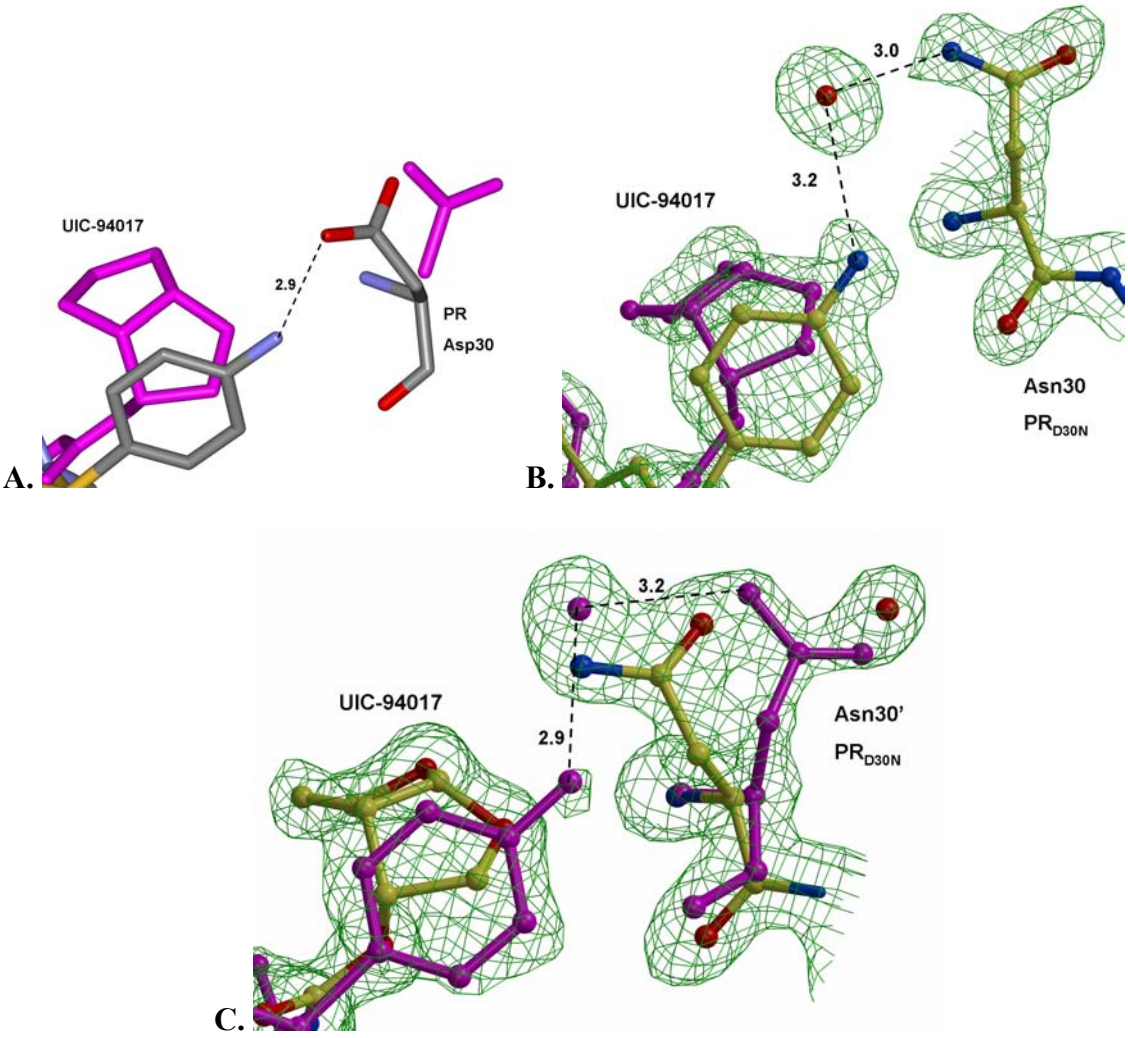


Figure 3.2.10 Comparison of the interactions of residue 30 with UIC-94017 in PR_{D30N} and PR structures.

Interatomic distances are indicated by dashed lines with the separation in Å. The major orientation of protease/inhibitor is colored by atom type and the minor orientation is shown in magenta. Maps were contoured at 1.6σ. Hydrogen bonds between the aniline ring of UIC-94017 and the main-chain atoms of residues 30/30', which were essentially same for PR and PR_{D30N}, are not shown.

A. Direct H-bond between carboxyl oxygen of Asp30 of PR with UIC-94017.
Similar interaction was observed in the minor orientation of inhibitor.

B. $2F_o$-F_c electron density map for residue Asn30 of PR_{D30N}.
A water-mediated H-bond was observed between the side-chain of Asn30 to aniline ring**.**

C. $2F_o$-F_c electron density map for residue Asn30' of PR_{D30N}.
The major orientation of protease/inhibitor is colored by atom type and the minor orientation is shown in magenta sticks. Two alternate conformations of Asn30' (56/44% occupancy distributions) and water-mediated contacts to the aniline moiety of UIC-94017 were suggested by clear maps. Interactions of *bis*-THF moiety are not shown.

hydrogen bonds to the amide nitrogen atoms of residues 29 and 30 (and 29', 30' for the other subunit) with the distances of 2.9-3.3 Å, which were similar to those of other four UIC-94017 complexes (Table 3.2.4). No significant difference in hydrophobic interactions was observed for PR_{D30N} and PR complexes. The changes in hydrogen bond network for PR_{D30N}/UIC-94017 may contribute to its 30 times worse Ki compared to PR.

In PR_{I50V} complex, a small bending of the major conformation of UIC-94017 was observed. This bending of the aniline ring induced a very strong hydrogen bond (2.5 Å) between *p*-NH_2 and the carboxylate of Asp30 and in turn, prevented the formation of other hydrogen bonds with the main-chain amide (3.4 Å) and carbonyl oxygen (4.1 Å) of the same residue (Figure 3.2.11); while for the other subunit, the interactions with the main-chain of Asp30' were similar to those of the other four complexes (Table 3.2.4). In addition to the changes in hydrogen bonding system, differences in hydrophobic interactions were observed at the mutation site caused by the loss of one methyl group. The sulfonamide oxygen atoms made close contacts with the terminal CH_3 group of Ile50 (3.4 Å). In the I50V mutant structure, there was a small movement of 0.2 Å on the Cα of Val50; yet, the valine isopropyl group was located opposite to the inhibitor and these interactions with the SO_2 moiety were even weaker (3.6-3.8 Å). Furthermore, the close contacts between Ile50 and aniline aromatic ring (3.6-4.0 Å) were totally interrupted by shorter side-chain of Val50 (>4.5 Å) (Figure 3.2.12A). On the other hand, Val50' in another subunit was oriented differently, so that the isopropyl group faced to the inhibitor. A larger shift of 0.4 Å on Cα of Val50' placed its two side-chain methyl groups at positions capable of forming two strong and almost symmetric interactions (2.6 and 2.8

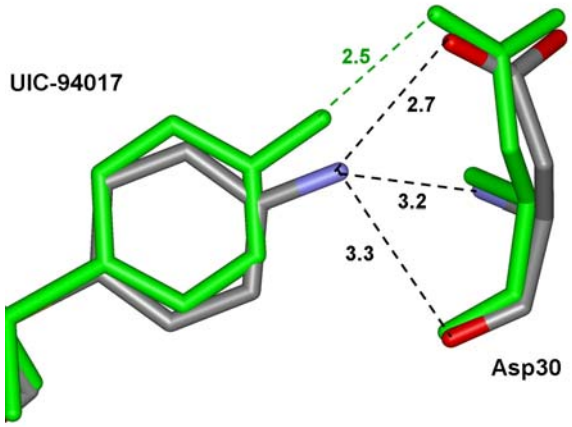


Figure 3.2.11 Interactions of the aniline group of UIC-94017 with residue Asp30 in PR and PR_{I50V} structures.
PR is colored by the atom type while PR_{I50V} in green sticks. The alternate conformation of the Asp30 side-chain in PR, and the second orientation of the inhibitor for both structures are omitted for clarity. The interactions of the other orientation of inhibitor with residue Asp30' of the other protease subunit are similar in all the structures.

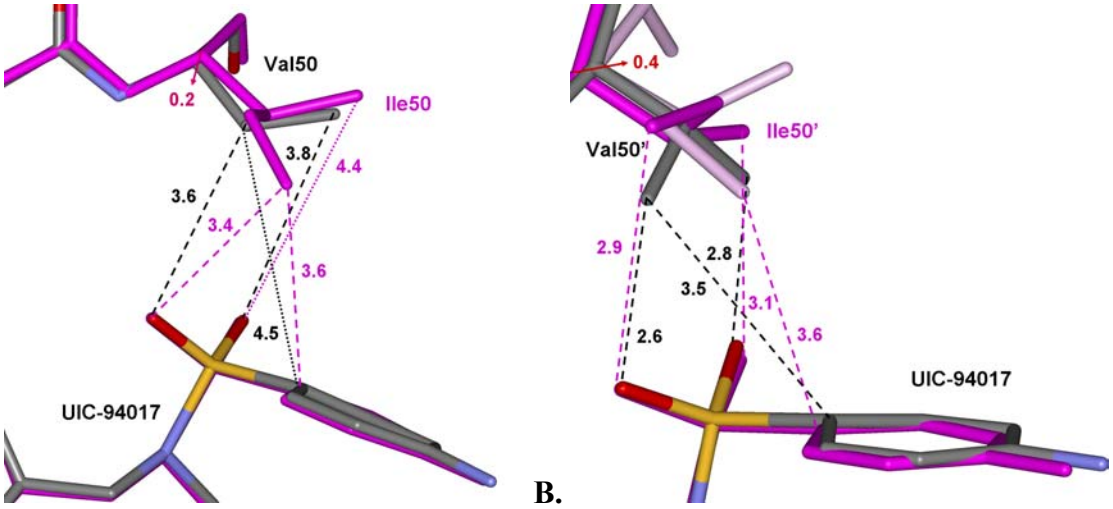


Figure 3.2.12 Interactions of mutated residue with UIC-94017 shown in superimposed PR_{I50V} and PR structures.
Interatomic distances are indicated by dashed lines with the separation in Å. PR is shown in magenta bonds (light magenta for the minor orientation of Ile50) and PR_{I50V} is colored by atom type. The interactions between CB atom of residue 50 and UIC-94017, which are very similar for PR or PR_{I50V}, and the second orientation of the inhibitor are omitted for clarity.
A) PR compared to PR_{I50V} near residue 50.
B) PR compared to PR_{I50V} near residue 50'.

Å) with each of the sulfonamide oxygen atoms (Figure 3.2.12B). The interactions from Ile50' CH_3 group to aniline group were also partially maintained by Val50' (~ 3.6 Å). Therefore, similar to PR_{V82A}, the small movement of main-chain atoms of residue 50' in PR_{I50V} towards inhibitor UIC-94017 partly compensated for the change from isoleucine in PR to a smaller amino acid valine in PR_{I50V}. Moreover, the observed 9 times lower inhibition constant for PR_{I50V} compared to PR was supported by the structural evidence of weaker hydrophobic contacts at mutation site 50/50' and the loss of hydrogen bonds with Asp30 in PR_{I50V}/UIC-94017 complex.

Geometry of catalytic aspartates and implications for reaction mechanism

The four carboxylate oxygen atoms of Asp 25 and 25' were arranged in almost the same plane and were at distances of 2.4-3.3 Å from the hydroxyl oxygen of UIC 94017 in all the structures (Table 3.2.4). Two inner carboxylate oxygens of the two catalytic aspartates lie within hydrogen bonding distance of each other in the HIV-1 protease structures (2.6-2.9 Å). Consequently, it has been assumed that one aspartate is protonated and the other has a negative charge. The catalytic aspartates have been shown to have different pKa values of 3.1 and 5.2, consistent with protonation of one catalytic residue (Hyland et al. 1991). NMR measurements have also indicated that one of the two catalytic aspartates was protonated in an HIV-1 protease-inhibitor complex (Wang et al. 1996). However, due to the symmetrical arrangement of the catalytic aspartates 25 and 25' in the dimer of HIV protease it has been difficult to determine the location of the expected proton. The existence and location of the proton has implications for the

reaction mechanism. Solvent isotope effects were consistent with the reaction proceeding by the rate-limiting step of dissociation of the enzyme-bound amide-hydrate intermediate (Hyland et al. 1991). Recent kinetic studies using a fluorogenic substrate at pH 6.0 and 1.25 M salt to obtain the solvent isotope effects have also suggested that cleavage of the carbon-nitrogen bond was the major contributor to k_{cat} (Porter et al. 2002). The proposed mechanism involves the catalytic aspartates together with a water molecule that acts as a nucleophile in attacking the carbonyl carbon of the scissile bond. This mechanism is difficult to support from the crystal structures of HIV protease with transition state analogs, since there is little space for a water molecule next to the catalytic aspartates. The closest water molecule cannot interact directly with the catalytic aspartates since it lies at the opposite side of the scissile peptide bond next to Ile 50 and 50' of the flaps.

The electron density for the catalytic aspartates can be examined for details relevant to the reaction mechanism. The crystal structures of PR, PR_{I84V}, PR_{D30N} and PR_{I50V} showed little or no Fo-Fc difference density around the Asp 25 and 25' side chains, in particular, there is no positive difference density suggesting a proton. These structures were refined at 1.2 to 1.5 Å resolution, and the lower resolution and disordered inhibitor may mask any density for a proton. However, the atomic 1.10 Å resolution of the PR_{V82A} structure and the presence of ordered density for inhibitor revealed differences in the electron density around the catalytic aspartates. In fact, significant Fo–Fc difference density was observed for the carboxylate oxygens, in contrast to the other Asp side chains in the structure (Figure 3.2.13). The two outer carboxylate oxygens (OD2) showed negative difference density. A streak of positive density was observed between

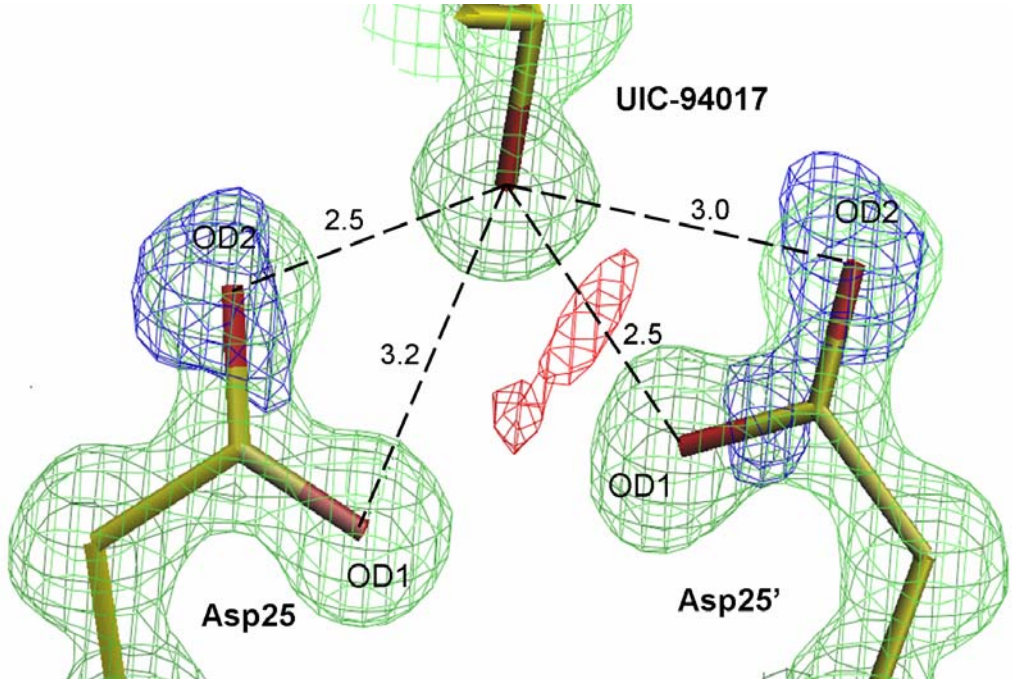


Figure 3.2.13 Electron density maps for catalytic aspartates of PR_{V82A}.
The 2Fo-Fc map is green and was contoured at a level of 1.7 σ, whereas the Fo-Fc map is contoured at 3.2 σ and colored red for positive and blue for negative. Polar interactions are indicated by dashed lines with the interatomic distances in Å. The carboxylate oxygen atoms of Asp 25 and 25' are labeled.

the hydroxyl group of the inhibitor and the inner and closest OD1 of Asp 25' (2.5 Å separation of OH and OD1), which may indicate a hydrogen atom. This positive density is reduced, but not eliminated, when hydrogen is added to the Asp 25' OD1. The unusual negative difference density can be interpreted as a redistribution of the electrons in the two C-O bonds in the Asp side chain to a less symmetric arrangement in which the C-O bond furthest from the center has more double bond character, and the other has less. Free atom refinement of the carboxylate groups gave distances of 1.27 Å for both C-O bonds in Asp 25, while Asp 25' showed asymmetry with 1.2 Å for C-OD1 and 1.3 Å for C-OD2, consistent with protonation of OD1 (Figure 3.2.13). However, the free atom refinement did not significantly change the negative density. Further analysis of the data by charge density analysis or quantum calculations will be necessary to understand the distribution of electrons at the active site. However, the difference densities suggest a more complex mechanism than protonation of a carboxylate oxygen.

Discussion

The UIC-94017 inhibitor is in clinical trials for further development as a new antiviral agent for treatment of primary and multi-PR inhibitor resistant HIV infections due to its favorable pharmacokinetics (Koh et al. 2003). The hydrogen bond interactions of UIC-94017 with HIV PR can be compared to those of substrate analogs. PR recognizes peptide substrates and peptidic inhibitors by means of a series of hydrogen bond interactions with the main chain atoms of the substrate, as observed for crystal structures of PR in complex with substrate analogs (Gustchina et al. 1994). These

interactions span about seven peptide residues lying in PR subsites S3-S4' (Louis et al. 2000). The clinical inhibitors lie primarily in subsites S2-S2' and generally show fewer hydrogen bond interactions. The UIC-94017 inhibitor is shorter than the substrate analogs and forms similar hydrogen bonds to PR main chain groups, including the conserved Asp 29 (Weber et al. 1997). In addition, new polar interactions are formed with the main chain amide, carbonyl oxygen and the carboxylate oxygen of Asp 29. These interactions resemble those of the P2' Gln or Glu side chain of peptide analogs (Weber et al. 1997; Mahalingam et al. 2001; Tie et al. 2005). These polar interactions of the *bis*-THF of UIC-94017 were not observed for the most closely related clinical inhibitor, amprenavir. Moreover, the PR/IDV structure contained only two direct hydrogen bonds to the PR main chain, while six such interactions were present in the wild type and mutant structures with UIC-94017. Hence, the UIC-94017 inhibitor more closely mimics many of the natural substrates in the interactions with the PR main chain atoms than observed for the other clinical inhibitors. Therefore, UIC-94017 may better adapt to mutations than other clinical inhibitors, making it more difficult for the virus to select resistant mutations.

