JCTC Journal of Chemical Theory and Computation

QM/MM Simulation on P450 BM3 Enzyme Catalysis Mechanism

Li Tian and Richard A. Friesner*

Department of Chemistry, Columbia University, New York, New York 10027

Received January 21, 2009

Abstract: Using a structure generated by induced fit modeling of the protein−ligand complex, the reaction path for hydrogen atom abstraction in P450 BM3 is studied by means of mixed QM/MM methods to determine the structures and energetics along the reaction path. The IFD structure is suitable for hydrogen atom abstraction at the $\omega-1$ position. The electronic structures obtained are similar to those observed in P450 cam. We show that the barrier for the hydrogen abstraction step from QM/MM modeling is 13.3 kcal/mol in quartet and 15.6 kcal/mol in doublet. Although there is some strain energy present in the ligand, the activation barrier is not dramatically affected. A crystal water molecule, HOH502, plays a role as catalyst and decreases the activation barrier by about 2 kcal/mol and reaction energy by about 3−4 kcal/mol. To achieve reactive chemistry at the remaining experimentally observed positions in the hydrocarbon tail of the ligand, other structures would have to be utilized as a starting point for the reaction. Finally, the present results still leave open the question of whether DFT methods provide an accurate computation of the barrier height in the P450 hydrogen atom abstraction reaction.

1. Introduction

Cytochrome P450 is an essential enzyme for biological functioning in a wide variety of organisms.[1] The primary function of the enzyme is the insertion of an oxygen atom (derived from dioxygen) into a hydrocarbon bond to form an alcohol; other reactions involving oxygen insertion, such as epoxidations, can be catalyzed as well. The insertion of oxygen increases the solubility of lipophilic compounds, and thus makes it easier to excrete them (e.g., through the kidneys in higher organisms). This ability to transform exogenous compounds is of vital importance in mediating the interaction of organisms with their environment.[2−5] Furthermore, interactions with various P450s in the liver often have a critical effect on drug metabolism, drug−drug interactions, and other events of pharmaceutical significance.[6] Consequently, the development of a detailed atomic level understanding of the mechanism of cytochrome P450 functioning is a high priority for computational biochemistry.

The reaction mechanism of P450 has been investigated in a large number of publications,[7−10] primarily focusing on P450 cam, a bacterial enzyme for which extensive structural, spectroscopic, and kinetic data are available.[11,12] Although a great deal of progress has been made in understanding the P450 cam catalytic cycle, there are still puzzling questions concerning the difficulty in trapping the proposed iron oxo intermediate, which is hypothesized to be the species that removes a hydrogen from the substrate and then inserts an oxygen atom.[13,14] Calculated values of the free energy barrier for this reaction are incompatible (too high) with the rapidity of a reaction that would be consistent with the challenges observed in trapping the intermediate. There are various possibilities that can explain this discrepancy, including inaccuracies in the density functional (DFT) electronic structure methods that have invariably been used to model the reaction, and alternative interpretations of the experimental data that do not require a low barrier (including the possibility that some other species is actually the key catalytic form of the enzyme). Investigation of these possibilities is currently ongoing in our laboratory as well as other research groups.

When one moves beyond P450 cam to investigate the reaction mechanisms of other P450s, an additional problem is encountered. In P450 cam, the camphor substrate is perfectly positioned for oxygen insertion relative to the

* Corresponding author e-mail: rich@chem.columbia.edu.

10.1021/ct900040n CCC: $40.75
Published on Web 04/17/2009

putative iron oxo intermediate, with the appropriate hydrogen of the camphor oriented appropriately toward the oxygen of the complex.[11] However, examination of the crystal structures of other P450 species reveals that in many, if not most, cases, the positioning of the ligand in the crystal structure is incompatible with the reported metabolic chemistry.[6,15−19] This observation is actually not terribly surprising. The crystal structures are typically obtained at low temperature, but it is clear that to accommodate a wide range of exogenous species in the active site, induced fit effects, in which the active site residues reorganize, must be quite important.

To investigate this hypothesis, we have carried out several studies of P450 BM3, a bacterial enzyme in which crystallographic studies of ligands have exhibited the discrepancies discussed above. Induced fit modeling[20−22] starting from the PDB structure 1jpz[15] reveals a gating mechanism in which a phenylalanine residue (Phe 87) flips upward, allowing the tail of the fatty acid ligand to move toward the Fe═O moiety.[23] Replica exchange molecular dynamics (REMD) calculations demonstrate that the new conformation (along with a second, closely related conformation generated during the MD simulations) increases in population as the temperature is increased. The results have been confirmed by comparisons with NMR and optical spectroscopic experiments carried out by McDermott and co-workers.[24] These experiments show a temperature-dependent activation of the P450 catalytic cycle, in which movement of Phe87 can be prominently detected. The temperature dependence in the experiments is in good semiquantitative agreement with that predicted from the replica exchange MD simulations.[25,26]

Although these calculations yield structures that can plausibly engage in metabolic chemistry with the catalytic site of P450 BM3, detailed quantum chemical modeling of the P450 BM3 reaction has not yet been performed. In the present paper, we address this question via DFT-based mixed quantum mechanical/molecular mechanics (QM/MM) modeling of the P450 BM3 active site. Starting from the induced fit structure for the complex, we show that the barrier for the rate-determining step of the hydroxylation reaction (hydrogen atom abstraction) is similar in magnitude to that calculated for P450 cam. This result suggests that the problem with calculating the barrier height is consistent for at least two rather different local environments of the catalytic species, suggesting that the explanation of an intrinsic problem with DFT modeling of the reaction may well be at fault. The results also validate the idea of using induced fit modeling, followed by quantum chemical computations, is a useful approach to investigating P450 metabolism in which induced fit effects play a significant role.

The paper is organized as follows. In section II, we discuss the initial setup of the P450 BM3 induced fit system and briefly outline the computational methods used to both generate the initial structures and carry out the QM/MM calculations. Section III presents both QM and QM/MM results for reactant, transition state, and product structures and for various electronic states of the reacting system and discusses the results from the point of view of both induced fit effects and energetics associated with P450 hydroxylation. Finally, in section IV, the conclusion, the results are summarized and future research directions are indicated.

2. Methods

The structure of the BMP domain of P450 BM3 bound with N-Palmitoylglycine, the substrate, has been determined by X-ray diffraction (1.65 Å, PDB Entry: 1jpz).[15] However, the ω end of substrate in this X-ray structure is too far from the reactive core of the porphyrin oxo species (presumed to be compound I) to make it possible to undergo reactive chemistry. Thus, the reaction cannot be performed starting from this "distal state" and another intermediate with the substrate being bound near the active site, the "proximal state", should exist along the path going from the distal state to the products. An induced fit docking (IFD) approach, a method that can accurately account for both ligand and receptor flexibility, is employed initially to find the possible conformations for the reactive chemistry. IFD, presented by Sherman et al.,[22] combines techniques for docking ligands into a rigid receptor with those for modeling receptor conformational changes, in an iterative fashion. The docking program Glide[20,21,27] is employed to account for ligand flexibility and the refinement module in the Prime program[27] is used to account for receptor flexibility. Tests of the IFD protocol on 21 challenging cross-docking cases taken from the PDB demonstrate that this induced fit method can dock the ligand properly as well as capture the key ligand−protein interactions.[22] In the study by Jovanovic et al.,[23] a docking structure was generated by the induced fit method, as shown in Figure 1a. In this induced fit structure, Phe87 rotates upward so that the ω end of the substrate is allowed to approach the Fe═O core, and the resulting structure has properties consistent with a proximal state. The Prime energy of this structure is only 5.8 kcal/mol higher than that of the nativelike prediction; and the ω end of the substrate is very close to heme iron: $\omega-1$, $\omega-2$, and $\omega-3$ carbons are at distances ranging from 3.6 to 5.8 Å from the iron. Thus, this structure, or structures very close to it, may be conjectured as reactive conformations.

