Protein dynamics and substrate protonation state mediate the catalytic action of *trans*-4-hydroxy-L-proline dehydratase

Zhongyue Yang[1] and Heather J. Kulik[1,*]

[1]*Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139*

ABSTRACT: The enzyme *trans*-4-Hydroxy-L-proline (Hyp) dehydratase (HypD) is among the most abundant glycyl radical enzymes (GREs) in the healthy human gut microbiome and is considered a promising antibiotic target for the prominent antibiotic-resistant pathogen *Clostridium difficile.* Although an enzymatic mechanism has been proposed, the role of the greater HypD protein environment in mediating radical reactivity is not well understood. To fill this gap in understanding, we investigate HypD across multiple time- and length- scales using electronic structure modeling and classical molecular dynamics. We observe that the Hyp substrate protonation state significantly alters both its enzyme-free reactivity and its dynamics within the enzyme active site. Accurate coupled cluster modeling suggests the deprotonated form of Hyp to be the most reactive protonation state for C5–$H_{pro\text{-}S}$ activation. In the protein environment, hydrophobic interactions modulate the positioning of Cys434 radical to enhance the reactivity of C5–$H_{pro\text{-}S}$ abstraction. Long-time dynamics reveal that changing Hyp protonation states triggers the switching of a Leu643-gated water tunnel, a functional feature that has not yet been observed for members of the GRE superfamily.

Keywords: Glycyl radical enzymes; molecular dynamics; substrate protonation states; water tunnel; substrate selectivity

1. Introduction

Human gut microbes have evolved to catalyze a broad variety of chemically challenging transformations that impact human health[1-7], metabolic,[8] cardiovascular,[9] immunological,[10] and neurological diseases.[11] A newly discovered enzyme in the human gut microbiome, *trans*-4-hydroxy-L-proline (Hyp) dehydratase (HypD) catalyzes the transformation of Hyp, an abundant nonproteinogenic amino acid,[12-14] to (*S*)-Δ^1-pyrroline-5-carboxylate (P5C) and water.[15] HypD is one of the most abundant glycyl radical enzymes (GREs) in the healthy human gut microbiome[15] and is considered a promising antibiotic target for the prominent antibiotic-resistant pathogen *Clostridium difficile.*[15-18] Along with glycerol dehydratase (GD), propanediol dehydratase (PD), choline trimethylamine-lyase (CutC), isethionate sulfite-lyase (IslA), 2(S)-dihydroxypropanesulfonate-sulfolyase/dehydratase (HpsG/HpfG),[19] HypD belongs to the 1,2-eliminase class of GREs, which are essential for anaerobic primary and secondary metabolism in bacteria and archaea.[16, 20-23] Common to all members of the GRE superfamily, HypD involves an evolutionarily conserved glycine-centered radical in the active site, which is installed by a cognate S-adenosylmethionine (AdoMet)-dependent activating enzyme.[24] Despite the characterization of stable glycyl radicals using electron paramagnetic resonance experiments,[25-26] the identity of subsequent radical intermediates generated by the glycyl radicals remains largely speculative. This is because the fleeting nature of these reactive intermediates is beyond the detection limit of common experimental techniques (i.e., X-ray crystallography and biochemical assays), making it extremely difficult to investigate the catalytic action of GREs.

Computational methods have played an essential role in unveiling the atomic-level detail of chemical reactions catalyzed by GREs. Quantum mechanics (QM) and multi-scale, quantum mechanics/molecular mechanics (QM/MM) calculations have been employed to elucidate the

direct elimination pathways for B_{12}-independent glycerol dehydratase[20, 27-28] and the mechanism of phenylacetate decarboxylase.[29] Molecular dynamics (MD) simulations have been performed to reveal the structural determinants of benzylsuccinate synthase and its mutants for substrate recognition.[30] Our previous studies combined large-scale quantum mechanical calculations with steered molecular dynamics[31] to investigate the catalytic mechanisms of CutC, revealing the spontaneous decomposition of choline to trimethylamine subsequent to substrate C–H activation.[32]

In contrast to other GRE 1,2-eliminases that involve the cleavage of a C–O or C–S bond in conformationally flexible substrates,[16] HypD catalyzes the oxidation of a C–N bond in a conformationally constrained pyrrolidine ring. Recently, the Balskus and Drennan groups have reported structural and biochemical characterization of HypD, identifying the critical structural features that enable the catalytic dehydration of Hyp.[33] Through crystallography and isotope labeling experiments, the studies unambiguously show that substrate positioning by the active-site residues mediates the C5–$H_{pro\text{-}S}$ abstraction of Hyp. Additionally, site-directed mutagenesis experiments confirm that residues coordinating Hyp are key for both substrate stabilization and catalysis. Despite considerable experimental guidance on HypD function, the nature of fleeting, reactive intermediates, including their favored protonation states, remain unclear. Although a crystal structure of HypD is available[33], the role of protein dynamics in catalysis remains particularly unknown.

To further uncover the catalytic actions of GREs, it is necessary to integrate simulation tools of different scales, ranging from correlated wavefunction theory (WFT) or density functional theory (DFT) calculations[34] and large-scale quantum mechanical computations[35-38] to classical MD simulations.[39] In this work, we employ all three techniques to determine how

different Hyp substrate protonation states influence energetics of dehydration and protein-substrate dynamics, resulting in distinct positioning and orientation of the reacting species. Our simulations reveal that a change in the Hyp protonation states triggers the switching of a Leu643-gated water tunnel, a phenomenon that has been widely reported in enzyme catalysis[40-46] but is unprecedented in the GRE superfamily.

2. Results and Discussion

2a. Deprotonated Hyp is the Most Reactive Protonation State for C5–$H_{pro\text{-}S}$ Activation by Thiyl Radical

HypD catalyzes the conversion of the *trans*-4-hydroxy-L-proline (Hyp) substrate to (*S*)-Δ^1-pyrroline-5-carboxylate (P5C) with the aid of radical chemistry (Figure 1).[33, 47] Across the family of glycyl radical enzymes (GREs), a convergent mechanism based on analysis of evolutionarily conserved residues has been proposed in which radical transfer from a C-terminus glycyl generates an active-site cysteine thiyl.[16, 48] Despite being a short-lived intermediate that is difficult to experimentally characterize, the cysteine thiyl radical is believed to be responsible for abstracting a hydrogen atom from the substrate Hyp.[16] Based on the crystal structure and biochemical assays,[33] the catalytic cycle has been proposed to involve three major steps after the formation of the Cys434 thiyl radical. First, the C5–$H_{pro\text{-}S}$ of Hyp is abstracted by the Cys434 thiyl radical to generate a substrate radical intermediate. Then the substrate radical intermediate undergoes dehydration to generate a pyrroline radical. Finally, it is believed that the pyrroline radical abstracts the Cys434 S–H to form the product, P5C (Figure 1). The cysteine thiyl radical further reacts with Gly765, generating the Gly765 radical to restart the catalytic cycle.

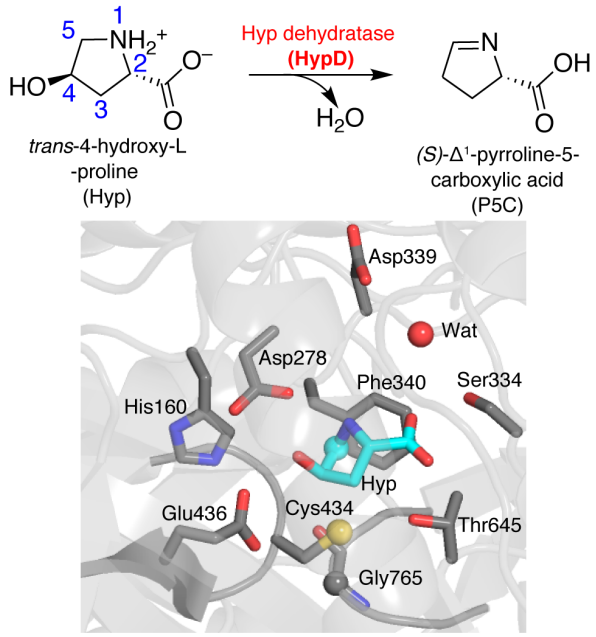

Figure 1. (top) The first step in the metabolic process of converting *trans*-4-hydroxy-L-proline (Hyp) to L-proline in human gut microbiota. The ring numbering for the Hyp substrate is shown in blue with Hyp in the zwitterionic protonation state. (bottom) Active-site structure of the HypD crystal structure (PDB ID: 6VXE). Representative active-site residues are labeled by three-letter codes and residue numbers. The protein cartoon is shown in translucent gray. The carbon atoms of substrate Hyp and representative active-site residues are colored in cyan and gray sticks, respectively. Reacting atoms are shown as spheres, including Gly765 α-carbon (gray), Cys434 thio-sulfur (yellow), and Hyp C5-carbon (cyan). The nitrogen and oxygen atoms are shown in blue and red, respectively.

