EDGE ARTICLE

Check for updates

On the engineering of reductase-based-monooxygenase activity in CYP450 peroxygenases†

Shalini Yadav, [ID] [a] Sason Shaik [ID] *[b] and Kshatresh Dutta Dubey [ID] *[a]

Recent bioengineering of CYP450$_{OleT}$ shows that peroxide-based CYP450$_{OleT}$ can be converted to a reductase-based self-sufficient enzyme, which is capable of showing efficient hydroxylation and decarboxylation activity for a wide range of substrates. The so-generated enzyme creates several mechanistic puzzles: (A) as CYP450 peroxygenases lack the conventional acid–alcohol pair, what is the source of two protons that are required to create the ultimate oxidant Cpd I? (B) Why is it only CYP450$_{OleT}$ that shows the reductase-based activity but no other CYP members? The present study provides a mechanistic solution to these puzzles using comprehensive MD simulations and hybrid QM/MM calculations. We show that the fusion of the reductase domain to the heme-binding domain triggers significant conformational rearrangement, which is gated by the propionate side chain, which constitutes a new water aqueduct *via* the carboxylate end of the substrate that ultimately participates in Cpd I formation. Importantly, such well-synchronized choreographies are controlled by remotely located Tyr359, which senses the fusion of reductase and communicates to the heme domain *via* non-covalent interactions. These findings provide crucial insights and a broader perspective which enables us to make a verifiable prediction: thus, the catalytic activity is not only limited to the first or second catalytic shell of an enzyme. Furthermore, it is predicted that reinstatement of tyrosine at a similar position in other members of CYP450 peroxygenases can convert these enzymes to reductase-based monooxygenases.

1. Introduction

The production of hydrocarbons in nature is restricted to a limited number of organisms.[1–7] CYP450$_{OleT}$ (CYP152L1 from *Jeotgalicoccus* sp. ATCC 8456), a member of CYP450 (Cytochrome 450) peroxygenase superfamily, is among such rare enzymes that have evolved to produce aliphatic hydrocarbons (α-alkenes) by the decarboxylation of the fatty acid substrate.[8,9] Aliphatic α-alkenes are important components that can be used to replace fossil fuels, as well as for the production of lubricants, detergents, and polymers.[10–12] This makes CYP450$_{OleT}$ a promising biocatalyst for biofuel production and commercial applications.

Unlike typical CYP450 monooxygenases that use molecular O$_2$ and redox partners, CYP450$_{OleT}$ uses H$_2$O$_2$ as the native oxygenating source as can be seen in Scheme 1.[13] Although

H$_2$O$_2$ is much cheaper than NADPH and other redox proteins, the large-scale production of α-alkenes demands high concentrations of H$_2$O$_2$ which constitute environmental hazards and cause deactivation of the oxidant of the biocatalyst. For this reason, Liu *et al.* used an alternative redox partner, RhFRED from *Rhodococcus.* sp., and observed that, even without H$_2$O$_2$ and presumably using O$_2$, peroxygenases still catalyze efficiently the decarboxylation of fatty acids, with carbon chain lengths of C$_{12}$ to C$_{20}$.[14] In a similar study, Lu *et al.* showed that the fusion of CYP450$_{OleT}$ with the reductase domain of bacterial CYP450$_{BM3}$ creates a self-sufficient enzyme, OleT-BM3R, capable of producing terminal alkenes from fatty acids of varying lengths using molecular O$_2$.[15] They further found that the fused OleT-BM3R enzyme has enhanced thermal stability compared to the isolated reductase domain BM3-R without CYP450$_{OleT}$. As such, these results demonstrate that CYP450$_{OleT}$ can function as a biocatalyst, similar to a conventional cytochrome P450, reducing O$_2$ instead of using hydrogen peroxide to generate the reactive intermediate in its catalytic cycle.

Nevertheless, this creates several questions. CYP450 monooxygenases require two electrons and two protons to form the ultimate oxidant Cpd I (*cf.* Schemes 1 and S1†).[16–19] The two electrons are normally supplied by the reductase partner,[20–22,55] while the two protons are provided by an organized water chain that is directed to the active site *via* acid–alcohol pairs (*cf.* states V and VI

[a]Department of Chemistry, School of Natural Science, Shiv Nadar Institution of Eminence, NH91 Tehsil Dadri, Greater Noida, Uttar Pradesh 201314, India. E-mail: kshatresh.dubey@snu.edu.in

[b]Institute of Chemistry, The Hebrew University, Edmond J. Safra Campus at Givat Ram, Jerusalem, 9190401, Israel. E-mail: sason@yfaat.ch.huji.ac.il

† Electronic supplementary information (ESI) available: Schemes of the reactions, RMSD and RMSF values, clustering of trajectories, variation of the angle and distance during simulations, energetics obtained using MMGBSA calculations, water flow, water occupancies and water channel, coordinates of the optimized geometry, *etc.* See DOI: https://doi.org/10.1039/d3sc06538c
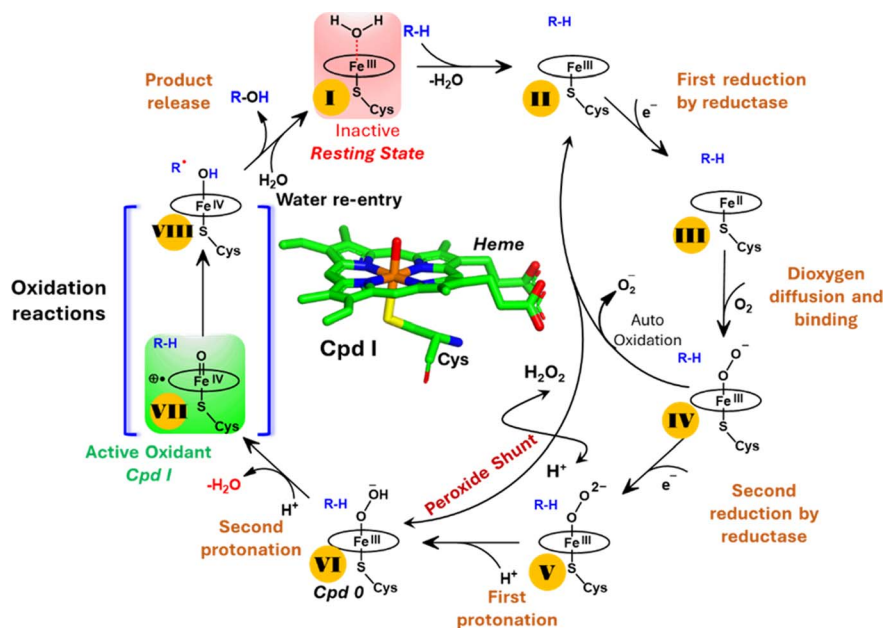
Scheme 1 A generic catalytic cycle for CYP450 enzymes using O_2, $2e^-$, and H^+ (steps: I → II → III → IV → V → VI → VII) and using directly H_2O_2 (peroxide shunt) as a co-substrate to generate the active oxidant Cpd I by the CYP450 peroxygenase family.

in Scheme 1).[23–27] For example, in CYP450$_{BM3}$, the acid–alcohol pair is Glu267–Thr268.[16,25] However, the experimental 3D structures of CYP450 peroxygenases do not reveal any acid–alcohol pairs in the respective heme domains.[13] This creates therefore the following mechanistic puzzle: (a) if CYP450 peroxygenases lack the conventional acid–alcohol pair, then what could be the possible protonation routes for the CYP450 peroxygenase in the fused CYP450$_{OleT-BM3R}$ complex to form the ultimate oxidant Cpd I? (*cf.* Scheme S2†). (b) Why is CYP450$_{OleT}$ the only peroxygenase family member, which exhibits such reductase-based monooxygenase activity? (c) Can we reproduce similar reductase-based monooxygenase activity in other members as well?

The present study addresses these issues using comprehensive MD simulation and hybrid QM/MM calculations. We show here that the fusion of the reductase domain to the heme binding domain triggers a well-synchronized large conformational rearrangement *via* the propionate side chain which forms a new water route through the carboxylate end of the substrate. Thus, the substrate instigates the formation of Cpd I that eventually causes its own oxidation. Such well-synchronized choreographies are mastered by far-located Tyr359, which can sense the fusion of reductase and communicates this to the heme domain *via* non-covalent interactions. Interestingly, unlike other CYP450 peroxygenase members, CYP450$_{OleT}$ possesses the key tyrosine 359 residue. As such, we propose that a similar reinstatement of the key residue may endow other members, as well, with a reductase-based monooxygenase activity.

2. Computational method and details

2.1 System preparation

The initial structure of the substrate bound heme-binding domain of CYP450$_{OleT}$ (PDB entry: 4L40)[9] and the FMN reductase domain

of CYP450$_{BM3}$ (PDB entry: 1BVY)[28] were taken from the protein data bank. The crystal structures of CYP450$_{OleT}$ and CYP450$_{BM3}$ were superimposed using Chimera software,[29] and the BM3 heme domain was deleted. The remaining combined CYP450$_{OleT}$-FMN structure was subsequently used as a single PDB (*cf.* Fig. 1).

