

Molecular Mechanism of Protein Arginine Deiminase 2: A Study Involving Multiple Microsecond Long Molecular Dynamics Simulations

Erdem Cicek, Gerald Monard,* and Fethiye Aylin Sungur*



Cite This: *Biochemistry* 2022, 61, 1286–1297



Read Online

ACCESS |



Metrics & More

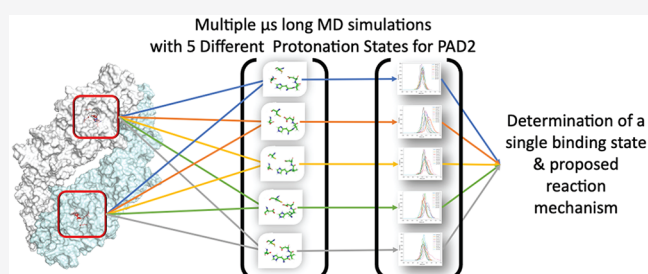


Article Recommendations



Supporting Information

ABSTRACT: Peptidylarginine deiminase 2 (PAD2) is a Ca^{2+} -dependent enzyme that catalyzes the conversion of protein arginine residues to citrulline. This kind of structural modification in histone molecules may affect gene regulation, leading to effects that may trigger several diseases, including breast cancer, which makes PAD2 an attractive target for anticancer drug development. To design new effective inhibitors to control activation of PAD2, improving our understanding of the molecular mechanisms of PAD2 using up-to-date computational techniques is essential. We have designed five different PAD2–substrate complex systems based on varying protonation states of the active site residues. To search the conformational space broadly, multiple independent molecular dynamics simulations of the complexes have been performed. In total, 50 replica simulations have been performed, each of 1 μs , yielding a total simulation time of 50 μs . Our findings identify that the protonation states of Cys647, Asp473, and His471 are critical for the binding and localization of the *N*- α -benzoyl-L-arginine ethyl ester substrate within the active site. A novel mechanism for enzyme activation is proposed according to near attack conformers. This represents an important step in understanding the mechanism of citrullination and developing PAD2-inhibiting drugs for the treatment of breast cancer.



Highly conserved positively charged histone proteins are primary protein components of chromatin fiber complexes serving as the scaffold for DNA in eukaryotic cells. Structural modifications in histone molecules cause loss of interaction with DNA and other nuclear proteins that affect major chromatin functions like transcriptional activation/inactivation, chromosome packaging, and DNA damage/repair.^{1,2}

Such structural modifications belong to a set of post-translational modifications, including phosphorylation, methylation, acetylation, ubiquitination, and citrullination.³ Thus, control of the regulation of gene expression within such a highly compact environment still represents a challenging question in cell biology.^{4,5} Peptidylarginine deiminase (PAD) enzymes, commonly found in mammalian cells, catalyze the hydrolysis of peptidylarginine to peptidyl citrulline in a reaction called deimination or citrullination.^{6–8} PADs are calcium-dependent enzymes that use a nucleophilic cysteine to hydrolyze guanidinium groups on arginine residues to form citrulline (Figure 1).⁹ This reaction results in the loss of a positive charge, thereby affecting protein function and altering protein–protein and protein–nucleic acid interactions.^{10,11}

The PAD family is composed of five calcium-dependent isozymes (PAD1–4 and PAD6), which share roughly 50% sequence similarity and have different tissue distributions and biological functions.^{12,13} PAD1 has an important role in

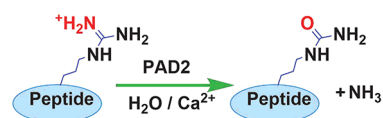


Figure 1. Citrullination reaction catalyzed by PAD2.

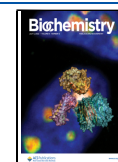
terminal differentiation of keratinocytes.^{14,15} PAD2 is involved in myelin loss.^{16,17} PAD3 enables hair growth by citrullination of trichohyalin.¹⁸ PAD4 has been reported to be involved in the regulation of gene expression^{19,20} and in the formation of extracellular DNA traps.²¹ A role in reproduction has been suggested for PAD6.²²

PAD enzymes have garnered significant attention over the past several years with regard to their dysregulated activity in cancer and involvement in a number of inflammatory (e.g., multiple sclerosis, rheumatoid arthritis, and ulcerative colitis) and autoimmune (lupus) diseases.^{7,10,13,23} Although it is unclear how PADs contribute to such a disparate number of

Received: March 18, 2022

Revised: June 9, 2022

Published: June 23, 2022



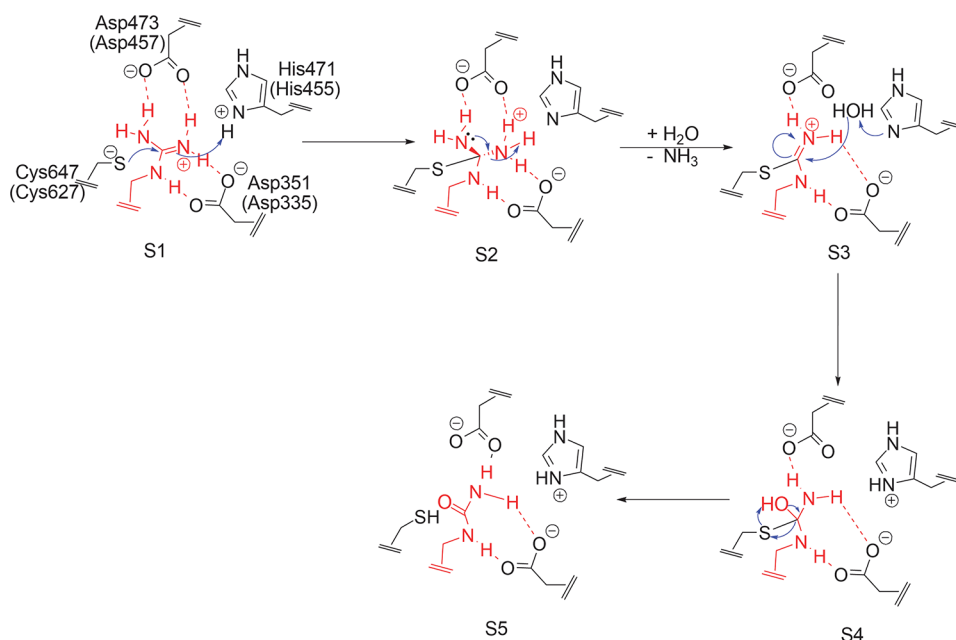


Figure 2. Experimentally proposed catalytic mechanism of PAD2 (from ref 36).

diseases, common links include a role for PAD4 in promoting neutrophil extracellular trap (NET) formation and regulating gene transcription.²⁴ Further evidence that upregulated PAD activity plays a role in these various diseases comes from the demonstration that Cl-amidine, a potent pan-PAD inhibitor, or its analogues show efficacy in animal models of cancer,^{25,26} rheumatoid arthritis,²⁷ lupus,²⁸ thrombosis, spinal cord injury,²⁹ and ulcerative colitis.

Although dysregulated PAD4 activity is typically associated with several diseases, more recent work suggests that PAD2 also plays an important role in both extracellular trap formation and gene regulation.^{30,31} Thus, it is possible that PAD4 and PAD2 carry out similar and/or related functions during disease progression. More recently, a detailed ChIP-chip study demonstrated that PAD2 also plays a critical role in ER target gene activation via the citrullination of histone H3Arg26 at ER target gene promoters.³² Additionally, it was found that PAD2 expression is highly correlated with HER2 expression across more than 60 breast cancer cell lines.³⁰ From a therapeutic standpoint, 75% and 15% of all breast cancers are ER and HER2+, respectively. Given that PAD2 likely plays an important role in the biology of both ER and HER2+ lesions, these observations suggest that PAD2 could represent a therapeutic target for 85–90% of all breast cancers in humans.³⁰ The high-resolution structures of PAD1–4 have been reported.^{9,33–35} The catalytic activity for PADs is known to be regulated by calcium ions through conformational changes, including the appropriate positioning of the catalytic cysteine in the active site cleft. Slade et al. have determined the X-ray structures of the apo and holo states for PAD2 (27 in total), including the wild type, and structures with mutations on the calcium binding residues.³⁴ The role of the six calcium ions and the activation of PAD2 has been revealed by these successive X-ray structures.

Experimental studies have identified the residues that play a role in the catalytic activity in addition to the mechanistic differences between PAD2 and PAD4.³⁶ The catalytic mechanism suggested for PAD2 by Dreyton et al. (see Figure 2) starts with the attack of a nucleophilic Cys647 on the

guanidinium carbon of the substrate arginine while His471 protonates the guanidinium group to form the S-alkyl tetrahedral intermediate, followed by the departure of an ammonia molecule. In the second part, His471 acts as a general base and activates the water molecule for a nucleophilic attack forming the second intermediate. The reaction is concluded by the formation of the citrullinated product.³⁶

When proteins are modeled using molecular dynamics (MD), obtaining a correct sampling of their different conformations can be quite a challenge. Given the size of the systems (i.e., the number of degrees of freedom), many configurations must be generated to provide a statistical representative sample. This is usually sought by simulating (very) large MD trajectories. The conjunction of these two requirements, large systems and long trajectories, adds up to yield computationally intensive simulations. In practice, the limited time scale of a protein MD leads to a sampling problem.^{37,38} To overcome it, increasing the simulation time and performing multiple independent simulations are possible solutions.³⁹ Recently, simulations of multiple copies with different initial parameters of the same system (e.g., initial speeds, hardware differences, shift multiplication accuracy, etc.) have been preferred over single long simulations to obtain reproducible results.^{38,40–43}

In this work, multiple independent MD simulations for the enzyme PAD2 complexed with *N*- α -benzoyl-L-arginine ethyl ester (PAD2–BAEE) were performed to understand the homodimer complex dynamics and particularly its active sites. In this context, five different PAD2–BAEE systems were designed to analyze the effect of protonation state alterations on the active site and ligand binding. For each system, 1 μ s long independent MD simulations were performed for 10 independent replicas. The results enabled us to determine the protonation state of the catalytic cysteine and the role of the active site histidine and to elucidate the possible initial attack structures based on the near attack conformation approach.