Before aiming to understand the role of the enzyme in catalyzing this reaction, we computed the free energy landscape of the uncatalyzed, radical-initiated conversion of Hyp to P5C. We characterized these reaction steps with accurate QM energetics obtained with the domain local pair natural orbital coupled cluster singles doubles and perturbative triples method (DLPNO-CCSD(T))[49] and a correction to account for an implicit solvent environment. Although the general mechanism has been proposed, the ionizable groups of the Hyp substrate have a range of possible protonation states that can drastically affect substrate reactivity (Figure 2). The possible states include a neutral state, Hyp_N, where both the amine and carboxylic acid functional groups of Hyp are neutral; the anionic state, Hyp_A, where the carboxylate has been deprotonated;

and a zwitterionic form, Hyp_{ZW}, in which the pyrrolidine amine is protonated and the carboxylate is deprotonated (Figure 2). The Hyp_{ZW} form is expected to predominate in aqueous conditions at pH 7 and has been proposed to be the form that binds to the enzyme active site. We estimated the free energy difference between Hyp_{ZW} and Hyp_A from the experimental pK_a of protonated amine of Hyp in aqueous solution at pH 7 (Supporting Information Text S1).

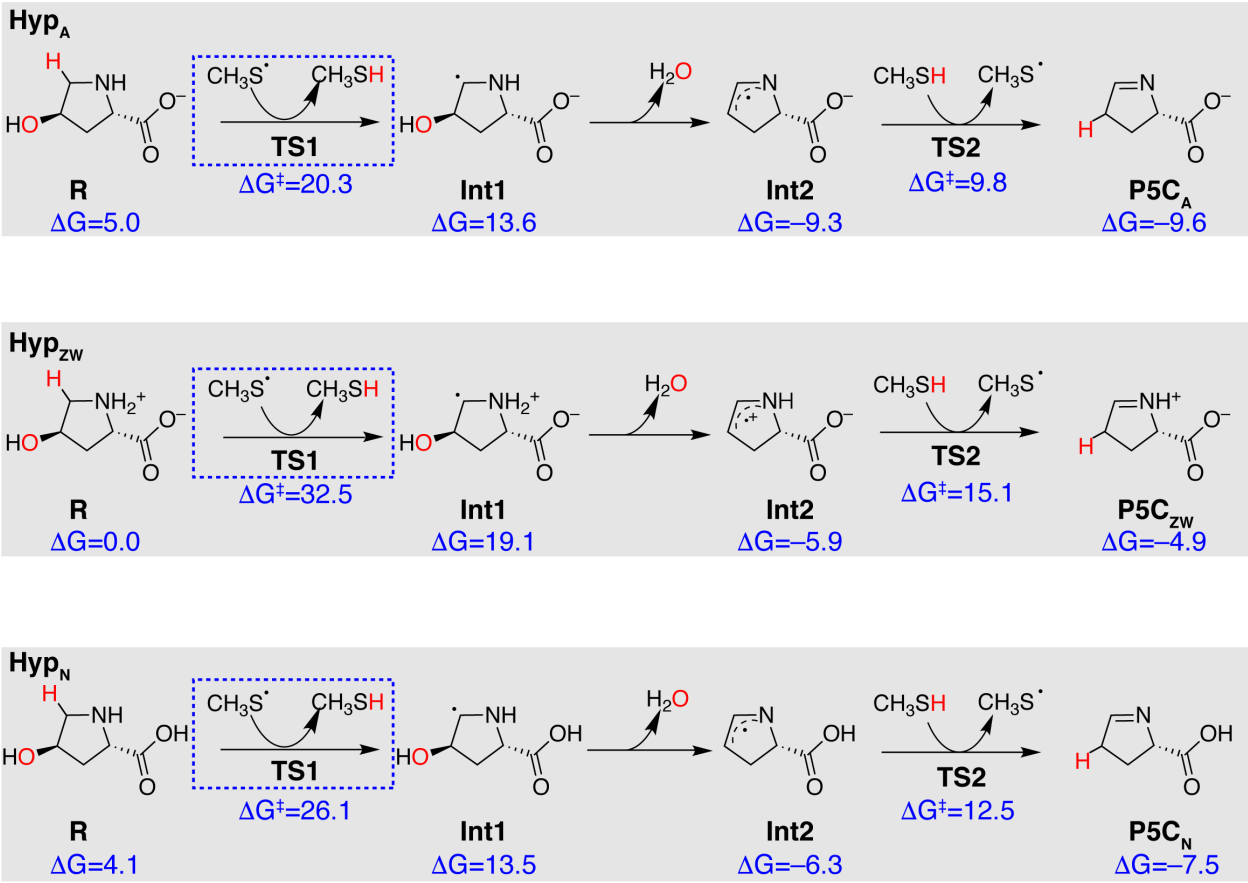

Figure 2. Non-enzymatic reaction pathways for Hyp conversion to P5C under different protonation states (i.e., Hyp_{ZW}, Hyp_N, and Hyp_A) with intermediate energies shown in kcal/mol. The reacting atoms are labeled in red. The rate-limiting steps are shown in blue dashed rectangles. The structures were computed using DLPNO-CCSD(T)-ΔC-PCM(MP2)/aug-cc-pVTZ//UB3LYP-SMD(water)/6-31G(d).

The reaction pathways of the Hyp_{ZW} and Hyp_A protonation states share the same rate-limiting step, the Hyp C5–$H_{pro\text{-}S}$ abstraction by a thiyl radical (Figure 2). Nevertheless, the barrier height varies substantially, with the anionic Hyp_A being more reactive ($\Delta G^{\ddagger}$ = 20.3

kcal/mol) than cases where the amine is protonated (i.e., Hyp_{ZW}: $\Delta G^{\ddagger}$ = 32.5 kcal/mol and Hyp_N: $\Delta G^{\ddagger}$ = 26.1 kcal/mol), despite comparable transition-state structures with similar S–H (1.4–1.5 Å) and C–H bond lengths (1.6–1.7 Å, Supporting Information Figure S1). Qualitatively similar observations are made for the alternative C5–$H_{pro\text{-}R}$ abstraction (Supporting Information Figure S1). The high activation energy for Hyp_{ZW} is due to the apparent destabilization of the partial radical character of C5 during H abstraction by the neighboring protonated amine (Figure 2). These observations suggest that the most favorable free energy pathway involves deprotonation of Hyp_{ZW} to Hyp_A which then undergoes C–H activation. However, the smaller differences between the forms with the deprotonated amine (i.e., Hyp_A and Hyp_N) suggest both should be studied further in the presence of the protein environment (see Sec. 2b).

Following C5–$H_{pro\text{-}S}$ activation, the resulting C5-centered radical intermediates (Int1) undergo exothermic dehydration to generate an enaminyl radical intermediate (Int2) with comparable reaction free energies for all substrate forms (ΔG = -6 to -9 kcal/mol, Figure 2 and Supporting Information Text S1). Finally, we modeled P5C product formation via the dehydrated enaminyl radical Int2 abstracting a hydrogen atom from the thiol to form the product and regenerate a thiyl radical (Figure 2). The barrier heights are comparable for all three substrate protonation forms ($\Delta G^{\ddagger}$ = 19–21 kcal/mol, Figure 2). In all cases, this step is relatively thermoneutral. These calculations indicate that along the reaction pathway the substrate protonation state affects the C5–$H_{pro\text{-}S}$ activation much more significantly than its following steps (Figure 2).

Alternatively, we considered pathways that start from the abstraction of C4–H, which was excluded experimentally based on isotope labeling experiments[33] and also appears unlikely from the crystal structure due to its longer distance to the Cys434 sulfur (by 2.1 Å, Supporting

Information Figure S2). Consideration of this alternate pathway was motivated for Hyp_{ZW}, which had a high barrier for C5–$H_{pro\text{-}S}$ abstraction but favorable reaction barriers for the remaining steps. We studied C4–H activation followed by a 1,2-shift as an alternate pathway to generate Int1 of Hyp in all three protonation forms (Supporting Information Figure S3). The barrier for C4–H activation of Hyp_{ZW} to form an α-hydroxyl radical intermediate (Int3) is reduced by 8.0 kcal/mol (to $\Delta G^{\ddagger}$ = 24.5 kcal/mol) with respect to abstraction at C5–$H_{pro\text{-}S}$. However, this barrier is still 4.2 kcal/mol higher in energy than the free energy barrier for the C5–$H_{pro\text{-}S}$ activation via the deprotonated substrate form Hyp_A. The subsequent 1,2-H shift involves a very high free energy barrier ($\Delta G^{\ddagger}$ = 60.3 kcal/mol for TS4), making formation of Int1 through this pathway improbable (Supporting Information Figure S3). Instead, it has been proposed that an α-hydroxyl radical (i.e., the structural moiety in Int3) can readily undergo deprotonation[50] to form a ketal radical anion, which may induce ring-expansion via direct elimination. A similar mechanism has been proposed for other GRE-catalyzed reactions, such as choline degradation,[32, 51-53] 1,2-propanediol glycerol dehydration,[54] and glycerol dehydration.[27]

These results suggest that the enzyme environment could be important for preferentially stabilizing one substrate protonation state. The most directional interactions likely to be sensitive to protonation and charge state differences are strong, charge-assisted hydrogen bonds[55-56] or distinct electrostatic interactions[38, 57], both of which are on the order of 5–10 kcal/mol. Because we observed reactivity at C5–$H_{pro\text{-}S}$ for Hyp_A to have the lowest barrier but the C5–$H_{pro\text{-}S}$ Hyp_N barrier to be only slightly higher, preferential directional hydrogen bonding and electrostatic interactions for one of the substrates by the enzyme have the potential to reverse this preference (Figure 2). Examining the substrate active site, we observe that there are several good hydrogen bond donors (i.e., conserved water, Ser334, and Thr645) that would favor binding a charged

carboxylate, although this would not distinguish Hyp_{ZW} from Hyp_A (Figure 1). Another residue, Asp278, is suited to stabilizing the protonated pyrrolidine amine in Hyp_{ZW} (Figure 1). Thus, we next investigated these substrates directly in the enzyme active site to answer how HypD stabilizes more reactive protonation states of Hyp.

