Modeling the ligand effect on the structure of CYP 450 within density functional theory

Matthew Dutra,[†] Shannon McElhenney,[†] Olivia Manley,[‡] Tom Makris,[‡] Vitaly Rassolov,[†] and Sophya Garashchuk[*,†]

[†]*Department of Chemistry & Biochemistry, University of South Carolina, Columbia, South Carolina 29208*

[‡]*Department of Molecular & Structural Biochemistry, North Carolina State University, Raleigh, North Carolina 27607*

E-mail: garashchuk@sc.edu

Abstract

An improved understanding of P450 structure is relevant to the development of biomimetic catalysts and inhibitors for controlled CH-bond activation, an outstanding challenge of synthetic chemistry. Motivated by the experimental findings of an unusually short Fe-S bond of 2.18 Å for the wild type (WT) OleT P450 decarboxylase relative to a cys-pocket mutant form (A369P), a computational model that captures the effect of the thiolate axial ligand on the iron-sulfur distance is presented. With computational efficiency and streamlined analysis in mind, this model combines a cluster representation of the enzyme – 40-110 atoms, depending on the heme and ligand truncation level – with a density functional theory (DFT) description of the electronic structure, and is calibrated against experimental data. The optimized Fe-S distances show a difference of 0.25 Å between the low and high spin states, in agreement with the crystallographic structures of the OleT WT and mutant forms. We speculate that this

difference is attributable to the packing of the ligand; the mutant is bulkier due to an alanine-to-proline replacement, meaning it is excluded from the energetically-favored low-spin minimum because of steric constraints. The presence of pure spin-state pairs and the intersection of the low/high spin states for the enzyme model is indicative of the limitations of single-reference electronic structure methods in such systems, and emphasizes the significance of using the proper state when modeling the hydrogen atom transfer (HAT) reaction catalyzed by OleT. At the same time, the correct characterization of both the short and the long Fe-S bonds within a small DFT-based model of 42 atoms paves the way for quantum dynamics modeling of the HAT step that initiates the OleT decarboxylation reaction.

1 Introduction

Theoretical studies of large molecular systems increasingly rely on electronic structure calculations for exploration of transition states and reaction intermediates, which are often inaccessible by experimental methods.[1–4] Computational studies of this type are frequently based on state-of-the-art multiscale – in space, time and theoretical rigor – simulations involving thousands of atoms described within a hybrid quantum mechanics/molecular mechanics (QM/MM) scheme. Within these approaches, the QM region itself can be treated at various levels of theory, ranging from the robust and efficient density functional theory (DFT)[5] for a few dozen to a hundred atoms to a more intensive coupled cluster[6] description of a few atoms within the active site, implemented via quantum embedding methods.[7,8] In this work we describe an alternative to the above approach, in which a practical electronic structure method (DFT) is used to develop a truncated molecular cluster model that is composed of a "small" number (~ 100) of atoms but is still representative of the full enzymatic system. These models are usually developed in conjunction with experimental data; the effect of long-range interactions can be incorporated through the experimental geometry of the selected atoms and molecular charge, while the choice of the density functional can be guided by

reproducing thermodynamic data, such as the temperature-dependent reaction rate, kinetic isotope effect (KIE), or the activation energy of a chemical process. Despite employing more conventional electronic structure methods, molecular cluster models remain useful tools for less conventional quantum dynamics studies of large systems, in which the impact of nuclear quantum effects on chemical reactivity can be elucidated.[9–11] These small models can also be readily adapted to new systems in conjunction with experiments.[12] Moreover, recent developments such as the rate constant expressions for the proton-coupled electron transfer reactions for the KIE in lipoxygenase,[13] a qualitative quantum rate model for hydrogen atom transfer (HAT) in soybean lipoxygenase,[14] and a measured KIE used as 'evidence for a Marcus-like model of hydride tunneling',[15] serve to highlight the role these various theoretical approaches can play in giving insight into the processes in extended molecular systems alongside state-of-the-art large-scale atomistic simulations.

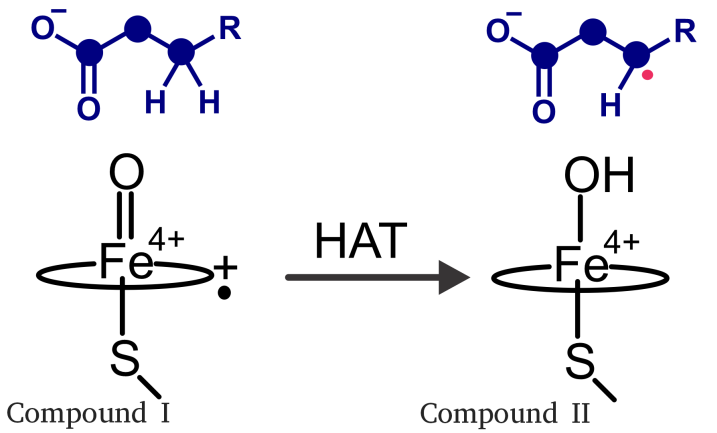
Controlled C-H bond activation is an outstanding challenge for synthetic chemistry, and enzymes often serve as a prototype for the "bio-inspired" design of small molecule catalysts. Cytochrome P450 enzymes (CYPs) often assume a central role due to their ability to perform efficient HAT on extremely challenging substrates. Despite over a half-century of intensive study, a number of fundamental questions remain regarding the mechanism for CYP-catalyzed HAT, and in particular, the factors that tune oxidative proficiency of the central ferryl-intermediate - an Fe(IV)-oxo $\pi-$cation radical species known as Compound I (CYP-I)[16] in this process. It is often thought that the chemical properties of this species can be 'tuned' by the thiolate axial ligand in a process often termed the "push" effect,[17–21] which is particularly relevant for HAT (Fig. 1).

