Published in final edited form as:

J Am Chem Soc. 2016 December 21; 138(50): 16283–16298. doi:10.1021/jacs.6b06856.

Dynamic and Electrostatic Effects on the Reaction Catalyzed by HIV-1 Protease

Agnieszka Krzemińska[1], **Vicent Moliner**[2], and **Katarzyna Świderek**[1,2]

[1]Institute of Applied Radiation Chemistry, Faculty of Chemistry, Lodz University of Technology, Zeromskiego 116, 90-924 Lodz, Poland [2]Departament de Química Física i Analítica, Universitat Jaume I, 12071 Castelló, Spain

Abstract

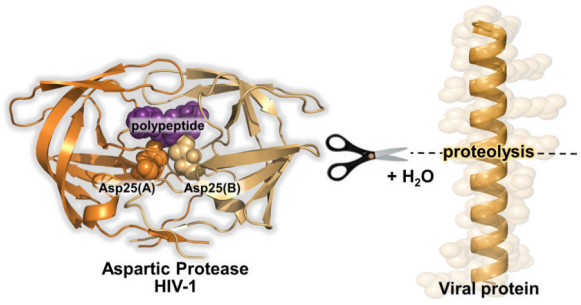
HIV-1 Protease (HIV-1 PR) is one of the three enzymes essential for the replication process of HIV-1 virus, which explains why it has been the main target for design of drugs against acquired immunodeficiency syndrome (AIDS). This work is focused on exploring the proteolysis reaction catalyzed by HIV-1 PR, with special attention to the dynamic and electrostatic effects governing its catalytic power. Free energy surfaces for all possible mechanisms have been computed in terms of potentials of mean force (PMFs) within hybrid QM/MM potentials, with the QM sub-set of atoms described at semiempirical (AM1) and DFT (M06-2X) level. The results suggest that the most favorable reaction mechanism involves formation of a gem-diol intermediate, whose decomposition into the product complex would correspond to the rate-limiting step. The agreement between the activation free energy of this step with experimental data, as well as kinetic isotope effects (KIEs), supports this prediction. The role of the protein dynamic was studied by protein isotope labelling in the framework of the Variational Transition State Theory. The predicted enzyme KIEs, also very close to the values measured experimentally, reveal a measurable but small dynamic effect. Our calculations show how the contribution of dynamic effects to the effective activation free energy appears to be below 1 kcal·mol^{-1}. On the contrary, the electric field created by the protein in the active site of the enzyme emerges as being critical for the electronic reorganization required during the reaction. These electrostatic properties of the active site could be used as a mould for future drug design.

GRAPHICAL ABSTRACT

SUPPORTING INFORMATION

The Supporting Information is available free of charge on the ACS Publications website at DOI: ***

Empirical estimation of pKa values calculated with PropKa; averaged values of KIE, BIE and EIEs obtained at AM1/MM level; Cartesian coordinates of QM subset of atoms for all TS structures optimized at M06-2X/MM level; AM1/MM PMFs obtained for all explored chemical steps; key interatomic distances and atomic charges of stationary point structures located and refined at M06-2X/MM level of theory; electric field created by the protein and water molecules in the center of the active site and in the C and N atoms of the scissile peptide bond; normalized autocorrelation function for light and heavy enzyme for two localized transition states TS3 and TS4 obtained at AM1/MM level; PMFs obtained in the vicinity of the transition state for two localized transition states TS3 and TS4 obtained at AM1/MM level.



Keywords

HIV-1 protease; proteolysis; peptide bond; KIEs; EIE; enzyme KIE; QM/MM; dynamic effects; heavy enzymes

Introduction

The role of HIV-1 PR in the viral life cycle of HIV-1 is to recognize the asymmetric shape of the peptide structure, rather than a particular amino acid sequence,[1] and later to catalyze the cleavage reaction leading to the hydrolysis of the peptide bond of the substrate.[2] HIV-1 PR is an enzyme of the aspartic proteases family which is characterized by the presence of two eponymous aspartyl residues in the active site located at the bottom of a substrate-binding cleft.[3, 4, 5, 6, 7, 8, 9] The structure of the HIV-1 PR, together with a detail of the active site with the substrate and the mentioned water molecule is presented on Figure 1. Kent and co-workers studied the protein conformational dynamics of HIV-1 PR by preparing series of analogues in which the flexibility of the 37–61 loop was systematically varied.[10] The characterization of these variants by means of different experimental techniques and molecular dynamics (MD) simulations suggested a more mobile L-domain, containing the protonated general acid Asp25(A), and a less mobile D-domain, containing the general base Asp25(B). In any case, the catalysis would not be rate-limited by opening or closing of these loops.[10] Other studies were also addressing the relevance of the HIV-1 PR large-scale protein motions involving the flaps and the cantilever,[11, 12, 13] and its conformational diversity.[14]

Despite that proteases are known since 1825,[15] as mentioned by van der Kamp and Mulholland, the amount of theoretical studies of the proteolysis catalyzed by HIV-1 PR are still incomplete to provide final conclusions about its mechanism and the origin of catalysis.[16] The relatively simple structure of the active site is nevertheless the origin of some problems, which are still the source of many scientific debates. Generally, experimental[17, 18, 19, 20] and theoretical[12, 13, 21, 22, 23, 24] studies have provided evidence that peptide bond breaking is the result of a mechanism where a water molecule is activated by an aspartate residue and it then attacks as a nucleophile on a carbonyl carbon of the substrate peptide chain. Nevertheless, in the past 30 years different mechanisms for the reaction catalyzed by aspartic proteases have been suggested. The first one, proposed by Davies and co-workers[25] based on a crystal structure of HIV-1 PR from *Rhizopy schinensis*, was by long-time the most accepted mechanism. It assumes that the water molecule is held and

activated by the two aspartate residues belonging to different monomers, Asp25(A) and Asp25(B), and it would act as a nucleophile by attacking a carbonyl carbon of the scissile peptide bond. This attack generates a tetrahedral covalent and metastable oxyanion (OA) intermediate as shown on mechanism [B] in Scheme 1. The presence of this OA intermediate was supported by theoretical studies using the approximate valence bond (AVB) method,[23, 26] suggesting that the hydroxyl anion intermediate is stable in the active site of HIV-1 PR. Nevertheless, it must be taken into account that the calculations were in fact calibrated assuming the existence of the OA intermediate.[27, 28]

In the early 90s Jaskólski et al. described a new possible mechanism, in which the reaction of proteolysis would be a one-step process,[29] initiated at Michaelis complex (MC) with protonated Asp25 residue from chain B, as presented in mechanism [A] in Scheme 1. In this chemical step the nucleophilic attack of the water molecule to the carbonyl carbon occurs simultaneously with the proton transfer to the nitrogen and cleavage of the peptide bond. However, early *ab initio* calculations based on small models did not confirm the synchronicity of both attacks but suggested that the reaction would take place in a step-wise manner, through an intermediate state.[30]

A step-wise mechanism, through a stable tetrahedral intermediate but with a gem-diol character (GD), was supported by hybrid Car-Parrinello/Classical MD simulations performed by Rothlisberger and co-workers,[13] and by experimental evidences obtained by Kent and co-workers.[10] This mechanism, presented as mechanism [C] in Scheme 1, is also in agreement with crystallographic studies of Weber and co-workers[31] who reported two high-resolution crystal structures of wild-type and the multi-drug-resistant variant with the I54V mutation in complex with a peptide in a gem-diol tetrahedral intermediate conformation. Hosur and co-workers[32] confirmed this observations using X-Ray snapshots technique that allowed following the reaction pathway. More recently, Schramm and co-workers proposed a model of transition state (TS) structure related with the decomposition of the GD intermediate, consistent with primary and secondary KIEs for native and I84V multidrug-resistant HIV-1 PR enzyme.[33] The conversion of the GD intermediate to the product complex should be rate limiting according to other kinetic studies.[17, 34, 35, 36] ^{15}N KIEs were previously used by Meek and co-workers to elucidate the reaction mechanism of HIV-1 PR arriving to the conclusion that the protonation of the nitrogen atom of the amide leaving group should contribute to the rate limiting step.[18]

Part of the uncertainty of the mechanism of HIV-1 PR is linked to the debate of the protonation state of Asp25(A) and Asp25(B). In fact, all four possibilities have been taken into consideration in the past. Meek and co-workers[17] demonstrated that pK_a values are different for Asp25(A) and Asp25(B), and equal to 3.1 and 5.1, respectively, concluding that only one of them should be protonated. On the other hand, Kent and co-workers[37] claimed that both residues are negatively charged at pH = 6 in the absence of an inhibitor, while Parrinello and co-workers,[38] based on computational results, suggested that both residues should be protonated. Nevertheless, from the first experimental evidence of Torchia and co-workers,[39] it is now well accepted, in agreement with a broad range of aspartic proteases, that the aspartic groups are in opposite states of protonation, such that the catalytically competent form of these enzymes is a “mono-protonated” one. This conclusion is based on

three-dimensional structural data of these enzymes, affinity labeling studies, and the observation of “bell-shaped” profiles of log V/K vs pH or log V vs. pH for oligopeptide substrates of porcine pepsin, [40, 41, 42, 43] penicillopepsin, [44] and human renin. [45] Moreover, this statement was recently confirmed by Kuroki and co-workers,[46] who observed a proton on only one aspartate residue using neutron diffraction technique. The resulting structural data showed that the catalytic residue Asp25(A) is protonated and that Asp25(B) is deprotonated.

The knowledge of the molecular mechanism can be crucial for the design of new HIV-1 PR inhibitors. Nevertheless, the definition of an enzymatic reaction mechanism implies not only a static but a dynamic description of the system in the different states involved along the reaction path. In this regard, in the last years experimental and theoretical studies have been focused in elucidating whether fast vibrational protein motions have a role in the barrier crossing of the chemical step of enzyme catalyzed reactions.[47] Fast dynamic motions of enzymes are recently a topic of great interest in the scientific community because of their possible role in driving the conversion of substrates into products. The use of enzymes with the normal isotopic distribution of their atoms modified by their corresponding heavier isotopes, the so-called “heavy enzymes”, has proved to be a useful tool to explore the role of these motions.[48] Changes in atomic masses can alter the vibrational frequencies while the evolution of the electronic properties of the system described by the Potential Energy Surface (PES) remains unaltered according to the Born-Oppenheimer approximation. These isotopic substitutions can have an effect in the vibrational contribution to the activation free energy but also in the tunneling probability when the chemical reaction involves the transfer of a light particle. Enzyme kinetic isotope effects (enzyme KIEs) can be measured and computed. The generally observed diminution in the rate constant of the chemical step in the ‘heavy enzyme’ relative to the ‘light enzyme’ (the isotopically unlabeled counterpart) reflects protein motions coupled to the reaction coordinate.

The pioneering studies using a mass-modified enzyme, which were based on just replacing the non-exchangeable hydrogen atoms by deuterium, showed a decrease in v_{max} by a factor equal to 1.8 with respect to the ordinary version of the enzyme.[49, 50] This difference was attributed to the diminution of the vibrational energy available in the active site of the deuterated enzyme. More recently, Schramm and co-workers performed kinetic studies on the human purine nucleoside phosphorylase[51] (PNP) and on the HIV-1 PR.[52] In these studies, the normal isotopic distribution of the protein atoms was modified replacing all carbon, nitrogen and non-exchangeable hydrogen atoms by ^{13}C, ^{15}N and ^{2}H, respectively. The identical circular dichroism spectra of heavy and light enzymes confirmed the structural integrity of the protein. Then, the fact that the “heavy” enzymes showed a diminution in the rate constant of the chemical step relative to the wild or ‘light enzyme” was interpreted as a dynamical link between femtosecond protein motions and events in the reaction coordinate in the barrier crossing.[51] However, the hypothesis of dynamics as a key driving force in enzyme catalysis has been questioned, and argued that the ability of an enzyme to sample an ideal electrostatic configuration is the source of the catalytic efficiency.[53, 54, 55]

As commented above, HIV-1 PR was the second enzymatic system, just after PNP,[51] used by Schramm and co-workers to generate a “heavy” enzyme in order to confirm their hypothesis

that enzyme motions on the femtosecond time scale play a role in transition state formation. Different than in case of PNP, analysis of catalytic process using aminobenzoyl-Thr-Ile-Nle*pNO2-Phe-Gln-Arg-NH2 peptide as a substrate showed that the steady-state rate constant k_{cat} is dominated by the final chemical step. The obtained experimental values of k_{cat} are equal to 3.28 ± 0.08 and 2.75 ± 0.08 s^{-1} for light and heavy enzyme, respectively, resulting in normal isotope effect of 1.19 ± 0.05. The observed reduction in k_{cat} of the reaction catalyzed by the heavy enzyme was interpreted by Schramm and co-workers as a proof of their hypothesis that reducing the frequency of the bond vibrations by isotopic substitution would lead to decreasing probability of crossing the chemical barrier.

The main goal of the present work is to explore the proposed mechanisms of proteolysis catalyzed by HIV-1 PR by means of MD simulations based on hybrid quantum mechanics/molecular mechanics (QM/MM) potentials. Once the rate limiting step of the most favorable mechanism was determined, primary, secondary and enzyme KIEs will be computed with a flexible model of the homodimer of HIV-1 PR to get information directly comparable with the experimental data. This treatment will allow the confirmation of the proposed molecular mechanism and identifying dynamic effects occurring in this system. Finally, the electrostatic effects of the protein will be analyzed in depth to complement the full picture of the HIV-1 PR catalyzed reaction.

Computational methods

Setup of the system

The system of HIV-1 PR was prepared based on a crystal structure of wild type in complex with 4-[2-(2-acetylamino-3-naphtalen-1-ylpropionylamino)-4-methylpentanoylamino]-3hydroxy-6-methyl-heptanoicacid[1-(1-carbanoyl-2-naphthalen-1-yl-ethylcarbamoyl)-propyl]-amide (LP-130) inhibitor with PDB ID 1ODY at 2.00 Å resolution.[56] The inhibitor was replaced by a peptide that consists of 6 alanine residues. The position of the peptide in the active site was based on the arrangement with LP-130, and was build using Discovery Studio program version 3.5.[57] Using SCit tool developed by Tufféry and co-workers,[58] the backbone of Met-46 from chain B has been revised automatically based on Hidden Markov model derived structural alphabet.[59]

Since the standard p*K*a values of ionizable groups can be shifted by local protein environments,[60] an accurate assignment of the protonation states of all these residues at pH = 7 was carried out. Recalculation of the p*K*a values of the titratable amino acids has been done using the empirical PROPKA 2.0 program of Jensen et al.[61] According to the results, most of the residues were found at their standard protonation state, except for the Asp25 from chain A whose pKa value indicates to be protonated. This is an interesting preliminary result since the pKa values of Asp25A and Asp25B should in principle be equivalent if both residues were surrounded by equal local environments. Keeping in mind that the pKa values have been estimated in the presence of the peptide that consists of 6 equal alanine residues, this result could suggest a non-symmetric environment in the active site of the protein. Nevertheless, after a deep insight into the geometry of the X-ray structure, no significant differences were detected in the local protein environments of both Asp. Instead, it appears that the carbonyl group of the peptide, closer to one aspartate than to the other, makes the

PROPKA calculations render different values of pKa. The prediction of a monoprotonated state in the catalytic aspartates has been already proposed by PROPKA calculations of Jensen and co-workers for HIV-1 PR with different inhibitors,[62] and by pH replica-exchange molecular dynamics of the apo and bound forms of HIV-1 PR with different protease inhibitors carried out by Roitberg and co-workers.[63] The obtained pKa values of all the pKa titratable residues can be found in the Supporting Information. Additionally, His-69 from chain A and B were protonated in δ-position. Subsequently, in order to neutralize the system a total of 3 counter ions (Cl^-) were placed into optimal electrostatic positions around the protein. Afterwards, series of optimization algorithms (steepest descent, conjugated gradient and L-BFGS-B)[64] were applied. To avoid a denaturation of the protein structure, all the heavy atoms of the protein and the inhibitor were restrained by means of a Cartesian harmonic umbrella with a force constant of 1000 $kJ \cdot mol^{-1} \cdot Å^{-2}$.