2b. Different Hyp Protonation States Result in Distinct Active-site Configurations and Substrate-Protein Dynamics

We performed classical MD to investigate specific active-site-substrate interactions that differentially stabilize the protonation states of Hyp (i.e., Hyp_{ZW}, Hyp_N, and Hyp_A). This also allowed us to characterize how the hydrogen bond network reorganizes upon changes in substrate protonation. We first investigated HypD dynamics with substrate Hyp_A, the most reactive protonation state identified by the uncatalyzed QM calculations. For the HypD/Hyp_A complex, the substrate pyrrolidine amine is neutral, and we protonated Asp278 in accordance with experimental mechanistic proposals.[33] Although the simulations were initialized with a hydrogen bond between the pyrrolidine amine and Asp278 as in the crystal structure, reorganization during dynamics orients protonated Asp278 toward neighboring Glu436 and His160 (Figure 3 and Supporting Information Figure S4). This shift of Asp278 is unsurprising due to the greater electrostatic attraction with Glu436 in comparison to the neutral pyrrolidine amine. At the same time, an active-site Ser334 H-bonds to the Hyp_A carboxylate and a water molecule bridges the Hyp_A hydroxyl group with Glu436, compensating for the loss of the interaction with Asp278 (Figure 3). With the exception of the rearrangement of Asp278, Hyp_A remains bound in the active site over the 200-ns MD trajectory in an orientation structure with only modest rearrangement (RMSD: 1.6 Å) with respect to the crystal structure (Supporting Information Figure S5–S7).

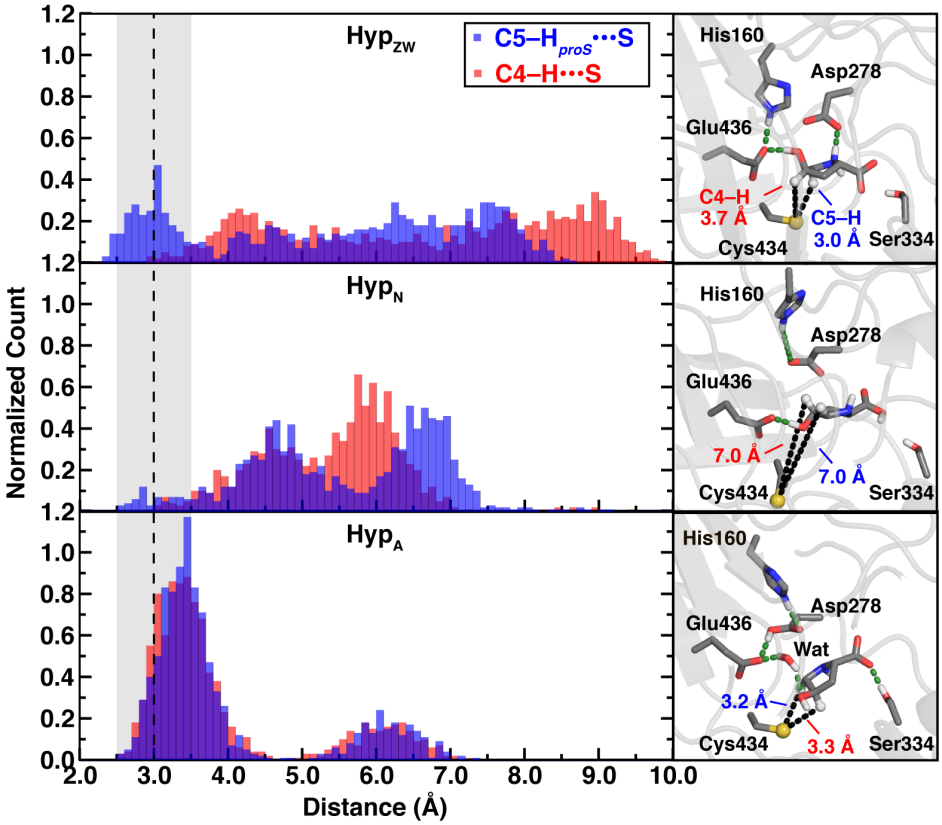

Figure 3. Distribution of distances from Cys434-S to C5–H (C5–H•••S, blue) and C4–H (C4–H•••S, red) along the MD trajectories, in which HypD is complexed with Hyp_{ZW}(top), Hyp_N (middle), and Hyp_A (bottom), respectively. The distances were taken along the trajectories with a 0.2 ns time interval. A reactive region 0.5 Å around a distance centered at 3.0 Å (sum of van der Waals radius of S and H, vertical dashed line) is indicated with gray shading. The final snapshot of the MD trajectory associated with each protonation state is shown on the right, where we labeled the C5/C4–H•••S distances and critical residues that interact with the substrates. The protein cartoon is shown in translucent gray. The carbon atoms are colored in gray sticks. Reacting atoms are shown as spheres, including the Cys434 thio-sulfur (yellow) and C5–H_{proS} and C4–H (white). The nitrogen, oxygen, and hydrogen atoms are shown in blue, red, and white, respectively. The polar H-bond and electrostatic interactions are shown in green dashed lines, and the distances between reacting atoms are shown in black dashed lines.

To characterize the conformations sampled by the Cys434 radical in the presence of Hyp_A, we computed the distribution of the distances between Cys434 sulfur atom and the substrate reacting hydrogen atoms (i.e., C5–$H_{pro\text{-}S}$···S and C4–H···S, Figure 3). We observe a bimodal distribution for both distances, with Cys434 sampling an orientation directed toward both hydrogen atoms most of the time (80%) and directed away a small fraction (20%) of the time

(Figure 3 and Supporting Information Figure S8). To extract information about the likelihood of a hydrogen abstraction reaction from non-reactive MD, we defined a region of reactive distances within 3 Å ± 0.5 Å (i.e., approximately the sum of the S and H van der Waals radii[58]). A significant fraction of the proximal Cys434 peak resides within this reactive region, leading to 51% of the overall trajectory sampling this orientation.

Our MD simulation shows that for the HypD/Hyp_A complex, the C5–$H_{pro\text{-}S}$ and C4–H have similar probabilities of being in a near-attack configuration[59] with the Cys434 thiyl radical. Although we see shifts in the quantitative populations of the two Cys434 orientations, the observation that the Cys434 radical distance is equivalent to both C5 and C4 is reproduced by additional independent simulations initialized from the configuration after 200 ns of dynamics (Supporting Information Figure S6). This observation is in contrast to the crystal structure,[33] where the C5–$H_{pro\text{-}S}$···S distance was observed to be much shorter than the C4–H···S distance (i.e., 2.7 Å vs 4.8 Å, Figure 3 and Supporting Information Figure S2).[33] In the absence of a strong positioning effect, our observations in the enzyme environment are consistent with the uncatalyzed (i.e., enzyme-free) QM reactivity differences for abstraction at C5 (i.e., 5.7 kcal/mol more favorable) and C4 for the Hyp_A substrate as a potential explanation for the experimentally observed exclusive activation[33] of C5–$H_{pro\text{-}S}$ (Figure 2 and Supporting Information Figure S3).

We next considered whether changing substrate protonation states alters these observations. For the HypD/Hyp_N complex, the only change from the HypD/Hyp_A complex is protonation of the substrate carboxylate and a compensating deprotonation of Asp278 to maintain a consistent number of protons and charges across the enzyme-substrate complexes. These changes cause spontaneous rearrangement of the Hyp_N substrate during dynamics. In this case, a direct hydrogen bond between the substrate hydroxyl and Glu436 is formed, but there is no H-bond

between the carboxyl group and Ser334 as was observed for the carboxylate in HypD/Hyp_A and no interaction with Asp278 (Figure 3). As a result, the substrate is free to rotate, causing it to orient into a distinct conformation (RMSD: 2.0 Å) with respect to that for the crystal structure (Supporting Information Figures S5 and S7). As a result of this substrate reorientation, the Cys434 radical remains equivalently distant from both C5–$H_{pro\text{-}S}$ and C4–H for the majority of the dynamics, seldom sampling the distances we deemed to be reactive (6% C5–$H_{pro\text{-}S}$···S and 3% C4–H···S distances, Figure 3).

Finally, we consider Hyp_{ZW}, which was suggested to be the protonation state of the substrate that binds to the HypD active site. The presence of both a protonated pyrrolidine amine and deprotonated carboxylate leads to a stable HypD/Hyp_{ZW} complex that indeed has the best overall (RMSD: 1.3 Å) agreement with the X-ray crystal structure (Supporting Information Figures S5 and S7). For the charged functional groups of Hyp_{ZW}, the protonated pyrrolidine amine strongly interacts with Asp278 and the carboxylate interacts with Ser334 (Figure 3). At the same time, the stabilizing Glu436 and substrate hydroxyl interaction is preserved, further supporting the hypothesis that Hyp_{ZW} represents the preferred protonation state for binding to the HypD active site.[33] To investigate the effect these interactions have on reactivity, we evaluated distances between the substrate and the Cys434 radical using MD. For Hyp_{ZW}, both C5–$H_{pro\text{-}S}$···S and C4–H···S distance distributions are broad, with most of the sampled distances far from the reactive region (Figure 3). Nevertheless, C5–$H_{pro\text{-}S}$···S distances are generally smaller than C4–H···S distances throughout, with a greater percentage (24.6% vs 2.5%) residing within the reactive region (Figure 3).

