Substrate Conformation Masquerades Two different Oxidants in Aromatic C-H and C-F bond activation by Heme-dependent Tyrosine Hydroxylase

Warispreet Singh*[a,b], Sonia F Santos[a,b], Shalini Yadav[d], Gary W Black [a,b], Vera Krewald [c], Kshatresh Dutta Dubey [d]*

[a] Department of Applied Sciences, Northumbria University, Newcastle upon Tyne, NE1 8ST, United Kingdom

[b] Hub for Biotechnology in Build Environment, Newcastle upon Tyne, NE1 8ST, United Kingdom

[c]Department of Chemistry, Theoretical Chemistry, Technical University of Darmstadt, Darmstadt, Germany

[d] Department of Chemistry and Center for Informatics, School of Natural Sciences, Shiv Nadar University, Gautam Buddha Nagar, Uttar Pradesh-201314, India

Email: kshatresh.dubey@snu.edu.in, w.singh@northumbria.ac.uk

ABSTRACT

A recently discovered heme-dependent enzyme TyrH offers a green approach for functionalizing the high-strength C-H and C-F bonds in aromatic compounds. However, there is ambiguity regarding the nature of intermediates (Cpd 0 or Cpd I) involved in activating these bonds. Herein, using comprehensive MD Simulations and hybrid QM/MM calculations, we reveal that it is Compound I (Cpd I) that acts as the primary oxidant involved in the functionalization of both the C-F and C-H bonds. Our study shows that the His88 close to the catalytic site acts as a modulator of aromaticity and is essential to initiate the C-H and C-F bond activation. The reaction proceeds via the proton abstraction from the 4-OH group of the substrate by His88, which is coupled with an electron transfer to form a single-electron reduced Cpd I-like intermediate. This species then inserts the oxygen atom into the sp^3 hybridized C-H and C-F bond of the substrate. Interestingly, our mechanistic study shows that the two different conformations of the substrate are actually responsible for the C-F and C-H functionalization, which masquerades as two different characters of oxidants involved in the reaction.

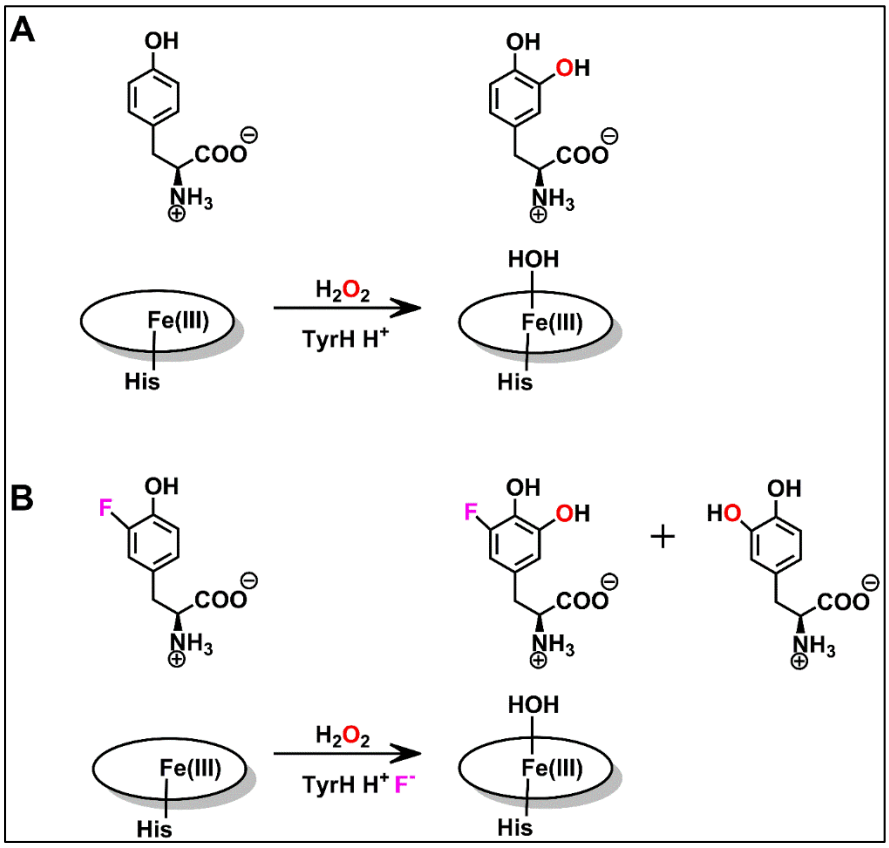
Keywords: Heme enzymes, MD, QM/MM, C-H bond, C-F bond, and arenes

1. INTRODUCTION

Organo-fluorine compounds have widespread usage in the pharmaceutical industry and material sciences due to the exceptionally high strength of the C-F bond (bond dissociation energy of C-F is 130 kcal/mol). [1-3] However, fluorochemicals in the form of per- and poly-fluoroalkyl substances are categorized as an environmental hazard,s, [4] [5], and therefore, the breakdown of these fluorochemicals in an eco-friendly manner has been an active area of research for the past few decades.[6] For instance, the C(sp^2)−H bond in an aromatic compound has a bond dissociation energy of 113 kcal/mol, which is significantly higher than any aliphatic C(sp^3)−H bond [7]. Therefore, the functionalization of the C-H bond in an aromatic substrate by greener and sustainable approaches are some of the most intriguing areas in biochemistry.

Traditionally, most synthetic metal complexes activate the C-H or C-F bond via a high-valent metal–oxo species under harsh chemical and environmental conditions.[8-10] Recently, a new heme-based biocatalyst, Tyrosine Hydroxylase (TyrH), has been reported to selectively activate both C(sp^2)−H and C-F bonds under mild reaction conditions.[11] In contrast to traditional CYP450 enzymes where a cysteine residue is ligated at the proximal position of heme, the TyrH enzyme is a histidine ligated heme-dependent tyrosine hydrolase that uses hydrogen peroxide for the C-H activation. TyrH catalyzes the oxidation of Tyr to produce L-3,4-dihydroxyphenylalanine (DOPA) using L-tyrosine (Tyr) as a native substrate. [11] However, in the presence of the 3-fluoro-L-tyrosine (3-F-Tyr) as a substrate, it oxidizes 3-F-Tyr to 3-fluoro-5-hydroxy-L-tyrosine and DOPA in 1.4:1 as shown in Scheme 1. Very recently, the X-ray structure of TyrH in complex with the Tyr and 3-F-Tyr was elucidated [11], which shows two identical binding conformations of the native and 3-F-Tyr substrates in the active site of TyrH. However, the fluorine atom of the C-F bond in the 3-F-Tyr showed two orientations. In a dominant conformation A, the fluorine atom faces upward from the heme plane and vice versa for conformation B. This alternative orientation of the fluorine atom is suggested to be a

plausible explanation for the different product distribution during the hydroxylation of 3-F-Tyr [11].

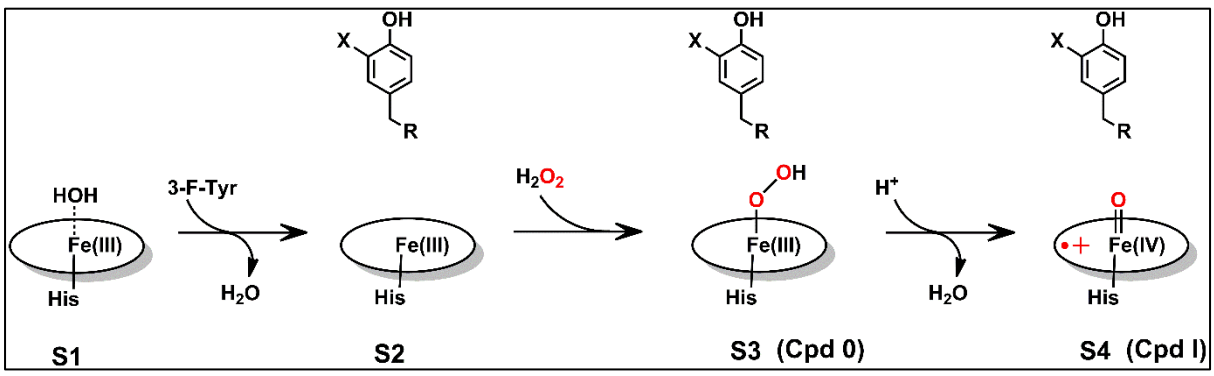


Scheme 1. Reaction catalysed by TyrH enzyme. (A) Oxidation of native substrate tyrosine to L-3,4-dihydroxyphenylalanine (DOPA) and (B) Hydroxylation of 3-fluoro-L-tyrosine substrate, to 3-Fluoro-5-hydroxy-L-tyrosine and DOPA. The proton and the fluoride anion are the leaving groups in the above reaction.

The presence of the 4-OH group in both the substrates is critical for the activation of the C-H and C-F bonds. The substrates analogs where the 4-OH group is replaced, such as phenylaniline and O-methyltyrosine, show no enzyme activity. [12] In the active site of TyrH, the 4-OH group is stabilized by hydrogen bonds with the side chain of highly conserved His88 and Tyr230 residues. The site-directed mutagenesis of both these residues suggests an indispensable role of His88 in catalysis. However, the juxtaposition of these residues is essential for C-F bond activation, as evident by the double variant H88Y/Y230H. Another interesting feature revealed by the X-ray structure is the out-of-plane conformation of one of the propionate side chains. [11]

The out-of-plane propionate is involved in hydrogen bonds with the side chain of S157 and S209 protein residues. The out-of-plane conformation is different from the other heme-dependent enzymes, where both the propionate side chains exhibit in-plane conformation relative to the heme group. How this out-of-plane conformation contributes to catalysis is an intriguing question that needs answering.

