doi:10.1006/jmbi.2000.4057 available online at http://www.idealibrary.com on IDEAL® *J. Mol. Biol.* (2000) **303**, 567–582

JMB

Free Energy Calculations on Dimer Stability of the HIV Protease using Molecular Dynamics and a Continuum Solvent Model

Wei Wang[1] and Peter A. Kollman[1,2]*

[1]*Graduate Group in Biophysics*

[2]*Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, CA 94143, USA*

**Corresponding author*

Dimerization of HIV-I protease (HIV PR) monomers is an essential prerequisite for viral proteolytic activity and the subsequent generation of infectious virus particles. Disrupting dimerization of the enzyme can inhibit its activity. We have calculated the relative binding free energies between different dimers of the HIV protease using molecular dynamics and a continuum model, which we call MM/PBSA. We examined the dominant negative inhibition of the HIV PR by a mutated form of the protease and found relative dimerization free energies of homo- and hetero-dimerization consistent with experimental data. We also developed a rapid screening method, which was called the virtual mutagenesis method to consider other mutations which might stabilize non-wild-type heterodimers. Using this approach, we considered the mutations near the dimer interface which might cause dominant negative inhibition of the HIV PR. The rapid method we developed can be used in studying any ligand-protein and protein-protein interaction, in order to identify mutations that can enhance the binding affinities of the complex.

Keywords: HIV protease; dimer stability; molecular dynamics; MM/PBSA; virtual mutagenesis method

Introduction

Molecular dynamics (MD) has provided dynamic and atomic insights to help understand complicated biological systems. Free energy calculation methods have become powerful tools in providing quantitative measurement of protein-ligand or protein-protein interactions (Kollman, 1993; Beveridge & Dicapua, 1989; van Gunsteren, 1989). The most rigorous approaches to evaluate binding free energies are free energy perturbation (FEP) and thermodynamic integration (TI) methods. Both MD and FEP/TI have been successfully applied to the study of many protein and nucleic acids systems. Nonetheless, these methods are computationally intensive. Thus, many semi-empirical methods have been developed to estimate binding free energies faster and with reasonable accuracy (Ajay & Murcko, 1995).

A new method, MM/PBSA, was proposed last year for evaluating solvation and binding free energies of macromolecules and their complexes (Srinivason *et al*., 1998). When this method is used to calculate binding free energy, the binding free energy is decomposed into contributions from van der Waals and electrostatic energies, non-polar and electrostatic solvation free energies, and relative solute entropy effects (Massova & Kollman, 1999). The van der Waals and electrostatic interactions between the components of the complex are calculated using molecular mechanics (MM) with an empirical force field (Cornell *et al*., 1995), the non-polar part of solvation free energy is estimated by empirical methods based on solvent accessible (SA) surface and the electrostatic contribution to solvation is calculated using a continuum model and solving the Poisson-Boltzmann (PB) equation. The entropy contribution has been estimated using normal mode analysis. An ensemble of different conformations is extracted from MD trajectories and each snapshot is analyzed using this MM/PBSA method. The binding free energies are obtained from this ensemble average. This method is able to

Abbreviations used: HIV PR, HIV protease; MD, molecular dynamics; VM, virtual mutagenesis; FEP, free energy perturbation; TI, thermodynamic integration; MM, molecular mechanics; CNAN, contact neighbor atom number; TNAN, total neighbor atom number.

E-mail address of the corresponding author: pak@cgl.ucsf.edu

0022-2836/00/040567–16 $35.00/0

calculate the free energy difference between any two states, even when the two states are quite dissimilar from each other. It is also significantly more computationally efficient than FEP or TI. Here, we used this MM/PBSA method in studying dimer stability of the HIV protease (PR).

The dimeric HIV-1 protease (HIV-1 PR) is crucial for the maturation of viral structural (gag) and enzymatic (pol) proteins of the AIDS virion (Debouck *et al*., 1987). This aspartyl protease has been the therapeutic target for the treatment of AIDS. However, the HIV-1 virus rapidly develops drug resistant variants. Therefore, it is critically important to understand the mechanism of the HIV PR for designing inhibitors to combat this resistance. The primary structure of the HIV-1 protease indicates that each monomer of the protease contributes one catalytic aspartic acid residue at the active site of the enzyme. Either mutating one of the two catalytic aspartic acid residues (Kohl *et al*., 1988; Babe *et al*., 1991; Krasslich, 1991) or disrupting the dimerization of active HIV PR monomers (Zhang *et al*., 1991; McPhee *et al*., 1996; Rozzelle *et al*., 1999) has been reported to eliminate the catalytic activity of the protease and thus block the infectivity of the virus.

Craik and co-workers have shown that mixing of wild-type (wt) and certain mutant protease monomers could lead to inactivation of HIV-1 virus. They concluded this upon monitoring accumulation of unprocessed polyproteins and the secretion of non-infectious virons, and inferred that this loss of activity of the protease was due to the formation of inactive heterodimers between wild-type and defective monomers (Babe *et al*., 1991; McPhee *et al*., 1996). The defective monomers used in their experiments were obtained by mutating the aspartic acid in the catalytic triad and several other residues of the PR flap region. The goal was to promote the formation of defective heterodimers and decrease the stability of the wild-type and mutant PR homodimers. In their studies, they found a triple-mutation, Asp25Lys, Gly49Trp and Ile50Trp, which significantly reduced the levels of PR activity and virus infectivity (McPhee *et al*., 1996). Due to the large interface between two PR monomers, this dominant-negative inhibition of the HIV PR by defective monomers may be less susceptible to the emergence of resistant mutations. It suggests a potential use of gene therapy as a treatment to AIDS.

Here, we examined the protonation state of the ligand free HIV protease and estimated relative binding free energies between wild-type homodimer and defective heterodimer or mutant homodimer using the MM/PBSA method. Since it is not trivial to measure the binding affinities of different dimers experimentally, computer simulations can provide useful insights to aid understanding of the interactions between HIV PR subunits.

We also present here a new method, which we called the virtual mutagenesis (VM) method, to identify mutations on the interface of two molecules which may enhance the binding between them. We applied this method to HIV PR dimer and identified a few more potential dominant negative mutations. This VM method is applicable to any set of interacting molecules.

Results and Discussion

Protonation state of the ligand-free HIV PR is dianionic

Several distances between pairs of atoms in the two catalytic aspartic acid residues were measured in the crystal structure and the 100 snapshots taken from the MD trajectory (see Figure 1 and Table 1). Since the two catalytic Asp residues are critical for proper function of the HIV PR, it is important to maintain their structures during the simulation. From Table 1 and Figure 1, the distances between the two CA atoms are close to that in the crystal structure in all three protonation states. However, the distances between heavy atoms on the side-chains in mono- and double-protonated states are much smaller than those in the doubly ionic state. While the distances in the doubly ionic state are maintained closest to the distances in the crystal structure, the side-chains of the two Asp residues came closer to each other in the mono- and double-protonated states. The reason is obvious: the anionic Asp residues electrostatically repel each other. If either one of the two catalytic residues is protonated, the repulsion between them is greatly reduced and hydrogen bonds can also form

Table 1. The distances between pairs of atoms in the two catalytic aspartic acid residues of the HIV protease

		PP[c]	HP	HH[d]
Atom pair[a]	$R_{cryt'1}$ [b](Å)	R_{MD} [d] (Å)	R_{MD} [d] (Å)	R_{MD} [d] (Å)
CA-CA	6.71	6.55 ± 0.23	6.32 ± 0.17	6.29 ± 0.20
CB-CB	7.60	7.31 ± 0.22	6.90 ± 0.22	6.87 ± 0.19
CG-CG	5.28	5.01 ± 0.23	4.50 ± 0.24	4.41 ± 0.22
OD1-OD1	3.01	3.68 ± 0.29	2.58 ± 0.14	3.14 ± 0.25
OD2-OD2	5.81	5.12 ± 0.29	4.71 ± 0.34	4.48 ± 0.28

[a] The first atom is in Asp25 and the second in Asp25′.
[b] $R_{cryt'1}$ is measured in the crystal structure.
[c] Protonation states of the HIV PR, PP represents double ionic states; HP represents protonated Asp25 and deprotonated Asp25′; HH represents double protonated state;
[d] R_{MD} is the average distance of the 100 snapshots taken from the MD trajectory.

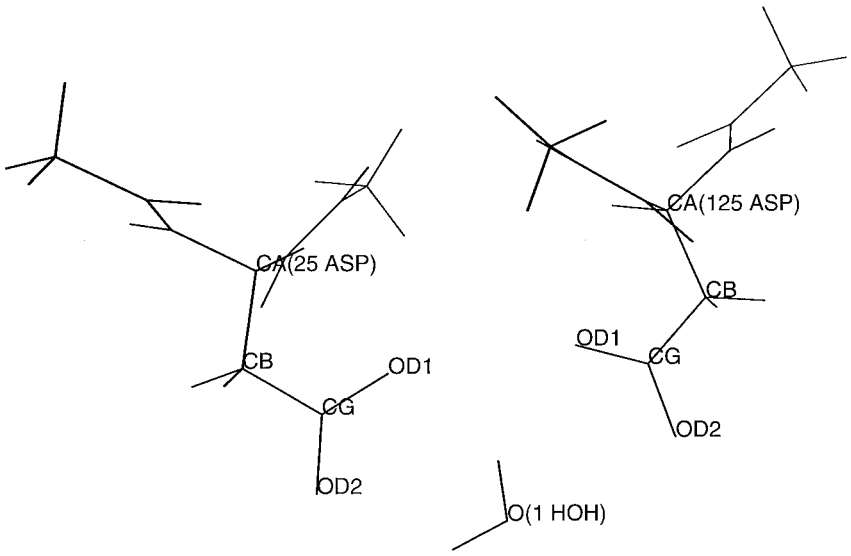

Figure 1. Fragments of the two catalytic aspartic acid residues capped with ACE and NME groups. A water molecule is proposed to interact with Asp in the way as shown in the Figure (Ido *et al.*, 1991).

between them directly or *via* nearby water molecules.

It is widely assumed that free HIV PR has a mono-protonated state and a water molecule is presumed to interact with the two catalytic Asp residues in the way shown in Figure 1 (Hyland *et al.*, 1991; Ido *et al.*, 1991). We investigated this assumption. One water molecule was put in the active site at the beginning of the MD simulations. The water molecule moved away in a few picoseconds of MD runs. But, other nearby water molecules moved into the active site and formed hydrogen bends with the catalytic Asp residues. It is assumed that a water molecule is crucial for the proteolystic reaction. Based on our simulations, this water molecule should be quite labile rather than fixed in the active site.

We calculated the binding free energies between the wild-type HIV PR dimers with different protonation states. The binding free energies and components for different dimers are shown in Table 2. In Table 2, the dianionic state has the most favorable binding free energy. Thus, the results of the binding free energy calculations and the distance measurement of pairs of atoms in the two catalytic Asp residues are consistent with the NMR data (Smith *et al.*, 1996), which suggests that both residues in the active site are deprotonated. It is worth pointing out that many experimental and theoretical works have been done to study the protonation states of the two catalytic Asp residues in the presence of HIV PR inhibitors (Yamazaki, *et al.*, 1994; Wang *et al.*, 1996; Luo *et al.*, 1998; Trylska *et al.*, 1999). The present and previous studies suggest that the binding of inhibitors has significant influence on the ionic states of the HIV PR.