To account for the sensitivity of these results to potential inaccuracies in the force field parameters,[32, 35, 60] we also optimized active-site structures for the cysteine radical using QM

cluster calculations. These large-scale QM calculations incorporate Hyp_{ZW}, one conserved water molecule, and 14 nearby HypD residues that were selected based on their proximity to the substrate and proposed roles[33] from mutagenesis experiments (Supporting Information Table S1 and Figure S9). In the two intermediates, C4–H···S distances are 4.2 Å, while C5–H···S distances are below 3.0 Å, supporting the substrate positioning effects of the active-site residues. Although this effect has been observed in the crystal structure, our results provide further support for this positioning effect in the enzyme radical intermediate.

Distinct substrate-protein interactions are observed between the complexes of HypD with the two most plausible substrate protonation states, Hyp_A and Hyp_{ZW}. In MD simulations, we observed the distance distribution of C5–$H_{pro\text{-}S}$···S for Hyp_{ZW} to be much broader than for Hyp_A. Some of this distinction comes from differences in positioning the substrate itself, while some arises from the rotamer favored by Cys434 (Supporting Information Figure S8). To investigate which contribution matters most, we next investigated the free energy surface of the Cys434 rotation.

2c. Repulsive Hydrophobic Interactions Modulate the Positioning of Cys434 Radical to Favor Hyp C5–$H_{pro\text{-}S}$ Activation

We performed free energy sampling along the Cys434 dihedral, χ, to investigate how the free energy profiles of this residue differ in the HypD/Hyp_A and HypD/Hyp_{ZW} complexes (Figure 4). The rotamers with χ around –60° correspond to a reactive conformation where the S atom of Cys434 orients toward the substrate reacting hydrogen atoms, while those around –180° or 180° correspond to unreactive conformers oriented away from the substrate. Based on the free energy profiles, Cys434 has distinct rotameric preferences in the two HypD complexes (Figure 4). The reactive conformation is favored by ca. 0.5 kcal/mol in the HypD/Hyp_A complex but the

unreactive conformation is preferred by ca. 1.5 kcal/mol in the HypD/Hyp$_{ZW}$ complex (Figure 4). This difference could suggest a further preference for the reaction to occur with the Hyp$_A$ substrate protonation state. In the case of Hyp$_{ZW}$, the higher population of the unreactive Cys434 orientation contributes to its broader distribution of C5–H$_{pro\text{-}S}$···S/C4–H···S distances (Figure 3).

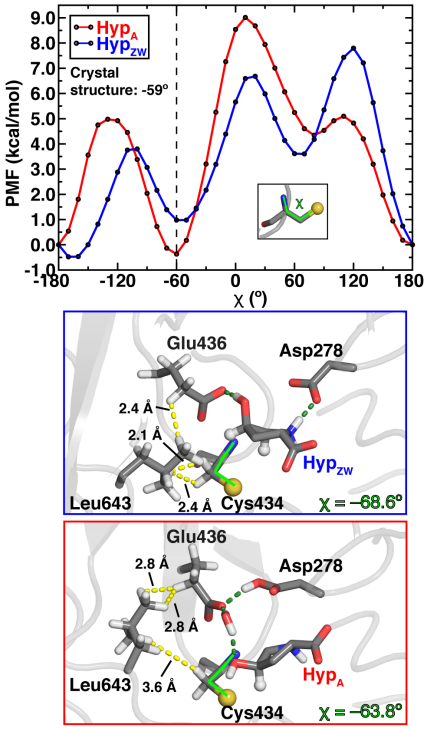

Figure 4. (Top) Potential of mean force (PMF) for the conformational changes of Cys434 radical (N–C$_\alpha$–C$_\beta$–S, labeled as χ and highlighted in inset) in HypD/Hyp$_{ZW}$ (blue) and HypD/Hyp$_A$ (red) complexes. For each complex, the free energies were computed relative to the conformational state at χ = –180°. The dihedral of the HypD crystal structure (-59º) is shown by a dashed line. (Bottom) Reactive conformers for the HypD/Hyp$_{ZW}$ (blue) and HypD/Hyp$_A$ (red) complex. The protein cartoon is shown in translucent gray. The carbon atoms are colored in gray sticks. Reacting atoms are shown in spheres, including Cys434 thio-sulfur (yellow) and C5–H$_{pro\text{-}S}$ and C4–H (white). The nitrogen, oxygen, and hydrogen atoms are shown in blue, red, and white sticks, respectively. The repulsive interactions between Leu643 and nearby residues (i.e., Glu436 and Cys434) are indicated by the H–H distances as shown as yellow dashed lines.

When comparing representative unreactive Cys434 conformations for the HypD/Hyp$_{ZW}$ and HypD/Hyp$_A$ complexes, we observe that Phe340, which contributes to substrate positioning through steric interactions, plays a distinct role in the dynamics for the two substrate protonation states. This observation is consistent with the severe deleterious effects of Phe340A mutation

reported in biochemical assay.[33] While Phe340 pushes Cys434 away from the substrate in the HypD/Hyp_{zw} complex, the residue is distant from Cys434 in the HypD/Hyp_A complex (Supporting Information Figure S10). Taken together, these observations suggest that both orientation preferences of Cys434 rotamers and positioning of the substrate play the role of repelling the Cys434 radical away from the Hyp_{ZW} substrate but closer to the Hyp_A substrate to initiate the C5–$H_{pro\text{-}S}$ activation.

To understand the origins of Cys434 conformer favorability, we investigated two representative structures in a reactive Cys434 orientation from the free energy simulations of the HypD/Hyp_A and HypD/Hyp_{ZW} complexes (Figure 4). We observe that Leu643 positioning is distinct in the two complexes. In HypD/Hyp_{ZW}, Leu643 tends to contribute to positioning Cys434 away from its reactive orientation, as suggested by the short distances (ca. 2.1–2.4 Å) between the β-methylene hydrogen atoms of Cys434 and the isopropyl hydrogen atoms of Leu643 (Figure 4). In this protonation state, the motion of Leu643 residue is also affected by a proximal Glu436 with short non-covalent distances (ca. 2.4 Å) between the isopropyl hydrogen atoms of Leu643 and the γ-methylene hydrogen atoms of Glu436 (Figure 4). In contrast for the HypD/Hyp_A complex, hydrogen bonding between protonated Asp278 and Glu436 positions Glu436 away from Leu643 (Figure 4). These distinct orientations leave room for Leu643 to rotate, removing crowding of Cys434 in its reactive state, as indicated by the longer non-bonded hydrogen atom distances (ca. 2.8–3.6 Å) between Leu643 and Glu436 or Cys434, respectively (Figure 4). This structural analysis suggests that the modulation of repulsive hydrophobic interactions in the active site plays an important role in making the reactive Cys434 rotamer the most favorable in the HypD/Hyp_A complex.

2d. Changing the Hyp Protonation States Triggers Leu643-Gated Water Tunnel Switching

During HypD/Hyp_A dynamics, bulk solvent appears to enter the enzyme active site. These water molecules in turn form hydrogen bonds with both the substrate hydroxyl group and Glu436 (Figure 5). To determine the path for solvent access to the active site, we carried out cavity analysis[61] on the HypD/Hyp_A complex. For 96% of the selected snapshots, a cavity with an average radius of 1.7 Å is large enough to accommodate a single chain of water (radius ca. 1.4 Å) molecules (Figure 5 and Supporting Information Table S2). In contrast, a narrow water tunnel is found in only 2% of the selected snapshots for the HypD/Hyp_{ZW} complex (Supporting Information Table S3). This infrequently occurring HypD/Hyp_{ZW} tunnel has a radius of 1.1 Å, which is narrower than that identified for the HypD/Hyp_A complex. To investigate why substrate protonation states (i.e., Hyp_{ZW} vs Hyp_A) influence the presence of a water tunnel in HypD, we re-examined the differences in substrate–protein hydrogen bonding in the two cases. For the HypD/Hyp_{ZW} complex with a closed water tunnel, the substrate forms a strong hydrogen bond to Glu436, whereas in HypD/Hyp_A complex Asp278 H-bonds to Glu436 (see Figure 3). As a result of this shift in hydrogen bonding, the bulky hydrophobic sidechain of Leu643 reorients, with its χ_1 dihedral favoring a distinct rotamer in HypD/Hyp_A (–60°) from that (60° or 180/-180°) in the HypD/Hyp_{ZW} complex (Supporting Information Figure S11). The reorientation of this bulky sidechain coincides with the opening of the active site to bulk solvent.

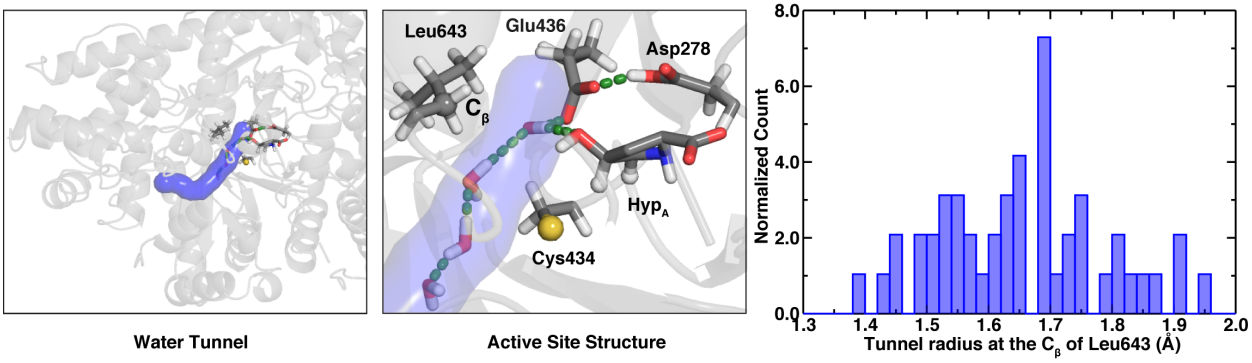

Figure 5. The distribution of tunnel radii at the C_β of Leu643. Water tunnels were found using CAVER in 48 of 50 snapshots that were taken evenly from the MD trajectory of the HypD/Hyp_A complex. The protein cartoon is shown in translucent gray and the water tunnel shown as a blue

transparent surface. The carbon atoms are colored in gray sticks with an exception of the C_β of Leu643 shown as a sphere. Reacting atoms are also shown as spheres, including Cys434 thio-sulfur (yellow) and C5–H_{proS} and C4–H (white). The nitrogen, oxygen, and hydrogen atoms are shown in blue, red, and white sticks, respectively. The polar H-bond and electrostatic interactions are shown in green dashed lines.

