Determination of the Relative Binding Free Energies of Peptide Inhibitors to the HIV-1 Protease

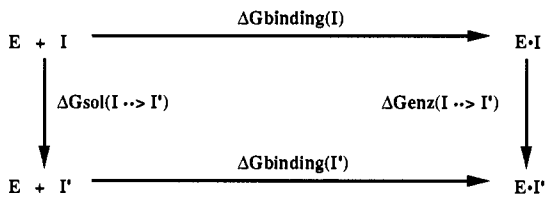
Introduction

The human immunodeficiency virus (HIV) presents one of the most challenging and medically important problems to confront the scientific community in many years. The process by which the virus replicates is only beginning to be understood, but some essential steps have been identified. One such step involves posttranslational processing, which activates structural proteins and enzymes. These are translated from the viral-derived RNA, forming a single polypeptide precursor containing the concatenated premature proteins. The HIV protease is responsible for cleaving the polypeptide precursor into the functional proteins and is essential for viral replication.[1] The protease has therefore become the target of many drug-development efforts.

One of the most significant events aiding drug development has been the determination of HIV-1 protein-inhibitor crystal structures[2,3] and the resulting confirmation that the enzyme is an aspartic protease, as suggested by the sequence. This not only allows the construction of inhibitors based on previous experiences with aspartic proteases but more importantly affords a template for building novel substrates. We are interested in developing theoretical approaches and applying these to determine (i) the factors that influence inhibitor-protein complex stabilities and (ii) how to take advantage of these factors and modify potential inhibitors to produce viable drugs for anti-AIDS therapy. This report describes a theoretical study of the basis for stereoselective binding of two isomeric inhibitors to the HIV-1 protease and predicts the binding affinity of a third, as of yet unreported, inhibitor.

The most effective protease inhibitors described to date are modified peptides that act as transition-state analogues. These types of inhibitors contain functional groups that interact favorably with the two aspartyl residues that catalyze peptide cleavage.[4-6] One such inhibitor reported by Rich et al. binds to the protease with impressive affinity.[4] This inhibitor contains a crucial hydroxyl group that mimics the tetrahedral intermediate in hydrolysis. The binding affinity, however, is dependent on the configuration of the chiral center to which this hydroxyl group is attached. The difference in the binding affinities of the two diastereomeric forms was not reported in the Rich et al. study and cannot be deduced from the X-ray crystal structure.[3] Moreover, predictions based on static structures cannot yield relative free energies for inhibitor-protein complexes.

Theoretical methods, however, can be employed to determine these energies and to recognize important structural features that stabilize the complex. Figure 1 depicts the interaction of the *S*-configured hydroxyl group with the active site residues after 10 ps of molecular dynamics starting from the X-ray crystal structure.[3] Although the structural interactions of the protein-inhibitor complex can be elucidated by using molecular dynamics, the relative binding free energies of inhibitors to the protease cannot be directly determined in this manner. Such energies have been reliably calculated by applying an advanced implementation of molecular dynamics, free energy perturbation (for a complete description of FEP see refs 7 and 8). This methodology allows the state of the system to be "mutated" to a second state by perturbing the parameters that define the system. Concurrently, the free energy change is accumulated over the course of the molecular dynamics simulation. This method can be employed to determine the enzyme binding free energy difference between two inhibitors (I and I') from a simple relationships that is derived from the following thermodynamic cycle.



Thermodynamic Cycle

In the present study, we will apply free energy perturbation to determine the relative binding free energy of *S* and *R* diastereomers of the Rich et al. inhibitor JG365.[4] For this calculation, the *S*-hydroxyl group will be mutated to a hydrogen while the hydrogen will be simultaneously mutated to the *R*-hydroxyl in the perturbing group (see Figure 2). The mutations will be applied in the thermodynamic cycle above from which the $\Delta\Delta G_{\text{binding}(I' \rightarrow I)}$ can be determined. The calculated free energy difference will then be compared to an experimental value for the diastereomers that has been determined independently and in parallel with our study. Finally, by only mutating the *S*-hydroxyl to a hydrogen, the relative binding free energy of an, as yet, unreported inhibitor (minus the key hydroxyl) will be calculated.