The two domains were further combined *via* a linker (provided in 1BVY) using MODELLER.[30] Missing hydrogen atoms were added by the leap module of the Amber18 package. As our primary goal is to identify protonation pathways, we have considered only protonated residues that are in close proximity to the heme.

Force field parameters for a dioxygen heme complex were taken from already published data by Cheatham and co-workers.[31] For the ferric-hydroperoxide species (Cpd 0), we generated parameters using the Metal Center Parameter Building (MCPB)[32] module in Amber 18. The partial atomic charges and missing parameters for the substrate (DCR) and FMN were obtained from the RESP[33,34] charge method at the HF/6-31G* level. Na^+ ions were added to neutralize the system. Finally, the resulting system was solvated within an octahedral box of TIP3P[35] water extending up to a minimum cutoff of 12 Å from the protein boundary. The Amber ff14SB[36] force field was used for protein in all the simulations performed.

2.2 MD simulation

After complete parametrization, the geometry of the entire system was minimized (5000 steps of steepest descent and 5000 steps of conjugate gradient) to find the local energy minimum of the structure and to remove poor contacts. The post minimization system was gently annealed under the NVT ensemble for 50 ps with a weak restraint of 4 kcal mol^{-1} Å^{-2}.

Subsequently, the system underwent 1 ns of density equilibration under the NPT ensemble at constant temperature (300
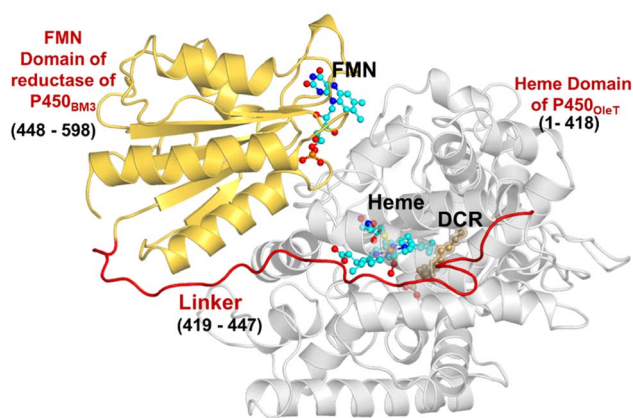
Fig. 1 The fused protein, OleT-BM3R, includes the heme domain (in grey) and the FMN binding domain (in gold), which are linked *via* the red linker. The DCR (the three letter code for the substrate icosanoic acid) substrate, heme, and FMN are drawn as ball and stick models in brown and light blue, respectively.

K) and pressure (1 atm), using the Langevin thermostat[37] with a collision frequency of 2 ps, and the Berendsen barostat[38] with a pressure relaxation time of 1 ps. This 1 ns density equilibration is a weakly restrained MD simulation, in which the system is released slowly to achieve a uniform density after heating under periodic boundary conditions.

After 1 ns of density equilibration, we removed all the afore-applied restraints (during heating and density equilibration) and the system underwent further equilibration for 3 ns to get a well-settled pressure and temperature for conformational and chemical analyses. This was followed by a MD production run for each system.

During all the MD simulations, covalent bonds containing hydrogens were constrained using the SHAKE[39] algorithm, and the particle-mesh Ewald (PME)[40] method was used to treat long-range electrostatic interactions. All the MD simulations were performed with the GPU version of the Amber 18 package.[41]

To understand this yet unclear source of protonation and conversion of dioxygen-state → Cpd 0 → Cpd I, MD simulation of 500 ns was performed on the following four systems, by using random velocity: (A) the peroxo-ferric state of the heme domain associated with the reductase domain; (B) the peroxo-ferric state of the heme domain only; (C) the Cpd 0 (ferric hydroperoxide) state of the heme domain associated with the reductase domain; (D) the Cpd 0 (ferric hydroperoxide) state of the heme domain only. Since our main objective is to find out a protonation source, we have kept all the titratable amino acid residues close (up to 5 Å) to the heme moiety in their protonated form, during all the simulation processes.

2.3 MMGBSA calculations

MMGBSA is a widely used method for studying protein–ligand interactions and protein–protein interactions. Here, the calculation was performed using the MMPBSA.py[51] module of the AMBER simulation package.[41] To check the energetics of the dynamic event that occurs after 200 ns, we have performed simulation in two sets. One set includes the geometry of protein before 200 ns and the second set involves the protein geometry after 200 ns. Water and ions were removed, and the dielectric constant was set to 1 for the solute and 80 for the solvent. The method combines molecular mechanics energy calculations, the solvent-accessible surface area, and the GB model (MMGBSA) to calculate solvation free energy.

2.4 QM/MM methodology

The mechanism for generating Cpd I from the peroxo-ferric species, $Fe–OO^{2-}$ → Cpd 0 → Cpd I, was investigated by QM/MM calculations on representative snapshots from the MD trajectory. The calculations were carried out using Chem-Shell[42,43] employing Turbomole[44] for the QM part and DL_POLY[45,46] using the FF14SB amber force field for the MM part.

The employed QM region contained the heme, the proton donor residues, and the water molecules, which shuttle protons towards the peroxo species. To account for the effect of the environment, during the QM/MM calculations, we included all the protein residues and water molecules present up to 8 Å of the heme surface, in the active region. The atoms of this active region interact with the QM atoms and lead to subsequent polarization effects through electrostatic and van der Waals interactions. An electronic embedding scheme was used to account for the polarizing effect of the protein residues in the QM region. Hydrogen link atoms with the charge-shift model[36] were applied to treat the QM-MM boundary.

The QM region was computed by use of the UB3LYP[47] functional with two basis sets Def2-SVP (denoted as B1) and Def2-TZVP (denoted as B2). B1 was used for geometry optimization, potential energy surface scanning, and frequency calculations. The transition states (TSs) were located by relaxed potential energy surface scans followed by full TS optimization using the partitioned rational function optimization (P-RFO)[48] method implemented in the HDLC code. The zero-point energy was calculated at the B1 level, for all the species. Grimme dispersion (G-D3)[49] was used to incorporate the dispersion correction in energetics. Energies were subsequently corrected by use of B2, and final energies are reported as UB3LYP/B2-D3+ZPE.

3. Results and discussion

3.1 Conformational dynamics in the heme binding domain (HBD) due to fusion of reductase

The study was initiated with simulation of the $CYP450_{OleT-BM3R}$ complex in the heme-peroxo-ferric state. The linker loop (Fig. 1) was found to be the most flexible region during the simulations (Root Mean Square Fluctuation ~4–6 Å in Fig. S1 and S2†). Additionally, during the simulation, the distance between the FMN and heme domains shrunk.

What factors drive the two domains to come closer to each other? To answer this question, we calculated the shortest path map (SPM),[50] a correlation-based approach to study the allosteric communications between two distant substructures. The shortest communication path is shown in Fig. 2 as beads (as
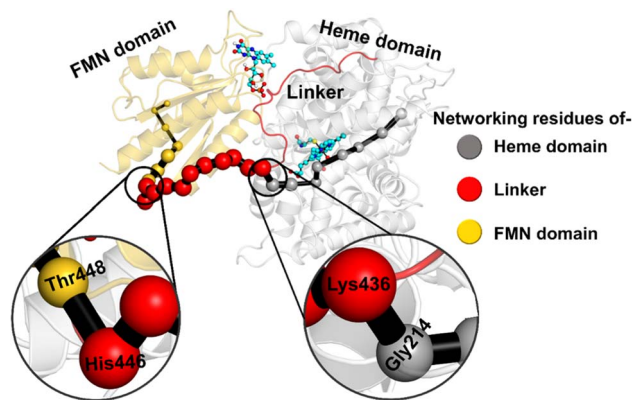
Fig. 2 Representation of the shortest path map (SPM) between the heme domain (grey) of OleT and the FMN domain (gold) of reductase. Key portions are magnified. Beads in grey show the portion belonging to the heme domain, gold the portion of the FMN domain and red the portion of the linker loop. Note that the size of the beads indicates the strength of the interaction, while the black edges represent the communication path among the residues.

small spheres) and threads (as black), where the size of the beads shows the magnitude of the interaction while the threads (in black) show the residues participating in the communications. As can be seen, the linker loop (shown in red), is key to the inter-domain talk that connects Gly214 of the heme domain to His446 of the FMN/reductase domain. We believe that this interaction network is crucial for bringing the two domains closer to each other during the simulations. In addition, the angle between the centers of mass of the HBD and the reductase also changes with the simulations (Fig. 3A), indicating topological rearrangement in the HBD due to reductase attachment.

Our simulations show that this conformational change is primarily orchestrated by the C-helix of the HBD and the α_1 helix of BM3-R.