This Leu643-gated water tunnel potentially serves a catalytic function. When open, the tunnel will enable water molecules to enter, shifting the hydrogen-bonding network to active-site residues and the substrate hydroxyl group. This extensive H-bond network could facilitate proton transfer from protonated Asp278 to the hydroxyl group of the substrate that would in turn enable dehydration to occur after formation of the radical Int1 (see Figure 2). Quantitative validation of this expectation will require further mechanistic study in future work. Alternatively, in the HypD/Hyp_{ZW} complex, the orientation of Leu643 closes the water tunnel and could instead prevent water molecules from disrupting the substrate–protein hydrogen bonds that are essential for reactivity. The protonation-state-dependent water tunnel computationally identified here represents a potentially unique feature of HypD catalysis not previously observed for other GRE family members.

3. Conclusions

We leveraged QM and MD simulations to investigate how the interplay between substrate protonation states (i.e., Hyp_{ZW}, Hyp_N, and Hyp_A) and protein dynamics contributes to the catalytic competence of *trans*-4-Hydroxy-L-proline dehydratase (HypD). Using coupled-cluster-level QM calculations, we showed that in the absence of enzyme the reaction pathways of possible substrate protonation states share the same rate-limiting step, i.e. the Hyp C5–$H_{pro\text{-}S}$ abstraction by a thiyl radical, and the most favorable free energy path is for Hyp_{ZW} to be deprotonated to Hyp_A and then undergo C–H activation.

We performed classical MD to investigate specific active-site noncovalent interactions that differentially stabilize the protonation states of Hyp. For the HypD/Hyp_A complex, the MD trajectory reveals reorganization that orients Asp278 toward neighboring Glu436 and His160. Hyp_A remains bound in the active site over the course MD trajectories in an orientation that allows thiyl radical to readily undergo substrate activation. The MD trajectory of the HypD/Hyp_{ZW} complex has the best overall agreement with the crystal structure. To compensate for potential sensitivity of these results to force field parameters, we employed large-scale QM simulations to optimize a key reactive intermediate in the presence of active-site residues and conserved a water molecule. The calculations reveal much closer spatial proximity of the C5–$H_{pro-S}\cdots$S distance than the C4–$H_{pro-S}\cdots$S, suggesting positioning that favors C5–H_{pro-S} abstraction in agreement with experimental studies.

Free energy profiles of the dihedral angle of Cys434, the residue bearing the radical before substrate activation, were observed to differ in the HypD/Hyp_A and HypD/Hyp_{ZW} complexes. We found that the positioning of the substrate plays the role of repelling the Cys434 radical away from the Hyp_{ZW} substrate and organizing the Cys434 radical closer to the Hyp_A substrate to initiate the C5–H_{pro-S} activation. We showed that substrate positioning is modulated by repulsive hydrophobic interactions, particularly by Leu643 and Phe340. We demonstrated that the reorientation of the Leu643 sidechain coincides with the opening of the active site to bulk solvent in the HypD/Hyp_A. These results reveal a cooperative effect for HypD – the deprotonation of Hyp_{ZW} to Hyp_A for facile substrate activation is coupled with the protonation of the Asp278 to trigger the active-site reorganization for substrate positioning and water-tunneling opening. This Leu643-gated water tunnel potentially serves a catalytic function and represents a

unique feature of HypD that has not been reported for other members of the GRE family. This finding motivates future investigation into water-tunnel-based inhibitor design.

4. Computational Details

Small-Molecule QM calculations. Electronic structure calculations were performed to investigate the energetics of substrates, intermediates, and products along the hypothesized intrinsic reaction pathways (i.e., without explicitly modeling the enzyme environment) using Q-Chem 5.2[62]. Geometry optimizations and Hessian-based transition state (TS) searches were performed using density functional theory (DFT) with the hybrid B3LYP[63-65] functional in conjunction with the 6-31G(d) basis. All intermediates and TSes were modeled and optimized in an aqueous environment (i.e., $\varepsilon = 78.39$) using the SMD implicit solvation model[66] with all defaults applied. The SMD model uses a series of parameterized radii for each element based on training solvation data.

The geometry optimizations were carried out in internal, *Z*-matrix coordinates using the eigenvector-following algorithm[67-69] for ground-state structures and the GDIIS algorithm[70] for TSes. Initial structures of intermediates were built by hand in Avogadro[71] v1.20, and TSes were modified from the optimized intermediates by stretching the forming and breaking bonds. Analytical Hessians were computed on the initial TS structures and provided as inputs to transition state search with partitioned rational function optimization (P-RFO)[72-73]. Default convergence thresholds of 3.0×10^{-4} hartree/bohr for the maximum gradient and 1.0×10^{-6} hartree for the change in energy were employed for both the geometry optimization and transition state search.

After convergence of TSes and intermediates, analytical Hessians were computed to ensure the presence of a single imaginary frequency for the TS corresponding to the expected

reaction coordinate motion and no imaginary frequencies for the intermediates. These Hessian calculations were also used to obtain entropic (T = 298.15 K) and zero-point vibrational energy contributions to the Gibbs free energies. All energies are provided in a spreadsheet in the Supporting Information .zip file.

To obtain more accurate reference energetics, corrected single-point energies were obtained with domain local pair natural orbital (PNO) coupled cluster calculations (i.e., DLPNO-CCSD(T)[49]) The DLPNO-CCSD(T)[49] calculations were carried out with the aug-cc-pVTZ basis set using ORCA[74] v.4.0.1.2. These calculations employed a TightPNO cutoff, which corresponds to a PNO occupation number of 1.0×10^{-7} and an estimated pair correlation energy of 1.0×10^{-5} hartree (Supporting Information Table S4). Since implicit solvent models are not implemented in DLPNO-CCSD(T), we corrected gas-phase DLPNO-CCSD(T) intermediate energies with SMD[66] (ε = 78.39) solvation energies obtained from MP2 in ORCA (Supporting Information Table S4).

QM Enzyme Cluster Calculations. QM cluster calculations were carried out with a developer version of TeraChem[75-77]. The cluster was extracted from the crystal structure. The cluster included the substrate, *trans*-4-hydroxy-L-proline (Hyp) and 14 residues selected due to their proximity to the substrate and proposed roles[33] from mutagenesis experiments (Supporting Information Table S1). The backbone atoms of all residues are included in the cluster and capped with hydrogen atoms for a cluster 269 atoms in size with a net charge of -2 and a spin multiplicity of 2 to model the Gly765, Cys434, or substrate radicals. The clusters were optimized in the gas phase with all C_α atoms held fixed using the ωPBEh[78] (ω=0.2 $bohr^{-1}$) exchange-correlation functional and the 6-31G(d)[79] basis set. This cluster was geometry optimized in translation-rotation internal coordinates (TRIC) with L-BFGS using the TRIC

optimizer[80]. The convergence criteria were loosened from their defaults by an order of magnitude due to the large size of the cluster, with a maximum for the gradient of 4.5×10^{-3} hartree/bohr and maximum change in energy between steps of 1.0×10^{-5} hartree. To approximate the protein environment, all intermediate energies were corrected with single-point evaluation using C-PCM[81-82], as implemented in TeraChem[83-84] with ε=4.0. All final optimized geometries are provided in the Supporting Information .zip file.

Protein Structure and Preparation. The crystal structure of substrate-bound HypD was obtained from the Protein DataBank (PDB ID: 6VXE), and all crystallizing agents were removed.[33] The protein was prepared in complex with the Hyp substrate in each of the zwitterionic, neutral, and anionic protonation states. In the anionic Hyp-bound HypD, Asp278 was hypothesized[33] to be protonated as aspartic acid (Ash278) based on mutagenesis experiments[33]. For the remaining residues, protonation states were assigned on the apoprotein using the H++ webserver[63, 85-86] assuming a pH of 7.0 and a dielectric constant of 10.0 with all other defaults applied. For all Hyp protonation states, this procedure produces a holoprotein with 12,435 atoms and –27 net charge.

Structures were prepared with the AMBER[87] tleap utility for classical molecular dynamics (MD) with the AMBER ff14SB force field[88] for the protein. Parameters for non-standard residues, i.e., the Gly765 and Cys434 radicals and Hyp substrate in three protonation states, were obtained using the generalized AMBER force field (GAFF)[89]. The R.E.D.S. web server[90-92] was employed to obtain restrained electrostatic potential (RESP) charges[64] calculated by GAMESS-US[93] at the Hartree–Fock/6-31G(d)[79] level of theory. The Hyp-bound protein complexes were solvated in a periodic rectangular box with at least a 10 Å buffer of TIP3P[94] water and neutralized with Na^+ counterions for a final system size of 88,500 atoms and a box

dimension of 90.5 Å x 108.6 Å x 106.4 Å. The AMBER MD input files (i.e. prmtop and inpcrd) are provided in the Supporting Information .zip file.