Although the "push" effect was described as a means to rationalize the difference in reactivity between cysteine and histidine-ligated heme enzymes many decades ago, direct reactivity data on CYP ferryl intermediates has been elusive until recently,[16,22–24] largely due to the highly transient nature of these fleeting species. Therefore, a large body of work has relied on inferring how structural/electronic changes on more readily accessible

redox states (ferric/ferrous) can be used as a predictive marker for ferryl reactivity. Of particular interest is the potential role that the cysteine-pocket architecture (composed of several potential hydrogen bond donors to the thiolate ligand) can serve to attenuate ferryl reactivity, as best evidenced through the differences in oxidative proficiencies in the thiolate-ligated chloroperoxidase (CPO) and CYPs.[19] Experiments from one of our groups have shown that the ferryl intermediates in CYPs that incorporate H_2O_2 as a terminal oxidant (a small subgroup of CYPs known as the CYP152 family) are amenable to transient kinetics studies.[25] In particular, the CYP152 decarboxylase known as OleT is a convenient scaffold to probe how the cys-pocket microenvironment can tune ferryl reactivity. Recently we have investigated the cys-pocket variant A369P OleT structure, which differs from its wild-type (WT) counterpart in that it contains a mutation which replaces alanine with proline at a position adjacent to the cys-thiolate. This in turn has both structural (increased Fe-S bond length and varied hydrogen bonding network) and electronic (high-spin stabilization) ramifications, as measured by various crystallographic and spectroscopic techniques that are described elsewhere[25].

Theoretical efforts to characterize CYPs have been numerous, and a full survey is beyond the scope of this work (the interested reader is directed to Refs[26,27] among many others). However, we note that a plethora of computational approaches have been applied to these systems, ranging from electronic structure characterizations of truncated small-model representations of the enzymatic active sites to full-scale molecular dynamics and hybrid QM/MM calculations simulating entire steps in the catalytic cycle.[28] From an electronic structure (ES) perspective, the B3LYP density functional method[5,29,30] paired with modestly sized basis sets has been a common approach in numerous studies of metalloenzymes.[31] One such application with particular relevance to this work has been the theoretical investigation of the "push effect" by various axial ligands on the formation and reactivity of key intermediate species. Indeed, the chemical properties of this ligand in OleT have been correlated to both the length of the iron-oxo bond and overall activation energy for the transfer step of the

a)



b)

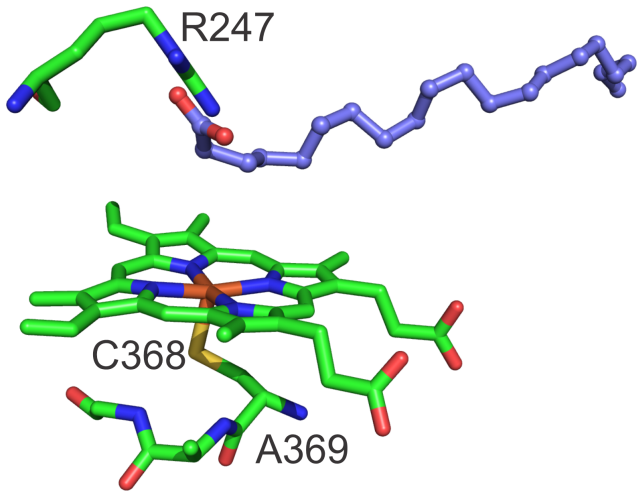
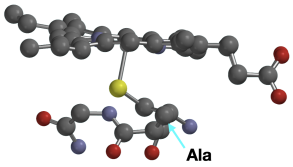


Figure 1: (a) Hydrogen atom transfer from the C3 position by Compound I that initiates decarboxylation in the cytochrome P450 OleT (b) active-site structure of ferric substrate-bound OleT. Compound I indicated in panel (a) is also referred to as 'Cpd I' in text.

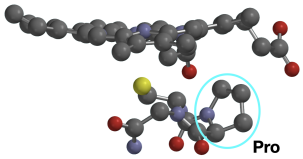
hydrogen abstraction process, emphasizing its apparent importance within the enzymatic complex.

The objectives of this work are to provide a theoretical explanation for the short (2.18 Å) and long (2.45 Å) Fe-S bond lengths observed in the ferric state of the active site in OleT WT and its mutant A369P, respectively (which are new experimental findings from one of our groups), and to develop a simple computational model which captures these two structures. This simplicity is essential for the subsequent theoretical modeling and analysis of the temperature-dependent kinetic isotope effect of the hydrogen abstraction from fatty acid substrates catalyzed by the WT enzyme. As is often the case in computational chemistry, we strive for a computational model which balances numerical efficiency with predictive power. Achieving such a balance is especially challenging for OleT (and large transition metal systems in general) due to its complex electronic structure consisting of nearly-degenerate doublet and quartet states. Such systems are best treated with multi-reference methods, although their unfavorable scaling drastically limits the overall computable system size. With this in mind, the cluster model developed here combines a cost-efficient DFT approach with an atomistic representation of the enzyme derived from experimental crystal structures of the WT and A369P forms of the OleT species, truncated at various levels (Fig. 2). This model reproduces key electronic and geometric features of the full enzymes via a tunable Fe-S bond stretch and – thanks to its modest size of 42 atoms – is amenable to higher levels of electronic structure theory, thus paving the way for theoretical investigation of the HAT step for OleT decarboxylation.

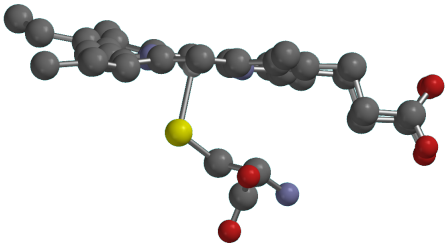
The remainder of the paper is organized as follows: the electronic structure method in the context of prior work is described in Section 2, the molecular models, results and analysis are presented in Section 3, and Section 4 concludes.



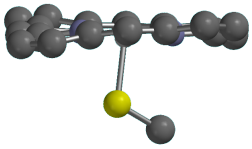
a) WT OleT



b) A369P



c) X:=S-Cys



d) X:=S-CH_3

Figure 2: Geometries of the OleT enzyme: (a) the wild type 'WT' and (b) the mutant A369P. The location of the mutated proline is specified by a circle in the latter, and the Fe-S distances are 2.18 Å and 2.45 Å respectively. The ligand, denoted as -X, is truncated in our models to X:=S-Cys (c) and X:=S-CH_3 (d). The carboxyl groups on the porphyrin ring have also been removed in (d).

