Molecular Dynamics Study of the Connection between Flap Closing and Binding of Fullerene-Based Inhibitors of the HIV-1 Protease†

Zhongwei Zhu and David I. Schuster

Department of Chemistry, New York University, New York, New York 10003

Mark E. Tuckerman*

Department of Chemistry and Courant Institute of Mathematical Sciences, New York University, New York, New York 10003

Received July 29, 2002

ABSTRACT: The complementary spatial relationship between fullerene C_{60} and the hydrophobic cavity region of the human immunodeficiency virus (HIV) protease, which houses the active site of the enzyme, has led to the suggestion that fullerene-based derivatives could have potential use as effective HIV protease inhibitors. The ability of such compounds to desolvate the cavity region leads to a strong hydrophobic interaction between the C_{60} moiety and residues in the cavity region. In this study, the connection between the motion of the so-called flexible flaps of the cavity and favorable binding of a fullerene-based protease inhibitor is explored using multiple-time scale molecular dynamics simulations and free energy techniques. In addition, the effect of the interaction between the C_{60} moiety and the residues in the cavity region on the water content of the cavity is also investigated. Conformational free energy profiles along a suitably chosen flap opening coordinate show a considerable barrier to flap opening in the presence of the inhibitor, while no such barrier exists for the protease alone. This result is interpreted in terms of a strong hydrophobic interaction between the C_{60} moiety and the flexible flaps, which cause the latter to close tightly around the inhibitor, thereby expelling water from the cavity and leading to a favorable binding interaction. This interpretation is rationalized by direct analysis of the water content in the cavity in the presence and absence of the inhibitor.

It is estimated that approximately 28 million people worldwide have thus far died from AIDS and that, currently, more than 36 million people are living with human immunodeficiency virus (HIV) or AIDS. These statistics underscore the urgent need for robust therapies in the fight against the spread of AIDS. The HIV-1 protease (PR), a small homodimeric enzyme consisting of two 99-residue strands that cleaves the *gag* and *pol* viral polyproteins into the enzymes and structural proteins of the virus (*1*−*8*), has emerged as an attractive target for novel therapies, and many of the currently approved drug regimens are based on inhibition of this enzyme (*8*−*12*). Other therapies are focused on inhibition of reverse transcriptase (see ref *13*) or combination ("cocktail") therapies based on inhibition of both HIV-1 PR and reverse transcriptase (*14*−*16*). The active site of the HIV-1 PR, which belongs to the class of aspartic proteases, is composed of two catalytic aspartic acid residues (Asp 25 and Asp 25′ contributed by the two monomers), which assist the attack of a water molecule on the scissile peptide bond of the polyprotein. The active site is contained within a largely hydrophobic cavity capped by two flexible flaps as shown in Figure 1. Although the precise role of these flaps is unknown, it is generally thought that they are important in facilitating substrate binding and product release (*17*, *18*).

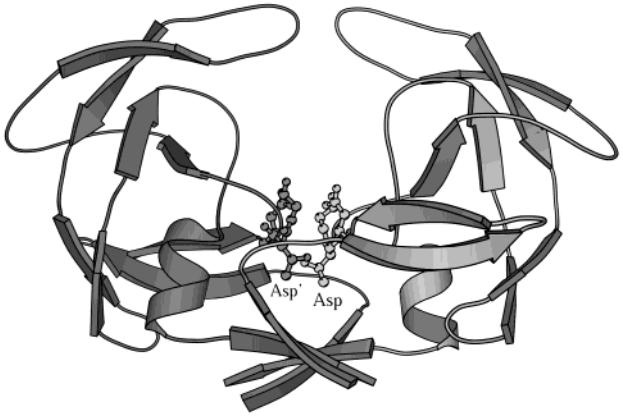

FIGURE 1: Three-dimensional rendering of the HIV-1 protease from the crystal structure of Spinelli et al. (*34*), showing the atomic detail of the catalytic Asp residues.

The hydrophobic cavity of the HIV-1 PR is a roughly open-ended cylinder, a fact that led Wudl, Friedman, and co-workers in 1993 (*19*, *20*) to propose the use of a fullerene-C_{60} derivative as an inhibitor of the HIV-1 PR on the basis of the former's structural complementarity to the cavity region. This complementarity translates into favorable binding to the active site, effective inhibition, and, therefore, potential for therapeutic application. Indeed, many antiviral fullerene drugs show good in vitro antiviral activity and lack of toxicity (at concentrations up to ∼100 μM) (*21*−*25*). Since

† We acknowledge support from NSF Grant CHE-9875824, NSF Grant EIA-0081307, and an NYU White Fellowship in Biomedical and Biological Sciences.

10.1021/bi020496s CCC: $25.00
Published on Web 01/15/2003

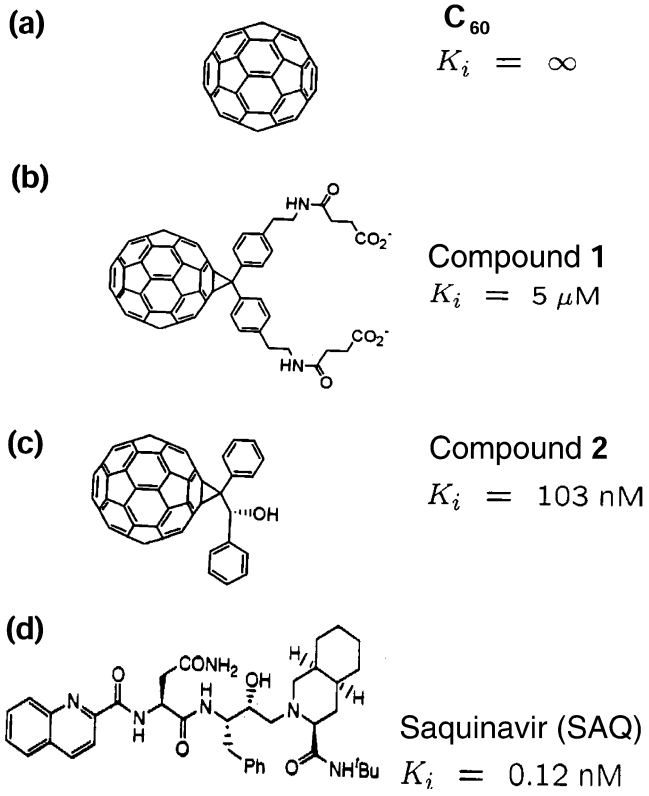

FIGURE 2: Schematic diagrams of (a) C_{60}, (b) compound **1**, (c) compound **2**, and (d) Saquinavir (SAQ). For each compound, the reported inhibition constant, K_i, is specified, where $K_i = [E][I]/[EI]$, where [E] and [I] are the concentrations of the uncomplexed enzyme and the inhibitor in solution, respectively, and [EI] is the concentration of the enzyme−inhibitor complex.