The tight binding of the *bis*-THF group to the main chain atoms of Asp 29 and 30 are expected to be important for the potency of UIC-94017 against multi-drug resistant HIV (Koh et al. 2003). Such interactions are probably less susceptible to resistant mutations than those with the side-chain atoms. However, inhibition data for PR_{V82A}, PR_{I84V}, PR_{D30N}, PR_{I50V} and PR_{L90M} have shown that UIC-94017 does lose some affinity for the mutated protease, especially in the case of I50V and D30N mutations. The

relative K_i values were 3.6, 5, 9 and 30 for mutants PR_{V82A}, PR_{I84V}, PR_{I50V} and PR_{D30N}, respectively (Table 3.2.1). In the case of PR_{I84V}, an enlarged binding site cavity had been observed and it showed fewer van der Waals interactions of residue 84 with UIC-94017, consistent with a four-fold reduction in inhibition as compared to PR. Similar effect has been reported for mutant I54V that also introduces a smaller amino acid (Clemente et al. 2004; Tie et al. 2004). In contrast, compensating conformational changes have been observed in the crystallographic analysis of PR_{V82A} and PR_{I50V}. PR_{V82A} showed rearrangements of the main chain atoms around residue 82 and 82', which permitted closer interaction of Ala 82 and 82' with UIC-94017, and partly compensated for the fewer side chain atoms in Ala compared to Val. The structural rearrangements were consistent with similar inhibition observed for both PR_{V82A} and PR. The crystal structure reported for PR_{V82A} with the inhibitor A-77003, also showed rearrangements of the backbone atoms in one subunit (Baldwin et al. 1995), while a recent study described fewer interactions of Ala 82 with inhibitors saquinavir and ritonavir compared to those of Val 82 (Prabu-Jeyabalan et al. 2003). On the other hand, there were more differences in the interactions between the inhibitor and the protease in the PR_{I50V} structure, in agreement with its bigger change in Ki. Although the inhibitor's aniline group (major orientation) bent considerably from the position observed in any other structures (Figure 3.2.11) and decreased its contacts with the main-chain of Asp30, the hydrogen bond to the side-chain carboxyl oxygen was shortened by 0.2 Å and became stronger. Therefore, the strength of the aniline-to-Asp30 interaction might not be significantly affected. However, PR_{I50V} lost favorable interactions between the side chain of Val50 to the π-

system of the inhibitor aniline moiety as observed for Ile50 in PR complex (Figure 3.2.12A), while the structural adjustment of Val50' main chain and different conformation of Val50 side chain partially maintained such hydrophobic contacts (Figure 3.2.12B). As expected, the mutation on residue 30, which can form tight hydrogen bonds with UIC-94017, induced some effects on protease-to-inhibitor interactions. The inhibitor had water-mediated contacts to the side chain of Asn30/30'; while in PR a direct hydrogen bond was observed with one of the alternate conformations of Asp30/30'. Since other interactions in the two structures (PR and PR_{D30N}) were very similar, the substitution of direct contacts by water-mediated interactions may help explain the 30 times higher K_i value for PR_{D30N} compared to the wild-type PR. Besides, as hydrogen bonds are stronger than hydrophobic contacts, it may explain why the K_i value increased more for PR_{D30N} than for PR_{I50V}.

Structural and kinetic studies have suggested that mutations V82A and I84V, which are common in drug resistant clinical isolates, had minor effects on the inhibition of UIC-94017. So, they are expected to show resistance only in combination with other mutations. Moreover, the kinetic measurements had shown that PR_{D30N} was only 10% as active as the wild type protease on the fluorescent substrate, while PR_{I50V} was 40% as active. On the other hand, the potency of UIC-94017 was reduced by about 30 and 9-fold, respectively, for PR_{D30N} and PR_{I50V}. It implies that the mutations PR_{I50V} and especially PR_{D30N} may be selected by the virus for resistance to UIC-94017. Consequently, UIC-94017 is not expected to be a good salvage therapy candidate for the

patients who have failed nelfinavir regimen due to the presence of the viral strains that already contain the D30N mutation in the protease.

These new crystal structures of HIV PR mutants with UIC-94017 will provide the framework for the design of new inhibitors that are more effective against resistant HIV. The highest resolution structure of PR_{V82A} mutant provides an intriguing view of the catalytic residues and further analysis will help understand the protonation states of the catalytic aspartates and the reaction mechanism. The results have been published in (Tie et al. 2004; Kovalevsky et al. 2006).

3.3 Structural analysis of complexes with saquinavir

Aim 3: To determine & analysis crystal structures of PR and mutants PR_{V82A} and PR_{I84V} with saquinavir (SQV).

Saquinavir, the first protease inhibitor proved by FDA, is a peptide-like substrate analog. The kinetic data and atomic resolution structures of wild-type PR, drug resistant mutants PR_{V82A} and PR_{I84V} in complex with saquinavir were studied. These structures will give the molecular based information to help understand the drug resistance to saquinavir and provide valuable reference to compare with other PI drugs. The comparison will help understand the structural advantage and disadvantage of compounds and further help the design of more efficient next generation inhibitors.

Saquinavir inhibition of PR_{V82A} and PR_{I84V}

PR and the drug-resistant mutant proteases, PR_{V82A} and PR_{I84V}, catalyze the hydrolysis of the fluorescent substrate, a synthetic peptide (Matayoshi et al. 1990; Wang et al. 1990) based on the HIV-1 MA-CA cleavage site in the Gag-Pol polyprotein, and their catalytic activities are competitively inhibited by saquinavir. Their saturation kinetics with the same fluorescence assays were published previously (Kovalevsky et al. 2006); PR_{V82A} had similar kcat/K_m value as PR while that of PR_{I84V} was 65% of PR (Kovalevsky et al. 2006). The measured inhibition constants for saquinavir were 3.9 ± 0.8 nM for PR, 4.3 ± 0.7 nM for PR_{V82A}, and 4.3 ± 0.4 nM for PR_{I84V}. Surprisingly, the mutants PR_{V82A} and PR_{I84V} had essentially the same Ki value as PR. In contrast, most of the other clinical inhibitors showed significantly poorer inhibition of the PR with V82A

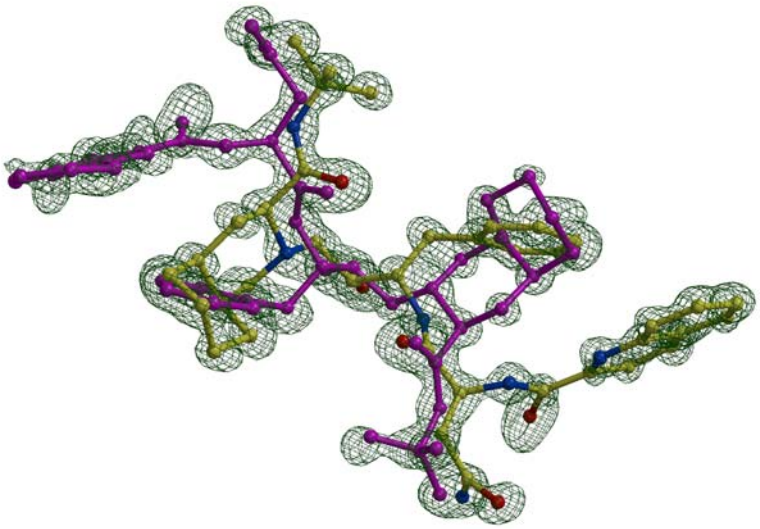
or I84V mutations relative to wild type PR (Gulnik et al. 1995; Klabe et al. 1998). However, this result is consistent with the rare appearance of mutations V82A or I84V, rather than L90M and G48V, as primary mutations in HIV clinical isolates exposed to saquinavir.

Crystal structures

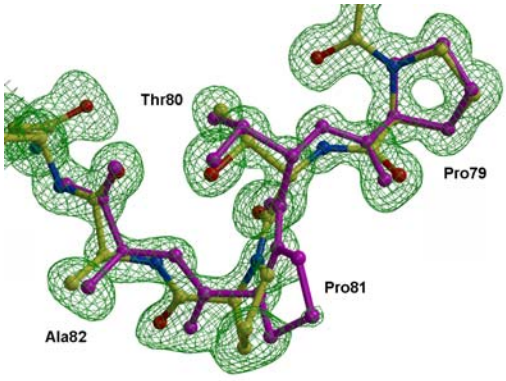
The crystal structures of PR and drug-resistant mutants PR_{V82A} and PR_{I84V} have been determined in their complexes with the clinical drug saquinavir. Crystallographic statistics are summarized in Table 3.3.1. The PR/SQV structure was determined in the $P2_12_12_1$ space group while the structures of the two mutants were determined in both space groups $P2_12_12$ and $P2_12_12_1$ with two different unit cells. Each crystallographic asymmetric unit contains one protease dimer. All structures were refined at resolutions of 0.97-1.25 Å and R factors of 0.12-0.15 including disordered residues, solvent molecules, anisotropic B-factors and hydrogen atoms. The electron density was excellent for all the atoms and especially for the sub-atomic resolution structure of PR_{V82A}/SQV complex. Saquinavir was observed in two alternate conformations in all structures, except for the PR_{I84V}/SQV complex in the $P2_12_12_1$ space group. The relative occupancies of the two conformations are listed in Table 3.3.1. The electron density map clearly showed two alternate orientations of saquinavir (Figure 3.3.1A). Alternate conformations were also observed for residues 79'-82' (Figure 3.3.1B), which had direct interactions with saquinavir, in all structures in the $P2_12_12_1$ space group. Similar to our other structures (Tie et al. 2004; Kovalevsky et al. 2006), the majority of residues with two alternate

Table 3.3.1 Crystallographic data statistics.

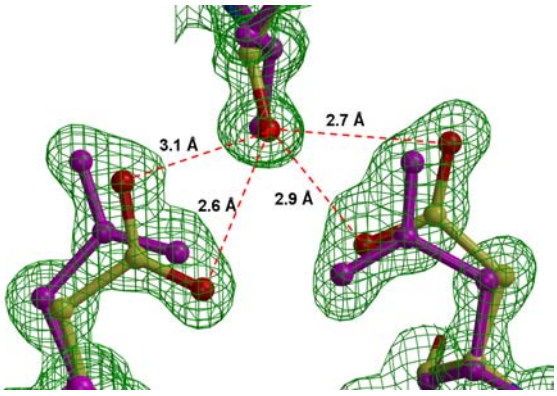
Protease		**PR**	**PR_{V82A}**	**PR_{V82A}**	**PR_{I84V}**	**PR_{I84V}**
Inhibitor		Saquinavir				
Space group		$P2_12_12_1$	$P2_12_12$	$P2_12_12_1$	$P2_12_12$	$P2_12_12_1$
Unit cell dimensions (Å)	**a**	52.10	59.99	51.13	58.58	51.05
	b	59.78	87.46	58.59	86.00	58.88
	c	62.49	47.19	61.61	45.83	61.50
Resolution range (Å)		50-1.16	50-1.10	50-0.97	50-1.25	50-1.20
Unique reflections		67949	95683	109869	64831	59108
Rmerge (%)		7.5	11	10	4.7	6.5
Overall (final shell)		(34.8)	(36.0)	(25.1)	(50.6)	(36.4)
<I/sigma>		14.6	10.5	20.5	11.3	11.1
Overall (final shell)		(3)	(2)	(4)	(2)	(2)
Data range for refinement (Å)		10-1.16	10-1.10	10-0.97	10-1.25	10-1.20
Rwork (%)		0.13	0.15	0.12	0.15	0.15
Rfree (%)		0.18	0.18	0.14	0.19	0.20
No. of waters (total occupancies)		185.5	217	182.2	166	138.2
Completeness (%)		96.6	94.6	89.3	92.8	92.6
Overall (final shell)		(72.6)	(64.5)	(53.8)	(61.7)	(64.3)
RMS deviation from ideality						
Bonds (Å)		0.017	0.020	0.016	0.013	0.013
Angle distance (Å)		0.041	0.045	0.035	0.033	0.033
Average B-factors (Å2)						
Main chain		12.3	13.7	8.8	15.8	15.0
Side chain		17.6	19.5	13.3	21.7	21.9
Inhibitor		11.3	13.3	9.4	14.2	16.7
Solvent		33.4	35.2	22.1	30.4	28.8
Relative occupancy of inhibitor		0.5/0.5	0.5/0.5	0.7/0.3	1	0.7/0.3



A.



B.



C

Figure 3.3.1 2*Fo*-*Fc* electron density map examples in PR_{V82A}/saquinavir complex structure.
Major conformation is colored by atom type while minor one by magenta.
A. Omit map of saquinavir (contoured at 3σ level)
B. 2*Fo*-*Fc* electron density at residue 79-82 (contoured at 1.8σ level)
C. 2*Fo*-*Fc* electron density at catalytic site (contoured at 1.8σ level)

conformations had long side chains and were located at the surface of the protein, while others were hydrophobic amino acids forming the inhibitor binding site, including Val82 and Ile84. The average B-factors were low for protein and inhibitor atoms, and decreased as the resolution of the data increased. The B-factors were in the ranges of 8.8-15.8, 13.3-21.9, 8.7-16.7 and 22.1-35.2 $Å^2$ for protein main chain atoms, side chain atoms, saquinavir, and solvent atoms, respectively.

Saquinavir, the catalytic triad, flap residues and 80's loop show correlated alternate conformations at sub-atomic resolution

The atom positions are defined very accurately with the sub-atomic resolution data for the PR_{V82A}/SQV complex in the $P2_12_12_1$ space group. Residues 24-26 and 24'-26' in the active site were observed in two conformations with the occupancy distribution of 0.55 to 0.45 (Figure 3.3.1C), which was the same as that of saquinavir. This suggests that the conformation of the active site residues is correlated with the alternate conformations of the inhibitor saquinavir. After refinement with the two conformations the partial positive and negative Fo-Fc density around the catalytic triad disappeared. As shown in Figure 3.3.1C, the hydroxyl group on the saquinavir formed 2.6, 3.1, 2.9 and 2.7 Å long hydrogen bonds to Asp25 OD1, OD2, Asp 25' OD1 and OD2 respectively. The other orientation of the inhibitor showed the same arrangement and distances. These more accurately determined atom positions of the active site residues may have not been correctly modeled before in other structures with two conformations of inhibitor. Thus,

the sub-atomic resolution data has more accurately described the detailed geometry of the PR active site.

The tips of the flap, residues 48-52 and 48'-52', were also modeled with two conformations with the same relative occupancy as those of saquinavir. The major orientation of the flap residues was correlated with the major conformation of saquinavir, while the minor one correlated with the minor conformation of saquinavir. The main chain peptide group of Ile50/50' had two conformations related by a 180 degree flip, and also the backbone of flap region residues 48-52/48'-52' was moved about 0.2 Å to provide more space for big P1' ring of saquinavir.

The third region with alternate conformations was the 80's loop in both subunits (residues 79-82 and 79'-82'), as shown in Figure 3.3.1B. In the PR/SQV and PR_{I84V}/SQV structures (in the $P2_12_12_1$ space group), the alternative conformations were observed only in one subunit for residues 79'-82'. Since this region plays an important role in binding inhibitor, it will be further discussed later.

Protease interactions with saquinavir

Saquinavir is a peptide-like inhibitor containing the main chain amides and carbonyl oxygen atoms of a peptide and a variety of groups corresponding approximately to the side chains at positions P3-P2' of a peptide substrate (Figure 3.3.2). The large hydrophobic groups at positions P3, P1 and P1' of saquinavir fit into the hydrophobic pockets S3, S1 and S1' of the PR. There is a smaller hydrophobic t-butyl group at P2' and the polar asparagine side chain at P2 of saquinavir. The P1 and P1' groups of

saquinavir formed van der Waals interactions with PR residues Leu23', Asp25/25', Gly27, Ala28, Gly49', Ile50, Thr80, Pro81, Val82' and Ile84/84'. The other groups of saquinavir interacted with residues Gly27, Ala28/28', Asp29, Asp30/30', Gly48/48', Gly49/49', Ile50, Pro81' and Ile84'. The residues 79' to 82' in the PR/SQV structure interact closely with the saquinavir and showed alternate positions with the occupancy distribution of 0.7 and 0.3. This suggests that the flexibility of this region allows the PR to adapt to various subgroups on the different inhibitors. Theoretically, the flexibility of the structure will increase the entropy and in turn reduce the interaction energy between protein and inhibitor.

The PR hydrogen bond interactions with saquinavir are shown in Figure 3.3.3. There were seven direct hydrogen bonds formed between protease and saquinavir. Four hydrogen bonds were formed between the hydroxyl group of saquinavir and the side chain carboxylate oxygen atoms of the catalytic Asp 25/25', and the other three were from O1 of saquinavir to the main chain amide of Asp29, from N3 to the amide nitrogen of Asp30, and from N3 to the main chain carbonyl oxygen of Gly48. Two water molecules mediated three more H-bond interactions. The conserved water linked the PR flaps and saquinavir, and the other water mediated the interactions between O1 of saquinavir and the main chain oxygen of Gly27 and side chain oxygen of Asp29. An equivalent hydrogen bond network was observed for both conformations of saquinavir.

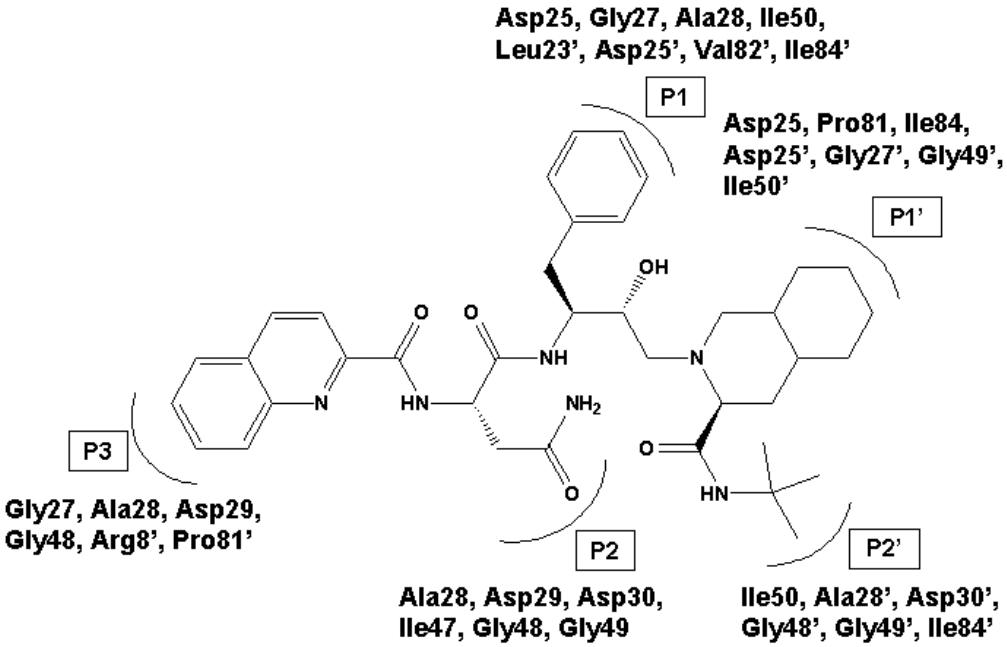


Figure 3.3.2 van der Waals interactions between PR and saquinavir at different subsites.

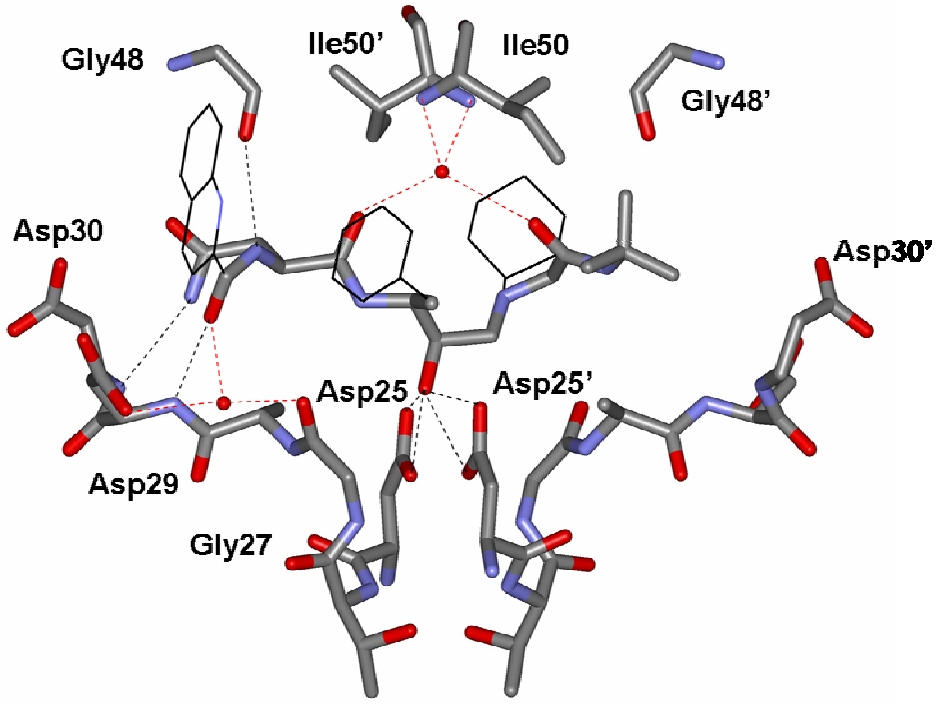
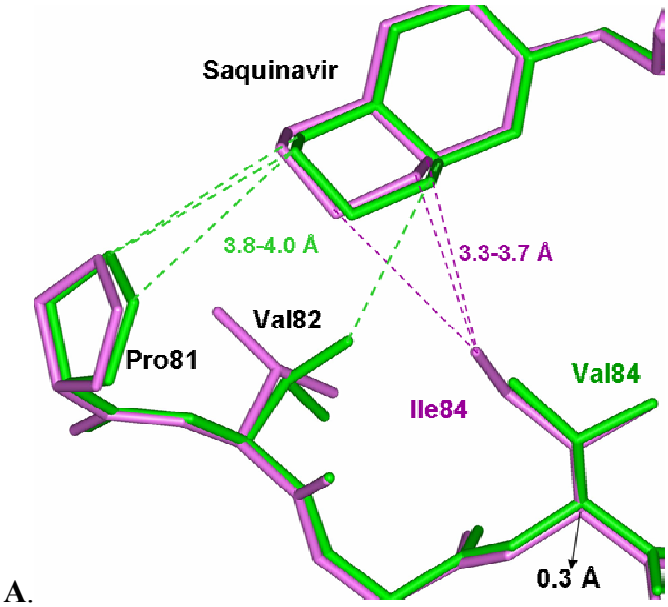


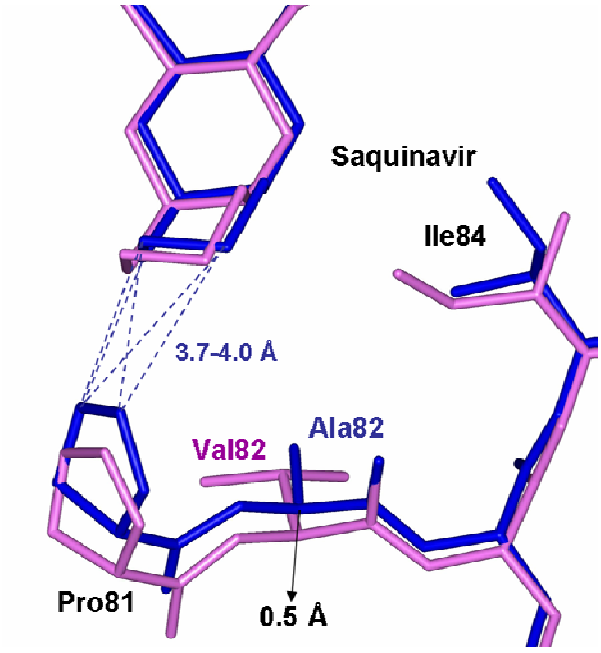
Figure 3.3.3 Hydrogen bond network between PR and inhibitor saquinavir.
Interactions mediated by water are shown in red.

