The role of hydrogen bonding in the enzymatic reaction catalyzed by HIV-1 protease

JOANNA TRYLSKA,^{1,2} PAWEŁ GROCHOWSKI,² AND J. ANDREW McCAMMON^{1,3}

¹Department of Chemistry and Biochemistry and Center for Theoretical Biological Physics, University of California at San Diego, La Jolla, California 92093, USA

²Interdisciplinary Centre for Mathematical and Computational Modelling, Warsaw University, Pawińskiego 5a, 02-106 Warsaw, Poland

³Howard Hughes Medical Institute and Department of Pharmacology, University of California at San Diego, La Jolla, California 92093, USA

(RECEIVED August 13, 2003; FINAL REVISION October 9, 2003; ACCEPTED October 10, 2003)

Abstract

The hydrogen-bond network in various stages of the enzymatic reaction catalyzed by HIV-1 protease was studied through quantum-classical molecular dynamics simulations. The approximate valence bond method was applied to the active site atoms participating directly in the rearrangement of chemical bonds. The rest of the protein with explicit solvent was treated with a classical molecular mechanics model. Two possible mechanisms were studied, general-acid/general-base (GA/GB) with Asp 25 protonated at the inner oxygen, and a direct nucleophilic attack by Asp 25. Strong hydrogen bonds leading to spontaneous proton transfers were observed in both reaction paths. A single-well hydrogen bond was formed between the peptide nitrogen and outer oxygen of Asp 125. The proton was diffusely distributed with an average central position and transferred back and forth on a picosecond scale. In both mechanisms, this interaction helped change the peptide-bond hybridization, increased the partial charge on peptidyl carbon, and in the GA/GB mechanism, helped deprotonate the water molecule. The inner oxygens of the aspartic dyad formed a low-barrier, but asymmetric hydrogen bond; the proton was not positioned midway and made a slightly elongated covalent bond, transferring from one to the other aspartate. In the GA/GB mechanism both aspartates may help deprotonate the water molecule. We observed the breakage of the peptide bond and found that the protonation of the peptidyl amine group was essential for the peptide-bond cleavage. In studies of the direct nucleophilic mechanism, the peptide carbon of the substrate and oxygen of Asp 25 approached as close as 2.3 Å.

Keywords: enzymatic reaction; valence bond method; hydrogen bond; proton transfer; molecular dynamics simulations

The human immunodeficiency virus type 1 protease (HIV-1 PR) is one of the three enzymes encoded by the viral genome. It exhibits its main role during the viral maturation, cleaving the peptide bonds of its polyprotein precursors. The enzyme itself has been studied for many years, as it was observed that virions that lack HIV-1 PR are noninfectious

(Kohl et al. 1988). Inhibitors of HIV-1 PR have been sought for use in chemotherapy of AIDS. A few have been approved for clinical therapy (de Clercq 2002; Kurup et al. 2003), but despite their high selectivity, they induce side effects, and drug-resistant strains of the virus are not uncommon. Therefore, there still is a constant demand for new inhibitors and for understanding how this enzyme really works.

HIV-1 PR is active as a 198-amino acid homodimer and belongs to a class of aspartic proteases containing a catalytic triad, Asp-Thr-Gly, in each monomer. Two closely apposed aspartic acids (nos. 25 and 125 [residues of the second

Reprint requests to: Joanna Trylska, Department of Chemistry and Biochemistry, University of California at San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA; e-mail: jtrylska@mccammon.ucsd.edu; fax: (858) 534-7042.

Article and publication are at <http://www.proteinscience.org/cgi/doi/10.1110/ps.03372304>.

monomer are often numbered 101–199)) are believed to be crucial for functioning of the enzyme (Fig. 1). On the basis of experimental and theoretical work, there have been several proposals of the reaction mechanism (Hyland et al. 1991a,b; Rodriguez et al. 1993; Chatfield and Brooks 1995; Lee et al. 1996; Silva et al. 1996; Chatfield et al. 1998; Park et al. 2000; Northrop 2001). However, there is not yet proof of which is the right one. Most studies indicate a general-acid/general-base (GA/GB) mechanism for the peptide-bond cleavage, in which hybridization of the peptide carbon is changed from sp^2 to sp^3 through a nucleophilic attack by a lytic water molecule (Fig. 2). One of the aspartates helps in ionizing the water molecule, taking away its proton. The resulting hydroxy anion attacks and bonds covalently to the peptide carbon, initiating the C–N bond breakage.

The aim of this project is to continue the investigation of the reaction mechanism and hydrogen-bond network in the active site of HIV-1 PR. By comparison with our previous study (Trylska et al. 2002), the current one focuses also on the GA/GB mechanism with a different starting protonation of the catalytic aspartates shown in Figure 3. The figure also shows the atom names used herein. This initial state was assumed in Piana and Carloni (2000) and Northrop (2001), and it differs from the starting conformation we used earlier, by the protonation of the OD1 atom instead of OD2 on Asp 25. The nucleophilic attack and the dissociation of the water molecule occur in the same manner, followed by the HW2 proton transfer onto the peptide nitrogen. The difference in the reaction path is that the HD1 proton stays hydrogen bonded to Asp 125:OD1 or transfers onto that atom, but does not transfer onto carbonyl oxygen of the peptide. Moreover, in this work, we tested a reaction mechanism that does not involve the water molecule, that is, a direct nucleophilic attack is carried out by one of the aspartic acids (Fig. 4).

Figure 1. Backbone model of the secondary structure of HIV-1 protease. β -strands are shown as ribbons, α -helices as cylinders.

For both mechanisms, we were interested in the nature of hydrogen bonds that are formed during the course of the reaction. Are they weak and long, or short and strong with a covalent character? Which ones are centered, which asymmetric, and what is their role in each reaction stage? The role of hydrogen bonds in enzymatic reactions is a well-known fact, but the proposal that low-barrier hydrogen bonds (LBHB) drive and contribute to enzymatic catalysis was first suggested in Gerlt and Gassman (1993), Cleland and Krevoy (1994), and Frey et al. (1994). This was followed by many controversies (Scheiner and Kar 1995; Warshel et al. 1995; Guthrie 1996; Warshel and Papazyan 1996), but the presence of LBHBs was proved experimentally in various enzymatic systems; for example, serine protease, ketosteroid isomerase, citrate synthase, and mandelate racemase (for review, see Cleland et al. 1998, and references therein). Recently, formation of a LBHB between the inner oxygens of aspartates in the HIV-1 PR active site was proposed in Northrop (2001) on the basis of the evidence of isotope effects and in Piana and Carloni (2000) on the basis of *ab initio* molecular dynamics (MD) study.

Classical MD simulations are not able to answer the above questions, because the molecular mechanics (MM) force field not only is unable to determine the type of hydrogen bonds, but also cannot account for the formation and breakage of bonds in the course of the reaction. Therefore, we used a quantum-classical MD/AVB method developed and described in detail in our previous work (Trylska et al. 2001, 2002). By this method, we were able to determine the formation of LBHB, their length, and the distribution of proton positions between the hydrogen-bond acceptor and donor, and discriminate between the single-well or asymmetric LBHB. In this study, we focused on the spontaneous proton transfer processes, and decomposition of the intermediate product of the nucleophilic attack occurring in a shorter time scale compared with the nucleophilic attack itself. The latter process, which is the rate-limiting step, was activated by an extra potential force as in the umbrella sampling technique (Torrie and Valleau 1974), but we assured that the steering force affected only one degree of freedom, and the process was conducted very slowly in comparison with atomic motions. The limitations of this approach do not allow one to estimate the energy barrier and time scale of the nucleophilic attack. However, all other processes in the system, that is, proton transfers, occur spontaneously, so we were able to determine the sequence of proton transfers and the types of hydrogen bonds formed in the active site.

Simulation methods

AVB active site representation

The approximate valence bond (AVB) method used for the HIV-1 PR active site has been described previously (Gro-

Figure 2. Reaction center showing GA/GB mechanism with the initial proton position on Asp 25:OD2. (A,B) Following the dissociation of H₂O, OH[−] acts as a nucleophile on the peptide carbon; (B) two proton transfers, from Asp 25 onto carbonyl O and from Asp 125 onto peptide N; (C) C–N bond breakage and return of proton to Asp 25; (D) final state. In the text, OD1 oxygens are named “inner” and OD2, “outer.”

chowski et al. 1996; Trylska et al. 2001, 2002). The modeled system is decomposed into two regions, AVB quantum part (16 or 18 atoms) and the rest of the protein with explicit solvent (~18700 atoms). The interactions in the latter region are described with a classical MM force field, and the coupling between the regions is accounted for. The electronic ground state for the AVB region is expressed as a linear combination of the electronic wave functions Ψ_I corresponding to various valence bond structures differing by dissociation or association of one chemical bond:

$$\Phi = \sum_{I=1}^N c_I \Psi_I \quad (1)$$

in which N is the number of valence bond structures, c_I are the combination coefficients and $\sum_I c_I^2 = 1$.

By applying the variational principle, one obtains a set of linear equations:

$$\sum_{J=1}^N (H_{IJ} - E S_{IJ}) c_J = 0, \quad I = 1, \dots, N \quad (2)$$

in which E is the ground-state energy, $H_{IJ} = \langle \Psi_I | \hat{H} | \Psi_J \rangle$ and $S_{IJ} = \langle \Psi_I | \Psi_J \rangle$ are the elements of the Hamiltonian and overlap matrices, respectively. In our method, following Warshel (Warshel and Weber 1980; Warshel 1991), the validity of the above equations is assumed, but Ψ_I do not need any mathematical representation.