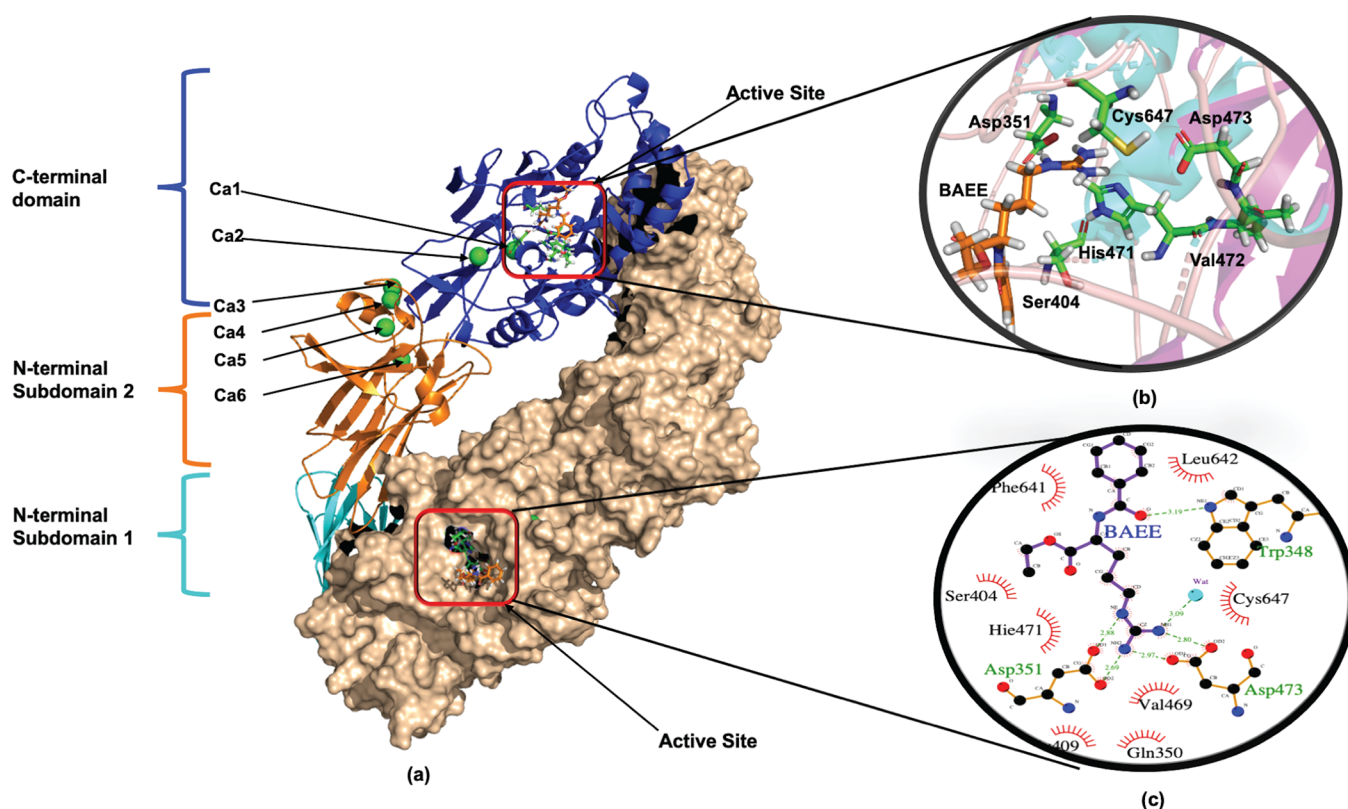


Figure 3. Structure of the PAD2–BAEE complex. (a) Chain A (cartoon) and chain B (molecular surface). (b) Close-up of the initial docking pose. (c) Two-dimensional LigPlot+ representation of the interactions in the active site of BAEE-bound PAD2.

COMPUTATIONAL DETAILS

The crystal structure of protein arginine deiminase 2 (PAD2) in its holo form, with a resolution of 3.02 Å, was obtained from the Protein Data Bank (PDB entry 4N2C).³⁴ The biological assembly of this protein consists of two identical monomers of 690 amino acid residues each. Six Ca²⁺ ions per monomer are included in the system to facilitate enzyme activation.³⁴ The monomers are labeled hereafter as chain A and chain B. This nomenclature will be used to distinguish between the monomers during the MD analysis. *N*- α -Benzoyl-L-arginine ethyl ester (BAEE) with known activity and affinity values from experimental studies by Dreyton et al.³⁶ was selected as the ligand due to its high selectivity for PAD2.

Generating the Enzyme–Substrate Complex. The molecular docking program AutoDock 4.2⁴⁴ was employed to generate docked conformations of BAEE as the ligand and 4N2C as the macromolecule host. The following docking protocol was applied. (i) The box was centered on the sulfur atom of the catalytic cysteine. (ii) The number of grid points was defined at 40 in all three dimensions with the default spacing (0.375 Å). (iii) The receptor was kept rigid while the ligand was allowed to move. (iv) Ten independent runs were performed with a population of 150 and a maximum number of energy evaluations of 2 500 000, with all remaining parameters being kept at their default values. The pose with the lowest binding energy was chosen as the initial ligand configuration and used as an initial structure for the subsequent MD simulation (see Figure 3). It contains many interactions already found for similar structures published for PAD4 (PDB entries 5N0M, 5N0Y, 5N0Z, and 5N1B⁴⁵). The substrate arginine is inserted into the active site cleft, between

His471 and Cys647; Asp473 interacts by hydrogen bonds with the guanidinium fragment as well as Asp351 (see Figure 3c).

Protonation States of Active Side Residues. The initial crystallographic structure from which we have built our initial system has been determined at a resolution of 3.02 Å. With such a resolution, the hydrogen atoms are of course missing in the reported PDB structure. Adding the right number of hydrogen atoms, especially around and inside the active site, represents a crucial step of this molecular modeling. To identify which protonation state(s) should be envisaged for PAD2 molecular dynamics, we have compared experimental pK_a values and suggested mechanisms reported in the literature with Propka3⁴⁶ computations.

First, Dreyton et al. have evaluated the pK_a of Cys647 at ~8.2 by measuring the rates of inactivation of PAD2 as a function of pH.³⁶ From their experimental results, they have suggested a catalytic mechanism that starts with a thiolate that proceeds to a nucleophilic attack on the guanidinium carbon of the substrate arginine. Within this step, a positively charged His471 acts as a general acid, while Asp473 and Asp351 provide electrostatic stabilization.

We report in Table 1 estimations of pK_a values by Propka3 for three different systems: the apoenzyme (PDB entry 4N2C), the holoenzyme (PDB entry 4N2C), and the ligand-bound enzyme (our docking structure). The computationally estimated pK_a value for Cys647 is much higher than the value reported by Dreyton et al.³⁶ From Propka3, this would be due to a strong desolvation shift of the “buried” Cys647 that is not compensated by the charge–charge interaction with the guanidinium fragment upon binding. This indicates that we should envisage for the initial stage of the catalysis mechanism that Cys647 could be in a neutral form instead of a thiolate

Table 1. pK_a Values of the Apoprotein, Holoenzyme, and Complex Forms Estimated by Propka3

	4N20 (apo)	4N2C (holo)	4N2C-BAEE (complex)
Cys647	14.25	16.01	15.38
Asp473	4.26	7.76	4.86
Asp351	3.23	4.10	6.19
His471	6.96	4.70	0.46

form. Adding a ligand inside the active site pocket yields a sharp decrease in the estimated pK_a of His471. This would indicate that the histidine would remain neutral and therefore could not act as a general catalyst. To clarify this discrepancy between the experimental and computational results, we think that multiple protonation states for His471 should be included in our MD models. While Propka3 predicts an acid behavior for Asp351, reinforced by our docking study that suggests a stabilizing role through hydrogen bond interactions with the substrate, Asp473 presents a high pK_a value (>7) in the case of the holoenzyme. In the docking structure, a possible hydrogen bond pattern between Cys647 and Asp473 is present. This suggests that this residue could be involved in proton transfer.