Methods and Model Development

The starting structures for the simulations involving the enzyme were built by adding hydrogens to the heavy-atom positions of an HIV-1 protease-inhibitor crystal structure (7HVP[3]) with AMBER 3.0 Revision A.[9] The pyrrolidine

(1) Kohl, N. E.; Emini, E. A.; Schleif, W. A.; Davis, L. J.; Heimbach, J. C.; Dixon, R. A.; Scolnick, E. M.; Sigal, I. S. Active Human Immunodeficiency Virus Protease is required for Viral Infectivity. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 4686-90.
(2) Miller, M.; Schneider, J.; Sathyanarayana, B. K.; Toth, M. V.; Marshall, G. R.; Clawson, L.; Selk, L.; Kent, S. B.; Wlodawer, A. Structure of Complex of Synthetic HIV-1 Protease with a Substrate-Based Inhibitor at 2.3 Å Resolution. *Science* **1989**, *246*, 1149-52.
(3) Swain, A. L.; Miller, M. M.; Green, J.; Rich, D. H.; Schneider, J.; Kent, S.; Wlodawer, A. X-ray Crystallographic Structure of a Complex Between a Synthetic Protease of Human Immunodeficiency Virus 1 and a Substrate-Based Hydroxyethylamine Inhibitor. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 8805-9.
(4) Rich, D. H.; Green, J.; Toth, M. V.; Marshall, G. R.; Kent, S. B. Hydroxyethylamine Analogues of the p17/p24 Substrate Cleavage Site are Tight-Binding Inhibitors of the HIV Protease. *J. Med. Chem.* **1990**, *33*, 1285-8.
(5) Dreyer, G. B.; Metcalf, B. W.; Tomaszek, T. A.; Carr, T. J.; Chandler, A. C.; Hyland, L.; Fakhoury, S. A.; Magaard, V. W.; Moore, M. L.; Strickler, J. E. Inhibition of Human Immunodeficiency Virus 1 Protease in Vitro: Rational Design of Substrate Analogue Inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 9752-6.
(6) 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. Rational Design of Peptide-Based HIV Proteinase Inhibitors. *Science* **1990**, *248*, 358-61.
(7) Bash, P. A.; Singh, U. C.; Langridge, R.; Kollman, P. A. Free Energy Calculations by Computer Simulation. *Science* **1987**, *236*, 564-8.
(8) Van Gunsteren, W. The Role of Computer Simulation Techniques in Protein Engineering. *Protein Eng.* **1988**, *2*, 5-13.

0022-2623/91/1834-2654$02.50/0

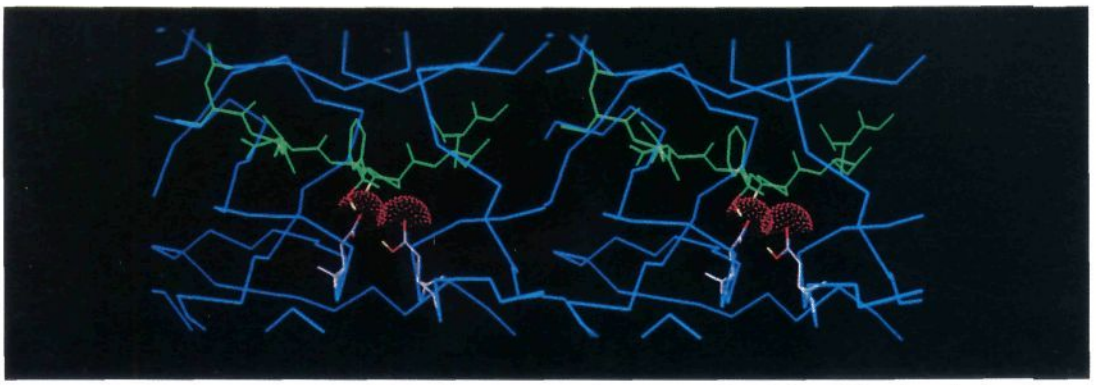


Figure 1. Stereoview of the HIV-1 active site with the *S* diastereomer of the heptapeptide inhibitor JG365 bound after 10 ps of molecular dynamics. The inhibitor is green, the protein backbone is blue, with the catalytic aspartyl residues highlighted in magenta. The oxygen of the *S*-hydroxyl group is red, while the hydrogen-bonding hydrogens are yellow.

amine group of the inhibitor is protonated in all of our studies since the pK_a of the protonated base is considerably greater than the pH at which binding studies are typically done. The stereochemistry of the quaternary nitrogen was chosen to allow interaction with the only hydrogen-bond acceptor (Asp25) within 4 Å in the enzyme crystal structure. Atomic charges for the nonstandard residues (inhibitors and protonated aspartic acid) were determined by electrostatic potential fitting[10,11] to wave functions with an STO-3G basis set, using quantum mechanically optimized (AM1) molecular geometries. Other molecular mechanical parameters for the structures were taken from previous work.[12-14]

Because the protonation state of the two catalytic aspartates (Asp25 and Asp125) has not been determined, three possible protonation models were considered. One model left both aspartyl residues anionic while the remaining models had one or the other aspartate protonated. In addition to the 95 crystallographic waters, a spherical cap containing 133 TIP3P[15] water molecules was added around the active site of the protein. This was done by placing water at all unoccupied positions within 18 Å of the perturbing group. The starting structure for the determination of $\Delta G_{sol(I \rightarrow I')}$ was a periodic box containing the *S* diastereomer in the conformation taken from the crystal coordinates and 1085 TIP3P waters.