The molecular fragments that constitute the valence bond structures are the side chains of catalytic aspartates 25 and 125, water molecule, and the cleaved peptide bond between methionines 203 and 204 (the region described with AVB is shown in Figs. 2–4). Depending on the type of the mechanism, or the stage of the reaction we wanted to model, appropriate fragments were chosen to build the valence bond structures (Tables 1 and 6, below). The number of structures serves as a basis set, allowing the system to evolve in a certain energy landscape. The potential energy surface of the AVB region is described by the lowest eigenvalue of the Hamiltonian matrix. The elements of the AVB Hamiltonian are approximated by analytical functions, depending on the positions of the atomic nuclei. The parameters of these functions were determined by fitting to the density functional (DFT) calculations performed for small molecular systems with the B3LYP exchange and

Figure 3. Reaction center showing GA/GB mechanism with the the initial proton position on Asp 125:OD1. (A) Following the dissociation of H₂O, OH⁻ acts as a nucleophile on the peptide carbon; (B) proton transfer from Asp 125 onto peptide N; (C) peptide bond breakage; (D) final state.

correlation functional (Lee et al. 1988; Becke 1993) and 6-31+G(d,p) basis sets. The parameterization procedure has been described in detail in Trylska et al. (2001).

Computational methodology

Most of the simulation methods have been described in our previous work (Trylska et al. 2002). The AVB code was

incorporated into the Gromos'96 MD package (van Gunsteren et al. 1996; Scott et al. 1999) by Piotr Bała. In the past, the method was applied successfully to phospholipase A₂ (Bała et al. 1998, 2000). Our calculations were based on a crystal structure of the HIV-1 PR:MVT-101 complex refined to a crystallographic R-factor of 15.4% at 2.0 Å resolution (Miller et al. 1997). The structure was reported previously at 2.3 Å resolution (Miller et al. 1989) with a PDB

Figure 4. Proposed direct nucleophilic mechanism not involving the lytic water molecule. In the first stage of this process, following proton transfer between the inner oxygens, Asp 25 is acting as a nucleophile.

Table 1. Valence bond structures used as a basis set in the quantum AVB region for the GA/GB mechanism

Number	Valence bond structure				
S1	$^{25}\text{RCOOH}_a$	$+\text{H}_2\text{O}$		$+^{125}\text{RCOO}^-$	$+\text{RCONHR}$
S2	$^{25}\text{RCOOH}_a$	$+\text{OH}^-$	$+\text{H}_w^+$	$+^{125}\text{RCOO}^-$	$+\text{RCONHR}$
S3	$^{25}\text{RCOOH}_a$	$+\text{OH}^-$		$+^{125}\text{RCOOH}_w$	$+\text{RCONHR}$
S4	$^{25}\text{RCOO}^- + \text{H}_a^+$	$+\text{H}_2\text{O}$		$+^{125}\text{RCOO}^-$	$+\text{RCONHR}$
S5	$^{25}\text{RCOO}^- + \text{H}_a^+$	$+\text{OH}^-$	$+\text{H}_w^+$	$+^{125}\text{RCOO}^-$	$+\text{RCONHR}$
S6	$^{25}\text{RCOO}^- + \text{H}_a^+$	$+\text{OH}^-$		$+^{125}\text{RCOOH}_w$	$+\text{RCONHR}$
S7	$^{25}\text{RCOOH}_a$	$+\text{H}_2\text{O}$		$+^{125}\text{RCOO}^-$	$+\text{RC}^+\text{O}^-\text{NHR}$
S8	$^{25}\text{RCOOH}_a$	$+\text{OH}^-$	$+\text{H}_w^+$	$+^{125}\text{RCOO}^-$	$+\text{RC}^+\text{O}^-\text{NHR}$
S9	$^{25}\text{RCOOH}_a$	$+\text{OH}^-$		$+^{125}\text{RCOOH}_w$	$+\text{RC}^+\text{O}^-\text{NHR}$
S10	$^{25}\text{RCOO}^- + \text{H}_a^+$	$+\text{H}_2\text{O}$		$+^{125}\text{RCOO}^-$	$+\text{RC}^+\text{O}^-\text{NHR}$
S11	$^{25}\text{RCOO}^- + \text{H}_a^+$	$+\text{OH}^-$	$+\text{H}_w^+$	$+^{125}\text{RCOO}^-$	$+\text{RC}^+\text{O}^-\text{NHR}$
S12	$^{25}\text{RCOO}^- + \text{H}_a^+$	$+\text{OH}^-$		$+^{125}\text{RCOOH}_w$	$+\text{RC}^+\text{O}^-\text{NHR}$
S13	$^{25}\text{RCOOH}_a$		$+\text{H}_w^+$	$+^{125}\text{RCOO}^-$	$+\text{RCO}^-\text{OHNNHR}$
S14	$^{25}\text{RCOOH}_a$			$+^{125}\text{RCOOH}_w$	$+\text{RCO}^-\text{OHNNHR}$
S15	$^{25}\text{RCOO}^- + \text{H}_a^+$		$+\text{H}_w^+$	$+^{125}\text{RCOO}^-$	$+\text{RCO}^-\text{OHNNHR}$
S16	$^{25}\text{RCOO}^- + \text{H}_a^+$			$+^{125}\text{RCOOH}_w$	$+\text{RCO}^-\text{OHNNHR}$
S17	$^{25}\text{RCOOH}_w$	$+\text{OH}^-$	$+\text{H}_a^+$	$+^{125}\text{RCOO}^-$	$+\text{RCONHR}$
S18	$^{25}\text{RCOOH}_w$	$+\text{OH}^-$		$+^{125}\text{RCOOH}_a$	$+\text{RCONHR}$
S19	$^{25}\text{RCOO}^-$	$+\text{H}_2\text{O}$		$+^{125}\text{RCOOH}_a$	$+\text{RCONHR}$
S20	$^{25}\text{RCOO}^-$	$+\text{OH}^-$	$+\text{H}_w^+$	$+^{125}\text{RCOOH}_a$	$+\text{RCONHR}$
S21	$^{25}\text{RCOOH}_w$	$+\text{OH}^-$	$+\text{H}_a^+$	$+^{125}\text{RCOO}^-$	$+\text{RC}^+\text{O}^-\text{NHR}$
S22	$^{25}\text{RCOOH}_w$	$+\text{OH}^-$		$+^{125}\text{RCOOH}_a$	$+\text{RC}^+\text{O}^-\text{NHR}$
S23	$^{25}\text{RCOO}^-$	$+\text{H}_2\text{O}$		$+^{125}\text{RCOOH}_a$	$+\text{RC}^+\text{O}^-\text{NHR}$
S24	$^{25}\text{RCOO}^-$	$+\text{OH}^-$	$+\text{H}_w^+$	$+^{125}\text{RCOOH}_a$	$+\text{RC}^+\text{O}^-\text{NHR}$
S25	$^{25}\text{RCOOH}_w$		$+\text{H}_a^+$	$+^{125}\text{RCOO}^-$	$+\text{RCO}^-\text{OHNNHR}$
S26	$^{25}\text{RCOOH}_w$			$+^{125}\text{RCOOH}_a$	$+\text{RCO}^-\text{OHNNHR}$
S27	$^{25}\text{RCOO}^- + \text{H}_a^+$			$+^{125}\text{RCOOH}_a$	$+\text{RCO}^-\text{OHNNHR}$
S28	$^{25}\text{RCOOH}_a$			$+^{125}\text{RCOO}^-$	$+\text{RCO}^-\text{OHNNH}_3^+\text{R}$
S29	$^{25}\text{RCOO}^- + \text{H}_a^+$			$+^{125}\text{RCOO}^-$	$+\text{RCO}^-\text{OHNNH}_3^+\text{R}$
S30	$^{25}\text{RCOOH}_a$			$+^{125}\text{RCOO}^-$	$+\text{RC}^+\text{O}^-\text{OH} + \text{RNH}_2$
S31	$^{25}\text{RCOO}^- + \text{H}_a^+$			$+^{125}\text{RCOO}^-$	$+\text{RC}^+\text{O}^-\text{OH} + \text{RNH}_2$
S32	$^{25}\text{RCOOH}_a$			$+^{125}\text{RCOO}^-$	$+\text{RCOOH} + \text{RNH}_2$
S33	$^{25}\text{RCOO}^- + \text{H}_a^+$			$+^{125}\text{RCOO}^-$	$+\text{RCOOH} + \text{RNH}_2$

R represents carbon atoms bonded to other fragments of the protein or the substrate (for DFT calculations and parameterization $\text{R} = \text{CH}_3$). $\text{RC}^+\text{O}^-\text{NHR}$, $\text{RC}^+\text{O}^-\text{OH}$, and $\text{RC}^+(\text{OH})_2$ are polarized molecules included for additional stages of the reaction (e.g., $\text{RCO}^-\text{OHNNH}_3^+\text{R}$ may break down through the $\text{RC}^+\text{O}^-\text{OH}$ and RNH_2 into RCOOH and RNH_2). C^+ indicates a +1 partial charge on the carbon atom. Superscripts denote the residue number, a and w stand for acid and water, respectively. H_a or H_w are either bound to Asp 25:OD1 or to Asp 125:OD2.

entry code of 4HVP. The MVT-101 inhibitor (N-acetyl-Thr-Ile-NLeu- Ψ [CH₂-NH₂]-NLeu-Gln-Arg-amide) was modified into a substrate with a proper peptide bond between the methionines (N-acetyl-Thr-Ile-Met-Met-Gln-Arg-amide); two α -amino-N-butyric acids were replaced with cysteines, and a water molecule was added (where applicable). The modified parts with the surrounding atom positions were energy minimized. The protein was solvated in an octahedral box of SPC water molecules (Smith and van Gunsteren 1994). The width of the box was set to 7 Å from the border of the solvent-accessible surface. Atomic charges used to describe all the protein and explicit solvent atoms, apart from those assigned to the AVB region, were taken from the Gromos'96 force field library (A-version; van Gunsteren et al. 1996). All of the MD/AVB simulations were performed in the canonical ensemble with the truncated octahedron periodic boundary conditions.