Then the protein was placed in a box of pre-equilibrated water molecules (100 × 80 × 80 $Å^3$), using the principal axis of the protein-inhibitor complex as the geometrical center. Any water with an oxygen atom lying in a radius of 2.8 Å from a heavy atom of the protein was deleted. The geometries of the remaining water molecules were then optimized. Later, an initial 500 ps classical MD simulation (at temperature 300 K), carried out to relax the system without significant changes in its geometry, was performed with the full model of protein with the short peptide and water molecules using the AMBER[65] and TIP3P[66] force fields for the protein and water molecules, respectively, as implemented in the fDYNAMO library.[67, 68, 69] The same force fields will be used to describe the MM region in the following hybrid QM/MM MD simulations. Due to the large amount of degrees of freedom, any water molecule 20 Å apart from any of the atoms of Asp-25 from chain A and chain B were kept frozen in the remaining calculations. Cutoffs for the non-bonding interactions were applied using a force switching scheme, within a range radius from 14.5 to 16 Å.

Afterwards, the system was optimized using hybrid QM/MM potentials, where a small part of the system consists of both Asp25 from chain A and B, one water molecule and two aminoacid residues of the peptide, Ala3 and Ala4 (see Figure 1), was described by quantum mechanics using the AM1[70] semiempirical Hamiltonian and by the M06-2X hybrid functional developed by Truhlar's and co-workers[71, 72] with the standard 6–31+G(d,p) basis set. To saturate the valence of the QM/MM frontier atoms, link atoms were placed in the peptide and in the two aspartate residues (see Figure 1). The rest of the protein and water molecules are described using the mentioned force fields. Stationary points were localized using the Berny algorithm at AM1/MM and M06-2X/MM level.

Free energy calculations

In order to get the free energy landscape of the full catalyzed reaction, PMFs have been traced[73, 74, 75] along selected coordinates using the weighted histogram analysis method (WHAM) combined with the umbrella sampling approach[75, 76] as implemented in fDYNAMO. The procedure for the PMF calculation is straightforward and requires series of molecular dynamics simulations in which the distinguished reaction coordinate variable, ξ, is constrained around particular values,[75] while the remaining degrees of freedom (including those of the water solvent molecules and protein environment) are conveniently sampled.

The distinguished reaction coordinate, that correspond to a reduced number of internal degrees of freedom considered as the most relevant ones in each particular chemical step, is obviously not the real reaction coordinate that involves many other degrees of freedom not only of the chemical system but from the environment. Nevertheless, our experience dictates that a good description of the key states and the differences in energy between them, can be deduced from the PMFs if a proper combination of internal coordinates are selected. The values of the variables sampled during the simulations are then pieced together to construct a distribution function from which the PMF is obtained as a function of the distinguished reaction coordinate (W(ξ)). The PMF is related to the normalized probability of finding the system at a particular value of the chosen coordinate by eq. 1:

$$W(\xi)=C - kT\ln\int\rho(r^N)\delta(\xi(r^N) - \xi)\mathrm{d}r^{N-1} \quad (1)$$

Then, the activation free energy of a chemical step can be expressed as:[77]

$$\Delta G^{\ddagger}(\xi)= \left\lfloor W(\xi^{\ddagger})+G_{\xi}(\xi^{R})\right\rfloor \quad (2)$$

where the superscripts indicate the value of the reaction coordinate at the reactants (R), and at the TS (‡), and $G_{\xi}(\xi^{R})$ is the free energy associated with setting the reaction coordinate to a specific value at the reactant state. Normally this last term makes a small contribution and the activation free energy is directly estimated from the PMF change between the maximum of the profile and the reactant's minimum:

$$\Delta G^{\ddagger}(\xi) \approx W(\xi^{\ddagger}) - W(\xi^{R})=\Delta W^{\ddagger}(\xi) \quad (3)$$

In the reaction studied in the present paper, some of the steps require the use of more than a single internal coordinate as the reaction coordinate (or the antisymmetric combination of two interatomic distances). In such cases, two dimensional PMFs (2D–PMF) are computed using two coordinates, ξ_1 and ξ_2, and then eq. 1 is transformed into eq. 4:

$$W(\xi_1,\xi_2)=C' - kT\ln\int\rho(r^N)\delta(\xi_1(r^N) - \xi_1)\delta(\xi_2(r^N) - \xi_2)\mathrm{d}r^{N-2} \quad (4)$$

To estimate the activation free energy from eq. 4, we recovered one-dimensional PMF changes tracing a maximum probability reaction path on the 2D–PMF surface and integrating over the perpendicular coordinate.

Spline corrections

The generation of PMFs requires a large number of structures that are generated from QM/MM MD calculations. Inevitably, we are restricted to the use of a semiempirical Hamiltonian, AM1 in this work. Then, in order to improve the quality of every single free energy surface, based on the work of Truhlar and co-workers[78, 79, 80] a spline under

tension[81, 82] is used to interpolate this correction term at any value of the reaction coordinate ξ1 (and ξ2 in the case of two dimensional PMFs; 2D PMFs) selected to generate the free energy surfaces.[83, 84, 85] In this way we obtain a continuous function in a new energy function to obtain corrected PMFs:

$$E = E_{\mathrm{LL/MM}} + S\left[\Delta E_{\mathrm{LL}}^{\mathrm{HL}}(\xi)\right] \quad (5)$$

where S refers to the spline function, and its argument $\Delta E_{\mathrm{LL}}^{\mathrm{HL}}(\xi)$ is a correction term evaluated from the single-point energy difference between a high-level (HL) and a low-level (LL) calculation of the QM subsystem. Herein the hybrid M06-2X functional with the 6–31+G(d,p) basis set, as suggested by Truhlar and co-workers,[86, 87] has been employed as the HL method. As our simulations will demonstrate, the energy difference between the LL/MM and HL/MM surfaces can be significant in some cases, thus justifying the corrections. Moreover, the fact that all optimized TS structures at M06-2X/MM were located on the quadratic region of the spline-corrected surface can be considered as an additional proof of the reliability of the selected technique.

Kinetic and Equilibrium Isotope Effects

Quasiclassical KIEs and equilibrium isotope effects (EIEs) have been computed for isotopic substitutions of key atoms from eleven couples of stationary structures, comparing rate limiting TSs and reactant complex (to compute KIEs) or by comparing two minimum energy structures (EIE), at M06-2X/MM level of theory. A special case of EIEs is the binding isotope effects (BIE) that refer to the binding step of the peptide from aqueous solution to the protein active site. Then, the ratio between the rate constants, or equilibrium constants, corresponding to the light atom “L” and the heavier isotope “H” can be computed using the Transition State Theory (TST), as:

$$\mathrm{IE} = \frac{\left(\frac{Q_b}{Q_a}\right)_L}{\left(\frac{Q_b}{Q_a}\right)_H} e^{-1/\mathrm{RT}(\Delta \mathrm{ZPE}_L - \Delta \mathrm{ZPE}_H)} \quad (6)$$

In eq. 6, the subscripts *H* and *L* refers to heavy and light isotopologs, respectively, the total partition function, *Q*, was computed as the product of the translational, rotational, and vibrational partition functions for the isotopologs in the two stationary structures under comparison, *a* and *b*. Thus, for KIE calculations *a* refers to MC (or INT) and *b* to TS, for EIE calculation *a* refers to MC and *b* to INT, and for BIE calculation *a* refers to the peptide in water and *b* to the peptide in the MC. Δ*ZPE* refers to the difference in the zero point energies between *a* and *b*. The subset of atoms used to define the Hessian for these IE calculations were those of the QM region, consistent with the “cut-off rule” and the local nature of isotope effects.[88] The Born–Oppenheimer, rigid-rotor, and harmonic oscillator approximations were considered to independently compute the different contributions. Keeping in mind that because both involved states, reactants and TS, are in a condensed media (the active site of a protein), contribution of translation and rotation to KIEs are

negligible. Nevertheless, the full 3N × 3N Hessians have been subjected to a projection procedure to eliminate translational and rotational components, which give rise to small nonzero frequencies, as previously described.[88] Thus, it has been assumed that the 3N – 6 vibrational degrees of freedom are separable from the 6 translational and rotational degrees of freedom of the substrate.

Enzyme Kinetic Isotope Effect (enzyme KIE)

In order to get an insight on non-equilibrium effects due to protein motions, the entire enzyme has been isotopically substituted by replacing carbon, nitrogen, and nonexchangeable hydrogen atoms by their corresponding heavier isotopologs ^{13}C, ^{15}N, and ^{2}H. The resulting "heavy" enzyme can then be compared with its natural, lighter, counterpart by computing the enzyme KIEs as the ratio between rate constant of the reaction catalyzed by light and heavy variant of the enzyme.[48] The rate constants of 'light' and 'heavy' PR-HIV1 were evaluated under the framework of Ensemble Averaged Variational Transition State Theory (EA-VTST), which was corrected for tunneling contributions and dynamic effects:[89, 90, 91]

$$k(T)=\Gamma(T,\xi)\frac{k_B T}{h}e^{-\left(\frac{\Delta G_{\mathrm{act}}^{\mathrm{QC}}(T,\xi)}{\mathrm{RT}}\right)}=\frac{k_B T}{h}e^{-\left(\frac{\Delta G_{\mathrm{eff}}(T,\xi)}{\mathrm{RT}}\right)} \quad (7)$$

where R is the ideal gas constant, T is the temperature, kB is the Boltzmann constant, h is Planck's constant, and ΔG_{eff} is the effective activation free energy, which includes all the contributions to the rate constant, and can be readily compared to the value derived from the experimental rate constant. $\Gamma(T, \xi)$ is the temperature-dependent transmission coefficient that contains dynamic and tunneling corrections to the classical rate constant:

$$\Gamma(T,\xi)=\gamma(T,\xi)\kappa(T) \quad (8)$$

where $\gamma(T, \xi)$ is the recrossing transmission coefficient that corrects the rate constant for the trajectories that recross the dividing surface from the product valley back to the reactant valley, and $\kappa(T)$ is the tunneling coefficient that accounts for reactive trajectories that do not reach the classical threshold energy. $\Delta G_{\mathrm{act}}^{\mathrm{QC}}$ is the quasiclassical activation free energy calculated along the reaction coordinate ξ:[92, 93]

$$\Delta G_{\mathrm{act}}^{\mathrm{QC}}(T,\xi)=\Delta G_{\mathrm{act}}^{\mathrm{CM}}(T,\xi)+\Delta G_{\mathrm{vib}}^{\mathrm{QM}}(T) \quad (9)$$

where $\Delta G_{\mathrm{act}}^{\mathrm{CM}}(T,\xi)$ is the activation free energy obtained from the classical PMF along the selected reaction coordinate, and $\Delta G_{\mathrm{vib}}^{\mathrm{QC}}(T)$ is the correction term due to the quantized nature of molecular vibrations (mainly zero-point energies).[94, 95, 96, 97]

As observed in previous computational studies, the only significant difference found when computing the ratio between the rate constant of the 'light' and 'heavy' enzymes is in their

recrossing coefficients.[98, 99, 100, 101, 102] Thus, while the mass modification can affect the zero point energies of the TSs and reactants to a different extent, and a slight change can appear in the corresponding $\Delta G_{\mathrm{vib}}^{\mathrm{QC}}(T)$ values if including residues around the substrate in the calculation of the Hessian, the observed differences are nevertheless too small to account for the experimental enzyme KIEs.[101] The tunneling coefficients of the 'light' and 'heavy' enzymes have been also demonstrated to be statistically identical and previous studies have also indicated that protein isotope labelling has a negligible effect on the potentials of the enzyme.[102] Consequently, the force fields of 'light' and 'heavy' HIV-1 PR and their classical activation free energy barriers can be considered to be identical. Therefore, the enzyme KIE can be approximated as the ratio of the recrossing transmission coefficients for the light and heavy enzyme:

$$\text{enzyme KIE} = \frac{k^{\mathrm{LE}}}{k^{\mathrm{HE}}} \approx \frac{\gamma^{\mathrm{LE}}}{\gamma^{\mathrm{HE}}} \quad (10)$$

Any non-equilibrium influence of protein dynamics (revealed as a variation of the rate constant due to the vibrational shift of protein motions caused by mass substitution) should be captured in this coefficient. In the present work, this recrossing transmission coefficient $\gamma(T, \xi)$ has been computed by means of the Grote-Hynes theory, as the ratio between the reactive frequency (ω_r) and the equilibrium frequency (ω_{eq}), the frequency obtained under the assumption of equilibrium between the reaction coordinate and the remaining degrees of freedom of the system:[103]

$$\gamma^{\mathrm{GH}} = \frac{\omega_r}{\omega_{\mathrm{eq}}} \quad (11)$$

The equilibrium frequency is obtained fitting the PMF previously performed to a parabolic function:

$$\Delta\mathrm{PMF} = -\frac{1}{2} k_{\mathrm{eq}} (\mathrm{RC} - \mathrm{RC}^{\mathrm{TS}})^2 \quad (12)$$

where ΔPMF is the potential of mean force difference with respect to the maximum in the profile, k_{eq} is the equilibrium force constant and RC^{TS} is the reaction coordinate of the maximum of the profile. So, the equilibrium frequency is

$$\omega_{\mathrm{eq}} = \frac{1}{2\pi c} \sqrt{\frac{k_{\mathrm{eq}}}{\mu_{\mathrm{RC}}}} \quad (13)$$

where μ_{RC} is the reaction coordinate reduced mass and c is the speed of light.

The reactive frequency is obtained by applying the Grote-Hynes equation. Thus, once the equilibrium frequency is known, the reactive frequency can be easily obtained from the following relationship:[103, 104]

$$\omega_r^2 - \omega_{\mathrm{eq}}^2 + \omega_r \int_0^\infty \xi_{\mathrm{TS}}(t) e^{-\omega_r t} \mathrm{d}t = 0 \quad (14)$$

The friction kernel (ξ^{TS} (t)) is obtained at the transition state (TS(t)) to determine the forces exerted during the passage over the top of the barrier. All force contributions, due to the reacting system and to the environment, are considered in the calculations. In this procedure, it is assumed that recrossings take place in the proximity of this dynamic bottleneck.[105] The friction kernel (ξ^{TS} (t)) gives the fluctuating forces acting on the reaction coordinate, then providing an efficient way to quantify the coupling of the rest of the degrees of freedom of the system with the selected reaction coordinate:[104]

$$\xi(t) = \frac{\langle F_{\mathrm{RC}}(0) F_{\mathrm{RC}}(t) \rangle}{\mu_{\mathrm{RC}} k_B T} \quad (15)$$

where $F_{RC}(t)$ is the force on the reaction coordinate, μ_{RC} is the associated reduced mass, k_B is the Boltzmann constant, T is the temperature, and $\xi^{TS}(t)$ quantifies the coupling of the reaction coordinate with the rest of the degrees of freedom of the system.

RESULTS AND DISCUSSION

Protonation state of Asp25(A) and Asp25(B)

The protonation state of Asp25(A) and Asp25(B), which has been a question of debate as mentioned in the Introduction, is crucial for the theoretical modeling of HIV-1 PR. Considering the possible alternatives, and the implications on the mechanism, we have first decided to calculate the pKa of the key titratable residues of the active site, and the relative free energies between the two possible “mono-protonation” states, the protonated Asp25(A) or Asp25(B). It is important to mention that in our case, pKa calculation includes not only the influence of the interaction with other residues from the protein being in close surrounding, but also the interaction with the polypeptide (Ala_6-substrate) bound to the active site. Nowadays, thanks to the efficient methods based on very fast empirical prediction of pK_a algorithm, the analysis of protonation state of these two key residues Asp25(A) and Asp25(B) is computationally feasible. The obtained result, shown in detail in the Supporting Information, are in agreement with the current opinions that suggest meaningfully different pK_a values for Asp25(A) and Asp25(B) (of 9.9 and 4.2, respectively). This result indicates that the catalytic aspartate residues must be in a monoprotonated state at physiological pH. Significantly different pK_a values were also predicted by Trylska et al.,[106] and more recently by Jensen and co-workers,[62] and by Roitberg and co-workers.[63]

Nevertheless, from the thermodynamic point of view, the free energy surface corresponding to the proton exchanged between Asp25(A) and Asp25(B) through one water molecule, shown in Figure 2, indicates that the two proposed MC do not differ in energy (the free

energy difference between MC AspH25(B) and MC AspH25(A) is 0.7 kcal·mol^{-1}). Moreover, the activation free energy of 2.5 kcal·mol^{-1} obtained for such proton jump suggest that both protonation states would be equally present in HIV-1 PR. Thus, in the present work both MC protonation states, whose representative structures are presented in right panels of Figure 2, were considered as starting point in the study of the HIV-1 PR reaction mechanism. The free energy surfaces computed at AM1/MM, for this and the rest of explored chemical steps, can be found in the Supporting Information.