All free energy simulations were carried out with the AMBER 4.0 GIBBS module.[16] For the enzyme simulations, all atoms of amino acid residues completely outside of a 9-Å region enclosing the inhibitor were fixed in Cartesian space to reduce the amount of computational time needed for the calculation.[17] The fixed region is conserved for all crystal structures of protease–inhibitor complexes and should not introduce significant phase-space sampling errors. To further constrain the structure to this conserved canonical form, the backbone atoms of the enzyme within the 9-Å region were harmonically restrained to the crys-

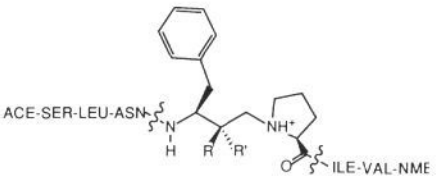


Figure 2. Heptapeptide inhibitors studied by using free energy perturbation. The perturbing group is noted and contains the atoms that will be "mutated" in the simulations. Structures **1** and **2** are the *S* and *R* diastereomers of the inhibitor JG365 previously reported.[4] Structure **3** is a modification of JG365 that does not contain the key hydroxyl group. The ACE and NME are acetyl and methylamide caps on the heptapeptide.

(9) Seibel, G.; Singh, U. C.; Weiner, P. K.; Caldwell, J.; Kollman, P. AMBER (UCSF) version 3.0, Revision A (1989), Department of Pharmaceutical Chemistry, University of California, San Francisco.

(10) Singh, U. C.; Kollman, P. A. An Approach for Computing Electrostatic Charges for Molecules. *J. Comput. Chem.* **1984**, *5*, 129–144.

(11) Charges listed by using 7HVP (Brookhaven Protein Data Bank Nomenclature) atom names (atom name:residue number,-charge). *S* and *R* diastereomers: N:204,–0.463; HN:204,0.252; CA:204,0.210; HA:204,0.050; CB:204,–0.310; HB:204,0.098; CG:204,0.151; CD1:204,–0.117; HD1:204,0.079; CE1:204,–0.057; HE1:204,0.066; CZ:204,–0.073; HZ1:204,0.068; CE2:204,–0.057; HE2:204,0.066; CD2:204,–0.117; HD2:204,0.079; C:204,0.666; HC:204,–0.092; O:204,–0.592; HO:204,0.344; CS:204,–0.535; HCS:204,0.177; N:204,0.387; HN:205,0.259; CD:205,–0.141; HD:205,0.112; CG:205,–0.050; HG:205,0.080; CB:205,–0.168; HB:205,0.089; CA:205,–0.117; HA:205,0.074; C:205,0.526; O:205,–0.500. No-OH inhibitor: N:204,–0.463; HN:204,0.252; CA:204,0.407; HA:204,0.025; CB:204,–0.336; HB:204,0.108; CG:204,0.221; CD1:204,–0.143; HD1:204,0.071; CE1:204,–0.047; HE1:204,0.064; CZ:204,–0.097; HZ1:204,0.069; CE2:204,–0.047; HE2:204,0.064; CD2:204,–0.143; HD2:204,0.071; C:204,0.296; HC:204,–0.032; CS:204,–0.441; HCS:204,0.174; N:205,0.221; HN:205,0.281; CD:205,–0.090; HCD:205,0.100; CG:205,–0.069; HCG:205,0.079; CB:205,–0.151; HCB:205,0.083; CA:205,–0.130; HCA:205,0.065; C:205,0.526; O:205,–0.500. Protonated (neutral) aspartic acid: N:125,–0.463; HN:125,0.252; CA:125,0.035; HA:125,0.048; CB:125,–0.086; HB1:125,0.039; CG:125,0.753; OD1:125,–0.582; HOD1:125,0.398; OD2:125,–0.545; C:125,0.616; O:125,–0.504.

(12) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. An All Atom Force Field for Simulations of Proteins and Nucleic Acids. *J. Comput. Chem.* **1986**, *7*, 230–52.

(13) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. A New Force Field for Molecular Mechanical Simulation of Nucleic Acids and Proteins. *J. Am. Chem. Soc.* **1984**, *106*, 765–84.

(14) Bash, P. A. Molecular Modeling and Free Energy Calculations in Molecular Biology. Ph.D. Dissertation, University of California, Berkley, 1986.

(15) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–35.

(16) AMBER (UCSF) version 4.0, GIBBS, by David A. Pearlman, and an extensive modification of AMBER 3.0, by U. C. Singh, P. K. Weiner, J. Caldwell, and P. Kollman, Department of Pharmaceutical Chemistry, University of California, San Francisco, CA, add the capability to carry out perturbations in which 10–12 hydrogen-bond parameters are mutated into 6–12 van der Waals parameters. It also includes the intraperturbed group free energies in the total free energies.