First, a 20-psec simulation of the solvent at 300 K temperature (with the protein atoms fixed) was performed with a time step of 0.25 fsec. Velocities were generated every 5 psec at a temperature of 100 K. The relaxation time of the thermal bath was set to 0.01 psec. The twin-range method was used for the nonbonded interactions (Berendsen 1985). The pair list was updated every 10 steps with a cutoff radius of 12 Å, and the cutoff used in long-range interactions and in the reaction-field calculations was set to 14 Å.

Next, ions were added to yield a neutral system. Ten 0.1-psec runs at 300 K temperature, with a time step of 0.25 fsec and velocities reassigned each time at 100 K were performed with only the solvent and ions free to move. Subsequently, a 20-psec dynamics of the solvent and ions at 300 K with velocities reassigned every 5 psec at temperature 100 K was carried out.

Next, we restrained the positions of solvent and ions with a default force constant of $10.5 \cdot 10^4$ kJ/mole·Å² and performed the thermalization of the solute. To reduce the computational cost, for each mechanism, we selected only those valence bond structures that modeled the first reaction step. The relaxation time of the thermal bath was changed to 0.1 psec, and the cutoff values were increased to 20 Å and 25 Å. A 40-psec run with velocities generated every 5 psec and temperature increased from 0 through 10 K, 50 K, and then every 50 K up to 300 K every 5 psec, was carried out with a time step of 0.5 fsec. The solute atoms were free to move; however, some additional restraints were applied (where applicable) to the active site atoms. The C...OW, Asp 25:OD1...Asp 125:OD1 distances (Fig. 3) were restrained with a one-sided harmonic potential (i.e., the potential was applied only when the actual distance was larger than the equilibrium) with a force constant of 500 kJ/mole·Å² and equilibrium distances 2.8 Å and 3.0 Å, respectively. In the next steps, all of the restraints were gradually decreased by a factor of two and four in two 10-psec runs. The last step included a 100-psec simulation with all solute and solvent atoms free to move. The root mean square deviation (rmsd) of the thermalized structure from the refined crystal structure equaled 1.99, and 2.33 Å for the backbone and the heavy protein atoms, respectively. All further production-phase simulations were performed with a 0.5-fsec time step and no restraints on any atoms. The cutoff values were set to 20 Å and 25 Å or 18 Å and 23 Å, for simulations with <20 and >20 valence bond structures, respectively.

In the activated MD/AVB simulations of the nucleophilic attack, the reaction coordinate (r), following this and other work (Liu et al. 1996; Piana et al. 2002; Trylska et al. 2002), was chosen as the distance between the C and OW nuclei. An external time-dependent harmonic potential, $U = k(r - r_0 + vt)^2/2$, was added to the AVB Hamiltonian, resulting in a force driving the two atoms toward each other. In all of the steered simulations, the stiffness of the harmonic spring k was set to 1000 kJ/mole·Å², the velocity with which the harmonic spring was pulled was 0.03 Å/psec or 0.04 Å/psec, and r_0 was the initial equilibrium distance of C...OW. All activated MD/AVB simulations were performed with the cutoff values of 16 Å and 18 Å. The nuclei were steered toward each other very slowly, and the total energy (AVB plus MM) was well behaved, it did not rise to very high values as might be expected in sampling a very unphysical path. The simulation enabled us to check whether the model accounts reliably for the nucleophilic attack, and in the mean time, the proton could dissociate spontaneously from the water molecule. However, this approach did not allow for the estimations of the energy barrier of the steered process itself. To do so, an umbrella sampling technique (Torrìe and Valleau 1974) should be performed in a series of long simulations to derive proper

statistical data. Such calculations are not yet computationally feasible using our MD/AVB approach.

Results

MD/AVB simulations with Asp 25 protonated at OD1

According to our previous pK_a calculations (Trylska et al. 2002) performed for HIV-1 PR complexed with a model substrate, Asp 125 is deprotonated, whereas Asp 25 is neutral. However, those pK_a calculations did not show which oxygen of Asp 25 is the protonated one, and moreover, they were performed only for one crystal conformation of the protein. In our previous studies, the hydrogen was assigned to the OD2 oxygen (for atom names in this section, see Fig. 3) as was proposed in Hyland et al. (1991a) and Silva et al. (1996). Following the observation that the expected strong hydrogen bond with the peptide carbonyl group did not occur (Fig. 2A,B), and therefore, the proton did not transfer onto the carbonyl oxygen, the hydrogen was initially assigned to the OD1 oxygen. Such assignment is in accord with a model presented in Northrop (2001) and Piana and Carloni (2000).

Preliminary stages of the reaction

First, we performed the MD/AVB simulations with a set of valence bond structures, accounting for the nucleophilic attack and proton transfer from the water molecule onto Asp 125 (Table 1, S1–S16). This limited basis set excluded the products of peptide hydrolysis, and allowed us to investigate the dynamics of preliminary stages of the reaction. To determine which states the system prefers and what hydrogen bonds are formed, a 150-psec simulation was performed with a starting structure taken from the equilibrated ensemble (see Simulation methods). During the simulation, the ionic structures S5 and S11 dominated the electronic state (both with a similar weight of 0.3), and a strong hydrogen bond was observed between the inner oxygens of the aspartyl dyad, as well as between Asp 125:OD2 and the lytic water molecule. These hydrogen-bond distances are reported in Table 2. One may see that the inner HD1 proton stayed at a closer distance to Asp 25:OD1 than to Asp 125:OD1. This is due to the strong interaction between Asp 125:OD2 and H₂O. The plot of rmsd of the active site and all solute atoms from the refined crystal structure in the course of this simulation is shown in Figure 5. The final rmsd of the backbone atoms of the flap region, residues 42–58 and 142–148 from the refined crystal structure, is 1.15 Å.

To check whether the HD1 proton may transfer onto Asp 125 before the nucleophilic attack, we added the corresponding valence structure including ¹²⁵RCOOH_a molecular fragment. The structures that take into account proton transfer from the water molecule onto Asp 25 were also

Table 2. Hydrogen-bond distances in the 150 psec MD/AVB simulation performed with S1–S16 valence-bond structures in tables

Distance [Å]	Minimal	Maximal	Mean $\pm$ stddev
Asp 25:OD1–Asp 125:OD1	2.20	2.98	2.45 $\pm$ 0.08
Asp 25:OD1–Asp 25:HD1	0.92	1.34	1.09 $\pm$ 0.05
Asp 25:HD1–Asp 125:OD1	1.21	2.32	1.51 $\pm$ 0.10
H ₂ O:OW–Asp 125:OD2	2.22	3.04	2.56 $\pm$ 0.11
H ₂ O:HW2–Asp 125:OD2	1.17	2.21	1.57 $\pm$ 0.12
H ₂ O:HW2–H ₂ O:OW	0.90	1.30	1.03 $\pm$ 0.04

For atom names, see Fig. 3.

included. The extended basis set was composed of structures S1–S27 of Table 1. During the 150-psec MD/AVB simulation, we observed that the proton transferred between the inner oxygens (from Asp 25:OD1 onto Asp 125:OD1), whereas the water molecule made a hydrogen bond with Asp 25. After this transfer, the OW $\cdots$ Asp 25:OD2 mean distance dropped to 2.55 ± 0.1 Å. The structures S5, S10, and S11 had their mean c_i^2 values of 0.25 ± 0.02 , 0.15 ± 0.02 , and 0.39 ± 0.04 , respectively. After the proton transfer between the inner oxygens, the contribution of the structure S24 including the $^{125}\text{RCOOH}_a$ molecular fragment grew significantly. The rmsd of the final conformation of the solute backbone from the refined crystal structure was 1.92 Å.

