PROTEINS: Structure, Function, and Genetics 36:147–156 (1999)

RESEARCH ARTICLES

The Effect of Inhibitor Binding on the Structural Stability and Cooperativity of the HIV-1 Protease

Matthew J. Todd and Ernesto Freire*
Department of Biology and Biocalorimetry Center, The Johns Hopkins University, Baltimore, Maryland

ABSTRACT **The effects of the peptide inhibitor acetyl pepstatin on the structural stability of the HIV-1 protease have been measured by high sensitivity calorimetric techniques. At 25°C and pH 3.6, acetyl pepstatin binds to HIV-1 protease with an affinity of 1.6×10^7 M^{-1} and an enthalpy of 7.3 ± 0.5 kcal/mol, indicating that binding is not favored enthalpically and that the favorable Gibbs energy originates from a large positive entropy. Since the binding of acetyl pepstatin is associated with a negative change in heat capacity (−450 cal/K•mol) the association reaction becomes enthalpically favored at temperatures higher than 40°C. The presence of the inhibitor stabilizes the dimeric structure of the protease in a fashion that can be quantitatively described by a set of thermodynamic linkage equations. The combination of titration and differential scanning calorimetry provides an accurate way of determining binding constants for high affinity inhibitors that cannot be determined by titration calorimetry alone. A structure-based thermodynamic analysis of the binding process indicates that the stabilization effect is not distributed uniformly throughout the protease molecule. The binding of the inhibitor selectively stabilizes those conformational states in which the binding site is formed, triggering a redistribution of the state probabilities in the ensemble of conformations populated under native conditions. As a result, the stability constants for individual residues do not exhibit the same change in magnitude upon inhibitor binding. Residues in certain areas of the protein are affected significantly whereas residues in other areas are not affected at all. In particular, inhibitor binding has a significant effect on those regions that define the binding site, especially the flap region which becomes structurally stable as a result of the additional binding free energy. The induced stabilization propagates to regions not in direct contact with the inhibitor, particularly to the strand between residues Pro9 and Ala22 and the helix between Arg87 and Gly94. On the other hand, the stability of the strand between Asp60 and Leu76 is not significantly affected by inhibitor binding. The structural distribution of binding effects define cooperative pathways within the protease molecule. Proteins 1999;36:147–156.** © 1999 Wiley-Liss, Inc.

Key words: HIV; protease; cooperative interactions; calorimetry

INTRODUCTION

The HIV-1 protease has been the most important target in the development of antiviral therapies against HIV-1 infection. In previous papers, we have investigated the binding thermodynamics of synthetic HIV-1 protease inhibitors for which high resolution structures are available;[1] the differences in the binding characteristics between substrates and inhibitors;[2] and the energetics of stabilization of the native dimeric structure of the protease molecule.[3] Structural stability and inhibitor or substrate binding, however, are not independent events but are processes coupled to one another by a set of well defined thermodynamic linkage equations. In this paper we present the results of a study aimed at evaluating the effects of inhibitor binding on the structural stability and the distribution of cooperative interactions within the protease molecule.

Under native conditions, the HIV-1 protease, as any other protein, cannot be considered as a single conformation but rather as a statistical ensemble of conformations populated according to their Gibbs energies. In particular, it has been shown experimentally and computationally that in the protease molecule the region known as the "flap" has little structural stability. In the native state this region experiences local folding/unfolding reactions that give rise to a number of conformational states.[1,3,4] The existence of multiple conformational states created by local unfolding events has been demonstrated for other proteins using nuclear magnetic resonance (NMR)-detected hydrogen/deuterium exchange experiments under native conditions. Those experiments have shown that

Grant sponsor: National Institutes of Health; Grant numbers: GM 51362 and GM 57144.
*Correspondence to: Ernesto Freire, Department of Biology and Biocalorimetry Center, The Johns Hopkins University, Baltimore, MD 21218. E-mail: ef@jhu.edu
Received 11 January 1999; Accepted 16 March 1999

under native conditions the majority of amide groups become exposed to solvent as a result of many independent local unfolding reactions rather than global cooperative events.[5–15] Under equilibrium conditions, the probability of any conformation created by independent local unfolding events is dictated by its Gibbs energy. Ligands, including H^+ and other ions, play an important role in modulating the distribution of states characterizing the protein under a given set of conditions by stabilizing or destabilizing certain conformations. Since different states might have different functional properties or different biological activities, ligands play the role of molecular switches by preferentially stabilizing certain conformations and redistributing the conformational ensemble.

The above observations have questioned the traditional view in which proteins behave as all-or-none cooperative entities. At the same time, these experiments have brought to light new fundamental questions. If the native state is a structure in which multiple regions are able to undergo small, independent local unfolding events, then cooperativity is regional rather than global. If cooperative interactions are regional, how can the effects of local perturbations be propagated from one region to another? How can the binding of a ligand induce changes in regions far away from the binding site? If cooperative interactions do not extend uniformly throughout the entire protein molecule, then some residues may have a more important role than others in the development of cooperative responses. In fact, molecular communication must occur through *cooperative pathways* that involve only a subset of residues within the protein molecule. The identification of those residues has been an important topic of investigation in this laboratory for the last few years.

In a recent publication,[16] we introduced a computational technique (single site thermodynamic mutation) by which the structural distribution of the effects due to thermodynamic perturbations originating at different locations in the protein can be evaluated. In this paper a new algorithm aimed at evaluating the effects of ligand binding on the structural stability and the structural distribution of cooperative interactions within a protein is presented. This algorithm (Core_Bind) has been applied to the inhibitor acetyl pepstatin binding to the HIV-1 protease. The energetics of acetyl pepstatin binding and its effect on the structural stability of the protease are also determined experimentally. The agreement between the computational and experimental analyses supports the statistical thermodynamic approach for identifying cooperative pathways in the protein molecule.

EXPERIMENTAL SECTION

Protease Purification

HIV-1 protease was prepared as described.[3] Briefly, plasmid-encoded mutant protease (Q7K/L33I/L63I) (a gift from Dr. A.G. Tomasselli), designed to remove three hypersensitve autolytic sites[40] was expressed as inclusion bodies in Escherichia coli 1458. Protease was purified from washed inclusion bodies using two ion exchange chromatographic steps, as described previously.[3] Following these two columns, protease was estimated to be $>$ 90% pure. Minor contaminants precipitated during folding.