Comparison of the structures of PR and mutants

The three structures in the same space group of $P2_12_12_1$ were very similar overall, with calculated RMS deviations of 0.43 and 0.41 Å for C alpha atoms in PR_{V82A} and PR_{I84V}, respectively, compared with PR. Saquinavir was bound in two conformations in all complexes, except for PR_{I84V}/SQV in the $P2_12_12_1$ space group (Table 3.3.1). Saquinavir has very similar conformation and interactions in the five crystal structures. The wild type PR and mutants PR_{I84V} and PR_{V82A} had almost identical hydrogen bond interactions with saquinavir, except that the hydroxyl of saquinavir was located more asymmetrically with respect to the side chains of catalytic Asp25 and 25' in the mutant complexes (Figure 3.3.1C). The distances from O4 to OD1, OD2 of Asp25 and Asp25' are 2.6, 3.1, 2.9 and 2.6Å in the mutants compared to 2.8, 2.9, 2.9 and 2.9 Å in the PR/SQV structure. Consequently, the N4 and O3 atoms of P2 were about 0.2 Å closer to residues 27 to 30 in the mutant structures. Both PR_{I84V} and PR_{V82A} showed shifts in the position of the main-chain atoms at the site of mutation compared to those observed for PR (Figure 3.3.4). The shifts of the alpha-carbon of Val84 were 0.25 Å and 0.3 Å in the two subunits (Figure 3.3.4A). However, this movement was perpendicular to the plane of the P1' decahydroisoquinoline ring and did not alter the overall interactions. In fact, the Ile84 side chain was located opposite to the inhibitor, and the close contacts of 84 CD1 with the C31, C32 and C33 atoms of saquinavir (3.5, 3.3, 3.7 Å, respectively) were eliminated in the PR_{I84V} mutant and substituted by the 3.8-4.0 Å contacts with residues 81 and 82 (Figure 3.3.4A). The PR and PR_{I84V} had almost identical numbers of van der



A.



B.

Figure 3.3.4 Shifts at residues 82-84 in mutant complex compared to PR/SQV. Structure of PR/SQV is colored in magenta while PR_{V82A}/SQV in blue and PR_{I84V}/SQV in green.

A. PR_{I84V}/SQV compared to PR/SQV.

B. PR_{V82A}/SQV compared to PR/SQV.

Waals contacts with saquinavir, except for the P2 position where the main chain O2 atom of saquinavir was 0.1-0.2 Å closer to Ile50' of the flap in PR_{I84V}.

A shift of 0.5 Å towards saquinavir was observed for the alpha-carbon of Ala82 in both subunits of PR_{V82A} (Figure 3.3.4B). Similar to PR_{I84V}, the Val82 side chain was located opposite to the P1' ring and was separated further from saquinavir (4.3-4.9 Å interatomic distances) in the PR structure than was the side chain of Ala82 in the mutant. In the mutant with Ala82, the CB atom was 0.2 Å closer to the P1' but the distances were still greater than 4 Å. Yet, the Pro81 also shifted and formed van der Waals contacts of 3.7-4.0 Å with the C29, C30 and C31 atoms of the decahydroisoquinoline group of saquinavir (Figure 3.3.4B), which were 0.3 Å closer compared to those of PR. The movement of Ala82 in the other subunit allowed 0.1-0.2 Å closer interactions with the P1 benzene ring. Similar movements of Ala82 in PR_{V82A} compared to PR have been reported previously for complexes with substrate analogs or UIC-94017 (Tie et al. 2004; Tie et al. 2005). In those structures the shifts partially compensated for the loss of van der Waals interactions between Ala82 and inhibitor caused by the absence of the two methyl groups of Val. No such compensating interactions were observed for saquinavir. All these observations suggest that the mutation to smaller amino acids at residues 82 and 84 can provide more space and freedom for the PR to accommodate the larger groups of saquinavir. The observed minor structural adaptation of the mutants to accommodate saquinavir was consistent with the very similar inhibition constants observed for mutants and PR.

Saquinavir complexes with PR_{V82A} or the $PR_{D25N/V82A}$ (1MTB) inactive mutant had similar structural differences as observed for comparison of the complexes with peptide analogs (Tie et al. 2005). The hydroxyl oxygen of saquinavir was 0.23 Å closer to the catalytic Asp25/25', which suggested that the active PR with Asp25 was closer to the transition state.

Comparison of PR/SQV with PR/UIC-94017

The PR complexes with saquinavir and UIC-94017 were compared to understand how PR binds the two different inhibitors. Both UIC-94017 and saquinavir have a central hydroxyl group which forms strong hydrogen bonds with all four carboxylate oxygens of catalytic Asp25 and Asp25'. Both inhibitors share a water mediated interaction with the flaps. UIC-94017 (Tie et al. 2004; Kovalevsky et al. 2006) formed seven other direct hydrogen bonds with PR, including six with the main chain and side chain atoms of Asp29 and Asp30/30' (Figure 3.2.6A). In contrast, saquinavir showed three direct and one water-mediated hydrogen bonds concentrated at one end of the inhibitor (P2 and P3) including only two interactions with Asp29/30 (Figure 3.3.3). The hydrogen bond to carbonyl oxygen of Gly48, which was observed in PR/SQV but not presented in PR/UIC-94017, was similar to the interactions observed for substrate analogs at the P2 position (Tie et al. 2005). The hydrophobic interactions dominated in PR/SQV. Saquinavir has larger hydrophobic groups at P1 and P1' that provide more van der Waals interactions with PR than observed for UIC-94017 (99 contacts of less than 4.0 Å for PR/SQV compared to 91 for PR/UIC-94017). The first generation PR inhibitor, saquinavir was

designed to mimic the hydrophobic interactions between substrate and PR, while the design goal for UIC-94017 was to reduce its susceptibility to mutations by making more hydrogen bonds to the main-chain atoms of PR.

In order to accommodate the large decahydroisoquinoline ring at P1' of saquinavir the 80's loop of residues 79-82 has undergone a conformational change by comparison to its position in the PR/UIC-94017 structure (Figure 3.3.5). The 80's loop has moved away from the decahydroisoquinoline ring with a 1.7 Å movement of the nitrogen atom of Pro81. This structural adjustment has allowed Pro81 to form favorable van der Waals interactions of 3.8-4.2 Å separation with the tip of the P1' ring. The other subunit was different; the residues 79'-82' have adopted two conformations. The major conformation has shifted by 0.8 Å compared to the corresponding region of PR/UIC-94017 and established 3.8-4.0 Å favorable contacts with the atoms at the tip of the P1 phenylalanine group of saquinavir. In the conformation with 49% population, the nitrogen atom of Pro81' was 1.8 Å away from its position in the PR/UIC-94017 complex, which was comparable to the position described for the other subunit, and interacted with the big P1' ring of the minor conformation of saquinavir with interatomic distances of 3.6-4.0 Å. Two alternate conformations of residues 79'-82' were observed in PR and both mutant structures, while the other subunit showed two conformations only in the PR_{V82A}/SQV complex (Figure 3.3.1B). This result suggests that the 80s loop is intrinsically flexible and its position may be influenced more by the properties of the inhibitor than by the PR mutations in this loop.

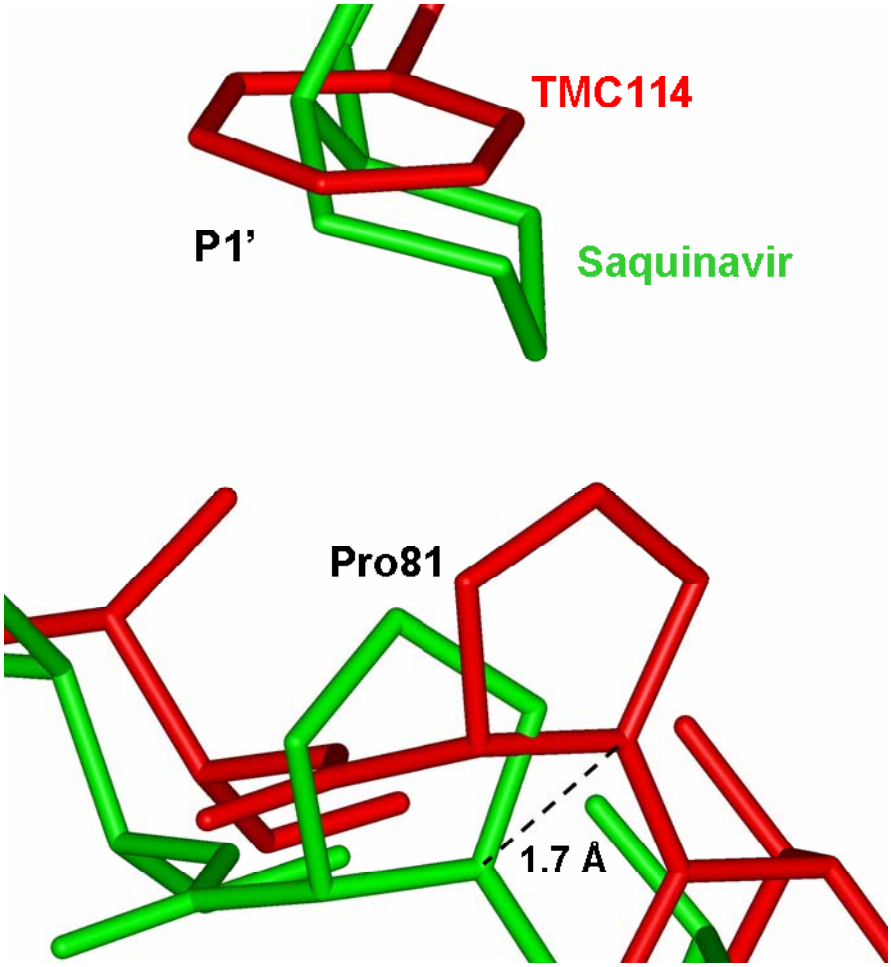


Figure 3.3.5 Comparison of P1' and S1' positions in PR/saquinavir and PR/UIC-94017 complex structures.
Structure of PR/UIC-94017 is colored in red while that of PR/saquinavir in green.

Comparison of the structures in two different space groups

Since the two mutant complexes were crystallized in both the $P2_12_12$ and $P2_12_12_1$ space groups, it was possible to analyze the structural differences due to different packing in the two crystal lattices. The structures of PR_{V82A}/SQV and PR_{I84V}/SQV in the two space groups showed RMS deviations on alpha carbons of 0.7 and 0.5 Å, respectively. There were no major differences in the protease interaction with saquinavir in the two space groups and the active site residues showed a very low RMS deviation of 0.1 Å. The biggest difference was for the surface residues 37-42, where the RMS deviation was over 1 Å. This region has lattice contacts with a symmetry related molecule in the $P2_12_12$ space group but not in the other space group. The flexibility of this loop did not affect the overall packing of the protease. The other major difference was that the 80's loop in the $P2_12_12$ structure had a single conformation, whereas this region of the structure in $P2_12_12_1$ had two alternate conformations. This comparison confirmed that the differences induced by lattice packing are mainly on the protein surface and have little effect on the details of protease-inhibitor interactions.

Torsion angles were calculated for each residue in the PR_{V82A}/SQV complex and compared for the structures in the space groups $P2_12_12$ and $P2_12_12_1$ with different unit cells to study the flexibility and packing of each PR residue. The dihedral angles of the backbone (shown in Figure 3.3.6) and specific main-chain hydrogen bond interactions are typically the key to protein folding. The observed range of the torsion angles and the rotational degrees of freedom studied here provide valuable information to apply in molecular models and predictions of protein structures. The average change in main

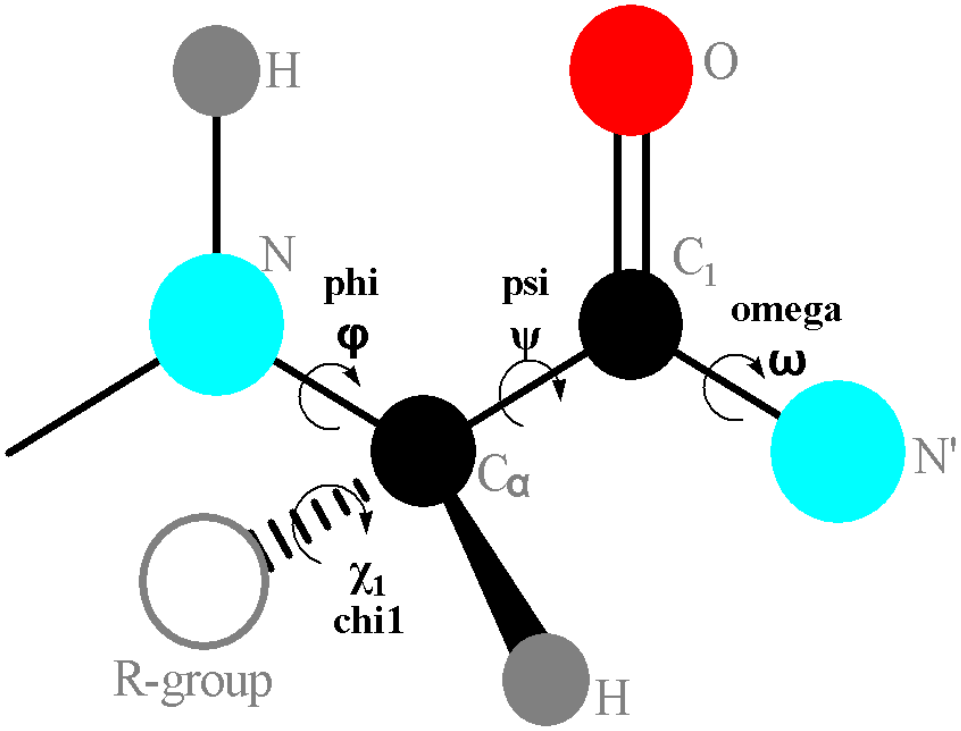


Figure 3.3.6 Definition of the dihedral angles of protein backbone.

Table 3.3.2 Comparison of two space group $P2_12_12_1$ and $P2_12_12$ for PR_{V82A}/SQV complex.

	Average changes for all residues			
Angle analyzed	psi	phi	omega	chi1
Degree (exclude 37-41/37-41')	**3.67**	**4.20**	**2.40**	**8.22**
Degree (include 37-41/37-41')	**4.27**	**4.53**	**2.49**	**9.34**

chain phi, psi and omega torsion angles, and side chain chi1 torsion angle for all residues in the PR_{V82A}/SQV complex in the two space groups is given in Table 3.3.2. The average changes were 3-4.5, 2.5 and 8-10 degree for phi/psi, omega and chi1, respectively. The dihedral phi and psi angles of the main chain are more restricted than the chi angles of the side chain, yet there were some variations. The difference in phi angle for residues Leu38 and Gly40' in the structures in the different space groups was over 50 degree, while other residues had differences of less than 20 degree. As mentioned above, loop 37-42 is located at the "shoulder" position leading to the flap of the PR and showed the biggest RMS deviation between structures in the two space groups. About 9% of the residues had differences (in phi angle) of 10 to 20 degree and most of these were located at the protein surface. Another 18% residues differed by 5 to 10 degree and were located at various positions: the dimer interface at both termini, the flaps and the residues in contact with the saquinavir, such as residues 28-31 and 80-82. The plots of differences in phi angles and in psi angles for each residue had similar shapes and distributions, which suggested that the changes in the two angles were coupled. Again, residue Ser37 had the biggest difference in psi angle of 68 degree, while Leu38' and Pro39'differed by about 25 degree, and all other residues varied by less than 20 degree. The big variation in phi and psi angles of loop 37-42 was correlated with its different lattice contacts in the two space groups. In $P2_12_12_1$, this loop interacts with residues within the same dimer, while in $P2_12_12$ it forms hydrogen bonds with a symmetry-related dimer. Around 10% residues had difference (in psi angle) of 10 to 20 degree and 20% residues had differences of 5 to 10 degree. There was less deviation in omega angles, where the change for all residues

was below 30 degree. 98% of protease residues showed difference of less than 10 degree in the omega angle; the exceptions were Met36/36', Ile64 and Gly49'.

The chi1 torsion angle of the side chain was much more flexible than the main chain dihedral angles. About 10% of the residues had differences of over 20 degree. A plot of the average change in torsion angles against amino acid type (excluding the loops 37-41/37'-41' which were involved in different lattice packing interactions) (Figure 3.3.7) further proved that the chi1 angle had the largest deviation, especially for residues with long side chains, such as Glu and Lys, and some hydrophobic residues, such as Pro and Ile. For residues at the protein surface with different chi1 angles, for example: loop residues 37-42, flap residue Lys55 and turn 64-65, the side chains can be perpendicular to each other or even flipped by 180 degree, depending on the environment (see examples in Figure 3.3.8). The difference in chi1 angle between structures in the two space groups was in the range of 0-15 degree for the core residues of the protein. The most important feature is that the changes in the chi1 angle showed no correlation with the changes in the psi and phi angles. The chi1 angle appears to depend more on the property of the amino acid and its chemical environment. Met was the only amino acid that had very similar and relatively high change in all four studied torsion angles. However, this result may not be significant since there are only four methionines in HIV-1 PR.

Comparison of the torsion angles for the structures solved in two space groups suggested that structural flexibility is an important feature of the PR and may be essential to its function. This analysis defines the level of freedom in torsion angles that should be applied during the molecular modeling process.