2 Computational Models and Methods

In general, prior theoretical studies of OleT have been centered on model systems that can be decomposed into three principal components: the ferryl porphyrin radical cation $PorFe^{\bullet+}$, the axial ligand -X, and – depending on the HAT reaction step – the oxo- (hydroxy-) group =O (-OH). Since one of the aims of this work is to generate a minimal representation of Compound I that is compatible with high-level ES theory, we briefly summarize several oxo-containing structures with -X represented by thiols, as computed by Shaik *et al.*,[32] Green *et al.*,[33] Kamachi and Yoshizawa,[34] de Visser *et al.*,[35] and Harvey *et al.*[36,37] in Table 1.

Table 1: Geometric data for the CYP OleT computed for the doublet and quartet spin states, with the values for the latter given in parentheses as appropriate. RB3LYP stands for the restricted open-shell B3LYP method.

Ref.	-X ID	Fe-X (Å)	Fe=O (Å)	Comments
[32]	-SH	2.581 (2.600)	1.651 (1.648)	B3LYP/ECP+LACVP-6-31G
[32]	-SH	2.620 (2.669)	1.658 (1.655)	RB3LYP/ECP+LACVP-6-311+G*
[32]	-SCys	2.690 (2.698)	1.643 (1.642)	RB3LYP/ECP+LACVP-6-31G
[32]	$-SCH_3$	2.764 (2.826)	1.643 (1.648)	RB3LYP/ECP+LACVP-6-31G
[32]	$-SCH_3$	2.909 (2.946)	1.643 (1.643)	RB3LYP/ECP+LACVP-6-311+G*
[33]	$-SCH_3$	2.69	1.65	UB3LYP/6-311+G//6-311+G*
[34]	$-SCH_3$	2.694 (2.686)	1.644 (1.644)	B3LYP/TZV-D95
[35]	$-SCH_3$	2.539 (2.557)	1.631 (1.630)	UB3LYP/SVP "Model A"
[36]	$-SCH_3$	2.63	1.65	UB3LYP/LACV3P-6-31G*
[37]	$-SCH_3$	2.55 (2.55)	1.65	RB3LYP/LACVP-6-31G (QM/MM I)

These structures largely agree upon optimal Fe-S and Fe=O bond lengths of 2.6-2.7 Å and 1.65 Å respectively, although the newer QM/MM study by de Visser *et al.*[35] predicts shorter values for both of these quantities. Interestingly, these results conflict somewhat with recent spectroscopic measurements by Green *et al.*,[19] which yield average bond lengths of 2.4 Å for Fe-S and 1.67 Å for Fe=O in CYP450 complexes.

Given the overwhelming prevalence of density functional-based approaches in the literature and the demonstrated importance of the axial thiolate ligand to the electronic properties of the enzyme, we opt to use both restricted and unrestricted DFT calculations (performed using the QChem-5.3[38] package visualized in IQmol,[39] and with Spartan18[40]) to probe the

effect of the ligand identity on the electronic properties of the system. To this end, we take the experimental crystal structure for the cationic apoenzyme (Por-Fe^{+}-X) and sequentially prune the -X group, resulting in three substituents considered for the molecular model: the full thiolate ("full"), the cysteine residue (cys), and a methyl-capped sulfur (-SCH_3). For the latter, the carboxyl groups on the porphyrin ring have been removed as well. This process is similar in its implementation to the truncation of the QM region in the QM/MM simulations of similar CYP450 structures by Harvey *et al.*;[37] we note that this study concludes that the overall characteristics of their resting state models are independent of the size of the QM region, which is encouraging for our own "smallest workable model" mindset. Furthermore, the positive charge on the porphyrin ring (which is seen even in solvent) reduces the hydrogen-abstraction barrier towards the range consistent with experiments.[41] Note that this description of the apoenzyme as a charged species results in a pair of triplet/quintet spin states, as opposed to the doublet/quartet pair seen in the full enzyme. This is a consequence of the removal of the environment surrounding the heme center, which would carry a balancing negative charge in nature. The ES for these models is then examined in gas phase employing the B3LYP functional – which has been demonstrated to be adequate to describe small organic molecules,[42–44] the hydrogen abstraction reaction $FeO^{+} + H_2$,[45] and the heme center of CYP450 structures[46] – paired with 6-31G* and 6-31+G** basis sets for both the triplet (low-spin) and quintet (high-spin) states. The D3 dispersion correction[47] is included into calculations with the larger basis. Solvent models were also briefly considered but ultimately omitted, as they had minimal impact on the system properties relative to the gas phase calculations. This aligns with the conclusions in Ref.,[41] where the observed solvent effect on the π-hole on the porphyrin ring of the heme was small compared to the solvent effect on the charged substituents explored by the authors.

First we examine the high/low-spin states for the WT OleT and A369P mutant with the full thiolate ligand (106 and 110 atoms), both at the experimental geometry and with the ligand fully relaxed. Optimizations are performed at the B3LYP/6-31G* level for the

hydrogen atoms in each case, and the ligand when appropriate; for the latter, the single point energy is also computed at the B3LYP-D3/6-31+G** level. The spin gaps, the Fe-S bond length and the shortest S...H hydrogen bond distance are listed in Table 2. The experimental Fe-S distance is significantly shorter for the native enzyme (2.18 Å) compared to the A369P mutant (2.45 Å). Based on the geometry of the relaxed ligand, the short and long Fe-S distances are consistent with the low and high spin states, respectively, for both species even though the ligands drift away from the porphyrin ring within the truncated molecular model. Nevertheless, A369P is still characterized by a shorter S$\cdots$H bond regardless of the spin state: approximately 2.8 vs 3.6 and 2.5 vs 2.2 Å for the optimized and relaxed geometries, respectively. The energies of the quintet state are lower than that of the triplet state in each case, though for the WT OleT this gap is quite small in two out of three calculations, pointing to the limitation of the chosen ES method.