(17) Only the following amino acid residues were allowed to move: 7–10, 21–34, 45–58, 74–90 (both monomers).

Table I. HIV-1 Protease–Inhibitor Binding Free Energy Differences[a]

perturbation	protonation state	$\Delta G_{enz(I\rightarrow I')}$ for.	$\Delta G_{enz(I\rightarrow I')}$ rev.	$\Delta G_{enz(I\rightarrow I')}$ ave[d]	$\Delta G_{sol(I\rightarrow I')}$ for.	$\Delta G_{sol(I\rightarrow I')}$ rev.	$\Delta G_{sol(I\rightarrow I')}$ ave[d]	$\Delta\Delta G_{binding}$[b]	$\Delta\Delta G_{exp}$[c]
$S \rightarrow R$	dianionic	2.2	1.7	2.0 ± 0.3	−1.8	−2.2	−2.0 ± 0.2	4.0 ± 0.4	2.6
$S \rightarrow R$	Asp25	4.8	4.8	4.8 ± 0.0	−1.8	−2.2	−2.0 ± 0.2	6.8 ± 0.2	2.6
$S \rightarrow R$	Asp125	1.0	−2.0	−0.5 ± 1.5	−1.8	−2.2	−2.0 ± 0.2	1.5 ± 1.5	2.6
$S \rightarrow R$[e]	Asp125	0.8	0.7	0.8 ± 0.1	−1.8	−2.2	−2.0 ± 0.2	2.8 ± 0.2	2.6
$S \rightarrow$ No-OH	dianionic	36.4	36.5	36.5 ± 0.1	31.0	28.8	29.9 ± 1.1	6.6 ± 1.2	
$S \rightarrow$ No-OH	Asp25	38.5	38.3	38.4 ± 0.1	31.0	28.8	29.9 ± 1.1	8.5 ± 1.2	
$S \rightarrow$ No-OH	Asp125	34.7	32.6	33.7 ± 1.1	31.0	28.8	29.9 ± 1.1	3.8 ± 1.5	

[a] Unless otherwise noted, variable constraints were used on the backbone atoms.[19] The simulations were all run for 20 ps in the forward (for.) and reverse (rev.) directions. [b] $\Delta\Delta G_{binding(I\rightarrow I')} = \Delta G_{enz(I\rightarrow I')} - \Delta G_{sol(I\rightarrow I')} = \Delta G_{binding(I')} - \Delta G_{binding(I)}$. [c] Reference 22. [d] Error is average deviation. [e] 10 kcal/Å^2 constraints were used on all backbone atoms.[21]

tallographically determined positions. The harmonic constraint energy decreased as the distance from the fixed region increased (for a description of the constraints, see note 19). The inhibitor was not constrained in any of the enzyme or periodic solvent simulations.

The structures were equilibrated for a minimum of 10 ps of molecular dynamics in both the solvent and the enzyme systems. The inhibitor in the enzyme simulations shifted ca. 1 Å from the position in the starting X-ray structure during the equilibration run and remained in this position during subsequent simulations. This movement allowed for the optimization of the charge–charge interactions of the protonated pyrrolidine with the catalytic aspartyl residues. This strong interaction can be seen in Figure 1, as well as the hydrogen-bonding interaction of the *S*-hydroxyl group with the aspartyl residues. Although the movement of the inhibitor has changed some of the nonbonded and hydrogen-bonded distances, the structural features and interactions that characterize the binding of the peptide are maintained during molecular dynamics. Furthermore, no internal conformational changes have taken place within the inhibitor during the equilibration. The structural changes that have occurred could be described as "normal" for an enzyme system that has been given the kinetic energy required to maintain an average temperature of 298.15 K. As noted previously, the starting structure for the peptide in solution was taken from the observed conformation in the X-ray structure. This structure was chosen for two reasons: First, it is an experimentally observed structure. Second, since it is computationally impractical to completely sample the free energies for flexible peptides in solution, the use of an arbitrary conformation may not allow the most relevant microstates to be sampled during the free energy perturbation calculation. The structure remained in the extended conformation throughout the periodic solvent simulations.

The free energy perturbations were run with the "slow growth" option of GIBBS[7] in the forward and reverse directions using the models described above. Several runs were initially performed with 20-, 30-, and 40-ps sampling times to examine the dependence of the free energy on this parameter. Our results indicated that the longer trajectories (30 and 40 ps) did not improve the convergence of the free energy, but increased the hysteresis in the reverse direction which is due to further "drift" of the structure from the X-ray geometry. The free energies in the forward direction, however, did not significantly change with respect to the length of the simulation, indicating that all the forward runs were comparable. While we cannot rule out that much longer simulations may effect better convergence, these do not appear to be necessary to determine consistent free energies and, furthermore, would be CPU prohibitive. The choice of 20-ps free energy simulations was found to be the best compromise all around and produced trajectories that remained close to the X-ray structure (less hysteresis) as well as reasonable free energies within a feasible amount of CPU time.

