Prediction of the Protonation State of the Active Site Aspartyl Residues in HIV-1 Protease–Inhibitor Complexes via Molecular Dynamics Simulation

William E. Harte, Jr.,* and David L. Beveridge†

Contribution from the Bristol-Myers Squibb Pharmaceutical Research Institute, 5 Research Parkway, Wallingford, Connecticut 06492, and the Chemistry Department, Hall-Atwater Laboratories, Wesleyan University, Middletown, Connecticut 06457

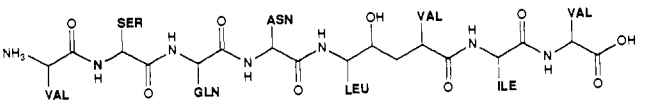
Received December 4, 1992

Abstract: A dynamical model for the structure of two human immunodeficiency virus-1 protease (HIV-1 PR)–inhibitor complexes, H-Val-Ser-Gln-Asn-LeuΨ[CH(OH)CH_2]Val-Ile-Val-OH (U85548e) and *N*-acetyl-Thr-Ile-Nle-Ψ[CH_2-NH]-Nle-Gln-Arg-amide (MVT-101), has been developed on the basis of molecular dynamics simulations. The study examines *all* possibilities of protonation state and compares the results with the corresponding crystal structures. The results indicate that only one of the four states for each inhibitor agrees well with the crystallographic data. This provides a theoretical prediction of the protonation state of the catalytic aspartic acids for each complex and may be useful for the evaluation of potential therapeutic targets.

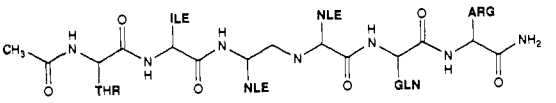
Introduction

HIV-1 protease (HIV-1 PR), responsible for polyprotein processing as new virions bud from an infected cell, is an essential enzyme in the lifecycle of the human immunodeficiency virus.[1] HIV-1 PR is classified as an aspartyl protease, the functional region of the enzyme active site consisting of a pair of amino acid triads beginning with an aspartyl residue. The molecular structure of HIV-1 PR apoenzyme determined by X-ray crystallography was first reported in the late 1980s[2–4] and now forms the basis for an extensive effort in the rational design of specific inhibitors that may be of use in AIDS therapy. This protein system quickly became the most structurally characterized enzyme with five HIV-1 PR inhibitor complexes reported[5–9] and more than 100 additional structures that remain proprietary.[10]

A molecular dynamics (MD) simulation on a hydrated model of the HIV-1 PR dimer was recently reported from this laboratory[11–13] and yielded a reasonable account of the crystal geometry, while providing evidence for dynamical correlations of possible functional significance within the enzyme structure. This simulation of the hydrated apoenzyme was based on the assumption that the carboxylate group of one of the aspartyl residues was anionic, while the other is neutral (protonated), henceforth denoted the (–1,0) state. This assumption was based on the interaction of aspartyl proteases with their substrates and has formed the basis of several hypotheses of catalysis[7,14,15] but does not address the protonation state in the active site of an enzyme–inhibitor complex. A knowledge of the ionization state within the active site is critical to an understanding of factors that are responsible for inhibitor binding at the submolecular level. In extending our MD studies to the leading HIV-1 PR inhibitor complexes, an initial simulation on HIV-1PR/U85548e assuming the (–1,0) protonation state resulted in a structure in which the position of the inhibitor in the binding pocket deviated considerably from that found in the crystal structure. This result led us to inquire further about the protonation state of the aspartyl groups as a possible source of the discrepancy.



U85548e



MVT-101

Figure 1. Sequences of U85548e and MVT-101.

* Address correspondence to author at Bristol-Myers Squibb Pharmaceutical Research Institute, 5 Research Parkway, Wallingford, CT 06492.
† Wesleyan University.

