Taylor & Francis
Taylor & Francis Group

Journal of Biomolecular Structure and Dynamics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tbsd20

Exploring the concerted mechanistic pathway for HIV-1 PR—substrate revealed by umbrella sampling simulation

Zainab K. Sanusi, Monsurat M. Lawal, Pancham Lal. Gupta, Thavendran Govender, Sooraj Baijnath, Tricia Naicker, Glenn E. M. Maguire, Bahareh Honarparvar, Adrian E. Roitberg & Hendrik G. Kruger

To cite this article: Zainab K. Sanusi, Monsurat M. Lawal, Pancham Lal. Gupta, Thavendran Govender, Sooraj Baijnath, Tricia Naicker, Glenn E. M. Maguire, Bahareh Honarparvar, Adrian E. Roitberg & Hendrik G. Kruger (2022) Exploring the concerted mechanistic pathway for HIV-1 PR—substrate revealed by umbrella sampling simulation, Journal of Biomolecular Structure and Dynamics, 40:4, 1736-1747, DOI: 10.1080/07391102.2020.1832578

To link to this article: https://doi.org/10.1080/07391102.2020.1832578

JOURNAL OF BIOMOLECULAR STRUCTURE AND DYNAMICS
2022, VOL. 40, NO. 4, 1736–1747
https://doi.org/10.1080/07391102.2020.1832578

Exploring the concerted mechanistic pathway for HIV-1 PR—substrate revealed by umbrella sampling simulation

Zainab K. Sanusi[a], Monsurat M. Lawal[a], Pancham Lal. Gupta[d], Thavendran Govender[b], Sooraj Baijnath[a], Tricia Naicker[a], Glenn E. M. Maguire[a,c], Bahareh Honarparvar[a], Adrian E. Roitberg[d] and Hendrik G. Kruger[a]

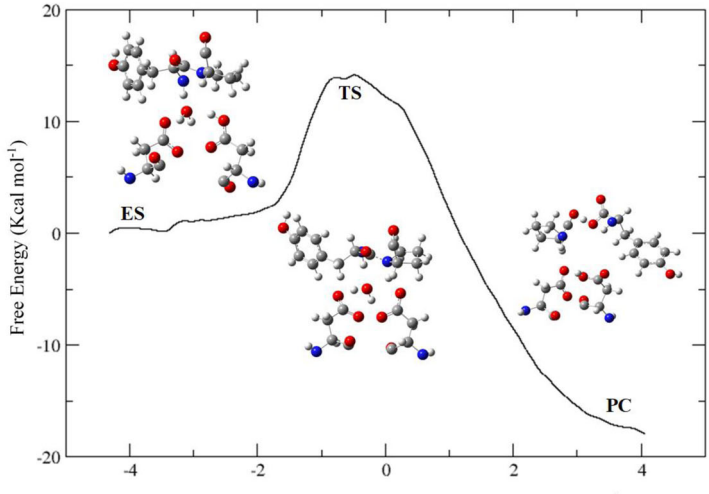
[a]Catalysis and Peptide Research Unit, School of Health Sciences, University of KwaZulu-Natal, Durban, South Africa; [b]Department of Chemistry, University of Zululand, KwaDlangezwa, South Africa; [c]School of Chemistry and Physics, University of KwaZulu-Natal, Durban, South Africa; [d]Department of Chemistry, University of Florida, Gainesville, Florida, USA

ABSTRACT

HIV-1 protease (HIV-1 PR) is an essential enzyme for the replication process of its virus, and therefore considered an important target for the development of drugs against the acquired immunodeficiency syndrome (AIDS). Our previous study shows that the catalytic mechanism of subtype B/C-SA HIV-1 PR follows a one-step concerted acyclic hydrolysis reaction process using a two-layered ONIOM B3LYP/6-31++G(d,p) method. This present work is aimed at exploring the proposed mechanism of the proteolysis catalyzed by HIV-1 PR and to ensure our proposed mechanism is not an artefact of a single theoretical technique. Hence, we present umbrella sampling method that is suitable for calculating potential mean force (PMF) for non-covalent ligand/substrate-enzyme association/dissociation interactions which provide thermodynamic details for molecular recognition. The free activation energy results were computed in terms of PMF analysis within the hybrid QM(DFTB)/MM approach. The theoretical findings suggest that the proposed mechanism corresponds in principle with experimental data. Given our observations, we suggest that the QM/MM MD method can be used as a reliable computational technique to rationalize lead compounds against specific targets such as the HIV-1 protease.

ARTICLE HISTORY
Received 20 June 2020
Accepted 30 September 2020

KEYWORDS
HIV-1 subtypes B and C-SA PR; HIV-1 PR cleavage mechanism; Umbrella sampling method; Steered molecular dynamics (SMD); Natural substrates; Activation free energy; Concerted transition state



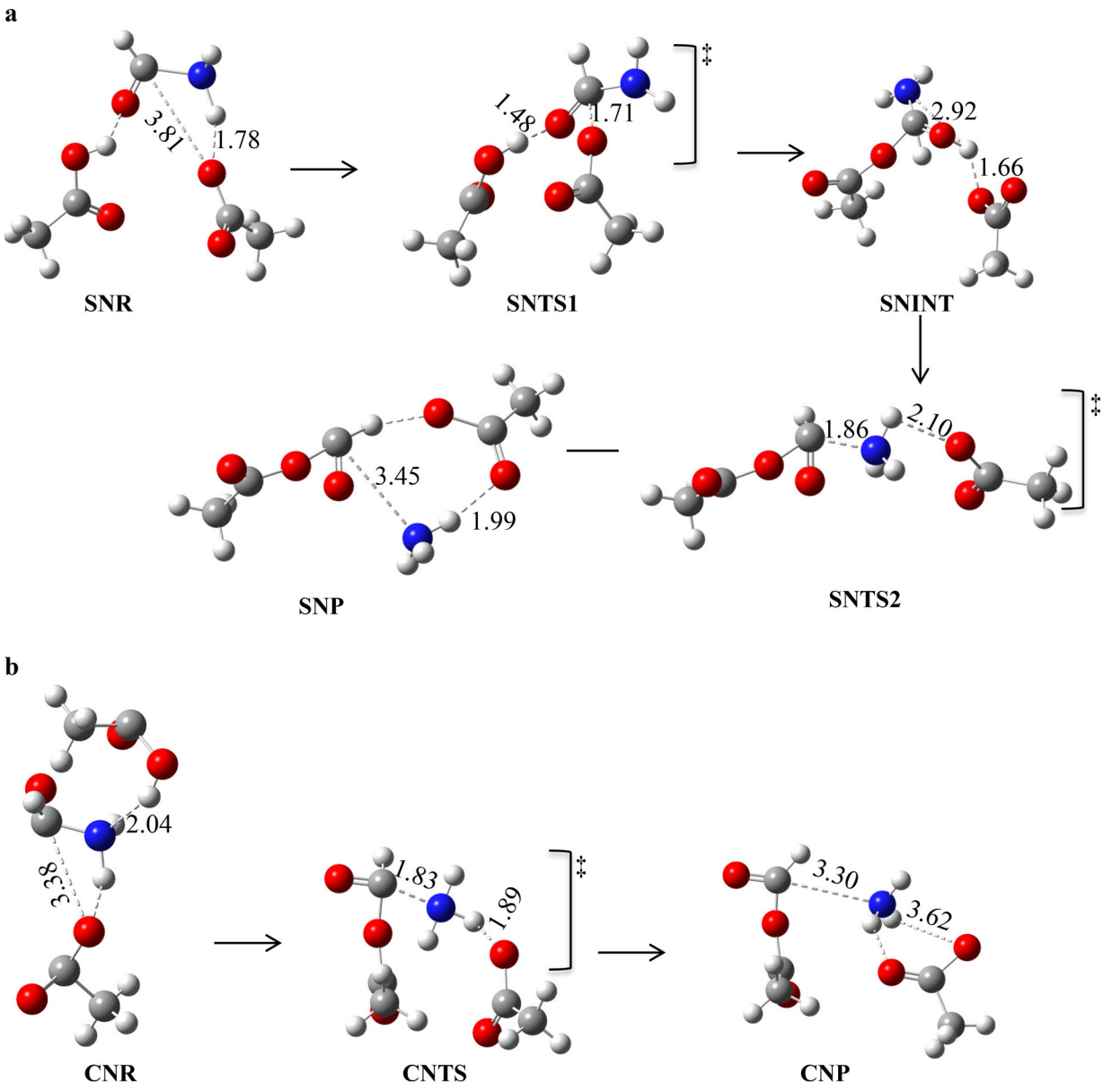
Free energy profile for the concerted acyclic mechanism

1. Introduction

In the quest for potential antiviral drugs for acquired immune deficiency syndrome (AIDS), the crystal structures of the active human immunodeficiency virus type 1 protease (HIV-1 PR) and its mutants have been investigated possibly more than any other protein structures (Rick, Erickson, & Burt, 1998; Wlodawer & Erickson, 1993). The virus encodes an