The possible key steps involved in the catalytic cycle of the TryH enzyme are shown in Scheme 2. In the resting state (S1) of the enzyme, Fe(III) is hexacoordinated and is bound to an axial aqua ligand. The EPR study and the theoretical investigations confirm that the binding of the substrate in the active site results in the displacement of the aqua ligand and the production of pentacoordinate high spin Fe(III) species (S2). [11] The vacant site is then occupied by molecular hydrogen peroxide, which deprotonates to form a ferric-hydroperoxo intermediate also known as Cpd 0 (S3). From a mechanistic point of view, it has been suggested that either a Cpd 0 or Cpd 1 can perform oxidation of both the native and 3-F-Ty substrates. [11-12] To shed light on the reaction mechanism of the C-H and C-F bond activation, we performed MD, computational UV-Vis spectroscopy and QM/MM calculations in the resting, Cpd 0 and Cpd I state of the TyrH in complex with the native Tyr and 3-F-Tyr substrate to address two main questions (a) what is the true oxidant Cpd 0 or Cpd I involved in the C-H and C-F bond activation (b) role of the out of plane propionate heme group in the formation of Cpd0 and Cpd I.



Scheme 2. Reaction intermediates involved in the C-H and C-F bond activation of the substrate in TyrH enzyme. X=F or X=H represents the 3-F-Tyr and native tyrosine substrate.

METHODS

2.1 Structure preparation for the resting, Cpd 0 and Cpd I

To understand the C-H and C-F bond functionalisation in TyrH, we focused our attention on the S1, S2, Cpd0, Cpd I intermediates of the TyrH in complex with the native and 3-F-Tyr substrate. The initial structure of the TyrH in complex with the Heme, Tyr-F(Conf A or B) and the hydroperoxo complex was obtained from the crystal structures (PDB: 7KQR, 7KQU).[11] To make the starting structure to be Fe(III)-hydroperoxo aka Cpd 0, we protonated the distal oxygen of the peroxo complex using the USCF chimera. [13] The Cpd I structure was prepared by first inspecting the crystal structure of TyrH in complex with native Tyr substrate and heme group. The water molecule (605 chain B) which was in the closed vicinity of the Fe atom was chosen as the oxygen atom of the Fe(IV)=O motif. The oxygen atom of this water molecule was moved close to the iron in the heme group. In an analogous manner to make the resting state of enzyme, we used this water molecule and transform it into a hydrogen peroxide. In all the setups of various states, we decided to use the chain B of the enzyme because two subunits structurally resemble each other with a rmsd value of 0.56Å. [11] The monomeric state is predominant in the solution and is functionally active so only single chain is sufficient to perform the computational analysis. Also, Chain A of the molecule was missing some residues in the loop region, so we decided to use chain B for all the subsequent modelling studies. The X-ray crystallographic waters were kept and were used to run the molecular dynamics simulations. The H++ server was used to assign the protonation states to the titratable residues at pH=7.0 and using the default settings of the H++ server. [14] This resulted in all the Glu, Asp to be negatively charged, all the Arg, Lys were positively charged. There were no disulphide bridges located and all the cysteines were in their reduced sate. The side chain of His196

coordinating the Heme group was protonated at the delta position. His88 residue was also protonated at delta position to facilitate hydrogen bonding with the hydroxyl group of the substrate.

2.2 Parameterization of the substrates and the active site

The parameters for native substrate tyrosine and 3-F-Tyr were developed using the general Amber force field (GAFF [15]). The partial charges were calculated according to the Merz–Singh–Kollman scheme using Gaussian 16 [16] at HF/6-31G* of theory using the RESP (restrained electrostatic potential) method.[17] The parameters for the resting state, Cpd 0 and Cpd 1 were obtained using MCPB.py [18], and the rest of the protein molecule was modelled using FF19SB force field.[19]

2.3 MD simulations

The enzyme substrate complexes in its various intermediate states were simulated in a truncated octahedral TIP3P (Transferable intermolecular potential 3P) water box.[20] The charge neutrality of each complex was maintained by adding appropriate counter ions using tleap module available in AmberTools20.[21] The protein molecule in the each simulation box was kept at a distance of 10 Å away from the box edges and periodic boundary conditions were used in each simulation run. The long-range electrostatic interactions were computed using Particle mesh Ewald with a cutoff value of 10 Å.[22] Prior to running productive MD simulations, each enzyme-substrate complex was relaxed by energy minimized using 5000 steepest descent and conjugate gradient method. Each system was then subjected to control heating from 0 to 298.15 K for 50 ps using Langevin thermostat with a collision frequency of 1 ps^{-1}. The harmonic potential of 5 kcal mol^{-1} $Å^2$ was applied to all the non-hydrogen atoms of the protein during the heating process. The entire system was then subjected to two rounds of equilibration at 298.15K for 50ps in an NPT ensemble. A Berendsen barostat was used to maintain the

pressure at 1 bar and the SHAKE algorithm was used to constrain bonds involving hydrogen. [23] A time step of 2 fs was used for all MD runs; production MD simulations were run 500ns in an NPT ensemble for each system. The Compute Unified Device Architecture (CUDA) version of particle mesh Ewald molecular dynamics (PMEMD) [24] was used to run all of the MD simulations using graphics processing units (GPUs) in Amber20.[21]

2.4 QM/MM calculations

The QM/MM calculations were performed using Chemshell 3.7.0 [25] to study the functionalization of the C-H and C-F bond of 3-F-Tyr substrate in complex with the TyrH enzyme. The QM calculations were performed using ORCA 4.2.0 [26 27] and the MM part with DL_POLY [28] using the FF19SB force field. [19] The def2-SVP def2/J auxiliary basis sets were used for all of the atoms except for transition metal iron, which was treated with def2-TZVP. DFT using the UB3LYP functional [29 30] with D3 dispersion correction and BJ damping [31] was used to run all the QM calculations at three spin states S=1/2, S=3/2, and S=5/2. The potential energy scans were run with a step size of 0.1Å using the def2-SVP basis set on all the atoms and def2-TZVP on iron with D3 dispersion correction and BJ damping. The effect of the protein environment on the polarization of the QM wavefunction was described by the electronic embedding scheme. The QM region consisted of His88, Tyr230, His198, Heme cofactor, hydrogen peroxide, 3-F-Tyr substrate, and water molecules. The choice of the QM region was based upon the benchmarking done by adding Y230 and then water molecules to the QM region. The addition of Y230 and water molecules near the iron reduced the activation energy barrier so were included in the final QMMM calculations. The residues which are within 10Å of the heme including water molecules were allowed to move freely and the rest of the system was frozen during the geometry optimization. The hydrogen link atoms were used to saturate the dangling bond at the QM/MM boundary. The transition state structures were

fully optimised using the dimer method. The transition state structures were validated by the presence of single imaginary frequency using the thermal keyword. The validation of local minima structures was conducted by the presence of no imaginary frequency. The single point corrections and Mulliken spin population analysis was carried on the stationery and saddle points using UB3LYP functional with def2-TZVP basis set on all the atoms with D3 dispersion correction and BJ damping.

The QM/MM simulated Ultraviolet-visible spectra were computed using the Time-dependent DFT (TDDFT) calculations on the optimized stationary points in the resting and Cpd 0 states of both the native and 3-F-Tyr substrates. The QM part was defined using ORCA 4.2.0 [26] [27] and the MM part with DL_POLY [28] using the FF19SB force field. [19]. The RIJCOSX and TDA approximation, def2-TZVP basis set on all the atoms with D3 dispersion correction and BJ damping and FinalGrid6 used in all the QM calculations. The excited state exited state properties, i.e., excitation energies, oscillator strengths, and electronic characters of the 150 excited states, were computed.

3. RESULTS AND DISCUSSION

3.1. Formation of active oxidants, Cpd0 and Cpd I from the resting state

The TyrH enzyme utilizes hydrogen peroxide to produce active oxidants (either Cpd 0 or Cpd I). Therefore, before any mechanistic elucidation, we must validate the feasibility of an efficient formation of both oxidants from H2O2 in TyrH. In addition, though the formation of Cpd I from H_2O_2 in cysteine-ligated CYP450 is well established [32], the mechanism of formation of active oxidants (Cpd 0 and/or Cpd I) in Histidine-ligated TyrH is yet to be elucidated. We, therefore, started our study to explore the mechanism of the formation of active oxidants in TyrH enzyme. In so doing, we performed several MD simulations for the **S2** (i.e., resting state)

with both Tyr and 3-F-Tyr substrates in the presence of H_2O_2, separately (See S.I figure S1 for RMSD analysis). In all simulations (for both substrates), the H_2O_2 molecule was flexible and rotated freely in the active site as anticipated by the large dihedral angle fluctuations (**Figure 1c**). However, the average distance of the H_2O_2 molecule from the heme iron is just ~ 3Å (See S.I figure S2 for trajectory distances) which shows that the H_2O_2 molecule may easily be accessible to react with heme-iron *en route* to the formation of the active oxidants.

