CHAPTER 1

1. Molecular recognition

In each cell of an organism, a myriad of fundamental molecular reactions, covalent and non-covalent, co-ordinate and regulate the way in which biological molecules interact with each other. There are also finely-tuned inter- and intra-molecular recognition mechanisms that form part of an intricate network of interdependent reactions crucial in a cellular environment. These interactions are paramount in regulating most of the biological processes, for example, protein-DNA interactions, protein-protein interactions, folding of a nascent polypeptide chain to its native and functional three-dimensional conformation, subunit-subunit interactions in oligomeric proteins, hormone-receptor and antibody-antigen interactions to name a few. As a corollary, interactions present in biomolecules necessary for biological processes and molecular recognition are governed by electrostatic and geometric complementarity between the surfaces of the interacting partners (Breckenridge, 1991). A quantitative understanding of these processes has a significant practical interest, since it will provide new insights in the development of structure-based molecular design strategies. These strategies are aimed at producing proteins with specific functional or stability properties, pharmaceutical drugs targeted to specific receptors, etc. Two types of molecular recognition mechanisms exist, the lock-and-key and induced-fit mechanisms.

1.1 Lock-and-key versus induced-fit mechanism of ligand binding

Proteins play an essential role in cellular activity, and their functions depend largely on their interactions with other molecules. Over 100 years ago, Emil Fischer described the classic lock-and-key model which states that the specificity of a molecular receptor such as a protein or enzyme (lock) for its ligand or substrate (key) arises from their geometric complementary shapes, where no structural optimization of the binding partners is required (for a review, see Koshland, 1994). The substrate binding site on the protein surface has a preformed cavity that is complementary in shape to the substrate and amino acid residues making up the binding site are structurally optimized to interact specifically with the substrate (Caret *et al.,* 1993). Therefore, enzymes can only accommodate molecules that are similar in shape and functional group, and hence improved specificity. The lock-and-key model assumes that the binding site is rigid and exists in the absence of a bound ligand. However,

there is a growing consensus that protein flexibility maybe required for optimal molecular recognition since proteins are flexible and dynamic molecules that often undergo structural reorganizations while perfoming specific functions (Ishima and Torchia, 2000). The importance of flexibility in the affinity and specificity of molecular interactions include substrate binding, product release, regulation and allosteric behaviour, contractile and motor functions, and the humoral immune system, where a limited set of proteins (antibodies) must bind virtually unlimited range of foreign molecules (antigens) (Patten *et al.,* 1996; Wedenmayer *et al.*, 1997).

The original model by Daniel Koshland that evoked flexibility, known as 'induced-fit', posits that the initial binding of a substrate induces conformational changes in the enzyme which result in optimized binding interactions (Koshland, 1994). In this model, the substrate binding site assumes a flexible binding pocket that approximates the shape of the substrate rather than a rigid pocket (Caret *et al,* 1993). Many studies have shown that the optimization of protein-based molecular recognition requires, to some extent, significant conformational adjustments or rearrangements of the interacting proteins, ligands or substrates (Sigrell *et al.*, 1999; Leavitt and Freire, 2001).

1.2 Protein-protein and protein-ligand interactions

The key role of protein binding interactions in cellular metabolism is well documented. Detailed knowledge of the forces or interactions involved in stabilization of protein complexes has also been an area of considerable interest. The non-covalent interactions responsible for inter- and intra-molecular recognition between protein-protein and protein-ligand include: hydrogen bonding, hydrophobic interactions, electrostatic interactions and steric interactions.

1.2.1 Polar interactions

Electrostatic interactions occur between oppositely charged groups that are within 4 Å of each other and referred to as salt-bridges or ion pairs (Barlow and Thornton, 1983; Dill, 1990). In proteins, electrostatic interactions normally occur at the amino and carboxyl termini and on ionizable side chains of lysine, arginine, histidine, glutamic acid and aspartic acid. The strength or magnitude of an electrostatic interaction is largely dependent on the dielectric constant of the medium. The interior of the protein

molecule, with a low dielectric constant, will lead to stronger charge interactions because of the absence of water molecules shielding the ions from each other (Murphy, 1995).

Hydrogen bonding (D-H----A) is a non-covalent interaction formed when a hydrogen atom is shared between two electronegative atoms. The hydrogen atom is linked covalently to one atom referred to as the donor (D-H), and bonded electrostatically to the acceptor atom (A) that contains a lone electron pair. The most energetically favorable hydrogen bond primarily assumes a linear configuration of donor, hydrogen and acceptor, and consists of electrostatic, dispersion, charge-transfer and steric repulsion interactions (Creighton, 1993). In proteins, hydrogen bond donors and acceptors include polar amino acid side chains, the carbonyl C=O and amide NH groups of the peptide backbone. Studies have shown that hydrogen bonding contributes favorably in the stabilization and function of protein molecules (Marqusee and Sauer, 1994; Makhatadze and Privalov, 1995; Habermann and Murphy, 1996; Jiang and Lai, 2002).

Studies on the role of hydrogen bonds at protein binding interfaces have also revealed that there is a higher frequency of hydrophilic pairs (hydrogen bonds, salt-bridges and polar-polar interactions) at the interfaces than in the protein core (Tsai *et al.*, 1997; Xu *et al.*, 1997; Jamin and Chothia, 1990). It must be noted, however, that these studies focused on the 'lock-and-key' paradigm, where no structural optimization for complex formation is required.

1.2.2 Hydrophobic interactions

Hydrophobic interactions occur when non-polar groups interact with each other due to the unfavorable interaction of these groups with water. The inability of non-polar groups to form hydrogen bonds with the surrounding water molecules forces water molecules to form ordered cages around the non-polar groups, as a way of satisfying their hydrogen bonding potential. Formation of this clathrate-like structure leads to an unfavorable entropy of water molecules around the non-polar molecules as compared to the water molecules in the bulk solvent (Franks, 1975). The contribution of hydrophobic interactions as the major driving force for protein folding, in protein stability and in ligand binding is well documented (Dill, 1990). A study by Tsai *et al.*

(1997) has suggested that the hydrophobic effect is not as dominant in protein-protein binding as it is in protein folding. This was based on the observation that binding interfaces of such protein-protein complexes are enriched with hydrophobic residues but have more charged and polar residues than protein cores (Xu *et al.,* 1997).

1.3 Protein binding energetics and isothermal titration calorimetry

The completion of the human genome project has catapulted the number of novel targets for drug development to great heights. Many of these targets belong to protein families with homologous structures and similar binding pockets which are crucial in regulating pathways and interaction networks describing cell function and inter-relation. It is also apparent that the basis of molecular recognition in drug discovery and in protein-ligand complexes requires complete structural and thermodynamic dissection of macromolecular interactions involved. Several techniques (fluorescence, absorbance, nuclear magnetic resonance, surface plasmon resonance and ultracentrifugation) have been used as premier tools for characterizing interactions of biomolecules. These techniques can only determine the binding affinity constant (K_a) and indirectly derive other thermodynamic parameters.

Due to its sensitivity, accuracy and precision, isothermal titration calorimetry (ITC) is the most rigorous and preferred method applied in a wide range of chemical and biochemical reactions. ITC measures directly the heat absorbed or released (ΔH) associated with a given binding process. A typical ITC experiment allows the dissection of a binding energy of a reaction into ligand-enzyme association constant (K_a), change in enthalpy (ΔH), change in entropy (ΔS), change in Gibbs free energy (ΔG) and the stoichiometry of association (N). The change in heat capacity (ΔC_p) is obtained from measurements of binding enthalpy over a range of temperature. The magnitude and signs of the thermodynamic parameters obtained provide insight into the nature of interactions involved in the binding process. The strength of an interaction and whether or not it is thermodynamically favorable is determined by the Gibbs free energy. The Gibbs free energy change of binding is an important thermodynamic descriptor of a binding reaction since it dictates the binding affinity:

$$\Delta G = -RT\ln K_{eq} \qquad (1)$$

where R is the universal gas constant (8.314 J/mol/K), T is the temperature in kelvin and K_{eq} is the equilibrium binding constant. ΔG is in turn defined by the enthalpy and entropy changes, and is expressed in the following equation:

$$\Delta G = \Delta H - T\Delta S \quad (2)$$

At the molecular level, this reflects the opposition of two fundamental effects – the tendency to fall to lower energy (bond formation, negative ΔH), offset by the equally natural tendency for thermal (Brownian) motion to disrupt things (bond breakage, positive ΔS) (Cooper, 1999).

A detailed description of the instrument and technique can be found elsewhere in the literature (see McKinnon *et al.,* 1984; Wiseman *et al.,* 1989; Freire *et al.,* 1990; Velazquez-Campoy *et al.,* 2004; Velazquez-Campoy and Freire, 2005 and references therein). Briefly, the titration calorimeter consists of the injector system, adiabatic shield, and matched reference and sample cells. The adiabatic shield ensures that no heat exchange occurs between the cells and the surroundings (Wiseman *et al.,* 1989). The two cells are maintained at a constant and identical temperature by a feedback system that supplies thermal power continuously. In the event of a reaction in the sample cell usually accompanied by heat (exothermic reaction), the system ensures that the feedback power is withdrawn in order to retain thermal equilibrium between the cells. The feedback power supplied or withdrawn to minimize temperature imbalances upon ligand injection is measured and converted into the heat of interaction. A sequence of injections is programmed and the ligand solution is injected at regular intervals into the sample cell. After each injection, the composition inside the sample cell changes causing the rearrangement of populations and complex formation (Velazquez-Campoy and Freire, 2005). The system will pass through different states of equilibrium each differing in composition, as the sequence of injections continues. The heat associated with each injection is proportional to the increase in complex concentration (advance of the reaction) and it is calculated by integrating the area under the deflection signal measured (amount of heat per unit of time provided to maintain thermal equilibrium in the sample and reference cells) (Velazquez-Campoy and Freire, 2005). At the end of the experiment, saturation of the macromolecule is reached and it is possible to estimate K_a, ΔH and N, (independent variables). If two binding processes are characterized by different enthalpic and entropic terms and have the same Gibbs free energy of binding, they correspond to

different binding modes and, therefore, the main underlying intermolecular interactions are different.

Since the native state of a protein exists as an ensemble of conformational states, the energy of stabilization of protein structure will not be evenly distributed throughout its three-dimensional structure (Luque and Freire, 2000) There are regions of the protein with high stability constants (e.g. the hydrophobic core) and regions with low stability constants (e.g. loops and turns) with the majority of proteins exhibiting a dual character as originally observed for the HIV-1 protease (Luque *et al.*, 1998; Todd and Freire, 1999). Since low molecular weight ligands are in general not found attached to the surface of proteins but are buried in crevices or cavities covered by loops or other structural elements of the protein, the number of contacts between ligand and protein are maximized and simultaneously buries a significant surface area from the solvent (Luque and Freire, 2000). This conformational rearrangement often permits the entry of the ligand into the binding site and its subsequent shielding from the solvent and; hence, makes favorable contributions to the Gibbs free energy of binding. If the rearrangements are only transient and the free and the bound states of the protein are similar, only binding kinetics are affected. If, however, the free and bound conformations of the protein are different, the binding affinity will be affected (Luque and Freire, 2000). The Gibbs free energy associated with the change in protein conformation from its free to its bound state is included in the computation of the effective Gibbs energy of binding and corresponding binding affinity:

$$\Delta G_{\text{bind}} = \Delta G^{\circ}_{\text{bind}} + \Delta G_{\text{conf}} \qquad (3)$$

where $\Delta G^{\circ}_{\text{bind}}$ is the Gibbs energy of binding obtained under the assumption that the free and bound states of the protein are the same, and ΔG_{conf} is the Gibbs energy associated with the change in protein conformation from its free to its bound form. In general, the Gibbs energy associated with a change from a less stable region to bound conformation will be smaller than that associated with a change from a stable conformation to the bound conformation (Luque and Freire, 2000). The presence of flexible regions in the protein molecule appears to facilitate the ligand-induced conformational changes if the putative binding site is not binding-competent in the ligand-free protein. The presence of regions with low stability also appears to provide a mechanism for achieving high binding affinity for low molecular weight ligands and

serves as a starting point for propagation of binding signals to distal sites (Luque and Freire, 2000).

Enthalpic and entropic contributions of the Gibbs energy originate from different types of interactions in the binding process. The binding enthalpy primarily reflects the energetic contribution of many individual interactions (hydrogen bonds, van der Waals interactions, polar and dipolar interactions) between the ligand and the protein during the binding process, the conformational changes associated with binding, including interactions associated with the solvent. A negative (favorable) ΔH occurs when the interactions between the interacting molecules (e.g. hydrogen bond formation and van der Waals interactions) over-compensate the interactions of the individual molecules with the bulk solvent; otherwise, it will be positive (unfavorable, as for non-specific hydrophobic interactions) (Ross and Subramanian, 1981). The observed binding ΔH measured from a single ITC experiment often includes contributions not only from the actual binding event, but also the heat that is due to buffer ionization (Gomez and Freire, 1995; Baker and Murphy, 1996; 1997). This is particularly true when the primary binding event is accompanied by the transfer of protons between the solvent and the protein-ligand complex. Thus, the determination of the intrinsic energetics of ligand binding requires experiments or measurements to be performed separately as a function of pH in buffers with different ionization enthalpies (Gomez and Freire, 1995; Doyle *et al.*, 1995; Baker and Murphy, 1996). From this, pK_a values of ionizable groups responsible for proton linkage in the free and bound states and the number of protons that are coupled to the binding reaction can be easily calculated (Murphy and Baker, 1996).

The binding entropy refers to the degree of disorder accompanying complex formation. Two major terms that contribute to the change in entropy are the solvation and conformational entropies. Solvation entropy arises from the gain in degrees of freedom of water molecules that, prior to the binding, are localized on the surface of the binding molecules and are released to the bulk solvent upon binding due to partial or complete desolvation of the two binding molecules. The change in solvation entropy is favorable (positive) if the surfaces that are buried upon binding are predominantly hydrophobic. It, therefore, originates from the burial of hydrophobic surfaces upon binding. Entropically driven ligand binding reactions are characterized

by a large positive entropic contribution driven by the tendency of the molecule to escape water rather than by favorable interactions with the target molecule. In addition, the burial of solvent-exposed molecular surface area upon binding also contributes substantially to the heat capacity change upon complex formation due to the release of electro-restricted water or "hydrophobic water" from the binding site (Sturtevant 1977). The conformational entropy, on the other hand, arises from changes in conformational degrees of freedom experienced by both the protein and the ligand upon binding. It is usually negative (unfavorable) due to the loss of degrees of freedom resulting from the reduction in the number of accessible conformations and configurations of both molecules (protein and ligand) upon binding.

1.4 Thermodynamic leads in drug design

Currently, development of lead compounds or drug design is centered on optimization of their binding affinity towards the intended target. The binding affinity of a compound can be improved by generating a favorable binding enthalpy, favorable solvation entropy and by minimizing the unfavorable conformational entropy. It is evident that simultaneous optimization of the three factors can achieve extremely high affinity. However, it is entirely feasible to design lead compounds that bind to the intended target with similar affinity but with different binding mechanisms, i.e., entropically or enthalpically driven ligands (Velazquez-Campoy *et al.*, 2001b).

Entropically driven ligand derives most of its binding energy from a non-specific hydrophobic effect, i.e., by making interactions of the drug with the solvent unfavorable, whilst enthalpically driven ligand derives its binding energy by establishing strong and specific hydrogen bonds with the target. Drug designers have long aimed at developing conformationally constrained ligands preshaped to the geometry of the selected binding site, which completes entropy optimization. Accordingly, a comformationally constrained molecule that is preshaped to the target achieves affinity, specificity and selectivity through hydrophobicity and shape complementarity (Ruben *et al.,* 2006). Perhaps, the most significant example is given by the development of the first-generation HIV-1 protease (HIV-1 PR) inhibitors (saquinavir, ritonavir, indinavir and nelfinavir). The binding of these HIV-1 protease inhibitors is entropically driven and their binding enthalpy is either unfavorable (saquinavir, indinavir and nelfinavir) or only slightly favorable (ritonavir) (Todd *et*

al., 2000; Velazquez-Campoy *et al.,* 2003a). In all cases, the dominant force for binding is a large positive entropy change that originates primarily from the burial of a large hydrophobic surface upon binding (Todd *et al.,* 2000). Moreover, since shape and hydrophobicity are non-specific interactions, a change in the target binding site would lead to a reduction in the binding affinity. A low binding affinity reflects the inability of these conformationally rigid ligands to adapt to changes in the target binding pocket due to mutations or naturally occurring polymorphisms arising from genetic diversity. Hydrophobicity has historically been the preferred variable in the pharmaceutical industry due to its ease in implementation (Lipinski, 2000).

A favorable enthalpic contribution is an indication of specific interactions between binding partners and a good way to measure ligand specificity, selectivity and adaptability. On the other hand, an unfavorable enthalpy contribution is an indication of non-specific interactions between the binding partners, making it very difficult to provide specificity, selectivity and adaptability. Despite apparent advantages of enthalpic interactions in achieving high affinity and improved selectivity, the optimization of the binding enthalpy has been more cumbersome to implement due to a large and unfavorable desolvation enthalpy of polar groups (Cabani *et al.,* 1981). Generally, a polar group needs to make a strong interaction with the target in order to compensate for the desolvation enthalpy. Energetic contributions to binding affinity are not simply localized to the direct interactions between molecules, but contain interactions from structural and dynamic changes propagated throughout the protein, and from counter ions and hydrating water molecules located at the binding site. To be effective, an inhibitor needs to exhibit an extremely high affinity for the intended target and be mildly affected by mutations. Ideally, an inhibitor should have a binding affinity in the 1 – 50 pM range against the wild-type and be affected by mutations by a factor of 100 or less (Velazquez-Campoy and Freire, 2001a; Velazquez-Campoy *et al.,* 2001b; Ohtaka *et al.,* 2002). Compounds that achieve high binding affinity or that maximize binding affinity have been shown to combine or balance the favorable entropic and enthalpic contributions to the overall Gibbs energy of binding (Ohtaka *et al.,* 2003; Ohtaka *et al.,* 2004; Ohtaka and Freire, 2005; Carbonell and Freire, 2005). Notably, drug design paradigms have to a large extent illustrated how enthalpic or entropic character of inhibitors is not dependent on the intended target, and that it is possible to develop entropically as well as enthalpically optimized inhibitors against

the same binding site (e.g. HIV-1 protease). It has, for example, taken over 10 years to optimize HIV-1 protease inhibitors from the entropically driven inhibitors to the new and more potent enthalpically driven inhibitors (Velazquez-Campoy *et al.,* 2003a; Ohtaka *et al.,* 2004; Muzammil *et al.,* 2007). The second-generation HIV-1 protease inhibitor, KNI-764 (AG-1776) for example, achieves the highest affinity (K_d = 32 pM) to the HIV-1 protease with a binding enthalpy (ΔH) of – 7.6 kcal/mol and an entropic contribution ($-T\Delta S$) of -6.7 kcal/mol and can still afford the presence of certain flexible elements (Velazquez-Campoy *et al.,* 2003a; Vega *et al.,* 2004). The introduction of flexible asymmetrical functional groups in regions facing or in close proximity to mutation-prone areas of the protein provides adaptability to the inhibitor and low susceptibility to mutations (Todd *et al.,* 2000; Ohtaka *et al.,* 2002; Velazquez-Campoy *et al.,* 2001b). The increased conformational flexibility found in the second-generation HIV-1 protease inhibitors can also allow the inhibitor to compensate for the loss of interactions as a result of mutations in the target by burying a comparable or even larger surface area from the solvent (Velazquez-Campoy *et al.,* 2001b).

New drug design strategies by calorimetric characterization have permitted the designers to recognize the nature of forces by which the HIV-1 protease inhibitors bind the target primarily because these forces originate from different interactions. ITC was particularly crucial at later stages, since it gave detailed description of the thermodynamic factors governing protein-inhibitor interactions essential for molecular recognition in HIV-1 protease binding and led to improvement in drug design. This task was also facilitated by structure-based algorithms able to predict the enthalpic and entropic consequences of introducing different functional groups in the lead compounds under investigation (Luque and Freire, 2002; Vega *et al.,* 2004).

Extensive studies using numerous techniques of molecular biology and the deepened understanding of drug-target at the molecular level have helped greatly in achieving rapid success in the area of drug development, especially in the treatment of AIDS (DeVita *et al.,* 1987; Fitzgerald and Springer, 1991; Kohlstaedt *et al.,* 1992; Martin, 1992; Fitzgerald, 1993; Jacobo-Molina *et al.,* 1993; Wlodawer and Erickson, 1993; Dyda *et al.,* 1994; Bujacz *et al.,* 1995; Winslow and Otto, 1995; Bujacz *et al.,* 1996; Vacca and Condra, 1997; Wlodawer and Vondrasek, 1998).

1.5 HIV and AIDS

The tremendous effort over the past decade to develop antiretroviral therapeutic targets for acquired immune deficiency syndrome (AIDS) has resulted in the discovery of drugs that have made important breakthroughs in the management and treatment of this chronic disease. AIDS remains a serious health problem, especially in the developing countries where it has reached dramatic proportions (Cohen, 2000). Of the 40 million people infected with the human immunodeficiency virus (HIV) worldwide, 28 million (70 %) are in Africa. HIV, the causative agent of AIDS (Weiss *et al.*, 1985; Gallo and Montagnier, 1988) is a highly mutable *Lentivirus*, a member of the retrovirus family *Retroviridae*, whose genome is encoded in RNA. HIV is a lentivirus, since it usually has a disease latency of several years, i.e., it is characterized by the slow development of disease after infection (White and Fenner, 1994). The mature HIV virion is almost spherical and is about one ten-thousandth of a millimeter across (ca. 100 nm) (Greene, 1993). The virus is enveloped by a lipid bilayer that is derived from the infected host cell. The outer surface is studded with surface glycoproteins (gp120) that are anchored to the virus via interactions with the transmembrane protein (gp41). These surface proteins play a pivotal role when HIV binds to and enters the host cells (see below).

HIV infects cells that carry the $CD4^+$ antigen on their surface (see Figure 1 for HIV life cycle). The infection of the viral and cellular membranes is a process that is mediated by the viral envelope glycoproteins (gp120, gp41) and receptors ($CD4^+$ and co-receptors, such as CCR5 or CXCR4) on the target cell. Membrane fusion is followed by a poorly understood uncoating event of the capsid that allows the release of the viral content into the host-cell cytosol. As the virus enters a cell, its single-stranded RNA is reverse-transcribed to yield a double-stranded DNA molecule by a virally encoded enzyme, the reverse transcriptase. The viral DNA enters the cell nucleus, where it is permanently integrated into the genetic material of the host cell by the catalytic activity of a second virally encoded enzyme, the integrase (IN). Activation of host cell results in the transcription of the viral DNA into messenger RNA, which is then translated into viral proteins (e.g. Tat, Rev, and Nef regulatory proteins and Gag and Gag-Pol polyproteins). HIV protease, the third virally encoded enzyme, is required in this step to cleave a viral polyprotein precursor (Gag and Gag-Pol) into mature functional enzymes and structural proteins.

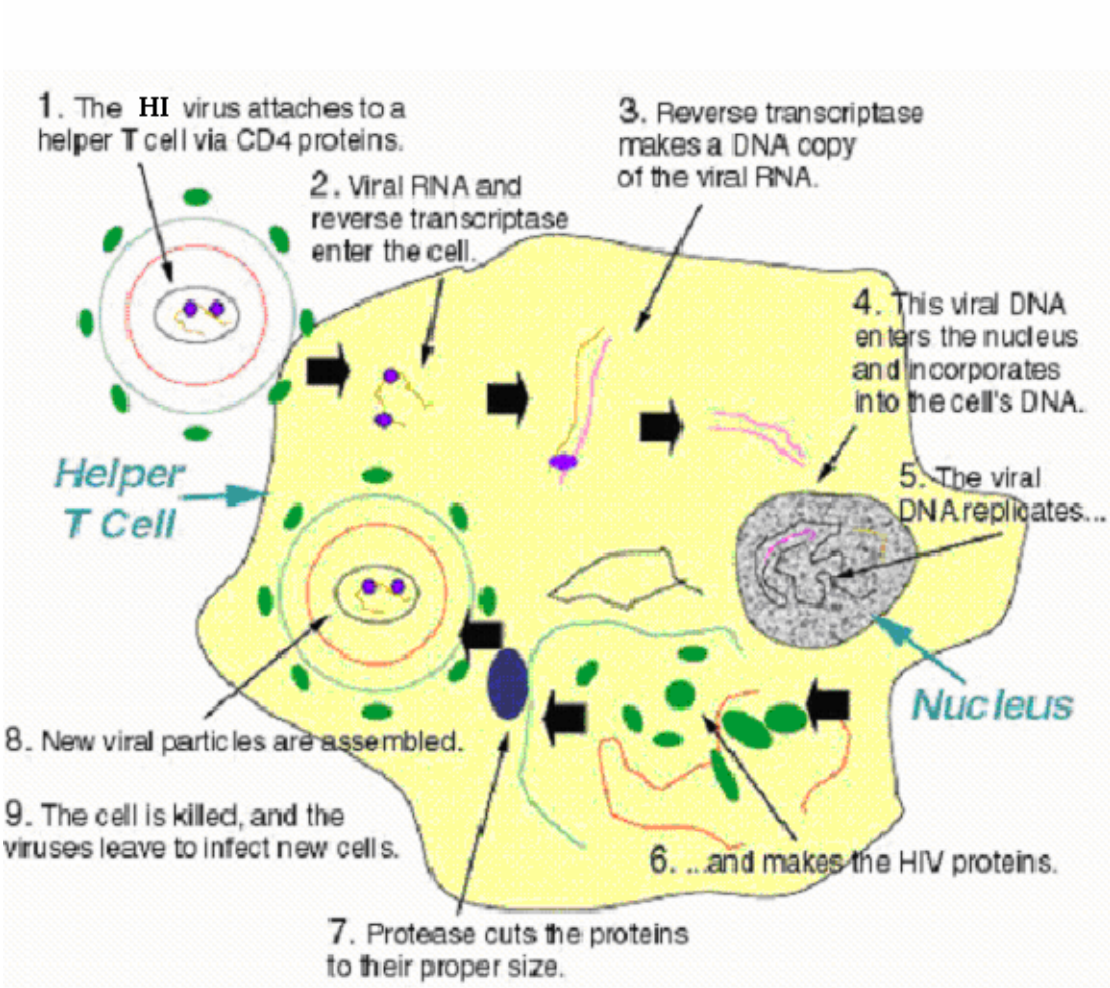


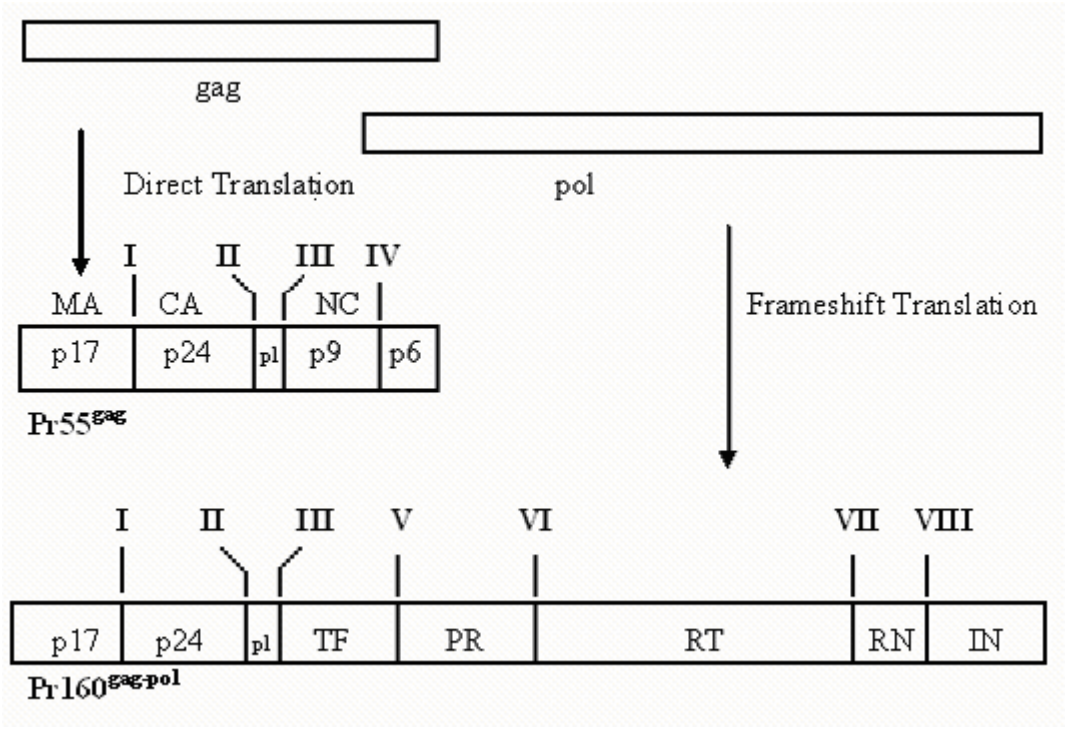
Figure 1. Life cycle of HIV. (1) HIV infects T cells that carry the $CD4^+$ antigen on their surface. (2, 3) As the virus enters a cell, its RNA is reverse-transcribed to DNA by a virally encoded enzyme, the reverse transcriptase. (4, 5) The viral DNA enters the cell nucleus, where it is integrated into the genetic material of the cell by a second virally encoded enzyme, the integrase. (6) Activation of host cell results in the transcription of the viral DNA into messenger RNA, which is then translated into viral proteins. (7) HIV protease, the third virally encoded enzyme, is required in this step to cleave a viral polyprotein precursor into individual mature proteins. (8) The viral RNA and viral proteins assemble at the cell surface into new virions, which then bud from the cell and are released to infect another cell. (9) The cell damage from the destruction of host's genetic system to the budding and release of virions leads to the death of the infected cells. The figure was adapted from Mugnaini *et al.*, 2005.

The viral RNA and viral proteins assemble at the cell surface into new virions, which then bud from the cell and are released to infect another cell and start a new replicative cycle. The extensive cell damage from the destruction of host's genetic system to the budding and release of virions leads to the death of the infected cells.

HIV has a wide range of genetic variation among distinct types, groups and clades. The two major distinct types of HIV, HIV-1 and HIV-2, are distinguished by their genomic organization and phylogenetic relationship. HIV-2 is, however, less widespread than HIV-1. Inspection of different strains of HIV-1 originating from diverse geographical locations, reveals that isolates can be subdivided into groups, subtypes, sub-subtypes and circulating recombinant forms (CRFs) according to phylogenetic sequence differences (Osmanov *et al.*, 2000). Groups refer to distinctive HIV-1 lineages, M (for Major), O (for Outlier) and N (for New, or non-M, non-O). Furthermore, the majority of strains found worldwide belong to group M. In addition, group M is further subdivided into nine subtypes namely: A, B, C, D, and F, G, H, J and K, as well as 14 CRFs (Osmanov *et al.*, 2002). Genomic differences between subtypes and recombinant forms can be as high as 30 % in the *env* gene and 15 % in the *pol* gene, which includes the coding region for protease (PR) and reverse transcriptase (RT).

HIV contains three open reading frames (ORFs), group-specific antigen (*gag*), *pol*, and *env* typical of other retroviruses. The gag ORF encodes the structural proteins that ultimately form the virus capsid, nucleocapsid and the matrix, the *pol* ORF encodes three essential retroviral enzymes (RT, PR, and IN), which have all become targets for structure-assisted "rational" drug design and discovery, and the env ORF, on the other hand, encodes the envelope surface proteins (Weiss *et al.*, 1985). The *gag* (p55) and *gag*/*pol* (p160) genes are expressed in a polyprotein format, and HIV protease is required to cleave this high molecular weight precursor, proteolytically at specific sites, into individual mature viral enzymes; the RT, IN and PR and functional core proteins (p6, p7, p17 and p24) (see Figure 2). The p17 lines the inner surface of the viral membrane, and a canonical capsid core particle constructed out of the capsid protein (p24) is located in the center of the virus. The capsid particle encapsulates two copies of the viral genome, stabilized by the nucleocapsid protein (p7) and also contains RT, IN and PR.

A.



B.

I	p17-p24	...SQNY↓PIVQ...
II	p24-p1	...ARVL↓AEAM...
III	p(24-p1)-p9	...ATIM↓MQRG...
IV	p9-p6	...PGNF↓LQSR...
V	TF-PR	...SFNF↓PQIT...
VI	PR-RT	...TLNF↓PISP...
VII	RT-RN	...AETF↓YVDG...
VIII	(RT-RN)-IN	...RKIL↓FLDG...

Figure 2. (**A**) Schematic representation of the HIV-1 gag and pol reading frames and their polyprotein translation products. Gag, group specific antigen; pol, polymerase; MA or p17, matrix protein; CA or p24, capsid protein; NC or p9, nucleocapsid protein; PR, protease; TF, transmembrane protein; RT, reverse transcriptase; RN, ribonuclease. (**B**) Processing sites in the gag and gag/pol by the HIV-1 protease. Adapted from Tomasselli and Heinrikson, 2000.

The HIV protease was first suggested as a potential target for AIDS therapy after it was shown that a frameshift mutation in the protease region of the pol-gene prevented cleavage of the Gag polyprotein precursor, which as stated before, is essential for the maturation of infectious HIV particles (Kramer *et al.,* 1986). Subsequently, inhibition of HIV polyprotein processing through the inactivation of the viral protease (either by mutation or chemical inhibition) leads to the production of budded, immature viral particles that cannot undergo maturation to an infective form (Kohl *et al.,* 1988; Schatz *et al.,* 1991). It is not surprising, then that HIV-1 protease was identified almost two decades ago as the prime target for structure-assisted or rational drug design. Compounds, having the ability to inhibit the protease have been studied intensively during the last decade and numerous reports of potent HIV-1 protease inhibitors have been published (see below and section 1.8).

The first drugs to have approval from the US Food and Drug Administration (FDA) were developed as nucleoside analogs; i.e., pseudo substrates designed to impede reverse transcription. RT, an enzyme unique to retroviruses and an important therapeutic target in AIDS treatment, transcribes retroviral RNA into DNA, which is then integrated into the host genome and expressed by the infected host cell under the control of viral genes. These drugs work by terminating further insertion of nucleotides into the elongating viral DNA chain carried out by the RT. The compounds of this class of drugs are: zidovudine (ZDV, or AZT), didanosine (ddI), stavudine (d4T), zalcitabine (ddc), lamivudine (3TC), abacavir (ABC), entricitabine (FTC), tenofovir (TDF) and zalcitabine (ddC). The success of these drugs in combating AIDS confirmed the importance of inhibiting RT and consequently led to the development and clinical use of another class of RT inhibitors that are not nucleosides, the non-nucleoside reverse transcriptase inhibitors (NNRTI). This class includes: nevirapine (NVP), delavirdine (DLV) and efavirnez (EFV). NNRTI mode of action involves inhibition of the reverse transcriptase by binding to an allosteric site of the enzyme.

The HIV-1 protease, therefore, represents another therapeutic target in AIDS treatment that is distinct in mechanism from RT. To date, efforts undertaken by scientists across a wide spectrum of disciplines have developed another distinct mechanistic class of protease inhibitors (PI) in the fight against AIDS. This class

includes: amprenavir (APV), atazanavir (ATV), darunavir (TMC-114), indinavir (IDV), lopinavir (LPV), nelfinavir (NFV), ritonavir (RTV), saquinavir (SQV) and tipranavir (TPV). Unlike RT inhibitors, which are only effective in blocking viral replication in acutely infected cells, protease inhibitors can inhibit infectious virus production in both acutely and chronically infected cells (Lambert *et al.,* 1992). The rational design of the PIs is to block the action of the PR, and taken together their synergy with the RT as well as with the recently approved entry inhibitors (EI) has brought new and efficient combination of highly-active antiretroviral therapies (HAART) for the treatment of AIDS. Although serious side effects and the development of resistance are still major unsolved problems, combination therapy involving members from the PR and RT inhibitors has transformed the once deadly disease into a treatable medical condition in many cases resulting in dramatic reductions in the viral load, and increases in $CD4^+$ cell counts (Collier *et al.,* 1996; Hammer *et al.,* 1997; Carpenter *et al.,* 1998). These effects have translated in the restoration of the functioning of the immune system and in the reduction, at least temporarily, of the morbidity and mortality caused by HIV infection. The early breakthrough in the development of HIV-PR inhibitors as drugs was based upon early recognition of the protease central role in the virus life cycle (see Figure 1) and an intensive effort was made toward gaining a comprehensive understanding of the protein's structure and function.

1.6 Structure of HIV-1 protease

Comparison of the genomic sequence of HIV-1 with that of several retroviruses, urged Ratner *et al.* (1985) to postulate that the HIV-1 genome encodes a protease. Subsequently, the availability of genomic information for HIV persuaded most researchers to consider the existence of an aspartyl protease encoded within the *pol* region of the viral genome, with a characteristic Asp-Thr-Gly active site signature sequence (Toh *et al.,* 1985). HIV-1 protease was found to be inhibited *in vitro* by pepstatin (Seelmeier *et al.,* 1988; Hansen *et al.,* 1988; Darke *et al.,* 1989), a general inhibitor of aspartic proteinases, whose mode of action operationally defines members of this class of enzymes (Umezawa *et al.,* 1970). Aspartic proteinases comprise a well-characterized family of widely distributed enzymes found in vertebrates, fungi, plants and retroviruses. Most of them are secreted as zymogens, inactive precursors, that are converted to the active enzymes by proteolytic cleavage of the propeptide at

the N terminus of the polypeptide chain. They are characterized by a low pH optimum and specificity for extended peptides (Fruton, 1976). The single protease encoded in the genome of HIV-1 was initially classified as an aspartic proteinase on the basis of the putative active-site homology to cellular aspartyl proteases (Toh *et al.*, 1985), the inhibition of the enzyme by pepstatin, the inactivation of the protease activity by mutagenesis of the highly conserved active site Asp-25 residue (Seelmeier *et al.*, 1988) and the determination of its crystal structure by Navia *et al.* in 1989.

PR is a 99-amino acid, symmetrical and obligate homodimer that is structurally similar to other aspartyl proteases of the pepsin family (Navia *et al.*, 1989; Wlodawer *et al.*, 1989), as well as those from other retroviruses including: Rous sarcoma virus (RSV) (Miller *et al.*, 1989), HIV-2 (Mulichak *et al.*, 1993; Tong *et al.*, 1993), simian immunodeficiency virus (SIV) (Rose *et al.*, 1993a; Zhao *et al.*, 1993), feline immunodeficiency virus (FIV) (Wlodawer *et al.*, 1995) and equine infectious anemia virus (EIAV) (Gustchina *et al.*, 1996). To date, more than 200 crystal structures of HIV-PR and of HIV-PR/inhibitor complexes have been determined. A comprehensive database providing structural information about the HIV-PR is available at the National Cancer Institute http://www.fbsc.ncifcrf.gov/HIVdb and http://srdata.nist.gov/hivdb. The homodimeric enzyme complexed with acetyl-pepstatin, the peptidic inhibitor, is depicted in Figure 3. Each monomer contains an extended β-sheet region (a glycine-rich loop) known as the flap region and one of the two catalytic aspartyl residues (Asp-25) which lie at the bottom of the cavity. The flap region contributes, in part, to the build-up of the substrate binding site and plays a crucial role in substrate binding and possibly in the exclusion of water from the catalytic site. Since each monomer contributes one of the catalytic aspartic residues, the monomers are catalytically inert and dissociation of the enzyme homodimer results in a complete loss of catalytic activity (Seelmeier *et al.*, 1988). A regulatory mechanism that controls activation the enzyme is derived from the dimerization process, since it is reversible. In a concentrated environment (e.g. in the budded virion), the protease is activated, and in a highly dilute environment (e.g. in the host cell), the protease is inactivated. This regulatory mechanism seems to be important for the virus in order to prevent premature breakdown of the polyproteins and to minimize damage of the host-cellular proteins (Tomasselli and Heinrikson, 2000).

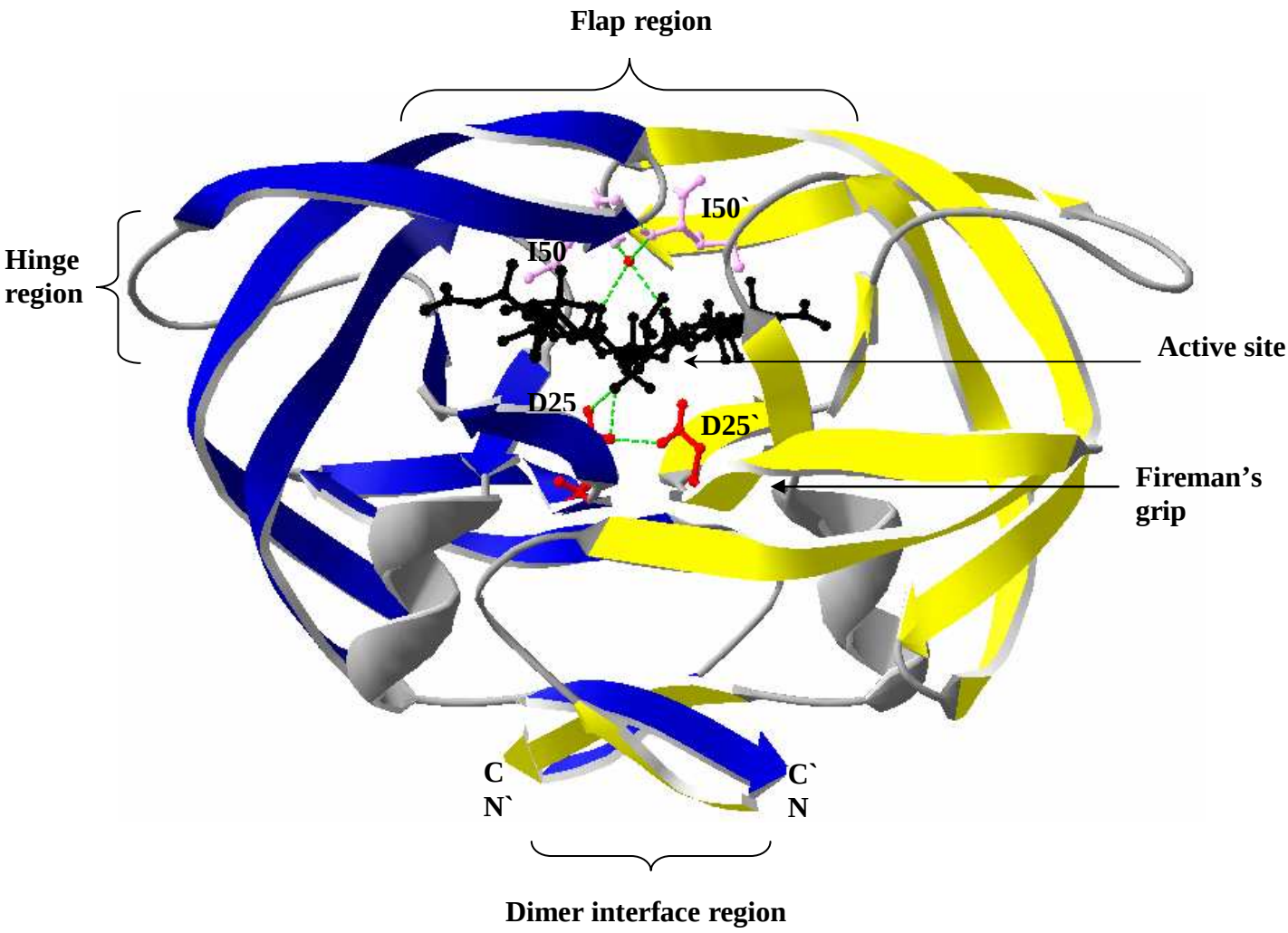


Figure 3. A ribbon diagram of the crystal structure of the homodimeric HIV-1 subtype B protease in complex with acetyl-pepstatin represented as ball-and-stick at the active site (Fitzgerald *et al.*, 1990). The pdb code for this structure is 5hvp. Each monomer contributes one catalytic aspartate (Asp-25) located on the floor of the substrate binding pocket. The carboxylate groups of Asp-25 from both chains are nearly co-planar and show close contacts. The conserved water molecule mediating the contacts between the P2/P1` carbonyl O atoms of the peptidic inhibitor and the amide groups of Ile50/Ile50` in the enzyme is also shown. This water molecule is tetrahedrally coordinated between the enzyme and inhibitor and is completely inaccessible to the bulk solvent. The dimer interface is formed by the N- and C-termini of each monomer forming a four-stranded, antiparallel β-sheet. The dimer interface is characterized by the presence of clusters of residues that contribute significantly to the subunit association. The protease monomers are shown as blue and yellow ribbons. The monomers interact with each other and the substrate in the flap region upon substrate binding. This figure was generated using Swiss-pdb Viewer (Guex and Peitsch, 1997).

The flaps are held in a closed position by hydrogen bonding from the flap residues Ile50/50` to a conserved water molecule, which in turn is hydrogen bonded to two carbonyl groups in the inhibitor/substrate peptides (see Figure 3). Comparison of unliganded wild-type protease crystal structures with its complexes with inhibitors shows a 7 -15 Å movement of the flap tips (around residues 50/50`) to allow the polyprotein substrate or inhibitor to enter the active site (Miller *et al.,* 1989b; Gustchina and Weber, 1990) (Figure 4). The substrate binds in an extended conformation and its interactions with different amino acid side chains determine the specificity of the enzyme for the substrate (Pettit *et al.,* 1993). The HIV-1 protease processes the Gag and Gag-Pol polyproteins proteolytically at specific sites as shown previously in Figure 2. The HIV-1 protease is specific for cleavage of these sites *in vivo*, although the general sequence homologies among these are small (Fitzgerald and Springer, 1991; Moore and Dreyer, 1993).

The dimer interface is formed by a four-stranded antiparallel β-sheet involving both the amino and the carboxyl-termini of each subunit. Upon substrate binding, both subunits form a long cleft where the catalytically essential aspartic acid residues are located in a co-planar orientation at the floor of the cleft. The main interface is formed by the N- and C-termini from each monomer (residues 1 – 4 and 96 – 99) joining to form a four stranded, antiparallel β-sheet. The dimers also interact in the flap region formed by residues 45-55 (Navia *et al.,* 1989; Wlodawer *et al.,* 1989). These flaps interact with each other and with substrate upon binding. The monomers become more closely packed together upon substrate binding, followed by several regions of the molecule moving inward toward the bound ligand. Subsequent to ligand or substrate binding, the protease dimer loses symmetry, based on the superimposition of the α-carbons of each monomer in the liganded and apoenzyme structures (Rose *et al.,* 1996). The presence of bound substrates also induces structural changes on both monomers that contribute in part to the stability of the HIV protease dimer. The result of these protein-ligand interactions is an extremely tight binding (high affinity) between individual subunits in an enzyme-substrate complex (low nanomolar range), that may depend greatly on the ligand concentration in addition to salt concentration and pH (Zhang *et al.,* 1991).