The free energy contributions from both the intermolecular and intramolecular contributions were accumulated during all of the simulations. The molecular dynamics trajectories were calculated from a 1fs time step and the SHAKE algorithm[18] was applied to all bonds to hydrogen. A constant dielectric of 1.0 was used with a nonbonded cutoff of 8.0 Å. The pair-list update was performed every 20 time steps. The systems were coupled to a constant temperature bath at 298.15 K with the addition of constant pressure (1 atm) coupling in the periodic solvent simulations.

Results and Discussion

The free energies reported in Table I are for the inhibitors bound to the protease in the three possible protonation states and the inhibitors in solvent alone. The free energies calculated for the enzyme-bound inhibitors ($\Delta G_{enz(I\rightarrow I')}$) predict the *S* diastereomer to interact more favorably with the enzyme in all three models. The relative binding free energies ($\Delta\Delta G_{binding(I'\rightarrow I)}$), derived from the combination of $\Delta G_{enz(I\rightarrow I')}$ and $\Delta G_{sol(I\rightarrow I')}$, also predict the *S* diastereomer to bind with higher affinity in all cases. The results show that $\Delta G_{sol(I\rightarrow I')}$ plays an important role in determining the relative binding of the diastereomers. Although this energy does not directly correspond to the relative solvation free energy, since contributions from the intramolecular interactions are also included,[20] it points out that differences in solvation can dramatically effect the relative binding of protease inhibitors. While the origin of these solvation differences cannot be ascertained, especially in the case of the diastereomers, the contributions to the relative binding free energy are accounted for in our free energy methodology.

The other contribution to the relative binding of the inhibitors to the protease is more tractable and is due to changes in binding interactions within the protein–inhibitor complex. In the case of the *R* and *S* inhibitors, the difference in binding free energy is related to the conformational change that is required for the *R* diastereomer to bind tightly to the active sight of the protease. This

(18) Van Gunsteren, W. F.; Berendsen, H. J. C. Algorithms for Macromolecular Dynamics and Constraint Dynamics. *Mol. Phys.* **1977**, *34*, 1311–27.

(19) Constraint on backbone atoms of amino acid residues: 7–10, 21–23, 33–34, 45–58, 74–76, 88–90 (2.0 kcal/Å^2), 24–26, 31–32, 77–79, 85–87 (1.0 kcal/Å^2), 27–28, 30–30, 80–84 (0.5 kcal/Å^2), (both monomers).

(20) To calculate the relative solution free energies the gas phase $\Delta G_{gas(I\rightarrow I')}$ would have to be calculated. This cannot be reliably determined in this case due to the conformational variability of the heptapeptide in the gas phase.