this proposal, a number of C_{60}-based compounds have been proposed and subsequently synthesized (*19*−*26*), and general discussions of the biological applicability of fullerene derivatives have entered the literature (*27*). Two such fullerene-based compounds are depicted in Figure 2 along with C_{60} (Figure 2a). Compound **1** (Figure 2b) (*19*) was found to have an inhibition constant (K_i) value of 53 μM and to desolvate approximately 333−352 Å^2 of hydrophobic surface area in the active site region (*19*), while compound **2** (Figure 2c) (*28*, *23*) was found to have a K_i value of 103 nM and to desolvate approximately 472 Å^2 of hydrophobic surface area (*23*). Both compounds were found to be sufficiently soluble to be assayed in the presence of 1% dimethyl sulfoxide solvent (*19*, *23*). It should be noted that compound **2** is actually synthesized as a mixture of the compound and its enantiomer (*23*). By comparison, C_{60} alone is capable of desolvating only ∼298 Å^2 of surface area. Friedman et al. (*19*, *20*, *23*) have proposed that a direct correlation exists between the activity of a compound and the amount of hydrophobic surface area it is capable of desolvating. Clearly, if this notion is correct, it should be associated with (and enhanced by) specific conformational rearrangements within the cavity region that create a "tight fit" around the spherical C_{60} moiety, thereby strengthening hydrophobic interactions between it and the active site residues.

In this paper, we employ molecular dynamics (MD) calculations as a means of investigating this hypothesis. In particular, we focus on the role played by the flexible flaps in facilitating the binding of an inhibitor in the active site region. Using free energy techniques, it will be shown that a barrier to flap opening exists in the presence of an inhibitor that is not found for the HIV-1 PR alone. To our knowledge, this is the first microscopic investigation of the connection between the flap motion and the binding of a compound to the active site of the protease. In addition, we investigate the effect on the free energy profile of the protonation state of the catalytic Asp residues. Finally, we show that the distribution of water molecules in the HIV PR cavity is strongly affected by the presence of a C_{60}-based inhibitor, a fact that supports the computed flap opening free energy profiles.

The MD simulations performed herein employ an all-atom force field and explicit solvent. In addition, a large number of relatively long trajectories is needed to address all of the issues posed above. It is, therefore, crucial that efficient simulation techniques be employed. In the calculations presented here, a novel simulation protocol is employed which is based on a number of state-of-the-art algorithms. These include (i) multiple-time step integration (*29*, *30*), which allows for a large gain in the MD time step; (ii) an efficient canonical dynamics approach (*31*), which enhances local fluctuations; and (iii) fast electrostatics (*32*) (details of the scheme will be presented in Methods).

This paper is organized as follows. A description of our system setup and simulation protocol, including the multiple-time scale scheme mentioned above, is given in Methods. The results and discussion are presented next. Specifically, conformational free energy profiles associated with the flexible flap motion are discussed in Free Energy Profiles for Flap Opening and Protonation Effects, average structural properties of the cavity are discussed in Cavity Structure, and an analysis of the water content in the cavity is presented in Water Content in the Active Site. Conclusions and future directions are presented in the final section.

METHODS

The initial structure for the HIV-1 PR was based on the 2.7 Å resolution X-ray data of Spinelli et al. (*34*) that are available in the Brookhaven Protein Data Bank (PDB entry 1hhp). The program InsightII (Molecular Simulations, Inc.) was used to add hydrogens to the X-ray structure; in addition, the docking module of InsightII was used to create an initial configuration of compound **2** docked into the active site. Since compound **2** is experimentally synthesized as a mixture with its enantiomer, we choose here the conformer that was shown by Friedman et al. (*23*) to cause optimal hydrophobic desolvation. The HIV-1 PR and HIV-1 PR−**2** (HIV-1 PR with compound **2**) complex were solvated in 50 Å × 50 Å × 75 Å and 70 Å × 70 Å × 70 Å boxes, respectively, subject to periodic boundary conditions. These box sizes were chosen according to the approximate shape of each complex obtained from previous gas-phase simulations (*35*). The boxes contained initially 6000 and 11 232 water molecules, respectively. Water molecules within 1.8 Å of any atom of the protease were deleted from the system, leaving 5068 water molecules remaining for the HIV-1 PR alone and 10 238 water molecules remaining for the HIV-1 PR−**2** complex. Finally, counterions were added to each system to ensure overall charge neutrality.

All MD simulations were performed using the PINY_MD software package developed by Tuckerman, Martyna, and

co-workers (*36*). Interactions were modeled using the fully flexible CHARMM22 force field (*37*) and the TIP3P water model (*38*). In particular, parameters for compound **2** were developed on the basis of the existing CHARMM22 atom types and interaction parameters (see section 1 of the Supporting Information). For comparison of its effect on the cavity structure, a similar model based on the CHARMM22 parameters for the commercial drug Saquinavir (*52*) (Invirase, Ro 31-8959, Hoffman-La Roche, denoted, here, as SAQ; see Figure 2) was constructed as well. Although these models are not comparable to, for example, a mixed ab initio/force field-based approach in which high-level quantum mechanical calculations are used to treat the complexation in the active site region and a force field is used to treat the remainder of the system (so-called quantum mechanical/molecular mechanical or QM/MM methodology), they are computationally efficient and allow long time scales to be accessed. Moreover, we believe our drug models are sufficient for the purposes of this study, which is concerned with the effect of the compound on flap opening. Indeed, despite the crudity of the drug models, we find that the compound generally remains within only a few angstroms of the active site, indicating that the model does allow for a favorable binding interaction.