MD/AVB simulations of the nucleophilic attack and formation of the RCO[−]OHNHR intermediate

The available time scale of the MD/AVB dynamics carried out in the previous section was too short for the complete

Figure 5. Plot of the rmsd (in Å) from the refined crystal structure during the 150-psec MD/AVB simulation with structures S1–S16 of Table 1. Filled symbols show the rmsd for the backbone and heavy atoms of the solute, and open symbols for the backbone of the active site region composed of the substrate, H₂O 301, Asp 25, Thr 26, Gly 27, Asp 125, Thr 126, and Gly 127.

reaction process to occur by itself, especially because the nucleophilic attack is usually the slow step of the cleavage process. HIV-1 PR is a much slower enzyme in comparison with other cellular proteases, and we did not expect that the first step would happen on such a short time scale. Therefore, to observe the nucleophilic attack an activated MD/AVB dynamics, simulation was performed (for details see Simulation methods) with the hydrogen initially placed on Asp 25:OD1. One degree of freedom was restrained, that is, the distance between the C and OW atoms, and no constraints were imposed on any other atoms. The initial equilibrium C $\cdots$ OW distance in the imposed harmonic potential was decreased by 0.00002 Å in each 0.5-fsec time step, from the initial value of 3.5 Å (trajectory I) and 3.95 Å (trajectory II) to 1.35 Å. The choice of valence bond structures allowed for the first reaction phase, that is, the dissociation of the water molecule onto Asp 125 and the nucleophilic attack (structures S1–S16 in Table 1). The starting configurations were chosen from one of the previous simulations with the water molecule initially hydrogen bonded to Asp 125.

We observed the HW2 proton transfer onto Asp 125:OD2 around 42nd psec (trajectory I, Fig. 6) and around 52nd psec (trajectory II; data not shown). The main contributions in these simulations came from the structure S11 comprising all ionized and polarized molecular fragments. Before the nucleophilic attack and the proton transfer, structures S5, S2, S4, S8, and S10 (given in the order of decreasing contributions) were significant. The contribution of the structure S8 increased during the simulation and after the formation of the transition state, structures S12, S14, S15, and S16 became dominant. This confirmed that the system underwent changes in the electronic structure, and both the proton transfer and nucleophilic attack occurred.

For both trajectories, a strong hydrogen bond was formed between the Asp 25 and Asp 125 inner oxygens (Table 3). However, the HD1 proton stayed bonded to Asp 25:OD1, even though the set of structures allowed for its dissociation. This was due to the fact that the other aspartate, Asp 125, acted as the proton acceptor from the lytic water molecule. During and after the nucleophilic attack, another short, strong hydrogen bond was formed through the HW2 proton, between the peptide N and Asp 125:OD2, with a distance dropping to 2.4–2.5 Å. This interaction helped deprotonate the water molecule its proton, the peptide bond to change its hybridization, and increased the partial charge on the peptide carbon, facilitating the attack of the hydroxy anion.

After the formation of the RCO[−]OHNHR intermediate fragment, the length of the C–N bond increased by 0.2 Å. The CA2 $\cdots$ CA3 distance, reflecting the change in the peptide carbon hybridization, decreased from the initial average value of 3.8 to 3.2 Å after the nucleophilic attack. We continued both simulations for an additional 100 psec with the restraint on the C $\cdots$ OW distance removed to ensure

Table 3. Hydrogen bond distances between the inner oxygens of the catalytic aspartates in the steered MD/AVB simulations of the nucleophilic attack carried out with S1–S16 valence-bond structures of Table 1

Distance [Å]	Trajectory I			Trajectory II		
	Minimal	Maximal	Mean $\pm$ stddev	Minimal	Maximal	Mean $\pm$ stddev
Asp 25:OD1–Asp 125:OD1	2.27	3.36	2.66 ± 0.15	2.19	3.16	2.54 ± 0.11
Asp 25:OD1–Asp 25:HD1	0.90	1.27	1.02 ± 0.04	0.89	1.25	1.04 ± 0.05
Asp 25:HD1–Asp 125:OD1	1.26	2.71	1.75 ± 0.18	1.23	2.49	1.62 ± 0.14

For atom names, see Fig. 3.

that the system is stable and to double-check that the reaction does not go back.

Further MD/AVB simulations of the intermediate state

To see how the system behaves after the nucleophilic attack, we continued simulating trajectories I and II. On the basis of negligible contributions of structures representing first steps of the reaction, we accounted only for those that involved further stages, S13–S16 and S28–S33 of Table 1. We selected a starting conformation from the last part of the previous trajectory, with the HW2 already bonded to Asp 125:OD2, removed the steering potential, and ran each of the simulations for 100 psec.

Most interestingly, both simulations showed the formation of a single-well hydrogen bond between Asp 125:OD2 and the peptide N (Fig. 3B,C). The proton was positioned midway and shared between the heavy atoms. Also, an asymmetric hydrogen bond was formed between the inner oxygens of the aspartates, with the HD1 proton forming a slightly elongated, but still covalent bond with Asp 25. Table 4 shows important distances between the active site

atoms for one of these simulations. The C–N bond was elongated in the intermediate state and the sp^3 hybridization on carbon was present during both simulations. The structure that was dominant in the simulations was S15 with all ionized or polarized molecular fragments. The rmsd of the final conformation of the protein backbone from the refined crystal structure was 1.80 Å.

MD/AVB simulations of the nucleophilic attack with formation of the $\text{RCO}^-\text{OHNH}_2^+\text{R}$ intermediate

In addition to the previous activated MD/AVB simulation (see above), this one also allowed for the formation of the $\text{RCO}^-\text{OHNH}_2^+\text{R}$ intermediate fragment and its breakage. The basis set was composed of structures S1–S16 and S28–S33 of Table 1. We chose the starting conformation with the $\text{C} \cdots \text{OW}$ distance equal to 3.5 Å, and decreased this distance by 0.00002 Å in each 0.5-fsec time step. Figure 7 presents the contributions of valence bond structures in the course of this simulation. In the first part, the S11 and S5 structures gave the highest contribution. During the dissociation of the lytic water molecule and the nucleophilic attack, the S15 structure started to dominate, but S11 was still significant. When the $\text{RCO}^-\text{OHNH}_2^+\text{R}$ molecular fragment was formed, the HW2 proton transferred onto N and the S28 and S29 structures, containing $\text{RCO}^-\text{OHNH}_2^+\text{R}$ frag-

Figure 6. Proton transfer from the water molecule onto Asp 125:OD2 in the activated MD/AVB simulation of the nucleophilic attack with valence bond structures S1–S16 of Table 1. For atom names, see Figure 3.**Table 4.** Hydrogen bond and other distances in one of the GA/GB MD/AVB simulations of the intermediate state with valence structures S13–S16 and S28–S33 of Table 1

Distance [Å]	Minimal	Maximal	Mean $\pm$ stddev
Asp 25:OD1–Asp 125:OD1	2.26	3.46	2.60 ± 0.11
Asp 25:HD1–Asp 25:OD1	0.91	1.37	1.06 ± 0.04
Asp 25:HD1–Asp 125:OD1	1.27	2.87	1.69 ± 0.14
C–N	1.43	1.81	1.58 ± 0.04
C–OW	1.35	1.65	1.49 ± 0.04
CA2 $\cdots$ CA3	3.00	3.75	3.34 ± 0.10
N–Asp 125:OD2	2.19	2.51	2.35 ± 0.07
N–HW2	1.00	1.45	1.14 ± 0.04
HW2–Asp 125:OD2	1.01	1.46	1.18 ± 0.05

For atom names, see Fig. 3.

ment, started to contribute. The contribution of S30 and S31 in the last part of the simulation showed that this fragment started to break down.

In this simulation, in contrary to our previous work, the hydrogen was positioned on Asp 25:OD1, not on Asp 25:OD2. Such initial configuration made a noticeable difference in this simulation when the system was allowed to select among 22 valence bond structures and had more available degrees of freedom. Due to a strong hydrogen bond formed between the inner oxygens of the aspartates, Asp 125 did not accept the HW2 proton from the lytic water molecule. The situation was more symmetric, and the HW2 atom first formed a hydrogen bond with Asp 125, but later it changed to hydrogen bond to Asp 25. The HW2 distance to both aspartates is illustrated in Figure 8. Such change in hydrogen bonding is also reflected in the contributions of valence bond structures; the ones with the $^{125}\text{RCOOH}_w$ molecular fragment were not present in the later stage of the simulation.

With the HD1 hydrogen positioned on the inner oxygen of Asp 25, the system is more symmetric, and both aspartic acids may help dissociate the water molecule. Which one actually does, most probably depends on the position of H_2O just before the nucleophilic attack. It should be noted, however, that with the current choice of structures, the HD1 hydrogen could be either in the ionic form (H_d^+) or bonded to Asp 25. The formation of the Asp 125:OD1–HD1 covalent bond was not accounted for. However, the HD1 dissociation from Asp 25 was enough to make that aspartate help dissociate H_2O through its OD2. The hydrogen bond distances between the inner oxygens are shown in Table 5.

Figure 7. Contributions (c_i^2) of valence bond structures in the steered MD/AVB dynamics of the nucleophilic attack. Basis set was composed of S1–S16 and S28–S33 structures (Table 1). Legend labels the most significant structures in this simulation. For atom names, see Figure 3.

Figure 8. Distance between the dissociable HW2 hydrogen of the lytic water molecule and OD2 oxygens of catalytic aspartates in the simulation with S1–S16 and S28–S33 valence bond structures as a basis set (Table 1). For atom names, see Figure 3.

We also noticed that during the nucleophilic attack, the Asp 25:OD2 $\cdots$ N distance decreased to 2.6 Å, and HW2 was positioned close to Asp 25:OD2 and N, suggesting that these two atoms together help deprotonate the water molecule. This is a similar case, as observed in our previous simulations with Asp 25 protonated at OD2, in which, preceding the nucleophilic attack, a strong hydrogen bond between Asp 125:OD2 and N was present. The HW2 proton was then also transferred onto Asp 125:OD2, and then immediately onto N.