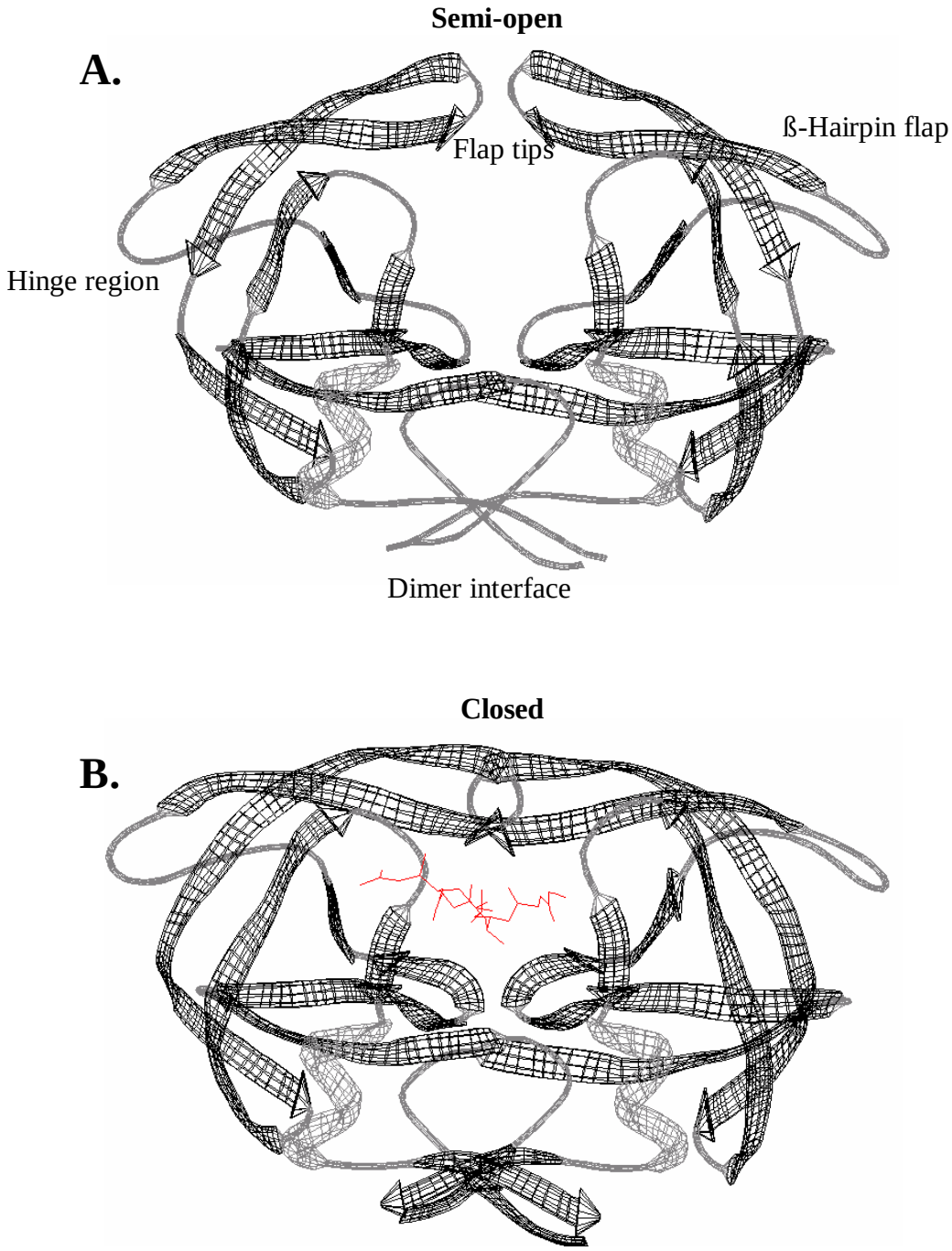


Figure 4. (**A**) Ribbon representations of the ligand-free homodimeric HIV-1 protease and (**B**) in complex with acetyl-pepstatin in the active site. The pdb code for the ligand-free structure is 1hhp. The pdb code for the HIV-1 protease in complex with acetyl-pepstatin is 5hvp. The structure of the unbound protein adopts a 'semi-open' conformation with the flaps shifted away from the active site, whereas in liganded form, the flaps are pulled in towards the bottom of the active site (i.e., the 'closed' form). This figure was generated using Swiss-pdb Viewer (Guex and Peitsch, 1997).

In addition, the proteases of both HIV-1 and HIV-2 contain cysteine residues at position 95 (Cys-95) in the dimer interface region. Structurally related proteases of other retroviruses including SIV, RSV and FIV also have cysteine or methionine residues at their dimer interfaces. HIV-1 has an additional solvent accessible cysteine residue at position 67 (Cys67) which is highly conserved among virus isolates. These two cysteine residues (Cys67 and Cys95) are proposed to provide a mechanism of regulation for HIV activity in response to oxidative stress within the host cell (Davis *et al.*, 1996; Davis *et al.*, 1999).

1.7 Catalytic mechanism of the HIV protease

Proteases, in general, are known to play crucial roles in various biological processes. They catalyze the hydrolysis of peptidic bonds with high sequence selectivity and catalytic proficiency. Two mechanisms of protease activity have been established or proposed thus dividing protease enzymes mechanistically into two different classes. The first class uses an activated water molecule as a nucleophile to attack the amide bond carbonyl carbon of the substrate scissile bond. The activation of this water molecule can be attained either by a zinc cation (the zinc metallo-proteinases) or by the aspartyl β-carboxylate group at the active site (in the case of aspartate proteases e.g., HIV-1 protease). In the second class of proteases, a nucleophilic atom of an amino acid side chain is used to hydrolyze the amide bond. By convention, the peptide bond that is hydrolyzed is referred to as the scissile bond or cleavage site and is positioned between P1 and P1` (Schechter and Berger, 1967).

The flanking amino acids proceeding, toward the amino terminus, are designated P1, P2, P3 and those going toward the carboxyl terminus are referred to as the P1`, P2`, P3` and provide space for water molecules necessary for substrate recognition and/or product release (Jacks *et al.*, 1998; Prabu-Jeyabalan *et al.*, 2002) (see Figure 5). The subsites of the enzyme interacting non-covalently with the corresponding side chains of the peptide (substrate or inhibitor) are termed, starting from the catalytic aspartic acid residues, S1, S2, S3 and S1`, S2`, S3`, respectively. Using the above standard nomenclature, the S1 and S1` (similarly S2 and S2` etc.) subsites are located at equivalent positions (see Figure 5). The two S1 subsites are highly hydrophobic with the exception of the active site aspartates. The residues which make contacts with the

P1/P1` side chains of the substrates or inhibitors include, Arg8, Leu23, Asp25, Gly27, Gly48, Gly49, Ile50, Thr80, Thr81 and Val82. The S2/S2` subsites are mostly hydrophobic (Ala28/28`, Val23/23`, Ile47/47`, Gly49/49`, Ile50/50`, Leu76/76` and Ile84/84`) except Asp29, Asp29`, Asp30 and Asp30`. The interior pockets, the S2 and S2` subsites, are smaller than the S1/S1` or the S3/S3` binding sites and have been shown to be more specific, restricting the type and size of residues at P2/P2` in substrates or inhibitors relative to other binding pockets in the protease molecule (Tözsér *et al.,* 1992; Dunn *et al.,* 1994). As such, these characteristics have been implicated for the lower tolerance in the drug-resistant mutants of the residues forming these binding sites (Rosin *et al.,* 1999).

Dreyer *et al.* (1992) have also conclusively shown that substrate specificity of the HIV-1 protease is significantly determined by S2-S2` subsites. Interestingly, all protease inhibitors are shorter than the natural substrates and contain hydrophobic moieties that interact primarily with the protease S2-S2` subsites. Thus, although protease inhibitors may be chemically different from each other, they occupy a similar space within the protease binding pocket, which explains how individual mutations may cause protease cross-resistance (see section 1.8). The S3 subsites adjacent to S1 subsites are also mostly hydrophobic and are known to have a rather broad specificity due to their ability to accept residues of different types and sizes (Wlodawer and Vondrasek, 1993). Residues forming the S3/S3` subsites include Arg8/8`, Leu23/23`, Asp29/29`, Gly48/48`, Gly49/49`, Pro81/81` and Val82/82` which form parts of the S1/S1` pockets. On the other hand, details on the S4/S4` and S5/S5` subsites are still sketchy due to the limited number of available structures of retroviral proteases in complex with ligands that extend beyond P3/P3` (Tözsér *et al.,* 1991; Vondrasek *et al.,* 1997). It has also been shown that the HIV protease most efficiently cleaves peptide substrates seven or eight amino acids long (P4-P3`) with the major processing subsites (S4-S3`).

It is generally accepted that, in the first step, the nucleophilic atom which can be a hydroxyl group or a thiol is subsequently activated by another amino acid side chain (Figure 6). The activated nucleophile attacks the carbonyl carbon of the scissile amide bond to form an ester or a thioester acyl intermediate.

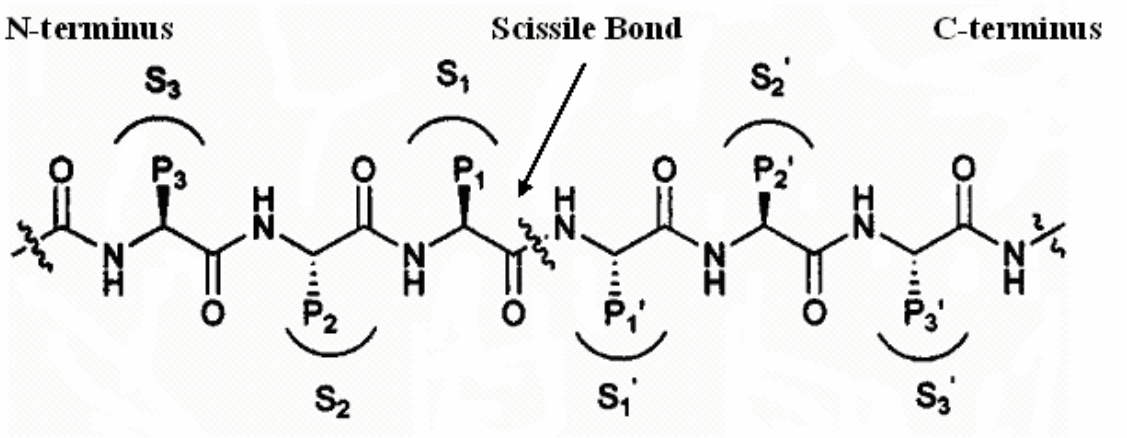


Figure 5. A schematic diagram showing the conventional nomenclature used to assign amino acid residues of peptide substrates (P) as well as the corresponding binding sites (S) on the protease molecule. This diagram was adapted from Wlodawer and Vondrasek, 1998.

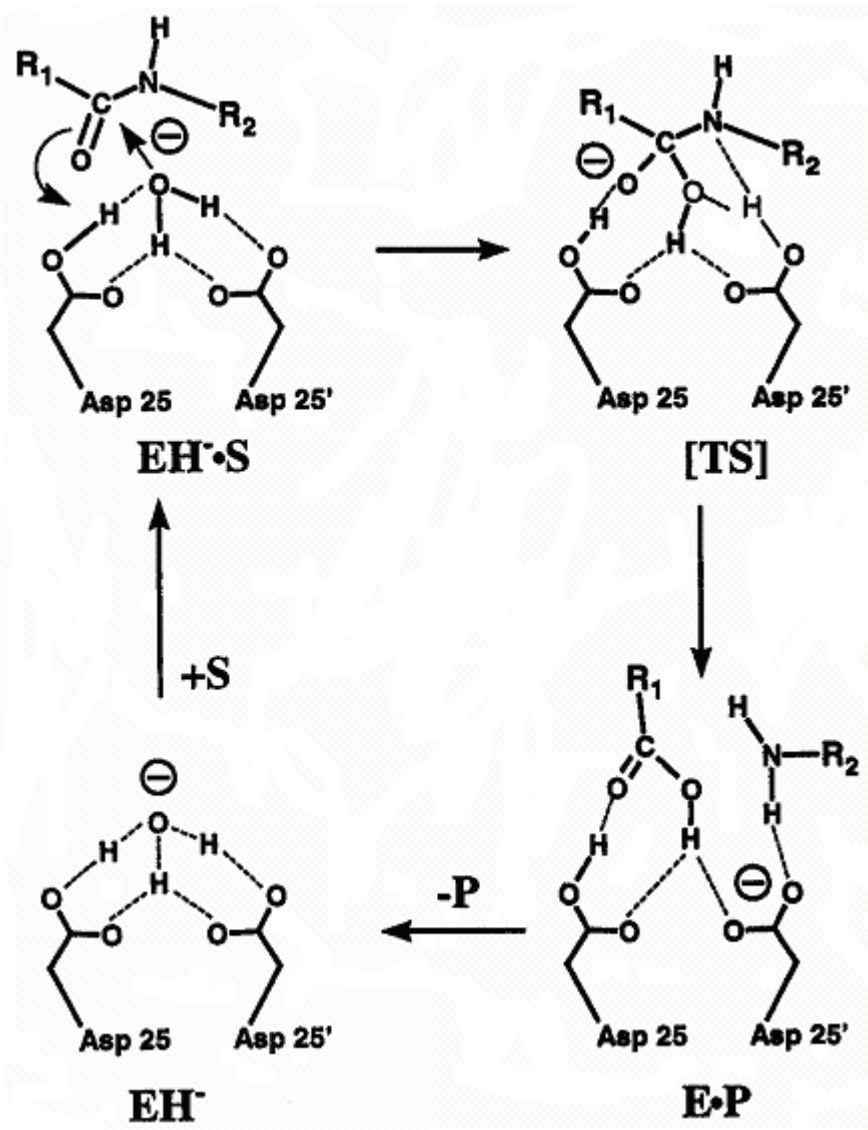


Figure 6. Schematic diagram showing the proposed kinetic mechanism of HIV protease. The two catalytic Asp residues are denoted 25 and 25` to identify their respective locations in individual subunits. The carboxylates activate the conserved water molecule, which then acts as a nucleophile for the cleavage of the peptide bond in the E•S complex. This diagram was adapted from Tomasselli and Heinrikson, 2000.

This acyl enzyme intermediate is then eventually hydrolyzed by a water molecule to the corresponding hydrolysis products. In the proposed mechanism of the aspartic protease, the Asp group that is closer to the nucleophilic water molecule is assigned a negative charge (Figure 6). The nucleophilic water molecule which is activated by the negative aspartate side chain is held between the two catalytic aspartate side chains and then attacks the carbonyl group in the substrate's scissile bond to generate an oxyanion tetrahedral intermediate. Protonation of the scissile amide N atom and rearrangement result in the breakdown of the tetrahedral intermediate to the hydrolysis products. Although the HIV PR catalytic mechanism is more or less similar to that of the aspartic protease family, its full detailed mechanism still remains elusive. Nonetheless, Jaskolski *et al.* (1991) have proposed a new model of the enzymatic mechanism for HIV PR enzyme based on the crystal structure of a complex between HIV PR and an octapeptide inhibitor U-85548e (H-Val-Ser-Gln-Asn-Leu-ψ-[CH(OH)-CH_2-Val-Ile-Val-OH). In this mechanism, the catalytic reaction is viewed as a one-step process during which the attack of the scissile bond by the nucleophilic water molecule and the acidic proton is achieved in a concerted manner.

1.8 Inhibition and development of drug resistance

A wealth of information and knowledge accumulated over the last two decades about the structure and the mechanism of HIV protease has indeed paved the way towards the development of effective drugs targeted against this enzyme. The effort of designing drugs with high efficacy towards HIV protease has to a large extent been aided by a huge amount of information regarding the design of inhibitors for other aspartic proteases, such as renin. The renin inhibitors were substrate-based peptidomimetic compounds containing non-cleavable inserts as replacements for the scissile-bond amino acid partner. However, extensive modifications were necessary to reduce the peptidic character and improve oral bioavailability. This *de novo* approach has been invaluable in accelerating the discovery of drugs that are effective against HIV protease. To date, there are nine FDA approved drugs that function as inhibitors of HIV protease: saquinavir, ritonavir, indinavir, nelfinavir, amprenavir, lopinavir, atazanavir, tipranavir and darunavir (see Figures 7, 8 and 9 for chemical structures).

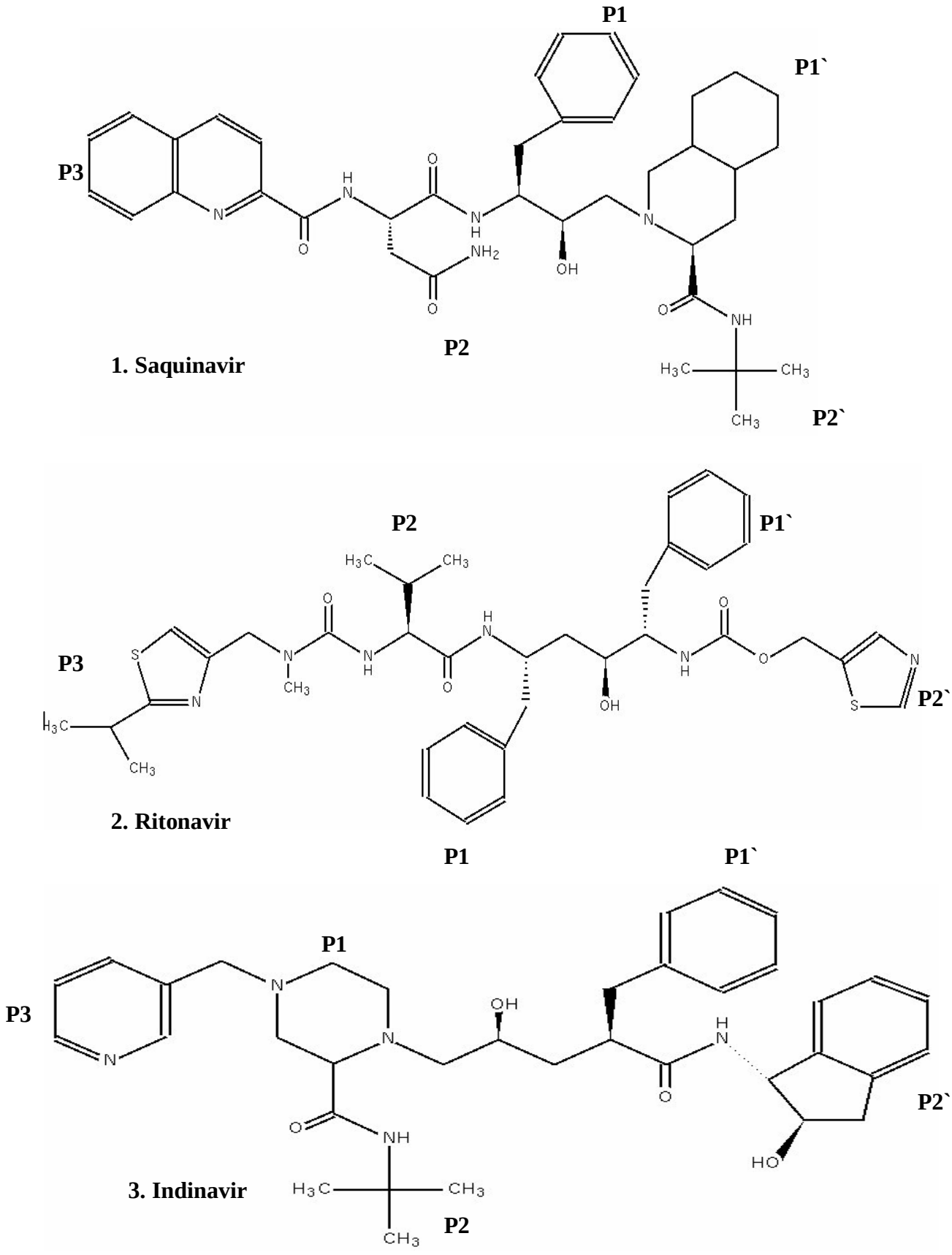
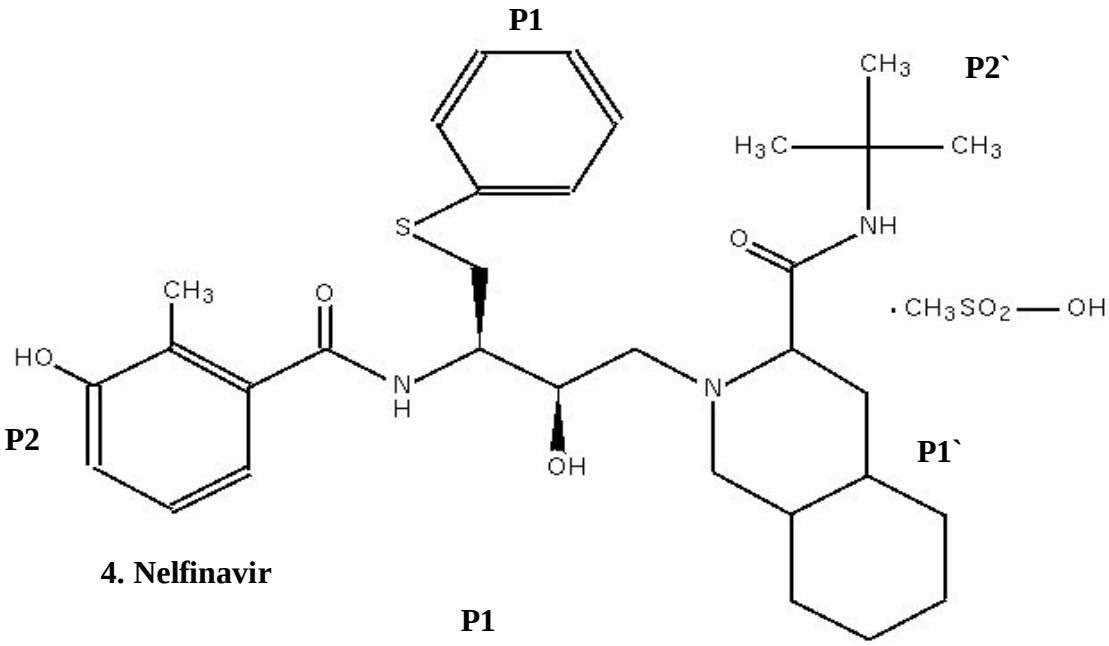
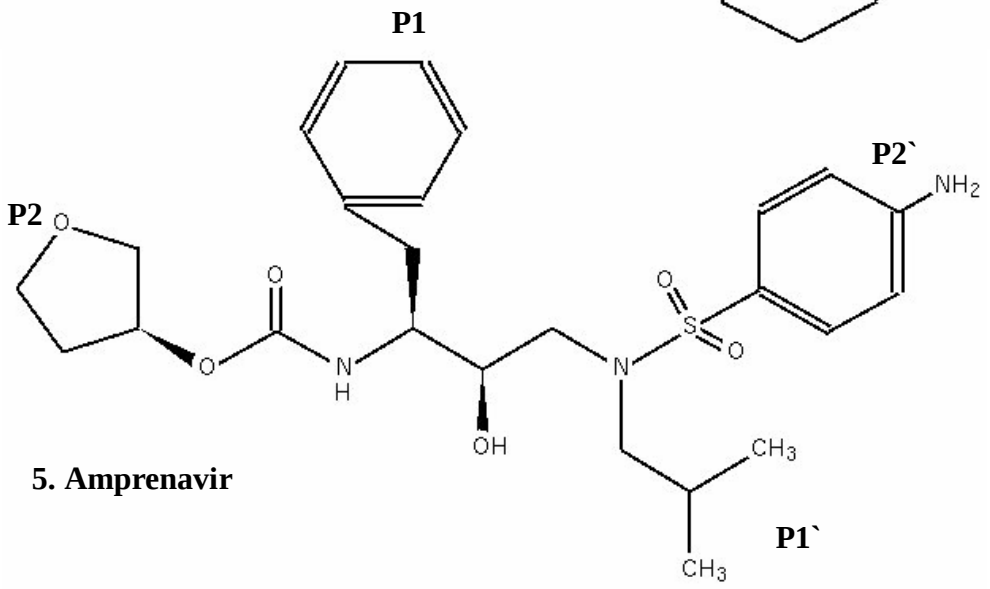


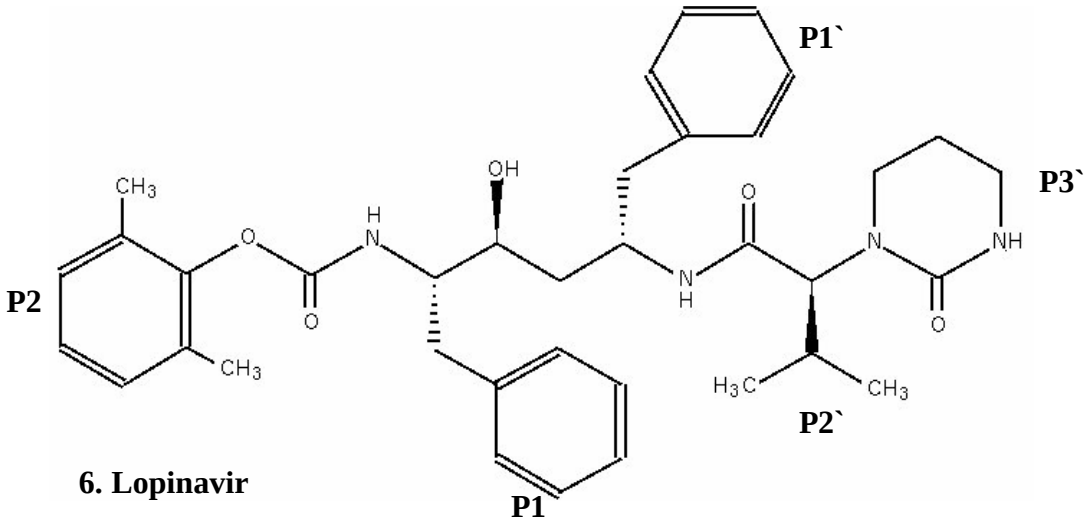
Figure 7. Chemical structures of saquinavir, ritonavir and indinavir showing the positions occupied by the interacting substituents.



4. Nelfinavir



5. Amprenavir



6. Lopinavir

Figure 8. Chemical structures of nelfinavir, amprenavir and lopinavir showing the positions occupied by the interacting substituents.

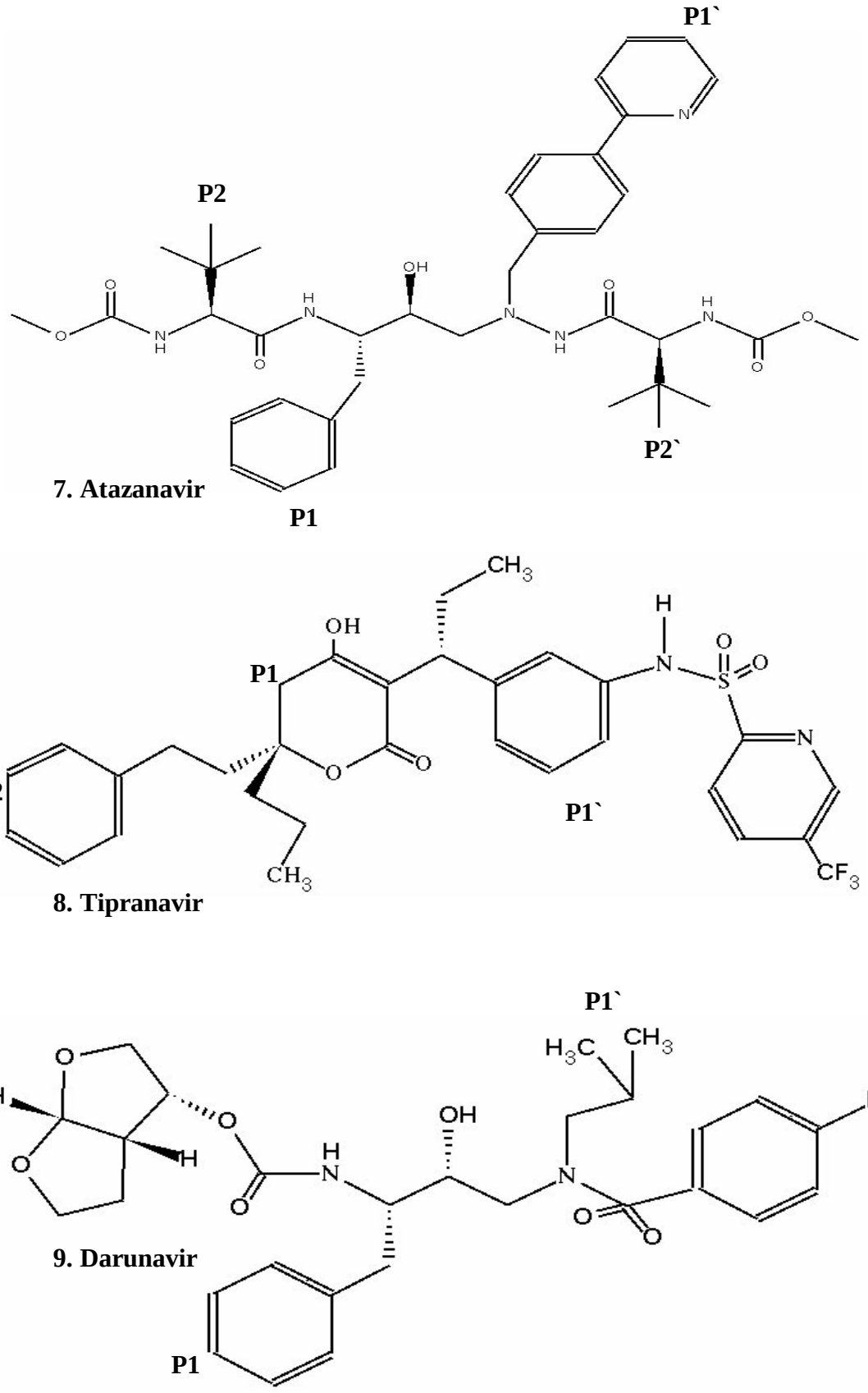


Figure 9. Chemical structures of atazanavir, tipranavir and darunavir showing the positions occupied by the interacting substituents.

HIV-1 protease inhibitors have been effective in suppressing viral levels in patients with AIDS. However, their therapeutic efficacy is hampered by the emergence of drug-resistant, cross-resistant mutants of the HIV-1 protease and by issues of patient compliance, rendering AIDS with no definitive cure. The high replication rate of the virus (10^8-10^9 virions/day) and the high error rate of HIV RT (about 1 in 10 000 bases) have been implicated in the rapid mutation and selection of drug-resistant viruses (Ji and Loeb, 1992; Coffin, 1995; Hertogs *et al.*, 2000; Wu, *et al.*, 2003). High level resistance can only develop if the virus is allowed to replicate in the presence of the drug. Thus, the occurrence of the mutations is not a direct consequence of drug action *per se*. Instead mutations, result from the viral replication process itself, and antiviral drug resistance is selected through the elimination of the most drug-susceptible viruses from the replicating pool, leaving only the fittest viruses to survive in its presence (Condra, 1998). In fact, because antiviral drugs inhibit viral replication, they inhibit the very processes responsible for the appearance of new mutations. Additionally, even in the absence of antiretroviral drugs, HIV-1 shows genetic diversity, especially in the protease gene (Vergne *et al.*, 2000). Analysis of clinical isolates has also revealed extensive protease mutations that confer cross-resistance to protease inhibitors (Hertogs *et al.*, 2000).

Drug resistant mutations of at least 27 of the 99 protease residues have been reported primarily from the B subtype of HIV-1. The loss of sensitivity usually occurs because the resistant viral strains encode for protease molecules containing specific amino acid mutations that lower the affinity for the protease inhibitors and yet still maintain a viable enzymatic profile and substrate affinity. In addition, more than 87 mutations have been observed in at least 49 positions within the HIV-1 protease monomer and shown to express resistance toward one or more inhibitors (Rose *et al.*, 1993b; Kozal *et al.*, 1996; Boden and Markowitz, 1998; Gulnik *et al.*, 1999; Hertogs *et al.*, 2000; Shafer *et al.*, 1999; Wu *et al.*, 2003). For specific inhibitors, these mutations are termed major or minor depending on whether they determine major phenotypic and genotypic resistance or whether they contribute as accessory mutations following the onset of resistance. These mutations have been classified as active site or primary, and non-active site or secondary mutations (Erickson and Burt, 1996). Active-site mutations are located directly in the binding site and usually do not involve residues that are directly involved in catalysis, for instance Asp25/25`, but involve residues

that have direct interactions with the inhibitors (see Table 1 and Figure 10). Non-active site mutations are usually located at the hinge of the flap region, the dimer interface and the β-sheet region (Ala *et al.*, 1998). Non-active site mutations may interfere with binding by (1) directly altering enzyme catalysis, dimer stability, inhibitor binding kinetics and conformational dynamics and (2) through long-range interactions that transmit distortions or perturbations in the active site rather than direct interactions with the inhibitors (see Table 1 and Figure 10) (Ala *et al.*, 1998; Shafer, 2002).

Drug resistance to HIV-1 protease inhibitors usually begins with the appearance of amino acid mutations within the binding cavity of the enzyme. These active site mutations are very conservative, i.e., they change the geometry of the active site without changing its polarity or charge. Non-active site mutations occur in later stages and have been assigned mostly compensatory functions aimed at neutralizing the detrimental effects of primary mutations on enzyme activity (Table 1) (Schock *et al.*, 1996; Velazquez-Campoy, *et al.*, 2001b; Shafer, 2002). Interestingly, compensatory mutations (M36I, M46I/L, L47V and L63P), located on the flap region or in close proximity, are thought to impose structural constraints on the flap elbows of the HIV-1 protease and cause a motion of the substrate towards the cleavage site, hence restoration of activity (Piana *et al.,* 2002). The mutations at positions 46 – 48, 50, 53 and 54 are the most commonly detected, and are usually accompanied by mutations at positions 82 and 84 in the inhibitor binding site. Combinations of these mutations arise during the treatment of infected patients with over three protease inhibitors leading to a production of multi-drug resistant variants of the HIV-1 protease.

Because active site mutations directly affect the drug/target interface, their mode of action has been interpreted in terms of lost interactions or steric effects created by altered site geometries (Velazquez-Campoy *et al.*, 2001b; Ohtaka *et al.*, 2002). The role of non-active site mutations in inhibitor binding has remained largely elusive, but a few recent studies have suggested that they affect the binding by altering the geometry of the binding site cavity through accumulation of mutations within the core of the protease molecule (Olsen *et al.,* 1999; Ohtaka *et al.,* 2003).

Table 1. Primary and secondary resistance mutations in HIV-1 subtype B protease associated with the development of resistance against currently approved protease inhibitors.

Primary	D30N, L33F, M46I, I47V, G48, I50V, V82A/F/T, I84V, L90M
Secondary	L10I/V, K20M/R, L24I, V32I, L33F, E35D, M36I, M46I/L, I47V, I54L/V, L63P, A71T/V, G73S, V77I, N88D, L90M

These data were taken from Tomasselli and Heinrikson, 2000

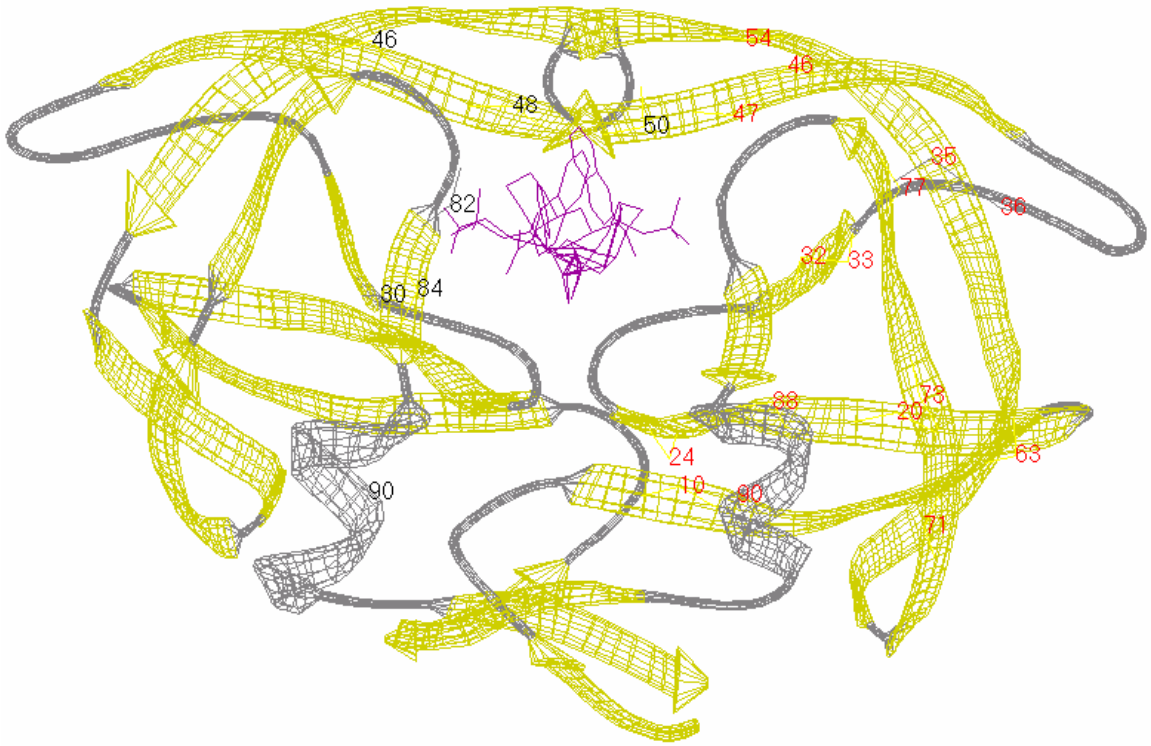


Figure 10. A ribbon representation of HIV-1 protease homodimer with positions of amino acid residues associated with clinical resistance to current protease inhibitors. The pdb code for this structure is 1hxb. Saquinavir, bound at the active site, is shown in purple. Primary and secondary mutations are indicated with black and red text, respectively. Mutations are shown on only one monomer for clarity. The figure was generated using Swiss-pdb Viewer (Guex and Peitsch, 1997).

Previous studies on the mutations occurring at distal positions strongly suggested that these changes likely play a compensatory role for other deleterious mutations (Schock *et al.,* 1996). But new insights by Muzammil *et al*. (2003), substantiate the notion that non-active site mutations are not only as a result of compensatory action; i.e., for activity restoration, but they in fact impact on ligand binding by either propagating distortions in the active site or by amplifying the effects of active site mutations, which affect inhibitor binding to the protein.

1.8.1 Saquinavir

Saquinavir, developed by Hoffmann-La Roche, was the first HIV PR to enter clinical trials, demonstrating that the protease enzyme was a valid target for the treatment of HIV infection (Jacobsen, 1994). Saquinavir is very effective as a protease inhibitor with a 90 % and 50 % inhibitory concentration (IC_{90} and IC_{50}) of 6 nM and 0.4 nM, respectively (Condra, 1996). Saquinavir is a hydroxyethylamine transition state analogue with a bulky decahydroquinoline moiety in P1` and quinoline in P3 (Figure 7). *In vitro*, the most frequent mutations or variants selected for by saquinavir are G48V and L90M and the less frequent G48V/L90M double mutant (Eberle *et al.,* 1995; Jacobsen *et al.,* 1995; Turriziani *et al.,* 1994). Amino acid residue 48 is located in the loop of the flexible flap of the protease, while residue 90 is found outside the binding pocket of the enzyme. Substitution at position 48 could result in a less flexible flap while the L90M mutation could induce conformational perturbations in the enzyme impacting on the binding of the inhibitor at the active site. The L90M mutation has also been shown to reduce catalytic activity and structural stability of the protease (Mahalingam *et al.,* 1999). Other mutations such as M36I, A71V/T and G73S have also been observed *in vitro* but at a lower frequency.

1.8.2 Ritonavir

Ritonavir, developed by Abbott Laboratories, was the second HIV protease inhibitor to be licensed in the United States. Ritonavir is a selective inhibitor of HIV protease that is derived from a C_2-symmetric, peptidomimetic inhibitor with the systematic modification of the heterocyclic P3 and P2` substituents (Ho, 1994). It is a potent antiretroviral agent *in vitro* with demonstrated activity against a variety of HIV-1 isolates, with IC_{50} and IC_{90} values of 10 nM and 70 nM, respectively and good oral bioavailability (Kempf *et al.,* 1995). Resistant variants associated with ritonavir

treatment are: K20R, L33F, M36I/L, M46L, I54V, A71V/T, V82A/F/T/S, I84V and L90M. Mutations at position 82, in particular, were highly associated with the initial loss of antiviral activity. However, this observation is not surprising since this mutation occurs within the immediate vicinity of the active site and could affect the protein-drug interactions.

Successful HAART relies on additive effects. Not long after the advent of the first antiretroviral drugs, it became clear that single agents used alone (monotherapy) could not suppress HIV over the long term, since the virus can mutate to develop drug resistance. Therefore, multiple agents from more than one drug class are now used to construct effective regimens. Additive and synergistic side effects are a major concern in anti-HIV therapy. When two or more drugs with overlapping toxicity profiles are used together, the combined toxicity may prove intolerable, even if the individual agents alone produce only mild side effects. For instance, ddC (zalcitabine, developed by Hoffmann-La Roche), ddI (didanosine, Bristol Myers-Squibb), and d4T (stavudine, Bristol Myers-Squibb), can all cause pancreatitis (inflammation of the pancreas), peripheral neuropathy (nerve damage), and mitochondrial toxicity. Thus, it is recommended that combinations of these drugs be avoided if possible.

Cytochrome P450 (CYP450) isoenzymes are classified into families with a small subset that carry out most of the drug metabolism. Many therapeutic agents are metabolized by a single isoenzyme, but others are processed by more than one. The most abundant isoenzyme, CYP3A4, is responsible for metabolizing about half of the drugs currently on the market. Drug metabolism is limited by the quantity of the CYP450 enzymes, and different agents compete for their use. Some CYP450 inhibitors known to retard the activity of these enzymes include ritonavir and erythromycin. Ritonavir works well as a pharmacological enhancer because it inhibits both the initial stages of protease inhibition metabolism in the intestines and CYP3A4 processing in the liver (Ernest *et al.*, 2005). The use of ritonavir with a protease inhibitor at therapeutic doses allows the achievement of a pharmacokinetic benefit that leads to higher drug plasma exposure, creating a higher genetic barrier to resistance (boosting). Boosting of protease inhibitors has become common practice and is recommended in treatment guidelines. Ritonavir's unique interaction profile allows it to be used to increase blood concentrations of other protease inhibitors. This

capacity is used by ritonavir as a low-dose (typically 100 mg) adjunct drug. At the same time, use of full-dose ritonavir has fallen out of favor because its propensity to raise levels of so many medications is more often harmful than beneficial. Low-dose ritonavir has also boosted the benefits of other first-generation protease inhibitors (e.g. indinavir and saquinavir), by dramatically reducing their high pill burdens, reducing intolerable side effects, loosening food requirements, and allowing them to be taken fewer times per day (Enerst *et al.,* 2005). Currently, protease inhibitor candidates are routinely tested in conjunction with ritonavir. Researchers are avidly searching for other drugs that can also be used as pharmacological boosters. A few of the approved protease inhibitors, including nelfinavir, have been linked to increased concentrations of other protease in some studies. But, to date none has demonstrated a strong enough or consistent enough effect compared to ritonavir.

1.8.3 Indinavir

The third HIV protease inhibitor licensed in the United States, indinavir, was developed by Merck & Co. Indinavir was derived as a peptidomimetic transition state analogue and belongs to the class of protease inhibitors known as hydroxyaminopentane amide (HAPA) compounds (Vacca *et al.,* 1994). This compound has several large, mainly hydrophobic groups that interact with the hydrophobic S2/S2` pockets in the active site of the protease and features a novel 2-aminoindanol substituent in P2`. It has a K_i value of 0.5 nM against wild-type HIV-1 protease and a good oral bioavailability (Dorsey *et al.,* 1994; Vacca *et al.,* 1994).

The data reported by Condra *et al.* (1995) from the early dose-ranging phase I/II clinical studies with indinavir demonstrated that the rapid development of resistance occurs at suboptimal doses of the drug. It was found that phenotypic resistance correlated with multiple substitutions among, at least, eleven sites in the protease: L10I/V/R, K20M/R, L24I, M46I/L, I54V/A, L63P, I64V, A71V/T, V82A/F/T, I84V and L90M (Condra, 1996). Interestingly, no single mutation was present in all resistant isolates. Also, there was no preferred order of appearance of any substitution, i.e., resistance was caused by the combined effect of multiple and variable combinations of amino acid alterations. Every viral isolate that conferred resistance to indinavir also conferred resistance to ritonavir, saquinavir and amprenavir.

1.8.4 Nelfinavir

Nelfinavir, developed by Agouron Pharmaceuticals, a selective nonpeptidic HIV-1 protease inhibitor was found to be a highly effective inhibitor with a K_i of 2 nM *in vitro* (Kaldor *et al.,* 1997). Nelfinavir has the same P1` moiety as saquinavir with an extended P1 substituent to increase lipophilicity (Kaldor *et al.,* 1997). The drug demonstrated antiviral activity with 50 % effective concentrations between 9 nM and 60 nM against HIV-1 and HIV-2 strains (Patick *et al.,* 1997). Sequence analysis of the clinical isolates from patients receiving nelfinavir monotherapy revealed a predominant D30N mutation in all phenotypically resistant variants. The D30N mutation had a 7-fold decrease in K_d compared to the wild-type. Other mutations were observed at much lower frequencies and these included: E35D, M36I, M46I, A71T/V, V77I, I84V, N88D and L90M.

1.8.5 Amprenavir

Amprenavir is a drug from Vertex and Glaxo Wellcome, which belongs to a class of sulfonamide protease inhibitors and has been shown to have antiviral activity against HIV-1 and HIV-2, with IC_{50} values of between 80 nM and 340 nM, respectively (Kim *et al.,* 1995). It is a small tetrapeptide analogue with a sulfonamide and a tetrahydrofuran moiety in P2`. DNA sequence analysis of the protease of HIV-1 resistant variants revealed a sequential accumulation of point mutations in amino acid substitutions L10F, M46I, I47V and I50V. For amprenavir, the predominant resistant mutation in the binding site appears to be the I50V. It confers a two- to three-fold susceptibility (Partaledis *et al.,* 1995). In addition, subsequent substitutions associated with the I50V mutation included changes at residues 54 and 84.

1.8.6 Lopinavir

Lopinavir was developed by Abbott Laboratories and is only available in combination with ritonavir. It has a very high antiviral activity against HIV-1 (K_i value of 1.3 pM) and HIV-2 (Sham *et al.,* 1998). It contains a simplified cyclic urea, which replaces the thiazolymethyl urea of ritonavir in P2 and a dimethylphenoxyacetyl moiety instead of the thiazoylmethyl carbamate in P2`. To date, reduced sensitivity to lopinavir has been mainly associated with mutations M46I, I54V, A71V, V82A and I84V (Sham *et al.,* 1998). Other mutations which have been associated with resistance to lopinavir include L10F/I/R/V and L90M, as well as mutations at codons 20, 24, 32, 33, 46, 47,

50, 53, 63 and 90 (Carrillo *et al.,* 1998). One study reported that having five or more mutations at these positions or positions 33, 36, or 48 is associated with complete resistance to lopinavir while having three or four of these mutations is linked to mild resistance (Wlodawer, 1999). V47A also causes resistance to lopinavir in patients infected with HIV-2.

1.8.7 Atazanavir

Atazanavir, a second generation HIV-1 protease inhibitor, was developed by Bristol-Myers Squibb as an alternative option for initial PI-based HAART. It is a 2-hydroxy-1,3-diaminopropane transition state isostere with an aza-dipeptide core and an extended P1` moiety (Robinson *et al.,* 2000). It displays a high potency against wild-type HIV-1 protease with a K_i value of 10 pM and good oral bioavailability (Gong *et al.,* 2000). It has been postulated to have a role in treatment simplification for patients with undetectable plasma HIV-1 RNA for > 6 months on an initial antiretroviral regimen to which intolerance has developed. It has minimal effects on lipids, but when taken in combination with other protease inhibitors in patients without underlying drug resistance, it is sensitive to the signature I50L mutation. Interestingly, unlike the common I50V mutation observed with amprenavir, this mutation is not associated with cross-resistance to other protease inhibitors. A number of primary and secondary mutations associated with atazanavir resistance have also been observed for patients who have previously been on one or more protease inhibitors before commencing with atazanavir (Collono *et al.,* 2003).