Classical MD Simulations. The enzyme complexes were equilibrated with MD using the GPU-accelerated form of AMBER, PMEMD.[95-96] The equilibration steps were as follows: i) 3000 minimization steps, ii) 10 ps NVT heating to 300 K with a Langevin thermostat with collision frequency of 1.0 ps^{-1} and a random seed, and iii) 250 ps NPT equilibration using the Berendsen barostat with a pressure relaxation time of 2 ps. These steps were followed by a 200-ns NPT production run. From steps i to iii, the substrate, protein scaffold, and conserved water molecules were restrained with 200 kcal/mol·$Å^2$ force constants to maintain the crystal structure atomic positions. The restraining potential was released during the production run to allow sampling of substrate orientation in the active site. The SHAKE algorithm[97] was applied to all bonds with hydrogen atoms in combination with a 2-fs timestep for all MD simulations. For the long-range electrostatics, the particle mesh Ewald method was used with a 10-Å real space electrostatic cutoff.

Umbrella Sampling. We employed umbrella sampling[98] to map out the free energy surfaces for rotations of residues Cys434 and Leu643. The appropriate reaction coordinates are changes in a residue dihedral angle. The selected windows were equally spaced with 10° interval over the entire dihedral range (−180° to 180°) for both Cys434 and Leu643. For each window, 250 ps equilibration and 750 ps production was carried out in the NpT ensemble with the targeted angle enforced with a 200 kcal/(mol·rad^2) force constant. To ensure a well-converged free energy surface, sampling windows were initiated sequentially, and the last production snapshot from the preceding window was used as input for the subsequent window. The weighted-histogram analysis method (WHAM)[99-100] was applied using the Grossfield software

package[101] to unbias the dihedral distributions. The WHAM free energy weights were converged to within 1×10^{-5}.

Tunnel Analysis. The protein structure was analyzed for cavities during the MD trajectory to identify "open" and "closed" orientations. These qualitative states were then quantitatively analyzed with CAVER[61]. For 50 representative snapshots of both the "open" and "closed" states, the analysis was run with a shell radius of 5.5 Å, probe radius of 1.0 Å, and shell depth of 7.0 Å, which were selected by trial and error. The search for tunnels started from the hydroxyl group of Hyp and visually verified to connect from the active site to the protein surface passing adjacent to Leu643. The overall tunnel bottleneck was extracted from CAVER output, and the sphere radius in the path closest to the C_β of Leu643 was used to determine the most relevant cavity radius (Supporting Information Table S2–S3). All analyzed snapshots and tunnels are provided in the Supporting Information .zip file.

ASSOCIATED CONTENT

Supporting Information. Benchmark and details for small-QM simulations.; cavity analysis for MD snapshots of the HypD/Hyp_A and HypD/Hyp_{ZW} complexes; free energies for the activation of the C5–$H_{pro\mathit{R}}$ and C4–H using small-QM simulations; typical configurations sampled from the classical MD trajectories; structural alignment of the active-site residues between the crystal structure and the end snapshots of MD for three Hyp protonation states; bond lengths distribution for additional sets of MD simulations; snapshot of crystal structure; RMSD of the active-site residues along the MD trajectories of different protonation states with respect the crystal structure; dihedral angle distributions of Cys434 and Leu643 for the HypD/Hyp_A and HypD/Hyp_{ZW} complexes; optimized structure for the Cys434 radical intermediate using QM cluster simulations. (PDF)

Starting structures for classical MD; optimized structures for small-QM; MD snapshots used for cavity analysis; optimized structures for the Cys434 radical intermediate using QM cluster simulations; absolute and relative energies for the small-QM and QM cluster simulations. (ZIP)

This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*email: hjkulik@mit.edu phone: 617-253-4584

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

This research was supported by the Department of Energy under grant number DE-SC0019112. Initial conception for this study was supported by an MIT Research Support Committee NEC Corporation Grant. This work was carried out in part using computational resources from the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562. H.J.K. holds a Career Award at the Scientific Interface from the Burroughs Wellcome Fund, an AAAS Marion Milligan Mason Award, and an Alfred P. Sloan Fellowship in Chemistry, which also supported this work. The authors thank Adam H. Steeves and Emily P. Balskus for providing a critical reading of the manuscript.

References

1. Shreiner, A. B.; Kao, J. Y.; Young, V. B., The gut microbiome in health and in disease. *Curr. Opin. Gastroenterol.* **2015,** *31* (1), 69.
2. Yatsunenko, T.; Rey, F. E.; Manary, M. J.; Trehan, I.; Dominguez-Bello, M. G.; Contreras, M.; Magris, M.; Hidalgo, G.; Baldassano, R. N.; Anokhin, A. P., Human gut microbiome viewed across age and geography. *Nature* **2012,** *486* (7402), 222-227.
3. Integrative, H. M. P.; Proctor, L. M.; Creasy, H. H.; Fettweis, J. M.; Lloyd-Price, J.; Mahurkar, A.; Zhou, W.; Buck, G. A.; Snyder, M. P.; Strauss Iii, J. F., The integrative human microbiome project. *Nature* **2019,** *569* (7758), 641-648.
4. Sharon, G.; Garg, N.; Debelius, J.; Knight, R.; Dorrestein, P. C.; Mazmanian, S. K., Specialized metabolites from the microbiome in health and disease. *Cell Metab.* **2014,** *20* (5), 719-730.
5. Joice, R.; Yasuda, K.; Shafquat, A.; Morgan, X. C.; Huttenhower, C., Determining microbial products and identifying molecular targets in the human microbiome. *Cell Metab.* **2014,** *20* (5), 731-741.
6. Koppel, N.; Rekdal, V. M.; Balskus, E. P., Chemical transformation of xenobiotics by the human gut microbiota. *Science* **2017,** *356* (6344).
7. Kang, D.-W.; Park, J. G.; Ilhan, Z. E.; Wallstrom, G.; LaBaer, J.; Adams, J. B.; Krajmalnik-Brown, R., Reduced incidence of Prevotella and other fermenters in intestinal microflora of autistic children. *PloS One.* **2013,** *8* (7), e68322.
8. Agus, A.; Planchais, J.; Sokol, H., Gut microbiota regulation of tryptophan metabolism in health and disease. *Cell Host Microbe* **2018,** *23* (6), 716-724.
9. Tang, W. H. W.; Hazen, S. L., The contributory role of gut microbiota in cardiovascular disease. *J. Clin. Investig.* **2014,** *124* (10), 4204-4211.
10. Fung, T. C.; Olson, C. A.; Hsiao, E. Y., Interactions between the microbiota, immune and nervous systems in health and disease. *Nat. Neurosci.* **2017,** *20* (2), 145.
11. Strati, F.; Cavalieri, D.; Albanese, D.; De Felice, C.; Donati, C.; Hayek, J.; Jousson, O.; Leoncini, S.; Renzi, D.; Calabrò, A., New evidences on the altered gut microbiota in autism spectrum disorders. *Microbiome* **2017,** *5* (1), 24.
12. Gorres, K. L.; Raines, R. T., Prolyl 4-hydroxylase. *Crit. Rev. Biochem. Mol. Biol.* **2010,** *45* (2), 106-124.
13. Verbeken, D.; Dierckx, S.; Dewettinck, K., Exudate gums: occurrence, production, and applications. *Appl. Microbiol. Biotechnol.* **2003,** *63* (1), 10-21.
14. Valiente, C.; Arrigoni, E.; Esteban, R. M.; Amado, R., Grape pomace as a potential food fiber. *J. Food Sci.* **1995,** *60* (4), 818-820.
15. Levin, B. J.; Huang, Y. Y.; Peck, S. C.; Wei, Y.; Martinez-Del Campo, A.; Marks, J. A.; Franzosa, E. A.; Huttenhower, C.; Balskus, E. P., A prominent glycyl radical enzyme in human gut microbiomes metabolizes trans-4-hydroxy-l-proline. *Science* **2017,** *355* (6325).
16. Backman, L. R. F.; Funk, M. A.; Dawson, C. D.; Drennan, C. L., New tricks for the glycyl radical enzyme family. *Crit. Rev. Biochem. Mol. Biol.* **2017,** *52* (6), 674-695.
17. Leffler, D. A.; Lamont, J. T., Clostridium difficile infection. *N. Engl. J. Med.* **2015,** *372* (16), 1539-1548.
18. Lessa, F. C.; Mu, Y.; Bamberg, W. M.; Beldavs, Z. G.; Dumyati, G. K.; Dunn, J. R.; Farley, M. M.; Holzbauer, S. M.; Meek, J. I.; Phipps, E. C., Burden of Clostridium difficile infection in the United States. *N. Engl. J. Med.* **2015,** *372* (9), 825-834.