Replica exchange molecular dynamics (REMD) simulations were also employed to study this system in order to explore the conformational equilibrium of P450 BM3 system. Replica exchange (or parallel tempering) molecular dynamics[28,29] involves running a series of constant temperature MD simulations at different specified temperatures for a certain number of time steps. Each simulation at a specific temperature is one replica system. Attempt to make exchange of conformations between a pair of replicas is made periodically, and the exchange is accepted according to the Metropolis transition probability. This method allows the system to experience jumps between potential basins separated by high barriers more efficiently. In the REMD study by Ravindranathen et al.,[25] the result has shown that proximal state has significant population at a temperature around 300 K, and that the IFD-generated structure is on the pathway of converting from the distal to free ligand proximal state. The free proximal state found in the REMD simulations is slightly different from the IFD-generated conformation; this will be discussed in more detail below.

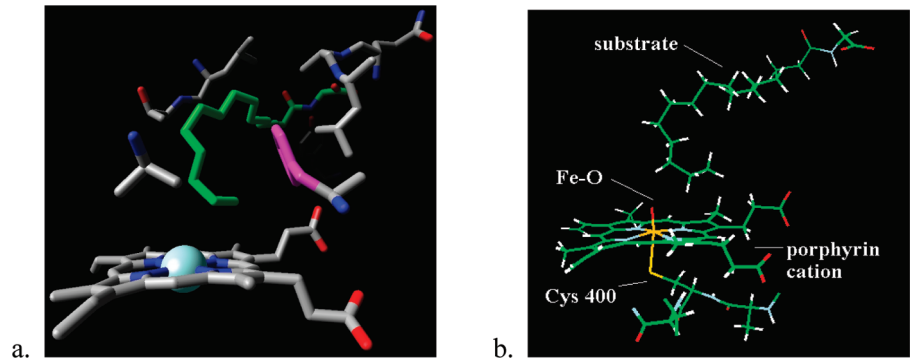

Figure 1. (a) Binding structure from Glide. The rotation of side chain of Phe87 (shown in magenta) allows the substrate, NPG (show in green), to approach the heme iron for reaction. (b) QM region in the QM/MM simulation, including the heme, the substrate NPG, and Cys400.

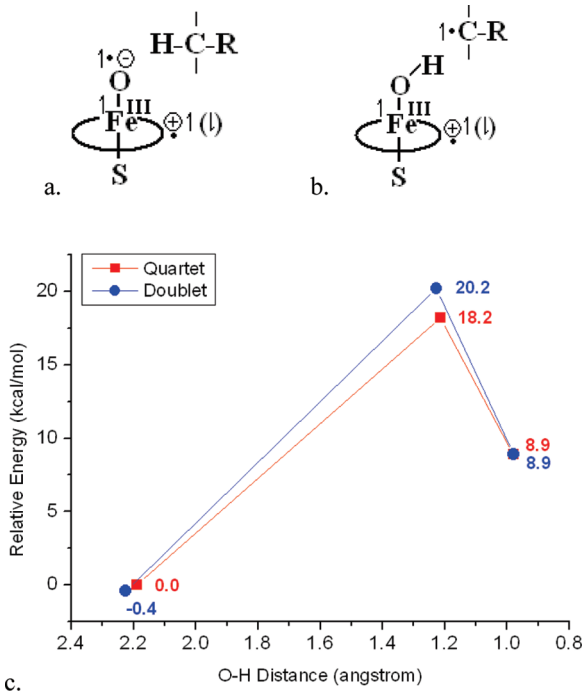

Figure 2. (a) Electronic configurations of reactants in doublet and quartet. (b) Electronic configurations of products in doublet and quartet. (c) Reaction path with relative energy. Energy values are relative to the reactant in quartet. Energy values are in kcal/mol, without ZPE.

On the basis of the reactive induced fit structure, the catalytic mechanism of hydrogen atom abstraction step was investigated in detail in this study. The complicated reaction catalyzed by this transition-metal-containing enzyme involves bond breaking, bond formation, and electron transfer, and there are important effects of the protein/water environment. Thus, a mixed quantum mechanics/molecular mechanics technology is mandatory to capture the detail of the chemistry at the active site and its interaction with the environment. In this study, the structure of our reactant, compound I, was built on the basis of the reactive docking structure from the induced fit method and then modeled with QM/MM as implemented in QSite.[27] Previous QM/MM simulations done in our laboratory, including both enzyme systems and extensive tests on model systems, have demonstrated that the QM/MM protocol in QSite has very small errors associated with the construction of the QM/MM interface.[8,30−36] Using standard protocols for locating reactant, transition state, and product structures and calculating zero point energies, we have also achieved good agreement with experimental free energy barriers (within 2−3 kcal/mol) for a variety of enzymatic reactions via simple transition state theory models.[37] Other effects (longer-range protein motions, quantum tunneling, etc.) are undoubtedly present, but in many if not most cases appear to make relatively minor contributions to experimentally observable energetics. Recent work by Thiel and co-workers has led to similar conclusions.[38]

From the docking structure, the distal oxygen atom was placed above the Fe atom according to the Fe−O distance found in our previous studies of P450 cam. The protein was then soaked in a water shell. As all the crystallographic waters were deleted during Glide docking, we reincorporated the structural waters because some close to the active site may have substantial effect. The water shell was then equilibrated through dynamics simulations using IMPACT.[39] Water molecules beyond 10 Å from the protein were deleted and the oxygen atoms of water molecules 6−10 Å from the protein were frozen throughout the following calculation. The whole system consists of about 22 000 atoms. The structure of the protein was carefully inspected to find titratable residues with unusual protonation states. At the end of this process, a total net charge of −16 e was found. The effect of this net charge on the energetics of the catalysis was tested with a model including counterions but the effect was found to be negligible. This is probably because the charged residues are mainly located at the surface of the protein, near the solvent, and their effect on the active site energetics is screened by solvent (reducing the effective field at long distance by roughly the dielectric constant of water, 80). The reaction was then modeled using QSite. Unrestricted density functional theory was used to model the QM region in conjunction with the B3LYP functional. A polarized basis set (LACVP*) was used for the QM region, as shown in Figure 1b, which includes the heme, the substrate *N*-palmitoylglycine (NPG), and the cysteine (∼160 atoms, ∼1000 basis functions). After geometry optimization was carried out with LACVP*, larger triple-ζ basis sets [cc-pVTZ(-f)] was used for some key atoms in QM region in a single point calculation. Because of the huge size of the QM region (more than 2300 basis functions with cc-pVTZ(-f) basis sets), it was difficult to apply triple-ζ basis sets on the

