JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

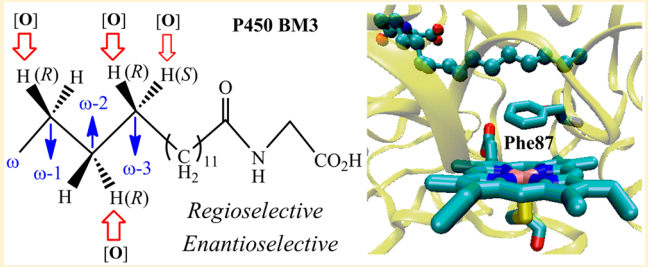
Molecular Dynamics and QM/MM Calculations Predict the Substrate-Induced Gating of Cytochrome P450 BM3 and the Regio- and Stereoselectivity of Fatty Acid Hydroxylation

Kshatresh Dutta Dubey,[†] Binju Wang,[†] and Sason Shaik*

Institute of Chemistry and The Lise Meitner-Minerva Center for Computational Quantum Chemistry, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel

Supporting Information

ABSTRACT: Theory predicts herein enzymatic activity from scratch. We show that molecular dynamics (MD) simulations and quantum-mechanical/molecular mechanics (QM/MM) calculations of the fatty acid hydroxylase P450 BM3 predict the binding mechanism of the fatty acid substrate and its enantio/regioselective hydroxylation by the active species of the enzyme, Compound I. The MD simulations show that the substrate's entrance involves hydrogen-bonding interactions with Pro25, Glu43, and Leu188, which induce a huge conformational rearrangement that closes the substrate channel by pulling together the A helix and the β_1 sheet to the F/G loop. In turn, at the bottom of the substrate's channel, residue Phe87 controls the regioselectivity by causing the substrate's chain to curl up and juxtapose its CH_2 positions ω-1/ω-2/ω-3 to Compound I while preventing access to the endmost position, ω-CH_3. Phe87 also controls the stereoselectivity by the enantioselective steric blocking of the pro-*S* C−H bond, thus preferring *R* hydroxylation. Indeed, the MD simulations of the mutant Phe87Ala predict predominant ω hydroxylation. These findings, which go well beyond the X-ray structural data, demonstrate the predictive power of theory and its insight, which can potentially be used as a partner of experiment for eventual engineering of P450 BM3 with site-selective C−H functionalization capabilities.



1. INTRODUCTION

The ability of theory to predict various aspects of chemical reactivity in partnership with experiment is constantly improving. In this article, we show that molecular dynamics (MD) simulations[1,2] in conjunction with atomistic quantum-mechanical/molecular mechanics (QM/MM) calculations[3−5] can predict, virtually from scratch, the substrate entry into the active pocket of a cytochrome P450 metalloenzyme;[6−11] specifically, the manner by which the substrate binds to and gates the enzyme as well as the substrate's dynamics of sliding toward the active oxidant moiety and its orientation by the key residues of the protein to eventually yield regio- and enantioselective products are all predicted by MD and QM/MM calculations.

Our choice of P450 stems from the importance of this superfamily of enzymes that utilizes molecular oxygen to oxygenate a great variety of organic molecules as means of detoxification, drug metabolism, and biosynthesis of vital compounds such as steroids, essential fatty acids, and brain chemicals.[6−11] Many of these enzymes employ specific substrate-binding machineries, and their hydroxylations and oxo-transfer reactions tend to be regio- and enantioselective.[5−11] Here we focus on P450 BM3 and the fatty acid derivative *N*-palmitoylglycine (NPG) substrate, which behaves much like the natural parent substrates. Moreover, NPG undergoes hydroxylation with the same regio- and stereoselectivity as the native fatty acid,[12] binds to P450 BM3 with higher affinity than any other known substrate, and was used to generate a high-resolution X-ray structure with its crystal waters.[12]

To facilitate the following discussions, we show in Figure 1a the enzyme with its embedded active species, the high-valent iron−oxo porphyrin radical cation species, so-called Compound I (Cpd I),[6−9,13−16] and the substrate. Also shown are key residues that play essential roles to be discussed later. In Figure 1b we show the substrate and its ω-1/ω-2/ω-3 methylene positions, which undergo C−H hydroxylation by the enzyme, as well as the pro-*R* and pro-*S* C−H bonds that are hydroxylated by the enzyme with *R* enantioselectivity.

The X-ray structural data[17−22] reveal that the substrate is bound at the carboxylate head by two H-bond-donor residues, Arg47 and Tyr51 (Figure 1a), near the surface of the substrate entrance channel, while the tail is constrained by the hydrophobic bulky residue Phe87 at the bottom of the channel, above the porphyrin moiety. Unusual features were revealed by the crystal structures of the heme domain in the substrate-free state[19,22] and the substrate-bound[20,21] resting state. As reviewed by Li and Poulos,[17,20] the substrate-free structure was found to have an open channel that stretches from the protein surface

Received: August 18, 2015
Published: December 30, 2015

DOI: 10.1021/jacs.5b08737
J. Am. Chem. Soc. 2016, 138, 837−845

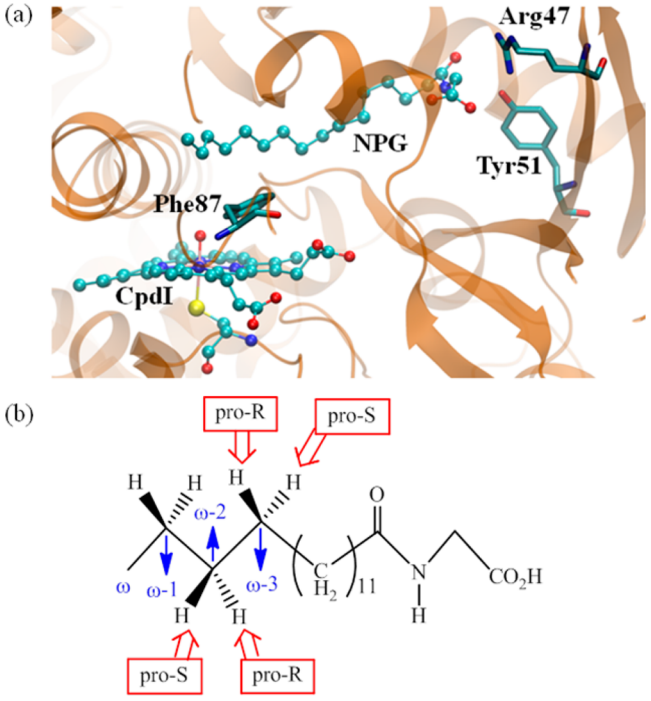


Figure 1. (a) Wild-type P450 BM3 with the active species Compound I (Cpd I), the *N*-palmitoylglycine (NPG) substrate, and some key protein residues. (b) Schematic drawing of NPG with labels of the methylene groups using the ω system, where ω is the endmost CH_3 group of the chain, and the pro-*R* and pro-*S* C−H bonds of ω-3.

deep to the heme iron. The residues Arg47, Tyr51, and Phe87, which constitute the substrate-binding machinery, line up this channel, and hence, the open channel could be identified as the substrate entrance channel. Upon substrate binding, the channel undergoes large conformational changes to a closed state.[17,20,23−25]