19. Liu, J.; Wei, Y.; Lin, L.; Teng, L.; Yin, J.; Lu, Q.; Chen, J.; Zheng, Y.; Li, Y.; Xu, R., Two radical-dependent mechanisms for anaerobic degradation of the globally abundant organosulfur compound dihydroxypropanesulfonate. *Proc. Natl. Acad. Sci. U.S.A.* **2020,** *117* (27), 15599-15608.
20. O'Brien, J. R.; Raynaud, C.; Croux, C.; Girbal, L.; Soucaille, P.; Lanzilotta, W. N., Insight into the mechanism of the B12-independent glycerol dehydratase from Clostridium butyricum: preliminary biochemical and structural characterization. *Biochemistry* **2004,** *43* (16), 4635-4645.
21. LaMattina, J. W.; Keul, N. D.; Reitzer, P.; Kapoor, S.; Galzerani, F.; Koch, D. J.; Gouvea, I. E.; Lanzilotta, W. N., 1, 2-Propanediol dehydration in Roseburia inulinivorans structural basis for substrate and enantiomer selectivity. *J. Biol. Chem.* **2016,** *291* (30), 15515-15526.
22. Kalnins, G.; Kuka, J.; Grinberga, S.; Makrecka-Kuka, M.; Liepinsh, E.; Dambrova, M.; Tars, K., Structure and function of CutC choline lyase from human microbiota bacterium Klebsiella pneumoniae. *J. Biol. Chem.* **2015,** *290* (35), 21732-21740.
23. Xing, M.; Wei, Y.; Zhou, Y.; Zhang, J.; Lin, L.; Hu, Y.; Hua, G.; Urs, A. N. N.; Liu, D.; Wang, F., Radical-mediated C–S bond cleavage in C2 sulfonate degradation by anaerobic bacteria. *Nat. Commun.* **2019,** *10* (1), 1-11.
24. Conradt, H.; Hohmann-Berger, M.; Hohmann, H.-P.; Blaschkowski, H. P.; Knappe, J., Pyruvate formate-lyase (inactive form) and pyruvate formate-lyase activating enzyme of Escherichia coli: isolation and structural properties. *Arch. Biochem. Biophys.* **1984,** *228* (1), 133-142.
25. Sun, X.; Ollagnier, S.; Schmidt, P. P.; Atta, M.; Mulliez, E.; Lepape, L.; Eliasson, R.; Gräslund, A.; Fontecave, M.; Reichard, P., The free radical of the anaerobic ribonucleotide reductase from Escherichia coli is at glycine 681. *J. Biol. Chem.* **1996,** *271* (12), 6827-6831.
26. Wagner, A. F.; Frey, M.; Neugebauer, F. A.; Schäfer, W.; Knappe, J., The free radical in pyruvate formate-lyase is located on glycine-734. *Proc. Natl. Acad. Sci. U.S.A.* **1992,** *89* (3), 996-1000.
27. Kovačević, B.; Barić, D.; Babić, D.; Bilić, L.; Hanževački, M.; Sandala, G. M.; Radom, L.; Smith, D. M., Computational tale of two enzymes: glycerol dehydration with or without B12. *J. Am. Chem. Soc.* **2018,** *140* (27), 8487-8496.
28. Feliks, M.; Ullmann, G. M., Glycerol dehydratation by the B12-independent enzyme may not involve the migration of a hydroxyl group: a computational study. *J. Phys. Chem. B* **2012,** *116* (24), 7076-7087.
29. Rodrigues, A. V.; Tantillo, D. J.; Mukhopadhyay, A.; Keasling, J. D.; Beller, H. R., Insight into the mechanism of phenylacetate decarboxylase (PhdB), a toluene‐producing glycyl radical enzyme. *ChemBioChem* **2020,** *21* (5), 663.
30. Salii, I.; Szaleniec, M.; Zein, A. A.; Seyhan, D.; Sekuła, A.; Schühle, K.; Kaplieva-Dudek, I.; Linne, U.; Meckenstock, R. U.; Heider, J., Determinants for substrate recognition in the glycyl radical enzyme benzylsuccinate synthase revealed by targeted mutagenesis. *ACS Catal.* **2021,** *11*, 3361-3370.
31. Izrailev, S.; Stepaniants, S.; Isralewitz, B.; Kosztin, D.; Lu, H.; Molnar, F.; Wriggers, W.; Schulten, K., Steered molecular dynamics. In *Computational molecular dynamics: challenges, methods, ideas*, Springer: 1999; pp 39-65.

32. Yang, Z.; Mehmood, R.; Wang, M.; Qi, H. W.; Steeves, A. H.; Kulik, H. J., Revealing quantum mechanical effects in enzyme catalysis with large-scale electronic structure simulation. *React. Chem. Eng.* **2019,** *4* (2), 298-315.
33. Backman, L. R. F.; Huang, Y. Y.; Andorfer, M. C.; Gold, B.; Raines, R. T.; Balskus, E. P.; Drennan, C. L., Molecular basis for catabolism of the abundant metabolite trans-4-hydroxy-L-proline by a microbial glycyl radical enzyme. *Elife* **2020,** *9*, e51420.
34. Gao, J. L.; Truhlar, D. G., Quantum mechanical methods for enzyme kinetics. *Annu. Rev. Phys. Chem.* **2002,** *53*, 467-505.
35. Kulik, H. J.; Zhang, J.; Klinman, J. P.; Martinez, T. J., How large should the QM region be in QM/MM calculations? The case of catechol O-methyltransferase. *J. Phys. Chem. B* **2016,** *120* (44), 11381-11394.
36. Ufimtsev, I. S.; Martinez, T. J., Quantum Chemistry on Graphical Processing Units. 3. Analytical Energy Gradients, Geometry Optimization, and First Principles Molecular Dynamics. *J. Chem. Theory Comput.* **2009,** *5* (10), 2619-2628.
37. Kulik, H. J.; Luehr, N.; Ufimtsev, I. S.; Martinez, T. J., Ab initio quantum chemistry for protein structures. *J. Phys. Chem. B* **2012,** *116* (41), 12501-12509.
38. Yang, Z.; Liu, F.; Steeves, A. H.; Kulik, H. J., Quantum mechanical description of electrostatics provides a unified picture of catalytic action across methyltransferases. *J. Phys. Chem. Lett.* **2019,** *10* (13), 3779-3787.
39. Karplus, M., Molecular dynamics of biological macromolecules: A brief history and perspective. *Biopolymers* **2003,** *68* (3), 350-358.
40. Kokkonen, P.; Bednar, D.; Pinto, G.; Prokop, Z.; Damborsky, J., Engineering enzyme access tunnels. *Biotechnol. Adv.* **2019,** *37* (6), 107386.
41. Marques, S. M.; Daniel, L.; Buryska, T.; Prokop, Z.; Brezovsky, J.; Damborsky, J., Enzyme tunnels and gates as relevant targets in drug design. *Med. Res. Rev.* **2017,** *37* (5), 1095-1139.
42. Kingsley, L. J.; Lill, M. A., Substrate tunnels in enzymes: structure–function relationships and computational methodology. *Proteins: Struct., Funct., Bioinf.* **2015,** *83* (4), 599-611.
43. Marques, S. r. M.; Dunajova, Z.; Prokop, Z.; Chaloupkova, R.; Brezovsky, J.; Damborsky, J., Catalytic cycle of haloalkane dehalogenases toward unnatural substrates explored by computational modeling. *J. Chem. Inf. Model.* **2017,** *57* (8), 1970-1989.
44. Brezovsky, J.; Chovancova, E.; Gora, A.; Pavelka, A.; Biedermannova, L.; Damborsky, J., Software tools for identification, visualization and analysis of protein tunnels and channels. *Biotechnol. Adv.* **2013,** *31* (1), 38-49.
45. Fink, M. J.; Syrén, P.-O., Redesign of water networks for efficient biocatalysis. *Curr. Opin. Chem. Biol.* **2017,** *37*, 107-114.
46. Kulkarni, Y. S.; Liao, Q.; Petrović, D.; Krüger, D. M.; Strodel, B.; Amyes, T. L.; Richard, J. P.; Kamerlin, S. C. L., Enzyme architecture: modeling the operation of a hydrophobic clamp in catalysis by triosephosphate isomerase. *J. Am. Chem. Soc.* **2017,** *139* (30), 10514-10525.
47. Levin, B. J.; Huang, Y. Y.; Peck, S. C.; Wei, Y.; Martinez-Del Campo, A.; Marks, J. A.; Franzosa, E. A.; Huttenhower, C.; Balskus, E. P., A prominent glycyl radical enzyme in human gut microbiomes metabolizes trans-4-hydroxy-l-proline. *Science* **2017,** *355* (6325), eaai8386.
48. Levin, B. J.; Balskus, E. P., Discovering radical-dependent enzymes in the human gut microbiota. *Curr. Opin. Chem. Biol.* **2018,** *47*, 86-93.