Interestingly, we found two conformations of the propionate side chain of the porphyrin ring during simulations. In a major conformation, both propionate side chains, Prop 6 and Prop7 (Figure 1 a), are lodged in mutually opposite orientations. On the other hand, both propionate side chains occupy the same plane of orientations in a minor conformation (Figure 1b). For the sake of simplicity, we refer former as '*out-plane*' and later as '*in-plane'* conformations in further discussions. In '*out-plane*' conformation, the propionate side chain, **Prop-7**, forms a strong hydrogen bond with S209, and connects proximal hydrogen of H_2O_2 via a bridging water molecule W1.

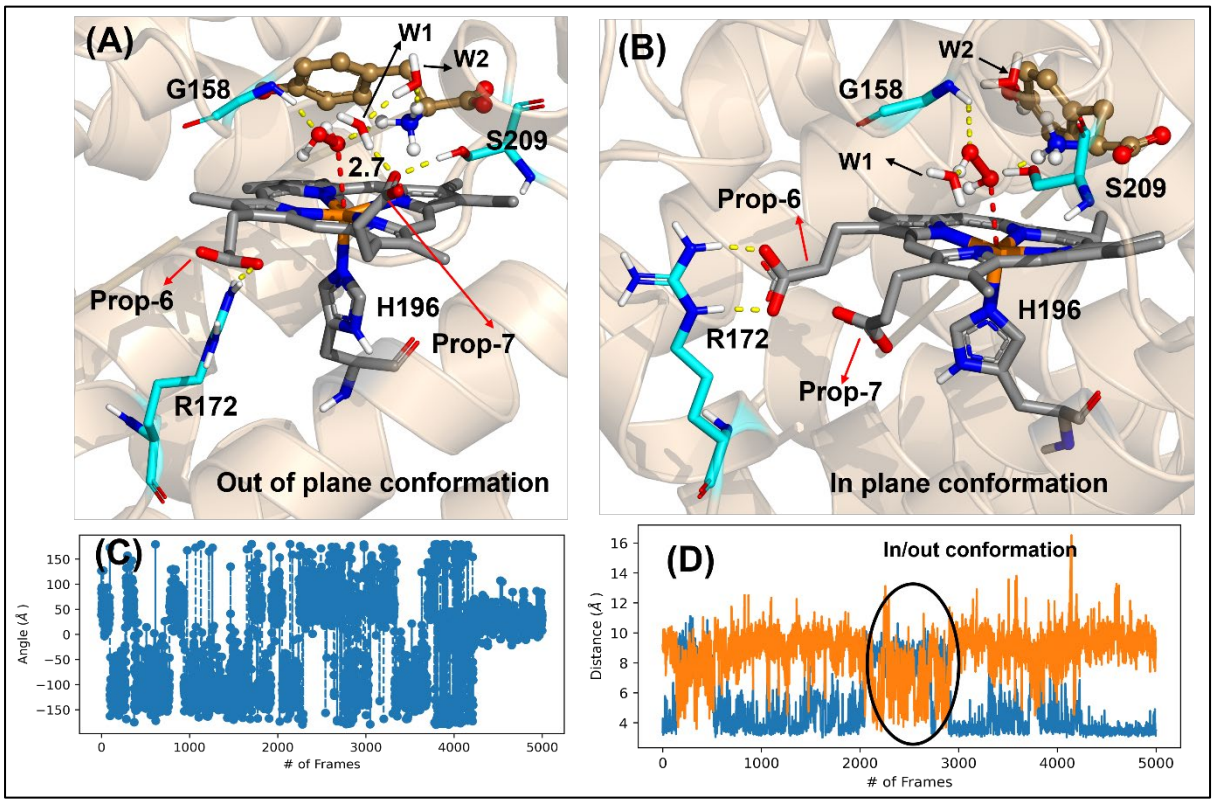


Figure 1 The interactions of hydrogen peroxide in the active site of TyrH in complex with native substrate Tyr. (A) 'Out plane' conformation of the propionate side chain, (B) In-plane conformation of the propionate side chain, yellow lines indicate hydrogen bonds. (C) dihedral angle O2-O1-Fe1-NA showing the rotation of hydrogen peroxide in the active site of TyrH enzyme and (D) distances between the carboxylate group of propionate and the side chain of S209 and R172 shown in blue and orange colors, respectively. The circle represents the transition to the in-plane conformation.

As mentioned earlier, we performed separate simulations with the other substrate, 3-F-Tyr to study the effect of substrate binding in the active site. Interestingly, during the MD simulations, we see two different conformations of substrate similar to the experimental co-crystal with the same substrate (cf. Figure 2). Concurrent with the experimental finding, our MD simulations also show that conformation A where the C-F bond is facing away from the imidazole ring of H88, is the dominant conformation. Furthermore, to repudiate any artifact of the simulation to overcome the local minima of conformation A, we performed a separate MD simulation starting from conformation B of 3-F-Tyr to see if the substrate flips to a more dominant conformation A. To our delight, the conformation B flipped to conformation A, and it maintained this conformation throughout the simulations (Figure 2) after 200ns of MD simulations. Therefore, our MD simulations show a strong correlation with the experimental observation. In addition, the propionate side chains also exit in two conformations where 'out plane' conformation dominates, as discussed for substrate Tyr.

In a nutshell, during the MD simulations of both substrates (i.e., Tyr and the 3-F-Tyr) in the resting state with H_2O_2, the active site is well hydrated, and the 'out plane' conformation of the propionate side chain is the major conformation. Therefore, we believe *this out-of-plane conformation may be of evolutionary importance with implications for the catalysis of TyrH activity.*