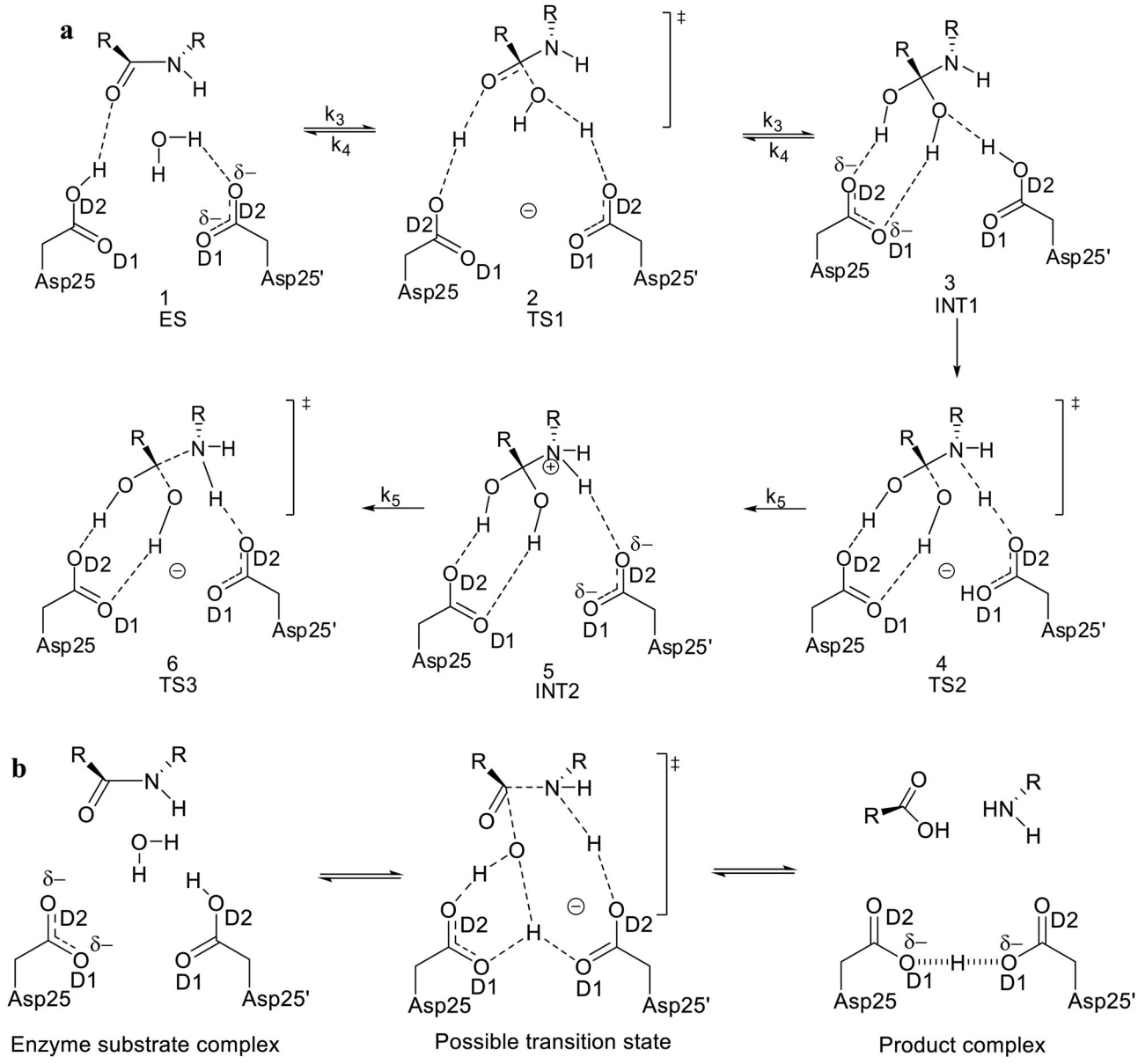
CONTACT H. G. Kruger kruger@ukzn.ac.za Catalysis and Peptide Research Unit, School of Health Sciences, University of KwaZulu-Natal, Durban, 4041, South Africa

Supplemental data for this article can be accessed online at https://doi.org/10.1080/07391102.2020.1832578



Scheme 1. Proposed (a) stepwise and (b) concerted nucleophilic reaction mechanism for HIV-1 PR—substrate system (redrawn from literature Park et al., 2000) involving a simplified peptide ($HCONH_2$) and active site. The legends are: SNR = stepwise nucleophilic reactant, SNTS1 = stepwise nucleophilic transition state 1, SNINT = stepwise nucleophilic intermediate, SNTS2 = stepwise nucleophilic transition state 2, SNP = stepwise nucleophilic product CNR=concerted nucleophilic reactant, CNTS = concerted nucleophilic transition state, CNP = concerted nucleophilic product. All distances are in Å.

aspartic protease (Asp-PR) known as the HIV PR, which is essential in replicating the infectious virus that causes AIDS (Kohl et al., 1988; Navia et al., 1989; Veronese et al., 1985). This enzyme catalyzes the cleavage of the viral Gag and Gag-Pol polypeptides necessary for the final maturation process for the HIV life cycle (Kohl et al., 1988; Navia et al., 1989; Veronese et al., 1985). The inhibition of this process results in the production of immature, non-infectious viral particles (Chatfield, Eurenius, & Brooks, 1998; Roberts et al., 1990; Weber, Cavanaugh, & Harrison, 1998; Wlodawer & Erickson, 1993). Early experiments (Coates et al., 2006; Hyland et al., 1991; Northrop, 2001; Rodriguez, Angeles, & Meek, 1993) and theoretical (Bjelic & Åqvist, 2006; Park et al., 2000; Piana et al., 2004; Piana, Carloni, & Parrinello, 2002; A. M. Silva et al., 1996; Trylska et al., 2004) studies of the HIV-1 PR have found that the transition state (TS) analogs of this enzyme-catalyzed reaction act as suitable inhibitors. Nevertheless, over the past decades, there is no consensus on the mechanistic pathway for the HIV-1 PR cleavage reaction. The structure and protonation of the catalytic aspartate dyad are ultimately linked to the HIV-1 PR mechanism's uncertainty and still a source of many scientific discussions (Krzemińska et al., 2016; Lawal et al., 2020). However, the reaction mechanism has been classified into two major groups; the nucleophilic (Hsu et al., 1977; Stanton, Hartsough, & Merz, 1993) and the general acid-general base (GA-GB) (Antonov et al., 1978; Jaskólski et al., 1991; Lapatto et al., 1989; Park et al., 2000; Trylska et al., 2004) process, with the latter being the most upheld mechanism for aspartic proteases, especially for HIV-1 PR (Scheme 2) (Antonov et al., 1978; Fruton, 1976; Hyland et al., 1991).

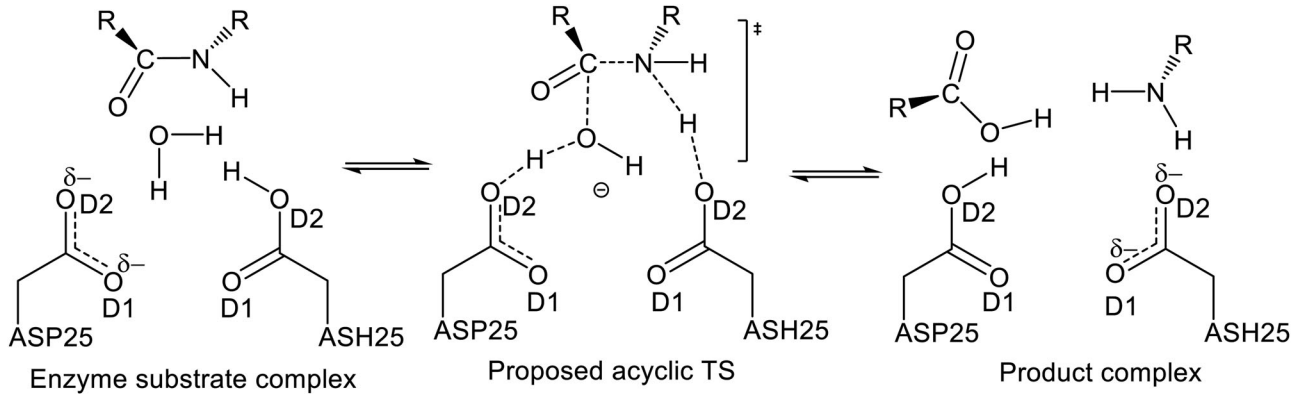


Scheme 2. Proposed (a) stepwise and (b) concerted general acid-base reaction mechanism for HIV-1 PR—substrate system, redrawn from literature(Lawal et al., 2020). Structures along the pathway of the stepwise are indicated as follows: (1) enzyme—substrate complex; (2) water attack TS; (3) tetrahedral gem-diol intermediate; (4) scissile N-protonation TS; (5) protonated amide intermediate; and (6) cleavage of scissile C-N bond TS leading to separated product complex.

These proposed mechanisms have been elaborated in our reported work (Lawal et al., 2020; Lawal et al., 2019; Z. K. Sanusi et al., 2019). In Summary, for the nucleophilic mechanism, the protonated catalytic Asp act as the nucleophile (Weber et al., 1998), which attacks the carbonyl atom of the scissile peptide (C—N) bond. This reaction yields either an anhydride intermediate (stepwise nucleophile reaction, Scheme 1a) or a transition state that leads to the product (concerted nucleophilic reaction, Scheme 1b) (Boross et al., 1999; Hsu et al., 1977). While, the general acid-base mechanism (Scheme 2) includes the catalytic water acting as the nucleophile, which attacks the carbonyl atom of the scissile peptide (C—N) at the active site (Krzemińska et al., 2016; Park et al., 2000). The reaction mechanism proceeds with the nitrogen atom of the scissile unit accepting hydrogen from the protonated Asp. At the same time, the catalytic water is ionized by donating its proton to the unprotonated Asp, this phase of the mechanism can either be a stepwise (Scheme 2a) or a concerted process (Scheme 2b) (Lawal et al., 2020; Trylska, Bała, Geller, & Grochowski, 2002; Trylska et al., 2004).

In 1991, Jaskólski et al (Jaskólski et al., 1991) put forward an experimentally supported perspective on the hydrolytic mechanism of HIV-1 PR and a substrate. They proposed a concerted catalytic model in which the nucleophilic water and the electrophilic aspartate (an acidic proton) attack the scissile bond synchronously. The reaction started with the acidic proton at the OD2 aspartate attacking the N atom (in proximal) of the approaching scissile unit. After the reaction, the catalytic aspartates are still bound by the acidic proton, which now resides between the inner OD1 atoms (Jaskólski et al., 1991). Based on our understanding, we have summarized this experimental process in Scheme 2b.

Regarding the concerted cleavage mechanism of HIV-1 PR, two theoretical investigations have reported transition



Scheme 3. Proposed concerted acyclic general acid-base reaction mechanism for HIV-1 PR—substrate system. The main difference between Scheme 2b and our proposed TS is that the acidic Asp became basic while the de-protonated Asp becomes acidic.

