Calculation of Relative Binding Free Energies of Peptidic Inhibitors to HIV-1 Protease and Its I84V Mutant

G. J. Tawa,* I. A. Topol, S. K. Burt, and J. W. Erickson

Contribution from the Frederick Biomedical Supercomputing Center and Structural Biochemistry Program, SAIC Frederick, NCI Frederick Cancer Research and Development Center, P.O. Box B, Frederick, Maryland 21702-1201

Received September 22, 1997

Abstract: A methodology is presented for calculating relative binding free energies of enzyme−inhibitor associations in aqueous solvent. The methodology uses synthesis of semiempirical quantum chemistry to determine the protonation state of important residues in the enzyme active site, molecular mechanics to determine the gas-phase energetic contributions to the relative binding free energy, and dielectric continuum solvation to calculate electrostatic hydration contributions. The methodology is then applied to the calculation of the relative binding free energy of the inhibitors KNI-272, Ro31-8959, L-735,524, and A-77003 to HIV-1 protease and its I84V mutant. The calculated relative binding free energy is sensitive to the active-site protonation state of the aspartic acid residues of HIV-1 protease. The protonation state is inhibitor dependent. Given a particular protonation state, it was found that quantitatively accurate relative binding free energies could only be achieved when solvent effects were included. Three categories of binding were found. In the first, the change in binding free energy due to mutation is mainly due to the change in enthalpic interactions within the inhibitor−enzyme complex (Ro31-8959). In the second (L-735,524 and A-77003), the change in affinity is caused both by a change in enthalpic interactions within the enzyme and by a change in the hydration energy of the enzyme and inhibitor−enzyme complexes. In the third case (KNI-272), the change in affinity is mainly a solvent effect—it is due to changes in hydration of the enzyme only. In all cases, it was found that the I84V mutant enzyme was more stable than the wild-type enzyme. This alone (without consideration of the inhibitor−enzyme complexes) can qualitatively explain the reduction in binding affinity due to mutation.

Introduction

The object of this paper is drug inhibition of HIV-1 protease, which is an important therapeutic target in the treatment of AIDS. Emergence HIV-1 mutants[1−7] that reduce the effectiveness of inhibitors is a severe problem. One way to study the nature of this reduced effectiveness (affinity) is to determine the change in binding free energy of inhibitors (ligands) to the enzyme (protein) when the enzyme is mutated. Theoretical approaches that assess ligand−protein binding affinity[8] prior to synthesis and testing of ligands are of obvious importance in the field of structure-based drug design.[9−16] However, an understanding of the principles of ligand−protein binding thermodynamics and the calculation of ligand−protein binding affinity are difficult problems for which there is no generally satisfactory solution.[17,18] At the most general level, this paper addressess each of these issues (principles and computation) that are a vital contribution in the area of structure-based drug design.

In aqueous solution, the ligand (L)−protein (P) binding affinity or absolute binding free energy, ΔG_b, is given by

$$\Delta G_b(\mathrm{LP}) = G_{aq}(\mathrm{LP}) - G_{aq}(\mathrm{L}) - G_{aq}(\mathrm{P}) \qquad (1)$$

This quantity is difficult to calculate because it is far smaller than the individual free energies of the ligand, $G_{aq}(L)$, the protein, $G_{aq}(P)$, and the ligand−protein complex, $G_{aq}(LP)$. Accurate calculations of the reactant and product free energies are then needed in order to obtain the nearly complete cancellation of energies necessary for estimates of ΔG_b that are typically on the order of a few kilocalories per mole. However, accurate calculation of absolute free energies for complex systems such as enzymes and substrates is currently beyond

(1) Anderson, B.; Kageyama, S.; Ueno, T.; Mitsuya, H. *Tenth International Converence on AIDS and STD*, Yokohama, Japan, 1994; Abstract 516B.
(2) Ho, D. D.; Toyoshima, T.; Mo, H.; Kempf, D. J.; Norbeck, D.; Chen, C,-M.; Wideburg, N. E.; Burt, S. K.; Erickson, J. W.; Singh, M. K. *J. Virol*. **1994**, *68*, 2016.
(3) El-Farrash, M. A.; Kuroda, M. J.; Kitazaki, T.; Masuda, T.; Kato, K.; Hatanaka, M.; Harada, S. *J. Virol*. **1994**, *68*, 233.
(4) Jacobsen, H.; Yasargil, K.; Winslow, D. L.; Craig, J. C.; Krohn, A.; Duncan, I. B.; Mous, J. *Virology* **1995**, *206*, 527.
(5) Kaplan, A. H.; Zack, J. A.; Knigge, M.; Paul, D. A.; Everitt, L.; Kempf, D. J.; Norbeck, D. W.; Swanstrom, R. *J. Virol.* **1993**, *67*, 4050.
(6) Markowitz, M.; Mo, H.; Kempf, D. J.; Norbeck, D. W.; Bhat, T. N.; Erickson, J. W.; Ho, D. D. *J. Virol*. **1995**, *69*, 701.
(7) Otto, M. J.; Garber, S.; Winslow, D. L.; Reid, C. D.; Aldrich, P.; Jadhav, P. K.; Paterson, C. E.; Hodge, D. N.; Cheng, Y.-S. E. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 7543.
(8) Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. *Biophys. J.* **1997**, *72*, 1047.
(9) Huff, J. R. *J. Med. Chem.* **1991**, *34*, 2305.
(10) Meek, T. D. *J. Enzyme Inhib.* **1992**, *6*, 65.
(11) Reich, S. H.; Fuhry, M. A.; Nguyen, D.; Pino, M. J.; et al. *J. Med. Chem*. **1992**, *35*, 847.
(12) Darke, P. L.; Huyff, J. R. *Adv. Pharmacol*. **1994**, *25*, 399.
(13) Verlinde, C. L.; Merritt, E. A.; den Akker, F. V.; Kim, H.; Feil, I.; Delboni, L. F.; Mande, S. C.; Sarfaty, S.; Petra, P. H.; Hol, W. G. *Protein Sci.* **1994**, *3*, 1670.
(14) Watson, K. A.; Mitchell, E. P.; Johnson, L. N.; Son, J. C.; Bichard, C. J.; Orchard, M. G.; et al. *Biochemistry* **1994**, *33*, 5748.
(15) Kim, E. E.; Baker, C. T.; Dwyer, M. D.; Murcko, M. A.; Rao, B. G.; Tung, R. D.; Navia, M. A. *J. Am. Chem. Soc.* **1995**, *117*, 1181.
(16) Veale, C. A.; Bernstein, P. R.; Bryant, C.; Ceccarelli, C.; Damewood, J. R., Jr.; Earley, R.; Feeney, S. W.; Gomes, B. G.; et al. *J. Med. Chem*. **1995**, *38*, 98.
(17) Ajay; Murco, M. A. *J. Med. Chem*. **1995**, *38*, 4953.
(18) Babine, R. E.; Bender, S. L. *Chem. Rev*. **1997**, *97*, 1359.

S0002-7863(97)03309-X CCC: $15.00
Published on Web 08/13/1998

the scope of computational methods. Therefore, if any reasonable estimate of the binding free energy is to be obtained, there must be a significant cancellation of errors as well.