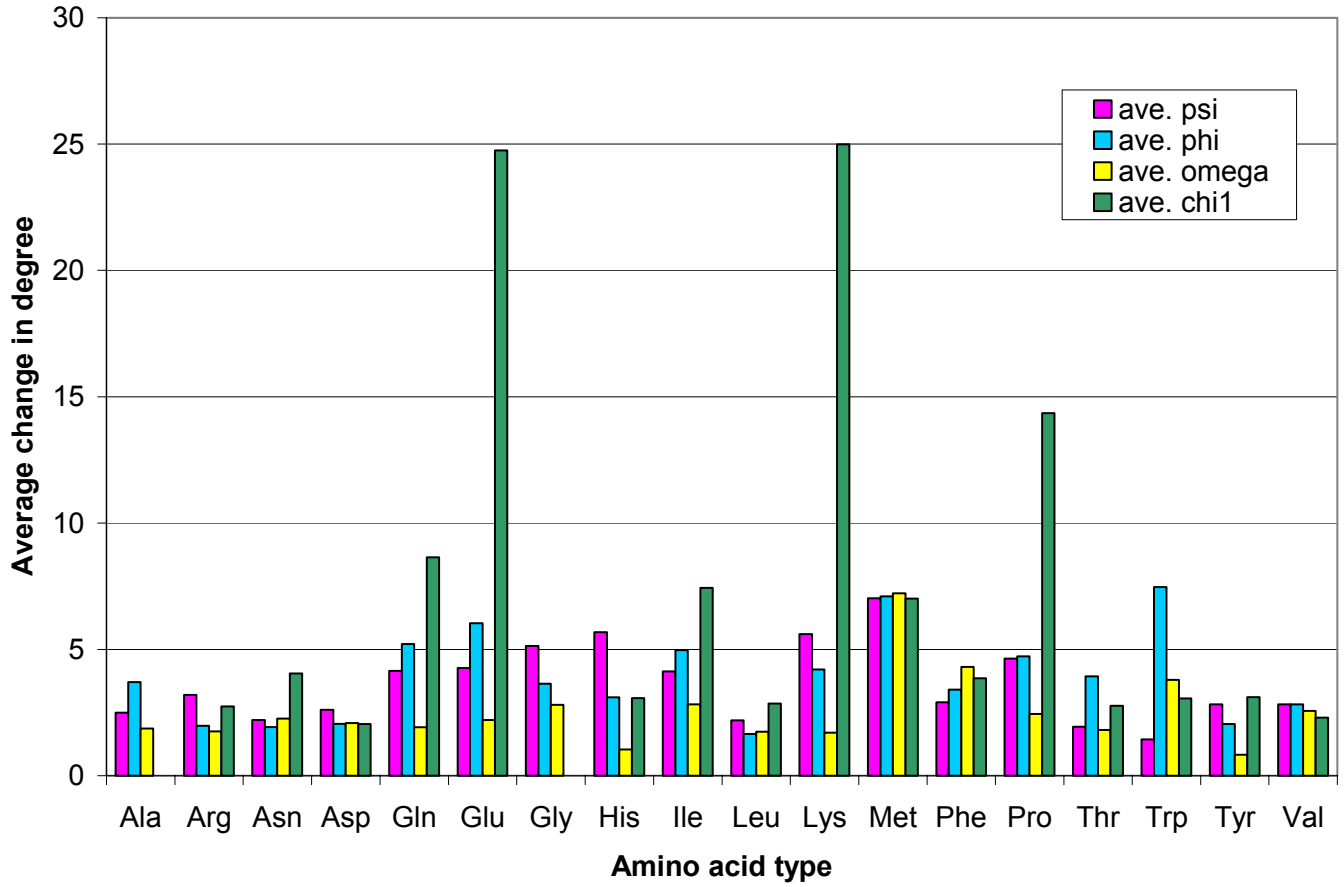
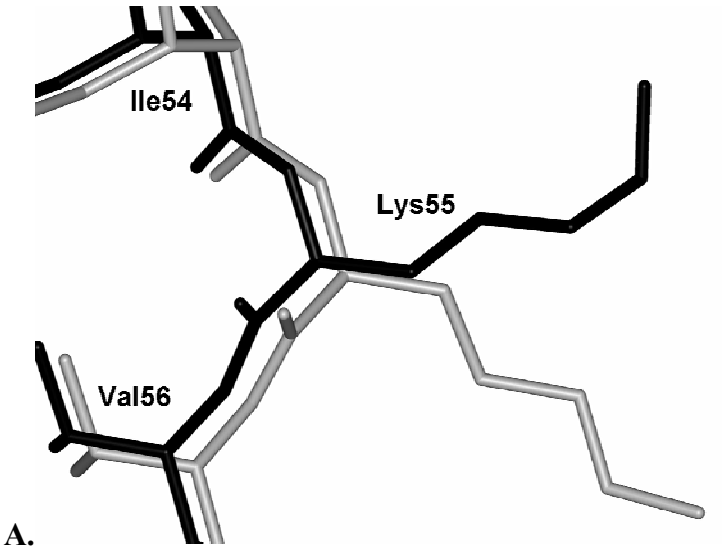
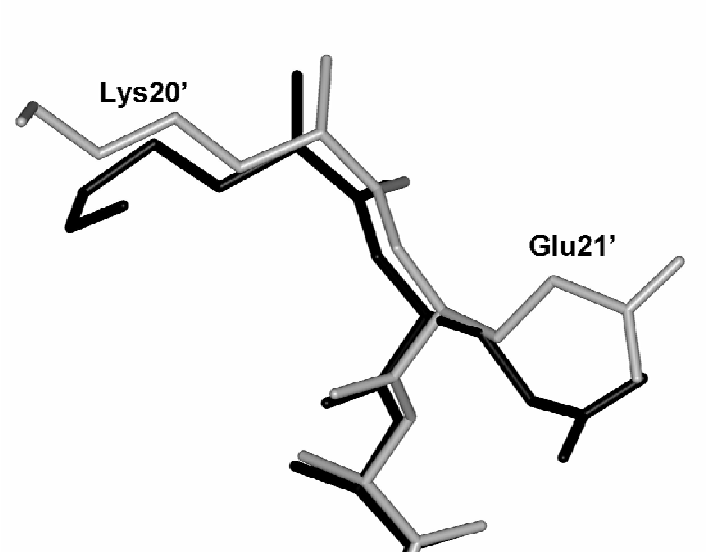


Figure 3.3.7 Average deviation of the torsion angle of different amino acid type between two space group $P2_12_12$ and $P2_12_12_1$ for PR_{V82A}/SQV complexes.



A.



B.

Figure 3.3.8 Different side chain orientation of same residue in two different space groups ($P2_12_12$ and $P2_12_12_1$) for PR_{V82A}/SQV complexes.
Positions in $P2_12_12_1$ is colored in black while those of $P2_12_12$ in grey.
A. Residues around Lys55.
B. Residues around Glu21'.

Discussion

These new atomic resolution structures have revealed details of the PR-saquinavir interactions and the structural changes in the presence of mutations or between different space groups. The PR interactions with saquinavir, the first FDA-approved anti-HIV drug, can be compared to those with peptide analogs or more recent clinical inhibitors. There were only seven direct and two water-mediated hydrogen bonds between saquinavir and PR, and they were focused at the P3-P2 end of saquinavir. The van der Waals interactions appeared to be the major contribution to the PR affinity for saquinavir. The binding affinity is determined by the Gibbs energy equation, so it depends on the enthalpy (ΔH) and entropy (ΔS) changes. Either making ΔH more negative or ΔS more positive will improve the PR-inhibitor binding affinity. Saquinavir binding to HIV-1 PR is enthalpically unfavorable and is driven by the large favorable entropy change (Velazquez-Campoy et al. 2000). The big rigid hydrophobic groups of saquinavir tend to maximize the gain in desolvation entropy with a small loss in conformational entropy due to the burial of these hydrophobic groups. Recent antiviral PR inhibitors, such as UIC-94017, were designed to enhance the strong and favorable hydrogen bonds with the main chain atoms of the PR. UIC-94017 shows an enthalpically favorable binding to PR with less dependence on the entropy. This design may also improve the specificity and provide better water solubility. Ideally, the inhibitor design should combine the two approaches, maintaining good hydrogen bonds with main chain atoms and improving the van der Waals contacts in the PR subsites, in order to provide tight binding and more effective inhibitors. In agreement with (Velazquez-Campoy et al. 2000), balancing the

enthalpically favorable polar interactions and desolvation entropy induced by hydrophobic groups will optimize the PR-inhibitor binding affinity and specificity.

Comparison of the complexes of saquinavir with the mutants has helped to understand drug resistance. Shifts of 0.3 and 0.5 Å on the alpha carbon atom of the mutated residue were observed for PR_{I84V} and PR_{V82A}, respectively, compared to the position in wild type PR. These changes did not significantly alter the PR-saquinavir interactions. The major difference at the active site was that the hydroxyl group on saquinavir was located more asymmetrically between the side chains of catalytic Asp25 and 25' in the mutant complexes. The observed minor structural changes in saquinavir complexes were consistent with the closely similar inhibition constants of mutants and PR. In contrast, other inhibitors have shown reduced inhibition and interactions with mutant compared to wild type PR. In these saquinavir complexes, the mutation from bigger to smaller amino acids allows more space to accommodate the bigger side chain at P1' of saquinavir.

The P1' decahydroisoquinoline group of saquinavir is very big for the S1' binding site. Compared to PR/UIC-94017, residues 79-82 have moved away from the active site while the residues 79'-82' have adopted two conformations in order to accommodate the big groups of saquinavir. The analysis further reveals that the 80s loop is intrinsically flexible. The flexibility of 80s loop was observed in all three structures, PR, PR_{V82A} and PR_{I84V}, which suggested that the structure of the 80s loop was more dependent on the specific inhibitor than on the mutations. Moreover, shifts of the 80s loop were observed for structures in both the $P2_12_12$ and $P2_12_12_1$ space groups, suggesting that these

structural movements were evidence of the flexible nature of the PR rather than a consequence of crystal lattice packing effects. Similar flexibility was described for the 9X mutant structure (Munshi et al. 2000). Other flexible areas in the protease have been suggested, such as the residues 15-21, the flap region and both termini (Rose et al. 1998; Swairjo et al. 1998; Ishima et al. 1999; Zoete et al. 2002). The flexibility of the PR structure may play an important role in its function in viral replication (Piana et al. 2002a; Piana et al. 2002b). Analysis of the non-active-site mutations, for example, mutations in the flap or dimer interface, has shown that altering these flexible regions of HIV-1 PR can influence the binding of inhibitors or substrates (Olsen et al. 1999; Liu et al. 2005).

The sub-atomic resolution structure of PR_{V82A}/SQV has provided greater structural detail and more accurate atomic positions to understand the subtle structural changes induced by one single mutation. The results presented here will help better understand the molecular mechanisms of drug resistance and will be useful in developing more potent PR inhibitors.

3.4 Object-oriented protein structure database integrated with analysis functions

About 240 HIV-1 protease structures are now available in PDB. In our lab, 50 or more high-resolution crystal structures have been finished and more are still in progress. It will be beneficial to design a powerful program to help organize these data as well as automate the detailed structural analysis. Thus, an object-oriented HIV-1 protease structure database was designed and tested. The 50 structures solved by our lab are the primary data source for the database, although there is no limitation to cover other data. The whole design is aimed to provide variant basic functions of structural analysis with a powerful and flexible computing architecture.

Data format and syntax

HIV-1 protease structure data can be fully expressed by PDB format (PDB 2002). Beside the general data available in PDB file, we extracted a few important items of information to store in a relational table and designed a search engine based on the combination:

a. Category: Each structure file will contain one protein component. It will be a particular kind of protein, for example, HIV-1 protease.

b. Protein: For each kind of protein, it can be either wild type or different mutant.

c. Inhibitor: For this test set of HIV-1 protease structures, the inhibitor is a substrate analog peptide with 7-11 amino acids or a synthesized small compound. Currently, there are 16 different inhibitors.

d. Space group: As some protein/inhibitor complexes can crystallize in different space groups, this information is added to identify them.

e. PDB ID: Each protein structure deposited to the Protein Data bank is assigned a unique four-digit ID code. The ID is unique so it can be used as a primary key to search the database. However, the ID code itself does not provide any information about the kind of protein or inhibitor involved and some structures which have not gone through this stage will not have an ID.

f. Year: Similar to space group, different research groups may have solved the structure for the same complex, or later on, a better resolution structure may be solved and updated for the same complex. The year was added to help distinguish them.

Each PDB file is saved in database with above information in a row (called a record). In order to make the design concise, each record for one structure was named by the defined naming syntax. Each available property described above was presented by one unique key and they are: C for category, P for protein, I for inhibitor, K for PDB ID, S for space group and Y for the year of publication. The value for each key is enclosed by a parenthesis and directly follows the key. A double mutant is indicated by two mutations separated by underscore. For instance, a sample string of one record is given as C(HIVPR)-P(D25N_V82A)-I(SQV)-K(1MTB). By the name, it tells that this is a HIV-1 protease double mutant D25N and V82A in complex with inhibitor saquinavir, which is coded as 1MTB in PDB.

Data searching, synchronization, storage and view

The search engine is used to narrow down the targets to be analyzed. Combined search criteria are provided to filter out needed files. If no criterion is given, all files will be selected. Once potential targets are defined, the system will automatically synchronize the target data between backend database and the client. There will be a folder in client program package which contains all the PDB files synchronized from the database. The PDB files are named by the syntax described above according to the database information. For the convenience of analysis, each PDB file will be parsed and represented by a PDB object, and all the PDB objects will be saved into a vector for future use. The PDB object reflects the construction of PDB file and was stored as a tree structure, as shown in Figure 3.4.1. The PDB, chain, residue and atom are all defined objects. The upper layer object has at least one lower layer object, and in most cases, it contains a series of lower layer objects. In other word, upper layer object can be considered as a container for lower layer objects. Therefore, they embody a parent-child relationship. The Model-Control-View (MCV) strategy was adapted to display the HIV-1 protease structure data in this program. All the data are saved in a vector of PDB objects.

After all the target files been stored in local computer and parsed as objects, they are arranged and displayed also as a tree on user's interface. Generally, each structural data runs across 6 levels or less in a tree and the last 4 level of nodes/leaves correspond to objects PDB, chain, residue and atom. The root lists all available kinds of view mode which can be used to display the tree and the arrangement of the second layer is determined once a view mode is chosen. Three view modes are set up in the system; and

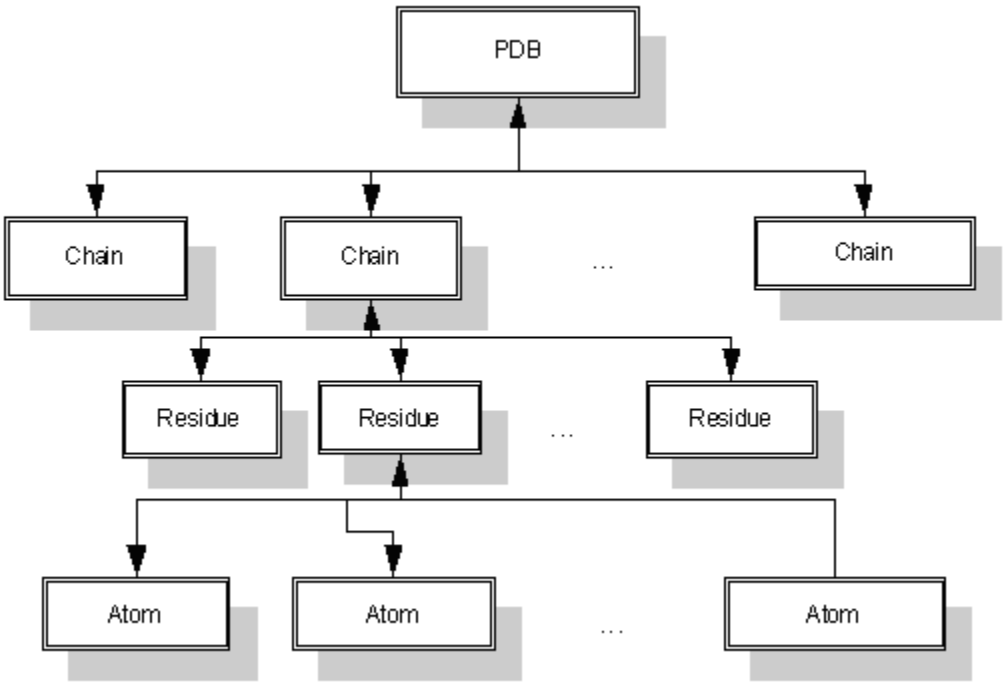


Figure 3.4.1 Hierarchical object composition.

they are View by Category, View by Proteins and View by Inhibitors. The mode of View by Category will divide all target structure files into different kinds of protein. This function is not fully developed and not applied now since only HIV-1 protease structures are stored. The mode of View by Proteins will classify all structures according to HIV-1 protease/mutant type, while the mode of View by Inhibitors will display file names by inhibitor type.

Analysis functions integrated within database

1. Data display and manipulation

Each PDB object will appear in the third layer of display tree with a link to text or graphic display choice. User can select to open the whole PDB text file to gather the detailed information needed, such as the structural factors, or open it with Rasmol (Sayle et al. 1995) to view the 3D structure and use the command lines for further manipulation.

2. Distance calculation

a) Calculation among different levels of objects

As mentioned in the Methods, the interactions between protease and inhibitor are generally characterized by the distances between atoms. Two atoms with separation of about 4 Å are considered to form good van der Waals interactions, while those of 2.6 to 3.2 Å in distance and between a hydrogen bond donor atom and a hydrogen bond acceptor atom are hydrogen bonds. The designed program employs the domain knowledge and applies the polymorphism to the calculation. Regularly only two atom

objects can be taken into the distance calculation. To make this function more efficient, different layers of objects are allowed to be selected, such as a residue and a few atoms. Once an object is selected, the program will automatically go to the lower level until the leaf of atom object and take all the atoms within that object. Distances will be measured between every pair of atoms out of all the atoms within two chosen objects. As shown in Table 3.4.1, the system can support 9 kinds of combination of objects. The calculated results will be filtered before being displayed on the screen. The filter is set up by two numbers: minimum and maximum distances which are specified by the user. Afterwards, the user can choose to store the results in excel format with user-defined file name for future use.

b) Group calculation

As we are studying the subtle changes induced by different mutant or inhibitor, sometimes it will be interesting to compare the interaction between the same two atoms from different complexes, for example, to examine the geometry of two catalytic Asps 25/25'.

Since all the HIV-1 protease structure files from our lab have the same kind of construction, it is possible to simplify this process. A group calculation is designed to allow user to input two atom objects within one PDB first, and then put a collection of selected PDBs that the user is interested in into the group. The system will retrieve the types and structure positions of two input atoms and apply the calculation to the structural equivalent atoms for all input PDBs. Again, the results can be saved as an excel file.

Table 3.4.1 Calculation among different levels of objects.

	Object1	Object2	Definition
1	Atom	Atom	
2	Atom	Residue	
3	Atom	Chain	
4	Atom	PDB	
5	Residue	Residue	
6	Residue	Chain	
7	Residue	PDB	
8	Chain	Chain	
9	Chain	PDB	

An example of distances between OD1 of Asp25 and OD1 of Asp25' in selected complexes is shown in Table 3.4.2.

3. Structure superimposing and visualization

Two structures can be compared by superimposing their main chain atoms using program FUD. The system provides the superimpose function by calling FUD program. A maximum of 8 HIV-1 protease structures can be selected to superimpose on one same reference file at the same time. For each pair of calculation, one statistics file and one superimposed PDB file will be generated and stored. For comparison convenience, these results can be visualized in one graph generated by result files. In this graph, the RMS deviations on Cα will be plotted against each residue. The global RMS deviation for each pair of superimposed structures will be shown at the top corner of the graph (Figure 3.4.2). User can also choose to display individual statistics file in text.

Application and future approaches

This database has been applied to the analysis of UIC-94017 complexes and saquinavir complexes discussed in 3.2 and 3.3. It provides an efficient tool to superimpose and calculate interactions, as well as to present more complete and quantified data to combine with the 3D structure visual program for analysis of detailed structural differences. For future development, more analytical tools can easily be incorporated into the database, such as a 3D structure visual program O or a more general superimposition program. In that case, the superimposing will no longer be limited to the

Table 3.4.2 Distances between OD1 of Asp25 and OD1 of Asp25' in selected complexes.