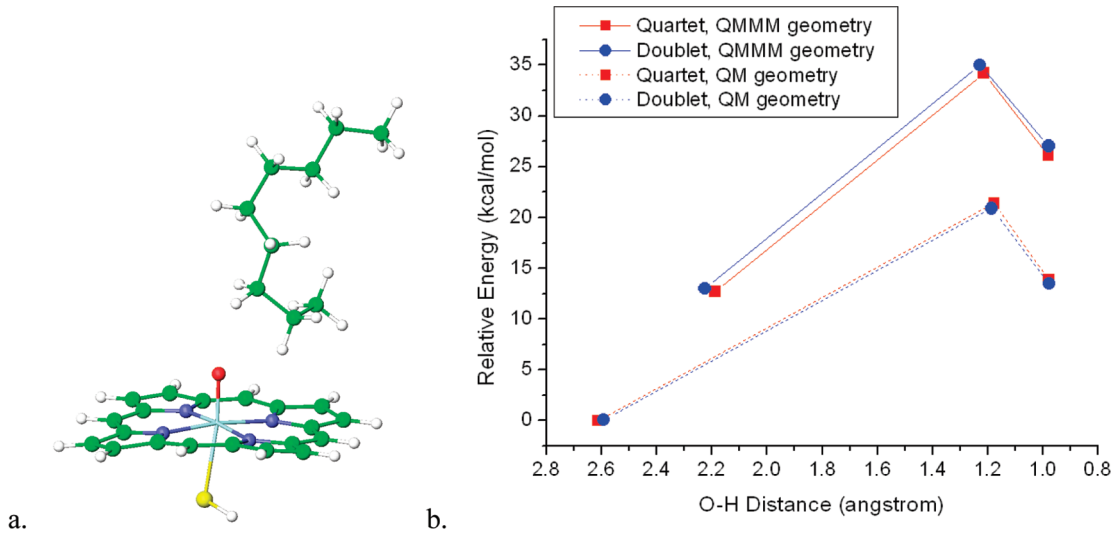

Figure 3. (a) QM model: C_9H_{20} hydrocarbon chain representing the substrate NPG; peripheral substitutes of the heme were all removed and S−H was the model for Cys400. (b) Energy profile.

whole QM region. Thus, we used cc-pVTZ(-f) for the oxygen and four nitrogen atoms bound to Fe as well as the ω terminus of the substrate, used LACV3P* for Fe, and used 6-31G* for others in the test. The energy difference between smaller and bigger basis sets is 0.8 out of 19.2 kcal/mol for activation barrier and 3.3 out of 7.3 kcal/mol for reaction energy. The effect of the larger basis set on the barrier height is minimal, but is somewhat larger for the reaction energy. However, the primary purpose of the present paper is not to quantitatively obtain energetics for the reaction path (particularly given that intrinsic errors in the DFT methods need to be investigated, as will be discussed below), but rather to establish the IFD generated structure as a reasonable model for a ligand pose leading to reactive chemistry. Consequently, to save computer time, we used smaller basis sets, LACVP*, in the following calculations. Finally, the OPLS 1999 force field was used for the MM region, which includes the protein and a shell of water molecules (~22 000 atoms).

Two different QM models were set up, the first to investigate the core active site quantum chemistry in the absence of perturbations, and second to incorporate some effects of protein and solvent environment in a controlled fashion. The most simplified QM model with truncated substrate and heme, as shown in Figure 3a, was used to describe the chemistry of the reaction core. A more extended QM model, the H-capped QM region in QM/MM set up with electrostatic potential from specified point charges, was used to evaluate the effect of protein and water environment. Unrestricted DFT implemented in Jaguar[27] was employed for the two sets of QM model calculations in both doublet and quartet spin states.

3. Results

Mass spectroscopy has shown that the position of hydroxylation is mainly at the $\omega-2$ and $\omega-3$ carbon of the ligand, but also at the $\omega-1$ carbon.[15] Because the $\omega-1$, $\omega-2$, and $\omega-3$ carbons are all reactive positions, the near attack binding structures for each may be different. Therefore, to determine to which reactive position our pose corresponds, we performed a scan along the hydrogen−oxygen distance coordinate for both $\omega-1$ and $\omega-2$ positions. The $\omega-3$ hydrogen was obviously too far for reaction in this binding structure. The $\omega-1$ and $\omega-2$ scan profiles suggest that the docking structure is probably more suitable for hydroxylation at position $\omega-1$, because an activation energy of more than 30 kcal/mol was found in the case of $\omega-2$ hydroxylation. Further sampling of the ligand would be needed to produce structures suitable for investigation of chemistry at the $\omega-2$ and $\omega-3$ positions; we reserve such investigations for another publication.

In the present paper, we report the results of $\omega-1$ hydrogen abstraction and analyze the effect of the protein environment. We first performed QM/MM simulation for reactant and product, and found the transition state, in both quartet and doublet states. The reasonable hydrogen atom abstraction activation barrier obtained in our calculation suggests that the structure obtained from the IFD simulation is realistic. We then discuss QM model calculation results and analyze the effect of the protein and water shell, such as strain energy, electrostatic interaction, and intermediate interaction. Because recent publications[40,41] have reported that a crystal water molecule plays an important role as a catalyst in the P450 cam model of hydrogen atom abstraction, the effect of an analogous water molecule was also analyzed to investigate the effect in the P450 BM3 model.

3.1. QM/MM Result in Quartet and Doublet. Relative energies, optimized geometry parameters, and spin density on critical atoms from QM/MM calculations are shown in Table 1. The two transition states were confirmed by additional minimization starting from points very close to the transition state toward the reactant and the product. Frequency calculation also confirmed the two transition states with negative frequencies (-1826 cm^{-1} in quartet and -1784 cm^{-1} in doublet) corresponding to O−H bond vibration mode. Several other negative frequencies with much smaller energy (mainly several tens to hundreds cm^{-1}), corresponding to small conformational fluctuations of the protein, were also found. This may be due to constraints from substrate binding

Table 1. QM/MM Energy, Optimized Geometry Parameters, and Spin Density on Critical Atoms[a]

Quartet, QM/MM						
	G	*E*	O−H	C−H	O−Fe	Fe−S
reactant	0.0	0.0	2.189	1.092	1.618	2.626
transition state	13.3	18.2	1.214	1.343	1.737	2.505
product	7.3	8.9	0.980	2.073	1.790	2.442

	spin					
	Fe	O	C	H	S	Porp.
reactant	1.15	0.85	0.01	0.00	0.30	0.69
transition state	0.93	0.61	0.53	−0.05	0.24	0.71
product	0.88	0.20	0.94	0.01	0.18	0.78

Doublet, QM/MM						
	G	*E*	O−H	C−H	O−Fe	Fe−S
reactant	−0.4	−0.4	2.225	1.092	1.630	2.548
transition state	15.2	20.2	1.228	1.327	1.748	2.480
product	6.7	8.9	0.979	2.080	1.806	2.425

	spin					
	Fe	O	C	H	S	Porp.
reactant	1.24	0.85	0.01	0.00	−0.26	−0.84
transition state	1.06	0.55	0.47	−0.06	−0.21	−0.84
product	1.00	0.11	0.87	0.01	−0.14	−0.85

[a] *G* is the relative free energy after ZPE correction, and *E* is the relative energy without ZPE; both sets of energy values are relative to quartet reactant and in kcal/mol. Distances are in Å and angles are in degrees. Porp. is the porphyrin π-cation radical.

and/or errors in frequencies calculation. The Gibbs free energy values including zero point energy correction are listed in Table 1, too.

Throughout the reaction, Fe is in its ferric state with net spin around 1.0, and the porphyrin ring is always a radical, as shown in configurations a and b in Figure 2. The sulfur in Cys400, as the fifth ligand of Fe, has localized spin density of about 0.2 and actually shares the third unpaired electron with the porphyrin ring. We will discuss this in detail later. During the reaction, the iron oxo moiety abstracts one electron from C−H bond in the substrate, generating a carbon radical. At the same time, the hydrogen atom is transferred from the carbon to the oxo atom. The difference between doublet and quartet states is mainly manifested in the spin of the porphyrin ring; whether or not the porphyrin ring's spin is parallel to the unpaired spin on Fe and oxo determines the multiplicity of the complex.

The reactants of the two spin states have approximately the same energy, with only a 0.4 kcal/mol difference. The free energy of activation barrier for H abstraction is 13.3 kcal/mol in the quartet, whereas that in the doublet is 15.6 kcal/mol, which is 2.3 kcal/mol higher. The reaction energy is 7.3 kcal/mol for the quartet and 7.1 kcal/mol for the doublet. The breaking of the O−H bond contributes similarly, about 5 kcal/mol, in zero point energy (ZPE) correction in both quartet and doublet cases. The energy values without ZPE are shown in Figure 2c. The comparison between the two spin states suggests that the energetics difference between quartet and doublet is noticeable but small, and the reaction paths should both be feasible in the reaction dynamics. This is consistent with simulation results in the P450 cam model.