In all of the crystal structures, the ω-1/ω-2/ω-3 methylene positions of the fatty acid chain (Figure 1b) are far away (7−8 Å) from the heme iron. Nevertheless, Cpd I hydroxylates the chain regioelectively at these methylene positions, but *not* at the ω-CH_3 position.[26−29] As pointed out by Li and Poulos,[17] the reduction of Fe^{III} to Fe^{II} causes the substrate to slide several angstroms closer to the iron. It was postulated that because of this translation of the chain, the bulky residue Phe87 pushes the fatty acid chain and causes it to curl, such that its ω position moves to a hydrophobic patch away from the heme iron while the ω-1/ω-2/ω-3 positions get exposed to the oxo ligand of Cpd I. In so doing, the substrate-binding machinery appears to dictate the regioselectivity of hydroxylation.[26] Moreover, the hydroxylation of these methylene groups is performed enantioselectively, yielding an *R* alcohol (see Figure 1b) as the major product. Single-site mutation of Phe87 changes the regioselectivity of oxidation, e.g., the F87A mutation leads to hydroxylation at the ω position.[27] Thus, substrate binding and its constraint by Phe87 may determine the observed regioselectivity and enantioselectivity of P450 BM3 oxidation by juxtaposing the ω-1, ω-2, and ω-3 carbon atoms vis-à-vis the iron−oxo moiety of Cpd I and by preferring the pro-*R* C−H bond over the pro-*S* one.

Can theory predict all of these physical and chemical features? As we show herein, MD simulation for up to 360 ns is sufficient to predict all of these features in both wild-type (WT) P450 BM3 and its F87A mutant. Furthermore, by the addition of QM/MM calculations on MD snapshots it is also possible to predict the hydroxylation regio- and enantioselectivity, their root causes, and control by the key residues. The complete match to experiment underscores the role of theory as a partner to experiment and the future role of MD simulations as a reliable predictor of the chemistry of P450 enzymes.

2. COMPUTATIONAL METHODS AND DETAILS

2.1. System Preparation and Setup. The initial structures of P450 BM3 were taken from the Protein Data Bank for substrate-bound (PDB entry 1JPZ)[12] and substrate-free (PDB entry 1BVY)[18] enzymes. Docking was performed by adding the substrate-bound coordinates into the substrate-free enzymes using Chimera.[30] Missing hydrogen atoms were added using the leap module of Amber 14.[31] The force field for the heme moiety was taken from a recent study.[32] Studies of the catalytic cycle revealed that iron prefers the high-spin state before the first reduction.[9] Accordingly, the heme parameters for the high-spin state were used during the resting-state simulations, while the hexacoordinate parameters were used for Cpd I. The partial atomic charges and missing parameters for NPG were obtained from the RESP method[33,34] at the HF/6-31G* level of theory.

A few Na^+ ions were added into the protein surface to neutralize the total charge of the system. Finally, the resulting system was solvated in a rectangular box of TIP3P[35] waters extending up to minimum cutoff of 10 Å from the protein boundary. The catalytic cycle of P450 has many states. Herein we focus on the two key states of interest: the resting state, where it is believed that the catalytic cycle is gated by the entrance of the substrate, and the Cpd I state, where the oxidation reactions actually happen. The Amber ff14SB force field was employed for the protein in all six of the following MD simulations: (1) resting state without substrate (open-state heme domain); (2) resting state with the substrate bound (closed-state heme domain); (3) resting state with the substrate docked in the open-state heme domain; (4) Cpd I state with the substrate bound (closed-state heme domain); (5) Cpd I state with the substrate docked in the open-state heme domain; and (6) Cpd I state with the substrate bound for the F87A mutant.

2.2. MD Simulations. After proper parametrizations and setup, the resulting systems' geometries were minimized (5000 steps of steepest descent and 10 000 steps of conjugate gradient) to remove the poor contacts and relax the systems. The systems were then gently annealed from 10 to 300 K under the *NVT* ensemble for 50 ps with a weak restraint of 5 kcal mol^{-1} $Å^{-2}$. Subsequently, the systems were maintained for 1 ns of density equilibration in the *NPT* ensemble at a target temperature of 300 K and a target pressure of 1.0 atm using the Langevin thermostat[36] and the Barendsen barostat[37] with collision frequency of 2 ps and a pressure relaxation time of 1 ps, with a weak restraint of 1 kcal mol^{-1} $Å^{-2}$. This 1 ns of density equilibration is not identical with conformational equilibration but rather is a weakly restrained MD simulation in which the system is slowly relaxed to achieve a uniform density after heating dynamics under periodic boundary conditions. Thereafter, we removed all of the restraints applied during heating and density dynamics and further equilibrated the systems for ∼3 ns to get a well-settled pressure and temperature for conformational and chemical analyses. This was followed by an MD production run for 360 ns for each of the six systems. During all of the MD simulations, the covalent bonds containing hydrogen were constrained using SHAKE,[38] and particle-mesh Ewald[39] was used to treat long-range electrostatic interactions. All of the MD simulations were performed with the GPU version of the Amber 14 package.[40]

2.3. QM/MM Methodology. Equilibrated snapshots from the MD simulations were used for the subsequent QM/MM calculations. All of the water molecules beyond the sphere of the enzyme were removed. The resulting enzyme was solvated with a 16 Å layer of TIP3P water molecules (yielding a total of ca. 29 000 atoms). All of the QM/MM calculations were performed using ChemShell,[41,42] combining Turbomole[43] for the QM part and DL_POLY[44] for the MM part. The CHARMM27 force field was employed for the MM region. Shifting force fields from Amber to CHARMM should have minor effects on the QM/MM results since the two force fields have similar

DOI: 10.1021/jacs.5b08737
J. Am. Chem. Soc. 2016, 138, 837−845

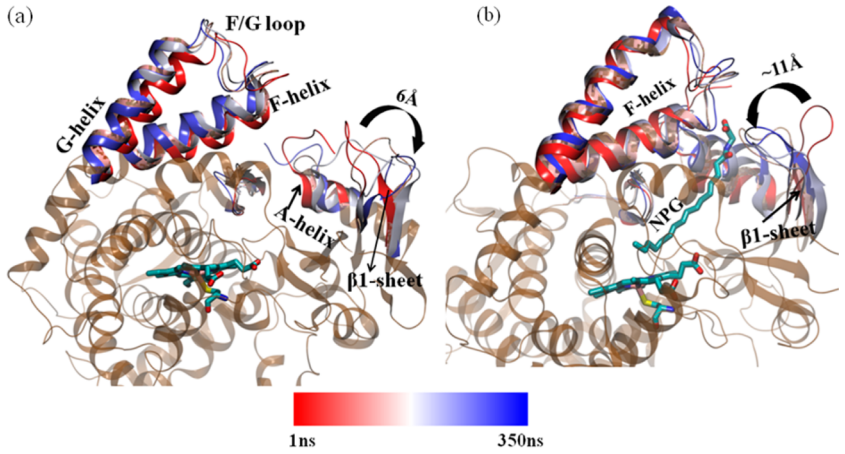


Figure 2. Superpositions of four snapshots of the enzyme (the heme and NPG are shown using stick models). The color bar at the bottom shows the simulation times of the snapshots during the MD simulations. (a) Substrate-free (open-cavity) form. Here the channel stays open. The 6 Å movement of the β1 sheet should be noted. (b) Substrate docked into the open-cavity form of the substrate-free enzyme. The 11 Å movement of the β1 sheet to close the channel should be noted.

atomic charges for the standard residues and water molecules and the MM part makes a minor contribution to the QM/MM barriers in P450. In fact, previous studies[5a,b] showed that the QM/MM results are not affected much by the identity of the force field.