ATOM1	ATOM2	DISTANCE
P(I84V)-I(SQV)-S(P21212).ent: A:ASP:25:OD1:1.0	P(I84V)-I(SQV)-S(P21212).ent: B:ASP:25:OD1:1.0	2.66
P(I84V)-I(SQV)-S(P212121).ent: A:ASP:25:OD1:1.0	P(I84V)-I(SQV)-S(P212121).ent: B:ASP:25:OD1:1.0	2.85
P(V82A)-I(SQV)-S(P21212).ent: A:ASP:25:OD1:1.0	P(V82A)-I(SQV)-S(P21212).ent: B:ASP:25:OD1:1.0	2.82
P(V82A)-I(SQV)-S(P212121).ent: A:ASP:25:OD1:0.55	P(V82A)-I(SQV)-S(P212121).ent: B:ASP:25:OD1:0.55	2.78
P(V82A)-I(SQV)-S(P212121).ent: A:ASP:25:OD1:0.45	P(V82A)-I(SQV)-S(P212121).ent: B:ASP:25:OD1:0.45	2.86
P(WT)-I(SQV)-S(P212121).ent: A:ASP:25:OD1:1.0	P(WT)-I(SQV)-S(P212121).ent: B:ASP:25:OD1:1.0	2.91
P(I50V)-I(SQV).ent: A:ASP:25:OD1:1.0	P(I50V)-I(SQV).ent: B:ASP:25:OD1:1.0	2.87
P(L90M)-I(IDV)-K(1SDU).pdb: A:ASP:25:OD1:1.0	P(L90M)-I(IDV)-K(1SDU).pdb: B:ASP:25:OD1:1.0	2.79
P(V82A)-I(IDV)-K(1SDV).pdb: A:ASP:25:OD1:1.0	P(V82A)-I(IDV)-K(1SDV).pdb: B:ASP:25:OD1:1.0	2.67
P(WT)-I(IDV)-K(1SDT).pdb: A:ASP:25:OD1:1.0	P(WT)-I(IDV)-K(1SDT).pdb: B:ASP:25:OD1:1.0	2.64
P(G86A)-I(TMC114).pdb: A:ASP:25:OD1:1.0	P(G86A)-I(TMC114).pdb: B:ASP:25:OD1:1.0	2.49
P(G86S)-I(TMC114).pdb: A:ASP:25:OD1:1.0	P(G86S)-I(TMC114).pdb: B:ASP:25:OD1:1.0	2.58
P(D30N)-I(TMC114).pdb: A:ASP:25:OD1:1.0	P(D30N)-I(TMC114).pdb: B:ASP:25:OD1:1.0	2.64
P(I50V)-I(TMC114).pdb: A:ASP:25:OD1:1.0	P(I50V)-I(TMC114).pdb: B:ASP:25:OD1:1.0	2.79
P(I84V)-I(TMC114)-K(1S6S).pdb: A:ASP:25:OD1:1.0	P(I84V)-I(TMC114)-K(1S6S).pdb: B:ASP:25:OD1:1.0	2.59
P(V82A)-I(TMC114)-K(1S65).pdb: A:ASP:25:OD1:1.0	P(V82A)-I(TMC114)-K(1S65).pdb:B:ASP:25:OD1:1.0	2.87
P(WT)-I(TMC114)-K(1S6G).pdb: A:ASP:25:OD1:1.0	P(WT)-I(TMC114)-K(1S6G).pdb: B:ASP:25:OD1:1.0	2.85

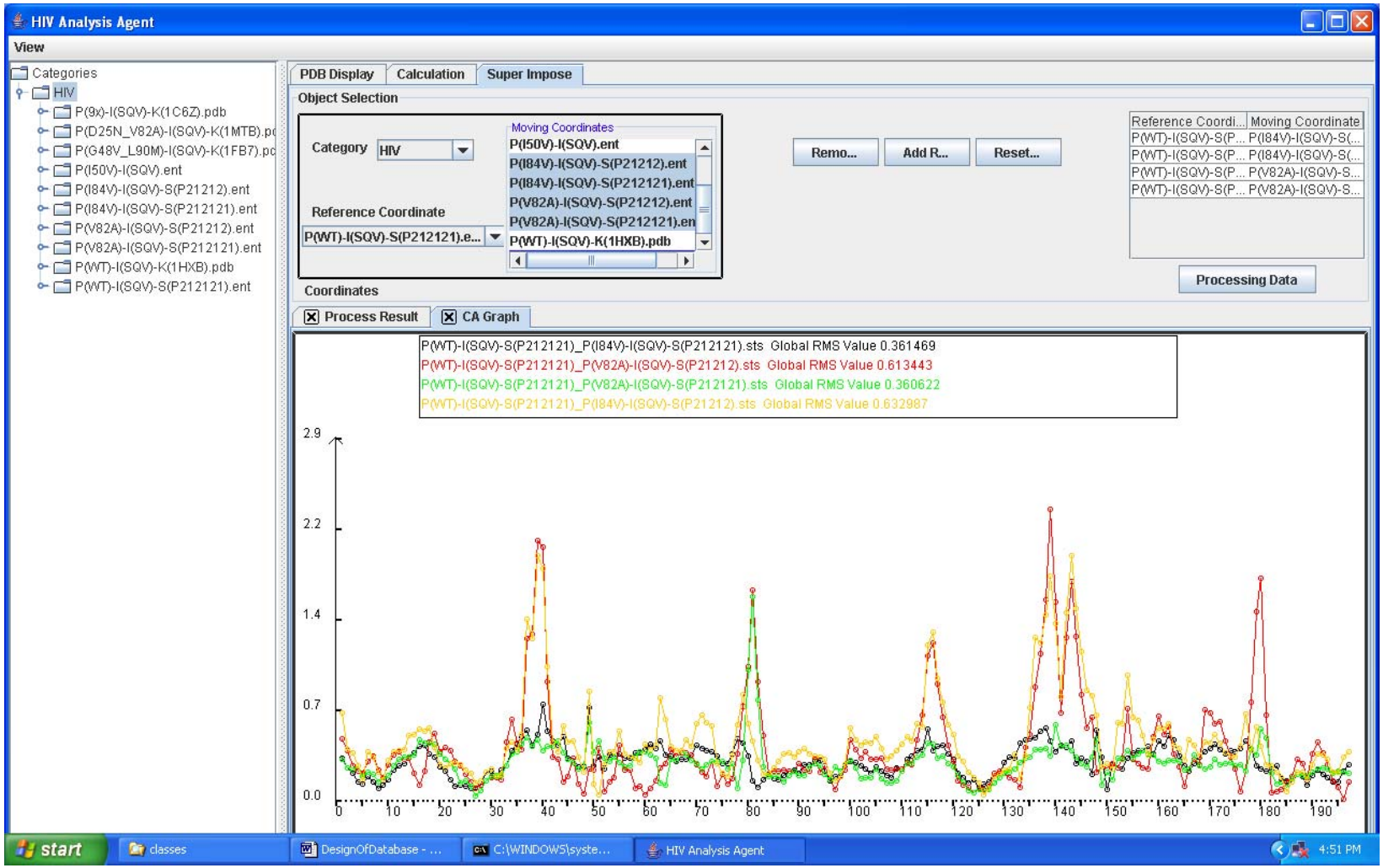


Figure 3.4.2 An example superimposing result for saquinavir complexes.

same kind/length of protein but based more on the structural similarity. A summary table of certain information may be generated for selected PDBs. Furthermore, it can be applied to proteins other than HIV-1 protease, for example, a large number of caspase structures are being studied in our lab.

4. General summary

HIV-1 PR is the most widely studied enzyme in the history of protein crystallography. The valuable structural information was used to guide the development of protease inhibitors; and indeed, these efforts have illustrated the enormous potential and utility of the rational structure-based drug design strategy. However, drug resistance is still a big obstacle in treating AIDS and its mechanism is not fully understood. To better understand the molecular basis for the altered enzymatic properties and the resistant phenotype and design more effective PIs to combat drug resistance, crystallographic and kinetic studies were applied to three series of complexes:

I. wild-type HIV-1 protease (PR) and drug-resistant mutants, PR_{V82A}, and PR_{I84V}, in complex with substrate analogs of five natural cleavage sites (CA-p2, p2-NC, $p6^{pol}$-PR, p1-p6 and NC-p1);
II. PR, PR_{V82A}, and PR_{I84V} with current drug saquinavir;
III. PR, PR_{V82A}, PR_{I84V}, PR_{D30N} and PR_{I50V}, with latest antiviral inhibitor design UIC-94017 (TMC-114).

The drug-resistant mutations V82A, I84V, D30N and I50V are all involved in forming the substrate-binding site. A total of eighteen crystal structures have been refined at resolutions of 0.97-1.60 Å. Observed structural changes were mostly consistent with the relative inhibition data.

Comparison of different inhibitors

For the first series, eight crystal structures of PR, PR_{V82A} and PR_{I84V} were determined with reduced peptide analogs that represent the CA-p2, p2-NC, $p6^{pol}$-PR, p1-p6 and NC-p1 cleavage sites to provide details of the PR interaction with reaction intermediates and better understanding of the substrate specificity. Substrate analogs bound to PR through amino acid positions P4 to P4' (P5 to P5' for longer analogs $p6^{pol}$-PR and p1-p6). Differences in the PR-analog interactions depended on the peptide sequence and were consistent with the relative inhibition. Analog $p6^{pol}$-PR formed more hydrogen bonds of P2Asn with PR and fewer van der Waals contacts at P1'Pro compared to those formed by CA-p2 or p2-NC in PR complexes. The P3Gly in p1-p6 provided fewer van der Waals contacts and hydrogen bonds at P2/P3 and more water-mediated interactions. The structures suggested that the binding affinity for mutants was modulated by the conformational flexibility of P1 and P1' side chains in the substrate analogs.

Unlike the longer substrate analogs which have more extended interactions within PR subsites S4 to S4', the clinical inhibitors maximize the contacts within S2 to S2'. The UIC-94017 is a potential next generation drug in clinical trial while saquinavir, indinavir and amprenavir have been used clinically for many years. All four inhibitors (UIC-94017, saquinavir, indinavir and amprenavir) have the same central hydroxyl group which formed strong hydrogen bonds with all four carboxylate oxygens of catalytic Asp25 and Asp25'. Besides these four contacts, seven direct hydrogen bonds were established between PR and UIC-94017, whilst three direct hydrogen bonds and three water-mediated interactions (except the one water between the PR flaps and the inhibitor)

for PR and indinavir. Three direct hydrogen bonds and one water-mediated interaction were observed between PR and only one end of saquinavir, and nevertheless, only two direct hydrogen bonds were observed for PR and amprenavir. The comparison showed that although UIC-94017 was smaller than the substrate analogs, it more closely mimics many of the natural substrates in the interactions with the PR main chain atoms than observed for the other clinical inhibitors, especially the additional polar interactions with the main chain atoms of Asp 29 and Asp 30. These interactions resembled those of the P2' Gln or Glu side chain of peptide analogs and have been proposed to be critical for the potency of this compound against multiple PI resistant HIV isolates. On the other hand, hydrophobic interactions were the major force for saquinavir binding. 10% more van der Waals interactions were observed for PR/SQV complex than in PR/UIC-94017. In brief, the design of early protease inhibitors, such as saquinavir, and the latest compound, UIC-94017 has different approaches. The earlier inhibitors are more similar to nature substrates. They mimic the hydrophobic side chains of substrate and tend to maximize the gain in desolvation entropy with a small loss in conformational entropy due to the burial of these big hydrophobic groups when they binds to protease. The latter approach of producing an enthalpically favorable process with less dependence on the entropy by introducing polar groups at inhibitor positions to form strong hydrogen bonds with the main chain of protease may improve the specificity, be less susceptible by the mutations and provide better water solubility. Ideally, the inhibitor design should combine both approaches, balancing the enthalpically favorable polar interactions and desalvation

entropy induced by hydrophobic groups, in order to provide tighter binding and more effective inhibitors.

The comparison of structures with saquinavir to those of UIC-94017 also showed an adjustment of the 80s loop (residues 79-82), which helped protease to accommodate the large groups of saquinavir. Analysis suggested that this 80s loop was intrinsically flexible and its conformation depended more on the nature of the inhibitor rather than the mutations in this loop.

Comparison of different mutant

Structural and kinetic studies have suggested that mutation V82A, which is common in drug resistant clinical isolates, had least effect on the inhibition of substrate analogs, saquinavir and UIC-94017. The reason is that PR_{V82A} was able to compensate for the loss of interactions with inhibitor caused by mutation by the shifts of the backbone of Ala82/82'. Furthermore, the complexes of PR_{V82A}/substrate analogs showed smaller shifts relative to PR, but more movement of the peptide analogs, compared to complexes with clinical inhibitors. Another common drug-resistant mutation I84V had only the minor effect on the inhibition of substrate analogs, saquinavir and UIC-94017. There were less movement observed for PR_{I84V} structures except for the saquinavir complex (0.3 Å for the alpha carbon of Val84), so PR_{I84V} simply lost favorable van der Waals interactions with inhibitor compared to PR. The structural comparison of these two mutant complexes suggested that substrate analogs have more flexibility than the drugs to accommodate the structural changes caused by mutation, which may explain how HIV

can develop drug resistance while retaining the ability of PR to hydrolyze natural substrates. Moreover, in saquinavir structures, the mutation from bigger to smaller amino acids allowed more space to accommodate the larger group at P1' of saquinavir rather than alter the protease-inhibitor interactions as seen in other inhibitors complexes.

In PR_{I50V}, the UIC-94017 lost favorable hydrophobic interactions with the side chain of Val50, while partial compensating conformational change had been observed on Val50' main chain. Similarly, in PR_{D30N} the inhibitor had a water-mediated interaction with the side chain of Asn30/30' rather than the direct interaction observed in PR. In addition, the kinetic measurements had shown that PR_{D30N} was only 10% as active as the wild type protease on the fluorescent substrate, while PR_{I50V} was 40% as active; and the potency of UIC-94017 was reduced by about 30 and 9-fold, respectively, for PR_{D30N} and PR_{I50V}. Both structural and kinetic studies implied that the mutations PR_{I50V} and especially PR_{D30N} might be selected by the virus for resistance to UIC-94017. Consequently, UIC-94017 is not expected to be a good salvage therapy candidate for the patients who have failed nelfinavir treatment due to the existence of the viral strains that already contain the D30N mutation in the protease.

In summary, the development of drug-resistance to PR inhibitors is complicated and can arise by more than one mechanism. No direct relationship was observed between relative catalytic activity, inhibition, and structural stability of the different PR mutants. Changes in any one of these parameters may cause drug resistance. Selected compensatory mutations may be developed in virus after the initial mutation which lowers the affinity for the inhibitor or the catalytic activity or dimer stability, to restore

the optimal polyprotein processing and replication capacity in the presence of PR inhibitor.

Advantage of atomic resolution data

With the high resolution crystal structures, the atom positions were more accurately defined. Three shells of water molecules, solvent ions (such as sodium ion) and molecules (such as glycerol and DMSO) were resolved. The 0.97 Å resolution data for the PR_{V82A}/SQV complex had shown more residues with alternative conformations than seen before, especially the two conformations were observed for the active site residues 24-26 and 24'-26' and they were correlated with the alternate conformations of saquinavir. The hydroxyl group on the saquinavir had similar arrangement and distances to catalytic Asp25/25' in both conformations. These precise positions of the active site residues may have not been correctly modeled before in other structures with alternate conformations of inhibitor. Thus, the sub-atomic resolution data has helped to describe the detailed geometry of the enzyme active site.

Furthermore, the 1.10 Å crystal structure of PR_{V82A} with one orientation of UIC-94017 showed an unusual distribution of electron density for the catalytic aspartate residues, which may be related to the reaction mechanism. A streak of positive density was observed between the hydroxyl group of the inhibitor and the inner and closest OD1 of Asp 25' (2.5 Å separation of OH and OD1), which might indicate a hydrogen atom. However, the observation can hardly be proved only by the extra electron density. Further analysis of the data by charge density analysis or quantum calculations will be

necessary to understand the distribution of electrons at the active site. Yet, the difference densities suggested a more complex mechanism than protonation of a carboxylate oxygen.

Advantage of the database incorporated with analysis programs

For proteins that serve as drug-design targets, it is important to study protease-inhibitor interactions accurately, completely and in as many complexes as possible. Careful analysis of the wild-type as well as drug-resistant mutants of HIV PR may also help in creating new drugs that would overcome the problem of resistance. The HIV-1 protease structure database offers an easy way to organize data and provides specialized tools to automate the superimposing and distance calculation. The presence of such tools and the ability for future development will help in better understanding the structural phenomena related to HIV PR.

Other designed inhibitors under development

With the goal to design more effective protease inhibitors to combat drug resistance, a series of similar compounds based on the structure of UIC-94017 had been synthesized by Dr. Arun Ghosh's lab and used for kinetic and crystallographic studies (as listed in Appendix I) with HIV-1 PR. The data for UIC-98065 and UIC-00072 has been completed. The detailed structural information would help evaluate the design goal and provide valuable reference for further drug development.

Overall, the large amount of kinetic and crystallographic studies has increased the knowledge of enzyme function, structure, and catalytic mechanism and led to a better understanding of how these drug-resistance mutations exert their effects at a molecular level. These insights are valuable for the design of new drugs and therapeutic strategies to combat drug resistance to AIDS.

References

AVERT. (2005). from http://www.avert.org/worldstats.htm.

Baca, M. and S. B. Kent (1993). "Catalytic contribution of flap-substrate hydrogen bonds in "HIV-1 protease" explored by chemical synthesis." *Proc Natl Acad Sci U S A* **90**(24): 11638-42.

Bagossi, P., J. Kadas, G. Miklossy, P. Boross, I. T. Weber and J. Tozser (2004). "Development of a Microtiter Plate Fluorescent Assay for Inhibition Studies on the HTLV-1 and HIV-1 Proteinases." *J. Virol. Met.* **119**: 87-93.

Baldwin, E. T., T. N. Bhat, B. Liu, N. Pattabiraman and J. W. Erickson (1995). "Structural basis of drug resistance for the V82A mutant of HIV-1 proteinase." *Nat. Struct. Biol.* **2**(3): 244-249.

Bally, F., R. Martinez, S. Peters, P. Sudre and A. Telenti (2000). "Polymorphism of HIV type 1 gag p7/p1 and p1/p6 cleavage sites: clinical significance and implications for resistance to protease inhibitors." *AIDS Res. Hum. Retroviruses.* **16**(13): 1209-1213.

Barre-Sinoussi, F., J. C. Chermann, F. Rey, M. T. Nugeyre, S. Chamaret, J. Gruest, C. Dauguet, C. Axler-Blin, F. Vezinet-Brun, C. Rouzioux, W. Rozenbaum and L. Montagnier (1983). "Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS)." *Science* **220**(4599): 868-71.

Barrie, K. A., E. E. Perez, S. L. Lamers, W. G. Farmerie, B. M. Dunn, J. W. Sleasman and M. M. Goodenow (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**(2): 407-416.

Berman, H. M., J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, N. ShindyalovI and P. E. Bourne (2000). "The Protein Data Bank " *Nucleic Acids Research* **28**(1): 235-242.

Boucher, C. (1996). "Rational approaches to resistance: using saquinavir." *Aids* **10 Suppl 1**: S15-9.

Cameron, C. E., B. Grinde, P. Jacques, J. Jentoft, J. Leis, A. Wlodawer and I. T. Weber (1993). "Comparison of the substrate-binding pockets of the Rous sarcoma virus and human immunodeficiency virus type 1 proteases." *J Biol Chem* **268**(16): 11711-20.

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

Chun, T. W., L. Stuyver, S. B. Mizell, L. A. Ehler, J. A. Mican, M. Baseler, A. L. Lloyd, M. A. Nowak and A. S. Fauci (1997). "Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy." *Proc Natl Acad Sci U S A* **94**(24): 13193-7.

Clapham, P. R. and R. A. Weiss (1997). "Immunodeficiency viruses. Spoilt for choice of co-receptors." *Nature* **388**(6639): 230-1.

Clemente, J. C., R. E. Moose, R. Hemrajani, L. R. Whitford, L. Govindasamy, R. Reutzel, R. McKenna, M. Agbandje-McKenna, M. M. Goodenow and B. M. Dunn (2004). "Comparing the accumulation of active- and nonactive-site mutations in the HIV-1 protease." *Biochemistry* **43**(38): 12141-51.

Coffin, J. M. (1995). "HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy." *Science* **267**(5197): 483-9.

Colonno, R. J., A. Thiry, K. Limoli and N. Parkin (2003). "Activities of atazanavir (BMS-232632) against a large panel of human immunodeficiency virus type 1 clinical isolates resistant to one or more approved protease inhibitors." *Antimicrob Agents Chemother* **47**(4): 1324-33.

Condra, J. H., D. J. Holder, W. A. Schleif, O. M. Blahy, R. M. Danovich, L. J. Gabryelski, D. J. Graham, D. Laird, J. C. Quintero, A. Rhodes, H. L. Robbins, E. Roth, M. Shivaprakash, T. Yang, J. A. Chodakewitz, P. J. Deutsch, R. Y. Leavitt, F. E. Massari, J. W. Mellors, K. E. Squires, R. T. Steigbigel, H. Teppler and E. A. Emini (1996). "Genetic correlates of in vivo viral resistance to indinavir, a human immunodeficiency virus type 1 protease inhibitor." *J Virol* **70**(12): 8270-6.

Craig, J. C., I. B. Duncan, D. Hockley, C. Grief, N. A. Roberts and J. S. Mills (1991). "Antiviral properties of Ro 31-8959, an inhibitor of human immunodeficiency virus (HIV) proteinase." *Antiviral Res* **16**(4): 295-305.

Croteau, G., L. Doyon, D. Thibeault, M. G., L. Pilote and D. Lamarre (1997). "Impaired fitness of human immunodeficiency virus type 1 variants with high-level resistance to protease inhibitors." *J. Virol.* **71**(2): 1089-1096.