Bonds between carbon atoms of the C_{60} moiety were treated as rigid using the SHAKE (*39*−*41*) and RATTLE (*42*) algorithms, and the three atoms in each water molecule were treated as a rigid 3×3 constraint group via a direct matrix equation technique. Each system was first equilibrated at a temperature of 300 K for approximately 30 ps at a constant volume and temperature (NVT ensemble) in conjunction with a "massive" Nosé-Hoover chain thermostating scheme (*31*), in which a Nosé-Hoover chain of length 4 was coupled to *each degree of freedom* or each group of atoms involved in a common constraint. This approach has been shown to be highly effective in enhancing the fluctuations in systems containing stiff intramolecular interactions (*43*). Following this initial NVT equilibration phase, each system was further equilibrated at a constant pressure (isotropic NPT ensemble) of 1 atm and a temperature of 300 K for approximately 50 ps using the Martyna−Tobias−Klein equations of motion (*44*) in conjunction with the massive thermostating scheme. Finally, production runs 2 ns in length for both compound **2** and SAQ at 300 K were carried out also using the NVT ensemble approach and massive Nosé-Hoover chain thermostating. In addition, a separate series of runs approximately 150 ps each in length using a reaction coordinate constraint (*45*, *46*) was performed on the HIV-1 PR−**2** complex to calculate the flap opening free energy profiles (see the Results and Discussion section).

In all cases, the equations of motion were integrated using the reversible multiple-time scale approach (r-RESPA) method of Tuckerman, Martyna, and Berne (*29*), wherein intramolecular forces are updated with a time step of 0.5 fs, short-range intermolecular forces with a time step of 3.0 fs, and long-range intermolecular forces with a time step of 6.0 fs. This leads to an overall factor of approximately 8 savings in CPU time over a single-time step simulation with a time step of 0.5 fs. The reversible integration algorithms for r-RESPA in the NVT and NPT ensembles (*30*) were employed. In addition, long-range electrostatic interactions were treated using the smooth particle-mesh Ewald technique (*32*, *33*), which saves a factor of 30 in CPU time over a simulation using the standard Ewald method for these systems. Finally, all hydrogens were assigned a mass of 10 amu, which has no effect on equilibrium properties but leads to an additional factor of 2 savings in CPU time. When all of the CPU savings factors are combined, the overall CPU savings for these calculations, compared to a single-time step simulation using standard Ewald and light hydrogens (and no bond constraints in the enzyme), is 480. It should be noted, however, that this number would be reduced by a factor of 3−5 if *all* bonds are constrained via the SHAKE/RATTLE algorithm.

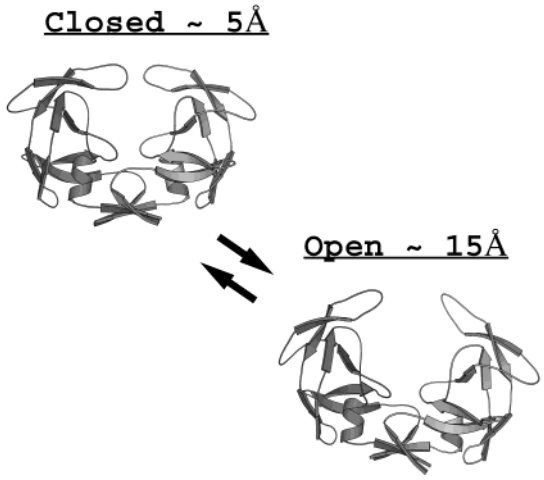

FIGURE 3: Three-dimensional rendering of the HIV-1 protease with the flaps in the closed and open states as they appear in terms of the chosen reaction coordinate (see the text).

RESULTS AND DISCUSSION

Free Energy Profiles for Flap Opening and Protonation Effects. As indicated in prior studies by Collins et al. (*17*) and by Mi et al. (*35*), the opening and closing of the flexible flaps is a rare event on the time scale of the simulations performed here. Therefore, to explore the conformational free energy change associated with the opening of the flaps in the presence and absence of an inhibitor compound in the active site, specialized free energy methodology is required along with additional simulations. In the study presented here, the so-called "bluemoon ensemble" approach (*45*, *46*) was employed in which a reaction coordinate characterizing the flap opening is defined, a constraint is imposed on this coordinate at each of a wide range of values, a simulation is performed at each value, and the free energy profile is obtained by subsequent thermodynamic integration along this coordinate.

To define a useful reaction coordinate, we chose the distance between the two C_α atoms of the Ile (Ile 50 and Ile 50′) residues at the tips of the flaps as a measure of the separation between the flaps. In terms of this coordinate, the "closed" and "open" conformations of the flaps are depicted in Figure 3. The distance constraint is maintained via a Lagrange multiplier (λ), and the free energy profile is, then, constructed via

$$\Delta F(R) = -\int_{R_{\min}}^{R} \mathrm{d}R' \, \langle \lambda \rangle_{R'}^{\mathrm{cond}} - F_0 \qquad (1)$$