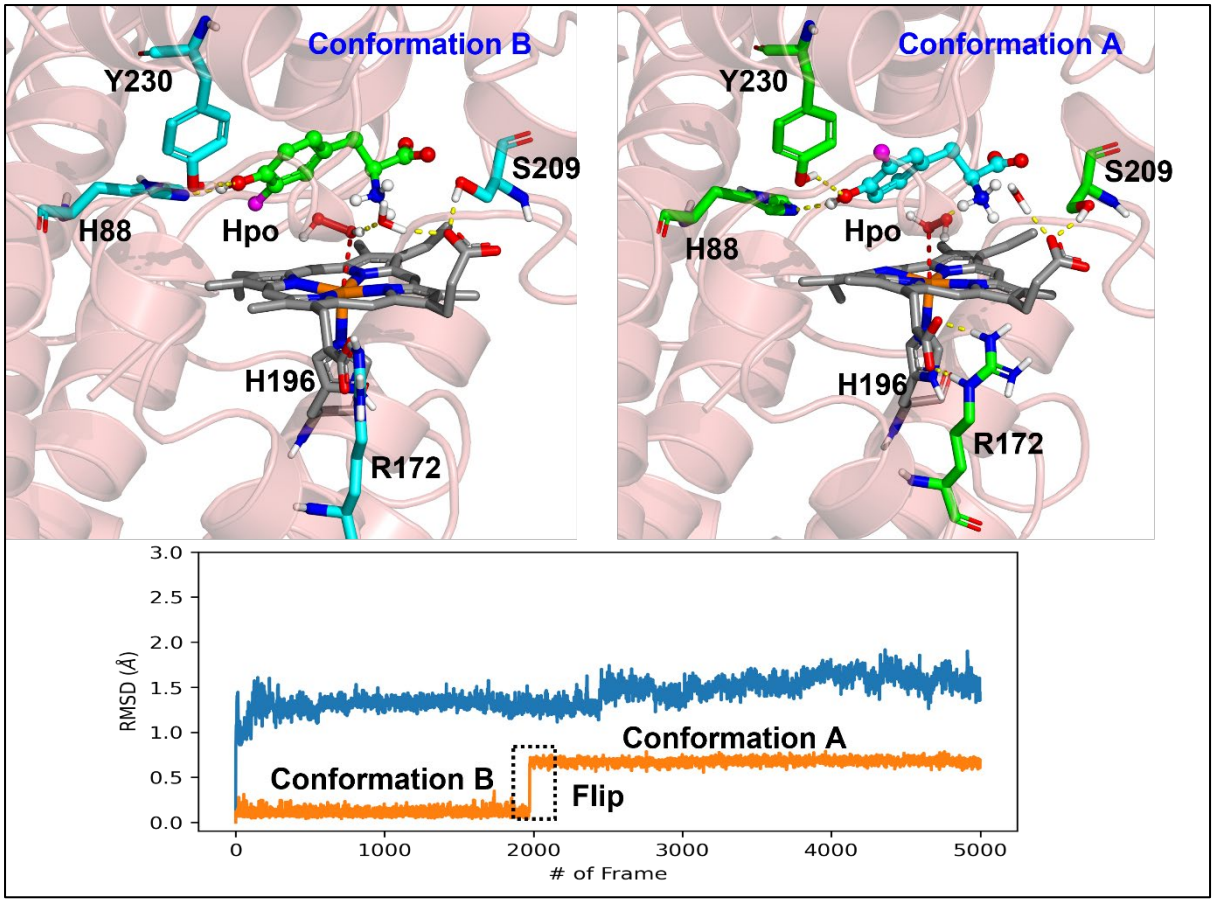


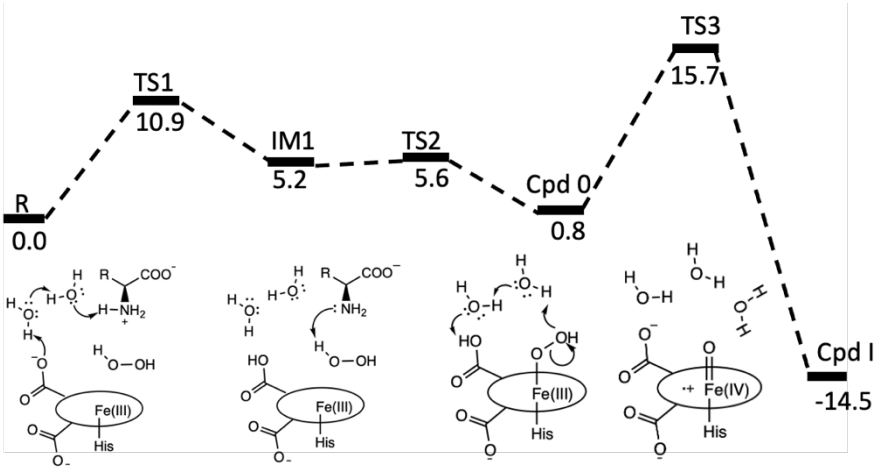
Figure 2 The most populated cluster obtained from the clustering of the TryH enzyme in the resting state in a complex with 3-F-Tyr. Here we have shown the two conformations of the 3-F-Tyr observed in the MD simulations. Yellow lines indicate hydrogen bonds. RMSD analysis of TyrH in complex with the 3-Tyr-F substrate. The blue and orange insert represent the C-alpha atoms of the protein and the heavy atoms of the 3-F-Tyr substrate

3.2. Mechanism of Cpd 0 and Cpd I formation:

In the previous section, we have seen that *out-plane* conformation is dominating, and proximal hydrogen of the H_2O_2 is connected with the propionate side chain via a well-organized water chain through the substrate. Hence, we propose that Cpd 0 may be formed via deprotonation of the hydrogen peroxide through the propionate side chain, as shown in Figure 3. As can be seen, the proton from the proximal oxygen of the hydrogen peroxide can be shuttled to the propionate side chain via the amine group of the substrate in association with nearby water molecules to form Cpd 0. To validate our proposed scheme, we performed the QM/MM calculations of a representative snapshot from the MD simulation's most populated

('out-plane' conformation). After geometry optimization of the reactant species, we performed a potential energy surface (PES) scanning for the deprotonation of the proximal oxygen of the H_2O_2. Figure 3 shows the reaction coordinate and the energy profile for the reactions from the QM/MM calculations. As can be seen, deprotonation of the amine group of the substrate is initiated by the propionate via the water molecule, and it went through the TS1 with a barrier of approximately 11 kcal/mol. In IM1, the propionate side chain is now protonated, and the amine of the substrate is in its neutral state. Since the amine nitrogen is just 3.32 Å away from the proximal hydrogen of the H_2O_2, it further abstracts the proximal proton of the H_2O_2 to form Cpd 0 in a barrier-less manner.

To explore whether the so-formed Cpd 0 can produce Cpd I, we extended the PES for the Cpd I formation using PC1 as the starting structure. As can be seen, the distance between the hydrogen atom of the protonated propionate and the distal oxygen of the Cpd 0 is just 4.1 Å, and it is well bridged through a water molecule W1; it can easily protonate the distal oxygen of the Cpd 0 via water W1. It results in the homolytic breakdown of the O-O bond of Cpd 0, which further leads to the formation of Cpd I and the release of the water molecule as shown in Figure 3. The formation of Cpd I from Cpd 0 goes through a TS barrier of 16.5 kcal/mol. These results match well with the experiment ^{18}O-labeling data that the oxygen in the product is from the hydrogen peroxide. [12]



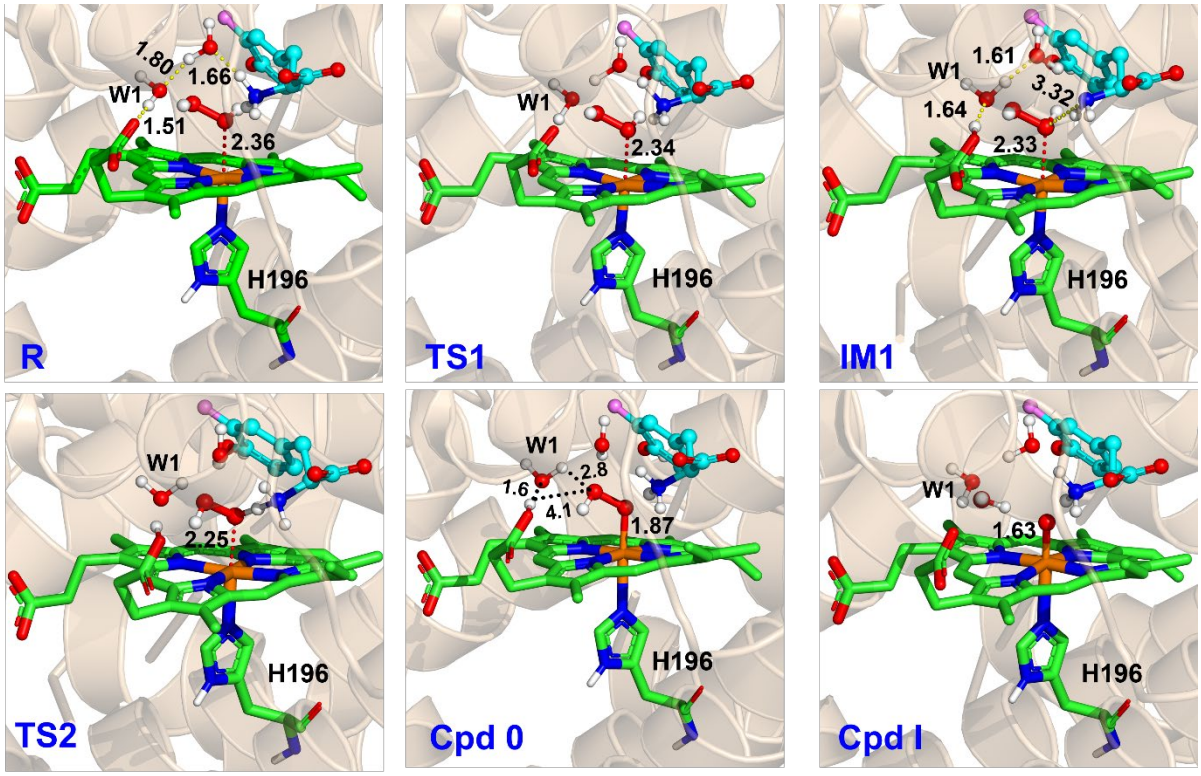


Figure 3 QM/MM reaction profile of the formation of Cpd 0 from the resting state of the TryH enzyme in complex with the 3-F-Tyr substrate at S=½.

3.3. Absorption spectroscopy of TyrH in complex with native and 3-F-Tyr

Wang et al. characterized the Cpd 0 formation by reacting the H_2O_2 with the native and 3-F-Tyr substrates. The absorption spectra indicated the formation of Cpd 0 at 525 and 554 nm compared to native and 3-F-Tyr without the presence of H_2O_2. Spectroscopy is a valuable link between experiment and theory; therefore, we computed the spectra of the S1, S2, and Cpd0 states in complex with both the native and 3-F-Tyr substrate. Our theoretically calculated spectra qualitatively reproduce the experimental spectroscopic finding of TyrH in complex with the native and 3-F-Tyr with and without the presence of H_2O_2, respectively (Figure 4A). The computed spectra also showed that the conformation B of the 3-F-Tyr substrate is not favored as it has low intensity compared to conformation A of the 3-F-Tyr substrate. So, the experimental spectra indicate the binding of the conformation A to H_2O_2 and subsequent formation of Cpd 0. This result is also in accordance with the obtained X-ray structure of Cpd

0 in complex with 3-F-Tyr, indicating the greater occupancy of Conf A in contrast to B. The computed spectra of Cpd 0 show two peaks at 493.2 nm and 557.6 nm, which are in reasonable agreement with experimental bands 525 and 554 nm for Cpd 0 formation (Figure 4B).[11] The transition at 493.2nm (see S.I figure) corresponds to the excitation of the beta electron from the porphyrin type orbital to the virtual dyz beta orbital of the iron, and the transition at 557.6nm corresponds to the excitation of the HOMO alpha electron from dxy to the LUMO+1 porphyrin type orbital.