In light of these findings, we have decided to create MD models that examine the different protonation states of the four main residues in the active site. The five different models that we have simulated are presented in Table 2. They are

Table 2. Charges of the Active Site Residues for Each Model System

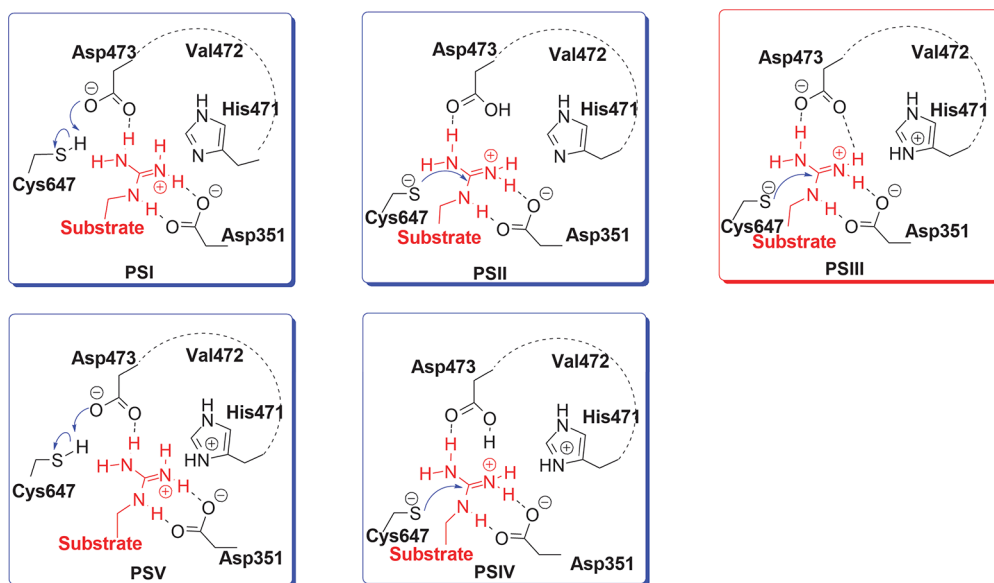
system	Cys647	Asp473	His471	Asp351
PS-I	0	-1	0	-1
PS-II	-1	0	0	-1
PS-III	-1	-1	+1	-1
PS-IV	-1	0	+1	-1
PS-V	0	-1	+1	-1

labeled hereafter PS-I–PS-V. Protonation states for Cys647, His471, and Asp473 vary, while Asp351 is always considered to

be in a deprotonated form. Figure 4 depicts the difference in active site protonation states for the different models of this study. PS-I represents an initial stage where Cys647 could be activated by Asp473. In PS-II, the initial stage of the catalytic reaction is an already deprotonated Cys647 with a neutral Asp473. PS-III corresponds to a papain-like initial stage, as suggested by Dreyton et al.,³⁶ where Cys647 is deprotonated and His471 is positively charged. PS-IV is similar to PS-II, but His471 is now doubly protonated. Finally, in PS-V, both Cys647 and His471 are protonated; therefore, only Asp473 can act as a proton acceptor. Except for these four residues, the protonation states of all remaining residues are identical for all MD models.

Molecular Dynamic Simulations. Each system was built with the `tLeap` module of AMBER18.⁴⁷ From the initial complex structure generated with AutoDock, the systems were protonated according to their respective definitions (e.g., PS-I, PS-II, etc.). Proteins and water molecules were described using the ff14SB29⁴⁸ and TIP3P⁴⁹ force fields, respectively. Periodic boundary conditions were applied by filling an isometric truncated octahedral box containing one complex system (a PAD2 dimer and two ligands, one for each active site) with water molecules. The minimum distance between any protein atom and the edge of the periodic box was set to 18 Å. The neutrality of the systems was imposed by adding the right number of chloride ions depending on the given system. For each system, the total number of solute atoms was around 20 500 atoms, while the number of water molecules was approximately 86 500. Energy minimization and MD simulation runs were performed using the GPU-supported `pmemd` module in AMBER18.⁴⁷ The particle mesh Ewald summation technique was used with the default 8 Å cutoff, and the SHAKE algorithm⁵⁰ for bonds involving hydrogen atoms was applied in addition to hydrogen mass repartitioning (HMR).⁵¹

The samples were equilibrated in three consecutive steps. First, 100 ps of molecular dynamics was performed in the *NPT* ensemble at a temperature of 10 K with harmonic positional restraints on the heavy atoms of the protein–ligand parts (50 kcal mol⁻¹ Å⁻²). The Andersen thermostat was used, and

**Figure 4.** Constructed models based on the different protonation states of the active site residues. Note that PS-III corresponds to the initial stage of the catalytic mechanism proposed by Dreyton et al.³⁶

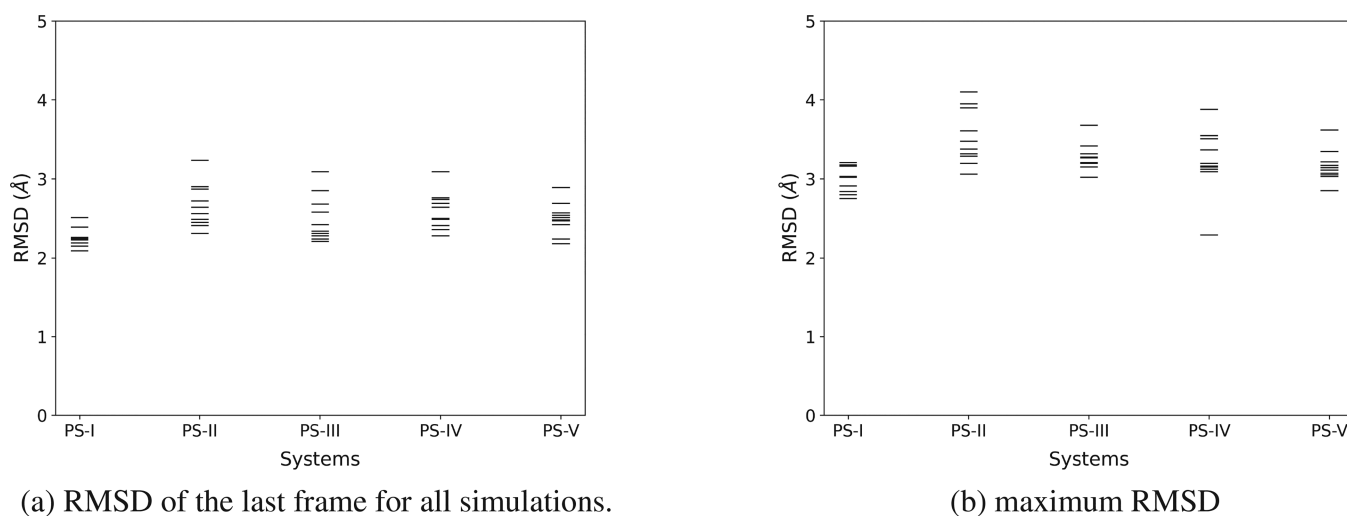


Figure 5. (a) RMSDs of the last frame and (b) maximum RMSDs for all 10 simulations and all five systems.

velocities were rescaled at 10 K every 10 steps to ensure a rapid decrease in the potential energy to an approximate local minimum. Then the systems were linearly heated to 300 K in the *NPT* ensemble during 2100 ps of molecular dynamics while the harmonic positional restraints on the heavy atoms of the protein–ligand parts were maintained. The increase in temperature was ensured again by using the Andersen thermostat with a velocity rescaling every 10 steps to the appropriate temperature. For this first two steps, constant-pressure dynamics at 1 bar was applied with isotropic position scaling and controlled by a Monte Carlo barostat with volume change attempts every 100 steps. The time step was 0.002 fs. Finally, the harmonic potential restraints were linearly lifted during a 50 ns molecular dynamics, the value of the restraints being decreased by $1 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ every nanosecond. This molecular dynamics and the production molecular dynamics were performed using the *NVT* ensemble, and the temperature was controlled by a Langevin thermostat with a collision frequency γ of 1.0 ps^{-1} . The time step was 0.004 fs. Production runs of $1 \mu\text{s}$ were carried without restraints in the *NVT* ensemble, and snapshots were saved every 40 ps (hence 25 000 snapshots per MD simulation).

For each of the five different constructed MD systems within the scope of this study (PS-I–PS-V), 10 independent simulations were performed using different initial random seeds. The 50 replica simulations, each of $1 \mu\text{s}$, represent a total simulation time of $50 \mu\text{s}$ as well as 100 different active site trajectories to analyze.

Analysis Methods. All analyses were performed using the *cpptraj* module of AMBER18.⁴⁷ In addition to various distance, angle, and dihedral analyses, root-mean-square deviations (RMSDs) were computed to assess the stability of the simulations. The interactions between Cys647 and water molecules were computed using radial distribution functions (RDFs). The solvent accessible surfaces of the protein itself (*P*), the ligand itself (*S*), and the protein–ligand complex (PS) were evaluated using the *surf* command by incrementing the van der Waals radii by 1.4 Å. The surface of contact (SC) between the protein and the ligand was simply defined by

$$\text{SC} = (P + S - \text{PS})/2 \quad (1)$$

All scripts and extended data analysis can be found in the [Supporting Information](#).

RESULTS AND DISCUSSION

Ten multiple $1 \mu\text{s}$ MD simulations using random initial velocities were applied to each system to enhance conformational samplings of the protein–ligand complexes. The presence of different types of interactions and movements of the protein domains obtained from the 10 replicas for each system enabled us to engage in a deeper discussion related to the structure and dynamics of these protein–ligand complexes.

Stability of the Simulated Systems. A first picture of the stability of the systems can be provided by the RMSDs of the atomic positions of the backbone atoms (C_α) for the dimer during the course of the $1 \mu\text{s}$ simulations with respect to the reference crystal structure. All RMSD variations for all simulations are given in the [Supporting Information](#). As a summary, we represent in [Figure 5](#) the maximum RMSD as well as the RMSD of each last frame of the 10 independent MD runs for the five considered systems. All observed RMSDs range between 2 and 3 Å. This indicates the good stability of the systems during the MD runs.