The HIV-1 protease was routinely folded (at concentrations up to 0.5 mg/ml) by dilution into 10 mM Formic acid at 0°C followed by an increase in pH to 3.8. Folded protease was desalted into 1 mM Na acetate, pH 5.0 and stored at either 4°C (no loss of activity in one month) or −20°C (at $\geq$2.5 mg/mL).

Spectrophotometric Enzymatic Assays

The specific activity of the HIV-1 protease preparations was measured by following the hydrolysis of the chromogenic substrate Lys-Ala-Arg-Val-Nle-nPhe-Glu-Ala-Nle-NH_2 (California Peptide Research, Inc., Napa, CA). Protease was added to a 120µL micro cuvette containing substrate at 25°C giving a final concentration of 100µM substrate, 100 mM NaOAc, 1 M NaCl, pH 5.0 The absorbance was monitored at 6 wavelengths (296–304 nm) using an HP 8452A diode array spectrophotometer (Hewlett Packard) and corrected for spectrophotometer drift by subtracting the average absorbance at 446–454 nm. 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 portion of the data, where [S] = 80% [So]. Typical protease preparations hydrolyzed chromogenic substrate at 4– 6 s^{-1} (per dimer) at 25°C.

Differential Scanning Calorimetry (DSC)

The heat capacity function of the HIV-1 protease was measured as a function of temperature with a high precision differential scanning calorimeter.[37] Protein samples and reference solutions were properly degassed and carefully loaded into the cells to avoid bubble formation. Exhaustive cleaning of the cells was undertaken before each experiment. Thermal denaturation scans were performed with freshly prepared protease that had been gel filtered into 10 mM glycine, pH 3.6 (PD-10, Pharmacia, Uppsala, Sweden). HIV-1 protease (at 0.3 mg/mL) and different concentrations of acetyl pepstatin (Peptides International, Louisville, KY) was examined for thermal denaturation at a scanning rate of 1°C/min, from 25°C to 85°C. Two heating scans were taken of each protease sample to confirm the reversibility of the transition. For all experiments, protease specific activity was determined in unheated samples and compared to samples that underwent thermal cycling. Reversibility for a single cycle was $>$80% under all conditions.

Isothermal Titration Calorimetry

Isothermal titration calorimetry experiments were carried out using a MCS titration calorimeter system (Microcal Inc., Northhampton, MA). In a typical calorimetric experiment, the enzyme solution (5–20µM dimer) in the calorimetric cell was titrated with acetyl pepstatin (100–1000µM in the same buffer). The heat evolved after each

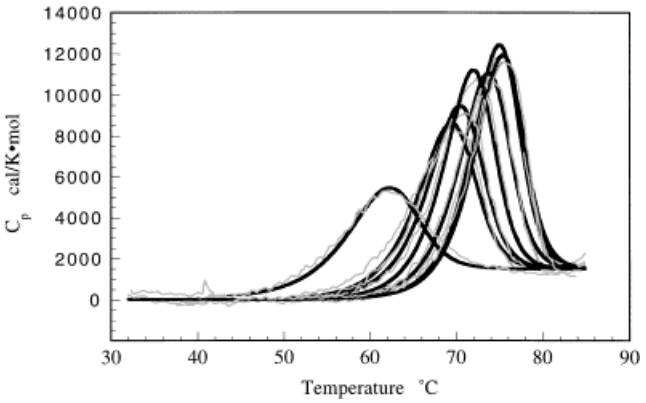

Fig. 1. The dependence of the structural stability of the HIV-1 protease on the concentration of the inhibitor acetyl pepstatin. Shown in the figure is the heat capacity function measured at a protease concentration of 0.3 mg/mL in 10 mM glycine pH 3.6. The transition temperature increases as a function of acetyl pepstatin concentration. In these experiments, acetyl pepstatin concentrations were 0, 15, 25, 50, 100, 150, and 200μM respectively. In each case, the solid line corresponds to the best fit obtained with the two-state folding/dimerization model.

ligand 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 reaction and the corresponding heat of dilution. Analysis of the data was performed using software developed in this laboratory as described previously.[38] Protease specific activity was determined before and after titration. No significant loss of activity was observed.

RESULTS AND DISCUSSION

Differential Scanning Calorimetry (DSC)

Figure 1 shows a series of calorimetric scans of the HIV-1 protease (pH 3.6) at increasing concentrations of the peptidic inhibitor acetyl pepstatin. It is clear from these experiments that the inhibitor stabilizes the native structure of the protease molecule as indicated by higher denaturation temperatures with increasing inhibitor concentrations. The stabilization of the HIV-1 protease is consistent with the preferential binding of the inhibitor to the native structure of the protease molecule. In a previous publication,[3] we demonstrated that the thermodynamic stability of the protease molecule can be accounted for by a model in which the dimeric native structure is in equilibrium with unfolded monomers with no detectable dimeric or monomeric intermediates (solid lines in Fig. 1):

$$N_2 \overset{K_d}{\rightleftarrows} 2U \tag{1}$$

According to this model, the fraction of molecules that are in the unfolded state, F_u, is given by the equation.[17,18]

$$F_u = \frac{A \bullet [(A^2 + 4)^{1/2} - A]}{2} \tag{2}$$

where A is a function of the equilibrium constant, K_d, and the total protease concentration, $[P_T]$, expressed in monomer units:

$$A = \left[\frac{K_d}{2[P_T]}\right]^{1/2} = \left[\frac{e^{-\Delta G/RT}}{2[P_T]}\right]^{1/2} \tag{3}$$