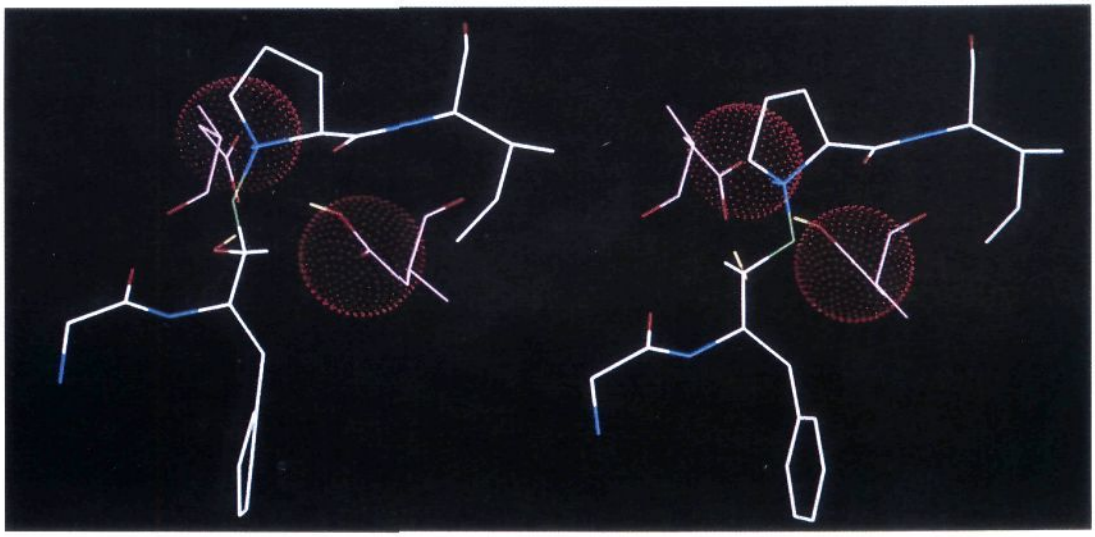


Figure 3. A view of the *S* diastereomer (left) and *R* diastereomer displaying the rotation of the dihedrals (center atom colored green) that are responsible for the large hysteresis in the free energy calculation. Note that the rotation allowed the *R*-hydroxyl group to be directed between both of the aspartyl residues. The conformation of the *R* diastereomer was taken at the end of the *S* to *R* free energy perturbation simulation.

conformational change is also responsible for the large hysteresis in the Asp125-protonated *S* to *R* simulation (see Table I). In order for the *R*-hydroxyl group to interact favorably with the aspartyl residues, several dihedral rotations occurred during the *S* to *R* FEP simulation, inducing the conformational change depicted in Figure 3. During the free energy calculation in the reverse direction (*R* back to *S*), the structure of the *S* diastereomer did not return to the original conformation (as seen in Figure 1). The free energy change calculated for the reverse process is therefore not an accurate measure of the relative binding free energy since the appropriate regions of conformational space were not sufficiently sampled for this particular simulation. This sampling problem for the reverse simulation was corrected when the backbone atoms of the protein were all equally constrained.[21] The inhibitor still underwent the conformational change during the forward perturbation to the *R* form and the corresponding free energy calculated agreed quite well with the previous determination (0.8 vs 1.0 kcal/mol). However, in the reverse direction, the structure returned to the original X-ray conformation of the *S* form and the free energy calculated was in close agreement with the forward value. While both constraint models sampled a similar set of microstates in the forward direction, tighter coupling to the X-ray structure of the protein enabled us to eliminate the hysteresis in the reverse simulation and further verify the dependence of the free energy on the conformation of the inhibitor.

While this study was in progress, the *S* diastereomer was determined experimentally to be the better inhibitor[22] in agreement with the results of our calculations. However, $\Delta\Delta G_{\text{binding}(I' \rightarrow I)}$ for the Asp25 protonated structure is an unrealistically high value (6.8 kcal/mol) compared to the experimental value of 2.6 kcal/mol; thus it does not appear to be the correct model for the enzyme. This is further supported by the fact the Asp25 hydrogen bonds to the protonated nitrogen of the inhibitor, so one would expect it to remain anionic. The results from the *R* and *S* calculations are still not conclusive as to whether Asp125 is protonated since the energies of the dianionic and Asp125 protonated structures are 4.0 and 2.8 kcal/mol,[20] both of which are reasonably close to the experimental value.

Although most studies suggest that the protease is singly protonated, indicating that the Asp125 protonated form is most likely the preferred state, we chose to further investigate the preferred protonation state by predicting the $\Delta\Delta G$ for a new inhibitor. The relative free energy of binding for the inhibitor with the hydroxyl group removed was predicted for all three protonation states. This specific inhibitor has not been studied experimentally, but analogous peptides without hydroxyl groups in this position[4] show 1200–20 000-fold decreases in affinity (4.2–5.9 kcal/mol). The total free energies reported in Table I for the perturbation of the *S*-hydroxyl group to a hydrogen have been determined from differences of relatively large ΔG_{enz} and ΔG_{sol} values. These energies include the intramolecular free energies of the perturbing group (see Figure 2) and can be quite large, especially when atoms are being annihilated during the simulation. Although we have derived relatively small free energy differences from these values, they can be determined rather accurately as shown by Singh,[23] since the absolute energies are predominantly intramolecular. Our results show that the dianionic and Asp25-protonated structures produce relatively high free energy differences ($K_{\text{i no hydroxyl}}/K_{\text{i hydroxyl}}$ of 59 000 and 1 500 000, respectively) for the ethylamine and hydroxyethylamine inhibitors. In contrast, the Asp125-protonated model yielded a free energy difference in the range of those observed, further supporting the single protonation model at Asp125. Although it would be difficult to unequivocally rule out the dianionic state from the free energies alone, this state is less likely since this would require two negative charges to be in close

(21) Constraints of 10.0 kcal/Å^2 were used on the backbone atoms of all amino acid residues.

(22) Rich, D. H.; Sun, C. Q.; Vara Prasad, J. V. N.; Pathiasseril, A.; Toth, M. V.; Marshall, G. R.; Clare, M.; Mueller, R. A.; Houseman, K. Effect of Hydroxyl Group Configuration in Hydroxyethylamine Dipeptide Isosteres on HIV Protease Inhibition—Evidence for Multiple Binding Modes. *J. Med. Chem.* **1991**, *34*, 1222–5.