1.8.8 Tipranavir

Tipranavir was developed as a second generation HIV-1 protease inhibitor by Boehringer Ingelheim from a class of compounds known as dihydropyrones that are structurally similar to coumadin and were found to inhibit protease (Chrusciel and Strohbach, 2004). It is a sulfonamide-based inhibitor shown to have antiviral activity against HIV-1 with K_i value of 5 pM and IC_{90} value of 100 nM (Turner *et al.,* 1998). It is effective against HIV-1 isolates that are resistant mainly to ritonavir, especially the M46I, L63P and I84V/A isolates which conferred resistance of (47- to 125-fold) to ritonavir. This resistant isolate conferred only a 6-fold resistance in Tipranavir (Rusconi *et al.,* 2000). It also exhibited robust activity against the V82A and V82F/I84V isolates with K_i values of 3.0 nM and 0.25 nM, respectively (Abbenante

and Fairlie, 2005). Nonetheless, resistance to tipranavir can still occur, although the acquisition of up to ten mutations is required to attain high levels of resistance (Doyon *et al.*, 2005). Of note, the L33F and I84V appear to be dominant substitutions responsible for the development of tipranavir resistance. L33F is a secondary resistance mutation remote from the active site and does not appear to affect tipranavir binding to the active site (Schake, 2005). The I84V mutation, on the other hand, affects tipranavir binding directly by altering hydrophobic interactions with the drug (Schake, 2004).

1.8.9 Darunavir

Darunavir is a second-generation HIV-1 protease non-peptide inhibitor with improved pharmacokinetics that is chemically related to the clinical inhibitor amprenavir. It was designed with a novel *bis*-tetrahydrofuranylurethane (*bis*-THF) group to incorporate additional polar interactions with main-chain atoms of the PR dimer. It was also developed by Tibotec Inc. (now Johnson and Johnson) and is a broad-spectrum potent inhibitor against HIV-1 isolates including multi-drug resistant clinical strains with minimal cyto-toxicity (Koh *et al.*, 2003). It also has higher genetic barrier for new drug resistance development compared to existing and other *bis*-THF containing protease inhibitors. *In vitro* resistance with strains harboring R41K and K70E mutations has been observed (King *et al.*, 2004b). In addition, isolates containing these two mutations were found to have 8- to 10-fold resistance to darunavir, 20-fold resistance to saquinavir, and 6-fold resistance to lopinavir (De Meyer *et al.*, 2005). For all other first-generation protease inhibitors, resistance remained less than 4-fold. To date, no additional resistant mutations have been observed with doses of darunavir.

1.9 Polymorphisms and resistance mutations in HIV-1 protease from African subtypes

The magnitude of the AIDS epidemic in Africa is well documented. It has been shown that Africa constitutes about 70 % of people infected with HIV worldwide. While the adult prevalence rate in North America and Western Europe is around 0.5 %, the overall figure is over 8 % in sub-Saharan Africa. However, the HIV-1 subtypes prevalent in the United States and Western Europe are not the same as those that are found in Africa. In the United States, Western Europe and Australia the subtype B is responsible for the greatest number of HIV infections whereas in sub-Saharan Africa

the A and C subtypes, and to a lesser extent the G subtype account for most of the infections. The A subtype is, however, predominant in the northern part of sub-Saharan Africa while the C subtype is prevalent in Southern Africa and the G subtype in Nigeria. It has been proposed that factors such as distinct transcriptional regulation could give rise to differences in transmission rates and pathogenesis among HIV-1 subtypes (Montano *et al.*, 1997; Essex, 1999). Protease-inhibiting drugs currently in use are optimized to bind tightly and specifically to the HIV-1 subtype B wild-type protease, and not the subtype C protease which carries amino acid polymorphisms, some of which have been associated with drug resistance (Becker-Pergola *et al.*, 2000). The high prevalence of certain resistance mutations in the non-B subtypes suggests that drug resistance may appear more rapidly, and that some non-B subtypes may be inherently resistant to certain, if not all, antiretroviral drugs. Presently studies involving the effects of the available HIV-1 protease inhibitors in people infected with non-B subtypes of HIV-1 are still at the early stages. Therefore, two crucial questions relate to the: (1) efficacy of the existing drugs against proteases from different HIV-1 subtypes and (2) mechanism of drug-resistant mutations within the scope of proteases from non-B subtypes.

Inspection of the HIV-1 sequence database (Benson *et al.*, 1998) shows that naturally occurring amino acid polymorphisms in the A and C HIV-1 subtype proteases occur at sites distinct from the active site (Figure 11). They appear to be clustered in specific regions including the hinge region of the flap, the loop connecting the β-strands, the α-helix and the opposite β-strand. In the absence of inhibitor pressure naturally occurring polymorphisms are most likely to occur in regions distal from the active site. As such, these polymorphisms do not affect the catalytic efficiency of the protease or viral fitness. As a corollary, naturally occurring polymorphisms do not occur in regions that could result in a loss of structural stability, catalytic activity or substrate binding affinity. In addition, studies by Velazquez-Campoy *et al.* (2001a; 2002) have shown that subtype A and C proteases with naturally occurring polymorphisms from drug-naïve individuals were associated with enhanced catalytic efficiency, decreased drug-binding affinity, improved biochemical fitness or vitality in the presence of existing inhibitors and an increase in structural stability. Under the same experimental conditions, the A, C and G subtype proteases showed a higher

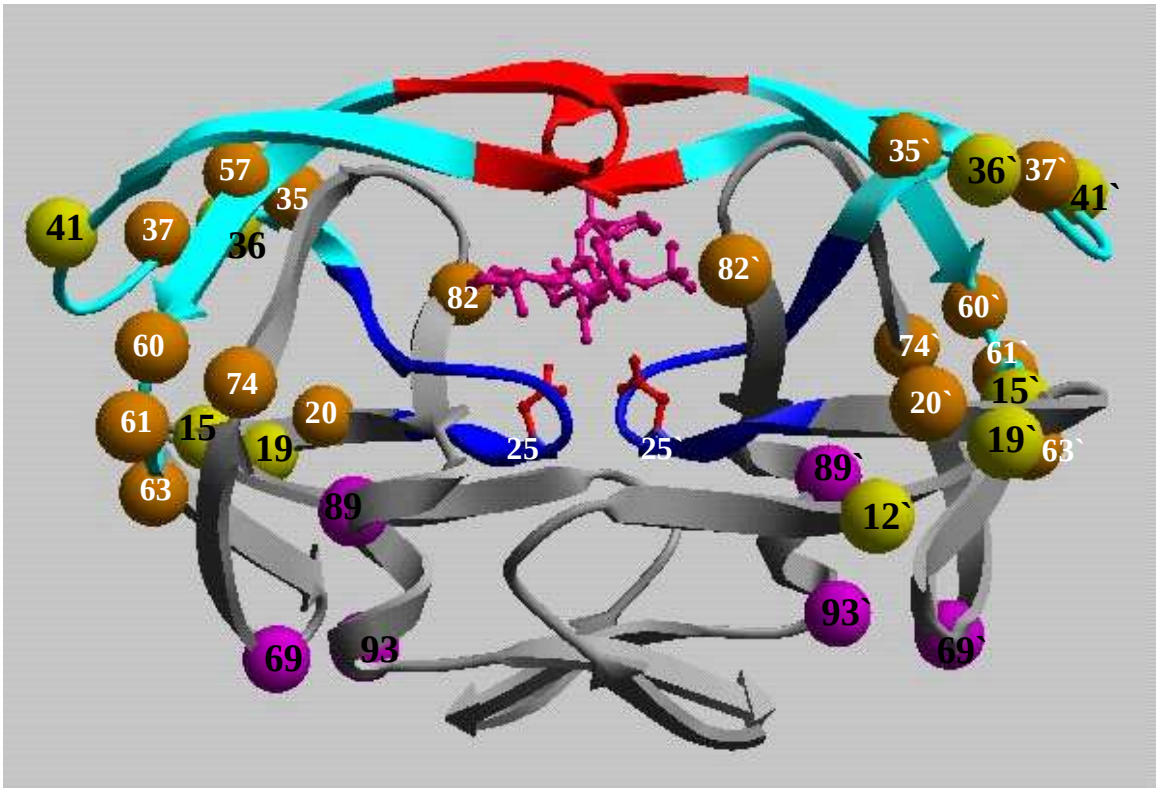
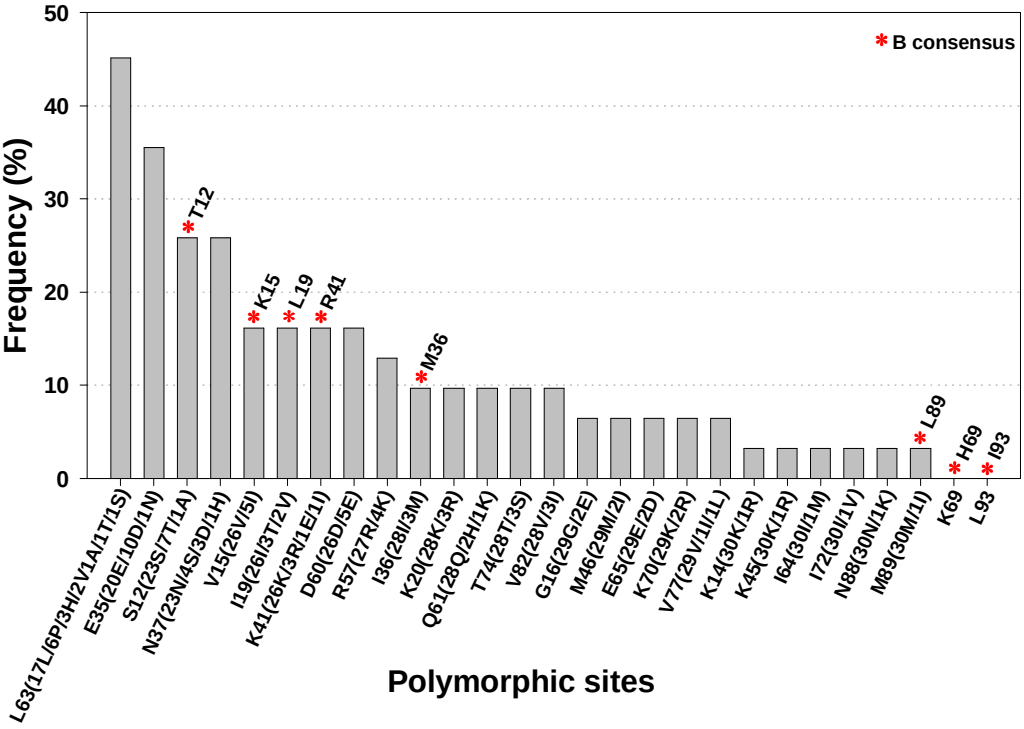


Figure 11. A ribbon representation of the homodimeric structure of HIV subtype B protease illustrating the active site loop (blue) with the catalytic Asp-25 (red ball-and-stick), the flap region (cyan) and flap tips (red). The protease inhibitor (purple ball-and-stick) is shown bound at the active site between the two subunits. The coloured spheres represent the polymorphic positions in the South African subtype C protease sequences: Brown, polymorphic frequency is equal to or greater than 10 % but same consensus for C and B; Green, polymorphic frequency is equal to or greater than 10 % and nonconsensus for C and B; Purple, no or low (<10 %) polymorphism and non-consensus for C and B. The topographical positions occupied by the 8 consensus South African subtype C amino acid residues: 12, 15, 19, 36, 41, 69, 89 and 93 are also shown.

structural stability than the subtype B protease (0.7, 1.3 and 2 kcal/mol, respectively) (Velazquez-Campoy *et al.*, 2002). Studies by Velazquez-Campoy *et al.* (2003b) have also shown that the South African C subtype (C-SA) protease is, however, slightly less stable than the B subtype. All HIV subtypes feature naturally occurring genetic variations. At the molecular level, the combined effects of naturally existing polymorphisms seem to have important consequences on the efficacy of the current HIV-1 protease inhibitors, since some of them influence the binding thermodynamics of inhibitors and may amplify the effects of drug-resistance mutations (Velazquez-Campoy *et al.*, 2001a; 2002). For example, a significant number of amino acid polymorphisms exist among subtypes A, B, C and G sequences. Natural polymorphisms of note within the subtype C protease include: L10I/V/F, K20R, M36I, S37N, R41K, M46I, H69K, A71V/T, V77I, L89M and I93L, some of which have been associated with or regarded as secondary resistant mutations in the subtype B virus (see Figure 12 and Appendix) (Holguin and Soriano, 2002; Kantor and Katzenstein, 2003).

Interestingly, none of these polymorphisms are directly located in the catalytic or active site. They are, however, located at sites that are pivotal for the conformational change associated with binding (e.g. the hinge region and the α-helix) (Rose *et al.*, 1998; Todd and Freire, 1999). Mutations occurring at these sites could affect the energetics of these changes and subsequently the binding affinity of inhibitors. As an example, the naturally occurring isoleucine at position 36 in the African subtypes is a secondary mutation in the B subtype virus (M36I) and may possibly act at a later stage of viral infection, following the onset of primary resistant mutations (Berker-Pergola *et al.*, 2000; Perno *et al.*, 2001). The I93L mutation has also been associated with drug resistance and classified as a secondary mutation (Cane *et al.*, 2001). As an increasing number of individuals with non-subtype B infection access HAART, these genotypic differences in non-subtype B consensus (wild-type) sequences may result in new drug resistant forms as well as differences in the long term outcomes of antiretroviral therapy.

A.



B.

```
           10          20          30          40          50
C: PQITLWQRPL VSIKVGGQIK EALLDTGADD TVLEEINLPG KWKPKMIGGI
B: .......... .T..K...L. .......... .....M.... R.........
           60          70          80          90
C: GGFIKVRQYD QILIEICGKK AIGTVLVGPT PVNIIGRNML TQLGCTLNF
B: .......... ........H. .......... ........L. ..I......
```

Figure 12. (**A**) A graphical representation showing the polymorphisms in the subtype C protease from drug-naïve South African patients. The subtype C-consensus residue and substitutions are shown along the x-axis. The 8 non-consensus positions between C and B subtypes are indicated with asterisks and the B consensus residue above the corresponding bars. (**B**) The sequence alignment for the protease from the HIV-1 B and C subtypes. South African subtype C consensus sequence is on top and the subtype B consensus sequence is the bottom sequence.

1.10 Objectives

The project focused on the effects arising from the combination of active site mutations in the C-SA HIV-1 protease. The polymorphic substitutions and active site mutations in the protease molecule were used to determine: (1) the energetics of acetyl-pepstatin and drug binding (2) the catalytic efficiency and the vitality/biochemical fitness of the protease. Studies by others (Velazquez-Campoy *et al.*, 2001a; 2002) have shown that amino acid polymorphisms by themselves do not appear to be sufficient enough to elicit drug resistance in the proteases from the C and A HIV-1 subtypes. But, because the C-SA protease has some amino acid polymorphisms which have been shown to be resistant to the clinical inhibitors in the subtype B HIV-1 protease, it is not known if the effects of these naturally occurring polymorphisms in combination with drug-resistant mutations will be additive or show some degree of positive or negative synergy. *In vitro,* selection experiments with HIV-1 protease inhibitors have identified combinations of active site mutations emerging from the viral population circulating in patients (Maschera *et al.*, 1996). Typically, active site mutations emerge first followed by compensatory mutations that seem to improve the efficiency of the protease by increasing *in vivo* viral replication in the presence of the inhibitor (Molla *et al.*, 1996). More so, residues altered at positions 82 and 84 have long been identified by passaging HIV-1 in the presence of increasing concentrations of HIV-1 protease inhibitors (Otto *et al.*, 1993; King *et al.*, 1995; Hodge *et al.*, 1996; Cane *et al.*, 2001). Thus, the V82A and V82F/I84V mutations were chosen. These mutations are well-characterized within the framework of the subtype B virus and are known to confer cross-resistance to most of the inhibitors currently in clinical use. Notably, the V82A mutation has also been observed at high frequencies in the South African treatment-experienced patients. The current project was aimed at addressing two questions: (1) Are existing drugs equally effective against the South African C subtype? and (2) How do active site resistant mutations operate within the framework of the South African C subtype? A study by *Todd et al.* (2000) on the resistant double mutant (V82F/I84V) of subtype B protease has shown that a key factor in the understanding of resistance is the relative balance of enthalpic and entropic contributions to the Gibbs energy of binding. Isothermal titration calorimetry was, therefore, used to determine the energetics of drug binding to the subtype C-SA proteases (wild-type and active site mutants).

CHAPTER 2

2. EXPERIMENTAL PROCEDURES

2.1 Materials

Acetyl-pepstatin was purchased from Bachem AG (Germany). Molecular grade urea was from Promega. The QuikChange™ Site-directed Mutagenesis kit was purchased from Stratagene (La Jolla, CA, USA). The oligonucleotide mutant primers and restriction enzymes *Bam*HI and *Nde*I were from Inqaba Biotech (South Africa). Other reagents were of an analytical grade. The expression vector, pET-11b, with the cDNA coding region for the wild-type B subtype HIV-1 protease was a kind gift from Dr J. Tang, University of Oklahoma Health Science Center, Oklahoma City (Ido *et al.*, 1991).

2.2 Source of sequence data

Protease sequence data were obtained from the AIDS Virus Research Unit (NICD, South Africa) for drug-naïve and drug-treated South African patients. This data was used to indicate baseline polymorphisms and the appearance of primary and secondary drug-resistance mutations during drug treatment. Information obtained from the Stanford HIV database (www.hivdb.stanford.edu) and from the National Institute for Communicable Diseases (NICD) was used to generate mutants (V82A and V82F/I84V) of the South African C subtype (C-SA) protease.

2.3 Source of HIV-1 protease and expression plasmid

A pET-11b expression vector with a gene encoding the B subtype HIV-1 protease was used. The pET-11b plasmid is under the control of the T7 promoter. The wild-type subtype C-SA HIV-1 protease naturally carries the mutations T12S, I15V, L19I, M36I, R41K, H69K, L89M and I93L with respect to the subtype B sequence. The oligonucleotide primers incorporating the T12S, I15V, L19I, M36I, R41K, H69K, L89M and I93L amino acid changes with one autocatalytic site protected, Q7K, were synthesized and purified. Site-directed mutagenesis was undertaken to systematically convert the subtype B protease to C-SA. Proteases (consensus wild-type and variants) were over-expressed in *Escherihia coli* BL21 (DE3) cells containing the pLysS plasmid and induced with 0.4 mM isopropyl-β-D-thiogalactoside (IPTG).

2.4 Verification of the wild-type pET-11b plasmid

Plasmid DNA preparations from *Escherichia coli* XL1-Blue cell cultures transformed with pET-11b were purified using the Qiagen mini-prep kit. Wild-type cDNA was digested using *Nde*I (10 U/µl) and *Bam*HI (10 U/µl) at 37 °C for at least 1 hour. The digested DNA was analyzed on a 1 % (w/v) agarose gel containing 0.2 µg/ml ethidium bromide and the sizes of the fragments were compared with standard DNA markers.

2.5 Protein engineering

2.5.1 Oligonucleotide primer design

The oligonucleotide primers were designed for use with the QuikChange™ Site-Directed Mutagenesis kit with the aid of the Gene Runner v3.04 software package. The oligonucleotides were designed in accordance with the published nucleotide sequence encoding HIV-1 protease (Ido *et al.,* 1991). The mutagenic primers used to generate the V82A and the V82F/I84V HIV-1 variant proteases had the following sequences:

V82A Forward primer: 5`- G GTT GGT CCG ACT CCG **GCT** AAC ATC ATC GGC-3`
V82A Reverse primer: 5`- GCC GAT GAT GTT **AGC** CGG AGT CGG ACC AAC C-3`

V82F/I84V Forward primer: 5` - GGT CCG ACT CCG **TTC** AAC **GTT** ATC GGC CGT AAC - 3`
V82F/I84V Reverse primer: 5` - GTT ACG GCC GAT **AAC** GTT **GAA** CGG AGT CGG ACC - 3`

Substituted residues are shown in bold.

2.5.2 Mutagenesis

The desired mutations at selected positions, V82A and V82F/I84V, were introduced into the cDNA insert using a commercial site-directed mutagenesis kit (QuikChange™ Stratagene; Papworth *et al.,* 1996). Site-directed mutagenesis was performed according to the manufacturer's instructions. Briefly, double stranded pET-

11b (Ido *et al.*, 1991) plasmid DNA was used as the template and PCR reactions were performed in a total volume of 50 µl. The PCR product was generated through sixteen amplification cycles of 30 seconds at 95 °C to denature the DNA, 30 seconds at 55 °C to anneal the primers and 120 seconds at 68 °C for DNA extension. The reaction contained 10 X reaction buffer, 5 – 50 ng of double stranded DNA template, 125 ng of each oligonucleotide primers and 10 mM dNTP mix. *Pfu* Turbo DNA polymerase (2.5 U/µl) was then added. The PCR mixture was treated using *Dpn*I (10 U/µl) to digest the methylated DNA template. The plasmid was recircularized at room temperature and used to transform *Escherichia coli* XL1-blue supercompetent cells supplied with the kit (Chung *et al.*, 1989). Cells were plated onto LB agar plates supplemented with ampicillin (100 µg/ml). Putative mutant plasmid DNA from picked colonies was prepared as described in section 2.3 and sent for sequencing to ensure that no other mutations were incorporated in the cDNA during the thermal cycling reaction.

2.5.3 Sequencing of mutant plasmid DNA

The entire cDNA encoding either the V82A or the V82F/I84V variant enzymes were sequenced to confirm the incorporation of the mutations into the pET-11b plasmid. DNA sequencing was performed by Inqaba Biotechnical Industries (Pty) Ltd (Pretoria, South Africa; www.inqaba.com) using the BigDye® Terminator v3.1 Cycle Sequencing Kit from Applied Biosystems (Foster City, CA, USA) and a SCE 2410 Genetic Analyser from Spectrumedix (State College, PA, USA).

2.5.4 Transformation of plasmid DNA into Escherichia coli BL21 (DE3) pLysS cells

The plasmid DNA encoding either the wild-type C-SA, the V82A or the V82F/I84V protease was transformed into *Escherichia coli* BL21 (DE3) pLysS cells by a one step transformation method described by Chung *et al.* (1989). *Escherichia coli* cells were grown in LB medium to the early exponential phase as measured by optical density (O.D) at 600 nm and diluted 1:1 with 2 x transformation and storage solution [LB medium with 10 % (w/v) polyethylene glycol, 5 % (v/v) DMSO and 30 mM $MgCl_2$, pH 6.5]. A 0.1 ml aliquot of cells was transferred into an ice-cold polypropylene tube, mixed with 1 µl (100 ng) of wild-type and mutant pET-11b plasmid DNA, respectively, and incubated for 30 minutes at 4 °C. LB medium (0.9 ml) containing 20 mM glucose was added and cells were grown at 37 °C in a shaker at 225 rpm for 1

hour to allow for expression of the antibiotic-resistance gene. Transformants were selected by plating the cells on LB plates supplemented with ampicillin (100 μg/ml) and chloramphenicol (35 μg/ml) and incubated at 37 °C for 17 – 20 hours.

2.6 Over-expression and purification of the HIV-1 protease

The HIV-1 protease was over-expressed as inclusion bodies (Ido *et al.*, 1991) in *Escherichia coli* BL21 (DE3) pLysS cells. Briefly, *Escherichia coli* cells harboring pET-11b plasmid DNA were grown at 37 °C in LB broth with ampicillin (100 μg/ml) and chloramphenicol (35 μg/ml). The overnight culture was diluted 100-fold into fresh LB medium containing ampicillin (100 μg/ml) and chloramphenicol (35 μg/ml). The culture was allowed to grow for 3 hours at 37 °C. When the optical density, measured at 600 nm, of the culture medium reached 0.4 – 0.5, IPTG was added to a final concentration of 0.4 mM. The cultures were allowed to grow for 4 hours. The cells were harvested by centrifugation at 5000 x g and stored at – 20 °C overnight. The cells were then resuspended in 100 ml of ice-cold extraction buffer A (10 mM Tris, 2 mM EDTA, 1 mM phenylmethylsulfonylfluoride [PMSF] pH 8.0) and disrupted by sonication. $MgCl_2$ and DNase I were then added to the final concentrations of 10 mM and 10 U/μl, respectively. When the viscosity of the mixture decreased, the cells were ruptured by sonication. The cell homogenate was centrifuged at 10 000 x g for 30 minutes. The pellet was resuspended in 100 ml of ice-cold buffer A containing 1 % (v/v) Triton X-100 and centrifuged at 10 000 x g for 30 minutes. Cell debris and protease-containing inclusion bodies were collected by centrifugation (10 000 x g for 30 minutes at 4 °C). The pellet, which contained the majority of the proteins expressed as inclusion bodies, was then dissolved in buffer B (10 mM Tris, 8 M urea, 10 mM DTT, pH 7.5). The solubilized protease was passed through a DEAE anion exchange column equilibrated with buffer B and acidified by adding formic acid to 25 mM immediately upon elution from the column. Precipitation of significant amount of contaminants occurred upon acidification. Protease-containing fractions, as determined by the Bradford dye reagent (Biorad), were pooled concentrated and stored at 4 °C in the concentration range of 5-10 mg/ml.

The HIV-1 protease was refolded by extensive dialysis into 10 mM formic acid at 4 °C. Thereafter, the HIV-1 protease was extensively dialyzed in 10 mM sodium acetate buffer containing 1 mM NaCl and 1 mM DTT, pH 5.0. Folded protease was desalted into 1 mM sodium acetate buffer at pH 5.0 using a PD-10 gel filtration column (Amersham, Pharmacia, Biotech, Sweden) and stored at either 4 or -20 °C (≥ 2.5 mg/ml). The homogeneity and molecular weight of the folded purified proteins was assessed by separation on 15 % acrylamide SDS-PAGE (Laemmli, 1970; see section 2.9), using proteins of known molecular size as markers. To assess to oligomeric state of the proteins, size exclusion high performance liquid chromatography (SEC-HPLC) was performed on the wild-type and variant HIV-1 proteases under non-denaturing conditions. The proteins were separated on an HPLC column at a flow rate of 0.5 ml/min and the column eluent monitored at absorbance of 280 nm. The column was calibrated with molecular standards which were used to determine the sizes of proteins.

2.7 Purification of clinical inhibitors

Clinical inhibitors, saquinavir, indinavir, ritonavir and nelfinavir were purified from commercially available capsules by HPLC using a semi-preparative C-18 reverse-phase column developed with 0-100 % acetonitrile in 0.05 % TFA (Velazquez-Campoy *et al.*, 2001a). Purified inhibitors were lyophilized and stored at -20 °C in the crystalline form (indinavir, nelfinavir) or as suspensions in DMSO (saquinavir, ritonavir). Only these four protease inhibitors were available in the country when the studies were conducted. Acetyl-pepstatin was used without further purification.

2.8 Protein concentration determination

The concentrations of the wild-type, V82A and V82F/I84V HIV-1 protease were determined spectrophotometrically using the Beer-Lambert Law:

$$A = \varepsilon_\lambda cl \qquad (4)$$

where A is the absorbance, ε_λ is the molar extinction of the absorbing species at a given wavelength λ, c is the concentration of the absorbing solution and l is the pathlength of light through the solution in the cuvette. The concentration of pure

protein was subsequently determined spectrophotometrically using an extinction coefficient $E_{1\%}$ of 11.8 at 280 nm (Polgar *et al.*, 1994).

2.8.1 HIV-1 protease active site concentration determination

The concentration of the active sites was determined calorimetrically using a high-precision VP-ITC titration calorimeter (Microcal Inc., Northampton, MA). Briefly, an enzyme solution (20 μM) in the calorimetric cell was titrated with an inhibitor solution, acetyl-pepstatin (Ac-Val-Val-Sta-Ala-Sta), dissolved in the same buffer (10 mM sodium acetate, pH 5.0). The inhibitor concentration in the experiment was in the region of 200 - 300 μM in the injection syringe. Certificate analysis from the supplier indicated that acetyl-pepstatin used in these studies was chemically pure. Therefore, no further purification of the ligand was needed. The protein concentration was corrected for light scattering effects (Winder and Gent, 1971). Protein spectra were obtained between 240 and 370 nm. Absorbance values were taken between 320 and 370 nm at 5 nm intervals. The resulting absorbance values were plotted against wavelength in order to obtain a linear fit with the equation: $y = mx + c$, where m and c are obtained from the linear regression curve fit. By substituting $x = 280$ nm into the above equation, the value of y is obtained. This value represents the contribution of light scattering effects to the protein sample. Subtraction of this value from the absorbance obtained at 280 nm yields the true absorbance in the absence of light scattering effects. The enzyme solution was then loaded into the calorimetric cell and the ligand loaded into the syringe. The system was then allowed to equilibrate and obtain a stable baseline prior to automatic initiation of the titrations. The reference power was set to 15 μcal/sec and the initial delay was 180 seconds. The stirring speed was kept constant at 480 rpm. The heat evolved after each injection was obtained from the integral of the calorimetric signal. The heat due to the binding reaction between the inhibitor and the enzyme was obtained as the difference between the heat of the reaction and the corresponding heat of dilution. All raw data obtained from these calorimetric experiments were analyzed using Origin v5.0 software package (Microcal Software Inc., Northampton, MA) (Muzammil *et al.*, 2003). The stoichoimetry (N) obtained from the titration data was then used to estimate the concentration of active sites in the protease molecule.

2.9 SDS-PAGE

SDS-PAGE according to the method of Laemmli (1970) was used to assess the purity of the DEAE purified protein samples. SDS-polyacrylamide separating gels of 15 % acrylamide and 0.4 % bisacrylamide and stacking gels of 5 % acrylamide and 0.1 % bisacrylamide were prepared. The protein samples were diluted five-fold in 5 X loading buffer (0.5 mM Tris-HCl, pH 6.8, 20 % (v/v) glycerol, 10 % (w/v) SDS, 100 mM ß-mercaptoethanol, 0.05 % (w/v) bromophenol blue) and boiled for 5 minutes before loading onto gels. The gels were run at 140 volts for 2 hours. Coomasie Blue (0.25 % Coomasie Brilliant Blue R250 in 45 % methanol and 10 % acetic acid) was used to stain the gels for 1 hour. This was followed by destaining in destain solution containing 85 % water, 15 % acetic acid and 10 % ethanol. The SDS-PAGE molecular mass markers used were: ß-galactosidase (116.0 kDa), phophorylase b (97.0 kDa), albumin (66.0 kDa), ovalbumin (45.0 kDa), carbonic anhydrase (30.0 kDa), trypsin inhibitor (20.1 kDa) and α-lactalbumin (14.4 kDa).

2.10 SE-HPLC

SE-HPLC was performed on a TSK G2000 SWXL size exclusion column equilibrated with 20 mM sodium phosphate buffer, pH 6.5 containing 100 mM NaCl, 1 mM EDTA and 0.02 % sodium azide. The buffer was filtered and degassed before use. An LKB 2150 pump (Pharmacia) was used to maintain the flow rate at 0.5 ml/min. Molecular weight marker proteins: thyroglobulin (670 kDa), bovine gammaglobulin (158 kDa), chicken ovalbumin (44 kDa) and equine myoglobulin (17 kDa) were used to construct a standard curve.

2.11 Spectroscopic studies of the wild-type, V82A and V82F/I84V HIV-1 protease

2.11.1 Circular dichroism

Circular dichroism (CD) is a technique used for assessing the structures of proteins and peptides in solution. CD gives information about differential absorption of left- and right- handed circularly polarized light in optically-active molecules. In proteins, the optically active groups are the amide interactions, peptide backbone and the aromatic side chains (Woody, 1995). The far-UV or amide region (250 – 200 nm) is dominated by secondary structural elements e.g., α-helices and ß-sheets, whereas CD bands in the near UV region (300 – 250 nm) originate from the aromatic side chains.

In particular, α-helices display two strong and characteristic minima at 222 nm and 208 nm, whilst the ß-sheets display a minimum at 216 nm.

2.11.2 Far-UV circular dichroism of the wild-type C-SA, V82A and V82F/I84V HIV-1 proteases

Far-UV (250 – 190 nm) CD spectra due to the peptide bond was used to give information about the backbone secondary structures of the wild-type and mutant proteins. The experiments were performed using 15 μM protein concentrations in 10 mM sodium acetate buffer, pH 5.0. CD spectra for native wild-type, V82A and V82F/I84V variant proteins were recorded from a Jasco model J-810 spectropolarimeter using a cuvette of 2 mm path length at 20 °C. Replicate scans were obtained at a 0.1 nm data pitch, 0.1 nm bandwidth and a scan speed of 50 nm/min. Spectra were averages of 10 scans with the baseline or buffer control subtracted from 250 to 190 nm in 0.1 nm increments.

2.11.3 Steady-state fluorescence

Fluorescence is an emission phenomenon whereby an energy transition from a higher to a lower energy state within the molecule is measured by the detection of emitted radiation. A molecule can be excited by an input of energy from electromagnetic radiation. Energy is released when the excited molecule falls back rapidly to a lower energy state. Thus, the emitted radiation occurs at a wavelength longer than the excitation wavelength. Fluorescence of the indole ring in the tryptophan residue and the phenyl ring of tyrosine is sensitive to the local environment and thus can be used as a valuable probe of tertiary structural changes occurring in the protein. This can also be used to monitor the thermodynamics and kinetics of protein unfolding/refolding reactions (Shirley, 1995).

Fluorescence measurements were performed using the Perkin Elmer Luminiscence Spectrometer LS 50B. The fluorescence spectral properties of the wild-type, V82A and V82F/I84V HIV-1 proteases were measured using 1 μM protein concentrations. The experiments were carried out in 10 mM sodium acetate buffer, pH 5.0. The intrinsic fluorescence emission spectra of native proteins (wild-type, V82A and V82F/I84V HIV-1 proteases) were obtained by selectively exciting the tryptophan residues at 295nm, and tyrosine and tryptophan residues at 280 nm. The emission

wavelengths were recorded from 300 – 420 nm at 20 °C. The slit widths on both the excitation and emission were set at 5 nm, and the spectra were obtained at 500 nm/min. Fluorescence data were corrected by subtracting the buffer scans from the sample scans.

2.12 Determination of kinetic parameters

The catalytic activity of the HIV-1 protease was monitored following the hydrolysis of the chromogenic peptide substrate Lys-Ala-Arg-Val-Nle-*p*-nitro-Phe-Glu-Ala-Nle-NH_2. This substrate mimics the conserved KARVL/AEAM cleavage site between the capsid protein and nucleopcapsid (CA-p2) in the Gag polyprotein precursor. Protease was added to a 120 µl microcuvette containing substrate. Final concentrations in the standard assays was 30-60 nM protease, 0-200 µM peptide substrate, 50 mM sodium acetate, and 0.1 M NaCl (pH 5.0). All experiments were performed at 20 °C. Protease hydrolytic activity was measured by monitoring the relative decrease in absorbance at 300 nm using a Jasco model V-550 spectrophotometer. An extinction coefficient for the difference in absorbance upon hydrolysis (1800 $M^{-1}cm^{-1}$ at 300 nm) was used to convert absorbance change to reaction rates. Hydrolysis rates were obtained from the initial part of the progress curves, where at least 80 % of the substrate remains unhydrolyzed. All reaction rates were corrected for corresponding non-enzymatic reaction rates.

2.13 Inhibition studies

To determine the concentration of the clinical inhibitors (saquinavir, ritonavir, indinavir, and nelfinavir) that result in 50 % inhibition of HIV-1 protease enzyme activity (IC_{50}), standard assays (i.e., 125 nM protein and 50 µM chromogenic substrate) were performed in the presence of increasing concentrations of inhibitor.

Inhibition constants, K_i, for the inhibitors (saquinavir, ritonavir, indinavir, and nelfinavir) were obtained at 20 °C by measuring the rate of chromogenic substrate hydrolysis using 15 - 40 nM protease in 50 mM sodium acetate, 0.1 M NaCl, pH 5.0, and (0 – 200 µM) substrate in increasing amounts of inhibitor.

The vitality of the drug-resistant variant protease compared to the wild-type in the presence of the inhibitor was calculated from (Velazquez-Campoy *et al.,* 2002; Ohtaka *et al.,* 2003; Gulnik *et al.,* 1995):

$$\text{vitality} = \frac{(K_d \times k_{cat}/K_m)_{mutant}}{(K_d \times k_{cat}/K_m)_{wild\text{-}type}} \quad (5)$$

2.14 Isothermal titration calorimetry

Isothermal titration calorimetry experiments were performed using a high precision VP-ITC titration calorimetry system (Microcal Inc., Northampton, MA). The enzyme solution was titrated with the appropriate inhibitor dissolved in the same buffer. The heat absorbed or released after each injection inhibitor solution was obtained from the integral of the calorimetric signal. The heat due to the binding reaction between the inhibitor and the enzyme was obtained as the difference between the heat of the reaction and the corresponding heat of dilution. This highly sensitive technique allowed for an accurate determination of the thermodynamic profile for a protein-ligand interaction. A thermodynamic profile includes the following parameters: N (stoichiometry), K_a (association constant), ΔH (change in enthalpy), ΔS (change in entropy), ΔG (change in Gibbs free energy) and ΔC_p (change in heat capacity). The independent variables of an ITC experiment include the stoichiometry of the interaction, the association constant, and the enthalpy of binding. The dependent variables are the Gibbs free energy and the entropy of binding.

Due to the extremely high affinity of the inhibitors for the HIV-1 protease, the binding thermodynamics were also measured in competition experiments with acetyl-pepstatin (a well-characterized inhibitor of low affinity and unfavorable binding enthalpy). In these displacement titration experiments, the inhibitor (saquinavir, ritonavir, indinavir or nelfinavir) was injected into the calorimetric cell containing the protein prebound to acetyl-pepstatin as previously described (Velazquez-Campoy *et al.,* 2001b). The selection of a weak binding inhibitor with a binding isotherm characteristic of positive binding enthalpy amplifies the magnitude of the signal in the displacement titration reaction due to the displacement of a weaker binding inhibitor by an inhibitor yielding an exothermic isotherm. Additionally, the binding isotherm has sufficient curvature

that is useful for the apparent binding affinity determination (Sigurskjold, 2000). The displacement experiments were performed at pH 5.0 using sodium acetate buffer with a negligible ionization enthalpy in order to minimize any proton coupling effects on the binding enthalpy. The binding enthalpy of saquinavir, ritonavir, nelfinavir or indinavir has previously showed no dependence on the ionization enthalpy of the buffer, indicating that there is no net proton exchange under the conditions used for these experiments (Todd *et al.,* 2000; Velazquez-Campoy *et al.,* 2000; Velazquez-Campoy *et al.,* 2001a; Velazquez-Campoy *et al.,* 2001b; Vega *et al.,* 2004). Analysis of the data was performed using the Origin 5.0 software.

2.14.1 Energetics of saquinavir binding to the wild-type, V82A and V82F/I84V HIV-1 proteases

The energetics of the wild-type HIV-1 protease-saquinavir and V82A HIV-1 protease-saquinavir interactions were investigated using displacement calorimetric titrations of saquinavir (300 μM) into a solution of the wild-type or V82A HIV-1 protease (20 μM) pre-bound to acetyl-pepstatin (200 μM). For the V82F/I84V HIV-1 protease, the energetics of saquinavir binding were investigated using direct calorimetric titrations of saquinavir (500 μM) into a 20 μM protein solution. The initial titration of saquinavir for all three proteins was 3 μl followed by twenty-four 10 μl injections. Between each injection, the baseline was allowed to stabilize for 6 minutes.

2.14.2 Energetics of ritonavir binding to the wild-type, V82A and V82F/I84V HIV-1 proteases

The energetics of the wild-type HIV-1 protease-ritonavir and/or V82A HIV-1 protease-ritonavir interactions were investigated using displacement of ritonavir (250 μM) into a solution of the wild-type C-SA or V82A HIV-1 protease (20 μM) pre-bound to acetyl-pepstatin (200 μM). The initial titration of ritonavir for the two proteins was 3 μl followed by twenty-four 10 μl injections. Between each injection, the baseline was allowed to stabilize for 6 minutes. For the V82F/I84V HIV-1 protease, the energetics of ritonavir binding were investigated using reverse calorimetric titrations of protein (300 μM) into a 20 μM ritonavir solution. This is because ritonavir has a low solubility and it precipitates at high concentrations. The initial titration of V82F/I84V HIV-1 protease was 3 μl followed by twenty-four 10 μl injections. Between each injection, the baseline was allowed to stabilize for 6 minutes.

2.14.3 Energetics of indinavir binding to the wild-type, V82A and V82F/I84V HIV-1 proteases

The energetics of the wild-type HIV-1 protease-indinavir interaction were investigated using displacement calorimetric titration of indinavir (250 μM) into a solution of the wild-type HIV-1 protease (20 μM) pre-bound to acetyl-pepstatin (200 μM). The energetics of the V82A HIV-1 protease-indinavir interaction were investigated using displacement calorimetric titrations of indinavir (400 μM) into a solution of the V82A HIV-1 protease (20 μM) pre-bound to acetyl-pepstatin (200 μM). For the V82F/I84V HIV-1 protease, the energetics of indinavir binding were investigated using direct calorimetric titrations of indinavir (400 μM) into a 20 μM protein solution. The initial titration of indinavir for all three proteins was 3 μl followed by twenty-four 10 μl injections. Between each injection, the baseline was allowed to stabilize for 6 minutes.

2.14.4 Energetics of nelfinavir binding to the wild-type, V82A and V82F/I84V HIV-1 proteases

The energetics of the wild-type HIV-1 protease-nelfinavir and the V82A HIV-1 protease-nelfinavir interactions were investigated using displacement calorimetric titration of nelfinavir (240 μM) into a solution of the wild-type or V82A HIV-1 protease (20 μM) pre-bound to acetyl-pepstatin (200 μM). For the V82F/I84V HIV-1 protease, the energetics of indinavir binding were investigated using direct calorimetric titrations of indinavir (400 μM) into a 20 μM protein solution. The initial titration of indinavir for all three proteins was 3 μl followed by twenty-four 10 μl injections. Between each injection, the baseline was allowed to stabilize for 6 minutes.

2.15 Protein crystallography

Crystals of the uncomplexed and inhibitor-complexed wild-type, V82A and V82F/I84V mutant HIV-1 proteases were grown at 20 °C by vapour-diffusion using the hanging-drop method (Fitzgerald *et al.*, 1990). The enzyme was prepared at 6 mg/ml in a pH 5 solution containing 10 mM MES, 1 mM EDTA, 1 mM DTT, 3 mM NaN_3. The reservoir solution contained 100 mM imidazole, 250 mM NaCl, 10 mM DTT, 3 mM NaN_3 at pH 7.0. Hanging drops were prepared by mixing the protein and reservoir solutions in 1:1 (v/v) ratio. Crystal appeared after 5 – 10 days.

2.16 Data analysis and molecular graphics

Unless otherwise stated, all linear and non-linear least-squares fitting of data was performed using the program Sigma Plot v8.0 (Jandel Corporation). All raw data obtained from calorimetric experiments were integrated using the ORIGIN v.5.0 software package supplied with the VP-ITC calorimeter. The images of the three-dimensional structures of the HIV-1 proteases displayed in this thesis were generated using Swiss-pdb Viewer v3.7 (Guex and Peitsch, 1997). Calculation of possible hydrogen bonding and inter-atomic distances were also performed using the Swiss-pdb Viewer.

CHAPTER 3

3. RESULTS

3.1 Verification of the pET-11b plasmid DNA

In Figure 13, the undigested wild-type pET-11b plasmid in lane 2 appears as one discrete band on the gel with a molecular size of 5.7 kb. The *Nde*I digested plasmid migrated as a single band to a position corresponding to approximately 6 kb. This is in agreement with the reported size of the wild-type pET-11b plasmid of 5.7 kb. The restriction map of pET-11b has single restriction sites for both *Nde*I and *Bam*HI which were used to verify the presence of the wild-type encoding insert (see Figure 13). The linearized insert migrated as a single band to a position corresponding to approximately 300 base pairs. The HIV-1 C-SA protease insert cDNA in the pET-11b plasmid was also verified by DNA sequencing using the universal primers of the T7 promoter (Inqaba Biotech, South Africa). The resultant sequence showed 100 % identity to the wild-type HIV-1 C-SA protease sequence.

Site-directed mutagenesis was used to incorporate the V82A and the V82F/I84V substitutions in the cDNA encoding the C-SA protease. Following the identification of the possible mutant plasmid DNA subclones, the entire cDNA encoding both the V82A and the V82F/I84V variants were sequenced to confirm the presence of the desired mutations, and to ensure that no other mutations were incorporated in the cDNA during the polymerase chain reaction. In Figure 14, segments of cDNA sequences containing the V82A and V82F/I84V codons are shown. The incorporation of the GTT → GCT (Val → Ala) and the GTT → TCC/ATC → GTT (Val → Phe/Ile → Val) mutations were verified.