where $\Delta F(R)$ denotes the relative Helmholtz free energy as

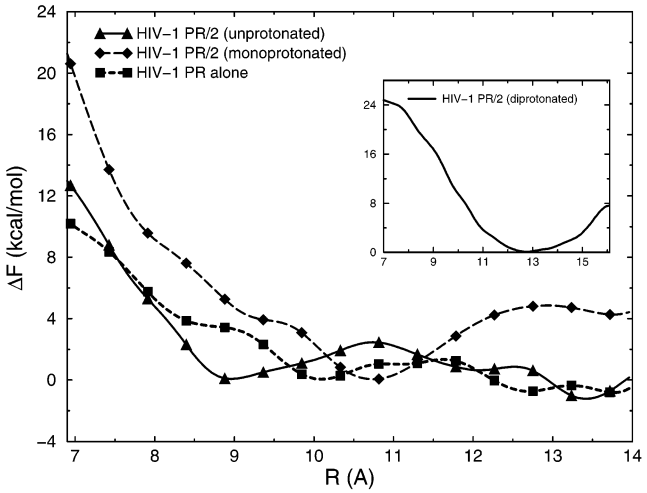

FIGURE 4: Conformational free energy profiles at 300 K corresponding to the flap opening process as functions of the flap separation reaction coordinate (R) defined in the text. The short dashed line with filled squares corresponds to data for the uncomplexed protease alone in water, the solid line with filled triangles to data for the unprotonated HIV-1 PR−**2** complex, and the long dashed line with filled diamonds to data for the monoprotonated HIV-1 PR−**2** complex. The inset shows the free energy profile obtained for the diprotonated HIV-1 PR−**2** complex.

a function of the distance, R, between the two C_α atoms, R_{min} is the minimum value of the reaction coordinate, F_0 is a constant reference free energy, and $\langle\lambda\rangle_{R'}^{cond}$ denotes a conditional ensemble average (*46*) of the Lagrange multiplier with the reaction coordinate fixed at the value R':

$$\langle\lambda\rangle_{R'}^{\mathrm{cond}} = \frac{\int \mathrm{d}^{\mathrm{N}}\mathbf{p}\,\mathrm{d}^{\mathrm{N}}\mathbf{r}\,\lambda(\mathbf{p},\mathbf{r})\mathrm{e}^{-\beta H(\mathbf{p},\mathbf{r})}\delta(R_{\mathrm{C}_\alpha(\mathrm{Ile}\,50)-\mathrm{C}_\alpha(\mathrm{Ile}\,50')} - R')}{\int \mathrm{d}^{\mathrm{N}}\mathbf{p}\,\mathrm{d}^{\mathrm{N}}\mathbf{r}\,\mathrm{e}^{-\beta H(\mathbf{p},\mathbf{r})}\delta(R_{\mathrm{C}_\alpha(\mathrm{Ile}\,50)-\mathrm{C}_\alpha(\mathrm{Ile}\,50')} - R')} \quad (2)$$

where $\mathbf{p}$ and $\mathbf{r}$ denote the momenta and coordinates, respectively, of the N-particle system (protease, compound, and solvent), $H(\mathbf{p},\mathbf{r})$ is the N-particle Hamiltonian, and $\beta = 1/k_BT$. In practice, the ensemble averages are evaluated as weighted averages over the constrained MD trajectory (*46*). [Note that the curvature term discussed by Ciccotti and Sprik (*46*) vanishes for a distance constraint and, therefore, does not appear in eq 1.]

The free energy profiles for flap opening for the HIV-1 PR−**2** complex and for HIV-1 PR alone are shown in Figure 4. Trajectories of 150 ps at each of 13 constraint values using the protocol described in Methods were performed. This corresponds to a total run length of approximately 2 ns for each case. Runs of this length for this system would not have been possible without the efficient methodology outlined in the Results and Discussion. As an independent check, we have reproduced several of these profiles using a recently introduced adiabatic dynamics approach (*47*, *48*). Note that there are two curves for the HIV-1 PR−**2** free energy profile in Figure 4. These represent two different protonation states of the catalytic oxygens in the active Asp residues, namely, unprotonated and monoprotonated. In addition, the inset in Figure 4 shows the free energy curve for the case of diprotonation of the active site. Although the protonation state is not known for the HIV-1 PR−**2** complex, recent plane wave-based ab initio molecular dynamics calculations (*49*) as well as NMR measurements (*50*) of the active site region of the HIV-1 PR complexed with pepstatin suggest that the system is at least monoprotonated; i.e., one of the two catalytic oxygens on an Asp residue is protonated. The connection between these studies and the present investigation lies in the fact that pepstatin, SAQ, and compound **2** all share a common design feature, namely, an OH group capable of forming hydrogen bonds with the catalytic oxygens of the Asp dyad in the active site. Indeed, the ab initio MD calculations strongly favor a diprotonated form for the pepstatin complex; however, given the fact that there are fewer possible hydrogen bonding sites on compound **2**, it is not clear whether the HIV-1 PR−**2** complex would most likely exist in the mono- or diprotonated form. It should be noted that other compounds chemically related to pepstatin tend to stabilize monoprotonated forms; see the discussion in ref *49* and references therein. Taking this result together with the fact that the free protease generally exists in a monoprotonated state (*51*, *53*), we chose to consider all three (unprotonated, monoprotonated, and diprotonated) forms to investigate the trend in the free energy profile as a function of protonation. In the diprotonated form, one of the catalytic oxygens on both Asp 25 and Asp 25′ is protonated.

The free energy curves (see Figure 4) show that the free energy profile for the free protease in water is flat in the region in which $R = 9.5-15$ Å, indicating a wide range of possible flap separation values and, hence, the flexibility of the flaps. In sharp contrast, the free energy profile for the unprotonated complex shows a barrier to flap opening of approximately 3 kcal/mol (for our particular choice of reaction coordinate) when $R \approx 11$ Å and a minimum when $R \approx 9$ Å. Monoprotonation of the active site leads to an even deeper free energy minimum when $R \approx 11$ Å and a barrier of 5 kcal/mol when $R \approx 13$ Å. Finally, diprotonation leads to yet a deeper minimum when $R \approx 12.5$ Å and a barrier of 8 kcal/mol when $R \approx 16$ Å. It should be noted that for R values of $\gtrsim$15 Å, the flaps are considered to be fully "open" (*17*). This shift in the location of the minimum as the protonation state is increased is due to the fact that the proton on one of the catalytic oxygens donates a hydrogen bond to the OH group of compound **2**. This additional "anchor" causes the compound to experience smaller fluctuations about its docked position in the active site. This leads to slight increases in the cavity dimensions and the flap separation as the compound "sits" better in the cavity and causes it to dilate slightly compared to the unprotonated state. Diprotonation further stabilizes the compound in the active site, causing still further dilation of the cavity. This trend is consistent with the observed averaged distances between the compound and the active site. If this distance measure is defined as the average distance between the OH oxygen on compound **2** and the four catalytic oxygens in the active site Asp dyad, it is found that the average of this distance measure over the bluemoon trajectory segment corresponding to the free energy minima in Figure 4 is 8.2 Å for the unprotonated case, 7.4 Å for the monoprotonated case, and 5.2 Å for the diprotonated case. These values indicate that increasing the level of the protonation state leads to more favorable binding.