The electronic embedding scheme[45] was used to account for the polarizing effect of the enzyme environment on the QM region. Hydrogen link atoms with the charge-shift model[42] were applied to treat the QM−MM boundary. In QM/MM geometry optimizations, the QM region was treated using the hybrid UB3LYP[46] functional with two basis sets. For geometry optimization and frequency calculations, we used a basis set B1 consisting of LACVP for iron and 6-31G(d) for all other atoms in the QM/MM calculations. In probing of the enantioselectivity of hydroxylation at the ω-3 site, we double-checked the B1 results by optimizing some of the key species with the all-electron def2-SVP basis set, denoted as B1′. All of the QM/MM transition states (TSs) were located by relaxed potential energy surface (PES) scans followed by full TS optimizations using the P-RFO optimizer implemented in the HDLC code.[47] The energies were further corrected with the large all-electron def2-TZVP basis set, denoted as B2. The zero-point energy (ZPE) was calculated for all species, and all of the final energies are reported as UB3LYP/B2+ZPE. For Cpd I-mediated reactions, the $S = {}^3/_2$ and $S = {}^1/_2$ states generally exhibit similar reactivities,[9,16] so no added information is provided by studying both states. Hence, all of the reactions of Cpd I were studied in the $S = {}^1/_2$ state.

3. RESULTS AND DISCUSSION

3.1. Substrate Gating and Mediation of Conformational Changes. The mechanism of substrate entry and closing and opening of the channel have not yet been fully studied. To see how the entrance of the substrate induces the major conformational change of the enzyme, we performed MD simulations of the substrate-free enzyme (open form), the substrate-bound enzyme (closed form), and the enzyme with the substrate docked into the open form. Figure 2a presents the superposition of four snapshots during the simulation of the substrate-free heme domain. It is quite clear that the substrate entry channel—formed by the A helix, the β1 sheet, and a few residues of the F helix—has a strong tendency to move outward in order to accommodate the substrate without any steric penalties. The β1 sheet and A helix, which form a significant part of the cavity, move outward and generate a wider channel that eases the access of the substrate into the active site of the enzyme. Although the substrate-free structure has to begin with a wide channel, it nevertheless further opens the pore wider to facilitate the substrate's entrance.

In order to ascertain whether it is indeed the presence of the substrate that closes the channel, we docked the substrate into the substrate-free open-channel structure and performed 360 ns simulations. The major changes are indicated in Figure 2b. Interestingly, the same substructure that moved outward to form the wider channel in the absence of NPG (Figure 2a) now moves inward to form a closed channel in the substrate-docked structure (Figure 2b). Similarly, the β1 sheet, which was far away from the F helix and formed the wide channel, moves as much as 11 Å after 100 ns of simulation and approaches the F helix, eventually forming a closed channel. This closure of the channel is irreversible, and the channel remains closed during the remaining 250 ns of the simulation. Animations of these movements of the subunits in the substrate-free and -bound states are shown the MD videos VS1 and VS2, respectively, in the Supporting Information.

As expected, the MD simulation of the already substrate-bound heme domain of P450 BM3 does not show any major conformational rearrangement, as the cavity is already closed and does not have more configurational space for further closing on the substrate (a superposition of three snapshots of the substrate-bound structure is shown in Figure S1). The MD simulations of the substrate-free and substrate-bound structures and of the structure in which the substrate was docked in the open channel clearly show that the open channel in the substrate-free structure is the initial structure of the heme domain and that it turns into a closed-cavity structure only after the substrate enters the cavity. Once the substrate binds optimally, thereafter there is no more conformational rearrangement. Therefore, the conformational transition from the open-channel to the closed-channel heme domain is entirely a substrate-induced process. *Thus, the convergence of the three structures reveals how the substrate gates the enzyme and closes its entrance channel, thereby preparing the enzyme to start its catalytic cycle.*

A closer inspection of the MD trajectory for the substrate docked into the open channel indicates interactions of the substrate with several protein residues, which thereby cause the closure of the channel. These interactions are explicitly indicated in Figure 3. Inspection of Figure 3 reveals that the

DOI: 10.1021/jacs.5b08737
J. Am. Chem. Soc. 2016, 138, 837−845

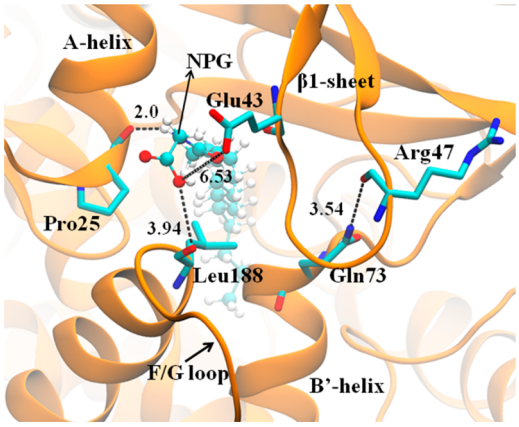


Figure 3. Snapshot showing the interactions of the substrate with specific protein residues. The interactions that cause the closing of the channel are shown as dotted lines. Distances are in Å. It should be noted that the carboxyl terminal of the substrate behaves as a bridge that connects the A helix, β1 sheet, and the F/G loop.

substrate behaves as a *three-way bridge* via Pro25 of the A helix, Glu43 of the β_1 sheet, and Leu188 of the F/G loop and thereby pulls these moieties in such a manner that they come closer to each other and form a closed channel. These interactions bring Arg47 of the β_1 sheet and Gln73 of the F/G loop close to each other, from an initial distance of 15 Å to 2.5 Å in the closed structure (see Figure S2 for comparisons of the distances between Arg47 and Gln73 in the substrate-free and -docked forms). The MD simulation in the absence of the substrate does not reveal any such interactions. A more global change is seen by inspection the hydrophobicity/hydrophilicity of the pore of the channel. Initially the pore is hydrophilic, but during the simulations all of hydrophobic residues that reside near the channel cluster together to attract the hydrophobic methyl end of the fatty acid substrate so that it moves toward the heme site for oxidation (see calculations of hydrophobic and hydrophilic surfaces in Figure S3). All of these interactions and changes vis-à-vis the substrate-free state reveal large conformational rearrangements of the channel and its surroundings as well as the fundamental role of the carboxyl group of the fatty acid in its oxidation.

3.2. Substrate Dynamics in the Resting State vis-à-vis the Cpd I State Predicts the Regioselectivity of C−H Hydroxylation. The catalytic machinery of P450 involves different states of the heme species. The ultimate oxidant is Cpd I, which carries out the C−H hydroxylations of the fatty acid. To elucidate the role of substrate juxtaposition vis-à-vis the nature of the catalytic state in dictating the observed regioselectivity for C−H hydroxylation, we performed comparative MD simulations for both the resting state and the Cpd I state in the presence of the substrate.