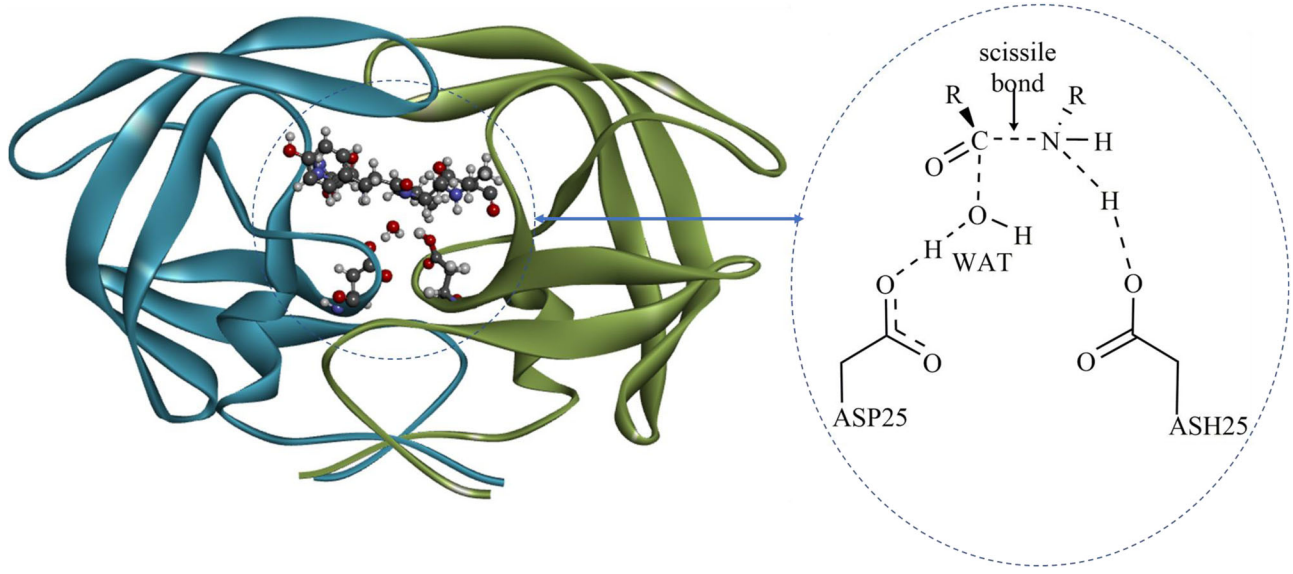


Figure 1. Graphical structure of the HIV-1 PR, showing the direction of the hydrogen at the catalytic active site, which includes the aspartates and peptide substrate. The circled-out part is the area of the system treated at the QM (DFTB) level of theory, while the rest of the system is treated at the MM level of theory. Free-energy plots, 3D-structures of ES, TSs, and PC are provided in the supporting information (SI), and these are listed in the table of content in the SI.

states with activation free energies of 30 (Park et al., 2000) and 43.5 (Krzemińska et al., 2016) kcal mol^{-1}. However, these are higher values than experimental data found in the literature (Boross et al., 1999; Hyland et al., 1991; Kipp, Silva, & Schramm, 2011; Rodriguez et al., 1993; Tözsér et al., 1996). The feasibility of the concerted general acid-base model involving a 6-membered ring TS with either one or two water molecule(s), and an acyclic TS with one water molecule, was reported in our previous ONIOM based study (Lawal et al., 2019). The concerted acyclic reaction mechanism (Scheme 3) of HIV-1 PR was observed to give an optimal basis for substrate cleavage (Lawal et al., 2019). Likewise, it was revealed that a more extended peptide sequence was necessary for favorable recognition (Z. K. Sanusi et al., 2019). The computed activation free energy values from these studies (Z. K. Sanusi et al., 2020; Lawal et al., 2019; Z. K. Sanusi et al., 2019) were found to be within the experimental range.

Studies of the HIV-1 PR structures show conformational variances in the apoenzyme and a variety of substrate/ligand-bound enzymes (Rick et al., 1998; Schulz, 1991). Hence, the energy difference between the different conformational states is an important factor. Umbrella sampling (G. Torrie & Valleau, 1977; G. M. Torrie & Valleau, 1974) has been applied to study reaction rates from free-energy barriers by establishing an appropriate reaction pathway (Rick et al., 1998). Once the reaction path is known the potential mean force (PMF) (Hazel & Gumbart, 2017) can be determined through a series of windows and combined with either umbrella integration (Kästner & Thiel, 2005) or the weighted histogram analysis method (WHAM) (S. Kumar et al., 1992; Souaille & Roux, 2001)

The application of the hybrid quantum mechanics/molecular mechanics (QM/MM) combined with molecular dynamics (MD) simulations have been used to study the catalytic mechanism of enzyme—substrate, and enzyme—intermediate complexes of the HIV-1 PR (Krzemińska et al., 2016) and other related biological systems (Di Martino et al., 2014; Di Palma, Bottaro, & Bussi, 2015; J. R. A. Silva et al., 2015; J. R. A. Silva, Roitberg, & Alves, 2014; Tsukamoto et al., 2018; Yildirim et al., 2013). This method provides improved

sampling in atomistic simulations by biasing potentials (one or more dimensional) along a reaction coordinate to drive a system from one thermodynamic state to another (reactant to transition state or product) (Kästner, 2011; Rick et al., 1998)

Similarly, the protonation state of the catalytic aspartates (25/25′) has been extensively studied (Meek et al., 1989; Northrop, 2001; Piana et al., 2001; Piana & Carloni, 2000; Smith et al., 1996). Ten possibilities (mono-protonation at different O2 of the Asp, di-protonation, un-protonated, low barrier hydrogen) have been considered in the past (Lawal et al., 2020). Nevertheless, the choice of protonation herein was influenced by Meek *et al* (Hyland et al., 1991; McGee, Edwards, & Roitberg, 2014) who demonstrated that the *pKa* of one of the catalytic Asp increases to 5.2, while the other is 3.1 when bound to a substrate/inhibitor. This *pKa* value suggests that the catalytic aspartate group at the active site is mono-protonated (Adachi et al., 2009; Wang et al., 1996)

The catalytic process in this study follows a concerted general acid-base mechanism with a mono-protonation of the catalytic aspartate at the active site (Z. K. Sanusi et al., 2020; Z. Sanusi et al., 2017; Z. K. Sanusi et al., 2018; Z. K. Sanusi et al., 2019). We applied a hybrid QM/MM MD umbrella sampling method to study the details of the concerted acyclic catalytic mechanism of the HIV-1 PR. In the QM/MM methodology, a part of the system (active site and the natural substrate) is described by quantum mechanics, while classical mechanics describes the remaining part (the whole system). The SCC-DFTB (Elstner et al., 1998) method is used to describe the QM region of the system, and this method has not only demonstrated an accurate level of theory for describing the energetics of chemical reactions (Woodcock, Hodoscek, & Brooks, 2007). It has also been established for large biological systems yielding minimum energy path results in good agreement with a high level of theory such as MP2 (Barnett & Naidoo, 2010). To ensure that the calculation is not an artefact of a single enzyme—substrate system, two PRs, HIV-1 subtype B complexed with a natural substrate (NC-p1), is set as the standard/control. Also, subtype C-South Africa (C-SA) HIV-1 PR complexed with three natural substrates Matrix-capsid (MA-CA), nucleocapsid-p1 (NC-p1), and protease-reverse transcriptase (PR-RT) were all considered.

Herein, a concerted acyclic TS reaction mechanism of the HIV-1 PR—substrate involving one water molecule with a more extended natural substrate sequence (P5-P5′) from our previous ONIOM report was employed (Z. K. Sanusi et al., 2020; Lawal et al., 2019; Z. K. Sanusi et al., 2019). Both the catalytic water molecule and the Asp (25/25′) dyad are involved in the reaction mechanism, the protonated Asp becomes basic by losing its proton, and the unprotonated Asp gains a proton and becomes acidic (Scheme 3). Since the active enzyme and its mutants recognize and cleave the same polypeptide precursors, this study aimed to explore the proposed mechanism of proteolysis catalyzed by HIV-1 PR *via* QM/MM umbrella sampling technique with an adequate reaction coordinate. The method involves simulations with restraining potentials based on a known estimation of the free energy profile derived mainly from experimental data. The present work intends to provide more information on the mechanistic landscape of the HIV-1 PR system. Figure 1 demonstrates the HIV-1 PR structure, showing graphical details of the catalytic Asp (25/25′) at the active site and a natural substrate. The recognition sequences cleaved by the HIV-1 PR is provided in Table S1 of the supporting information (SI).

2. Computational methods

2.1. Structure preparation of the HIV-1 natural substrate complexes

The initial X-ray coordinates of the subtypes B and C-SA HIV-1 PR were obtained from the Protein Data Bank (PDB code: 2P3B resolution of 1.8 Å (Kempf et al., 1990) and 3U71 at 2.7 Å resolution (Naicker et al., 2013) respectively), which contain a bound inhibitor fragment. These bound-inhibitors were exchanged for three of the HIV-1 natural substrates (MA-CA, PR-RT, and NC-p1) also derived from PDB bank code: 1KJ4, (Prabu-Jeyabalan et al., 2002) 3OU3, (Z. Liu et al., 2011) and 1TSU (Prabu-Jeyabalan et al., 2004) (Table S1) respectively. The contribution of water is an essential aspect for the HIV-1 PR proteolysis; therefore, a catalytic water molecule close to the active binding site of HIV-1 PR was taken from PDB code:1LV1 (M. Kumar et al., 2002). Our model complexes used in this study were obtained by overlaying the crystal structures of both subtypes B, and C-SA HIV-1 PR with the natural substrates and catalytic water using the same superimposition method explained in our previous papers (Z. Sanusi et al., 2017; Z. K. Sanusi et al., 2018). The root mean square (RMS) values for the overlaid substrates after superimposition ranged between 1.2-1.5 Å, and was measured using PyMOL (DeLano, 2002), which helps in evaluating how well the substrates are overlaid in the enzyme—substrate complex systems. An optimal superimposition is considered satisfactory if the RMS is less than 2 Å (Cole et al., 2005; Gohlke, Hendlich, & Klebe, 2000; Kontoyianni et al., 2004). Hence, the complexed natural substrates are well superimposed and maintain the same position at the active binding sites of the subtype B/C-SA HIV-1 PR as the substituted inhibitors. The structures of all the enzyme—substrate systems were refined by removing all non-related enzyme atoms from the complexes except for the catalytic water at the active site. These enzyme—substrate (ES) models serve as a starting structure for all calculations carried out in this study.