3.2. QM Model in Vacuum Result. A simplified QM model, as shown in Figure 3a, was investigated in a vacuum using QM/MM-optimized geometries and vacuum-optimized geometries. Because it has been reported that S−H was a better model for cysteine than methyl mercaptan (CH_3−SH),[42] we built the small QM model including S−H, porphyrin with no substituents, and a 9-carbon hydrocarbon chain. The energy profiles of the quartet and doublet states, QM/MM-optimized geometries, and vacuum-optimized geometries are shown in Figure 3b and listed in Table 2.

We assessed the effect of strain imposed by the protein in the small QM model simulation. From QM/MM-optimized geometries, optimization in a vacuum decreases the energy of the system by about 13 kcal/mol. This indicates that the steric strain and other interactions within the binding cavity impose some constraints on the substrate conformation. We also noticed that several negative vibrational frequencies of QM/MM-optimized geometries are eliminated in the vacuum optimization. These vibrational modes include jiggling of the two ends of the substrate chain and wiggling of S−H. The comparison of conformational change during the reaction in both spin states is shown in Figure 4.

The activation barrier values with and without vacuum optimization in the quartet are 21.4 and 21.5 kcal/mol, respectively, and this result suggests that the optimization does not have a significant effect on the activation barrier in quartet, even though the conformational change through the reaction is much more significant in the optimization. In the case of the doublet, the same conformational change was observed in the optimization; however, the activation barrier was decreased from 22.0 to 20.8 kcal/mol, by removing steric strain in the protein pocket. Although constrained on conformation, optimization in a vacuum does not decrease the activation energy very much, which is almost zero in the quartet (21.5 and 21.4 kcal/mol) and 1.2 kcal/mol in the doublet (22.0 and 20.8 kcal/mol). This suggests that the protein pocket does not facilitate the reaction by conformational constraints or imposed strain. The constraints imposed by the protein do not destabilize or stabilize a particular point on the reaction path, but instead limit the conformation to a smaller phase space without changing the relative energetics.

Comparing the QM model results in a vacuum to the QM/MM results within the protein environment, we observe the main differences in the energy profile and spin distributions. The protein decreases the activation barrier by 3.3 kcal/mol in the quartet but only 1.4 kcal/mol in doublet, and the reaction energy is decreased by 4.5 kcal/mol in the quartet and similarly in the doublet by 4.7 kcal/mol. This indicates that the catalysis mechanism in the quartet and doublet share some common characteristics but also have differences, accounting for about 2 kcal/mol. The main difference in spin distribution is how the third unpaired electron is shared between the porphyrin ring and the sulfur atom in Cys400. Because of the stabilization of the sulfur by protein pocket,

Table 2. QM Model in Vacuum: Relative Energy, Optimized Geometry Parameters, and Spin Density on Critical Atoms[a]

Quartet QM/MM						
	E	O−H	C−H	C−O	C−H−O	relative *E*
reactant	0.0	2.189	1.092	3.020	131.0	12.7
transition state	21.5	1.214	1.343	2.502	156.4	34.2
product	13.4	0.980	2.073	2.960	149.6	26.1

	spin					
	Fe	O	C	H	S	Porp.
reactant	1.09	0.92	0.01	0.00	0.52	0.48
transition state	0.89	0.70	0.53	−0.07	0.39	0.56
product	0.83	0.25	0.95	0.00	0.35	0.63

Doublet QM/MM						
	E	O−H	C−H	C−O	C−H−O	relative *E*
reactant	0.0	2.225	1.092	3.055	131.0	13.0
transition state	22.0	1.228	1.327	2.510	158.3	35.0
product	14.0	0.979	2.080	2.963	149.0	27.0

	spin					
	Fe	O	C	H	S	Porp.
reactant	1.18	0.91	0.01	0.00	−0.53	−0.58
transition state	1.04	0.60	0.44	−0.07	−0.36	−0.68
product	0.99	0.10	0.92	0.01	−0.22	−0.79

Quartet QM Model						
	E	O−H	C−H	C−O	C−H−O	relative *E*
reactant	0.0	2.612	1.097	3.534	141.2	0.0
transition state	21.4	1.177	1.362	2.527	168.8	21.4
product	13.9	0.979	2.334	3.251	155.7	13.9

	spin					
	Fe	O	C	H	S	Porp.
reactant	1.08	0.94	0.00	0.00	0.54	0.45
transition state	0.97	0.67	0.57	−0.06	0.46	0.38
product	0.93	0.25	0.98	0.00	0.47	0.40

Doublet QM Model						
	E	O−H	C−H	C−O	C−H−O	relative *E*
reactant	0.0	2.593	1.097	3.527	142.6	0.1
transition state	20.8	1.185	1.328	2.502	169.3	20.9
product	13.4	0.977	2.297	3.205	154.5	13.5

	spin					
	Fe	O	C	H	S	Porp.
reactant	1.20	0.89	0.00	0.00	−0.58	−0.52
transition state	0.99	0.54	0.46	−0.06	−0.42	−0.54
product	0.98	0.11	0.96	0.01	−0.44	−0.60

[a] *E* is energy relative to corresponding reactant in kcal/mol. Relative *E* is energy relative to quartet reactant with QM model optimized geometry.

the net spin on the sulfur is about 0.15−0.3, which is 0.1−0.2 less compared with the net spin in QM model in a vacuum.

4. Discussion

On the basis of the transition state and the energy profile, the proximal binding structure from the IFD calculation is a good conjecture about the starting point (reactant) for hydroxylation on the $\omega-1$ position of the ligand, and both quartet and doublet are feasible reaction paths. In the QM/MM simulation of a similar system P450 cam, Schöneboom et al.[43] reported energy profiles and spin distribution as listed in Table 3. The result is consistent with what we found in P450 BM3, with similar activation barrier and endothermicity. However, the slight but noticeable difference between quartet and doublet energetics is not reported in P450 cam system.

Besides the distal conformational state observed in X-ray structure and the proximal state we used in this paper to predict catalysis, another proximal state was identified in molecular dynamics calculation by Ravindranathan et al.[25] The structure is shown in Figure 5. The ω end of the substrate is very close to the iron atom. The conformational entropy stabilizes this state, and surrounding residues give the substrate the ability to move more freely. However, on the basis of our scanning of hydrogen abstraction potentials, significant rearrangement is needed before hydroxylation can happen on any of $\omega-1$, $\omega-2$, and $\omega-3$ positions.

To evaluate the effect of various contributions to the electrostatic potential of the protein, we performed single-point calculation with QM/MM geometries. The QM model in this series of calculation was the same as the QM region, with H-capped Cys400. Formal charge values and geometry coordinates of all the water molecules and protein residues are extracted from QM/MM-optimized geometries, whereas only the backbone of Ala399 and Ile401, which are connected to QM region, are removed. A single-point QM calculation is then performed with the fully implemented electrostatic potential as in the protein environment. The corresponding energy profile and spin distribution data are shown in Table 4. It is clear from the comparison that modeling the electrostatic potential by point charges can recover almost all the effect of the protein on the wave functions and energetics of species in the catalytic cycle, from decreasing the activation barrier to tuning the spin distribution between the Cys400 sulfur and the porphyrin ring. This indicates that in the case of P450 BM3 hydrogen atom abstraction, the main effect of protein catalysis is to provide an appropriate electrostatic potential. We will discuss several of the most important components of this significant electrostatic potential and the responsible residues. Key residues around the active site are shown pictorially in Figure 6.