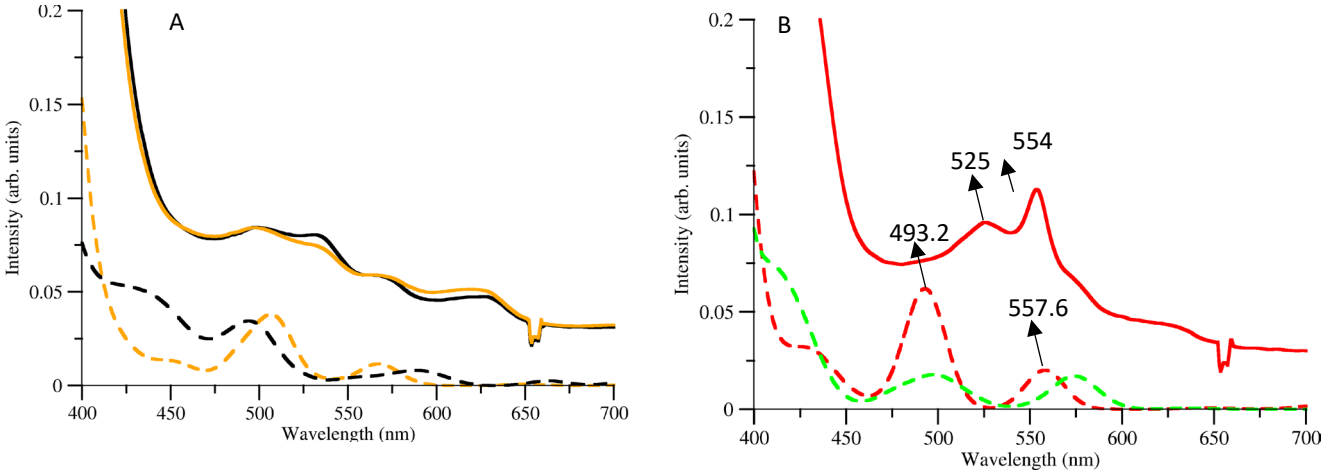


Figure 4 (A) Absorption spectroscopy of the TyrH enzyme in complex with the native Tyr and 3-F-Tyr substrates. The experimental (solid black line) and computed (dashed black line) indicate the spectra of 3-F-Tyr without the presence of H_2O_2, respectively. The experimental (solid orange line) and calculated (dashed orange line) show the spectra of native Tyr substrate reacting with H_2O_2. (B) Absorption spectra of the formation of Cpd 0. The experimental (solid red line) and computed (red dash line, conformation A) indicate the reaction of 3-F-Tyr with H2O2 and the formation of Cpd 0. The dashed green line indicates the conformation B of 3-F-Tyr reacting with H_2O_2.

3.4. Can Cpd 0 be an efficient active oxidant?

In the previous section, we have seen that both oxidants (putative), *i.e.,* Cpd 0 and Cpd I, can be easily formed by H_2O_2 in the substrate TyrH. Moreover, the study of Wang et al. *proposed the involvement of both oxidants, Cpd 0 and Cpd I.* They suggested that Cpd 0 acts as an oxidant in the C-F activation while Cpd I involves in the C-H bond activations.[11 12] We,

therefore, explored the possibility of Cpd 0 activating the C-F bond using MD simulations and QM/MM calculations of 3-F-Tyr in Cpd 0 state.

The MD simulations with 3-F-Tyr as substrate show several fluctuations of the hydroperoxo moiety (Figure S7, and see S.I for Figure S6 for RMSD values) and exhibit two different orientations of Fe-O1-O2-H. Similar to resting state simulations, the propionate side chain also shows 'out-plane' as the dominant conformation (see SI Figure S6 for more details).

3.5. QMMM calculations of Cpd 0 activation of C-H and C-F

As discussed earlier, the experimental study by Wang et al. proposed Cpd 0 as an active oxidant for the C-F activation; therefore, we performed QMMM calculation for C-F and C-H bond activations to test this possibility. We started our calculations with the conformation of Cpd 0, where the hydroperoxo points towards the C-H bond (cf. Figure 6). After geometry optimization of the reactant complexes, we performed the PES scanning for the C-F and C-H bond activation (see S.I Figure S8). As can be seen, the activation of both the C-H and C-F bond in 3-F-TyrH using Cpd 0 as oxidant is an energy-demanding step due to a significantly high barrier (37- 49 kcal/mol). It indicates Cpd 0 is not a feasible oxidant involved in C-F and C-H activation process. Therefore, *our QM/MM calculations of 3-F-Tyr show that Cpd 0 could not be an active oxidant for the C-H and C-F activation as proposed by Wang et al.*

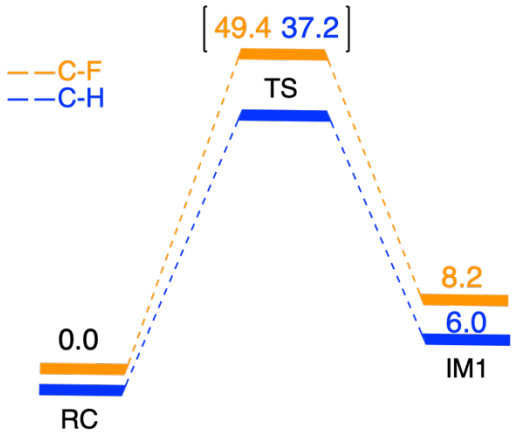


Figure 6 The QMMM potential energy profile for the nucleophilic attack of the hydroperoxo on the C-H and C-F bond of the 3-F-Tyr substrate at S=½.

3.6. Cpd I is the ultimate oxidant for C-H and C-F activation

As we see, Cpd 0 is not a preferred oxidant; therefore, we explored Cpd I as a possible candidate for the C-H and C-F bond activation for native TyrH and 3-F-Tyr substrates. The MD simulations of native substrate Tyr in Cpd I state of heme show flipping of aromatic side chain, which was *not seen earlier in the Co-crystal structure of the substrate with the enzyme*. Figure 7 shows that the CA-CB-CG angle was 150° during the initial stage of the simulations, while it changed to -150° after 100 ns of the simulation, and it maintained this orientation for the rest of the simulations. The dihedral angle can also support this flippling of the conformation and indicates that the substrate is flexible and rotates around the CG-CB bond in the active site of the TyrH enzyme. *In contrast, we didn't see such flipping in the simulations with the 3-F-Tyr substrate.* Thorough monitoring of the MD trajectory shows that the extra stability of the 3-F-Tyr substrate could be attributed to the hydrogen bond between the side chain of Y225 and the Fluoride atom of the substrate, which was missing in the natural substrate TyrH.

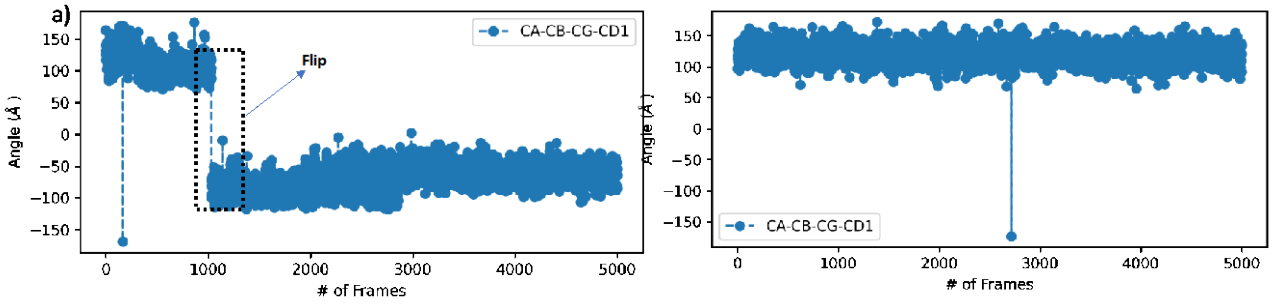


Figure 7 The dihedral angle of the substrates in TyrH for 500ns trajectory. a) the native and b) 3-F-Tyr substrate.

Regardless of the aromatic ring flip of the native substrate, the hydroxyl groups of the native Tyr substrate and 3-F-Tyr are stabilized by hydrogen bonds with the side chain of His88 and Y230. It is noteworthy that His88 is considered an essential catalytic residue in the reaction

mechanism.[11] The MD simulations (see Figure S1 and S9 for rmsd) of the native Tyr and 3-F-Tyr substrates sampled two conformations where the 'out plane' conformation is dominant. This observation agrees with the experimental structure (PDB id: 7KQR, a substrate-bound Tyr complex), where one of the propionate side chains of the heme group was in 'out plane' conformation.[11] Regardless of the conformation, there is always a water molecule in the vicinity of the iron oxo complex (see S.I in figure S10), making hydrogen bonds with Cpd I, which might have important implications for catalysis. In addition, the amine group of both substrates also forms hydrogen bonds with the iron oxo complex (Figure 8).