As earlier mentioned, a significant area of debate in modeling the HIV PR is the protonation state of the catalytic Asp (25/25′) at the active binding site, and this has been discussed at length in the literature (Garrec, Sautet, & Fleurat-Lessard, 2011; Kontoyianni et al., 2004; Sussman et al., 2013). Herein, the protonation states for catalytic aspartates and other residues were accurately assigned at pH 7 using the empirical propKa (Li, Robertson, & Jensen, 2005) server by recalculating the standard *pKa* values of titratable amino acids. As a result, a mono-protonation model of the catalytic aspartate at the binding site was adopted, and also for the

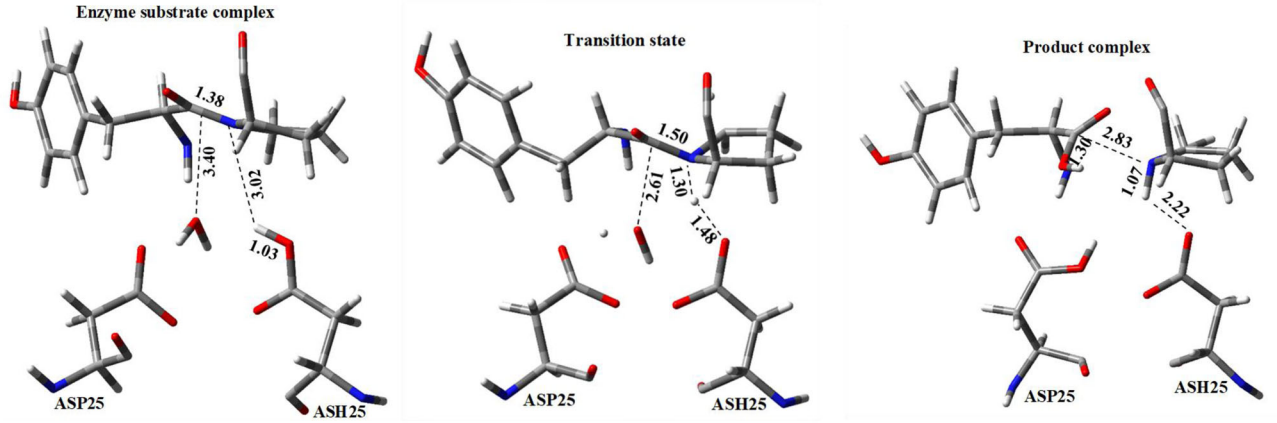


Figure 2. Proposed concerted acyclic general acid-base reaction mechanism for HIV-1 PR—substrate system. Showing the distances in the chosen reaction coordinates, from the ES to TS and product complex for one of the enzymes—substrate complex. Values are reported in Å.

majority of related studies (Jaskólski et al., 1991; Krzemińska et al., 2016; H. Liu, Müller-Plathe, & van Gunsteren, 1996; Park et al., 2000; Piana & Carloni, 2000). The charge of the natural substrate was determined using the restrained electrostatic potential (RESP) method by applying the HF/6-31G* level of theory/basis set using the Gaussian16 package (Frisch et al. 2016).

To prepare the enzyme—substrate complexes for MD simulations and the free energy calculations, the proton (H^+) of the protonated Asp in the ES system complexes is positioned midway to facilitate its donation to the nitrogen atom of the scissile peptide bond. This position has been measured to be energetically stable *via* MD simulations (Fakhar et al., 2017) Subsequently, missing atoms were added to the system complexes using the Leap module as implemented in the AMBER 18 molecular dynamics package (Case et al., 2018). All the enzyme—substrate systems considered were solvated in an octahedral box of TIP3P (Jorgensen et al., 1983) water molecules, extending 12 Å outside the protein on all sides. Chlorine ions (Cl^-) were positioned around the protein to ensure the system is neutralized; the protein and natural substrate were described using the AMBER force field 99SB (Hornak et al., 2006).

The ES system complexes were initially energy-minimized by performing minimization at 5000 steps (of each steepest descent and conjugate gradient methods) with a 20 kcal $mol^{-1}.Å^{-2}$ restraint applied to only the backbone of the protein and natural substrate peptide residues. The system is then gradually heated from 0 to 300 K using the Langevin thermostat (Pastor, Brooks, & Szabo, 1988) at a constant volume, over 1000 ps, with a restraint of 10 kcal $mol^{-1}.Å^{-2}$ on the backbone atoms. Afterward, the system was equilibrated at constant pressure for 1000 ps at 300 K with 5 kcal $mol^{-1}.Å^{-2}$ harmonic restraints and subsequently relaxed with weak restraints of 2 kcal $mol^{-1}.Å^{-2}$ on the backbone atoms. All minimizations, heating, and equilibration simulation steps utilized a non-bonded cutoff of 8 Å. The long-range Coulomb forces were calculated using the particle mesh Ewald (PME) method, and the time step was set to 2 fs, the SHAKE algorithm (Ryckaert, Ciccotti, & Berendsen, 1977) was used to constrain the bond length between heavy atom and hydrogen atoms to its equilibrium value during the MD simulations.

2.2. Umbrella sampling and free energy calculations

For the hybrid QM/MM MD calculations, the catalytic water and Asp 25/25′, with the peptide amino acid residues at the scissile bond (P1-P1′) were selected at the QM region (a total of 62 atoms) for the ES complexes, only MA-CA has 63 atoms. The QM region is described using the self-consistent charge density-functional tight-binding (SCC-DFTB) Hamiltonian (Seabra et al., 2007) as implemented in AMBER18 (Case et al., 2018). To satisfy the valence of the QM/MM boundary atoms, link atoms were placed in the substrate peptide and the catalytic aspartates' residues (Figure 1). The rest of the system (protein and water molecules), as mentioned earlier, is described using the AMBER ff99SB (Hornak et al., 2006) and TIP3P (Jorgensen et al., 1983) force fields, respectively.

The free energy profile for the reaction mechanism catalyzed by HIV-1 PR (subtypes B/C-SA) was first evaluated with steered molecular dynamics (SMD) (Jarzynski, 1997a, 1997b) to sample the reaction path and drive the reaction process to completion. Afterward, the QM/MM MD umbrella sampling was performed. The SMD method is very similar to the umbrella sampling in which the restraint center is time-dependent, by integrating the force over distance or time, a generalized simulation can be performed (Case et al., 2018). To set up this feature for the SMD method, the NMR-specific option (nmropt) in the umbrella run is replaced with 'jar', which are used to generate the distance restraints based on the reference coordinates in both instances. The reaction coordinate (RC) for the acylation step that was chosen came from our previous study (Z. K. Sanusi et al., 2020; Lawal et al., 2019; Z. K. Sanusi et al., 2019), having observed a lower energy barrier in the concerted acyclic TS reaction mechanism involving one water molecule. The RC involves the breaking of the C-N at the scissile region and O_{asp}-H_{asp} bonds, and the formation of HO_{wat}-C and N-H_{asp} bonds, hence, the acyclic TS step is defined as; RC = d(C-N) + d(O_{asp}-H_{asp}) – d(HO_{wat}-C) – d(N-H_{asp}) (see Figure 2, Scheme 3).

The simulation starts by sampling the RC from −4.00 Å to +4.00 Å using SMD calculation for 50 ps. Then, structures were obtained at every 0.75 ps from the trajectories produced with the SMD simulation, which serves as the initial configuration for each window in the umbrella sampling. Later, we perform an umbrella sampling of 100 ps, and the scan was done in

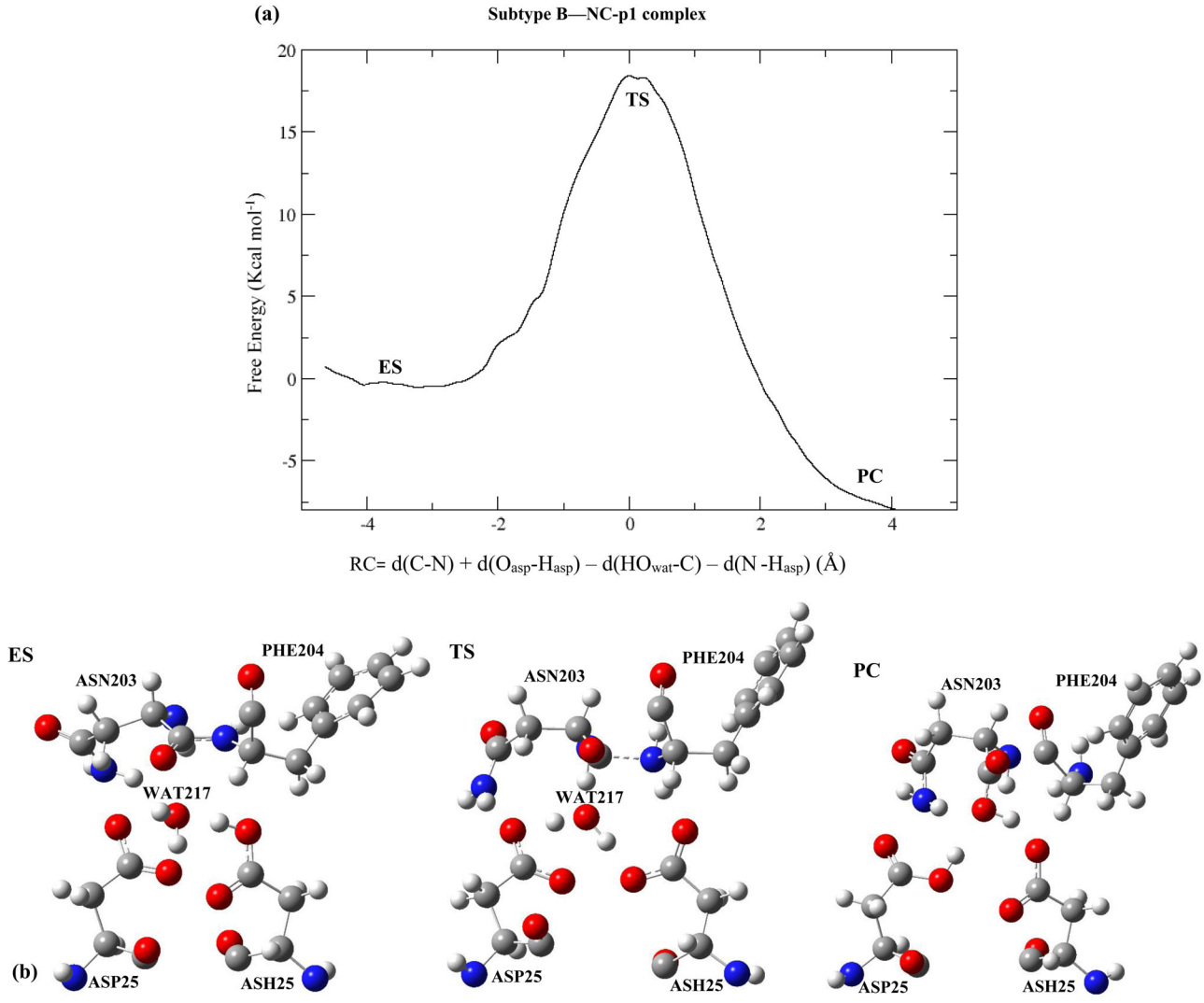


Figure 3. (a) Free energy profile for the concerted acyclic mechanism of subtype B HIV-1 PR with NC-p1 natural substrate. (b) 3D structures of enzyme—substrate (ES), transition state (TS), and product complex (PC) for the acyclic mechanistic process taken from trajectories. Free-energy plots, 3D-structures of ES, TSs, and PC are provided supporting information (SI), and these are listed in the table of content in the SI.