In general, the Gibbs energy of stabilization is a function of temperature and solvent conditions. In our previous paper[3] we experimentally obtained the following master equation for the Gibbs energy as a function of temperature, pH and urea concentration:

$$\begin{aligned}\Delta G^0 \equiv \Delta G(T, pH, urea) = \\ -9{,}000 + 3{,}200 \bullet (T - 298.15) \\ - T \bullet (-79.15 + 3{,}200 \bullet \ln (T/298.15)) \\ - R \bullet T \bullet 4 \ln \left(\frac{(1 + 10^{4.3-pH})}{(1 + 10^{2.9-pH})}\right) - 2{,}100 \bullet [urea]\end{aligned} \tag{4}$$

where the reference state is the native state at pH 5.5 and 25°C. The protein concentration dependence of the equilibrium is given by equations 2 and 3. In the presence of the inhibitor, the Gibbs energy will also be affected in a manner that can be accounted for by a set of linkage equations (see for example Wyman and Gill[19]). In general, if the inhibitor, X, does not bind to the unfolded state, the Gibbs energy of stabilization will be modified according to the following equation:

$$\Delta G = \Delta G^0 + RT \ln (1 + K_a[X]) \tag{5}$$

where K_a is the association constant of the inhibitor to the protease molecule and ΔG^0 the Gibbs energy of stabilization in the absence of inhibitor (ΔG^0 is given by equation 4). Since K_a is temperature dependent, the application of equation 5 requires knowledge of ΔH, ΔC_p and ΔG for inhibitor binding at the reference temperature. It must be noted also that [X] is the free and not the total concentration of inhibitor. The free concentration of inhibitor is equal to the difference between the total concentration, $[X_T]$, and the bound concentration, $[X_B]$, of inhibitor ($[X] = [X_T] - [X_B]$) and is given by the following equation:

$$\begin{aligned}[X] = \frac{\sqrt{(K_a[P_N] - K_a[X_T] + 1)^2 + 4K_a[X_T]}}{2K_a} \\ \frac{- (K_a[P_N] - K_a[X_T] + 1)}{2K_a}\end{aligned} \tag{6}$$

where $[P_N]$ is the molar concentration of protease molecules in the dimeric native state ($[P_N] = 0.5(1 - F_U)[P_T]$). In general, the free and total ligand concentrations differ significantly when the inhibitor and protease concentrations are of the same magnitude. Under those conditions, using the total ligand will yield erroneous results not only with respect to the transition temperature but also with

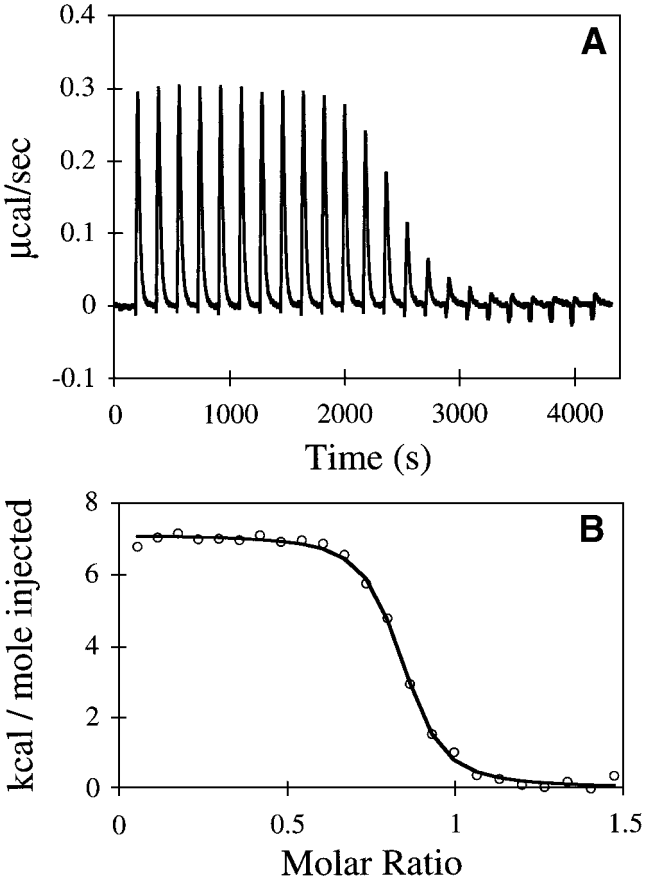

Fig. 2. Calorimetric titration of HIV-1 protease with the peptide acetyl pepstatin. **A** shows the heat effects associated with the injection of acetyl pepstatin (10µL per injection of a 280 µM solution) into the calorimetric cell (1.4 mL) containing HIV-1 protease at a concentration of 20µM. The experiment was performed at 25°C. **B** shows the binding isotherm corresponding to the data in A and the best fitted curve.

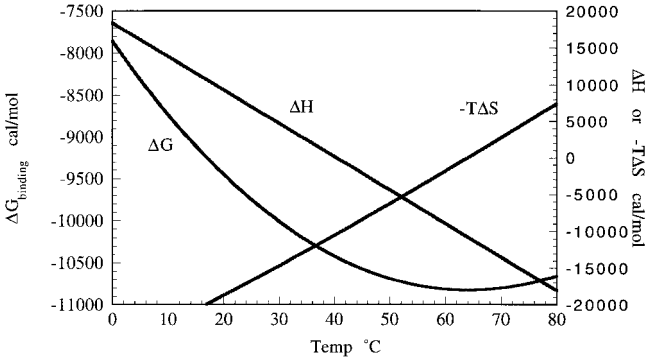

Fig. 3. Temperature dependence of the thermodynamic parameters for acetyl pepstatin binding to the HIV-1 protease. Note that at temperatures lower than 40°C, the Gibbs energy of binding is not favored enthalpically and that at higher temperatures the situation is reversed. At physiological temperatures and below the binding affinity increases with temperature.

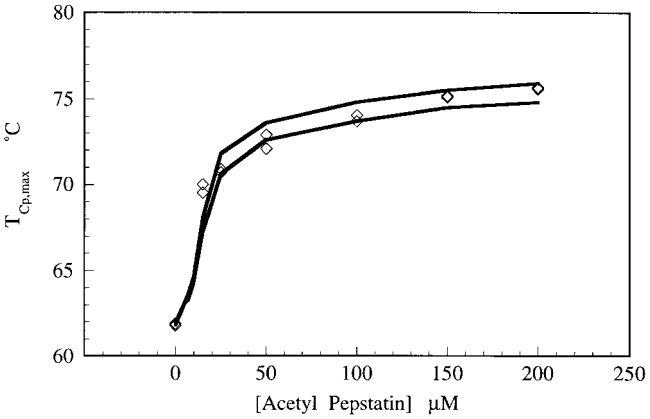

Fig. 4. The dependence of the denaturation temperature of the protease molecule on acetyl pepstatin concentration. The solid lines were calculated with the master equation for the Gibbs energy for the structural stability of the HIV-1 protease (equation 4) and the linkage equation (equations 5, 6) with the thermodynamic values indicated in the text.

respect to the shape of the transition curve itself.[20] Also, for a dimeric protein, like the HIV-1 protease, the fraction of molecules in the native state may be significantly less than one at low protein concentrations (see Todd et al.[3]).