Next, we turn to further truncation of the model, focusing on the WT OleT system. At the cheaper B3LYP/6-31G* level of theory the quintet is consistently found to be the ground state by approximately 1.5-2.5 kJ/mol. Increasing the size of the basis and incorporating dispersion corrections changes the ordering, although the spin splitting still falls within $\sim$3 kJ/mol, as seen in Table 3. Based on these minimal changes in the spin gaps, we argue that -SCH_3 is an adequate stand-in for the full thiolate group, though in this case the exact ordering is beyond the accuracy achievable with DFT. Note that the $\langle S^2 \rangle$ values are consistent across the models, but the spin-contamination – somewhat pronounced for the triplet state ($\langle S^2 \rangle = 2.17$) – indicates that a fully accurate description of these states requires a multireference method. Within single-reference DFT, however, one can further analyze the enzyme spin states by comparing the restricted and unrestricted orbital descriptions.

Table 2: Properties of the spin states for the native (WT) and mutant (A369P) OleT enzyme (charge +1). Optimization of the hydrogen and ligand atoms are performed at B3LYP/6-31G* level; the B3LYP-D3/6-31+G** energies are single point calculations at the geometry with the optimized ligand. Unrestricted Kohn-Sham orbitals are used in all cases. The low/high spin (2 vs 4 unpaired electrons) energy gap is denoted $\Delta_s = E_2 - E_4$. The distances are listed in Å.

method	6-31G* exp		6-31G* opt		D3/6-31+G**	
system	Fe-S	S$\cdots$H	Fe-S	S$\cdots$H	Fe-S	S$\cdots$H
WT triplet	2.180	2.515	2.206	3.647	2.206	3.647
WT quintet	2.180	2.514	2.437	3.527	2.437	3.527
Δ_s [kJ/mol]	1.88		17.07		2.69	
A369P triplet	2.452	2.193	2.198	2.851	2.198	2.851
A369P quintet	2.452	2.195	2.431	2.740	2.431	2.740
Δ_s [kJ/mol]	39.89		20.62		11.60	

Table 3: The energy gaps between the triplet and quintet spin states for three molecular models of the WT OleT (charge +1). The molecular models, labeled CH_3, Cys and 'full', are shown in Fig. 2(d,c,a). Positions of the heavy atoms are taken at the experimental geometry (Fe-S=2.18 Å.); positions of the hydrogen atoms are optimized at the B3LYP/6-31G* level. The B3LYP-D3/6-31+G** energies are single point calculations. Unrestricted Kohn-Sham orbitals are used in all cases. The low/high (triplet/quintet) spin gap is $\Delta_s = E_2 - E_4$.

method	B3LYP/6-31G*			B3LYP-D3/6-31+G**		
	$\langle S^2 \rangle$		Δ_s	$\langle S^2 \rangle$		Δ_s
model	2e	4e	[kJ/mol]	2e	4e	[kJ/mol]
CH_3	2.17	6.07	2.23	2.15	6.07	0.15
Cys	2.17	6.07	1.96	2.14	6.07	-2.36
full	2.17	6.07	1.88	2.14	6.07	-3.15

3 Results and Discussion

In light of the wide range of reported Fe-S distances (e.g. Table 1), we first compute the energy profile of the Fe-S stretch using the methyl-capped ligand model. The positions of the heavy atoms on the ring are taken at the experimental geometry of OleT, while the Fe-S bond length is scanned and the positions of the -CH_3 atoms optimized. The energies as a function of the Fe-S stretch, computed at the B3LYP-D3 level of theory with two sizes of basis set, are shown in Fig. 3. The smaller basis set is again taken to be 6-31G*, while the larger basis adds diffuse functions on the nonmetal atoms and adopts the m6-31G functions[48,49] for the iron center. The restricted open-shell Kohn-Sham densities converge to two energetically-separated sets of triplet (low-spin) and quintet (high-spin) states, shown in Fig. 3(a). In the low-spin cases, the spin density is localized on S and Fe in the low-energy state, and additionally distributed across the porphyrin ring in the high-energy state (see panels (a) and (b) of Fig. 4). These two states are separated in energy by approximately 30 kJ/mol near their minima (Fe-S = 2.2 Å), and the curves are essentially parallel (Fig. 3(b)). The high-spin states are separated by approximately 38 kJ/mol near the low-energy minimum at Fe-S = 2.4 Å, and the energy splitting between the low-energy triplet and quintet states at their respective minima is $\sim$ 15 kJ/mol. For completeness, we note that a similar four-state energy splitting is seen in the model Cpd I enzyme of Shaik *et al.*,[32,50] where the ground state is composed of a degenerate doublet/quartet pair corresponding to a structure with an unpaired electron in the porphyrin ring and the higher-energy state is a degenerate doublet/quartet pair with an unpaired electron allotted to the sulfur atom. In contrast to the present work, however, that model (still using the -SCH_3 axial ligand) displays a very small low/high energy splitting (<2.1 kJ/mol) between the triplet states and an elongated Fe-S bond (2.7-2.9 Å).

We have also observed two triplet states in the unrestricted calculations, characterized by $\langle S^2 \rangle \sim 2.4$ and $\langle S^2 \rangle \sim 2.2$. The former, which is lower in energy by $\sim$ 28 kJ/mol, is shown as the low-lying black curve with circles Fig. 3(b). The low-spin curves exhibit minima at $\sim$ 2.25