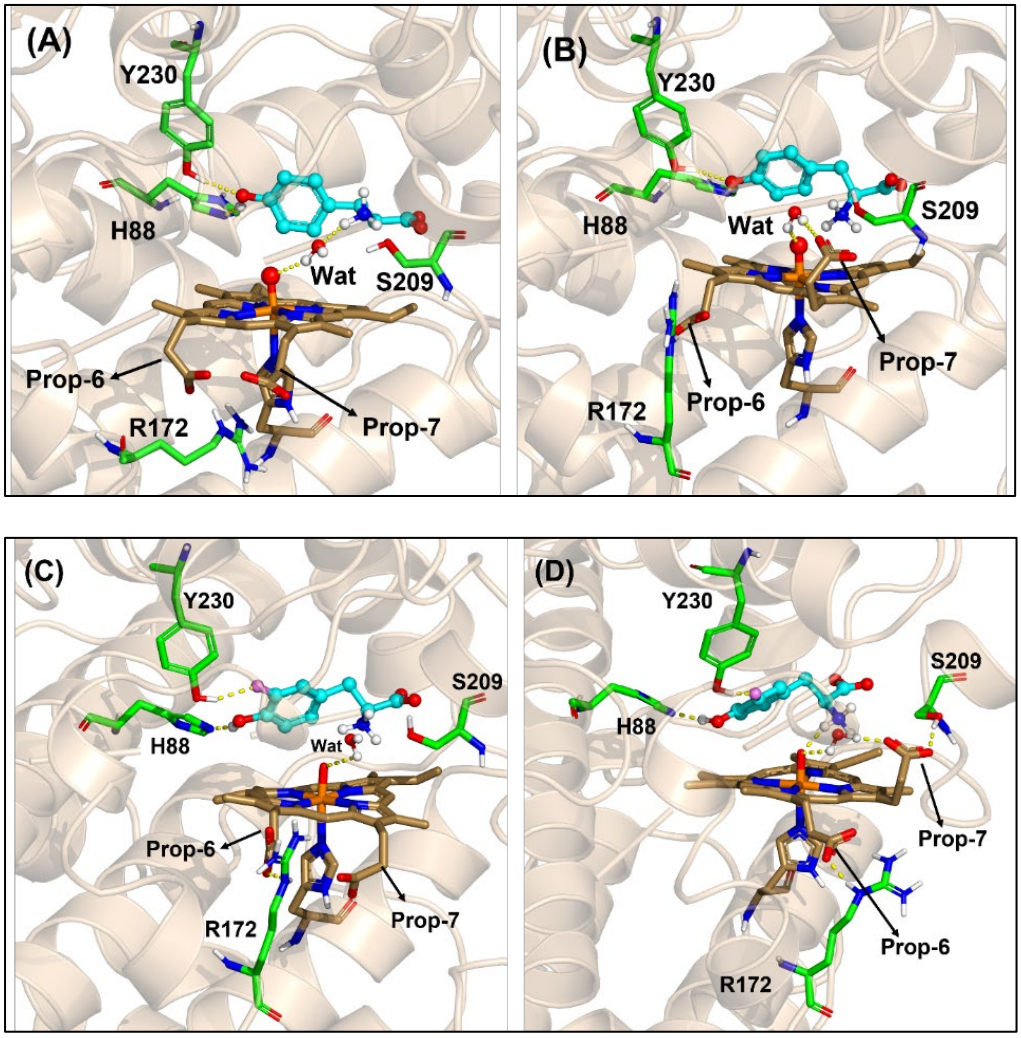


Figure 8 The alternative conformation of propionate side in the Cpd 1 simulation of TyrH in complex with native Tyr and 3-F-Tyr substrate.

3.7 Mechanism of the C-H bond activation using Cpd I

As we demonstrated that the Cpd 0 is not the primary oxidant in TyrH, we focused our attention on the Cpd I. The energetic calculations for Cpd I with a His196 axial ligand show an isoenergic S=1/2 and S=3/2 spin states, while spin S=5/2 is significantly higher in energy and thus was discarded for the subsequent calculations. The orbital diagram for the same is shown in Supplementary Information Figure S11. We explored two possible routes of C-H and C-F bond activation by Cpd I. In the first route, we explored the hydrogen atom abstraction (HAA) as a first route, and in the second route, we dug through the insertion of an oxygen atom into the aromatic ring of the 3-F-Tyr. The QMMM potential energy scan of the HAA shows a high activation energy barrier (see supporting information Figure S12); therefore, this route was discarded. Hence, we highlight the second route for the addition of the oxygen atom into the sp^2 hybridized carbon atom of the aromatic ring of 3-F-Tyr using QMMM calculations in both conformations A and B (See figure 9 for Conf A and B).

The QM/MM calculations show that H88 initiates the functionalization of the C-H and C-F bond through the proton abstraction from the hydroxyl group of the substrate in both conformation A and B. In conformation A, this process goes through a barrier-less process and results in the formation of exothermic IM1 at -5.2 and -4.9 kcal/mol for S=1/2 and S=3/2, respectively (Figure 10 and see S.I figure S13 for S=3/2). In conformation B, we again obtained a similar barrier-less process (Figure 11). Proton transfer from substrate to His88 has coupled with the electron transfer from the 3-F-Tyr to a single occupied a2u porphyrin orbital of the Cpd I. This PCET reaction results in the formation of IM1 could be best described as a one-electron reduced Cpd I and 3-F-Tyr as a quinone with a radical centered on the C-X bond. Next, the oxygen atom is inserted into the radical-centered C-F or C-H bond of the 3-F-Tyr. In conformation A, this process goes through TS-C-H or TS-C-F, which has an activation energy barrier of 18.0 and 16.9 kcal/mol, respectively, resulting in the formation of IM2-C-H (-13.6

kcal/mol) and IM2-C-F (-22.8 kcal/mol) exothermic intermediates. The oxidation state of iron in IM2 has changed from Fe(IV) to Fe(III). IM2 is best described as a tetrahedral complex with a newly formed C-O bond between the oxygen of the heme and the C-sp3 of the 3-F-Tyr. The computed barrier is in good agreement with the experimental kcat (4.7 ± 0.2 min-1 and 2.6 ± 0.1 min-1 C-H and C-F bond cleavage, respectively), which corresponds to 18.96 and 19.31 kcal/mol, respectively using the Eyring equation. Similar complexes are also reported in the aromatic hydroxylation for the P450 enzyme, albeit with a cationic nature.[33-34]

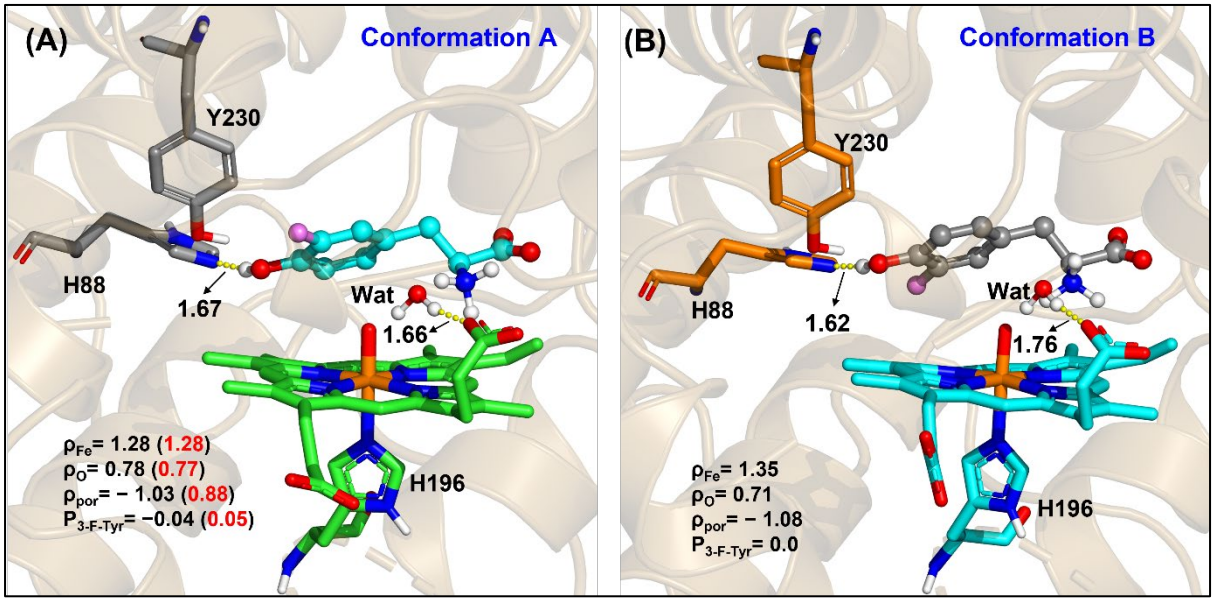
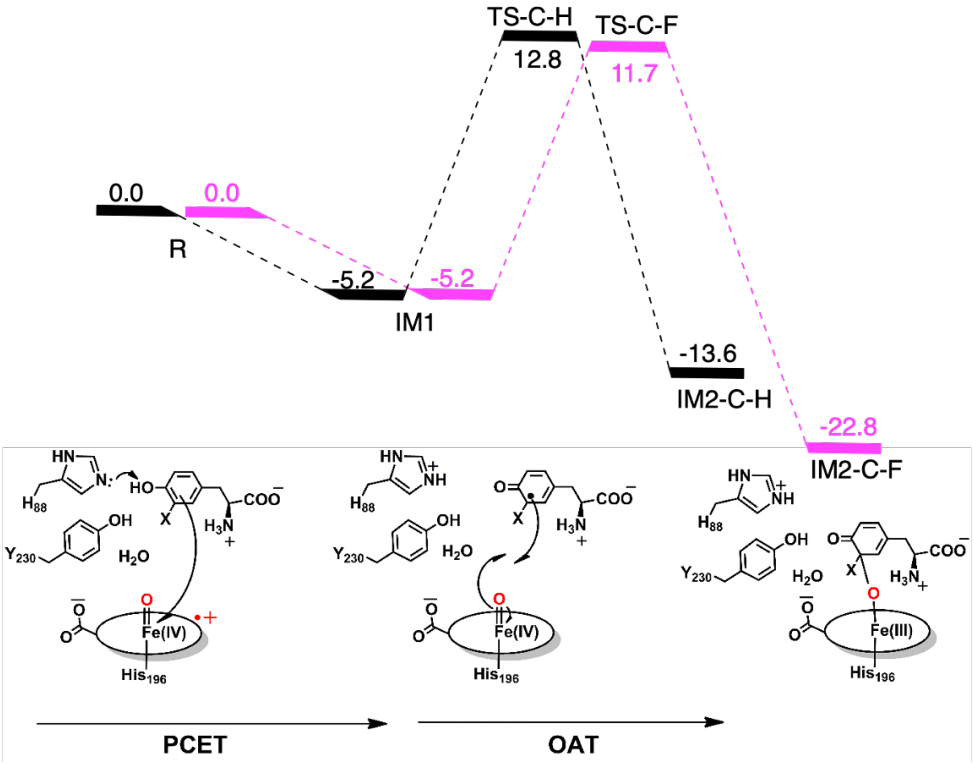


Figure 9. QMMM optimized structure of low spin S=1/2 complex in conformation A and B. ρ represents the calculated spin density for S=1/2and S=3/2 colored black and red, respectively.

