

DOI: 10.1002/jcc.21922

Quantum Mechanically Derived AMBER-Compatible Heme Parameters for Various States of the Cytochrome P450 Catalytic Cycle

Kiumars Shahrokh,^[a] Anita Orendt,^[b] Garold S. Yost,^[a] and Thomas E. Cheatham, III^{*[c]}

Molecular mechanics (MM) methods are computationally affordable tools for screening chemical libraries of novel compounds for sites of P450 metabolism. One challenge for MM methods has been the absence of a consistent and transferable set of parameters for the heme within the P450 active site. Experimental data indicate that mammalian P450 enzymes vary greatly in the size, architecture, and plasticity of their active sites. Thus, obtaining X-ray-based geometries for the development of accurate MM parameters for the major classes of hepatic P450 remains a daunting task. Our previous work with preliminary gas-phase quantum mechanics (QM)-derived atomic partial charges greatly improved the accuracy of docking studies of raloxifene to CYP3A4. We have therefore developed and tested a consistent set of transferable MM

parameters based on gas-phase QM calculations of two model systems of the heme—a truncated (T-HM) and a full (F-HM) for four states of the P450 catalytic cycle. Our results indicate that the use of the atomic partial charges from the F-HM further improves the accuracy of docked predictions for raloxifene to CYP3A4. Different patterns for substrate docking are also observed depending on the choice of heme model and state. Newly parameterized heme models are tested in implicit and explicitly solvated MD simulations in the absence and presence of enzyme structures, for CYP3A4, and appear to be stable on the nanosecond simulation timescale. The new force field for the various heme states may aid the community for simulations of P450 enzymes and other heme-containing enzymes. © 2011 Wiley Periodicals, Inc. *J Comput Chem* 33: 119–133, 2012

Keywords: cytochrome P450 enzymes · heme force field parameters · molecular mechanics · RESP charges · AMBER · drug metabolism

Introduction

Biomolecular simulation is playing an increasingly important role as an adjunct to experiment in understanding the structural and dynamic factors involved in many biological processes at the atomic level.^[1] With the ever increasing number of X-ray structures available for drug-metabolizing P450 enzymes, the development and refinement of structure-based computational tools for predicting metabolic pathways for lead compounds continues to be an area of intensive research.^[2] The identification of sites of metabolism and further, the identification of the structural factors that can contribute to competing P450-catalyzed reaction mechanisms are important goals of rational drug design. Competing P450 reaction mechanisms can lead either to the formation of safe products for excretion or to the formation of unstable products that can lead to P450-related toxicities.^[3] Thus, increasing the predictive power of computational tools to improve the accuracy of “structural alerts” for identifying compounds that can form reactive intermediates can have a very direct impact on overall drug safety and human health.^[2,4]

The diversity in P450 substrate selectivity, sites of metabolism, and reaction mechanism is in part due to the large number of different subfamilies of hepatic P450s that are involved in drug metabolism. Each P450 subfamily exhibits unique functional characteristics in the metabolism of a variety of endogenous and exogenous compounds.^[5] The ever growing number of experimentally derived structures of hepatic P450 clearly demonstrates that structure–function relationships can be

established for these subfamilies. Even though hepatic P450 structures share a common fold,^[5] the subclasses of hepatic P450 enzymes vary significantly in size and architecture. Furthermore, many of the experimentally derived structures of hepatic P450s both with and without substrates and ligands suggest significantly varying degrees of the plasticity of the active site between the different subclasses.^[6]


Another major contributor to the diversity of reaction mechanisms catalyzed by P450 is the highly reactive nature of the catalytic heme oxyferryl species. Heme prosthetic groups, present in the P450 active site, are specialized protein cofactors that incorporate porphyrin rings interacting with various metals, ligands, and protein side chains. The unique chemical

[a] K. Shahrokh, G. S. Yost
Department of Pharmacology and Toxicology, University of Utah, Salt Lake City, Utah 84112

[b] A. Orendt
Center for High Performance Computing, University of Utah, Salt Lake City, Utah 84112

[c] T. E. Cheatham, III
Department of Medicinal Chemistry, University of Utah, Salt Lake City, Utah 84112
tec3@utah.edu

Contract/grant sponsor: NIH; Contract/grant number: GM079383 (TEC3); Contract/grant sponsor: NIH; Contract/grant number: GM0742249 (G. Y.).

 Additional Supporting Information may be found in the online version of this article.

structure of the heme facilitates a wide chemistry including—but not limited to—oxidation, dehydrogenation, and oxygen transport.^[7] Pioneering work in the crystallization of bacterial cytochrome P450, specifically P450cam,^[8] and the application of quantum mechanics (QM)-based methods to the heme has structurally identified and electronically characterized a number of the heme species populated during the P450 catalytic cycle.^[9] Moreover, although this work has clearly shown that that gas-phase density functional theory (DFT) methods can provide reasonable results in agreement with experiment depending on the model system used,^[10] and that only with the inclusion of the protein environment, from structures derived from experiment, combined with mixed quantum and molecular mechanics (QM/MM) simulations it is possible to accurately model the electronic state, and specifically the spin densities that are critical for understanding the contribution of structural factors during catalysis to these open-shell species to within experimental accuracy.^[11] For an extensive review of the applications of different computational methods and utility of different model systems, see these excellent reviews.^[10(a),11]

To better characterize and screen putative drug candidates, it is useful to understand how the drugs interact with P450s. Unfortunately, the acquisition of P450-bound X-ray structures for each new substrate is potentially very time consuming and not always possible. Thus, obtaining accurate crystal structures to facilitate the development of accurate QM/MM parameters for the major classes of hepatic P450 across the complete catalytic cycle, as previously performed for bacterial P450cam, remains a daunting task. Even given sufficient data, the use of QM/MM methods remains currently very computationally expensive to meet the demand for routine *in silico* structure-based screening of large libraries of compounds against multiple P450 structures. MM-based force field methods are computationally much less expensive than QM/MM methods and allow for adequate sampling of diverse chemical libraries of substrates and enzyme structures. With the use of molecular docking and scoring schemes, these MM-based methods can potentially identify the site of, and reaction mechanism of, drug metabolism. However, despite the availability of increasing numbers of P450 structures, the predictions of metabolic pathways for novel agents using MM-based computational tools and crystal structures have frequently been incorrect.^[3,12] Although the modeling of enzyme–substrate interactions and substrate metabolism are complex processes involving additional factors such as substrate ingress, binding, catalysis, and product release and egress, one obstacle for these methods has been the absence of a consistent set of MM parameters for the heme species during the P450 catalytic cycle.

Even though the electronic changes to the heme within the P450 active site during ligand binding and substrate turnover have been extensively experimentally studied, the development of a consistent set of MM-based parameters for key species of the heme during the P450 catalytic cycle faces significant challenges due to open-shell states of this iron-containing prosthetic group. Earlier, simple and less computationally expensive approaches for the development of MM

parameters for the resting heme state of the P450 catalytic cycle and also for the resting heme in cytochrome *c*^[13] have been attempted using AM-1, Hartree-Fock (HF), and DFT levels of theory with a number of different basis sets.^[14] Some of the limitations encountered with these models were the treatment of the heme side chains, the assignment of simple, but inaccurate, Mulliken charges, and the use of only a very small set of different heme states. Even with more accurate and consistent methods, it must be acknowledged that the development of a generally applicable heme MM model faces limitations because we cannot fully account for the entire set of potential influences on the heme structure and properties. Although we can control direct or covalent influences on the heme by developing separate models, that include additional bonded atoms, we cannot easily include indirect influences. These indirect influences include changes in the heme properties caused by the protonation states of nearby active site residues, changes in active site geometries or heme environment after substrate binding, and changes in active site hydration. Despite this inherent limitation of pairwise nonadditive MM models, such residue-based parameterizations have proven extremely useful to understand biomolecular structure, dynamics, and interactions.^[15] Specifically in the case of metalloproteins, MM parameters were developed and successfully used to understand dynamic processes for iron-containing systems.^[16] The bonded plus electrostatics model of a metal ion allows for the incorporation of the metal ion into MM force fields, such as AMBER and is based on the method first developed for zinc metalloproteins.^[17]

Recent work from our lab has demonstrated that the assignment of more reasonable atomic partial charges for the resting ferric state of the heme greatly improves docking studies for raloxifene to the CYP3A4 P450 enzyme.^[18] Further, the results facilitated the identification of key residues involved in shunting between two distinct P450 metabolic pathways.^[18] Raloxifene is an anticancer therapeutic with a significant number of degrees of structural rotational freedom and heteroatoms capable of polar interactions. CYP3A4 is the most biomedically relevant subfamily of hepatic P450 studied, as CYP3A4 makes the largest contribution to overall P450-mediated hepatic drug metabolism.^[9] CYP3A4 enzymes are able to perform this critical role because of promiscuity in substrate selectivity due to the size and flexibility of the active site.^[19] These characteristics provide a very significant challenge to structure-based methods for the prediction of metabolism. The incorporation of a preliminary set of atomic partial charges for the penta-coordinate ferric state to the heme in the X-ray structure of CYP3A4 (PDB ID: 1W0E)^[20] developed with a gas-phase model of an extended heme structure with a X-ray structure vastly improved the modeling of raloxifene within the CYP3A4 active site^[18] over models lacking any charge assignment to the heme.