We begin with a discussion of the propionate substituents of the heme and their local protein environment. Positively charged residues Lys69 and Arg398 next to the heme peripheral propionates function as a structural anchor for the porphyrin. The Lys and Arg are also important components of the electrostatic environment, lowering the energy levels of the peripheral carboxylate moieties by stabilizing the negative charge on the two propionate groups. When a carboxylate is left as COO− in vacuum without proper screening from either positively charged residue or an equivalent point charge, it cannot maintain the full −1 charge and becomes partly radical with 0.3−0.5 net spin, because compound I is a strong oxidant. Although this destabilization of peripheral substitutes might not show a significant effect on the activation barrier or endothermicity in a single-point

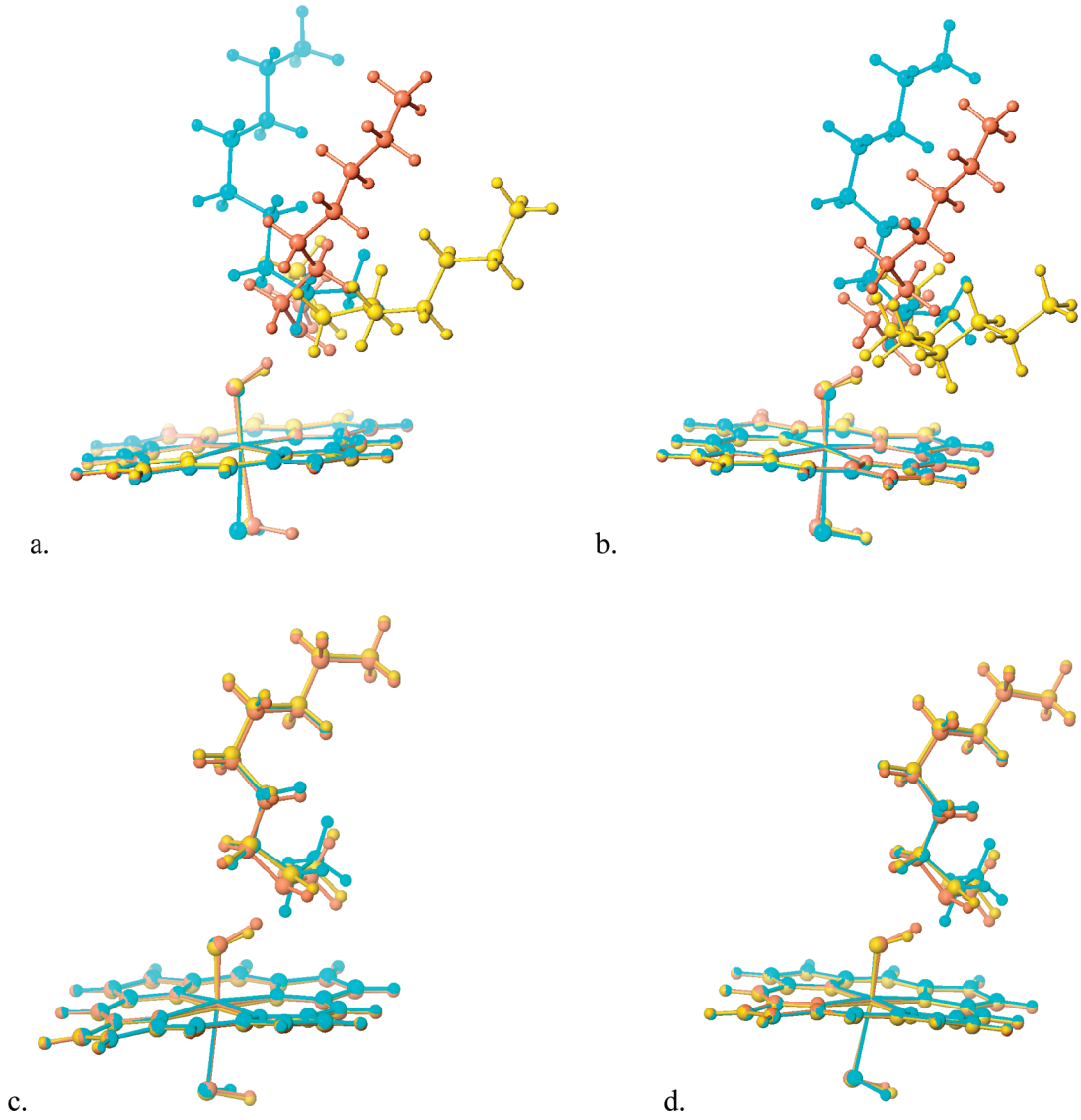

Figure 4. Superposition of reactant, transition state, and product geometries. Reactant is in cyan, transition state is in coral, and product is in goldenrod. Optimized QM model geometries: (a) quartet, (b) doublet. QM/MM geometries: (c) quartet, (d) doublet.

Table 3. Comparisons of Energy Profiles and Spin Distribution in QM/MM Calculation of P450 cam and P450 BM3[a]

		A. E.	R. E.	O−H	spin Fe	spin O	spin C	spin S	spin Porp.
P450 BM3	quartet	18.2	8.9	1.214	0.93	0.61	0.53	0.24	0.71
P450 BM3	doublet	20.6	9.3	1.228	1.06	0.55	0.47	−0.21	−0.84
P450 cam[b]	quartet	21.8	10.5	1.188	0.95	0.61	0.55	0.21	0.72
P450 cam[c]	doublet	21.1	9.4	1.195	1.16	0.55	0.50	−0.26	−0.89
P450 cam[c]	quartet	18.7	6.5	1.225	0.98	0.61	N/A	N/A	0.77

[a] A. E. is the activation barrier in kcal/mol and R. E. is the reaction energy in kcal/mol. Both are before ZPE corrections. Distances and spin density values are of transition state. [b] R2/B2, P450 cam result based on snapshot 40, which did not consider the catalysis effect of crystal water w903; heme peripheral substitute are replaced by H atoms, in QM region.[43] [c] R2/B2, P450 cam result based on snapX incorporating the effect of water903, without heme peripheral substitute in QM region, protonated Asp297.[41]

calculation, it is not natural and not indicated by the experimental protein structure of this model. Thus it is preferable, as a general guideline in QM/MM, to prepare the protein such that the QM part stays globally neutral as to avoid unphysical electron transfer.

However, although mainly because of electrostatic interaction, the H-bond component in the interaction between positive residues and propionates is not negligible. The QM/MM energy difference between two orientations of the amine group in Lys69 is about 8 kcal/mol. The orbital shown in Figure 7a indicates a π orbital of the heme propionate, which forms the π cation H-bond with positively charge amino group. In the lower-energy geometry, the π orbital is almost perpendicular to C═O plane and parallel with N−H σ bond in the amino group of Lys69, whereas in the higher-energy conformation, the angle is about 40°.

There is one crystal water molecule, HOH502, located around the reaction core, with position shown in Figure 6.