As shown in Fig. 3A–C, the helix angle between C and $\alpha1$ helices changes from $\sim79°$ to $101°$ and remains constant, throughout the simulation. We found that the interdomain interactions of Val438$_{HBD}$ with Ser474$_{BM3-R}$ and Tyr359$_{HBD}$ with Gly459$_{BM3-R}$ (subscript HBD is for the heme domain and BM3-R is for the reductase domain) are key for this change in the helix angle. As can be seen, initially (Fig. 3B), Gly459 of the $\alpha1$-helix is far away (~7.1 Å) from the Tyr359 residue of the heme domain. However, at the later stage of simulation (Fig. 3C), the same Gly459 moves closer to Tyr359 and forms a strong hydrogen bond with Tyr359. We believe that this alteration in the interaction between Gly450 and Tyr359 is the root cause of the tilting of C and $\alpha1$ helices. The energetic calculations of the two different conformations *i.e.*, Fig. 3B and C, also show that the final conformations of C and $\alpha1$ (*i.e.*, Fig. 3C) are stabilized by ~33 kcal mol^{-1} relative to the initial conformation (see Table S2† for the MMPBSA analysis).[51] The thermodynamic stability of conformation 3B by ~33 kcal mol^{-1} provides the driving force needed for the change of orientation of the C and $\alpha1$-helices. In the subsequent section, we will see that this conformational change due to the Tyr359–Gly459 interaction works as a signal that triggers the necessary water inflow into the heme site and propagates the catalytic cycle.

3.2 The mechanism of water inflow into the heme active site

Water inflow is crucial in conventional CYP450 monooxygenases for proton shuttling during the catalytic cycle,[52] and it has been documented in several articles; however, the
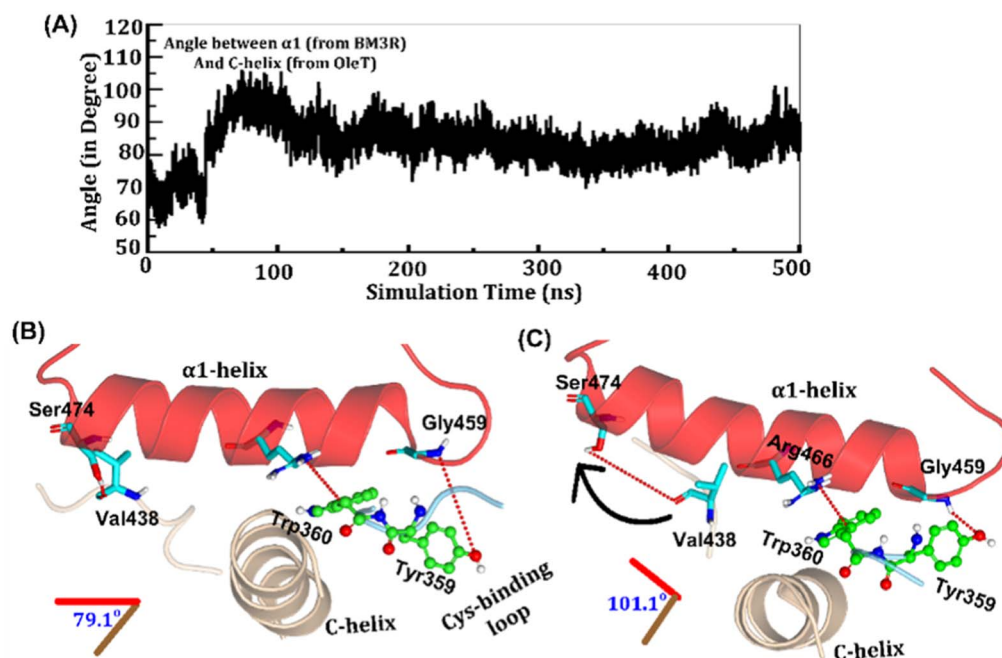

Fig. 3 (A) Variation of the angle between $\alpha1$-helix and C-helix during the MD simulations. (B and C) The key interdomain interactions where B shows initial conformation, while C is the final positioning of the C-helix, cysteine binding loop and $\alpha1$-helix.

formation of a proton channel in engineered CYP450$_{OleT}$ fused with reductase is nevertheless intriguing to study. Therefore, we performed two separate sets of simulations: (a) without fused reductase and (b) with fused reductase, and we carefully monitored the enzyme conformation *via* MD simulations to study the pattern of water inflow. In (a) without reductase, we did not see, during the entire 500 ns of simulation, much water inflow *via* the propionate side chain as is necessary for the first protonation (heme-peroxo → Cpd 0). However, when the reductase was included in (b), there appeared a consistent water inflow after ∼200 ns of the simulation (see Fig. S3†).

To unravel the mechanism of water inflow, we carefully looked at the active site topology of the heme-binding domain in both cases, and we found that the conformation of the propionate side chain (prop7), Tyr59 and Arg66 is crucial for the water flow. In the absence of reductase, both the guanidium hydrogen atoms of Arg66 and the hydroxy group of Tyr59 interact strongly with prop7 of the heme prosthetic group (see Fig. 4A). These interactions form a closed doorway that blocks the water egress through heme-peroxo. Interestingly, in case (b), where reductase is fused, Tyr59 and Arg66 move apart and open a water channel that stretches through the heme-peroxo *via* propionate (Fig. 4B).

So how does the enzyme sense the opening of the water gateway when the reductase binds to the heme-binding domain? We found that residue Tyr359 is the key and the transducer to open and close the water gate by releasing a signal through non-covalent interactions. When there is no reductase, Tyr359 is weakly associated with the interaction network of Tyr359–Arg364–Gly354–Gly355–Tyr59–Prop7. However, when the reductase gets attached, it undergoes a conformational rearrangement, which is thermodynamically driven by the Gly459–Tyr359 interaction (see the previous section and a discussion on the movement of the C-helix due to reductase attachment), thus pulling the interaction network towards Tyr359, and thereby opening the gate for water inflow.

In summary, the interaction between the two domains in the fused protein OleT-BM3R disrupts the H-bond interaction between the heme propionate and the Tyr59 and Arg66 residues in the ferric-peroxo species of heme. This disruption leads to the opening of the gate *via* Prop7, which in turn enables bulk water to enter the active site *via* the distal site of the heme domain *i.e.*, from the FMN domain (see also Fig. S4† for the loop dynamics).

3.3 Why is P450$_{OleT}$ unique in showing the reductase-based monooxygenase activity?

It is important to note, CYP450$_{OleT}$ is unique in the peroxygenase family that shows the reductase-based monooxygenase
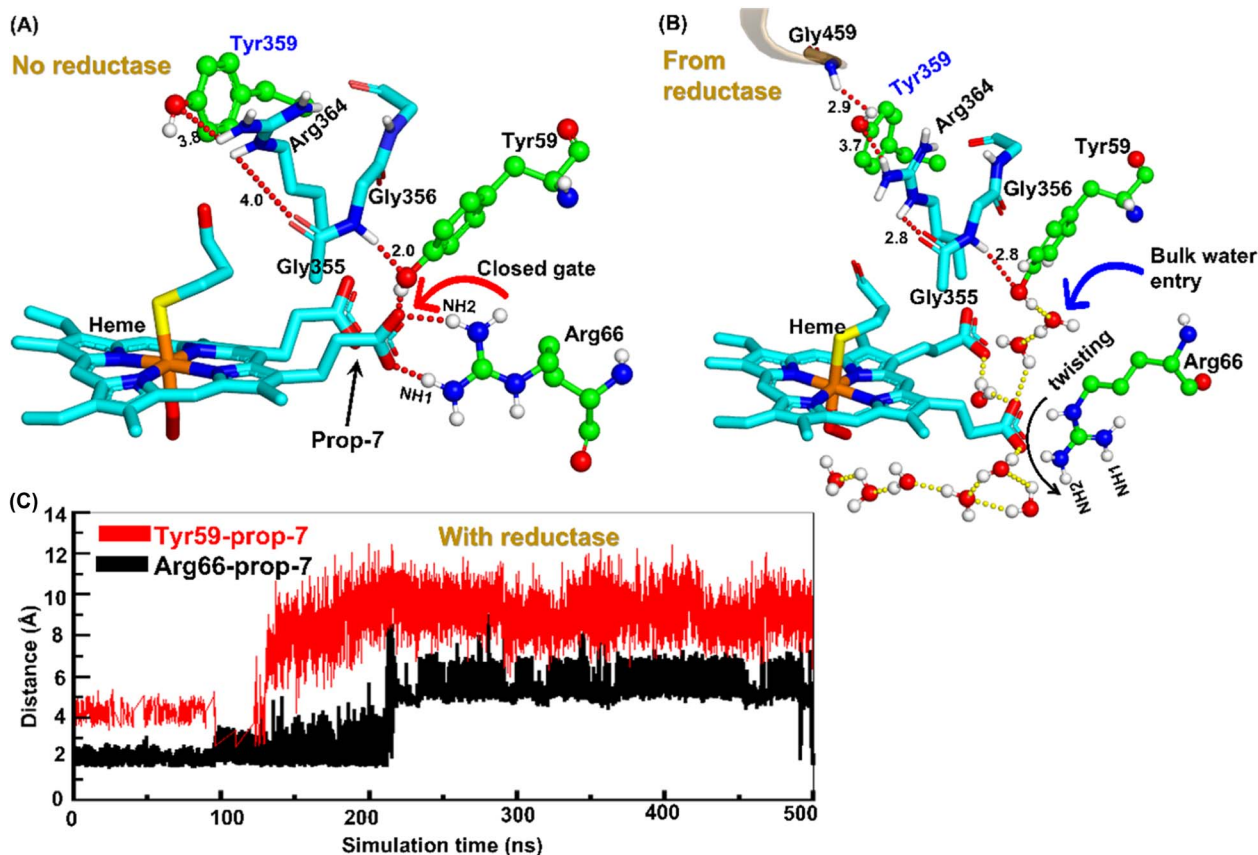