Here, we describe the development of a consistent, transferable, and robust set of MM parameters for different heme states of the P450 catalytic cycle using two distinct model systems of the heme, specifically a minimal “truncated” heme (T-HM) compared to the full-heme model (F-HM). Because of

computational cost, we have investigated only four commonly accepted heme states in the P450 catalytic cycle, as schematically shown in Figure 1, with the two separate model systems

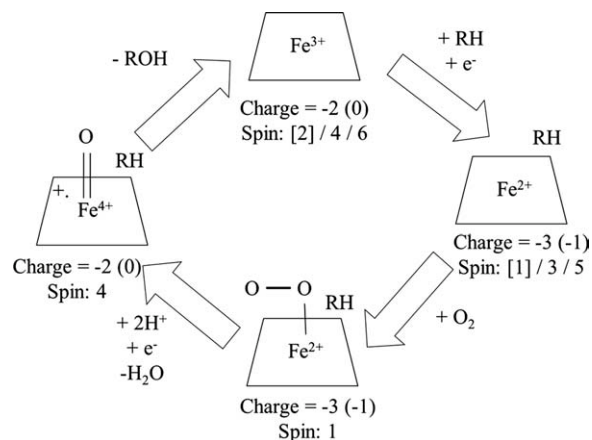


Figure 1. Schematic for describing the P450 catalytic cycle. Charge states for the extended F-HM are shown without parentheses. Charge values used for the T-HM are shown in parentheses. Spin values used for both T- and F-HM are shown without parentheses. Spin values only used for T-HM calculations are shown in brackets.

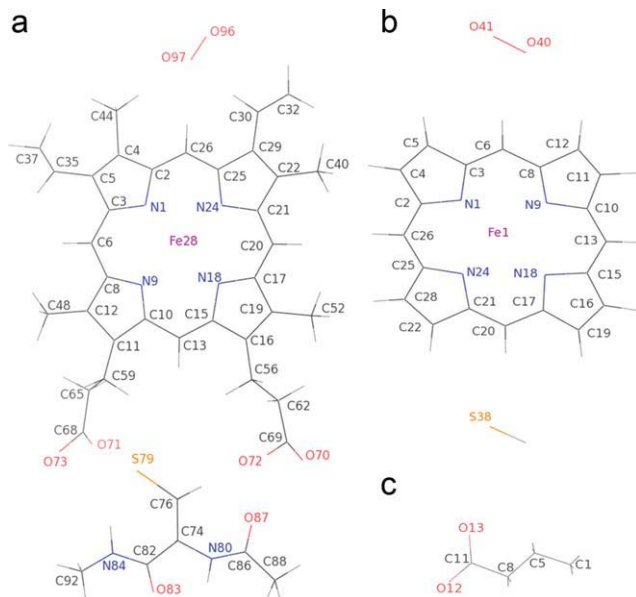


Figure 2. Atomic numbering scheme for a) the extended F-HM that includes a cysteine dipeptide (F-HM), b) the T-HM with a thiolate ligand, and c) propionic acid models. Hydrogen labels have been omitted for clarity.

illustrated in Figure 2. The goal was to identify the most appropriate, or most consistent, set of heme parameters for MM simulations of enzyme–substrate interactions. Our results indicate that the use of the RED-III-derived atomic partial charges^[21] from the gas-phase F-HM, despite some charge-transfer artifacts due to the absence of full shielding of the propionates in these gas-phase-based calculations, improves the accuracy of cytochrome P450 enzymes/substrate-docked

predictions for raloxifene by CYP3A4 in docking experiments with AutoDock3 (Scripps Research Institute, La Jolla, CA)^[22] using a distance-based scoring. Furthermore, docking with the new heme parameters and heme states suggests a strong influence of the choice of heme state on the docking results. Specifically, the nature and population of the binding mode changes depend on the applied heme state model. MD simulations on the heme molecules in solution suggest that the distinctive heme structure is maintained and motion about rotatable groups is not inhibited.

Methods

Calculations

Calculations on the heme models were based on two different model systems: the truncated porphine/thiolate (T-HM) and a full protoporphyrin macrocycle model with all its substituents including propionate side chains and a bound cysteine dipeptide ligand ligated to the iron (F-HM) as shown in Figure 2. The calculations were performed in a buildup manner from less (T-HM) to more complicated (F-HM). After optimization at the resting ferric high-spin state, interacting ligands (oxygen, etc.) and different iron spin states were added to represent the various stages of the catalytic cycle (Fig. 1). The final geometries were compared to experimental structures where known, and where unknown, careful consideration of homologous theoretical structures was given to eliminate potential errors and identify limitations.

A chief concern centered on the treatment and influence of the macrocycle substituents on the results, especially the propionate groups and the ligated proximal cysteine dipeptide. The model systems chosen were investigated in a manner consistent with QM-based MM parameter development approaches used with the current AMBER ff9X and ff10 force fields.^[23] For charge derivation, this typically involves HF optimization of one or more geometries of the residue of interest using the overly polarized 6-31G* Pople-type basis set followed by restrained electrostatic potential (RESP) charge fitting.^[24] To avoid influences due to the initial molecule orientation, the automated RED-III procedures were used.^[21] The standard HF/6-31G* level of theory used for small organic molecules in AMBER force field development is not applicable to the iron–protoporphyrin heme model. Instead, DFT hybrid functionals such as B3LYP with LACVP were applied, as they have been previously shown to produce sufficiently accurate results for geometries at a reasonable computational cost.^[10(a),11]

Geometry optimizations

Calculations were begun with a porphyrin ring at the HF/3-21//B3LYP/6-31G* level with the imposition of C_{4v} symmetry constraints. The iron atom was then added along with the thiolate ligand and symmetry constraints were removed. Optimizations were performed at the B3LYP/LACVP (with 6-31G*^[25] on the organic atoms and the LANL2DZ^[26] basis set with frozen core electrons, as implemented in Gaussian 03) level for the high-spin ferric species. This species was then used to sequentially model the changes in the P450 catalytic cycle and the spin states of interest.

The representative species for the heme were selected based on the P450 catalytic cycle. This cycle begins with a

penta-coordinate ferric species (IC*), a species that is then reduced to a penta-coordinate ferrous species (OUS*), which then binds molecular oxygen. This dioxygen-bound species (DIOXY) is then reduced further through the addition of a proton and electron leading to expected cleavage of the oxygen–oxygen bond to form the catalytically active oxyferryl species, compound I (CPDI). Many of these species can exist in a low-, medium-, or high-spin state. Because of the reduced computational costs for the smaller truncated heme models, a number of these spin states were investigated for the penta-coordinate truncated heme models. The dioxygen-bound species is a closed shell singlet. For CPDI, MM parameters were only derived for the high-spin state, as it has been shown that the low-spin doublet state requires significantly more computationally demanding multireference configuration interaction^[27] or CASSCF/CASPT2^[28] methods to describe accurately. Additionally, some ambiguity remains as to whether this low-spin state is degenerate or lower in energy than the quartet state.^[11]

The optimized T-HM penta-coordinate ferric high-spin sextet state (IC6) geometry was used to develop the F-HM model. The substituent groups were then added to the heme, including the propionate groups and a more extended model of the bound sulfur, through the inclusion of a cysteine dipeptide (i.e., a model *N*-methyl and acetyl blocked cysteine residue consistent with the AMBER force field^[23]). The dihedral angles on the propionates were restricted to values consistent with the available set of published X-ray structures of mammalian P450. Restraints were applied because the positioning of the propionates in most of these experimental structures differs significantly from those observed in the P450cam structures. The values were calculated using the heme in the A chain of the following P450 structures for CYP3A4 (PDB ID: 1W0E, 1W0F, 1W0G, 1TQN, 2JOD, and 2VOM),^[19,20,29] CYP2D6 (PDB ID: 2F9Q),^[30] and CYP2B4 (PDB ID: 1PO5, 2BDM, and 1SUO).^[6(b,e),31] These structures were chosen, as they show the plasticity of the P450 active site in the absence and presence of a number of structurally diverse ligands. The average dihedral angle defined by C11–C59–C65–C68 was thus frozen at 176.00° and the dihedral defined by C16–C56–C62–C69 was frozen at 174.50° (see Fig. 2 for atomic numbering). Optimizations were performed with the cysteine dipeptide in a number of different starting positions relative to the propionate side chains at the B3LYP/LACVP level for the ferric high-spin species. Because of the vagaries of Gaussian 03 implementation, only one structure converged. This structure was then used as the starting point for the other species in the P450 catalytic cycle.