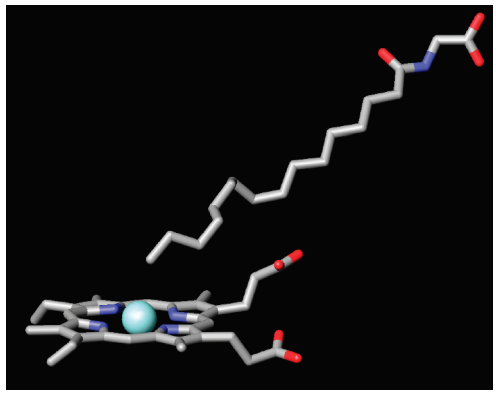

Figure 5. Active site of another proximal state binding conformation predicted by Ravindranathen et al.[25]

Table 4. QM Model with Point Charges: Energy Profiles and Spin Distribution. This Data Can Be Compared with QM/MM Results Listed in Table *1*[a]

Quartet, QM						
	E	spin Fe	spin O	spin C	spin S	spin Porp.
reactant	0.0	1.15	0.85	0.01	0.34	0.65
transition state	19.4	0.94	0.61	0.54	0.29	0.65
product	9.8	0.88	0.20	0.94	0.20	0.77

Doublet, QM						
	E	spin Fe	spin O	spin C	spin S	spin Porp.
reactant	−0.2	1.26	0.85	0.01	−0.31	−0.80
transition state	19.8	1.07	0.55	0.47	−0.25	−0.80
product	9.3	1.02	0.10	0.85	−0.17	−0.80

[a] *E* is the energy relative to quartet reactant, in kcal/mol.

To evaluate the effect of this water molecule, we performed single-point calculation of both QM/MM and QM models in the protein electrostatic environment without HOH502. In the QM/MM calculation, the HOH502 was moved from active site to 60 Å away, which ensured that this water does not impact the active site energetics in a meaningful way. In the QM model calculation, point charges from all of the atoms in the protein and water shell were applied as the electrostatic environment, except for the three points from HOH502. The comparison of the results is shown in Table 5. From these results, it is clear that HOH502 decreases the activation barrier by about 2−3 kcal/mol and the reaction energy by about 3−4 kcal/mol, in both quartet and doublet states, with slightly different quantitative effects depending on calculational settings (this level of uncertainty is, however, negligible compared to other possible errors in the full QM/MM calculations).

The H abstraction reaction, by its nature, increases the number of electrons in heme, the oxidant. The oxo part starts as a negative radical, abstracts one electron and forms an O−H bond; along with the increasing of electrons, the localized spin of the oxo group is decreased and the partial charge on the oxo site becomes more negative. The water HOH502, which is very close to the active site, stabilizes the whole system through its hydrogen bond with oxo, by about 2 kcal/mol in the QM/MM calculations. Moreover, this interaction is stronger with the transition state and product as compared to the reactant. This is because oxo gains more electron density, and becomes more negative, during the course of the reaction, which is favored by this hydrogen-bonded structure. It is clear from data in Table 5 that the electrostatic interaction in the H-bond with HOH502 decreases the spin on oxo by 0.02−0.07 and draws in 0.04−0.08 more negative charge, which is in line with the trend of the reaction. Thus this crystal water functions as a catalyst by stabilizing the transition state and product. From our result in QM model, it is clear that this catalysis effect is mainly an electrostatic interaction: substituting point charges for water molecule can recover the effect.

In the X-ray structure 1JPZ, the B factor of the oxygen atom in HOH502 is 19.98, which is comparable to most atoms in porphyrin ring and significantly smaller than most atoms in the substrate. This indicates that HOH502 is a well-ordered water molecule and its position, as well as interaction with it, should be stable. This observation is also consistent with recent result of calculations on P450 cam. Altun et al.[40,41] reported that in P450 cam, a similar water molecule w903 decreases barrier and endothermicity by about 4 and 6 kcal/mol, respectively, in QM/MM, and 3.0 and 5.2 kcal/mol in QM model calculation. Results of P450 cam system in quartet with similar QM region settings are listed in Table 3. This water provides the most significant catalytic factor of the protein. In the doublet, the electrostatic effect of HOH502 can solely recover the activation barrier difference between the full QM/MM calculation and QM model in a vacuum. In the quartet, besides the catalysis of HOH502, there is still 1.1 kcal/mol stabilization from protein in addition.

In the QM/MM simulation, different electronic configurations with very close energy values under the same geometries were found, indicating that this system has an extremely complicated electronic structure. The main difference between these configurations is how the third single electron is shared between the porphyrin ring and the sulfur atom in Cys400. Our QM/MM results show that lower net spin on sulfur does not always lead to lower energy, and net spin on sulfur in the lowest energy configuration is about 0.2 to 0.3, decreasing along the reaction coordinate. This result is consistent with reported P450 cam values.[43]The spin delocalization, in this case as the electron transfers from the thiolate ligand to the conjugated porphyrin ring, has some effects on the energy profiles, although these are limited from a quantitative point of view. The net spin on the sulfur atom can be as high as 0.5 in the QM model calculation and it can also be as low as about zero in some electronic configurations in the QM/MM calculation, which both lead to higher activation barrier. We estimate that the effect of tuning the net spin on the sulfur atom is less than 1 kcal/mol, on both the activation barrier and the reaction energy.

It is interesting to find out factors stabilizing the sulfur, tuning the spin distribution between the sulfur and the porphyrin, and finally lowering the activation barrier. As we have discussed before, electrostatic potential is the key effect from protein, and thus the question will be which residues or electrostatic factors are responsible. In P450 cam system, the stabilization effect for sulfur is reported to be the

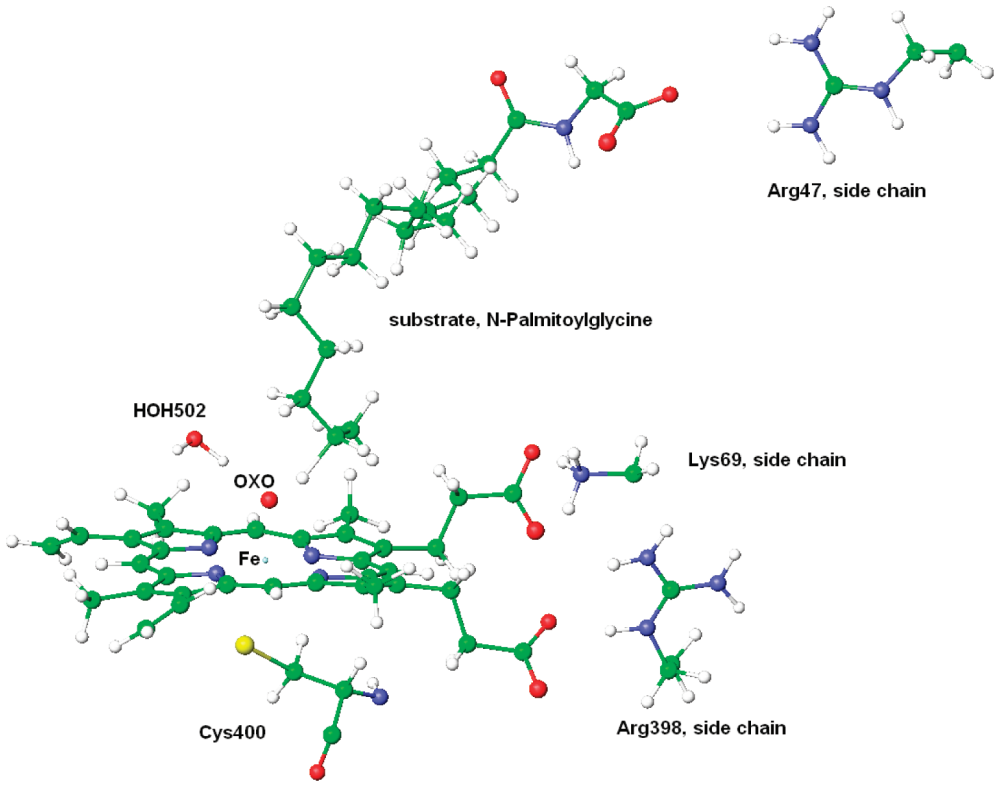

Figure 6. Key residues around the active site of P450 BM3.