49. Riplinger, C.; Neese, F., An efficient and near linear scaling pair natural orbital based local coupled cluster method. *J. Chem. Phys.* **2013,** *138* (3), 034106.
50. Laroff, G. P.; Fessenden, R. W., Equilibrium and kinetics of the acid dissociation of several hydroxyalkyl radicals. *J. Phys. Chem.* **1973,** *77* (10), 1283-1288.
51. Levin, B. J.; Huang, Y. Y.; Peck, S. C.; Wei, Y.; Martínez-del Campo, A.; Marks, J. A.; Franzosa, E. A.; Huttenhower, C.; Balskus, E. P., A prominent glycyl radical enzyme in human gut microbiomes metabolizes trans-4-hydroxy-l-proline. *Science* **2017,** *355* (6325), eaai8386.
52. Craciun, S.; Marks, J. A.; Balskus, E. P., Characterization of choline trimethylamine-lyase expands the chemistry of glycyl radical enzymes. *ACS Chem. Biol.* **2014,** *9* (7), 1408-1413.
53. Bodea, S.; Funk, M. A.; Balskus, E. P.; Drennan, C. L., Molecular basis of C–N bond cleavage by the glycyl radical enzyme choline trimethylamine-lyase. *Cell Chem. Biol.* **2016,** *23* (10), 1206-1216.
54. Levin, B. J.; Balskus, E. P., Characterization of 1,2-propanediol dehydratases reveals distinct mechanisms for B12-dependent and glycyl radical enzymes. *Biochemistry* **2018,** *57* (23), 3222-3226.
55. Gilli, G.; Gilli, P., Towards an unified hydrogen-bond theory. *J. Mol. Struct.* **2000,** *552* (1-3), 1-15.
56. Cleland, W. W.; Kreevoy, M. M., Low-barrier hydrogen bonds and enzymic catalysis. *Science* **1994,** *264* (5167), 1887-1890.
57. Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H.; Olsson, M. H. M., Electrostatic basis for enzyme catalysis. *Chem. Rev.* **2006,** *106* (8), 3210-3235.
58. Bondi, A. v., van der Waals volumes and radii. *J. Phys. Chem.* **1964,** *68* (3), 441-451.
59. Hur, S.; Bruice, T. C., The mechanism of catalysis of the chorismate to prephenate reaction by the Escherichia coli mutase enzyme. *Proc. Natl. Acad. Sci. U.S.A.* **2002,** *99* (3), 1176-1181.
60. Mehmood, R.; Qi, H. W.; Steeves, A. H.; Kulik, H. J., The protein's role in substrate positioning and reactivity for biosynthetic enzyme complexes: the case of SyrB2/SyrB1. *ACS Catal.* **2019,** *9* (6), 4930-4943.
61. Petřek, M.; Otyepka, M.; Banáš, P.; Košinová, P.; Koča, J.; Damborský, J., CAVER: a new tool to explore routes from protein clefts, pockets and cavities. *BMC Bioinformatics* **2006,** *7* (1), 316.
62. Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T. B.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X., Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **2015,** *113* (2), 184-215.
63. Anandakrishnan, R.; Aguilar, B.; Onufriev, A. V., H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res.* **2012,** *40* (W1), W537-W541.
64. Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A., A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **1993,** *97* (40), 10269-10280.
65. Becke, A. D., Density-functional thermochemistry. III. The role of exact exchange (1993) J. *Chem. Phys 98*, 5648.
66. Marenich, A. V.; Cramer, C. J.; Truhlar, D. G., Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *J. Phys. Chem. B* **2009,** *113* (18), 6378-6396.

67. Cerjan, C. J.; Miller, W. H., On finding transition states. *J. Chem. Phys.* **1981,** *75* (6), 2800-2806.
68. Banerjee, A.; Adams, N.; Simons, J.; Shepard, R., Search for stationary points on surfaces. *J. Phys. Chem.* **1985,** *89* (1), 52-57.
69. Simons, J.; Joergensen, P.; Taylor, H.; Ozment, J., Walking on potential energy surfaces. *J. Phys. Chem.* **1983,** *87* (15), 2745-2753.
70. Farkas, Ö.; Schlegel, H. B., Methods for optimizing large molecules. Part III. An improved algorithm for geometry optimization using direct inversion in the iterative subspace (GDIIS). *Phys. Chem. Chem. Phys.* **2002,** *4* (1), 11-15.
71. Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R., Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminformatics* **2012,** *4* (1), 17.
72. Sheppard, D.; Terrell, R.; Henkelman, G., Optimization methods for finding minimum energy paths. *J. Chem. Phys.* **2008,** *128* (13), 134106.
73. Heyden, A.; Bell, A. T.; Keil, F. J., Efficient methods for finding transition states in chemical reactions: Comparison of improved dimer method and partitioned rational function optimization method. *J. Chem. Phys.* **2005,** *123* (22), 224101.
74. Neese, F., The ORCA program system. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012,** *2* (1), 73-78.
75. Ufimtsev, I. S.; Martínez, T. J., Quantum chemistry on graphical processing units. 3. analytical energy gradients, geometry optimization, and first principles molecular dynamics. *J. Chem. Theory Comput.* **2009,** *5*, 2619-2628.
76. Ufimtsev, I. S.; Martínez, T. J., Quantum chemistry on graphical processing units. 1. strategies for two-electron integral evaluation. *J. Chem. Theory Comput.* **2008,** *4*, 222-231.
77. Ufimtsev, I. S.; Martinez, T. J., Quantum chemistry on graphical processing units. 2. direct self-consistent-field implementation. *J. Chem. Theory Comput.* **2009,** *5* (4), 1004-1015.
78. Rohrdanz, M. A.; Martins, K. M.; Herbert, J. M., A long-range-corrected density functional that performs well for both ground-state properties and time-dependent density functional theory excitation energies, including charge-transfer excited states. *J. Chem. Phys.* **2009,** *130* (5), 054112.
79. Harihara, P. C.; Pople, J. A., Influence of polarization functions on molecular-orbital hydrogenation energies. *Theor. Chim. Acta.* **1973,** *28* (3), 213-222.
80. Wang, L.-P.; Song, C., Geometry optimization made simple with translation and rotation coordinates. *J. Chem. Phys.* **2016,** *144* (21), 214108.
81. Lange, A. W.; Herbert, J. M., A smooth, nonsingular, and faithful discretization scheme for polarizable continuum models: The switching/Gaussian approach. *J. Chem. Phys.* **2010,** *133* (24), 244111.
82. York, D. M.; Karplus, M., A smooth solvation potential based on the conductor-like screening model. *J. Phys. Chem. A.* **1999,** *103* (50), 11060-11079.
83. Liu, F.; Luehr, N.; Kulik, H. J.; Martinez, T. J., Quantum Chemistry for Solvated Molecules on Graphical Processing Units Using Polarizable Continuum Models. *J. Chem. Theory Comput.* **2015,** *11* (7), 3131-3144.
84. Liu, F.; Sanchez, D. M.; Kulik, H. J.; Martinez, T. J., Exploiting graphical processing units to enable quantum chemistry calculation of large solvated molecules with conductor-like polarizable continuum models. *Int. J. Quantum Chem.* **2019,** *119* (1), e25760.

85. Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A., H++: a server for estimating pKas and adding missing hydrogens to macromolecules. *Nucleic Acids Res.* **2005,** *33* (suppl 2), W368-W371.
86. Myers, J.; Grothaus, G.; Narayanan, S.; Onufriev, A., A simple clustering algorithm can be accurate enough for use in calculations of pKs in macromolecules. *Proteins: Struct., Funct., Bioinf.* **2006,** *63* (4), 928-938.
87. D.A. Case, J. T. B., R.M. Betz, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, T. Luchko, R. Luo, B. Madej, K.M. Merz, G. Monard, P. Needham, H. Nguyen, H.T. Nguyen, I. Omelyan, A. Onufriev, D.R. Roe, A. Roitberg, R. Salomon-Ferrer, C.L. Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, D.M. York and P.A. Kollman AMBER 2015, University of California, San Francisco. 2015.
88. 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* (8), 3696-3713.
89. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A., Development and testing of a general amber force field. *J. Comput. Chem.* **2004,** *25* (9), 1157-1174.
90. F. Wang, J.-P. B., P. Cieplak and F.-Y. Dupradeau R.E.D. Python: Object oriented programming for Amber force fields, Université de Picardie - Jules Verne, Sanford|Burnham Medical Research Institute, Nov. 2013. http://q4md-forcefieldtools.org/REDServer-Development/ (accessed 04/20/2021).
91. Vanquelef, E.; Simon, S.; Marquant, G.; Garcia, E.; Klimerak, G.; Delepine, J. C.; Cieplak, P.; Dupradeau, F.-Y., R.E.D. Server: a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments. *Nucleic Acids Res.* **2011,** *39* (suppl 2), W511-W517.
92. Dupradeau, F.-Y.; Pigache, A.; Zaffran, T.; Savineau, C.; Lelong, R.; Grivel, N.; Lelong, D.; Rosanski, W.; Cieplak, P., The R.E.D. tools: advances in RESP and ESP charge derivation and force field library building. *Phys. Chem. Chem. Phys.* **2010,** *12* (28), 7821-7839.
93. Gordon, M. S.; Schmidt, M. W., Advances in electronic structure theory: GAMESS a decade later. *Theory and Applications of Computational Chemistry: the first forty years* **2005**, 1167-1189.
94. 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* (2), 926-935.
95. Götz, A. W.; Williamson, M. J.; Xu, D.; Poole, D.; Le Grand, S.; Walker, R. C., Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. generalized Born. *J. Chem. Theory Comput.* **2012,** *8* (5), 1542-1555.
96. Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C., Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. explicit solvent particle mesh ewald. *J. Chem. Theory Comput.* **2013,** *9* (9), 3878-3888.
97. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C., Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977,** *23* (3), 327-341.
98. Torrie, G. M.; Valleau, J. P., Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* **1977,** *23* (2), 187-199.

99. Souaille, M.; Roux, B., Extension to the weighted histogram analysis method: combining umbrella sampling with free energy calculations. *Comput. Phys. Commun.* **2001,** *135* (1), 40-57.
100. Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A., The weighted histogram analysis method for free‐energy calculations on biomolecules. I. The method. *J. Comput. Chem.* **1992,** *13* (8), 1011-1021.
101. Grossfield, A. WHAM: the weighted histogram analysis method, version 2.0.9. http://membrane.urmc.rochester.edu/ (accessed 04/20/2021).

TOC Graphic