Reaction mechanisms

As already shown in the introduction section, despite a consensus on many of its features, a clear statement about the reaction mechanism catalyzed by HIV-1 PR is still missing. Thus, in this study we are using several tools to explore and test the proposed mechanisms of proteolysis catalyzed by HIV-1 PR starting from the two conformations of the MC deduced from the above preliminary calculations on protonation state of the aspartate residues located in the active site.

Concerted mechanism—As shown on Scheme 1, there is only one proposal corresponding to a concerted mechanism that starts from the protonated Asp25H(B) and a negatively charged Asp25(A). In this proposal the nucleophilic attack of water is combined with a simultaneous proton attack to the scissile C-N peptide bond. Thus, the free energy surface of this step was explored by controlling the simultaneous nucleophile attack of oxygen of activated water molecule O^{WAT} on carbonyl carbon C^{Ala3} of the peptide and the proton transfer from the same water molecule H^{WAT} to the nitrogen N^{Ala4} of the scissile peptide bond. Our calculations suggest that, despite the initial free energy surface could indicate the possibility of this mechanism, any attempt to localize a TS structure at M06-2X/MM level for this concerted mechanism was unsuccessful (see Supporting Information for details), what would be in agreement with the predictions of Pederson and co-workers.[30] On the contrary, when exploring the mechanism from the protonated AspH25(A), the resulting surface, presented in Figure 3, and the localized TS (reported in the Supporting Information) show that this process can take place in a concerted manner. Nevertheless, the free energy barrier appears to be significantly high (43.5 kcal·mol^{-1}), by comparison with barriers that can be derived from experimental data such as 16.4 kcal·mol^{-1} (k_{cat} 3.28 ± 0.08 s^{-1}) obtained for proteolysis of aminobenzoyl-Thr-Ile-Nle*pNO_2-Phe-Gln-Arg-NH_2 substrate,[52] or in the range between 17.9 and 15.1 kcal·mol^{-1} (k_{cat} ranged between 0.24 and 0.29 s^{-1}) for Ac-Ser-Gln-Asn-Tyr-Pro-Val-Val-NH_2, Ac-Arg-Ala-Ser-Gln-Asn-Tyr-Pro-Val-Val-NH_2, Ac-Ser-Gln-Ser-Tyr-Pro-Val-Val-NH_2 and Ac-Arg-Lys-Ile-Leu-Phe-Leu-Asp-Gly-NH_2 oligopeptides.[107]

Mechanism via an oxyanion intermediate—The first explored step-wise mechanism, *via* an OA intermediate, is labeled in Scheme 1 as mechanism [B]. It is assumed that in the first step of this mechanism the active site water molecule, activated by AspH25(A) and Asp25(B), attacks the carbonyl carbon C^{Ala3} of peptide bond and a proton H^{WAT} from this water is transferred to Asp25(B) leading to a metastable OA intermediate. Subsequently, the proton H^{WAT} is transferred from Asp25H(B) to nitrogen N^{Ala4} and the peptide bond is broken. In this mechanism the role AspH25(A) would be to activate the water molecule, not

being directly involved in the covalent breaking or forming bonds along the reaction pathway.

In order to explore the free energy surfaces associated to these two steps, one 2D-PMFs and one 1D-PMF were computed, which are shown in Figure 4. The first step was generated from the MC with protonated AspH25(A), by controlling the distance between the oxygen atom of the water molecule and the carbonyl carbon atom of the Ala4, d(O^{WAT}-C^{Ala3}), and the antisymmetric combination of distances defining the position of the transferring proton from the water molecule to Asp25(B): d(O^{WAT}-H^{WAT}) and d(H^{WAT}-$O^{Asp25(B)}$). The obtained free energy surface, presented on Figure 4A, shows a possible reaction path leading to the OA intermediate that is reached after crossing a free energy barrier of 11.5 kcal·mol^{-1}. It is important to note that an activation free energy of 2 kcal·mol^{-1} is obtained for the inverse reaction, the decomposition of the OA back to the MC, thus indicating the metastable character of this intermediate. Interestingly, the 2D-PMF of Figure 4A shows another local minimum in the upper left corner of the surface, at similar free energy value than the MC AspH25(A), which corresponds to the alternative protonation state of the MC in the active site: unprotonated Asp25(A) and protonated AspH25(B). The presence of this local minimum in the free energy surface confirms that the reaction from this MC to the OA intermediate is also feasible, with also similar free energy of activation.

From the **OA** intermediate, a 1D-PMF corresponding to the transfer of H^{WAT} from AspH25(B) to the nitrogen atom of the scissile peptide bond, together with the breaking of this peptide bond, has been generated and it is shown in Figure 4B. The obtained free energy profile clearly indicates the appearance of a zwitterion intermediate (ZW), in which the nitrogen atom is doubly protonated but the C^{Ala3} - N^{Ala4} peptide bond is not broken, yet. This structure, that appeared as a stable species in previous MD/AMD simulations,[23] presents a small but noticeable barrier to proceed to product complex according to our results.

Within the information of the full free energy profile of this mechanism, the overall rate limiting step is the one controlled by the TS2(OA) that corresponds to the transformation from OA to ZW, involving a proton transfer from the Asp25(B) oxygen atom $O^{Asp25(B)}$ to the nitrogen N^{Ala4}. Considering the MC as the reference state, the free energy barrier for this step is equal to 21.1 kcal·mol^{-1}, which is still rather high by comparison with the experimental observations. The activation free energy of the N-C peptide bond breaking from the ZW intermediate is really small. A transient but relevant behavior of a zwitterion species was already observed in the inverse process, the peptide bond formation, catalyzed by ribosome,[108] and also in the counterpart non-enzymatic reaction in solution.[109, 110]

Mechanism via a gem-diol intermediate—The last explored mechanism corresponds to the one proceeding *via* a gem-diol (GD) intermediate, depicted in Scheme 1 as mechanism [C]. It is generally assumed that in the first step, that leads to the GD intermediate, a water molecule activated by AspH25(A) and Asp25(B), attacks the carbonyl carbon C^{Ala3} of the peptide bond concomitant with the transfer of two hydrogen atoms: H^{WAT} from water to Asp25(B) and $H^{Asp25(A)}$ from AspH25(A) to the carbonyl oxygen atom O^{Ala3} of Ala3. The second step, from GD to products, corresponds to the transfer of H^{WAT}

from AspH25(B) to the nitrogen N^{Ala4}, the $H^{Asp25(A)}$ transfer back from O^{Ala3} oxygen to Asp25(A) and the C-N peptide bond breaking. The explored free energy surfaces are shown in Figure 5.

The first conclusion derived from these results is that the formation of the GD intermediate can take place in a single step with a barrier of 8.5 kcal·mol^{-1}. Incidentally, a feasible path connecting OA to GD intermediates was obtained with an activation free energy barrier of 2.0 kcal·mol^{-1}. The corresponding PMFs can be found in the Supporting Information. This result means that an alternative step-wise reaction path for the formation of GD intermediate can take place through the OA intermediate. Nevertheless, it would be kinetically less favorable considering that the energy required for surmounting the TS that connects MC and the OA intermediate (13.5 kcal·mol^{-1}) is already higher than the activation energy to reach the TS of the concerted process that connects MC and GD. Moreover, the OA intermediate is 15.9 kcal·mol^{-1} less stable than the GD intermediate, which in turn is also more stable than the MC by 3 kcal·mol^{-1}. Our predicted activation free energy for the formation of the GD intermediate, 8.5 kcal·mol^{-1}, is lower than the ones previously obtained by Rothlisberger and co-workers[13] (18.1 kcal·mol^{-1}) and the potential energy barrier obtained by Tsuda and co-workers (16.26 kcal·mol^{-1}).[111] Anyway, this step is not the rate limiting step of the full chemical process since, according to the PMF depicted in Figure 5B, the decomposition of the GD intermediate into products presents a free energy barrier of 15.0 kcal·mol^{-1}.

Once the free energy surfaces of every single step have been computed, a complete picture of the possible reaction paths from MC to product complex can be drawn. The schematic representation of all possible chemical transformations is presented in Figure 6, where the free energy barriers for every single step are also reported. A detail of representative snapshots of the different TSs, the OA and the GD intermediates are reported in Figure 7, while a table containing key interatomic distances is deposited in the Supporting Information. According to our results, the proteolysis catalyzed by HIV-1 PR would proceed in a step-wise manner through the formation of a stable GD intermediate. The rate limiting step would be the decomposition of this intermediate into products, taking place through the TS2(GD) with an activation free energy of 15.0 kcal·mol^{-1}. This value is in very good agreement with experimental data that, as mentioned before, oscillates between 15.1 and 17.9 kcal·mol^{-1}, depending on the peptide.[52, 107] The other explored mechanisms, through the OA intermediate or through a concerted mechanism, present significantly higher energy barriers that preclude their participation. The obtained reaction free energy describes an exergonic process, which is consistent with an enzymatic process.

Isotope effects

Once the reaction mechanism of proteolysis catalyzed by HIV-1 PR has been explored, the next step involves testing our predicted results by computing magnitudes that can be directly compared with experimental measurements. These are, apart from the already obtained free energy barriers that can be related with rate constants in the framework of the TST, the EIEs due to the different interactions established between the Ala_6-peptide and the aqueous solution or the active site of HIV-1 PR, or the KIEs due to the substitution of atoms involved in the forming and breaking bonds. A schematic representation of the definition of the

different isotope effects is shown in Figure 8, while the obtained results are listed in Table 1. Due to computer limitations, only one value of BIEs, EIEs and KIEs was computed at M06-2X/MM level for each mechanism and, consequently, no measure of uncertainty can be computed for the values listed in Table 1. Nevertheless, according to the standard deviations of 10 different values of BIEs, EIEs and KIEs computed at AM1/MM level (see the Supporting Information), similar low uncertainty (on the third decimal place) can be expected to the values reported in Table 1. The kinetic effect of converting the naturally abundant isotope protein into its “heavy” counterpart will be treated in a separately section since, while it is true that enzyme KIEs can be used to compare theoretical predictions and experimental data, they will be used primarily to get insights into enzyme dynamics.

Equilibrium Isotope Effects—In our previous work focused on the special case of EIEs related to the binding step of the substrate in the HIV-1 RT, i.e. BIEs,[112] it was concluded that BIEs are expected to differ from 1 (value that corresponds to no isotope effect) only in cases when ligands bind to hydrophilic cavities.[69] Since active site of HIV-1 PR can be considered as hydrophilic, we have decided to examine BIEs for ^{14}C-Ala3, ^{18}O-Ala3 and ^{15}N-Ala4 to check whether the obtained values can be useful in the interpretation of the interactions between the active site and the substrate. Traditionally, BIEs could be measured experimentally and used to compare known inhibitors with predicted new ones.[113] Our calculations reflect that, as expected, negligible BIEs were obtained for ^{14}C-Ala3 and ^{15}N-Ala4 (0.993 and 0.991, respectively) indicating the lack of hydrogen bond type interactions with surrounding residues both in water and in the active site of HIV-1 PR. Different situation occurs in the case of ^{18}O-Ala3, where normal BIEs of 17% are obtained. This value reveals a meaningful change in the interactions between this residue of the peptide and the water molecules of the aqueous solution and the active site of HIV-1 PR. In particular, the normal value indicates a weaker hydrogen bond interaction in the active site than in the aqueous solution.

Two sets of EIE have been computed in the present work taking the structures generated along the stepwise mechanisms by comparing the MCs with the intermediate OA or GD (see Table 1). According to the average structures of both intermediates, no significant differences would be expected for the computed isotope effects. As observed in Figure 7, the scissile peptide bond in the OA intermediate is slightly more elongated than in the GD intermediate (1.53 vs 1.49 Å) which can explain the slightly higher normal ^{15}N and ^{14}C EIE observed when considering the equilibrium between the MC and the OA intermediate (1.034 and 1.013 Å, respectively) than when considering the equilibrium with the GD intermediate (1.008 and 1.005 Å, respectively). This small difference is also related with the higher charges on the C and N atoms of the scissile peptide bond in the OA than in the GD; 0.606 and −0.511 a.u. in the former vs 0.701 and −0.798 a.u. in the later. A table with atomic charges for all key atoms in all localized states is reported in the Supporting Information. The protonation of the carbonyl oxygen atom of Ala3 in the GD intermediate is also responsible of getting a small but inverse ^{18}O EIE when considering the equilibrium of the GD intermediate, but a small normal effect in the case of formation of the OA intermediate. This new hydroxylic group presents a strong hydrogen bond interaction with the Asp25A in the GD, producing a tighter force constant associated to the oxygen atom and explains the

inverse value of the EIE. If the oxygen atom is not protonated, as occurring in the OA intermediate, the interaction with the Asp25(A) is slightly weaker and a small but normal EIE is obtained.

Kinetic Isotope Effects (KIEs)—Analysis of KIEs can render more interesting conclusions not only because they can provide indirect information of the TSs, but also because experimental data is available for our system, despite they were measured for reactions with different peptides. KIEs have been computed for the isotopically substitution of the three atoms directly involved in the peptide bond breaking; ^{14}C-Ala3, ^{18}O-Ala3 and ^{15}N-Ala4. The TSs used to compute these KIEs where those localized in the rate limiting steps of the three explored mechanisms. As shown in Table 1, and in agreement with the conclusions derived from the analysis of the energetics of the alternative reaction paths, only the KIEs obtained from the GD mechanism are in qualitative and almost quantitative agreement with the experimental data.[33, 114]

The normal ^{14}C KIE obtained with the corresponding TSs of the three mechanisms means a tighter force constant associated to the C atom in the MC than in the TS, which can not be explained by the hybridization state (changing from sp2 to sp3) but by the interactions established with this atom. As can be observed by comparing the corresponding structures presented in Figure 2 and Figure 7, the scissile C-N bond is dramatically elongated in all the three TSs by comparison with the initial state in the MC.

The primary ^{15}N KIE and the secondary ^{18}O KIE are even more symptomatic. Thus, while the former is normal in the concerted and in the OA mechanism, the inverse value obtained with the GD mechanism would be the only one in agreement with experiments of Schramm and co-workers[33] or Meek and co-workers.[107] This inverse KIE can be rationalized based on the fact that the nitrogen atom establishes an additional interaction with the proton from Asp25(B) (1.20 Å) while the C-N bond is still, at some extent, preserved (1.56 Å). Thus, much tighter surrounding of the nitrogen is observed on TS than in MC. In the case of the TSs of the concerted and the step-wise mechanism through the OA intermediate, the distance between the N and the transferring proton are significantly larger (1.32 and 1.26 Å, respectively), and the scissile peptide bond more elongated in the case of the TS2(OA) (1.58 Å).

The secondary ^{18}O KIE is inverse in the concerted and in the GD mechanism (0.997 and 0.993, respectively). In the concerted mechanism, this effect can be explained by the strong interaction that is established with the proton of AspH25(A) in the TS (1.52 Å) while in the GD mechanism this proton participates in the reaction since it is being transferred from the oxygen atom to the Asp25(A), thus showing an even shorter distance is detected (1.45 Å) that reveals a strong interaction. On the contrary, the TS of the rate limiting step of the OA mechanism does not involve this proton transfer and this interaction is not so strong: the distance between the oxygen atom of the peptide and the hydrogen atom of AspH25(A) is significantly larger (1.82 Å). As a consequence, a normal KIE is obtained (1.007), not in accordance with the experimental data of Schramm and co-workers.[33]

A final conclusion that can be derived from our calculations is that, if the KIEs are computed using the GD intermediate as the reference state (see the $KIE^{(INT-TS)}$ column on Table 1), the agreement with the experimental data is much worse than when KIEs are computed between MC and the rate limiting step of the GD mechanism.