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

Debouck, C. (1992). "The HIV-1 protease as a therapeutic target for AIDS." *AIDS Res Hum Retroviruses* **8**(2): 153-64.

Deminie, C. A., C. M. Bechtold, D. Stock, M. Alam, F. Djang, A. H. Balch, T. C. Chou, M. Prichard, R. J. Colonno and P. F. Lin (1996). "Evaluation of reverse transcriptase and protease inhibitors in two-drug combinations against human immunodeficiency virus replication." *Antimicrob Agents Chemother* **40**(6): 1346-51.

Doranz, B. J., J. Rucker, Y. Yi, R. J. Smyth, M. Samson, S. C. Peiper, M. Parmentier, R. G. Collman and R. W. Doms (1996). "A dual-tropic primary HIV-1 isolate that uses fusin and the beta-chemokine receptors CKR-5, CKR-3, and CKR-2b as fusion cofactors." *Cell* **85**(7): 1149-58.

Erickson, J. W., S. V. Gulnik and M. Markowitz (1999). "Protease inhibitors: resistance, cross-resistance, fitness and the choice of initial and salvage therapies." *AIDS Res Hum Retroviruses.* **13**(Suppl A): S189-S204.

Esnouf, R. M. (1997). "An extensively modified version of MolScript that includes greatly enhanced coloring capabilities." *J Mol Graph Model* **15**(2): 132-4, 112-3.

Esnouf, R. M. (1999). "Further additions to MolScript version 1.4, including reading and contouring of electron-density maps." *Acta Crystallogr D Biol Crystallogr* **55 (Pt 4)**: 938-40.

FDA. (2005). from http://www.fda.gov/oashi/aids/virals.html.

Feher, A., I. T. Weber, P. Bagossi, P. Boross, B. Mahalingam, J. M. Louis, T. D. Copeland, I. Y. Torshin, R. W. Harrison and J. Tozser (2002). "Effect of sequence polymorphism and drug resistance on two HIV-1 Gag processing sites." *Eur. J. Biochem.* **269**(16): 4114-4120.

Feng, Y., C. C. Broder, P. E. Kennedy and E. A. Berger (1996). "HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor." *Science* **272**(5263): 872-7.

Finzi, D., M. Hermankova, T. Pierson, L. M. Carruth, C. Buck, R. E. Chaisson, T. C. Quinn, K. Chadwick, J. Margolick, R. Brookmeyer, J. Gallant, M. Markowitz, D. D. Ho, D. D. Richman and R. F. Siliciano (1997). "Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy." *Science* **278**(5341): 1295-300.

Fouchier, R. A., B. E. Meyer, J. H. Simon, U. Fischer and M. H. Malim (1997). "HIV-1 infection of non-dividing cells: evidence that the amino-terminal basic region of

the viral matrix protein is important for Gag processing but not for post-entry nuclear import." *Embo J* **16**(15): 4531-9.

Freed, E. O. (2001). "HIV-1 replication." *Somat Cell Mol Genet* **26**(1-6): 13-33.

Freed, E. O. and M. A. Martin (1995). "Virion incorporation of envelope glycoproteins with long but not short cytoplasmic tails is blocked by specific, single amino acid substitutions in the human immunodeficiency virus type 1 matrix." *J Virol* **69**(3): 1984-9.

Gallo, R. C., S. Z. Salahuddin, M. Popovic, G. M. Shearer, M. Kaplan, B. F. Haynes, T. J. Palker, R. Redfield, J. Oleske, B. Safai and et al. (1984). "Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS." *Science* **224**(4648): 500-3.

Ghosh, A. K. and Y. Chen (1995). "Synthesis and optical resolution of high affinity P2-ligands for HIV-1 protease inhibitors." *Tetrahedron Lett.* **3**: 505-508.

Ghosh, A. K. and S. Fidanze (1998a). "Transition-State mimetics for HIV protease inhibitors: stereocontrolled synthesis of hydroxyethylene and hydroxyethylamine isosteres by ester derived titanium enolate syn- and anti-aldol reactions." *J. Org. Chem.* **63**(18): 6146-6152.

Ghosh, A. K., J. F. Kincaid, W. Cho, D. E. Walters, K. Krishnan, K. A. Hussain, Y. Koo, Cho, H., C. Rudall, L. Holland and J. Buthod (1998b). "Potent HIV protease inhibitors incorporating high sffinity P2-ligands and (R)-hydroxyethylaminosufonamide isostere." *Bioorg. Med. Chem. Lett.* **8**(6): 979-982.

Ghosh, A. K., J. F. Kincaid, D. E. Walters, Y. Chen, N. C. Chaudhuri, W. J. Thompson, C. Culberson, P. M. D. Fitzgerald, S. P. McKee, P. M. Munson, M. G. Axel and J. R. Huff (1996). "Nonpeptidal P2-ligands for HIV protease inhibitors: Structure-based design, synthesis and biological evaluations." *J. Med. Chem.* **39**: 3278-3294.

Goff, S. P. (1990). "Retroviral reverse transcriptase: synthesis, structure, and function." *J Acquir Immune Defic Syndr* **3**(8): 817-31.

Grzesiek, S., A. Bax, L. K. Nicholson, T. Yamazaki, P. T. Wingfield, S. J. Stahl, C. J. Eyermann, D. A. Torchia, C. N. Hodge, P. Y. S. Lam, P. K. Jadhav and C.-H. Chang (1994). "NMR Evidence for the Displacement of a Conserved Interior Water Molecule by a Non-peptide Cyclic Urea Based HIV Protease Inhibitor." *J. Am. Chem. Soc.* **116**: 1581-1582.

Gulick, R. M., J. W. Mellors, D. Havlir, J. J. Eron, C. Gonzalez, D. McMahon, D. D. Richman, F. T. Valentine, L. Jonas, A. Meibohm, E. A. Emini and J. A. Chodakewitz (1997). "Treatment with indinavir, zidovudine, and lamivudine in adults with human immunodeficiency virus infection and prior antiretroviral therapy." *N Engl J Med* **337**(11): 734-9.

Gulnik, S. V., L. I. Suvorov, B. Liu, B. Yu, B. Anderson, H. Mitsuya and J. W. Erickson (1995). "Kinetic characterization and cross-resistance patterns of HIV-1 protease mutants selected under drug pressure." *Biochemistry* **34**(29): 9282-7.

Gustchina, A., J. Kervinen, D. J. Powell, A. Zdanov, J. Kay and A. Wlodawer (1996). "Structure of equine infectious anemia virus proteinase complexed with an inhibitor." *Protein Sci* **5**(8): 1453-65.

Gustchina, A., C. Sansom, M. Prevost, J. Richelle, S. Y. Wodak, A. Wlodawer and I. T. Weber (1994). "Energy calculations and analysis of HIV-1 protease-inhibitor crystal structures." *Protein Eng.* **7**(3): 309-317.

Hammond, J., C. Calef, B. Larder, S. Schinazi and J. W. Mellors. "Mutations in retroviral genes associated with drug resistance." from http://hiv-web.lanl.gov/content/hiv-db/COMPENDIUM/1999/6/resistance.pdf.

Harrison, R. R. and C. Koch (1999). "An analog VLSI implementation of a visual interneuron: enhanced sensory processing through biophysical modeling." *Int J Neural Syst* **9**(5): 391-5.

Harrison, R. W. (1993). "Stiffness and energy conservation in the molecular dynamics: an improved integrator." *J. Comp. Chem.* **14**: 1112-1122.

Harrison, R. W. and I. T. Weber (1994). "Molecular dynamics simulations of HIV-1 protease with peptide substrate." *Protein Eng* **7**(11): 1353-63.

Hayakawa, T., Y. Misumi, M. Kobayashi, Y. Yamamoto and Y. Fujisawa (1992). "Requirement of N- and C-terminal regions for enzymatic activity of human T-cell leukemia virus type I protease." *Eur J Biochem* **206**(3): 919-25.

Henderson, L. E., M. A. Bowers, R. C. Sowder, 2nd, S. A. Serabyn, D. G. Johnson, J. W. Bess, Jr., L. O. Arthur, D. K. Bryant and C. Fenselau (1992). "Gag proteins of the highly replicative MN strain of human immunodeficiency virus type 1: posttranslational modifications, proteolytic processings, and complete amino acid sequences." *J Virol* **66**(4): 1856-65.

Herger, B. E., V. L. Mariani, K. Dennison and S. B. Shuker (2004). "The 10 C-terminal residues of HTLV-I protease are not necessary for enzymatic activity." *Biochem Biophys Res Commun* **320**(4): 1306-8.

Hertogs, K., S. Bloor, S. D. Kemp, C. Van den Eynde, T. M. Alcorn, R. Pauwels, M. Van Houtte, S. Staszewski, V. Miller and B. A. Larder (2000). "Phenotypic and genotypic analysis of clinical HIV-1 isolates reveals extensive protease inhibitor cross-resistance: a survey of over 6000 samples." *AIDS.* **14**(9): 1203-1210.

Ho, D. D., A. U. Neumann, A. S. Perelson, W. Chen, J. M. Leonard and M. Markowitz (1995). "Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection." *Nature* **373**(6510): 123-6.

Hodge, C. N., P. E. Aldrich, L. T. Bacheler, C. H. Chang, C. J. Eyermann, S. Garber, M. Grubb, D. A. Jackson, P. K. Jadhav, B. Korant, P. Y. Lam, M. B. Maurin, J. L. Meek, M. J. Otto, M. M. Rayner, C. Reid, T. R. Sharpe, L. Shum, D. L. Winslow and S. Erickson-Viitanen (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**(4): 301-14.

Hong, L., X. C. Zhang, J. A. Hartsuck and J. Tang (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**(10): 1898-1904.

Hyland, L. J., T. A. J. Tomaszek and T. D. Meek (1991). "Human immunodeficiency virus-1 protease. 2. Use of pH rate studies and solvent kinetic isotope effects to elucidate details of chemical mechanism." *Biochemistry.* **30**(34): 8454-8463.

Ishima, R., D. I. Freedberg, Y. X. Wang, J. M. Louis and D. A. Torchia (1999). "Flap opening and dimer-interface flexibility in the free and inhibitor-bound HIV protease, and their implications for function." *Structure Fold Des* **7**(9): 1047-55.

Jacobsen, H., M. Hanggi, M. Ott, I. B. Duncan, S. Owen, M. Andreoni, S. Vella and J. Mous (1996). "In vivo resistance to a human immunodeficiency virus type 1 proteinase inhibitor: mutations, kinetics, and frequencies." *J Infect Dis* **173**(6): 1379-87.

Jacobsen, H., K. Yasargil, D. L. Winslow, J. C. Craig, A. Krohn, I. B. Duncan and J. Mous (1995). "Characterization of human immunodeficiency virus type 1 mutants with decreased sensitivity to proteinase inhibitor Ro 31-8959." *Virology* **206**(1): 527-34.

Jarvis, B. and D. Faulds (1998). "Nelfinavir. A review of its therapeutic efficacy in HIV infection." *Drugs* **56**(1): 147-67.

Ji, J. P. and L. A. Loeb (1992). "Fidelity of HIV-1 reverse transcriptase copying RNA in vitro." *Biochemistry* **31**(4): 954-8.

Johnson, V. A., F. Brun-Vezinet, B. Clotet, B. Conway, D. R. Kuritzkes, D. Pillay, J. M. Schapiro, A. Telenti and D. D. Richman (2005). "Update of the Drug Resistance Mutations in HIV-1: Fall 2005." *Top HIV Med* **13**(4): 125-31.

Jones, T. A., J. Y. Zou, S. W. Cowan and M. Kjeldgaard (1991). "Improved methods for building protein models in electron density maps and the location of errors in these models." *Acta Cryst.* **A47**: 110-119.

Katz, R. A. and A. M. Skalka (1994). "The retroviral enzymes." *Annu Rev Biochem* **63**: 133-73.

Kempf, D. J., J. D. Isaacson, M. S. King, S. C. Brun, Y. Xu, K. Real, B. M. Bernstein, A. J. Japour, E. Sun and R. A. Rode (2001). "Identification of genotypic changes in human immunodeficiency virus protease that correlate with reduced susceptibility to the protease inhibitor lopinavir among viral isolates from protease inhibitor-experienced patients." *J Virol* **75**(16): 7462-9.

Kim, E. E., C. T. Baker, M. D. Dwyer, M. A. Murcko, B. G. Rao, R. D. Tung and M. A. Navia (1995). "Crystal structure of HIV-1 protease in complex with VX-478, a potent and orally bioavailable inhibitor of the enzyme." *J. Am. Chem. Soc.* **117**(3): 1181-1182.

King, N. M., M. Prabu-Jeyabalan, E. A. Nalivaika, P. Wigerinck, M. P. de Bethune and C. A. Schiffer (2004). "Structural and thermodynamic basis for the binding of TMC114, a next-generation human immunodeficiency virus type 1 protease inhibitor." *J Virol* **78**(21): 12012-21.

Klabe, R. M., L. T. Bacheler, P. J. Ala, S. Erickson-Viitanen and J. L. Meek (1998). "Resistance to HIV protease inhibitors: a comparison of enzyme inhibition and antiviral potency." *Biochemistry* **37**(24): 8735-42.

Koh, Y., H. Nakata, K. Maeda, H. Ogata, G. Bilcer, T. Devasamudram, J. F. Kincaid, P. Boross, Y. F. Wang, Y. Tie, P. Volarath, L. Gaddis, R. W. Harrison, I. T. Weber, A. K. Ghosh and H. Mitsuya (2003). "A novel bis-tetrahydrofuranylurethane-containing nonpeptidic protease inhibitors (PI) UIC-94017 (TMC114) potent against multi-PI-resistant HIV in vitro." *Antimicrob. Agents Chemother.* **47**(10): 3123-3129.

Kohl, N. E., E. A. Emini, W. A. Schleif, L. J. Davis, J. C. Heimbach, R. A. Dixon, E. M. Scolnick and I. S. Sigal (1988). "Active human immunodeficiency virus protease is required for viral infectivity." *Proc. Natl. Acad. Sci.* **85**(13): 4686-4690.

Konvalinka, J., P. Strop, J. Velek, V. Cerna, V. Kostka, L. H. Phylip, A. D. Richards, B. M. Dunn and J. Kay (1990). "Sub-site preferences of the aspartic proteinase from the human immunodeficiency virus, HIV-1." *FEBS Lett* **268**(1): 35-8.

Kovalevsky, A. Y., Y. Tie, F. Liu, P. I. Boross, Y. F. Wang, S. Leshchenko, A. K. Ghosh, R. W. Harrison and I. T. Weber (2006). "Effectiveness of Nonpeptide Clinical Inhibitor TMC-114 on HIV-1 Protease with Highly Drug Resistant Mutations D30N, I50V, and L90M." *J. Med. Chem.* **Accepted.**

Kraulis, P. J. (1991). "MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures." *J. Appl. Cryst.* **24**: 946-950.

Krohn, A., S. Redshaw, J. C. Ritchie, B. J. Graves and M. H. Hatada (1991). "Novel binding mode of highly potent HIV-proteinase inhibitors incorporating the (R)-hydroxyethylamine isostere." *J Med Chem* **34**(11): 3340-2.

Lam, P. Y., P. K. Jadhav, C. J. Eyermann, C. N. Hodge, Y. Ru, L. T. Bacheler, J. L. Meek, M. J. Otto, M. M. Rayner, Y. N. Wong and et al. (1994). "Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors." *Science* **263**(5145): 380-4.

Lam, P. Y., Y. Ru, P. K. Jadhav, P. E. Aldrich, G. V. DeLucca, C. J. Eyermann, C. H. Chang, G. Emmett, E. R. Holler, W. F. Daneker, L. Li, P. N. Confalone, R. J. McHugh, Q. Han, R. Li, J. A. Markwalder, S. P. Seitz, T. R. Sharpe, L. T. Bacheler, M. M. Rayner, R. M. Klabe, L. Shum, D. L. Winslow, D. M. Kornhauser, C. N. Hodge and et al. (1996). "Cyclic HIV protease inhibitors: synthesis, conformational analysis, P2/P2' structure-activity relationship, and molecular recognition of cyclic ureas." *J Med Chem* **39**(18): 3514-25.

Levy, J. A., A. D. Hoffmann, S. M. Kramer, J. A. Landis, J. M. Shimabukuro and L. S. Oshiro (1984). "Recovery of AIDS-associated retroviruses from patients with AIDS or AIDS-related conditions and from clinically healthy individuals." *J. Infect. Dis.* **152**: 734-738.

Li, M., G. S. Laco, M. Jaskolski, J. Rozycki, J. Alexandratos, A. Wlodawer and A. Gustchina (2005). "Crystal structure of human T cell leukemia virus protease, a novel target for anticancer drug design." *Proc Natl Acad Sci U S A* **102**(51): 18332-7.

Liu, F., P. I. Boross, Y. F. Wang, J. Tozser, J. M. Louis, R. W. Harrison and I. T. Weber (2005). "Kinetic, stability, and structural changes in high-resolution crystal structures of HIV-1 protease with drug-resistant mutations L24I, I50V, and G73S." *J Mol Biol* **354**(4): 789-800.

Louis, J. M., N. T. Nashed, K. D. Parris, A. R. Kimmel and D. M. Jerina (1994). "Kinetics and mechanism of autoprocessing of human immunodeficiency virus type 1 protease from an analog of the Gag-Pol polyprotein." *Proc. Natl. Acad. Sci.* **91**(17): 7970-7974.

Louis, J. M., C. A. Smith, E. M. Wondrak, P. T. Mora and S. Oroszlan (1989). "Substitution mutations of the highly conserved arginine 87 of HIV-1 protease result in loss of proteolytic activity." *Biochem Biophys Res Commun* **164**(1): 30-8.

Louis, J. M., I. T. Weber, J. Tozser, G. M. Clore and A. M. Gronenborn (2000). "HIV-1 protease: maturation, enzyme specificity, and drug resistance." *Adv. Pharmacol.* **49**: 111-146.

Louis, J. M., E. M. Wondrak, T. D. Copeland, C. A. Smith, P. T. Mora and S. Oroszlan (1989). "Chemical synthesis and expression of the HIV-1 protease gene in E. coli." *Biochem. Biophys. Res. Commun.* **159**(1): 87-94.

Maguire, M., D. Shortino, A. Klein, W. Harris, V. Manohitharajah, M. Tisdale, R. Elston, J. Yeo, S. Randall, F. Xu, H. Parker, J. May and W. Snowden (2002). "Emergence of resistance to protease inhibitor amprenavir in human immunodeficiency virus type 1-infected patients: selection of four alternative viral protease genotypes and influence of viral susceptibility to coadministered reverse transcriptase nucleoside inhibitors." *Antimicrob Agents Chemother* **46**(3): 731-8.

Mahalingam, B., P. Boross, Y. F. Wang, J. M. Louis, C. C. Fischer, J. Tozser, R. W. Harrison and I. T. Weber (2002). "Combining mutations in HIV-1 protease to understand mechanisms of resistance." *Proteins: Structure, Function, and Genetics.* **48**(1): 107-116.

Mahalingam, B., J. M. Louis, J. Hung, R. W. Harrison and I. T. Weber (2001). "Structural implications of drug resistant mutants of HIV-1 protease: High resolution crystal structures of the mutant protease/substrate analogue complexes." *Proteins: Structure, Function, and Genetics.* **43**(4): 455-464.

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

Mahalingam, B., Y. F. Wang, P. I. Boross, J. Tozser, J. M. Louis, R. W. Harrison and I. T. Weber (2004). "Crystal structures of HIV protease V82A and L90M mutants reveal changes in the indinavir-binding site." *Eur. J. Biochem.* **271**(8): 1516-1524.

Markowitz, M., M. Conant, A. Hurley, R. Schluger, M. Duran, J. Peterkin, S. Chapman, A. Patick, A. Hendricks, G. J. Yuen, W. Hoskins, N. Clendeninn and D. D. Ho

(1998). "A preliminary evaluation of nelfinavir mesylate, an inhibitor of human immunodeficiency virus (HIV)-1 protease, to treat HIV infection." *J Infect Dis* **177**(6): 1533-40.

Matayoshi, E. D., G. T. Wang, G. A. Krafft and J. Erickson (1990). "Novel fluorogenic substrates for assaying retroviral proteases by resonance energy transfer." *Science* **247**(4945): 954-8.

Miller, M., M. Jaskolski, J. K. Rao, J. Leis and A. Wlodawer (1989). "Crystal structure of a retroviral protease proves relationship to aspartic protease family." *Nature* **337**(6207): 576-9.

Miller, M. D., C. M. Farnet and F. D. Bushman (1997). "Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition." *J Virol* **71**(7): 5382-90.