The change in the system may also be seen by looking at the changes in the length of the peptide bond. Before the nucleophilic attack, its mean length was 1.39 ± 0.03 Å, but during and after it increased to an average of 1.62 ± 0.05 Å. The maximum length of the C–N bond observed in this simulation was 1.82 Å. During the nucleophilic attack, the hybridization of the peptide bond changed, what may be seen in the shortening of the CA2 $\cdots$ CA3 distance. Its initial average value was 3.8 Å; during the formation of the RCO^-OHNHR molecular fragment, it achieved a minimum of 2.9 Å, but later it relaxed to ~ 3.4 Å.

Table 5. Hydrogen bond distances in GA/GB mechanism in one of the MD/AVB simulations of the formation of the RCO^-OHNHR intermediate

Distance [Å]	Minimal	Maximal	Mean $\pm$ stddev
Asp 25:OD1–Asp 125:OD1	2.21	2.96	2.48 ± 0.08
Asp 25:OD1–Asp 25:HD1	0.89	1.75	1.13 ± 0.07
Asp 25:HD1–Asp 125:OD1	1.17	2.14	1.48 ± 0.10

The basis set was formed of structures S1–S16 and S28–S33 of Table 1. For atom names, see Fig. 3.

Peptide bond breakage

We continued the above simulation after the 51st psec without any constraints applied to the C...OW distance. To save the computer time, on the basis of the analysis of the valence bond structure contributions, some structures, that is, those with negligible c_i^2 coefficients in the last part of the simulation, were removed. Therefore, further simulations included only structures S13–S16 and S28–S33 of Table 1. First, we observed a strong hydrogen bond between the inner oxygens, with Asp 25:OD1–HD1 bond elongated to a mean value of 1.08 ± 0.06 Å. Second, we noticed that the C–N bond was extended to an average length of 1.63 ± 0.04 Å. Next, around 90th psec, a complete breakage of the peptide bond occurred and the mean distance between the C and N atoms reached a value of 3.57 ± 0.18 Å.

In this simulation, the peptide N formed a hydrogen bond with the OD2 oxygen of Asp 25 or Asp 125 or, from time to time, to both. This interaction helped the C–N bond to rupture, because the peptide nitrogen atom was held by the outer oxygens, allowing the peptide carbon with its bonded fragments to dissociate. It is worth noting that the breakage of the C–N bond was not observed in those simulations of the GA/GB mechanism, in which the initial hydrogen was placed at Asp 25:OD2 (data not shown). Previously, in two 200-psec simulations of various intermediate states, only an elongation of the C–N bond was observed with a maximum value of 1.8 Å. Also, in all of these earlier simulations, the peptide N formed a strong, short hydrogen bond with Asp 125:OD2, with an average proton distance of 1.14 ± 0.04 Å to N and 1.21 ± 0.05 Å to OD2, but it did not form such a bond with Asp 25.

Proton transfer from water molecule onto Asp 25

Observing that due to the mobility of the HD1 proton, the water molecule may hydrogen bond with either of the aspartates, we performed two more simulations that additionally allowed the HW2 proton to transfer onto Asp 25:OD2. The HD1 was also enabled to form a covalent bond with either Asp 25:OD1 or Asp 125:OD1. The structures forming the basis set for these simulations were S1–S27 of Table 1. We added $^{125}\text{RCOOH}_w$ and $^{125}\text{RCOOH}_a$ fragments, but to save the computer time, we did not account for the breakage of the peptide bond. The starting structures were taken from the equilibrated ensemble. The initial distance of the C...OW atoms was 2.9 Å and 2.8 Å for trajectories I and II, respectively, and it was steered by 0.000155 Å in each activated MD/AVB step. After 50 psec, this constraint was removed.

In both simulations, the inner oxygens formed a strong hydrogen bond and when the C...OW distance reached 2 Å, the HD1 proton transferred from Asp 25:OD1 onto Asp 125:OD1. Simultaneously with the nucleophilic attack, the

HW2 proton was donated to Asp25:OD2. For trajectory I, this proton transfer is presented in Figure 9. Therefore, with the hydrogen positioned at one of the inner oxygens, not only Asp 125, but also Asp 25 may accept the proton from the lytic water molecule.

The valence bond structure that gave the most contribution in both simulations was S11, again the one containing all polarized or ionized molecular fragments. Before the nucleophilic attack and the HW2 proton transfer, its c_{11}^2 grew from 0.3 to 0.8, dropping afterward to a value of 0.5. Other valence structures contributing in the early stages were in decreasing order of significance S5, S10, and S8. Later on, when the transition-state intermediate was formed, S15 gained a contribution of 0.4. Also, structures containing the $^{25}\text{RCOOH}_w$ intermediate fragment became significant.

Direct nucleophilic mechanism

Early stages of the direct nucleophilic mechanism in which Asp 25:OD2 is interacting directly onto peptide carbon were also subject to tests. The x-ray structures cocrystallized with inhibitors that are available in the Protein Data Bank do not contain the lytic water molecule. It may not be captured, possibly due to the presence of various bulky groups in the place of the cleaved peptide bond. However, the hypothesis that the lytic H₂O is not present in the active site when the substrate binds cannot be ruled out. Therefore, we studied how the system behaves in the presence of the substrate alone, as presented in Figure 4.

Primary simulations were performed for two protonated aspartic acids. We kept the proton that hydrogen bonded to Asp 125:OD2 in most of the simulations involving the water molecule. Therefore, one hydrogen was assigned to Asp

Figure 9. Proton transfer from the water molecule onto Asp 25:OD2 in the activated MD/AVB simulation of the nucleophilic attack with S1–S27 valence bond structures as a basis set (Table 1). For atom names, see Figure 3.

125:OD2 and the other to Asp 25:OD1 (Fig. 4). First, we carried out two MD/AVB simulations; one, 100 psec, with a basis set formed by structures S1, S2, S4–S7, S9, S10, S12, S13 (simulation I), and the other (simulation II), 200 psec, with structures S1–S13 (Table 6). In simulation II, Asp 25:HD1 could not only dissociate from Asp 25, but could also transfer onto the inner oxygen of Asp 125.

In simulation I, a short, strong hydrogen bond was observed between the peptide N and Asp 125:OD2 as presented in Figure 10. The HD2 proton transferred back and forth on a picosecond timescale to finally bond to N around the 46th psec, with the N...Asp125:OD2 distance reaching a value of 2.7 Å. We observed that before the HD2 proton finally bonded to N, the C...Asp 25:OD2 average distance was 2.64 ± 0.10 Å with a minimum of 2.35 Å. After the final HD2 transfer, the above mean distance grew by 0.4 Å, but did not exceed 3 Å. These short distances suggest that Asp 25 may alternatively act as a nucleophile, and in the conformation conducive to cleavage, the hydrogen bond between Asp 125:OD2 and N is short and strong.

In simulation II, the HD2 proton of Asp 125 was also very mobile and it transferred back and forth onto the pep-

Figure 10. Fragment of the MD/AVB simulation of the direct nucleophilic mechanism showing HD2 proton transfer from Asp 125:OD2 onto peptide N in the simulation with a basis set of S1–S10 valence bond structures (Table 6). For atom names, see Figure 4.

tide N, implying a strong low-barrier or even single-well hydrogen bond. For the distribution of proton positions, see Figure 11. The average distances of HD2 to Asp 125:OD2 and to N were 1.31 ± 0.21 and 1.41 ± 0.23 , respectively. The mean C...Asp 25:OD2 distance was 2.63 ± 0.09 Å, with the minimal value of 2.36 Å. The inner oxygens formed an asymmetric nearly linear LBHB in both simulations. Appropriate distances are shown in Table 7. We observed only the lengthening of the covalent bond, and the proton did not transfer between the aspartates.

The basis set was further extended to structures S1–S22 and 200-psec simulation was carried out (simulation III).