Is the Active Site Cysteine Preactivated? Two different mechanisms were proposed for the conversion of a peptidyl-arginine into peptidyl-citrulline by different isoforms of PADs. The experimental and theoretical studies have revealed that the active site of PAD4 involves a Cys645 (Cys647 in PAD2) in thiolate form in the apo and holo structures stabilized by a protonated His471 just like the thiolate–imidazolium pair in papain.⁵² The reaction mechanism is denoted as reverse protonation.⁵³ In contrast, a substrate-assisted mechanism has been proposed in the case of PAD2. Experimental studies suggest that the pK_a of the active site cysteine residue (Cys647) is shifted after the binding of the 2-chloroacetamide molecule that is analogous to the positively charged substrate guanidinium group.³⁶ Meanwhile, the active site histidine (His471) is suggested to be protonated. Variations on the protonation states of an enzyme active site have consequences on its possible reaction mechanisms; this is why it is important, before any further mechanistical studies, to clearly assess the protonation states of the active site residues.^{54–57}

To elucidate the structure of the Michaelis complex, five different systems were designed on the basis varying protonation states. As presented in [Table 2](#), for the PS-II–PS-IV systems, the catalytic cysteine (Cys647) is deprotonated.

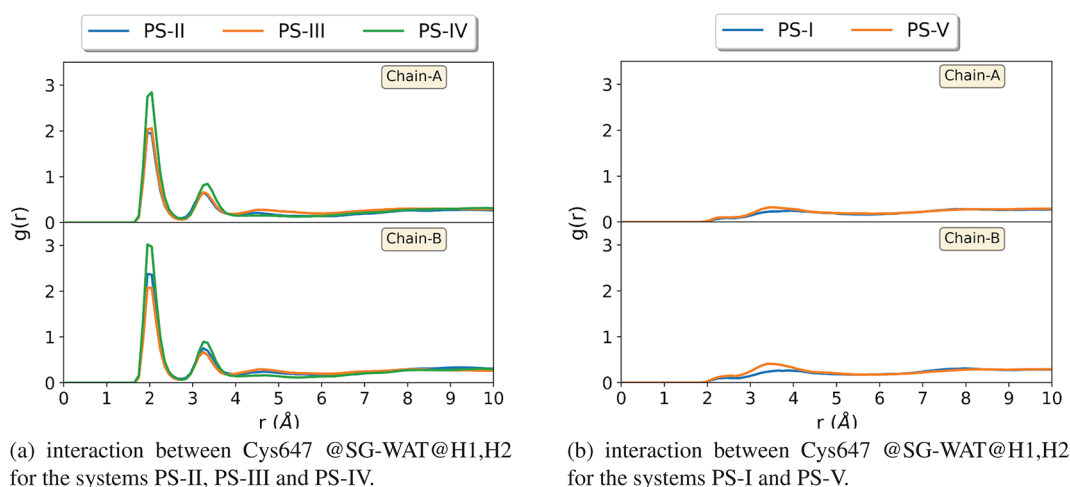


Figure 6. Radial distribution function, $g(r)$, between the sulfur atom of Cys647 (Cys647@SG) and water hydrogen atoms for all five considered systems.

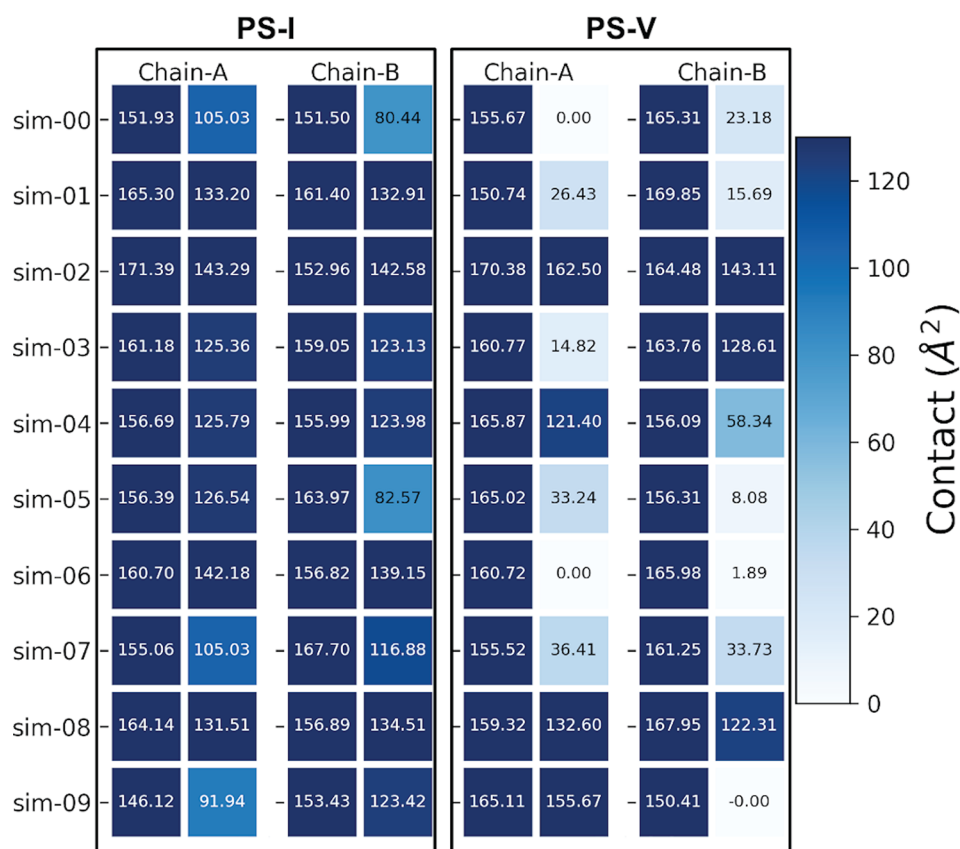


Figure 7. Surface contact area, in square angstroms, between the BAEE substrate and the active site residues for both monomer chains (A and B) of the PS-I (left) and PS-V (right) systems. For each chain, the left column represents the maximum surface contact area and the right column represents the surface contact area of the last MD frame.

nated, whereas in PS-I and PS-V, it must be activated through a proton transfer to Asp473. As a first analysis, we have investigated the interactions between the Cys647 side chain and solvent water molecules for all PAD2–BAEE samples by RDF analysis. According to the radial distribution function [$g(r)$] patterns depicted in Figure 6, the probability of finding a water molecule at a certain distance r from the S_γ atom varies depending on the protonation state of Cys647. A strong peak around 2 Å indicates a hydrogen bond interaction between the sulfur atom and water. Moreover, the average number of water

molecules around this sulfur atom during the simulations can be estimated by the integration of RDF curves until the first minima. The $g(r)$ functions in systems having a deprotonated cysteine (systems PS-II–PS-IV) are depicted in Figure 6a. The estimated numbers of water molecules obtained by the integrations of the $g(r)$ curves are found to be approximately two molecules for PS-II and 1.5 molecules for PS-III and PS-IV. This shows that the negatively charged Cys647 side chains are surrounded by water molecules that prevent the attack of the sulfur atom on the carbon atom of the substrate

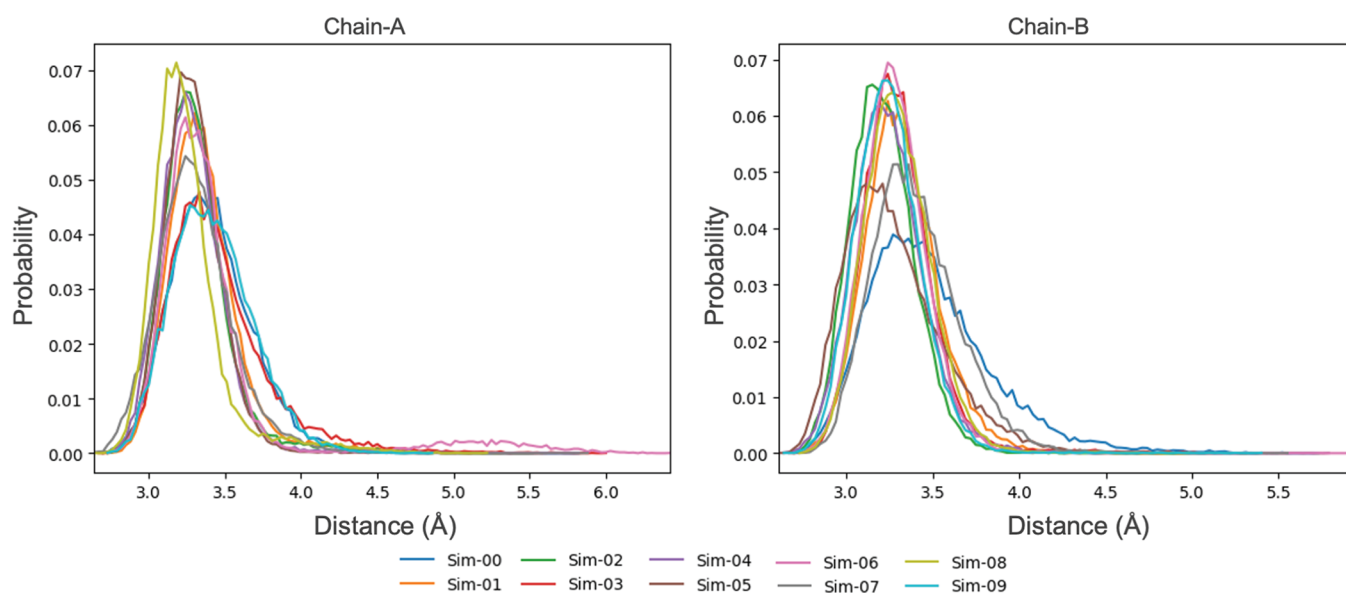


Figure 8. Minimum distance probabilities between His471 and the substrate amino groups for PS-I system subunits A and B.