(1) Kohl, N.; Emini, E. A.; Schlief, W. A.; Davis, L. J.; Heimbach, J.; Dixon, R. A. F.; Scolnik, E. M.; Sigal, I. S. *Proc. Natl. Acad. Sci. U.S.A* **1988**, *85*, 4686–4690.
(2) Navia, M. A.; Fitzgerald, P. M. D.; McKeever, B. M.; Leu, C.; Heimbach, J. C.; Herber, W. K.; Segal, I. S.; Darke, P. L.; Springer, J. P. **1989**, *337*, 615–620.
(3) Wlodawer, A.; Miller, M.; Jaskolski, M.; Sathyanarayana, B. K.; Baldwin, E.; Weber, I. T.; Selk, L. M.; Clawson, L.; Schneider, J.; Kent, S. B. H. **1989**, *245*, 616–621.
(4) Lipatto, R.; Blundell, T.; Hemmings, A.; Overington, J.; Wilderspin, A.; Wood, S.; Merson, J. R.; Whittle, P. J.; Danley, D. E.; Geogharan, K. F.; Hawrylik, S. J.; Lee, S. E.; Scheld, K. G.; Hobart, P. M. **1989**, *342*, 299–301.
(5) Miller, M.; Schneider, J.; Sathyanarayana, B. K.; Toth, M. V.; Marshall, G. R.; Clawson, L.; Selk, L.; Kent, S. B. H.; Wlodawer, A. **1989**, *246*, 1149–1151.
(6) Swain, A. L.; Miller, M.; Green, J.; Rich, D. H.; Schneider, J.; Kent, S. B. H.; Wlodawer, A. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 8805–8809.
(7) Jaskolski, M.; Tomasselli, A. G.; Sawyer, T. K.; Staples, D. G.; Heinrikson, R. L.; Schneider, J.; Kent, S. B. H.; Wlodawer, A. *Biochemistry* **1991**, *30*, 1600–1609.
(8) Fitzgerald, P. M. D.; McKeever, B. M.; VanMiddleworth, J. F.; Springer, J. P.; Heimbach, J. C.; Leu, C.; Herbert, W. K.; Dixon, R. A. F.; Darke, P. L. *J. Bio. Chem.* **1990**, *265*, 14209–14219.
(9) Erickson, J.; Neidhart, D. J.; VanDrie, J.; Kempf, D. J.; Wang, X. C.; Norbeck, D. W.; Plattner, J. J.; Rittenhouse, J. W.; Turon, M.; Widenburg, N.; Kohlbrenner, W. E.; Simmer, R.; Helfrich, R.; Paul, D. A.; Knigge, M. *Science* **1990**, *249*, 527–533.
(10) Tomasselli, A. G.; Howe, W. J.; Sawyer, T. K.; Wlodawer, A.; Heinrikson, T. L. *Chimicaoggi* **1991**, May.
(11) Harte, W. E., Jr.; Swaminathan, S.; Mansuri, M. M.; Martin, J. C.; Rosenberg, I. E.; Beveridge, D. L. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 8864–8868.
(12) Swaminathan, S.; Harte, W. E., Jr.; Beveridge, D. L. *J. Am. Chem. Soc.* **1990**, *113*, 2717–2721.
(13) Harte, W. E., Jr.; Swaminathan, S.; Beveridge, D. L. *Proteins: Struct., Funct., Genet.* **1992**, *13*, 175–194.
(14) Suguna, K.; Padlan, E. A.; Simth, C. W.; Carlson, W. D.; D., D. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 7009–7013.
(15) Pearl, L. H. *FEBS Lett.* **1987**, *214*, 8–12.