Table 6. Valence-bond structures used as a basis set in the AVB quantum region for the direct nucleophilic mechanism not involving the water molecule

Number	Valence-bond structure			
S1	$^{25}\text{RCOO}^1\text{H}_a$	$+^{125}\text{RCOO}^2\text{H}_{a'}$	+RCONHR	
S2	$^{25}\text{RCOO}^-$	$+ \text{H}_a^+$	$+^{125}\text{RCOO}^2\text{H}_{a'}$	+RCONHR
S3	$^{25}\text{RCOO}^-$	$+^{125}\text{RCOO}^1\text{H}_a$		$+ \text{H}_{a'}^+$ +RCONHR
S4	$^{25}\text{RCOO}^1\text{H}_a$	$+^{125}\text{RCOO}^-$	$+ \text{H}_{a'}^+$	+RCONHR
S5	$^{25}\text{RCOO}^-$	$+ \text{H}_a^+$	$+^{125}\text{RCOO}^-$	$+ \text{H}_{a'}^+$ +RCONHR
S6	$^{25}\text{RCOO}^1\text{H}_a$	$+^{125}\text{RCOO}^2\text{H}_{a'}$	+RC ⁺ O [−] NHR	
S7	$^{25}\text{RCOO}^-$	$+ \text{H}_a^+$	$+^{125}\text{RCOO}^2\text{H}_{a'}$	+RC ⁺ O [−] NHR
S8	$^{25}\text{RCOO}^-$	$+^{125}\text{RCOO}^1\text{H}_a$	$+ \text{H}_{a'}^+$	+RC ⁺ O [−] NHR
S9	$^{25}\text{RCOO}^1\text{H}_a$	$+^{125}\text{RCOO}^-$	$+ \text{H}_{a'}^+$	+RC ⁺ O [−] NHR
S10	$^{25}\text{RCOO}^-$	$+ \text{H}_a^+$	$+^{125}\text{RCOO}^-$	$+ \text{H}_{a'}^+$ +RC ⁺ O [−] NHR
S11	$^{25}\text{RCOO}^-$	$+^{125}\text{RCOO}^1\text{H}_a$	+RCONH ₂ ⁺ R	
S12	$^{25}\text{RCOO}^1\text{H}_a$	$+^{125}\text{RCOO}^-$	+RCONH ₂ ⁺ R	
S13	$^{25}\text{RCOO}^-$	$+ \text{H}_a^+$	$+^{125}\text{RCOO}^-$	+RCONH ₂ ⁺ R
S14	$^{25}\text{RCOO}^2\text{H}_{a'}$	$+^{125}\text{RCOO}^2\text{H}_{a'}$	+RCONHR	
S15	$^{25}\text{RCOO}^2\text{H}_{a'}$	$+^{125}\text{RCOO}^1\text{H}_a$	+RCONHR	
S16	$^{25}\text{RCOO}^2\text{H}_{a'}$	$+^{125}\text{RCOO}^-$	$+ \text{H}_a^+$	+RCONHR
S17	$^{25}\text{RCOO}^2\text{H}_{a'}$	$+^{125}\text{RCOO}^-$	$+ \text{H}_{a'}^+$	+RCONHR
S18	$^{25}\text{RCOO}^2\text{H}_{a'}$	$+^{125}\text{RCOO}^2\text{H}_{a'}$	+RC ⁺ O [−] NHR	
S19	$^{25}\text{RCOO}^2\text{H}_{a'}$	$+^{125}\text{RCOO}^1\text{H}_a$	+RC ⁺ O [−] NHR	
S20	$^{25}\text{RCOO}^2\text{H}_{a'}$	$+^{125}\text{RCOO}^-$	$+ \text{H}_a^+$	+RC ⁺ O [−] NHR
S21	$^{25}\text{RCOO}^2\text{H}_{a'}$	$+^{125}\text{RCOO}^-$	$+ \text{H}_{a'}^+$	+RC ⁺ O [−] NHR
S22	$^{25}\text{RCOO}^2\text{H}_{a'}$	$+^{125}\text{RCOO}^2\text{H}_{a'}$	+RCONH ₂ ⁺ R	

R = CH₃ for DFT calculations and for fitting of parameters, and in the MD/AVB simulations represents carbon atoms bonded to other fragments of the protein or the substrate. RC⁺O[−]NHR is a polarized molecule included for additional stage of the reaction. Superscripts 25 and 125 denote the residue number, subscripts *a* or *a'* differentiate between the protons. Superscripts 1 and 2 denote proton binding to OD1 and OD2 oxygens, respectively. H_a⁺ was initially placed on Asp 25:OD1 and H_{a'}⁺ on Asp 125:OD2.

Figure 11. Proton position in the single-well hydrogen bond formed between N and Asp 125:OD2 in one of the MD/AVB simulations of direct nucleophilic mechanism with structures S1–S13 (Table 6). For atom names, see Figure 4.

Table 7. Hydrogen bond distances in the two MD/AVB simulations of the direct nucleophilic attack including S1–S10 or S1–S13 valence-bond structures (see Table 6)

Distance [Å]	Simulation I			Simulation II		
	Minimal	Maximal	Mean $\pm$ stddev	Minimal	Maximal	Mean $\pm$ stddev
Asp 25:OD1–Asp 125:OD1	2.25	3.01	2.53 $\pm$ 0.10	2.37	2.99	2.58 $\pm$ 0.06
Asp 25:OD1–Asp 25:HD1	0.89	1.54	1.09 $\pm$ 0.07	0.90	1.24	1.03 $\pm$ 0.04
Asp 25:HD1–Asp 125:OD1	1.17	2.22	1.51 $\pm$ 0.12	1.43	2.11	1.59 $\pm$ 0.06

For atom names, see Fig. 3.

Both protons were allowed to dissociate, and they could transfer parallelly onto the other aspartate. It was also possible for the Asp 125:HD2 to transfer onto the peptide N. This time, we observed that the HD1 proton transferred between the inner oxygens around 108 psec as presented in Figure 12. This was possible because, earlier, the Asp 125:HD2 made a covalent bond with the peptide N forming the RCONH_2^+R molecular fragment (Fig. 13). In the first 100 psec, we observed that the inner proton of Asp 25 tried to move onto Asp 125:OD1, but it could not form a stable bond there, because the other Asp 125 was protonated at OD2. Before the transfer, Asp 125:HD2 formed a strong, single-well hydrogen bond with the peptide N and searched the conformational space close to N, as was visible in the fluctuating structures including the RCONH_2^+R molecular fragment. When the HD2 transferred onto N, Asp 25:HD1 could move onto the inner oxygen of Asp 125.

In the first half of simulation III, the most significant contribution came from the structures S5, S9, and S10. The structures with fluctuating contributions were S12 and S13, and the dominant ones after the proton transfers were S10, S13, S8, and S11. Therefore, the changes in atom positions

were reflected in the changes of the electronic structure. The average distance between the peptide carbon and Asp 25:OD2 was 2.75 ± 0.24 Å with a minimum value reaching as close as 2.31 Å. This number suggests that Asp 25 may act as a nucleophile in the mechanism with direct nucleophilic attack. However, to further confirm this conclusion, additional calculations have to be performed, in which only one aspartic acid is protonated in the active site.

Discussion

Hydrogen-bond network in the active site

In all quantum-classical MD/AVB simulations performed in this study, the formation of short, strong hydrogen bonds was observed in the active site region. One such bond was formed between the peptide nitrogen and the outer oxygen of Asp 125, displaying features of a strong single-well hydrogen bond, and another one between the inner oxygens—a short, but asymmetric hydrogen bond. Whenever the water molecule was included, it also formed a hydrogen bond with oxygens of one of the aspartic acids. The length of that bond was between 2.5 Å and 2.6 Å, but it was asymmetric with the proton bonded covalently to the water oxygen.

In the GA/GB mechanism, during and after the nucleophilic attack, we observed a strong interaction (through the HW2 proton) between N and Asp 125:OD2, displaying features of a single-well hydrogen bond. The same applies to the reaction path without the lytic water molecule, a strong centered hydrogen bond comprising N, HD2, and Asp 125:OD2 was observed. The proton was diffusely distributed, what may be seen in one of the exemplar graphs presented in Figure 11. The character of this graph confirms a LBHB, with a substantial distribution of proton positions approximately midway between the heavy heteroatoms and the proton moving freely between them. One may see an equal mixture of $\text{N}^+-\text{H}\cdots\text{O}^-$, $\text{N}\cdots\text{H}^+\cdots\text{O}^-$ and $\text{N}\cdots\text{H}-\text{O}$ structures. According to our previous DFT calculations (Trylska et al. 2001) performed for a proton transfer between N and Asp 125:OD2 for model molecular fragments of the active region in vacuum, at a distance of 2.4 Å, this hydrogen bond becomes a single-well one. In the pro-

Figure 12. Fragment of the MD/AVB simulation showing proton transfer between the inner oxygens of aspartate side chains with S1–S22 valence bond structures as a basis set and no water molecule (Table 6). For atom names, see Figure 4.

Figure 13. Fragment of the MD/AVB simulation showing proton transfer from Asp 125:OD2 onto the peptide N with S1–S22 valence bond structures as a basis set and no water molecule (Table 6). For atom names, see Figure 4.

tein environment, the N and OD2 atom distances become nearly that short. Our observations are also supported by the fact that LBHB's are observed in the absence of bulk water, and when the inhibitor or substrate binds, the aspartic dyad is inaccessible to solvent. Therefore, the enzyme environment helps to form such strong hydrogen bonds. Similar LBHB's were reported in chymotrypsin, a serine protease, between His 57 nitrogen and Asp 102 oxygen, (Frey et al. 1994; Tobin et al. 1995) and also in other enzymes (Gerlt and Gassman 1993; Cleland and Krevoy 1994; Cleland et al. 1998).

After the formation of the intermediate state in the GA/GB reaction, the peptide N made a hydrogen bond with one or both aspartates. This interaction helped in the breaking of the C–N bond, because the peptide nitrogen served as a holding point and allowed the molecular fragment on the other side of the bond [i.e., RCOOH or $\text{RC}(\text{OH})_2^+$] to dissociate. This fact was observed in one of our simulations in which the breakage of the peptide bond occurred. During that process, the nitrogen made a hydrogen bond with an aspartate oxygen, and after the breakage, this interaction was weakened.