Dynamic effects: enzyme KIEs

Once the GD mechanism has been confirmed as the most favourable one employed by HIV-1 PR to break the peptide bond, the study of the dynamic effects has focused on the rate limiting step of this mechanism. According to the kind of chemical reaction involved in this step, no dramatic tunnelling effects should be expected since, despite a proton transfer takes place, this chemical steps involves the breaking of a carbon-nitrogen bond. In any case, as revealed in previous studies, the tunnelling transmission coefficients obtained in light and heavy enzymes are equivalent within the standard deviations.[98, 99, 100, 101, 102] Then, as commented in the Computational Methods section, since in our QM/MM simulations the reaction coordinate does not depend on coordinates of the protein, any non-equilibrium influence of protein dynamics (revealed as a variation of the rate constant due to the vibrational shift of protein motions caused by mass substitution) should therefore be captured in the recrossing transmission coefficient and the enzyme KIEs were computed as the ratio of these coefficients in the light and heavy enzyme (eq. 10). As also explained in previous section, these were computed by means of the Grote Hynes theory[103] that requires MD simulations starting from TSs structures. Then, the recrossing transmission coefficients are necessary based on AM1/MM simulations and the corresponding TSs structures selected from the quadratic region of the 2D PMF generated at this level of theory. The corresponding AM1/MM free energy surface, presented in Figure 9, shows a stepwise mechanism for this chemical step, taking place through a stable ion-pair like intermediate. Consequently, the transmission coefficients were independently computed from these two TSs.

The results of the recrossing transmission coefficients, the equilibrium and the reactive frequencies computed for light and heavy variant, together with the enzyme KIEs obtained from the calculations from TS3(GD) and TS4(GD), are reported in Table 2. As observed, the first conclusion that can be obtained is that there are more non-reactive recrossing trajectories in the heavy enzyme than in the light enzyme, either for the TS3(GD) and TS4(GD). This is revealed in the values of the recrossing coefficient, being further than unity in the heavy enzyme than in the light enzyme. The heavy enzyme would not follow the chemical transformation as well as the light enzyme, which in turn is reflected in lower recrossing coefficients. These values, different from unity, would support a contribution of the protein into the real reaction coordinate, not considered in the generation of the classical PMFs. Nevertheless, this contribution that can be interpreted as a dynamic contribution of the enzyme, when transformed into an effective free energy contribution reveals a small contribution in reducing the effective energy barrier by 0.77 to 0.28 $kcal \cdot mol^{-1}$, depending on the TS. Table 2 provides also the equilibrium frequency ω^{eq}, derived from the curvature of the free energy profile in the region around the barrier top (see Supporting Information), and the reaction frequency (ω^{react}) that is the actual frequency of the system motion along the reaction coordinate. The ratio between these two magnitudes is, in fact, the transmission

coefficient. By comparison of the reaction frequency in the light and heavy enzyme, it seems that the friction, or the coupling between the reaction coordinate and the remaining degrees of freedom of the system, is significantly larger in the later. The normalized autocorrelation functions (ACFs) of the forces acting on the reaction coordinate in the light and heavy enzymes, displayed in the Supporting Information, shows a very rapid relaxation followed by fast oscillations in both systems, although slightly slower in the heavy than in the light enzymes. The behavior of the normalized ACFs at the transition state are in agreement with previous studies where a very rapid relaxation is observed. The decay is associated with a fine structure that can be attributed to the coupling of the reaction coordinate with other less relevant fast vibrational motions of the environment or the chemical system (i.e. X-H stretchings). This autocorrelation function, in fact, determines the friction kernel as expressed in eq. 15.

With the obtained values of the recrossing transmission coefficients, the resulting enzyme KIEs are 1.228 and 1.042, computed from TS3(GD) and TS4(GD), respectively. Interestingly, the experimental enzyme KIE of 1.19, measured by Schramm and co-workers,[33] is in between these two theoretical predictions than can be considered as the limiting cases. In fact, when the AM1/MM free energy surface of the transformation from GD into products is corrected at M06-2X/MM level, the reaction is described as a concerted process (see Figure 5b). Despite there are precedents on the formation of stable ion pair stationary state structures before peptide bond breaking in high-level DFT PESs,[115] the change in the topology of PESs or free energy surfaces after high level QM/MM corrections is an effect that has been already observed in previous studies.[116, 117, 118] In particular, intermediates located for enzyme catalyzed reactions involving a double proton and hydride transfers, when the QM subset of atoms were described by a semiempirical method, were transformed into just a shallow minimum when improving the description of the QM region at DFT level,[116] or even disappeared when the correction was at DFT[117] or MP2[118] level. Then, it is reasonable to believe that high level M06-2X/MM QM/MM MD trajectories could perfectly predict enzyme KIE that would fit within the two values obtained at AM1/MM level, supporting a reaction mechanism through a GD intermediate as shown in Figure 6.

Electrostatic effects

The analysis of the electrostatic effects of the HIV-1 PR in catalysing the proteolysis of the Ala6 peptide has been based on the calculations of the charges of key atoms involved in the reaction and the electric field generated by the protein inside the active site in different states along the reaction path. Tables with a complete list of computed values can be found in the Supporting Information. In order to proceed with the analysis of the results, two imaginary planes have been drawn. The first one dividing the active site such that the peptide would remind on one side and on the other side the two active site aspartate residues (horizontal dashed lines in Figure 10). The second plane, perpendicular to this one, would cross the scissile peptide bond (vertical dashed lines in Figure 10). Then, the sum of the atomic charges of the QM atoms in each side of both planes were computed and reported in Figure 10.

As Figure 10 reveals, an important charge transfer takes place across the vertical plane when comparing the two MCs. This is basically associated with the proton transfer from Asp25(A) to Asp25(B). Nevertheless, the energy required for this charge transfer appears to be negligible, according to the activation free energy and the almost equal free energy that present both states in the free energy surface on Figure 2. This is in agreement with the small value of the projection of the electric field created by the protein on the vector defined between the two aspartate residues, $\overrightarrow{E}_{\parallel}$, depicted as green arrows in Figure 10. According to the results, the electric field created by the protein would stabilize both complexes and the work in moving a positive charge between both aspartate residues would be small. This process does not involve a significant charge transfer across the perpendicular plane, being a much more negative charge located in the lower side of the active site (−0.890 and −0.938 a.u.) than in the upper side (−0.110 and −0.062 a.u.) in both MCs. On the contrary, when comparing the MCs with the OA intermediate, an important charge transfer takes place through the horizontal plane (ca. −0.638 a.u.). We must keep in mind that this step requires a significant activation free energy (12.5 kcal·mol^{-1}) and it is endergonic (10.5 kcal·mol^{-1}). Considering the magnitude of the projection of the electric field in this perpendicular direction, $\overrightarrow{E}_{\perp}$, a negative charge transfer from top to bottom, equivalent to the transformation from MC to the OA intermediate, would require a significant work. On the contrary, the step from the MC to the GD intermediate, that presented a lower barrier (8.5 kcal·mol^{-1}) and was an exergonic process (−3.5 kcal·mol^{-1}), involves a significantly lower charge transfer of ca. 0.147 a.u. and, consequently, despite the evolution of the electric field from MC to GD behaves similarly, the electrostatic work is significantly smaller than from MC to OA. Finally, when analysing the transformation from the GD to products, which was the rate limiting step of the most favourable mechanism, a small charge transfer is also observed from the bottom to the top side of the active site (0.140 a.u.). All in all, it appears that the generation of a negative charge in the side where the peptide is located represent a non-favourable process for the enzyme, while the transfer in the direction of the scissile peptide bond does not requires an energetic penalty. An analysis based on electrostatic arguments in the first step of the reaction catalyzed by HIV-1 PR (the formation of the GD intermediate) has been recently carried out by Fernandes and co-workers to explain the effect of the conformational fluctuations in the activation free energy barriers.[14] The authors postulated a flow of negative charge from the unprotonated active site aspartate residue in reactant complex (Asp25B in our Figure 6) to the substrate (around the carbon atom of the scissile peptide bond) and the hydroxide ion in the TS of the GD formation step. This charge transfer was used to explain the role of protein residues on stabilizing this TS and, since the proton was not transferred to the peptide in the TS of the GD formation, their conclusions were valid. Nevertheless, our simulations reveal that the hydroxide transfer is associated with a proton transfer from the other protonated aspartate to the carbonyl oxygen atom of the peptide and, in fact, the charge on the carbon atom increases from 0.467 a.u. to 0.701 a.u. in the GD. According to our results, the carbonyl oxygen atom of the scissile peptide bond would be the center where the largest amount of negative charge would be accumulated (from −0.569 a.u. to −0.728 a.u. in the GD). This charge transfer would be, nevertheless, more dramatic if the process would take place through the OA intermediate, as can be observed in the analysis of the evolution of atomic charges. A positive value of the electric

field in the perpendicular direction, from bottom to top, is in accordance with an electrostatic work required if a negative charge would be displaced in this direction.

Finally, in order to focus on the C-N breaking bond, the projection of the electric field created by the protein in the direction of this bond has been computed in the two involved atoms. Now, considering the atomic charges on the carbon and nitrogen atoms in the GD (0.491 and −0.141 a.u., respectively) and the electric field on both atoms (0.5×10^{-3} and 4.8×10^{-3} a.u. in the C and N, respectively), it appears that the protein would be stabilizing this intermediate. Interestingly, when computing the charges (0.364 and 0.011 a.u. in the C and N, respectively) and the electric field (-0.8×10^{-3} and 3.6×10^{-3} a.u. in the C and N, respectively) in the TS of the rate limiting step (TS2(GD) from GD to products), the combination of electrostatic forces generated by the protein in the two involved atoms would be helping to break the C-N peptide bond (see Figure 11). This behaviour is not observed in the rate limiting step of the mechanism through the OA intermediate that, while protein was creating similar electric fields on the C and N atoms (-1.7×10^{-3} and 2.8×10^{-3} a.u. in the C and N, respectively), the resulting electrostatic forces on the two atoms would not favour the breaking of the bond due to the negative charge on the N atom computed in this TS (TS2(OA) from OA to products).

Conclusions

In the present study the mechanisms of proteolysis catalyzed by HIV-1 PR proposed until now have been explored, i.e. the concerted mechanism and two different step-wise mechanisms through an oxyanion or a gemdiol intermediate. Free energy surfaces for each step of the proteolysis of a 6 alanines peptide have been generated in terms of PMFs within hybrid QM/MM potentials at AM1/MM and M06-2X/6–31+G(d,p)/MM levels. Based on the obtained results it has been concluded that the reaction pathway of proteolysis catalyzed by HIV-1 PR must proceed as a two-step process. In the first step a nucleophilic attack of a water molecule on the carbon atom C^{Ala3} of Ala3 is accompanied by a hydrogen transfer from this water molecule to Asp25(B) resulting in a gemdiol intermediate. Subsequently, the peptide bond is broken concertedly with a double proton transfer, from the oxygen of the protonated Asp25(B) to the nitrogen atom of the scissile peptide bond and from one of the hydroxyl groups of the carbon atom of the peptide bond to the Asp25(A). The rate-limiting step would correspond to the gemdiol decomposition into the products complex, with a computed free energy barrier, 15.0 kcal·mol^{-1}, very close to the experimentally measured values of the hydrolysis of different peptides catalyzed by HIV-1 PR, that are in the range between 15.1 and 17.9 kcal·mol^{-1}.[52, 107] Moreover, calculation of primary ^{14}C and ^{15}N KIEs and secondary ^{18}O KIEs based on structures of the transition state of this step (1.022, 0.995 and 0.993 for ^{14}C ^{15}N, and ^{18}O isotopic substitutions, respectively) are in perfect agreement with experimental data (1.029 ± 0.003, 0.987 ± 0.004 and 0.993 ± 0.003, respectively). KIEs computed with structures of intermediates appearing along a step-wise enzymatic mechanism as the reference state can be used to confirm whether the experimental measurements correspond to an overall or intrinsic KIE. In this case, the lack of agreement between the experimental KIEs and the KIEs computed with structures of the OA and GD as the reactant state confirms that the measurements correspond to an overall KIE.

Once the mechanism has been solved, the role of the protein dynamics has been studied by computing the rate constant after increasing the mass of the enzyme by 10.24%, as the result of substituting all carbon, nitrogen, and nonexchangeable hydrogen atoms with ^{13}C, ^{15}N, and ^{2}H, respectively. The enzyme KIEs has been estimated as the ratio of the recrossing transmission coefficients in the light and heavy enzyme, computed by means of the Grote Hynes theory[103] from the rate-limiting TSs structures. Since these calculations require long MD simulations, we were restricted to the use of AM1/MM potentials. The corresponding AM1/MM free energy surface, when exploring the decomposition of the GD intermediate into products, shows a stepwise mechanism for this chemical step, taking place through a stable ion-pair like intermediate. Due to the similar energetics of the two involved AM1/MM TSs, the transmission coefficients were independently computed from both, TS3(GD) and TS4(GD). The results of the recrossing transmission coefficients, the equilibrium and the reactive frequencies computed for light and heavy variant obtained from the calculations from TS3(GD) and TS4(GD), show that there are more non-reactive recrossing trajectories in the heavy enzyme than in the light enzyme. The heavy enzyme would not follow the chemical transformation as well as the light enzyme does, which in turn is reflected in lower recrossing coefficients. These values, different from unity, would support a contribution of the protein into the real reaction coordinate, not considered in the generation of the classical mechanical PMFs. Nevertheless, these values, that can be interpreted as a dynamic contribution of the enzyme, when transformed into an effective free energy, reveal a small contribution of 0.77 and 0.28 kcal·mol^{-1}, for TS3(GD) and TS4(GD), respectively. Consequently, within these results we can claim that dynamic effects have a low impact in the chemical step of the reaction catalyzed by HIV-1 PR.

From the obtained values of the recrossing transmission coefficients, the resulting enzyme KIEs are 1.228 and 1.042, computed on TS3(GD) and TS4(GD), respectively. Interestingly, the experimental enzyme KIE of 1.19, measured by Schramm and co-workers,[52] is in between these two theoretical predictions that can be considered as the limiting cases. In fact, when the AM1/MM free energy surface of the transformation from GD into products (the rate limiting step of the reaction) is corrected at M06-2X/MM level, the reaction is described as a concerted process. This suggests that the enzyme KIE, if computed from high level QM/MM MD trajectories, could perfectly fit within the values deduced from the two AM1/MM TSs, whose average (1.135) agrees with the experimental data. Once again, the results support a reaction mechanism through a GD intermediate and our analysis of the dynamic effects.

Finally, the electrostatic effects of HIV-1 PR have been estimated by analysing the evolution of charge transfer in the active site of the protein and by computing the electric field generated by the enzyme in center of the active site. The electric field has been projected in two directions defined by the scissile C-N bond, and the one perpendicular to this bond into the direction of the two active site aspartate residues. Thus, the electric field obtained in these directions suggests that the transfer of a negative charge from the aspartate residues to the region where the peptide is located represents a non-favourable process for the enzyme, while the transfer along the scissile peptide bond does not requires a significant energetic penalty.

The value of the projection of the electric field in the direction of the scissile C-N bond is small, which means that the electric work to move a charge in this direction is negligible while a positive value of the electric field in the perpendicular direction, from bottom to top, means that an electric work would be required if a negative charge would be displaced in this direction. In fact, the decomposition of the electrostatic forces generated by the protein in the scissile peptide bond on the rate limiting transition state would favour the peptide bond cleavage. Thus, it appears that electrostatic effects on HIV-1 PR are definitely favouring the reaction to take place. These electrostatic properties of the active site could be used as a mould for future drug design.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

AK and KŚ acknowledge the Polish Ministry of Science and Higher Education (“Iuventus Plus” program project no. 0478/IP3/2015/73, 2015–2016). V.M. acknowledges the Spanish Ministerio de Economía y Competitividad (project number CTQ2015-66223-C2-1-P), by Generalitat Valenciana (PrometeoII/2014/022), by Universitat Jaume I (Project P1 1B2014-26) and the USA National Institute of Health (ref. NIH R01 GM065368). Authors acknowledge computational resources from the Servei d’Informàtica of Universitat Jaume I and Informatic Center on the ‘Blueocean’ supercomputer of Lodz University of Technology.