(23) Singh, U. C. Probing the Salt Bridge in the Dihydrofolate Reductase–methotrexate Complex by Using the Coordinate-Coupled Free-Energy Perturbation Method. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 4280–84.

proximity to each other. A dianionic state would also preclude the formation of a hydrogen bond between the aspartyl residues and an overall network of hydrogen bonds with the inhibitor (as can be seen in Figure 1).

Our calculations suggest that the protease is singly protonated. The implication that the protease is singly protonated when active would agree with the established model for pepsin, which is known to be singly protonated in an acidic pH range. While the active site of pepsin contains serine and threonine residues that can hydrogen bond to the aspartyl residues,[24] lowering their basicity, no such mechanism exists in the HIV-1 protease. In this case, the aspartyl residues would be more basic, functioning in an analogous fashion to those in pepsin, but at higher pH values. It is clear from our study that the ethylamine hydroxyl group produces marked stabilization of the protein–inhibitor complex due to hydrogen bonding to the aspartyl residues. This interaction was shown to induce a conformational change in the *R* diastereomer that resulted in a decrease in binding affinity. It therefore should be more favorable to build inhibitors in which the hydroxyl group (or equivalent functionality) can be placed between the two aspartyl residues without inducing strain or unfavorable conformational changes. The addition of a second hydroxyl group to the inhibitor may help avoid the conformational requirements for binding that depend on the configuration of the inhibitor. Kempf et al. have shown that the binding of symmetric glycol containing inhibitors is not dependent on the configuration of the two carbon centers that the hydroxyl groups are attached.[25] As noted in that study, this may be due to flexibility at the hydroxyl–carbon centers of the inhibitors, but may also be a result of the availability of multiple binding modes for the diastereomers, afforded by the presence of a second hydroxyl group in the active site.

Caveats

The major limitation of free energy perturbation approaches is the multiple minima problem that is the bane of theoretical chemistry on complex molecules.[26] It is clear that we cannot carry out the free energy calculations over a long enough period of time to completely sample all of the microstates; all we can hope is that our sampling be reasonably representative. While free energy calculations tend to converge quite rapidly if they are electrostatically dominated, this occurs much more slowly if they have a significant van der Waals component. For example, one can determine the relative free energies of methanol and ethane and Ne and Na^+ accurately with simulations of 30 ps.[27,28] But on the other hand, the mutation of threonine to alanine has significant error bars unless one carries out the mutation over 100 ps.[29] Simulations of this length would require ca. 100 Cray YMP hours for each inhibitor analogue bound to a particular protonation state of the HIV-1 protease, which is impractical. Furthermore, such long simulations may not be warranted since our mutation is electrostatically dominated. One also has the problem that the force field and the representation of the system (e.g. limited solvation of the enzyme) is imperfect and the system has a tendency to drift from the X-ray structure under these conditions unless constrained. The constraints we employed were designed to minimize the "drift" from the experimental structure, while allowing the efficient sampling of phase space, but no constraint set is perfect. Thus, our results should be taken with a "grain of salt" because they are only relevant if the system retains a geometry that is "close" to that observed in the crystal structure during the free energy perturbation.

While these above caveats were present in an early version of this paper, one referee still took issue with the conditions we applied to conduct the free energy simulations and had "mixed feelings" about the paper. In particular, the reviewer suggested that the work has several technical problems and cited the use of numerous constraints, short cutoffs, little water, STO-3G charges, and molecular simplifications, as well as the use of short free energy simulation times. We have discussed the rational for employing the protocol we did (vide anti), but we should comment on those that we have not mentioned. The cutoff of 8.0 Å used is quite typical and reasonable in macromolecular simulations. The use of STO-3G charges for the inhibitor is appropriate because that is the basis on which the charges for the protein were derived.[12,13] The use of a limited solvation sphere is an approximation to "solvent effects" but is also standard practice in simulations performed that involve perturbations in the active site.[30,31] Moreover, our perturbation is local to the center of the active site in the protease that is buried in the protein where solvation is not likely to be the key issue. The referee was most concerned about the shortness of the simulation time, but as noted in the paragraph above, such short times have been found to be adequate in many cases when the perturbation is electrostatic dominated, as this one is. We agree that the simplifications may seem extreme, but in our opinion are required to produce a *model* that can be used to address the questions posed in this study given current computing resources. (Note that each 20-ps free energy simulation exhausts 20 h of Cray YMP time.)

Conclusions

Despite the caveats noted above, it is also important to emphasize what we can glean from this study. To our knowledge, this is only the second example of the use of free energy perturbation theory to predict relative free energies of enzyme–inhibitor binding where the affinity

(24) THR 218 and SER 35 in the crystal structure 5PEP: Cooper, J. B.; Khan, G.; Taylor, G.; Tickle, I. J.; Blundell, T. L. X-ray Analysis of Aspartic Proteinases—The 3-Dimensional Structure at 2.1 Å Resolution of Endothiapepsin. *J. Mol. Biol.* **1990**, *214*, 199–222.