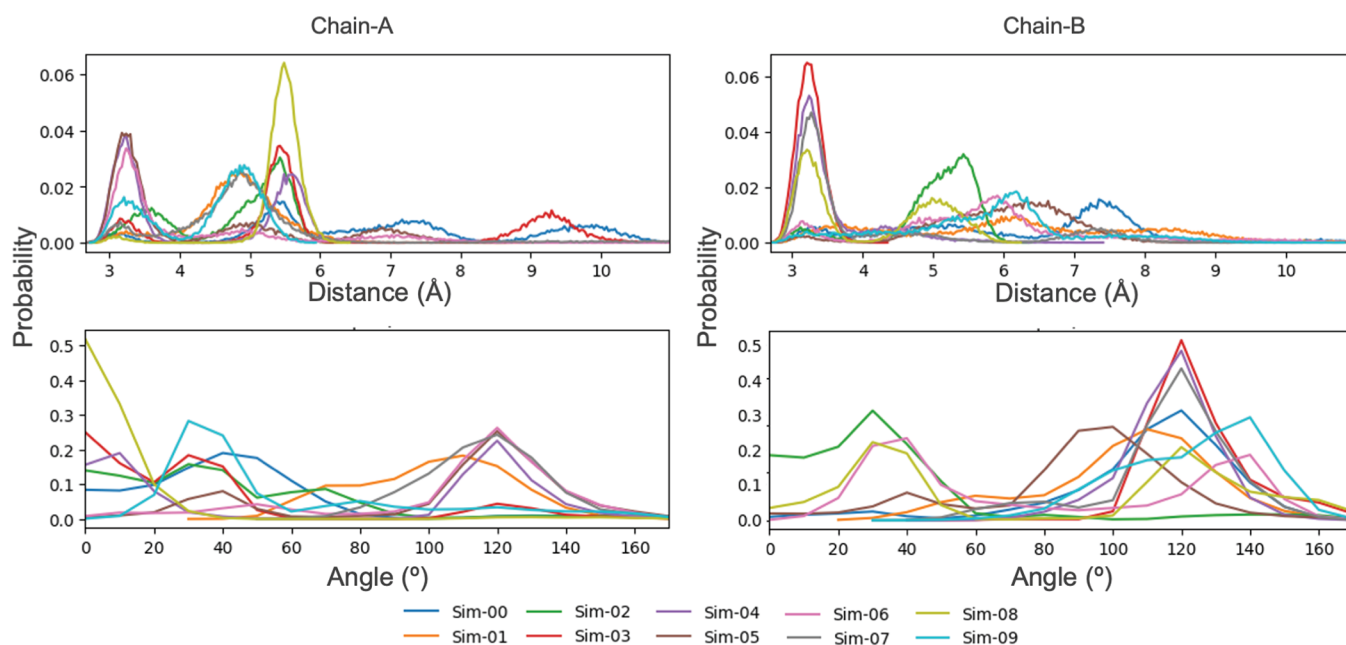


Figure 9. Minimum distance probabilities between His471 and the substrate amino groups (top) and angle probabilities formed by His471@ND1-His471@HD1-Arg623@NH (bottom) for the PS-V system subunits A (left) and B (right).

guanidinium. On the contrary, the RDF patterns given in Figure 6b for the PS-I and PS-V systems show no typical hydrogen bond peak. This indicates that when the Cys647 side chain is neutral, the number of water molecules is much smaller in the active site and they do not form hydrogen bond interactions with Cys647. Integrations of $g(r)$ until 2.75 Å for both systems indicate that only ~ 0.3 water molecule is in the vicinity of the sulfur atom.

These results can be interpreted in two ways. First, the presence of strong water interactions between the deprotonated Cys647 and approximately two water molecules in average for systems PS-II–PS-IV prevents Cys647 from attacking the substrate. These systems can therefore be considered as nonreactive. Second, the usage of a molecular mechanics force field here does not allow proton transfer. It

seems that the strong hydrogen bond interaction could be seen as the mark of a potential proton transfer between Cys647 and solvent water, which would indicate that the pK_a of Cys647 is much larger than 7 (i.e., the stable form of Cys647 tends to a thiol form rather than a thiolate one).

Role of Histidine. Radial distribution function analysis revealed that a neutral Cys647 should be favored in the PAD2 system for the Michaelis complex structure. Hence, we will rule out systems PS-II–PS-IV in the rest of this analysis and focus now on only PS-I and PS-V. The difference between the two latter systems comes from His471. In PS-I, His471 is neutral, while it is protonated in PS-V.

To gain further insight into the differences between the two systems, substrate–active site surface contact calculations, trajectory analysis, and a linear interaction energy (LIE)

analysis between the His471 residue and the substrate molecule have been performed.

Visual inspections of the MD reveal some major differences between the two types of systems in the behavior of the substrate. To illustrate these differences, we have computed the contact surface area using eq 1 with P being the solvent accessible surface (SAS) around the set of the four active site residues (taken alone), S the SAS of the BAEE substrate, and PS the SAS of the complex.

The calculated values for the surface contact areas for the two systems, 10 simulations, and two monomers are illustrated in Figure 7. The higher the number (and the more blue), the larger the contact area between the substrate and the active site. For each chain, the first column represents the average contact while the second column represents the surface contact at the end of the microsecond simulations. The striking point here is that, for PS-V, the surface contact area is mainly close to zero at the end of the simulations. This means that the substrate has left the active site and bears no contact with it anymore. Roughly, a surface contact area of $<60 \text{ \AA}^2$ can be considered as a substrate that does not maintain contact with the main atoms of the catalytic site, while a value of $>100 \text{ \AA}^2$ can be interpreted as a good attachment of the substrate inside the active site. In PS-I, large surface contact areas are maintained for all simulations and for both chains. In contrast, with a protonated His471 in PS-V, the contact is not maintained for a majority of the complexes. We further extrapolate that for the few PS-V simulations where the substrate stays inside the active site pocket, there would be a great chance that the substrate would depart if the MD were to be extended beyond the current microsecond length. It is noteworthy that, as opposed to the case for PS-V, the substrate remains stuck to the active site for PS-II–PS-IV (see Figure S7), but the solvation of the negative Cys647 diminishes the surface contact area between the host and the ligand.

Another way to illustrate the differences between PS-I and PS-V is to check the distance between the substrate and His471. The minimum distance probabilities between the arginine side chain of the substrate, BAEE, and the imidazole ring of His471 are reported for PS-I and PS-V in Figures 8 and 9, respectively. In the case of a deprotonated His471, the interaction is stable with a peak at $\sim 3.3\text{--}3.4 \text{ \AA}$ for both monomers. When His471 is protonated, large distances between the imidazole ring and the substrate arginine can be observed, and this confirms a departure of the substrate from the active site during most of the MD trajectories.

Also, the positioning of His471 in PS-V does not favor a proton transfer to the substrate as illustrated in Figure 9 by the computation of the angle between the ND1 and HD1 atoms of His471 and the closest NH atom of the substrate arginine. An angle of $\sim 180^\circ$ would favor a proper proton transfer from the imidazole ring of His471 to the guanidinium group of BAEE. For the few MD runs in which the substrate remains inside the active site, this angle has a maximum probability of $\sim 120^\circ$, away from being at the appropriate value to facilitate the proton transfer.

To understand the differences between PS-I and PS-V, and especially to provide a reason for the departure of the substrate when His471 is protonated, we have computed the LIEs between the His471 side chain and the substrate arginine side chain for all 10 independent replicas of the PS-I and PS-V systems. The calculated LIE values are listed in Table 3. A higher affinity between His471 and the substrate in all copies

Table 3. Average Linear Interaction Energies (LIEs) in Kilocalories per Mole between His471 and the BAEE Substrate for Chains A and B of PS-I and PS-V

	PS-I		PS-V	
	chain A	chain B	chain A	chain B
Sim-00	−6.8 (2.1)	−6.3 (2.4)	21.9 (17.5)	22.4 (13.5)
Sim-01	−9.4 (2.6)	−9.5 (1.9)	28.0 (9.0)	20.7 (15.0)
Sim-02	−9.5 (2.4)	−10.3 (1.3)	44.9 (4.3)	41.1 (5.3)
Sim-03	−8.2 (3.2)	−9.7 (2.1)	34.5 (16.5)	49.3 (3.6)
Sim-04	−9.9 (2.1)	−9.9 (1.9)	46.8 (4.2)	45.1 (9.3)
Sim-05	−10.0 (1.6)	−8.0 (3.1)	43.5 (11.0)	20.9 (12.1)
Sim-06	−8.7 (4.3)	−10.0 (1.5)	31.0 (20.5)	26.3 (13.9)
Sim-07	−9.0 (2.8)	−9.7 (1.7)	26.3 (10.3)	42.2 (13.9)
Sim-08	−9.1 (3.3)	−10.0 (1.4)	43.3 (3.3)	47.0 (4.5)
Sim-09	−6.3 (2.2)	−10.3 (1.5)	44.6 (5.7)	21.9 (10.3)

of the PS-I system facilitates the positioning of the substrate in the active site. In contrast, the presence of a repulsive interaction between His471 and the substrate molecule can be seen as the primary reason for the escape of the substrate from the active site of the PS-V system.

The obtained results suggest that the protonation state of His471 plays a critical role in the proper positioning of the substrate molecule in the active site. Due to the charge distribution inside the active site that creates an electrostatic repulsion between the substrate and His471 within the PS-V system, the substrate molecule is not stable in the active site during most of the simulations. Thus, the PS-V system has also been considered to represent an inappropriate protonation state for PAD2 reactivity and has been eliminated from further analysis.