Molla, A., M. Korneyeva, Q. Gao, S. Vasavanonda, P. J. Schipper, H. M. Mo, M. Markowitz, T. Chernyavskiy, P. Niu, N. Lyons, A. Hsu, G. R. Granneman, D. D. Ho, C. A. Boucher, J. M. Leonard, D. W. Norbeck and D. J. Kempf (1996). "Ordered accumulation of mutations in HIV protease confers resistance to ritonavir." *Nat Med* **2**(7): 760-6.

Moore, J. P. (1997). "Coreceptors: implications for HIV pathogenesis and therapy." *Science* **276**(5309): 51-2.

Munshi, S., Z. Chen, Y. Yan, Y. Li, D. B. Olsen, H. B. Schock, B. B. Galvin, B. Dorsey and L. C. Kuo (2000). "An alternate binding site for the P1-P3 group of a class of potent HIV-1 protease inhibitors as a result of concerted structural change in the 80s loop of the protease." *Acta Crystallogr D Biol Crystallogr* **56 (Pt 4)**: 381-8.

Navaza, J. (1994). "AMoRe: An automated package for molecular replacement." *Acta Cryst.* **A50**: 157-163.

Nicholson, L. K., T. Yamazaki, D. A. Torchia, S. Grzesiek, A. Bax, S. J. Stahl, J. D. Kaufman, P. T. Wingfield, P. Y. Lam, P. K. Jadhav and et al. (1995). "Flexibility and function in HIV-1 protease." *Nat Struct Biol* **2**(4): 274-80.

Nie, Z., D. Bergeron, R. A. Subbramanian, X. J. Yao, F. Checroune, N. Rougeau and E. A. Cohen (1998). "The putative alpha helix 2 of human immunodeficiency virus type 1 Vpr contains a determinant which is responsible for the nuclear translocation of proviral DNA in growth-arrested cells." *J Virol* **72**(5): 4104-15.

Olsen, D. B., M. W. Stahlhut, C. A. Rutkowski, H. B. Schock, A. L. vanOlden and L. C. Kuo (1999). "Non-active site changes elicit broad-based cross-resistance of the HIV-1 protease to inhibitors." *J Biol Chem* **274**(34): 23699-701.

Oroszlan, S. and R. B. Luftig (1990). "Retroviral proteases." *Curr. Top. Microgiol. Immunol.* **157**: 153-185.

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

Palella, F. J., Jr., K. M. Delaney, A. C. Moorman, M. O. Loveless, J. Fuhrer, G. A. Satten, D. J. Aschman and S. D. Holmberg (1998). "Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators." *N Engl J Med* **338**(13): 853-60.

Partaledis, J. A., K. Yamaguchi, M. Tisdale, E. E. Blair, C. Falcione, B. Maschera, R. E. Myers, S. Pazhanisamy, O. Futer, A. B. Cullinan and et al. (1995). "In vitro selection and characterization of human immunodeficiency virus type 1 (HIV-1) isolates with reduced sensitivity to hydroxyethylamino sulfonamide inhibitors of HIV-1 aspartyl protease." *J Virol* **69**(9): 5228-35.

Patick, A. K., M. Duran, Y. Cao, D. Shugarts, M. R. Keller, E. Mazabel, M. Knowles, S. Chapman, D. R. Kuritzkes and M. Markowitz (1998). "Genotypic and phenotypic characterization of human immunodeficiency virus type 1 variants isolated from patients treated with the protease inhibitor nelfinavir." *Antimicrob Agents Chemother* **42**(10): 2637-44.

Pazhanisamy, S., C. M. Stuver, A. B. Cullinan, N. Margolin, B. G. Rao and D. J. Livingston (1996). "Kinetic characterization of human immunodeficiency virus type-1 protease-resistant variants." *J Biol Chem* **271**(30): 17979-85.

PDB. (2002). "An Information Portal to Biological Macromolecular Structures ", from http://www.pdb.org/pdb/file_formats/pdb/pdbguide2.2/guide2.2_frame.html.

PDB. (2005). from www.pdb.org.

Pearl, L. and T. Blundell (1984). "The active site of aspartic proteinases." *FEBS Lett* **174**(1): 96-101.

Pettit, S. C., M. D. Moody, R. S. Wehbie, A. H. Kaplan, P. V. Nantermet, C. A. Klein and S. R. (1994). "The p2 domain of human immunodeficiency virus type 1 Gag regulates sequential proteolytic processing and is required to produce fully infectious virions." *J. Virol.* **68**(12): 8017-8027.

Phylip, L. H., A. D. Richards, J. Kay, J. Kovalinka, P. Strop, I. Blaha, J. Velek, V. Kostka, A. J. Ritchie, A. V. Broadhurst and et al. (1990). "Hydrolysis of synthetic chromogenic substrates by HIV-1 and HIV-2 proteinases." *Biochem Biophys Res Commun* **171**(1): 439-44.

Piana, S., P. Carloni and M. Parrinello (2002a). "Role of conformational fluctuations in the enzymatic reaction of HIV-1 protease." *J Mol Biol* **319**(2): 567-83.

Piana, S., P. Carloni and U. Rothlisberger (2002b). "Drug resistance in HIV-1 protease: Flexibility-assisted mechanism of compensatory mutations." *Protein Sci* **11**(10): 2393-402.

Polgar, L. (1987). "The mechanism of action of aspartic proteases involves 'push-pull' catalysis." *FEBS Lett* **219**(1): 1-4.

Polgar, L., Z. Szeltner and I. Boros (1994). "Substrate-dependent mechanisms in the catalysis of human immunodeficiency virus protease." *Biochemistry* **33**(31): 9351-7.

Porter, D. J., M. H. Hanlon and E. S. Furfine (2002). "HIV-1 protease: characterization of a catalytically competent enzyme-substrate intermediate." *Biochemistry.* **41**(4): 1302-1307.

Prabu-Jeyabalan, M., E. Nalivaika and C. A. Schiffer (2002). "Substrate shape determines specificity of recognition for HIV-1 protease: analysis of crystal structures of six substrate complexes." *Structure* **10**(3): 369-381.

Prabu-Jeyabalan, M., E. A. Nalivaika, N. M. King and C. A. Schiffer (2003). "Viability of a drug-resistant human immunodeficiency virus type 1 protease variant: structural insights for better antiviral therapy." *J. Virol.* **77**(2): 1306-1315.

Prabu-Jeyabalan, M., E. A. Nalivaika, N. M. King and C. A. Schiffer (2004). "Structural basis for coevolution of a human immunodeficiency virus type 1 nucleocapsid-p1 cleavage site with a V82A drug-resistant mutation in viral protease." *J. Virol.* **78**(22): 12446-12454.

Prado, J. G., T. Wrin, J. Beauchaine, L. Ruiz, C. J. Petropoulos, S. D. Frost, B. Clotet, R. T. D'Aquila and J. Martinez-Picado (2002). "Amprenavir-resistant HIV-1 exhibits lopinavir cross-resistance and reduced replication capacity." *Aids* **16**(7): 1009-17.

Precigoux, G., S. Geoffre, R. Leonard, S. Llido, A. Dautant, B. L. d'Estaintot, P. Picard, A. Menard, B. Guillemain and M. Hospital (1993). "Modelling, synthesis and biological activity of a BLV proteinase, made of (only) 116 amino acids." *FEBS Lett* **326**(1-3): 237-40.

Roberts, J. D., K. Bebenek and T. A. Kunkel (1988). "The accuracy of reverse transcriptase from HIV-1." *Science* **242**(4882): 1171-3.

Robinson, L. H., R. E. Myers, B. W. Snowden, M. Tisdale and E. D. Blair (2000). "HIV type 1 protease cleavage site mutations and viral fitness: implications for drug susceptibility phenotyping assays." *AIDS Res. Hum. Retroviruses.* **16**(12): 1149-1156.

Rose, R. B., C. S. Craik and R. M. Stroud (1998). "Domain flexibility in retroviral proteases: structural implications for drug resistant mutations." *Biochemistry* **37**(8): 2607-21.

Sayle, R. A. and E. J. Milner-White (1995). "RASMOL: biomolecular graphics for all." *Trends Biochem Sci* **20**(9): 374.

Shafer, R. W. (2002). "Genotypic testing for human immunodeficiency virus type 1 drug resistance." *Clin. Microbiol. Rev.* **15**(2): 247-277.

Shafer, R. W., M. A. Winters, S. Palmer and T. C. Merigan (1998). "Multiple concurrent reverse transcriptase and protease mutations and multidrug resistance of HIV-1 isolates from heavily treated patients." *Ann. Intern. Med.* **128**(11): 906-911.

Sheldrick, G. M. and T. R. Schneider (1997). "SHELXL: High resolution refinement." *Methods in Enzymology* **277**: 319-343.

Shurtleff, A. C. (2004). "TMC-114 (Tibotec)." *Curr Opin Investig Drugs* **5**(8): 879-86.

Silva, A. M., R. E. Cachau, H. L. Sham and J. W. Erickson (1996). "Inhibition and Catalytic Mechanism of HIV-1 Aspartic Protease." *J. Mol. Biol.* **255**(2): 321-346.

Skalka, A. M. and S. P. Goff (1993). "Reverse Transcriptase."

Slee, D. H., K. L. Laslo, J. H. Elder, I. R. Ollmann, A. Gustchina, J. Kervinen, A. Zdanov, A. Wlodawer and C. H. Wong (1995). "Selectivity in the Inhibition of HIV and FIV Protease: Inhibitory and Mechanistic Studies of Pyrrolidine-Containing .alpha.-Keto Amide and Hydroxyethylamine Core Structures." *J. Am. Chem. Soc.* **117**(48): 11867-11878.

Sperka, T., G. Miklóssy, Y. Tie, P. Bagossi, G. Zahuczky, P. Boross, R. W. Harrison, I. T. Weber and J. Tözsér (2006). "Bovine Leukemia Virus Protease: Comparison with Human T-cell Leukemia Virus and Human Immunodeficiency Virus Proteases." Submitted.

Suguna, K., R. R. Bott, E. A. Padlan, E. Subramanian, S. Sheriff, G. H. Cohen and D. R. Davies (1987). "Structure and refinement at 1.8 A resolution of the aspartic proteinase from Rhizopus chinensis." *J Mol Biol* **196**(4): 877-900.

Surleraux, D. L., A. Tahri, W. G. Verschueren, G. M. Pille, H. A. de Kock, T. H. Jonckers, A. Peeters, S. De Meyer, H. Azijn, R. Pauwels, M. P. de Bethune, N. M. King, M. Prabu-Jeyabalan, C. A. Schiffer and P. B. Wigerinck (2005). "Discovery and selection of TMC114, a next generation HIV-1 protease inhibitor." *J Med Chem* **48**(6): 1813-22.

Swairjo, M. A., E. M. Towler, C. Debouck and S. S. Abdel-Meguid (1998). "Structural role of the 30's loop in determining the ligand specificity of the human immunodeficiency virus protease." *Biochemistry* **37**(31): 10928-36.

Tamalet, C., C. Pasquier, N. Yahi, P. Colson, I. Poizot-Martin, G. Lepeu, H. Gallais, P. Massip, J. Puel and J. Izopet (2000). "Prevalence of drug resistant mutants and virological response to combination therapy in patients with primary HIV-1 infection." *J. Med. Virol.* **61**(2): 181-186.

Telesnitsky, A. and S. P. Goff (1997). "Reverse transcriptase and the generation of retroviral DNA." *In Retroviruses* **1**: 121-160.

Tessmer, U. and H. G. Krausslich (1998). "Cleavage of human immunodeficiency virus type 1 proteinase from the N-terminally adjacent p6* protein is essential for efficient Gag polyprotein processing and viral infectivity." *J Virol* **72**(4): 3459-63.

Tie, Y., P. I. Boross, Y. F. Wang, L. Gaddis, A. K. Hussain, S. Leshchenko, A. K. Ghosh, J. M. Louis, R. W. Harrison and I. T. Weber (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**(2): 341-52.

Tie, Y., P. I. Boross, Y. F. Wang, L. Gaddis, F. Liu, X. Chen, J. Tozser, R. W. Harrison and I. T. Weber (2005). "Molecular basis for substrate recognition and drug resistance from 1.1 to 1.6 angstroms resolution crystal structures of HIV-1 protease mutants with substrate analogs." *Febs J* **272**(20): 5265-77.

Tisdale, M., R. Myers, S. Randall, M. Maguire, M. Ait-Khaled, R. Elston and W. Snowden (2000). "Resistance of the HIV Protease Inhibitor Amprenavir In Vitro and in Clinical Studies." *Clin. Drug Invest.* **20**: 267-285.

Tozser, J., I. Blaha, T. D. Copeland, E. M. Wondrak and S. Oroszlan (1991a). "Comparison of the HIV-1 and HIV-2 proteinases using oligopeptide substrates representing cleavage sites in Gag and Gag-Pol polyproteins." *FEBS Lett* **281**(1-2): 77-80.

Tozser, J., A. Gustchina, I. T. Weber, I. Blaha, E. M. Wondrak and S. Oroszlan (1991b). "Studies on the role of the S4 substrate binding site of HIV proteinases." *FEBS Lett* **279**(2): 356-60.

Tozser, J., I. T. Weber, A. Gustchina, I. Blaha, T. D. Copeland, J. M. Louis and S. Oroszlan (1992). "Kinetic and modeling studies of S3-S3' subsites of HIV proteinases." *Biochemistry* **31**(20): 4793-800.

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

Tozser, J., G. Zahuczky, P. Bagossi, J. M. Louis, T. D. Copeland, S. Oroszlan, R. W. Harrison and I. T. Weber (2000). "Comparison of the substrate specificity of the human T-cell leukemia virus and human immunodeficiency virus proteinases." *Eur J Biochem* **267**(20): 6287-95.

Trylska, J., P. Grochowski and J. A. McCammon (2004). "The role of hydrogen bonding in the enzymatic reaction catalyzed by HIV-1 protease." *Protein Sci* **13**(2): 513-28.

Turner, B. G. and M. F. Summers (1999). "Structural biology of HIV." *J Mol Biol* **285**(1): 1-32.

Velazquez-Campoy, A., Y. Kiso and E. Freire (2001). "The binding energetics of first- and second-generation HIV-1 protease inhibitors: implications for drug design." *Arch Biochem Biophys* **390**(2): 169-75.

Velazquez-Campoy, A., M. J. Todd and E. Freire (2000). "HIV-1 protease inhibitors: enthalpic versus entropic optimization of the binding affinity." *Biochemistry* **39**(9): 2201-7.

Wang, G. T., E. D. Matayoshi, J. Huffaker and G. A. Krafft (1990). "Design and synthesis of new fluorogenic HIV protease substrates based on resonance energy transfer." *Tetrahedron Lett.* **31**: 6493-6496.

Wang, Y. X., D. I. Freedberg, T. Yamazaki, P. T. Wingfield, S. J. Stahl, J. D. Kaufman, Y. Kiso and D. A. Torchia (1996). "Solution NMR evidence that the HIV-1 protease catalytic aspartyl groups have different ionization states in the complex formed with the asymmetric drug KNI-272." *Biochemistry.* **35**(31): 9945-9950.

Weber, I. T. (1991). "Modeling of structure of human immunodeficiency virus-1 protease with substrate based on crystal structure of Rous sarcoma virus protease." *Methods Enzymol* **202**: 727-41.

Weber, I. T., J. Wu, J. Adomat, R. W. Harrison, A. R. Kimmel, E. M. Wondrak and J. M. Louis (1997). "Crystallographic analysis of HIV-1 protease with an analog of the conserved CA-p2 substrate: Interactions with frequently occurring Glu at P2' position of substrate." *Eur. J. Biochem.* **249**(2): 523-530.

Wei, X., S. K. Ghosh, M. E. Taylor, V. A. Johnson, E. A. Emini, P. Deutsch, J. D. Lifson, S. Bonhoeffer, M. A. Nowak, B. H. Hahn and et al. (1995). "Viral dynamics in human immunodeficiency virus type 1 infection." *Nature* **373**(6510): 117-22.

Whitcomb, J. M. and S. H. Hughes (1992). "Retroviral reverse transcription and integration: progress and problems." *Annu Rev Cell Biol* **8**: 275-306.

Wiegers, K., G. Rutter, H. Kottler, U. Tessmer, H. Hohenberg and H. G. Krausslich (1998). "Sequential steps in human immunodeficiency virus particle maturation revealed by alterations of individual Gag polyprotein cleavage sites." *J Virol* **72**(4): 2846-54.

Winters, M. A., J. M. Schapiro, J. Lawrence and T. C. Merigan (1998). "Human immunodeficiency virus type 1 protease genotypes and in vitro protease inhibitor susceptibilities of isolates from individuals who were switched to other protease inhibitors after long-term saquinavir treatment." *J Virol* **72**(6): 5303-6.

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

Wlodawer, A. and A. Gustchina (2000). "Structural and biochemical studies of retroviral proteases." *Biochim Biophys Acta* **1477**(1-2): 16-34.

Wlodawer, A., A. Gustchina, L. Reshetnikova, J. Lubkowski, A. Zdanov, K. Y. Hui, E. L. Angleton, W. G. Farmerie, M. M. Goodenow, D. Bhatt and et al. (1995). "Structure of an inhibitor complex of the proteinase from feline immunodeficiency virus." *Nat Struct Biol* **2**(6): 480-8.

Wlodawer, A., M. Miller, M. Jaskolski, B. K. Sathyanarayana, E. Baldwin, I. T. Weber, L. M. Selk, L. Clawson, J. Schneider and S. B. Kent (1989). "Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease." *Science* **245**(4918): 616-21.

Wondrak, E. M. and J. M. Louis (1996). "Influence of flanking sequences on the dimer stability of human immunodeficiency virus type 1 protease." *Biochemistry.* **35**(39): 12957-12962.

Wondrak, E. M., J. M. Louis, H. de Rocquigny, J. C. Chermann and B. P. Roques (1993). "The gag precursor contains a specific HIV-1 protease cleavage site between the NC (P7) and P1 proteins." *FEBS Lett* **333**(1-2): 21-4.

Wong, J. K., M. Hezareh, H. F. Gunthard, D. V. Havlir, C. C. Ignacio, C. A. Spina and D. D. Richman (1997). "Recovery of replication-competent HIV despite prolonged suppression of plasma viremia." *Science* **278**(5341): 1291-5.

Wu, J., J. M. Adomat, T. W. Ridky, J. M. Louis, J. Leis, R. W. Harrison and I. T. Weber (1998). "Structural basis for specificity of retroviral proteases." *Biochemistry* **37**(13): 4518-26.

Wu, T. D., C. A. Schiffer, M. J. Gonzales, J. Taylor, R. Kantor, S. Chou, D. Israelski, A. R. Zolopa, W. J. Fessel and R. W. Shafer (2003). "Mutation patterns and structural correlates in human immunodeficiency virus type 1 protease following different protease inhibitor treatments." *J. Virol.* **77**(8): 4836-4847.

Yamazaki, T., A. P. Hinck, Y. X. Wang, L. K. Nicholson, D. A. Torchia, P. Wingfield, S. J. Stahl, J. D. Kaufman, C. H. Chang, P. J. Domaille and P. Y. Lam (1996). "Three-dimensional solution structure of the HIV-1 protease complexed with DMP323, a novel cyclic urea-type inhibitor, determined by nuclear magnetic resonance spectroscopy." *Protein Sci* **5**(3): 495-506.

Zhang, Y. M., H. Imamichi, T. Imamichi, H. C. Lane, J. Falloon, M. B. Vasudevachari and N. P. Salzman (1997). "Drug resistance during indinavir therapy is caused by mutations in the protease gene and its Gag substrate cleavage sites." *J. Virol.* **71**(9): 6662-6670.

Zoete, V., O. Michielin and M. Karplus (2002). "Relation between sequence and structure of HIV-1 protease inhibitor complexes: a model system for the analysis of protein flexibility." *J Mol Biol* **315**(1): 21-52.

APPENDIX I. List of all HIV-1 protease structures I have studied.