steps of 0.12 Å on each window with the structure generated through the SMD simulation. The outputs of each MD in the sampled umbrella window serve as the initial/starting structure to perform the next one for which the RC's value is restrained to the next position. All the QM/MM MD simulations were carried out, restraining the reaction coordinate distance at a 100 kcal $mol^{-1}.Å^{-2}$ force constant of harmonic potential. The free energy curve was obtained by unbiasing the system along the reaction coordinate through variation free energy profile (vFEP) developed by Lee *et al*. (Lee et al., 2013). 2014;

3. Results and discussion

3.1. The concerted acyclic general acid-base catalytic mechanism of HIV-1 PR

Recent computational research (Z. K. Sanusi et al., 2020; Lawal et al., 2020; Lawal et al., 2019; Z. K. Sanusi et al., 2019) from our group has buttressed the feasibility of the earlier suggested (Jaskólski et al., 1991) existence of the concerted process (Scheme 2b). Based on our previous theoretical investigations (Z. K. Sanusi et al., 2020; Lawal et al., 2019; Z. K. Sanusi et al., 2019), we have applied the umbrella sampling algorithm to explore the proposed concerted process (Scheme 3). The outcome follows the general acid-base model in a synchronous mechanism in which water donates one of its protons to the unprotonated Asp25, and the protonated Asp25 loses its proton to the scissile nitrogen atom; thus, the initial acidic Asp becomes basic and vice versa. Meanwhile, the nucleophilic water (OH) attacks the scissile carbon resulting in substrate cleavage (Scheme 3). In the concerted acylation mechanism (Scheme 3, Figure 2), generally, three major types of conformational states occurred. One is the formation of the HO_{wat}-C bond, the $d(HO_{wat}\text{-}C)$ decreases considerably from 3.40 Å in ES to 2.61 Å in TS, and then to 1.30 Å in product complex (PC). This indicates the nucleophilic attack of the OH group of water on the carbonyl carbon (C) of the peptide at the scissile bond.

Another significant conformational change is the cleavage of the C-N bond. The d(C-N) increases from 1.38 Å in ES to 1.50 Å in the transition state model, and then to 2.83 Å in PC, which indicates the breaking of the natural substrate's

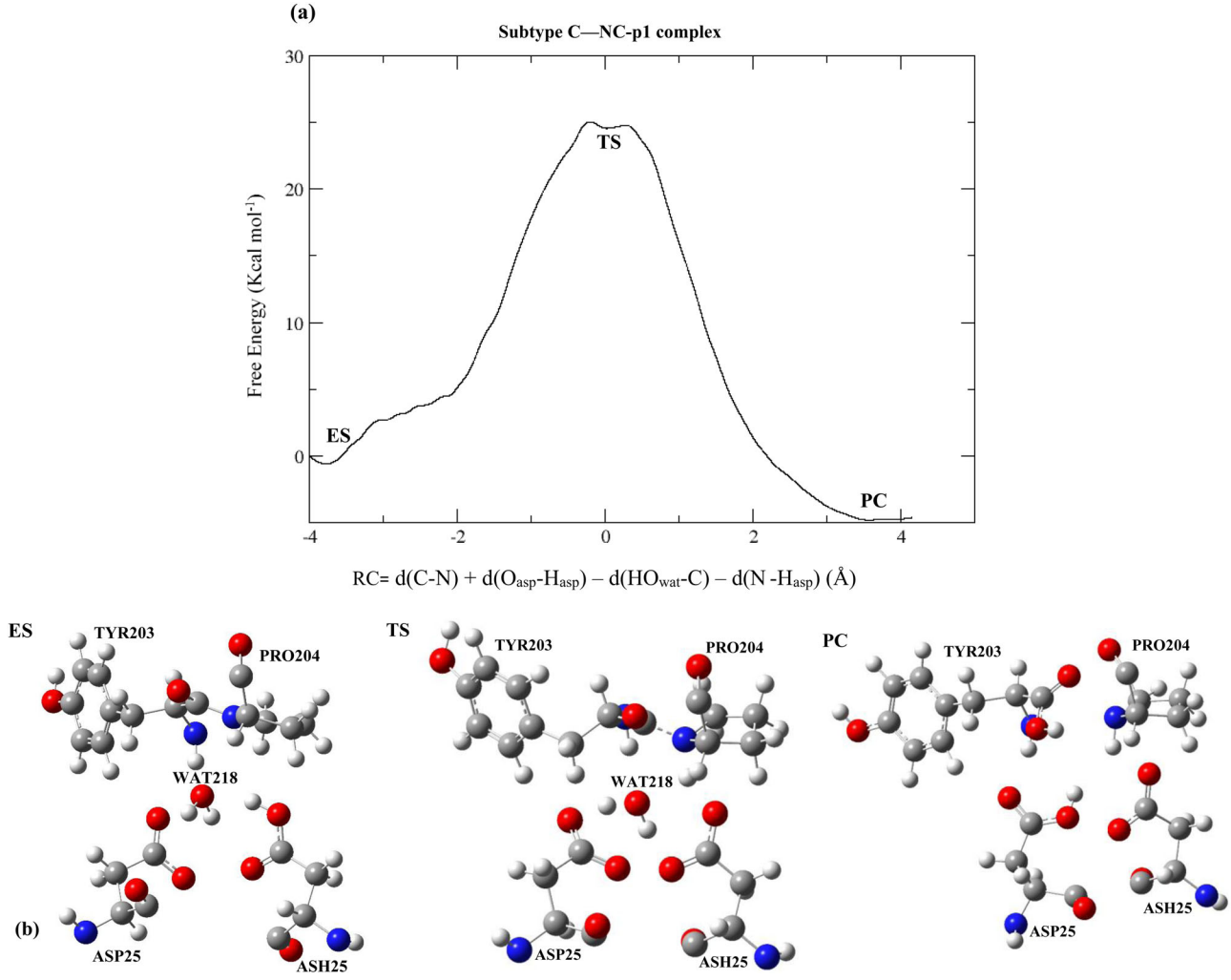


Figure 4. (a) Free energy profile for the concerted acyclic mechanism of subtype C-SA HIV-1 PR with MA-CA natural substrate. (b) 3D enzyme—substrate (ES), transition state (TS), and product complex (PC) for the acyclic mechanistic process taken from trajectories. Free-energy plots, 3D-structures of ES, TSs, and PC are provided in supporting information (SI), and these are listed in the table of content in the SI.

peptide bond. The last occurrence is the proton transfer from the protonated aspartate (H_{asp}) to the nitrogen atom (N) of the substrate at the scissile bond, the d(N-H_{asp}) and d(O_{asp}-H_{asp}), respectively, changes from 3.02 and 1.03 Å in ES to 1.30 and 1.48 Å in TS, and then to 1.07 and 2.22 Å in PC. The measured distances are summarized in Table 1, S2, S3, and S4 for all the enzyme—substrate complexes considered.

The plausibility of this mechanistic process was evaluated by calculating the activation free energy ($\Delta G^{\ddagger}$) value associated with each transition state (TS) structure of the respective complex system (Table 2). It was noticed that the $\Delta G^{\ddagger}$ obtained from the umbrella sampling approach is between 14.2 and 24.7 kcal mol^{-1}. The calculated values follow the same trend as experimental results (Fehér et al., 2002; Lawal et al., 2019; Maschera et al., 1996) and previously calculated values (Z. K. Sanusi et al., 2020) in which an ONIOM approach was used. The overall reaction process (Figures 3a and 4a) is exergonic with negative free energy values for the product complexes (PC).

3.2. Free energy analysis

It should be noted that since the active enzyme cleaves the same scissile precursor as its mutant, the computed results

Table 1. Averaged distances for ES, TS, and PC structures for the concerted mechanism obtained by DFTB/MM 1 D-PMF (Umbrella sampling) for subtype C-SA—MA-CA complex.

DFTB/MM	ES	TS	PC
d(HO_{wat}-C)	3.40	2.61	1.30
d(C-N)	1.38	1.50	2.83
d(N-H_{asp})	3.02	1.30	1.07
d(O_{asp}-H_{asp})	1.03	1.48	2.22

Values are reported in Å. Provided in the supplementary information is the table for the average distances for ES, TS, and PC of the remaining considered complexes. Free-energy plots, 3D-structures of ES, TSs, and PC are provided in the supporting information (SI), and these are listed in the table of content in the SI.

(Table 2) were compared to the activation free energies of the HIV-1 PR found in the literature (Fehér et al., 2002; Lawal et al., 2019; Maschera et al., 1996). For the subtype B—NC-p1 system, $\Delta G^{\ddagger}$ of 18.3 kcal mol^{-1} was obtained for the cleavage of the scissile bond Phe*Asn (Table 2, Figure 3b), which is −0.4 kcal mol^{-1} lower than experimental HIV-1 PR hydrolysis data of the same substrate. (Fehér et al., 2002; Lawal et al., 2019; Maschera et al., 1996)

However, the subtype C-SA—NC-p1 system gave a higher value of 24.7 kcal mol^{-1}, approximately 6 and 2.1 kcal mol^{-1} greater than experimental (subtype B HIV-1 PR) and

Table 2. Free energy value of the concerted acyclic general acid-base cleavage mechanism of the natural substrate by HIV-1 PR (subtypes B and C-SA) using QM(DFTB)/MM MD (Umbrella sampling).

	B—NC-p1	C-SA—MA-CA	C-SA—NC-p1	C-SA—PR-RT
$\Delta G^{\ddagger}$	18.3	14.2	24.7	15.5
PC	−8.0	−18.0	−4.6	−2.5
$\Delta G^{o\ddagger}$	19.7	13.5	22.6	16.4
$\Delta G^{e\ddagger}$	18.7	15.7	18.7	15.8

Values are reported in kilocalories per mole (kcal mol^{-1}).
$\Delta G^{\ddagger}$ = calculated activation free energy (Umbrella sampling), PC product complex, $\Delta G^{o\ddagger}$ = computed activation free energy from previous paper using ONIOM approach (Z. K. Sanusi et al., 2020), $\Delta Ge^{\ddagger}$ = experimental activation free energy from literature for subtype B HIV-1 PR (Fehér et al., 2002; Lawal et al., 2019; Maschera et al., 1996). Free-energy plots, 3D-structures of ES, TSs, and PC are provided in supporting information (SI), and these are listed in the table of content in the SI.

previously calculated data (Z. K. Sanusi et al., 2020) respectively. This could be as a result of the eight-point mutation that occurs in the subtype C-SA HIV-1 PR. Likewise, NC-p1 natural substrate has been suggested to have the slowest mechanistic process as well as being the rate-limiting cleavage step in the maturation of Gag. (Prabu-Jeyabalan et al., 2006) It is noticeable from all components in Table 2 that this system (C-SA—NC-p1) produced the highest $\Delta G^{\ddagger}$ values for both calculated theoretical methods. Hence, it could be that our calculated result is accurate but needs experimental verification for HIV-1 subtype C-SA PR. From the analysis of the calculated $\Delta G^{\ddagger}$ obtained through an enhanced sampling approach herein, it is tenable to say that the computed values are comparable with experimental results. Therefore, the the theoretical mechanism gives a reasonable description of reality. Future studies will involve advanced computational techniques to determine if alternative (stepwise) routes do exist.