According to our previous pK_a calculations in the enzyme–model substrate and in the enzyme–MVT-101 inhibitor complex, the pK_a s of the two aspartates did not match, and the pK_a of Asp 25 was shifted up (Trylska et al. 1999). Moreover, they showed that the nitrogen of the reduced link $\Psi[\text{CH}_2\text{-NH}_2]$ of the peptide-mimicking inhibitor is fully protonated over the whole pH range in which the enzyme is stable. Such observation also confirms the importance of protonation of the nitrogen in the cleaved peptide bond. During the formation of the transition-state intermediate, the difference in the pK_a values of the N and Asp 125:OD2 heteroatoms has to decrease if they are to form the observed

single-well hydrogen bond. This is in accord with previous findings (e.g., for review, see Cleland et al. 1998) that an enzyme converts a weak hydrogen bond in the initial state into a strong one by eliminating the pK_a mismatch.

The hydrogen bond between the inner oxygens is also a LBHB, but it does not display the characteristics of a single-well one. It is asymmetric, and the proton is either bound to one oxygen or the other. Figure 14 shows the distribution of proton positions before and after the transfer from one exemplar MD/AVB simulation, in which the initial placement of HD1 was on Asp 25:OD1, and the proton transfer was accounted for in the set of valence bond structures. The covalent bond is only stretched, but the proton cannot move freely between oxygens and is not shared between them. It is due to the fact that the Gly 27 and Gly 127 residues are donating their peptide NH groups, which interact with the inner oxygens. Additionally, the proton may choose to hydrogen bond to the outer OD2 oxygen, as happens in our MD/AVB simulation from time to time. In the GA/GB mechanism, one of the aspartates interacts with the water molecule and the peptide nitrogen. Hence, there are many other, also favorable interactions to compete with. Also, in simulations including the lytic water molecule, one of the aspartic acids must help its dissociation; therefore, the HD1 proton bonds to that aspartate, namely its OD1, which is not interacting and making a hydrogen bond with H_2O .

Generally, with the proton positioned at the inner oxygen, the functionality of the aspartic acids displays more symmetric features. MD/AVB simulations show that the hydrogen bonds with Gly 27 and Gly 127 are still formed (at least two at a time), however, the one with the protonated aspartic oxygen is much weaker, with distances on average ~ 3.5 Å. The possibility of the proton transferring between the inner oxygens allows either of the aspartates to be crucial for the

Figure 14. Exemplar graph from one of the MD/AVB simulations showing proton position in the asymmetric hydrogen bond between the inner oxygens before and after its transfer.

first reaction step, which in the case of GA/GB mechanism, is the dissociation of the water molecule.

A water molecule (nr 301) that links the flaps with the substrate through a set of hydrogen bonds and is present in the crystal structures complexed with inhibitors, was included in our MD/AVB simulations. At least two hydrogen bonds at a time are present in every simulation, and the flaps stay closed over the substrate. However, this water molecule and its hydrogen bonds were modeled with the MM force field, so it was not possible to determine the type of hydrogen bond formed in this case. It may happen that the lytic water molecule is not captured between the aspartates during the substrate binding (as observed in the crystal structures), and then the only possible reaction path would be a direct nucleophilic attack. The later stage of this process most probably involves water molecule 301, which helps to reprotonate the aspartic acid for the next reaction round and donate the OH^- anion to the peptide carbon.

Conclusions

With the MD/AVB method, we were able to simulate the formation and breakage of bonds in the reaction process and probe different reaction paths, which is impossible with conventional, classical MD. Additionally, the classical MM force fields do not model single-well or low-barrier hydrogen bonds properly. Our parameterization was performed with a level of quantum theory good enough for describing short, strong, and LBHBs. Their presence in the active site of HIV-1 PR was confirmed by our simulations.

In the GA/GB mechanism, with hydrogen initially positioned at Asp 25:OD1, both aspartate side chains may deprotonate the water molecule. The proton transfer from the lytic water molecule onto one of the aspartates and the nucleophilic attack are concerted processes. Two stages of the catalytic process, the nucleophilic attack and the protonation of the peptide nitrogen, are enough for the formed intermediate state ($\text{RCO}^-\text{OHNH}_2^+\text{R}$) to break down. This is in accord with our previous studies, in which we postulated that Asp 25:OD2 $\cdots$ O hydrogen bond is very weak and seldom present in the trajectory; therefore, the fully protonated intermediate state, $\text{RC}(\text{OH})_2\text{NH}_2^+\text{R}$, as in Figure 2C, did not form. Hence, the proton position at the inner oxygen of Asp 25 is more favorable for the reaction, because a higher number of hydrogens bonds is present in stabilizing the substrate.

In principle, the proton of Asp 25 may be placed initially either on the OD1 or OD2 oxygen atoms, because this aspartate side chain is able to flip around the CB–CG axis and also change to donate its proton to either oxygen of the other Asp 125. However, for the water molecule to be able to come underneath the peptide carbon, again the favorable position of the proton for the reaction would be at Asp 25:OD1. If it is positioned at OD2, the water cannot easily

come in between the aspartate side chains, and we observed fewer conformations that could lead to the nucleophilic attack. Moreover, we observed fewer hydrogen bonds that are formed in the active center in the course of the reaction in comparison to the initial configuration with Asp 25:OD1 protonated. Therefore, the most probable configurations that lead to the nucleophilic attack are with either Asp protonated at OD1.

The inner oxygens still form hydrogen bonds with Gly 27 and Gly 127 of the catalytic triad, but they are much weaker than without the proton between the OD1 atoms. With the hydrogen positioned at Asp 25:OD1, the system is more symmetric, and the water molecule may select which aspartic side chain would take its proton. Both Asp 25 and Asp 125 may become proton acceptors, because the hydrogen positioned between the inner oxygens may transfer on a picosecond time scale. Moreover, it has been shown that the substrate is quite mobile in the active site cleft (Piana et al. 2002), as is also seen in our simulation, and it can adjust its position to be the most favorable one for cleavage.

We observed the formation of a short hydrogen bond between the inner oxygens of the aspartic dyad. However, the hydrogen bond was not a single-well type and asymmetry was observed, that is, the HD1 hydrogen stayed bonded to Asp 25 in most cases, or it transferred and bonded covalently to Asp 125. This short distance and the possibility of proton transfer on a short time scale, indicate that this bond is a LBHB, including covalent-bonding contributions between hydrogens. Our earlier DFT calculations in vacuum showed a lack of barrier for the proton transfer between the oxygens in two RCOO^- molecules at a distance of 2.4 Å, and the barrier was <8 kJ/mole for the 2.6 Å distance.

We also observed that whenever a hydrogen bond was formed between the peptide N and Asp 125:OD2, it displayed properties of a short and single-well hydrogen bond. The strong interaction of the peptide N with Asp 125 helps to change the hybridization of the peptide bond during the nucleophilic attack. It also increases the charge on peptide carbon, facilitating the hydroxy anion attack. This interaction is also important for the cleavage of the intermediate state keeping one side of the peptide bond in place. The presence of a protonated NH_2 group in the intermediate state is crucial for the peptide breakage that occurs for the $\text{RCO}^-\text{OHNH}_2^+\text{R}$.

In the active region, valence bond structures that the system chooses as dominant are always the ones that contain either ionized or polarized molecular fragments that are not taken into account in classical MD. Enzymes, through favorable interactions, are believed to stabilize ionized forms of the substrate or catalytic residues and to allow atoms to approach at a closer distance for the reaction to take place, which is consistent with our observations.

Moreover, our simulations suggest that this reaction is driven by strong interactions between the atoms forming

LBHBs or even single-well ones in the active site. Therefore, we conclude that LBHBs are important interactions for the HIV-1 PR-catalyzed reaction. Our quantum-classical MD/AVB simulations proved their existence during the course of the reaction, and spontaneous proton transfers in a picosecond time scale were observed.

For both mechanisms, there are still some initial protonation states of the aspartates to analyze. In the GA/GB mechanism, it is the one with Asp 125 protonated at OD2. The proton of Asp 125:OD2 could form a stable hydrogen bond with the peptide nitrogen, enabling the substrate to acquire a conformation conducive to cleavage. In the direct nucleophilic attack, tests need to be performed with only one proton positioned in the active site.

Our simulations did not allow for estimation of the energy barriers of the nucleophilic attack and other stages of the reaction, due to relatively short simulations not giving enough statistical data. In the case of the activated molecular dynamics, the statistics is poor, especially at the top of the energy barrier. A full conformational free-energy calculation would involve an umbrella sampling technique, but this is still too computationally demanding for the quantum MD/AVB method. However, a larger number of hydrogen bonds in the active site that appear in the GA/GB mechanism with the proton positioned at Asp 25:OD1, suggests that the enzyme would most probably favor this path. Thus, with the limitations of this method, it is at this stage impossible to discriminate between the GA/GB and direct nucleophilic mechanism on the basis of the energy profile in the protein.