Alternatively, one may consider the relative binding free energy of two related ligands to a protein,

$$\Delta(\Delta G_b) = [G_{aq}(L_2P) - G_{aq}(L_1P)] + [G_{aq}(L_1) - G_{aq}(L_2)] \quad (2)$$

or of two related proteins to a ligand,

$$\Delta(\Delta G_b) = [G_{aq}(LP_2) - G_{aq}(LP_1)] + [G_{aq}(P_1) - G_{aq}(P_2)] \quad (3)$$

In these cases, the double subtraction may cause a nearly complete cancellation of errors, resulting in an accurate $\Delta(\Delta G_b)$. In the former case, one must treat the isolated ligands, whereas in the latter one must treat the isolated proteins. One may expect that two closely related proteins, P_1 and P_2 (different only in a single residue, for example), will be structurally related.[19] Therefore, a consistent treatment of P_1 and P_2 or LP_1 and LP_2 will result in cancellation of errors for both terms of eq 3 and term 1 of eq 2. On the other hand, small changes in L_1 and L_2 may result in structurally unrelated[20] ligands, mainly due to the fact that these changes will be large relative to the size of the ligands. Therefore, cancellation of errors is not as likely for term 2 of eq 2. Current theoretical methods may then be most accurate for situations governed by eq 3.

The drug inhibition of HIV-1 protease and its mutants is a perfect example of the type of situation governed by eq 3. HIV-1 is responsible for the posttranslational processing of the polyprotein gene products of gag and gag-pol to yield the structural proteins and enzymes of the viral particle.[21] It is a member of the aspartic proteinase family and is composed of two structurally identical monomers. The active site of the enzyme contains two aspartyl residues, one from each monomer.[22] Figure 1 shows a generalized representation of the protease with an inhibitor in the active site. HIV-1 protease is essential for infection; hence, it is a target for the design of drugs for AIDS. However, its rapid replication rate favors the emergence of drug-resistant mutants.

In this paper, consideration is given to the binding of inhibitors KNI-272,[23] Ro31-8959,[24] L-735,524,[25] and A-77003[26] (see Figures 1−3) to HIV-1 protease. The goal is to calculate the change in binding free energy of these inhibitors to the protease when the protease is altered from its wild-type form to the I84V mutant (simply changing a methyl group to hydrogen, for residues 84 and 184). The accuracy of the calculations is then determined by comparison to experimental relative binding free energies derived from the K_i ratios of Gulnick et al.[27] (these are reproduced in Table 1). Conclusions are then drawn regarding some of the details of the binding.

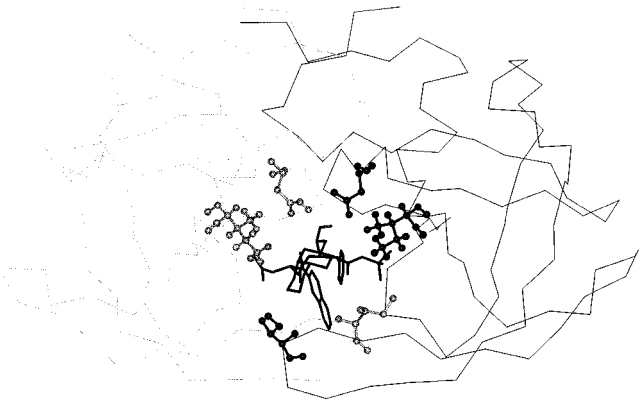

Figure 1. HIV-1 protease complexed with Ro31-8959. The active-site region is detailed, showing the inhibitor in stick format and important active-site residues in ball-and-stick format. The rest of the enzyme is shown as a linear trace along the backbone atoms. In the active site, the catalytic ASP 25 (right) and ASP 125 (left) pair is located directly above the inhibitor, the ILE 84 (right) and ILE 184 (left) pair is located to either side of the inhibitor, and the ILE 50 (left) and ILE 150 (right) pair is located below the inhibitor. The mutation studied in this paper, I84V, is obtained by changing ILE 84 and ILE 184 to valine. The ILE 50, ILE 150 pair is associated with the flap region of HIV-1.

Table 1. Experimental Relative Binding Free Energies (kcal/mol) at 298 K Derived from Experimental K_i Ratios (Gulnick et al.[27]) and Eq 12

	inhibitor			
protease	Ro31-8959	KNI-272	L-735,524	A-77003
WT (wild type)	0.0	0.0	0.0	0.0
R8Q	0.86, 0.77[d] **0.09**[e]	0.69	1.25	3.02, 2.44[b], 2.05[a] **0.97**[e]
V321	0.47, 1.17[d] **0.70**[e]	0.99	1.23, 1.19[c] **0.04**[e]	1.66, 1.20[b] **0.45**[e]
M46I	0.16	1.11	0.86	0.93, 0.20[a] **0.72**[e]
V82A	0.77	0.98	1.82	1.58
V82F	0.21	0.54	2.63	1.77
V82I	0.57, 0.23[d] **0.80**[e]	0.35	1.14, 1.12[c] **2.26**[e]	1.87, 0.00[b] **1.87**[e]
I84V	*1.04*	*2.05*	*1.36, 1.28[c]* ***0.09***[e]	*1.31*
M46I/V82F	0.0	0.06	1.14	1.15
M46I/I84V	0.89	1.99	1.82	1.76
V32I/I84V	1.57	2.99	2.60	2.77
V32I/K45I/F53L/A71V/I84V/L89M	1.56	3.29	2.38	3.04

[a] Ho et al.[2] [b] Kaplan et al.[54] [c] Vacca et al.[55] [d] Sardana et al.[56] [e] Bold face entries are the differences between the largest and smallest quoted experimental numbers.

The methodology used to calculate the relative binding free energy is a synthesis of molecular mechanics,[28] dielectric continuum solvation,[29] and surface area based methods.[30,31] It provides a broad description the ligand−protein binding process and enables the study of binding at various levels of detail.

(19) Weng, Z.; Vajda, S.; Delisi, C. *Protein Sci.* **1996**, *5*, 614.
(20) Vajda, S.; Weng, Z.; Rosenfeld, R.; Delisi, C. *Biochemistry* **1994**, *33*, 13977.
(21) Debouck, C. *Aids, Res. Hum. Retroviruses* **1992**, *8*, 153.
(22) Wlodawer, A.; Erickson, J. W. *Annu. Rev. Biochem.* **1993**, *62*, 543.
(23) Baldwin, E. T.; Bhat, T. N.; Liu, B.; Gulnik, S.; Topol, I. A.; Kiso, Y.; Mimoto, T.; Mitsuya, H.; Erickson, J. W. *Structure* **1995**, *3*, 581.
(24) 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. *Science* **1990**, *248*, 358.
(25) Chen, Z.; Li, Y.; Chen, E.; Hall, D. L.; Darke, P. L.; Culberson, C.; Shaffer, J. A.; Kuo, L. C. *J. Biol. Chem.* **1994**, *269*, 26344.
(26) Hosur, M. V.; Bhat, T. N.; Kemp, D.; Baldwin, E. T.; Liu, B.; Gulnik, S.; Wideburg, N.; Norbeck, D. W.; Erickson, J. W. *J. Am. Chem. Soc.* **1994**, *116*, 847.
(27) Gulnik, S. V.; Suvorov, L. I.; Liu, B.; Yu, B.; Anderson, B.; Mitsuya, J.; Erickson, J. W. *Biochemistry* **1995**, *34*, 9282.
(28) The Amber force field (Weiner, S. J., Kollman, P. A., Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230) was used within the InsightII 95.0/Discover 2.9.7 software package from Biosym/MSI, 9685 Scranton Rd., San Diego, CA 92121-2777. Force field parameters for standard residues are published in Weiner et al., see eq 1, Figure 1, and the appendix of this reference for details. In nonstandard cases, InsightII/Discover determines the AMBER atom types, connectivity, and partial charges. From the atom types and connectivity, the rest of the force field parameters can be obtained from the appendix of Weiner et al. See the Supporting Information for details.