(25) Kempf, D. J.; Norbeck, D. W.; Codacovi, L. M.; Wang, X. C.; Kohlbrenner, W. E.; Wideburg, N. E.; Paul, D. A.; Knigge, M. F.; Vasavanonda, S.; Craig-Kennard, A. Structure-Based, C_2 Symmetric Inhibitors of HIV Protease. *J. Med. Chem.* **1990**, *33*, 2687–9.

(26) Kollman, P. A.; Merz, K. M. Computer Modeling of the Interactions of Complex Molecules. *Acc. Chem. Res.* **1990**, *23*, 246–52.

(27) Singh, U. C.; Brown, F. K.; Bash, P. A.; Kollman, P. A. An Approach to the Application of Free Energy Perturbation Methods Using Molecular Dynamics: Applications to the Transformations of $CH_3OH \rightarrow CH_3CH_3$, $H_2O^+ \rightarrow NH_4^+$, Glycine→Alanine, Alanine→Phenylalanine in Aqueous Solution and to $H_3O^+(H_2O)_3 \rightarrow NH_4^+(H_2O)_3$ in the Gas Phase. *J. Am. Chem. Soc.* **1987**, *109*, 1607–14.

(28) Straatsma, T. P.; Berendsen, H. J. C. Free Energy of Ionic Hydration: Analysis of a Thermodynamic Integration Technique to Evaluate Free Energy Differences by Molecular Dynamics Simulations. *J. Chem. Phys.* **1988**, *89*, 5879–86.

(29) Mitchell, M. J.; McCammon, J. A. Free Energy Difference Calculations by Thermodynamic Integration: Difficulties in Obtaining a Precise Value. *J. Comput. Chem.* **1991**, *12*, 271–275.

(30) Bash, P. A.; Singh, U. C.; Brown, F. K.; Langridge, R.; Kollman, P. A. Calculation of the Relative Binding Free Energy of a Protein–Inhibitor Complex. *Science* **1987**, *235*, 574–6.

(31) Rao, R. N.; Singh, U. C.; Bash, P. A.; Kollman, P. A. Free Energy Perturbation Calculations on Binding and Catalysis: Mutating Asn155 in Subtilisin. *Nature* **1987**, *328*, 551–4.

has been subsequently tested experimentally.[32] Moreover, this system poses an additional problem because, unlike our previous work with thermolysin,[32] there is ambiguity on the protonation state of the two catalytic aspartyl residues in the active site. The results of our study are all *qualitatively* consistent with the preference of the *S* over the *R* diastereomer for the heptapeptide inhibitor JG365, and this is consistent with the experimental results carried out in another lab concurrently with our calculations. But our relative free energies can be used in a quantitative fashion to predict the protonation state of the HIV-1 protease–JG365 complex, a prediction that can be tested experimentally by neutron diffraction. Furthermore, our predicted relative free energy for the *No-OH* compound provides yet another result that can be tested experimentally and related to the probable protonation state. Thus, our calculations have been useful, and predictive, within the limitations of the methodology noted above.

Acknowledgment. We thank David A. Pearlman for providing a preliminary copy of AMBER 4.0 and George L. Seibel for helpful comments. Thanks are also due to the Computer Graphics Laboratory at UCSF (supported by NIHRR-1081), Robert Langridge, P.I., and the San Diego Super Computer Center for computer time. We are grateful to Greg Couch for modifications to the graphics program MIDAS that allowed us to create the color stereoviews. This work was supported by grant GM-39552, George Kenyon, P.I., and grant GM-29072 (to P.A.K.).

Note Added in Revision: Our original drafts of this paper reported the experimental relative free energy for the (*R*)- and (*S*)-hydroxyethylamine inhibitors JG-365 to be 3.1 kcal/mol from a personal communication with Garland Marshall. More recently Rich and Marshall et al. have determined this value to be 2.6 kcal/mol.[22] We were unaware of this difference in the relative binding free energy until we received and manuscript after review. This more recent value lends even stronger support to our predictions and conclusions.

* To whom correspondence should be addressed.

David M. Ferguson, Randall J. Radmer
Peter A. Kollman*

University of California
Department of Pharmaceutical Chemistry
San Francisco, California 94143-0446

Received April 29, 1991

(32) Merz, K. M.; Kollman, P. A. Free Energy Perturbation Simulations of the Inhibition of Thermolysin: Prediction of the Free Energy of Binding of a New Inhibitor. *J. Am. Chem. Soc.* **1989**, *111*, 5649–58.