Isothermal Titration Calorimetry (ITC)

The binding thermodynamics of acetyl pepstatin to the HIV-1 protease was measured directly by titration calorimetry under identical buffer conditions as those used in the scanning calorimetric experiments (Fig. 2). These experiments yield a binding constant of $1.6 \times 10^7\ M^{-1}$ ($\Delta G(25) = -9.8$ kcal/mol) and a binding enthalpy of 7.3 ± 0.5 kcal/mol, indicating that the binding of acetyl pepstatin is not favored enthalpically at 25°C and that the favorable Gibbs energy of binding originates from a large positive entropy (56.5 cal/K●mol). We have previously determined that the binding of the inhibitor is associated with a negative change in heat capacity,[2] ΔC_p, of -452 ± 50 cal/K●mol. Figure 3 shows the temperature dependence of the thermodynamic parameters of binding. This figure illustrates the characteristic curvature associated with processes involving a heat capacity change. In this case, the enthalpy change is positive at low temperatures but progressively diminishes until it changes sign at about 40°C. The entropy contribution, $-T\Delta S$, is favorable at low temperatures but increases with temperature and eventually changes sign at 64°C. At that temperature the Gibbs energy of binding reaches a minimum. The maximum in the binding affinity constant occurs when $\Delta H = 0$ and corresponds to 40.5°C.

Determination of Binding Constants From DSC Data

As described by equation 5, in the presence of acetyl pepstatin the stability of the protease is a function of the binding constant, K_a, suggesting that the change in protease stability can be used to evaluate the binding affinity of the inhibitor. Figure 4 shows the dependence of the denaturation temperature of the HIV-1 protease on the concentration of acetyl pepstatin. In this figure the temperature at which the heat capacity function is maximal

($T_{Cp,max}$) is shown. For a monomeric folding/unfolding reaction, the midpoint of the transition is independent of protein concentration and is located at the temperature where $\Delta G = 0$. At that temperature half of the molecules are unfolded and the heat capacity function is very close to its maximum. For a transition involving dissociation this is not the case. For a dimer undergoing unfolding coupled to dissociation, like the HIV-1 protease, the temperature at which half the molecules are unfolded is a function of the protein concentration and occurs at the point in which $\Delta G = -RT\ln([P_T])$, where $[P_T]$ is the total protein concentration in monomer units. Similarly, the temperature of the maximum in the heat capacity does not coincide with the temperature at which half of the molecules are unfolded but to the temperature at which $F_U = (2 - 2^{1/2}) \approx 0.586$.[21] At this point $\Delta G = -RT(\ln(4(3 - 2\sqrt{2})/(\sqrt{2} - 1)) + \ln [P_T]) \approx -RT(0.5 + \ln [P_T])$. Since the Gibbs energy function in the absence of inhibitor (ΔG^0) is known from the DSC experiments, application of equation 5 permits calculation of the transition temperature as a function of inhibitor concentration. To be accurate this calculation requires knowledge of the enthalpy and heat capacity of binding since the binding is temperature dependent (see Fig. 3). It must be noted that the binding enthalpy and binding heat capacity changes can be determined in ITC experiments even for those cases in which the binding affinity is too strong to be measurable by this technique.[22] Using the enthalpy and heat capacity changes of binding, the observed dependence of the transition temperature on inhibitor concentration can be used to estimate the inhibitor binding constant at 25°C, or at other convenient reference temperature. The two limiting lines in Figure 4 are drawn using $K_a(25) = 2.4 \times 10^7$ and 5.4×10^7 M^{-1}. These values represent the range of values that are consistent with the experimental transition temperatures measured at different inhibitor concentrations and are in close agreement with the binding constant measured by isothermal titration calorimetry at that temperature (1.6×10^7 M^{-1}; Fig. 2).

Previously, Brandts and Lin[23] demonstrated that differential scanning calorimetry could be used to estimate binding constants for ultratight binding to monomeric proteins. The studies presented here show that a similar approach can be used to evaluate inhibitor binding affinities to a dimeric system like the HIV-1 protease. Since the binding of HIV-1 inhibitors are characterized by significant enthalpy and heat capacity changes, the differential scanning calorimetry measurements must be supplemented by isothermal titration calorimetry determinations of the enthalpy and heat capacity changes of binding.

Structure-Based Analysis of Binding Energetics

In a previous paper,[2] we presented the results of a structure-based thermodynamic analysis of acetyl pepstatin binding to the HIV-1 protease. Those studies employed the structural parameterization of the energetics developed in this laboratory.[24–29] According to that analysis, at pH 3.6 the expected Gibbs energy of binding is −10,400 cal/mol at 25°C which compares favorably with the experimental value of −9,750 cal/mol. In addition, the structural parameterization correctly predicts a positive enthalpy change for binding (10,500 cal/mol compared to the experimental value of 7,300 cal/mol) and a negative change in heat capacity (−427 cal/K●mol compared to an experimental value of −452 cal/K●mol). In agreement with the experimental results, the structure-based analysis indicates that the binding of acetyl pepstatin is an entropically driven process at 25°C.

The good agreement between experimental and calculated binding parameters indicates that the structural parameterization correctly captures the nature of the forces that determine binding. According to the structural parameterization, the following residues: Arg8, Asp25, Gly27, Ala28, Asp29, Asp30, Ile47, Gly48, Gly49, Ile50, Pro81, Val82, and Ile84 establish the most significant van der Waals contacts with the inhibitor and define the main attachment points for acetyl pepstatin. The backbone of other residues (notably Met46) also contribute even though their side chains are exposed to the solvent in both the free and bound states. Figure 5 shows the structure of the HIV-1 protease color-coded according to the relative residue contributions to the Gibbs energy of binding. It is clear that the residues that contribute the most to the binding energetics are located in three distinct regions of the protease molecule: 1) Arg8, Asp25, Gly27, Ala28, Asp29, and Asp30 define the bottom of the binding site and are located at the dimer interface; 2) Met46, Ile47, Gly48, Gly49, and Ile50 define the top of the binding site and are located in the flap region; and, 3) Pro81, Val82, and Ile84 define the sides of the binding site. These three regions define the critical structural determinants of the binding site.