Methods

Theory. The change in ligand−protease binding free energy due to mutation of the protease is given by eq 3, with P_1 and P_2 replaced by the designations P_{WT} and P_{I84V} for wild-type and I84V mutant proteins,

$$\Delta(\Delta G_b) = \Delta G_b(LP_{I84V}) - \Delta G_b(LP_{WT}) \quad (4a)$$

$$= [G_{aq}(LP_{I84V}) - G_{aq}(LP_{WT})] + [G_{aq}(P_{WT}) - G_{aq}(P_{I84V})] \quad (4b)$$

and L specifies the inhibitor (ligand). The free energies, G_{aq}, that go into eq 4b are partitioned into gas-phase and hydration components as

$$G_{aq} = G_{gas} + \Delta G_{hyd} \quad (5)$$

The gas-phase component is approximated by

$$G_{gas} = E - TS \quad (6)$$

E can be calculated using standard molecular mechanics potentials[28] and is interpreted to be the gas-phase ground-state electronic energy of the system. The entropic contribution $-TS$ is considered to arise only from side-chain degrees of freedom of the protein; main-chain entropic contributions are neglected, as well as entropic contributions due to the inhibitor. The side-chain entropy is given by

$$-TS(F) = \sum_i \left(1 - \frac{A_i(F)}{A_i(U)}\right)\sigma_i \quad (7)$$

where the sum runs over all residues of the protein, $A_i(F)$ is the van der Waals surface area of the residue i side chain in folded protein F = LP_{WT}, LP_{I84V}, P_{WT}, or P_{I84V}, and $A_i(U)$ is the van der Waals surface area of the residue i side chain in isolation, i.e., the residue alone with the rest of the protein removed. The parameter σ_i is taken from the entropy scale of Pickett and Sternberg.[31] It is an estimate of the side-chain folding entropy of residue "i" or the entropic burden incurred when residue i is taken from an exposed state at the surface of the protein and buried within the protein. Pickett and Sternberg[31] estimated the σ_i values for all 20 amino acids by analyzing the torsional preferences of surface-exposed side chains relative to those of buried side chains. This analysis was performed over a set of 50 nonhomologous protein crystal structures which included HIV-1 protease.

The hydration component of eq 5 is given by

$$\Delta G_{hyd} = \Delta G_{el} + \Delta G_{np} \quad (8)$$

where ΔG_{el} is the electrostatic hydration free energy determined in the dielectric continuum approximation[29] using boundary element methods.[32−42] ΔG_{el} represents the change in ground-state electronic energy of the system due to the water environment. More will be said about its calculation later. The nonpolar contribution, ΔG_{np}, is the hydration free energy associated with the creation of the uncharged solute cavity within the solvent. It consists of the van der Waals contribution to the solute−solvent interaction and any entropy change of the solvent due to the presence of the cavity. Changes in the solute translational, rotational, and vibrational energies or solute entropies due to hydration are explicitly neglected. ΔG_{np} is given by

$$\Delta G_{np} = 1.17 + 0.00164 \times A(\text{vdW}) \quad (9)$$

where A(vdW) is the van der Waals surface area of the solute and the constants 1.17 and 0.00164 are derived from Rashin et al.[43] Equation 9 was derived by averaging four linear fits of the hydration entropy as a function of surface area taken from the literature[43] and then adding to this a linear fit representing the nonpolar solute−solvent interaction as a function of surface area (eq 7 from ref 43).

Substituting eq 5 into eq 4b, one obtains the relative binding free energy in terms of its gas-phase and hydration components:

$$\Delta(\Delta G_b) = [G_{gas}(LP_{I84V}) - G_{gas}(LP_{WT})] + [G_{gas}(P_{WT}) - G_{gas}(P_{I84V})] + [\Delta G_{hyd}(LP_{I84V}) - \Delta G_{hyd}(LP_{WT})] + [\Delta G_{hyd}(P_{WT}) - \Delta G_{hyd}(P_{I84V})] \quad (10)$$

A further substitution of eqs 6 and 8 into eq 10 gives the relative binding free energy in terms of subcomponents of the gas-phase and hydration terms:

$$\Delta(\Delta G_b) = [E(LP_{I84V}) - E(LP_{WT})] + [E(P_{WT}) - E(P_{I84V})] - [TS(LP_{I84V}) - TS(LP_{WT})] - [TS(P_{WT}) - TS(P_{I84V})] + [\Delta G_{el}(LP_{I84V}) - \Delta G_{el}(LP_{WT})] + [\Delta G_{el}(P_{WT}) - \Delta G_{el}(P_{I84V})] + [\Delta G_{np}(LP_{I84V}) - \Delta G_{np}(LP_{WT})] + [\Delta G_{np}(P_{WT}) - \Delta G_{np}(P_{I84V})] \quad (11)$$

Equation 11 gives the most detailed partitioning of the relative binding free energy. Its first four terms comprise the gas-phase component and the last four terms comprise the hydration contribution to the free energy.

Experimentally determined relative binding affinities are often given as a ratio of K_i values. Therefore, to compare our calculated results with experiment, it is important to know how the K_i ratio is related to the relative binding free energy. This relationship is given by

$$\Delta(\Delta G_b) = RT \ln[K_i(\text{I84V})/K_i(\text{WT})] \quad (12)$$

where the quantities K_i(I84V) and K_i(WT) are equilibrium constants for the dissociation of the ligand−protein complexes. The definition of a standard state for the dissociation process is not necessary here because the ratio of equilbrium constants will remove any reference to it.

Computational Protocol. This section details the precise steps necessary for the calculation of the relative binding free energy, $\Delta(\Delta G_b)$, and its components as given by eq 11. First, the active-site protonation state is determined. It is difficult to experimentally determine a specific protonation state for each aspartic acid (ASP 25, ASP 125)[44] present in the isolated HIV-1 protease binding pocket. That is because, in the isolated enzyme, these two aspartic acids are chemically equivalent and proximate, so the protonation state of one is inexplicably correlated with that of the other. Titration studies yield

(29) (a) Chothia, C. H.; Janin, J. *Nature* **1975**, *256*, 705. (b) Gilson, M. K.; Sharp, K. A.; Honig, B. *J. Comput. Chem.* **1988**, *9*, 327. (c) Novotny, J.; Bruccoleri, R. E.; Saul, F. A. *Biochemistry* **1989**, *28*, 4735. (d) Sharp, K. A.; Honig, B. *Annu. Rev. Biophys. Chem.* **1990**, *19*, 301. (e) Karshikov, A.; Bode, W.; Tulinsky, A.; Stone, S. R. *Protein Sci.* **1992**, *1*, 727. (f) Miertus, S. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 2105. (g) Smith, K. C.; Honig, B. *Proteins: Struct. Funct. Genet.* **1994**, *18*, 119. (h) Shen, J.; Quiocho, F. A. *J. Comput. Chem.* **1995**, *16*, 445. (i) Jackson, R. M.; Sternberg, J. E. *J. Mol Biol.* **1995**, *250*, 258.

(30) Connolly, M. L. *J. Appl. Crystallogr.* **1985**, *18*, 499.

(31) Pickett, S. D.; Sternberg, J. E. *J. Mol. Biol.* **1993**, *231*, 825.

(32) Mohan, V.; Davis, M. E.; McCammon, J. A.; Pettitt, B. M. *J. Phys. Chem.* **1992**, *96*, 6428.

(33) Rashin, A. A.; Namboodiri, K. *J. Phys. Chem.* **1987**, *91*, 6003.

(34) Tawa, G. J.; Pratt, L. R. *J. Am. Chem. Soc.* **1995**, *117*, 1625.

(35) Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027 (see section III in particular).

(36) Simonson, T.; Brunger, A. T. *J. Phys. Chem.* **1994**, *98*, 4863.

(37) Zauhar, R. J.; Morgan, R. S. *J. Comput. Chem.* **1988**, *9*, 171.

(38) Yoon, B.; Lenhoff, A. M. *J. Comput. Chem.* **1990**, *11*, 1080.