3.2 Over-expression and purification of the wild-type and variant HIV-1 proteases

The wild-type C-SA, V82A and V82F/I84V variant proteins were expressed as inclusion bodies in *Escherichia coli* BL21 (DE3) pLysS cells and purified by washing with two buffers (buffer A and buffer B). Each wash consisted of resuspension (glass homogenizer, sonication) and centrifugation (10 000 x g for 30 min at 4°C). Lastly, the protease was solubilized in a buffer containing 8 M urea, pH 8.0, clarified by centrifugation and purified using DEAE anion exchange chromatography. Purity and monomeric size was assessed by SDS-PAGE (Figure 15). The monomeric sizes were

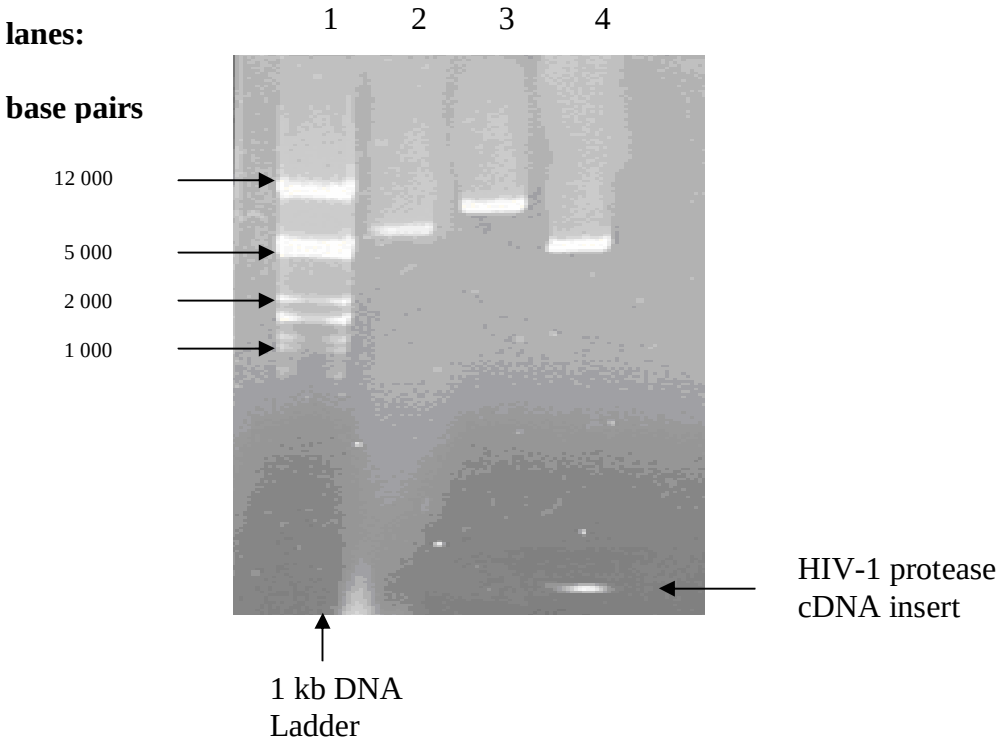


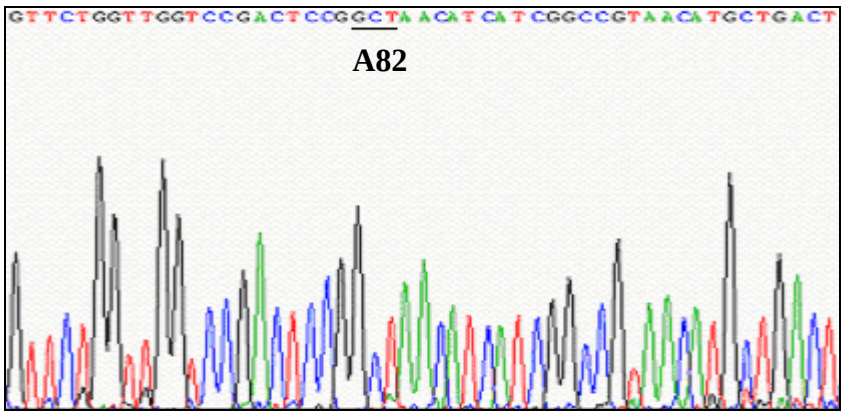
Figure 13. 1 % agarose gel showing separation of pET-11b plasmid DNA restricted with *Nde*I/*Bam*HI. Lane 1: 1kb Plus marker (Gibco, BRL®). Lane 2 is the uncut pET-11b control plasmid. Lanes 3 and 4: single digest of pET-11b plasmid DNA with *Nde*I, and double digest of pET-11b plasmid DNA with *Bam*HI and *Nde*I, respectively.

A. wild-type C-SA: GGT CCG ACT CCG **GTT** AAC **ATC** ATC GGC CGT

Val82 **Ile84**

B.

V82A variant:



V82F/I84V variant:

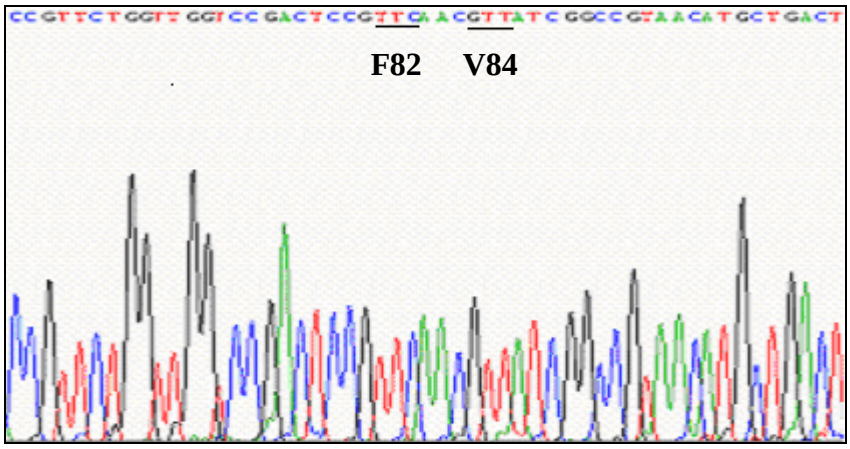


Figure 14. (**A**) A portion of the wild-type pET-11b nucleotide sequence encoding HIV-1 subtype C protease showing the position of the Val82 and Ile84 codons in bold. (**B**) Segments of the nucleotide sequences obtained from DNA sequencing of the V82A and the V82F/I84V substitutions introduced into pET-11b coding region of the HIV-1 protease (underlined).

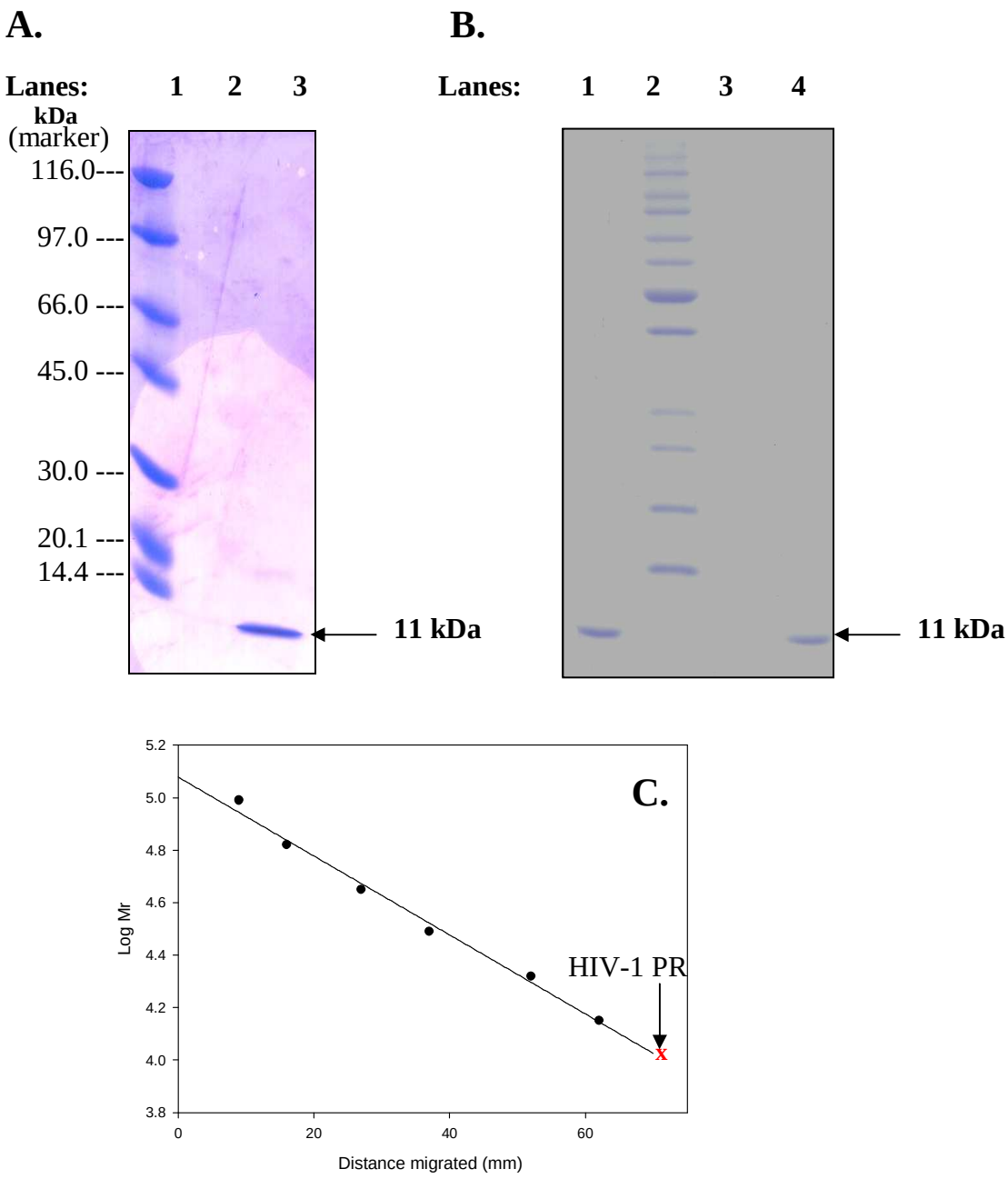


Figure 15. **Determination of the subunit molecular mass from SDS-PAGE.**

(**A**) Lane 1 represents the molecular weight standards: ß-Galactosidase (116.0 kDa), Phophorylase b (97 kDa), Albumin (66 kDa), Ovalbumin (45 kDa), Carbonic anhydrase (30 kDa), Trypsin inhibitor (20.1 kDa) and α-Lactalbumin (14.4 kDa). Lane 3 represents the purified wild-type C-SA HIV-1 protease (**B**) Lanes 1 and 4 represent the purified V82A and V82F/I84V variant HIV-1 proteases, respectively. (**C**) Calibration curve for SDS-PAGE molecular weight markers. The arrow indicates the molecular mass of the monomeric subunit of purified wild-type, V82A and V82F/I84V HIV-1 proteases. The correlation coefficient is 0.99.

determined to be very similar and corresponding to a molecular mass of 11 kDa (Figure 15) which is consistent with the value published previously (Ido *et al.,* 1991). The monomeric size of the column-purified protein (wild-type C-SA, V82A and V82F/I84V) at 11 kDa corresponds to the size of the HIV-1 protease subunit expected from the amino acid composition.

Analysis by SE-HPLC (Figure 16) indicates that the protein eluted as a single peak with an apparent oligomeric molecular mass of 22 kDa. This suggests that the hydrodynamic volume of the mutant proteins (V82A and V82F/I84V) was similar to that of the wild-type and that the homodimeric integrity of the protein was maintained in both proteases.

3.3 Spectroscopic properties of the wild-type and variant HIV-1 proteases

3.3.1 Far UV-CD

Circular dichroism was used as a probe to compare the secondary structure of the V82A and V82F/I84V variants with the wild-type C-SA protease. Far-UV CD (250 – 190 nm) spectra of the wild-type C-SA, V82A and V82F/I84V proteases were very similar and exhibited a minimum at 216 nm, typical for predominantly ß-sheeted content (Figure 17).

3.3.2 Spectral properties of the wild-type C-SA, V82A and V82F/I84V HIV-1 proteases

Intrinsic tryptophan fluorescence is often used as a probe to measure the conformational stability of proteins, as it is very sensitive to its local environments. Depending on their number and location, tryptophan residues could be used as probes for local or global changes. Single or localized tryptophan residue(s) monitor local conformational changes whereas multiple tryptophan residues distributed across three-dimensional structure would monitor global conformational changes. The tryptophan residues were selectively excited at 295 nm (Figure 18 (A)). Excitation of tryptophan and tyrosine residues was at 280 nm (Figure 18 (B)). Fluorescence studies, with excitation at 295 nm and 280 nm for all native proteins gave spectra with the emission maxima at 355 nm. However, the fluorescence intensity of the V82F/I84V HIV-1 protease is significantly enhanced (40 – 60 %) relative to the wild-type and the V82A HIV-1 proteases.

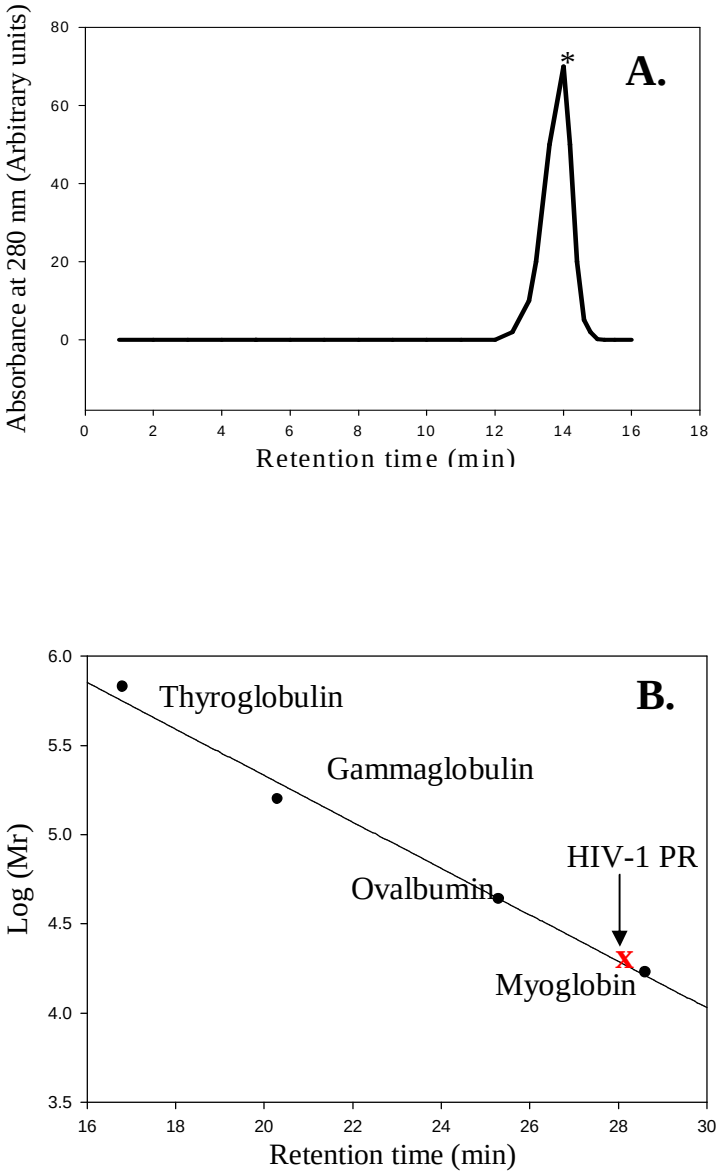


Figure 16. SE-HPLC elution profile of the HIV-1 protease

(**A**) SE-HPLC column elution profile of the native wild-type, V82A and V82F/I84V HIV-1 protease. The column was equilibrated with a 20 mM sodium phosphate buffer, pH 6.5 containing 100 mM NaCl, 1 mM EDTA and 0.02 % sodium azide at a flow rate of 0.5 ml/min. The asterisk in figure A denotes the position of the HIV-1 PR (**B**) Calibration curve of SE-HPLC molecular weight markers: thyroglobulin (670 kDa), gammaglobulin (158 kDa), ovalbumin (44 kDa) and myoglobin (17 kDa). The arrow in figure B indicates the position of the eluted wild-type and mutant HIV-1 proteases with a corresponding molecular mass of 22 kDa. The correlation coefficient was 0.99.

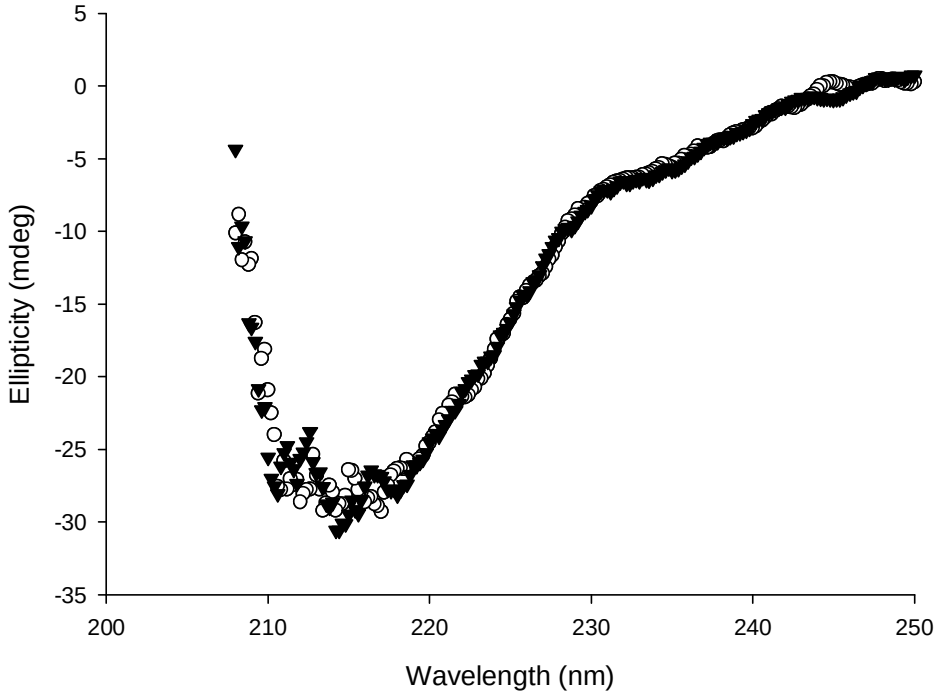


Figure 17. Far-UV circular dichroism spectra of the wild-type (●), the V82A (○) and the V82F/I84V (▼) HIV-1 proteases. The data from 10 runs were collected from 250 nm to 200 nm and averaged. Experiments were performed in 10 mM sodium acetate buffer, pH 5.0, containing 15 μM protein solutions.

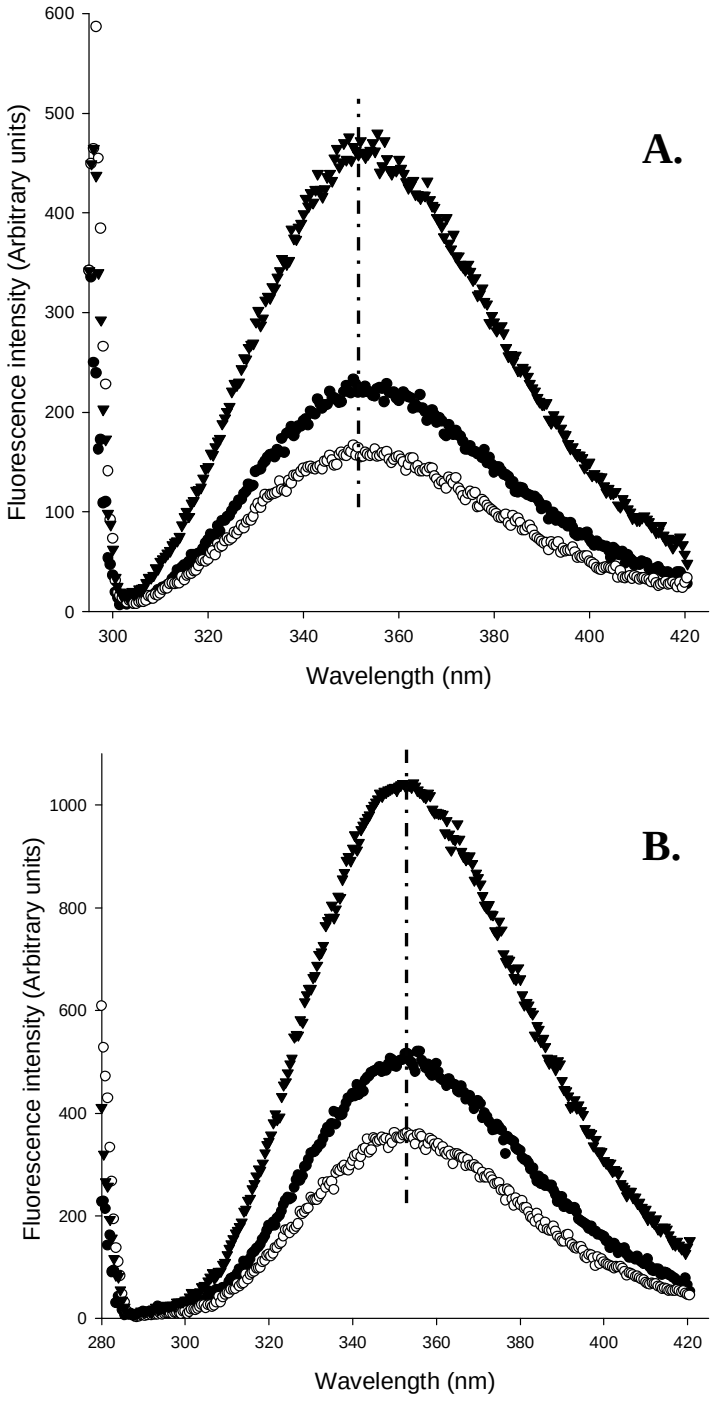


Figure 18. Spectral characterization of the wild-type (●), V82A (○) and V82F/I84V (▼) HIV-1 proteases in 10 mM sodium acetate buffer, pH 5.0. (**A**) The intrinsic trytophan emission (excitation at 295 nm) spetrum. (**B**) Represents Tyr and Trp emission spectra of HIV-1 proteases (excitation at 280 nm). The vertical lines represent emission maxima.

3.4 Characterization of the wild-type C-SA, V82A and V82F/I84V variant proteases

3.4.1 Effects of the V82A and V84F/I84V mutations on the catalytic activity of the C-SA HIV-1 protease

From an enzymatic point of view, no major differences were observed between the wild-type C-SA, V82A and V82F/I84V variant proteases. Table 2 shows the kinetic parameters of all the proteins against the same substrate (chromogenic peptide mimicking the highly conserved cleavage site KARVL/AEAM between the capsid protein and p2 in the gag protein precursor). Specific activities of the wild-type and the V82F/I84V proteases were determined to be 0.91 and 0.95 μmol/min/mg, respectively, whilst that of the V82A protease increased 3-fold to about 2.75 μmol/min/mg (Figure 19). The K_m for the substrate remains essentially unchanged for all proteases, with a 27 % and a 17 % loss in catalytic efficiency for the V82A and V82F/I84V mutants, respectively. These results are in agreement with values obtained previously (Gulnik *et al.,* 1995) (Figures 20, 21 and 22).

3.4.2 Enzyme inhibition

The inhibition constants (IC_{50} and K_i) of four inhibitors in clinical use (saquinavir, ritonavir, indinavir and nelfinavir) for the wild-type C-SA, V82A and V82F/I84V proteases were determined using enzyme kinetic assays with the chromogenic substrate (Lys-Ala-Arg-Val-Nle-nPhe-Glu-Ala-Nle-NH_2). Tables 3 and 4 summarize the IC_{50} and K_i values for all inhibitors studied. For the wild-type and the V82A proteases the IC_{50} and K_i values are in the nanomolar range, in agreement with previous data (Velazquez-Campoy *et al.,* 2001a; Klabe *et al.,* 1998). IC_{50} values for ritonavir and indinavir for the V82A mutation were slightly affected with a 5- and 9-fold increase, respectively.

Against the V82F/I84V mutant, however, the inhibition constants for all inhibitors were drastically weaker and characterized by IC_{50} and K_i ratios ranging from 50 to 450 (see Tables 3 and 4). Saquinavir was the least affected of the protease inhibitors used in these studies with 50-fold increase in the IC_{50} value and 100-fold increase in the K_i value whereas indinavir incurred the greatest increase (500- and 450-fold, respectively).

Table 2: Kinetic parameters for the wild-type, V82A and V82F/I84V HIV-1 protease

Protease	Specific activity (μmol/min/mg)	K_m (μM)	k_{cat} (1/s)	k_{cat}/K_m (1/μM.s)
wild-type C-SA	0.91	76	2.74	0.036
V82A	2.70	71	0.71	0.010
V82F/I84V	0.95	81	0.49	0.006

Activity was obtained by measuring the decrease in absorbance associated with the cleavage of the chromogenic substrate, using 30 – 60 nM protease in 50 mM sodium acetate buffer, pH 5.0 and 0 - 200 μM substrate. Protease hydrolytic activity was measured by monitoring the relative decrease in absorbance at 300 nm using a Jasco model V-550 spectrophotometer. All experiments were performed at 20 °C.

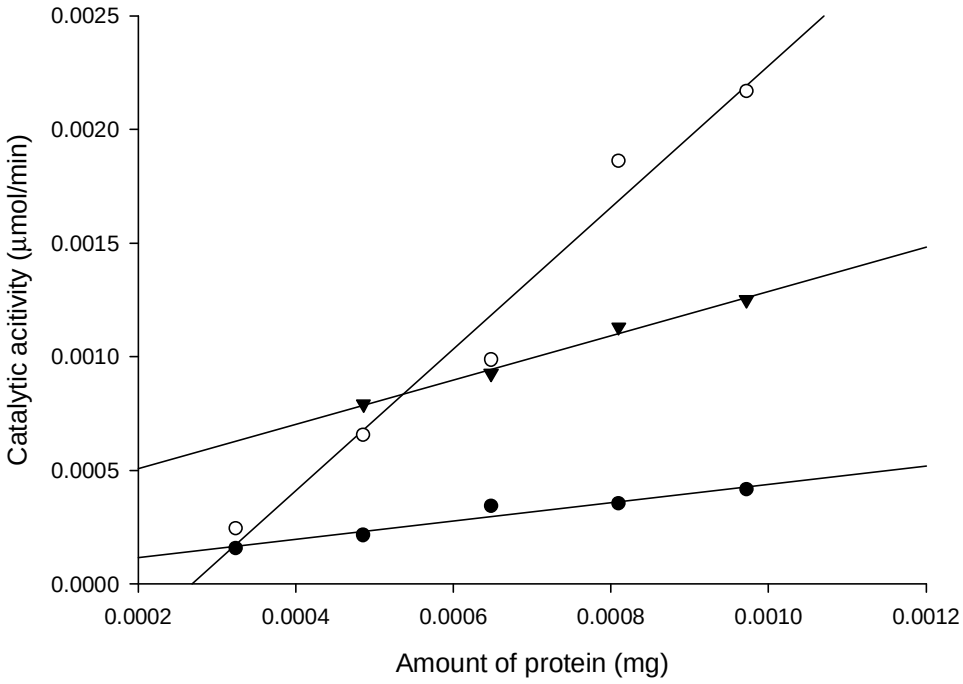


Figure 19. Specific activities of the wild-type C-SA (●), V82A (○) and V82F/I84V (▼) proteases. Catalytic activity was obtained by measuring the decrease in absorbance at 300 nm associated with the cleavage of the chromogenic substrate. Reaction conditions were: 5 – 40 nM protease in 50 mM sodium acetate buffer, pH 5.0. The reaction was initiated using 40 µM substrate. The solid lines represent linear fits to the experimental data. The specific activity of the HIV-1 protease was determined from the slope of each line. The correlation coefficients are 0.98, 0.97 and 0.99 for wild-type, V82A and V82F/I84V HIV-1 protease, respectively.

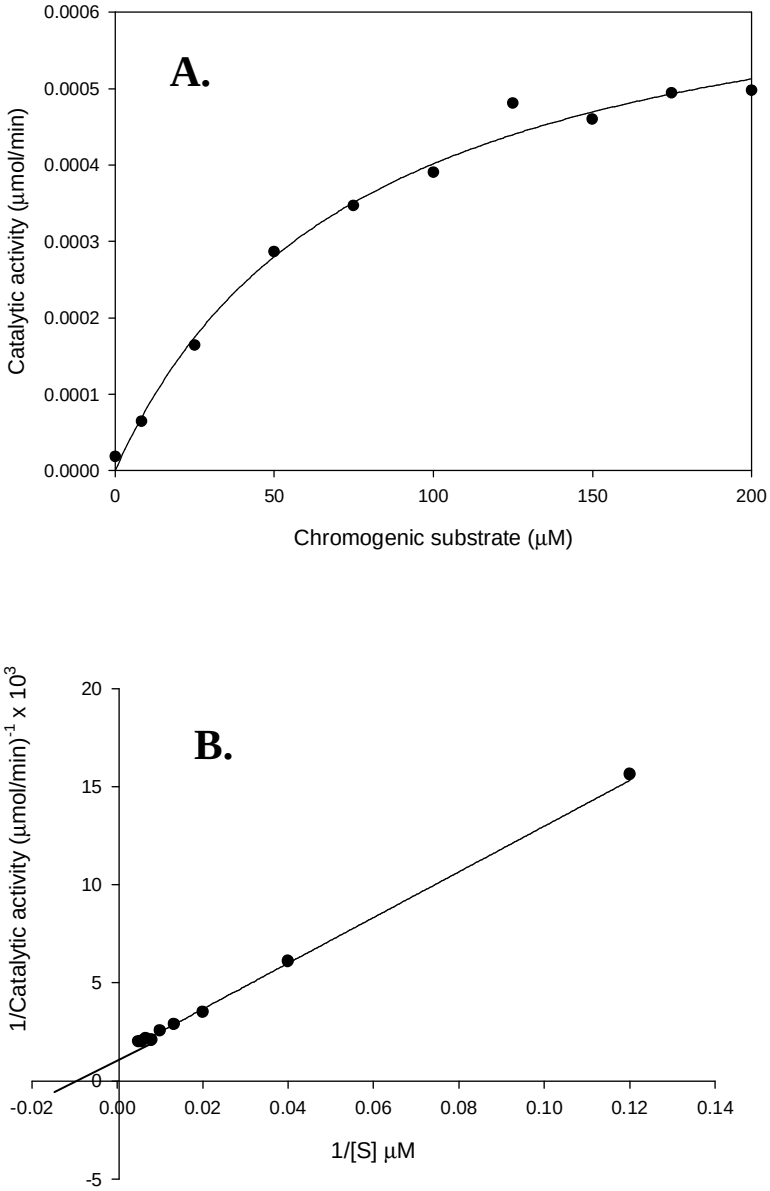


Figure 20. (**A**) Michaelis-Menten plot for the determination of K_m for wild-type C-SA HIV-1 protease towards the chromogenic substrate, where $V_{max}/2 = K_m$. (**B**) Lineweaver-Burke plot of 1/velocity versus 1/[chromogenic substrate] used to calculate K_m, where the x-intercept is $-1/K_m$. The solid line represents a linear fit to the experimental data. Correlation coefficient is 0.99.

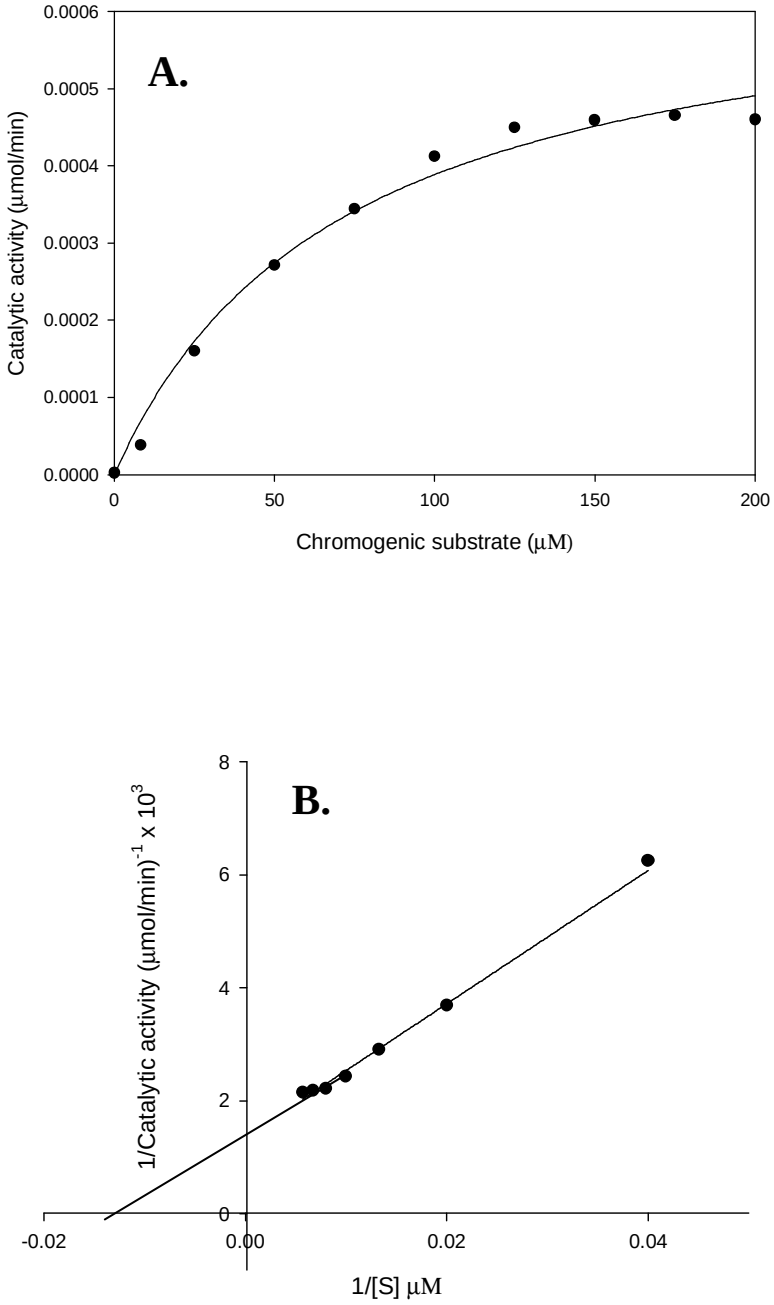


Figure 21. (**A**) Michaelis-Menten plot for the determination of K_m for V82A HIV-1 protease towards the chromogenic substrate, where $V_{max}/2 = K_m$. (**B**) Lineweaver-Burke plot of 1/velocity versus 1/[chromogenic substrate] used to calculate K_m, where the x-intercept is $-1/K_m$. The solid line represents a linear fit to the experimental data. Correlation coefficient is 0.98.

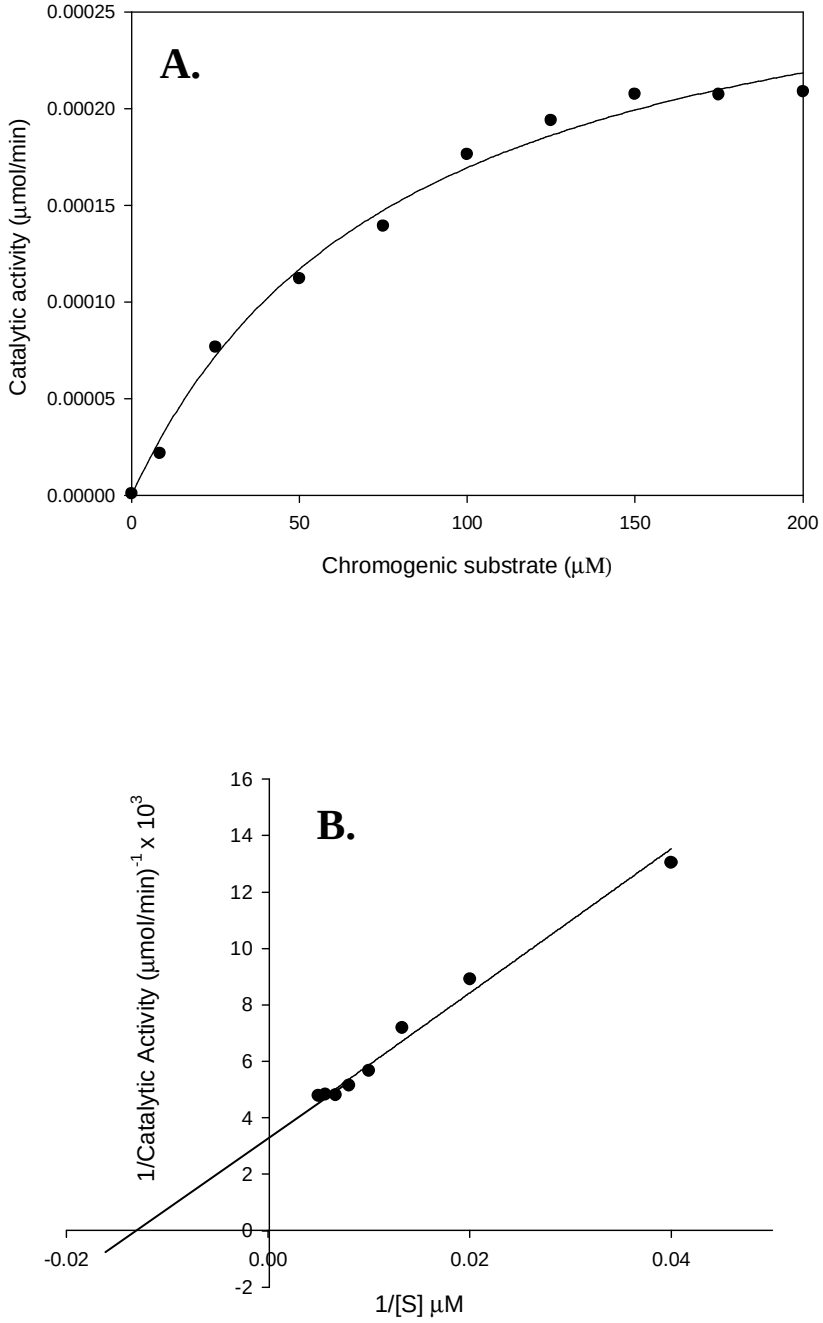


Figure 22. (**A**) Michaelis-Menten plot for the determination of K_m for V82F/I84V HIV-1 protease towards the chromogenic substrate, where $V_{max}/2 = K_m$. (**B**) Lineweaver-Burke plot of 1/velocity versus 1/[chromogenic substrate] used to calculate K_m, where the x-intercept is $-1/K_m$. The solid line represents a linear fit to the experimental data. Correlation coefficient is 0.98.

Table 3. Comparison of the inhibition characteristics (IC_{50}) of the wild-type C-SA, V82A and V82F/I84V HIV-1 protease

Inhibitor	wild-type C-SA (nM)	V82A (nM)	V82F/I84V (nM)
Saquinavir	0.12 (1)	0.28 (2.3)	12 (100)
Ritonavir	1.25 (1)	6.26 (5.0)	250 (200)
Indinavir	3.02 (1)	26.01 (8.6)	1500 (497)
Nelfinavir	2.34 (1)	2.87 (1.2)	240 (103)

The inhibition constants (IC_{50}) were obtained at 25 °C by measuring the decrease in absorbance at 300 nm associated with the cleavage of the chromogenic substrate. The experimental conditions were: 125 nM protease in 50 mM sodium acetate buffer, pH 5.0 including 50 μM substrate and increasing amounts of each inhibitor. The IC_{50} ratios ($IC_{50,\ mutant}/IC_{50,\ wild\text{-}type}$) for the mutant proteases are shown in parentheses.

Table 4: Inhibition constants (K_i) for the wild-type C-SA, V82A and V82F/I84V HIV-1 protease

Inhibitor	wild-type C-SA (nM)	V82A (nM)	V82F/I84V (nM)
Saquinavir	0.62 (1)	0.97 (1.6)	29 (47)
Ritonavir	0.50 (1)	1.08 (2.2)	97 (194)
Indinavir	0.49 (1)	1.20 (2.4)	219 (447)
Nelfinavir	0.52 (1)	0.90 (1.7)	56 (108)

The inhibition constants (K_i) were obtained at 25 °C by measuring the decrease in absorbance at 300 nm associated with the cleavage of the chromogenic substrate. The experimental conditions were: 15 – 40 nM protease in 50 mM sodium acetate buffer, pH 5.0, (0 - 200 μM) substrate and increasing amounts of each inhibitor. The K_i ratios ($K_{i,\ mutant}/K_{i,\ wild\text{-}type}$) for the mutant proteases are shown in parentheses.

3.4.3 Energetics of peptide inhibitor binding

Energetics of acetyl-pepstatin binding to the HIV-1 protease were measured by isothermal titration calorimetry (ITC). Acetyl-pepstatin is a synthetic nonpeptide inhibitor for HIV-1 proteinases (Luque *et al.*, 1998; Velazquez-Campoy *et al.*, 2000; Velazquez-Campoy *et al.*, 2001b). ITC experiments were designed to satisfy a c-value between 1 and 1000, which allows for accurate determinations of binding constants (Wiseman *et al.*, 1989). The c-value is defined as the product of association constant, K_a, and the macromolecule concentration, P, according to the following equation: c = K_a x [P]. The dissociation constant, K_d, which is an inverse of K_a, for association of acetyl-pepstatin (440 nM) with HIV-1 subtype B protease (Luque *et al.*, 1998) was used as a reference point for experimental design so as to satisfy a c-value of 50 for acetyl-pepstatin. Figures 23, 24 and 25 show typical calorimetric titrations of acetyl-pepstatin into the sample cell containing the wild-type, V82A and V82F/I84V HIV-1 C-SA proteases, respectively, in sodium acetate buffer, pH 5.0 at 25 °C. Each peak, in the top panel, represents the binding of acetyl-pepstatin to the protein and as the binding reaction progresses, the area under each peak decreases due to increased occupancy of available binding sites on HIV-1 protease. The bottom panel represents the binding isotherm obtained by plotting the integrated heat obtained after each injection versus the ratio of concentration of the inhibitor and protein added, corrected for heats of dilution (Figures 23, 24 and 25). The binding isotherms are best described as monophasic with a sigmoidal fit to the data representing the decrease in available binding sites on the protein. Association of acetyl-pepstatin with the wild-type, V82A and V82F/I84V HIV-1 proteases is characterized by successive endothermic heats representative of an unfavorable change in enthalpy. Thermodynamic parameters obtained from the data fitting as described by Luque *et al.* (1998) are given in Table 5. Briefly, analysis of the titration data for the wild-type HIV-1 C-SA yielded a Gibbs free energy of binding (ΔG) of – 8.7 kcal/mol, binding enthalpy (ΔH) of 7.7 kcal/mol, a binding entropy ($-T\Delta S$) of -16.4 kcal/mol and a K_d of 400 nM. These values are in agreement with those previously published by others (Luque *et al.*, 1998; Ohtaka *et al.*, 2003; Muzammil *et al.*, 2003). The stoichoimetry (N) obtained from the titration data is used to estimate the concentration of active sites in the protease molecule. For example, a stoichiometry of 1 corresponds to an active site concentration of 100 %. For all proteins, the concentrations of the active sites were over 80 % (see Table 5).

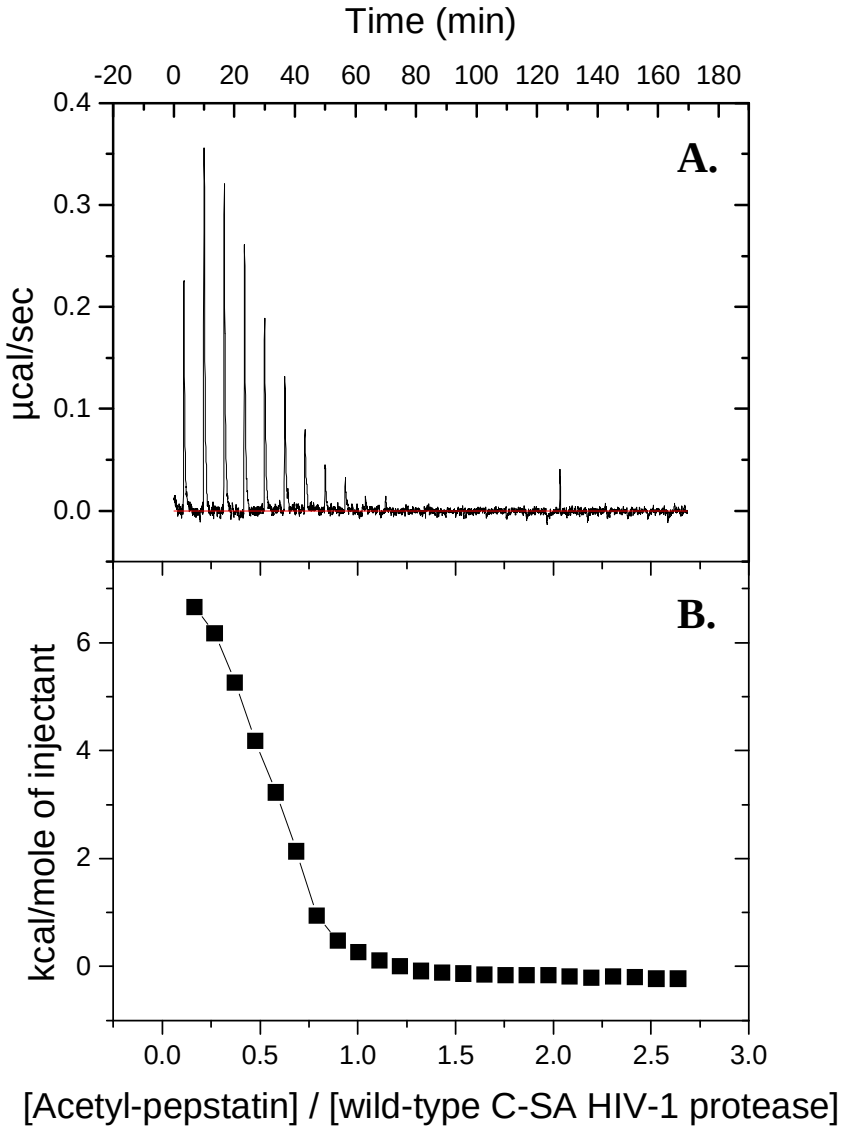


Figure 23. Representative calorimetric profile for the titration of wild-type HIV-1 C-SA protease with acetyl-pepstatin. The experiments were performed in 10 mM sodium acetate buffer containing 2 % DMSO, pH 5.0, at 25 ºC. (**A**) Top panel shows the raw data generated by titration of 20 µM protein with 10 µl injections of 300 µM acetyl-pepstatin stock in the experiment. (**B**) Bottom panel shows the integrated heats for the above peaks plotted against molar ratio of acetyl-pepstatin to protein molecule. The solid line through the data represents the best non-linear least squares fit to the experimental data using one-site binding model and ORIGIN v5.0 software.

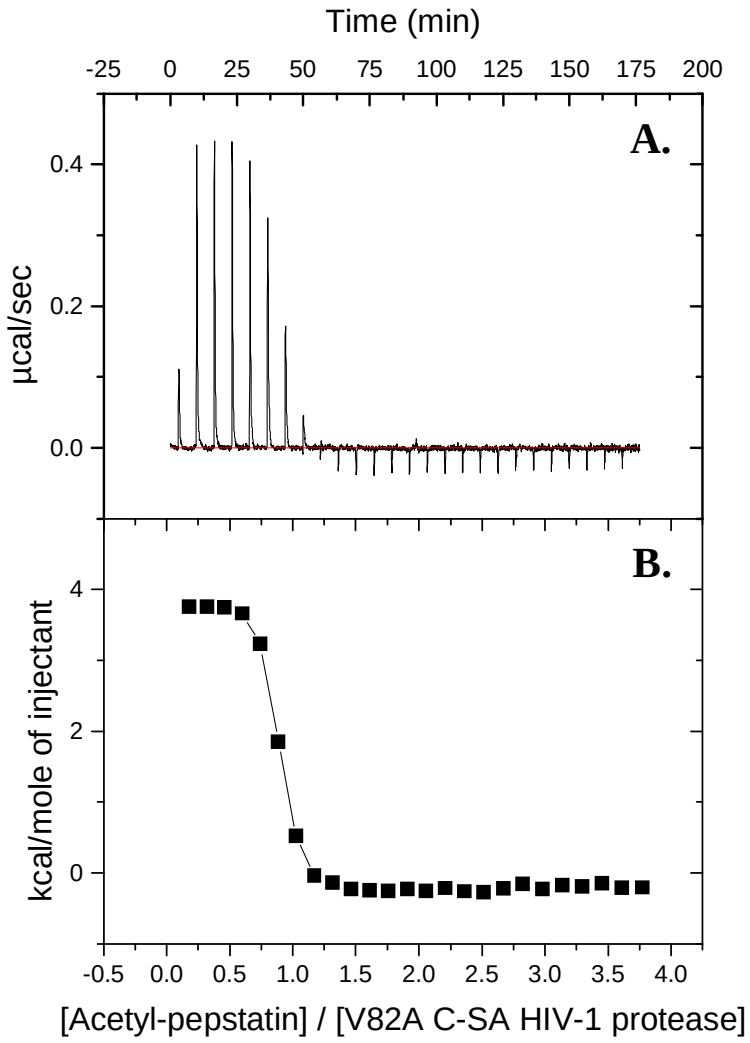


Figure 24. Representative calorimetric profile for the titration of the V82A HIV-1 protease with acetyl-pepstatin. The experiments were performed in 10 mM sodium acetate buffer containing 2 % DMSO, pH 5.0, at 25 ºC. (**A**) Top panel shows the raw data generated by titration of 20 µM protein with 10 µl injections of 300 µM acetyl-pepstatin stock in the experiment. (**B**) Bottom panel shows the integrated heats for the above peaks plotted against molar ratio of acetyl-pepstatin to protein molecule. The solid line through the data represents the best non-linear least squares fit to the experimental data using one-site binding model and ORIGIN v5.0 software.