As mentioned in Methods, the effect of conformational entropy upon dimerization is not included in equations (1)-(3). The absolute values of the binding free energies thus will overestimate the strength of binding. We estimated the free energy due to the conformational entropy using normal mode analysis. Due to the heavy computational demand of this analysis, we only carried out a single calculation as to estimate the order of magnitude of the conformational entropy contribution to the binding free energy. In our calculation, the conformational free energy is +71.1 kcal/mol at 298 K. If this value was included in our calculations, the values of binding free energies of the HIV dimer would fall into the range of −9 to −13 kcal/mol. The binding free energy measured experimentally varies with experimental conditions, such as pH (Zhang *et al.*, 1991; Cheng *et al.*, 1990; Grant *et al.*, 1992; Jordan *et al.*, 1992). At pH7, the K_d was measured as 50 nM (Cheng *et al.*, 1990), which corresponds to a binding free energy −10.0 kcal/mol at 298 K. The order of magnitude of our results is consistent with the experimental data. It is also worth pointing out that we assume that the two monomers of the HIV PR are already fully folded before forming the dimer. Because we are interested in calculating relative binding free energies between different dimers, the free energies of folding the two monomers are likely to cancel out.

If we examine each component of the binding free energy in Table 2, we can see that the order of binding free energies is the same as the order of the van der Waals interaction energies. In other words, van der Waals interactions are dominant in the HIV PR dimer binding. Non-polar solvation terms are similar in all protonation states, which is not unexpected. Electrostatic interaction energies, $\Delta G_{\text{int}}^{\text{ele}}$, and difference of electrostatic contribution to the solvation energy term, $\Delta G_{\text{sol}}^{\text{ele}}$, are quite different in the three protonation states. However, the sums of these two terms, $\Delta G_{\text{int + sol}}^{\text{ele}}$, are quite similar, suggesting that the two terms compensate each other. It helps to rationalize why considering solvation energy can greatly improve ranking ligands in drug design (Zou *et al.*, 1999).

Using a value of 1 underestimates the interior dielectric constant of proteins. Thus, we also calcu-

Table 2. Influence of the protonation states of the two aspartic acid residues at the active site to the binding free energies

				ε_{in}=1, ε_{out}=80				ε_{in}=2, ε_{out}=80			
Dimers	ΔG_{int}^{vdw} (kcal/mol)	ΔG_{int}^{ele} (kcal/mol)	ΔG_{sol}^{nonpol} [d] (kcal/mol)	ΔG_{sol}^{ele} [e] (kcal/mol)	$\Delta G_{int+sol}^{ele}$ [g] (kcal/mol)	ΔG_b [h] (kcal/mol)	$\Delta\Delta G_b$ [i] (kcal/mol)	ΔG_{sol}^{ele} [g] (kcal/mol)	$\Delta G_{int+sol}^{ele}$ [f] (kcal/mol)	ΔG_b [i] (kcal/mol)	$\Delta\Delta G_b$ [j] (kcal/mol)
WTP-	−183.0	−365.1	−16.6	+480.4	+115.3	−84.3	0	+236.7	+54.2	−145.5	0
WTP[a]	±0.2	±4.2	±0.8	±5.1	±0.9	±1.9		±2.5	±0.7	±1.4	
WTH-	−180.0	−332.6	−17.0	+447.1	+114.5	−82.5	+1.8	+219.9	+53.6	−143.4	+2.1
WTP[b]	±0.2	±14.4	±0.1	±12.4	±2.0	±1.8		±6.1	±1.1	±0.9	
WTH-	−173.6	−307.3	−16.6	+417.0	+109.7	−80.5	+3.8	+205.0	+51.4	−138.8	+6.7
WTH[c]	±0.5	±5.5	±0.2	±6.0	±0.5	±0.1		±3.0	±0.3	±0.4	

[a] The wild-type HIV PR dimer at double ionic state.
[b] The wild-type HIV PR dimer at mono-ionic state.
[c] The wild-type HIV PR dimer at double protonated state.
[d] $\Delta G_{sol}^{nonpol} = \Delta G_{sol\ D}^{nonpol} - \Delta G_{sol\ M1}^{nonpol} - \Delta G_{sol\ M2}^{nonpol}$.
[e] $\Delta G_{sol}^{ele}(\varepsilon_{in}=1, \varepsilon_{out}=1) = \Delta G_{RFE\ D}^{1\text{-}80} - \Delta G_{RFE\ M1}^{1\text{-}80} - \Delta G_{RFE\ M2}^{1\text{-}80}$.
[f] $\Delta G_{sol}^{ele}(\varepsilon_{in}=2, \varepsilon_{out}=1) = \Delta G_{RFE\ D}^{2\text{-}80} - \Delta G_{RFE\ M1}^{2\text{-}80} - \Delta G_{RFE\ M2}^{2\text{-}80}$.
[g] $\Delta G_{int+sol}^{ele} = \Delta G_{int}^{ele} + \Delta G_{sol}^{ele}$.
[h] $\Delta G_b = \Delta G_{int}^{vdw} + \Delta G_{int}^{ele}(\varepsilon_{in}=1, \varepsilon_{out}=1) + \Delta G_{sol}^{nonpol} + \Delta G_{sol}^{ele}(\varepsilon_{in}=1, \varepsilon_{out}=1)$.
[i] $\Delta G_b = \Delta G_{int}^{vdw} + (1/2)\Delta G_{int}^{ele}(\varepsilon_{in}=2, \varepsilon_{out}=1) + \Delta G_{sol}^{nonpol} + \Delta G_{sol}^{ele}(\varepsilon_{in}=2, \varepsilon_{out}=1)$.
[j] $\Delta G_b = \Delta G_b(\text{dimer}) - \Delta G_b(\text{WT})$.

lated binding free energies using a dielectric constant of 2 (see Table 2). It is encouraging that the ranking order is same as that obtained using dielectric constant 1. Therefore, we only use value 1 to calculate binding free energies for other dimers in the rest of this paper. We noticed that the absolute values of the binding free energies are too negative using dielectric constant 2. This is probably due to the fact that the parameterization of the force field we used was carried out with dielectric constant 1. We can see in Table 2 that the sums of the electrostatic interaction energy and the electrostatic contribution to solvation in three protonation states are +54.2, +53.6, and +51.4 kcal/mol, respectively, for dielectric constant 2, and +115.3, +114.5, and +109.7 kcal/mol, respectively, for dielectric constant 1 Using dielectric constant other than 1 reduced the influence of the overall electrostatic contribution to the binding free energy.

MM/PBSA can differentiate stabilities of different HIV PR dimers

Craik and co-workers found certain mutants which could inhibit the infectivity of AIDS (Babe *et al.*, 1991; McPhee *et al.*, 1996). Among these mutants, one triple mutation, Asp25Lys, Gly49Trp, Ile50Trp, had the most significant effect. Asp25Lys mutation caused the HIV PR loss of proteolytic

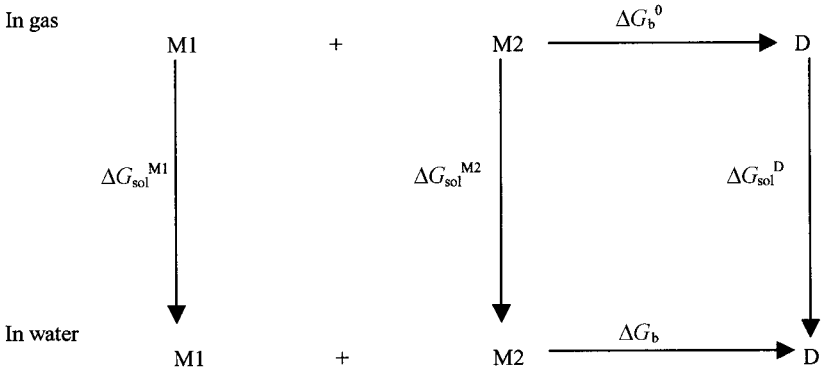

Figure 2. Thermodynamic cycle for calculating binding free energies ΔG_b^0 and ΔG_b^0 are binding free energies in gas and in water, respectively, ΔG_{sol}^{M1}, ΔG_{sol}^{M2} and ΔG_{sol}^{D} are solvation free energies for the monomer 1, monomer 2 and dimer of the HIV PR, respectively:

$$\Delta G_b = \Delta G_b^0 + \Delta G_{sol}^{D} - \Delta G_{sol}^{M1} - \Delta G_{sol}^{M2}$$

$$\Delta G_{sol} = \Delta G_{sol}^{ele} + \Delta G_{sol}^{nonpolar}$$

where ΔG_{sol}^{ele} was obtained from PB calculations and $\Delta G_{sol}^{nonpolar}$ was calculated from solvent accessible surface:

$$\Delta G_b^0 = \Delta G_{int}^{ele} + \Delta G_{int}^{vdw}$$

where ΔG_{int}^{ele} and ΔG_{int}^{vdw} were calculated from molecular mechanics energies.

activity. Gly49 and Ile50 are in the flap region of the HIV PR (see Figure 3). They were mutated to two Trp residues. Trp has larger side-chain group than either Gly or Ile. The purpose is to enhance formation of dimer between wild-type and mutant monomers but prevent dimerization between the mutant monomers. It also known that the flexibility of the ''flap'' region is crucial for the activity of the protease (Ishima *et al.*, 1999). The residues 1-27 and 60-99 in each monomer are defined as the ''core'' region and 28-59 as the ''flap'' region (Collins *et al.*, 1995). Two Trp residues in the ''flap'' region could reduce the flexibility and thus reduced the activity.

The average structure for each dimer during the 120 ps data collection period in MD simulations was calculated. The MD trajectory was superimposed with the average structure and the RMSD of heavy atoms on the backbone was calculated (Figure 4 and Table 3). The ''flap'' region of the wild-type homodimer, WTP-WTP, is most flexible. This is shown by the ratio between the deviation of the RMSD, σ, and the RMSD in Table 3. This ratio is 0.223 for WTP-WTP and about 0.140 for other dimers. In Figure 4(a), the fluctuation of the WTP-WTP is also obviously larger than others. For comparison, the RMSD, its deviation σ and the σ/RMSD ratio of the core regions are also listed in Table 3 and the RMSD values are plotted in Figure 4(b). Although the core region of the WTP-WTP dimer still has the largest σ/RMSD ratio, the difference between different dimers is not as large as the flap region. Thus, the triple mutations mainly influence the flexibility of the flap region of the HIV PR.

Binding free energies for different dimers are shown in Table 4; entropy terms are not included. As mentioned above, it is assumed that entropy terms are similar for the different dimers. The triple mutant KWW monomer bound most tightly to the monomer wt whose catalytic Asp is ionic. This dimer, WTP-KWW, is much more favorable than all other dimers. The mutant KWW homodimer, KWW-KWW, is most unfavorable, even worse than the wild-type homodimer, WTP-WTP. The wild-type monomer with protonated catalytic Asp binds to the mutant KWW monomer with an intermediate binding free energy (WTH-KWW), however, which is still more negative than that of the wild-type homodimer. The ranking order of the different dimers is consistent with experimental data. The heterodimer formed between wild-type and KWW mutant monomers is observed to have a higher melting temperature than the wild-type dimer (Rozzelle *et al.*, 1999). The defective homodimer, KWW-KWW, was not obtained in that experiment due to aggregation.