The stability of the minima was confirmed with normal mode analysis and the stability of the wave functions was confirmed with addition of the “stable = opt” keyword.^[32] Spin contamination was also verified to be in agreement with values in the literature^[33] for the open-shell species (see Supporting Information Table 1).

For the T-HM, thiolate–porphyrine heme species, “tight” convergence criteria were used for another round of optimizations and the stability of the wave function and minima were confirmed again. Differences in energies were compared between structures optimized with these different criteria, differences of ~0.1 kcal/mol were observed for the smaller, more symmetric heme model. For the F-HM, the optimizations were carried out to the default G03 criteria.

Charge fitting

RESP charges were derived using RED-III^[21] for the optimized geometries for all F-HM and T-HM species at the B3LYP/LACVP level. Intramolecular charge restrictions and constraints were applied to the blocked cysteine residue (consistent with the AMBER ff9X force field)^[23] and molecular electrostatic potential computations were performed on multiple orientations of a single conformation of each system, based on atoms N1, Fe28, and S79 (see Fig. 2 for atom nomenclature).

Docking

AutoDock3.05 and AutoDockTools (Scripps Research Institute)^[22] were used for all docking studies. Docking was performed consistent with our previous work.^[34]

Using UCSF Chimera, the heme coordinates from the B3LYP/LACVP optimized structure for the F-HM and the PDB ID: 1W0E^[20] coordinates were matched at the pyrrole nitrogen atoms of the heme to translate the F-HM and 1W0E^[20] coordinates into a common frame of reference and saved. The coordinates for the heme from the X-ray study were removed and the coordinates for each charge and spin state of the F-HM were added to the prepped and matched 1W0E. In this manner, 1W0E files were generated to represent CYP3A4 during different phases of the P450 catalytic cycle for the ferric high and intermediate, O₂-bound, and oxyferryl states. The B3LYP/LACVP RED-III-derived partial charges were then assigned to the atoms in each modified 1W0E template. The van der Waals nonbonded parameters for these values were taken from the AMBER ff99SB^[35] and GAFF force fields,^[36] so that the heme force field would be transferable between the AMBER ff9X force fields. The Fe values used in the previous docking study^[34] and recommended in the AutoDock package for iron atoms were assigned: 1.3 Å radii (Rii) and well-depth 0.01 (epsii) and solvent parameters assigned with AutoDockTools (Scripps Research Institute).

AutoGrid3.05 (Scripps Research Institute)^[22] was used to define the active site space of CYP3A4. A grid box was assigned to encompass the whole active site of CYP3A4 and to closely match the one used in the previous study^[34] within the new frame of reference of the matched system. Within this frame of reference, the grid box was recentered at –4.225, 0.944, and 5.434 displacement from the default center assigned by AutoGrid3.05 with the dimensions of 68 × 52 × 62 and a resolution of 0.375 Å in each dimension.

In this study, we scored all docking results with less stringent criteria for identification of the site of metabolism for the penta-coordinate state from that commonly used with X-ray coordinates for the heme. This was done because in the QM-based optimizations the Fe atom is located below the plane of the heme in the high-spin state, whereas in X-ray structure assignments of mammalian P450, the Fe atom is determined to lie approximately in the plane of the heme. Thus, for the penta-coordinate states, we permitted a distance of <6.0 Å from the Fe to the known site of metabolism as a positive. For the hexa-coordinate states, we increased the stringency of this criteria for the catalytic oxygen present and scored a positive as a heavy atom at a known site of metabolism of the substrate lying within <4.5 Å from the proximal/activated oxygen atom.

Molecular dynamics

AMBER parameter development. GAFF atom types were used to build the heme frcmod files containing the necessary force field parameters and atom types were selected by similarity of the bond length and angle parameters in the force field that were closest to the QM structures. Specifically, cc and cd atom types were used to define C α and C β ; nc and nd atom types were used to define the pyrrole nitrogen atoms; cg was used to define C_{meso} based on bond length and was modified with existing cc–cc–cf, cd–cd–ha, and ce–ce–cf for missing angle parameters: cc/cd–cc/cd–cg, cc/cd–cg–ha, and cd–cg–cc, respectively. For the hexa-coordinate states, oa was used for the proximal and ob, for the distal axial oxygen atom. The propionates were built with c3, h1, c, and o atom types, and a CYP residue type was introduced for the axial proximal cysteinate ligand. This CYP residue was a standard CYS residue with the thiolate H atom removed, and RED-III atomic partial charges assigned. The CYP and HEM residues had the Fe–S and Fe–N bonds added manually in tleap.

QM-based coordinate scans Fe–S, Fe–N, and Fe–O bond and Fe–N–C, N–Fe–N, S–Fe–N, N–Fe–O, and Fe–O–O force constant parameters of the T-HM were developed by fitting B3LYP/LACVP scan results. Scans of bond coordinates were performed in 0.05 Å steps and scans of angle coordinates were performed in 2° increments. However, for certain coordinates for different heme states, these steps had to be reduced or not performed for the same number of steps from the equilibrium distance/angle. This we believe was caused by the asymmetry of the model system for certain heme states. The F-HM compound I model due to its size was only used for one scan along an increasing Fe–S–C angle coordinate, as the T-HM did not possess an equivalent coordinate.

$$V_{\text{MMbond}} = k(r - r_0) \quad (1)$$

$$V_{\text{MMangle}} = k(\theta - \theta_0) \quad (2)$$

Energies of QM scans of the T-HM along key bond and angle coordinates and the MM potential energy (V_{MM}) were fit using either eq. (1) or (2), and the force constant (k) was fit to minimize the sum of the square of the residual differences between the QM and MM energies. The barrier height for torsion terms involving the Fe that were absent from GAFF and ff99SB force fields was assigned as has been done previously for transition metals using the bonded model for metal ions.^[37] An alternative approach that could be used is the MTK++^[38] from AmberTools which creates a model structures of the metal ion's first coordination sphere from experimentally determined structures, and can then be used to determine charges, and bond and angle constants with *ab initio* calculations.

Implicit solvent simulations. To loosely demonstrate the stability of the various heme states, 10 ns-generalized Born implicit solvent^[39] MD simulations of the various heme states with parameters derived from the DFT B3LYP optimized structures with AMBER-compatible RESP charges were performed. Langevin temperature coupling with a 1.0/ps coupling time at 300 K was applied with a 1-fs time step and 16-Å nonbonded interaction cutoff with AMBER.^[40]

Explicit solvent simulations. All the MD simulations were carried out using AMBER 11^[40] with the ff99SBildn force field.^[41] Additional Na⁺ and Cl[−] ions^[42] were added to net-neutralize

each system, and then the model was solvated in an octahedral box of TIP3P^[43] water with an extension of at least 12 Å from each side. Periodic boundary conditions were applied with the particle mesh Ewald (PME) method^[44] to calculate the full electrostatic energy. The 10-Å nonbonded pair list was updated whenever any atom moved more than 0.5 Å since the last list update. SHAKE^[45] constraints were applied on the bonds involving hydrogen (tolerance: 10^{−5} Å), and the time step was set to 2 fs. A cutoff of 9 Å was applied to the Lennard-Jones and direct space electrostatic interactions with a uniform density approximation included to correct for the long range van der Waals.