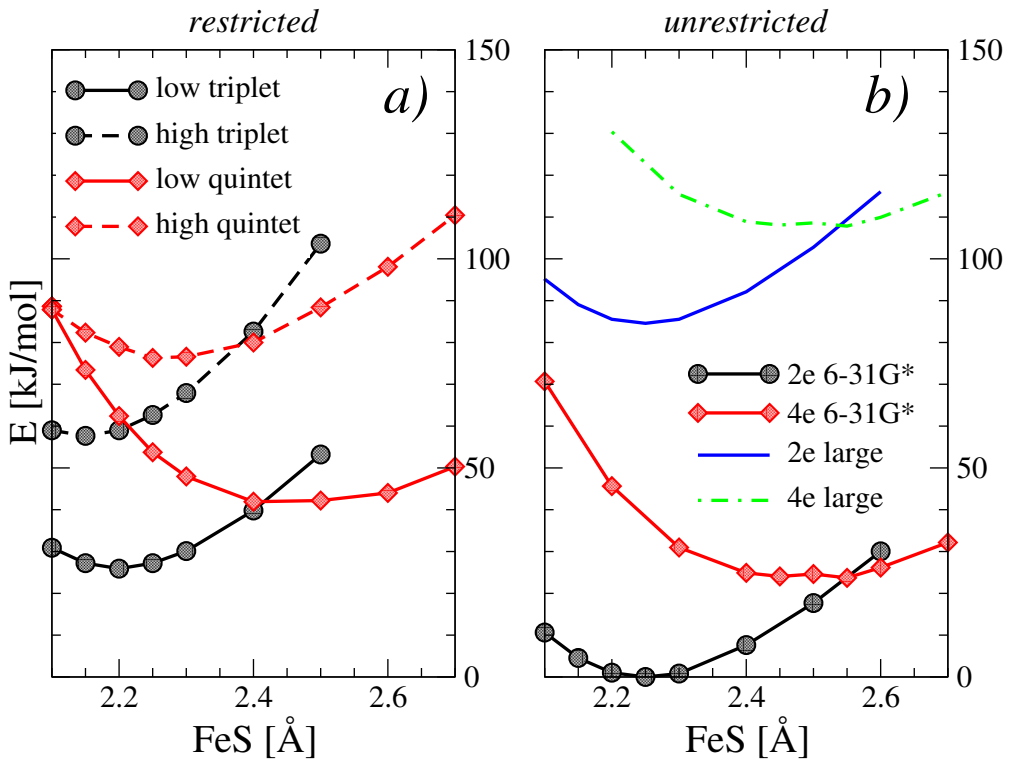
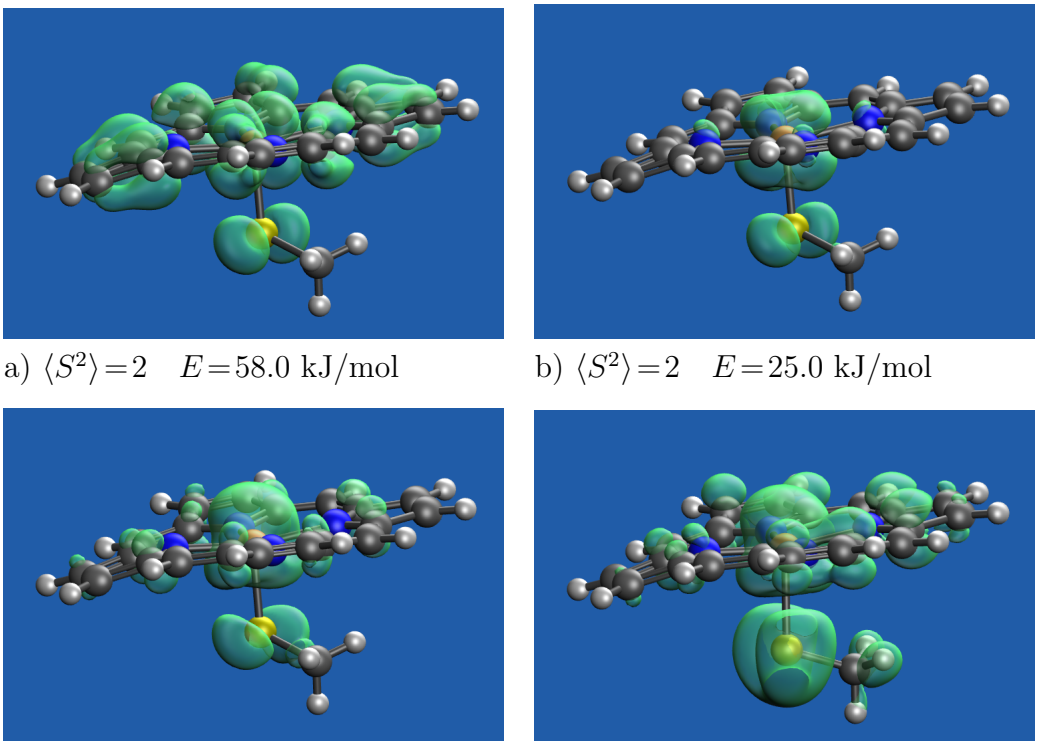


Figure 3: Model enzyme (heme^{+}-SCH_3): (a) the energy profiles as functions of Fe-S distance for the triplet and quintet states obtained with restricted B3LYP-D3/6-31G*. (b) Unrestricted calculations using B3LYP-D3/6-31G* and B3LYP-D3/m6-31G(6-31+G*); the latter composite basis is labeled as 'large' in the legend. The curves in both panels are shifted by the same energy value to place the lowest minimum in (b) at zero kJ/mol. The spin states in (b) are labeled by the number of unpaired electrons.

Å, and their high-spin counterparts exhibit minima at a stretched distance of $\sim 2.4 - 2.55$ Å for both sizes of basis used. This is consistent with many experimental measurements – including those of Makris, whose crystal structures are used here – falling into 'short' and 'long' Fe-S bond categories (e.g. Ref.[19]). In our case the native and mutant forms of OleT are experimentally characterized by the Fe-S distances of 2.18 Å and 2.45 Å, respectively. Within the single determinant ES theory, the positions of these minima correlate with the low- and high-spin states, their energy curves crossing around 2.5 Å. Both the low- and high-spin structures obtained from the unrestricted calculations display an 'intermediate' character, as seen in Fig. 4(c,d); the similarity of the spin density distributions between these cases illustrates spin-state mixing and a breakdown of the single-reference description. Overall, the energy profile of the unrestricted calculation correlates with the low-energy pure triplet and quintet profile quite well. Thus, within this ES model, the difference in the Fe-S bond lengths measured for the WT enzyme and the A369P mutant can be attributed to the packing of their ligands, the latter of which replaces alanine with a bulkier proline residue. This results in the A369P structure falling into the quintet minimum of 2.45 Å instead of the measured 2.18 Å for the WT. Hydrogen bonding between the ligand and the sulfur atom – expected to be stronger in the mutant than the WT structure, based on S-H distances[51] – may also contribute to the stabilization of the high-spin/longer Fe-S enzyme geometry.