We recall that cytochrome P450 BM3 shows special regioselectivity for C−H hydroxylation at the ω-1, ω-2, and ω-3 positions of the substrate (Figure 1b) but never at the ω terminal! However, the crystal structure of NPG-bound P450 BM3 in the resting state shows that ω-1, ω-2, and ω-3 are far away from the heme site. Therefore, there must exist a trigger that shifts the ω-1, ω-2, and ω-3 positions of NPG toward the heme at the expense of distancing the ω position. Furthermore, this trigger must leave the ω-4 position out of reach of Cpd I.

Figure 4 shows the distances of the ω/ω-1/ω-2/ω-3 positions of the substrate to the Fe atom of (a) the resting state and (b) Cpd I. As can be seen in Figure 4a, initially the ω position is the closest to the iron (4−5 Å), while the ω-2 and ω-3 positions are approximately 6−8 Å away. However, after 275 ns, the substrate moves downward, bringing its ω-2/ω-3 positions closer to the heme center (∼4 Å), while the ω terminal moves away from the heme iron to a distance of 8−12 Å.

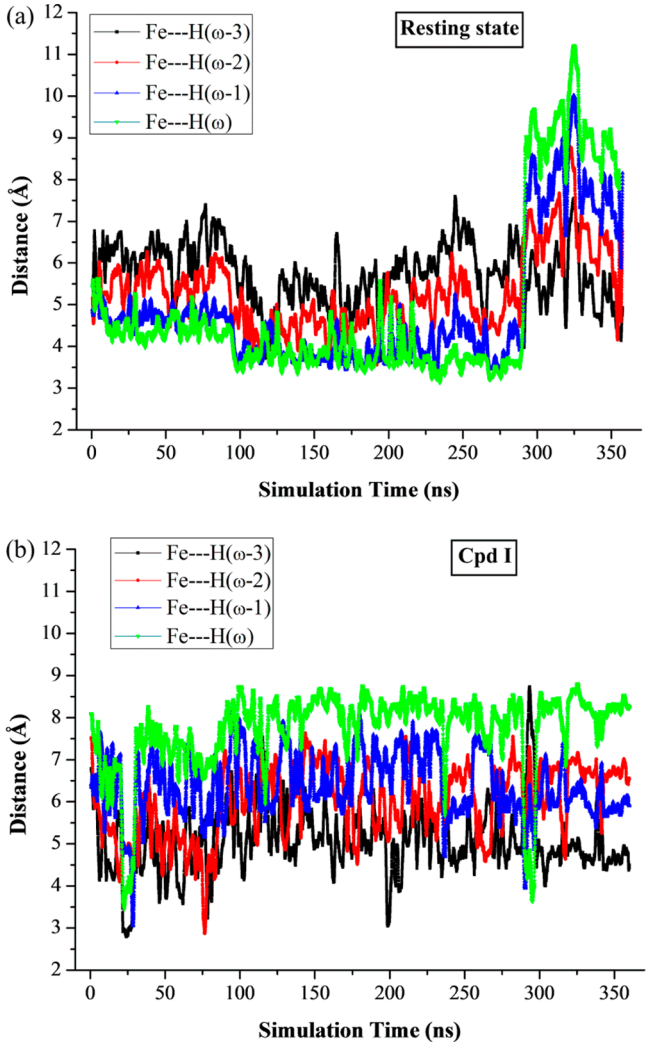


Figure 4. Distance fluctuations of different C−H groups to the Fe center for (a) the resting state simulations and (b) the Cpd I state simulations. The color code is green for ω, blue for ω-1, red for ω-2, and black for ω-3.

The simulations of the substrates in the presence of Cpd I (Figure 4b) further underscore these trends by preferring the juxtaposition of Cpd I and the ω/ω-1/ω-2/ω-3 positions, which exhibit damped fluctuations compared with the situation in the resting state. The steady distances of the C−H terminals from the Fe−O moiety (Figure 4b) clearly indicate a *less dynamic nature of NPG in the Cpd I state relative to the resting state*. Furthermore, the population distribution of Fe···H(ω-4) from the MD trajectory reveal that ω-4 remains more distant from Fe−O relative to ω-3 (Figure S4). On average, ω-3 is the closest position to Fe with a distance of 4.72 Å, and the farthest is the ω position with a distance of 8 Å. Clearly, the long-time-scale MD simulation juxtaposes the ω-1/ω-2/ω-3 sites of the fatty acid to Cpd I while distancing the ω terminal. These juxtapositions match the experimentally observed regioselectiv-

DOI: 10.1021/jacs.5b08737
J. Am. Chem. Soc. 2016, 138, 837−845

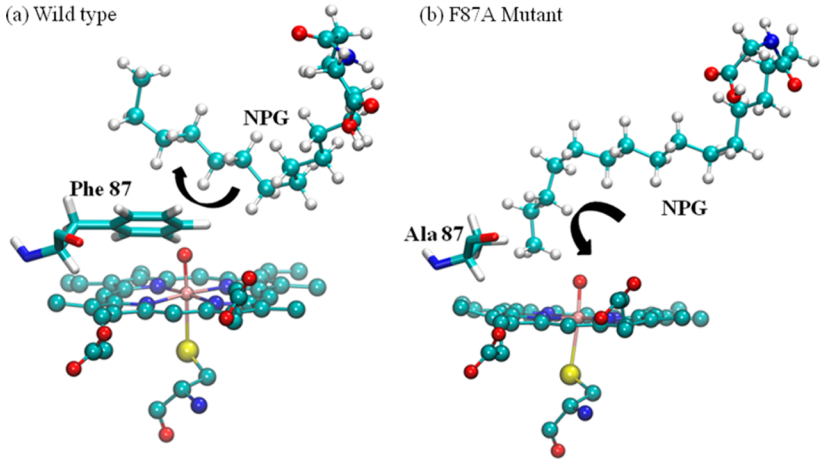


Figure 5. (a) In the wild-type enzyme, the residue Phe87 resides between the heme plane and the substrate's tail, causing the substrate to curl upward because of steric repulsion with Phe87. Consequently the ω end moves away from Fe−O. (b) In the F87A mutant, there is sufficient room for the substrate to slide downward and thereby bring the ω end to an optimum position for C−H hydroxylation.

ity for C−H hydroxylation by the WT enzyme and the fact that hydroxylation of the ω position has never been observed in the WT. It is notable that all of the above productive changes vis-à-vis the X-ray structure are mostly caused by temperature differences. Thus, the crystal structure is characterized at a very low temperature while the MD simulations are performed at a biologically relevant temperature (300 K). Such changes in the position of the substrate at different temperatures have also been observed by others.[48]

Figure S5 further characterizes the substrate dynamics during the MD simulations by inspecting the distance fluctuations of the centers of masses of the heme and the NPG substrate. During all six MD simulations, no motion of heme was observed. Therefore, any large changes in these distances would indicate that the entity that moves is NPG. Thus, the center-of-mass distance starts at almost 15 Å for the initial 250 ns of simulations and is reduced, due to NPG movement, to less than 8 Å after 250 ns of simulation in the resting state. Furthermore, the center-of-mass distance of the substrate and Cpd I is almost constant, hence providing further support for the lesser flexibility of NPG in the presence of Cpd I (see videos VS3 and VS4 for the NPG dynamics in the resting and Cpd I states, respectively). The observations in Figure 4 (further augmented by Figure S5) clearly show that it is the dynamics of NPG in the substrate channel that is responsible for the ultimate oxidation at the ω-1/ω-2/ω-3 positions, while the lesser mobility in the Cpd I state implies the role of Cpd I in the oxidation process. Still, however, the sequestering by Phe87 is the prime cause for the juxtaposition of these sites vis-à-vis Cpd I, as shown next. Figure S6 shows a non-covalent interaction plot[49] demonstrating that the sequestering effect is due to non-covalent attractive and repulsive interactions of Phe87 with the substrate and the heme.