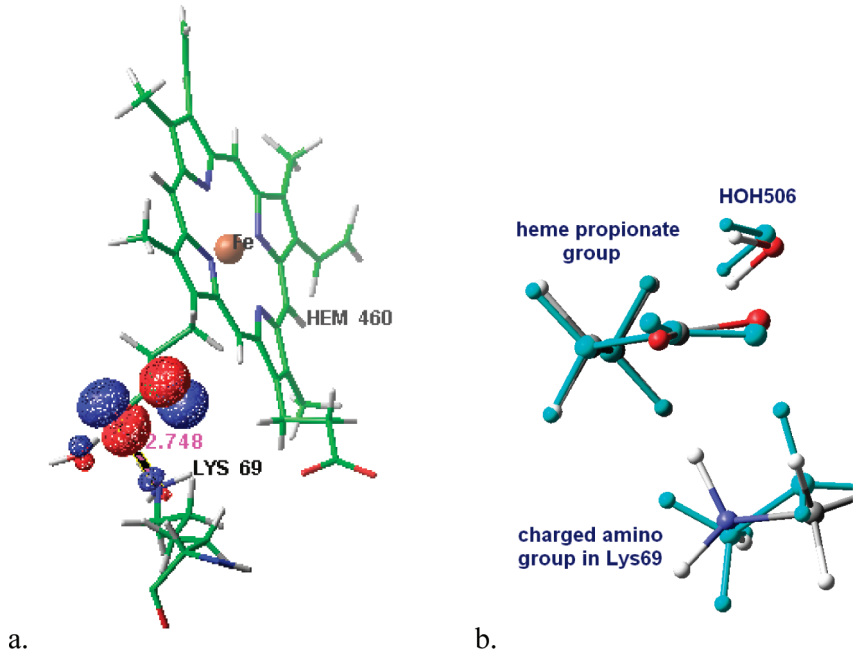

Figure 7. (a) π Orbital of propionate. (b) Superposition of two side chain conformations of Lys69. Higher-energy geometry is in cyan.

H-bonding with neighboring residues Gly359 and Gln360 and the polarization of protein environment, e.g., higher dielectric constant ($\varepsilon = 5.7$).[44,45] In P450 BM3, there are two residues at very similar position showing H-bonding with the sulfur: Gly402 and Gln403. To study whether the same factor plays a role in P450 BM3 system, we investigated the extended QM model as shown in Figure 1b with applied electrostatic potential and dielectric constant. However, neither applying the point charges of the backbone N−H groups of Gly402 and Gln403 nor implementing higher dielectric constant could recover both the spin distribution and the energetics. Moreover, non-negligible energy fluctuations were found in calculations with different model settings. These results indicate that the problem may be more complicated than expected. The propionates connected to conjugated porphyrin ring, as well as the interaction between sulfur atom and backbone of G402Q403 and other surrounding residues may all play a role in this stabilization. The

Table 5. Catalysis Effect of HOH502[a]

Quartet, QM/MM				
	E	spin O	charge O	O−O
reactant	0.0/0.0	0.85/0.87	−0.49/−0.45	2.719
transition state	18.2/20.4	0.61/0.63	−0.67/−0.62	2.667
product	8.9/12.3	0.20/0.22	−0.78/−0.74	2.637

Doublet, QM/MM				
	E	spin O	charge O	O−O
reactant	0.0/0.0	0.85/0.87	−0.49/−0.45	2.715
transition state	20.6/22.6	0.55/0.62	−0.67/−0.59	2.660
product	9.3/13.0	0.11/0.14	−0.79/−0.71	2.632

Quartet, QM			
	E	spin O	charge O
reactant	0.0/0.0	0.85/0.91	−0.46/−0.40
transition state	19.4/22.3	0.61/0.68	−0.65/−0.57
product	9.8/13.9	0.20/0.24	−0.77/−0.69

Doublet, QM			
	E	spin O	charge O
reactant	0.0/0.0	0.85/0.90	−0.47/−0.40
transition state	20.0/23.1	0.55/0.60	−0.64/−0.56
product	9.5/14.0	0.10/0.12	−0.77/−0.69

[a] Values before slash are results with HOH502 and values after slash are results without HOH502. O−O is the distance between oxo and O atom in HOH502, in Å. *E* is energy in kcal/mol relative to reactant of each set. Comparison indicates that HOH502 decreases the activation barrier and endothermicity by about 2−3 and 3−4 kcal/mol, respectively, depending on the different systems.

detail of stabilization of sulfur needs more investigation and will be explored in future research.

Finally, we briefly discuss a recent publication that has studied hydroxylation of a number of alkanes by P450-BM3 using docking, molecular dynamics, and quantum chemical techniques.[46] Although the systems studied in ref 46 and methods used have some similarities to those of the present paper, there are also major differences, e.g., the use of gas phase semiempirical quantum chemistry to compute barrier heights, details of the sampling approaches to generate candidate reactive structures, study of different ligands, and investigation of mutations of the protein. These differences, particularly the very different quantum chemical methods used to estimate barrier heights, make any direct comparisons of the results extremely difficult. As more calculations on flexible P450-based systems are performed, and compared quantitatively to experiment (in ref 46 the comparisons were only qualitative, whereas our own such comparisons presented several difficulties as discussed in the text), it will be possible to achieve a better picture of how well various approaches work in rationalizing P450 chemistry.

5. Conclusion

We have calculated the reaction path for hydrogen atom abstraction in P450 BM3 using a structure generated by induced fit modeling of the protein−ligand complex, and applying mixed QM/MM methods to determine the structures and energetics along the reaction path. The electronic structures obtained are similar to those observed in P450 cam. The IFD structure is suitable for hydrogen atom abstraction at the $\omega-1$ position; although there is some strain energy present in the ligand, it is distributed more or less uniformly in the reactant, transition state, and product, so that the activation barrier is not dramatically affected.

To achieve reactive chemistry at the remaining experimentally observed positions in the hydrocarbon tail of the ligand, other structures would have to be utilized as a starting point for the reaction. The next step in understanding the full range of chemistry in P450 BM3 is to carry out a combination of conformational search and molecular dynamics simulations to produce a suitable ensemble of conformations, and establish that all of the experimentally observed chemistry can be explained via simulations. There are also quantitative questions with regard to the activation free energy required to produce the initial conformations. We have not investigated this issue in the present paper, but a fully quantitative analysis of the thermodynamics of the reactive process would require an assessment of this contribution.

Finally, the present results still leave open the question of whether DFT methods provide an accurate computation of the barrier height in the P450 hydrogen atom abstraction reaction. The fact that the barrier computed here is very similar to that in P450 cam suggests that the explanation for the experimental observation of a very fast reaction is not due to effects of the protein that are somehow being missed, or miscalculated, in the model. Rather, it may be the case that the B3LYP estimation of the barrier height is problematic. The performance of DFT methods for transition-metal-containing systems, and for transition states generally, is far from uniform; systematic errors can be detected in a number of different systems and applications. We have recently been investigating these errors via a novel scheme for introducing empirical localized orbital corrections.[47−49] Application of such methods to the P450 hydrogen atom abstraction is at present not straightforward, but is an area we expect to investigate in future work.

Acknowledgment. This work was supported in part by a grant to R.A.F. from the NIH (GM40526). The authors thank Dr. David Rinaldo for helpful discussions.

References

(1) Ortiz de Montellano, P. R.; De Voss J. J. Substrate Oxidation by Cytochrome P450 Enzymes. In *Cytochrome P450: Structure, Mechanism, and Biochemistry*, 3rd ed.; Ortiz de Montellano, P. R., Ed.; Springer: New York, 2004; pp 183−245.

(2) Capdevila, J. H.; Falck, J. R.; Estabrook, R. W. *Faseb J.* **1992**, *6*, 731–6.

(3) Guengerich, F. P. *J. Biol. Chem.* **1991**, *266*, 10019–22.