It is important to note that the actual values of the free energy barrier obtained through the bluemoon ensemble approach are not as relevant as their value *relative* to each other. The reason for this is that the actual barrier value will

depend on the choice of reaction coordinate. Although the chosen distance is a physically meaningful reaction coordinate, it is not the only possible distance choice. Moreover, other reaction coordinates could have been selected, and the details of the free energy profiles would be expected to depend on this choice, although the overall qualitative trend observed here should be the same. Once a coordinate is chosen, however, free energy profiles obtained along this coordinate for different systems relative to each other are meaningful.

The free energy curves in Figure 4 indicate a connection between inhibitor binding and the motion of the flaps. Specifically, they suggest a tight closing of the flaps in the presence of the inhibitor. They also indicate that there is an overall attractive interaction between the C_{60} moiety of compound **2** and the active site residues in the cavity region of the protease that is responsible for this tight closing of the flaps. As will be shown in the next two subsections, the structure of the cavity adjusts to create a "tight fit" to the fullerene compound, which has the effect of "squeezing out" water from the active site region, leading to a strong hydrophobic interaction between the drug and the interior of the cavity, a notion that is consistent with the suggestion that activity is directly connected to hydrophobic desolvation of the cavity.

Cavity Structure. We begin our investigation of the cavity structure by examining the ability of the potential model to describe the active site region. To this end, we report the average values of the distances between the two C_γ atoms in Asp 25 and Asp 25′ as well as the distances between the $O_{\delta 1}$ atoms and $O_{\delta 2}$ atoms in the Asp dyad over the MD run of the protease alone in water. These distances are 6.1, 5.1, and 5.6 Å, respectively. These numbers are in reasonable agreement with those reported in ref *53* and indicate that the model describes the active site accurately enough for the purposes of this study.

Next, to quantify how the shape of the cavity changes upon complexation with fullerene compound **2**, we introduce the following measures of cavity height and width. The cavity height is defined as the distance between C_β of Ile 50 and the catalytic carboxylate oxygen OD2 of Asp 25; the cavity width is taken to be the distance between C_γ of Val 32 and that of Val 32′. These measures are the same as those used in previous gas-phase studies (*35*). The distance measures are averaged over the MD trajectories described in Methods for the HIV-1 PR alone and for the HIV-1 PR−**2** and HIV-1 PR−SAQ complexes. (In this section, all complexes refer to the monoprotonated state of the active site.) In the HIV-1 PR−SAQ complex, a particular crystallographic water (W301) that serves as an anchor for the drug to Ile 50 and Ile 50′ in the flap region was explicitly included in the initial setup; however, it is observed that this water is often exchanged with other waters in the bath. The results obtained for the average cavity dimensions are summarized in Table 1.

The table clearly indicates that, for HIV-1 PR alone, the cavity width exceeds its height considerably, suggesting a more elliptical rather than circular cross section for the cavity. Interestingly, in the HIV-1 PR−SAQ complex, the dimensions experience only very minor changes. In contrast, in the HIV-1 PR−**2** complex, the cavity width and cavity height are much closer in value, suggesting a more circular cross section; i.e., the cavity has adjusted its shape to accommodate the spherical C_{60} moiety of compound **2**. The substantial change in the structure of the cavity upon complexation with compound **2** underscores the importance of performing MD simulations in which all degrees of freedom are allowed to evolve and no restraints are imposed. Since the diameter of C_{60} is approximately 7 Å, the approximate space between C_{60} and the inner surface of the cavity is roughly between 3 and 4 Å all around (accounting for the fact that the atoms used to define the height and width of the cavity are approximately 2 Å from the cavity's inner surface). While this narrow space could, in principle, be occupied by hydrogen-bonded strands of water molecules, the strong hydrophobic nature of the C_{60} moiety and its interaction with the hydrophobic residues lining the surface of the cavity tend to exclude water from this region. This point is investigated in the next subsection.

Table 1: Average Dimensions of the HIV-1 PR Cavity for HIV-1 PR Alone and for Complexes with Compound **2** (HIV-1 PR−**2**) and with Saquinavir (HIV-1 PR−SAQ)

	HIV-1 PR	HIV-1 PR−**2**	HIV-1 PR−SAQ
cavity width (Å)	20.69	19.23	20.60
cavity height (Å)	12.32	17.29	13.18

Table 2: Average Cavity Box Dimensions (depth × width × height), Excluded Volume (V_{excl}), and Number Density (ρ) of Water in the Cavity Region

	box dimensions (Å)	V_{excl} (Å^3)	ρ (Å^{-3})
HIV PR	20.5 × 20.5 × 10.8	0	0.0200
HIV PR−**2** (monoprotonated)	21.4 × 19.0 × 14.9	1232	0.0156
HIV PR−**2** (diprotonated)	17.8 × 19.4 × 14.3	1232	0.0125
HIV PR−SAQ	20.8 × 20.2 × 13.0	841	0.0178

Water Content in the Active Site. An important question that arises in the context of drug binding studies for the HIV PR concerns the distribution of water molecules in the cavity region in the absence and presence of a particular ligand. The exclusion of water from the cavity region in the presence of the C_{60}-based drugs can significantly enhance hydrophobic interactions between the compound and the inner surface of the cavity, particularly in the flap region. Moreover, inhibition of the catalytic mechanism of the protease is most likely a combination of favorable binding of a particular compound and exclusion of water from the active site region. For this reason, we have investigated the water content and distribution of water molecules in the cavity region in the absence and presence of the different inhibitors.