We can see that in Table 4 the wild-type homodimer, WTP-WTP, has the least favorable van der Waals interaction energy compared with the other dimers. This is because the triple mutant KWW has two Trp residues in the flap region and they provide stronger van der Waals interactions between the monomers. The total electrostatic contribution to the binding free energy, $\Delta G^{ele}_{int+sol}$, of the heterodimer with ionic Asp in the active site, WTP-KWW, is most favorable. The wild-type homodimer WTP-WTP, and the defective heterodimers WTP-KWW and WTH-KWW, have similar $G^{ele}_{int+sol}$. The values are +115.3, +114.8 and +117.3 kcal/mol, respectively. The non-polar part of solvation free energy is also similar for different dimers. For WTP-WTP, WTP-KWW and WTH-KWW, the van der Waals interaction energy determines the rank order of the binding free energy. For the mutant homodimer, KWW-KWW, however, although the van der Waals is much more favorable, the $\Delta G^{ele}_{int+sol}$ term is over 20 kcal/mol less favorable compared with the other three dimers. The KWW-KWW has less

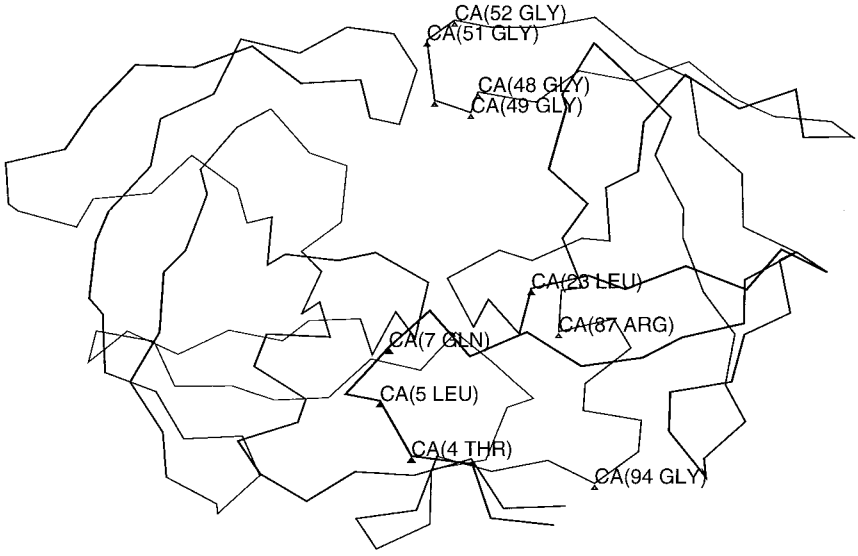

Figure 3. Locations of the mutations on the HIV protease.

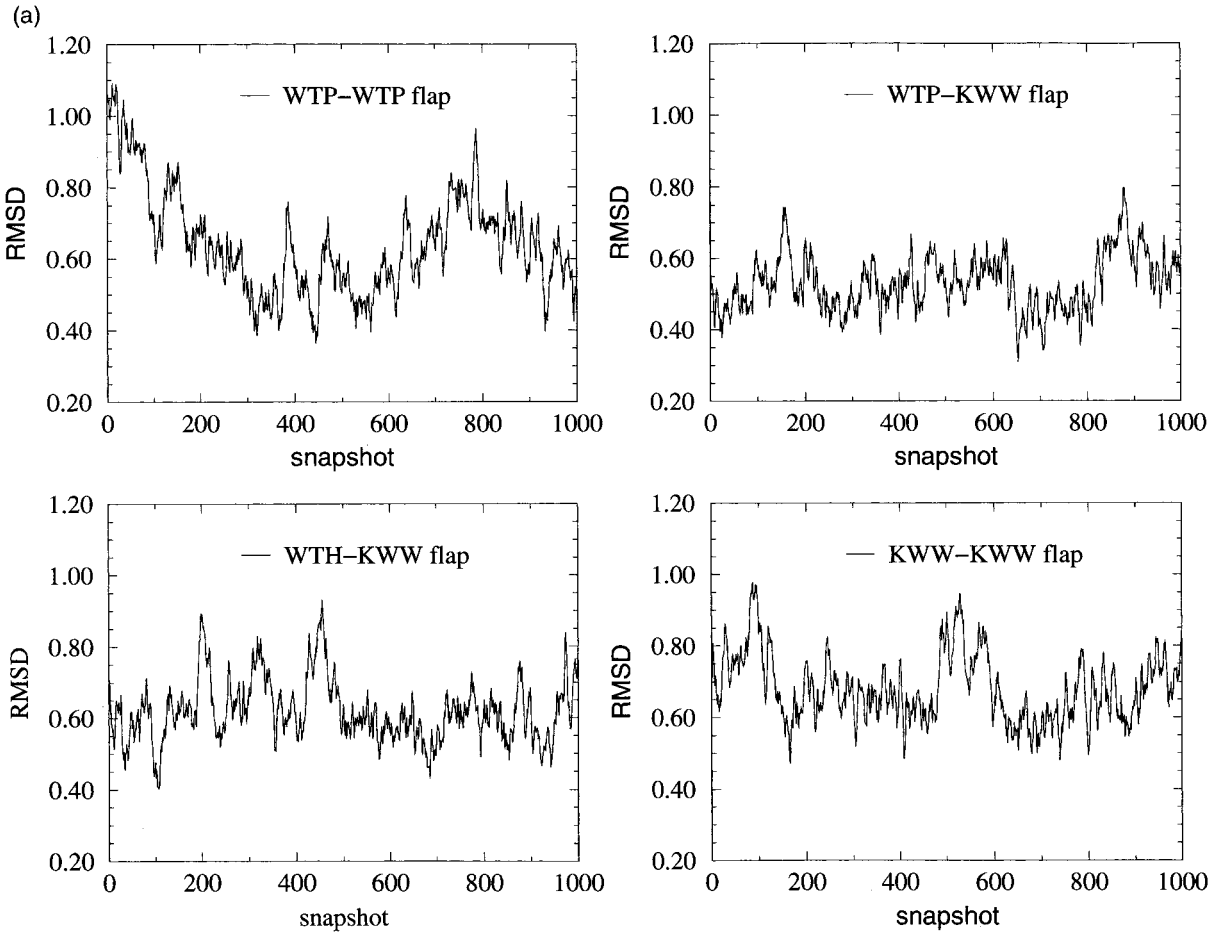

Figure 4 (*legend opposite*)

favorable electrostatic energy and more unfavorable electrostatic solvation energy than the wild-type homodimer. The less favorable electrostatic energy is due to the stronger repulsive interaction between the two Lys residues in KWW-KWW than that between the two Asp residues in WTP-WTP. The average distance between the two NZ atoms in the two Lys residues in KWW-KWW is 4.61(±0.31) Å compared with 5.01(±0.23) Å between the two CG atoms in the two Asp residues in WTP-WTP. The Lys in one monomer also repulsively interacts with Trp49 and Trp50 in another monomer in KWW-KWW, but there are no such repulsive interactions in WTP-WTP in which Gly49 and Ile50 are further away from the catalytic Asp in another monomer. If we examine the structure of KWW-KWW, we can see that the aromatic rings of Trp49A and Trp49B are partially buried by Trp50A and Trp50B, respectively. This burial of polar groups gives a larger solvation penalty to the KWW-KWW dimerization than the WTP-WTP dimerization. This explains why the KWW-KWW dimer has more unfavorable electrostatic solvation energy than the WTP-WTP dimer.

The VM method can predict several new dominant negative mutants

Encouraged by the above results, we tried a simpler but faster approach, which we named virtual

Table 3. RMSD and its deviation σ of the flap region (residue 28-60 in each monomer) and the core region (residue 1-27 and 61-99) for different HIV PR dimers

Dimer	Flap RMSD[a]	Flap RMSD deviation σ	Flap σ/RMSD	Core RMSD[a]	Core RMSD deviation σ	Core σ/RMSD
WTP-WTP	0.645	0.144	0.223	0.571	0.074	0.130
WTP-KWW	0.529	0.078	0.147	0.534	0.046	0.086
WTH-KWW	0.620	0.087	0.140	0.523	0.060	0.115
KWW-KWW	0.681	0.095	0.140	0.578	0.058	0.100

[a] RMSD was calculated for all heavy atoms on the backbone compared with average structure obtained from the MD.

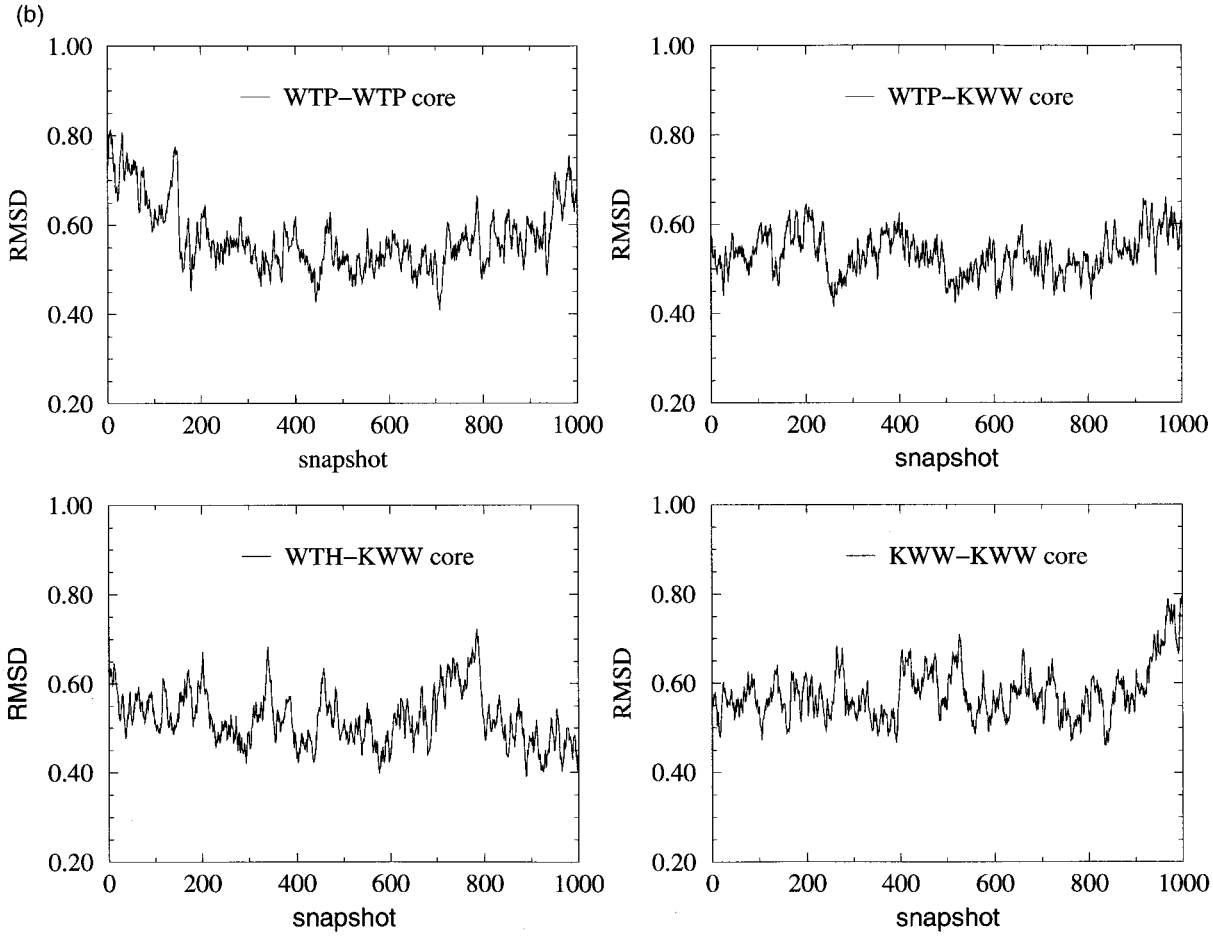

Figure 4. (a) RMSD of the flap region of different HIV protease dimers. Snapshot is taken every 0.120 ps. (b) RMSD of the ''core'' region of different HIV protease dimers. Spanshot is taken every 0.120 ps.

mutagenesis (VM) method, to estimate the relative binding free energies between different dimers. We took one snapshot that has the closest binding free energy to the average binding free energy value obtained from the MD trajectory. We can see from Table 5 that the binding free energy of the snapshot we chose is −84.5 kcal/mol, which is very close to the average value −84.3 kcal/mol calculated from the MD trajectory (Table 3). Mutations were suggested by a fast screening procedure (see below) and then made on this snapshot. For each mutation, a systematic conformation search for total 100 conformations was performed. Only those conformations with no steric clash with other atoms in the molecule were further investigated (see Methods). Each surviving conformation was minimized with a distance dependent dielectric constant while all other residues in the molecule were fixed. The binding free energy was then calculated using MM/PBSA. The final binding free energy for each mutation is the average value for all rotamers.