0002-7863/93/1515-3883$04.00/0

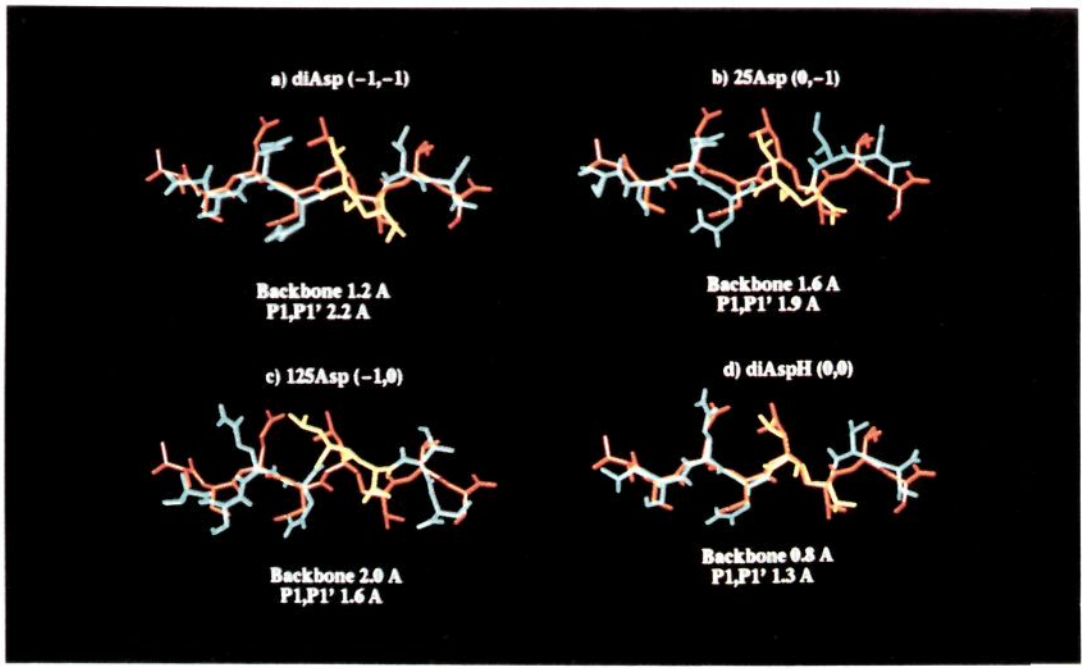


Figure 2. Superposition of the average structures from each HIV/U85548e MD simulation (cyan) with the X-ray structure (red, ref 7). (Protease backbone not shown.) The P1/P1′ site is highlighted in yellow. The RMS deviation is reported for the backbone atoms of the inhibitor as well as all atoms in the P1/P1′ site. Structural comparison with the crystallographic model was accomplished by a RMS superposition of the protease backbone residues (1–99,101–199).

There are four possible protonation states of the two catalytic aspartates (Asp25 and Asp125 in the notation of ref 3): the dianionic form (–1,–1), the two monoanionic forms (–1,0) and (0,–1), in which one of the catalytic Asp groups is anionic and the other protonated, and the diprotonated or neutral form (0,0). We report herein a series of MD simulations on solvated HIV-1 PR complexed with the substrate-based inhibitors, H-Val-Ser-Gln-Asn-LeuΨ[CH(OH)CH_2]Val-Ile-Val-OH (U85548e, Figure 1) and *N*-acetyl-Thr-Ile-Nle-Ψ[CH_2-NH]-Nle-Gln-Arg-amide (MVT-101, Figure 1), examining *all* possibilities of protonation state and comparing the results with the corresponding crystal structures. The results indicate that only one of the four states for each inhibitor agrees well with the crystallographic data. This provides a theoretical prediction of the protonation state of the catalytic aspartic acids for each complex.

Calculations and Results

A simulation protocol analogous to that utilized previously for the uncomplexed HIV-1 PR was followed rigorously for each of the eight simulations of this study. The point of departure for each MD was either the X-ray crystal structure of HIV-1 PR complexed with U85548e[7] or the complex with MVT-101.[5] Molecular dynamics calculations in this study were performed with the Monte Carlo (MC) and MD simulation program WESDYN 1.0,[16] using the GROMOS86 force field[17] and the SPC model for water.[18] All ionizable residues were assumed to be in the form prevalent at physiological pH, with the obvious exception of Asp25 and Asp125. Switching functions were used to make the long-range nonbonded interactions go smoothly to zero between 7.5 and 8.5 Å and applied on a group by group basis to avoid artificially splitting dipoles. The protein was solvated with approximately 7000 molecules of water in a hexagonal prism cell and treated under periodic boundary conditions in the MD. The volume of the system was chosen to produce a solvent density of 1 g/cc. The simulations in each case involved an initial solvent relaxation of 3 K passes of Metropolis MC (~21 M configurations), followed by 50 steps of conjugate gradient minimization on the total system. The MD involved heating to 300 K over 1.5 ps, a Gaussian equilibration step of 2.5 ps, and 96 ps of free MD under a temperature window of 5 K. No rescaling was required beyond 25 ps into any of the runs, indicative of intrinsic stability in the calculated dynamical structure. Analysis of results was based on the interval between 50 and 100 ps of the trajectory and included the determination of the average structure over the last 50 ps of the trajectory along with the isotropic temperature factors for each atom. Comparison of the MD results with the corresponding crystal structure data was accomplished by a RMS superposition of the protease backbone residues (1–99,101–199).

U85548e. All four simulations of the HIV-1 PR-U85548e complex resulted in protein structures that agreed well with the corresponding crystal forms. The RMS deviation of the protease backbone was between 1.0 and 1.2 Å for all models of the protonation state, consistent with that obtained previously on the simulation on uncomplexed HIV-1 PR. The calculated temperature factors for the protease backbone were also in good accord in each case. Comparisons of the average MD structure for the inhibitor geometry superimposed on the inhibitor structure determined from crystallography are shown in Figure 2. The forms (–1,–1), (–1,0), and (0,–1), Figure 2, parts a–c, respectively, show evident deviations from the corresponding crystal structure. Only the diprotonated form (0,0), Figure 2d, coincides well with the experimentally observed inhibitor geometry. The residual RMS deviation of the inhibitor backbone in this case is 0.8 Å, 0.36 Å less than the next closest case and 0.9 Å less for the all atom comparison within the active site (P1 and P1′ sites). The distance of close contacts in the active site are collected in Table I. The diprotonated model not only reproduces the inhibitor geometry experimentally observed in this critical region but also provides the best overall description of the protease active site.