Near Attack Conformers. The remaining system, PS-I, is the only system that stabilizes clearly the substrate inside the active site pocket. In this system, BAEE is sandwiched between the neutral catalytic cysteine, Cys647, and the histidine residue, His471, in a neutral form also. This histidine residue is too far from Cys647 and is blocked by BAEE to serve as a general base to activate the nucleophilicity of Cys647. RDF analysis (Figure S17) between the carboxylate oxygen atoms of Asp473 and the hydrogen atoms of the guanidinium fragment of BAEE indicates that, on average, Asp473 makes 0.85 hydrogen bond with BAEE (integration of the RDF until the first minimum at 2.3 \AA). Therefore, Asp473 is properly positioned and partially free to abstract, directly or indirectly, the proton from Cys647. Detailed distance analysis shows that the amino groups of the substrate molecule form a stable salt bridge with the embedded Asp351 carboxylate ion during the MD simulations. With this interaction, the guanidinium substrate fragment is properly positioned for Cys647 to attack (Supporting Information).

Possible initial reaction mechanisms from PS-I are depicted in Figure 10. The first possible mechanism starts with a direct proton transfer from Cys647 to Asp473 (Figure 10a). The second possible mechanism involves a water-assisted proton transfer to activate Cys647 by Asp473 (Figure 10b). A direct concerted mechanism can be imagined between Cys647 and BAEE. This is a third possible mechanism (Figure 10c). Finally, the attack of Cys647 on BAEE without the assistance of Asp473 could be water-assisted (Figure 10d). To check which mechanisms could be the most plausible, we can use the concept of near attack conformers (NACs) that states that

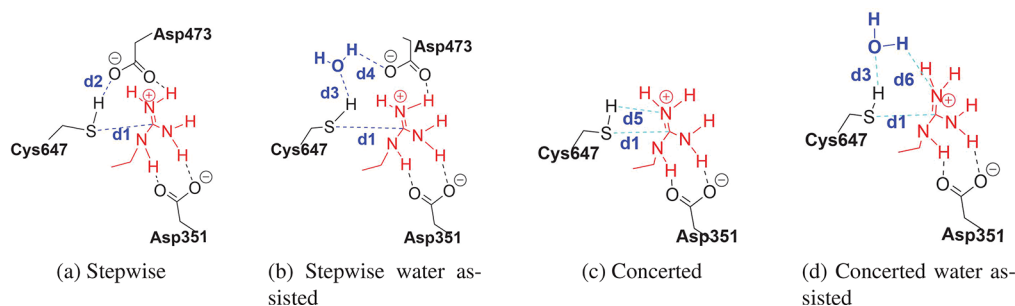


Figure 10. NAC reaction coordinate definitions.

ground state structures tend to adopt favorable conformations that can convert easily to transition states.^{58,59}

To analyze the probabilities of having NAC structures in our simulations, we have defined six critical distances based on the four possible initial attacks for the reaction pathway (Figure 10). The active site cysteine, Cys647, should perform a nucleophilic attack to the substrate guanidinium. The distance between Cys647@SG and guanidinium@CZ for the nucleophilic attack is defined as d_1 . According to our ongoing QM studies (data not shown), d_1 should be <3.5 Å to reflect a van der Waals contact between the sulfur atom and the guanidinium π orbitals. Distance d_2 reflects the direct proton abstraction of Cys647 by Asp473. Distances d_3 and d_4 defined a water bridge between Cys647 and Asp473 that could facilitate a water-assisted mechanism. These three distances are critical for proton transfer, and a threshold of 2.5 Å was chosen to mimic hydrogen bond character. For the concerted mechanism, an additional distance, d_5 , was defined between the sulfur-bound hydrogen and the closest of the guanidinium nitrogen atoms. In conjunction with a proper d_1 distance, a d_5 distance of <2.5 Å is a synonym of NAC for the concerted mechanism. Finally, in the case of a water-assisted concerted mechanism, distance d_6 is defined between the closest hydrogen of the assisting water molecule and the guanidinium nitrogen atoms with a threshold of 2.5 Å.

The distance analysis was carried out on 500 000 frames starting with the distance criterion d_1 , which is common to all mechanisms. The results are given as the number and percentage of NAC structures in Figure 11; 18.2% of the frames in PS-I contain a short distance between the Cys647 sulfur atom at SG and the center carbon of the guanidinium fragment at CZ. These frames are candidates for a possible initiation of a reaction between Cys647 and BAEE. From these 91 180 frames, 27 042 (5.41%) are compatible with a concerted mechanism. Then, 4651 frames can be regarded as the start of a stepwise mechanism. For the water-assisted mechanism, only 684 and 188 frames are found for the stepwise and concerted mechanisms, respectively.

While NACs have been found for all mechanisms, those who do not involve water assistance are the most prominent. The concerted mechanism that consists of the attack of the sulfur atom on the guanidinium carbon while the sulfur hydrogen is transferred to the nitrogen atom has probably a high energetic barrier as found in similar reaction mechanisms studied by quantum mechanics.^{60–62} Therefore, it would seem that the stepwise mechanism, which involves an activation of Cys647 by Asp473, is the most probable initial reaction of the catalytic mechanism in PAD2 (Figure 12).

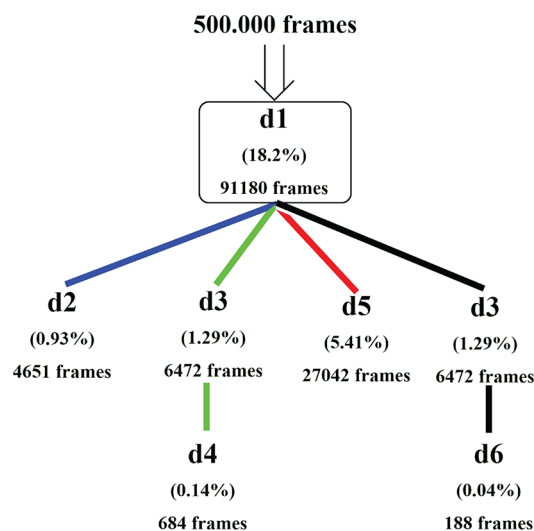


Figure 11. Percentage and number of selected NAC structures based on the predefined distance criteria.

CONCLUSIONS

Five different systems were defined on the basis of the different possible protonation states of the active site residues: Cys647, His471, Asp351, and Asp473. To describe the conformational space as broadly as possible, 10 independent MD runs with a length of 1 μ s were performed for each defined system. A total of 50 independent MD runs of 1 μ s were carried out to provide a deeper understanding of the structural changes within the PAD2 homodimer complex.

Radial distribution functions of the interactions of Cys647 with solvent water clearly show that for the three systems containing a negatively charged Cys647 (PS-II–PS-IV), two water molecules on average form strong hydrogen bonds with the catalytic cysteine and prevent the latter from attacking the substrate. This demonstrates that Cys647 must be protonated during the initial step of the catalytic reaction.

PS-I and PS-V differ from the protonation state of His471. In PS-I, His471 is neutral, while in PS-V, the imidazole ring is positively charged. The stabilities of the substrate during the MD simulations were very different. For many PS-V trajectories, the substrate lost surface contact with the active site residues. In contrast, the guanidinium fragment of BAEE maintained a stable interaction with the neutral His471 for all PS-I simulations. The results have been emphasized by LIE analysis that shows that the interactions between the substrate and His471 are repulsive for PS-V and attractive for PS-I. We interpret this findings by stating that His471 should be neutral

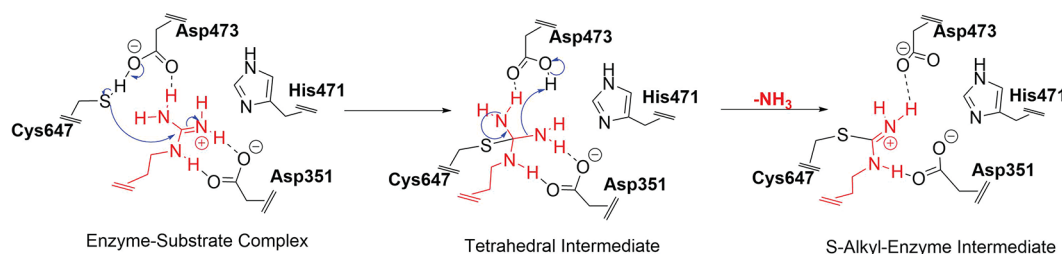


Figure 12. Proposed reaction mechanism for the formation of the S-alkylation intermediate in PAD2.

to accommodate a PAD2 substrate and that PS-I represents the only valid protonation state to start the catalytic reaction.

From the PS-I state, several reaction mechanisms can be envisaged. We have used NAC to evaluate the different probabilities of finding conformations that could initiate the catalytic reaction. Overall, we have found frames compatible with all possible mechanisms. Further quantum mechanical or QM/MM works will have to determine the different energetic barriers of every possible mechanisms starting from neutral Cys647 and His471. However, given the presence in the vicinity of Cys647 of an aspartate residue, Asp473, that could play the role of a general base to activate the catalytic cysteine, we hypothesize that the PAD2 catalytic mechanism should be very similar to the proposed reaction mechanism depicted in Figure 12.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.biochem.2c00158>.