We first applied this VM method to several mutations for which experimental data were available. The screening procedure is not necessary here. No binding free energy or disassociate equilibrium constant K_d has been measured on any of those mutant dimers. However, there is experimental evidence indicating that thermal denaturation of single chain heterodimers, D25K, G49W/I50W and D25K/G49W/I50W, have a 1.5°C to 7.2°C higher thermal stability than single chain wild-type HIV PR (Rozzelle *et al.*, 1999). The accumulation of unprocessed polyproteins and the secretion of non-infectious virons display the same trend as the thermal stability. Thus, we assume that the binding affinities between different mutant and wild-type monomer are in the same order as viral infectivity. With this assumption in mind, we found that the calculated binding free energies are consistent with the experimental data (Table 5).

Let us further examine some of these mutations. For the 49W mutation, it has a −1.8 kcal/mol more favorable van der Waals interaction than wild-type homodimer because the Trp has a much larger side-chain than Gly. The non-polar solvation free energy difference, ΔG_{sol}^{nonpol} is −0.1 kcal/mol more favorable for the mutant heterodimer. However, the total electrostatic contribution, $\Delta G_{int+sol}^{ele}$, is +1.4 kcal/mol less favorable for the mutant heterodimer. In net, the total binding free energy

Table 4. MM/PBSA results on the binding free energies of different HIV PR dimers

Dimers	ΔG_{int}^{vdw} (kcal/mol)	ΔG_{int}^{ele} (kcal/mol)	ΔG_{sol}^{nonpol} [e] (kcal/mol)	ΔG_{sol}^{ele} [f] (kcal/mol)	$\Delta G_{int+sol}^{ele}$ [g] (kcal/mol)	ΔG_b [h] (kcal/mol)	ΔG_b [i] (kcal/mol)	Expt'l ranking order
WTP-WTP[a]	-183.0 ± 0.2	-365.1 ± 4.2	-16.6 ± 0.8	$+480.4 \pm 5.1$	$+115.3 \pm 0.9$	-84.3 ± 1.9	0	2
WTP-KWW[b]	-190.6 ± 0.9	-445.4 ± 6.6	-17.8 ± 0.2	$+560.2 \pm 4.7$	$+114.8 \pm 1.9$	-93.6 ± 0.8	-9.3	1
WTH-KWW[c]	-187.8 ± 1.6	-293.7 ± 10.6	-18.1 ± 0.2	$+411.0 \pm 8.8$	$+117.3 \pm 1.9$	-88.6 ± 0.5	-4.3	1
KWW-KWW[d]	-199.6 ± 2.2	-159.0 ± 9.0	-19.4 ± 0.1	$+298.2 \pm 5.9$	$+139.2 \pm 3.0$	-79.8 ± 0.9	$+4.3$	N/A

[a] The wild-type homodimer with double ionic catalytic Asp residues.
[b] The heterodimer between the wild-type monomer with ionic catalytic Asp and the triple mutation monomer.
[c] The heterodimer between the wild-type monomer with protonated catalytic Asp and the triple mutation monomer.
[d] The triple mutation homodimer.
[e] $\Delta G_{sol}^{nonpol} = \Delta G_{sol\ D}^{nonpol} - \Delta G_{sol\ M1}^{nonpol} - \Delta G_{sol\ M2}^{nonpol}$;
[f] $\Delta G_{sol}^{ele}(\varepsilon_{in} = 1, \varepsilon_{out} = 1) = \Delta G_{RFE\ D}^{1\text{-}80} - \Delta G_{RFE\ M1}^{1\text{-}80} - \Delta G_{RFE\ M2}^{1\text{-}80}$.
[g] $\Delta G_{int+sol}^{ele} = \Delta G_{int}^{ele} + \Delta G_{sol}^{ele}$.
[h] $\Delta G_b = \Delta G_{int}^{vdw} + \Delta G_{int}^{ele}(\varepsilon_{in} = 1, \varepsilon_{out} = 1) + \Delta G_{sol}^{nonpol} + \Delta G_{sol}^{ele}(\varepsilon_{in} = 1, \varepsilon_{out} = 1)$.
[i] $\Delta\Delta G_b = \Delta G_b(\text{dimer}) - \Delta G_b(\text{WT})$.

is -0.5 kcal/mol more favorable for mutant heterodimer.

The 49W50W mutant is similar to 49W. The 49W50W heterodimer has larger favorable van der Waals interaction energy, -193.2 kcal/mol, compared with -191.6 kcal/mol for 49W heterodimer and -189.8 kcal/mol for wild-type homodimer. Ile50 mutated to Trp provides -1.6 kcal/mol van der Waals interaction energy to binding free energy *versus* -1.8 kcal/mol while Gly49 is mutated to Trp. This is not unexpected because Ile has a larger side-chain than Gly so that the mutation from Ile to Trp has smaller effect than that of Gly to Trp. The non-polar solvation energy has a small but favorable contribution to the stability of the 49W50W heterodimer. The total electrostatic contribution, $\Delta G_{int+sol}^{ele}$, is unfavorable compared with 49W and wild-type dimer and it cancels part of the favorable van der Waals interactions. The unfavorable $\Delta G_{int+sol}^{ele}$ terms in 49W and 49W50W are due to the unfavorable electrostatic interactions between Asp25 and 49W/50W. The aromatic ring of Trp49 is partially buried by Trp50 in the 49W50W heterodimer. This helps to explain why 49W50W has a more unfavorable $\Delta G_{int+sol}^{ele}$ term than 49W. The binding free energy of 49W50W is -0.8 kcal/mol more favorable than wild-type homodimer and -0.3 kcal/mol than the 49W heterodimer. Obviously, favorable van der Waals interaction is dominant in the 49W50W heterodimer.

The total electrostatic contribution term, $\Delta G_{int+sol}^{ele}$, is not unfavorable in the 25K49W50W heterodimer case. Instead, it is -2.0 kcal/mol more favorable than wild-type homodimer. This is due to favorable electrostatic interactions between Lys25 in the mutant monomer and Asp25 in the wild-type monomer. The van der Waals and non-polar solvation terms have similar values as in the 49W50W heterodimer. This suggests that the mutation of Asp25 to a positive charged residue with stronger binding than the wild-type homodimer is mainly due to favorable electrostatic interaction.

The above conclusion is consistent with the D25K data where the van der Waals and non-polar solvation only are -0.3 kcal/mol and -0.1 kcal/mol more favorable than wild-type homodimer but $\Delta G_{int+sol}^{ele}$ term contributes -1.0 kcal/mol. However, in the D25R heterodimer, the $\Delta G_{int+sol}^{ele}$ term is unfavorable relative to the wild-type dimer and it is the van der Waals and non-polar solvation energies that make the total binding free energy of the D25R heterodimer -1.6 kcal/mol more favorable than the wild-type homodimer. This is interesting. It is worthy pointing out that the calculated bind-

Table 5. Binding free energies of several dimers calculated using the VM method

Mutation	ΔG_{int}^{vdw} (kcal/mol)	ΔG_{int}^{ele} (kcal/mol)	ΔG_{sol}^{nonpol} [b] (kcal/mol)	ΔG_{sol}^{ele} [c] (kcal/mol)	$\Delta G_{int+sol}^{ele}$ [d] (kcal/mol)	ΔG_b [e] (kcal/mol)	$\Delta\Delta G_b$ [f] (kcal/mol)	Expt'l viral infectivity ranking order[a]
25K49W50W	-193.1 ± 3.0	-425.3 ± 1.9	-17.8 ± 0.3	$+546.0 \pm 2.8$	$+120.7 \pm 2.4$	-90.2 ± 1.7	-5.7	1
25R	-191.7 ± 2.1	-401.8 ± 10.0	-17.7 ± 0.2	$+525.1 \pm 9.5$	$+123.3 \pm 3.4$	-86.1 ± 3.5	-1.6	2
25K	-190.1 ± 1.6	-401.3 ± 15.0	-17.5 ± 0.1	$+523.0 \pm 14.4$	$+121.7 \pm 2.8$	-85.9 ± 3.5	-1.4	3
49W50W	-193.2 ± 2.8	-350.6 ± 1.8	-17.8 ± 0.3	$+476.3 \pm 2.11$	$+125.7 \pm 2.1$	-85.3 ± 1.5	-0.8	4
49W	-191.6 ± 1.2	-347.2 ± 1.3	-17.5 ± 0.1	$+471.3 \pm 2.5$	$+124.1 \pm 1.82$	-85.0 ± 1.1	-0.5	5
Wild-type	-189.8	-344.7	-17.4	$+467.4$	$+122.7$	-84.5	0	6

[a] The smaller ranking order, the weaker the viral infection, it is assumed, the more favorable the binding free energy for the dimer.
[b] $\Delta G_{sol}^{nonpol} = \Delta G_{sol\ D}^{nonpol} - \Delta G_{sol\ M1}^{nonpol} - \Delta G_{sol\ M2}^{nonpol}$;
[c] $\Delta G_{sol}^{ele}(\varepsilon_{in} = 1, \varepsilon_{out} = 1) = \Delta G_{RFE\ D}^{1\text{-}80} - \Delta G_{RFE\ M1}^{1\text{-}80} - \Delta G_{RFE\ M2}^{1\text{-}80}$.
[d] $\Delta G_{int+sol}^{ele} = \Delta G_{int}^{ele} + \Delta G_{sol}^{ele}$.
[e] $\Delta G_b = \Delta G_{int}^{vdw} + \Delta G_{int}^{ele}(\varepsilon_{in} = 1, \varepsilon_{out} = 1) + \Delta G_{sol}^{nonpol} + \Delta G_{sol}^{ele}(\varepsilon_{in} = 1, \varepsilon_{out} = 1)$.
[f] $\Delta\Delta G_b = \Delta G_b(\text{dimer}) - \Delta G_b(\text{WT})$.

ing free energies of the D25R and the D25K heterodimers have larger error bars. This is because Asp25 is in the active site and has empty space around it. Many conformations can be considered for the mutant residues and this causes the relatively larger variation. Comparing D25R and D25K we can see D25R has more favorable van der Waals interaction than D25K. This is reasonable because the Arg side-chain is larger than the Lys side-chain. The unfavorable $\Delta G^{ele}_{int+sol}$ term is unexpected. Compared with D25 K, D25R has similar electrostatic interaction energy, −401.8 kcal/mol *versus* −401.3 kcal/mol, but more unfavorable PB solvation energy, +525.1 kcal/mol *versus* +523.0kcal/mol. This unfavorable $\Delta G^{ele}_{int+sol}$ term may be due to the additional burial of NH1 or NH2 polar groups in the Arg upon dimerization.

With the above encouraging results, we can try to predict some new mutations which may enhance binding of defective heterodimers. Since the interface between the HIV PR monomers is large and the interface between the two monomers is well packed, it is difficult to determine which residue can be mutated if one just visualizes the structure of the HIV PR. We exploited the simple method which is described in Methods to scan all possible residues which are close to interface but still have enough surrounding space to allow larger side-chain replacement. Each residue which satisfies the scanning criteria is evaluated by the VM method.