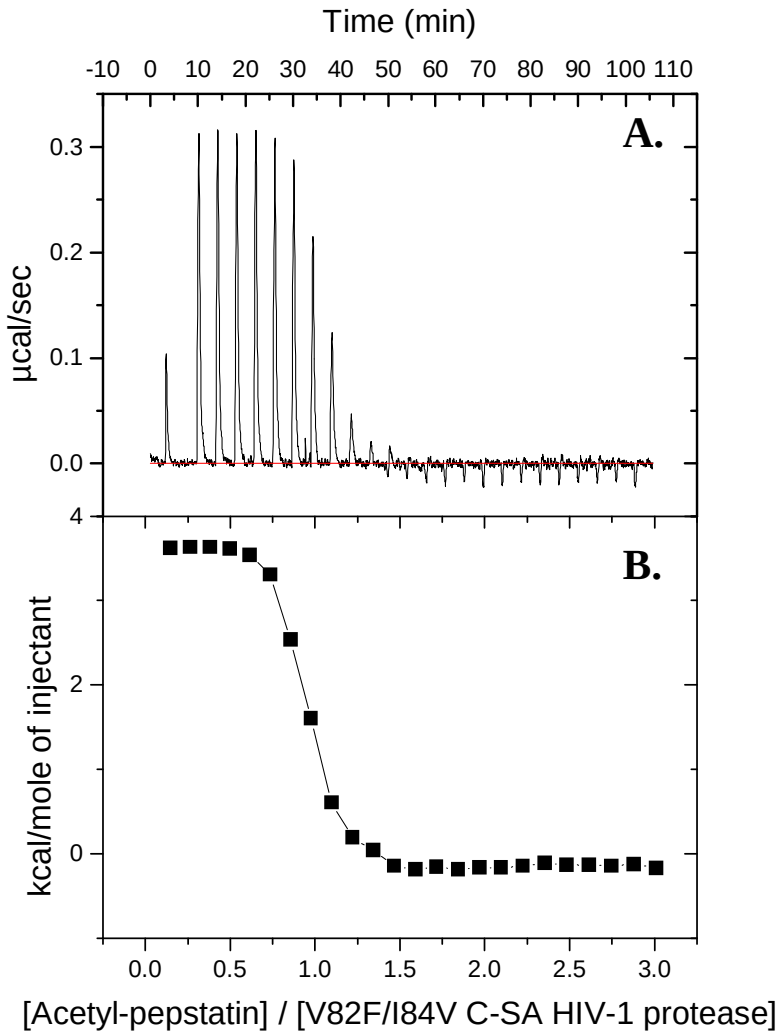


Figure 25. Representative calorimetric profile for the titration of the V82F/I84V HIV-1 protease with acetyl-pepstatin. The experiments were performed in 10 mM sodium acetate buffer containing 2 % DMSO, pH 5.0, at 25 ºC. (**A**) Top panel shows the raw data generated by titration of 20 µM protein with 10 µl injections of 300 µM acetyl-pepstatin stock in the experiment. (**B**) Bottom panel shows the integrated heats for the above peaks plotted against molar ratio of acetyl-pepstatin to protein molecule. The solid line through the data represents the best non-linear least squares fit to the experimental data using one-site binding model and ORIGIN v5.0 software.

Table 5: Thermodynamic parameters for the binding of acetyl-pepstatin to the HIV-1 proteases: wild-type B, wild-type C-SA, V82A and V82F/I84V

	wild-type B[a]	**wild-type C-SA**	**V82A**	**V82F/I84V**
K_a (M^{-1})	2.3 x 10^6	2.34 x 10^6	2.00 x 10^6	6.05 x 10^6
K_d (nM)	440	400	500	165
N	0.98	0.83	0.95	0.98
ΔG (kcal/mol)	-8.70	-8.70	-8.70	-9.30
ΔH (kcal/mol)	8.00	7.70	7.00	3.90
$-T\Delta S$ (kcal/mol)	-16.70	-16.40	-15.70	-13.20

Experiments were performed in 10 mM sodium acetate, pH 5.0, at 25 °C. [a]Values were obtained from studies done by Ohtaka *et al.* (2003)

A positive enthalpy indicates that the binding of acetyl-pepstatin is enthalpically unfavorable and that the favorable energy of binding originates from a large positive and favorable entropy of binding ($-T\Delta S$) with stoichiometry of 0.83 (Luque *et al.*, 1998). At 25 °C, the entropy change contributes 16.4 kcal/mol to the overall Gibbs energy of binding for the wild-type C-SA protease. The binding of acetyl-pepstatin to the V82A and V82F/I84V variant proteases is also endothermic and characterized by favorable entropy of binding (Table 5). The binding affinity for acetyl-pepstatin is also not significantly affected, with K_d ratios of 1.25 and 0.41 for the V82A and V82F/I84V variant proteases, respectively, relative to the wild-type C-SA HIV-1 protease (Table 5).

3.4.4 Energetics of inhibitor binding to wild-type C-SA, V82A and V82F/I84V HIV-1 proteases

Isothermal titration calorimetry experiments were performed in order to determine the binding affinity, binding enthalpy, Gibbs free energy of binding and stoichiometry of different inhibitors to the wild-type C-SA, V82A and V82F/I84V proteases. The four inhibitors used in this study (saquinavir, ritonavir, indinavir and nelfinavir) are inhibitors that bind the wild-type HIV-1 protease with high affinity as shown in section 3.4.2 (with $K_i \leq 1$ nM). Therefore, standard titration experiments do not provide accurate measurements of the binding affinity, even though the binding enthalpy can be determined with high accuracy. To overcome this hurdle, calorimetric displacement titrations were carried out, allowing for the determination of the affinity and enthalpy of binding, as described previously (Sigurskjold, 2000; Velazquez-Campoy *et al.*, 2001b; Ohtaka *et al.*, 2002). This technique allows complete determination of binding thermodynamics of high-affinity ligands ($K_a \geq 10^9$ M^{-1}) that are beyond the range of determination by direct titration.

In calorimetric displacement titrations, the high-affinity inhibitor is titrated into a protease sample prebound to a weaker binding inhibitor (acetyl-pepstatin), a well-characterized inhibitor of lower binding affinity and unfavorable binding enthalpy (Todd and Freire, 1999). The selection of a weak binding inhibitor with a binding isotherm of opposite sign (positive ΔH) produces a larger signal during the displacement reaction due to the displacement of the weaker binding inhibitor by an inhibitor yielding an exothermic isotherm (negative ΔH) (see Figure 26 for a diagrammatic representation of a displacement titration).

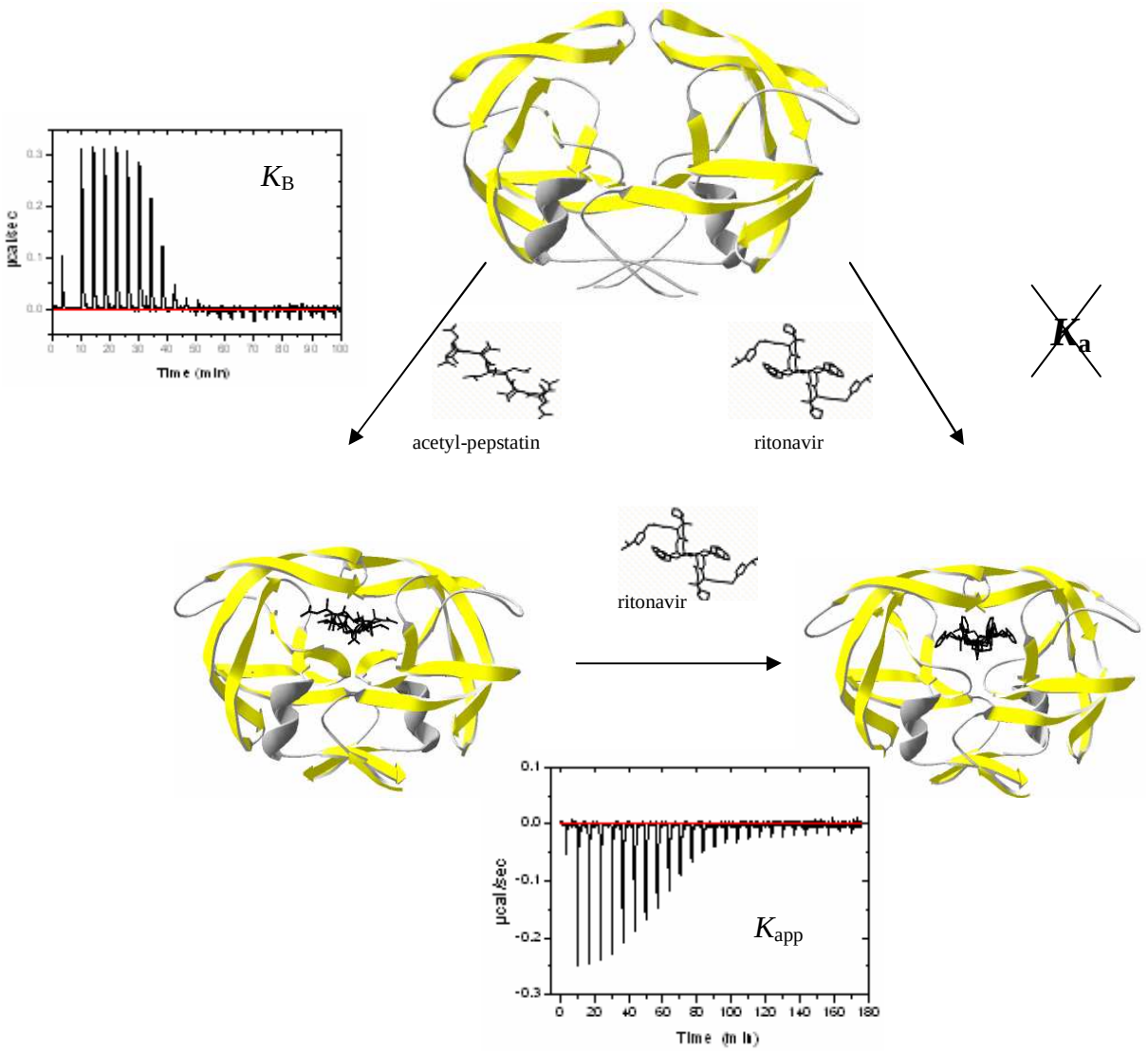


Figure 26. Overview of a displacement titration assay for HIV-1 protease. The binding affinity of ritonavir, K_a, is beyond the limit of direct calorimetric determination. The displacement titration experiment is performed in the presence of the weak binding inhibitor acetyl-pepstatin (K_B = 2.0 x 10^6 M^{-1}). The selection of a weak binding inhibitor with a binding enthalpy of opposite isotherm (positive ΔH) to that of the inhibitor of interest (negative ΔH) produces a larger signal (bottom) due to the displacement of a weak binding inhibitor by an inhibitor yielding an exothermic isotherm. In the presence of the weak binding inhibitor, the apparent binding constant for the inhibitor which binds tightly, K_{app}, falls within the range required for ITC determination. K_{app} is given by the equation: $K_{app} = K_a/(1 + K_B[B])$, where B is the concentration of the weaker binding inhibitor. In addition, K_{app} can be lowered to the desired level by increasing the concentration of the weak inhibitor. This diagram was adapted from Velazquez-Campoy *et al.* (2001b).

In addition, the binding isotherm, in this case, has sufficient curvature to allow for the calorimetric measurement of the apparent binding affinity of the stronger binding ligand (Sigurskjold, 2000 and references therein). Two calorimetric titrations are required to solve the binding competition equations and determine the binding affinity and binding enthalpy of the tight binding inhibitor: (1) titration of the weak binding inhibitor into the protease and (2) titration of the inhibitor of interest into the protease prebound to the weak binding inhibitor. The competition experiments were also performed at pH 5.0 using an acetate buffer with negligible binding enthalpy in order to minimize any proton coupling effect on the observed binding enthalpy (Velazquez-Campoy *et al.*, 2001b). Figures 27, 28, 29 and 30 show typical displacement titrations for saquinavir, ritonavir, indinavir and nelfinavir into the wild-type C-SA HIV-1 protease, respectively, in the presence of acetyl-pepstatin, pH 5.0. Each peak in the top panel represents the displacement of a weaker binding inhibitor (acetyl-pepstatin) from the active site of the protein by the tight-binding inhibitor (e.g. saquinavir, with high binding affinity). As the titration progresses, the area under each peak becomes smaller due to increased occupancy of the available binding sites on the enzyme by the inhibitor of interest. The bottom panel in each respective figure shows the calorimetric binding isotherm obtained by plotting integrated heats obtained after each injection as a function of inhibitor concentration of interest per protein dimer. For the wild-type C-SA protease, the experimental data fit best to a single-site displacement binding model; i.e., with stoichiometry of 1:1. The binding isotherms are monophasic with a sigmoidal fit to the data representing the decrease in available binding sites on the protein as the reaction progresses to completion. Thermodynamic parameters obtained from the fit are summarized in Table 6. Clinical inhibitors, saquinavir, ritonavir, indinavir and nelfinavir bind to the wild-type C-SA HIV-1 protease with high binding affinities, K_a, (ranging from 0.2 x 10^9 M^{-1} to 0.3 x 10^9 M^{-1}) in a process strongly favored by entropic contributions, contributing about 12 kcal/mol for saquinavir and indinavir and 9 kcal/mol for ritonavir and 13 kcal/mol for nelfinavir to the overall Gibbs energy of binding at 25 °C. At 25 °C, the stoichiometry of binding for HIV-1 protease inhibitors ranges from 0.9 to 1.3 interpreted as one molecule of each inhibitor bound per dimer of HIV-1 protease and is consistent with crystallographic data (Baldwin *et al.*, 1995b; Chen *et al.*, 1995; Hong *et al.*, 1996; 1997; 1998). Accordingly, the binding of saquinavir, ritonavir, indinavir and nelfinavir is favored by entropic contributions of -18.8 kcal/mol, -9.1 kcal/mol, -15.0

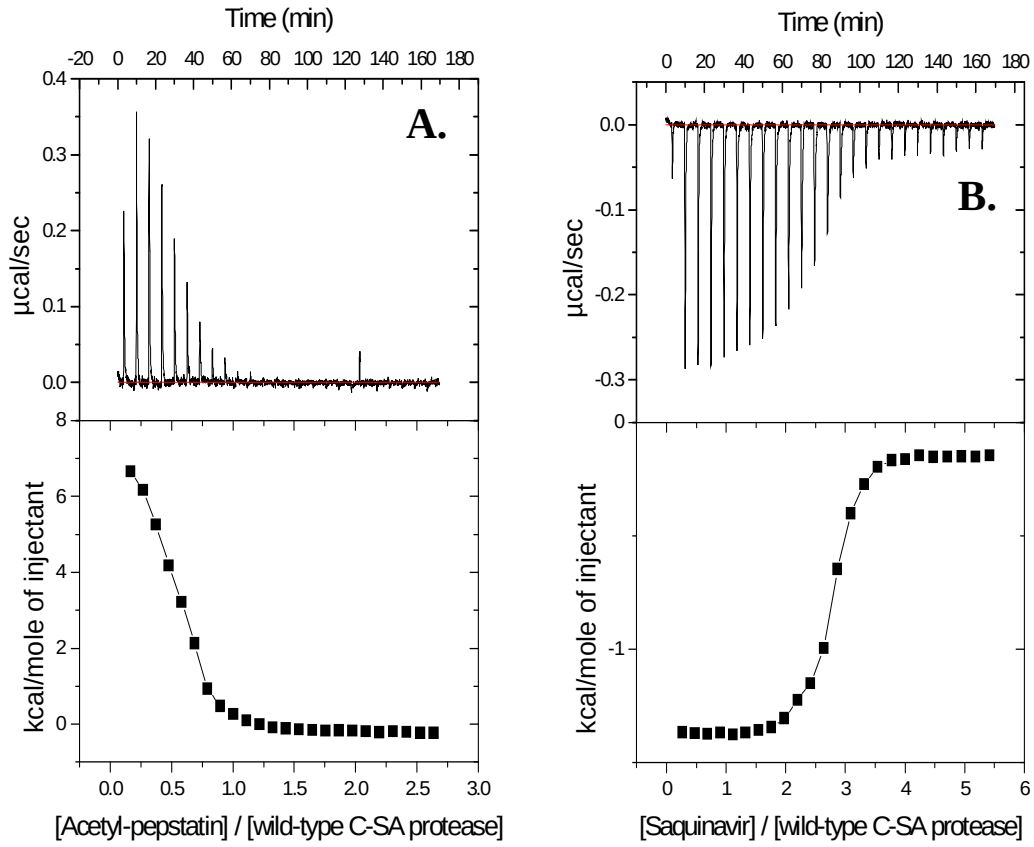


Figure 27. (**A**) A representative calorimetric profile of the titration of the wild-type C-SA HIV-1 protease with acetyl-pepstatin. Titrations of acetyl-pepstatin (300 µM) into protease solution (20 µM). The top panel shows raw endothermic heats generated upon each injection of acetyl-pepstatin into wild-type C-SA HIV-1 protease. The bottom panel shows the integrated heats for the above peaks plotted against molar ratio of acetyl-pepstatin to protein molecule. The solid line through the data represents the best fit using a one-site binding model. (**B**) ITC displacement calorimetric titration of saquinavir (300 µM) into a solution of the wild-type C-SA HIV-1 protease (20 µM) pre-bound to acetyl-pepstatin (200 µM). The top panel shows the exothermic heat effects associated with the addition of saquinavir into the wild-type C-SA HIV-1 protease pre-bound to acetyl-pepstatin. The bottom panel shows the integrated heats for the above peaks plotted against molar ratio of saquinavir to protein dimer. The solid line through the data represents the best fit using the displacement binding model. The experiments were performed in 10 mM sodium acetate buffer containing 2 % DMSO, pH 5.0, at 25 ºC. The data were analyzed using ORIGIN v5.0 software as explained previously (Velazquez-Campoy *et al.*, 2001b).

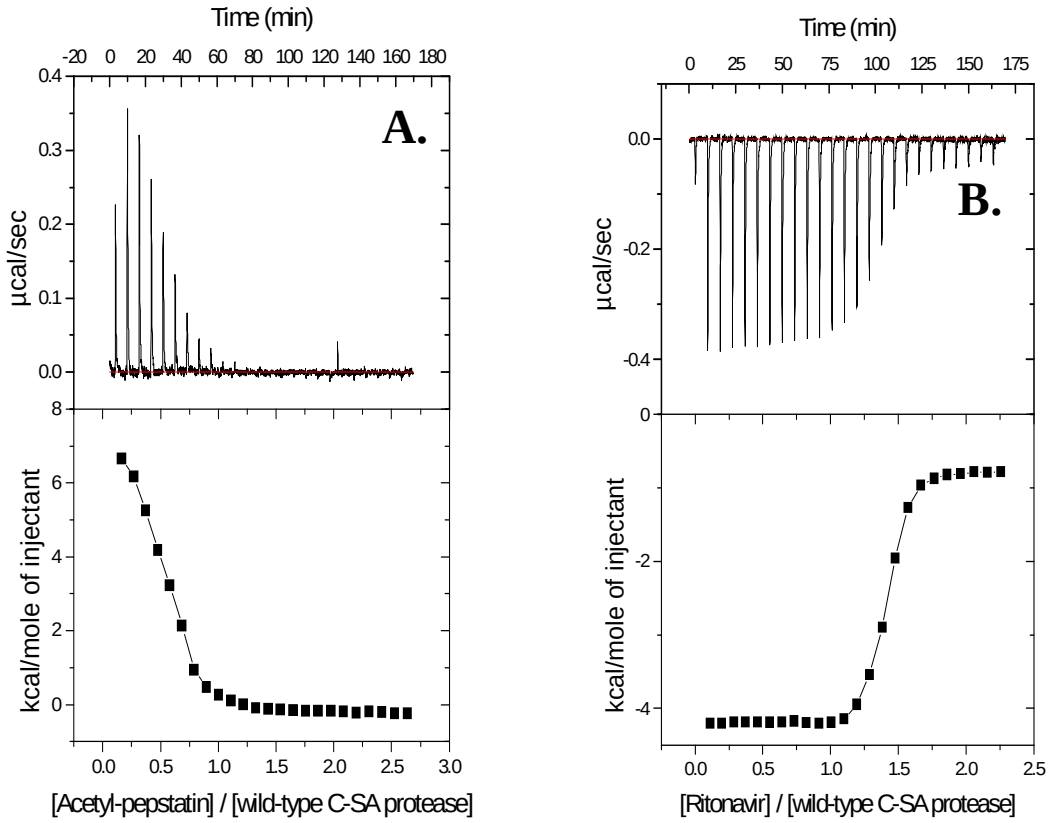


Figure 28. (**A**) A representative calorimetric profile of the titration of the wild-type C-SA HIV-1 protease with acetyl-pepstatin. Titrations of acetyl-pepstatin (300 µM) into protease solution (20 µM). The top panel shows raw endothermic heats generated upon each injection of acetyl-pepstatin into wild-type C-SA HIV-1 protease. The bottom panel shows the integrated heats for the above peaks plotted against molar ratio of acetyl-pepstatin to protein molecule. The solid line through the data represents the best fit using a one-site binding model. (**B**) ITC displacement calorimetric titration of ritonavir (250 µM) into a solution of the wild-type C-SA HIV-1 protease (20 µM) pre-bound to acetyl-pepstatin (200 µM). The top panel shows the exothermic heat effects associated with the addition of ritonavir into the wild-type C-SA HIV-1 protease pre-bound to acetyl-pepstatin. The bottom panel shows the integrated heats for the above peaks plotted against molar ratio of ritonavir to protein dimer. The solid line through the data represents the best fit using the displacement binding model. The experiments were performed in 10 mM sodium acetate buffer containing 2 % DMSO, pH 5.0, at 25 ºC. The data were analyzed using ORIGIN v5.0 software as explained previously (Velazquez-Campoy *et al.*, 2001b).

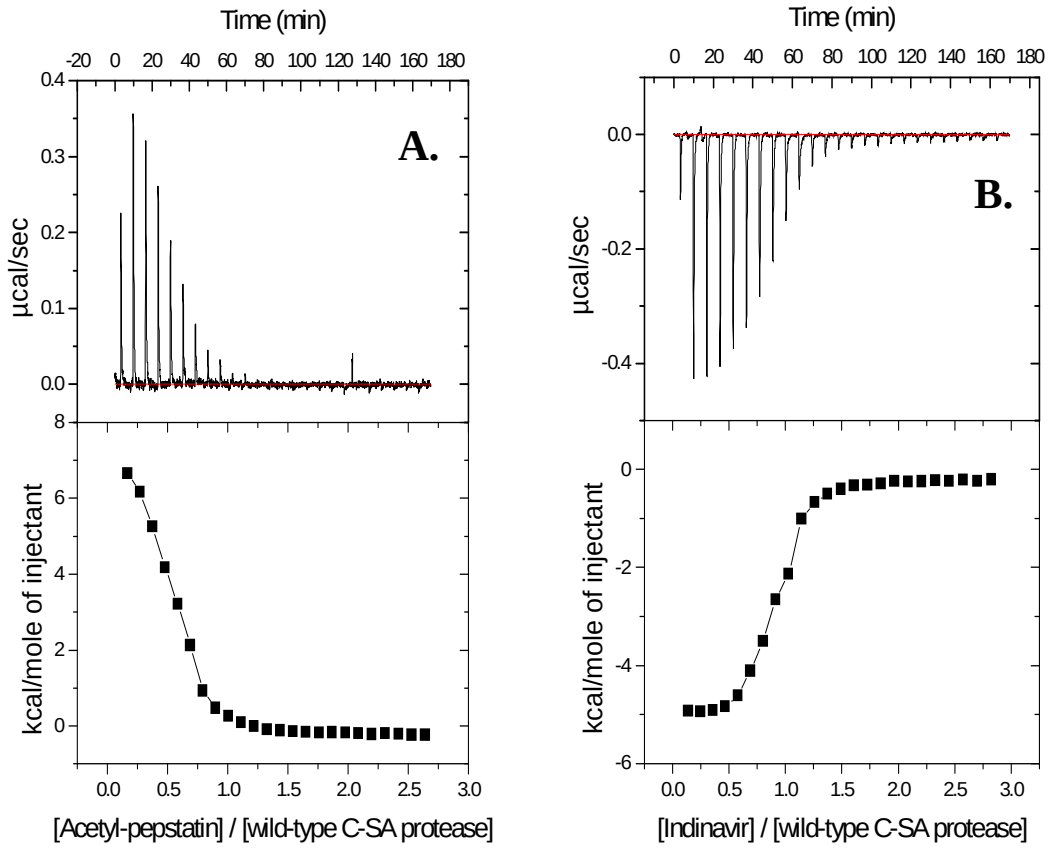


Figure 29. (**A**) A representative calorimetric profile of the titration of the wild-type C-SA HIV-1 protease with acetyl-pepstatin. Titrations of acetyl-pepstatin (300 µM) into protease solution (20 µM). The top panel shows raw endothermic heats generated upon each injection of acetyl-pepstatin into wild-type C-SA HIV-1 protease. The bottom panel shows the integrated heats for the above peaks plotted against molar ratio of acetyl-pepstatin to protein molecule. The solid line through the data represents the best fit using a one-site binding model. (**B**) ITC displacement calorimetric titration of indinavir (250 µM) into a solution of the wild-type C-SA HIV-1 protease (20 µM) pre-bound to acetyl-pepstatin (200 µM). The top panel shows the exothermic heat effects associated with the addition of indinavir into the wild-type C-SA HIV-1 protease prebound to acetyl-pepstatin. The bottom panel shows the integrated heats for the above peaks plotted against molar ratio of indinavir to protein dimer. The experiments were performed in 10 mM sodium acetate buffer containing 2 % DMSO, pH 5.0, at 25 ºC. The solid line through the data represents the best fit using the displacement binding model. The data were analyzed using ORIGIN v5.0 software as explained previously (Velazquez-Campoy *et al.*, 2001b).

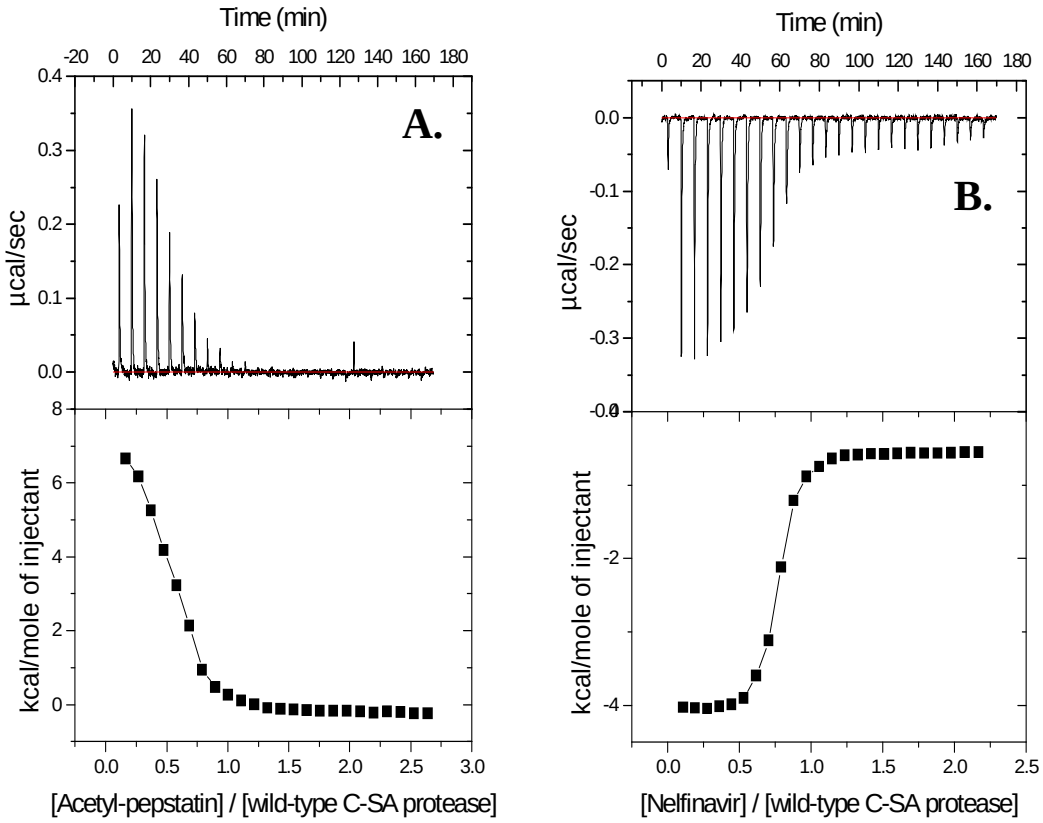


Figure 30. (**A**) A representative calorimetric profile of the titration of the wild-type C-SA HIV-1 protease with acetyl-pepstatin. Titrations of acetyl-pepstatin (300 µM) into protease solution (20 µM). The top panel shows raw endothermic heats generated upon each injection of acetyl-pepstatin into wild-type C-SA HIV-1 protease. The bottom panel shows the integrated heats for the above peaks plotted against molar ratio of acetyl-pepstatin to protein molecule. The solid line through the data represents the best fit using a one-site binding model. (**B**) ITC displacement calorimetric titration of nelfinavir (240 µM) into a solution of the wild-type C-SA HIV-1 protease (20 µM) pre-bound to acetyl-pepstatin (200 µM). The top panel shows the exothermic heat effects associated with the addition of nelfinavir into the wild-type C-SA HIV-1 protease pre-bound to acetyl-pepstatin. The bottom panel shows the integrated heats for the above peaks plotted against molar ratio of nelfinavir to protein dimer. The solid line through the data represents the best fit using the displacement binding model. The experiments were performed in 10 mM sodium acetate buffer containing 2 % DMSO, pH 5.0, at 25 ºC. The data were analyzed using ORIGIN v5.0 software as explained previously (Velazquez-Campoy *et al.*, 2001b).

Table 6: Thermodynamics parameters for the binding of inhibitors to the wild-type B[a], wild-type C-SA, V82A and V82F/I84V HIV-1 proteases

Inhibitor	wild-type B	wild-type C-SA	V82A	V82F/I84V
***K*$_d$ (nM)**				
Saquinavir	0.40	0.84	1.85	98
Ritonavir	0.03	0.21	2.00	230
Indinavir	0.48	0.95	12.70	450
Nelfinavir	0.26	0.30	0.66	110
***K*$_d$ ratio**				
Saquinavir	0.48	1.0	2.2	117
Ritonavir	0.14	1.0	9.5	1095
Indinavir	0.51	1.0	13.4	474
Nelfinavir	0.87	1.0	2.2	367
Δ*G* (kcal/mol)				
Saquinavir	-12.80	-12.40	-12.00	-9.60
Ritonavir	-14.40	-13.20	-11.90	-9.10
Indinavir	-12.70	-12.30	-10.80	-8.70
Nelfinavir	-13.10	-13.00	-12.50	-9.50
Δ*H* (kcal/mol)				
Saquinavir	1.90	6.40	5.50	0.80
Ritonavir	-3.70	-4.20	-2.20	0.60
Indinavir	2.10	2.70	2.90	3.80
Nelfinavir	2.60	4.10	4.30	4.00
-*T*Δ*S* (kcal/mol)				
Saquinavir	-14.70	-18.80	-17.50	-10.40
Ritonavir	-10.70	-9.10	-9.70	-9.60
Indinavir	-14.80	-15.00	-13.70	-12.50
Nelfinavir	-15.70	-17.10	-16.90	-13.50

Experiments were performed in 10 mM sodium acetate and 2 % DMSO, pH 5.0 at 25°C. [a]Thermodynamic parameters for the wild-type B HIV-1 protease were taken from Velazquez-Campoy *et al.* (2002).

kcal/mol and -17.1 kcal/mol, respectively. The binding of saquinavir, indinavir and nelfinavir to the wild-type HIV-1 C-SA protease is characterized by a positive (unfavorable) enthalpy change of 6.4 kcal/mol, 2.7 kcal/mol and 4.1 kcal/mol, respectively, whilst the binding of ritonavir is characterized by a slightly favorable enthalpy change of -4.2 kcal/mol. This in agreement with the thermodynamic data obtained previously, which showed entropically controlled binding affinities and unfavorable or slightly favorable binding enthalpies (Todd *et al.,* 2000). For these inhibitors, (saquinavir, ritonavir, indinavir and nelfinavir), entropy ($-T\Delta S$) contributions as large as -16 kcal/mol were required to compensate for the unfavorable binding enthalpies (Todd *et al.,* 2000). Figures 31, 32, 33 and 34 show the results of displacement calorimetric titrations of saquinavir, ritonavir, indinavir and nelfinavir, respectively into the V82A HIV-1 protease prebound to acetyl-pepstatin under the same conditions used for the wild-type C-SA HIV-1 protease. The V82A mutation lowers the Gibbs energy of binding for the respective four clinical inhibitors by 0.4 kcal/mol, 1.3 kcal/mol, 1.5 kcal/mol and 0.6 kcal/mol, respectively, relative to the wild-type C-SA HIV-1 protease (see Table 6). The V82A mutation does not seem to significantly affect the binding of inhibitors when compared to the V82F/I84V mutation. The binding affinity for saquinavir and nelfinavir is reduced 2.2 fold, while the binding affinity for ritonavir and indinavir is reduced 10- and 13-fold, respectively. In addition, the stoichiometry of binding ranges from 0.9 to 1.1 for the four clinical inhibitors (saquinavir, ritonavir, indinavir and nelfinavir) in association with the V82A HIV-1 protease. The effect of the V82A mutation on ΔH amounts to 2 kcal/mol for ritonavir, 0.2 kcal/mol for both indinavir and nelfinavir. There was an observable gain in enthalpy change of 0.9 kcal/mol for saquinavir. On the other hand, the effect on $-T\Delta S$ amounts to 1.3 kcal/mol for saquinavir and indinavir, and 0.2 kcal/mol for nelfinavir. There was a slight gain in $-T\Delta S$ of 0.6 kcal/mol for ritonavir.

Interestingly, the V82F/I84V mutation causes a drastic drop in the binding affinity of the four inhibitors used in this study sufficient to allow measurement of their thermodynamic energetics by direct calorimetric titrations (see Figures 35, 36, 37 and 38). All data from each experiment also fit well to a one-site binding model. The binding isotherms are best described as monophasic with a sigmoidal curve representing the decrease in available binding sites on the protein molecule as the titration progresses to completion.

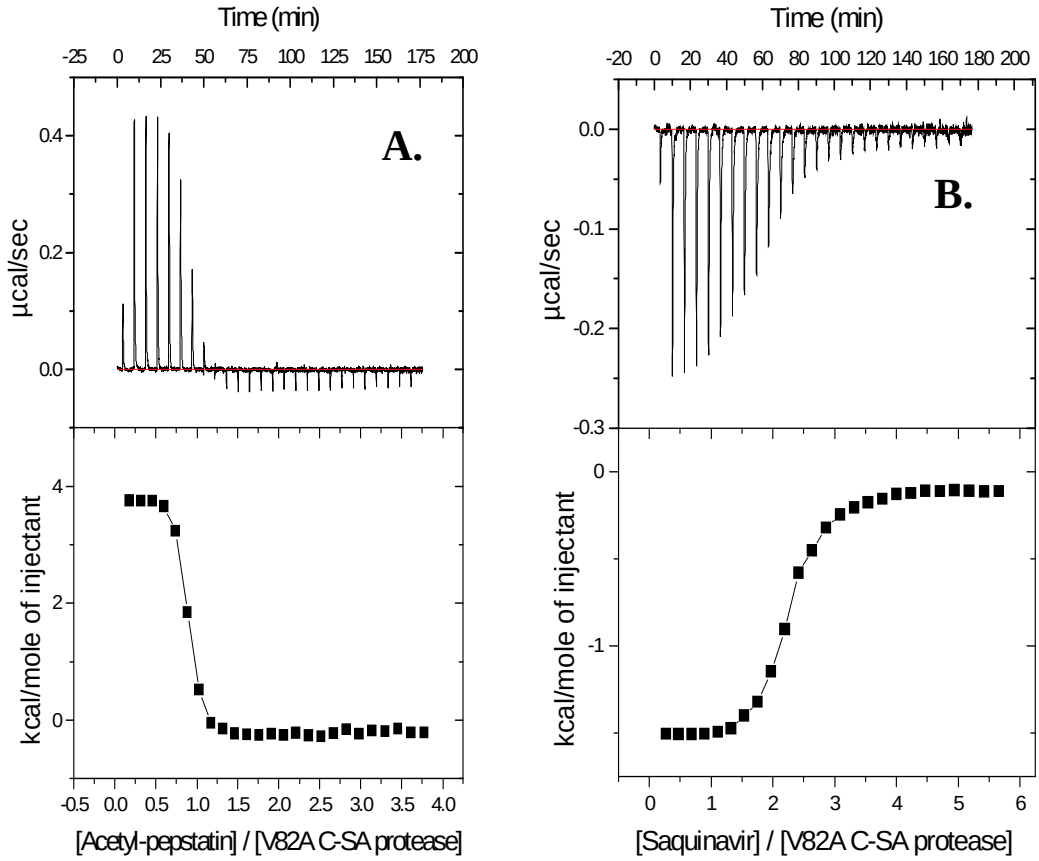


Figure 31. (**A**) A representative calorimetric profile of the titration of the V82A HIV-1 protease with acetyl-pepstatin. Titrations of acetyl-pepstatin (300 µM) into protease solution (20 µM). The top panel shows raw endothermic heats generated upon each injection of acetyl-pepstatin into V82A HIV-1 protease. The bottom panel shows the integrated heats for the above peaks plotted against molar ratio of acetyl-pepstatin to protein molecule. The solid line through the data represents the best fit using a one-site binding model. (**B**) ITC displacement calorimetric titration of saquinavir (300 µM) into a solution of the V82A HIV-1 protease (20 µM) pre-bound to acetyl-pepstatin (200 µM). The top panel shows the exothermic heat effects associated with the addition of saquinavir into the V82A HIV-1 protease pre-bound to acetyl-pepstatin. The bottom panel shows the integrated heats for the above peaks plotted against molar ratio of saquinavir to protein dimer. The solid line through the data represents the best fit using the displacement binding model. The experiments were performed in 10 mM sodium acetate buffer containing 2 % DMSO, pH 5.0, at 25 ºC. The data were analyzed using ORIGIN v5.0 software as explained previously (Velazquez-Campoy *et al.*, 2001b).

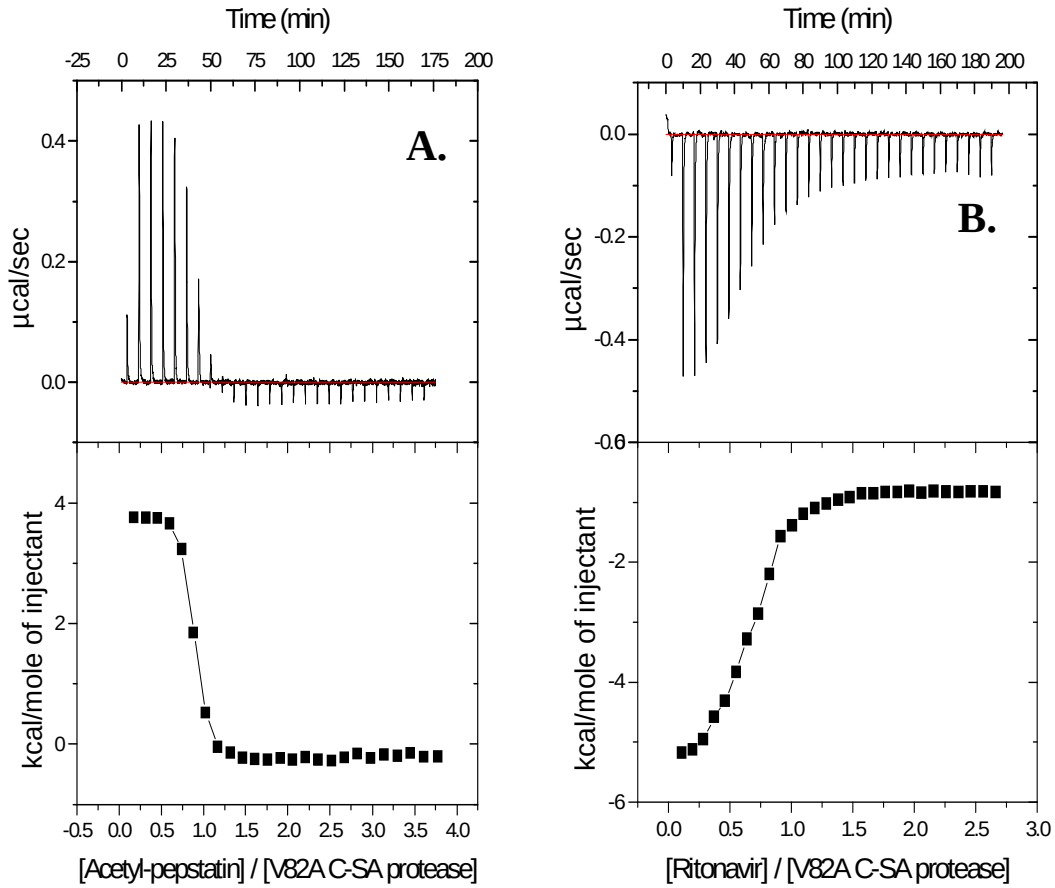


Figure 32. (**A**) A representative calorimetric profile of the titration of the V82A HIV-1 protease with acetyl-pepstatin. Titrations of acetyl-pepstatin (300 µM) into protease solution (20 µM). The top panel shows raw endothermic heats generated upon each injection of acetyl-pepstatin into V82A HIV-1 protease. The bottom panel shows the integrated heats for the above peaks plotted against molar ratio of acetyl-pepstatin to protein molecule. The solid line through the data represents the best fit using a one-site binding model. (**B**) ITC displacement calorimetric titration of ritonavir (250 µM) into a solution of the V82A HIV-1 protease (20 µM) pre-bound to acetyl-pepstatin (200 µM). The top panel shows the exothermic heat effects associated with the addition of ritonavir into the V82A HIV-1 protease pre-bound to acetyl-pepstatin. The bottom panel shows the integrated heats for the above peaks plotted against molar ratio of ritonavir to protein dimer. The solid line through the data represents the best fit using the displacement binding model. The experiments were performed in 10 mM sodium acetate buffer containing 2 % DMSO, pH 5.0, at 25 ºC. The data were analyzed using ORIGIN v5.0 software as explained previously (Velazquez-Campoy *et al.*, 2001b).

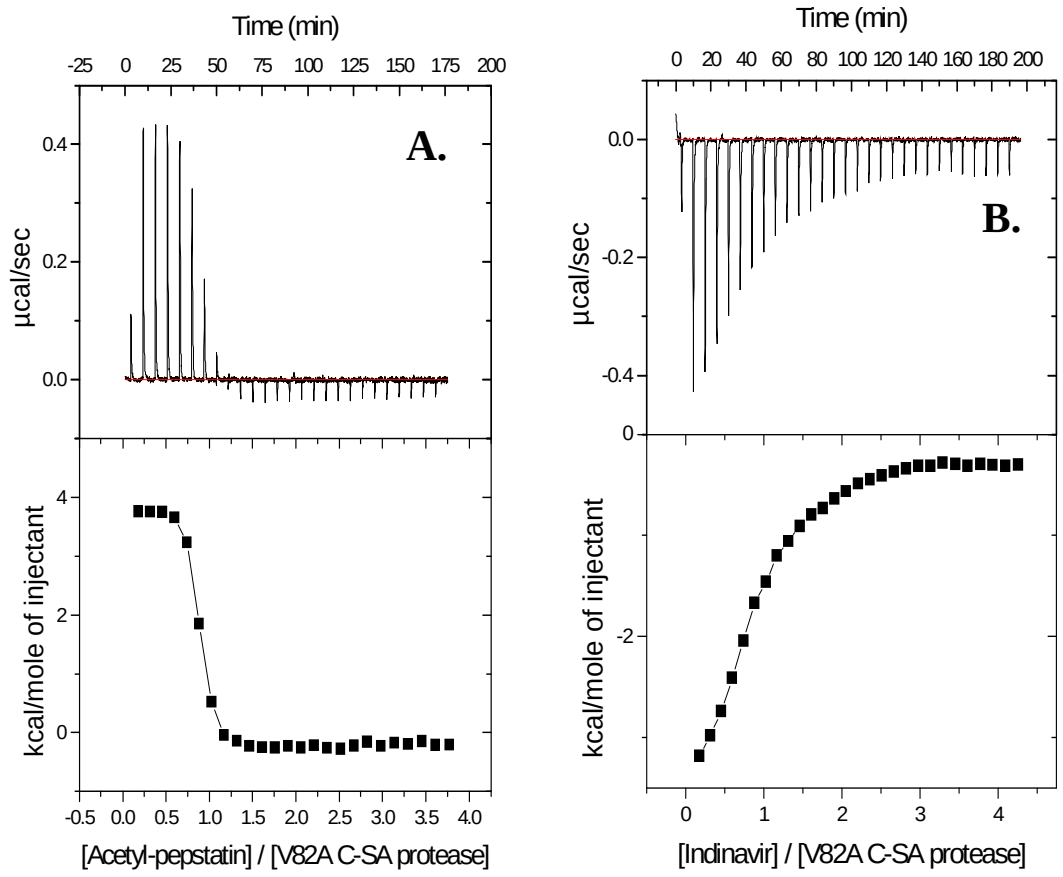


Figure 33. (**A**) A representative calorimetric profile of the titration of the V82A HIV-1 protease with acetyl-pepstatin. Titrations of acetyl-pepstatin (300 µM) into protease solution (20 µM). The top panel shows raw endothermic heats generated upon each injection of acetyl-pepstatin into V82A HIV-1 protease. The bottom panel shows the integrated heats for the above peaks plotted against molar ratio of acetyl-pepstatin to protein molecule. The solid line through the data represents the best fit using a one-site binding model. (**B**) ITC displacement calorimetric titration of indinavir (400 µM) into a solution of the V82A HIV-1 protease (20 µM) pre-bound to acetyl-pepstatin (200 µM). The top panel shows the exothermic heat effects associated with the addition of indinavir into the V82A HIV-1 protease pre-bound to acetyl-pepstatin. The bottom panel shows the integrated heats for the above peaks plotted against molar ratio of indinavir to protein dimer. The solid line through the data represents the best fit using the displacement binding model. The experiments were performed in 10 mM sodium acetate buffer containing 2 % DMSO, pH 5.0, at 25 ºC. The data were analyzed using ORIGIN v5.0 software as explained previously (Velazquez-Campoy *et al.*, 2001b).