First, we consider the average number density of water molecules in the cavity region. This quantity is obtained by defining an orthorhombic box that fits just inside the cavity. Clearly, the shape of the box depends on the average shape of the cavity. The dimensions of this box are chosen according to the algorithm described in section 2 of the Supporting Information. Table 2 gives the dimensions of the box for each system that has been considered. Note that compound **2** corresponds to the largest box dimensions because of the dilation of the cavity due to the C_{60} moiety. Once the box dimensions are obtained, the volume within the box available to water molecules is computed. In the case of the protease alone, this is the entire box volume; however, in the presence of a drug, the volume excluded by the drug needs to be subtracted. The excluded volumes (V_{excl})

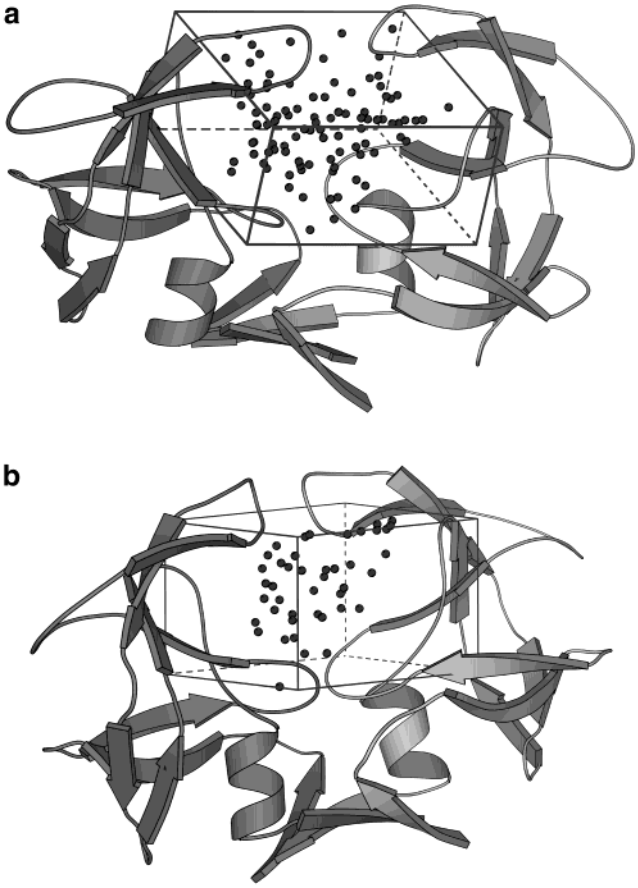

FIGURE 5: Typical snapshot configurations from the MD simulations of the distribution of water molecules in the cavity region of (a) the uncomplexed protease and (b) the HIV-1 PR−**2** complex. The isolated spheres are water oxygens, and hydrogens are not explicitly shown. Also, for easier viewing, the drug is not shown in panel b.

are also reported in Table 2. The number density of water molecules is then calculated as the ratio of the number of water molecules within the box divided by the available volume within the box. The number densities (ρ) thus obtained are reported in Table 2. Note that, here, both the monoprotonated and diprotonated states of the HIV-1 PR−**2** complex are studied. The table shows that, despite the large box volume of compound **2**, the number density (0.0156 Å^{-3} for monoprotonated form and 0.0125 Å^{-3} for the diprotonated form) is lowest for this compound. Each of the number densities for water in the cavity is lower than the number density for bulk water ($\rho = 0.033$ Å^{-3}). To illustrate where the largest water deficit occurs spatially, we show typical configurations from the MD simulation of the water molecules contained within the box in panels a and b of Figure 5. For the HIV-1 PR−**2** complex, it can be seen from the configuration in Figure 5b that, in addition to the exclusion of water in the 3−4 Å region around the C_{60} moiety, water is largely excluded from the region between the C_{60} moiety and the flaps. The exclusion of water from this region allows for a strong hydrophobic interaction between the C_{60} moiety and residues in the flap region, thus enhancing flap closing, consistent with the picture implied by the free energy profiles of Figure 4.

Finally, the distribution of water around the catalytic Asp residues is investigated by computing a radial probability distribution function between the catalytic carboxylate oxygens of the active site Asp residues and the water oxygens. The results are shown in Figure 6. The figure shows that there is a sharp peak at $r = 3$ Å for the protease alone in the water bath. This distance corresponds to the length of a hydrogen bond and indicates that water is fairly structured around these catalytic residues. For the HIV-1 PR−**2** complex, we consider, again, both the monoprotonated and diprotonated (*49*) states of the catalytic Asp oxygens. In the monoprotonated state, there is some reduction in the first peak, but the most dramatic reduction occurs in the diprotonated state, suggesting that it is this state which most effectively excludes water from the region near the active site. Given the crudeness of the drug model, the reduction in the peak must be viewed only as a general trend. We are currently developing a more accurate model to refine this qualitative picture. Although no definite conclusions about the protonation state of the active site can be drawn directly, the results described in this section together with the flap opening free energy profiles in Figure 4 strongly suggest that it is the diprotonated state that leads to the most favorable binding in the HIV-1 PR−**2** complex.