Fig. 4 Formation of an ordered H-bond network between residues Tyr359 of the heme domain to heme prop7 (A) before reductase attaches and (B) after reductase attaches. Note here how a new interaction between Gly459 of the reductase with Tyr359 perturbs the entire H-bond network after the fusion of reductase. (C) A plot showing the increasing Prop7–Tyr59 and Prop7–Arg66 distances, which signifies a gate opening.
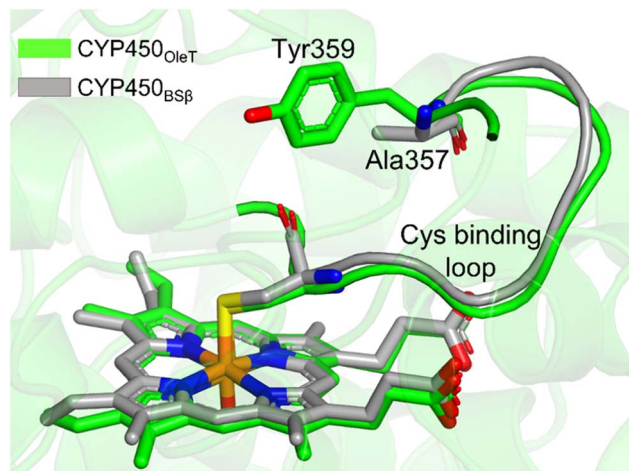
Fig. 5 The superimposed structure of CYP450$_{OleT}$ with CYP450$_{BSβ}$. Note the lack of Tyr359 in CYP450$_{BSβ}$.

activity. The similar experiments for the fusion of the reductase with other members of the peroxygenase family, such as in CYP450$_{BSβ}$ (CYP152A1 from *Bacillus subtilis*), did not appear to function for their native reactions (β-hydroxylation of fatty

acid).[53] To investigate the root cause we inspected the availability of the key residues in all three members. Interestingly, as shown in Fig. 5, CYP450$_{BSβ}$ lacks Tyr359 or an analogous residue, which can sustain water inflow. As mentioned above, Tyr359 is the key residue that opens the gate to facilitate the water inflow in CYP450$_{OleT}$. Thus, the absence of Tyr359 in CYP450$_{BSβ}$ prevents the water inflow and consequently, CYP450$_{BSβ}$ does not show reductase-based monooxygenase activity. This provides solid experimental support of the role of Tyr359 in the reductase activity.

3.4 The protonation route in CYP450$_{OleT}$ fused with an FMN domain

The water inflow is crucial for the first protonation of the heme-peroxo complex. In previous studies it is found that the first protonation could be facilitated *via* propionate or other sources than an acid–alcohol pair.[25] Therefore, we carefully monitored the different water pathway that is connected with a possible base.

As can be seen from Fig. 6, there are three possible water networks present in the active site, which can be potential proton donors to the distal oxygen (O$_d$) in order to convert dioxygen Fe(III) to Cpd 0, *i.e.*, state 5 → 6: (i) protonation
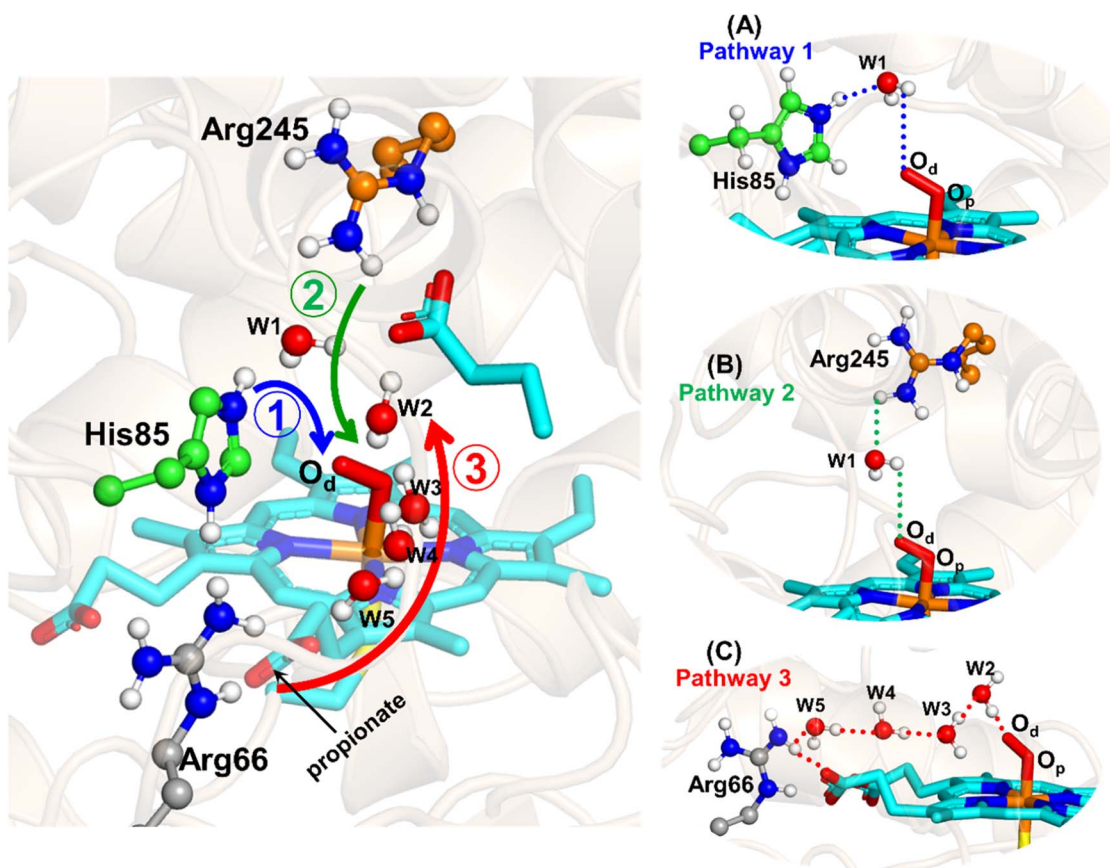

Fig. 6 Hydrogen bonding network between water and potential proton–donor amino acid residues, which connect the distal oxygen (O$_d$) of heme. Note that the representative snapshot has been taken from the most populated frame obtained from cluster analysis of a 500 ns MD trajectory. (A–C) The three possible pathways (shown by the three different colored residues) for plausible proton transfer to the distal oxygen (O$_d$) to form Cpd 0 through Arg66/His85/Arg245.

through His85 → (H_2O) → O_d; (ii) protonation through Arg245 → (H_2O) → O_d; (iii) protonation through Arg66 → (H_2O)$_n$ → O_d. To check the respective potencies of protonation through these residues, we have performed QM/MM calculation.

3.4.1 Mechanism of the first protonation *via* QM/MM calculations. Unlike classical MD simulations, which cannot handle bond breaking/making, hybrid QM/MM calculations can explore the mechanism of Cpd 0 formation in fused OleT-BM3R. Here, initially, we investigated the mechanism of the first protonation *i.e.*, the formation of the oxidant Cpd 0. In doing so we have explored all three possible pathways to generate Fe–OO^{2-} → Fe–OOH^-. As can be seen in Fig. 6, the distal oxygen (O_d) of the dioxygen moiety is connected to His85 and Arg245 *via* a common water molecule w1 (*cf.* Fig. 6A and B). Meanwhile in the other plausible pathway, Arg66 connects the distal oxygen (O_d), *via* several water molecules (*cf.* Fig. 6C). For a clear observation we have performed QM/MM separately for all three possibilities.

To get the respective reaction coordinates and energy profiles for Pathways 1 and 2, we performed the potential energy surface (PES) scanning of the QM/MM optimized geometries. The QM regions used for each pathway are shown in Fig. 7A–C. Interestingly, in the paths A and B, the dioxygen moiety of ferric-peroxide species (strongly basic in nature) accepts the proton from the nearby source His85 and Arg245 in a barrierless manner and generates Cpd 0 (see Fig. 8A and B). Meanwhile the third pathway, which is supposed to be facilitated *via* distant Arg66 through organized water chains, is found to be highly unfavourable for Cpd 0 formation (*e.g.*, the energy barrier is more than 32.5 kcal mol^{-1}, see Fig. S5† for details). This indicates that Arg66 may not directly participate in the proton transfer mechanism, but instead, it serves as a gatekeeper to facilitate the entry of water molecules towards the active site, as we discussed in an earlier section (*cf.* Fig. 6). In essence, while Arg66 is not directly involved in the conversion of dioxygen to Cpd 0, its role is pivotal in facilitating water entry to the heme active site to promote the process.