Ligand Binding and Structural Cooperativity

The energy of stabilization of the protease molecule is not uniformly distributed throughout the protease structure. In particular, in the free enzyme the flaps are known to be structurally unstable as indicated by high temperature factors in the crytallographic structure.[4] The absence of folded monomers is an indication that the dimerization interface plays a critical role in the stabilization of the native structure.[3,30] Consistent with the experimental results, a structure-based thermodynamic analysis of the protease structure also indicates the critical role of the dimerization interface. Residues located at the carboxy and amino terminus (Cys 95, Thr 96, Leu 97, Asn 98 and Phe 99, and Pro 1, Ile 3, Leu 5) contribute significantly to the total dimerization energy. Residues at the base of the active site (Thr 26, Gly 27, and Asp 29) and the tip of the flap (Gly 49, Ile 50, Gly 51) are also important.

The global stability of the protease molecule is increased by inhibitor binding as described by equation 5. This stabilization originates at the binding site but extends to many other residues in the molecule. How does this

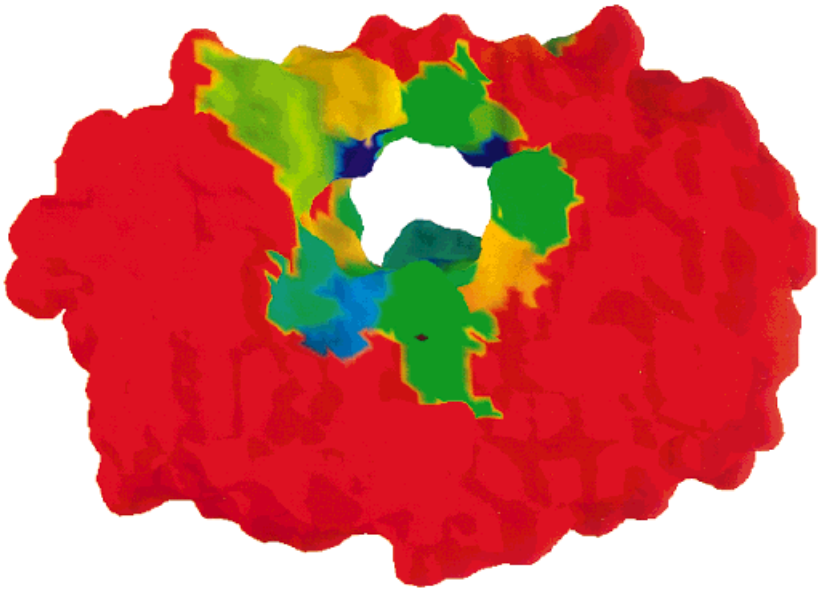

Fig. 5. The surface of the HIV-1 protease color-coded according to the magnitude of the residue contributions to the binding energy of acetyl pepstatin. In this figure blue identifies the residues that make the most significant contributions to binding and the areas colored red the regions with the least significant contributions. The residues that contribute the most to the binding energetics are located in three distinct regions of the protease molecule: 1) Arg8, Asp25, Gly27, Ala28, Asp29 and Asp30 define the bottom of the binding site and are located at the dimer interface; 2) Met46, Ile47, Gly48, Gly49 and Ile50 define the top of the binding site and are located in the flap region; and, 3) Pro81, Val82, and Ile84 define the sides of the binding. The figure was made using the program GRASP.[39]

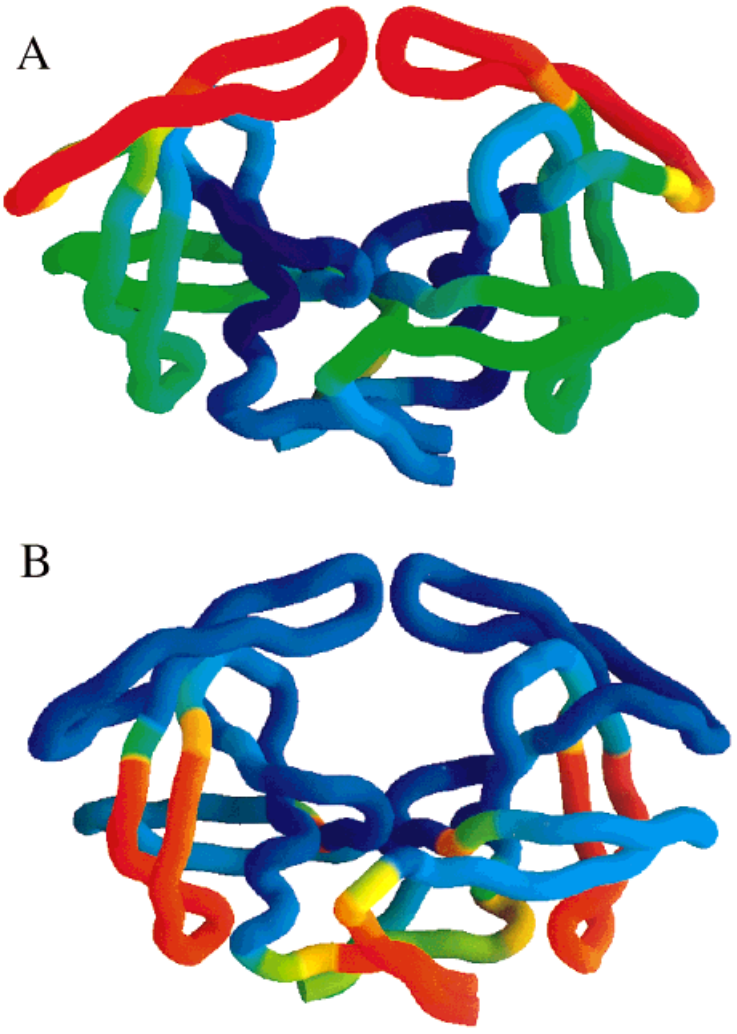

Fig. 6. (**A**) shows the structure of the HIV-1 protease color-coded according to the magnitude of the stability constants per residue. In this figure blue denotes the most stable regions of the protein and the areas colored red the least stable regions. The stability constants were calculated with the COREX algorithm.[31,34,35] (**B**) shows the calculated effects of acetyl pepstatin binding on the magnitude of the stability constants per residue. Blue denotes the residues most affected and red the residues least affected by binding. The figure was made using the program GRASP.[39]