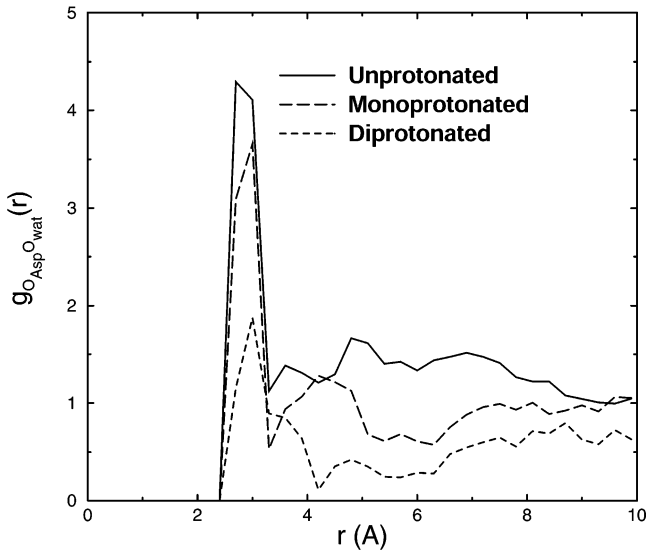

FIGURE 6: Radial distribution functions of the catalytic Asp oxygens with respect to water oxygens computed over the MD trajectories. The solid line corresponds to data for the uncomplexed protease alone in water, the long dashed line to data for the monoprotonated HIV-1 PR−**2** complex, and the short dashed line to data for the diprotonated HIV-1 PR−**2** complex.

CONCLUSIONS

We have carried out molecular dynamics simulations of the HIV-1 PR complexed with a novel C_{60}-based inhibitor compound to investigate the connection between flap motion, binding affinity, and hydrophobic desolvation. The simulations provide the first microscopic level prediction of the tight closing of the flexible flaps when an effective binder is docked into the active site. The simulations also show that significant conformational adjustments occur within the cavity to accommodate the spheroidal C_{60} moiety of the inhibitor compound, thereby creating a relatively "tight fit" between the cavity and the inhibitor. The result is an exclusion of water in the roughly 3−4 Å region between the C_{60} moiety and the residues lining the inner surface of the cavity, particularly near the flap regions and around the

catalytic Asp residues. This exclusion of water leads to an enhancement of the hydrophobic interaction between the C_{60} moiety and the interior of the cavity, including the flaps. Our finding is consistent with that of Friedman et al., who showed that the activity of an inhibitor compound can be improved by maximizing the amount of hydrophobic surface area desolvated by the compound. Moreover, the finding presented here provides a molecular mechanism for this desolvation by connecting it with the flap motion, thereby providing important insights into how the fullerene-based inhibitors interact with the active site of the HIV-1 PR. Finally, the simulations indicate that the most effective binding of the C_{60}-based inhibitor to the active site occurs when the Asp dyad is diprotonated. Future work will focus on refinement of the drug and active site models and on estimation of binding free energies using novel adiabatic alchemical transition techniques currently being developed by us (*54*).

ACKNOWLEDGMENT

We gratefully acknowledge a generous grant of computer time from the Hewlett-Packard Co. facility in Richardson, TX, where many of the calculations were performed on the site's high-performance computing facilities. The remainder of the calculations were performed on the HP V-Class system at Caltech under the support of an NPACI allocation (NYU 210). We also acknowledge D. J. Tobias for useful discussions on the construction of drug models. Finally, we acknowledge G. J. Martyna for many useful discussions.