The system was first minimized without electrostatics for 300 steps, then with a restraint of 25 kcal/(mol Å²) applied on the C α (500 steepest descent cycles followed by 500 conjugate gradient cycles). This minimization was followed by 100-ps MD simulation with 25 kcal/(mol Å²) positional restraints applied on the backbone C α atoms, and the temperature was slowly increased from 0 to 300 K. The temperature was regulated using Langevin dynamics^[46] with a collision frequency of 0.2/ps. Then, followed by five cycles of equivalent 500 steps of steepest descent cycles followed by 500 steps of conjugate gradient minimization, and 50-ps equilibrations with a restraint force constant of 5, 4, 3, 2, and 1 kcal/(mol Å²), followed by final 500 ps equilibration with a force constant of 0.5 kcal/(mol Å²) to equilibrate the density. The above equilibration steps were all carried out at constant volume, 300 K (Langevin dynamics, collision frequency: 1/ps) with 1-atm constant pressure, and this was followed by at least 25-ns production MD simulation for each model under equivalent conditions. All calculations were performed on the SGI Pople machine at the Pittsburgh Supercomputing Center and also at the University of Utah's Center for High Performance Computing various clusters.

Results

Geometry optimizations

As an initial test of the calculations, we compared the geometries obtained for the smaller T-HM with previously published values for a larger set of heme states due to the lower computational cost than for the more extended F-HM. Energy optimizations of the T-HM structure model were performed at the B3LYP^[47]/LACVP level for high-, intermediate-, and low-spin states of two penta-coordinated T-HM structures, the ferric (Fe³⁺) and ferrous (Fe²⁺). Key geometries such as the Fe–S distances, average Fe–N distances, and Fe out of N-plane values are in good quantitative agreement with previously published work,^[48] as shown in Figure 3. General trends that can be observed for these states are that the maximum Fe–S bond length is observed for the intermediate-spin state for both of these penta-coordinate states, whereas the average Fe–N distance and deflection of Fe from the N-plane both increase from low to high spin. The differences in energies of the different ferric penta-coordinate spin states were all within 1.5 kcal/mol, whereas for the ferrous species, the intermediate triplet spin state was found to be 9.2 kcal/mol higher, with the low-spin singlet state 14.0 kcal/mol higher than the high-spin

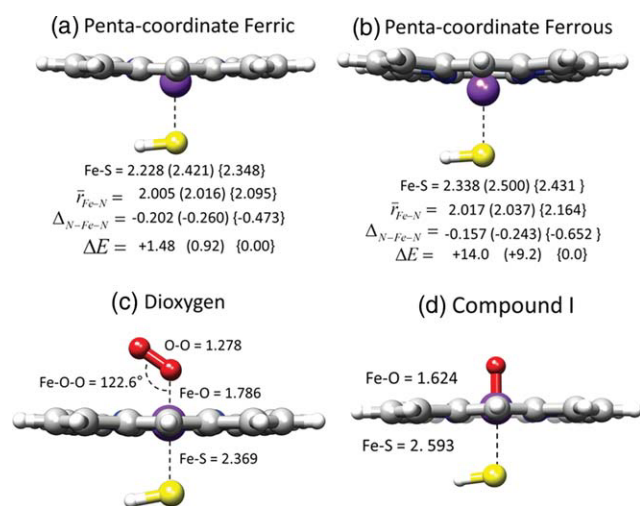


Figure 3. B3LYP/LACVP optimized geometry parameters of key heme species during P450-catalytic cycle (a–d) using the T-HM. Differences in parameters for different spin states are indicated as follows: low spin (medium/high spin). Calculated distances are in Angstroms. The relative difference in the sum of electronic and thermal energies is indicated for each species for which calculations were made using different spin states (kcal/mol). Adapted from Meunier et al. (2004).

quintuplet state again in good agreement with that in the literature,^[48] as shown in Figures 3a and 3b.

For the T-HM hexa-coordinate states, the dioxygen-bound ferrous form of the T-HM, calculations were performed with the singlet state^[48] and for compound I for the high-spin quar-

tet state of compound I. Key distances and angles for the hexa-coordinate T-HM are shown in Figures 3c and 3d and are in good quantitative agreement with previously published work.^[48]

Geometry optimizations at the same level of theory were then applied to the more extended model system. Comparison of the results for these calculations with previously published work is more difficult because the treatment of the substituents and the contribution to the free energy of the system by them can vary significantly due to the increased number of degrees of freedom of the extended F-HM system compared to the truncated T-HM.^[11] Furthermore, the computational cost for the treatment of the larger system is much greater than that for the T-HM. Thus, the number of heme P450 catalytic cycle states that were modeled was limited to what could be performed at a reasonable computational expense for our largest model system, F-HM. For the F-HM penta-coordinate ferric and ferrous species, geometry optimizations were performed for the high- and intermediate-spin states. As we could not compare these differences directly to other published work, we compared observed trends in key geometries for the T-HM and also to that of published X-ray structure for P450cam. The heme in P450cam has the same substituents as observed in human hepatic P450, but X-ray data show a different orientation of one of the propionates. Key distances and angles from the gas-phase geometry optimizations for the F-HM penta-coordinate states are shown in Figure 4 alongside values from P450cam X-ray structures: PDB ID: 1DZ4 (ferric) and 1DZ6 (ferrous) for reference in blue. The trends in the geometries of the F-HM system are in agreement with those

in the literature for similar more extended systems.

In brief, for the F-HM ferric intermediate- and high-spin states, the Fe–S distance is longer than that for the T-HM and the X-ray structure, the average Fe–N distances are shorter than the T-HM and in agreement with the X-ray structure value, deflection of the Fe from the N-plane is less for the intermediate spin and greater for the high-spin state than the T-HM while the X-ray value is within the limits of the values for both models. For the ferrous intermediate- and high-spin states, the Fe–S distance is longer but in closer agreement with the T-HM than the ferric state, and larger than that of the X-ray structure; the average Fe–N distances are shorter than the T-HM and in closer agreement with the X-ray structure value; the deflection of the Fe atom from the N-plane is less for the intermediate-

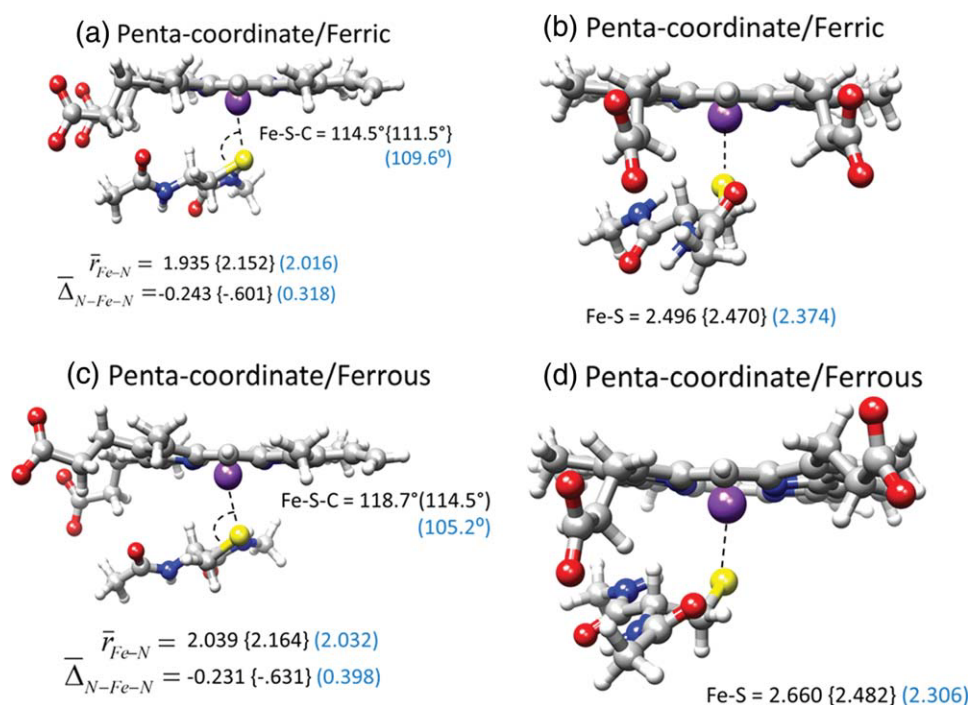


Figure 4. B3LYP/LACVP optimized geometry parameters of key heme species during P450-catalytic cycle (a–d) using the extended F-HM. Differences in parameters for different spin states are indicated as follows: intermediate spin (measurements from X-ray data shown in [blue a and b, PDB code: 1DZ4 and c and d, PDB code: 1DZ6]). Calculated distances are in Angstroms. Adapted from Meunier et al. (2004).

and high-spin states than the T-HM, while the X-ray value is in agreement and within the limits of the values from both models.