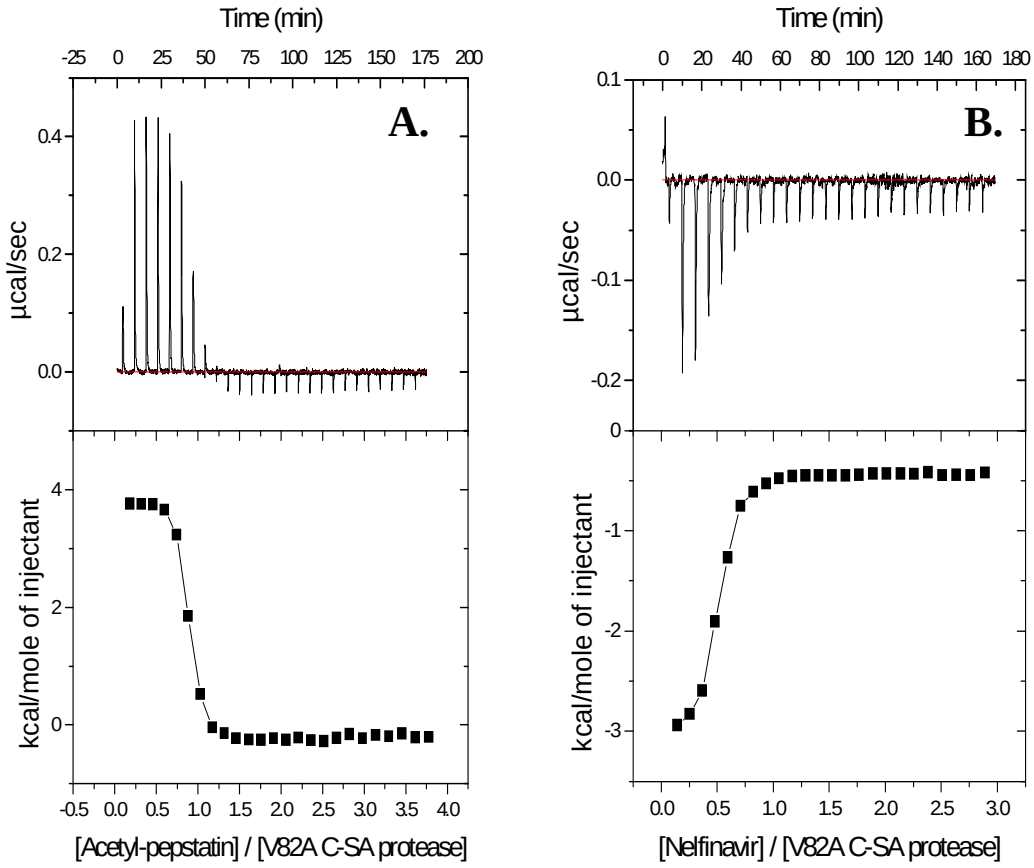


Figure 34. (**A**) A representative calorimetric profile of the titration of the V82A HIV-1 protease with acetyl-pepstatin. Titrations of acetyl-pepstatin (300 µM) into protease solution (20 µM). The top panel shows raw endothermic heats generated upon each injection of acetyl-pepstatin into V82A HIV-1 protease. The bottom panel shows the integrated heats for the above peaks plotted against molar ratio of acetyl-pepstatin to protein molecule. The solid line through the data represents the best fit using a one-site binding model. (**B**) ITC displacement calorimetric titration of nelfinavir (240 µM) into a solution of the V82A HIV-1 protease (20 µM) pre-bound to acetyl-pepstatin (200 µM). The top panel shows the exothermic heat effects associated with the addition of nelfinavir into the V82A HIV-1 protease pre-bound to acetyl-pepstatin. The bottom panel shows the integrated heats for the above peaks plotted against molar ratio of nelfinavir to protein dimer. The solid line through the data represents the best fit using the displacement binding model. The experiments were performed in 10 mM sodium acetate buffer containing 2 % DMSO, pH 5.0, at 25 ºC. The data were analyzed using ORIGIN v5.0 software as explained previously (Velazquez-Campoy *et al.*, 2001b).

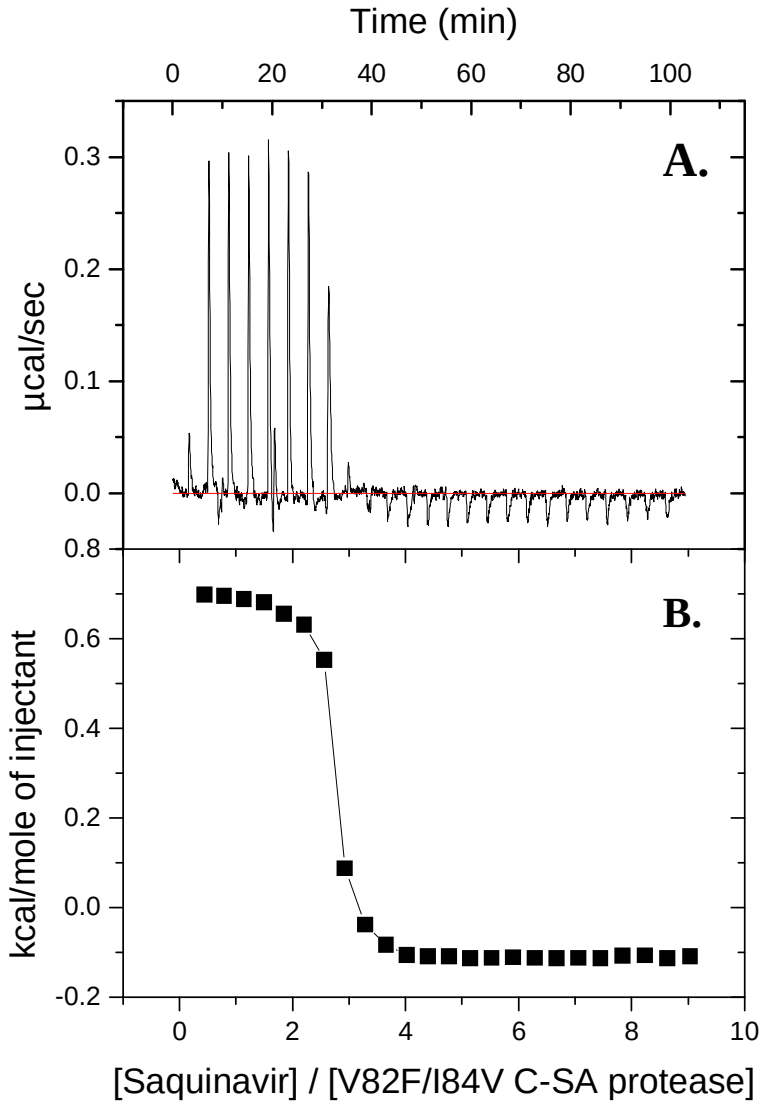


Figure 35. Representative calorimetric profile for the direct titration of the V82F/I84V HIV-1 protease with saquinavir. The experiments were performed in 10 mM sodium acetate buffer containing 2 % DMSO, pH 5.0, at 25 ºC. (**A**) Top panel shows the raw data generated by titration of 20 µM protein with 10 µl injections of 500 µM saquinavir stock in the experiment. (**B**) Bottom panel shows the integrated heats for the above peaks plotted against molar ratio of saquinavir to protein molecule. The solid line through the data represents the best non-linear least squares fit to the experimental data using one-site binding model and ORIGIN v5.0 software.

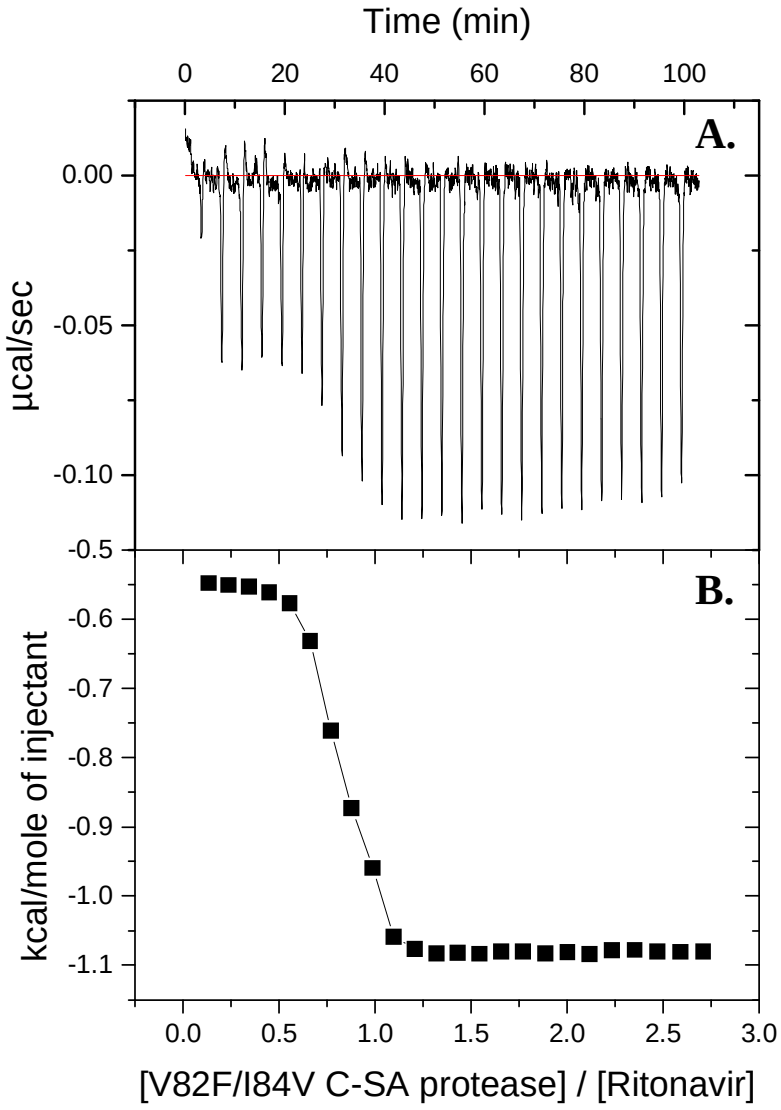


Figure 36. Representative calorimetric profile for the direct titration of ritonavir with the V82F/I84V HIV-1 protease. The experiments were performed in 10 mM sodium acetate buffer containing 2 % DMSO, pH 5.0, at 25 ºC. Because of its low solubility, the binding of ritonavir was measured by titration of 20 µM ritonavir in the calorimetric cell with 300 µM of the protein solution. (**A**) Top panel shows the raw data generated by titration of ritonavir with 10 µl injections of protein stock in the experiment. (**B**) Bottom panel shows the integrated heats for the above peaks plotted against molar ratio of protein molecule to ritonavir. The solid line through the data represents the best non-linear least squares fit to the experimental data using one-site binding model and ORIGIN v5.0 software.

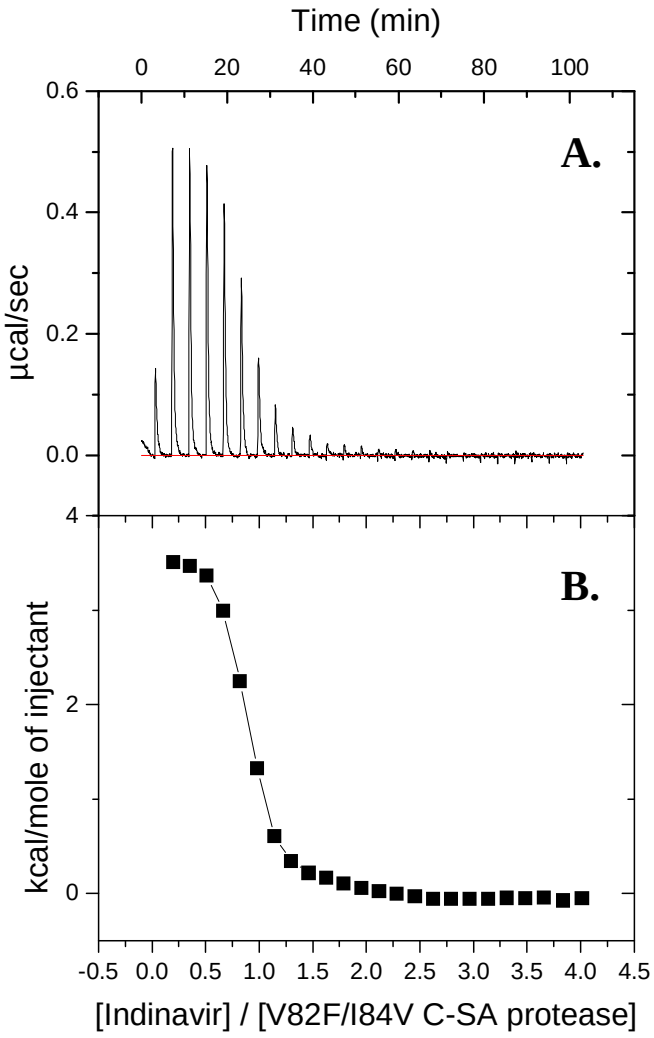


Figure 37. Representative calorimetric profile for the direct titration of the V82F/I84V HIV-1 protease with indinavir. The experiments were performed in 10 mM sodium acetate buffer containing 2 % DMSO, pH 5.0, at 25 ºC. (**A**) Top panel shows the raw data generated by titration of 20 µM protein with 10 µl injections of 400 µM indinavir stock in the experiment. (**B**) Bottom panel shows the integrated heats for the above peaks plotted against molar ratio of indinavir to protein molecule. The solid line through the data represents the best non-linear least squares fit to the experimental data using one-site binding model and ORIGIN v5.0 software.

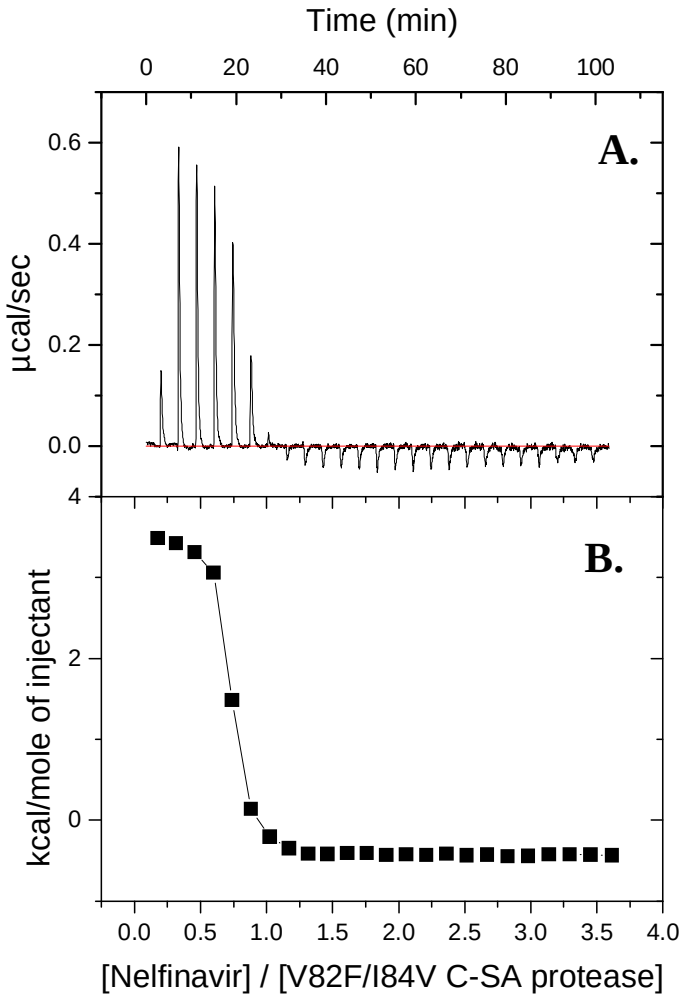


Figure 38. Representative calorimetric profile for the direct titration of the V82F/I84V HIV-1 protease with nelfinavir. The experiments were performed in 10 mM sodium acetate buffer containing 2 % DMSO, pH 5.0, at 25 ºC. (**A**) Top panel shows the raw data generated by titration of 20 µM protein with 10 µl injections of 400 µM nelfinavir stock in the experiment. (**B**) Bottom panel shows the integrated heats for the above peaks plotted against molar ratio of nelfinavir to protein molecule. The solid line through the data represents the best non-linear least squares fit to the experimental data using one-site binding model and ORIGIN v5.0 software.

As shown in Figures 35, 36, 37 and 38, the association of saquinavir, ritonavir, indinavir and nelfinavir with the V82F/I84V HIV-1 protease is characterized by endothermic heats representative of unfavorable change in binding enthalpy (i.e., positive ΔH). Each peak in the top panel represents the binding of an inhibitor to the protease molecule and as the binding reaction progresses, the area under each peak decreases due to increased occupancy of available binding sites on the protein. The bottom panel in each figure shows the binding isotherm obtained by integrating the heats evolved or absorbed after each injection as a function of inhibitor concentration per dimer. Thermodynamic parameters obtained from data fitting are given in Table 6. The clinical inhibitors bind to the V82F/I84V HIV-1 protease with a stoichiometry ranging from 0.9 to 1.3. This mode of inhibitor binding is consistent with crystallographic data and has been observed previously in calorimetric studies (Baldwin *et al.,* 1995; Chen *et al.,* 1995; Todd *et al.,* 2000). At 25 °C the binding of all four clinical inhibitors (saquinavir, ritonavir, indinavir and nelfinavir) is characterized by a positive (unfavorable) enthalpy change (see Table 6). Under the conditions used in calorimetric experiments, the Gibbs free energy of binding is -9.6 kcal/mol for saquinavir, -9.1 kcal/mol for ritonavir, -8.7 kcal/mol for indinavir and -9.5 kcal/mol for nelfinavir. This indicates that the $-T\Delta S$ contributes between -10 kcal/mol and -13.5 kcal/mol to the Gibbs energy to compensate for the unfavorable enthalpy change (positive ΔH) and consequently providing the favorable binding affinity for saquinavir, ritonavir, indinavir and nelfinavir ($K_d \approx 98 - 450$ nM). However, the double mutation V82F/I84V reduces the binding affinity of the clinical inhibitors 117- to 1095-fold, corresponding to a loss in Gibbs energy of binding ranging from 2.8 kcal/mol to 4.1 kcal/mol (Table 6). The K_d values measured for saquinavir, ritonavir, indinavir and nelfinavir agree well with the inhibition constants, IC_{50} and K_i, measured in the enzyme inhibition experiments (see Tables 3 and 4).

3.4.5 Viral fitness

The empirical parameter, vitality, initially introduced by Gulnik *et al.* (1995), was used to compare the selective advantage of the V82A and V82F/I84V mutations in the C-SA subtype in the presence of clinical inhibitors. Vitality provides a measurement of the biochemical or enzymatic fitness of a particular mutant in the presence of a given inhibitor. Table 7 summarizes the vitality values of the V82A and V82F/I84V mutations. The values of the V82A mutation for saquinavir and nelfinavir are about

Table 7: Table showing vitality values of V82A and V82F/I84V HIV-1 proteases

Protease	**Vitality**			
Inhibitor :	**Saquinavir**	**Ritonavir**	**Indinavir**	**Nelfinavir**
wild-type C-SA	1	1	1	1
V82A	0.61	2.6	3.7	0.61
V82F/I84V	20	180	80	60

0.6, implying that this mutation is less likely to be selected in the presence of these inhibitors. There was a marginal, 2.6- and 3.7-fold, increase in vitalities of the V82A variant in the presence of indinavir and ritonavir, consistent with results obtained previously for the B subtype V82A variant (Gulnik *et al.,* 1995). The double mutation V82F/I84V appears to be the most sensitive mutation, with increased vitality values, from 20 to 180, with the four protease inhibitors used. This mutation seems to be particularly selective for indinavir and ritonavir with vitalities of 80 and 180, respectively (Table 7). This result is consistent with a large decrease in sensitivity of the protease for the two drugs as was shown with inhibition and thermodynamics studies in sections 3.4.2 and 3.4.4.

3.5 Crystallographic studies

Figure 39 shows representative tetragonal bipyramidal crystals (30 µm x 25 µm x 25 µm) of the C-SA HIV-1 protease. Crystals appeared in 4 to 10 days. Similar crystals have also been obtained for HIV-1 protease by other groups (Rose *et al.,* 1996; Hong *et al.,* 1996; Prabu-Jeyabalan *et al.,* 2004; Heaslet *et al.,* 2007). In future, these crystals will be harvested and taken for three-dimensional structural determinations of C-SA protease, including the V82A and the V82F/I84V variants. This will give us a deepened understanding of the interactions involved between the C-SA proteins and inhibitors.

Figure 39. Tetragonal bipyramidal crystals (30 µm x 25 µm x 25 µm) ofthe wild-type C-SA HIV-1 protease obtained in 4 – 10 days. The enzyme was prepared at 6 mg/ml in a pH 5 solution containing 10 mM MES, 1 mM EDTA, 1 mM DTT, 3 mM NaN_3. The reservoir solution contained 100 mM imidazole, 250 mM NaCl, 10 mM DTT, 3 mM NaN_3 at pH 7.0. Hanging drops were prepared by mixing the protein and reservoir solutions in 1:1 (v/v) ratio.

CHAPTER 4

4. Discussion

4.1 Spectral properties of the wild-type, V82A and V82F/I84V HIV-1 proteases

4.1.1 Far-UV circular dichroism

Far-UV CD (250 – 190 nm) is sensitive to the secondary structural conformation of proteins. As such, it was used as a probe to compare the secondary structural content of the V82A and V82F/I84V variants to the wild-type C-SA protease. The overall secondary conformation of the wild-type and variant proteins was similar. The three proteins each exhibited minima at 216 nm, which is typical for predominantly ß-sheeted proteins. This is consistent with crystal structure evidence and previous CD studies on HIV-1 protease, which report that the protein is about 4 % alpha-helical, 48 % ß-sheeted and 48 % aperiodic structured (Navia *et al.*, 1989; Dash and Rao, 2001; Foulkes *et al.*, 2006). The results obtained from far-UV CD spectra report on the non-disruptive nature of the V82A and V82F/I84V mutations on the secondary structural content of the C-SA HIV-1 protease, i.e., the mutations did not induce gross conformational changes in the secondary structure.

4.1.2 Fluorescence spectral properties

Intrinsic tryptophan fluorescence is often used as a probe to monitor the conformational changes of proteins, as it is very sensitive to local environments. Depending on the number and location, tryptophan residues could be used as probes for local or global changes. Single or localized tryptophan residue(s) monitor local conformational changes whereas multiple tryptophan residues distributed across three-dimensional structure would monitor global conformational changes. The dimeric HIV-1 protease has 4 intrinsic tryptophan residues, located at positions 6 and 42 in each monomer (Figure 40). The presence of intrinsic tryptophan probes makes the protease a candidate for fluorescence studies. However, these four tryptophan residues are quite exposed to the solvent and are located at positions where they would be insensitive to major unfolding events like dimer dissociation or domain unfolding. As a result, they are expected to be sensitive to local structural changes. It is for this reason that conformational stability studies of the HIV-1 protease have traditionally been performed by differential scanning calorimetry (DSC) or by activity assays under denaturing conditions (Velazquez-Campoy *et al.*, 2000; Velazquez-Campoy *et al.*, 2001a; Muzammil, *et al.*, 2003; Mahalingam, *et al.*, 1999; 2004; Liu *et*

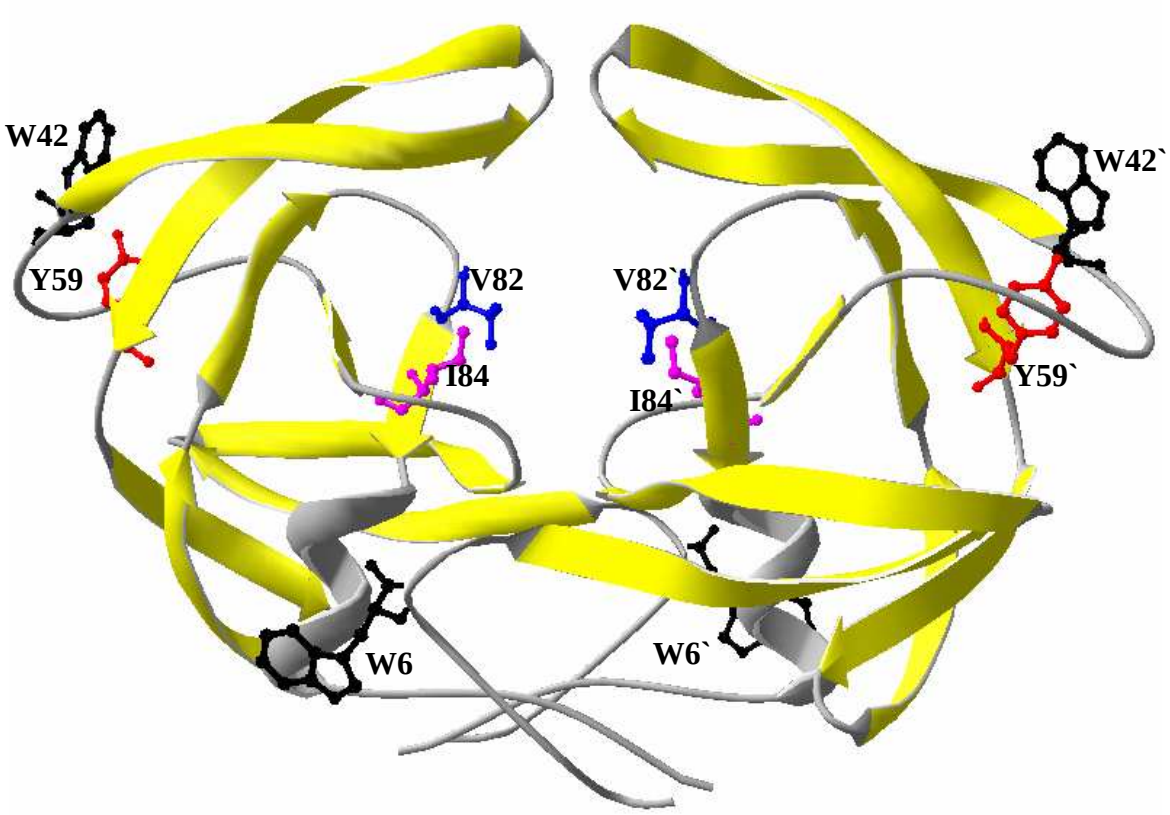


Figure 40. Ribbon representation of the structure of a ligand-free HIV-1 protease. The pdb code for this structure is 1hhp. Tryptophan, tyrosine and amino acid residues at positions 82 and 84 are rendered as balls and sticks. Tryptophan residues are partially solvent exposed. This figure was generated using Swiss-pdb Viewer (Guex and Peitsch, 1997).

al., 2006). Fluorescence studies, with excitation at 295 nm and 280 nm, gave spectra with emission maxima at 355 nm (Figure 18). Additionally, tyrosine emission is often not visible in tryptophan-containing proteins because most of the tyrosine-excited state energy is transferred by fluorescence resonance energy transfer to nearby tryptophan residues. These emission maxima indicate that the tryptophan residues are highly solvent exposed as the emission spectrum of tryptophan in water is between 350 – 355 nm (Lacowicz, 1999). Importantly, the results are consistent with the determined HIV-1 protease crystal structure in which Trp6 and Trp42 are partially solvent-exposed (Figure 40). Due to the fact that Trp6 is located ± 22 Å from the substituted residues at positions 82 and 84 and ± 30 Å from the flap region, the dominant fluorophore appears to be Trp42 which is located at elbow or hinge of the flexible flap region (± 18 Å) (Figure 40). Not only is the fluorescence of Trp42 sensitive to changes occurring at the active site, it is also sensitive to changes in the conformational dynamics of the flap region HIV-1 protease (Ullrich *et al.,* 2000). Iodide quenching fluorescence could also be used to monitor the changes in tryptophan residues occurring at the surface of the protein (Quay *et al.,* 1985). The absence of Stokes'or wavelength shift suggests that the tryptophanyl environments of all proteins are more or less similar. However, there is a major difference in the intensities at the emission maxima for all three proteins. The fluorescence intensity of the V82F/I84V HIV-1 protease is significantly enhanced (40 – 60 %) relative to the wild-type and the V82A HIV-1 proteases, suggestive of local changes in the V82F/I84V variant protease molecule.

Upon inspection of the environment of Trp42 there are several charged residues within 5 Å (Arg41, Lys43 and Arg57). The only residue, found on the structural element preceding the flap region, with its side chain pointing towards the indole ring of Trp42 is Arg57 (see Figure 40). Therefore, it may be possible that in the V82A HIV-1 protease, Arg57 approaches Trp42 even closer and quenches its fluorescence whilst in the V82F/I84V variant HIV-1 protease this residue (Arg57) does not come into contact with the Trp42. The structural element (the hinge region) leading to the flap region in the V82F/I84V HIV-1 protease could assume a different 'semi-open' conformation relative to the wild-type and the V82A HIV-1 proteases and hence an enhanced fluorescence. This is consistent with molecular dynamic simulation studies by Aruksakunwong *et al.* (2006) which showed that the V82F/I84V mutation in the B

subtype results in a 1 – 1.5 Å shift of the inter-residue distance of monomer B, suggesting that the flap structure may have changed because of the mutation. These authors propose that the flap tip of the double mutant has a greater degree of opening than that of the wild-type and may possibly lead to a loss of interactions between the drug and the protein (see sections 4.4 and 4.5).

4.2 Effects of the V82A and V82F/I84V mutations on substrate binding

The V82A and V82F/I84V HIV-1 proteases did not show a large variation in their affinity for the chromogenic substrate. Both variants had the same K_m values comparable to that of the wild-type C-SA protease. The catalytic efficiencies and k_{cat} values for the cleavage of the CA-p2 substrate were lowered 3.6-fold for the V82A mutation and 6-fold for the V82F/I84V mutation relative to the wild-type C-SA protease (Table 2). Although similar K_m values were obtained for all proteins, it is not clear why Ca-p2 peptide is cleaved less efficiently by the V82A and V82F/I84V mutants. However, calorimetric studies by Todd *et al.* (2000) have shown that the free or open conformation of the subtype B V82F/I84V variant is more stable (i.e., it has a higher melting temperature) than the wild-type protease. These authors propose that the V82F/I84V mutation stabilizes the flaps in the open conformation and as a result will lower the binding affinity of substrates or inhibitors due to increased energy required to bring the flaps into a closed conformation. Furthermore, according to studies by Klabe *et al.* (1998), the V82F/I84V mutation resulted in a 1.1 kcal/mol decrease in substrate binding which is close to the value (1.4 kcal/mol) of the difference between the conformational energy of the wild-type protease and that of the variant V82F/I84V protease (Todd *et al.,* 2000). It is possible that the same scenario occurs in the V82F/I84V C-SA protease, where productive binding of the substrate and efficient release of proteolytic products is dampened possibly by a more stable semi-open conformation.

Subtle changes in the specific activity and K_m values observed for all proteins could also be due to the peptide-substrate's intrinsic flexibility. The substrate is more amenable to adapt to changes induced by mutations (e.g., V82A and V82F/I84V) in the binding pocket and is able to maintain strong van der Waals and hydrophobic contacts without losing significant affinity (Tie *et al.,* 2005). It is this property that accounts for the loss in sensitivity to conformationally constrained protease inhibitors

in resistant viral strains encoding for protease molecules containing specific amino acid mutations that lower the affinity for the inhibitors (see section 4.4). Clinical inhibitors with fewer degrees of freedom or with conformational rigidity are less able to accommodate to a geometrically rearranged binding pocket. It is for this reason that inhibitors with low flexibility will have a higher specificity toward a selected target but loose their potency, even with minor changes in the geometry of the binding site. As a result, reduction in inhibitor susceptibility will usually be as a consequence of a combination of decreased inhibitor potency coupled with sustained or enhanced catalytic capability and substrate affinity. Additionally, mutations occurring within the binding site are expectedly conservative and, therefore, introduce geometric distortions without changing its polarity in order to maintain reasonable affinity for the substrate and to achieve a viable catalytic activity.

Thus, although binding affinity of the natural substrates is in the micromolar range, the introduction of mutations conferring resistance in the protease molecule or conformational rearrangements in the binding pocket does not seem to affect the binding of the natural substrate (see Table 2). On the other hand, preshaped and structurally rigid clinical inhibitors which bind in the nanomolar range are less able to accommodate changes incurred in the binding site. The conformational rearrangement of the active site and inability of the inhibitors adapting to the geometrically distorted binding site in the mutated protease lead to reduced affinity (see sections 4.4 and 4.5).

Furthermore, crystal structures of natural substrate peptides complexed to either wild-type or V82A variant subtype B protease have also not demonstrated any significant changes in the contacts formed between the two proteins and the substrate (King *et al.,* 2004a; Tie *et al.,* 2005). Molecular interactions between protease and the longer natural substrates consist mainly of extensive backbone-backbone hydrogen bonds as well as more extensive van der Waals interactions. This suggests that the side chain substitution in V82A could have little effect on substrate binding. This observation can also be used to explain almost similar kinetic parameters obtained between the wild-type HIV-1 protease and the V82A variant (Table 2). King *et al.* (2004a) demonstrated that despite the chemical differences of the drugs, all compounds occupied a similar volume within the active site that is termed the ‘inhibitor envelope’. If the inhibitor envelope is compared with the ‘substrate envelope’ (the

space within the protease that is occupied by a natural substrate), the inhibitors protrude from the substrate envelope in very distinct locations. At these positions, protease inhibitors may have van der Waals interactions with amino acids such as G48, I50, V82 and I84. Mutations at these residues are known to result in protease inhibitor primary resistance and, therefore, these mutations likely disrupt the inhibitor and protease interactions.

Studies on structural distributions of protease interactions with synthetic inhibitors and substrate have shown that the protease interacts with both, the inhibitor and substrate, at similar positions, albeit with different magnitudes towards the overall binding energy (Luque *et al.*, 1998). These authors (Luque *et al.*, 1998) propose that the origin of resistance resides in the thermodynamic distribution of the binding or interaction energetics rather than the type of residues involved in inhibitor or substrate binding. Their studies have also shown that the binding of the chromogenic substrate to the wild-type subtype B HIV-1 protease results in a positive or unfavorable enthalpy of 12 kcal/mol, indicating that the binding of this substrate is also entropically driven ($-T\Delta S$ = -28.3 kcal/mol). This favorable entropy change amounts to Gibbs free energy (ΔG) of -16.3 kcal/mol. The unfavorable binding enthalpy seems to arise from the rearrangement of the flap region in each subunit upon binding. This conformational rearrangement involves primarily residues 40 – 60 on the flap itself and residues 60 – 84 found on the adjacent ß-strand (Luque *et al.*, 1998).

4.3 Effect of C-SA natural polymorphisms on inhibitor efficiency

It is important to note that the effect of the natural polymorphisms in the C-SA subtype on inhibitor binding affinity are not large enough to elicit drug resistance and suggests that current inhibitors can, without doubt, be effective against the wild-type form of this subtype. The affinity of clinical inhibitors used in this study (saquinavir, ritonavir, indinavir and nelfinavir) for the wild-type C-SA protease is in the nanomolar range and comparable to that of the wild-type B subtype. Results presented here are also in agreement with inhibition studies reported earlier (Velazquez-Campoy *et al.*, 2001a; 2002; 2003b) indicating that the affinity of clinical inhibitors is reduced by a factor of 2 – 7.5 in the African subtype C HIV-1 protease (see Table 6 for comparison). However, the subtype C in Velazquez-Campoy *et al.* (2001a) studies differed from the C-SA at 4 positions (S12T, V15I, I19L and L93I). Although by

itself this reduction in the binding affinity is not sufficient to induce drug resistance, it suggests that the efficacy of the existing protease inhibitors against the wild-type South African subtype C protease is not hampered. Thus, by themselves, the amino acid polymorphisms found in the South African subtype C protease are not enough to induce drug resistance. However, it should be noted, that the genetic barrier to resistance might be decreased in comparison with the subtype B virus. Furthermore, the binding affinity for saquinavir, ritonavir, indinavir and nelfinavir is significantly reduced in the presence of V82A and V82F/I84V mutations in the C-SA HIV-1 protease as shown in sections 3.4.2 and 3.4.4 (Table 6).

4.4 Structural implications of the V82A and V82F/I84V mutations on clinical inhibitor binding

Over 200 crystal structures of HIV-1 protease including specific drug-resistant protease inhibitor complexes have been solved in the past several years (Baldwin *et al.*, 1995; Chen *et al.*, 1995; Hong *et al.*, 1996; 1997; 1998; Kervinen *et al.*, 1996; Silva *et al.*, 1996; Ala *et al.*, 1997, 1998; Swairjo *et al.*, 1998; Mahalingam *et al.*, 1999; Vondrasek and Wlodawer, 2002; Tie *et al.*, 2007). Comparisons of these variant proteins with the wild-type protease-inhibitor complexes have been invaluable in the analysis of mechanisms by which these mutations reduce or affect the binding of the drugs especially in the B subtype. Although, to date, no crystal structure of the subtype C HIV-1 protease has been elucidated, crystallographic comparison of the V82A and V82F/I84V subtype B mutants in complex with first-generation protease inhibitors can be used to provide a structural basis for understanding the observed effects.

Crystal structure studies of the V82A mutations in the subtype B have shown that the structure of the wild-type and V82A mutant in complex with saquinavir are very similar with calculated rms deviations of 0.43 Å for the main-chain atoms (Tie *et al.*, 2007). This value is within the range observed for identical proteins (Bett and Sternberg, 1999). Saquinavir, as described in section 1.8, is a peptidomimetic inhibitor containing the main-chain amides, the carbonyl oxygen atoms of a peptide and a variety of groups corresponding to the side chains at positions P3-P2` of a peptide substrate (see Figure 7). The large hydrophobic groups at positions P3, P1 and P1` of saquinavir are accommodated in the hydrophobic pockets of S3, S1 and S1` of

the HIV-1 protease. There is a smaller hydrophobic tertiary-butyl or t-butyl group at P2` and the polar asparagine side chain at P2 of saquinavir, both of which occupy the S2` and S2 binding sites, respectively. The P1 and P1` groups of saquinavir form van der Waals contacts with the protein residues: Leu23`, Asp25/25`, Ala28, Gly49`, Ile50, Thr80, Pro81, Val82` and Ile84/84`. The other residues that interact with saquinavir include Gly27, Ala28/28`, Asp29, Asp30/30`, Gly48/48`, Gly49/49`, Ile50, Pro81` and Ile84` as well as residues forming the 80's loop (79-82, 79`-82`). The wild-type and mutant V82A HIV-1 proteases have almost identical hydrogen bond interactions with saquinavir, except that the hydroxyl group of saquinavir is positioned more asymmetrically with respect to the side chains of the two catalytic aspartic acid residues, 25 and 25`, in the mutant protein. The role of hydrogen bond is recognized as being of fundamental importance in determining the structures of proteins as well as their complexes with ligands (Connelly *et al.,* 1994). Figure 41 shows the hydrogen bond network established between the protease and saquinavir. Seven direct hydrogen bonds are formed between the protease and saquinavir. Four hydrogen bonds are formed between the hydroxyl group of saquinavir and the carboxylate oxygen atoms of the catalytic Asp25/25` side chains, and the other three are from O1 of saquinavir to the main-chain of Asp29 from N3 (of saquinavir) to amide nitrogen of Asp30 and from N3 to the main-chain carbonyl oxygen of Gly48. In this structure, two water molecules mediate three more hydrogen bond interactions between protease and the inhibitor. The conserved water molecule links the PR flaps and saquinavir through interactions with Ile50/50`, and the other water molecule mediates interactions between saquinavir and the main-chain oxygen of Gly27 and side chain of Asp29 (Figure 41).

Of note, the V82A mutation results in small shifts (0.5 Å towards saquinavir) in the position of the main-chain atoms at the site of mutation, in both subunits (Figure 42). In the wild-type protease, the Val82 side chain seems to be opposite the P1` ring and further from saquinavir (by about 4.5 Å) than the side chain of Ala82 in the mutant protease. Instead, the $C_{ß}$ atom of Ala82 seems to move 0.2 Å nearer to the P1` group, although the inter-atomic distances are still over 4 Å. This subtle conformational change shifts Pro81 to establish van der Waals contacts of 3.7 – 4.0 Å with the carbon atoms of the decahydroisoquinoline group of saquinavir, which are absent in the wild-type protease (Figure 41).

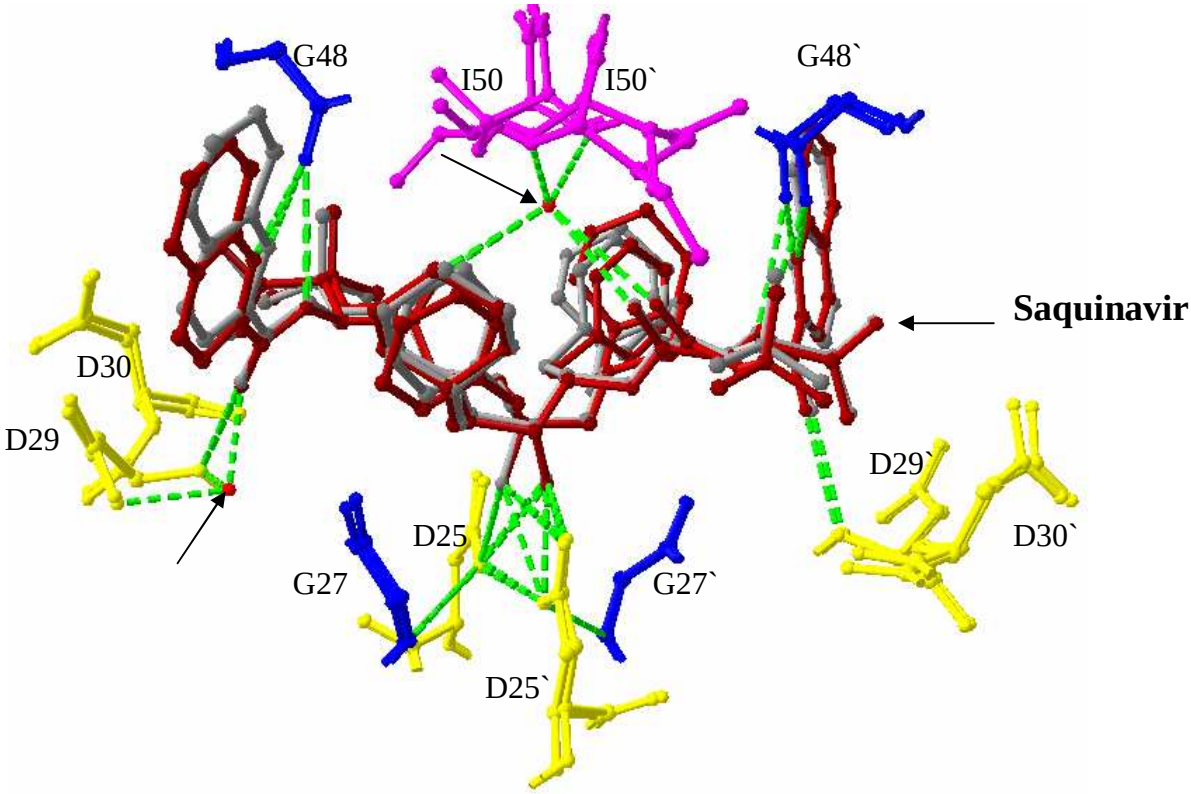


Figure 41. Hydrogen bond interactions between the wild-type, V82A HIV-1 subtype B protease and saquinavir shown in brown and grey. Hydrogen bond interactions are indicated by dashed lines. Interactions mediated by water molecules are indicated by arrows. This figure was generated using Swiss-pdb Viewer (Guex and Peitsch, 1997).

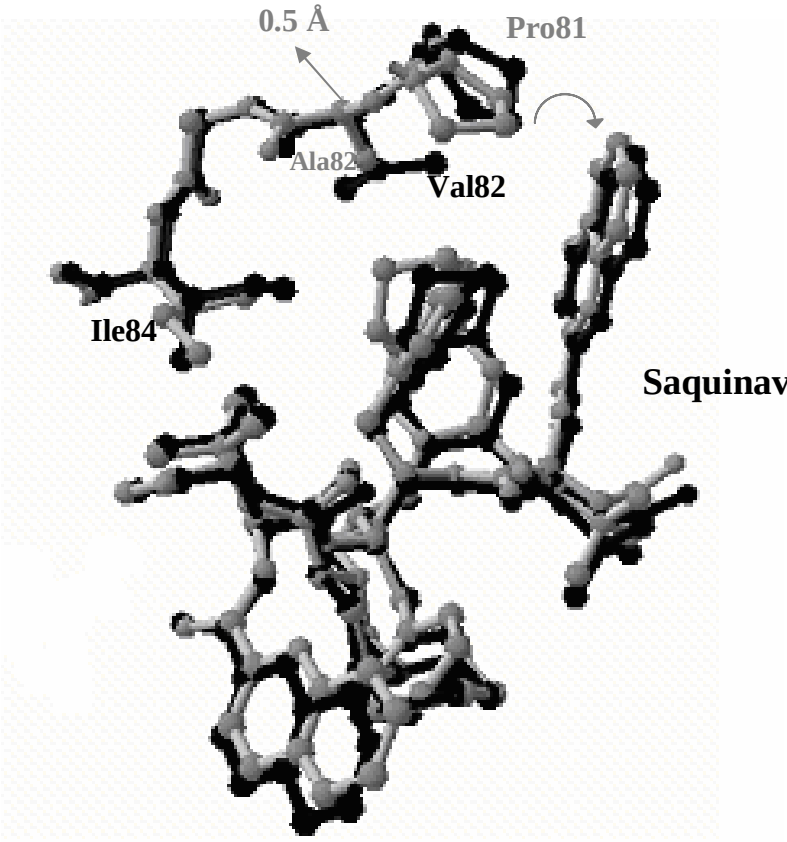


Figure 42. Structural superimposition of residues 80-84 in the wild-type and V82A subtype B HIV-1 protease in complex with saquinavir. The wild-type HIV-1 PR is colored in black and the V82A mutant is coloured in grey. The V82A mutation results in small shifts (0.5 Å towards saquinavir) in the position of the main-chain atoms at the site of mutation, in both subunits. There is also a subtle conformational change which shifts Pro81 to establish van der Waals contacts of 3.7 – 4.0 Å with the carbon atoms of the decahydroisoquinoline group of saquinavir which are absent in the wild-type PR. This figure was generated using Swiss-pdb Viewer (Guex and Peitsch, 1997).

The movement of Ala82 in the other subunit allows the P1 benzyl ring of the inhibitor to move 0.1 – 0.2 Å closer to the S1 subsite within the protein's binding pocket. Similar conformational rearrangements have been reported for the V82A mutants in complex with substrate analogues (Tie *et al.*, 2005). These studies also suggest that a mutation at position 82 to a smaller amino acid or cavity forming amino acid provides more space for the mutant HIV-1 protease to accommodate the bulkier hydrophobic groups of saquinavir. Furthermore, previous studies on the V82F mutation in the subtype B HIV-1 protease have shown that the protease accommodates the bulky side chain of Phe82 within the binding pocket without forming steric interactions with the inhibitor (Ala *et al.,* 1997). Small reductions in inhibition constants and affinity constants obtained for saquinavir and nelfinavir against the V82A HIV-1 variant protease are, therefore, suggestive of minor structural adaptation of the V82A variant to accommodate the inhibitor (saquinavir or nelfinavir) as seen in the B subtype. Our studies, which are in agreement with the previously published data, show that the presence of a single V82A mutation affect the binding of indinavir and ritonavir (Chen *et al.,* 1994; Baldwin *et al.,* 1995a). Other studies have also implicated the V82A mutation as the most common resistance mutation to arise in patients receiving indinavir or ritonavir, consistent with reduced inhibitory effects and reduced affinity (Shafer *et al.,* 1999). The effect of this mutation in the B subtype, as analyzed by X-ray crystallography, revealed a widespread rearrangement of the backbone around the binding pocket, and not merely a removal of a single methyl group in the protease molecule (Baldwin, *et al.,* 1995a). Close examination of the wild-type subtype B HIV-1 protease in complex with ritonavir and indinavir, revealed that Val82 forms van der Waals contacts with six or more atoms of the inhibitor (Chen *et al.,* 1994; Kaldor *et al.,* 1997). Indinavir's pyridyl and piperidine groups occupy the S3 and S1 subsites, whereas the benzyl moiety occupies the S1` subsite. The t-butyl group and the indanol moiety occupy the S2 and S2` subsites, respectively (Figure 7). The V82A mutation results in a loss of these van der Waals interactions between the protease and the inhibitor at the P1` site, and hence a reduction in the inhibition of the mutant protease (Baldwin, *et al.,* 1995a; Prabu-Jeyalbalan *et al.,* 2003; 2004; Clemente *et al.,* 2004; Mahalingam *et al.,* 2004). These studies also revealed a loss of protein-inhibitor hydrogen bonds in the V82A-inhibitor complexes due to these subtle structural rearrangements.