The product complexes gave free energy values within −18 to −2.5 kcal mol^{-1} (Table 2), thus indicating that the entire mechanism is energetically favorable. Provided in Figure 3a is the free energy pathway involving subtype B catalysis of NC-p1, showing the scanned reaction coordinates for the acyclic TS model. The zoomed-in structural details of the reaction center are provided in Figure 3b. The free energy profile for the concerted acyclic mechanism of subtype C-SA HIV-1 PR cleavage of MA-CA natural substrate is also provided in Figure 4a with the 3D moieties along its reaction path, Figure 4b.

4. Conclusion

The enzymatic mechanism catalyzed by HIV-1 PR in the presence of its natural substrate was the focus of this study. The mechanism of the proteolysis catalyzed by subtype B/C-SA HIV-1 PR proposed in our previous ONIOM study is explored and this is described utilizing enhanced umbrella sampling calculations at a hybrid DFTB/MM level of theory. Based on the results obtained, we were able to compute a complete profile of the possible reaction pathway that corresponds to the proposed concerted acylation mechanism. The reaction process is synchronous and involves a nucleophilic attack from the catalytic water on the carbonyl atom of the scissile bond and a concurrent proton transfer from the protonated catalytic Asp25 to the nitrogen atom of the scissile bond which results in substrate's peptide bond cleavage.

The present investigation provides a single rate-limiting step with a computed free energy barrier of 14.2, 24.7 (18.3 for subtype B—NC-p1), and 15.5 kcal mol^{-1} for the cleavage of MA-CA, NC-p1, and PR-RT by subtype C-SA, respectively. These calculated energies correspond well with the experimental data for the above natural substrates catalyzed by HIV-1 protease. Since these calculations require moderately high computational cost (long MD simulations), we were restricted to performing a 1-dimensional calculation. However, the accuracy of the results in comparison with experimental data is reasonable; the present investigation shows that the QM/MM MD umbrella sampling is a viable method to predict the free energy barrier. The applied approach requires fewer parameters, convergence is easy to achieve with minimal or no error, and the reaction pathway can be easily determined. However, the system setup and the appropriate choice of the RC along which the simulation is biased and can be challenging to ascertain.

In contrast, the ONIOM calculation in principle has a straightforward calculation setup but the number of input parameters and errors makes achieving convergence more difficult and time-consuming. Nevertheless, both theoretical techniques have proved to give results that are comparable to experimental data. Since umbrella sampling is one of the most applied methods to calculate free energies along a reaction pathway relative to certain reaction coordinate, it can be concluded that the reaction mechanism catalyzed by HIV-1 PR follows the proposed one-step concerted process. The plausibility of the mechanism (Scheme 3) has been established using a multidimensional theoretical approach with substantial outcomes.

Most HIV-1 PR inhibitors are developed to resemble the molecular structure that occurs during the chemical process, particularly the transition state model. We believe electronic modification of the amide bond is possible (for example, via N-methylation)(Rauf et al., 2015), and the QM/MM MD umbrella sampling technique will be useful to determine if the activation energy for hydrolysis increases. If the activation energy of the 'scissile' bond of a natural peptide can be sufficiently increased, it may potentially be used as substrate-based inhibitors (with limited toxicity and side effects) for HIV-1 PR and other enzymes.

Acknowledgements

The authors thank the College of Health Sciences, University of KwaZulu-Natal, Asphen Pharmacare, Medical Research Council, and the National Research Foundation (all in South Africa) for financial support. The authors also acknowledge the University of Florida Research Computing for providing computational resources and support that contributed to the research results reported in this publication. GK thank AR for hosting ZS and ML at UF during the project. We are also grateful to the Centre for High-Performance Computing (www.chpc.ac.za) for computational resources.

Disclosure statement

The authors declare that they have no competing interests.

ORCID

Thavendran Govender http://orcid.org/0000-0003-2511-2503
Sooraj Baijnath http://orcid.org/0000-0001-7860-1779
Tricia Naicker http://orcid.org/0000-0002-7134-6258
Glenn E. M. Maguire http://orcid.org/0000-0002-9135-5370
Hendrik G. Kruger http://orcid.org/0000-0003-0606-2053

References

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. (2009). Structure of HIV-1 protease in complex with potent inhibitor KNI-272 determined by high-resolution X-ray and neutron crystallography. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(12), 4641–4646. https://doi.org/10.1073/pnas.0809400106

Antonov, V., Ginodman, L., Kapitannikov, Y. V., Barshevskaya, T., Gurova, A., & Rumsh, L. (1978). Mechanism of pepsin catalysis: General base catalysis by the active-site carboxylate ion . *FEBS Letters*, *88*(1), 87–90. https://doi.org/10.1016/0014-5793(78)80613-9

Barnett, C. B., & Naidoo, K. J. (2010). Ring puckering: A metric for evaluating the accuracy of AM1, PM3, PM3CARB-1, and SCC-DFTB carbohydrate QM/MM simulations. *The Journal of Physical Chemistry. B*, *114*(51), 17142–17154. https://doi.org/10.1021/jp107620h

Bjelic, S., & Åqvist, J. (2006). Catalysis and linear free energy relationships in aspartic proteases. Biochemistry, *45*(25), 7709–7723. https://doi.org/10.1021/bi060131y

Boross, P., Bagossi, P., Copeland, T. D., Oroszlan, S., Louis, J. M., & Tözsér, J. (1999). Effect of substrate residues on the P2′ preference of retroviral proteinases. *European Journal of Biochemistry*, *264*(3), 921–929. https://doi.org/10.1046/j.1432-1327.1999.00687.x

Case, D., Ben-Shalom, I., Brozell, S., Cerutti, D., Cheatham III, T., Cruzeiro, V., Darden, T., Duke, R., Ghoreishi, D., & Gilson, M. (2018). AMBER 2018. University of California.

Case, D., Cerutti, D., Cheatham, T., III, Darden, T., Duke, R., Giese, T., … Homeyer, N. (2018). AMBER reference manual. University of California.

Chatfield, D. C., Eurenius, K. P., & Brooks, B. R. (1998). HIV-1 protease cleavage mechanism: A theoretical investigation based on classical MD simulation and reaction path calculations using a hybrid QM/MM potential. *Journal of Molecular Structure: Theochem*, *423*(1-2), 79–92. https://doi.org/10.1016/S0166-1280(96)04875-0

Coates, L., Erskine, P. T., Mall, S., Gill, R., Wood, S. P., Myles, D. A., & Cooper, J. B. (2006). X-ray, neutron and NMR studies of the catalytic mechanism of aspartic proteinases. *European Biophysics Journal : Ebj*, *35*(7), 559–566. https://doi.org/10.1007/s00249-006-0065-7

Cole, J. C., Murray, C. W., Nissink, J. W. M., Taylor, R. D., & Taylor, R. (2005). Comparing protein-ligand docking programs is difficult. Proteins, *60*(3), 325–332. https://doi.org/10.1002/prot.20497

DeLano, W. L. (2002). The PyMOL molecular graphics system.

Di Martino, G. P., Masetti, M., Cavalli, A., & Recanatini, M. (2014). Mechanistic insights into P in1 peptidyl-prolyl cis-trans isomerization from umbrella sampling simulations. *Proteins: Structure, Function, and Bioinformatics*, *82*(11), 2943–2956. https://doi.org/10.1002/prot.24650

Di Palma, F., Bottaro, S., & Bussi, G. (2015). Kissing loop interaction in adenine riboswitch: Insights from umbrella sampling simulations. *BMC Bioinformatics*, *16*(Suppl 9), S6. https://doi.org/10.1186/1471-2105-16-S9-S6

Elstner, M., Porezag, D., Jungnickel, G., Elsner, J., Haugk, M., Frauenheim, T., Suhai, S., & Seifert, G. (1998). Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. *Physical Review B*, *58*(11), 7260–7268. https://doi.org/10.1103/PhysRevB.58.7260

Fakhar, Z., Govender, T., Lamichhane, G., Maguire, G. E., Kruger, H. G., & Honarparvar, B. (2017). Computational model for the acylation step of the β-lactam ring: Potential application for l, d-transpeptidase 2 in mycobacterium tuberculosis. *Journal of Molecular Structure*, *1128*, 94–102. https://doi.org/10.1016/j.molstruc.2016.08.049

Fehér, A., Weber, I. T., Bagossi, P., Boross, P., Mahalingam, B., Louis, J. M., Copeland, T. D., Torshin, I. Y., Harrison, R. W., & Tözsér, J. (2002). Effect of sequence polymorphism and drug resistance on two HIV-1 Gag processing sites. *European Journal of Biochemistry*, *269*(16), 4114–4120. https://doi.org/10.1046/j.1432-1033.2002.03105.x

Frisch, M., Trucks, G., Schlegel, H., Scuseria, G., Robb, M., Cheeseman, J., … Nakatsuji, H. (2016). Gaussian 16 Revision B. 01. 2016 Gaussian Inc. Wallingford CT.

Fruton, J. S. (1976). The mechanism of the catalytic action of pepsin and related acid proteinases. *Advances in Enzymology and Related Areas of Molecular Biology*, *44*, 1–36. Volume https://doi.org/10.1002/9780470122891.ch1

Garrec, J., Sautet, P., & Fleurat-Lessard, P. (2011). Understanding the HIV-1 protease reactivity with DFT: What do we gain from recent functionals? *The Journal of Physical Chemistry. B*, *115*(26), 8545–8558. https://doi.org/10.1021/jp200565w

Gohlke, H., Hendlich, M., & Klebe, G. (2000). Knowledge-based scoring function to predict protein-ligand interactions. *Journal of Molecular Biology*, *295*(2), 337–356. https://doi.org/10.1006/jmbi.1999.3371

Hazel, A., & Gumbart, J. C. (2017). Methods for calculating potentials of mean force. *School of Physics, Georgia Institute of Technology*.

Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., & Simmerling, C. (2006). Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics*, *65*(3), 712–725. https://doi.org/10.1002/prot.21123

Hsu, I.-N., Delbaere, L. T., James, M. N., & Hofmann, T. (1977). Penicillopepsin from penicillium janthinellum crystal structure at 2.8 A and sequence homology with porcine pepsin. Nature, *266*(5598), 140–145. https://doi.org/10.1038/266140a0

Hyland, L. J., Tomaszek, T. A., Jr, & Meek, T. D. (1991). Human immunodeficiency virus-1 protease. 2. Use of pH rate studies and solvent kinetic isotope effects to elucidate details of chemical mechanism. Biochemistry, *30*(34), 8454–8463. https://doi.org/10.1021/bi00098a024

Jarzynski, C. (1997a). Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach. *Physical Review E*, *56*(5), 5018–5035. https://doi.org/10.1103/PhysRevE.56.5018

Jarzynski, C. (1997b). Nonequilibrium equality for free energy differences. *Physical Review Letters*, *78*(14), 2690–2693. https://doi.org/10.1103/PhysRevLett.78.2690

Jaskólski, M., Tomasselli, A. G., Sawyer, T. K., Staples, D. G., Heinrikson, R. L., Schneider, J., Kent, S. B., & Wlodawer, A. (1991). Structure at 2.5-A resolution of chemically synthesized human immunodeficiency virus type 1 protease complexed with a hydroxyethylene-based inhibitor . Biochemistry, *30*(6), 1600–1609. https://doi.org/10.1021/bi00220a023

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, *79*(2), 926–935. https://doi.org/10.1063/1.445869

Kästner, J. (2011). Umbrella sampling. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *1*(6), 932–942. https://doi.org/10.1002/wcms.66

Kästner, J., & Thiel, W. (2005). Bridging the gap between thermodynamic integration and umbrella sampling provides a novel analysis method: "Umbrella integration":". *The Journal of Chemical Physics*, *123*(14), 144104 https://doi.org/10.1063/1.2052648

Kempf, D. J., Norbeck, D. W., Codacovi, L., Wang, X. C., Kohlbrenner, W. E., Wideburg, N. E., Paul, D. A., Knigge, M. F., Vasavanonda, S., & Craig-Kennard, A. (1990). Structure-based, C2 symmetric inhibitors of HIV protease. *Journal of Medicinal Chemistry*, *33*(10), 2687–2689. https://doi.org/10.1021/jm00172a002

Kipp, D. R., Silva, R. G., & Schramm, V. L. (2011). Mass-dependent bond vibrational dynamics influence catalysis by HIV-1 protease. *Journal of the American Chemical Society*, *133*(48), 19358–19361. https://doi.org/10.1021/ja209391n

Kohl, N. E., Emini, E. A., Schleif, W. A., Davis, L. J., Heimbach, J. C., Dixon, R. A., Scolnick, E. M., & Sigal, I. S. (1988). Active human immunodeficiency virus protease is required for viral infectivity. *Proceedings of the National Academy of Sciences of the United States of America*, *85*(13), 4686–4690. https://doi.org/10.1073/pnas.85.13.4686

Kontoyianni, M., McClellan, L. M., & Sokol, G. S. (2004). Evaluation of docking performance: Comparative data on docking algorithms. *Journal of Medicinal Chemistry*, *47*(3), 558–565. https://doi.org/10.1021/jm0302997

Krzemińska, A., Moliner, V., & Świderek, K. (2016). Dynamic and electrostatic effects on the reaction catalyzed by HIV-1 protease. *Journal of the American Chemical Society*, *138*(50), 16283–16298. https://doi.org/10.1021/jacs.6b06856

Kumar, M., Kannan, K., Hosur, M., Bhavesh, N. S., Chatterjee, A., Mittal, R., & Hosur, R. (2002). Effects of remote mutation on the autolysis of HIV-1 PR: X-ray and NMR investigations. *Biochemical and Biophysical Research Communications*, *294*(2), 395–401. https://doi.org/10.1016/S0006-291X(02)00482-5

Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., & Kollman, P. A. (1992). The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *Journal of Computational Chemistry*, *13*(8), 1011–1021. https://doi.org/10.1002/jcc.540130812

Lapatto, R., Blundell, T., Hemmings, A., Overington, J., Wilderspin, A., Wood, S., Merson, J. R., Whittle, P. J., Danley, D. E., & Geoghegan, K. F. (1989). X-ray analysis of HIV-1 proteinase at 2.7 A resolution confirms structural homology among retroviral enzymes. Nature, *342*(6247), 299–302. https://doi.org/10.1038/342299a0

Lawal, M. M., Sanusi, Z. K., Govender, T., Maguire, G. E., Honarparvar, B., & Kruger, H. G. (2020). From recognition to reaction mechanism: An overview on the interactions between HIV-1 protease and its natural targets. *Current Medicinal Chemistry*, *27*(15), 2514–2549. https://doi.org/10.2174/0929867325666181113122900

Lawal, M. M., Sanusi, Z. K., Govender, T., Tolufashe, G. F., Maguire, G. E., Honarparvar, B., & Kruger, H. G. (2019). Unraveling the concerted catalytic mechanism of the human immunodeficiency virus type 1 (HIV-1) protease: A hybrid QM/MM study. *Structural Chemistry*, *30*(1), 409–417. https://doi.org/10.1007/s11224-018-1251-9

Lee, T.-S., Radak, B. K., Huang, M., Wong, K.-Y., & York, D. M. (2014). Roadmaps through free energy landscapes calculated using the multidimensional vFEP approach. *Journal of Chemical Theory and Computation*, *10*(1), 24–34. https://doi.org/10.1021/ct400691f

Lee, T.-S., Radak, B. K., Pabis, A., & York, D. M. (2013). A new maximum likelihood approach for free energy profile construction from molecular simulations. *Journal of Chemical Theory and Computation*, *9*(1), 153–164. https://doi.org/10.1021/ct300703z

Li, H., Robertson, A. D., & Jensen, J. H. (2005). Very fast empirical prediction and rationalization of protein pKa values. *Proteins: Structure, Function, and Bioinformatics*, *61*(4), 704–721. https://doi.org/10.1002/prot.20660

Liu, H., Müller-Plathe, F., & van Gunsteren, W. F. (1996). A combined quantum/classical molecular dynamics study of the catalytic mechanism of HIV protease. *Journal of Molecular Biology*, *261*(3), 454–469. https://doi.org/10.1006/jmbi.1996.0476

Liu, Z., Wang, Y., Brunzelle, J., Kovari, I. A., & Kovari, L. C. (2011). Nine crystal structures determine the substrate envelope of the MDR HIV-1 protease. *The Protein Journal*, *30*(3), 173–183. https://doi.org/10.1007/s10930-011-9316-2

Maschera, B., Darby, G., Palú, G., Wright, L. L., Tisdale, M., Myers, R., Blair, E. D., & Furfine, E. S. (1996). Human immunodeficiency virus mutations in the viral protease that confer resistance to saquinavir increase the dissociation rate constant of the protease-saquinavir complex. *The Journal of Biological Chemistry*, *271*(52), 33231–33235. https://doi.org/10.1074/jbc.271.52.33231

McGee, T. D., Jr., Edwards, J., & Roitberg, A. E. (2014). pH-REMD simulations indicate that the catalytic aspartates of HIV-1 protease exist primarily in a monoprotonated state. *The Journal of Physical Chemistry B*, *118*(44), 12577–12585. https://doi.org/10.1021/jp504011c

Meek, T. D., Dayton, B. D., Metcalf, B. W., Dreyer, G. B., Strickler, J. E., Gorniak, J. G., Rosenberg, M., Moore, M. L., Magaard, V. W., & Debouck, C. (1989). Human immunodeficiency virus 1 protease expressed in escherichia coli behaves as a dimeric aspartic protease. *Proceedings of the National Academy of Sciences of the United States of America*, *86*(6), 1841–1845. https://doi.org/10.1073/pnas.86.6.1841

Naicker, P., Achilonu, I., Fanucchi, S., Fernandes, M., Ibrahim, M. A. A., Dirr, H. W., Soliman, M. E. S., & Sayed, Y. (2013). Structural insights into the South African HIV-1 subtype C protease: Impact of hinge region dynamics and flap flexibility in drug resistance. *Journal of Biomolecular Structure & Dynamics*, *31*(12), 1370–1380. https://doi.org/10.1080/07391102.2012.736774

Navia, M. A., Fitzgerald, P. M., McKeever, B. M., Leu, C. T., Heimbach, J. C., Herber, W. K., Sigal, I. S., Darke, P. L., & Springer, J. P. (1989). Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. Nature, *337*(6208), 615–620. https://doi.org/10.1038/337615a0

Northrop, D. B. (2001). Follow the protons: A low-barrier hydrogen bond unifies the mechanisms of the aspartic proteases. *Accounts of Chemical Research*, *34*(10), 790–797. https://doi.org/10.1021/ar000184m

Park, H., Suh, J., & Lee, S. (2000). Ab initio studies on the catalytic mechanism of aspartic proteinases: Nucleophilic versus general acid/general base mechanism. *Journal of the American Chemical Society*, *122*(16), 3901–3908. https://doi.org/10.1021/ja992849p

Pastor, R. W., Brooks, B. R., & Szabo, A. (1988). An analysis of the accuracy of Langevin and molecular dynamics algorithms. *Molecular Physics*, *65*(6), 1409–1419. https://doi.org/10.1080/00268978800101881

Piana, S., Bucher, D., Carloni, P., & Rothlisberger, U. (2004). Reaction mechanism of HIV-1 protease by hybrid car-Parrinello/classical MD simulations. *The Journal of Physical Chemistry B*, *108*(30), 11139–11149. https://doi.org/10.1021/jp037651c

Piana, S., & Carloni, P. (2000). Conformational flexibility of the catalytic Asp dyad in HIV-1 protease: An ab initio study on the free enzyme. *Proteins: Structure, Function, and Genetics*, *39*(1), 26–36. https://doi.org/10.1002/(SICI)1097-0134(20000401)39:1<26::AID-PROT3>3.0.CO;2-N

Piana, S., Carloni, P., & Parrinello, M. (2002). Role of conformational fluctuations in the enzymatic reaction of HIV-1 protease. *Journal of Molecular Biology*, *319*(2), 567–583. https://doi.org/10.1016/S0022-2836(02)00301-7

Piana, S., Sebastiani, D., Carloni, P., & Parrinello, M. (2001). Ab initio molecular dynamics-based assignment of the protonation state of pepstatin A/HIV-1 protease cleavage site. *Journal of the American Chemical Society*, *123*(36), 8730–8737. https://doi.org/10.1021/ja003145e

Prabu-Jeyabalan, M., Nalivaika, E. A., King, N. M., & Schiffer, C. A. (2004). Structural basis for coevolution of a human immunodeficiency virus type 1 nucleocapsid-p1 cleavage site with a V82A drug-resistant mutation in viral protease. *Journal of Virology*, *78*(22), 12446–12454. https://doi.org/10.1128/JVI.78.22.12446-12454.2004

Prabu-Jeyabalan, M., Nalivaika, E. A., Romano, K., & Schiffer, C. A. (2006). Mechanism of substrate recognition by drug-resistant human immunodeficiency virus type 1 protease variants revealed by a novel structural intermediate. *Journal of Virology*, *80*(7), 3607–3616. https://doi.org/10.1128/JVI.80.7.3607-3616.2006

Prabu-Jeyabalan, M., Nalivaika, E., & Schiffer, C. A. (2002). Substrate shape determines specificity of recognition for HIV-1 protease: Analysis of crystal structures of six substrate complexes. *Structure (London, England : 1993)*, *10*(3), 369–381. https://doi.org/10.1016/s0969-2126(02)00720-7

Rauf, S. M. A., Arvidsson, P. I., Albericio, F., Govender, T., Maguire, G. E., Kruger, H. G., & Honarparvar, B. (2015). The effect of N-methylation of amino acids (Ac-X-OMe) on solubility and conformation: A DFT study. *Organic & Biomolecular Chemistry*, *13*(39), 9993–10006.