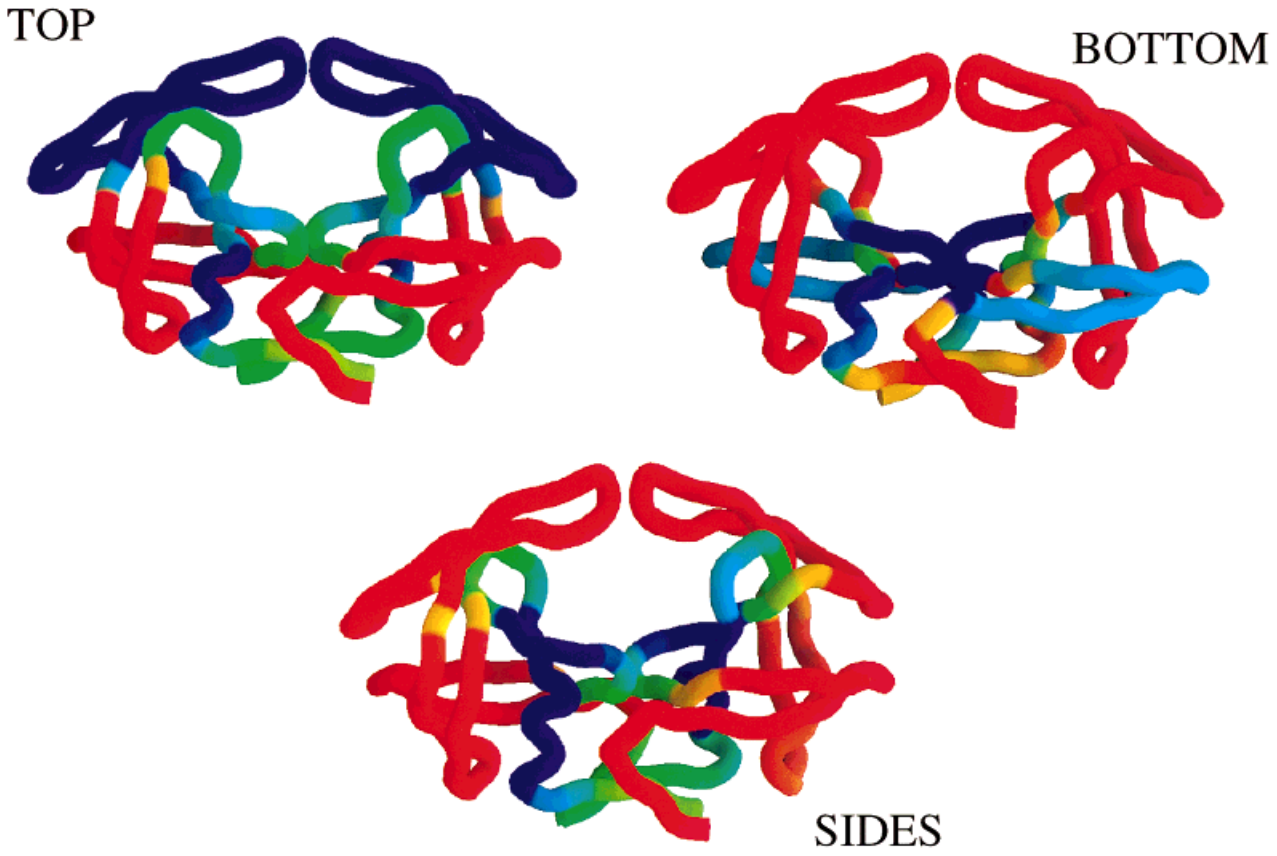

Fig. 7. Dissection of the effects of inhibitor binding on the magnitude of the stability constants per residue. **Top** shows the effects attributed to the binding perturbation on the residues at the top of the binding site (Met46, Ile47, Gly48, Gly49, and Ile50); **Bottom** shows the effects due to the perurbation at the bottom of the binding site (Arg8, Asp25, Gly27, Ala28, Asp29, and Asp30); and **Sides** shows the effects due to the perturbation at the sides of the binding sites (Pro81, Val82, and Ile84). Blue denote the residues most affected and red the residues least affected by binding. The figure was made using the program GRASP.[39]

additional energy of stabilization propagate throughout the molecule? Does it affect every residue in the molecule by an equal amount, or some residues more than others? The answer to these questions depends on the way in which cooperative interactions are distributed within the protease structure, i.e., the network of intramolecular interactions within different structural elements in the protease molecule. This cooperative network can be studied by examining the distribution of residue stability constants, κ_f, and their change upon inhibitor binding. The residue stability constants are an important statistical descriptor of protein stability because they explicitly account for the population of states with different degrees of folding.[31] For any particular residue j, the stability constant per residue, $\kappa_{f,j}$, is calculated as the ratio of the summed probability of all states in which residue j is folded, to the summed probability of all states in which that residue is not folded:

$$\kappa_{f,j} = \frac{\sum_i P_{f,j,i}}{\sum_i P_{nf,j,i}} \qquad (7)$$

In a previous paper, we performed a structure-based analysis of the protease molecule and calculated the entire set of residue stability constants.[3] Figure 6 **(A)** shows the structure of the HIV-1 protease color-coded according to the magnitude of $\kappa_{f,j}$. There are two regions with extreme stabilities. The central part of the molecule, including the dimer interface and adjacent areas (shown in blue), defines the most stable part of the protease while the region corresponding to the flap (shown in red) defines the least stable part of the molecule. While this analysis permits identification of regions with different stabilities, a characterization of cooperativity requires an analysis of the statistical correlation in the behavior of different residues, i.e., a quantitation of the extent by which a perturbation to any given residue will affect other residues in the molecule. Within this context, the presence of a ligand molecule can be interpreted as a direct or specific "perturbation" to those residues that define the binding site. The effects of this perturbation, however, are not limited to those amino acids that come in direct contact with the ligand. The effects propagate to other residues through cooperative interactions and manifest themselves in changes in the magnitude of the stability constants per

residue. Residues that are stabilized by the binding of a ligand molecule will show enhanced stability constants. From a statistical point of view, this effect occurs because ligand binding increases the probability of those states that are binding competent (i.e., the binding site is present). Experimentally, this phenomenon may be observed as an increase in the magnitude of the protection factors for hydrogen/deuterium exchange of those residues that show enhanced stability constants.[32]

Ligand or inhibitor binding will affect the stability of different protein conformations by an amount given by a set of linkage equations similar to equation 5. In the presence of a ligand X, the Gibbs energy of an arbitrary state (ΔG_i) is given by:

$$\Delta G_i = \Delta G_i^0 - RT \ln \frac{(1 + K_{a,i}[X])}{(1 + K_{a,0}[X])} \quad (8)$$

where $K_{a,o}$ is the binding constant to the native state and $K_{a,i}$ the binding constant to state i. The Gibbs energy of each conformational state will be affected in a manner dependent on the magnitude of the binding constant for that state. Accordingly, those states that are able to bind the ligand will be stabilized with respect to those states that are not able to bind the ligand, causing a change in the probability distribution of states.