To comprehend the electronic structure of the resulting product (Cpd 0), the spin populations of crucial atoms and groups were computed. In both pathways, the total spin densities are comparable, with the Fe center having one unpaired electron with an occupancy of approximately 1.0. This

is a typical characteristic of Cpd 0 states observed in CYP450 enzymes.[16,18]

3.5 MD result of OleT-BM3R in the ferric hydroperoxide Fe–OOH– state

The second protonation occurs at Cpd 0, which is commonly thought to be a necessary step for the eventual formation of the active oxidant Cpd I. Furthermore, though the Cpd I formation from Cpd 0 in cysteine-ligated conventional CYP450s is well established,[16,24–26] the mechanism of Cpd I formation through Cpd 0 in peroxygenase CP450 is yet to be elucidated. In doing so, we performed 500 ns of MD simulations (along with 2 replicas; 2 × 500 ns) of OleT-BM3R in the Cpd 0 state of heme. The root mean square deviation of the system converged after a simulation time of about 70 ns (*cf.* Fig. S6†). In all the simulations, the hydroperoxide ($OOH-$) moiety of heme was flexible and rotated freely in two possible orientations A and B as shown in Fig. 9A and B. This flexibility and rotation ability of the OOH– moiety can be anticipated from the drastic changes (+180° to −180°) in the dihedral angle of Fe–O_p–O_d–H in Cpd 0 of OleT-BM3R (*cf.* Fig. 9C).

Similar to the ferric-peroxide state, a careful inspection of the trajectory of the active site (near to Fe–OOH^-) shows a consistent water flow, *via* propionate. This flow is maintained throughout the 500 ns simulation time in the Cpd 0 state (*cf.* Fig. S7†). Surprisingly, in addition to ferric-peroxide, Cpd 0 also exhibits a disruption of the salt bridge between Arg66 and the propionate group, as shown in Fig. S8,† which reinforces the idea of Arg66 being a regulator of the water gate. Upon examining the hydrogen bonding network near the Fe–OOH– site, we noted the presence of the amino acid residues His85 and Arg245, as well as the substrate interacting with the hydroperoxide moiety, as depicted in Fig. 9D.

As can be seen in Fig. 9D, the presence of two water molecules close to the OOH– moiety connects His85, Arg245, and the substrate (DCR). Thus, water w_1 is connected with the imidazolium side chain of His85 and both oxygens (proximal O_p and distal O_d) of the hydroperoxide of the heme moiety. On the other side, Arg245 and substrate's carboxylic acid end are connected to the heme *via* water w_1 and w_2. This interaction persists for most of the simulation time. Since protonation *via* water molecules is more common in P450 chemistry,[16] we have
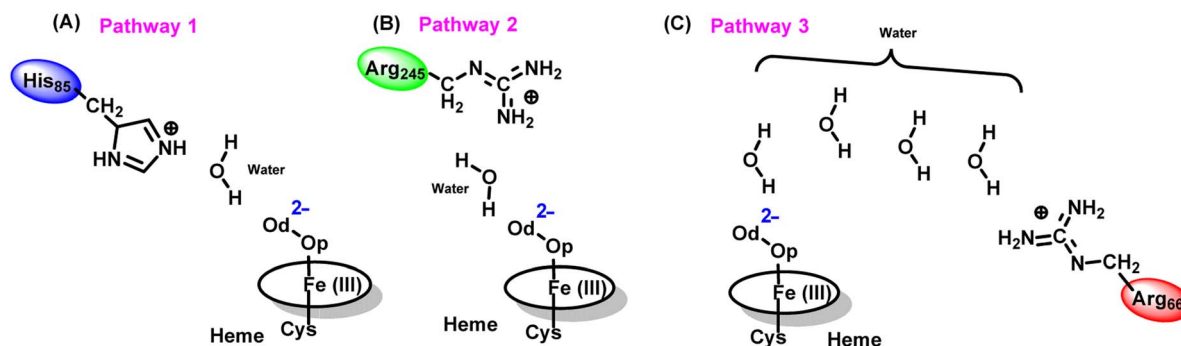

Fig. 7 QM region selected for all three pathways (A to C) to proceed the QM/MM calculation.
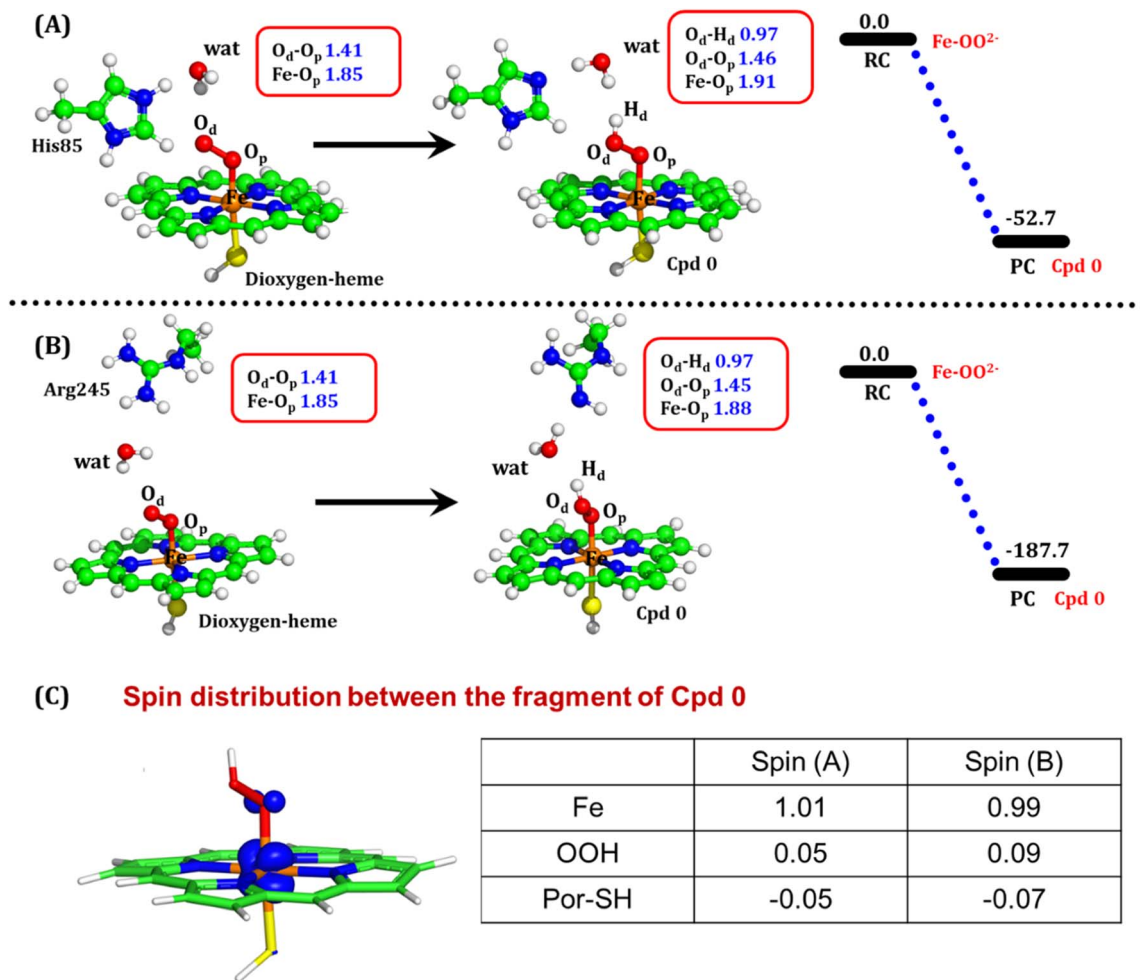
Fig. 8 QM/MM optimized geometries for the Cpd 0 formation when (A) His85 acts as a proton donor and (B) Arg245 acts as a proton donor. (C) Electronic spin distribution of the Cpd 0 state. Blue isosurface depicts the presence of one alpha electron on Fe. Note that the energies are reported at the B2-D3 + ZPE level of theory. All energies are in kcal mol^{-1}.
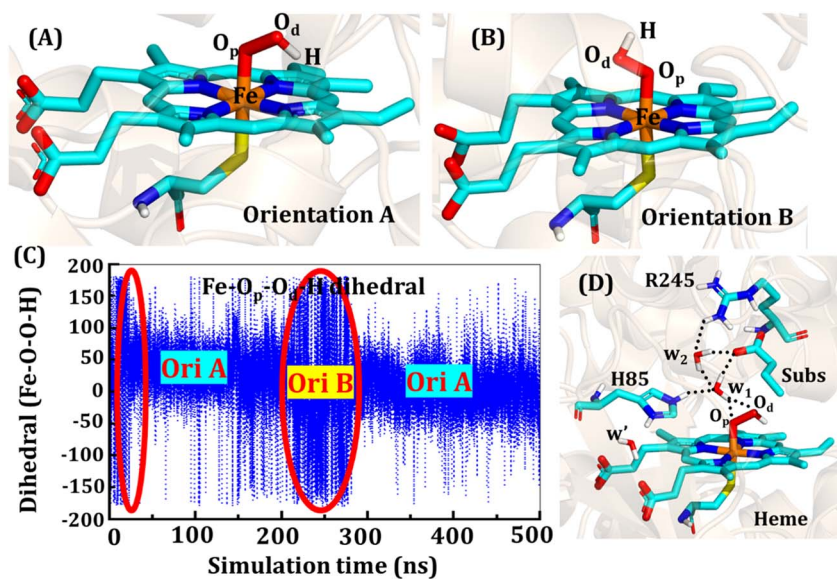

Fig. 9 Two prominent orientations (A and B) of the hydroperoxide moiety in Cpd 0. (C) Dihedral angle of Fe−O_p−O_d−H showing the rotation of OOH in the active site. The area encircled by the red ellipsoids shows the less prominent conformation in the time span of 500 ns. (D) Active site network near the active site of heme in the Cpd 0 state.

considered the amino acid residue of the H-bonding network for the protonation source. In this state, we have tested the nearest protonation source: (a) protonation *via* His85; (b) protonation *via* substrate carboxylic acid; (c) protonation *via* Arg245.