Rick, S. W., Erickson, J. W., & Burt, S. K. (1998). Reaction path and free energy calculations of the transition between alternate conformations of HIV-1 protease. *Proteins: Structure, Function, and Genetics*, *32*(1), 7–16. https://doi.org/10.1002/(SICI)1097-0134(19980701)32:1<7::AID-PROT3>3.0.CO;2-K

Roberts, N. A., Martin, J. A., Kinchington, D., Broadhurst, A. V., Craig, J. C., Duncan, I. B., Galpin, S. A., Handa, B. K., Kay, J., & Kröhn, A. (1990). Rational design of peptide-based HIV proteinase inhibitors. *Science (New York, N.Y.)*, *248*(4953), 358–361. https://doi.org/10.1126/science.2183354

Rodriguez, E. J., Angeles, T. S., & Meek, T. D. (1993). Use of nitrogen-15 kinetic isotope effects to elucidate details of the chemical mechanism of human immunodeficiency virus 1 protease. Biochemistry, *32*(46), 12380–12385. https://doi.org/10.1021/bi00097a015

Rodriguez, E. J., Debouck, C., Deckman, I. C., Abu-Soud, H., Raushel, F. M., & Meek, T. D. (1993). Inhibitor binding to the Phe53Trp mutant of HIV-1 protease promotes conformational changes detectable by spectrofluorometry. Biochemistry, *32*(14), 3557–3563. https://doi.org/10.1021/bi00065a006

Ryckaert, J.-P., Ciccotti, G., & Berendsen, H. J. (1977). Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *Journal of Computational Physics*, *23*(3), 327–341. https://doi.org/10.1016/0021-9991(77)90098-5

Sanusi, Z. K., Govender, T., Maguire, G. E., Maseko, S. B., Lin, J., Kruger, H. G., & Honarparvar, B. (2018). An insight to the molecular interactions of the FDA approved HIV PR drugs against L38L↑ N↑ L PR mutant. *Journal of Computer-Aided Molecular Design*, *32*(3), 459–471. https://doi.org/10.1007/s10822-018-0099-9

Sanusi, Z., Govender, T., Maguire, G., Maseko, S., Lin, J., Kruger, H., & Honarparvar, B. (2017). Investigation of the binding free energies of FDA approved drugs against subtype B and C-SA HIV PR: ONIOM approach. *Journal of Molecular Graphics & Modelling*, *76*, 77–85. https://doi.org/10.1016/j.jmgm.2017.06.026

Sanusi, Z. K., Lawal, M. M., Govender, T., Baijnath, S., Naicker, T., Maguire, G. E. M., Honarparvar, B., & Kruger, H. G. (2020). Concerted hydrolysis mechanism of HIV-1 natural substrate against subtypes B and C-SA PR: Insight through molecular dynamics and hybrid QM/MM studies. *Physical Chemistry Chemical Physics: PCCP*, *22*(4), 2530–2539. https://doi.org/10.1039/c9cp05639d

Sanusi, Z. K., Lawal, M. M., Govender, T., Maguire, G. E., Honarparvar, B., & Kruger, H. G. (2019). Theoretical model for HIV-1 PR that accounts for substrate recognition and preferential cleavage of natural substrates. *The Journal of Physical Chemistry. B*, *123*(30), 6389–6400. https://doi.org/10.1021/acs.jpcb.9b02207

Schulz, G. E. (1991). Domain motions in proteins. *Current Opinion in Structural Biology*, *1*(6), 883–888. https://doi.org/10.1016/0959-440X(91)90082-5

Seabra, G. d M., Walker, R. C., Elstner, M., Case, D. A., & Roitberg, A. E. (2007). Implementation of the SCC-DFTB method for hybrid QM/MM simulations within the Amber molecular dynamics package. *The Journal of Physical Chemistry A*, *111*(26), 5655–5664. https://doi.org/10.1021/jp070071l

Silva, A. M., Cachau, R. E., Sham, H. L., & Erickson, J. W. (1996). Inhibition and catalytic mechanism of HIV-1 aspartic protease. *Journal of Molecular Biology*, *255*(2), 321–340. https://doi.org/10.1006/jmbi.1996.0026

Silva, J. R. A., Govender, T., Maguire, G. E., Kruger, H. G., Lameira, J., Roitberg, A. E., & Alves, C. N. (2015). Simulating the inhibition reaction of mycobacterium tuberculosis L,D-transpeptidase 2 by carbapenems . *Chemical Communications (Cambridge, England))*, *51*(63), 12560–12562. https://doi.org/10.1039/c5cc03202d

Silva, J. R. A., Roitberg, A. E., & Alves, C. u N. (2014). Catalytic mechanism of L,D-transpeptidase 2 from Mycobacterium tuberculosis described by a computational approach: insights for the design of new antibiotics drugs . *Journal of Chemical Information and Modeling*, *54*(9), 2402–2410. https://doi.org/10.1021/ci5003069

Smith, R., Brereton, I. M., Chai, R. Y., & Kent, S. B. (1996). Ionization states of the catalytic residues in HIV-1 protease. *Nature Structural & Molecular Biology*, *3*(11), 946–950.

Souaille, M., & Roux, B. (2001). Extension to the weighted histogram analysis method: Combining umbrella sampling with free energy calculations. *Computer Physics Communications*, *135*(1), 40–57. https://doi.org/10.1016/S0010-4655(00)00215-0

Stanton, R. V., Hartsough, D. S., & Merz, K. M. Jr, (1993). Calculation of solvation free energies using a density functional/molecular dynamics coupled potential. *The Journal of Physical Chemistry*, *97*(46), 11868–11870. https://doi.org/10.1021/j100148a005

Sussman, F., Villaverde, M. C., L., Dominguez, J., & Danielson, U. H. (2013). On the active site protonation state in aspartic proteases: Implications for drug design. *Current Pharmaceutical Design*, *19*(23), 4257–4275. https://doi.org/10.2174/1381612811319230009

Torrie, G. M., & Valleau, J. P. (1974). Monte Carlo free energy estimates using non-Boltzmann sampling: Application to the sub-critical lennard-Jones fluid. *Chemical Physics Letters*, *28*(4), 578–581. https://doi.org/10.1016/0009-2614(74)80109-0

Torrie, G., & Valleau, J. (1977). Monte Carlo study of a phase-separating liquid mixture by umbrella sampling. *The Journal of Chemical Physics*, *66*(4), 1402–1408. https://doi.org/10.1063/1.434125

Tözsér, J., Bagossi, P., Weber, I. T., Copeland, T. D., & Oroszlan, S. (1996). Comparative studies on the substrate specificity of avian myeloblastosis virus proteinase and lentiviral proteinases. *The Journal of Biological Chemistry*, *271*(12), 6781–6788. https://doi.org/10.1074/jbc.271.12.6781

Trylska, J., Bała, P., Geller, M., & Grochowski, P. (2002). Molecular dynamics simulations of the first steps of the reaction catalyzed by HIV-1 protease. *Biophysical Journal*, *83*(2), 794–807. https://doi.org/10.1016/S0006-3495(02)75209-0

Trylska, J., Grochowski, P., & McCammon, J. A. (2004). The role of hydrogen bonding in the enzymatic reaction catalyzed by HIV-1 protease. *Protein Science : a Publication of the Protein Society*, *13*(2), 513–528. https://doi.org/10.1110/ps.03372304

Tsukamoto, S., Sakae, Y., Itoh, Y., Suzuki, T., & Okamoto, Y. (2018). Computational analysis for selectivity of histone deacetylase inhibitor by replica-exchange umbrella sampling molecular dynamics simulations. *The Journal of Chemical Physics*, *148*(12), 125102. https://doi.org/10.1063/1.5019209

Veronese, F. D., DeVico, A. L., Copeland, T. D., Oroszlan, S., Gallo, R. C., & Sarngadharan, M. (1985). Characterization of gp41 as the transmembrane protein coded by the HTLV-III/LAV envelope gene. *Science (New York, N.Y.).)*, *229*(4720), 1402–1405. https://doi.org/10.1126/science.2994223

Wang, Y. X., Freedberg, D. I., Yamazaki, T., Wingfield, P. T., Stahl, S. J., Kaufman, J. D., Kiso, Y., & Torchia, D. A. (1996). Solution NMR evidence that the HIV-1 protease catalytic aspartyl groups have different ionization states in the complex formed with the asymmetric drug KNI-272. Biochemistry, *35*(31), 9945–9950. https://doi.org/10.1021/bi961268z

Weber, I. T., Cavanaugh, D. S., & Harrison, R. W. (1998). Models of HIV-1 protease with peptides representing its natural substrates. *Journal of Molecular Structure: THEOCHEM*, *423*(1-2), 1–12. https://doi.org/10.1016/S0166-1280(96)04869-5

Wlodawer, A., & Erickson, J. W. (1993). Structure-based inhibitors of HIV-1 protease. Annu. Rev. Biochem, *62*(1), 543–585. https://doi.org/10.1146/annurev.bi.62.070193.002551

Woodcock, H. L., Hodoscek, M., & Brooks, B. R. (2007). Exploring SCC-DFTB paths for mapping QM/MM reaction mechanisms. *The Journal of Physical Chemistry. A*, *111*(26), 5720–5728. https://doi.org/10.1021/jp0714217

Yildirim, I., Park, H., Disney, M. D., & Schatz, G. C. (2013). A dynamic structural model of expanded RNA CAG repeats: A refined X-ray structure and computational investigations using molecular dynamics and umbrella sampling simulations. *Journal of the American Chemical Society*, *135*(9), 3528–3538. https://doi.org/10.1021/ja3108627