In general, those protein conformational states in which the critical determinants of the binding site are formed will be able to bind the ligand with full or almost full affinity and will be preferentially stabilized with respect to those states in which the binding site is not formed. These ligand-induced changes in the probability distribution of conformational states will be reflected in the magnitude of the stability constants per residue (see Freire[33] for a general treatment). In the presence of a ligand the probability of any arbitrary state of the protein will be given by the equation:

$$P_i = \frac{\left(e - \frac{\Delta G_i^0}{RT}\right) \bullet \frac{(1 + K_{a,i}[X])}{(1 + K_{a,0}[X])}}{\sum_j \left(e - \frac{\Delta G_j^0}{RT}\right) \bullet \frac{(1 + K_{a,j}[X])}{(1 + K_{a,0}[X])}} \quad (9)$$

The Core_Bind algorithm calculates the expected changes in the probability distribution of different protein states induced by the presence of a ligand. This algorithm generates a large ensemble of partially folded conformations and calculates the intrinsic Gibbs energy (ΔG_i^0) of each state using procedures developed earlier.[16,31,33–35] For each state, Core_Bind evaluates if the binding site is structurally intact and according to that criteria determines whether the conformation is able to bind the ligand or not. In our calculations, it is assumed that those states in which the binding site is intact will exhibit a binding affinity similar to that found experimentally. This approach was used to investigate the effects of acetyl pepstatin on the structural cooperativity of the HIV-1 protease. Figure 6A shows the structure of the protease molecule color-coded according to the magnitude of the stability constants per residue (blue most stable, red least stable). Figure 6B shows the expected effects of acetyl pepstatin binding on the magnitude of the stability constants per residue (blue largest effect, red no effect). It is clear that under native conditions inhibitor binding does not affect all residues in the protein with the same intensity. Some areas are affected significantly (colored blue), whereas other areas are not affected at all (colored red). In particular, the binding of the inhibitor has a significant effect on those regions that define the binding site, especially the flap region which becomes structurally stable as a result of the additional binding free energy. The stabilization, however, propagates to regions not in direct contact with the inhibitor, specifically the strand between residues Pro9 and Ala22 and the helix between residues Arg87 and Gly94. On the other hand, the stability of the strand between Asp60 and Leu76 and the amino and carboxy terminii are the least affected by inhibitor binding. Consistent with these calculations, NMR studies of the HIV-1 protease bound to two inhibitors (DMP323 and P9941)[36] indicated that residues in the β strand (Asp 60, Ile 62, Val 75, and Leu 76) exhibited chemical exchange in the μs – ms range suggesting that these residues are not restricted to a single conformation in the inhibitor bound state. Unfortunately, the same type of information is still unavailable for the free protease and, therefore, a direct comparison of experimental differences between free and bound proteins cannot be done.

Cooperative Pathways

As described above, the binding energy is not uniformly distributed throughout the binding site (Fig. 5). Three distinct groups of amino acids located at the base, top, and sides of the binding site ("hot spots") contribute most of the binding energy. Those amino acids are located in regions of the protein that are distant in sequence and characterized by vastly different stabilities. These sites can be considered as the points of origin (initiation points) of thermodynamic perturbations on the protease molecule. When the ligand molecule binds, the perturbation is exerted simultaneously at all initiation points. While the overall effect is not additive (i.e., it is not equal to the sum of the effects of the isolated perturbations at each hot-spot), it is evident that the effects originating at individual hot-spots tend to concentrate preferentially in some specific areas. This is illustrated in Figure 7. Top shows the propagation of the effects attributed to the interactions of the inhibitor with the residues located in the flap region (top of the binding site: Met46, Ile47, Gly48, Gly49, and Ile50). Bottom shows the propagation of the effects attributed to the interactions of the inhibitor with the residues located at the bottom of the binding site (residues Arg8, Asp25, Gly27, Ala28, Asp29, and Asp30) and Sides the propagation of the effects due to the interactions with the lateral Sides of the binding site (residues Pro81, Val82, and Ile84). It is evident in the figure that the most extensive effects arise from the interactions of the ligand with the flap region. These

effects propagate significantly to the bottom of the binding site. Conversely, the effects arising from the interactions of the inhibitor with the bottom of the binding site are largely localized to that region. This is expected because the flap is a region with very low structural stability, whereas the bottom of the binding site belongs to the most stable core of the protease. In general, stabilizing perturbations in low stability regions trigger a much larger redistribution in the conformational ensemble than similar perturbations in more stable regions.[16]

The structure-based analysis indicates that isolated perturbations at discrete locations within the binding site do not trigger the complete array of effects induced by the ligand. This result suggests that the effects of a ligand depend not only on the magnitude of the affinity constant but also on the exact location of the residues that provide the main attachment points. It is conceivable that two ligands that bind to the same binding site with similar binding affinities could elicit different effects because they have a different structural distribution of binding interactions. This is particularly important for allosteric systems in which the binding of a ligand is required to elicit a response at a distant site. Unless the proper structural distribution of interactions is satisfied, ligand binding would not necessarily elicit the expected effect.

Traditionally, the main goal in ligand design has been the achievement of high affinity and specificity. The analysis presented here suggests that other considerations, like the structural distribution of binding interactions, need to be taken into account in order to elicit the required response. In addition, our analysis provides a framework for understanding the mechanism by which mutations distant from the active site affect substrate or inhibitor binding, an issue of importance for the elucidation of drug resistance.