3.5.1 The mechanism of Cpd I formation. The second protonation at Cpd 0 leads to dehydration (removal of one water molecule) and Cpd I formation. In the previous section, as we have seen that, close to the active site (Fe–OOH$^-$) the hydrogen bonding network still suggests that for His85, Arg245 or the substrate might act as a proton donor source for Cpd 0.

3.5.2 Pathway 1: protonation *via* His85. To explore whether the conversion of Cpd 0 → Cpd I is initiated *via* a protonation from His85, we have performed QM/MM calculations. As can be seen in Fig. 10, the proton transfer from His85 to the heme moiety first generates Fe–H$_2$O$_2$ (RC), in a barrierless manner.

It is well documented in CYP450 s that Fe–H$_2$O$_2$ cleavage of proximal and distal oxygen (O$_p$ and O$_d$), either homolytic or heterolytic, generates the active oxidant Cpd I *via* dehydration.[54] A PES scanning was performed to test a similar mechanistic pathway for Cpd I generation. However, in the present case, the O$_p$–O$_d$ cleavage occurred with a feasible energy barrier of 16.6 kcal mol^{-1}, generating a free hydroxyl (O$_d$–H$_d$) at intermediate IM. The state IM was unstable on the potential energy surface, and recombination of proximal Hydrogen (Hp) to the free hydroxyl moiety generates a new water molecule (H$_p$–O$_d$–H$_d$) and active oxidant Cpd I (PC in Fig. 10). Note that the generated

Cpd I is 5.4 kcal mol^{-1} endothermic to RC (Fe–H$_2$O$_2$). Similarly, the Cpd 0 → Cpd I formation through Arg245 was also found to be energetically unfavorable (*cf.* Fig. S9†). Therefore, we deemed this pathway to be unfavourable.

3.5.3 Pathway2: protonation *via* the substrate. The second possibility of Cpd I generation involves the carboxylic acid hydrogen of the substrate, where the proton transfer from the substrate to the proximal oxygen (O$_p$) occurred during geometry optimization only and generated Fe–H$_2$O$_2$ (RC). Further cleavage of the O$_p$–O$_d$ step occurred at an energy expense of 8.46 kcal mol^{-1} and generated the intermediate IM (*cf.* Fig. 11), along with a free hydroxyl group. It is noteworthy that here the IM is −1.06 kcal mol^{-1} lower than Fe–H$_2$O$_2$ and is stable on the potential energy surface. Further recombination of proximal hydrogen (H$_p$) to the free hydroxyl (O$_d$–H$_d$) leads to the generation of a new water molecule (H$_p$–O$_d$–H$_d$) and generates the energetically stable active oxidant Cpd I (*cf.* Fig. 11).

It is interesting to note that the Cpd I species, generated *via* this pathway, is highly stable (exothermic by −24.6 kcal mol^{-1}) relative to RC. Thus, we can conclude that, in the fused protein OleT-BM3R, the Cpd 0 → Cpd I formation step is mediated by proton donation from the substrate. It is noteworthy that unlike other CYP450 monooxygenases, *e.g.*, CYP450$_{BM3}$ where the hydrophobic terminal of the fatty acid substrate occupies the heme site, in the CYP450 peroxygenase family enzyme, the acidic terminal occupies the heme site. Therefore, this finding provides an evolutionary rationale that the present orientation of the
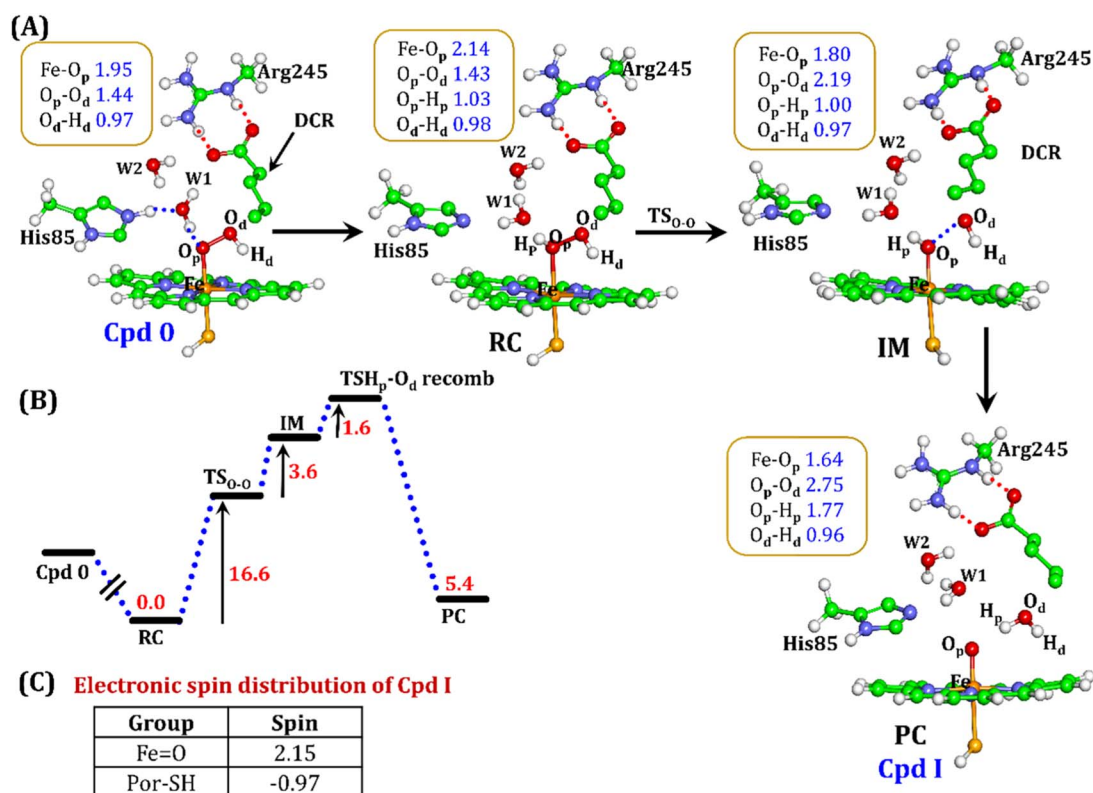

Fig. 10 (A) QM/MM optimized geometries for the Cpd 0 → Cpd I formation. Note that His85 has been considered as the proton source. (B) Energy profile diagram for Cpd 0 → Cpd I formation. Energies are reported at the B2-D3 + ZPE level of theory. All energies are in kcal mol^{-1}. (C) The spin distribution of the ultimate oxidant Cpd I.
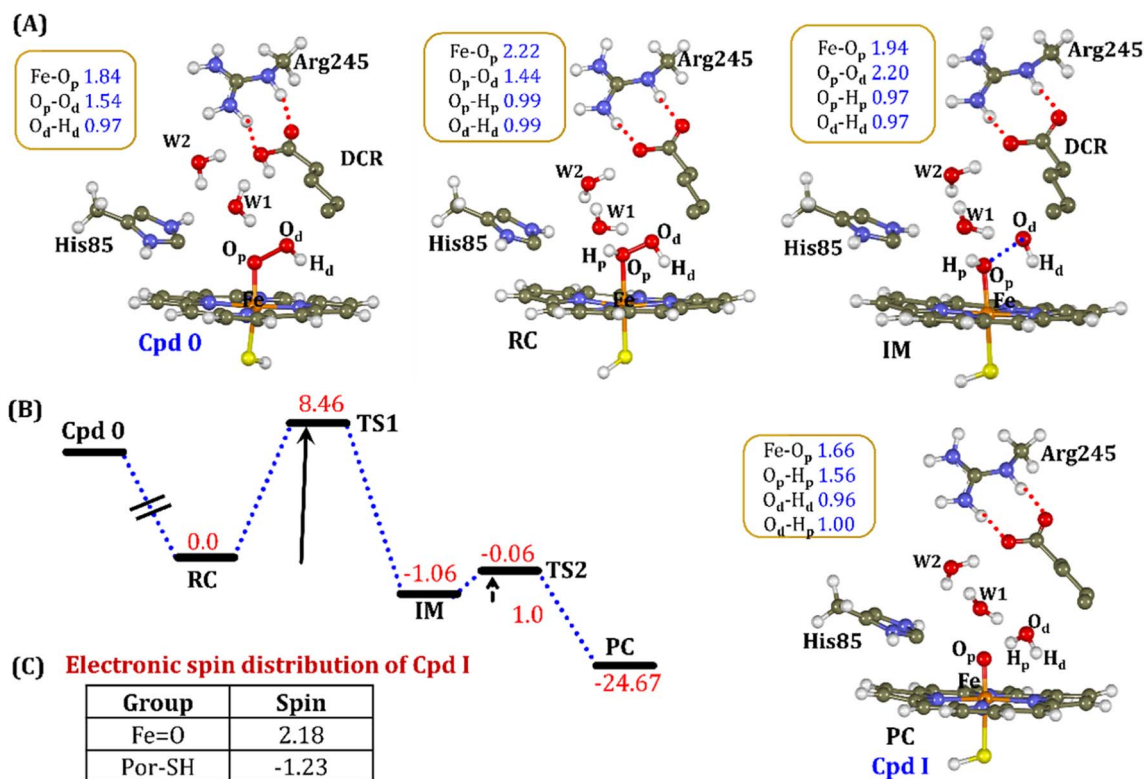
Fig. 11 (A) QM/MM optimized geometries during the Cpd 0 → Cpd I formation. Note that the carboxylic proton of the substrate has been considered as the proton source. (B) Energy profile diagram for the Cpd 0 → Cpd I formation. Energies are reported at the B2-D3 + ZPE level of theory and relative to RC-opt (Fe−H$_2$O$_2$). All energies are in kcal mol^{-1}. (C) The spin distribution of the ultimate oxidant Cpd I.
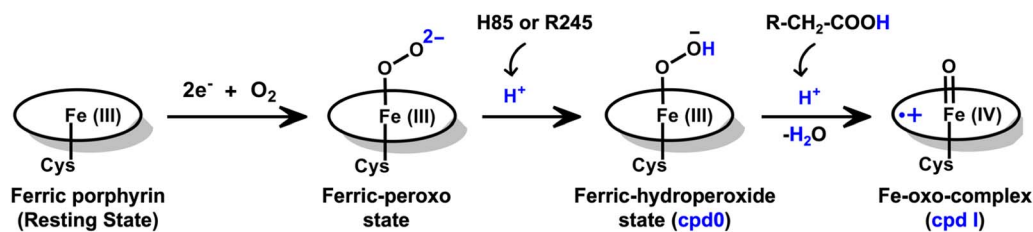
Scheme 2 An overall reaction scheme for the formation of Cpd I species in the fused enzyme CYP450$_{OleT}$-BM3R. The QM/MM calculations indicate that the initial protonation originates from either H85 or Arg245. Furthermore, the fatty acid substrate compensates for the absence of an acid−alcohol pair to facilitate the second protonation, ultimately leading to the formation of Cpd I within the fused CYP$_{OleT}$-BM3R enzyme.