Reference

1. Prabu-Jeyabalan M, Nalivaika E, Schiffer CA. Structure. 2002; 10:369–381. [PubMed: 12005435]
2. Tzoupis H, Leonis G, Megariotis G, Supuran CT, Mavromoustakos T, Papadopoulos MG. J. Med. Chem. 2012; 55:5784–5796. [PubMed: 22621689]
3. Coates L, Tuan HF, Tomanicek S, Kovalevsky A, Mustyakimov M, Erskine P, Cooper J. J. Am. Chem. Soc. 2008; 130:7235–7237. [PubMed: 18479128]
4. Freedberg DI, Ishima R, Jacob J, Wang YX, Kustanovich I, Louis JM, Torchia DA. Protein Sci. 2002; 11:221–232. [PubMed: 11790832]
5. Hornak V, Okur A, Rizzo RC, Simmerling C. Proc. Natl. Acad. Sci. U.S.A. 2006; 103:915–920. [PubMed: 16418268]
6. Hornak V, Okur A, Rizzo RC, Simmerling C. J. Am. Chem. Soc. 2006; 128:2812–2813. [PubMed: 16506755]
7. Nicholson LK, Yamazaki T, Torchia DA, Grzesiek S, Bax A, Stahl SJ, Kaufman JD, Wingfield PT, Lam PY, Jadhav PK. Nat. Struct. Biol. 1995; 2:274–280. [PubMed: 7796263]
8. Rahuel J, Priestle P, Grutter MG. J. Struct. Biol. 1991; 107:227–236. [PubMed: 1807356]
9. Jensen C, Herold P, Brunner HR. Nat. Rev. Drug Discovery. 2008; 7:399–410. [PubMed: 18340340]
10. Torbeev VY, Raghuraman H, Hamelberg D, Tonellie M, Westlere WM, Perozo E, Kent SBH. Proc. Natl. Acad. Sci. U.S.A. 2011; 108:20982–20987. [PubMed: 22158985]
11. Piana S, Carloni P, Rothlisberger U. Protein Sci. 2002; 11:2393–2402. [PubMed: 12237461]
12. Piana S, Carloni P, Parrinello M. J. Mol. Biol. 2002; 319:567–583. [PubMed: 12051929]
13. Piana S, Bucher D, Carloni P, Rothlisberger U. J. Phys. Chem. B. 2004; 108:11139–11149.
14. Ribeiro AJM, Santos-Martins D, Russo N, Ramos MJ, Fernandes PA. ACS Catal. 2015; 5:5617–5626.
15. Gillespie, AL. The Natural History of Digestion. London: W. Scott, Ltd; 1898.
16. Van Der Kamp MW, Mulholland AJ. Biochemistry. 2013; 52:2708–2728. [PubMed: 23557014]
17. Hyland LJ, Tomaszek TA, Meek TD. Biochemistry. 1991; 30:8454–8463. [PubMed: 1883831]
18. Rodriguez EJ, Angeles TS, Meek TD. Biochemistry. 1993; 32:12380–12385. [PubMed: 8241126]

19. Coates L, Erskine PT, Mall S, Gill R, Wood SP, Myles DA, Cooper JB. Eur. Biophys. J. 2006; 35:559–566. [PubMed: 16673078]
20. Northrop DB. Acc. Chem. Res. 2001; 34:790–797. [PubMed: 11601963]
21. Silva AM, Cachau RE, Sham HL, Erickson JW. J. Mol. Biol. 1996; 255:321–346. [PubMed: 8551523]
22. Park H, Suh J, Lee S. J. Am. Chem. Soc. 2000; 122:3901–3908.
23. Trylska J, Bala P, Geller M, Grochowski P. Biophys. J. 2002; 83:794–807. [PubMed: 12124265]
24. Liu H, Müller-Plathe F, Van Gunsteren WF. J. Mol. Biol. 1996; 261:454–469. [PubMed: 8780786]
25. Segun K, Padlan EA, Smith CW, Carlson WD, Davies DR. Proc. Natl. Acad. Sci. U.S.A. 1987; 84:7009–7013. [PubMed: 3313384]
26. Trylska J, Grochowski P, McCammon JA. Protein Sci. 2004; 13:513–528. [PubMed: 14739332]
27. Bjelic S, Aqvist J. Biochemistry. 2006; 45:7709–7723. [PubMed: 16784222]
28. Barrett, AJ.; Rawlings, ND. Handbook of Proteolytic Enzymes. second. Woessner, JF., editor. London: Elsevier; 2004. p. 1417-1439.
29. Jaskólski M, Tomasselli AG, Sawyer TK, Staples DG, Heinrikson RL, Schneider J, Kent SBH, Wlodawer A. Biochemistry. 1991; 30:1600–1609. [PubMed: 1993177]
30. Lee H, Darden TA, Pedersen LG. J. Am. Chem. Soc. 1996; 118:3946–3950.
31. Kovalevsky AY, Chumanevich AA, Liu F, Louis JM, Weber IT. Biochemistry. 2007; 46:14854–14864. [PubMed: 18052235]
32. Das A, Mahale S, Prashar V, Bihani S, Ferrer JL, Hosur MV. J. Am. Chem. Soc. 2010; 132:6366–6373. [PubMed: 20397633]
33. Kipp DR, Hirschi JS, Wakata A, Goldstein H, Schramm VL. Proc. Natl. Acad. Sci. U.S.A. 2012; 109:6543–6548. [PubMed: 22493227]
34. Szeltner Z, Polgar L. J. Biol. Chem. 1996; 271:32180–32184. [PubMed: 8943273]
35. Polgar L, Szeltner Z, Boros I. Biochemistry. 1994; 33:9351–9357. [PubMed: 8049236]
36. Porter DJ, Hanlon MH, Furfine ES. Biochemistry. 2002; 41:1302–1307. [PubMed: 11802730]
37. Smith R, Brereton IM, Chai RY, Kent SBH. Nat.Struct.Biol. 1996; 3:946–950. [PubMed: 8901873]
38. Piana S, Sebastiani D, Carloni P, Parrinello M. J. Am. Chem. Soc. 2001; 123:8730–8737. [PubMed: 11535077]
39. Wang YX, Freedberg DI, Yamazaki T, Wingfield PT, Stahl SJ, Kaufman JD, Kiso Y, Torchia DA. Biochemistry. 1996; 35:9945–9950. [PubMed: 8756455]
40. Clement GE, Snyder SL, Price H, Cartmell R. J. Am. Chem. Soc. 1968; 90:5603–5610. [PubMed: 4878300]
41. Cornish-Bowden AJ, Knowles JR. Biochem J. 1969; 113:353–362. [PubMed: 4897198]
42. Hunkapiller MW, Richards JH. Biochemistry. 1972; 11:2829–2839. [PubMed: 4557517]
43. Clement GE. Prog. Bioorg. Chem. 1973; 2:177–238.
44. Hofmann T, Hodges RS, James MN. Biochemistry. 1984; 23:635–643. [PubMed: 6424704]
45. Green DW, Ayknet S, Gierse JK, Zupec ME. Biochemistry. 1990; 29:3126–3133. [PubMed: 2186807]
46. Adachi M, Ohhara T, Kurihara K, Tamada T, Honjo E, Okazaki N, Arai S, Shoyama Y, Kimura K, Matsumura H, Sugiyama S, Adachi H, Takano K, Mori Y, Hidaka K, Kimura T, Hayashi Y, Kiso Y, Kuroki R. Proc. Natl. Acad. Sci. U.S.A. 2009; 106:4641–4646. [PubMed: 19273847]
47. See: Tuñón I, Laage D, Hynes JT. Arch. Biochem. Biophys. 2015; 582:42–55. and references cited in. [PubMed: 26087289]
48. Świderek K, Ruiz-Pernía JJ, Moliner V, Tuñón I. Curr. Opin. Chem. Biol. 2014; 21:11–18. [PubMed: 24709164]
49. Rokop S, Gajda L, Parmerter S, Crespi HL, Katz JJ. Biochim. Biophys. Acta. 1969; 191:707–715. [PubMed: 4903503]
50. Di Costanzo L, Moulin M, Haertlein M, Meilleur F, Christianson DW. Arch. Biochem. Biophys. 2007; 465:82–89. [PubMed: 17562323]
51. Silva RG, Murkin AS, Schramm VL. Proc. Natl. Acad. Sci. U. S. A. 2011; 108:18661–18665. [PubMed: 22065757]

52. Kipp DR, Silva RG, Schramm VL. J. Am. Chem. Soc. 2011; 133:19358–19361. [PubMed: 22059645]
53. Adamczyk AJ, Cao J, Kamerlin SCL, Warshel A. Proc. Natl. Acad. Sci. U.S.A. 2011; 108:14115–14120. [PubMed: 21831831]
54. Pisliakov AV, Cao J, Kamerlin SCL, Warshel A. Proc. Natl. Acad. Sci. U.S.A. 2009; 106:17359–17364. [PubMed: 19805169]
55. Boekelheide N, Salomón-Ferrer R, Miller TF. Proc. Natl. Acad. Sci. U.S.A. 2011; 108:16159–16163. [PubMed: 21930950]
56. Kervinen J, Lubkowski J, Zdanov A, Bhatt D, Dunn BM, Hui KY, Powell DJ, Kay J, Wlodawer A, Gustchina A. Protein Sci. 1998; 7:2314–2323. [PubMed: 9827997]
57. Accelrys Software Inc. Discovery Studio Modeling Environment, Release 3.5. San Diego: Accelrys Software Inc; 2012.
58. Gautier R, Camproux AC, Tufféry P. Nucleic Acid Res. 2004; 32:W508–W511. [PubMed: 15215438]
59. Camproux AC, Gautier R, Tufféry P. J. Mol. Biol. 2004; 339:591–605. [PubMed: 15147844]
60. Antosiewicz J, McCammon JA, Gilson MK. J. Mol. Biol. 1994; 238:415–436. [PubMed: 8176733]
61. Hui L, Robertson AD, Jensen JH. Proteins. 2005; 61:704–721. [PubMed: 16231289]
62. Hui L, Robertson AD, Jensen JH. Proteins. 2008; 73:765–783. [PubMed: 18498103]
63. McGee TD Jr, Edwards J, Roitberg AE. J. Phys. Chem. B. 2014; 118:12577–12585. [PubMed: 25340507]
64. Byrd RH, Lu P, Nocedal J, Zhu C. J. Sci. Comp. 1995; 16:1190–1208.
65. Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang J, Kollman P. J. Comput. Chem. 2003; 24:1999–2012. [PubMed: 14531054]
66. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. J. Chem. Phys. 1983; 79:926–935.
67. Field, MJ. A practical Introduction to the Simulation of Molecular Systems. Cambridge, U.K: Cambridge University Press; 1999.
68. Field MJ, Albe M, Bret C, Proust-de Martin F, Thomas A. J. Comput. Chem. 2000; 21:1088–1100.
69. Krzemińska A, Paneth P, Moliner V, Świderek K. J. Phys. Chem. B. 2015; 119:917–927. [PubMed: 25132465]
70. Dewar MJS, Zoebisch EG, Healy EF. J. Am. Chem. Soc. 1985; 107:3902–3909.
71. Zhao Y, Truhlar DG. Theor. Chem. Acc. 2008; 120:215–241.
72. Zhao Y, Truhlar DG. Acc. Chem. Res. 2008; 41:157–167. [PubMed: 18186612]
73. Kirkwood JG. J. Chem. Phys. 1935; 3:300–313.
74. McQuarrie, DA. Statistical Mechanics. New York: Harper and Row; 1976. p. 266-267.
75. Roux B. Comput. Phys. Commun. 1995; 91:275–282.
76. Torrie GM, Valleau JP. J. Comput. Phys. 1977; 23:187–199.
77. Schenter GK, Garrett BC, Truhlar DG. J. Chem. Phys. 2003; 119:5828–5833.
78. Nguyen KA, Rossi I, Truhlar DG. J. Chem. Phys. 1995; 103:5522–5530.
79. Corchado JC, Coitiño EL, Chuang Y, Fast PL, Truhlar DG. J. Phys. Chem. A. 1998; 102:2424–2438.
80. Chuang YY, Corchado JC, Truhlar DG. J. Phys. Chem. A. 1999; 103:1140–1149.
81. Renka RJ. SIAM J. Stat. Comput. 1987; 8:393–415.
82. Renka RJ. ACM Trans. Math. Software. 1993; 19:81–94.
83. Ruiz-Pernia JJ, Silla E, Tuñón I, Martí S, Moliner V. J. Phys. Chem. B. 2004; 108:8427–8433.
84. Roca M, Moliner V, Ruiz-Pernía JJ, Silla E, Tuñón I. J. Phys. Chem. A. 2006; 110:503–509. [PubMed: 16405322]
85. Ruiz-Pernía JJ, Silla E, Tuñón I, Martí S. J. Phys. Chem. B. 2006; 110:17663–17670. [PubMed: 16942112]
86. Zhao Y, Truhlar DG. Theor. Chem. Acc. 2008; 120:215–241.
87. Lynch BJ, Zhao Y, Truhlar DG. J. Phys. Chem. A. 2003; 107:1384–1388.

88. Ruggiero GD, Guy SJ, Martí S, Moliner V, Williams IH. J. Phys. Org. Chem. 2004; 17:592–600.
89. Glasstone, S.; Laidler, KJ.; Eyring, H. Diffusion and Electrochemical Phenomena, The Theory of Rate Processes: The Kinetics of Chemical Reactions, Viscosity. New York: McGraw-Hill; 1941.
90. Keck, JC. Advances in Chemical Physics. John Wiley & Sons, Inc; 2007. p. 85-121.
91. Truhlar DG, Garrett BC, Klippenstein SJ. J. Phys. Chem. 1996; 100:12771–12800.
92. Alhambra C, Corchado J, Sánchez ML, Garcia-Viloca M, Gao J, Truhlar DG. J. Phys. Chem. B. 2001; 105:11326–11340.
93. Garcia-Viloca M, Alhambra C, Truhlar DG, Gao J. J. Comput. Chem. 2003; 24:177–190. [PubMed: 12497598]
94. Alhambra C, Corchado J, Sanchez ML, Garcia-Viloca M, Gao J, Truhlar DG. J. Phys. Chem. B. 2001; 105:11326–11340.
95. Pu JZ, Gao JL, Truhlar DG. Chem. Rev. 2006; 106:3140–3169. [PubMed: 16895322]
96. Truhlar DG, Gao JL, Alhambra C, Garcia-Viloca M, Corchado J, Sanchez ML. Villa, J. Acc. Chem. Res. 2002; 35:341–349. [PubMed: 12069618]
97. Truhlar DG, Gao JL, Garcia-Viloca M, Alhambra C, Corchado J, Sanchez ML, Poulsen TD. International Journal of Quantum Chemistry. 2004; 100:1136–1152.
98. Luk LYP, Ruiz-Pernía JJ, Dawson MW, Roca M, Loveridge EJ, Glowacki DR, Harvey JN, Mulholland AJ, Tuñón I, Moliner V, Allemann RK. Proc. Natl. Acad. Sci. U. S A. 2013; 110:16344–16349. [PubMed: 24065822]
99. Ruiz-Pernia JJ, Luk LYP, García-Meseguer R, Martí S, Loveridge EJ, Tuñón I, Moliner V, Allemann RK. J. Am. Chem. Soc. 2013; 135:18689–18696. [PubMed: 24252106]
100. Luk LYP, Ruiz-Pernía JJ, Loveridge EJ, Adesina AS, Loveridge EJ, Tuñón I, Moliner V, Allemann RK. Angew. Chem. 2015; 54:9016–9020. [PubMed: 26079622]
101. Ruiz-Pernía JJ, Behiry E, Luk LYP, Loveridge EJ, Tuñón I, Moliner V, Allemann RK. Chem. Sci. 2016; 7:3248–3255.
102. Luk LYP, Ruiz-Pernía JJ, Dawson WM, Loveridge EJ, Tuñón I, Moliner V, Allemann RK. J. Am. Chem. Soc. 2014; 136:17317–17323. [PubMed: 25396728]
103. Grote RF, Hynes JT. J. Chem. Phys. 1980; 73:2715–2732.
104. Gertner BJ, Wilson KR, Hynes JT. J. Chem. Phys. 1989; 90:3537–3558.
105. Kim HJ, Hynes JT. J. Am. Chem. Soc. 1992; 114:10508–10528.
106. Trylska J, Antosiewicz J, Geller M, Hodge CN, Klabe RM, Head MS, Gilson MK. Protein Sci. 1999; 8:180–195. [PubMed: 10210196]
107. Hyland LJ, Tomaszek TA, Roberts GD, Carr SA, Magaard VW, Bryan HL, Fakhoury SA, Moore ML, Minnich MD, Culp JS, DesJarlais RL, Meek TD. Biochemistry. 1991; 30:8441–8453. [PubMed: 1883830]
108. Świderek K, Marti S, Tuñón I, Moliner V, Bertran J. J. Am. Chem. Soc. 2015; 137:12024–12034. [PubMed: 26325003]
109. Świderek K, Marti S, Tuñón I, Moliner V, Bertran J. Chem. Commun. 2012; 48:11253–11255.
110. Świderek K, Marti S, Tuñón I, Moliner V, Bertran J. J. Am. Chem. Soc. 2013; 135:8708–8719. [PubMed: 23679067]
111. Okimoto N, Tsukui T, Hata M, Hoshino T, Tsuda M. J. Am. Chem. Soc. 1999; 121:7349–7354.
112. Świderek K, Paneth P. Chem. Rev. 2013; 113:7851–7879. [PubMed: 23848598]
113. Gawlita E, Caldwell WS, O'Leary MH, Paneth P, Anderson VE. Biochemistry. 1995; 34:2577–2583. [PubMed: 7873538]
114. Rodriguez EJ, Angeles TS, Meek TD. Biochemistry. 1993; 32:12380–12385. [PubMed: 8241126]
115. Brás NF, Ramos MJ, Fernandes PA. Phys. Chem. Chem. Phys. 2012; 14:12605–12613. [PubMed: 22796659]
116. Ferrer S, Ruiz-Pernia JJ, Tuñón I, Moliner V, Garcia-Viloca M, Gonzalez-Lafont A, Lluch JM. J. Chem. Theory Comput. 2005; 1:750–761. [PubMed: 26641696]
117. Świderek K, Tuñón I, Martí S, Moliner V. ACS Catal. 2015; 5:1172–1185. [PubMed: 25705562]
118. Ferrer S, Silla E, Tuñón I, Oliva M, Moliner V, Williams IH. Chem. Commun. 2005; 1:5873–5875.