The significance of the short vs long Fe-S bond lies in its impact on the energetics of hydrogen abstraction by Cpd I, sketched in Fig. 1. As discussed in Ref.,[19] a shorter Fe-S bond results in greater electron donation from the ligand into the Fe π^* antibonding orbitals, which in turn yields a slightly longer ($\sim$0.01Å) and weaker Fe=O bond. This then correlates with more efficient HAT in the hydrogen abstraction step of the OleT enzymatic cycle (see Fig. 1). Within our DFT description we observe a similar trend. As seen in Table 4, the O-optimized Fe-O bond for both the truncated and full WT structures is $\sim$ 0.01(0.05) Å longer than that of their A369P counterparts in the low(high) spin state. Furthermore, the difference in Fe-O values between the WT and A369P forms in their energetically favored



a) $\langle S^2 \rangle = 2 \quad E = 58.0$ kJ/mol

b) $\langle S^2 \rangle = 2 \quad E = 25.0$ kJ/mol

c) $\langle S^2 \rangle = 2.44 \quad E = 0$

d) $\langle S^2 \rangle = 6.07 \quad E = 23.6$ kJ/mol

Figure 4: Spin density computed at the RB3LYP/6-31G* level for the (a) low and (b) high energy triplet states, and using UB3LYP/6-31G* for the (c) triplet and (d) quintet states. Positions of CH_3 are optimized using Fe-S=2.2 Å for the triplet and Fe-S=2.5 Å for the quintet at the UB3LYP/6-31G* level. The porphyrin positions are frozen at the experimental geometry of the OleT WT.

states is on the order of 0.005-0.01Å in both models. Therefore, we expect that HAT facilitated by A369P will be characterized by a higher activation energy and proceed less efficiently than via WT OleT. Optimized Cartesian coordinates for Cpd I, along with those for the apoenzyme models, are given in SI.

4 Conclusions

In summary, we have presented a theoretical model of the OleT apoenzyme and its A369P mutant derived from experimental crystal structures based on the DFT/B3LYP electronic structure method. This model, composed of a central iron atom ligated by a hydrated porphyrin ring and an axial $-SCH_3$ group, exhibits a triplet/quintet spin gap comparable (i.e. within the accuracy of the method) to that obtained for a similar structure with the full axial ligand; we thus argue that the electronic structure of the enzyme is adequately captured by a smaller and more computationally practical model. Subsequent calculations mapping the Fe-S stretching coordinate show a four-state splitting of the enzyme into low- and high-energy triplet/quintet states – a result corroborated by other computational studies of similar systems.

Using the location of the energy minima along this stretching motion, we then identify the low-spin, low-energy state (Fe-S $\approx$ 2.25 Å) as the one preferred by the wild-type apoenzyme, and the high-spin, low energy state (Fe-S $\approx$ 2.5 Å) as the one preferred by the A369P mutant, based on comparison to experimental data ($\text{OleT}_{\text{Fe-S}} = 2.18$ Å , $\text{A369P}_{\text{Fe-S}} = 2.45$ Å). These results further validate our minimal model, as it is capable of capturing the experimentally-measured low- vs. high-spin preference of the WT and A369P species, respectively, by simply elongating the Fe-S bond.

Given the relative energies of the triplet and quintet states, we suggest that the proclivity of the mutant to a lengthened Fe-S bond is due to the bulkiness of its ligand, which substitutes proline in place of alanine as the sulfur-adjacent residue and thereby prohibits relaxation

into the triplet minimum. It is also possible that the shortened S-H distances in the mutant produce stronger hydrogen bonding than is found in the wild-type, which would further stabilize this elongated Fe-S configuration, although this effect cannot be captured using the truncated model.

Finally, the Compound I structure corresponding to the WT OleT is characterized by a slightly longer (and therefore weaker) Fe=O bond than that of the mutant, according to the distances computed for both the full and the methyl-capped ligands across both spin states. The native enzyme is thus expected to be more proficient at the hydrogen atom transfer step of the OleT decarboxylation reaction than its A369P counterpart. With respect to the computational model itself, while the system truncation and reliance on the DFT methods limit the accuracy of the electronic structure analysis, the efficiency of the model paves the way for future quantum dynamics studies and modeling of HAT and associated KIE for the OleT decarboxylation reaction.

Table 4: The optimal Fe=O distances in Å for the -SCH_3 model and for the full ligand (heme truncated to porphyrin) at the experimental geometry of the enzyme computed at the B3LYP-D3/6-31G* level. The spin gap, given in kJ/mol, is defined as $\Delta_s = E_2 - E_4$. The lowest-energy Fe-O distance for each model is set in bold face.

model	CH_3			full ligand		
enzyme	Fe=O 2e	Fe=O 4e	Δ_s	Fe=O 2e	Fe=O 4e	Δ_s
WT	**1.651**	1.714	-2.6	**1.656**	1.698	-2.4
A369P	1.635	**1.647**	3.6	1.642	**1.647**	9.7

5 Supporting Information

Wild-type and A369P mutant enzyme structures as .xyz files, and additional tabulated electronic structure properties.

6 Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. CHE-1565985 and No. CHE-1955768 (SG), and Grant No. CHE-1555066 (TMM), as well as an ASPIRE award from UofSC VP for research office; calculations were performed in part under an XSEDE[52] allocation TG-DMR110037.

References

(1) van der Kamp, M. W.; Mulholland, A. J. Combined Quantum Mechanics/Molecular Mechanics (QM/MM) Methods in Computational Enzymology. *Biochemistry* **2013**, *52*, 2708.

(2) Daniels, A. D.; Campeotto, I.; van der Kamp, M. W.; Bolt, A. H.; Trinh, C. H.; Phillips, S. E. V.; Pearson, A. R.; Nelson, A.; Mulholland, A. J.; Berry, A. Reaction Mechanism of N -Acetylneuraminic Acid Lyase Revealed by a Combination of Crystallography, QM/MM Simulation, and Mutagenesis. *ACS Chem. Biol.* **2014**, *9*, 1025.