For the hexa-coordinate DIOXY state, we observe a longer Fe—S and Fe—O distances for the F-HM than the T-HM and the X-ray structure. In general, the values for the T-HM are closer to the X-ray structure as observed previously. The Fe—S—C angle more acute in the X-ray structures than in the F-HM-based calculations. Also, for the DIOXY-bound F-HM structure model, the Fe—O and O—O bond lengths are longer than what is observed in the X-ray structure PDB ID: 1DZ8.^[8] The Fe—O—O bond angle is more acute in both the T-HM and F-HM than observed in the X-ray structure, whereas the Fe—S—C is more obtuse. The average Fe—N distance is very similar in both models and close to that of the X-ray structure.

For the F-HM high-spin quartet CPDI structure, key distances and geometries are provided and are compared to the values from the P450cam structure PDB ID: 1DZ9^[8] in Figures 5c and 5d. Despite the presence of some ambiguity of the electron density data,^[49] this structure provides the only available X-ray structural data to date for CPDI in a P450 enzyme and the values from our calculations are in good agreement with key geometries inferred from this experimental data,^[8,50] for F-HM. With the T-HM structure, Fe—O, R_{avg} Fe—N distances are also in good agreement with experiment. Both structure models underestimate the observed Fe—S bond length although the T-HM value is closer. This is not surprising because the models lack context of the protein which may alter bond lengths. Also, the Fe—S—C angles are slightly more acute in the F-HM than are observed in experiment. Both of our gas-phase model structures predict that in the absence of active site constraints the Fe will be slightly above the N-plane, which agrees with related calculations,^[11] noting that this is contrary to what is found in the X-ray structure of P450cam.^[8]

Charge and spin

As the results of our geometry optimizations were in reasonable agreement with the results of other published calculations and X-ray structures, we

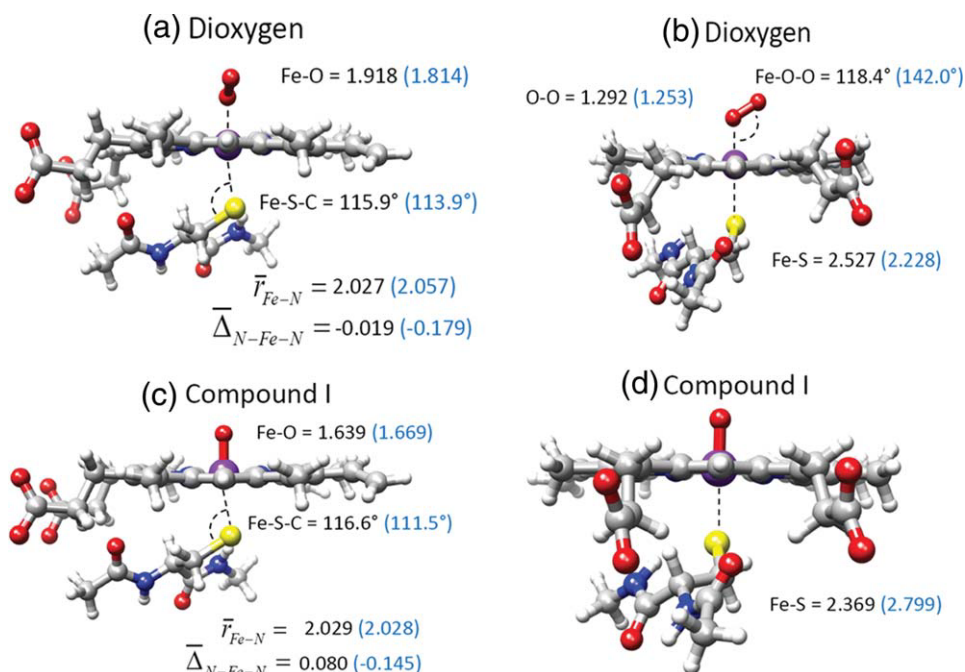


Figure 5. B3LYP/LACVP optimized geometry parameters of hexa-coordinate heme species during P450-catalytic cycle (a–d) with extended F-HM. Distances are in Angstroms. Measurements from X-ray data shown in blue (a and b, PDB code: 1DZ8 and c and d, PDB code: 1DZ9.). Adapted from Meunier et al. (2004). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

then used these geometries to derive atomic partial charges and compare the atomic partial charge values for these open-shell states for the T-HM and F-HM. Atomic partial charges were calculated with the RED-III^[21] program that provides for multiple conformations and orientations in the fit, as described in greater detail in the ‘Methods’ section. A few general trends for the RED-III calculated atomic partial charges for the two-different model systems are given in Table 1. For all the heme states, a greater asymmetry is observed in the distribution of atomic partial charges for the F-HM structure models than is observed with the T-HM, due to the greater asymmetry of the F-HM.

For the penta-coordinate, high- and intermediate-spin states, similar positive charges are observed on the iron atom, along

Table 1. RED-III atomic partial charges at nonbackbone heteroatoms for (A) the extended F-HM and (B) the T-HM systems.

	CPDI	DIOXY	IC6	IC4	OUS5	OUS3
(A) F-HM						
N (sum)	0.0676	1.1168	−0.5112	−0.8098	−0.3531	0.0746
Fe1	0.262	0.0289	0.2492	0.048	0.1462	0.0297
S79	−0.4381	−0.1587	−0.4602	−0.3723	−0.5562	−0.6037
O71 and O73 (avg)	−0.65515	−0.71915	−0.5791	−0.6017	−0.7117	−0.7181
O70 and O72 (avg)	−0.65075	−0.7168	−0.64265	−0.6851	−0.72285	−0.7212
O (proximal)	−0.3729	−0.1174				
O (distal)		−0.1945				
(B) T-HM						
N (sum)	0.4311	0.5947	−0.5441	−0.0863	−1.4568	−0.3553
Fe1	0.033	0.131	0.2769	0.047	0.6999	0.0592
S38	−0.3835	−0.6807	−0.379	−0.424	−0.7309	−0.6977
O (proximal)	−0.3422	−0.2078				
O (distal)		−0.2295				

with similar net negative charges on the nitrogen atoms with both T-HM and F-HM structure models. For the sulfur atom, a greater negative charge is calculated for the F-HM structure model for the sextet, and a smaller negative charge is calculated for the quartet state compared to the T-HM. For the reduced penta-coordinate ferrous quintet and triplet states, greater charges are observed for the sulfur and iron atoms, and net negative charges are observed for the nitrogen atoms comparing the T-HM to the F-HM. The F-HM structure model predicts a net positive charge on the nitrogen atoms for the ferrous triplet state (0.0746) as opposed to the T-HM that predicts a net negative charge (−0.3553).

For the hexa-coordinate states, the dioxygen-bound structure is a singlet and therefore this F-HM structure model is in principle a good “control” to look out for potential changes in the atomic partial charges due to shifts in unpaired electron-spin density due to the absence of shielding of the propionates by basic active site residues in these gas-phase calculations. However, with this state, we observe significant differences between the partial charges calculated with the different structural models. For the nonbackbone heteroatoms, we observe much greater net atomic partial charge on the sulfur, iron, and molecular-oxygen atoms with the T-HM than the F-HM. The nitrogen atoms exhibit greater asymmetry in their charge distribution while bearing positive charges in both models with the sum of these charges being almost double for the F-HM (1.1168) than the T-HM (0.5947).

For compound I, the F-HM predicts a greater positive charge on the iron atom, slightly greater negative charges on the sulfur and axial oxygen atoms, and a much lower net positive charge on the nitrogen atoms than the T-HM.

It has been previously shown that for gas-phase calculations of extended heme models, charge-transfer artifacts can be introduced due the absence of full shielding of the propionates by active site basic residues, and thus spin-density delocalization to the oxygen atoms of the propionate groups.^[51] As we observed significant differences in atomic partial charges for T-HM and F-HM systems for each equivalent state of the catalytic cycle, we more closely investigated if the differences were due in major part to electron-spin delocalization to the propionates. As the DIOXY state is a closed shell system it should not suffer from any charge-transfer artifacts due to unpaired spin delocalization to the propionate oxygen atoms and should so provide a good reference for the values calculated for the other open-shell heme states. The ferrous open-shell penta-coordinate high- and intermediate-spin states produced similar negative atomic partial charge values for the propionate oxygen atoms as observed in the DIOXY state. However, the ferric ic4, ic6, and the CPDI states show differences in the atomic partial charge values compared with the DIOXY and ferrous states (Table 1).