Acknowledgments

We thank Dr. Piotr Bała for implementing the AVB code into the Gromos'96 MD package. This work was supported in part by grants from HHMI, NSF, NIH, NSF Supercomputer Centers, W.M. Keck Foundation, and Accelrys, Inc. J.T. and P.G. were supported by Polish State Committee for Scientific Research with Grant no 115/E-343/S/BST-783 ICM/2002. J.T. was also supported by the NSF Center for Theoretical Biological Physics at UCSD.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

- Bała, P., Grochowski, P., Lesyng, B., and McCammon, J.A. 1998. Quantum dynamics of proton transfer processes in enzymatic reactions. Simulations of phospholipase A₂. *Ber. Bunsenges. Phys. Chem.* **102**: 580–586.
- Bała, P., Grochowski, P., Nowiński, K., Lesyng, B., and McCammon, J.A. 2000. Quantum-dynamical picture of a multi-step enzymatic process. Reaction catalyzed by phospholipase A₂. *Biophys. J.* **79**: 1253–1262.
- Becke, A.D. 1993. Density functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **98**: 5648–5652.
- Berendsen, H.J.C. 1985. Treatment of long-range forces in molecular dynamics. In *Molecular dynamics and protein structure* (ed. J. Hermans), pp. 18–22. Polycrystal Book Service, Western Springs, IL.
- Chatfield, D.C. and Brooks, B.R. 1995. HIV-1 protease cleavage mechanism elucidated with molecular dynamics simulation. *J. Am. Chem. Soc.* **117**: 5561–5572.
- Chatfield, D.C., Eurenus, K.P., and Brooks, B.R. 1998. HIV-1 protease cleavage mechanism: A theoretical investigation based on classical MD simulation and reaction path calculations using a hybrid QM/MM potential. *J. Mol. Struct. (THEOCHEM)* **423**: 79–92.
- Cleland, W.W. and Krevoy, M.M. 1994. Low-barrier hydrogen bonds and enzymatic catalysis. *Science* **264**: 1887–1890.
- Cleland, W.W., Frey, P.A., and Gerlt, J.A. 1998. The low barrier hydrogen bond in enzymatic catalysis. *J. Biol. Chem.* **273**: 25529–25532.
- de Clercq, E. 2002. New anti-HIV agents and targets. *Med. Research. Rev.* **22**: 531–565.
- Frey, P.A., Whitt, S.A., and Tobin, J.B. 1994. A low-barrier hydrogen bond in the catalytic triad of serine proteases. *Science* **264**: 1927–1930.
- Gerlt, J.A. and Gassman, P.G. 1993. Understanding the rates of certain enzyme-catalyzed reactions: Proton abstraction from carbon acids, acyl-transfer reactions, and displacement reactions of phosphodiesteres. *Biochemistry* **32**: 11943–11952.
- Grochowski, P., Lesyng, B., Bała, P.B., and McCammon, J.A. 1996. Density functional based parametrization of a valence bond method and its applications in quantum-classical molecular dynamics simulations of enzymatic reactions. *Int. J. Quant. Chem.* **60**: 1143–1164.
- Guthrie, J.P. 1996. Short strong hydrogen bonds: Can they explain enzymatic catalysis? *Chem. and Biol.* **3**: 163–170.
- Hyland, L.J., Tomaszek Jr., T.A., and Meek, T.D. 1991a. Human immunodeficiency virus-1 protease. 2. Use of pH rate studies and solvent kinetic isotope effects to elucidate details of chemical mechanism. *Biochemistry* **30**: 8454–8463.
- Hyland, L.J., Tomaszek Jr., T.A., Roberts, G.D., Carr, S.A., Magaard, V.W., Bryan, H.L., Fakhoury, S.A., Moore, M.L., Minnich, M.D., Culp, J.S., et al. 1991b. Human immunodeficiency virus-1 protease. 1. Initial velocity studies and kinetic characterization of reaction intermediates by ¹⁸O isotope exchange. *Biochemistry* **30**: 8441–8453.
- Kohl, N.E., Emini, E.A., Schleif, W.A., Davis, L.J., Heimbach, J.C., Dixon, R.A., Scolnick, E.M., and Sigal, I.S. 1988. Active human immunodeficiency virus protease is required for viral infectivity. *Proc. Natl. Acad. Sci.* **85**: 4686–4690.
- Kurup, A., Mekapati, S.B., Garg, R., and Hansch, C. 2003. HIV-1 protease inhibitors: A comparative QSAR analysis. *Curr. Med. Chem.* **10**: 1679–1688.
- Lee, C., Yang, W., and Parr, R.G. 1988. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **37**: 785–789.
- Lee, H., Darden, T.A., and Pedersen, L.G. 1996. An ab initio quantum mechanical model for the catalytic mechanism of HIV-1 protease. *J. Am. Chem. Soc.* **118**: 3946–3950.
- Liu, H., Müller-Plathe, F., and van Gunsteren, W.F. 1996. A combined quantum/classical molecular dynamics study of the catalytic mechanism of HIV protease. *J. Mol. Biol.* **261**: 454–469.
- Miller, M., Schneider, J., Sathyanarayana, B.K., Toth, M.V., Marshall, G.R., Clawson, L., Selk, L., Kent, S.B., and Wlodawer, A. 1989. Structure of complex of synthetic HIV-1 protease with a substrate-based inhibitor at 2.3 Å resolution. *Science* **246**: 1149–1152.
- Miller, M., Geller, M., Swanson, S.M., and Maizel, J. 1997. Analysis of the structure of chemically synthesized HIV-1 protease complexed with a hexapeptide inhibitor. Part I: Crystallographic refinement of 2 Å data. *Proteins: Struct. Funct. Genet.* **27**: 184–194.
- Northrop, D.B. 2001. Follow the protons: A low-barrier hydrogen bond unifies the mechanisms of the aspartic proteases. *Acc. Chem. Res.* **34**: 790–797.
- Park, H., Suh, J., and Lee, S. 2000. Ab initio studies on the catalytic mechanism of aspartic proteinases: Nucleophilic versus general acid/general base mechanism. *J. Am. Chem. Soc.* **122**: 3901–3908.
- Piana, S. and Carloni, P. 2000. Conformational flexibility of the catalytic aspartate dyad in HIV-1 protease: An ab initio study on the free enzyme. *Proteins: Struct. Funct. Genet.* **39**: 26–36.
- Piana, S., Carloni, P., and Parrinello, M. 2002. Role of conformational fluctuations in the enzymatic reaction of HIV-1 protease. *J. Mol. Biol.* **319**: 567–583.
- Rodriguez, E.J., Angeles, T.S., and Meek, T.D. 1993. Use of nitrogen-15 kinetic isotope effects to elucidate details of the chemical mechanism of human immunodeficiency virus 1 protease. *Biochemistry* **32**: 12380–12385.
- Scheiner, S. and Kar, T. 1995. The nonexistence of specially stabilized hydrogen-bonds in enzymes. *J. Am. Chem. Soc.* **117**: 6970–6975.
- Scott, W.R.P., Hünenberger, P.H., Tirion, I.G., Mark, A.E., Billeter, S.R., Fennen, J., Huber, T., Kruger, P., and van Gunsteren, W.F. 1999. The

- GROMOS biomolecular simulation program package. *J. Phys. Chem. A* **103**: 3596–3607.
- Silva, A.M., Cachau, R.E., Sham, H.L., and Erickson, J.W. 1996. Inhibition and catalytic mechanism of HIV-1 aspartic protease. *J. Mol. Biol.* **255**: 321–346.
- Smith, P.E. and van Gunsteren, W.F. 1994. Consistent dielectric properties of the simple point charge and extended simple point charge water models at 277 and 300K. *J. Chem. Phys.* **100**: 3169–3174.
- Tobin, J.B., Whitt, S.A., Cassidy, C.S., and Frey, P.A. 1995. Low-barrier hydrogen bonding in molecular complexes analogous to histidine and aspartate in the catalytic triad of serine proteases. *Biochemistry* **34**: 6919–6924.
- Torrie, G.M. and Valleau, J.P. 1974. Monte Carlo free energy estimates using non-Boltzmann sampling: Application to the sub-critical Lennard-Jones fluid. *Chem. Phys. Lett.* **28**: 578–581.
- Trylska, J., Antosiewicz, J., Geller, M., Hodge, C.N., Klabe, R.M., Head, M.S., and Gilson, M.K. 1999. Thermodynamic linkage between the binding of protons and inhibitors to HIV-1 protease. *Protein Sci.* **8**: 180–195.
- Trylska, J., Grochowski, P., and Geller, M. 2001. Parameterization of the Approximate Valence Bond (AVB) method to describe potential energy surface in the reaction catalyzed by HIV-1 protease. *Int. J. Quant. Chem.* **82**: 86–103.
- Trylska, J., Bała, P., Geller, M., and Grochowski, P. 2002. Molecular dynamics simulations of the first steps of the reaction catalyzed by HIV-1 protease. *Biophys. J.* **83**: 794–807.
- van Gunsteren, W.F., Billeter, S.R., Eising, A.A., Hünenberger, P.H., Kruger, P., Mark, A.E., Scott, W.R.P., and Tironi, I.G. 1996. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*. Hochschulverlag AG an der ETH Zurich and BIOMOS b.v. Zurich, Groningen.
- Warshel, A. 1991. Computer modeling of chemical reactions in enzymes and solutions. John Wiley & Sons, Inc., USA.
- Warshel, A. and Papazyan, A. 1996. Energy considerations show that low-barrier hydrogen bonds do not offer a catalytic advantage over ordinary hydrogen bonds. *Proc. Natl. Acad. Sci.* **93**: 13665–13670.
- Warshel, A. and Weber, R.M. 1980. An empirical valence bond approach for comparing reactions in solutions and in enzymes. *J. Am. Chem. Soc.* **102**: 6218.
- Warshel, A., Papazyan, A., and Kollman, P.A. 1995. On low-barrier hydrogen bonds and enzyme catalysis. *Science* **269**: 102–104.