SUPPORTING INFORMATION AVAILABLE

Potential parameters for compound **2** in the form of a list of CHARMM22 atom types and partial charges; in addition, this material describes the algorithm used to construct the rectangular box in the cavity used to compute the water number densities in Water Content in the Active Site. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Fruton, J. S. (1976) in *Advances in Enzymology* (Meister, A., Ed.) Vol. 44, pp 1−36, Wiley, New York.
2. Swain, A. L., Miller, M. M., Green, J., Rich, D. H., Schneider, J., Kent, S. B. H., and Wlodawer, A. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8805−8809.
3. Jaskolski, M., Tomasselli, A. G., Sawyer, T. K., Staples, D. G., Heinrikson, R. I., Schneider, J. S., Kent, S. B. H., and Wlodawer, A. (1991) *Biochemistry 30*, 1600−1609.
4. Ido, E., Han, H. P., Kezdy, F. J., and Tang, J. (1991) *J. Biol. Chem. 266*, 24359−24366.
5. Rodrigues, E. J., Angeles, T. S., and Meek, T. D. (1993) *Biochemistry 32*, 12380−12385.
6. Chatfield, D. C., and Brooks, B. R. (1995) *J. Am. Chem. Soc. 117*, 5561−5572.
7. Liu, H., Florian, M.-P., and van Gunsteren, W. F. (1996) *J. Mol. Biol. 261*, 454−469.
8. Guldnik, S., Erickson, J. W., and Xie, D. (2000) *Vitam. Horm. 58*, 213−256.
9. Wlodawer, A., and Erickson, J. W. (1993) *Annu. Rev. Biochem. 62*, 543−585.
10. Vacca, J. P. L., and Condra, J. H. (1997) *Drug Discovery Today 2*, 261−272.
11. Flexner, C. (1998) *N. Engl. J. Med. 338*, 1281−1292.
12. Wlodawer, A., and Vondrasek, J. (1998) *Annu. Rev. Biophys. Biomed. Struct. 27*, 249−284.
13. Fischl, M. A., Richman, D. D., Grieco, M. H., et al. (1987) *N. Engl. J. Med. 317*, 185−197.
14. Craig, J. C., Whittaker, L., Duncan, I. B., and Roberts, N. (1994) *Antiviral Chem. Chemother. 5*, 380−386.
15. Gulick, R., Mellors, J., Havlir, D., Eron, J., Gonzales, C., McMahon, D., Richman, D., Valentine, F., Jonas, L., Meibohm, A., Emini, E., and Chodakewitz, J. (1997) *N. Engl. J. Med. 337*, 734−739.
16. Autran, B., Carcelain, G., Li, T. S., Blanc, C., Mathez, D., Tubiana, R., Katalaba, C., Debre, P., and Leibowitch, J. (1998) *Science 277*, 112−116.
17. Collins, J. R., Burt, S. K., and Erickson, J. W. (1995) in *Aspartic Proteinases: Structure, Function, Biology, and Biomedical Implications* (Takahashi, K., Ed.) pp 455−460, Plenum Press, New York.
18. Ishima, R., Freedberg, D. I., Wang, Y. X., Louis, J. M., and Torchia, D. A. (1999) *Struct. Folding Des. 7*, 1047−1055.
19. Friedman, S. H., Decamp, D. L., Sijbesma, R. P., Srdanov, G., Wudl, F., and Kenyon, G. L. (1993) *J. Am. Chem. Soc. 115*, 6506−6509.
20. Sijbesma, R., Srdanov, G., Wudl, F., Castoro, J. A., Wilkins, C., Friedman, S. H., DeCamp, D. L., and Kenyon, G. L. (1993) *J. Am. Chem. Soc. 115*, 6510−6512.
21. Schinazi, R. F., Sijbesma, R., Srdanov, G., Hill, C. I., and Wudl, F. (1993) *Antimicrob. Agents Chemother. 37*, 1707−1710.
22. Schuster, D. I., Wilson, S. R., and Schinazi, R. F. (1996) *Bioorg. Med. Chem. Lett. 6* 1253−1256.
23. Friedman, S. H., Ganapathi, P. S., Rubin, Y., and Kenyon, G. L. (1998) *J. Med. Chem. 41*, 2424−2429.
24. Friedman, S. H., Schinazi, R. F., Wudl, F., Hill, C. I., DeCamp, D. L., Sijbesma, R., and Kenyon, G. L. (1998) U.S. Patent 5,811,460.
25. Schuster, D. I., Wilson, S. R., Kirschner, A. N., Schinazi, R. F., Schleuter-Wirtz, S., Tharnish, P., Barnett, T., Ermolieff, J., Tang, J., Brettreich, M., and Hirsch, A. (2000) *Proc. Electrochem. Soc. 11*, 267−270.
26. Marcorin, G. L., Ros, T. D., Castellano, S., Stefancich, G., Bonin, I., Miertus, S., and Prato, M. (2000) *Org. Lett. 2*, 3955−3958.
27. Da Ros, T., Spalluto, G., and Prato, M. (2001) *Croat. Chem. Acta 74*, 743−755.
28. Bingel, C. (1993) *Chem. Ber. 126*, 1957−1959.
29. Tuckerman, M. E., Martyna, G. J., and Berne, B. J. (1992) *J. Chem. Phys. 97*, 1990−2001.
30. Martyna, G. J., Tuckerman, M. E., Tobias, D. J., and Klein, M. L. (1996) *Mol. Phys. 87*, 1117−1157.
31. Martyna, G. J., Tuckerman, M. E., and Klein, M. L. (1992) *J. Chem. Phys. 97*, 2635−2643.
32. Essman, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) *J. Chem. Phys. 103*, 8577−8593.
33. Procacci, P., Darden, T., and Marchi, M. (1996) *J. Phys. Chem. 100*, 10464−10468.
34. Spinelli, S., Liu, Q. Z., and Alzari, P. M. (1991) *Biochimie 73*, 1391−1396.
35. Mi, H., Schuster, D. I., Wilson, S. R., and Tuckerman, M. E. (1999) *Proc. Electrochem. Soc. 99*, 256−262.
36. Tuckerman, M. E., Yarne, D. A., Samuelson, S. O., Hughes, A. L., and Martyna, G. J. (2000) *Comput. Phys. Commun. 128*, 333−376.
37. MacKerell, A. D., Jr., Bashford, D., Bellott, M., Dumbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, O., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., III, Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) *J. Phys. Chem. B 102*, 3586−3616.
38. Jorgensen, W., Chandrasekhar, J., Madrua, J., Impey, R., and Klein, M. L. (1983) *J. Chem. Phys. 79*, 926−935.
39. Ryckaert, J. P., Ciccotti, G., and Berendsen, H. J. C. (1977) *J. Comput. Phys. 23*, 327−341.
40. Ciccotti, G., Ferrario, M., and Ryckaert, J. P. (1982) *Mol. Phys. 47*, 1253−1264.
41. Ryckaert, J. P., and Ciccotti, G. (1986) *Comput. Phys. Rep. 4*, 345−392.
42. Andersen, H. C. (1983) *J. Comput. Phys. 52*, 24−34.
43. Martyna, G. J., Tobias, D. J., and Klein, M. L. (1993) *J. Phys. Chem. 97*, 12959−12966.

44. Martyna, G. J., Tobias, D. J., and Klein, M. L. (1994) *J. Chem. Phys. 101*, 4177−4189.
45. Carter, E. A., Ciccotti, G., Hynes, J. T., and Kapral, R. (1989) *Chem. Phys. Lett. 156*, 472−477.
46. Sprik, M., and Ciccotti, G. (1998) *J. Chem. Phys. 109*, 7737−7744.
47. Rosso, L., and Tuckerman, M. E. (2002) *Mol. Simul. 28* (1−2), 91−112.
48. Rosso, L., Minary, P., Zhu, Z., and Tuckerman, M. E. (2002) *J. Chem. Phys. 116*, 4389−4402.
49. Piana, S., Sebastiani, D., Carloni, P., and Parrinello, M. (2001) *J. Am. Chem. Soc. 123*, 8730−8737.
50. Smith, R., Brereton, I. M., Chai, R. Y., and Kent, S. B. H. (1996) *Nat. Struct. Biol. 3*, 946−950.
51. Hyland, L. J., Tomaszek, T. A., Roberts, G. D., Carr, S. A., Magaard, V. W., Bryan, H. L., Fakhoury, S. A., Moore, M. L., Minnich, M. D., Culp, J. S., Desjarlais, R. L., and Meek, T. D. (1991) *Biochemistry 30*, 8441−8453.
52. Roberts, N. A., Martin, J. A., Kinchington, D., Broadhurst, A. V., Craig, J. C., Duncan, I. B., Galpin, S. A., Handa, B. K., Kay, J., Krohn, A., Lambert, R. W., Merritt, J. H., Mills, J. S., Parkes, K. E. B., Redshaw, S., Ritchie, A. J., Taylor, D. L., Thomas, G. J., and Machin, P. J. (1990) *Science 248*, 358−361.
53. Piana, S., and Carloni, P. (2000) *Proteins: Struct., Funct., Genet. 39*, 26−36.
54. Rosso, L., and Tuckerman, M. E. (2003) manuscript in preparation.

BI020496S