substrate compensates for the lack of the acid–alcohol pair, which is needed for the second protonation, and thus the acid–alcohol pair is absent in CYP450 peroxygenases (Scheme 2).

4. Conclusion

The present study disentangles the mechanistic puzzles during the conversion of peroxide-based CYP450$_{OleT}$ to a reductase-based monooxygenase. The work reveals a fascinating choreography of the enzyme, and its capability to sense and respond to external stimuli (here the fusion of reductase). Using extensive MD simulations, we showed here the manner in which a far-located (from catalytic center) Tyr359 residue transduces an interaction with Gly459 of the reductase domain into a signal

that triggers the inflow of water and supplies the two protons needed for the reductase-based monooxygenase activity. These water chains further mediate the first protonation for the efficient formation of ferric-hydroperoxide (Cpd 0) necessary for the catalytic cycle. We demonstrate here that the acidic end of the fatty acid compensates for the lack of the acid–alcohol pair needed for the second protonation. This rationalizes the evolutionary mechanism whereby, unlike CYP450$_{BM3}$, CYP450 peroxygenase positions the fatty acid substrates with its acidic terminal close to heme-porphyrin. The present study provides a broader perspective with two important take-home lessons:

● The regulation of the catalytic activity is not only limited to the first or second catalytic shell of an enzyme. In fact, far residing residues regulate the present catalytic mechanism.

• The positioning of tyrosine (Tyr359) at a similar position in other members of CYP450 peroxygenases can convert these enzymes to reductase-based monooxygenases.

Data availability

Data necessary to produce the results are provided in the ESI.† It contains the QM optimized geometries for all species used in the calculations.

Author contributions

SY performed the calculation. KDD supervised the project. SS and KDD wrote the manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

References

1 P. P. Peralta-Yahya, F. Zhang, S. B. Del Cardayre and J. D. Keasling, Microbial Engineering for the Production of Advanced Biofuels, *Nature*, 2012, **488**(7411), 320–328.

2 S. K. Lee, H. Chou, T. S. Ham, T. S. Lee and J. D. Keasling, Metabolic Engineering of Microorganisms for Biofuels Production: From Bugs to Synthetic Biology to Fuels, *Curr. Opin. Biotechnol.*, 2008, **19**(6), 556–563.

3 D. F. Savage, J. Way and P. A. Silver, Defossiling Fuel: How Synthetic Biology Can Transform Biofuel Production, *ACS Chem. Biol.*, 2008, **3**(1), 13–16.

4 A. L. Demain, Biosolutions to the Energy Problem, *J. Ind. Microbiol. Biotechnol.*, 2009, **36**(3), 319–332.

5 F. Zhang, S. Rodriguez and J. D. Keasling, Metabolic Engineering of Microbial Pathways for Advanced Biofuels Production, *Curr. Opin. Biotechnol.*, 2011, **22**(6), 775–783.

6 Y. Qiu, C. Tittiger, C. Wicker-Thomas, G. Le Goff, S. Young, E. Wajnberg, T. Fricaux, N. Taquet, G. J. Blomquist and R. Feyereisen, An Insect-Specific P450 Oxidative Decarbonylase for Cuticular Hydrocarbon Biosynthesis, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**(37), 14858–14863.

7 T. D. Niehaus, S. Okada, T. P. Devarenne, D. S. Watt, V. Sviripa and J. Chappell, Identification of Unique Mechanisms for Triterpene Biosynthesis in Botryococcus Braunii, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**(30), 12260–12265.

8 M. A. Rude, T. S. Baron, S. Brubaker, M. Alibhai, S. B. Del Cardayre and A. Schirmer, Terminal Olefin (1-Alkene) Biosynthesis by a Novel P450 Fatty Acid Decarboxylase from Jeotgalicoccus Species, *Appl. Environ. Microbiol.*, 2011, **77**(5), 1718–1727.

9 J. Belcher, K. J. McLean, S. Matthews, L. S. Woodward, K. Fisher, S. E. J. Rigby, D. R. Nelson, D. Potts, M. T. Baynham, D. A. Parker, D. Leys and A. W. Munro, Structure and Biochemical Properties of the Alkene Producing Cytochrome P450 OleTJE (CYP152l1) from the Jeotgalicoccus Sp. 8456 Bacterium, *J. Biol. Chem.*, 2014, **289**(10), 6535–6550.

10 M. Alibhai, M. Rude and A. Schirmer, Methods and Compositions for Producing Olefins, *US Pat.*, US8183028B2, May 22, 2012.

11 Y. V. Popov, Y. L. Uskach, S. M. Ledenev, D. A. Kroman and V. A. Pavlova, Synthesis of a Sulfur-Containing Additive to Transmission Lubricants Based on the C 20-C 26 Fraction of α-Olefins, *Russ. J. Appl. Chem.*, 2013, **86**, 447–449.

12 L. I. Kioupis and E. J. Maginn, Molecular Simulation of Poly-α-Olefin Synthetic Lubricants: Impact of Molecular Architecture on Performance Properties, *J. Phys. Chem. B*, 1999, **103**(49), 10781–10790.

13 A. W. Munro, K. J. McLean, J. L. Grant and T. M. Makris, Structure and Function of the Cytochrome P450 Peroxygenase Enzymes, *Biochem. Soc. Trans.*, 2018, **46**(1), 183–196.

14 Y. Liu, C. Wang, J. Yan, W. Zhang, W. Guan, X. Lu and S. Li, Hydrogen Peroxide-Independent Production of α-Alkenes by OleT JE P450 Fatty Acid Decarboxylase, *Biotechnol. Biofuels*, 2014, **7**(1), 1–12.

15 C. Lu, F. Shen, S. Wang, Y. Wang, J. Liu, W. J. Bai and X. Wang, An Engineered Self-Sufficient Biocatalyst Enables Scalable Production of Linear α-Olefins from Carboxylic Acids, *ACS Catal.*, 2018, **8**(7), 5794–5798.

16 S. Shaik, S. Cohen, Y. Wang, H. Chen, D. Kumar and W. Thiel, P450 Enzymes: Their Structure, Reactivity, and Selectivity - Modeled by QM/MM Calculations, *Chem. Rev.*, 2010, **110**(2), 949–1017.

17 S. Shaik and K. D. Dubey, The Catalytic Cycle of Cytochrome P450: A Fascinating Choreography, *Trends Chem.*, 2021, **3**(12), 1027–1044.

18 S. Shaik, D. Kumar, S. P. de Visser, A. Altun and W. Thiel, Theoretical Perspective on the Structure and Mechanism of Cytochrome P450 Enzymes, *Chem. Rev.*, 2005, **105**(6), 2279–2328.

19 B. Meunier, S. P. De Visser and S. Shaik, Mechanism of Oxidation Reactions Catalyzed by Cytochrome P450 Enzymes, *Chem. Rev.*, 2004, **104**(9), 3947–3980.