(3) Chen, Z.; Yin, G. The Reactivity of the Active Metal Oxo and Hydroxo Intermediates and Their Implications in Oxidations. *Chem. Soc. Rev.* **2015**, *44*, 1083–1100.

(4) Haatveit, K.; Garcia-Borràs, M.; Houk, K. Computational Protocol to Understand P450 Mechanisms and Design of Efficient and Selective Biocatalysts. *Front. Chem.* **2019**, *6*, 1–6.

(5) Parr, R. G.; Yang, W. *Density-Functional Theory of Atoms and Molecules*; New York: Oxford University Press, 1989.

(6) Bartlett, R. J.; Dykstra, C. E.; Paldus, J. *Advanced Theories and Computational Approaches to the Electronic Structure of Molecules*; Springer, 1984; Vol. 133; Chapter Coupled-Cluster Methods for Molecular Calculations.

(7) Bennie, S. J.; van der Kamp, M. W.; Pennifold, R. C. R.; Stella, M.; Manby, F. R.; Mulholland, A. J. A Projector-Embedding Approach for Multiscale Coupled-Cluster Calculations Applied to Citrate Synthase. *J. Chem. Theory Comput.* **2016**, *12*, 2689–2697.

(8) Yu, K.; Carter, E. A. Extending Density Functional Embedding Theory for Covalently Bonded Systems. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, E10861–E10870.

(9) Pu, J.; Gao, J.; Truhlar, D. G. Multidimensional Tunneling, Recrossing, and the Transmission Coefficient for Enzymatic Reactions. *Chem. Rev.* **2006**, *106*, 3140–3169.

(10) Jakowski, J.; Huang, J.; Garashchuk, S.; Luo, Y.; Hong, K.; Keum, J.; Sumpter, B. G. Deuteration As a Means to Tune Crystallinity of Conducting Polymers. *J. Phys. Chem. Lett* **2017**, *8*, 4333–4340.

(11) Klinman, J. P.; Kohen, A. Hydrogen Tunneling Links Protein Dynamics to Enzyme Catalysis. *Annu. Rev. Biochem.* **2013**, *82*, 471–496.

(12) Klinman, J. P.; Offenbacher, A. R.; Hu, S. Origins of Enzyme Catalysis: Experimental Findings for C–H Activation, New Models, and Their Relevance to Prevailing Theoretical Constructs. *J. Am. Chem. Soc.* **2017**, *139*, 18409–18427.

(13) Soudackov, A. V.; Hammes-Schiffer, S. Proton-Coupled Electron Transfer Reactions: Analytical Rate Constants and Case Study of Kinetic Isotope Effects in Lipoxygenase. *Faraday Discuss.* **2016**, *195*, 171–189.

(14) Jevtic, S.; Anders, J. A Qualitative Quantum Rate Model for Hydrogen Transfer in Soybean Lipoxygenase. *J. Chem. Phys.* **2017**, *147*, 114108.

(15) Howe, G. W.; van der Donk, W. A. Temperature-Independent Kinetic Isotope Effects As Evidence for a Marcus-like Model of Hydride Tunneling in Phosphite Dehydrogenase. *Biochemistry* **2019**, *58*, 4260–4268.

(16) Rittle, J.; Green, M. T. Cytochrome P450 Compound I: Capture, Characterization, and C-H Bond Activation Kinetics. *Science* **2010**, *330*, 933–937.

(17) Ehudin, M. A.; Gee, L. B.; Sabuncu, S.; Braun, A.; Moënne-Loccoz, P.; Hedman, B.; Hodgson, K. O.; Solomon, E. I.; Karlin, K. D. Tuning the Geometric and Electronic Structure of Synthetic High-Valent Heme Iron(IV)-Oxo Models in the Presence of a Lewis Acid and Various Axial Ligands. *J. Am. Chem. Soc.* **2019**, *141*, 5942–5960.

(18) Kang, Y.; Chen, H.; Jeong, Y. J.; Lai, W.; Bae, E. H.; Shaik, S.; Nam, W. Enhanced Reactivities of Iron(iv)-Oxo Porphyrin π-Cation Radicals in Oxygenation Reactions by Electron-Donating Axial Ligands. *Chem. Eur. J.* **2009**, *15*, 10039–10046.

(19) Krest, C.; Silakov, A.; Rittle, J.; Yosca, T.; Onderko, E.; Calixto, J.; Green, M. Significantly Shorter Fe-S Bond in Cytochrome P450-I is Consistent with Greater Reactivity Relative to Chloroperoxidase. *Nature Chem.* **2015**, *7*, 696–702.

(20) Ohno, T.; Suzuki, N.; Dokoh, T.; Urano, Y.; Kikuchi, K.; Hirobe, M.; Higuchi, T.; Nagano, T. Remarkable Axial Thiolate Ligand Effect on the Oxidation of Hydrocarbons by Active Intermediate of Iron Porphyrin and Cytochrome P450. *J. Inorg. Biochem.* **2000**, *82*, 123–125.

(21) Aldag, C.; Gromov, I. A.; García-Rubio, I.; von Koenig, K.; Schlichting, I.; Jaun, B.; Hilvert, D. Probing the Role of the Proximal Heme Ligand in Cytochrome P450cam by Recombinant Incorporation of Selenocysteine. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 5481–5486.

(22) Grant, J. L.; Hsieh, C. H.; Makris, T. M. Decarboxylation of Fatty Acids to Terminal Alkenes by Cytochrome P450 Compound I. *J. Am. Chem. Soc.* **2015**, *137*, 4940–4943.

(23) Grant, J. L.; Mitchell, M. E.; Makris, T. M. Catalytic Strategy for Carbon-Carbon Bond Scission by the Cytochrome P450 OleT. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113*, 10049–10054.

(24) Hsieh, C. H.; Huang, X.; Amaya, J. A.; Rutland, C. D.; Keys, C. L.; Groves, J. T.; Austin, R. N.; Makris, T. M. The Enigmatic P450 Decarboxylase OleT Is Capable Of, but Evolved to Frustrate, Oxygen Rebound Chemistry. *Biochemistry* **2017**, *56*, 3347–3357.