MVT-101. The four MD simulations for the various aspartyl group protonation states of the HIV-1 PR/MVT-101 complex

(16) Swaminathan, S. Wesdyn, Wesleyan University, 1990.
(17) van Gunsteren, W. F.; Berendsen, H. J. C. Gromos, University of Groningen, 1986.
(18) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces*; Pullman, B., Ed.; D. Reidel: Dordrecht, 1981.

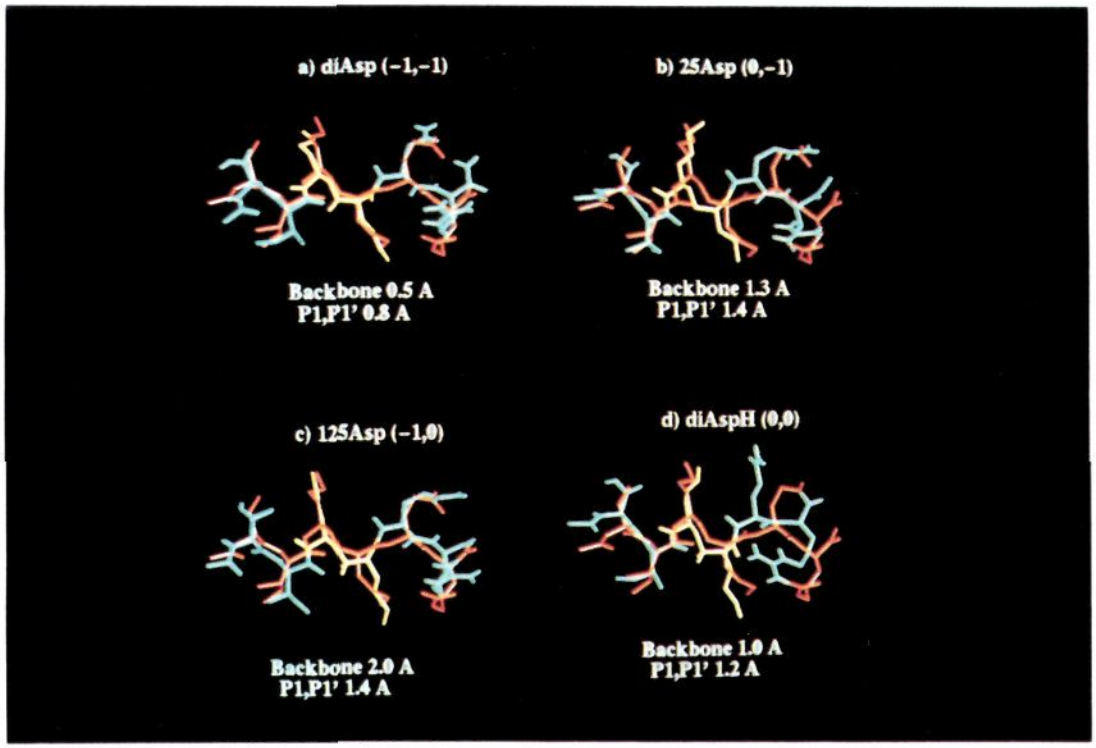


Figure 3. Superposition of the average structures from each HIV/MVT-101 MD simulation (cyan) with the X-ray structure (red, ref 5). (Protease backbone not shown.) The P1/P1′ site is highlighted in yellow. The RMS deviation is reported for the backbone atoms of the inhibitor as well as all atoms in the P1/P1′ site. Structural comparison with the crystallographic model was accomplished by a RMS superposition of the protease backbone residues (1–99,101–199).

Table I. Close Contacts in the Active Site

	distance (Å)				
contact	exp	diAsp	25Asp	125Asp	diAspH
Protease–Inhibitor					
OS(P1)–OD1(Asp25)	2.41	3.80	4.64	3.03	2.95
OS(P1)–OD2(Asp25)	2.87	2.76	2.66	3.40	2.86
OS(P1)–OD1(Asp125)	3.07	5.77	5.25	4.13	3.00
OS(P1)–OD2(Asp125)	2.72	7.70	4.82	2.68	2.68
N(P1)–O(Gly27)	3.51	5.74	5.45	4.88	2.88
N(P2′)–O(Gly127)	3.37	7.13	6.75	5.37	4.61
Protease–Protease					
OD1(Asp25)–OD1(Asp125)	2.79	5.02	2.95	2.90	2.93
N(Gly27)–OD1(Asp25)	3.28	3.86	3.03	3.11	3.19
N(Gly127)–OD1(Asp125)	2.78	2.96	3.19	3.12	3.22
N(Thr126)–OG1(Thr26)	3.01	4.22	3.30	3.07	2.97
N(Thr26)–OG1(Thr126)	3.15	4.58	3.05	3.45	2.99
OG1(Thr26)–O(Leu124)	2.94	3.42	2.74	2.84	2.65
OG1(Thr126)–O(Leu24)	2.48	3.70	2.75	2.76	2.58
N(Ala28)–O(Asp25)	3.51	4.56	3.96	4.73	3.28
N(Ala128)–O(Asp125)	2.68	5.97	4.61	4.06	3.21