In Table 6, contact neighbor atom number (CNAN) and total neighbor atom number (TNAN) of all residues whose C^{α}-C^{β} vector points to the interface are listed. CNAN counts how many contacts one residue has with another monomer and TNAN indicates how crowded a given residue is. In the present study, we used two distance cutoffs, 3 Å and 6 Å. The shell contact neighbor atom number ratio (SCNANR) was calculated for each residue. SCNANR shows how many contacts one residue can make with another monomer between a 3 Å and 6 Å shell around it. In order to find mutation to enhance binding, one wants to identify residues which have small TNAN with a 3 Å distance cutoff and a large SCNANR. A small TNAN in the 3 Å distance cutoff means the residue has empty surrounding space so that larger side-chain replacement is possible. A large SCNANR shows that a larger side-chain has the potential to have more contacts or stronger van der Waals interactions with another monomer.

In this study, we chose 20 as the threshold for TNAN using the 3 Å distance cutoff and 20 % as the cutoff for SCNANR. The VM calculations were only performed on those residues whose TNAN using 3 Å distance cutoff was less than 20 and for which SCNANR was larger than 20 %. Some 11 residues which satisfy these two criteria are printed as bold in Table 6. Five of them, Gln2,

Table 6. Average contact neighbor atom number (CNAN) and total neighbor atom number (TNAN) of residues whose C^{α}-C^{β} vector points toward the HIV PR dimer interface

Sequence number[a]	Residue name	3 Å distance cutoff		6 Å distance cutoff		SCNANR[b] (X100 %)
		CNAN	TNAN	CNAN	TNAN	
2A	**GLN**	**8**	**19**	**39**	**91**	**46.0**
4A	**THR**	**2**	**15**	**29**	**111**	**28.1**
5A	LEU	17	31	129	204	64.7
6A	**TRP**	**2**	**13**	**31**	**81**	**42.7**
7A	**GLN**	**2**	**17**	**22**	**97**	**25.0**
9A	PRO	6	26	51	192	27.1
23A[c]	LEU	3	30	24	200	12.4
24A	LEU	7	32	49	229	21.3
25A[c]	ASP	2	22	37	170	23.7
26A	THR	13	28	105	205	52.0
27A	**GLY**	**6**	**17**	**65**	**122**	**56.2**
48A	**GLY**	**4**	**18**	**28**	**100**	**29.3**
49A	**GLY**	**3**	**17**	**35**	**113**	**33.3**
50A	**ILE**	**4**	**14**	**51**	**90**	**61.8**
51A	**GLY**	**7**	**16**	**56**	**95**	**62.0**
52A	**GLY**	**2**	**16**	**26**	**100**	**28.6**
67A	ABA	1	15	6	115	5.0
69A	HIS	2	21	17	129	13.9
87A	ARG	9	29	79	219	36.8
90A	LEU	2	34	37	239	17.1
93A	ILE	3	27	21	164	13.1
94A	**GLY**	**0**	**14**	**22**	**95**	**27.2**
96A	THR	11	23	104	158	68.9
97A	LEU	24	38	163	236	70.2
99A	PHE	18	22	148	174	85.5

[a] Sequence number is according to the wild-type HIV PR dimer. The PDB entry is 3hvp.
[b] SCNANR is the shell contact neighbor atom number ratio (see Methods).
[c] 23A is listed due to special interests (see text).

Thr4, Trp6, Gln7 and Gly94 are in the core region. Gln2 is too close to the N-terminal of the chain. Therefore no mutation was made. Trp6 was not mutated because no natural amino acid residue with a larger side-chain exists. Gly27 is in the catalytic triad. The remaining five, Gly48, Gly49, Ile50, Gly51 and Gly52, are in the flap region. Gly27 has ϕ and ψ, angles in the right side of the Ramachandran map. Thus, we did not mutate Gly27 either.

Since we are interested in finding new mutant monomers which can inhibit virial infectivity, we first mutated Asp25 to Lys. This D25K mutation can also reduce binding affinities between defective homodimers (see Table 4). Residues identified by our scanning method were then also mutated. The binding free energies calculated by the VM method are listed in Table 7. Mutations with more favorable binding free energies than D25K are printed in bold and those with binding free energies between the wild-type homodimer and the D25K heterodimer are printed in italics.

The most interesting mutations are those in the flap region. As discussed above, compared with mutations in other regions, these mutations can reduce the flexibility of the flap region and, thus, can further reduce the activity of the HIV PR. Since these residues are exposed to water, we mutated them to Trp so that they can have stronger van der Waals interactions with another monomer but do not get dramatic unfavorable solvation free energy penalties. 25K48W, 25K49W, 25K50W and 25K52W have more favorable binding free energies, even compared with the D25K heterodimer. Among these four mutations, 25K48W and 25K50W have the most and least favorable van der Waals interaction energy, respectively. The reason is that the mutated residue, Trp, can pack well with Ile50 from another chain in the 25K48W case. This packing also explains the most favorable non-polar solvation energy for 25K48W because the solvent-accessible surface of Ile50 from another chain is reduced. This deeper burial of the hydrophobic residue favors binding. For 25K50W, the Ile is much larger than Gly in the wild-type protease. Thus, the mutation to Trp from Ile does not create as many more van der Waals contacts than the wild-type dimer as the mutation from Gly to Trp. The 25K49W and the 25K52W have intermediate van der Waals interaction energies, as one would expect. In terms of total electrostatic contribution to binding free energy, $\Delta G^{ele}_{int+sol}$, the 25K50W has the least unfavorable value. This is also not unexpected because the hydrophobic residue Ile50 is exposed to water. If it is replaced by Trp, Trp has large aromatic ring and, therefore, a more favorable solvation energy. However, the 25K51W has a more unfavorable binding free energy than the wild-type homodimer. We can see that this is due to an very unfavorable van der Waals interaction energy. It implies that there are steric clashes. Thus, we mutated Gly51 to Ala instead of Trp. The van der Waals interaction becomes more favorable than the 25K51W but still unfavorable if compared with the wild-type homodimer. We examined the structure and found that this is due to the fact that Gly51 is flanked by Ile50 in the same chain and Phe53 in another chain. The C^{β} atom in the substituted Ala or Trp has unfavorable contacts with these two residues.

Three other mutations, 25K4Y, 25K7W and 25K94W are in the core region of the HIV PR. 25K7W has the least unfavorable $\Delta G^{ele}_{int+sol}$ term. Since Gln7 is on the surface of the HIV PR, mutation to Trp can provide more favorable solvation energy. So can the 25K4Y, which also has a less unfavorable $\Delta G^{ele}_{int+sol}$ term than the 25K mutation. After Thr4 and Gln7 are mutated to Tyr and Trp, respectively, Tyr4 and Trp7 can reduce

Table 7. Binding free energies of different dimers calculated using the VM method

Mutation	ΔG^{vdw}_{int} (kcal/mol)	ΔG^{ele}_{int} (kcal/mol)	ΔG^{nonpol}_{sol} [a] (kcal/mol)	ΔG^{ele}_{sol} [b] (kcal/sol)	$\Delta G^{ele}_{int+sol}$ [c] (kcal/mol)	ΔG_b [d] (kcal/mol)	ΔG_b [e] (kcal/mol)
Wild-type	−189.8	−344.7	−17.4	+467.4	+122.7	−84.5	0
25K	−190.1 ± 1.6	−401.3 ± 15.0	−17.5 ± 0.1	+523.0 ± 14.4	+121.7 ± 2.8	−85.9 ± 3.5	−1.4
25K4Y	**−191.4 ± 1.7**	**−421.9 ± 0.5**	**−17.6 ± 0.1**	**+542.1 ± 3.3**	**+120.2 ± 3.0**	**−88.8 ± 1.8**	**−4.3**
25K5F	−186.8 ± 0.0	−386.5 ± 0.1	−17.5 ± 0.0	±514.1 = 0.3	±127.6 ± 0.0	−76.7 ± 0.0	+7.8
25K7W	**−191.1 ± 0.5**	**−423.4 ± 0.5**	**−17.6 ± 0.0**	**+538.0 ± 1.6**	**+114.6 ± 1.3**	**−94.1 ± 1.1**	**−9.6**
25K23Y	−190.2 ± 1.7	−394.8 ± 4.3	−17.5 ± 0.0	+518.4 ± 4.1	+123.6 ± 2.6	−84.1 ± 2.9	+0.4
25K48W	**−195.7 ± 0.5**	**−422.5 ± 0.4**	**−18.5 ± 0.0**	**+545.6 ± 0.8**	**+123.1 ± 0.4**	**−91.1 ± 0.2**	**−6.6**
25K49W	**−192.4 ± 0.3**	**−393.9 ± 0.5**	**−17.6 ± 0.1**	**+517.1 ± 0.7**	**+123.2 ± 0.1**	**−86.8 ± 0.3**	**−2.3**
25K50W	**−191.6 ± 2.8**	**−425.1 ± 1.0**	**−17.7 ± 0.3**	**+545.2 ± 2.8**	**+120.1 = 2.4**	**−89.2 ± 0.8**	**−4.7**
25K51W	−171.2 ± 0.2	−420.9 ± 0.0	−17.8 ± 0.0	+540.0 ± 0.5	+119.1 ± 0.7	−69.9 ± 0.7	+14.6
25K51A	−177.5 ± 4.6	−422.5 ± 1.3	−17.6 ± 0.0	+540.6 ± 1.6	+118.1 ± 0.3	−77.0 ± 4.3	+7.5
25K52W	**−192.4 ± 0.9**	**−422.4 ± 0.9**	**−17.7 ± 0.0**	**+544.4 ± 2.9**	**+122.0 ± 2.5**	**−88.1 ± 1.9**	**−3.6**
25K87W	−187.4 ± 0.2	−428.7 ± 0.4	−17.8 ± 0.1	+553.0 ± 0.2	+124.3 ± 0.4	−80.9 ± 0.3	+3.6
25K94W	*−192.7 ± 1.8*	*−424.2 ± 4.5*	*−17.8 ± 0.2*	*+549.7 ± 6.4*	*+125.5 ± 2.7*	*−85.0 ± 3.9*	*−0.5*
23Y	−189.7 ± 2.0	−348.3 ± 3.9	−17.4 ± 0.0	+472.2 ± 4.1	+123.9 ± 1.3	−83.2±2.1	+1.3

[a] $\Delta G^{nonpol}_{sol} = \Delta G^{nonpol}_{sol\ D} - \Delta G^{nonpol}_{sol\ M1} - \Delta G^{nonpol}_{sol\ M2}$.
[b] $\Delta G^{ele}_{sol}(\varepsilon_{in}=1, \varepsilon_{out}=1) = \Delta G^{1\text{-}80}_{RFE\ D} - \Delta G^{1\text{-}80}_{RFE\ M1} - \Delta G^{1\text{-}80}_{RFE\ M2}$.
[c] $\Delta G^{ele}_{int+sol} = \Delta G^{ele}_{int} + \Delta G^{ele}_{sol}$.
[d] $\Delta G_b = \Delta G^{vdw}_{int} + \Delta G^{ele}_{int}(\varepsilon_{in}=1, \varepsilon_{out}=1) + \Delta G^{ele}_{sol}(\varepsilon_{in}=1, \varepsilon_{out}=1)$.
[e] $\Delta\Delta G_b = \Delta G_b$(dimer)-$\Delta G_b$(WT).

solvent-accessible surface of some nearby hydrophobic residues, such as Leu10, Leu5 and Ile31. This deeper burial of hydrophobic residues is also favorable for binding. 25K4Y and 25K7W also have more favorable van der Waals interaction energies than wild-type, which is due to larger side-chain replacement. However, the main contributions to the binding are from the $\Delta G^{ele}_{int+sol}$ term. The 25K94W heterodimer has the most favorable van der Waals interaction in these three dimers. This is because when Gly94 is mutated to Trp, the Trp can pack well with Trp6 from another chain. It is the van der Waals interaction that leads to the total binding free energy of 25K94W being more favorable than the wild-type dimer.