Scripts and figures for the parametrization of the BAE ligand molecule, RMSD analysis, RDF analysis, surface contact computations, LIE analysis, NAC analysis, RDF analysis between Asp473 and the substrate, and analysis of the distance between Asp351 and the substrate (PDF)

Coordinates of the last frames of all 10 simulations and five systems with the 100 closest water molecules from each active site (ZIP)

■ AUTHOR INFORMATION

Corresponding Authors

Gerald Monard – *Université de Lorraine, CNRS, LPCT, F-54000 Nancy, France*; orcid.org/0000-0002-3378-3481; Email: gerald.monard@univ-lorraine.fr

Fethiye Aylin Sungur – *Informatics Institute, Computational Science and Engineering, Istanbul Technical University, TR-34469 Istanbul, Turkey*; orcid.org/0000-0003-3245-742X; Email: aylin.sungur@itu.edu.tr

Author

Erdem Cicek – *Informatics Institute, Computational Science and Engineering, Istanbul Technical University, TR-34469 Istanbul, Turkey*; orcid.org/0000-0002-9391-7496

Complete contact information is available at:

<https://pubs.acs.org/doi/10.1021/acs.biochem.2c00158>

Funding

This work was supported by the Tübitak (Project 115Z860). E.C., F.A.S., and G.M. are grateful to Campus France and Tubitak for travel grants.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

High Performance Computing resources were partially provided by the EXPLOR centre hosted by the Université de Lorraine and the Centre de Calcul ROMEO of the Université de Reims Champagne-Ardenne.

■ REFERENCES

- (1) Luger, K.; Richmond, T. J. The Histone Tails of the Nucleosome. *Current Opinion in Genetics & Development* **1998**, *8*, 140–146.
- (2) Kornberg, R. D.; Lorch, Y. *Twenty-Five Years of the Nucleosome. Fundamental Particle of the Eukaryote Chromosome*; 1999.
- (3) Strahl, B. D.; Allis, C. D. The Language of Covalent Histone Modifications. *Nature* **2000**, *403*, 41–45.
- (4) Cavalieri, V. The Expanding Constellation of Histone Post-Translational Modifications in the Epigenetic Landscape. *Genes* **2021**, *12*, 1596.
- (5) Zee, B. M.; Levin, R. S.; Dimaggio, P. A.; Garcia, B. A. Global Turnover of Histone Post-Translational Modifications and Variants in Human Cells. *Epigenet. Chromatin* **2010**, *3*, 22.
- (6) Thompson, P. R.; Fast, W. Histone Citrullination by Protein Arginine Deiminase: Is Arginine Methylation a Green Light or a Roadblock? *ACS Chem. Biol.* **2006**, *1*, 433–441.
- (7) Bicker, K. L.; Thompson, P. R. The Protein Arginine Deiminases: Structure, Function, Inhibition, and Disease. *Biopolymers* **2013**, *99*, 155–163.
- (8) Mondal, S.; Thompson, P. R. Chemical Biology of Protein Citrullination by the Protein A Arginine Deiminases. *Curr. Opin. Chem. Biol.* **2021**, *63*, 19–27.
- (9) Arita, K.; Hashimoto, H.; Shimizu, T.; Nakashima, K.; Yamada, M.; Sato, M. Structural Basis For Ca(2+)-Induced Activation of Human PAD4. *Nat. Struct. Mol. Biol.* **2004**, *11*, 777–783.
- (10) Christophorou, M. A.; Castelo-Branco, G.; Halley-Stott, R. P.; Oliveira, C. S.; Loos, R.; Radziszheuskaya, A.; Mowen, K. A.; Bertone, P.; Silva, J. C.; Zernicka-Goetz, M.; Nielsen, M. L.; Gurdon, J. B.; Kouzarides, T. Citrullination Regulates Pluripotency and Histone H1 Binding to Chromatin. *Nature* **2014**, *507*, 104–108.
- (11) Slade, D. J.; Subramanian, V.; Thompson, P. R. Pluripotency: Citrullination Unravels Stem Cells. *Nat. Chem. Biol.* **2014**, *10*, 327–328.
- (12) Mondal, S.; Thompson, P. R. Protein Arginine Deiminases (PADs): Biochemistry and Chemical Biology of Protein Citrullination. *Accounts of chemical research* **2019**, *52*, 818–832.
- (13) Vossenaar, E. R.; Zendman, A. J. W.; Van Venrooij, W. J.; Puij, G. J. M. PAD, a Growing Family of Citrullinating Enzymes: Genes, Features and Involvement in Disease. *BioEssays* **2003**, *25*, 1106–1118.
- (14) Ishida-Yamamoto, A.; Senshu, T.; Eady, R. A.; Takahashi, H.; Shimizu, H.; Akiyama, M.; Iizuka, H. Sequential Reorganization of Cornified Cell Keratin Filaments Involving Filaggrin-Mediated Compaction and Keratin 1 Deimination. *Journal of Investigative Dermatology* **2002**, *118*, 282–287.