REFERENCES

1. Bardi JS, Luque I, Freire E. Structure-based thermodynamic analysis of HIV-1 protease inhibitors. Biochemistry 1997;36:6588–6596.
2. Luque I, Todd MJ, Gomez J, Semo N, Freire E. The molecular basis of resistance to HIV-1 protease inhibition: a plausible hypothesis. Biochemistry 1998;37:5791–5797.
3. Todd M, Semo N, Freire E. The structural stability of the HIV-1 protease. J Mol Biol 1998;283:475–488.
4. Spinelli S, Liu QZ, Alzari PM, Hirel PH, Poljak RJ. The three-dimensional structure of the aspartyl protease from the HIV-1 isolate BRU. Biochimie 1991;73:1391–1396.
5. Jeng M-F, Englander SW. Stable submolecular folding units in a non-compact form of cytochrome *c*. J Mol Biol 1991;221:1045–1061.
6. Radford SE, Buck M, Topping KD, Dobson CM, Evans PA. Hydrogen exchange in native and denatured states of hen egg-white lysozyme. Proteins 1992;14:237–248.
7. Loh SN, Prehoda KE, Wang J, Markley JL. Hydrogen exchange in unligated and ligated staphylococcal nuclease. Biochemistry 1993; 32:11022–11028.
8. Kim K-S, Woodward C. Protein internal flexibility and global stability: Effect of urea on hydrogen exchange rates of bovine pancreatic trypsin inhibitor. Biochemistry 1993;32:9609–9613.
9. Woodward C. Is the slow-exchange core the protein folding core? TIBS. 1993;18:359–360.
10. Clarke J, Fersht AR. An evaluation of the use of hydrogen exchange at equilibrium to probe intermediates on the folding pathway. Folding Design 1996;1:243–254.
11. Swint-Kruse L, Robertson AD. Temperature and pH dependences of hydrogen exchange and global stability for ovomucoid third domain. Biochemistry. 1996;35:171–180.
12. Bai Y, Sosnick TR, Mayne L, Englander SW. Protein folding intermediates: native-state hydrogen exchange. Science. 1995;269: 192–197.
13. Jacobs MD, Fox RO. Staphylococcal nuclease folding intermediate charaterized by hydrogen exchange and NMR spectroscopy. Proc Natl Acad Sci USA 1994;91:449–453.
14. Morozova LA, Haynie DT, Arico-Muendel C, Van Dael H, Dobson CM. Structural basis of the stability of a lysozyme molten globule. Nat Struct Biol 1995;2: 871–875.
15. Schulman BA, Redfield C, Peng Z, Dobson CM, Kim PS. Different subdomains are most protected from hydrogen exchange in the molten globule and native states of human α-lactalbumin. J Mol Biol 1995;253:651–657.
16. Hilser VJ, Dowdy D, Oas TG, Freire E. The structural distribution of cooperative interactions in proteins: analysis of the native state ensemble. Proc Natl Acad Sci USA 1998;95:9903–9908.
17. Thompson K, Vinson C, Freire E. Thermodynamic characterization of the structural stability of the coiled-coil region of the bZIP transcription factor GCN4. Biochemistry 1993;32:5491–5496.
18. Johnson CR, Freire E. Structural stability of oligomeric proteins. Techniques Protein Chem 1996;VII:459–467.
19. Wyman J, Gill SJ. Binding and linkage: the functional chemistry of biological macromolecules. Mill Valley: University Science Books; 1990.
20. Straume M, Freire E. Two-dimensional differential scanning calorimetry: simultaneous resolution of intrinsic protein structural energetics and ligan binding interactions by global linkage analysis. Anal Biochem 1992;203:259–268.
21. Freire E. Statistical thermodynamic analysis of the heat capacity function associated with protein folding-unfolding transitions. Comments Mol Cell Biophys 1989;6:123–140.
22. Gomez J, Freire E. Thermodynamic mapping of the inhibitor site of the aspartic protease endothiapepsin. J Mol Biol 1995;252:337–350.
23. Brandts JF, Lin LN. Study of strong to ultratight protein interactions using differential scanning calorimetry. Biochemistry 1990; 29:6927–6940.
24. Murphy KP, Freire E. Thermodynamics of structural stability and cooperative folding behavior in proteins. Adv Protein Chem 1992; 43:313–361.
25. Gomez J, Hilser JV, Xie D, Freire E. The heat capacity of proteins. Proteins 1995;22:404–412.
26. Hilser VJ, Gomez J, Freire E. The enthalpy change in protein folding and binding. refinement of parameters for structure based calculations. Proteins 1996;26:123–133.
27. Lee KH, Xie D, Freire E, Amzel LM. Estimation of changes in side chain configurational entropy in binding and folding: general methods and application to helix formation. Proteins 1994;20: 68–84.
28. DAquino JA et al. The magnitude of the backbone conformational entropy change in protein folding. Proteins. 1996;25:143–156.
29. Luque I, Mayorga O, Freire E. Structure based thermodynamic scale of α-helix propensities in amino acids. Biochemistry 1996;35: 13681–13688.
30. Grant SK et al. Use of protein unfolding studies to determine the conformational and dimeric stabilities of HIV-1 and SIV proteases. Biochemistry 1992;31:9491–9501.
31. Hilser VJ, Freire E. Structure based calculation of the equilibrium folding pathway of proteins. Correlation with hydrogen exchange protection factors. J Mol Biol 1996;262:756–772.
32. Williams DC, Benjamin DC, Poljak RJ, Rule GS. Global changes in amide hydrogen exchange rates for a protein antigen in complex with three different antibodies. J Mol Biol 1996;257:866–876.
33. Freire E. The statistical thermodynamic linkage between conformational and binding equilibrium. Adv Protein Chem 1998;51:255–279.
34. Hilser VJ, Townsend BD, Freire E. Structure-based statistical thermodynamic analysis of t4 lysozyme mutants: structural mapping of cooperative interactions. Biophysical Chem. 1997;64: 69–79.

35. Hilser VJ, Freire E. Predicting the equilibrium protein folding pathway: structure-based analysis of staphylococcal nuclease. Proteins. 1997;27:171–183.
36. Nicholson LK et al. Flexibility and function in HIV-1 protease. Nat Struct Biol 1995;2:274–280.
37. Privalov G, Kavina V, Freire E, Privalov PL. Precise scanning calorimeter for studying thermal properties of biological macromolecules in dilute solution. Anal Biochem 1995;232:79–85.
38. Freire E, Mayorga OL, Straume M. Isothermal titration calorimetry. Anal Chem . 1990;62:950A–959A.
39. Nicholls A, Bharadwaj R, Honig B. GRASP: graphical representation and analysis of surface properties. Biophys J 1993;64:166–170.
40. Mildner AM, Rothrock DJ, Leone JW et al. The HIV-1 protease as enzyme and substrate: mutagenesis of autolysis sites and generation of a stable mutant with retained kinetic properties. Biochemistry 1994;33:9405–9413.