also produced protease structures that agreed well with the crystal structure, with an RMS deviation for the protease backbone between 1.0 and 1.2 Å. The calculated temperature factors for the protease backbone were also in good accord for each of the systems. Comparisons of the average inhibitor geometries obtained from each MD simulation with the polypeptide backbone of the inhibitor from the crystal structure is shown in Figure 3. The results based on the forms (0,0), (–1,0), and (0,–1), Figure 3, parts b–d, respectively, show considerable deviation with the crystal structure data. Only the dianionic model (–1,–1), Figure 3a, accurately reproduces the experimentally observed inhibitor geometry. The residual RMS deviation of the inhibitor backbone and all atom comparison within the active site (P1 and P1′ sites) are both less than 1.0 Å overall and 0.5 Å better than the next closest model.

Discussion

The simulation results clearly support the diprotonated state (0,0) for U85548e and the dianionic state (–1,–1) for MVT-101. As with any theoretical result, however, the question of plausibility must be addressed. The pK_a of an aspartic acid side chain in aqueous solution is typically 3.9. Both crystals were grown from a solution buffered at pH 5.4, yet the model for U85548e has both aspartic acids protonated some 1.5 pH units above the typical pK_a for this side chain. Additionally, a different substrate-based inhibitor, MVT-101, grown from identical conditions to U85548e proports it to be in a dianionic state. Is it reasonable to have a diprotonated species at pH 5.4 as well as to observe two different protonation states that were determined from experimentally similar systems? Both of these concerns can be addressed by examining the local environment surrounding the active site.

The environment surrounding the active site in each of these enzyme–inhibitor complexes is clearly hydrophobic with the inhibitor side chains P1 and P1′ interacting with Pro 81/181 and Val 82/182. Neither the crystal structure nor the dynamical models have any ordered water within 6 Å of the aspartates. Any charge on the acid side chains would presumably exist only if a compensatory interaction with a complementary charge on the inhibitor was possible. The hydroxyethylene peptide isostere of **U85548e** (Figure 1) contains no ionizable groups to neutralize a resultant charge. It is therefore plausible that both acids remain protonated. Furthermore, the hydroxyethylene isostere resembles the transition state in the hydrolysis of the amide bond. While there are several postulates of a mechanism,[7,14,15] each contains a diprotonated acid species at the transition state; one acid accepting a proton from the attacking water molecule, and one acid ready to donate a proton to the amide in order to make it a better leaving group.

Conversely, MVT-101 inhibitor complex contains a secondary amine as the peptide isostere (Figure 1), which will carry a positive

charge, thus creating a drastically different environment than the charge neutral U85548e. There is no reason to expect the aspartyl side chains to have the same protonation state in such different surroundings. Thus the observation of different ionization states between the two enzyme–inhibitor complexes should not be unexpected.

Finally, these results have potentially significant implications to theoreticians engaging in free energy calculations aimed at accounting for the relative affinities of various HIV-Pr–inhibitor complexes. Several investigators have examined inhibitors in an effort to elucidate factors that lead to binding;[19-22] however none of these reports demonstrate that the structure of the resultant model agrees with the experimentally determined crystal structure This research clearly points out the need to validate the assumptions with all of the experimental evidence that is available.

We have developed the first model that demonstrates substantive agreement with the available structural data while providing leading evidence for the protonation state in the active site and may be useful for the evaluation of potential therapeutic targets.

(19) Reddy, M. R.; Viswanadhan, V. N.; Weinstein, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 10287–10291.
(20) Ferguson, D. M.; Radmer, R. J.; Kollman, P. A. *J. Med. Chem.* **1991**, *34*, 2654–2659.
(21) Tropsha, A.; Hermans, J. *Prot. Eng.* **1992**, *5*, 29–33.
(22) Rao, B. G.; Tilton, R. F.; Singh, U. C. *J. Am. Chem. Soc.* **1992**, *114*, 4447–4452.