In addition to those 11 residues which were identified by our scanning method, we also performed calculations on other residues.

Craik and co-workers proposed that the L23Y mutation might enhance the binding for the defective heterodimer. On the basis of analyzing the structure of the HIV PR, the L23Y mutation may form new hydrogen bonds and therefore enhance binding (McPhee *et al.*, 1996). We did mutations for L23Y alone and in combination with D25K and L23Y, i.e. 25K23Y. We can see in both cases, they do have more favorable electrostatic interaction energies, −394.8 kcal/mol in 25K23Y and −348.3 kcal/mol in 23Y *versus* −344.7 kcal/mol in wild-type. However, they also have larger solvation penalties compared with wild-type, +518.4 kcal/mol in 25K23Y and +472.2 kcal/mol *versus* +467.4 kcal/mol in wild-type. Therefore, the total electrostatic contribution to the binding free energy, $\Delta G^{ele}_{int+sol}$, is more unfavorable than the wild-type dimer and so are the binding free energies.

We investigated another two double mutations, 25K5F and 25K87W as well. Leu5 and Arg87 have large SCNANR, 64.7% and 36.8%, respectively, and they are not close to termini. The TNAN values in 3 Å distance cutoff are 31 and 29, respectively, which means they are crowded. We can see in Table 7 that both of them have less favorable van der Waals interaction energies than wild-type, i.e. −186.8 kcal/mol in 25K5F and −187.4 kcal/mol in 25K87W *versus* −189.8 kcal/mol in the wild-type dimer. In addition, their $\Delta G^{ele}_{int+sol}$ terms are also more unfavorable. This is due to larger solvation penalty. Arg87 is exposed to water. The solvation energy is unfavorable if this charged residue is mutated to a neutral one. In 25K5F, the aromatic ring of Phe5 is totally buried which is also unfavorable.

Conclusion

Here, we investigated the protonation state of the free HIV PR and calculated the relative binding free energies of different dimers of HIV PR to wild-type dimer using the MM/PBSA method. We also suggest several dominant negative inhibition mutations on the basis of binding free energies calculated using the VM method.

We compared the average distances between several pair atoms in the two catalytic Asp residues obtained from the MD trajectory with those in the crystal structure. These calculations indicated that the dianionic state had the closest structure to the crystal structure. According to the binding free energy calculations on the wild-type dimers with different protonation states, the dianionic dimer structure was also suggested to be most stable. These results are consistent with NMR data on the ligand free HIV protease (Smith *et al.*, 1996).

The heterodimer formed between the triple mutation monomer, Asp25Lys/Gly49Trp/Iles50Trp (25K49W50W), and the wild-type monomer was shown experimentally to have higher thermal stability than the wild-type dimer (Rozzelle *et al.*, 2000). We calculated the binding free energies on the 25K49W50W bound to wild-type monomer with deprotonated and protonated catalytic Asp. Both of these heterodimers are more stable than the wild-type homodimers. The homodimer of the triple mutations is the least stable dimer. The ranking order of dimer stability is consistent with the experimental observations.

The Virtual Mutagenesis (VM) method was developed to identify mutations which can enhance binding between two subunits of a macromolecule. With the assumption that local mutations will not change the overall structure of a protein, this method was used to calculate binding free energies for different HIV protease dimers. The ranking order of calculated binding free energies is consistent with that of viral infectivity. Moreover, several new dominant negative mutations were suggested by this method. Four of them, 25K48W, 25K49W, 25K50W and 25K52W, are similar to the triple mutations in terms of mutated residues. However, another two, 25K4Y and 25K7W, are novel mutations and are not obvious choices if one just visualizes the HIV PR structure. These results await experimental verification. Because these mutations are not close to each other in space (except 48w, 49w, 50w and 52w), we speculate that the multiple mutations may have stronger dominant negative inhibition effects.

In summary, the MM/PBSA method is able to calculate binding free energies on systems for which more rigorous methods such as free energy perturbation (FEP) and thermodynamic integration (TI) can not be efficiently applied. The VM method can quickly identify residues on which mutations can be made to enhance the binding between protein-protein or protein-ligand. Caveats include the assumptions of similar entropy change upon dimerization of different mutants, additivity of free energy terms, and adequacy of sampling of conformation space. With these caveats in mind, the results obtained using the MM/PBSA and the VM methods are promising and worthy of further development and experimental testing.

Methods

Protonation state of the HIV PR and the MM/PBSA method

All molecular dynamics simulations presented in this work were performed using the AMBER5.0 simulation package (Pearlman *et al*., 1995) and the Cornell *et al*. (1995) force field with TIP3P water model (Jorgensen *et al*., 1983). The starting structure for the wild-type homodimer of the HIV protease was taken from the RCSB Protein Data Bank. The PDB entry is 3hvp (Wlodawer *et al*., 1989). Mutations were made manually using SYBYL6.5 (Tripos Associates Inc., 1998) and MidasPlus (Ferrin *et al*., 1988). The molecules were solvated in a 80 Å × 80 Å × 80 Å box of water. All systems were neutralized by adding counter ions close to the solute surface. The number of counter ions varied with different HIV PR dimer. Particle Mesh Ewald (PME) (Darden *et al*., 1993) was exploited to consider the long-range electrostatic interactions. All structures were minimized first using SANDER module in AMBER5.0. Molecular dynamics simulations were carried out thereafter. The temperature of the system was raised gradually from 50 K to 298 K in 50 ps followed by 120 ps equilibration at 298 K. Another 120 ps MD simulation was performed for data collection and 100 snapshots were saved for the consequent analysis. The SHAKE procedure (Rychaert *et al*., 1977) was employed to constrain all bonds. The time step of the simulations was 2 fs. A 8.5 Å cut-off was used for the non-bonded van der Waals interactions and no cutoff was used for nonbonded electrostatic interactions. The non-bonded pairs were updated every 15 steps.

The binding free energy between the two monomers of the HIV PR was calculated according to the thermodynamic cycle shown in Figure 2:

$$\Delta G_b = \Delta G_b^0 + \Delta G_{sol}^D - \Delta G_{sol}^{M1} - \Delta G_{sol}^{M2} \quad (1)$$

where G_b^0 and ΔG_b are the binding free energies in gas and in water, respectively, ΔG_{sol}^{M1}, ΔG_{sol}^{M2} and ΔG_{sol}^{D} are solvation free energies for the monomer 1, monomer 2 and dimer of the HIV PR, respectively. G_b^0 is calculated from molecular mechanics (MM) interaction energies:

$$\Delta G_b^0 = \Delta G_{int}^{elc} + \Delta G_{int}^{vdw} \quad (2)$$

where ΔG_{int}^{ele} and ΔG_{int}^{vdw} are electrostatic and van der Waals interaction energies between the two monomers in gas which were calculated using the CARNAL and ANAL modules in AMBER5.0 software suite.

The solvation energy, ΔG_{sol}, is divided into two parts, the electrostatic contributions, ΔG_{sol}^{ele}, and all other contributions, $\Delta G_{sol}^{nonpolar}$.

$$\Delta G_{sol} = \Delta G_{sol}^{ele} + \Delta G_{sol}^{nonpolar} \quad (3)$$

The electrostatic contribution to the solvation free energy, ΔG_{sol}^{ele}, was calculated using the DelPhiII software package (Gilson *et al*., 1987), which solves the Poisson-Boltzmann equations numerically and calculates the electrostatic energy according to the electrostatic potential. The grid size we used was 0.5 Å. Potentials at the boundaries of the finite-difference lattice were set to sum of the Debye-Huckel potentials. The interior dielectric constant was set to 1 in our primary simulations in order to be consistent with the molecular mechanics force field. Other value for the interior dielectric constant was also examined (see below). The dielectric constant of water was set to 80 and the dielectric boundary was taken as the solvent accessible surface defined by a 1.4 Å probe sphere. The radii of atoms were taken from PARSE parameter set (Sitkoff *et al*., 1994). Partial charges were taken from Cornell *et al*. force field for standard amino acid residues. One non-standard amino acid in the 3hvp was ABA and its partial charges were calculated using *ab initio* calculations and the RESP method (Bayly *et al*., 1993).

As mentioned above, the value 1 was first used for the interior dielectric constant originally in MM/PBSA. Since the dielectric constants for the interior of proteins is considered to be in the range from 2 to 4, we examined the case where the interior dielectric constant had values other than 1. As shown in the Appendix, the binding free energy was calculated slightly different from equation (1):

$$\begin{aligned}\Delta G_b &= \Delta G_{int}^{vdw} + (\Delta G_{sol}^{nonpolar}{}_{D} - \Delta G_{sol}^{nonpolar}{}_{M1} - \Delta G_{sol}^{nonpolar}{}_{M2}) \\ &\quad + (1/n)\Delta G_{int}^{ele} + (\Delta G_{RFE}{}^{D}_{n-80} - \Delta G_{RFE}{}^{M1}_{n-80}\Delta G_{RFE}{}^{M2}_{n-80} \\ &= \Delta G_{int}^{vdw} + \Delta G_{sol}^{nonpolar} + (1/n)\Delta G_{int}^{ele} + \\ &\quad (\Delta G_{RFE}{}^{D}_{n-80} - \Delta G_{RFE}{}^{M1}_{n-80} - \Delta G_{RFE}{}^{M2}_{n-80}) \quad (4)\end{aligned}$$

where n is the interior dielectric constant. In this study, it was set to 2. $\Delta G_{RFE\ n\text{-}80}^{D}$, $\Delta G_{RFE\ n\text{-}80}^{M1}$ and $\Delta G_{RFE\ n\text{-}80}^{M2}$ are reaction field energies obtained from DelPhi for dimer, monomer 1 and monomer 2 of the HIV PR, respectively, with interior and exterior dielectric constants set to n and 80, respectively.

The solvent accessible surfaces (SAS) were calculated using the MSMS program (Sanner *et al*., 1996). The nonpolar contribution to the solvation free energy, $\Delta G_{sol}^{nonpolar}$, was calculated as 0.00542 × SAS + 0.92 kcal/mol (Sitkoff *et al*., 1994).

It is worth pointing out that in equation (1), no solute entropy contribution is included. We estimated the conformational entropy contribution (translation, rotation and vibration) to the binding free energy using normal mode analysis (Case, 1994). This is only an estimate for the order of magnitude of the entropy contribution. We assumed that the entropy contributions are similar for different HIV protease dimers. When we calculate the relative binding free energies between them, the entropy contribution is assumed to cancel. The normal mode analysis was carried out using the NMODE module in AMBER5.0. The structure used for normal mode analysis was obtained by minimizing the crystal structure of the wild-type HIV PR dimer using a distance dependent dielectric constant which is proportional to 4r, where r is the distance between the atoms.

Virtual mutagenesis method

Mutations which might enhance binding between the two monomers of the HIV PR can only be made on residues which satisfy the following three qualitative criteria.

(1) The vector from C^α to C^β, $\mathbf{n}_{\alpha\beta}$, points toward the dimer interface.

(2) The residue is close to the dimer interface.

(3) The residue has some extra space around it and a number of atoms in another monomer are a short distance from this residue.