- (15) Nachat, R.; Méchin, M. C.; Takahara, H.; Chavanas, S.; Charveron, M.; Serre, G.; Simon, M. Peptidylarginine Deiminase Isoforms 1–3 are Expressed in the Epidermis and Involved in the Deimination of K1 and Filaggrin. *Journal of Investigative Dermatology* **2005**, *124*, 384–393.
- (16) Musse, A. A.; Li, Z.; Ackerley, C. A.; Bienzle, D.; Lei, H.; Poma, R.; Harauz, G.; Moscarello, M. A.; Mastronardi, F. G. Peptidylarginine Deiminase 2 (PAD2) Overexpression in Transgenic Mice Leads to Myelin Loss in the Central Nervous System. *Disease models & mechanisms* **2008**, *1*, 229–240.
- (17) Lamensa, J. W. E.; Moscarello, M. A. Deimination of Human Myelin Basic Protein by a Peptidylarginine Deiminase from Bovine Brain. *Journal of Neurochemistry* **1993**, *61*, 987–996.
- (18) Steinert, P. M.; Parry, D. A.; Marekov, L. N. Trichohyalin Mechanically Strengthens the Hair Follicle: Multiple Cross-Bridging Roles in the Inner Root Sheath. *J. Biol. Chem.* **2003**, *278*, 41409–41419.
- (19) Cuthbert, G. L.; Daujat, S.; Snowden, A. W.; Erdjument-Bromage, H.; Hagiwara, T.; Yamada, M.; Schneider, R.; Gregory, P. D.; Tempst, P.; Bannister, A. J.; Kouzarides, T. Histone Deimination Antagonizes Arginine Methylation. *Cell* **2004**, *118*, 545–553.
- (20) Zhang, X.; Gamble, M. J.; Stadler, S.; Cherrington, B. D.; Causey, C. P.; Thompson, P. R.; Roberson, M. S.; Kraus, W. L.; Coonrod, S. A. Genome-Wide Analysis Reveals PADI4 Cooperates with Elk-1 to Activate C-Fos Expression in Breast Cancer Cells. *PLoS Genet.* **2011**, *7*, e1002112.
- (21) Li, P.; Li, M.; Lindberg, M. R.; Kennett, M. J.; Xiong, N.; Wang, Y. PAD4 is Essential for Antibacterial Innate Immunity Mediated by Neutrophil Extracellular Traps. *Journal of Experimental Medicine* **2010**, *207*, 1853–1862.
- (22) Horibata, S.; Coonrod, S. A.; Cherrington, B. D. Role for Peptidylarginine Deiminase Enzymes in Disease and Female Reproduction. *Journal of Reproduction and Development* **2012**, *58*, 274–282.
- (23) Mohanan, S.; Cherrington, B. D.; Horibata, S.; McElwee, J. L.; Thompson, P. R.; Coonrod, S. A. Potential Role of Peptidylarginine Deiminase Enzymes and Protein Citrullination in Cancer Pathogenesis. *Biochem. Res. Int.* **2012**, *2012*, 1.
- (24) Lewis, H. D.; Liddle, J.; Coote, J. E.; Atkinson, S. J.; Barker, M. D.; Bax, D.; Bicker, K. L.; Bingham, R. P.; Campbell, M.; Chen, Y. H.; Chung, C.-w.; Craggs, P. D.; Davis, R. P.; Eberhard, D.; Joberty, G.; Lind, K. E.; Locke, K.; Maller, C.; Martinod, K.; Patten, C.; Polyakova, O.; Rise, C. E.; Rüdiger, M.; Sheppard, R. J.; Slade, D. J.; Thomas, P.; Thorpe, J.; Yao, G.; Drewes, G.; Wagner, D. D.; Thompson, P. R.; Prinjha, R. K.; Wilson, D. M. Inhibition of PAD4 Activity is Sufficient to Disrupt Mouse and Human NET Formation. *Nat. Chem. Biol.* **2015**, *11*, 189–191.
- (25) Knuckley, B.; Causey, C. P.; Jones, J. E.; Bhatia, M.; Dreyton, C. J.; Osborne, T. C.; Takahara, H.; Thompson, P. R. Substrate Specificity and Kinetic Studies of PADs 1, 3, and 4 Identify Potent and Selective Inhibitors of Protein Arginine Deiminase 3. *Biochemistry* **2010**, *49*, 4852–4863.
- (26) Lewallen, D. M.; Bicker, K. L.; Madoux, F.; Chase, P.; Anguish, L.; Coonrod, S.; Hodder, P.; Thompson, P. R. A FluoPol-ABPP PAD2 High-Throughput Screen Identifies the First Calcium Site Inhibitor Targeting the PADs. *ACS Chem. Biol.* **2014**, *9*, 913–921.
- (27) Willis, V. C.; Banda, N. K.; Cordova, K. N.; Chandra, P. E.; Robinson, W. H.; Cooper, D. C.; Lugo, D.; Mehta, G.; Taylor, S.; Tak, P. P.; Prinjha, R. K.; Lewis, H. D.; Holers, V. M. Protein Arginine Deiminase 4 Inhibition is Sufficient For the Amelioration of Collagen-Induced Arthritis. *Clin. Exp. Immunol.* **2017**, *188*, 263–274.
- (28) Knight, J. S.; Subramanian, V.; O'Dell, A. A.; Yalavarthi, S.; Zhao, W.; Smith, C. K.; Hodgins, J. B.; Thompson, P. R.; Kaplan, M. J. Peptidylarginine Deiminase Inhibition Disrupts NET Formation and Protects Against Kidney, Skin and Vascular Disease in Lupus-Prone MRL/lpr Mice. *Annals of the Rheumatic Diseases* **2015**, *74*, 2199–2206.
- (29) Lange, S.; Gögel, S.; Leung, K. Y.; Vernay, B.; Nicholas, A. P.; Causey, C. P.; Thompson, P. R.; Greene, N. D.; Ferretti, P. Protein Deiminases: New Players in the Developmentally Regulated Loss of Neural Regenerative Ability. *Dev. Biol.* **2011**, *355*, 205–214.
- (30) Cherrington, B. D.; Zhang, X.; McElwee, J. L.; Morency, E.; Anguish, L. J.; Coonrod, S. A. Potential Role for PAD2 in Gene Regulation in Breast Cancer Cells. *PLoS one* **2012**, *7*, e41242.
- (31) Holmes, C. L.; Shim, D.; Kernien, J.; Johnson, C. J.; Nett, J. E.; Shelef, M. A. Insight Into Neutrophil Extracellular Traps Through Systematic Evaluation of Citrullination and Peptidylarginine Deiminases. *J. Immunol. Res.* **2019**, *2019*, 1.
- (32) Zhang, X.; Bolt, M.; Guertin, M. J.; Chen, W.; Zhang, S.; Cherrington, B. D.; Slade, D. J.; Dreyton, C. J.; Subramanian, V.; Bicker, K. L.; et al. Peptidylarginine Deiminase 2-Catalyzed Histone H3 Arginine 26 Citrullination Facilitates Estrogen Receptor Alpha Target Gene Activation. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 13331–13336.
- (33) Saijo, S.; Nagai, A.; Kinjo, S.; Mashimo, R.; Akimoto, M.; Kizawa, K.; Yabe-Wada, T.; Shimizu, N.; Takahara, H.; Unno, M. Monomeric Form of Peptidylarginine Deiminase Type I Revealed by X-ray Crystallography and Small-Angle X-ray Scattering. *J. Mol. Biol.* **2016**, *428*, 3058–73.
- (34) Slade, D. J.; Fang, P.; Dreyton, C. J.; Zhang, Y.; Fuhrmann, J.; Rempel, D.; Bax, B. D.; Coonrod, S. A.; Lewis, H. D.; Guo, M.; Gross, M. L.; Thompson, P. R. Protein Arginine Deiminase 2 Binds Calcium in an Ordered Fashion: Implications for Inhibitor Design. *ACS Chem. Biol.* **2015**, *10*, 1043–1053.
- (35) Funabashi, K.; Sawata, M.; Nagai, A.; Akimoto, M.; Mashimo, R.; Takahara, H.; Kizawa, K.; Thompson, P. R.; Ite, K.; Kitanishi, K.; Unno, M. Structures of Human Peptidylarginine Deiminase Type III Provide Insights Into Substrate Recognition and Inhibitor Design. *Arch. Biochem. Biophys.* **2021**, *708*, 108911.
- (36) Dreyton, C. J.; Knuckley, B.; Jones, J. E.; Lewallen, D. M.; Thompson, P. R. Mechanistic Studies of Protein Arginine Deiminase 2: Evidence for a Substrate-Assisted Mechanism. *Biochemistry* **2014**, *53*, 4426–4433.
- (37) Grossfield, A.; Patrone, P. N.; Roe, D. R.; Schultz, A. J.; Siderius, D.; Zuckerman, D. M. Best Practices for Quantification of Uncertainty and Sampling Quality in Molecular Simulations [Article v1.0]. *Living Journal of Computational Molecular Science* **2019**, *1*, 1–24.
- (38) Zuckerman, D. M. Equilibrium Sampling in Biomolecular Simulations. *Annual Review of Biophysics* **2011**, *40*, 41–62.
- (39) Knapp, B.; Ospina, L.; Deane, C. M. Avoiding False Positive Conclusions in Molecular Simulation: The Importance of Replicas. *J. Chem. Theory Comput.* **2018**, *14*, 6127–6138.
- (40) Wang, R.; Zheng, Q. Multiple Molecular Dynamics Simulations of the Inhibitor GRL-02031 Complex with Wild Type and Mutant HIV-1 Protease Reveal the Binding and Drug-Resistance Mechanism. *Langmuir* **2020**, *36*, 13817–13832.
- (41) Çifci, G.; Aviyente, V.; Akten, E. D.; Monard, G. Assessing Protein-Ligand Binding Modes with Computational Tools: The Case of PDE4B. *Journal of Computer-Aided Molecular Design* **2017**, *31*, 563–575.
- (42) Vassaux, M.; Wan, S.; Edeling, W.; Coveney, P. V. Ensembles Are Required to Handle Aleatoric and Parametric Uncertainty in Molecular Dynamics Simulation. *J. Chem. Theory Comput.* **2021**, *17*, 5187–5197.
- (43) Childers, M. C.; Daggett, V. Validating Molecular Dynamics Simulations Against Experimental Observables in Light of Underlying Conformational Ensembles. *J. Phys. Chem. B* **2018**, *122*, 6673–6689.
- (44) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. Autodock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* **2009**, *30*, 2785–2791.
- (45) Muth, A.; Subramanian, V.; Beaumont, E.; Nagar, M.; Kerry, P.; McEwan, P.; Srinath, H.; Clancy, K.; Parelkar, S.; Thompson, P. R. Development of a selective inhibitor of protein arginine deiminase 2. *Journal of medicinal chemistry* **2017**, *60*, 3198–3211.
- (46) Olsson, M. H.; Sondergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface

Residues in Empirical pKa Predictions. *J. Chem. Theory Comput.* **2011**, *7*, 525–537.

(47) Case, D. A.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T. E., III; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Gohlke, H.; Goetz, A. W.; Greene, D.; Harris, R.; Homeyer, N.; Izadi, S.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D. J.; Merz, K. M.; Miao, Y.; Monard, G.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Smith, J.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. *AMBER 2018*; University of California, San Francisco: San Francisco, 2018.

(48) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.

(49) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison Of Simple Potential Functions For Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(50) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. Numerical Integration of the Cartesian Equations of Motion of a System With Constraints: Molecular Dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.

(51) Hopkins, C. W.; Le Grand, S.; Walker, R. C.; Roitberg, A. E. Long-time-step Molecular Dynamics Through Hydrogen Mass Repartitioning. *J. Chem. Theory Comput.* **2015**, *11*, 1864–1874.

(52) Lewis, S. D.; Johnson, F. A.; Shafer, J. A. Effect of cysteine-25 on the ionization of histidine-159 in papain as determined by proton nuclear magnetic resonance spectroscopy. Evidence for a histidine-159-cysteine-25 ion pair and its possible role in catalysis. *Biochemistry* **1981**, *20*, 48–51.

(53) Knuckley, B.; Bhatia, M.; Thompson, P. R. Protein arginine deiminase 4: evidence for a reverse protonation mechanism. *Biochemistry* **2007**, *46*, 6578–6587.

(54) Park, H.; Lee, S. Determination of the active site protonation state of β -secretase from molecular dynamics simulation and docking experiment: implications for structure-based inhibitor design. *J. Am. Chem. Soc.* **2003**, *125*, 16416–16422.

(55) Warshel, A. Calculations of enzymic reactions: calculations of pKa, proton transfer reactions, and general acid catalysis reactions in enzymes. *Biochemistry* **1981**, *20*, 3167–3177.

(56) Condic-Jurkic, K.; Smith, A.-S.; Zipse, H.; Smith, D. M. The protonation states of the active-site histidines in (6–4) photolyase. *J. Chem. Theory Comput.* **2012**, *8*, 1078–1091.

(57) Bello, A. M.; Wasilewski, E.; Wei, L.; Moscarello, M. A.; Kotra, L. P. Interrogation of the active sites of protein arginine deiminases (PAD1,-2, and-4) using designer probes. *ACS Med. Chem. Lett.* **2013**, *4*, 249–253.

(58) Hur, S.; Bruice, T. C. Near Attack Conformation Approach to the Study of the Chorismate to Prephenate Reaction. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12015–12020.

(59) Giraldo, J.; Roche, D.; Rovira, X.; Serra, J. The Catalytic Power of Enzymes: Conformational Selection or Transition state Stabilization? *FEBS Lett.* **2006**, *580*, 2170–2177.

(60) Catak, S.; Monard, G.; Aviyente, V.; Ruiz-Lopez, M. F. Reaction Mechanism of Deamidation of Asparaginyl Residues in Peptides: Effect of Solvent Molecules. *J. Phys. Chem. A* **2006**, *110*, 8354–8365.

(61) Catak, S.; Monard, G.; Aviyente, V.; Ruiz-López, M. F. Deamidation of Asparagine Residues: Direct Hydrolysis versus Succinimide-Mediated Deamidation Mechanisms. *J. Phys. Chem. A* **2009**, *113*, 1111–1120.

(62) Wang, C.; Guo, Q.-X.; Fu, Y. Theoretical Analysis of the Detailed Mechanism of Native Chemical Ligation Reactions. *Chemistry-An Asian Journal* **2011**, *6*, 1241–1251.