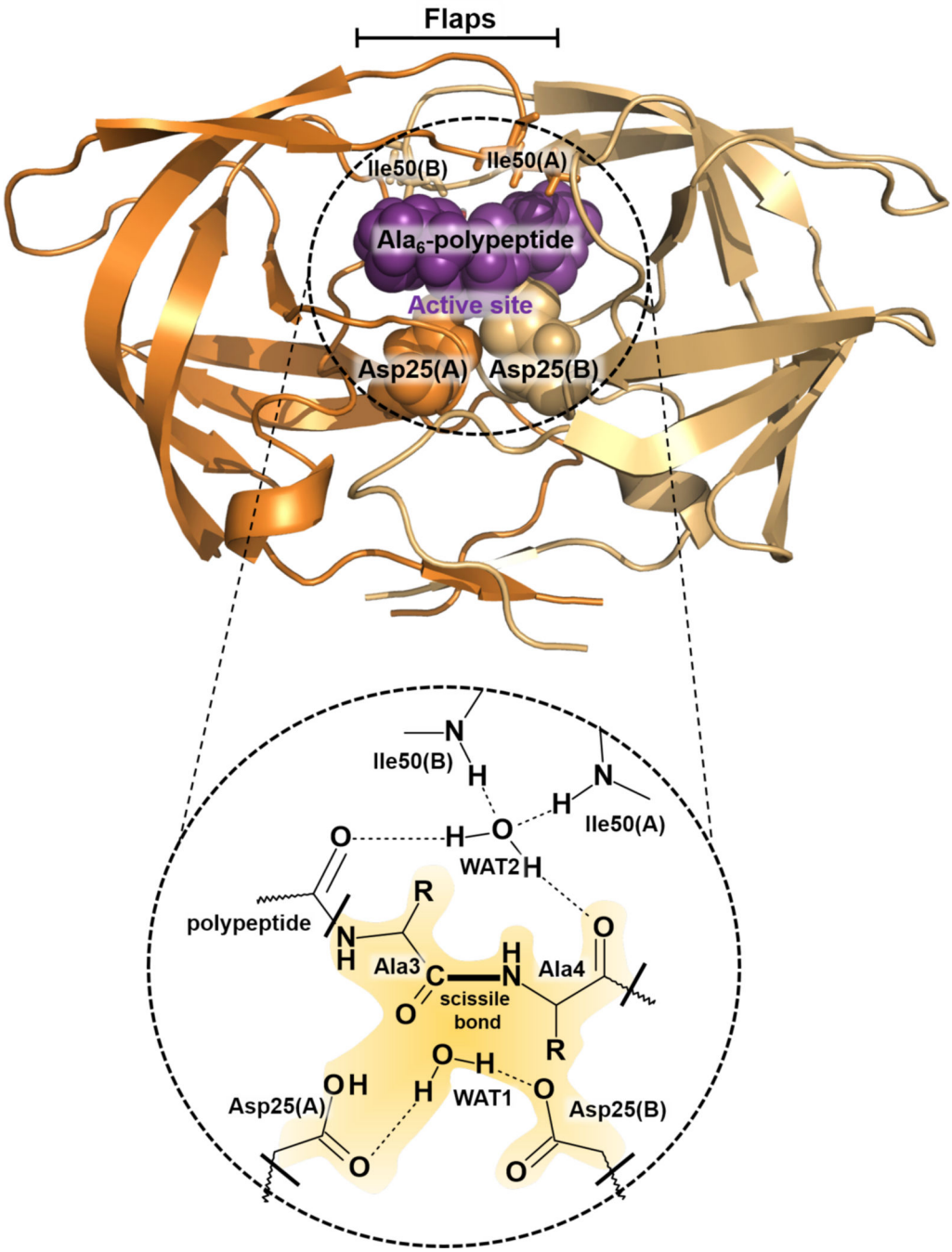


Figure 1.

Structure of HIV-1 PR and detail of the active site, with protonated Asp-25(A), and Ala_6 peptide as substrate. Yellow area in the bottom panel contains the part of the system treated at QM level of theory during all the QM/MM calculations. 4 link atoms are indicated as thick black lines.

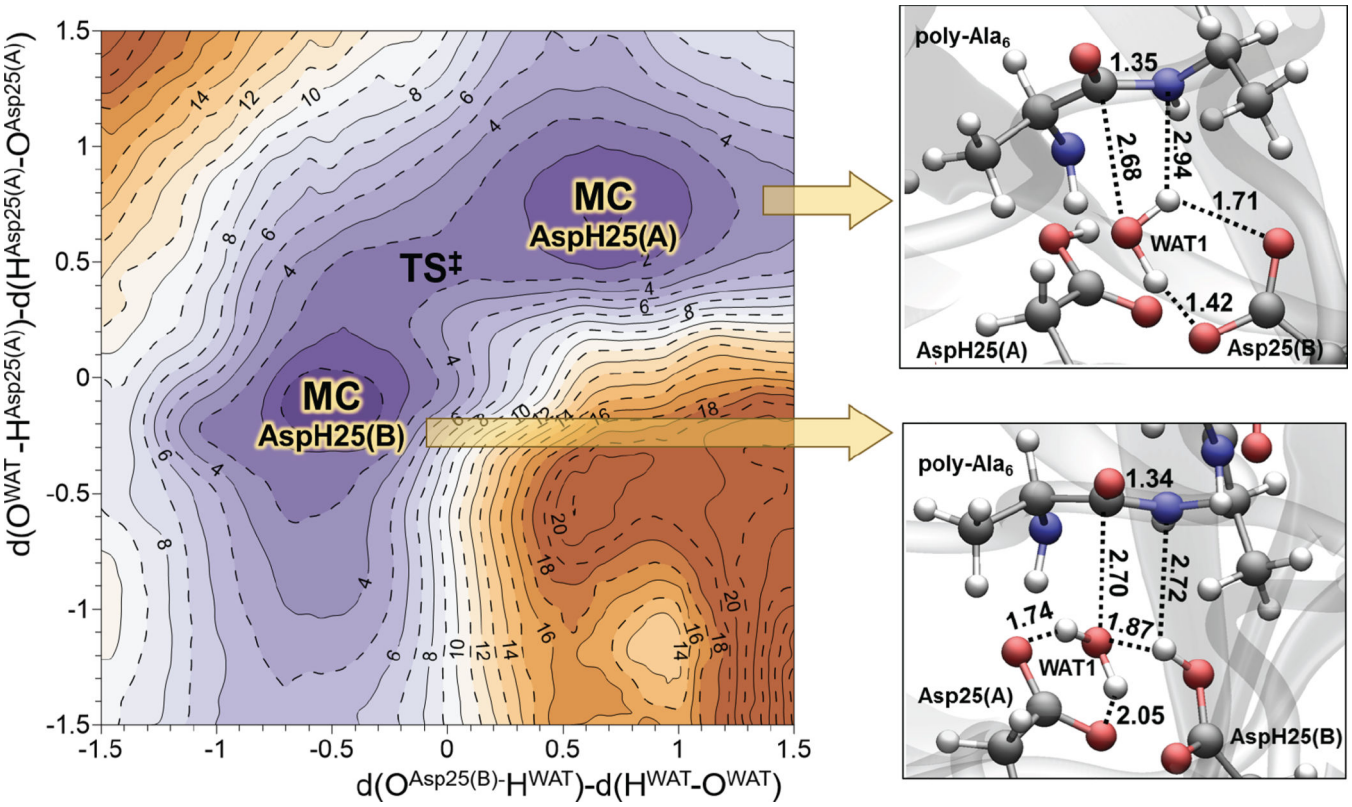


Figure 2.

AM1/MM 2D–PMF corrected at M06-2X:AM1/MM level for proton exchange process between Asp25(A) and Asp25(B). Detail of representative structures of the two MCs located at M06-2X/MM level are displayed in the right panels. Key distances are given in Å and values of isoenergetic lines in kcal·mol^{-1}.

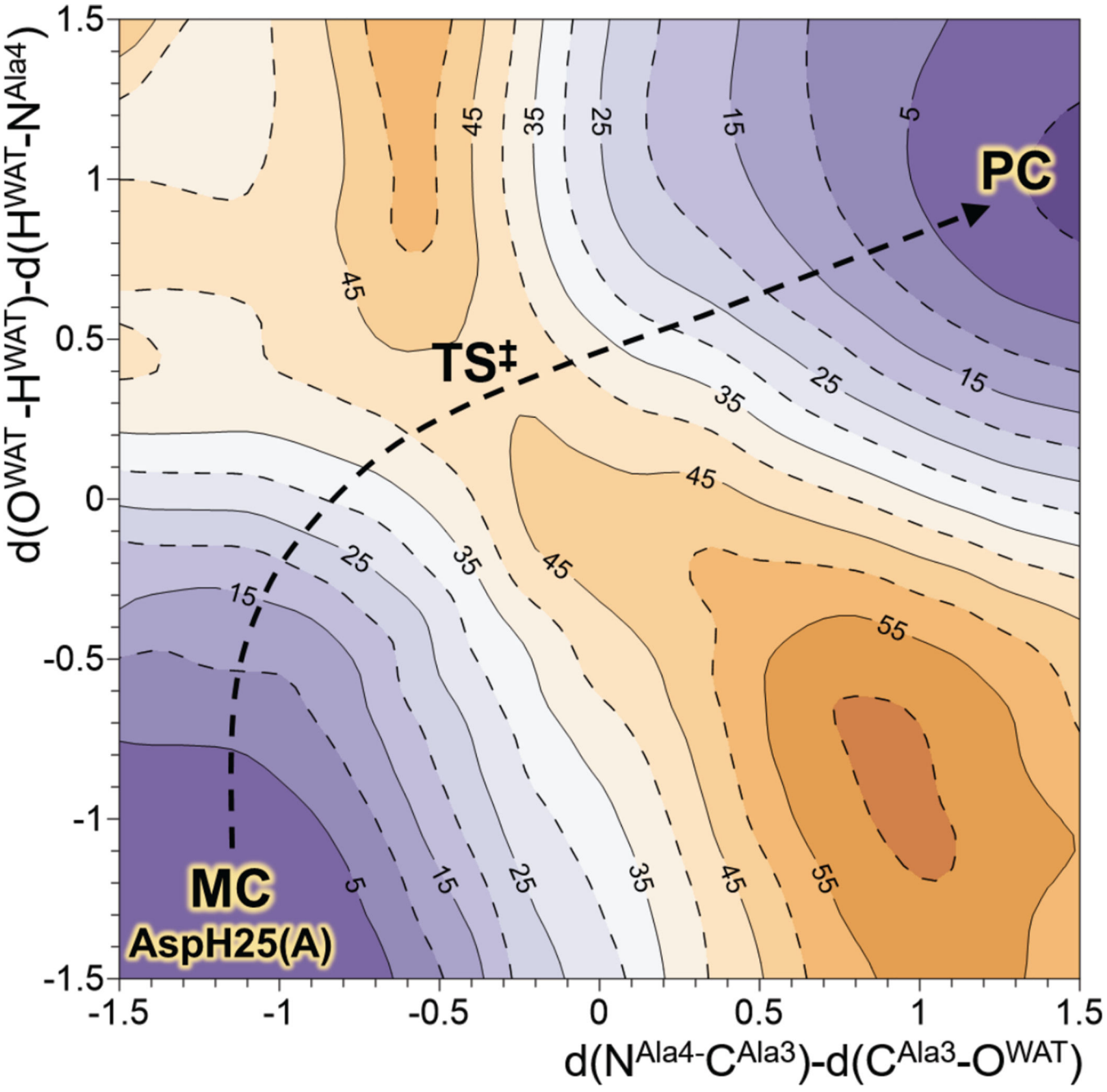


Figure 3.

AM1/MM 2D–PMF, corrected at M06-2X:AM1/MM level of theory, for mechanism [A] of proteolysis of Ala6 peptide catalyzed by HIV-1 PR starting from conformation of active site with protonated AspH25(A). Energy values are given in kcal·mol^{-1} and distances of the axis in Å.