(39) Corcelli, S. A.; Kress, J. D.; Pratt, L. R.; Tawa, G. J. *Pacific Symosium on Biocomputing '96*; World Scientic: River Edge, NJ, 1995; p 143.

(40) Bharadway, R.; Windemuth, A.; Sridharan, S.; Honig, B.; Nicholls, A. *J. Comput. Chem.* **1995**, *16*, 898.

(41) Pratt, L. R.; Tawa, G. J.; Hummer, G.; Garcia, A. E.; Corcelli, S. *A. Int J. Quantum Chem.* **1997**, *64*, 121.

(42) Tawa, G. J.; Pratt, L. R. In *Structure and Reactivity in Aqueous Solution: Characterization of Chemical and Biological Systems*; Cramer, C. J., Truhlar, D. G., Eds.; ACS Symposium Series 568; American Chemical Society: Washington, DC, 1994; p 60.

(43) Rashin, A. A.; Young, L.; Topol, I. A. *Biophys. Chem.* **1994**, *51*, 359.

(44) Hyland, L. J.; Tomaszek, T. A., Jr.; Meek, T. D. *Biochemistry* **1991**, *30*, 8454.

two pK_a values;[44,45] however, these are associated with the ASP 25, ASP 125 pair as a whole.

In the case of HIV-1 protease bound to an inhibitor, there is plenty of evidence to suggest that the protonation state is inhibitor dependent.[23,46−48] The work of Baldwin et al.[23] on HIV-1 protease bound to KNI-272 suggested that the carboxyl group of ASP 25 was protonated while that of ASP 125 was not. These findings were later confirmed in the NMR experiments of Wang et al.[48] In the work of Baldwin et al.,[23] the determination of the protonation state was based on a detailed analysis of the crystal structure of the complex, augmented with semiempirical quantum chemical calculations. A similar analysis is performed here for the L-735,524,[24] Ro31-8959,[25] and A-77003[26] complexes. Molecular models of each inhibitor−wild-type enzyme complex active site were built on the basis of the reported crystallographic coordinates of the complexes. The active-site models[49] included the inhibitor and six residues from the crystal structure. The six residues were ASP 25, THR 26, GLY 27, ASP 125, THR 126, and GLY 127. The GLY residues were terminated with neutral sp^2-hybridized amine groups. Energy convergence studies in aspartic proteases[49] have shown that this size of active-site model is sufficient for reliable determination of the protonation state in such enzymes (see also below). The protonation state of the aspartic acid residues was then determined by considering all single protonated configurations of the HIV-1 protease active site. For each trial configuration, the proton positions were fully optimized using the MNDO/H[50−52] method, while the non-hydrogen atom coordinates were restrained to their crystal structure values. The configuration having the lowest heat of formation was chosen as the preferred protonation state. The protonation states of HIV-1 protease complexed with the various inhibitors are shown in Figures 2 and 3. Note that the protonation state is assumed to be the same whether the inhibitor is bound to the wild-type or the I84V mutant HIV-1 protease.

Given the active-site protonation states as determined above, the structures of LP_{WT}, LP_{I84V}, P_{WT}, and P_{I84V} that go into the free energy calculations defined by eq 10 or 11 were determined using the following protocol (this optimization protocol is also displayed in Figure 4):

(i) Start with the crystal structure of the ligand complexed to the wild-type HIV-1 protease, LP_{WT} (KNI-272, ref 23, Ro31-8959, ref 24, L-735,524, ref 25, A-77003, ref 26). Fix the protonation state to that determined by the MNDO/H[50−52] method. Optimize using the AMBER[28] force field to obtain LP_{WT} and $E(LP_{WT})$.

(ii) Take the inhibitor out of LP_{WT} and reoptimize to obtain P_{WT} and $E(P_{WT})$.

(iii) Take LP_{WT} from step i and manually mutate isolucene 84 and 184 to valine, and then reoptimize to obtain LP_{I84V} and $E(LP_{I84V})$.

(iv) Take the inhibitor out of LP_{I84V} and reoptimize to obtain P_{I84V} and $E(P_{I84V})$.

All optimizations involved a three-step procedure: (a) optimize hydrogens only, (b) optimize hydrogens and side chains, and then (c) optimize the complete molecule.

A distance-dependent dielectric constant of 4R (R being the distance between interacting atomic charges and 4 being the internal dielectric constant of the molecule) was used to calculate the Coulombic portion of the energy during these optimizations. All of the nonbonded 1−4 interactions were scaled by a factor of 0.50.[28] Note that the force field parameters for standard residues are given in ref 28. Force field parameters for nonstandard cases, e.g., the protonated catalytic aspartate and the inhibitors, can be obtained from the Supporting Information referred to at the end of this work.

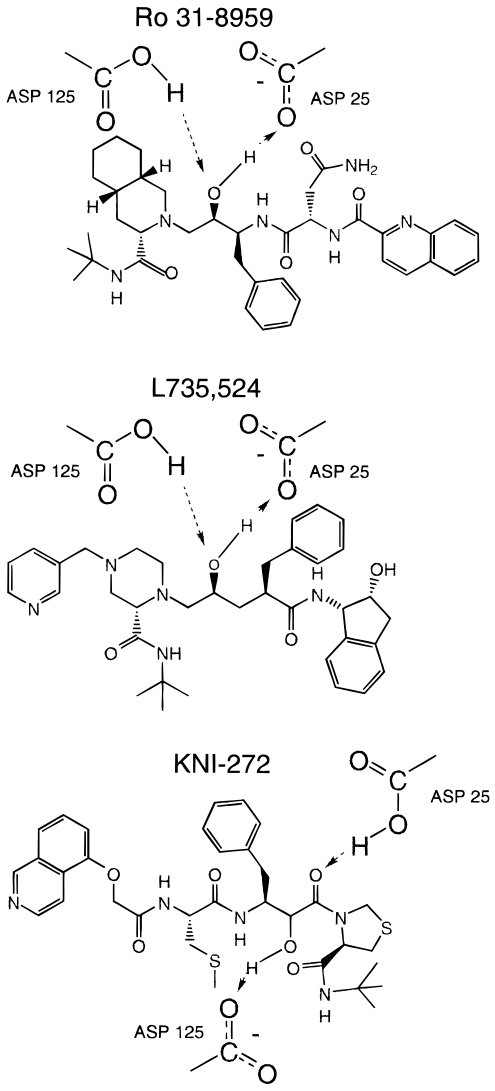

Figure 2. Active-site structures and protonation states of the HIV-1 protease inhibitors Ro31-8959, L-735,524, and KNI-272 studied in this work.

This particular optimization scheme (shown in Figure 4) is chosen so that the set of molecules considered spans a limited region of conformational space. This minimizes the chance that a spuriously large $\Delta(\Delta G_b)$ will occur due to incomplete cancellation of energies and/or errors in eq 4b. If LP_{I84V} and LP_{WT} are obtained via optimization from different crystal structures (rather than deriving one from the other, as is done here), the resultant energy difference can be large (10−50 kcal/mol or more). The large difference can arise if the crystal structures that are used as starting points for the optimizations are members of different space groups.

The energies alone comprise the first two terms of eq 11 for the relative binding free energy. With the structures in hand, the rest of eq 11 can now be calculated. First, the van der Waals surfaces are constructed for LP_{WT}, LP_{I84V}, P_{WT}, and P_{I84V} using van der Waals radii from the AMBER[28] force field database and the Connolly surface program[30] with probe radius set to 0.0. The contributions to the total van der Waals surface area, $A_i(LP_{WT})$, $A_i(LP_{I84V})$, $A_i(P_{WT})$, $A_i(P_{I84V})$, and $A_i(U)$ from the various residue side chains i are then determined, and eq 7 is used with the σ_i values from Table 6 of Pickett and Sternberg[31] to construct the side-chain entropies $TS(F)$. The energies and side-chain entropies comprise the first four terms of eq 11.