Mulliken atomic spin densities

To better understand the delocalization of spin densities in these gas-phase calculations for the F-HM, we compared these values at key atoms to the T-HM. The T-HM cannot suffer from spin delocalization to propionates and has been observed to

be a better model system for determining quantum effects such as spin densities in gas-phase calculations.^[11] As given in Table 2, the sum of the Mulliken atomic spin densities is

Table 2. Mulliken atomic spin densities for the T-HM and the extended F-HM systems from the B3LYP/LACVP geometry optimizations.

	IC6	IC4	OUS5	OUS3	CPDI
F-HM					
Fe	3.662	2.453	3.661	1.969	1.187
S	0.116	0.393	0.108	0.108	−0.019
N (sum)	0.176	−0.025	0.171	−0.063	−0.007
Porph (−N)	−0.017	0.025	0.050	−0.018	−0.128
Cys (−S)	0.012	0.014	0.011	0.005	0.003
Prop O (sum)	1.052	0.140	0.001	−0.002	1.062
O					0.901
T-HM					
Fe	3.904	2.465	3.645	1.911	1.075
S	0.474	0.485	0.141	0.199	0.554
N (sum)	0.531	−0.013	0.166	−0.023	0.384
Porph (−N)	0.094	0.070	0.046	−0.080	0.059
O					0.943

N (sum) is the sum of the Mulliken spin densities on all the heme nitrogen atoms, Porph (−N) is the sum of the spin densities on the protoporphyrin (F-HM) or porphyrin (T-HM) minus the N (sum) value, Cys (−S) is the sum of the spin densities on the blocked cysteine ligand minus the spin density on the sulfur atom (F-HM), Prop O (sum) is the sum of the spin densities on the propionate oxygen atoms.

increased on propionate oxygen atoms for the open-shell species that show a difference in partial charges, and that spin density is delocalized on the propionate oxygen atoms more for the ic4, ic6, and CPDI than for ferrous states as previously observed.^[51] Closer examination of the spin densities for the penta-coordinate states from the optimization calculations reveals very similar values for the iron atom predicted for the all the penta-coordinate states. For the ferric high-spin, we observe less spin density on the sulfur and greater spin density on the porphyrin for the F-HM compared with the T-HM. The ferric states show greater spin density on the nitrogen atoms and porphyrin for the T-HM, whereas for the ferrous high spin shows similar values between the two structure models and the ferrous intermediate values differ between the two structure models. For CPDI, we observe similar spin on the iron and oxygen with the two models but different spin on the sulfur, nitrogen atoms, and porphyrin between the two models.

As we calculated the atomic partial charges with RED-III, and this method uses multiple reorientations of the molecule to “average” out the charges to make them more reproducible, we examined the Mulliken spin densities for each of these calculations. We found that there was an orientation dependence for the Mulliken atomic spin calculations (see Supporting Information Table 2 for the averages and standard deviation of Mulliken charges and spin for T-HM and F-HM CPDI). Therefore, we calculated the averages for Mulliken atomic spin densities for all reorientations made for each heme states during the RED-III charge fitting for both T-HM and F-HM and these are presented in Table 3. Comparison of the agreement between the average spin-density values from six reorientations for each

Table 3. Average Mulliken atomic spin densities from RED-III reorientation RESP calculations for (A) extended F-HM and (B) the T-HM.

	IC6 Avg	IC6 Std Dev	IC4 Avg	IC4 Std Dev	OUS5 Avg	OUS5 Std Dev	OUS3 Avg	OUS3 Std Dev	CPDI Avg	CPDI Std Dev
(A) F-HM										
Fe	3.814942	0.068763	2.632201	0.000106	3.660980	0.000067	1.968843	0.000066	1.077921	0.005778
S	0.224528	0.050995	0.181283	0.000103	0.107627	0.000052	0.107839	0.000026	0.316469	0.005664
Cys (w/NME&ACE)-	0.014580	0.000227	0.018785	0.000014	0.010580	0.000016	0.005450	0.000007	0.010583	0.000351
N (sum)	0.344437	0.078408	0.104677	0.000044	0.170377	0.000019	-0.062941	0.000081	0.481340	0.015486
Porph (-Fe)-[N (sum)]	0.000416	0.028386	0.159447	0.000023	0.048458	0.000017	-0.016633	0.000033	0.105598	0.082849
Prop (-O)	-0.042027	0.001366	-0.021686	0.000014	0.001281	0.000016	-0.002306	0.000026	0.066800	0.000222
O-prop (sum)	0.643122	0.152473	-0.074708	0.000023	0.000697	0.000001	-0.001761	0.000003	0.104380	0.099501
O (Opdl)									0.915887	0.004121
Spin annihilation										
Before	8.830883	0.033076	3.982900	0.000155	6.008900	0.000000	2.033867	0.000052	3.969400	0.070334
After	8.750933	0.000408	3.755733	0.000052	6.000000	0.000000	2.000600	0.000000	3.764667	0.007101
(B) T-HM										
Fe	3.903714	0.000022	2.465124	0.000166	3.644619	0.000013	1.910712	0.000094	1.075067	0.000080
S	0.474990	0.000104	0.484387	0.000031	0.141229	0.000030	0.199167	0.000086	0.554511	0.000058
(S)H	-0.003301	0.000005	-0.006393	0.000001	0.002452	0.000003	-0.006957	0.000003	-0.015107	0.000003
N (sum)	0.530832	0.000113	-0.013365	0.000133	0.166183	0.000039	-0.022582	0.000170	0.383719	0.000107
Porph (-Fe)-[N (sum)]	0.093765	0.000005	0.070246	0.000037	0.045518	0.000004	-0.080337	0.000145	0.059097	0.000010
									0.942711	0.000051
Spin annihilation										
Before	8.766000	0.000000	3.790833	0.000103	6.008900	0.000000	2.056233	0.000052	3.791667	0.000052
After	8.750100	0.000000	3.750500	0.000000	6.000000	0.000000	2.001333	0.000052	3.750800	0.000000

state to that of the original optimization indicates a closer agreement when comparing the average values, for the IC6 and Cpdl states. However, the agreement between the F-HM and T-HM values for the ferric intermediate state appears to have worsened while differences between the ferrous penta coordinate remain relatively unchanged.

To identify the contribution of these differences in spin densities to the partial atomic charges was, we calculated the difference in charges and spin for key atoms and moieties of the heme molecule (see Table 4) and plotted the difference in charge versus the difference in spin between the T-HM and F-HM (T-HM value – F-HM value) for the iron, sulfur, and nitrogen atoms (Fig. 6) and fit these lines with a linear regression. No clear correlation is apparent between the differences of the RED-III charge and average Mulliken spin density, nor for Mulliken charges and Mulliken spin values from the optimizations, for any of these atoms.

For the propionate oxygen atoms, however, there does appear to be a decrease in negative charge from the average charge for the dioxy and ferrous states (average O71 and O73 value is -0.7163 and the average O70 and O72 value is -0.7203); this may be due in part to the delocalization of unpaired electron density on the propionate oxygen atoms. However, the effect of this charge transfer to the rest of the system is unclear. Changes in charge for the iron, sulfur, nitrogen atoms, and other key moieties vary to such a degree between the T-HM and F-HM systems and heme states, even for closed shell systems, that the contribution of this charge-transfer amount to the rest of the system is impossible to identify.

Molecular docking

As we could not observe a direct correlation between the spin-density delocalization and the atomic partial charges

calculated with RED-III, we tested the use of the charges empirically derived from the T-HM and the F-HM. Previous docking studies from our lab have shown that the incorporation of more accurate atomic partial charges for the heme with the x-ray structure of CYP3A4 (PDB ID: 1W0E)^[20] improved the docking results and identified specific residues involved in substrate binding that were verified experimentally.^[18] Specifically, we observed that with the better charge models, there was an increase in the number of conformations found through docking that are consistent with the observed metabolism of raloxifene. We further identified and experimentally confirmed a single residue (Phe215) within the active site that significantly contributed to competing metabolic pathways. Thus, to test the influence of the current set of B3LYP/LACVP optimized structure models and RED-III calculated RESP atomic partial charges for different heme states during the P450 catalytic cycle on the molecular docking results, we used the same CYP3A4 template modified with these new QM-based heme coordinates and charges for additional docking simulations. Here, dockings were performed with the F-HM penta-coordinate ferric high-spin [IC6 (F-HM)] and intermediate-spin [IC4 (F-HM)] geometries and RED-III-derived charges, F-HM penta-coordinate ferric high-spin [IC6 (Oda)] and intermediate-spin [IC4 (Oda)] geometries with previously published^[14(a)] charges for the ferric penta-coordinate state, F-HM hexa-coordinate molecular-oxygen-bound [DIOXY (F-HM)] geometries with RED-III-derived partial charges, F-HM hexa-coordinate catalytic compound I [CPDI (F-HM)] geometries with RED-III-derived partial charges, and T-HM hexa-coordinate catalytic compound I [CPDI (T-HM)] RED-III-derived atomic partial charges assigned to the F-HM-derived geometry modified with gas-phase propionic acid-derived charges at the B3LYP/6-31G* level. In this work, we scored with slightly less stringent criteria

Table 4. Differences in charge and spin for key atoms and moieties of the T-HM and extended F-HM.