Although some authors have previously implicated these movements of the main-chain atoms of residues 81-82 (Figure 43) as compensatory readjustments for the loss of a methyl group in the V82A mutation, residue 82 still remains a major contributor to the binding pocket by establishing contacts with the substrate or inhibitor group at the P1/P1` sites (Tie *et al.*, 2004). In the wild-type HIV-1 protease, Val82 is part of the hydrophobic S1/S1` subsite, which also contains Leu23, Gly27, Ile50, Pro81 and Ile84 residues. Thus, Val82/82` stabilizes the binding of hydrophobic residue at P1/P1` through favourable interactions. The V82A variant substitutes the Val side chain with a smaller cavity-forming alanine. Hence, favorable hydrophobic contacts with the inhibitor are lost in the variant protein. This loss in inhibitor interactions with residue 82 could also lead to a more flexible 80's loop and, in turn, affect the binding of the flap residues to the inhibitors. For example, Ile50, which contributes significantly to the hydrophobic interactions at the flap region/dimer interface and lead to poor inhibition and/or poor stability. Other residues which have been shown to be to be impacted upon by mutation at residue 82 include: Trp6, pro9, Gly17 and Pro39, which are distant from this site. Recent studies by Liu *et al.* (2006) have corroborated these effects by showing that mutations that produce conformationally more flexible flaps, especially at the tips, can indirectly transmit changes in the 80's loop which open up the active site cavity even further and reduce inhibitor binding affinity.

In addition, recent studies by Heaslet *et al.* (2006) have shown that the introduction of alanine at position 82 weakened the packing contact with the P1` phenyl ring of TL-3 inhibitor (a $C_2$ symmetric broad-based HIV protease inhibitor) resulting in an increase in $IC_{50}$ value for TL-3 by a factor of 4 (6 nM to 22 nM). TL-3 has (1*S*,2*R*,3*R*,4*S*)-1,4-diamino-1,4-dibenzyl-2,3-butanediol core (P1/P1` residues); valine residues at P2/P2`; alanine residues at P3/P3` and a carboxy-benzyl moiety at the P4/P4`. In this V82A-TL-3 complex, the side-chain of Pro81 appears to be isomerizing so that the pyridyl ring is flipped or displaced away from the P1` phenyl ring to avoid unfavorable interactions. Interestingly, other packing contacts between the TL-3 inhibitor and the protein, remained essentially unchanged comparable to those observed in the wild-type structure.

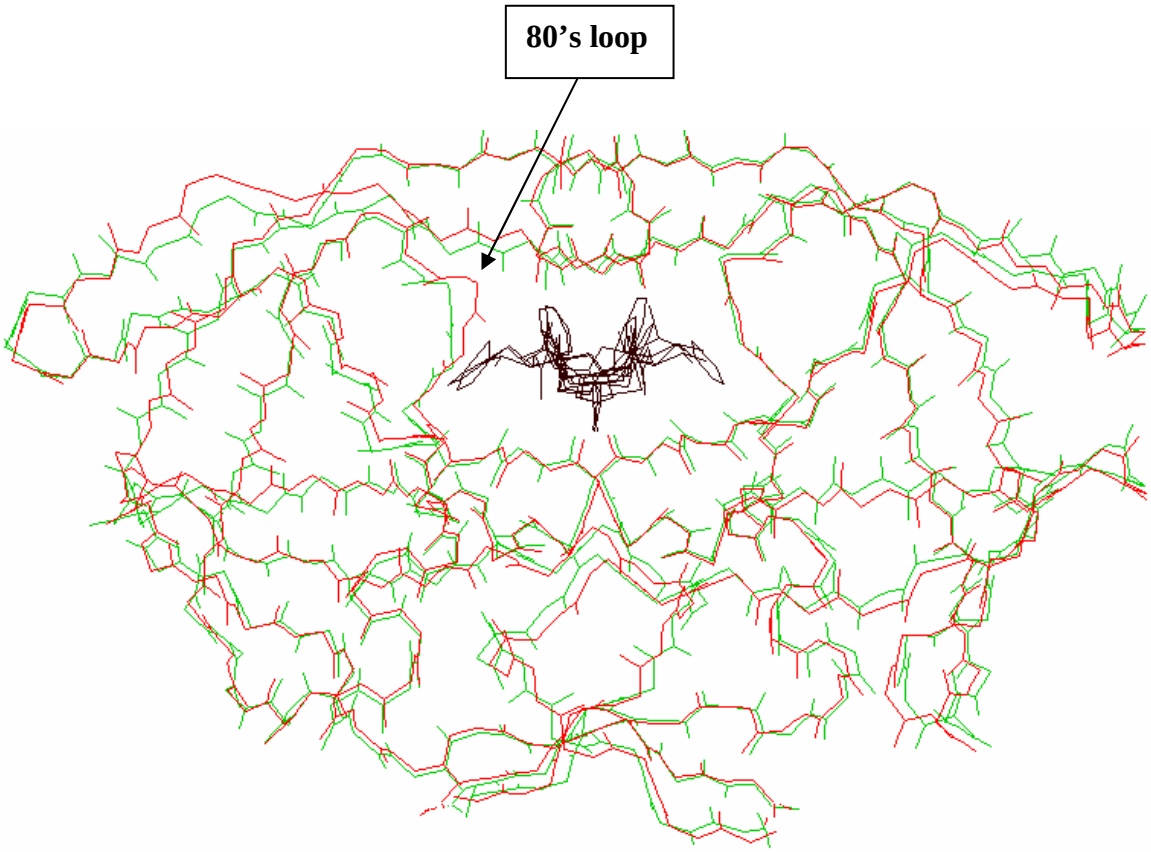


Figure 43. Backbone alpha-carbon traces for the structural superposition of wild-type (1HXW; green) and V82A mutant (1RL8; red) subtype B HIV-1 protease in complex with ritonavir (black) calculated using Swiss pdb Viewer (Guex and Peitsch, 1997). Although both structures superimpose very well with each other, subtle rearrangement of the main-chain atoms of residues 81-82 around the binding pocket is evident. The rmsd values for wild-type-V82A HIV-1 protease backbone superpositions is 0.78 Å (calculated using Swiss-pdb Viewer).

To date, the subtype B V82F/I84V variant in complex with cyclic urea-based inhibitors (DMP323 and DMP450) and with second-generation inhibitors (KNI-577 and KNI-764) are the only three-dimensional structures of this mutant form to be determined by X-ray crystallography (Ala *et al.*, 1997; 1998; Vega *et al.*, 2004). To our knowledge, the crystal structure of the V82F/I84V HIV-1 protease in complex with the first-generation inhibitors which includes saquinavir, ritonavir, indinavir, nelfinavir, amprenavir and lopinavir, is not yet known. To understand the structural basis of reduced sensitivity of the V82F and I84V double substitutions in the subtype B HIV-1 protease, a model of ritonavir bound to the HIV-1 protease was constructed on the basis of the crystal structure of the related inhibitor A-78791 (Hosur *et al.*, 1994). At the S1` subsite, both Val82` and Ile84` make favorable and direct contacts with the benzyl moiety of the inhibitor at the P1` subsite, whereas the Phe82` and Val84` have different effects because they distort the binding site geometry without changing its polarity and chemical composition. The V82F` results in a spatial overlap with the phenyl ring of ritonavir-related inhibitor at the P1` site, while the I84V` decreases van der Waals contacts with the Cß group of the benzyl side chain of the inhibitor. Avoidance of the spatial group overlap requires a considerable conformational change in either the inhibitor or the enzyme that will likely result in further decreased number of interactions with the P1 benzyl moiety of the inhibitor. First-generation inhibitors are conformationally constrained and, therefore, can not adapt to changes in the target thus causing a reduced affinity. Ile84 also makes an additional contact with the Cγ methyl group of the valine side chain in P2 of the inhibitor through its Cγ2 methyl. The latter interaction may also be lost in the V82F/I84V double mutant as a result of conformational rearrangements due to the V82F mutation.

Studies by Ala *et al.* (1997; 1998), have also shown that the wild-type and I84V, V82F and V82F/I84V mutant HIV-1 proteases in complexes with cyclic urea inhibitors (DMP323 and DMP450) are very similar with rms deviations of 0.2 – 0.3 Å. The marked reduction in binding affinity for DMP323 and DMP450 is not due to a mutation-induced change in the hydrogen-bonding patterns between the protease and the inhibitors at the active site, as these are not altered in the crystal structure of the V82F/I84V variant in complex with either DMP323 or DMP450. Rather, the only observable differences were in the patterns of van der Waals contacts, which appeared

to modulate the K_i values of the mutant proteins. The distribution of contacts was observed to be over three regions of the protein: (1) residues 23-28 at the base of the substrate binding site, (2) residues 47-50 at the flap region, and (3) residues 29-32 and 81-84 at the site of the binding pocket. Notably, Val82 and Ile84, in the wild-type protein make up ~14 % of the total number of contacts with the inhibitors and appear to modulate the binding affinities of the mutants, i.e., substitutions at these respective positions. As an example, about 70 % of the contacts formed between residues in positions 82 and 84 and DMP323 is lost in the V82F/I84V variant (Ala *et al.*, 1998). These studies also suggest that the Val82 in the V82F/I84V variant has created a cavity in the S1 subsite sufficient to accommodate the bulky benzyl group of Phe82. In this case, the P1 of the inhibitor would have to sterically force the Phe82 out of the subsite to be accommodated in the binding pocket, consistent with the spatial displacement of the phenyl ring of ritonavir-related inhibitor at the P1` site as suggested by Hosur *et al.* (1994) and more recently by Aruksakunwong *et al.* (2006). The increment of K_i values for all inhibitors used in the V82F/I84V mutation in the C-SA protease could, therefore, be attributed to the inability of the Phe82 to form favorable contacts with the inhibitor. It was also proposed from these studies (Ala *et al.*, 1998) that the simultaneous substitution of Val82 and Ile84 with Phe and Val, respectively, causes subtle shifts of Ile47 and Ile50 which appear to form an intricate network that can communicate changes between the S1 and S2 subsites. Recent molecular dynamics studies on the wild-type and the V82F/I84V subtype B HIV-1 protease complexed with ritonavir have also shown similar global tertiary structures (Aruksakunwong *et al.*, 2006). The major observable difference was on the flaps of chain B (49`-54`) in the V82F/I84V mutant. These authors propose that the flap tip of the double mutant has a greater degree of opening than that of the wild-type and that this structural opening may possibly lead to a loss of interactions between the drug and the protease. It also appears that the Ile50` side chain is in close proximity with P1` and P2 subsites of ritonavir and residues 81-84. As such, the displacement of the flap tip in the V82F/I84V variant results in a loss of favorable contacts between the protein and the drug and hence a decrease in the binding affinity. The most significant difference between the wild-type and the V82F/I84V variant was found at the P1` subsite of ritonavir, resulting in the positional or conformational displacement of the phenyl ring of Phe82 out of the S1` subsite. Similar to observations by Ala *et al.* (1997; 1998) on the V82F/I84V variant in complex with DMP323 inhibitor, the Val84

seems to bury itself deeper into the larger cavity of the S1` binding pocket whilst the phenyl ring of Phe82 moves away from the S1` subsite allowing the P1` phenyl ring of ritonavir to rotate or reorient and establish contacts with Val84. This apparent volume change in the S1` subsite may decrease direct contacts between ritonavir and predominant residues in the binding pocket and hence a decrease in inhibitor potency. Additionally, van der Waals contacts between the protease and the inhibitor contribute ~70 % to the overall binding free energy.

Previous studies on the solved protease-indinavir complexes; wild-type and quadruple mutant (M46I, L63P, V82T, I84V) by Chen *et al.* (1994; 1995) revealed that the conformations of residue 82 in the S1` loop were the same as those in the triple mutant (L63P, V82T, I84V). The CG1 atom of residue 82 makes van der Waals contacts with the benzyl ring of indinavir. In these two mutant complexes with indinavir, the side chain oxygen of Thr82 faces away from the active site, thus reducing its inhibitory effect, while the position occupied by the side chain of Thr82` in the other subunit of the dimer is similar to that in the wild-type structure. The I84V mutation, on the other hand, physically changes the volume of the S1 and S1` subsites within the active site. It is evident that the introduction of a bulkier Phe side chain at position 82 and smaller cavity-forming Val at position 84 most likely distorts the geometry of the binding site and results in important losses of van der Waals contacts and other favorable interactions with the inhibitors. Significant reductions in the inhibitory effects observed for saquinavir, ritonavir, indinavir and nelfinavir against the V82F/I84V C-SA HIV-1 protease could, therefore, in part be attributed to the altered binding pocket's geometry.

Analysis of the structure of HIV-1 protease shows that, of the eight mutated residues in the C-SA HIV-1 protease, six appear to be largely buried from the solvent (Thr12, Ile15, Leu19, His69, Leu89 and Ile93). With the exception for Met36 and Arg41, which are located in the hinge region of the flaps, the conservatively substituted amino acids in the C-SA protease are located at the core of the protein (Figure 44). It is possible that these substitutions induce atomic packing rearrangements of the amino acids at the protease core immediately beneath the active site. This effect, coupled with a slight change at the hinge region, due to the M36I and R41K mutations can affect the energetics of conformational change associated with binding.

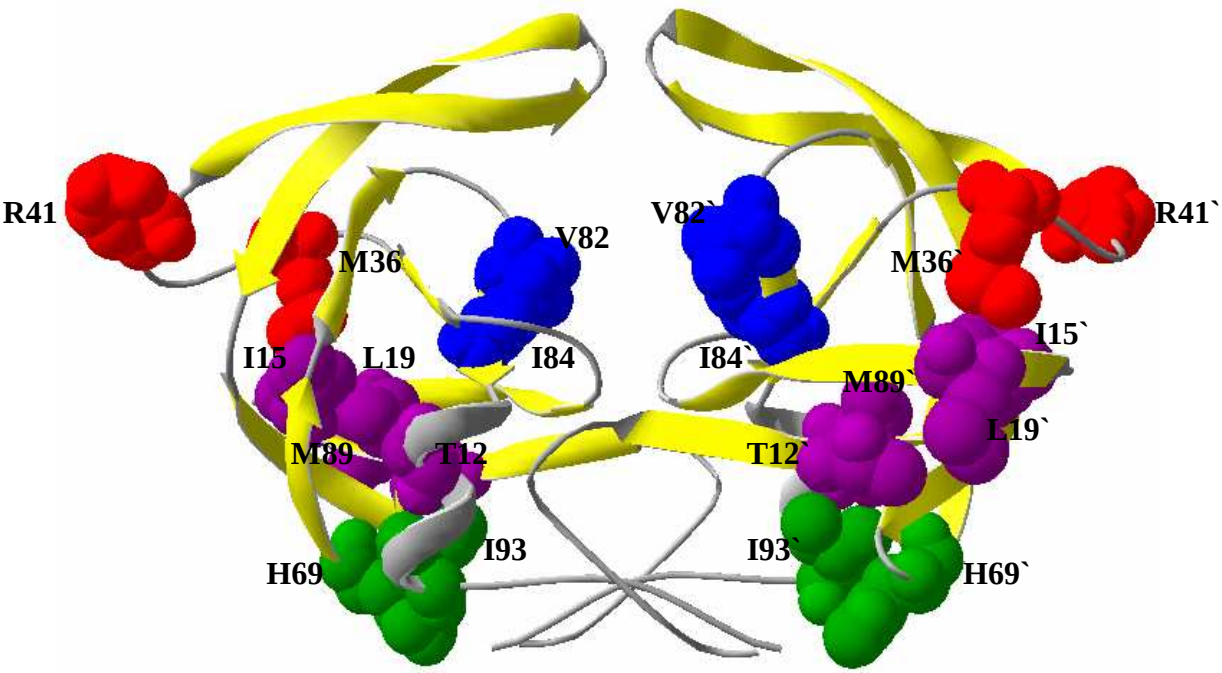


Figure 44. A ribbon representation of the homodimeric structure of HIV-1 protease indicating the topographical positions occupied by the 8 consensus amino acid residues in the South African C subtype: T12S, I15V, L19I and L89M (purples spheres), M36I and R41K (red spheres) and H69K and I93L (green spheres). Spheres occupied by mutations at positions 82 and 84 are shown in blue. The pdb code for this structure is 1HHP. This figure was generated using Swiss-pdb Viewer (Guex and Peitsch, 1997).

In addition, the largest observable difference between the wild-type subtype B and subtype F HIV-1 protease was in the loop region (between residues 33 and 42), with a deviation of 2.85 Å between the Cα positions of residues at position 35 (Sanches *et al.*, 2007). The loop region containing residue 36 shifts towards the protein core and the end of the side chain in position 36 tends to occupy equivalent positions in both the wild-type B subtype and the wild-type F subtype in order to maintain contacts with the neighboring interacting residues (Leu33, Leu38 and Lys20). The subtype F HIV-1 protease used in these studies differed in twelve positions (I15V, E35D, M36I, S37A, R41K, R57K, D60E, Q61N, I62V, L63S, I64L, L89M) in relation to the consensus subtype B protease (Sanches *et al.*, 2007). It is proposed that the M36I mutation which is a natural polymorphism for the subtype F HIV-1 protease results in the stiffening of the flap hinges for the protein. Similar structural effects have also been observed in the structure of the subtype B multi-drug resistant variant containing the M36I mutation complexed with either ritonavir or indinavir (Clemente *et al.*, 2004). Recent studies by Ode *et al.* (2007) have also shown that the M36I decreases the volume of the active site of subtype B HIV-1 protease by approximately 8 %.

The L89M mutation on the other hand is proposed to be the structural equivalent of the well-characterized L90M mutation, which makes additional van der Waals contacts with the main chain atoms of the active site residues 24 through 26. This increased number of interactions subsequently leads to a decrease of the volume of the substrate-binding pocket and reduces activity and the structural flexibility of the S1/S1` pockets, relative to the wild-type protease. This effect has been implicated in rendering resistance to saquinavir and nelfinavir (Hong *et al.*, 2000). The longer methionine side-chain at position 89 in the subtype F seems to displace the neighboring Leu90 toward the main-chain atoms of the catalytic loop, thereby mimicking the effects of the L90M mutation in the subtype B. Therefore, these aforementioned effects, coupled with substitutions at positions 82 and 84 could also induce geometric distortions in the active site cavity and consequently lead to a substantial loss in binding affinity of inhibitors with limited adaptability.

4.5 Thermodynamics of clinical inhibitor binding

HIV protease inhibitors are competitive inhibitors, and as such their potency is directly related to the affinity with which they bind to the protease molecule. From a thermodynamic point of view, the level of binding affinity, K_a, is determined by the Gibbs free energy of binding, ΔG, which is an important descriptor of a binding reaction and in turn a function of the enthalpy, ΔH, and entropy, ΔS, changes ($K_a = \exp(-\Delta G / RT)$, where $\Delta G = \Delta H - T\Delta S$). It is clear that different combinations of ΔH and ΔS values can elicit same ΔG and therefore same K_a. The binding affinity of a ligand is, to a significant extent, proportional to the solvent-accessible surface area that becomes buried from the solvent upon binding because the energetic contributions are not simply localized to the direct interactions between molecules, but contain contributions from structural and dynamic changes transmitted throughout the protein and from counter ions and hydrating water molecules located at the binding site (Ladbury and Williams, 2004).

The free energy of HIV-1 protease-ligand complex formation is determined by the ΔH_{int} of the interactions formed at the binding interface, the ΔH arising from the rearrangement of the flap region in each subunit upon binding, the ΔH_{solv} associated with solvation and desolvation events, the ΔS_{conf} resulting from changes in conformational degrees of freedom of the HIV-1 protease inhibitor, the flap region as well as the protein backbone and side-chains involved in binding, the ΔS_{solv} of ordered water molecules that are released from the surface of the binding molecules into bulk solvent and water molecules that become ordered during the binding and localization events due to burial of hydrophobic groups (Murphy, 1999). This conformational change involves the closing of the flaps and a slight rotation of each monomer around the dimer interface (Miller *et al.*, 1998b; Rose *et al.*, 1998).

Two main forces thus, contribute to the entropy change: (1) an unfavorable term which originates from the loss in conformational degrees of freedom from the ligand and from residues in the protein (ΔS_{conf}) and (2) a favorable term which originates from the release of water molecules upon complex formation (ΔS_{solv}). Other entropic effects, like the loss of translational degrees of freedom, are common to all binding processes. It is thus, evident that the most favorable binding entropy is obtained if the conformational entropy losses are minimized and the solvation gain is maximized.

The conformational entropy losses are minimized if the inhibitor is engineered to be relatively rigid in solution and preshaped to the geometry of the binding site (Ruben *et al.,* 2006). The solvation entropy, on the other hand, is maximized if the inhibitor is made extremely hydrophobic.

The enthalpic term (ΔH) primarily reflects the energetic contribution of many individual interactions (hydrogen bonds, van der Waals and ionic interactions) between the ligand and the protein during complex formation, considering the interactions with the solvent as a reference and, therefore, it also reflects the solvation enthalpy (Ross and Subramanian, 1981; Velazquez-Campoy and Freire, 2005). The binding enthalpy is negative (favorable) if the interactions between the binding molecules over-compensate the interactions of the individual molecules with the solvent; otherwise, it will be positive (unfavorable; as for non-specific hydrophobic interactions). The first-generation PR inhibitors in clinical use have been entropically optimized and their high binding affinity is due to the combination of (1) an extremely large positive entropy of solvation (high hydrophobicity) and (2) a minimal loss of conformational entropy (due to preshaped design of inhibitors or conformationally constrained molecules) that more than overcome their unfavorable or slightly favorable (e.g., ritonavir) binding enthalpy.

The first-generation HIV-1 protease inhibitors: saquinavir, ritonavir, indinavir and nelfinavir bind wild-type HIV-1 C-SA, V82A and V82F/I84V proteases with a stoichiometry of one inhibitor molecule bound per dimer, consistent with the crystallograpically determined stoichiometry for the subtype B HIV-1 protease (Baldwin *et al.,* 1995b; Chen *et al.,* 1995; Hong *et al.,* 1996, 1997, 1998; Kervinen *et al.,* 1996; Silva *et al.,* 1996; Ala *et al.,* 1997, 1998; Swairjo *et al.,* 1998; Mahalingam *et al.,* 1999; Tie *et al.,* 2007). Clinical inhibitors bind to the wild-type C-SA HIV-1 protease with high binding affinities (ranging from 0.2×10^9 M^{-1} and 3.0×10^9 M^{-1}). The binding process is favored by entropic contributions ($T\Delta S$) and thus contributing favorably to the Gibbs free energy of binding. The V82A mutation does not seem to affect the binding of saquinavir and nelfinavir significantly (Table 6). The affinity of V82A HIV-1 protease for saquinavir, ritonavir, indinavir and nelfinavir (K_d = 1.85 nM, 2.00 nM, 12.70 nM and 0.66 nM, respectively, at 25 °C) is in the range of 2- to 13-fold of magnitude weaker than that of the wild-type C-SA protein (K_d = 0.84 nM,

0.21 nM, 0.95 nM and 0.30 nM, respectively, at 25 °C). The clinical inhibitors exhibit the highest binding affinity to both the wild-type and the V82A enzymes, but are extremely sensitive to the V82F/I84V mutation (Table 6). The V82F/I84V mutant reduced the binding of saquinavir, ritonavir, indinavir and nelfinavir 117-, 1095-, 474- and 367-fold, respectively. A drop in K_d values obtained for the V82F/I84V in association with saquinavir, ritonavir, indinavir and nelfinavir is consistent with a decrease of between 2.8 - 4.2 kcal/mol in ΔG, which is equivalent to at least 2 to 3 orders of magnitude in binding affinity. The higher affinity of the inhibitor towards the wild-type HIV-1 protease is evident in the more negative free Gibbs energy of binding (ΔG), thus confirming that the first-generation HIV-1 protease inhibitors have been optimized for their entropic contribution; i.e., the inhibitors are more hydrophobic and structurally rigid.

The hydrophobic nature of the ligand is the dominant driving force in the binding process because the inhibitors have the optimal shape and geometry for the target's binding site. It, therefore, appears that the dominant driving force for binding is a large favorable entropic change that originates from large positive solvation entropy (a consequence of the burial of a large hydrophobic surface area upon binding, and a loss of conformational entropy due to the little or reduced flexibility of these conformationally constrained inhibitors preshaped to the geometry of the putative binding site) (Luque *et al.*, 1998; Todd *et al.*, 2000). It has previously been shown that all known first-generation inhibitors of the HIV-1 protease for which data are available including; saquinavir, ritonavir, indinavir and nelfinavir used in this study, bind to the active site with an unfavorable or only slightly favorable enthalpy change, in a process driven by a large positive (favorable) entropy change (Luque *et al.*, 1998; Velazquez-Campoy *et al.*, 2001b; Ohtaka *et al.*, 2003; Muzammil *et al.*, 2003). These HIV-1 protease inhibitors span the entire binding site and interact almost equally with the two protease monomers, inducing a larger conformational rearrangement of the flap region.

The interactions of saquinavir, indinavir and nelfinavir with the wild-type C-SA and V82A HIV-1 protease are, therefore, enthalpically unfavorable and only slightly favorable for ritonavir (Table 6). On the other hand, the interaction of all four inhibitors with the V82F/I84V variant is enthalpically unfavorable. This is consistent

with previous studies indicating that first-generation HIV-1 protease inhibitors, in particular, overcome unfavorable binding enthalpy by a large positive (favorable) entropy change (Todd *et al.,* 2000; Velazquez-Campoy *et al.,* 2001b; Velaquez-Campoy *et al.,* 2002; Muzammil *et al.,* 2003; Ohtaka *et al.,* 2003). The individual contributions of the entropic term to the Gibbs free energy of binding of all four inhibitors suggest that in all four cases, the binding process is entropically driven (Table 6). The majority of the binding energy for saquinavir, ritonavir, indinavir and nelfinavir, therefore, could originate from the burial of a large hydrophobic surface area upon binding. Since solvent contributions towards entropy are primarily a function of changes in solvent-exposed non-polar surfaces and the ordered water molecules solvating these surfaces, it is possible that inhibitor-binding site of the V82A and V82F/I84V variants results in a potential increase in the exposure of previously buried groups. This is consistent with the presence of additional water molecules mediating hydrogen bond interactions between the V82A protease from the B subtype and saquinavir (Tie *et al.,* 2005; Tie *et al.,* 2007). When different contributions to the Gibbs energy of binding are considered for the V82A HIV-1 C-SA protease, it becomes clear that saquinavir, indinavir and nelfinavir lose entropic contributions when compared to the wild-type C-SA protease, while there is a slight gain in entropy with the binding of ritonavir (Table 6). The V82A mutant compensates for the loss in binding entropy by a small gain in binding enthalpy for saquinavir and nelfinavir. Indinavir and more so ritonavir also lose the enthalpic contributions to the binding energy, resulting in a higher loss in binding affinity. For indinavir, the effect of the V82A mutation is almost exclusively entropic. The reduction in the binding affinity is almost entirely reflected in a loss of entropic contributions (Table 6). The inhibitors that are not significantly affected by the V82A mutation either lose very little or no binding enthalpy energy (nelfinavir) or are able to compensate enthalpically for the entropic losses with enthalpic gain (saquinavir) consistent with subtle structural rearrangements observed in the subtype B V82A protease which establish van der Waals contacts with saquinavir (Tie *et al.,* 2007). These two inhibitors, saquinavir and nelfinavir, appear to be less affected by the V82A mutation, suggesting that some functional characteristics of the compounds render them less susceptible to the binding site distortions or changes associated with the V82A mutation. Against the V82F/I84V variant it is apparent that saquinavir, indinavir and nelfinavir lose significant entropic contributions relative to both the

wild-type and V82A C-SA HIV-1 proteases, while ritonavir slightly gains entropic contributions. This loss in binding entropy is also partially compensated for by a slight gain in binding enthalpy for saquinavir and nelfinavir; while indinavir and ritonavir, on the other hand, lose enthalpic contributions to the binding energy and consequently suffer higher affinity loses (see Table 6).

Comparison of the binding energetics obtained for the wild-type C-SA, V82A and V82F/I84V enzymes, indicate that the major difference is in the magnitude of the entropy change (see Table 6). The dominant binding force, which is, a large positive entropy, decreases drastically in the V82F/I84V variant. This higher entropy loss accounts for the observed lower binding affinity of saquinavir, ritonavir, indinavir and nelfinavir towards the V82F/I84V variant. It may thus be possible that the V82F/I84V mutation results in a different free conformation, particularly around the flap region, relative to the wild-type C-SA and V82A mutation proteases. This would impose a larger energetic penalty associated with the obligatory conformational change coupled to inhibitor binding and thus decreased binding affinity. In addition to the loss of favorable contacts with the inhibitors, active site mutations may indirectly affect the geometry of the binding cavity by energetically favoring a slightly distorted conformation. Notably, studies by Todd *et al.* (2000) have conclusively shown that the V82F/I84V mutation in the HIV-1 subtype B protease lowers the binding affinity of saquinavir, ritonavir, indinavir and nelfinavir by making the binding enthalpy more positive and making the entropy change less favorable. Conformational changes due to these substitutions resulted in fewer enzyme-inhibitor interactions which were associated with loss of drug potency. Studies by Perryman *et al.* (2004) and more recently by Aruksakunwong *et al.* (2006) have conclusively shown, albeit by molecular dynamics simulations, that the V82F/I84V variant in the B subtype induced rapid and frequent flap tip curling, which suggested a more open conformation assumed by this mutant. This is consistent with low binding affinities for most inhibitors currently in clinical use, primarily as a consequence of a larger energetic penalty for the burial of a larger exposed hydrophobic surface area. It was also apparent from these studies that the flap region is more dynamic with an increased flexibility compared to the wild-type protease. Moreover, recent crystallographic studies by Liu *et al.* (2006) have shown that mutation producing conformationally more flexible flaps, especially at the tips, can indirectly transmit changes in the 80's

loop structure which opens up the active site cavity even further. Simultaneous substitution of the two active site residues in the C-SA protease, V82F/I84V, appears to result in a conformational change of the flap region as suggested by fluorescence studies in section 3.3.2. It is possible that the V82F/I84V mutation results in a loss of intra-flap, inter-flap or flap-inhibitor contacts in relation to the wild-type C-SA and the V82A variant proteases. In addition, considering that ligand binding is enthalpically unfavorable or only slightly favorable (ritonavir), (this study and previous studies by Velazquez-Campoy, *et al.*, 2002), due partly to the closing and rearrangement of the flaps (Luque *et al.*, 1998), it is plausible to assume that the V82F/I84V mutation in the C-SA subtype induces an even more open conformation at the flap region, which contributes to the observed inhibition properties. This is because all the clinical inhibitors have to pay a larger enthalpic cost to close the flap region or a higher binding energy is required to bring the flap regions of the V82F/I84V mutant into a bound conformation (Todd *et al.*, 2000). Consequently, the mobility of the flaps in the V82F/I84V variant is enhanced and this flap region could move further away from the active site, thus exposing a larger non-polar surface to the solvent.

Previous calorimetric studies have also shown that the changes in binding affinity induced by the V82F/I84V mutation are similar, independent of the subtype background (Velazquez-Campoy *et al.*, 2002). In the B subtype, for example, the double mutation V82F/I84V lowers the binding affinity for saquinavir and nelfinavir by a factor ~ 20, indinavir by 68 and ritonavir by 368 (Velazquez-Campoy *et al.*, 2002). When the binding energetics of clinical inhibitors for the wild-type B subtype are compared to those of the wild-type C-SA, V82A and the V82F/I84V HIV-1 proteases (Table 8) it becomes clear that the inhibitory potency is affected for the V82A mutation and more so for the V82F/I84V mutation. In the C-SA subtype, the V82F/I84V mutation lowers the binding affinity for the clinical inhibitors by factors ranging from ~120 to ~1100 (Table 6). But when the effects of the V82F/I84V mutation in the C-SA subtype are compared to the B subtype as a standard, the efficacy of the clinical inhibitors is significantly hampered. Within the C-SA subtype, the V82F/I84V mutation lowers the affinity of inhibitors by factors ranging from ~ 240 to 7670 (see Table 8). The affinity of saquinavir, ritonavir, indinavir and nelfinavir against the V82F/I84V is weakened far below the level required for

Table 8: Relative changes for the binding inhibitors to the wild-type C-SA, V82A and V82F/I84V proteases. The wild-type B subtype is used as a reference. The K_d values for the wild-type B subtype were obtained from Velazquez-Campoy *et al.* (2002).

	wild-type B	wild-type C-SA	V82A	V82F/I84V
Inhibitor		**K_d ratio**		
Saquinavir	1.0	2.1	4.6	245
Ritonavir	1.0	7.0	67	7667
Indinavir	1.0	2.0	26.5	938
Nelfinavir	1.0	1.2	2.5	423

effective inhibition. Thus, the background polymorphisms in the C-SA protease in combination with drug-resistant mutations can drastically affect the inhibitory potency of the first-generation inhibitors currently in use. Amplification of drug resistance which is generally additive, i.e., the effect of a drug resistant mutation on the potency of the inhibitors is amplified by a factor that roughly corresponds to the pre-existing loss in binding affinity due to the background polymorphisms, can thus have serious implications in the long term viability of protease inhibition therapy in the C-SA subtype. In addition, higher vitality ratios obtained for the V82F/I84V protease in this study signify that this set of mutations will adversely affect the binding of all four clinical inhibitors (see Table 7). It would, thus, be more advantageous for the virus to incorporate the V82F/I84V mutation as a drug resistant mutation when exposed to all four of these inhibitors. Previous studies have shown that conformationally rigid ligands, such as the first-generation protease inhibitors in clinical use, have or no capacity to adapt to changes in the target binding site and lose affinity significantly even when confronted by conservative mutations (Velazquez-Campoy *et al.,* 2001b; Velazquez-Campoy *et al.,* 2002; Ohtaka *et al.,* 2002). Due to the absence of a crystal structure of the HIV-1 C-SA protease in complex with first-generation inhibitors, it is difficult to interpret the thermodynamic parameters in terms of structural data, but the results obtained in these studies can only be corroborated in relation to the HIV-1 subtype B protease-inhibitor complexes with the corresponding V82A and V82F/I84V mutations as shown in section 4.4. Lastly, detailed knowledge of the thermodynamics of ligand binding to the wild-type C-SA and its mutant forms and the elucidation of the three-dimensional structures of the proteases (ligand-free and ligand-bound) will serve as invaluable tools in the design of new potent drugs specific to the C-SA subtype and drug-resistant mutants. The crystal structures of the wild-type C-SA, V82A and V82F/I84V HIV-1 proteases can also be used to gain insight into the molecular basis of drug resistance. Darunavir and tipranavir, as second-generation protease inhibitors, have been shown to bind to the HIV-1 subtype B protease in a process strongly favored by enthalpic contributions primarily due to the addition of stronger interactions established between the inhibitors and main-chain atoms of the protein at the bottom of the active site (King *et al.,* 2004b; Muzammil *et al.,* 2007). Although resistance mutations do cause a drop in binding affinity, second-generation inhibitors still bind to the mutant enzymes with an affinity that is about 1.5 order of magnitude higher than the first-generation protease inhibitors (King *et al.,* 2004b).

Therefore, it will be interesting to see how the C-SA subtype and drug-resistant mutations function in the presence of second-generation inhibitors which have been designed for optimized target binding and inhibition even if the target is polymorphic and prone to evolutionary change as a consequence of various selection pressures.

4.6 Conclusions

These results demonstrate that the natural amino acid polymorphisms found in the HIV-1 subtype C-SA protease in combination with small changes in the active site might have significant effects on binding affinities of structurally rigid inhibitors, without a similar effect on the binding affinity of more flexible molecules like peptide substrates. From a thermodynamic standpoint, these results are consistent with previous studies that indicated that not all inhibitors respond identically to mutations (Todd *et al.,* 2000; Velazquez-Campoy *et al.,* 2001b; Velazquez-Campoy *et al.,* 2002; Ohtaka *et al.,* 2002; Vega *et al.,* 2004). Overall, a reduction in binding affinity is reflected in a loss of binding enthalpy and entropic contributions. From a structural point of view, however, the thermodynamic data indicate that for all four inhibitors (saquinavir, ritonavir, indinavir and nelfinavir) used in this study, the protein-inhibitor interactions are not strong enough to overcome (1) the unfavorable enthalpy associated with the desolvation of the hydrophobic inhibitor and the binding site, and (2) the unfavorable enthalpy associated with the change in the flap region upon complex formation. The first-generation clinical inhibitors exhibit the highest binding affinity to both the wild-type and the V82A enzymes, but are extremely sensitive to the V82F/I84V mutation. The affinity of saquinavir, ritonavir, indinavir and nelfinavir against the V82F/I84V mutation is weakened far below the level required for effective inhibition. Thus, the background polymorphisms in the C-SA protease in combination with drug-resistant mutations can drastically affect the inhibitory potency of the first-generation inhibitors currently in use. This can, therefore, have serious implications in the long term viability of protease inhibition therapy in the clade C-SA virus. Lastly, the elucidation of the three-dimensional structures of the proteases (ligand-free and ligand-bound) will serve as an invaluable tool in the design of new potent drugs specific to the C-SA subtype and drug-resistant mutants. It will also be interesting to see how the C-SA subtype and drug-resistant mutations function in the presence of second-generation adaptive inhibitors which have been designed for optimized target binding and inhibition.

REFERENCES

Abbenante, G. and Fairlie, D. P. (2005) Protease inhibitors in the clinic. *Medicinal Chemistry* **1**, 71-104.

Ala, P. J., Huston, E. E., Klabe, R. M., McCabe, D. D., Duke, J. L., Rizzo, C. J., Korant, B. D., Deloskey, R. J., Lam, P. Y. S., Hodge, C. N. and Chang, C. H. (1997) Molecular basis of HIV-1 protease drug resistance: Structural analysis of mutant proteases complexed with cyclic urea inhibitors. *Biochemistry* **36**, 1573-1580.

Ala, P. J., Huston, E. E, Klabe, R. M., Jadhav, P. K., Lam, P. Y. and Chang, C. H. (1998) Counteracting HIV-1 protease drug resistance: structural analysis of mutant proteases complexed with XV638 and SD146, cyclic urea amides with broad specificities. *Biochemistry* **37,** 15042-15049.

Aruksakunwong, O., Wolschann, P., Hannongbua, S. and Sompornpisut, P. (2006) Molecular dynamics and free energy studies of primary resistance mutations in HIV-1 protease-ritonavir complexes. *J. Chem. Inf. Model* **46**, 2085-2092.

Baker, B. M. and Murphy, K. P. (1996) Evaluation of linked protonation effects in protein binding reactions using isothermal titration calorimetry. *Biophys. J.* **71**, 2049-55.

Baker, B.M. and Murphy, K. P. (1997) Dissecting the energetics of a protein-protein interaction: the binding of ovomucoid third domain to elastase. *J. Mol. Biol.* **268**, 557-569.

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

Baldwin, E. T., Bhat, T .N., Gulnik, S., Liu, B., Topol, I. A., Kiso, Y., Mimoto, T., Mitsuya, H., Erickson, J. W. (1995b) Structure of HIV-1 protease with KNI-272, a

tight-binding transition-state analog containing allophenylnorstatine. *Structure* **3**, 581-590.

Barlow, D. J. and Thornton, J. M. (1983) Ion-pairs in proteins. *J. Mol. Biol.* **168**, 867-885.

Becker-Pergola, G., Kataaha, P., Johnston-Dow, L., Fung, S., Jackson, J. B., Eshleman, S. H. (2000) Analysis of HIV type 1 protease and reverse transcriptase in antiretroviral drug-naive Ugandan adults. *AIDS Res. Hum. Retroviruses* **16**, 807-813.

Benson, D. A., Boguski, M. S., Lipman, D. J, Ostell, J. and Ouellette, B. F. (1998) GenBank. *Nucleic Acids Res.* **26**, 1-7.

Betts, M. J. and Sternberg, M. J. (1999) An analysis of conformational changes on protein-protein association: implications for predictive docking. *Protein Eng.* **12**, 271-283.

Boden, D. and Markowitz, M. (1998) Resistance to immunodeficiency virus type 1 protease inhibitors. *Antimicrob. Agents Chemotherap.* **42**, 2775-2783.

Breckenridge, R. J. (1991) Molecular recognition: models for drug design. *Experientia* **47**, 1148-1161.

Bujacz, G., Jaskolski, M., Alexandratos, J., Wlodawer, A., Merkel, G. and Skalka, A. M. (1995) High resolution structure of the catalytic domain of the avian sarcoma virus integrase. *J. Mol. Biol.* **253**, 333-246.

Bujacz, G., Alexandratos, J., Zhou-Liu, Q., Clement-Mella, C. and Wlowader, A. (1996) The catalytic domain of human immunodeficiency virus integrase: ordered active site in the F185H mutant. *FEBS Lett.* **398**, 175-178.

Cabani, S., Gianni, P., Mollica, V. and Lepori, L. (1981) Group distribution to the thermodynamic properties of non-ionic solutes in dilute aqueous solution. *J. Solution Chem.* **10**, 563-595.

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

Carbonell, T. and Freire, E. (2005) Binding thermodynamics of statins of HMG-CoA reductase. *Biochemistry* **44**, 11741-11748.

Caret, R. L., Denniston, K. J. and TopPing, J. J. (1993) Enzyme catalysis, in *Principles and applications of inorganic, organic and biological chemistry*. pp 465-466, Wm C. Brown Communications, Inc, U.S.A.

Carpenter, C. J., Fischl, M. A., Hammer, S. M., Hirsch, M. S., Jacobsen, D. M., Katzeinstein, D. A. et al. (1998) Antiretroviral therapy for HIV infection in 1998: updated recommendations of the International AIDS Society-USA panel. *J. Am. Med. Assoc.* **280**, 78-86.

Carrillo, A., Stewart, K. D., Sham, H. L., Norbeck, D. W., Kohlbrenner, W. E., Leonard, J. M., Kempf, D. J. and Molla, A.(1998) In vitro selection and characterization of human immunodeficiency virus type 1 variants with increased resistance to ABT-378, a novel protease inhibitor. *J. Virol.* **72**, 7532-7541.

Chen, Z., Li, Y., Chen, E., Hall, D. L., Darke, P. L., Culberson, C., Shafer, J. A. and Kuo, L. C. (1994) Crystal structure at 1.9 Å resolution of human immunodeficiency virus (HIV) II protease complexed with L-735, 524, an orally bioavailable inhibitor of the HIV proteases. *J. Biol. Chem.* **269**, 26344-26348.

Chen, Z., Schock, H. B., Hall, D., Chen, E. and Kuo, L. C. (1995) Three-dimensional structure of a mutant HIV-1 protease displaying cross-resistance to all protease inhibitors in clinical trials. *J. Biol. Chem.* **270**, 21433-21436.

Chung, C. T., Niemela, S. L. and Miller, R.H. (1989) One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution. *Proc. Natl. Acad. Sci. U. S. A.* **86**, 2172-2175.

Chrusciel, R. A. and Strohbach, J. W. (2004) Non-peptidic HIV protease inhibitors. *Curr. Top. Med. Chem.* **4**, 1097-1114.

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

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

Cohen J. (2000) Ground zero: AIDS research in Africa. *Science* **288**, 2150-5153.

Collier, A. C., Coombs, R. W., Schnoefeld, D. A., Bassett, R. L., Timpone, J., Baruch, A., Jones. M., Facey, K., Whitacre, C., McAuliffe, V. J., Friedman, H. M., Merigan, T. C., Reichman, R. C., Hooper, C. and Corey, L. (1996) Treatment of human immunodeficiency virus infection with saquinavir, zidovudine and zalcitabine. *New Engl. J. Med.* **334**, 1101-1017.

Collono, R. J., Thiry, A. and Limoli, K. (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**, 1324-1333.

Condra, J. H. W. A., Schleif, O. M., Blaby, L. J., Gabryelski, D. J., Graham, J. C., Quintero, A., Rhodes, H. L., Robbins, E., Roth, M., Shivaprakash, D., Titus, T., Yang, H., Teppler, K. E., Squires, P. J. and Emini, E. A. (1995). *In vivo* emergence of HIV-1 variants resistant to multiple protease inhibitors. *Nature* **374**, 569-571.

Condra, J. H. D. J., Holder, W. A., Schleif, O. M., Blaby, R. M. Danovich, I. 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, K. E. and

Emini, E. A. (1996). Genetic correlates of in vivo viral resistance to indinavir, a human immunodeficiency virus type 1 protease inhibitor. *J. Virol.* **70,** 8270-8276.

Condra, J. H. D. J. (1998) Resistance to HIV protease inhibitors. *Haemophilia* **4**, 610-615.

Connelly, P. R., Aldape, R. A., Bruzzese, F. J., Chambers, S. P., Fitzgibbon, M. J., Fleming, M. A., Itoh, S., Livingston, D. J., Navia, M. A., Thornton, J. A. and et al., (1994) Enthalpy of hydrogen bond formation in a protein-ligand binding reaction. *Proc. Natl. Acad. Sci. U S A* **91**, 1964-1968.

Cooper, A. (1999) Thermodynamic analysis of biomolecular interactions. *Curr. Opin. Chem. Biol.* **3**, 557-563.

Creighton, T. E. (1993) Physical interactions that determine the properties of proteins, in *Protein structures and molecular properties*. pp 142-157, W.H. Freeman and company, USA.

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

Dash, C. and Rao, M. (2001) Interactions of a novel inhibitor from an extremophilic *Bacillus* sp. with HIV-1 protease. *J. Biol. Chem.* **276**, 2487-2493.

Davis, D. A., Dorsey, K., Wingfield, P. T., Stahl, S. J., Kaufman, J., Fales, H. M., Levine, R. L. (1996) Regulation of HIV-1 protease activity through cysteine modification. *Biochemistry* **35**, 2482-2488.

Davis, D. A., Yusa, K., Gillim, L. A., Newcomb, F. M., Mitsuya, H. and Yarchoan, R. (1999) Conserved cysteines of the human immunodeficiency virus type 1 protease are involved in regulation of polyprotein processing and viral maturation of immature virions. *J. Virol.* **73**, 1156-1164.

De Meyer, S., Azjin, H., Surleraux, D. Jochmans, D., Tahri, A., Pauwels, R., Wigerinck, P. and de Bethume, M. P. (2005) TMC114, a novel human deficiency virus type 1 protease inhibitor active against protease inhibitor-resistant viruses, including a broad range of clinical isolates. *Antimicrob. Agents Chemother.* **49**, 2314-2321.

DeVita V. T. Jr., Broder, S., Fauci, A. S., Kovacs, J. A. and Chabner, B. A. (1987) Developmental therapeutics and the acquired immunodeficiency syndrome. *Ann. Intern. Med.* **106**, 568-581.