(45) Ido, E.; Ping, H. H.; Kezdy, F. J.; Tang, J. *J. Biol. Chem.* **1991**, *266*, 24349.

(46) Harte, W. E., Jr.; Beveridge, D. L. *J. Am. Chem. Soc.* **1993**, *115*, 3883.

(47) Geller, M.; Miller, M.; Swanson, S. M.; Maizel, J. *Proteins: Struct. Funct. Genet.* **1997**, *27*, 184.

(48) Wang, Y.-X.; Freedberg, D. I.; Yamazaki, T.; Wingfiled, P. T.; Stahl, S. J.; Kaufman, J. D.; Kiso, Y.; Torchia, D. A. *Biochemistry*, **1996**, *35*, 9945.

(49) Topol, I. A.; Cachau, R. E.; Burt, S. K.; Erickson, J. W. In *Aspartic Proteinases: Structure, Function, Biology, and Biomedical Implications*; Takahashi, K., Ed.; Plenum Press: New York, 1995; pp 549−554.

(50) Burstein, K. Ya.; Isaev, A. N. *Theor. Chim. Acta* **1984**, *64*, 397.

(51) Goldblum, A. *J. Comput. Chem.* **1987**, *8*, 835.

(52) The MNDO94 code implemented in UniChem, version 3.0, was used. The UniChem software environment is available from Cray Research Inc, Eagan, MN.

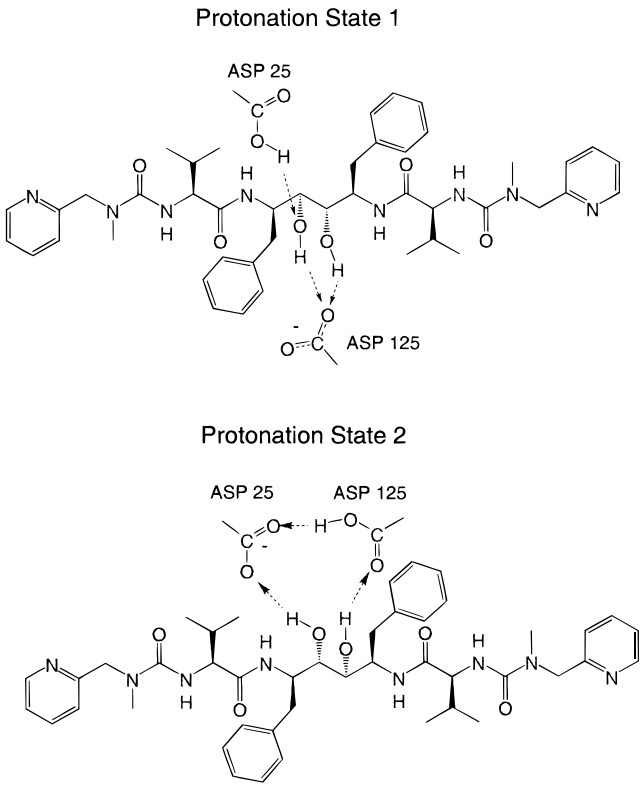

Figure 3. Active-site structure and protonation states of the HIV-1 protease inhibitor A-77003. Upper and lower structures represent the lowest and second lowest energy protonation states, respectively.

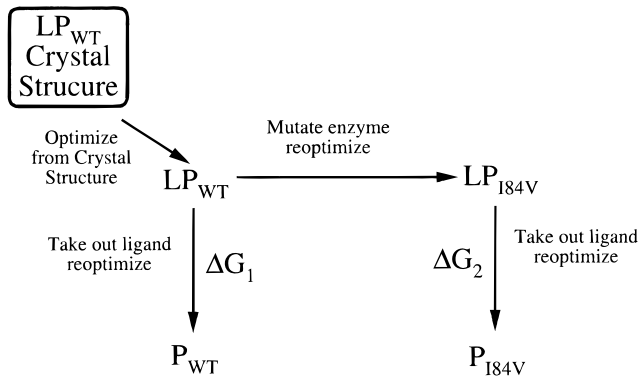

Figure 4. Scheme for generating structures used in the relative free energy calculations.

The classical electrostatic solvation terms are calculated next for LP_{WT}, LP_{I84V}, P_{WT}, and P_{I84V} solutes. The atomic charge distributions used in the electrostatic calculations come from the AMBER[28] force field database. The solvent response to the solute charge distribution is obtained by solving an integral form of the Poisson equation ($\epsilon_{out} = 78.5$, $\epsilon_{in} = 4$) using boundary element methods.[32−42] The reaction field of the solvent is obtained as a set of polarization charges placed on the solute van der Waals surface, and the electrostatic solvation free energy is calculated as

$$\Delta G_{el}(F) = \frac{1}{2}\sum_{s=1}^{N_s}\sum_{A=1}^{N_A}\left[\frac{Q_A q_s}{R_{As}}\right] \qquad (14)$$

where Q_A is an atom charge, q_s is a surface polarization charge, the sums run over all of the atoms N_A and polarization charges N_s of the solute, and F = LP_{WT}, LP_{I84V}, P_{WT}, or P_{I84V}.

Table 2. Calculated Change in Binding Free Energy (kcal/mol) Due to Mutation for the Inhibitors Ro31-8959, A-77003, L-735,524, and KNI-272 Binding to Wild-Type HIV-1 Protease and Its I84V Mutant

inhibitor	$H(\text{Prot}_2) - H(\text{Prot}_1)$[a]	$\Delta(\Delta G_b)$ calcd[b]	$\Delta(\Delta G_b)$expt[c]
Ro31-8959	9.60	**1.15**	**1.04 ± 0.53**[e]
A-77003	**1.93**	**2.44**	**1.30 ± 1.00**[e]
	(1.57)	**6.34**[d]	
L-735,524	7.90	**2.25**	**1.36 ± 0.80**[e]
KNI-272		**2.13**	**2.05**[e]

[a] $H(\text{Prot}_i)$ is the heat of formation of the inhibitor wild-type HIV-1 complex with the enzyme being in protonation state i, $i = 1$ or 2. The heat of formation is determined using the MNDO/H[50−52] method on a truncated system containing the inhibitor and six residues in the binding pocket (25−27, 125−127). The number in parentheses for A-77003 is the difference in molecular mechanics energy (AMBER[28]) between the A-77003 wild-type complex in protonation state 2 versus the A-77003 wild-type complex in protonation state 1. The whole protein is used in the molecular mechanics calculation. [b] Equation 4a was used to calculate the relative binding free energy. [c] Equation 12 was used to obtain the experimental relative binding free energy from the published K_i ratios in ref 27. See Table 1. [d] Second lowest protonation state was utilized to calculate the relative binding free energy. See Figure 3b. [e] These are rough error estimates determined by averaging the boldface entries in Table 1 for each of the inhibitors considered.

The energies, side-chain entropies, and electrostatic solvation terms comprise the first six terms of eq 11. The last two terms of eq 11 are calculated using eq 9 and the total van der Waals surface areas of each of the solutes LP_{WT}, LP_{I84V}, P_{WT}, or P_{I84V}. With the theoretically determined relative binding free energy in hand, eq 12 is used to transform experimental K_i ratios into relative binding free energies, and subsequent comparison with experiment can then be made. The experimentally determined relative binding free energies are given in Table 1; these are derived from Gulnick et al.[27] The portion of Table 1 shown in italics highlights the experimental data that are of relevance to our study.