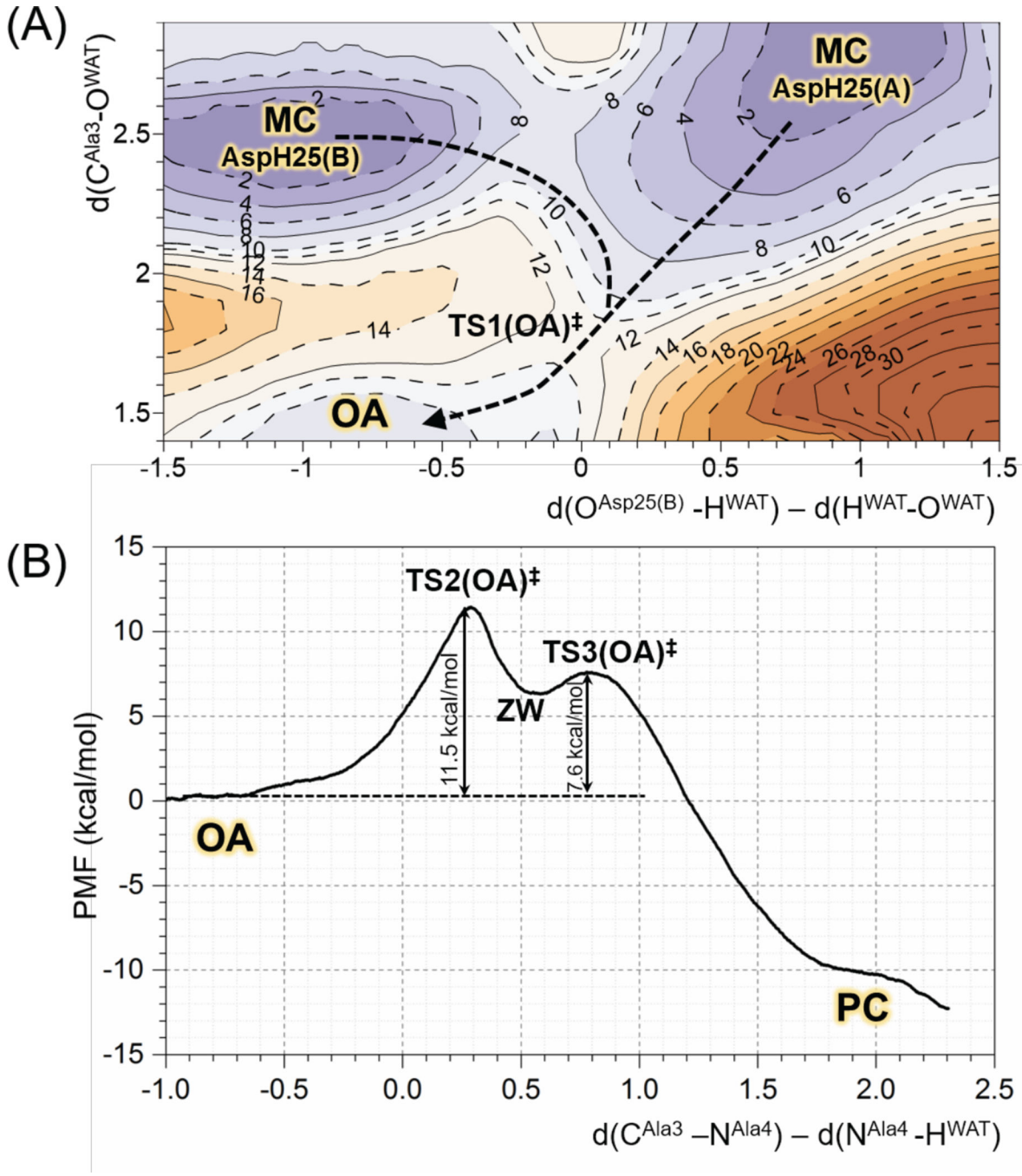


Figure 4.

(**A**) AM1/MM 2D–PMF, corrected at M06-2X:AM1/MM level, for the oxyanion intermediate formation step initiated at MC with protonated AspH25(A), or with protonated AspH25(B). (**B**) AM1/MM 1D–PMF, corrected at M06-2X/MM level, for the decomposition of the oxyanion intermediate into products. Values given in the isoenergetic lines of the 2D–PMFs are in kcal·mol^{-1} and distances of axis in Å.

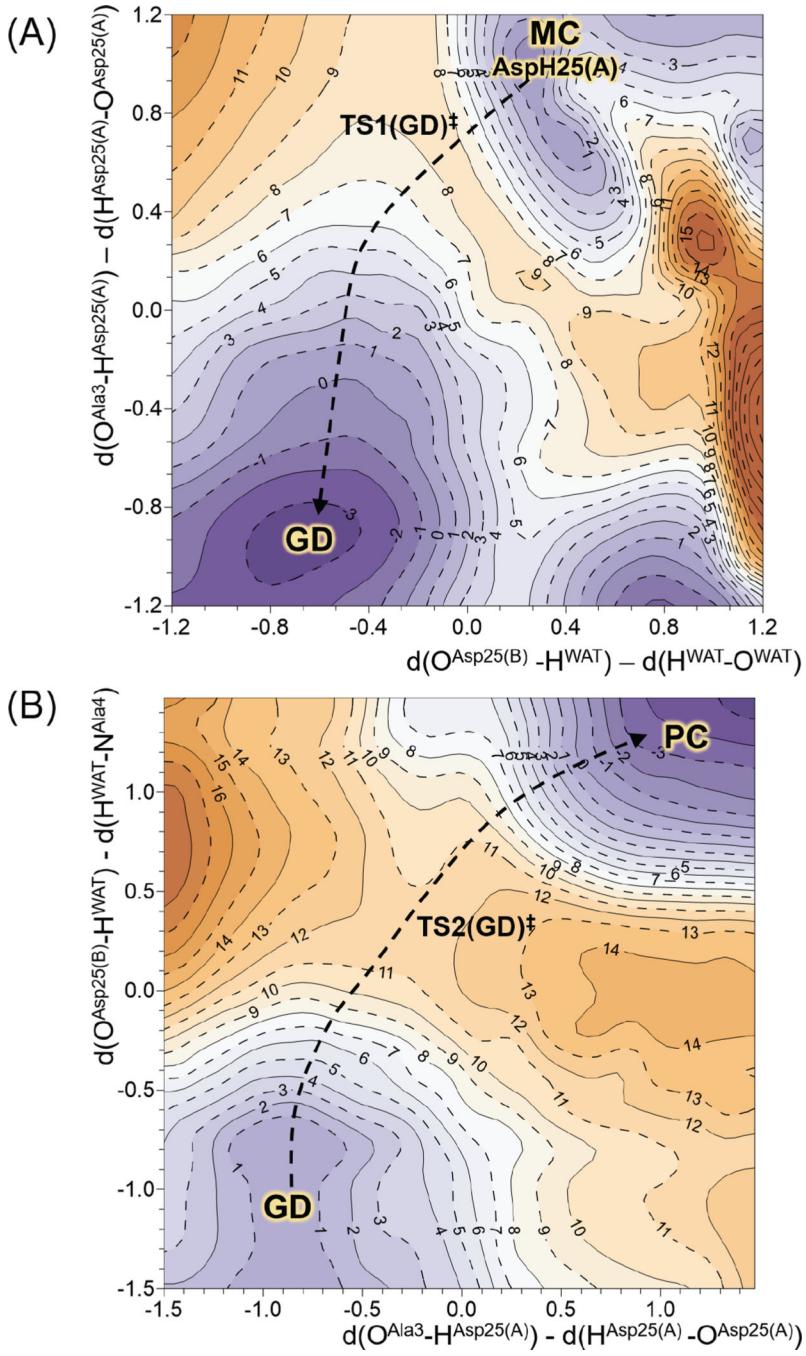


Figure 5.

AM1/MM 2D–PMFs, corrected at M06-2X:AM1/MM level, for the **(A)** transformation from MC to GD intermediate through a concerted process mechanism, and **(B)** decomposition of GD into products. Values given in the isoenergetic lines are in kcal·mol^{-1} and distances of the axis in Å.

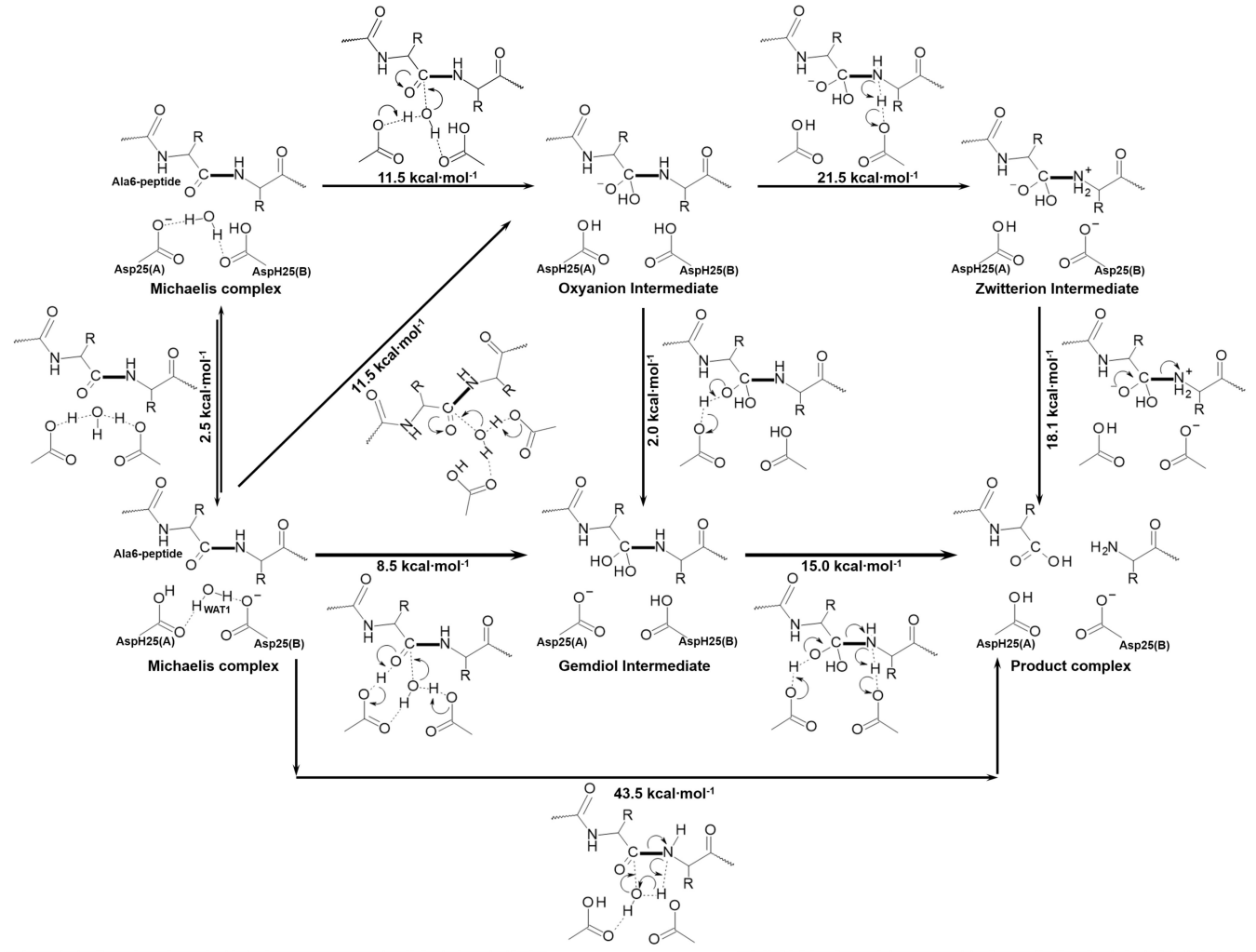


Figure 6.

Schematic representation of the overall studied mechanisms for HIV-1 PR-catalyzed reaction of peptide bond cleavage. Activation free energies, in relation to the initial reactant state, were derived from the AM1/MM PMFs corrected at M06-2X:AM1/MM level. All values are reported in kcal·mol^{-1}.

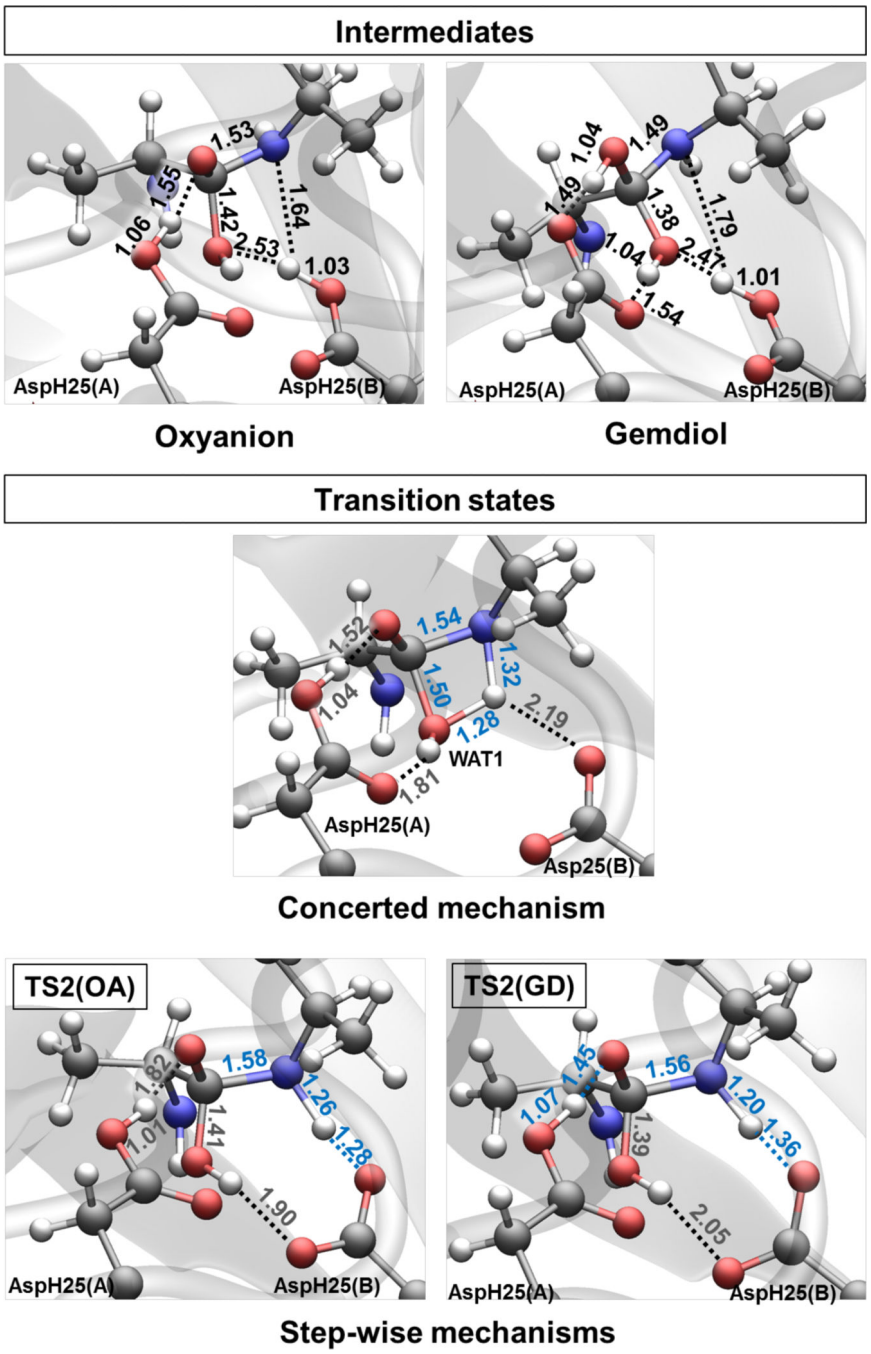


Figure 7.

Detail of representative snapshots of OA and the GD intermediates and the rate limiting TSs of the three explored mechanism. Key distances (in Å) involved in TSs are marked in blue.

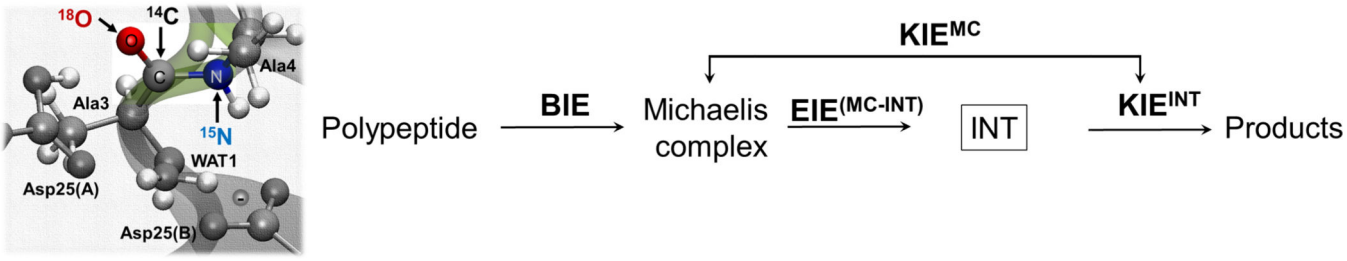


Figure 8.

Detail of the Ala6 peptide in the active site of the HIV-1 PR with indication of atoms isotopically substituted for the calculation of the isotope effect studies (left panel), and schematic representation of the definition of the different isotope effect along the reaction path (right panel).