Inhibitor	PR	Reso. (Å)	Rwork	Rfree	Occu. Of Inh.	Collected at	Crystal	Solved by	Code	
CA-p2	V82A	1.54	0.12	0.19	0.65/0.35	ANL-1202			2AOE	
p2-NC	WT	1.40	0.15	0.19	1.0	ANL-1003	#	#	2AOD	$
	I84V	1.30	0.12	0.16	1.0	BNL-1102		#	2AOC	
	V82A	1.10	0.13	0.17	1.0	ANL-0903			2AOG	
p6-PR	WT	1.60	0.15	0.23	0.6/0.4	ANL-0703			2AOJ	
	V82A	1.42	0.18	0.23	0.8/0.2	BNL-1102		#	2AOH	
p1-p6	WT	1.38	0.16	0.21	0.6/0.4	ANL-0702			2AOI	
	V82A	1.32	0.13	0.18	0.6/0.4	BNL-0602			2AOF	
UIC-94017	WT	1.30	0.13	0.17	0.55/0.45	BNL-1102	#	#	1S6G	$
	I84V	1.53	0.13	0.20	0.69/0.31	ANL-1202	#	#	1S6S	$
	V82A	1.10	0.12	0.15	partial	ANL-0702			1S65	
	I50V	1.22	0.15	0.22	0.6/0.4	ANL-1003	#		2F8G	
	D30N	1.45	0.13	0.19	0.52/0.48	ANL-1004	#	*	2F80	
	G86A	1.60	0.19	0.26	0.67/0.33	ANL-0705				
	G86S	1.80	0.22	0.25	0.6/0.4	ANL-0305				
DMP	G86A	1.80	0.22	0.29	1.0	ANL-1104				
SQV	WT	1.16	0.13	0.18	0.51/0.49	ANL-1104				
	I84V	1.25	0.15	0.19	0.69/0.31	ANL-0305				
	I84V	1.20	0.15	0.20	1.0	ANL-0305				
	V82A	1.10	0.15	0.18	0.66/0.34	ANL-1104				

	V82A	0.97	0.12	0.14	0.48/0.52	ANL-1104				
UIC98065	V82A	1.11	0.13	0.16	1.0	ANL-0703		#		
	I50V	1.28	0.16	0.20	0.57/0.43	ANL-1003		#		
UIC00122	WT	1.29	0.14	0.17	1.0	BNL-1102		#		
	I84V	1.40			1.0	ANL-0703		#		
	I50V	1.25			1.0	ANL-1003		#		
UIC00072	WT	1.37	0.13	0.18	1.0	BNL-1102				
	V82A	1.25	0.16	0.20		ANL-1105				%
UIC02070	WT	1.40	0.12	0.18	0.81/0.19	ANL-0303				
UIC02029	WT	1.45	0.13	0.19	0.61/0.39	ANL-0303				
UIC02059	WT	1.10	0.16	0.19	1.0	ANL-1003				%
UIC02038	WT	1.15	0.14	0.19	0.51/0.49	ANL-0704	#			

#: Done by other labmates; $: analyzed only; *: partial contribution; %: further refinement needed

APPENDIX II. Structure prediction of BLV PR

Introduction

All replication competent retroviruses, including bovine leukemia virus (BLV), code for an aspartic protease (PR). The function of the mature PR is critical for virion replication (Oroszlan et al. 1990). As discussed above, the HIV-1 PR has proved to be the most effective target of antiviral therapy; however, the application of PR inhibitors was largely limited by rapidly development of drug resistance variants (Tamalet et al. 2000). The residues observed in HIV-1 PR drug resistant frequently can be found in structurally equivalent positions in other wild type proteases, as indicated for BLV PR in Figure 5.1. Comparison of various retroviral proteases will reveal the common features of their specificity; and furthermore, help the design of new antiviral inhibitors that are efficient against several different retroviral proteases.

The amino acid sequences of different retroviral proteases have been determined (Oroszlan et al. 1990). The 3-D structures of seven proteases from HIV-1, HIV-2, SIV, RSV, FIV, EIAV and HTLV-1 are also available (Wlodawer et al. 2000; Li et al. 2005). The alignment of the primary and secondary structures of all retroviral proteases suggested a single domain of the cellular aspartic proteases. Comparison of the predicted PR structures of HIV-1, FIV and EIAV proteases with their latest crystal structures proved that the models were basically correct in prediction of the substrate binding sites (Weber 1991; Wlodawer et al. 1995; Gustchina et al. 1996). Therefore, the molecular modeling of proteases can serve as an important tool in the absence of crystal structures.

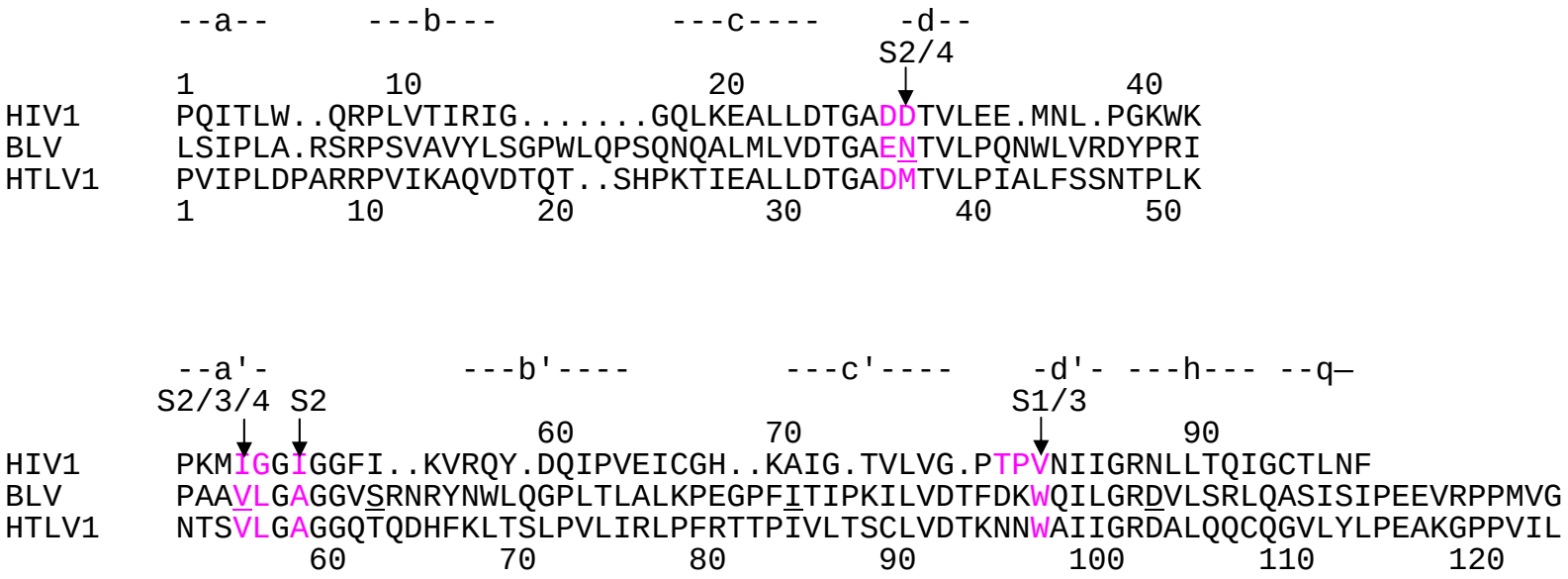


Figure 5.1 Sequence comparison of BLV, HTLV-1 and HIV-1 proteases.
Residues in BLV PR which appear in HIV-1 drug resistance at the equivalent position in HIV-1 PR are underlined (analyzed based on hivdb.stanford.edu).

Method

The sequence alignment of the BLV PR, HIV-1 PR and RSV PR was shown in Figure 5.1. The crystal structure of the HTLV-1 PR with a substrate-based statine inhibitor (Li et al. 2005) was the basis for the model for BLV PR. The BLV dimer was modeled with the peptide substrates TKVL-VVQP using the program AMMP (Harrison 1993), as described previously for HTLV PR (Tozser et al. 2000). A conserved water molecule was included between the flaps and the peptide and a proton was used to stabilize the charged oxygen atoms of the two catalytic Asp residues. The positions for all new atoms were generated with the sp4 potential set using 15 cycles of the Kohonen algorithm (Harrison et al. 1999), followed by minimization of the non-bonded and geometrical terms using 320 steps of conjugate gradients. Finally, the entire PR-substrate complex was minimized using 10 cycles of total ~ 4000 steps of conjugate gradients. Structural models were examined using the molecular graphic programs Sybyl (Tripos Inc.) run on Silicon Graphics workstations, 'O' (Jones et al. 1991) or RasMol (Sayle et al. 1995) run on Linux PCs.

Results

Modeled structure of the BLV protease and comparison with the structures of HTLV-1 and HIV-1 proteases.

In the absence of a crystal structure for the BLV PR, a model was built with a peptide substrate in order to understand the molecular basis for the specificity. The amino acid sequence of BLV PR was aligned with the sequences of HIV-1 PR and HTLV-1 PR,

the structures of which are known, to find the best starting structure for building the model, as shown in Figure 5.1. The lengths of the BLV, HTLV-1 and HIV-1 proteases are different. The HIV-1 PR is the shortest one with 99 residues, followed by the HTLV-1 PR (125 residues), while BLV PR has 126 residues. Since BLV PR and HTLV-1 PR have comparable length and higher similarity, HTLV-1 PR was chosen as the initial model. Regardless of the difference lengths, the modeled BLV PR structure shared a conserved core region with the crystal structures of HIV-1 and HTLV-1 proteases, including the substrate binding site and the dimer interface (Figure 5.2). As shown in Figure 5.2, BLV and HTLV-1 proteases formed longer loop structures on the surface of the proteins. Besides, the C-termini extended by 10 additional amino acid residues for BLV and HTLV proteases compared to the HIV-1 PR. The function of the C-terminal extension is still controversial. Based on previous studies, these residues are not required for activity (Precigoux et al. 1993; Herger et al. 2004), while other group suggested that five additional C-terminal residues (116-120) appeared to be important for HTLV-1 protease (Hayakawa et al. 1992). The final model of the BLV PR dimer with substrate had a RMS deviation of 0.79 Å for 223 pairs C_α atoms of compared to the crystal structure of HTLV-1 PR and 1.72 Å for 175 pairs of C_α atoms of compared to the crystal structure of HIV-1 PR (PDP code 2AOD). These RMS deviation values are comparable to that of 1.25 Å observed for 194 pairs of C_α atoms in the dimers of RSV S9 PR and HIV-1 PR with inhibitor (Wu et al. 1998). Residues forming the substrate binding sites of BLV, HTLV-1 and HIV-1 proteases are listed in Table 5.1 and a schematic representation with a HTLV-1 capsid-nucleocapsid substrate is shown in Figure 5.3.

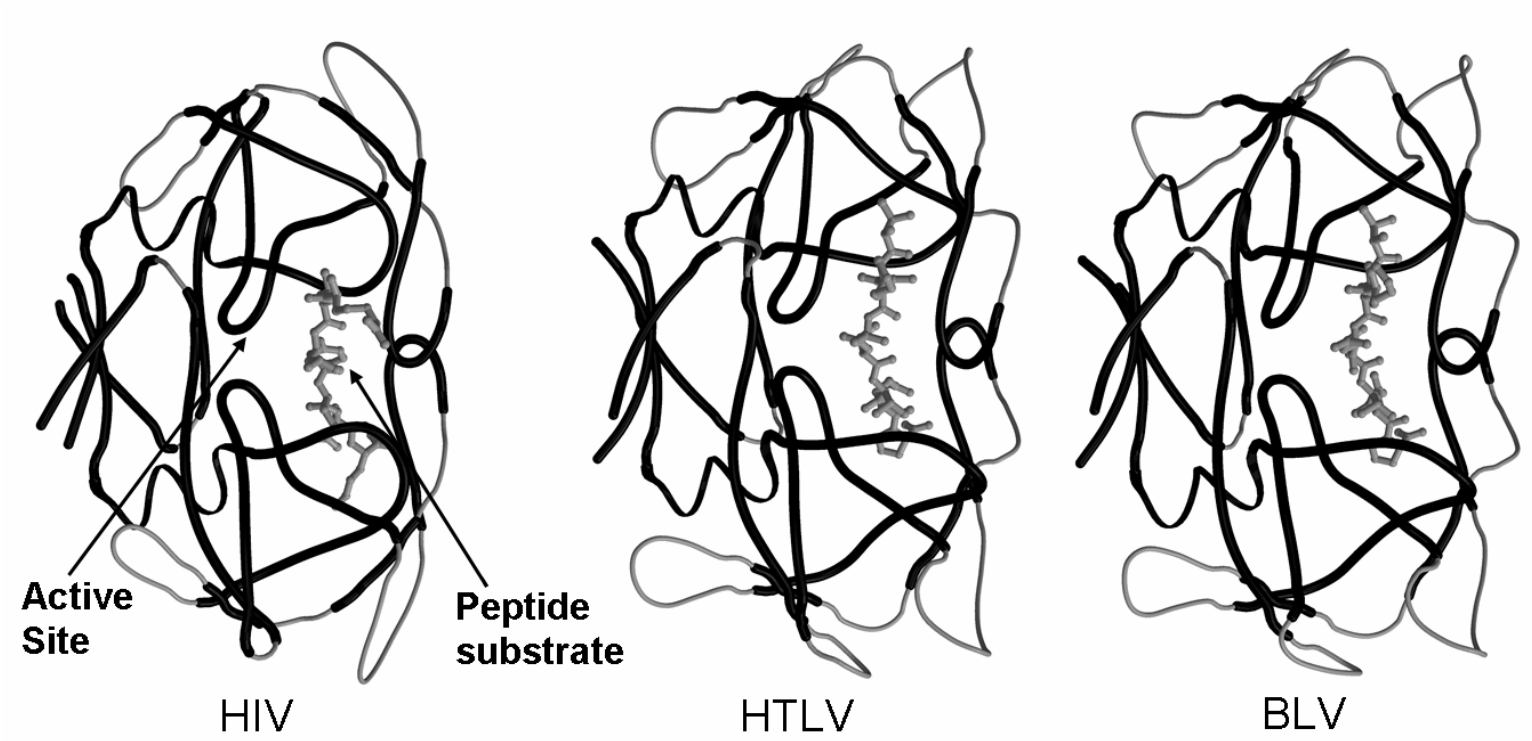


Figure 5.2 Ribbon models of the BLV, HTLV-1 and HIV-1 PR structures.

Table 5.1 Residues predicted to form the subsites of BLV, HTLV and HIV proteases.

Subsite	BLV/HTLV-1/HIV-1 Residues[a]
S_4:	*Glu_{37}/Asp_{36}/Asp_{29}*, **Asn_{38}/Met_{37}/Asp_{30}**, **Ala_{56}/Ser_{55}/Met_{46}**, *Val_{57}/Val_{56}/Ile_{47}*, *Leu_{58}/Leu_{57}/Gly_{48}*, Tyr_{68}/Phe_{67}/Val_{56}, **Trp_{70}/Leu_{69}/Gln_{58}**, Leu_{92}/Leu_{91}/Leu_{76}
S_3:	$Arg_{9'}$/$Arg_{10'}$/$Arg_{8'}$, $Leu_{31'}$/$Leu_{30'}$/$Leu_{23'}$, *Glu_{37}/Asp_{36}/Asp_{29}*, *Leu_{58}/Leu_{57}/Gly_{48}*, *-/-/$Thr_{80'}$*[b], **-/-/$Pro_{81'}$**, *$Trp_{99'}$/$Trp_{98'}$/$Val_{82'}$*
S_2:	(Gly_{35}/Gly_{34}/Gly_{27}), Ala_{36}/Ala_{35}/Ala_{28}, (*Glu_{37}/Asp_{36}/Asp_{29}*), **Asn_{38}/Met_{37}/Asp_{30}**, Val_{40}/Val_{39}/Val_{32}, *Val_{57}/Val_{56}/Ile_{47}*, *Leu_{58}/Leu_{57}/Gly_{48}*,(Gly_{59}/Gly_{58}/Gly_{49}), *$Ala_{60'}$/$Ala_{59'}$/$Ile_{50'}$*, Tyr_{68}/Phe_{67}/Val_{56}, Leu_{92}/Leu_{91}/Leu_{76}, Ile_{101}/Ile_{100}/Ile_{84}
S_1:	$Arg_{9'}$/$Arg_{10'}$/$Arg_{8'}$, $Leu_{31'}$/$Leu_{30'}$/$Leu_{23'}$, (Asp_{33}/Asp_{32}/Asp_{25}), ($Asp_{33'}$/$Asp_{32'}$/$Asp_{25'}$) (Gly_{35}/Gly_{34}/Gly_{27}), (Gly_{59}/Gly_{58}/Gly_{49}) *Ala_{60}/Ala_{59}/Ile_{50}*, *-/-/$Thr_{80'}$*[b], **-/-/$Pro_{81'}$**, *$Trp_{99'}$/$Trp_{98'}$/$Val_{82'}$*, $Ile_{101'}$/$Ile_{100'}$/$Ile_{84'}$

[a]Residues that are different in all three proteases are indicated in bold, while residues differing in two proteases are indicated in italics. Amino acid residues in the second subunit are indicated by a prime. Only the residues forming the S_4-S_1 subsites are given. Primed binding sites (like S_1') have the same composition as the nonprimed ones, but they are built from residues of the other subunit. Residues involved only in interactions with the main chain atoms of the peptide substrate are shown in parenthesis.
[b]While $Thr_{80'}$ and $Pro_{81'}$ of HIV-1 PR participate in the formation of the binding sites, the corresponding residues of BLV and HTLV-1 proteases do not interact with ligand; their position is very different, due to an upstream loop insertion.

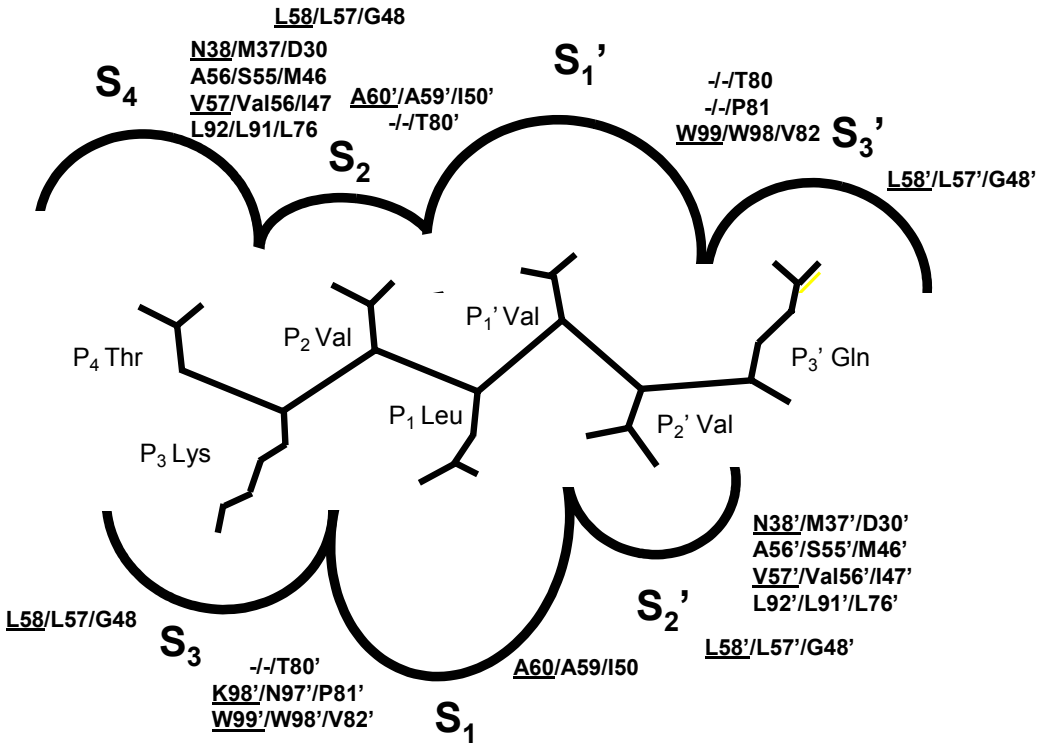


Figure 5.3 A schematic representation of the HTLV-1 capsid-nucleocapsid substrate in the S4-S3' subsites of BLV, HTLV-1 and HIV-1 proteases.
The relative size of each subsite is indicated approximately by the area enclosed by the *curved line* around each substrate side chain. PR residues forming the subsites are shown only for those that differ between at least two PRs as BLV/HTLV-1/HIV-1 residues. Residues of BLV PR that were mutated in this study are underlined.

The specificity of the BLV PR was studied with a set of oligopeptides representing naturally occurring cleavage sites in various retroviruses by Dr. Tozser's group. They have concluded that the BLV PR had a broad specificity, which is different from HTLV-1 PR but similar to HIV-1 PR, and a great sensitivity towards mutations as previously observed for the HTLV-1 PR. Furthermore, the BLV PR was more susceptible towards inhibitors designed against either HIV-1 PR or HTLV-1 PR compared to the susceptibility of the HTLV-1 PR (Sperka et al. 2006).