With all of the various contributions to eq 11 in hand, one can collect the gas-phase terms together into G_{gas} and all of the hydration terms together into ΔG_{hyd} and reduce eq 11 to eq 10. A further collection of G_{gas} and ΔG_{hyd} into G_{aq} reduces eq 10 to eq 4b. Finally, at the simplest level, with no component information present, the relative binding free energy can be given by eq 4a. This hierarchy of equations, eqs 4a, 4b, 10, and 11, provides a broad description of the ligand−protein binding process. The relative binding free energies and components as based on eqs 4a, 4b, 10, and 11 are given in Tables 2−5, respectively.

Results and Discussion

The protonation states (lowest heat of formation) for HIV-1 complexed with the various inhibitors considered are shown in Figures 2 and 3. The relative heats of formation between the first and second lowest protonation states are given in column 2 of Table 2. For KNI-272 complexed with wild-type HIV-1 protease, it was previously determined[23,48] that the active site was charged−1, with protonated ASP 25 donating a hydrogen bond to the carbonyl oxygen of KNI-272 (see Figure 2). ASP 125 is not protonated but accepts a hydrogen bond from the hydroxy group of the inhibitor. When considering L-735,524 and Ro31-8959, we found only one relevant protonation state (see Figure 2). In these cases, the preferred protonation state (as determined using the MNDO/H[50−52] method) had a heat of formation more than 7 kcal/mol lower than that of any other protonation state considered (see column 2 of Table 2). In particular, for the L-735,524 and Ro31-8959 complexes with wild-type HIV-1 protease, ASP 125 is protonated, donating a hydrogen bond to the central hydroxy oxygen of the inhibitors, and ASP 25 accepts a hydrogen bond from the central hydroxy oxygen.

The situation is markedly different for HIV-1 complexed to A-77003. In this case, the heats of formation for the lowest and second lowest protonation states differ by only 1.93 kcal/mol (column 2 of Table 2). In the lowest energy configuration, ASP 25 is protonated, donating a hydrogen bond to one of the central hydroxy units on the inhibitor. ASP 125 is negatively charged, and it accepts two hydrogen bonds, one from each hydroxy unit of the inhibitor. In the second lowest protonation state, ASP 125 is protonated. It accepts a hydrogen from one of the central hydroxy units, and it donates a hydrogen to ASP 25. ASP 25 is negatively charged, and it accepts two protons, one from a central hydroxy group of the inhibitor and the other from ASP 125. These two protonation states are shown in Figure 3.

We note the number in parentheses in column 2 of Table 2 for A-77003. This is the molecular mechanics energy difference (AMBER[28]) between the protonation state 2 and protonation state 1 complexes for A-77003. The whole enzyme was used to determine this energy difference. The molecular mechanics energy difference utilizing the whole inhibitor−enzyme complex is consistent with the heat of formation difference determined using the MNDO/H method[50−52] for a truncated system which includes the inhibitor and only six residues in the active site. This gives us confidence that there will be no large errors in the calculated protonation state due to the fact that a truncated system was used in its determination.

The relative binding free energies calculated at the simplest level, using eq 4a (binding to I84V relative to that of wild-type HIV-1 protease), are given in column 3 of Table 2. The associated experimental values are given in column 4 of the table. These are taken from the I84V row of Table 1. All calculated relative binding free energies are positive, showing that the binding is worse to the I84V mutant than it is to the wild-type protease. This is in perfect agreement with the experimental findings. The calculated relative binding free energies exhibit an average deviation of 0.56 kcal/mol from experiment. In the cases of Ro31-8959 and KNI-272, the calculated results are essentially identical to experiment. The largest deviations from experiment occur for A-77003 and L-735,524; these are +1.14 and +0.89 kcal/mol, respectively. However, when one considers the precision of the experimental results (±0.5 to ±1.0 kcal/mol), the differences between the A-77003 and L-735,524 results and experiment are reduced significantly. The precision of the experimental results is obtained from the experimental relative binding free energies of Table 1. We consider each inhibitor separately. Going down each column of Table 1, we find those mutants for which there is more than one experimental estimate of the relative binding free energy. In those cases, we simply evaluate the magnitude of the difference between the largest and smallest experimental values. These differences are the boldface entries in Table 1. The differences are then averaged for each inhibitor, and the average value is the quoted error bar in column 3 of Table 2. Since the experiments were all performed under slightly different conditions, the quoted error bars are, at best, crude estimates. Nonetheless, looking at columns 3 and 4 of Table 2 and considering the experimental error bars (there are error bars associated with the theoretical values as well, yet there is no simple way to evaluate these), we can then say that the current theoretical relative binding free energies are quite close to the experimental ones for the systems considered here.

In the case of A-77003, we also report the relative binding free energy evaluated using the second lowest protonation state of Figure 3b. The value of 6.34 kcal/mol is far outside the error bars of the experimental result. This illustrates the fact that protonation state is very important in any quantitative evaluation of the relative binding free energy. The second lowest protonation state for the other systems was not considered as these were more than 7 kcal/mol higher than the lowest protonation state.

Table 3. Calculated Change in Binding Free Energy (kcal/mol) Due to Mutation for the Inhibitors Ro31-8959, A-77003, L-735,524, and KNI-272 Binding to Wild-type HIV-1 Protease and Its I84V Mutant[a]

inhibitor	$G_{aq}(LP_{I84V}) - G_{aq}(LP_{WT})$	$G_{aq}(P_{WT}) - G_{aq}(P_{I84V})$	$\Delta(\Delta G)^b$ calcd	$\Delta(\Delta G_b)^{27}$ expt
Ro31-8959	0.41	0.74	**1.15**	**1.04**
A-77003	−0.68	3.12	**2.44**	**1.30**
L-735,524	1.54	0.71	**2.25**	**1.36**
KNI-272	1.04	1.09	**2.13**	**2.05**

[a] The inhibitor−enzyme and isolated enzyme components of the relative binding free energy are given in columns 2 and 3. [b] Equation 4b was used to calculate the relative binding free energy and its components.

At a slightly more complicated level, we use eq 4b to obtain the inhibitor−enzyme and enzyme components of the relative binding free energy. These are given in columns 2 and 3 of Table 3. In the case of Ro31-8959, the wild-type complex is 0.41 kcal/mol more stable than the mutant complex; hence, the mutant complex is more likely to dissociate. The I84V enzyme is 0.74 kcal/mol more stable than the wild-type enzyme; hence, the I84V enzyme is less likely to associate with the inhibitor. Both of these factors contribute to a reduction of binding affinity of 1.15 kcal/mol when the protease is mutated. The situation is similar for both L-735,524 and KNI-272. However, A-77003 is a little different. In this case, it is the wild-type complex that is more likely to dissociate because it is 0.68 kcal/mol less stable than the I84V mutant complex. However, the isolated I84V enzyme is over 3 kcal/mol more stable than the wild-type enzyme and is far less likely to associate with the inhibitor than the wild-type enzyme. This completely destroys any chance that binding will be better to the mutant, even though the mutant complex is the most stable.

One may have noticed in column 3 of Table 3 that the free energy difference between the wild-type enzyme and the mutant enzyme is different for each case considered. This is true because, in the optimization scheme of Figure 4, the isolated wild-type enzyme and I84V enzyme structures are generated from the crystal structure of the inhibitor bound to the wild-type enzyme. Since the inhibitors are all different, the isolated enzyme structures will be different as well. Furthermore, we are calculating the binding free energy using single static structures, and no averaging is being done. That being the case, we must interpret the results of Table 3 and the tables that follow as being true for one particular set of inhibitor−enzyme and isolated enzymes conformations. There may very well be a multitude of binding conformations (in which the structures of the relevant species are all different than the ones currently used) that give similar relative binding free energies, yet the details of the binding will be slightly different.