The idea is that the mutation will not change the HIV PR structure dramatically (criterion (1)) and more favorable contacts with another monomer can be created if this residue is mutated to another residue with a larger side chain (criteria (2) and (3)).

First, in order to identify those residues satisfying these criteria in the HIV PR, vector **n**, which was perpendicular to the plane defined by three atoms in the B chain of the HIV PR, N in Gly49B, C^{α} in Asn98B and C^{α} in Arg8B, was constructed. The plane defined by the three atoms was parallel to the dimer interface. $\mathbf{n}_{\alpha\beta}$ for each residue in the A chain of the HIV PR was also calculated. By examining the sign of $\mathbf{n} \times n_{\alpha\beta}$ we could determine whether $\mathbf{n}_{\alpha\beta}$ pointed to or away from the interface.

Second, TNAN and CNAN values were calculated for two different distance cutoffs. The TNAN is the number of atoms within the distance cutoff, r_{thr}, of any atom of the residue being investigated. CNAN is the number of atoms in another subunit of the molecule within the distance cutoff r_{thr}. The value of TNAN reflects how crowded the residue being investigated is surrounded by other residues and the value of CNAN represents how many contacts this residue has with another subunit of the molecule. Obviously, the values of TNAN and CNAN depend on the distance cutoff, r_{thr}. In the present study, we used two distance cutoffs, 3 Å and 6 Å. A ratio, SCNANR was calculated as:

$$\mathrm{SCNANR} = (\mathrm{CNAN}_2 - \mathrm{CNAN}_1)/(\mathrm{TNAN}_2 - \mathrm{TNAN}_1) \quad (5)$$

where $CNAN_1$ and $CNAN_2$ are CNAN using 3 Å and 6 Å distance cutoff, respectively, and so are $TNAN_1$ and $TNAN_2$ for TNAN. This ratio, SCNANR, reflects how many contacts may form if the current side-chain is replaced by a larger one.

A residue which satisfies the second criterion must have a relatively large CNAN value at least at the 6 Å distance cutoff if not already at the 3 Å distance cutoff. The third criterion requires residues which have small TNAN values in the 3 Å distance cutoff range and large SCNANR values.

Glycine residues were specifically considered. In addition to the three criteria, ϕ and φ torsion angles for each glycine residue were also examined. Only those whose ϕ and φ, were not unique for glycine residues, i.e. were in the left half of the Ramachandran map, were considered for mutation.

Any residues satisfying the above three criteria were mutated to one of those amino acid residues which has larger side-chain, such as Trp, Tyr. A systematic conformation search for 100 conformations was carried out for the mutant residue. For each conformation, a steric bump check was executed first to avoid serious steric clash between the mutant residue and any other residue in the molecule. If any atom of the built-in residue was closer than 1 Å to any atom of any other residue in the molecule, that conformation of the built-in residue was discarded. Each surviving conformation was then minimized using the SANDER module in AMBER5.0. Only the mutant residue was allowed to move and no explicit water was added. The solvent effect was considered roughly by using a distance dependent dielectric constant which was proportional to $4r$, where r is the distance between atoms. MM/PBSA was used to evaluate the energy of each conformation. The final energy for each mutation was the average energy of all conformations. Multiple mutations were made one by one. For example, the triple mutation D25 K/G49W/I50W was obtained by mutating Asp25 to Lys first. A conformation of Lys25 with closest binding free energy to the average value was chosen and Gly49 was mutated to Trp. The third mutation, I50W, was made on the conformation whose binding free energy is closest to the average one for the D25 K/I49W mutations. Since no full MD simulations were carried out, the computational efficiency is quite high using such an approach.

Acknowledgments

This work was supported in part by grants GM-56531 (P. Ortiz de Montellano, P.I.) and GM-56609 (E. Arnold, P.I.). Part of this investigation was completed using the facilities in the UCSF Computer Graphics Lab, T. Ferrin, director, supported by RR-1081 from the NIH. W.W. is grateful for kind help and stimulating discussions from Drs Yong Duan, Bemd Kuhn and Taisung Lee. We gratefully acknowledge supercomputer time at NCSA, University of Illinois, supported by the National Science Foundation

References

Ajay, & Murcko, M. A. (1995). Computational methods to predict binding free energy in ligand-receptor complexes. *J. Med. Chem.* **38**, 4953-4967.

Babé, L. M., Rose, J. R. & Craik, C. S. (1995). Transdominant inhibitory human immunodeficiency virus type I protease monomers prevent protease activation and virion maturation. *Proc. Natl Acad. Sci. USA,* **92**, 10069-10073.

Bayly, C. I., Cieplak, P., Cornell, W. D. & Kollman, P. A. (1993). A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges - the RESP model. *J. Phys. Chem.* **97**, 10269-10280.

Beveridge, D. L. & Dicapua., F. M. (1989). Free energy *via* molecular simulations: application to chemical and biochemical system. *Annu. Rev. Biophys. Biophys. Chem.* **18**, 431-492.

Case, D. A. (1994). Normal mode analysis of protein dynamics. *Curr Opin. Struct. Biol.* **4**, 285-290.

Cheng, Y. S. F., Yin, F. H., Foundling, S., Blomstrom, D. & Kettner, C. A. (1990). Stability and activity of human immunodeficiency virus protease - comparison of the natural dimer with a homologous, single-chain tethered dimer. *Proc. Natl Acad. Sci. USA,* **87**, 9660-9664.

Collins, J. R., Burt, S. K. & Erickson, J. W. (1995). Flap opening in HIV-1 protease simulated by ''activated'' molecular dynamics. *Nature Struct. Biol.* **2**, 334-338.

Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I., Merz, K. M., Ferguson, D., Spellmeyer, D. C., Fox, T., Caldwell, J. W. & Kollman, P. A. (1995). A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **117**, 5179-5197.

Darden, T., York, D. M. & Pedersen, L. (1993). Particle Mesh Ewald - an Nlog(N) method for Ewald sums in large systems. *J Chem. Phys.* **98**, 10089-10092.

Debouck, C., Gorniak, J. G., Strickler, J. E., Meek, T. D., Metcalf, B. W. & Rosenberg, M. (1987). Human immunodeficiency virus protease expressed in *Escherichia coli* exhibits autoprocessing and specific

maturation of the gag precursor. *Proc. Natl Acad. Sci. USA,* **84**, 8903-8906.

DiIanni, C. L., Davis, L. J., Holloway, M. K., Herber, W. K., Darke, P. L., Kohl, N. E. & Dixon, R. A. (1990). Characterization of an active single polypeptide form of the human immunodeficiency virus type-I protease. *J. Biol. Chem.* **265**, 17348-17354.

Ferrin, T. E., Huang, C. C., Jarvis, L. E. & Langridge, R. (1988). The MIDAS display system. *J. Mol. Graph.* **6**, 13-27.

Gilson, M. K., Sharp, K. A. & Honig, B. H. (1987). Calculating electrostatic interactions in biomolecules: method and error assessment. *J. Comput. Chem.* **9**, 327-335.

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

Hyland, L. J., Tomaszek, T. A., Jr, Roberts, D. G., Carr, S. A., Magaard, V. W., Bryan, H. L., *et al*. (1991). Human immunodeficiency virus-1 protease. 2. Use of pH rate studies and solvent kinetic isotope effects to elucidate details of chemical mechanism. *Biochemistry,* **30**, 8454-8463.

Ido, E., Han, H. P., Kezdy, F. J. & Tang, J. (1991). Kinetic studies of human immunodeficiency virus type I protease and its active site hydrogen bond mutant A285. *J. Biol. Chem.* **266**, 24359-24366.

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

Jordan, S. P., Zugay, J., Darke, P. L. & Kuo, L. C. (1992). Activity and dimerization of human immunodeficiency virus protease as a function of solvent composition and erzyme concentration. *J. Biol. Chem.* **267**, 20028-20032.

Jorgensen, W. L., Chandrasekhar, J., Madura, J., Impey, R. W. & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926-935.

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

Kollman, P. A. (1993). Free energy calculations: applications to chemical and biochemical phenomena. *Chem. Rev.* **93**, 2395-2417.

Krausslich, H. G. (1991). Human immunodeficiency virus proteinase dimer as component of the viral polyprotein prevents particle assembly and viral infectivity. *Proc. Natl Acad. Sci. USA,* **88**, 3213-3217.

Luo, R., Head, M. S., Moult, J. & Gilson, M. K. (1998). pK(a) shifts in small molecules and HIV protease: Electrostatics and conformation. *J. Am. Chem. Soc.* **120**, 6138-6146.

Massova, I. & Kollman, P. A. (1999). Computational alanine scanning to probe protein-protein interactions: a novel approach to evaluate binding free energies. *J. Am. Chem. Soc.* **121**, 8133-8143.

McPhee, F., Good, A. C., Kuntz, I. D. & Craik, C. S. (1996). Engineering human immunodeficiency virus 1 protease heterodimers as macromolecular inhibitors of viral maturation. *Proc. Natl Acad. Sci. USA,* **93**, 11477-11481.

Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham, T. E., Debolt, S., Ferguson, D., Seibel, G. & Kollman, P. A. (1995). AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comp. Phys. Comm.* **91**, 1-41.

Rick, S. W., Erickson, J. W. & Burt, S. K. (1998). Reaction path and free energy calculations of the transition between alternate conformations of HIV-1 protease. *Proteins: Struct. Funct. Genet.* **32**, 7-16.

Rozzelle, J. E., Dauber, D. S., Todd, S., Kelly, R. & Craik, C. S. (2000). Macromolecular inhibitors of HIV-1 protease: characterization of designed heterodimers. *J. Biol. Chem.* **275**, 7080-7086.

Rychaert, J. P., Ciccotti, G. & Berendsen, H. J. C. (1977). Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327-341.

Sanner, M. F., Olson, A. J. & Spehner, J. (1996). Reduced surface - an efficient way to compute molecular surfaces. *Biopolymers,* **38**, 305-320.

Sitkoff, D., Sharp, K. A. & Honig, B. (1994). Accurate calculation of hydration free energies using macroscopic solvent models. *J. Phys. Chem.* **98**, 1978-1988.

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

Srinivasan, J., Cheatham, T. E., III, Cieplak, P., Kollman, P. A. & Case, D. A. (1998). Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate - DNA helices. *J. Am. Chem. Soc.* **120**, 9401-9409.

Trylska, J., Antosiewicz, J., Geller, M., Hodge, C. N., Klabe, R. M., Head, M. S. & Gilson, M. K. (1999). Thermodynamic linkage between the binding of protons and inhibitors to HIV-1 protease. *Protein Sci.* **8**, 180-195.

van Gunsteren, W. F. (1989). Methods for calculation of free energies and binding constants: successes and problems. In *Computer Simulation of Biomolecular Systems* (van Gunsteren, W. F. & Weiner, P. K., eds), pp. 27-59, ESCOM, Leiden, Germany.

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

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

Yamazaki, T., Nicholson, L. K., Torchia, D. A., Wingfield, P., Stahl, S. J., *et al*. (1994). NMR and X-ray evidence that the HIV protease catalytic aspartyl groups are protonated in the complex formed by the protease and a non-peptide cyclic urea-based inhibitor. *J. Am. Chem. Soc.* **116**, 10791-10972.

Zhang, Z. Y., Poorman, R. A., Maggiora, L. L., Heinrikson, R. L. & 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.,* 15591-15594.

Zou, X. Q., Sun, Y. X. & Kuntz, I. D. (1999). Inclusion of solvation in ligand binding free energy calculations using the generalized Born model. *J. Am. Chem. Soc.* **121**, 8033-8043.