20 S. Govindaraj and T. L. Poulos, The Domain Architecture of Cytochrome P450BM-3, *J. Biol. Chem.*, 1997, **272**(12), 7915–7921.

21 J. T. Hazzard, S. Govindaraj, T. L. Poulos and G. Tolling, Electron Transfer between the FMN and Heme Domains of Cytochrome P450BM- 3: Effects of Substrate and CO, *J. Biol. Chem.*, 1997, **272**(12), 7922–7926.

22 K. D. Dubey and S. Shaik, Choreography of the Reductase and P450BM3 Domains Toward Electron Transfer Is Instigated by the Substrate, *J. Am. Chem. Soc.*, 2018, **140**(2), 683–690.

23 J. A. Amaya, D. Batabyal and T. L. Poulos, Proton Relay Network in the Bacterial P450s: CYP101A1 and CYP101D1, *Biochemistry*, 2020, **59**(31), 2896–2902.

24 D. Fishelovitch, S. Shaik, H. J. Wolfson and R. Nussinov, How Does the Reductase Help to Regulate the Catalytic Cycle of Cytochrome P450 3A4 Using the Conserved Water Channel?, *J. Phys. Chem. B*, 2010, **114**(17), 5964–5970.

25 S. Kalita, S. Shaik, H. K. Kisan and K. D. Dubey, A Paradigm Shift in the Catalytic Cycle of P450: The Preparatory Choreography during O2Binding and Origins of the Necessity for Two Protonation Pathways, *ACS Catal.*, 2020, **10**(19), 11481–11492.

26 O. de Montellano, Structure and Function of Cytochrome P450 Enzymes, *Encycl. Drug Metab. Interact.*, 2012, **6**, 1–16.

27 T. I. Oprea, G. Hummer and A. E. García, Identification of a Functional Water Channel in Cytochrome P450 Enzymes, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**(6), 2133–2138.

28 I. F. Sevrioukova, H. Li, H. Zhang, J. A. Peterson and T. L. Poulos, Structure of a Cytochrome P450–Redox Partner Electron-Transfer Complex, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**(5), 1863–1868.

29 E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, UCSF Chimera—a Visualization System for Exploratory Research and Analysis, *J. Comput. Chem.*, 2004, **25**(13), 1605–1612.

30 A. Šali and T. L. Blundell, Comparative Protein Modelling by Satisfaction of Spatial Restraints, *J. Mol. Biol.*, 1993, **234**(3), 779–815.

31 K. Shahrokh, A. Orendt, G. S. Yost and T. E. Cheatham III, Quantum Mechanically Derived AMBER-compatible Heme Parameters for Various States of the Cytochrome P450 Catalytic Cycle, *J. Comput. Chem.*, 2012, **33**(2), 119–133.

32 P. Li and K. M. Merz Jr, *MCPB. Py: A Python Based Metal Center Parameter Builder*, ACS Publications, 2016.

33 C. I. Bayly, P. Cieplak, W. Cornell and P. A. Kollman, 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.

34 W. D. Cornell, P. Cieplak, C. I. Bayly and P. A. Kollman, Application of RESP Charges to Calculate Conformational Energies, Hydrogen Bond Energies, and Free Energies of Solvation, *J. Am. Chem. Soc.*, 2002, **115**(21), 9620–9631.

35 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, Comparison of Simple Potential Functions for Simulating Liquid Water, *J. Chem. Phys.*, 1983, **79**(2), 926–935.

36 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB, *J. Chem. Theory Comput.*, 2015, **11**(8), 3696–3713.

37 J. A. Izaguirre, D. P. Catarello, J. M. Wozniak and R. D. Skeel, Langevin Stabilization of Molecular Dynamics, *J. Chem. Phys.*, 2001, **114**(5), 2090–2098.

38 H. J. C. Berendsen, J. P. M. van Postma, W. F. Van Gunsteren, A. DiNola and J. R. Haak, Molecular Dynamics with Coupling to an External Bath, *J. Chem. Phys.*, 1984, **81**(8), 3684–3690.

39 J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, 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.

40 T. Darden, D. York and L. Pedersen, Particle Mesh Ewald: An N Log (N) Method for Ewald Sums in Large Systems, *J. Chem. Phys.*, 1993, **98**(12), 10089–10092.

41 R. Salomon-Ferrer, A. W. Gotz, D. Poole, S. Le Grand and R. C. Walker, Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald, *J. Chem. Theory Comput.*, 2013, **9**(9), 3878–3888.

42 P. Sherwood, A. H. de Vries, M. F. Guest, G. Schreckenbach, C. R. A. Catlow, S. A. French, A. A. Sokol, S. T. Bromley, W. Thiel, A. J. Turner, S. Billeter, F. Terstegen, S. Thiel, J. Kendrick, S. C. Rogers, J. Casci, M. Watson, F. King, E. Karlsen, M. Sjøvoll, A. Fahmi, A. Schäfer and C. Lennartz, QUASI: A General Purpose Implementation of the QM/MM Approach and Its Application to Problems in Catalysis, *J. Mol. Struct.: THEOCHEM*, 2003, **632**(1), 1–28.

43 S. Metz, J. Kästner, A. A. Sokol, T. W. Keal and P. Sherwood, ChemShell-a Modular Software Package for QM/MM Simulations, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2014, **4**(2), 101–110.

44 S. G. Balasubramani, G. P. Chen, S. Coriani, M. Diedenhofen, M. S. Frank, Y. J. Franzke, F. Furche, R. Grotjahn, M. E. Harding, C. Hättig, A. Hellweg, B. Helmich-Paris, C. Holzer, U. Huniar, M. Kaupp, A. Marefat Khah, S. Karbalaei Khani, T. Müller, F. Mack, B. D. Nguyen, S. M. Parker, E. Perlt, D. Rappoport, K. Reiter, S. Roy, M. Rückert, G. Schmitz, M. Sierka, E. Tapavicza, D. P. Tew, C. Van Wüllen, V. K. Voora, F. Weigend, A. Wodyński and J. M. Yu, TURBOMOLE: Modular Program Suite for Ab Initio Quantum-Chemical and Condensed-Matter Simulations, *J. Chem. Phys.*, 2020, **152**(18), 1–37.

45 R. Ahlrichs, M. Bär, M. Häser, H. Horn and C. Kölmel, Electronic Structure Calculations on Workstation Computers: The Program System Turbomole, *Chem. Phys. Lett.*, 1989, **162**(3), 165–169.

46 W. Smith and T. R. Forester, DL_POLY_2.0: A General-Purpose Parallel Molecular Dynamics Simulation Package, *J. Mol. Graphics*, 1996, **14**(3), 136–141.

47 A. Becke, Density-Functional Thermochemistry. III. The Role of Exact Exchange, *J. Chem. Phys.*, 1993, **98**, 5648.

48 J. Kästner, J. M. Carr, T. W. Keal, W. Thiel, A. Wander and P. D. L.-F. I. N. D. Sherwood, An Open-Source Geometry Optimizer for Atomistic Simulations, *J. Phys. Chem. A*, 2009, **113**(43), 11856–11865.

49 S. Grimme, Semiempirical GGA-type Density Functional Constructed with a Long-range Dispersion Correction, *J. Comput. Chem.*, 2006, **27**(15), 1787–1799.

50 A. Romero-Rivera, M. Garcia-Borras and S. Osuna, Role of Conformational Dynamics in the Evolution of Retro-Aldolase Activity, *ACS Catal.*, 2017, **7**(12), 8524–8532.

51 B. R. I. I. I. Miller, T. D. J. McGee, J. M. Swails, N. Homeyer, H. Gohlke and A. E. Roitberg, MMPBSA.Py: An Efficient

Program for End-State Free Energy Calculations, *J. Chem. Theory Comput.*, 2012, **8**(9), 3314–3321.

52 K. D. Dubey and S. Shaik, Cytochrome P450 - The Wonderful Nanomachine Revealed through Dynamic Simulations of the Catalytic Cycle, *Acc. Chem. Res.*, 2019, **52**(2), 389–399.

53 D.-S. Lee, A. Yamada, H. Sugimoto, I. Matsunaga, H. Ogura, K. Ichihara, S. Adachi, S.-Y. Park and Y. Shiro, Substrate Recognition and Molecular Mechanism of Fatty Acid Hydroxylation by Cytochrome P450 from Bacillus Subtilis: Crystallographic, Spectroscopic, and Mutational Studies, *J. Biol. Chem.*, 2003, **278**(11), 9761–9767.

54 R. Ramanan, K. D. Dubey, B. Wang, D. Mandal and S. Shaik, Emergence of Function in P450-Proteins: A Combined Quantum Mechanical/Molecular Mechanical and Molecular Dynamics Study of the Reactive Species in the H2O2-Dependent Cytochrome P450SPα and Its Regio-and Enantioselective Hydroxylation of Fatty Acids, *J. Am. Chem. Soc.*, 2016, **138**(21), 6786–6797.

55 Z. Wang, S. Shaik and B. Wang, Conformational Motion of Ferredoxin Enables Efficient Electron Transfer to Heme in the Full-Length P450TT, *J. Am. Chem. Soc.*, 2021, **143**(2), 1005–1016.