Looking at column 3 of Table 3, we see that the free energy of the I84V enzyme is always lower than that of the wild-type enzyme, even though each entry in that column was derived using different structures for the enzymes. That being the case, it can be said that the I84V mutant enzyme is more stable than the wild-type enzyme, and this will always hinder binding to the mutant enzyme. All that can be said about the magnitude

Table 4. Gas-Phase and Solvation Components of the Change in Binding Free Energy Due to Mutation, $\Delta(\Delta G_b)$, for Ro31-8959, L-735,524, KNI-272, and A-77003 Binding to Wild-Type HIV-1 Protease and Its I84V Mutant

	$X(LP_{I84V}) - X(LP_{WT})$	$X(P_{WT}) - X(P_{I84V})$	$\Delta(\Delta X)^a$
	Ro31-8959		
$X = G_{gas}$	1.36	−0.07	1.29
$X = \Delta G_{hyd}$	−0.95	0.81	−0.14
$X = G_{aq}$	0.41	0.74	**1.15**
$= G_{gas} + \Delta G_{hyd}$			**1.04**[27]
	A-77003		
$X = G_{gas}$	0.55	5.61	6.16
$X = \Delta G_{hyd}$	−1.23	−2.49	−3.72
$X = G_{aq}$	−0.68	3.12	**2.44**
$= G_{gas} + \Delta G_{hyd}$			**1.30**[27]
	L-735,524		
$X = G_{gas}$	0.03	−0.73	−0.70
$X = \Delta G_{hyd}$	1.51	1.44	2.95
$X = G_{aq}$	1.54	0.71	**2.25**
$= G_{gas} + \Delta G_{hyd}$			**1.36**[27]
	KNI-272		
$X = G_{gas}$	0.62	−0.45	0.17
$X = \Delta G_{hyd}$	0.42	1.54	1.96
$X = G_{aq}$	1.04	1.09	**2.13**
$= G_{gas} + \Delta G_{hyd}$			**2.05**[27]

[a] $\Delta(\Delta X) = X(LP_{I84V}) - X(LP_{WT}) + X(P_{WT}) - X(P_{I84V})$, where $X = G_{gas}$, ΔG_{hyd}, or G_{aq}, see eq 10.

of the free energy difference between relevant wild-type and mutant enzymes is that it roughly varies between 0.71 and 3.12 kcal/mol.

At a yet more complicated level, we use eq 10 to obtain the inhibitor−enzyme and enzyme components of the relative binding free energy. These are further divided into their gas-phase and solvation components. These results are compiled in Table 4, where columns 2 and 3 give the inhibitor−enzyme and isolated enzyme components of the relative binding free energy. These components are subdivided into gas-phase and hydration terms. Rows 1, 2, and 3 beneath each inhibitor give the gas-phase term, the hydration term, and their sum, respectively. Column 4 of the table gives the total gas-phase and hydration components as well as the final relative binding free energy. The entries for a given row in column 4 are simply obtained by summing columns 1 and 2 for that row. Analysis of the gas-phase component of column 4 shows that, aside from the Ro31-8959 case, the gas-phase portion of the relative binding free energy is poorly correlated with experiment. For A-77003, a gas-phase relative binding free energy, $\Delta(\Delta G_{gas})$, of 6.16 kcal/mol is in error by +4.86 kcal/mol. For L-735,524, $\Delta(\Delta G_{gas}) = -0.70$, which is in error by 2.06 kcal/mol and is of the wrong sign. For KNI-272, $\Delta(\Delta G_{gas}) = 0.17$ kcal/mol, which is in error by 1.88 kcal/mol. Analysis of the hydration component of column 4 shows that it is not at all correlated with experiment, either. However, the sum of the gas-phase and hydration components correlates nicely with experiment. Furthermore, the hydration component always corrects the gas-phase relative binding free energy in a direction closer to the experimental values. Clearly, quantitative accuracy in the relative binding free energies can be obtained only if hydration effects are included.

A closer analysis shows that there are three types of binding scenarios. In the case of Ro31-8959, the hydration component of the relative binding free energy is small relative to that of the gas-phase component ($\Delta[\Delta(\Delta G_{hyd})] = -0.14$ kcal/mol compared to $\Delta(\Delta G_{gas}) = 1.29$ kcal/mol); therefore, the relative binding free energy is determined mainly by the change in the solute−solute interactions due to mutation (here, the solutes are defined as either the isolated enzymes or the inhibitor−enzyme complexes). For A-77003 and L-735,524, it appears that both the gas-phase and hydration components are large (for A-77003, $\Delta(\Delta G_{gas}) = 6.16$ kcal/mol; $\Delta[\Delta(\Delta G_{hyd})] = -3.72$ kcal/mol; for L-735,524 $\Delta(\Delta G_{gas}) = -0.70$ kcal/mol, $\Delta[\Delta(\Delta G_{hyd})] = 2.95$ kcal/mol). The relative binding free energy is then defined by the change of both solute−solute and solute−solvent interactions due to enzyme mutation. For KNI-272, the hydration component is large relative to the gas-phase component $\Delta(\Delta[\Delta G_{hyd}]) = 1.96$ kcal/mol compared to $\Delta(\Delta G_{gas}) = 0.17$ kcal/mol); therefore, the relative binding free energy is primarily determined by the change in the solute−solvent interaction due to mutation.

Table 5. Molecular Mechanics, Electrostatic Solvation, van der Waals Solvation, and Entropic Components of the Change in Binding Free Energy Due to Mutation, $\Delta(\Delta G_b)$, for the Inhibitors Ro31-8959, L-735,524, KNI-272, and A-77003 Binding to Wild-Type HIV-1 Protease and Its I84V Mutant

	$X(LP_{I84V}) - X(LP_{WT})$	$X(P_{WT}) - X(P_{I84V})$	$\Delta(\Delta X)^a$
	Ro31-8959		
$X = E$	0.97	0.08	1.05
$X = -TS_{solute}$	0.39	−0.15	0.24
$X = \Delta G_{el}$	−0.22	0.16	−0.06
$X = \Delta G_{np}$	−0.73	0.65	−0.08
$X = G_{aq}$	**0.41**	**0.74**	**1.15**[b]
			0.99[c]
			1.04[27]
	A-77003		
$X = E$	0.62	4.96	5.58
$X = -TS_{solute}$	−0.07	0.65	0.58
$X = \Delta G_{el}$	−1.18	−2.50	−3.68
$X = \Delta G_{np}$	−0.05	0.01	−0.04
$X = G_{aq}$	**−0.68**	**3.12**	**2.44**[b]
			1.90[c]
			1.30[27]
	L-735,524		
$X = E$	−0.40	−0.26	−0.66
$X = -TS_{solute}$	0.43	−0.47	−0.04
$X = \Delta G_{el}$	2.01	1.05	3.06
$X = \Delta G_{np}$	−0.50	0.39	−0.11
$X = G_{aq}$	**1.54**	**0.71**	**2.25**[b]
			2.40[c]
			1.36[27]
	KNI-272		
$X = E$	0.47	−0.08	0.39
$X = -TS_{solute}$	0.15	−0.37	−0.22
$X = \Delta G_{el}$	0.54	1.10	1.64
$X = \Delta G_{np}$	−0.12	0.44	0.32
$X = G_{aq}$	**1.04**	**1.09**	**2.13**[b]
			2.03[c]
			2.05[27]

[a] $\Delta(\Delta X) = X(LP_{I84V}) - X(LP_{WT}) + X(P_{WT}) - X(P_{I84V})$, where $X = E$, ΔG_{el}, $-TS_{solute}$, ΔG_{np}, or G_{aq}, see eq 11. [b] $G_{aq} = E - T\Delta S_{solvent} + \Delta G_{el} + \Delta H_{np} - T\Delta S_{solvent}$. [c] $G_{aq} = E + \Delta G_{el}$.