In conformation B, the TS-C-H is slightly higher (21 kcal/mol vs. 18 kcal/mol) for C-H bond activation than conformation A. However, the transition state barrier of C-F bond activation is significantly higher in conformation B (27.6 kcal/mol vs. 16.9 kcal/mol in conformation A, c.f see SI). These results are in accordance with the computed absorbance spectra that the conformation A of 3-F-Tyr is the most dominant (see figure 4B). This data indicates that the C-F bond activation is not feasible in conformation B; however, both the conformations can

activate the C-H bond. Therefore, the product distribution is dependent on the conformation of the 3-F-Tyr. In conformation A, both the C-F and C-H bonds are activated, and in conformation B, only the C-H bond is oxidized, and the C-F bond goes through a higher activation energy barrier. Therefore, both the conformations can result in C-H bond oxidation; only conformation A gives the C-F bond functionalization.



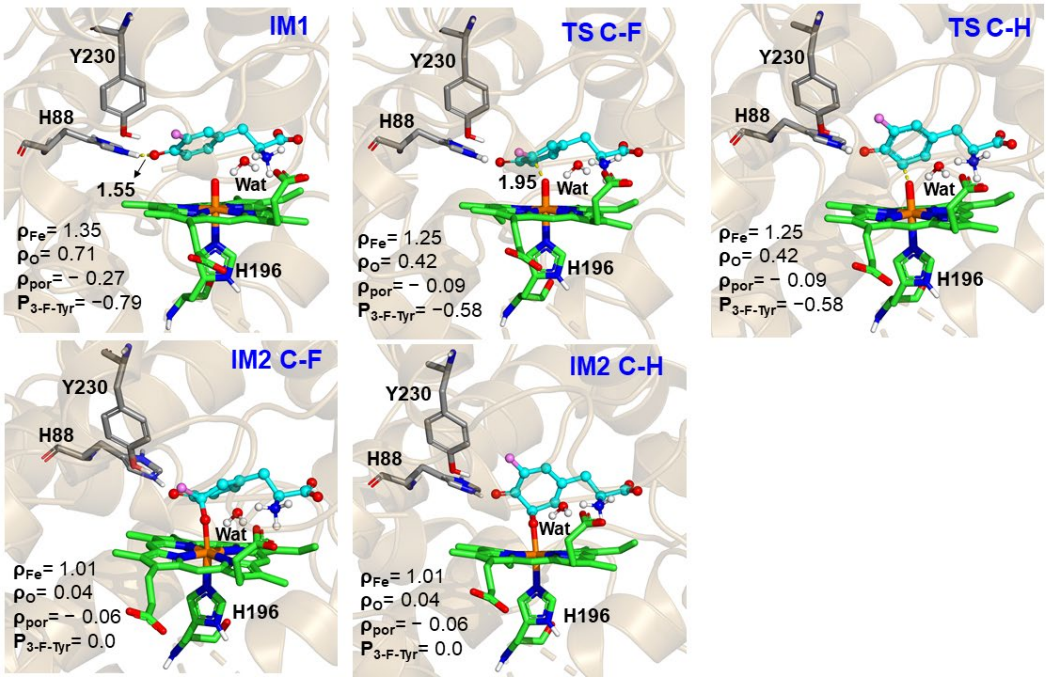


Figure 10. QMMM reaction profile and stationary and saddle point structures of C-H and C-F bond activation in 3-F-Tyr for conformation A.

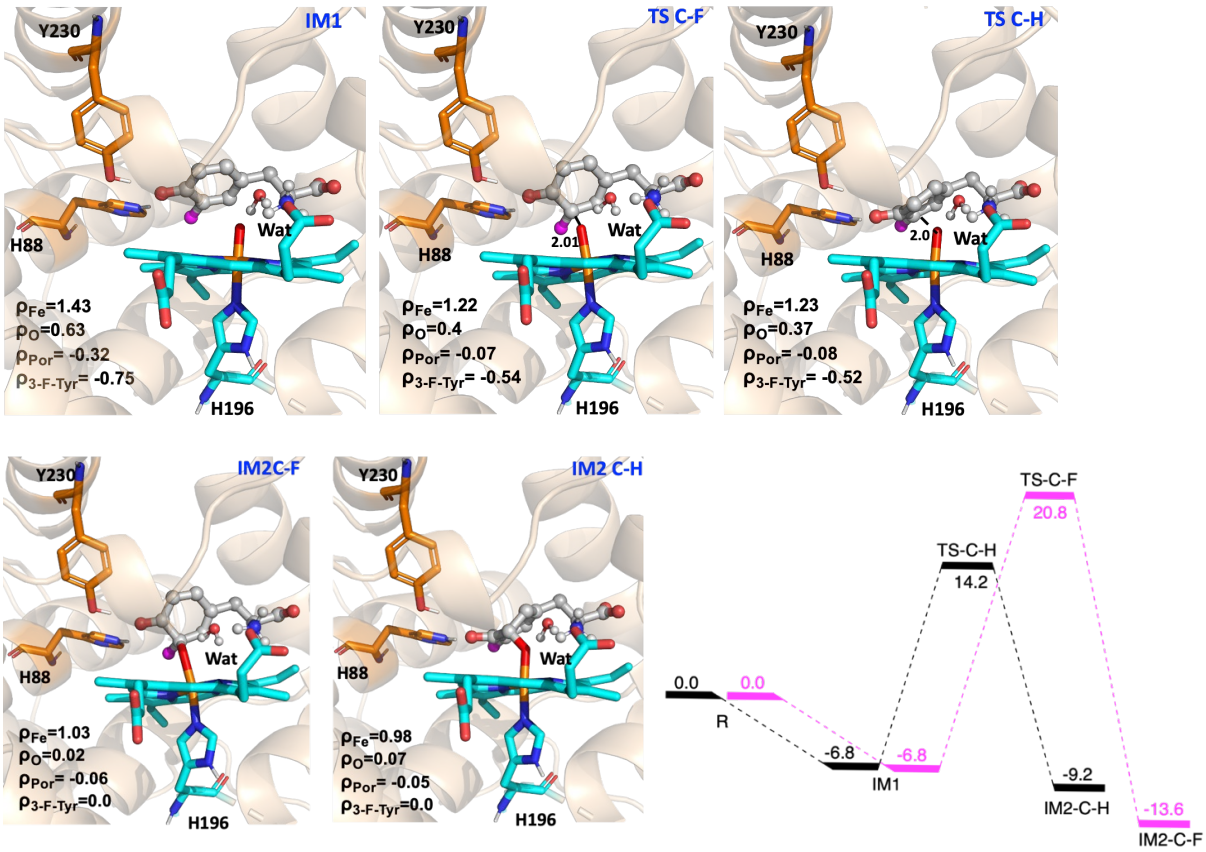


Figure 11. QMMM reaction profile and stationary and saddle point structures of C-H and C-F bond activation in TyrH for conformation B of the 3-F-Tyr.

3.8. Importance of His88 in catalysis

It has been shown experimentally that the presence of a hydroxyl group in the substrate is essential for enzyme activity. The substrate analogs where the hydroxyl group is replaced with either methoxy or hydrogen atom abolish the enzyme activity.[12] To probe the functional role of the OH group in catalysis, we performed two sets of QMMM calculations; one in the protonated and the other in the deprotonated form of fluorinated substrate. In the protonated substrate, we kept His88 in its neutral state, while in the deprotonated substrate, His88 was kept in the protonated form. The reaction profile for C-H bond activation, where the OH of the 3-F-Tyr substrate remained protonated, and His88 in its neutral state goes through an activation

energy barrier of 26 kcal/mol (See Figure S14 of SI). However, when the His88 abstracts a proton from the OH group of 3-F-Tyr and gets protonated, the reaction barrier goes through 18 kcal/mol (c.f Figure 10). Based on the spin population of the TS C-H neutral, there has been a one-electron reduction of the Cpd I, and at the same time, the 3-F-Tyr has an unpaired beta electron delocalized on the aromatic ring. As the reaction *proceeds, alpha electron from the Fe=O is transferred to the substrate coupled with the proton transfer to the His88.* The proton transfer to His88 results in the breakdown of the aromaticity of the substrate, and it allows the formation of tetrahedral intermediate IM2 C-H where the oxygen atom of iron oxo group make a covalent bond with the sp3 hybridised carbon atom the 3-F-Tyr. In essence, the presence of the His88 in the vicinity of the hydroxyl group of 3-F-Tyr substrate allows for the breakdown of aromaticity which allows for the oxygen atom to be inserted into the substrate by Cpd II and reduces the activation free energy barrier for the overall process.