3.3. Dynamics of the Phe87Ala Mutant. Experimentally it has been observed that the WT enzyme leads to ω-1/ω-2/ω-3 hydroxylations without a trace of ω oxidation, whereas the F87A mutant, where Phe87 is replaced by the smaller Ala residue, gives >90% ω hydroxylation with lauric and myristic acids.[26] We therefore performed MD simulations of the F87A mutant to ascertain the role of the phenyl group of Phe87 in affecting the regioselectivity of oxidation in the WT enzyme and contrast it with the mutant where Phe87 is replaced by the smaller Ala residue.

In the initial stage of the simulation, the ω terminal was farther away compared with the ω-2 and ω-3 ends. However, at later stages, the unwinding of the chain caused the ω end to move toward the Fe−O moiety and become the closest position of the substrate to the oxidant (see the 200 ns snapshot in Figure S7 and video VS5). This is in contrast with the WT complex, where the ω end is distant.

Clearly, the modes of substrate juxtaposition for the WT and F87A enzymes account for the observed regioselectivities in the two enzymes, as shown in Figure 5, which depicts the mechanical scenarios for the positional oxidation. In the WT enzyme (Figure 5a), Phe87 hinders the approach of the terminal part of the substrate's chain to Cpd I, and consequently, the chain curls up and distances the ω terminal from Cpd I. By contrast, the mutation of Phe to Ala (Figure 5b) eliminates this steric penalty, allowing the substrate's tail to slide down and enter deeper toward the iron, thus bringing the ω terminal to an optimal position for hydroxylation.

3.4. QM/MM Calculations for Enantioselectivity. As shown above, the MD simulations successfully predict the observed hydroxylation of the C−H bonds in the ω-1/ω-2/ω-3 sites of NPG. Since all these positions have pro-*R* and pro-*S* C−H bonds, we might wonder whether theory can also account correctly for the enantioselectivity of hydroxylation. A detailed MD trajectory analysis shows that the pro-*R* C−H bonds are invariably closer to Cpd I for the ω-1 and ω-2 sites (Figure S8), suggesting that ω-1 and ω-2 will be selectively oxidized into pro-*R* products. These *R*-enantioselective juxtapositions are in good agreement with the experimental findings of pro-*R* hydroxylations at the ω-1 and ω-2 positions. Indeed, the QM/MM barriers for ω-1 hydroxylation (see Figure S11) further support this association that the closer pro-*R* C−H bonds undergo more facile H abstraction compared with the more distant pro-*S* C−H bonds.

By contrast, the simulations for the ω-3 site (Figure S9) show that in the majority of trajectories it is the pro-*S* C−H bond that is closer to the oxo ligand of Cpd I. However, the observed enantioselectivity, 75% *R* versus 25% *S*, does not reflect this proximity of the pro-*S* C−H bond. Therefore, it was deemed essential to complement the MD simulations by QM/MM calculations, as shown in Figure 6. Figure 6a displays the barriers for the H-abstraction steps and Figure 6b presents the QM/MM optimized reactant clusters and transition-state