	CPDI	DIOXY	IC6	IC4	OUS5	OUS3
Diff charge						
Fe	−0.2290	0.1021	0.0277	−0.0010	0.5537	0.0295
S	0.0546	−0.5220	0.0812	−0.0517	−0.1747	−0.0940
N (sum)	0.3635	−0.5221	−0.0329	0.7235	−1.1037	−0.4299
Propionate (Diff chg)						
O71 and O73 (each)	−0.0612	0.0028	−0.1372	−0.1146	−0.0046	0.0018
O70 and O72 (each)	−0.0695	−0.0035	−0.0776	−0.0352	0.0026	0.0009
Sum of diff for all O	−0.1307	−0.0007	−0.2149	−0.1498	−0.0021	0.0027
Propionate (charge) non-O	1.2519	1.2191	1.2539	1.1751	1.2376	1.2495
Porph (−Fe) − [N (sum)] (CHG)						
F-HM	−0.2110	−1.7144	−0.1459	0.3702	−0.5986	−0.8131
T-HM	0.1028	−0.7450	0.4774	0.3013	0.0998	0.3404
Diff	0.3138	0.9694	0.6233	−0.0689	0.6984	1.1535
CYP (CHG)						
Sidechain − (S)	0.1666	−0.1933	0.1722	0.2766	0.1079	0.0558
Sidechain + (S)	−0.2715	−0.352	−0.288	−0.0957	−0.4483	−0.5479
CYP	−0.3859	−0.4664	−0.4024	−0.2101	−0.5627	−0.6623
Diff avg spin						
Fe	−0.0029		0.0888	−0.1671	−0.0164	−0.0581
S	0.2380		0.2505	0.3031	0.0336	0.0913
Cys ((w/NME & ACE)-(S))	−0.0257		−0.0179	−0.0252	−0.0081	−0.0124
N (sum)	−0.0976		0.1864	−0.1180	−0.0042	0.0404
Porph (−Fe) − [N (sum)]	−0.0465		0.0933	−0.0892	−0.0029	−0.0637
Propionate (spin)						
Prop (−O)	0.0668		−0.0420	−0.0217	0.0013	−0.0023
O-prop (sum)	0.1044		0.6431	−0.0747	0.0007	−0.0018
Porph (−Fe) − [N (sum)] (spin)						
F-HM	0.1056		0.0004	0.1594	0.0485	−0.0166
T-HM	0.0591		0.0938	0.0702	0.0455	−0.0803
Diff	−0.0465		0.0933	−0.0892	−0.0029	−0.0637
CYP (spin)						
S	0.3165		0.2245	0.1813	0.1076	0.1078
CYP ((w/NME & ACE)-(S))	0.0106		0.0146	0.0188	0.0106	0.0055
CYP ((w/NME & ACE)-(S))	0.3271		0.2391	0.2001	0.1182	0.1133

for identification of the site of metabolism for all model systems and atomic partial charges, allowing a distance of <6.0 Å from the Fe to the substrate heavy atom for the ferric states and <4.5 Å from the proximal/activated oxygen atom for predicting the site of metabolism. We relaxed the criteria for the “resting” penta-coordinate state, as the QM calculations predict that the Fe atom is below the plane of the heme, whereas the

x-ray structure used in our previous work assigns that Fe atom within the plane of the heme. We also increased the stringency of the criteria, correspondingly, for the hexa-coordinate heme parameters with the catalytic oxygen present.

For the F-HM ferric high-spin state, our results indicate that the use of the new RED-III-derived atomic partial charges produces results consistent with our previous work (see Supporting

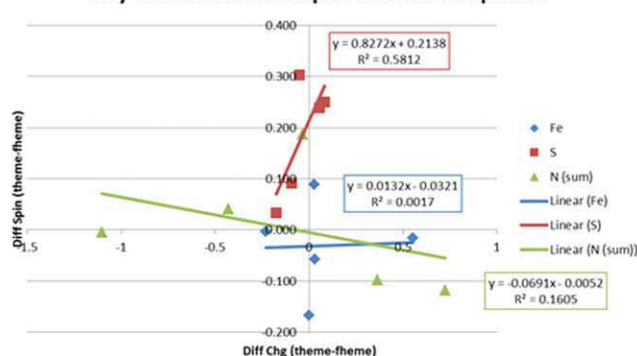
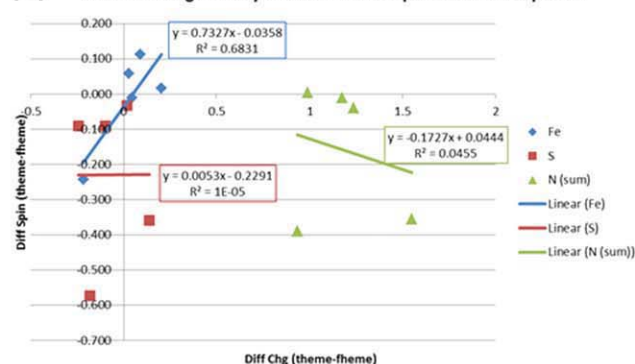
(a) Difference in spin versus difference in charge for key atoms of all five open shell heme species**(b)** Difference in Mulliken atomic spin densities versus difference in Mulliken charge for key atoms of all five open shell heme species

Figure 6. Differences in a) RED-III atomic partial charges versus average Mulliken atomic spin densities and b) Mulliken charges and atomic spin densities from geometry optimization calculations for the truncated heme (T-HM) and the full heme (F-HM) models. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Information Tables 3A–3C). Furthermore, the number of conformations that bind in a manner consistent with the observed metabolism is increased. Use of the ferric intermediate-spin parameters produced a greater number of large clusters containing only nonproductive binding states compared to the high-spin state, and this model also failed to produce any large clusters that predicted the observed hydroxylation of raloxifene (see Table 5 and Supporting Information Tables 3A–3C).

Docking results using the dioxygen-bound hexa-coordinate full heme geometries and RED-III-derived charges [DIOXY (F-HM)] also did not produce any significant population predicting the observed hydroxylation and dramatically reduced the number of nonproductive binding modes (see Table 5 and

Table 5. Percentage of binding modes of AutoDock3.0 results for CYP3A4 (PDB: 1W0E) and raloxifene for known site of metabolism with different heme parameters.

Overall	Dehydrogenation	3'OH	Nonproductive
IC4 (F-HM)	0.474	0.000	0.526
IC6 (F-HM)	0.662	0.042	0.296
IC4 (Oda)	0.269	0.000	0.731
IC6 (Oda)	0.325	0.000	0.675
DIOXY (F-HM)	0.846	0.000	0.154
CPDI (F-HM)	0.521	0.068	0.411
CPDI (T-HM)	0.329	0.032	0.639

Supporting Information Tables 3A–3C). The use of the hexa-coordinate heme parameters for the catalytic oxyferryl CPDI produced as the highest populated and lowest energy clusters, binding modes supporting the observed dehydrogenation of raloxifene, in addition to one major cluster predicting nonproductive binding modes and a minor cluster predicting the observed hydroxylation of raloxifene (see Table 5 and Supporting Information Tables 3A–3C).

Comparison of docking studies performed with these new F-HM charges against previously calculated charges^[14(a)] assigned to either of our QM-optimized penta-coordinate ferric heme geometries (ic4-Oda: our intermediate-spin ferric and ic6-Oda: our high-spin ferric with the previously published charges^[14a]) shows that with this model system, we see improved results with our new charges (see Table 5 and Supporting Information Tables 3A–3C).

Comparison of docking studies with the RED-III-derived atomic partial charges for the hexa-coordinate full heme model of compound I [CPDI (F-HM)] versus the RED-III-derived atomic partial charges for the truncated heme compound I with charges for the gas-phase propionates added and the other macrocycle substituents left neutral [CPDI (T-HM)], indicate again that the CPDI (F-HM) charges produce improved docking results with this model system. The results suggest a profound dependence on the choice of heme model and catalytic state.