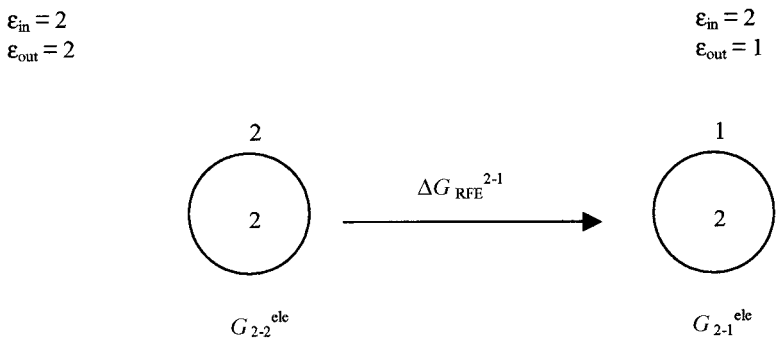

Figure A1. How the reaction field energy is calculated in DelPhi.

Appendix

In DelPhi, the reaction field energy of a molecule is defined as the energy of taking the molecule from a solvent of dielectric equal to that of the interior, to that of the exterior under the condition that there is no salt present and the molecule lies entirely within the box (see DelPhi manual). For example, if the interior dielectric constant ε_{in} equals 2 and the exterior dielectric constant ε_{out} equals 1, the reaction field energy G_{RFE}^{2-1} is calculated as the difference between electrostatic energies in ($\varepsilon_{in} = 2$, $\varepsilon_{out} = 1$) and ($\varepsilon_{in} = 2$, $\varepsilon_{out} = 2$) environments (Figure A1):

$$\Delta G_{RFE}^{2-1} = G_{2-1}^{ele} - G_{2-2}^{ele} \quad \text{(A1)}$$

where G_{2-1}^{ele} and G_{2-2}^{ele} are the electrostatic energies in the ($\varepsilon_{in} = 2$, $\varepsilon_{out} = 1$) and ($\varepsilon_{in} = 2$, $\varepsilon_{out} = 2$) environments, respectively.

Thus:

$$G_{2-1}^{ele} = G_{2-2}^{ele} + \Delta G_{RFE}^{2-1} = (1/2) \times G_{1-1}^{ele} + \Delta G_{RFE}^{2-1} \quad \text{(A2)}$$

where G_{1-1}^{ele} is the electrostatic energy in gas. Therefore, G_{2-2}^{ele} equals half of G_{1-1}^{ele}.

The binding free energy, ΔG_b^0, of the HIV protease dimer in the ($\varepsilon_{in} = 2$, $\varepsilon_{out} = 1$) environment is calculated as:

$$\Delta G_b^0 = \Delta G_{int}^{vdw} + G_{2-1\ D}^{ele} - G_{2-1\ M1}^{ele} - G_{2-1\ M2}^{ele} \quad \text{(A3)}$$

where ΔG_{int}^{vdw} is the van der Waals interaction energy between the two monomers, $G_{2\text{-}1\ M1}^{ele}$, $G_{2\text{-}1\ M2}^{ele}$, and $G_{2\text{-}1\ D}^{ele}$ are electrostatic energies of monomer 1, monomer 2 and the dimer respectively. Substitute equation (A2) into equation (A3), we get:

$$\begin{aligned}\Delta G_b^0 =& \Delta G_{int}^{vdw} + (1/2) \times (G_{1-1D}^{ele} - G_{1-1\ M1}^{ele} - G_{1-1\ M2}^{ele} \\ &+ (\Delta G_{RFE\ D}^{2-1} - \Delta G_{RFE\ M1}^{2-1} - \Delta G_{RFE\ M2}^{2-1} \\ &= \Delta G_{int}^{vdw} + (1/2) \times \Delta G_{int}^{ele} \\ &+ (\Delta G_{REF\ D}^{2-1} - \Delta G_{RFE\ M1}^{2-1} - \Delta G_{RFE\ M2}^{2-1} \quad \text{(A4)}\end{aligned}$$

where ΔG_{int}^{ele} is the electrostatic interaction energy between the two monomers.

If one wants to calculate the binding free energy ΔG_b of the HIV PR dimer in water, one has to calculate the solvation energies ΔG_{sol}^{M1}, ΔG_{sol}^{M2} and ΔG_{sol}^{D} for the monomer 1, monomer 2 and dimer of

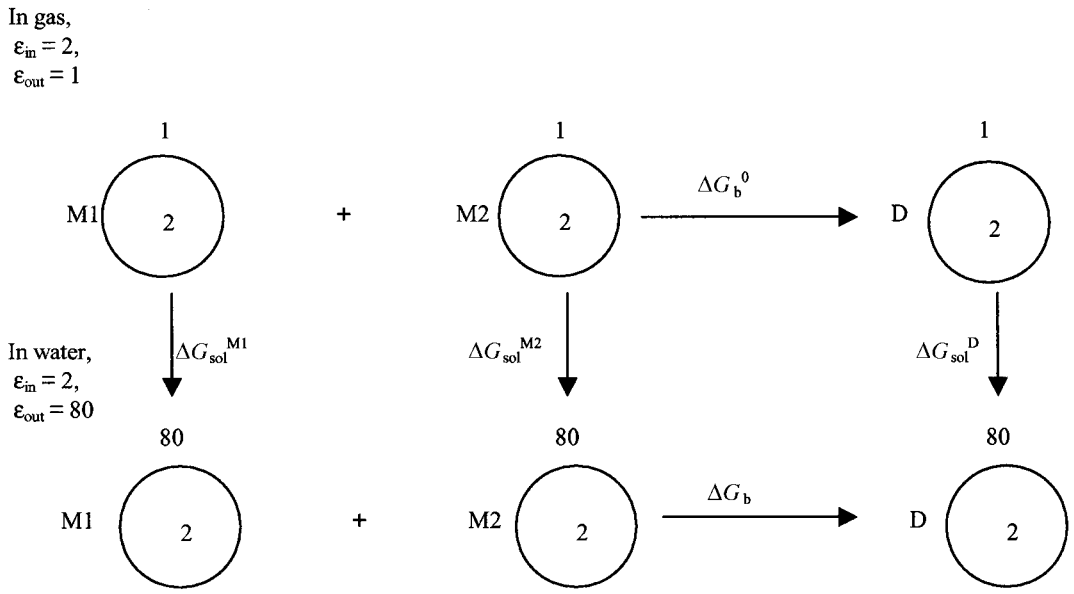

Figure A2. Thermodynamic cycle for calculating the binding free energy of the HIV PR dimer.

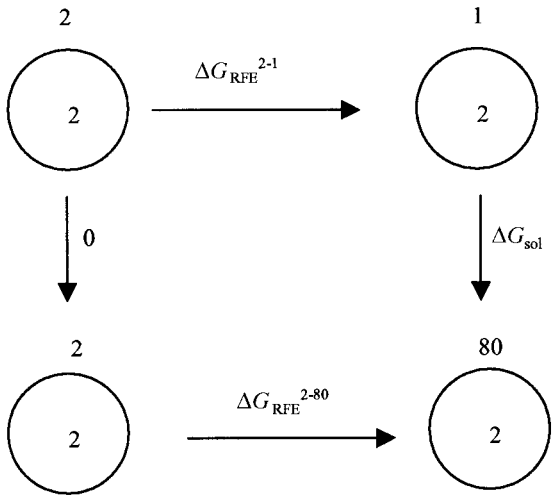

Figure A3. Thermodynamic cycle for calculating electrostatic interaction contribution to solvation free energy.

the HIV PR (see Figure A2), respectivel:

$$\Delta G_b = \Delta G_b^0 + \Delta G_{sol}^D - \Delta G_{sol}^{M1} - \Delta G_{sol}^{M2} \quad (A5)$$

The solvation energy can be decomposed to two parts, electrostatic contribution ΔG_{sol}^{ele} and all other contributions $\Delta G_{sol}^{nonpolar}$:

$$\Delta G_{sol} = \Delta G_{sol}^{ele} + \Delta G_{sol}^{nonpolar} \quad (A6)$$

According to the thermodynamic cycle shown in Figure A3, the electrostatic solvation energy ΔG_{sol}^{ele} of taking a molecule from gas ($\varepsilon_{out} = 1$) to water ($\varepsilon_{out} = 80$) is:

$$\Delta G_{sol}^{ele} = \Delta G_{RFE}^{2-80} - \Delta G_{RFE}^{2-1} \quad (7a)$$

Substitute equation (A6) and (A7) into equation (A5), we get:

$$\begin{aligned}\Delta G_b = \Delta G_b^0 &+ (\Delta G_{sol\ D}^{nonpolar} - \Delta G_{sol\ M1}^{nonpolar} - \Delta G_{sol\ M2}^{nonpolar}) \\ &+ (\Delta G_{RFE\ D}^{2-80} - \Delta G_{RFE\ M1}^{2-80} - \Delta G_{RFE\ M2}^{2-80}) \\ &- (\Delta G_{RFE\ D}^{2-1} - \Delta G_{RFE\ M1}^{2-1} - \Delta G_{RFE\ M2}^{2-1})\end{aligned} \quad (A8)$$

Substitute equation (A4) into equation (A8), we get the formula to calculate binding free energy of the HIV PR dimer whose interior dielectric constant ε_{in} equals 2.

$$\begin{aligned}\Delta G_b = \Delta G_{int}^{vdw} &+ (\Delta G_{sol\ D}^{nonpolar} - \Delta G_{sol\ M1}^{nonpolar} \\ &- \Delta G_{sol\ M2}^{nonpolar}) + (1/2) \times (G_{1-1\ D}^{ele} - G_{1-1\ M2}^{ele}) \\ &+ (\Delta G_{RFE\ D}^{2-80} - \Delta G_{RFE\ M1}^{2-80} - \Delta G_{RFE\ M2}^{2-80}) \\ &= \Delta G_{int}^{vdw} + \Delta G_{sol}^{nonpolar} + (1/2) \times \Delta G_{int}^{ele} \\ &+ (\Delta G_{RFE\ D}^{2-80} - \Delta G_{RFE\ M1}^{2-80} - \Delta G_{RFE\ M2}^{2-80})\end{aligned} \quad (A9)$$

where

$$\Delta G_{sol}^{nonpolar} = \Delta G_{sol\ D}^{nonpolar} - \Delta G_{sol\ M1}^{nonpolar} - \Delta G_{sol\ M2}^{nonpolar}.$$

It is easy to generalize the above derivation to the case where the interior dielectric constant value equals n for any ligand-protein system. If ε_{in} equals n, equation (A9) becomes:

$$\begin{aligned}\Delta G_b = \Delta G_{int}^{vdw} &+ (\Delta G_{sol\ LP}^{nonpolar} - \Delta G_{sol\ L}^{nonpolar} - \Delta G_{sol\ P}^{nonpolar}) \\ &+ (1/n) \times (G_{1-1\ LP}^{ele} - G_{1-1\ L}^{ele} - G_{1-1\ P}^{ele}) \\ &+ (\Delta G_{RFE\ LP}^{2-80} - \Delta G_{RFE\ L}^{2-80} - \Delta G_{RFE\ P}^{2-80}) \\ &= \Delta G_{int}^{vdw} + \Delta G_{sol}^{nonpolar} + (1/n) \times \Delta G_{int}^{ele} \\ &+ (\Delta G_{RFE\ LP}^{2-80} - \Delta G_{RFE\ L}^{2-80} - \Delta G_{RFE\ P}^{2-80})\end{aligned} \quad (A10)$$

Edited by B. Honig

(Received 29 December 1999; received in revised form 18 April 2000; accepted 21 July 2000)