DOI: 10.1021/jacs.5b08737
J. Am. Chem. Soc. 2016, 138, 837−845

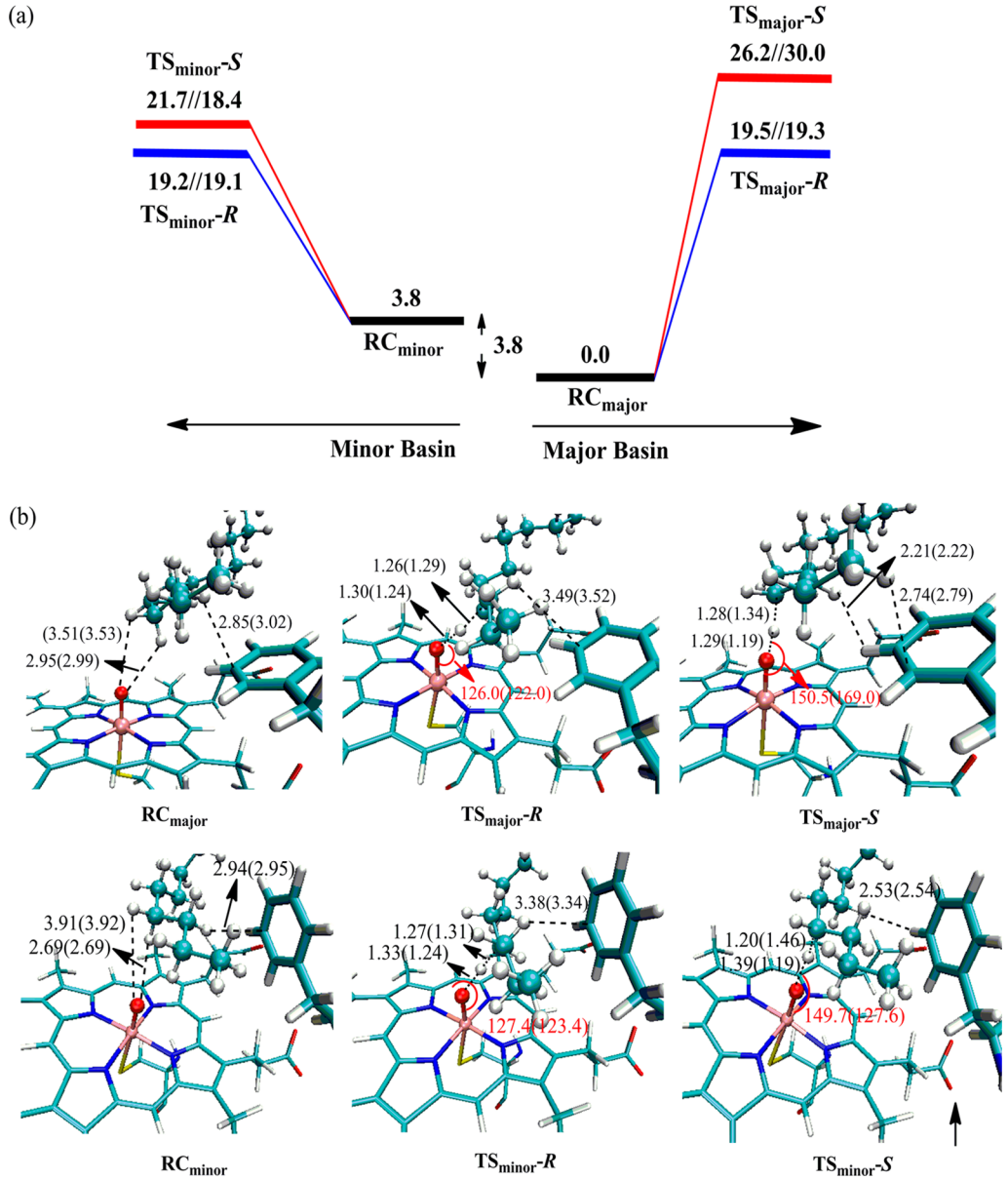


Figure 6. (a) Comparison of the pro-*R* (blue profiles) and pro-*S* (red profiles) H-abstraction processes at the ω-3 position in the major and minor conformational basins. The MD simulations showed that the major basin is more frequently visited than the minor basin. The relative numbers of visits translate, by the Boltzmann distribution, to an energy difference of 3.8 kcal/mol between the reactant clusters of the two basins (denoted as **RC$_{major}$** and **RC$_{minor}$**, respectively). The relative energies (in kcal/mol) of the TSs are given in the order (B2//B1′)//(B2//B1). (b) Key geometric features of species involved in the H-abstraction processes in the major and minor conformational basins. The geometric data (distances in Å and angles in deg) are given as B1′(B1).

species for two snapshots taken from the MD trajectory. Since the trajectory displayed major and minor conformational basins, marked by the juxtaposition of Phe87 vis-à-vis the substrate, we selected two snapshots, one from each of these two basins. The reactant clusters are accordingly marked in Figure 6 as **RC$_{major}$** and **RC$_{minor}$**. As can be seen from **RC$_{major}$** in Figure 6b, in the major conformational basin Phe87 is parallel to the heme and hence blocks the substrate, whereas **RC$_{minor}$** shows that in the minor conformational basin Phe87 is further away from the substrate and perpendicular to the heme. A detailed analysis of the MD trajectory (see pp S11−S12 in the Supporting Information) shows that the frequencies of visits are 9978 and 22 for the major and minor basins, respectively. Translating the relative visit frequencies into energy units at 310 K indicates that the major conformational basin is on average lower in energy by 3.8 kcal/mol compared with the minor basin. Accordingly, Figure 6a shows that the reactant cluster **RC$_{minor}$** for the minor basin lies at 3.8 kcal/mol relative to **RC$_{major}$**. Thus, all of the energies in the figure are given relative to **RC$_{major}$** as the zero level of the energy profile. The TS energies relative to these clusters are shown in two sets of data in Figure 6a. The left-hand set corresponds to the QM/MM energies obtained using the B2 and B1′ basis sets, while the right-hand set corresponds to the B2//B1 energy.

Starting from **RC$_{major}$** and going to the right, we present the energy profiles for H abstraction (in the $S = {}^1/_2$ spin state of Cpd I[9]) for the major basin. It can be seen from the structure of **RC$_{major}$** in Figure 6b that the pro-*S* C−H bond is closer to the oxo moiety of Cpd I compared with the pro-*R* C−H. Nevertheless, irrespective of the basis set combination, the H

DOI: 10.1021/jacs.5b08737
J. Am. Chem. Soc. 2016, 138, 837−845

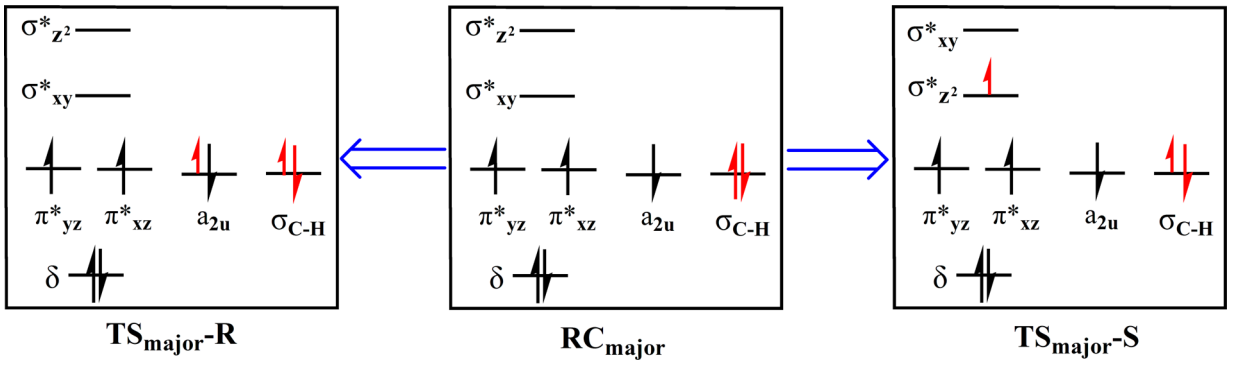


Figure 7. Electron-shift diagrams for *R* and *S* H abstractions at the UB3LYP/B2 level for the major basin of the *ω*-3 site. The short arrows represent the partial electron occupation in the TSs.

abstraction of the pro-*S* hydrogen atom proceeds with a much higher barrier than for the pro-*R* atom. The root cause of this pro-*R* preference is Phe87. As shown by the QM/MM optimized structures of the major basin ($\mathbf{RC}_{major}$, $\mathbf{TS}_{major}$-***R***, and $\mathbf{TS}_{major}$-**S** in Figure 6b), Phe87 is nestled just below the substrate. During the abstraction of the pro-*R* hydrogen, the substrate tilts away from Phe87, so that the pro-*R* H abstraction is not so much affected by steric blocking of Phe87. In contrast, the H abstraction of the pro-*S* C−H bond brings the substrate closer to Phe87. The penalty for this enhanced steric blocking is that $\mathbf{TS}_{major}$-**S** is forced to assume an unnatural trajectory with an Fe−O−H angle of 150.5° (169.0°), which reduces the overlap of the C−H bond orbital with the $\pi^*_{xz/yz}$ orbitals of the Fe−O moiety[9] and raises the energy of this transition state by populating the high-lying $\sigma^*_{z^2}$ orbital (Figure 7). Consequently, despite the closer initial distance of the pro-*S* C−H bond to the iron−oxo moiety of Cpd I, Phe87 enforces a higher energy barrier for activating this bond compared with the pro-*R* C−H bond.

Shifting now in Figure 6a to $\mathbf{RC}_{minor}$ and going in the left-hand direction, we present the energy profiles for H abstraction of the pro-*S* and pro-*R* C−H bonds at the *ω*-3 position in the minor basin. As Phe87 is now away, the pro-*S* and pro-*R* barriers for H abstraction are closer in energy than in the major basin. Nevertheless, the B2//B1′ datum indicates that the pro-*S* H-abstraction process still experiences some steric repulsion from Phe87 even in the minor basin.