4. CONCLUSIONS

TryH enzyme has a great potential in biocatalytic as it is able to activate both C-H and C-F bond in a regioselective manner in aromatic substrates. The X-ray structure of TyrH provided a valuable information on the binding of the native tyrosine and 3-F-Tyr substrate in the active site and important residues in involved in catalysis. We used MD simulations, computed absorption spectroscopy and QM/MM calculations to elucidate that the out of plane conformation of the propionate side chain plays an important role in the formation of Cpd 0 and Cpd I. The main oxidant involved in the functionalisation of both the C-F and C-H bond is Cpd I and His88 act as a modulator of aromaticity which is essential to initiate the C-H and C-F bond activation. His88 acts as a catalytic base and trigger the proton coupled electron transfer for the formation of one electron reduced CpdI like species which then add the oxygen atom into the C-H and C-F bond of the 3-F-Tyr substrate. The conformation A of 3-F-Tyr is

the dominant conformation and the product distribution in TyrH enzyme is dependent on the conformation of the 3-F-Tyr substrate. Both the conformations of the 3-F-Tyr can activate the C-H bond but only conformation A allows for the C-F activation.

5. ACKNOWLEDGMENT

KDD acknowledges the Department of Biotechnology, Govt. of India for the Ramalingamswami re-entry research grant (BT/RLF/Re-entry/10/2017). WS, SFS and GWB acknowledge the support by Research England's Expanding Excellence in England (E3) Fund.

6. Reference List

1. Berger, R.; Resnati, G.; Metrangolo, P.; Weber, E.; Hulliger, J., Organic fluorine compounds: a great opportunity for enhanced materials properties. *Chemical Society Reviews* **2011,** *40* (7), 3496-3508.
2. Gillis, E. P.; Eastman, K. J.; Hill, M. D.; Donnelly, D. J.; Meanwell, N. A., Applications of Fluorine in Medicinal Chemistry. *Journal of Medicinal Chemistry* **2015,** *58* (21), 8315-8359.
3. Cui, B.; Jia, S.; Tokunaga, E.; Shibata, N., Defluorosilylation of fluoroarenes and fluoroalkanes. *Nature Communications* **2018,** *9* (1).
4. Camargo, J. A., Fluoride toxicity to aquatic organisms: a review. *Chemosphere* **2003,** *50* (3), 251-64.
5. Jagtap, S.; Yenkie, M. K.; Labhsetwar, N.; Rayalu, S., Fluoride in drinking water and defluoridation of water. *Chem Rev* **2012,** *112* (4), 2454-66.
6. Wang, Y.; Liu, A., Carbon–fluorine bond cleavage mediated by metalloenzymes. *Chemical Society Reviews* **2020,** *49* (14), 4906-4925.
7. Münch, J.; Püllmann, P.; Zhang, W.; Weissenborn, M. J., Enzymatic Hydroxylations of sp3-Carbons. *ACS Catal.* **2021**, 9168-9203.
8. de Ruiter, G.; Carsch, K. M.; Takase, M. K.; Agapie, T., Selectivity of C−H versus C−F Bond Oxygenation by Homo- and Heterometallic Fe4, Fe3Mn, and Mn4 Clusters. *Chemistry – A European Journal* **2017,** *23* (45), 10744-10748.
9. Jana, A.; Samuel, P. P.; Tavčar, G.; Roesky, H. W.; Schulzke, C., Selective Aromatic C−F and C−H Bond Activation with Silylenes of Different Coordinate Silicon. *J. Am. Chem. Soc.* **2010,** *132* (29), 10164-10170.
10. Zámostná, L.; Sander, S.; Braun, T.; Laubenstein, R.; Braun, B.; Herrmann, R.; Kläring, P., Synthesis and structure of rhodium(i) silyl carbonyl complexes: photochemical C–F and C–H bond activation of fluorinated aromatic compounds. *Dalton Transactions* **2015,** *44* (20), 9450-9469.

11. Wang, Y.; Davis, I.; Shin, I.; Xu, H.; Liu, A., Molecular Rationale for Partitioning between C–H and C–F Bond Activation in Heme-Dependent Tyrosine Hydroxylase. *J. Am. Chem. Soc.* **2021**.
12. Wang, Y.; Davis, I.; Shin, I.; Wherritt, D. J.; Griffith, W. P.; Dornevil, K.; Colabroy, K. L.; Liu, A., Biocatalytic Carbon–Hydrogen and Carbon–Fluorine Bond Cleavage through Hydroxylation Promoted by a Histidyl-Ligated Heme Enzyme. *ACS Catal.* **2019,** *9* (6), 4764-4776.
13. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., UCSF Chimera—A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **2004,** *25* (13), 1605-1612.
14. 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* (Web Server), W368-W371.
15. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A., Development and testing of a general amber force field. *Journal of Computational Chemistry* **2004,** *25* (9), 1157-1174.
16. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16 Rev. C.01*, Wallingford, CT, 2016.
17. 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.
18. Li, P.; Merz, K. M., MCPB.py: A Python Based Metal Center Parameter Builder. *J. Chem. Inf. Model.* **2016,** *56* (4), 599-604.
19. Tian, C.; Kasavajhala, K.; Belfon, K. A. A.; Raguette, L.; Huang, H.; Migues, A. N.; Bickel, J.; Wang, Y.; Pincay, J.; Wu, Q.; Simmerling, C., ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *J. Chem. Theory Comput.* **2020,** *16* (1), 528-552.
20. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **1983,** *79* (2), 926-935.
21. D.A. Case, H. M. A., K. Belfon, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, G.A. Cisneros, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, G. Giambasu, M.K. Gilson, H. Gohlke, A.W. Goetz, R. Harris, S. Izadi, S.A. Izmailov, C. Jin, K. Kasavajhala, M.C. Kaymak, E. King, A. Kovalenko, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, M. Machado, V. Man, M. Manathunga, K.M. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, K.A. O'Hearn, A. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rahnamoun, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C.L. Simmerling, N.R. Skrynnikov, J. Smith, J. Swails, R.C. Walker, J. Wang, H. Wei, R.M. Wolf, X. Wu, Y. Xue, D.M. York, S. Zhao, and P.A. Kollman, Amber 2021. *University of California, San Francisco.* **2021**.

22. Darden, T.; York, D.; Pedersen, L., Particle mesh Ewald: An *N* ·log(*N*) method for Ewald sums in large systems. *The Journal of Chemical Physics* **1993,** *98* (12), 10089-10092.
23. 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. *Journal of Computational Physics* **1977,** *23* (3), 327-341.
24. 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.
25. Sherwood, P.; de Vries, A. H.; Guest, M. F.; Schreckenbach, G.; Catlow, C. R. A.; French, S. A.; Sokol, A. A.; Bromley, S. T.; Thiel, W.; Turner, A. J.; Billeter, S.; Terstegen, F.; Thiel, S.; Kendrick, J.; Rogers, S. C.; Casci, J.; Watson, M.; King, F.; Karlsen, E.; Sjøvoll, M.; Fahmi, A.; Schäfer, A.; Lennartz, C., QUASI: A general purpose implementation of the QM/MM approach and its application to problems in catalysis. *Journal of Molecular Structure: THEOCHEM* **2003,** *632* (1), 1-28.
26. Neese, F., The ORCA program system. *WIREs Computational Molecular Science* **2012,** *2* (1), 73-78.
27. Neese, F., Software update: the ORCA program system, version 4.0. *WIREs Computational Molecular Science* **2018,** *8* (1), e1327.
28. Smith, W.; Forester, T. R., DL_POLY_2.0: A general-purpose parallel molecular dynamics simulation package. *Journal of Molecular Graphics* **1996,** *14* (3), 136-141.
29. Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H., Results obtained with the correlation energy density functionals of becke and Lee, Yang and Parr. *Chemical Physics Letters* **1989,** *157* (3), 200-206.
30. Lee, C.; Yang, W.; Parr, R. G., Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988,** *37* (2), 785-789.
31. Grimme, S.; Ehrlich, S.; Goerigk, L., Effect of the damping function in dispersion corrected density functional theory. *Journal of Computational Chemistry* **2011,** *32* (7), 1456-1465.
32. Zhang, X.; Jiang, Y.; Chen, Q.; Dong, S.; Feng, Y.; Cong, Z.; Shaik, S.; Wang, B., H-Bonding Networks Dictate the Molecular Mechanism of H2O2 Activation by P450. *ACS Catal.* **2021**, 8774-8785.
33. Bathelt, C. M.; Ridder, L.; Mulholland, A. J.; Harvey, J. N., Aromatic Hydroxylation by Cytochrome P450: Model Calculations of Mechanism and Substituent Effects. *J. Am. Chem. Soc.* **2003,** *125* (49), 15004-15005.
34. a) Shaik, S.; Kumar, D.; De Visser, S. P.; Altun, A.; Thiel, W., Theoretical Perspective on the Structure and Mechanism of Cytochrome P450 Enzymes. *Chemical Reviews* **2005,** *105* (6), 2279-2328 b) Shaik, S and Dubey, K.D. The Catalytic Cycle of Cytochrome P450: A Fascinating Choreography. *Trend Chem.* **2022**, *3*, 1027-1044. c) Dubey, K.D. and Shaik, S. Cytochrome P450- the Wonderful Nanomachine Revealed through the Dynamic Simulations of the Catalytic Cycle. *Acc. Chem. Res*, **2019**, *52*, 289-399.