(25) Amaya, J. M. Mechanisms of Decarboxylation in the CYP152 Family of Cytochrome P450S. Ph.D. thesis, University of South Carolina, 2018.

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

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

(28) Dubey, K. D.; Shaik, S. Cytochrome P450 – the Wonderful Nanomachine Revealed Through Dynamic Simulations of the Catalytic Cycle. *Acc. Chem. Res.* **2019**, *52*, 389–399.

(29) 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*, 785–789.

(30) Becke, Axel D., Density-Functional Thermochemistry. III. the Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(31) Edler, E.; Stein, M. Spin-State Dependent Properties of an Iron(III) Hydrogenase Mimic. *Eur. J. Inorg. Chem.* **2014**, *2014*, 3587–3599.

(32) Ogliaro, F.; Harris, N.; Cohen, S.; Filatov, M.; De Visser, S. P.; Shaik, S. A Model 'Rebound' Mechanism of Hydroxylation by Cytochrome P450: Stepwise and Effectively

Concerted Pathways, and Their Reactivity Patterns. *J. Am. Chem. Soc.* **2000**, *122*, 8977–8989.

(33) Green, M. T. Evidence for Sulfur-Based Radicals in Thiolate Compound I Intermediates. *J. Am. Chem. Soc.* **1999**, *121*, 7939–7940.

(34) Kamachi, T.; Yoshizawa, K. A Theoretical Study on the Mechanism of Camphor Hydroxylation by Compound I of Cytochrome P450. *J. Am. Chem. Soc.* **2003**, *125*, 4652–4661.

(35) Faponle, A. S.; Quesne, M. G.; de Visser, S. P. Origin of the Regioselective Fatty-Acid Hydroxylation Versus Decarboxylation by a Cytochrome P450 Peroxygenase: What Drives the Reaction to Biofuel Production? *Chem. Eur. J.* **2016**, *22*, 5478–5483.

(36) Bathelt, C. M.; Ridder, L.; Mulholland, A.; Harvey, J. N. Mechanism and Structure-Reactivity Relationships for Aromatic Hydroxylation by Cytochrome P450. *Org. Biomol. Chem.* **2004**, *2*, 2998–3005.

(37) Bathelt, C. M.; Zurek, J.; Mulholland, A. J.; Harvey, J. N. Electronic Structure of Compound I in Human Isoforms of Cytochrome P450 from QM/MM Modeling. *J. Am. Chem. Soc.* **2005**, *127*, 12900–12908.

(38) Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T. B.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X.; et al, Advances in Molecular Quantum Chemistry Contained in the Q-Chem 4 Program Package. *Mol. Phys.* **2015**, *113*, 184–215.

(39) Gilbert, A. IQmol 2.14. `http://iqmol.org/`.

(40) Spartan'18 Parallel Suite. Wavefunction, Inc. Irvine, CA, `http://www.wavefun.com/products/spartan.html`.

(41) Ricciardi, G.; Baerends, E. J.; Rosa, A. Charge Effects on the Reactivity of Oxoiron(IV) Porphyrin Species: A DFT Analysis of Methane Hydroxylation by Polycationic Compound I and Compound II Mimics. *ACS Catalysis* **2016**, *6*, 568–579.

(42) Bauschlicher Jr., C. W. A Comparison of the Accuracy of Different Functionals. *Chem. Phys. Lett.* **1995**, *246*, 40–44.

(43) Halls, M. D.; Schlegel, H. B. Comparison of the Performance of Local, Gradient-Corrected, and Hybrid Density Functional Models in Predicting Infrared Intensities. *J. Chem. Phys.* **1998**, *109*, 10587–10593.

(44) DiLabio, G. A.; Pratt, D. A.; LoFaro, A. D.; Wright, J. S. Theoretical Study of X−H Bond Energetics (X = C, N, O, S): Application to Substituent Effects, Gas Phase Acidities, and Redox Potentials. *J. Phys. Chem. A* **1999**, *103*, 1653–1661.

(45) Altun, A.; Breidung, J.; Neese, F.; Thiel, W. Correlated Ab Initio and Density Functional Studies on H2 Activation by FeO+. *J. Chem. Theory Comput.* **2014**, *10*, 3807–3820.

(46) Chen, H.; Song, J.; Lai, W.; Wu, W.; Shaik, S. Multiple Low-Lying States for Compound I of P450cam and Chloroperoxidase Revealed from Multireference Ab Initio QM/MM Calculations. *J. Chem. Theory Comput.* **2010**, *6*, 940–953.

(47) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.

(48) Mitin, A. V.; Baker, J.; Pulay, P. An Improved 6-31G* Basis Set for First-Row Transition Metals. *J. Chem. Phys.* **2003**, *118*, 7775–7782.

(49) Pritchard, B. P.; Altarawy, D.; Didier, B.; Gibsom, T. D.; Windus, T. L. A New Basis

Set Exchange: An Open, Up-To-Date Resource for the Molecular Sciences Community. *J. Chem. Inf. Model.* **2019**, *59*, 4814–4820.

(50) Filatov, M.; Shaik, S. Theoretical Investigation of Two-State-Reactivity Pathways of H−H Activation by FeO+: Addition-Elimination, "Rebound", and Oxene-Insertion Mechanisms. *J. Phys. Chem. A* **1998**, *102*, 3835–3846.

(51) Amaya, J. A. Mechanisms of Decarboxylation in the CYP152 Family of Cytochrome P450s. Ph.D. thesis, University of South Carolina, ProQuest LLC. 789 East Eisenhower Parkway, P.O. Box 1346 Ann Arbor, MI 48106 - 1346, 2018.

(52) Towns, J.; Cockerill, T.; Dahan, M.; Foster, I.; Gaither, K.; Grimshaw, A.; Hazlewood, V.; Lathrop, S.; Lifka, D.; Peterson, G. D.; et al, XSEDE: Accelerating Scientific Discovery. *Comput. Sci. Eng.* **2014**, *16*, 62–74.