At the most complex level, we use eq 11 to present the results in the form of Table 5. Table 5 shows the inhibitor−enzyme (column 2) and enzyme components (column 3) of the relative binding free energy. These are then divided into subcomponents. The subcomponents are the molecular mechanics terms (E, row 1 beneath inhibitor), solute entropy terms ($-TS_{solute}$, row 2 beneath inhibitor), electrostatic solvation free energy terms (ΔG_{el}, row 3 beneath inhibitor), nonpolar solvation terms (ΔG_{np}, row 4 beneath inhibitor), and the total solution-phase free energy (G_{aq}, row 5 beneath inhibitor). Column 4 of the table gives the total values for all subcomponents; these are obtained by summing columns 2 and 3 for a particular subcomponent.

If we average the values of the various subcomponents in column 4 across inhibitors, we obtain $\langle\Delta(\Delta E)\rangle = 1.92$ kcal/

mol, $\langle\Delta[\Delta(\Delta G_{el})]\rangle = 2.11$ kcal/mol, $\langle -T\Delta(\Delta S_{solute})\rangle = 0.27$ kcal/mol, and $\langle\Delta[\Delta(\Delta G_{np})]\rangle = 0.13$ kcal/mol. This analysis reveals that the molecular mechanics energy terms and the electrostatic solvation terms are, by far, the most important contributors to the relative binding free energy. Conformational entropy and solute−solvent van der Waals interaction play little role in the relative binding thermodynamics for these systems.

Utilizing only the molecular mechanics energy terms and the electrostatic solvation terms, the relative binding free energies for the various inhibitors become 0.99, 1.90, 2.40, and 2.03 kcal/mol for Ro31-8959, A-77003, L-735,524, and KNI-272, respectively (see footnote c, Table 5). The average deviation is 0.43 kcal/mol, and the maximum is 1.04 kcal/mol for L-735,524. Comparing this to an average deviation of 0.56 kcal/mol and maximum of 1.14 kcal/mol when all the terms are included, we find that the results are a little better. For calculations of the type considered here, molecular mechanics augmented with electrostatic solvation is sufficient for calculating the relative binding thermodynamics.

Conclusions

HIV-1 protease is an important therapeutic target in the treatment of AIDS. Inhibitor-resistant mutants of HIV-1 (I84V considered in this paper) limit the effectiveness of drug therapy. The methodology presented in this paper is used to understand some of the details of drug resistance by *computation* of the relative binding free energy, that is, the binding free energy of an inhibitor to the mutant protease relative to that of the wild-type protease.

The method used is a combination of semiempirical quantum chemistry, molecular mechanics, and dielectric continuum solvation. It is capable of achieving quantitative accuracy in relative binding free energies. All of the results are within about 1 kcal/mol of experiment. However, there is difficulty in being more exact concerning this agreement due to experimental uncertainty in the relative binding free energies, which is on the order of 0.5−1 kcal/mol (Table 1). This quantitative accuracy in the calculations can be obtained only by first examining the protonation states of the catalytic ASP 25, ASP 125 pair located in the HIV-1 protease binding pocket. All protonation states examined were singly protonated states (one ASP protonated and neutral, the other deprotanated and negatively charged); however, the position of the proton was found to be extremely important. In the case of A-77003 and KNI-272, ASP 25 is protonated, and for Ro31-8959 and L-735,524, ASP 125 is protonated (see Figures 2 and 3). If the protonation state is incorrect, quantitative accuracy is not possible, even if the mistaken protonation state is very close in energy to the actual one, e.g., A-77003 (Figure 3 and Table 2).

Given the protonation states, the relative binding free energies must include contributions due to the solvent (Table 4). It appears that, for the systems studied here, molecular mechanics augmented with the electrostatic portion of the solvation term alone is sufficient for a quantitative description of the relative binding thermodymanics (Table 5). Apparently, van der Waals solute−solvent contributions and configurational entropy effects are small and can safely be neglected.

Analysis of the components of the relative binding free energy reveals some interesting details regarding the loss in binding affinity due to mutation. It is found that the I84V mutant enzyme in all cases is more stable than the wild-type enzyme (Table 3, column 2). Since the mutant enzyme is more stable (by 0.7−3 kcal/mol), it will have less of a tendancy to bind than the wild-type enzyme. This fact alone (independent of the inhibitor association with the enzyme) can be used to qualitatively explain the reduced binding affinity due to mutation. The 0.70−3 kcal/mol deficit must then be made up for when the inhibitor associates with the enzyme during complexation. However, in all cases it does not (see Table 3 column 2).

Three types of binding scenarios are found: solute driven, solute−solvent driven, and solvent driven. The reduced binding affinity of Ro31-8959 is solute driven; that is, the solvent is not a player, and the reduction in affinity is due to a change in the enthalpic contacts between inhibitor and enzyme due to mutation. For KNI-272, the reduction in binding affinity is solvent driven. It has nothing to do with a change in enthalpic contacts between the inhibitor and the enzyme but rather has to do with a change in hydration state of the enzymes and the enzyme inhibitor complexes (Table 3). For A-77003 and L-735,524, the reduction in binding affinity has significant solute and solute−solvent components.

The current methodology has provided some useful insights into the nature of drug resistance (in a thermodynamic sense). However, it will be useful to relate the thermodynamics to important structural aspects of the binding. Also, the I84V mutant enzyme is only one of many that may occur.[53] The current methodology will be used to obtain theoretical estimates of all of the relative binding free energies shown in Table 1 to determine precisely where the methodology works and where it does not. This information will then be used to make improvements. The final method will be useful not only in gaining a conceptual understanding of how the mutation process works but also as a quantitative tool for screening new HIV inhibitors to evaluate possible drug resistance patterns.

Acknowledgment. We thank the staff and administration of the Frederick Biomedical Supercomputing Center and the National Cancer Institute for their support of this project. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organization imply endorsement by the U. S. Government.

Supporting Information Available: AMBER atom types, partial charges, and connectivity for nonstandard residues, e.g., protonated catalytic aspartic acids and the inhibitors Ro31-8959, L735-524, KNI-272, and A-77003; discussion on how to obtain all other force field parameters for nonstandard cases, based on the atom type and connectivity (14 pages, print/PDF). See any current masthead page for ordering information and Web access instructions.

JA9733090

(53) Erickson, J. W. *Nature Struct. Biol.* **1995**, *2*, 523.

(54) Kaplan, A. H.; Michael, S. F.; Wehbie, R. S.; Knigge, M. F.; Paul, D. A.; Everitt, L.; Kempf, D. J.; Norbeck, D. W.; Erickson, J. W.; Swanstrom, R. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 5597.

(55) Vacca, J. P.; Dorsey, B. D.; Schleif, W. A.; Levin, R. B.; McDaniel, S. L.; Darke, P. L.; Zugai, J.; Quintero, J. C.; Blahy, O. M.; Roth, E.; Sardana, V. V.; Schlabach, A. J.; Graham, P. I.; Condra, J. H.; Gotlib, L.; Holloway, M. K.; Lin, J.; Chen, I.-W.; Vastag, K.; Ostovic, D.; Anderson, P. S.; Emini, E. A.; Huff, J. R. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 4096.

(56) Sardana, V. V.; Schlabach, A. J.; Graham, P.; Bush, B. L.; Condra, J. H.; Culberson, J. C.; Gotlib, L.; Graham, D. J.; Kohl, N. E.; LaFemina, R. L.; Schneider, C. L.; Wolanski, B. S.; Wolfgang, J. A.; Emini, E. A. *Biochemistry* **1994**, *33*, 2004.