Relative to $\mathbf{RC}_{major}$, the B2//B1 barriers in the major/minor basins are 19.3/19.1 kcal/mol for pro-*R* abstraction and 30.0/18.4 kcal/mol for pro-*S* abstraction. For the B2//B1′ data, the barriers correspond to 19.5/19.2 kcal/mol for pro-*R* and 26.2/21.7 for pro-*S*. When the relative barriers are used in the Eyring equation, the B2//B1 barriers correspond to an enantioselectivity of 24% *R* versus 76% *S*, while those for B2//B1′ correspond to an enantioselectivity of 98% *R* versus 2% *S*. Since the experimental enantioselectivity corresponds to a difference of 0.7 kcal/mol, it is clear that theory is doing the best it can, predicting, in accord with experiment,[26,50] that the reaction will mostly lead to *R* enantioselectivity of the hydroxylated product with an *S*-enantioselective minor product. Thus, all of these calculations support the analysis that the Phe87 steric sequestering effects at the *ω*-3 position render the pro-*R* H abstraction more favorable than the pro-*S* H-abstraction even though the pro-*S* C−H bond is closer to Cpd I. Let us note that we followed the reaction beyond the H abstraction all the way to the hydroxylation and found that the intermediate radical has no barriers for rebound (Figures S10−S12); hence, the enantioselectivity, which is determined in the H-abstraction step, is conserved.

3.5. Summary of the Enantioselectivity at the *ω*-3 Position versus the *ω*-1 and *ω*-2 Positions. Clearly, the combined MD simulations and QM/MM calculations rationalize nicely the enantioselectivity for the *ω*-3 site and show *that Phe87 exerts enantioselective steric blocking on the H abstraction of the pro-S C−H bond*. Thus, despite the closer proximity of the pro-*S* C−H bond to Cpd I in the majority of the MD frames, the *S* hydroxylation at *ω*-3 remains unfavorable. The scenario for the *ω*-1 and *ω*-2 sites is different. For these sites the pro-*R* C−H bond is closer to Cpd I relative to the pro-*S* C−H bond (Figure S8), and the latter C−H bond still suffers steric repulsion from Phe87. These two factors make the pro-*R* H abstraction more favorable than the pro-*S* H abstraction at *ω*-1 and *ω*-2. This conclusion is fully supported by our QM/MM calculations (Figure S11), which show that the pro-*R* H abstraction is favored over the pro-*S* H abstraction by 6.3 kcal/mol. These findings, as predicted by the MD simulations through the use of the proximity argument and reinforced by the QM/MM calculations, show good agreement with the experiments.

4. CONCLUSIONS

Long-time-scale molecular dynamic simulations of P450 BM3 supplemented by QM/MM calculations predict the active-site gating mechanism as well as the regio- and enantioselectivity of fatty acid hydroxylation by the enzyme and its F87A mutant:

(a) The simulations clearly show that the major conformational transition from the open state to the closed state is exclusively induced by the NPG substrate, which pulls together the A helix, the *β*1 sheet, and F/G loop by hydrogen-bonding interactions and thereby closes the channel.

(b) The Phe87 residue, which is nestled between the heme plane and the substrate, causes the substrate to avoid steric repulsion by bending its chain upward, thus bringing the *ω*-1/*ω*-2/*ω*-3 sites closer to the heme and moving away the *ω* end. These positional proximities accord with the regioselectivity.

(c) Upon mutation of Phe87 to Ala, the steric constraint is removed, and the substrate is able to juxtapose its *ω* end exclusively toward the active species. This is in accord with the observed >90% *ω* hydroxylation in the mutant enzyme.

(d) Our QM/MM calculations show that Phe87 applies enantioselective steric control on the hydroxylation of

DOI: 10.1021/jacs.5b08737
J. Am. Chem. Soc. 2016, 138, 837−845

the C−H bonds at the ω-3 position. Thus, despite the closer proximity of the pro-*S* C−H bond to the active species, the H abstraction of the pro-*S* hydrogen requires a much higher barrier than the pro-*R* abstraction because of the steric hindrance of the phenyl ring of Phe87 in the transition state of the pro-*S* abstraction.

(e) The enantioselective *R* hydroxylations at the ω-1 and ω-2 positions are predicted directly by the MD simulations (and supported by the QM/MM calculations), which position the pro-*R* C−H bonds closer to the active compound without being perturbed by Phe87.

The present MD simulations appear also to reflect the key stages in the catalytic cycle. Thus, in the resting state of the enzyme, the substrate is highly dynamic, since much movement is needed for the substrate to adapt to and cause conformational changes in the entrance channel. On the other hand, in the presence of the active species Cpd I that has to perform the substrate oxidation, the substrate has to be less dynamic in order to lower the free energy barrier for oxidation, and indeed, the MD simulation reveals less substrate movement.

Finally, by articulating such fundamental features of enzymatic reactivity, we have uncovered clues about the evolution of this enzyme to do its particular tasks as a fatty acid hydroxylase. It is remarkable how theory is capable of picking up the salient features of this machinery, and in so doing it also provides insight for eventual engineering of P450 BM3 with site-selective C−H functionalization capabilities. The good match to experiment underscores the role of theory as a partner to experiment and the future role of MD as a reliable predictor of the chemistry of P450 enzymes.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.5b08737.

Supporting figures, calculation of populations for the major and minor basins, total QM/MM and ZPE energies, and Cartesian coordinates of all computed species (PDF)

Video files (ZIP)

■ AUTHOR INFORMATION

Corresponding Author

*sason@yfaat.ch.huji.ac.il

Author Contributions

†K.D.D. and B.W. contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This article is dedicated to Christian Amatore. The work was supported by the Israel Science Foundation (Grant ISF-1183/13 to S.S.).

■ REFERENCES

(1) McCammon, J. A.; Gelin, B. R.; Karplus, M. *Nature* **1977**, *267*, 585−590.

(2) Karplus, M.; McCammon, J. A. *Nat. Struct. Biol.* **2002**, *9*, 646−652.

(3) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227−249.

(4) Warshel, A. *Angew. Chem., Int. Ed.* **2014**, *53*, 10020−10031.

(5) (a) Senn, H. M.; Thiel, W. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198−1229. (b) Hu, H.; Elstner, M.; Hermans, J. *Proteins: Struct., Funct., Genet.* **2003**, *50*, 451−463. (c) Lin, H.; Truhlar, D. G. *Theor. Chem. Acc.* **2007**, *117*, 185−199. (d) Ryde, U. *Curr. Opin. Chem. Biol.* **2003**, *7*, 136−142. (e) Olah, J.; Mulholland, A. J.; Harvey, J. N. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 6050−6055. (f) Chung, L. W.; Sameera, W. M. C.; Ramozzi, R.; Page, A. J.; Hatanaka, M.; Petrova, G. P.; Harris, T. V.; Li, X.; Ke, Z.; Liu, F.; Li, H.-B.; Ding, L.; Morokuma, K. *Chem. Rev.* **2015**, *115*, 5678−5796.

(6) (a) *Cytochrome P450: Strcture, Mechanism and Biochemistry*, 2nd ed.; Ortiz de Montellano, P. R., Ed.; Plenum Press: New York, 1995. (b) Sono, M.; Roach, M. P.; Coulter, E. D.; Dawson, J. H. *Chem. Rev.* **1996**, *96*, 2841−2881. (c) Ortiz de Montellano, P. R. *Chem. Rev.* **2010**, *110*, 932−948. (d) Groves, J. T. *Nat. Chem.* **2014**, *6*, 89−91. (e) van Eldik, R. *Chem. Rev.* **2005**, *105*, 1917−1922.