(4) Kubota, M.; Sogawa, K.; Kaizu, Y.; Sawaya, T.; Watanabe, J.; Kawajiri, K.; Gotoh, O.; Fujii-Kuriyama, Y. *J. Biochem. (Tokyo)* **1991**, *110*, 232–6.

(5) Porter, T. D.; Coon, M. J. *J. Biol. Chem.* **1991**, *266*, 13469–72.

(6) Williams, P. A.; Cosme, J.; Vinkovic, D. M.; Ward, A.; Angove, H. C.; Day, P. J.; Vonrhein, C.; Tickle, I. J.; Jhoti, H. *Science* **2004**, *305*, 683–686.

(7) Guallar, V.; Baik, M.-H.; Lippard, S. J.; Friesner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 6998–7002.

(8) Guallar, V.; Friesner, R. A. *J. Am. Chem. Soc.* **2004**, *126*, 8501–8508.

(9) Kamachi, T.; Yoshizawa, K. *J. Am. Chem. Soc.* **2003**, *125*, 4652–4661.

(10) Shaik, S.; Kumar, D.; de Visser, S. P.; Altun, A.; Thiel, W. *Chem. Rev.* **2005**, *105*, 2279–2328.

(11) Schlichting, I.; Berendzen, J.; Chu, K.; Stock, A. M.; Maves, S. A.; Benson, D. E.; Sweet, R. M.; Ringe, D.; Petsko, G. A.; Sligar, S. G. *Science* **2000**, *287*, 1615–1622.

(12) Poulos, T. L.; Finzel, B. C.; Howard, A. J. *J. Mol. Biol.* **1987**, *195*, 687–700.

(13) Davydov, R.; Makris, T. M.; Kofman, V.; Werst, D. E.; Sligar, S. G.; Hoffman, B. M. *J. Am. Chem. Soc.* **2001**, *123*, 1403–1415.

(14) Groenhof, A. R.; Ehlers, A. W.; Lammertsma, K. *J. Am. Chem. Soc.* **2007**, *129*, 6204–6209.

(15) Haines, D. C.; Tomchick, D. R.; Machius, M.; Peterson, J. A. *Biochemistry* **2001**, *40*, 13456–13465.

(16) Schoch, G. A.; Yano, J. K.; Wester, M. R.; Griffin, K. J.; Stout, C. D.; Johnson, E. F. *J. Biol. Chem.* **2004**, *279*, 9497–9503.

(17) Williams, P. A.; Cosme, J.; Ward, A.; Angove, H. C.; Vinkovic, D. M.; Jhoti, H. *Nature (London)* **2003**, *424*, 464–468.

(18) Guengerich, F. P. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 13565–13566.

(19) Hritz, J.; de Ruiter, A.; Oostenbrink, C. *J. Med. Chem.* **2008**, *51*, 7469–7477.

(20) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. *J. Med. Chem.* **2004**, *47*, 1739–1749.

(21) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. *J. Med. Chem.* **2004**, *47*, 1750–1759.

(22) Sherman, W.; Day, T.; Jacobson, M. P.; Friesner, R. A.; Farid, R. *J. Med. Chem.* **2006**, *49*, 534–553.

(23) Jovanovic, T.; Farid, R.; Friesner, R. A.; McDermott, A. E. *J. Am. Chem. Soc.* **2005**, *127*, 13548–13552.

(24) Jovanovic, T.; McDermott, A. E. *J. Am. Chem. Soc.* **2005**, *127*, 13816–13821.

(25) Ravindranathan, K. P.; Gallicchio, E.; Friesner, R. A.; McDermott, A. E.; Levy, R. M. *J. Am. Chem. Soc.* **2006**, *128*, 5786–5791.

(26) Ravindranathan, K. P.; Gallicchio, E.; McDermott, A. E.; Levy, R. M. *J. Am. Chem. Soc.* **2007**, *129*, 474–475.

(27) *Glide, Prime, QSite, Jaguar*; Schrodinger, LLC: Portland, OR.

(28) Felts, A. K.; Harano, Y.; Gallicchio, E.; Levy, R. M. *Proteins: Struct., Funct., Bioinf.* **2004**, *56*, 310–321.

(29) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141–151.

(30) Friesner, R. A. *Drug Discovery Today: Technol.* **2004**, *1*, 253–260.

(31) Friesner, R. A.; Baik, M.-H.; Gherman, B. F.; Guallar, V.; Wirstam, M.; Murphy, R. B.; Lippard, S. J. *Coord. Chem. Rev.* **2003**, *238*−*239*, 267–290.

(32) Gherman, B. F.; Goldberg, S. D.; Cornish, V. W.; Friesner, R. A. *J. Am. Chem. Soc.* **2004**, *126*, 7652–7664.

(33) Gherman, B. F.; Lippard, S. J.; Friesner, R. A. *J. Am. Chem. Soc.* **2005**, *127*, 1025–1037.

(34) Guallar, V.; Jacobson, M.; McDermott, A.; Friesner, R. A. *J. Mol. Biol.* **2004**, *337*, 227–239.

(35) Murphy, R. B.; Philipp, D. M.; Friesner, R. A. *Chem. Phys. Lett.* **2000**, *321*, 113–120.

(36) Philipp, D. M.; Friesner, R. A. *J. Comput. Chem.* **1999**, *20*, 1468–1494.

(37) Friesner, R. A.; Guallar, V. *Annu. Rev. Phys. Chem.* **2005**, *56*, 389–427.

(38) Thiel, W. Presented at the *Theoretical Biochemistry: Methods and Applications Symposium*, Stockholm, May 14−17, 2008.

(39) Banks, J. L.; Beard, H. S.; Cao, Y.; Cho, A. E.; Damm, W.; Farid, R.; Felts, A. K.; Halgren, T. A.; Mainz, D. T.; Maple, J. R.; Murphy, R.; Philipp, D. M.; Repasky, M. P.; Zhang, L. Y.; Berne, B. J.; Friesner, R. A.; Gallicchio, E.; Levy, R. M. *J. Comput. Chem.* **2005**, *26*, 1752–1780.

(40) Altun, A.; Guallar, V.; Friesner, R. A.; Shaik, S.; Thiel, W. *J. Am. Chem. Soc.* **2006**, *128*, 3924–3925.

(41) Altun, A.; Shaik, S.; Thiel, W. *J. Comput. Chem.* **2006**, *27*, 1324–1337.

(42) Ogliaro, F.; Cohen, S.; Filatov, M.; Harris, N.; Shaik, S. *Angew. Chem., Int. Ed.* **2000**, *39*, 3851–3855.

(43) Schoneboom Jan, C.; Cohen, S.; Lin, H.; Shaik, S.; Thiel, W. *J. Am. Chem. Soc.* **2004**, *126*, 4017–34.

(44) Ogliaro, F.; Cohen, S.; de Visser, S. P.; Shaik, S. *J. Am. Chem. Soc.* **2000**, *122*, 12892–12893.

(45) de Visser, S. P.; Ogliaro, F.; Sharma, P. K.; Shaik, S. *J. Am. Chem. Soc.* **2002**, *124*, 11809–11826.

(46) Feenstra, K. A.; Starikov, E. B.; Urlacher, V. B.; Commandeur, J. N. M.; Vermeulen, N. P. E. *Protein Sci.* **2007**, *16*, 420–431.

(47) Friesner, R. A.; Knoll, E. H.; Cao, Y. *J. Chem. Phys.* **2006**, *125*, 124107/1–124107/24.

(48) Knoll, E. H.; Friesner, R. A. *J. Phys. Chem. B* **2006**, *110*, 18787–18802.

(49) Rinaldo, D.; Tian, L.; Harvey, J. N.; Friesner, R. A. *J. Chem. Phys.* **2008**, *129*, 164108/1–164108/23.

CT900040N