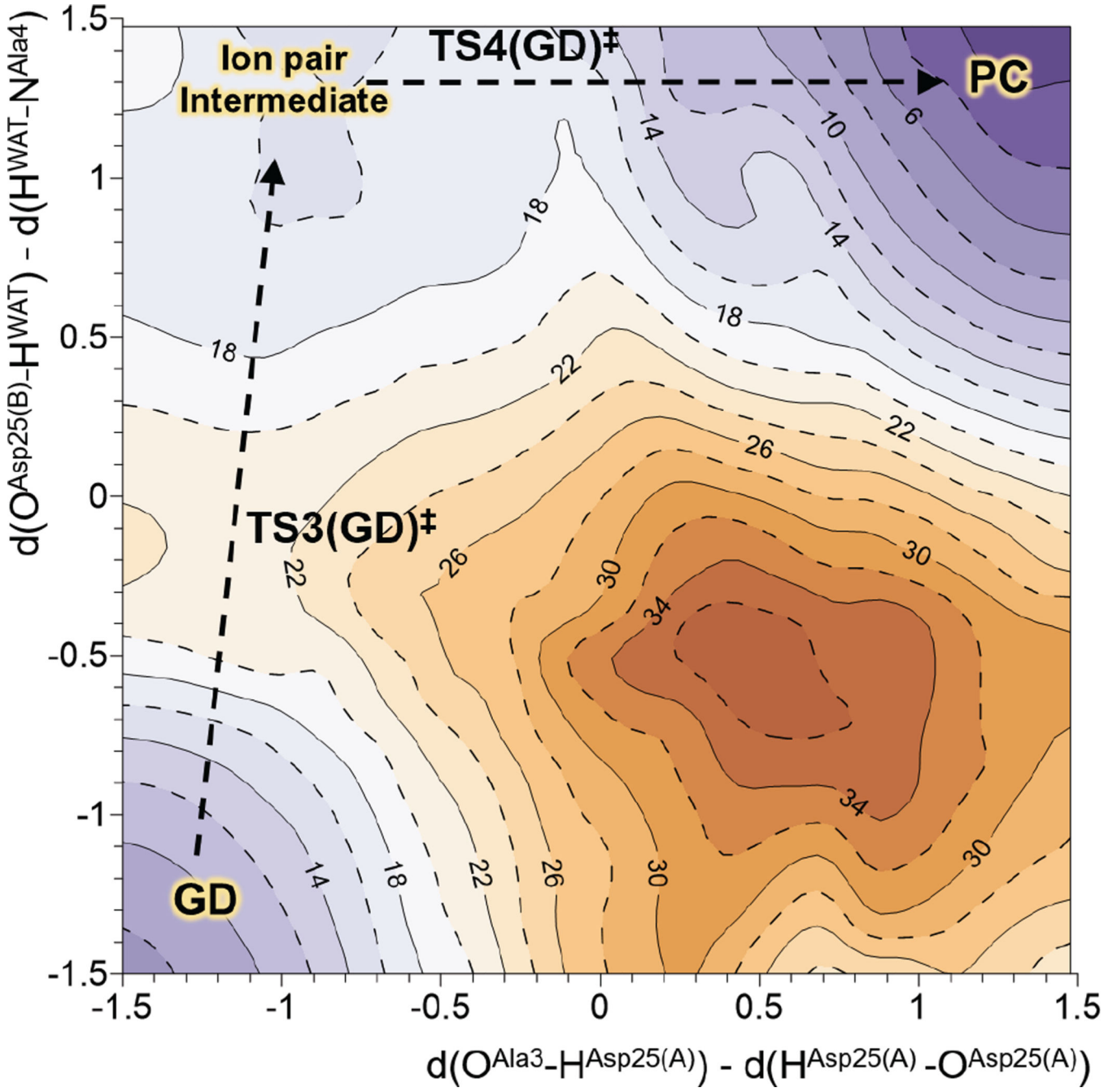


Figure 9.

2D–PMFs for the decomposition of the GD into products obtained at AM1/MM level. Energies of isoenergetic lines are reported in kcal·mol^{-1} and distances of the axis in Å.

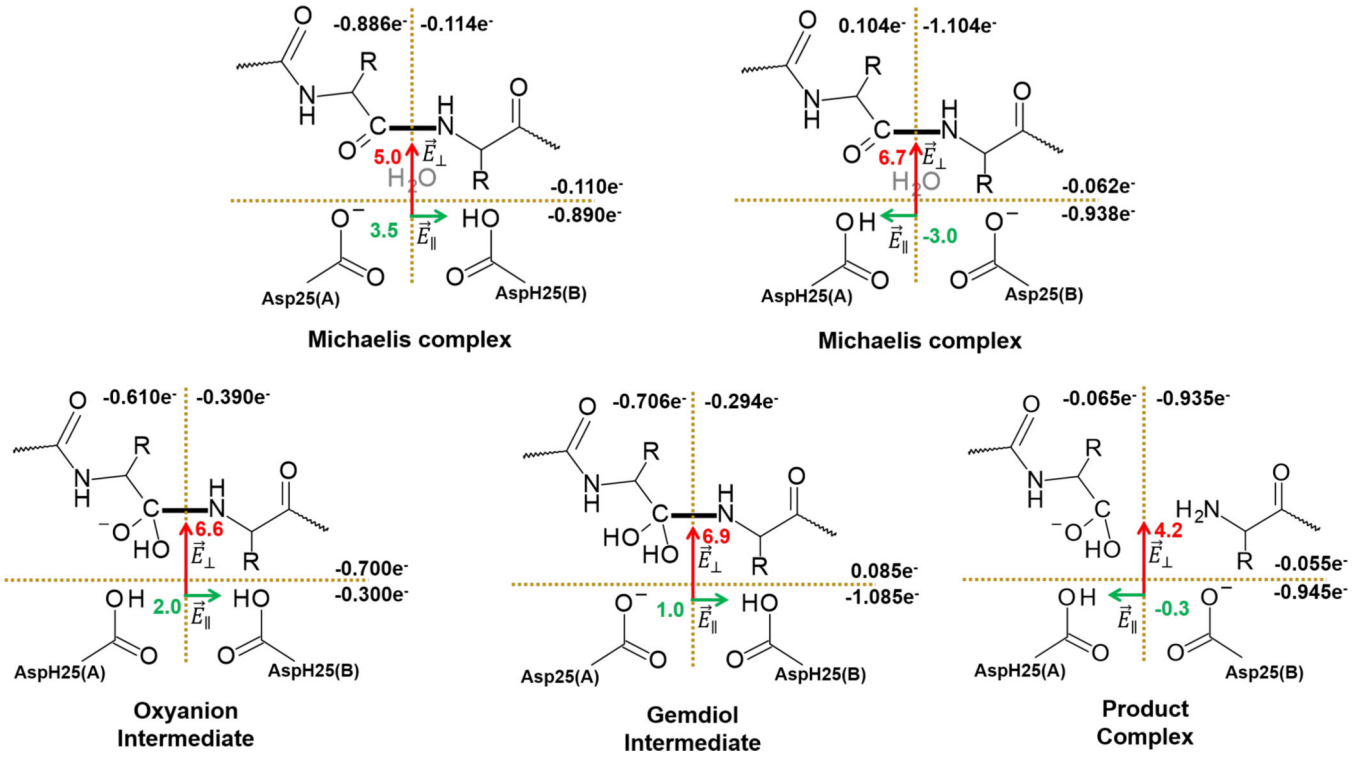


Figure 10.

Electronic charge (in a.u.) summed into the different fragments separated by the two imaginary planes represented by dashed lines. Projection of the electric field (in a.u. × 10^3) created by the protein and water molecules in the C-N peptide bond direction ($\vec{E}_{\parallel}$) and in a perpendicular direction ($\vec{E}_{\perp}$), computed in the center of the active site, are shown as green and red arrows, respectively. See text for details.

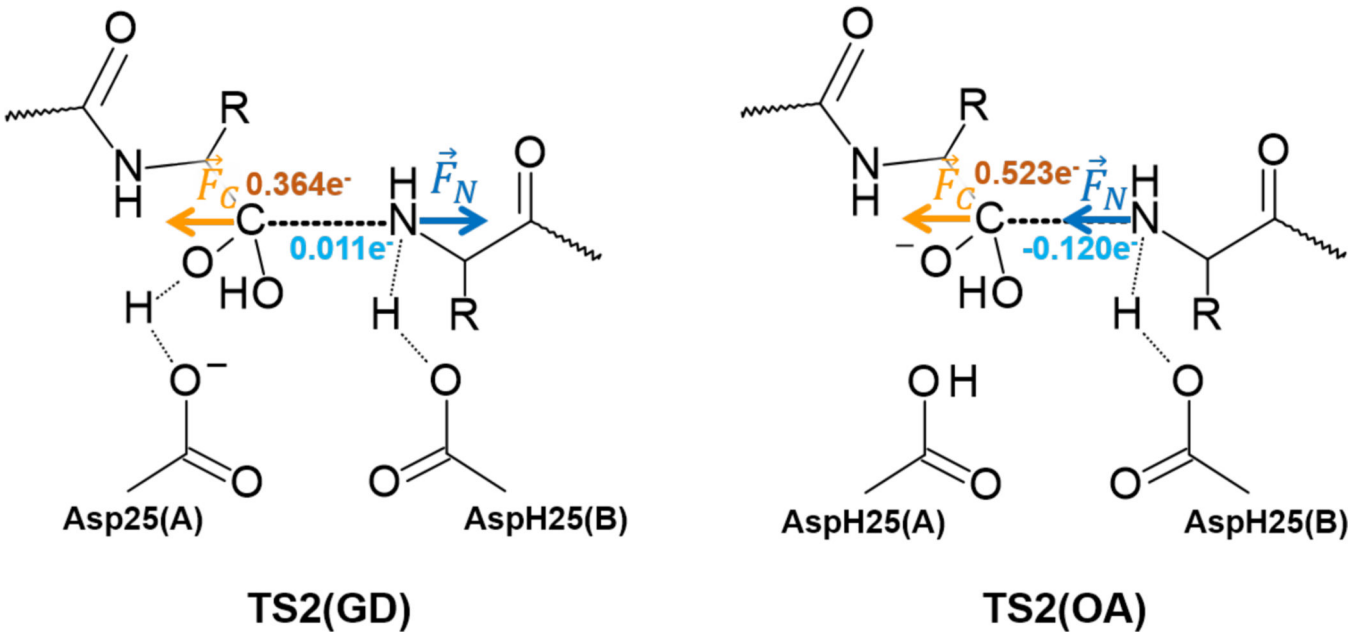
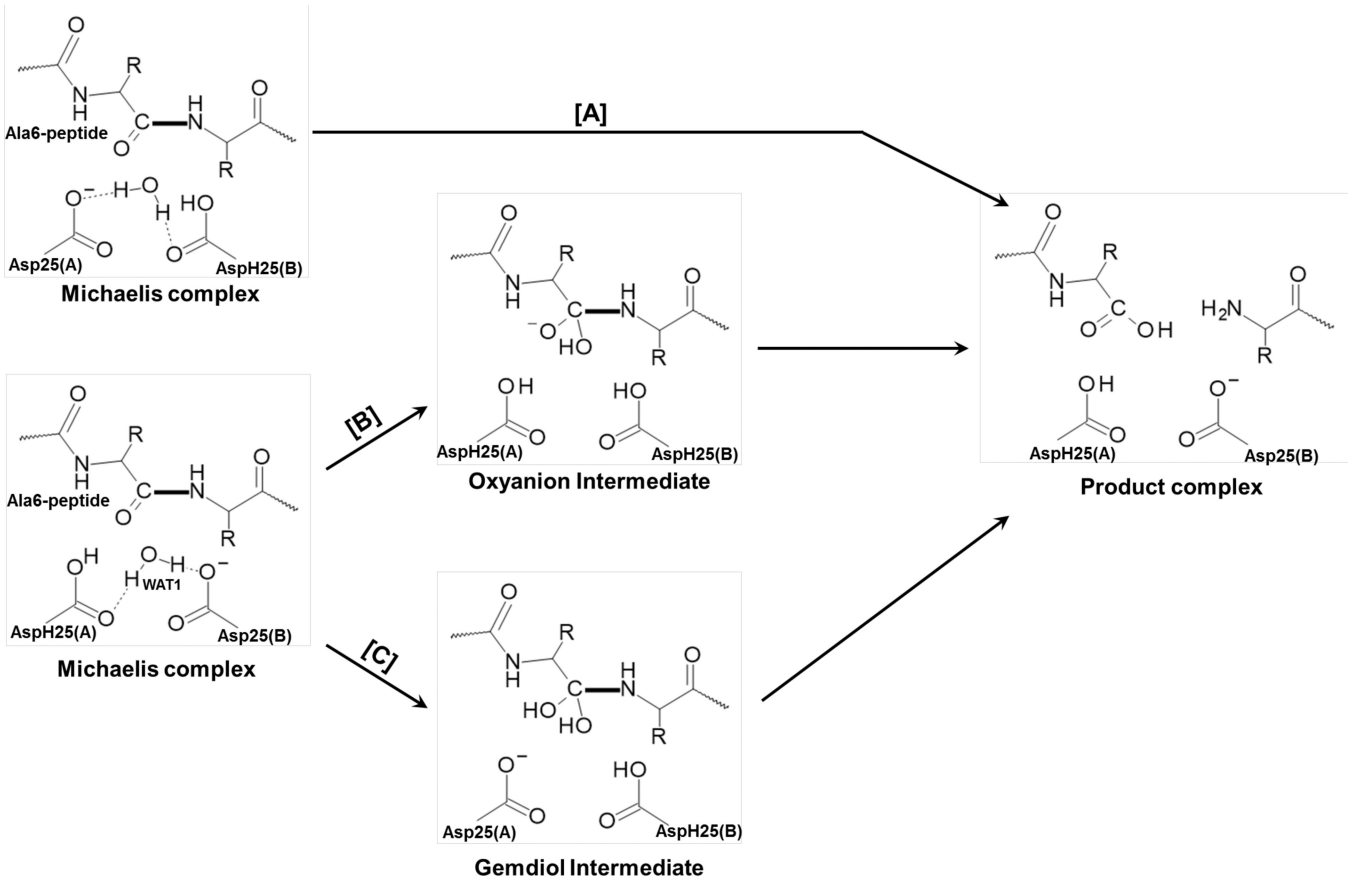


Figure 11.

Atomic charges computed using CHelpG method for structures of TS2(GD) and TS2(OA) localized at M06-2X/MM level (in a.u.) and schematic representation of the electrostatic forces created by the protein on the C and N atoms, projected on the C-N scissile peptide bond direction.



Scheme 1.

Proposed mechanisms for the hydrolysis of peptides catalyzed by HIV-1 PR.

Table 1

M06-2X/MM Binding, Equilibrium and Kinetic Isotope Effects (BIEs, EIEs and KIEs) for the concerted and step-wise mechanisms with oxyanion (OA) and gem-diol (GD) intermediate.

Atoms	Intermediate	BIE	$EIE^{(MC-INT)}$	$KIE^{(MC-TS)}$	$KIE^{(INT-TS)}$	KIE^{exp}
^{14}C-Ala3	concerted	0.993	-	1.040	-	1.029 ± 0.003[a]
	OA		1.034	1.039	1.005	
	GD		1.008	1.022	1.014	
^{15}N-Ala4	concerted	0.991	-	1.012	-	0.987 ± 0.004[a] 0.995 ± 0.002[b]
	OA		1.013	1.006	0.993	
	GD		1.005	0.995	1.000	
^{18}O-Ala3	concerted	1.017	-	0.997	-	0.993 ± 0.003[a]
	OA		1.009	1.007	0.998	
	GD		0.982	0.993	1.011	

[a] from ref. 33;

[b] from ref. 18.

Table 2

Enzyme kinetic isotope effect (enzyme KIE), recrossing coefficients (γ), equilibrium (ω^{eq}) and reactive frequencies (ω^{react}) computed from two different TSs for light and heavy variant of HIV-1 PR at AM1/MM level.

		γ	ω^{eq} (cm^{-1})	ω^{react} (cm^{-1})	enzyme KIE	
					theor.	exp.
TS3(GD)	Light	0.570 ± 0.018	1638.76	934.52	1.228 ± 0.055	1.19 ± 0.05[a]
	Heavy	0.464 ± 0.006		761.13		
TS4(GD)	Light	0.791 ± 0.014	1806.04	1427.78	1.042 ± 0.028	
	Heavy	0.759 ± 0.007		1370.64		

[a] experimental value from ref. 52.