(7) Denisov, I. G.; Makris, T. M.; Sligar, S. G.; Schlichting, I. *Chem. Rev.* **2005**, *105*, 2253−2278.

(8) Groves, J. T. *J. Chem. Educ.* **1985**, *62*, 928−931.

(9) (a) Shaik, S.; de Visser, S. P.; Kumar, D.; Altun, A.; Thiel, W. *Chem. Rev.* **2005**, *105*, 2279−2328. (b) Shaik, S.; Cohen, S.; Wang, Y.; Chen, H.; Kumar, D.; Thiel, W. *Chem. Rev.* **2010**, *110*, 949−1017. (c) Yoshizawa, K. *Coord. Chem. Rev.* **2002**, *226*, 251−259.

(10) Poulos, T. L. *Chem. Rev.* **2014**, *114*, 3919−3962.

(11) Kille, S.; Zilly, F. E.; Acevedo, J. P.; Reetz, M. T. *Nat. Chem.* **2011**, *3*, 738−743.

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

(13) Rittle, J.; Green, M. T. *Science* **2010**, *330*, 933−937.

(14) Mak, P. J.; Luthra, A.; Sligar, S. G.; Kincaid, J. R. *J. Am. Chem. Soc.* **2014**, *136*, 4825−4828.

(15) Boaz, N. C.; Bell, S. R.; Groves, J. T. *J. Am. Chem. Soc.* **2015**, *137*, 2875−2885.

(16) Wang, B.; Li, C.; Dubey, K. D.; Shaik, S. *J. Am. Chem. Soc.* **2015**, *137*, 7379−7390.

(17) Li, H.; Poulos, T. L. *Biochim. Biophys. Acta, Mol. Cell Biol. Lipids* **1999**, *1441*, 141−149.

(18) Sevrioukova, I. F.; Li, H.; Zhang, H.; Peterson, J. A.; Poulos, T. L. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 1863−1868.

(19) Ravichandran, K. G.; Boddupalli, S. S.; Hasermann, C. A.; Peterson, J. A.; Deisenhofer, J. *Science* **1993**, *261*, 731−736.

(20) Li, H.; Poulos, T. L. *Biochimie* **1996**, *78*, 695−699.

(21) Li, H.; Poulos, T. L. *Nat. Struct. Biol.* **1997**, *4*, 140−146.

(22) Li, H.; Poulos, T. L. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1995**, *51*, 21−32.

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

(24) Munro, A. W.; Leys, D. G.; McLean, K. J.; Marshall, K. R.; Ost, T. W. B.; Daff, S.; Miles, C. S.; Chapman, S. K.; Lysek, D. A.; Moser, C. C.; Page, C. C.; Dutton, P. L. *Trends Biochem. Sci.* **2002**, *27*, 250−257.

(25) Whitehouse, C. J. C.; Bell, S. G.; Wong, L.-L. *Chem. Soc. Rev.* **2012**, *41*, 1218−1260.

(26) Miura, Y.; Fulco, A. J. *Biochim. Biophys. Acta, Lipids Lipid Metab.* **1975**, *388*, 305−317.

(27) Oliver, C. F.; Modi, S.; Sutcliffe, M. J.; Primrose, W. U.; Lian, L.-Y.; Roberts, G. C. K. *Biochemistry* **1997**, *36*, 1567−1572.

(28) Cryle, M. J.; De Voss, J. J. *Angew. Chem., Int. Ed.* **2006**, *45*, 8221−8223.

(29) Cryle, M. J.; De Voss, J. J. *ChemBioChem* **2008**, *9*, 261−266.

(30) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *25*, 1605−1612.

(31) Case, D. A.; Babin, V.; Berryman, J. T.; Betz, R. M.; Cai, Q.; Cerutti, D. S.; Cheatham, T. E., III; Darden, T. A.; Duke, R. E.; Gohlke, H.; Götz, A. W.; Gusarov, S.; Homeyer, N.; Janowski, P.; Kaus, J.; Kolossváry, I. J.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Luchko, T.; Luo, R.; Madej, B.; Merz, K. M.; Paesani, F.; Roe, D. R.; Roitberg, A.; Sagui, C.; Salomon-Ferrer, R.; Seabra, G.; Simmerling, C. L.; Smith, W.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.;

DOI: 10.1021/jacs.5b08737
J. Am. Chem. Soc. 2016, 138, 837−845

Kollman, P. A. *Amber 14*; University of California: San Francisco, 2015.
(32) Shahrokh, K.; Orendt, A.; Yost, G. S.; Cheatham, T. E. *J. Comput. Chem.* **2012**, *33*, 119−133.
(33) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269−10280.
(34) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollmann, P. A. *J. Am. Chem. Soc.* **1993**, *115*, 9620−9631.
(35) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(36) Izaguirre, J. A.; Catarello, D. P.; Wozniak, J. M.; Skeel, R. D. *J. Chem. Phys.* **2001**, *114*, 2090−2098.
(37) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(38) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(39) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(40) Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. *J. Chem. Theory Comput.* **2013**, *9*, 3878−3888.
(41) Sherwood, P.; de Vries, A. H.; Guest, M. F.; Schreckenbach, G.; Catlow, C. R. A.; French, S. A.; Sokol, A. A.; Bromley, S. T.; Thiel, W.; Turner, A. J.; Billeter, S.; Terstegen, F.; Thiel, S.; Kendrick, J.; Rogers, S. C.; Casci, J.; Watson, M.; King, F.; Karlsen, E.; Sjøvoll, M.; Fahmi, A.; Schäfer, A.; Lennartz, C. *J. Mol. Struct.: THEOCHEM* **2003**, *632*, 1−28.
(42) Metz, S.; Kästner, J.; Sokol, A. A.; Keal, T. W.; Sherwood, P. *Comput. Mol. Sci.* **2014**, *4*, 101−110.
(43) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165−169.
(44) Smith, W.; Forester, T. R. *J. Mol. Graphics* **1996**, *14*, 136−141.
(45) Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996**, *100*, 10580−10594.
(46) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(47) Billeter, S. R.; Turner, A. J.; Thiel, W. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2177−2186.
(48) (a) Jovanovic, T.; Farid, R.; Friesner, R. A.; McDermott, A. E. *J. Am. Chem. Soc.* **2005**, *127*, 13548−13552. (b) Ravindranathan, K. P.; Gallicchio, E.; Friesner, R. A.; McDermott, A. E.; Levy, R. M. *J. Am. Chem. Soc.* **2006**, *128*, 5786−5791.
(49) Johnson, E. R.; Keinan, S.; Mori-Sánchez, P.; Contreras-García, J.; Cohen, A. J.; Yang, W. *J. Am. Chem. Soc.* **2010**, *132*, 6498−6506.
(50) Cryle, M. J.; De Voss, J. J. *Tetrahedron: Asymmetry* **2007**, *18*, 547−551.

DOI: 10.1021/jacs.5b08737
J. Am. Chem. Soc. 2016, 138, 837−845