Dill, K. A. (1990) Dominant forces in protein folding. *Biochemistry* **29**, 7133-7155.

Dorsey, B. D., Levin, R. B., McDaniel, S. L., Vacca, J. P., Guare, J. P. and Darke, P. L., Dorsey, B. D., Schleif, W. A., McDaniel, S. L., Zugay, J., Quintero, J. C., Blahy, O. M., Roth, E., Sardana, V. V., Schlabach, A. J., Graham, P. I., Condra, J. H., Gotlib, L., Holloway, M. K., Lin, J., Chen, I. W., Vastag, K. V., Oatovich, D., Anderson, P. S., Emini, E. A. and Huff, J. R. (1994) L-735,524: the design of a potent and orally bioavailable HIV protease inhibitor. *J. Med. Chem.* **37**, 3443-3451.

Doyle, M. L., Louie, G., Dal Monte P. R. and Sokoloski, T. D. (1995) Tight binding affinities determined from thermodynamic linkage to protons by titration calorimetry. *Methods Enzymol.* **259**, 183-194.

Doyon, L., Tremblay, S., Bourgon, L., Wardrop, E. and Cordingley, M. G. (2005) Selection and characterization of HIV-1 showing reduced susceptibility to the non-peptidic protease inhibitor tipranavir. *Antiviral. Res.* **68**, 27-35.

Dreyer, G. B., Lambert, D. M., Meek, T. D., Carr, T. J., Tomaszek, T. A. Jr, Fernandez, A. V., Bartus, H., Cacciavillani, E., Hassell, A. M. and Minnich, M. (1992) Hydroxyethylene isostere inhibitors of human immunodeficiency virus-1 protease: structure-activity analysis using enzyme kinetics, X-ray crystallography, and infected T-cell assays. *Biochemistry* **31**, 6646-6659.

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

Dyda, F., Hickman, A. B., Jenkins, T. M., Engelman, A., Craigie, R. and Davies, D. R. (1994) Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. *Science* **266**, 1981-1986.

Eberle, J., Bechowsky, B, Rose, D., Hauser, U., vonder Helm, K., Guertler, L. and Nitschko, H. (1995) Resistance of HIV type 1 to proteinase inhibitor Ro31-8959. *AIDS Res. Hum. Retroviruses* **11,** 671-676.

Ernest II, C. S., Hall, S. D. and Jones, D. R. (2005) Mechanism-based inactivation of CYP3A by HIV protease inhibitors. *J. Pharmacol. Exp. Ther.* **312**, 583-591.

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

Essex, M. (1999) Human immunodeficiency viruses in the developing world. *Adv. Virus Res.* **53,** 71-88.

Fitzgerald, P. M., McKeever, B. M., VanMiddlesworth, J. F., Springer, J. P., Heimbach, J. C., Leu, C. T., Herber, W. K., Dixon, R. A. and Darke, P. L. (1990) Crystallographic analysis of a complex between human immunodeficiency virus type 1 protease and acetyl-pepstatin at 2.0-A resolution. *J. Biol. Chem.* **265**, 14209-14219.

Fitzgerald, P. M. and Springer, J. P. (1991) Structure and function of retroviral proteases. *Annu. Rev. Biophys. Biophys. Chem.* **20**, 299-320.

Fitzgerald, P. M. D. (1993) HIV protease-ligand complexes. *Curr. Opin. Struct. Biol.* **3**, 868-874.

Foulkes, J. E., Prabu-Jeyabalan, M., Cooper, D., Henderson, G. J., Harris, J., Swanstrom, R. and Schiffer, C. A. (2006) Role of invariant Thr80 in human

immunodeficiency virus type 1 protease structure, function, and viral infectivity. *J. Virol.* **80**, 6906-6916.

Franks, F. (1975) *The hydrophobic effect. In water: A comprehensive treatise.* Vol. 4 Plenum Press, New York.

Frater, A. J., Beardall, A., Ariyoshi, K., Churchill, D., Galpin, S., Clarke, J. R., Weber, J. N. and McClure, M. O. (2001) Impact of baseline polymorphisms in RT and protease on outcome of highly active antiretroviral therapy in HIV-1-infected African patients. *AIDS* **15**, 1493 – 1502.

Freire, E., Mayorga, O. L. and Straume, M. (1990) Isothermal titration calorimetry. *Anal. Biochem.* **62**, 950A-959A.

Fruton, J. S. (1976) The mechanism of catalytic action of pepsin and related acid proteinases. *Advan. Enzymol*. **44**, 1-36.

Gallo, R. C. and Montagnier, L. (1988) AIDS in 1988. *Sci. Am.* 25941-25948.

Gomez, J. and Freire, E. (1995) Thermodynamic mapping of the inhibitor site of the aspartic protease endothiapepsin. *J. Mol. Biol.* **252**, 337-350.

Gong, Y. F., Robison, B. S., Rose, R. E., Deminie, C., Spicer, T. P., Stock, D. Colonno, R. J. and Lin, P. F. (2000) *In vitro* resistance profile of the human immunodeficiency virus type 1 protease inhibitor BMS-232632 *Antimicrob. Agents Chemother.* **44**, 2319-2326.

Greene, W. C. (1993) Aids and the immune-system. *Scientific AmScientific American* **269**, 98-105.

Guex, N. and Peitsch, M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714-2723.

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

Gulnik, S., Erickson, J. W. and Xie, D. (1999) HIV protease: enzyme function and drug resistance. In: Litwack, G. (Ed.), Vitamins and Hormones. Academic Press, New York.

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

Gustchina, A. and Weber, I. T. (1990) Comparison of inhibitor binding in HIV-1 protease and in non-viral aspartic proteases: the role of the flap. *FEBS Lett.* **269**, 269-272.

Gustchina, A., Kervinen, J., Powell, D. J., Zdanov, A., Kay, J. and Wlodawer, A. (1996) Structure of equine infectious anemia virus proteinase complexed with an inhibitor. *Protein Sci.* **8**, 1453-1465.

Habermann, S. M. and Murphy, K. P. (1996) Energetics of hydrogen bonding in proteins: a model compound study. *Protein Sci.* **5**, 1229-1239.

Hammer, S. M., Squires, K. E., Hughes, M. D., Grimes, M. J., Dementer, L. M., Currier, J. S., Eron. J. J., Feinberg, J. E., Balfour, H. H., Deyton, L. R., Chodakewitz, J. A. and Fischl, M. A. (1997) A controlled trial of two nucleosides analogues plus indinavir in persons with human immunodeficiency virus infection and CD4 cell counts of 200 per cubic millimeter or less. AIDS Clinical Trials Group 320 Study Team. *New Engl. J. Med.* **337**, 725-733.

Hansen, J., Billich, S., Schulze, T., Sukrow, S. and Moelling, K. (1988) Partial purification and substrate analysis of bacterially expressed HIV protease by means of monoclonal antibody. *EMBO. J.* **6**, 1785-1791.

Heaslet, H., Kutilek, V., Morris, G. M., Lin, Y., Elder, J. H., Torbett, B. E. and Stout, D. C. (2006) Structural insights into the mechanisms of drug resistance in HIV-1 protease NL4-3. *J. Mol. Biol*. **356**, 967-981.

Heaslet, H., Rosenfeld, R., Giffin, M., Lin, Y., Tam, K., Torbett, B. E., Elder, J. H., McRee, D. E. and Stout, C. D. (2007) Conformational flexibility in the flap domains of ligand-free HIV protease. *Acta Cryst.* **D 63**, 866-875.

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

Ho, D. D. (1994) HIV-1 dynamics in vivo. *J. Biol. Regul. Homeost. Agents.* **9**, 76-77.

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

Holguin, A. and Soriano, V. (2002) Resistance to antiretroviral agents in individuals with HIV-1 non-B subtypes. *HIV Clin Trials.* **3**, 403-411.

Hong, L., Treharne, A., Hartsuck, J. A., Foundling, S. and Tang, J. (1996) Crystal structures of complexes of a peptidic inhibitor with wild type and two mutant HIV-1 proteases. *Biochemistry* **35**, 10627-10633.

Hong, L., Zhang, X. J., Foundling, S., Hartsuck, J. A. and Tang, J. (1997) Structure of a G48H mutant of HIV-1 protease explains how glycine-48 replacements produce mutants resistant to inhibitor drugs. *FEBS Lett.* **420**, 11-16.

Hong, L., Hartsuck, J. A., Foundling, S., Ermolieff, J. and Tang, J. (1998) Active site mobility in human immunodeficiency virus, type 1, protease as demonstrated by crystal structure of A28S mutant. *Protein Sci.* **7**, 300-305.

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

Hosur, M. V., Bhat, T. N., Kempf, D., Baldwin, E. T., Liu, B., Gulnik, S., Wideburg, N. E., Norbeck, D. W., Appelt, K. and Erickson, J. W. (1994) Influence of Stereochemistry on Activity and Binding Modes for C2 Symmetry-Based Diol Inhibitors of HIV-1 Protease. *J. Am. Chem. Soc.* **116**, 847-851.

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

Ishima, R. and Torchia, D. A. (2000) Protein dynamics from NMR. *Nature Struct. Biol*. **7**, 740-743.

Jacks, T., Power, M. D., Masiarz, F. R., Luciw, P. A., Barr, P. J. and Varmus, H. E. (1998) Characterisation of ribosomal frameshifting in HIV-1 gag-pol expression. *Nature* **331**, 280-283.

Jacobo-Molina, A., Ding, J., Nanni, R. G., Clark, A. D. Jr., Lu, X., Tantillo, C., Williams, R.L., Kamer, G., Ferris, A.L., Clark, P., Hizi, A., Hughes, S. H. and Arnold, E. (1993) Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 Å resolution shows bent DNA. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6320-6324.

Jacobsen, H., Brun-Vezinet, F., Duncan, I., Hanggi, M., Ott., M., Vella, S., Weber, J. and Mous, J. (1994) Genotypic characterization of HIV-1 from patients after prolonged treatment with increased resistance to a C2-symmetric protease inhibitor. *J. Virol.* **68**, 2016-2020.

Jacobsen, H. Yasargil, K., Winslow, D. L., Craig, J. C., Kroehn, A., Duncan, I. B. and Mous, J. (1995) Characterisation of human immunodeficiency virus type 1 mutant with decreased sensitivity to proteinase inhibitor Ro31-8959, *Virology* **206**, 527-534.

Jamin, J. and Chothia, C. (1990) The structure of protein-protein recognition sites. *J. Biol. Chem.* **265**, 16027-16030.

Jaskolski, M., Tomasselli, A. G. Sawyer, T. K., Staples, D. G., Heinrikson, R. L., Schneider, J., Kent, S. B. and Wlodawer, A. (1991) Structure at 2.5 Á resolution of chemically synthesized human immunodeficiency virus type 1 protease complexed with a hydroxyethylene-based inhibitor. *Biochemistry* **30**, 1600-1609.

Ji, J. P. and Loeb, L. A. (1992) Fidelity of HIV-1 reverse transcriptase copying RNA in vitro. *Biochemistry* **31**, 954-958.

Jiang, L. and Lai, L. (2002) CH...O hydrogen bonds at protein-protein interfaces. *J. Biol. Chem.* **277**, 37732-37740.

Kaldor, S. W., Kalish, V. J., Davies, J. F., Shetty, B. V., Fritz, J. E., Applet, K., Burgess, J. A., Campanale, K. M., Chirgadze, N. Y., Clawson, D. K., Dressman, B. A., Hatch, S. D., Khalil, D. A., Kosa, M. B., Lubbehusen, P. P., Muesing, M. A., Patick, A. K., Reich, S. H., Su, K. S. and Tatlock, J. H. (1997) Viracept (nelfinavir mesylate, AG1343): a potent, orally bioavailable inhibitor of HIV-1 protease. *J. Med. Chem.* **40**, 3979-3985.

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

Kempf, D. J., Marsh, K. C., Denissen, J. F., McDonald, E., Vasavanonda, S. and Flentge, C. A. (1995) ABT-538 is a potent inhibitor of human immunodeficiency virus protease and has high oral bioavailability in humans. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 2484-2488.

Kervinen, N., Thanki, N., Zdanov, A., Tino, J., Barrish, J., Lin, P., Colonno, F., Riccardi, K., Samanta, H. and Wlodawer, A. (1996) Structural analysis of the native and drug-resistant HIV-1 proteinases complexed with an amino-diol inhibitor. *Protein Pept. Lett.* **3**, 399.

Kim, E. E., Baker, C. T., Dwyer, M. D., Murcko, M. A., Rao, B. G. and Tung, R. D. (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.* 1181-1182.

King, R. W., Winslow, D. L., Garbe, S., Bacheler, L. T., Stack, S. and Otto, M. J. (1995) Identification of a clinical isolate of HIV-1 with an isoleucine at position 82 of the protease which retains susceptibility to protease inhibitors. *Antiviral Res.* **28**, 13-24.

King, N. M., Prabu-Jeyabalan, M., Nalivaika, E. A. and Schiffer, C. A. (2004a) Combating susceptibility to drug resistance: lessons from HIV-1 protease. *Chem. Biol.* **11**, 1333-1338.

King, N. M., Prabu-Jeyabalan, M., Nalivaika, E. A. and Schiffer, C. A. (2004b) Structural and thermodynamics basis for the binding of TMC114, a next generation human immunodeficiency virus type 1 protease inhibitor. *J. Virol.* **78**, 12012-12021.

Klabe, R. M., Bacheler, L. T., Ala, P. J., Erickson-Viitanen, S. and Meek, J. L. (1998) Resistance to HIV protease inhibitors: a comparison of enzyme inhibition and antiviral potency. *Biochemistry* **37**, 8735-8742.

Koh, Y., Nakata, H., Maeda, K., Ogata, H., Bilcer, G., Devasamudram, T., Kincaid, J. F., Boross, P., Wang, Y. F., Tie., Y., Volarath, P., Gaddis, L., Harrison, R. W., Weber, I. T., Ghosh, A. K. and Mitsuva, H. (2003) Novel bis-tetrahydrofuranylurethane-containing nonpeptidic protease inhibitor (PI) UIC-94017 (TMC114) with potent activity against multi-PI-resistant human immunodeficiency virus in vitro. *Antimicrob. Agents Chemother.* **47**, 3123-3129.

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

Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A. and Steitz, T. A. (1992) Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science* **256**, 1783-1790.

Koshland, D. (1994) Key-lock and induced fit theory. *Angew. Chem. Int. Ed. Engl.* **33**, 2375-2378.

Kozal, M. J., Shah, N., Shen, N., Yang, R., Fucini, R., Merigan, T. C., Richman, D. D., Morris, D., Hubbell, E., Chee, M. and Gingeras, T. R. (1996) Extensive polymorphisms observed in HIV-1 clade B protease gene using high-density oligonucleotide arrays. *Nat. Med.* **2**, 753-759.

Kramer, R. A., Schaber, M. D., Skalka, A. M., Ganguly, K., Wong-Staal, F. And Reddy, E. P. (1986) HTLV-III *gag* protein is processed in yeast cells by the virus *pol*-protein. *Science* **231**, 1580-1585.

Kumar, M., Prashar, V., Mahale, S. and Hosur, M. V. (2005) Observation of a tetrahedral reaction intermediate in the HIV-1 protease-substrate complex. *Biochem. J.* **389**, 365-371.

Lacowicz, J. R. (1999) *Principles of fluorescence spectroscopy.* Plenum Press, New York.

Ladbury, J. E. and Williams, M. A (2004) The extended interface: measuring non-local effects in biomolecular interactions. *Curr. Opin. Struct. Biol.* **14**, 562-569.

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

Lambert, D. M., Petteway, Jr., S. R., McDanal, C. E., Hart, T. K., Leary, J. J., Dreyer, G. B., Meek, T. D., Buigelski, P. J., Bolognesi, D. P., Metcalf, B. W. and Matthews, T. J. (1992) Human immunodeficiency virus type 1 protease inhibitors irreversibly

block infectivity of purified virions from chronically infected cells. *Antimicrob. Agents Chemother.* **36**, 982-988.

Leavitt, S. and Freire, E. (2001) Direct measurement of protein binding energetics by isothermal titration calorimetry. *Curr. Opin. Struct. Biol*. **11**, 560-566.

Lipinski, C. A. (2000) Drug-like properties and the causes of poor solubility and poor permeability. *J. Pharmacol. Toxicol. Methods.* **44**, 235-249.

Liu, F., Kovalevsky, A. Y., Louis, J. M., Boross, P. I., Wang, Y. –F., Harrison, R. W. and Weber, I. T. (2006) Mechanism of drug resistance revealed by the crystal structure of the unliganded HIV-1 protease with F53L mutation. *J. Mol. Biol*. **358**, 1191-1199.

Luque I, Todd, M. J., Gomez, J., Semo, N. and Freire, E. (1998) Molecular basis of resistance to HIV-1 protease inhibition: a plausible hypothesis. *Biochemistry* **37**, 5791-5797.

Luque, I. and Freire, E. (2000) Structural stability of binding sites: consequences for binding affinity and allosteric effects. *Proteins* **4**, 63-71.

Luque, I. and Freire, E. (2002) Structural parameterization of the binding enthalpy of small ligands. *Proteins* **49**, 181-190.

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

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

Makhatadze, G. I. and Privalov, P. L. (1995) Energetics of protein structure. *Adv. Protein Chem.* **47**, 307-425.

Marqusee, S. and Sauer, R. T. (1994) Contributions of a hydrogen bond/salt bridge network to the stability of secondary and tertiary structure in lambda repressor. *Protein Sci.* **3**, 2217-2225.

Martin, J. A. (1992) Recent advances in the design of HIV proteinase inhibitors. *Antiviral Res.* **17**, 265-278.

Maschera, B., Darby, G., Palú, G., Wright, L. L., Tisdale, M., Myers, R., Blair, E. D. and Furfine, E. S.(1996) Human immunodeficiency virus. Mutations in the viral protease that confer resistance to saquinavir increase the dissociation rate constant of the protease-saquinavir complex. *J. Biol. Chem.* 271, 33231-33235.

McKinnon, I. R., Fall, L., Parody-Morreale, A. and Gill, S. J. (1984) A twin titration microcalorimeter for the study of biochemical reactions. *Anal. Biochem.* **139**, 134-139.

Miller, M., Jaskolski, M., Rao, J. K., Leis, J. and Wlodawer, A. (1989a) Crystal structure of a retroviral protease proves relationship to aspartic protease family. *Nature* **337**, 576-579.

Miller, M., Schneider, J., Sathyanarayana, B. K., Toth, M. V., Marshall, G. R., Clawson, L., Selk, L., Kent, S. B. H. and Wlodawer, A. (1989b) Structure of complex of synthetic HIV-1 protease with a substrate-based inhibitor at 2.3 Å resolution. *Science* **246**, 1149-1152.

Molla, A., Korneyeva, M., Gao, Q., Vasavanonda, S., Schipper, P. J., Mo, H. M., Markowitz, M., Chernyavskiy, T., Niu, P., Lyons, N., Hsu, A., Granneman, G. R., Ho, D. D., Boucher, C. A., Leonard, J. M., Norbeck, D. W. and Kempf, D. J. (1996) Ordered accumulation of mutations in HIV protease confers resistance to ritonavir. *Nat. Med.* **2**, 760-766.

Montano, M. A., Novitsky, V. A., Blackard, J. T., Cho, N. L., Katzenstein, D. A. and Essex, M. (1997) Divergent transcriptional regulation among expanding human immunodeficiency virus type 1 subtypes. *J. Virol.* **71,** 8657-8665.

Moore, M. L. and Dreyer, G. B. (1993) Substrate-based inhibitors of HIV-1 protease. *Prospect. Drug Discovery Des.* **1**, 85-108.

Mugnaini, C., Petricci, E., Corelli, F. and Botta, M. (2005) Combinatorial chemistry as a tool for targeting different stages of the replicative HIV-1 cycle. *Comb. Chem. and High Thr. Scr.* **8**, 387-401.

Mulichak, A. M., Hui, J. O., Tomasselli, A. G., Heinrikson, R. L., Curry, K. A., Tomich, C. S., Thaisrivongs, S., Sawyer, T. K. and Watenpaugh, K. D. (1993) The crystallographic structure of the protease from human immunodeficiency virus type 2 with two synthetic peptidic transition state analog inhibitors. *J. Biol. Chem.* **268,** 13103-13119.

Murphy, K. P. (1999) Predicting binding energetics from structure: looking beyond ΔG°. *Med. Res. Rev.* **19**, 333-339.

Muzammil, S., Ross, P. and Freire, E. (2003) A major role for a set on non-active site mutations in the development of HIV-1 protease drug resistance. *Biochemistry* **42**, 631-638.

Muzammil, S., Armstrong, A. A., Kang, L. W., Jakalian, A., Bonneau, P. R., Schmelmer, V., Amzel, L. M. and Freire, E. (2007) Unique thermodynamic response of tipranavir to human immunodeficiency virus type 1 protease drug resistance mutations. *J. Virol.* **81**, 5144-5154.

Navia, M. A., Fitzgerald, P. M., McKeever, B. M., Leu, C. T., Heimbach, J. C., Herber, W. K., Sigal, I. S., Darke, P. L. and Springer, J. P. (1989) Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. *Nature* **337**, 615-620.

Ode, H., Matsuyama, S., Hata, M., Neya, S., Kakizawa, J., Sugiura, W. and Hoshino, T. (2007) Computational characterization of structural role of the non-active site

mutation M36I of human immunodeficiency virus type 1 protease. *J. Mol. Biol.* **370**, 598-607.

Ohtaka, H., Schon, A. and Freire, E. (2002) Overcoming drug resistance in HIV-1 chemotherapy: the binding thermodynamics of Amprenavir and TMC-126 to wild-type and drug-resistant mutants of the HIV-1 protease. *Protein Sci.* **11**, 1908-1916.

Ohtaka, H., Velazquez-Campoy, A., Xie, D. and Freire, E. (2003) Multidrug resistance to HIV-1 protease inhibition requires cooperative coupling between distal mutations. *Biochemistry* **42**, 13659-13666.

Ohtaka, H., Muzammil, S., Schon, A., Velazquez-Campoy, A., Vega, S. and Freire, E. (2004) Thermodynamic rules for the design of high affinity HIV-1 protease inhibitors with adaptability to mutations and high selectivity towards unwanted targets. *Int. J. Biochem. Cell Biol.* **36**, 1787-1799.

Ohtaka, H. and Freire, E. (2005) Adaptive inhibitors of the HIV-1 protease. *Prog Biophys Mol Biol.* **88**, 193-208.

Olsen, D. B., Stahlhut, M. W., Rutkwoski, C. A., Schock, H. B., vanOlden, A. L. and Kuo, L. C. (1999) Non-active site changes elicit broad-based cross-resistance of the HIV-1 protease to inhibitors *J. Biol. Chem.* **274**, 23699-23701.

Osmanov, S., Pattou, C., Walker, N., Schwardlander, B. and Esparza, J. (2002) Estimated global distribution and regional spread of HIV-1 genetic subtypes in the year 2000. *J. Acquir. Immune Defic. Syndr.* **29**, 184-190.

Otto, M. J., Garber, S. Winslow, D. L., Ried, C. D., Aldrich, P., Jadhav, P. K., Patterson, C. E., Hodge, C. N. and Cheng, Y. S. E. (1993) In vitro isolation and identification of human immunodeficiency virus (HIV) variants with reduced sensitivity to C-2 symmetrical inhibitors of HIV type 1 protease. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 7543-7547.

Papworth, C., Bauer, J. C., Braman, J. and Wright, D. A. (1996) Site-directed mutagenesis using double-stranded plasmid DNA templates. *Strategies* **9**, 3-4.

Partadelis, J. A., Yamaguchi, K., Tisdale, M., Blair, E. E., Falcione, C., Maschera, B., Myers, R. E., Pazhanisamy, S., Futer, O., Cullinan, A. B., Stuver, C. M., Byrn, R. A. and Livingston, D. J. (1995) In vitro selection and characterization of human immunodeficiency virus type 1 (HIV-1) isolates with reduced sensitivity to hydroxyethylamino sulfonamide inhibitors of HIV-1 aspartyl protease. *J. Virol.* **69**, 5228-5235.

Patick, A. K., Boritzki, T. J. and Bloom, L. A. (1997) Activities of human immunodeficiency virus type 1 (HIV-1) protease inhibitor nelfinavir mesylate in combination with reverse transcriptase and protease inhibitors against acute HIV-1 infection in vitro. *Antimicrob. Agents Chemother.* **41**, 2159-2164.

Patten, P. A., Gray, N. S., Yang, P. L., Marks, C. B., Wedemayer, G. J., Boniface, J. J., Stevens, R. C. and Schultz, P. G. (1996) The immunological evolution of catalysis. *Science* **271**, 1086-1091.

Perno, C. F., Cozz-Lepri, A., Balotta, C., Forbici, F., Violin, M., Bertoli, A., Facchi, G., Pezzotti, P., Cadeo, G., Tositti, G., Pasquinucci, S., Pauluzzi, S., Scalzini, A., Salassa, B., Vincenti, A., Phillips, A. N., Dianzani, F., Appice, A., Angarano, G., Monno, L., Ippolito, G., Moroni, M. and Monforte, A. (2001) Secondary mutations in the protease region of human immunodeficiency virus and virologic failure in drug-naïve patients treated with protease inhibitor based therapy. *J. Infect. Dis.* **184**, 983–991.

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

Pettit, S. C., Moody, M. D., Wehbie, R. S., Kaplan, A. H., Nantermet, P. V., Klein, C. A. and Swanstrom, R. (1993) 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**, 8017-8027.

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

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

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

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

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

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

Quay, S. C., Heropoulos, A., Commes, K. and Dzau, V. J. (1985) Probing the renin active site by collisional quenching of endogenous fluorescence. *J. Biol. Chem.* **260**, 15055-15058.

Ratner, L., Starcich, B., Josephs, S. F., Hahn, B. H., Reddy, E. P., Livak, K. J., Petteway, S. R. Jr., Pearson, M. L., Haseltine, W. A., Arya, S. K., et al. (1985) Polymorphism of the 3' open reading frame of the virus associated with the acquired immune deficiency syndrome, human T-lymphotropic virus type III. *Nucleic Acids Res.* **13**, 8219-8229.

Robinson, B. S., Riccardi, K. A., Gong, Y. F., Guo, Q., Stock D. A., Blair, W. S., Terry, B. J., Deminie, C. A., Djang, F., Colonno, R. J. and Lin, P. F. (2000) BMS-232632, a highly potent human immunodeficiency virus protease inhibitor that can be used in combination with other available antiretroviral agents. *Antimicrob. Agents Chemother.* **44**, 2093-2099.

Rose, R. B., Rose, J. R., Salto, R., Craik, C. S. and Stroud, R. M. (1993a) Structure of the protease from simian immunodeficiency virus: complex with an irreversible nonpeptide inhibitor. *Biochemistry* **32**, 12498-12507.

Rose, J. R., Salto, R. and Craik, C. S. (1993b) Regulation of autoproteolysis of the HIV-1 and HIV-2 proteases with engineered amino acid substitution. *J. Biol. Chem.* **268**, 11939-11945.

Rose, R. B., Craik, C. S., Douglas, N. L. and Stroud, R. M. (1996) Three-dimensional structures of HIV-1 and SIV protease product complexes. *Biochemistry* **39**, 12933-12944.

Rose, R. B., Craik, C. S. and Stroud, R. M. (1998) Three-dimensional structures of HIV-1 and SIV protease product complexes. *Biochemistry* **37**, 2607-2621.

Rosin, C. D., Belew, R. K., Walker, W. L., Morris, G. M., Olson, A. J. and Goodsell, D. S. (1999) Coevolution and subsite decomposition for the design of resistance-evading HIV-1 protease inhibitors. *J. Mol. Biol.* **287**, 77-92.

Ross, P. D. and Subramanian, S. (1981) Thermodynamics of protein association reactions: forces contributing to stability. *Biochemistry* **20**, 3096-3102.

Ruben, A. J., Kiso, Y. and Freire, E. (2006) Overcoming roadblocks in lead optimization: a thermodynamic perspective. *Chem. Biol. Drug Des.* **67**, 2-4.

Rusconi, S., La Seta, C. S., Citterio, P., Kurtagic, S., Violi, M., Balotta, C., Moroni, M., Galli, M. and d'Arminio-Monforte, A. (2000). Susceptibility to PNU-140690 (Tipranavir) of human immunodeficiency virus type 1 isolates derived from patients

with multidrug resistance to other protease inhibitors. *Antimicrob. Agents Chemother.* **44**, 1328-1332.

Sanches, M., Krauchenco, S., Martins, N. H., Gustchina, A., Wlodawer, A., and Polikarpov, I. (2007) Structural characterization of B and non-B subtypes of HIV-protease: insights into the natural susceptibility to drug resistance development. *J. Mol. Biol.* **369**, 1029-1040.

Schake, D. (2004) How flexible is tripanavir in complex with the HIV-1 protease active site? *Aids* **18**, 579-580.

Schake, D. (2005) Molecular significance of tipranavir related codon 33 protease gene mutations. *Aids* **19**, 218-219.

Schatz, H., Gelderbloom, H. R., Nitschko, N. and von der Helm, K. (1991) Analysis of non-infectious HIV-1 particles in the presence of HIV proteinase inhibitor. *Arch. Virol.* **120**, 71-81.

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

Schock, H. B., Garsky, V. M. and Kuo, L. C. (1996) Mutational anatomy of an HIV-1 protease variant conferring cross-resistance to protease inhibitors in clinical trials. Compensatory modulations of binding and activity. *J. Biol. Chem.* **271**, 31957-31962.

Seelmeier, S., Schmidt, H., Turk, V. and von der Helm, K. (1988) Human immunodeficiency virus has an aspartic-type protease that can be inhibited by pepstatin A. *Proc. Natl. Acad. Sci. U. S. A.* **85**, 6612-6616.

Shafer, R. W., Stevenson, D. and Chan, B. (1999) Human immunodeficiency virus reverse transcriptase and protease sequence database. *Nucl. Acids Res.* **27**, 348-352.

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

Sham, H. L., Kempf, D. J., Molla, A., Marsh, K. C., Kumar, G. N., Chen, C. M., Kati, W., Stewart, K., Lal, R., Hsu, A., Betebenner, D., Korneyeva, S., McDonald, E., Saldivar, A., Wideburg, N., Chen, X., Niu, P., Park, C., Jayanti, V., Grabowski, B., Granneman, G. R., Sun, E., Japour, A. J., Leonard, J. M., Platter, J. J. and Norbeck, D. W. (1998) ABT-378, a highly potent inhibitor of the human immunodeficiency virus protease. *Antimicrob. Agents Chemother.* **42**, 3218-3224.

Shirley, B. A. (1995) Urea and guanidine hydrochloride denaturation curves. *Methods Mol. Biol.* **40**, 177-190.

Sigrell, J. A., Cameron, A. D., and Mowbray, S. L. (1999) Induced fit on sugar binding activates ribokinase. *J. Mol. Biol.* **290**, 1009-1018.

Sigurskjold, B. W. (2000) Exact analysis of competition ligand binding by displacement isothermal titration calorimetry. *Anal. Biochem.* **277**, 260-266.

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

Sturtevant, J. M. (1977) Heat capacity and entropy changes in processes involving proteins. *Proc. Natl. Acad. Sci. U. S. A.* **74**, 2236-2240.

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

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

Tie, Y., Boross, P. I., Wang, Y. –F., Gaddis, L., Liu, F., Chen, X., Tozser, J., Harrison, R. W. and Weber, I. T. (2005) Molecular basis for substrate recognition and

drug resistance from 1.1 to 1.6 angstroms resolution crystal structures of HIV-1 protease mutants with substrate analogs. *FEBS J.* **272**, 5265-5277.

Tie, Y., Kovalevsky, A. Y., Boross, P. I., Wang, Y. –F., Ghosh, A. K., Tozser, J., Harrison, R. W. and Weber, I. T. (2007) Atomic resolution crystal structures of HIV-1 protease and mutants V82A and I84V with saquinavir. *Proteins* **67**, 232-242.

Todd, M. J. and Freire, E. (1999) The effect of inhibitor binding on the structural stability and cooperativity of the HIV-1 protease. *Proteins* **36**, 147-156.

Todd, M. J., Luque, I., Velazquez-Campoy, A. and Freire, E. (2000) Thermodynamic basis of resistance to HIV-1 protease inhibition: calorimetric analysis of the V82F/I84V active site resistant mutant. *Biochemistry* **39**, 11876-11883.

Toh, O., Ono, M. and Miyata, T. (1985) Retroviral gag and DNA endonuclease coding sequences in IgE-binding factor gene. *Nature* **318**, 388-389.

Tomasselli, A. G. and Heinrikson, R. L. (2000) Targeting the HIV-protease in AIDS therapy: a current clinical perspective. *Biochimica et Biophysica Acta.* **1477**, 189-214.

Tong, L., Pav, S., Pargellis, C., Do, F., Lamarre, D. and Anderson, P. C. (1993) Crystal structure of human immunodeficiency virus (HIV) type 2 protease in complex with a reduced amide inhibitor and comparison with HIV-1 protease structures. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 8387-8391.

Tözsér, J., Gustchina, A., Weber, I. T., Blaha, I, Wondrak, E. M. and Oroszlan, S. (1991) Studies on the role of the S4 substrate binding site of HIV proteinases. *FEBS Lett.* **279**, 356-360.

Tözsér, J., Weber, I. T., Gustchina, A., Blaha, I, Copeland, T. D., Louis, J. M., Oroszlan, S. (1992) Kinetic and modeling studies of S3-S3` subsites of HIV proteinases. *Biochemistry* **31**, 4793-4800.

Tsai, C. J., Xu, D. and Nussinov, R. (1997) Structural motifs at protein-protein interfaces: protein cores versus two-state and three-state model complexes. *Protein Sci.* **6**, 1793-1805.

Turner, S. R., Strohbach, J. W., Tommasi, R. A., Aristoff, P. A., Johnson, P. D., Skulnick, H. I., Dolak, L. A., Seest, E. P., Tomich, P. K., Bohanon, M. J., Horng, M., Lynn, J. C., Chong, K., Hinshaw, R. R., Watenpaugh, K. D., Janakiraman, M. N. and Thaisrivongs, S. (1998) Tipranavir (PNU-140690): a potent, orally bioavailable nonpeptidic HIV protease inhibitor of the 5,6-dihydro-4-hydroxy-2-pyrone sulfonamide class. *J. Med. Chem.* **41**, 3467-3476.

Turriziani, O., Antonelli, G., Jacobsen, H., Mous, J., Riva, E., Pistello, M. and Dianzani, F. (1994) Identification of an amino acid substitution involved in the reduction of sensitivity of HIV-1 to an inhibitor of viral proteinase. *Acta Virol.* **38**, 297-298.

Ullrich, B., Laberge, M., Tölgyesi, F., Szeltner, Z., Polgár, L. and Fidy, J. (2000) Trp42 rotamers report reduced flexibility when the inhibitor acetyl-pepstatin is bound to HIV-1 protease. *Protein Sci.* **9**, 2232-2245.

Umezawa, H., Aoyagi, T., Morishima, H., Matsuzaki, M. and Hamada, M. (1970) Pepstatin, a new pepsin inhibitor produced by Actinomycetes. *J. Antibiot.* (Tokyo) 1970 **23**, 259-262.

Vacca, J. P., Dorsey, B. D., Schleif, W. A., Levin, R. B., McDaniel, S. L., Darke, P. L., Zugay, J., Quintero, J. C., Blahy, O. M., Roth, E., Sardana, V. V., Schlabach, A. J., Graham, P. I., Condra, J. H., Gotlib, L., Holloway, M. K., Lin, J., Chen, I. W., Vastag, K. V., Oatovich, D., Anderson, P. S., Emini, E. A. and Huff, J. R. (1994) L-735,524, an orally bioavailable human immunodeficiency virus type 1 protease inhibitor. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 4096-4100.

Vacca, J. P. and Condra, J. H. (1997) Clinically effective HIV-1 protease inhibitors. *Drug Discov. Today* **2**, 261-272.

Vega, S., Kang, L., Velazquez-Campoy, A. Kiso, Y., Amzel, L. M. and Freire, E. (2004) A structural and thermodynamic escape mechanism from a drug resistant mutation of the HIV-1 protease. *Proteins* **55**, 594-602.

Velazquez-Campoy, A., Luque, I., Todd, M. J., Milutinovich, M., Kiso, Y. and Freire, E. (2000) Thermodynamic dissection of the binding energetics of KNI-272, a powerful HIV-1 protease inhibitor. *Protein Sci.* **9**, 1801-1809.

Velazquez-Campoy, A. and Freire, E. (2001) Incorporating target heterogeneity in drug design. *J. Cell Biochem. Suppl*. **37**, 82-88.

Velazquez-Campoy, A., Todd, M. J., Vega, S. and Freire, E. (2001a) Catalytic efficiency and vitality of HIV-1 proteases from African viral subtypes. *Proc. Natl. Acad. Sci. U. S. A.* **98,** 6062-6067.

Velazquez-Campoy, A., Kiso, Y. and Freire, E. (2001b) The binding energetics of first- and second-generation HIV-1 protease inhibitors: Implications fro drug design. *Arch. Biochem. Biophys.* **390**, 169-175.

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

Velazquez-Campoy, A., Muzammil, S., Ohtaka, H., Schon, A., Vega, S. and Freire E. (2003a) Structural and thermodynamic basis of resistance to HIV-1 protease inhibition: implications for inhibitor design. *Curr. Drug Targets Infect. Disord.* **3**, 311-328.

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

Velazquez-Campoy, A., Leavitt, S. A and Freire E. (2004) Characterization of protein-protein interactions by isothermal titration calorimetry. *Methods Mol. Biol.* **261**, 35-54.

Velazquez-Campoy, A. and Freire E. (2005) ITC in the post-genomic era…? Priceless. *Biophys. Chem.* **115**, 115-124.

Vergne, L. Peeters, M., Mpoudi-Ngole, E., Bourgeois, A., Liegeois, F. J., Reynes, J. and Delaporte, E. (2000) Genetic diversity of protease and reverse transcriptase sequences in non-subtype-B human immunodeficiency virus type 1 strains: evidence of many minor drug resistant mutations in treatment-naïve patients. *J. Clin. Microbiol.* **38**, 3919-3925.

Vondrasek, J., van Buskirk, C. P., Wlodawer, A. (1997) Database of three-dimensional structures of HIV proteinases. *Nat. Struct. Biol.* **4**, 8.

Vondrasek, J. and Wlodawer, A. (2002) HIVdb: a database of the structures of human immunodeficiency virus protease. *Proteins* **49**, 429-431.

Wedenmayer, G. J., Patten, P. A., Wang, L. H., Schultz, P. G. and Stevens, R. C. (1997) Structural insights into the evolution of an antibody combining site. *Science* **276**, 1665-1669.

Weiss, A., Hollander, H. and Stobo, J. (1985) Acquired immunofefeciency syndrome: epidemiology, virology, and immunology. *Annu. Rev. Med.* **36**, 545-562.

White, D. O. and Fenner, F. J. (1994) *Medical Virology.* Fourth Ed. Academic Press, Sydney.

Winder, A. F. and Gent, W. L. (1971) Correction of light-scattering errors in spectrophotometric protein determinations. *Biopolymers* **10**, 1243-1251.

Winslow, D. L. and Otto, M. J. (1995) HIV protease inhibitors. *AIDS* **9**, S183-S192.

Wiseman, T., Williston, S., Brandts, J. F. and Lin, L. N. (1989) Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal. Biochem.* **179**, 131-37.

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

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

Wlodawer, A., Gustchina, A., Reshetnikova, L., Lubkowski, J., Zdanov, A., Hui, K. Y., Angleton, E. L., Farmerie, W. G., Goodenow, M. M. and Bhatt, D. (1995) Structure of an inhibitor complex of the proteinase from feline immunodeficiency virus. *Nat. Struct. Biol.* **6,** 480-488.

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

Wlodawer, A. (1999) ABT-378 Abbott Laboratories. *Curr. Opin. Anti-infect. Investig.* **1**, 246-250.

Woody, R. W. (1995) Circular dichroism. *Methods Enzymol.* **246**, 34-71.

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

Xu, D., Lin, S. L., and Nussinov, R. (1997) Protein binding versus protein folding: the role of hydrophilic bridges in protein associations. *J. Mol. Biol.* **265**, 68-84.

Zhang, Z. Y., Poorman, R. A., Maggiora, L. L., Heinrikson, R. L. and Kezdy, F. J. (1991) Dissociative inhibition of dimeric enzymes. Kinetic characterization of the inhibition of HIV-1 protease by its COOH-terminal tetrapeptide. *J. Biol. Chem.* **266**, 15591-15594.

Zhao, B., Winborne, E., Minnich, M. D., Culp, J. S., Debouck, C.and Abdel-Meguid, S. S. (1993) Three-dimensional structure of a simian immunodeficiency virus protease/inhibitor complex: Implications for the design of human immunodeficiency virus type 1 and 2 protease inhibitors. *Biochemistry* **32**, 13054-13060.

Appendix

Resistance in HIV-1 protease subtypes

Substrate cleft mutations

- V82A/T/F/S occur predominantly in HIV-1 isolates from patients receiving treatment with indinavir and ritonavir
- V82A also occurs in isolates from patients receiving prolonged therapy with saquinavir following the development of the mutation G48V.
- By themselves, mutations at codon 82 confer decreased *in vitro* susceptibility to indinavir, ritonavir, and lopinavir but not to nelfinavir, saquinavir, or amprenavir. However, when present with other PI mutations, V82A/T/F/S contributes phenotypic and clinical resistance to each of the PIs.
- V82A is the most common mutation at this position; V82S, the least common. The phenotypic and clinical significance of the differences between each of these mutations has not been studied.
- V82I occurs in about 1% of untreated individuals with subtype B HIV-1 and in 5-10% of untreated individuals with non-B isolates. Preliminary data suggest that V82I confers minimal or no resistance to the available PIs.

- I84V has been reported in patients receiving indinavir, ritonavir, saquinavir, and amprenavir and causes phenotypic and clinical resistance to each PI.
- I84V tends to develop in isolates that already have the mutation L90M and is rarely the first major mutation to develop in patients receiving a PI.

- G48V occurs primarily in patients receiving saquinavir and rarely in patients receiving indinavir.
- This mutation causes 10-fold resistance to saquinavir and about 3-fold resistance to indinavir, ritonavir, and nelfinavir.
- Its affect on amprenavir and lopinavir is not known. G48V often occurs with mutations at positions 54 and 82; and isolates with this set of three mutations are phenotypically resistant to each of the PIs .

- D30N occurs solely in patients receiving nelfinavir and confers no *in vitro* or clinical cross-resistance to the other PIs.
- Cross-resistance to indinavir, ritonavir, and saquinavir has been observed in isolates that have D30N along with mutations at positions 88 and 90.

- I50V has been reported only in patients receiving amprenavir as their first PI.
- In addition to causing reduced amprenavir susceptibility, it has been shown to increase K_i values to ritonavir, indinavir, and nelfinavir in biochemical studies and to cause *in vitro* cross-resistance to ritonavir and lopinavir.
- Possibly because of the rarity of this mutation, there have been few reports of multidrug- resistant isolates containing this mutation.

- V32I occurs in patients receiving indinavir, ritonavir, and amprenavir.

- It usually occurs only in association with other PI resistance mutations in the substrate cleft or flap and by itself appears to cause minimal resistance to any one drug.
- R8K and R8Q are substrate cleft mutations that cause high-level resistance to one of the precursors of ritonavir (A-77003) but they have not been reported with the current PIs.

Protease flap mutations

- The protease flaps (residues 33-62) extends over the substrate-binding cleft and must be flexible to allow entry and exit of the polypeptide substrates and products. The tips of the flap tips residues 46-54 are particularly mobile and contain many drug resistance mutations.
- In addition to G48V and I50V, which are also in the substrate cleft, mutations at positions 46, 47, 53, and 54 make important contributions to drug resistance.
- Mutations at position 54 (generally I54V, less commonly I54T/L/M) contribute resistance to each of the six approved PIs and have been commonly reported during therapy with indinavir, ritonavir, amprenavir, saquinavir, and lopinavir.
- I54L and I54M are particularly common in persons receiving amprenavir and have a greater effect on amprenavir than I54V.
- Mutations at position 46 contribute to resistance to each of the PIs except saquinavir and have been commonly reported during therapy with indinavir, ritonavir, amprenavir, and nelfinavir. M46I mutation has been shown to stabilize the flaps in a closed conformation and thus affect the binding of the PIs in the active site.
- Mutations at codon 47 have been reported in patients receiving amprenavir, indinavir, and ritonavir, and often occur in conjunction with the nearby substrate cleft mutation, V32I.
- F53L has been reported rarely in patients receiving PI monotherapy, but it occurs in more than 10% of patients treated with multiple PIs. In a multivariate analysis F53L has been associated with phenotypic resistance to lopinavir.

Protease mutations at other conserved residues

- L90M has been reported in isolates from patients treated with saquinavir, nelfinavir, indinavir, and ritonavir.
- L90M either contributes to or directly confers in vitro resistance to each of the six approved PIs and plays a role in causing clinical cross-resistance to each of the PIs.
- Crystal structures with and without the mutant have shown that the Leu90 side chain lies next to Leu24 and Thr26 on either side of the catalytic Asp25 but the mechanism by which L90M causes PI resistance is not known.

- Mutations at codon 73, particularly G73S, have been reported in 10% of patients receiving indinavir and saquinavir monotherapy and occasionally

during nelfinavir monotherapy. However, this mutation occurs most commonly in patients failing multiple PIs, usually in conjunction with L90M.

- Mutations at position 88 (N88D and N88S) commonly occur in patients receiving nelfinavir and occasionally in patients receiving indinavir. By itself, a mutation at this position causes low-level nelfinavir resistance. However, a mutation at this position causes high-level nelfinavir resistance in the presence of D30N or M46I.
- N88S (but not N88D) has been shown to hypersensitize isolates to amprenavir but the clinical significance of this finding is not known.
- L24I has been reported only in HIV-1 isolates from patients receiving indinavir and has not been shown to confer cross-resistance to other PIs, except possibly lopinavir.

Polymorphic sites contributing to resistance

- Amino acid variants at seven polymorphic positions, including codons 10, 20, 36, 63, 71, 77, and 93 also make frequent contributions to drug resistance. These mutations do not cause drug resistance by themselves. Some contribute to drug resistance when present together with other protease mutations; others compensate for the decrease in catalytic efficiency caused by other mutations.
- Mutations at codons 10, 20, 36, and 71 occur in up to 5 to 10% of untreated persons infected with subtype B viruses. However, in heavily treated patients harboring isolates with multiple mutations in the substrate cleft, flap, or at codon 90, the prevalence of mutations at these positions increases dramatically. Mutations at codon 10 and 71 increase to 60 to 80%, whereas mutations at codons 20 and 36 increase to 30 to 40%.
- Codon 63 is the most polymorphic protease position. In untreated persons about 45% of isolates have 63L (considered the subtype B consensus), about 45% have 63P, and about 10% have other residues at this position. However, the prevalence of amino acids other than L increases to 90% in heavily treated patients.
- Mutations at codons 77 and 93 double in prevalence from 15 to 20% in untreated persons to 30 to 40% in heavily treated persons.
- In some HIV-1 subtypes, mutations at codons 10, 20, and 36 occur at higher rates than they do in subtype B isolates. It has been hypothesized that individuals harboring isolates containing multiple accessory mutations may be at a greater risk of virologic failure during PI therapy. However, most studies to date have not supported this hypothesis.