In summary of our docking studies, our results demonstrate that the incorporation of atomic partial charges for the heme in each of the different states calculated for the P450 catalytic cycle produce different results for each set of atomic partial

charges for each heme state and model system, including different spin states for the same penta-coordinate ferric charge state, as well as for both hexa-coordinate, dioxygen-bound and the catalytic CPDI species. Surprisingly, for this particular substrate-P450 combination with this force field, despite charge-transfer artifacts identified earlier, the open-shell F-HM-derived charges for the resting ferric high-spin and the catalytic CPDI species produce results most closely in agreement with experimentally observed metabolism of raloxifene by CYP3A4. This may be in part due to the fact that the RED-III multiple reorientation-derived charges for F-HM-unshielded propionates exhibit a smaller charge transfer than observed elsewhere.^[51] This would suggest that the extended model F-HM produces a better representation for the electrostatics of the heme than the T-HM despite the absence of shielding of basic active site residues in these gas-phase calculations. Interestingly, the docking study was also sensitive to the presence of molecular oxygen with the DIOXY (F-HM) parameters not supporting both of the observed mechanisms of P450 metabolism, whereas with the CPDI (F-HM) parameters for the catalytic compound I species produced binding modes suggestive of both hydroxylation and dehydrogenation of raloxifene.

Bond force constants and molecular dynamics

For performing molecular dynamics simulations of the heme alone or of the P450 enzymes, heme force constant parameters were required that were absent from the GAFF parameters. Specifically, these heme parameters were for the porphyrin moiety, and key chelations to the iron atom. Thus, scans for all the missing parameters were performed with the T-HM, except the Fe—S—C describing the angle for the iron bond angle to the proximal cysteine which was performed with the F-HM (Fig. 7).

The Fe—S and Fe—N force constants were obtained for the MM bond parameters by fitting the QM scan results and are relatively similar for the penta-coordinate and dioxygen-bound states (Fig. 7). By contrast, the estimated force constant values for these bonds for CPDI differed substantially. Furthermore, the Fe—O force constant for CPDI was also much greater than the force constant for the DIOXY state (Table 6).

For the angle terms, the force constants varied more between the different heme species. The hexa-coordinate states were similar in value for the Fe—N—C and N—Fe—O. The penta-coordinate and the dioxygen-bound states were closer in value for the N—Fe—N than CPDI. The C—S—Fe angle term was determined with a scan only due to computational cost using the F-HM of CPDI and fit in a similar manner (Table 6). An alternative strategy, when a complete set of bond and angle parameters are desired could be used by using the MTK++ code in AmberTools to mechanically derive the force matrix quantum.^[38] This approach has the advantage over the method used here in that it avoids defining internal coordinates; however, the influence of small changes in bond and angle parameters introduced to the force constants that we have already developed and tested to be compatible with the AMBER ff99SB force field and GAFF in both MD and docking,

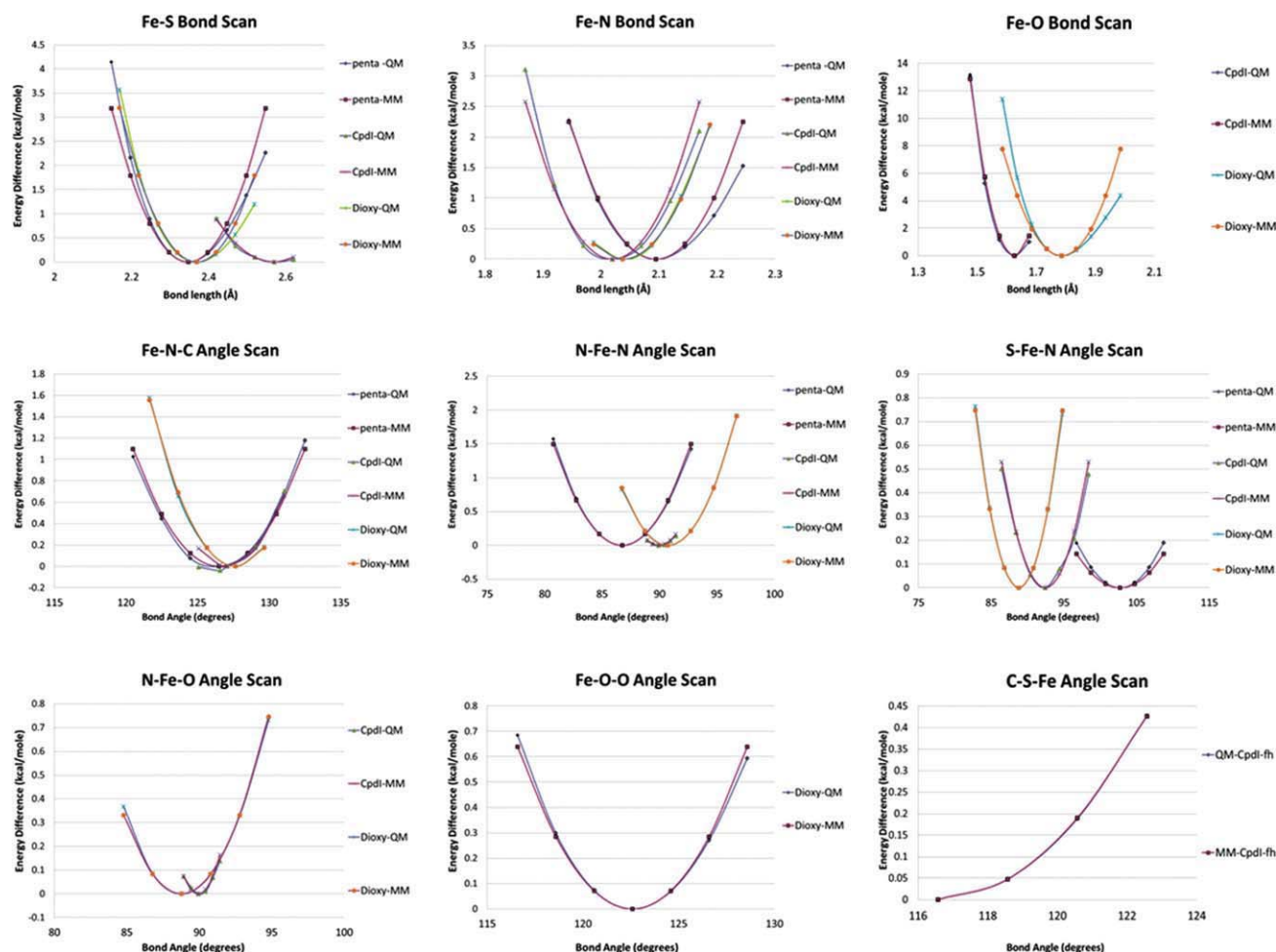


Figure 7. Fit of QM scans along key coordinates for the development of MM force constants. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

is likely minimal. We do recommend the use of this approach for the next generation of heme parameters.

To determine if the GAFF^[36]-based heme models, with F-HM-based RESP charges and with the additional T-HM-based force constants were sufficient to maintain the heme structure, molecular dynamics simulations were performed on the various heme models with a generalized Born implicit solvent treatment (GB-MD). The initial GB-MD runs indicated that the penta-coordinate state required the addition of a number of dihedral angle energy terms to maintain the iron in an out-of-plane state as observed in the QM geometry optimizations. These dihedral angles were X-S-Fe-X, X-cg-cc-X, X-cg-cd-X, cc-nc-fe-nc, cd-nd-fe-nd, cc-nc-fe-SH, and cd-nd-fe-SH (see Supporting Information). For the hexa-coordinate states, the only extra dihedral angle energy term that was included was for X-cg-cd-X and X-cg-cc-X for CPDI and the DIOXY states.

Molecular dynamics of the F-HM in implicit solvent

Visual inspection of the average structures generated for the heme models over the entire 10-ns simulation appears reasonable in that the heme maintains its overall “planarity” and the

expected movement is occurring with axial substituents and the cysteinate ligand (Figs. 8a–8c). A visual inspection of the MD trajectory determined that the larger fluctuations in the RMSD occur during movement of the cysteinate ligand. RMSD plots of 10-ns GB-MD simulations for each of the three different heme states (Fig. 8d) suggest that these parameters are sufficient to maintain the expected heme structure.

Table 6. Calculated MM force constant parameters for bond stretching and angle bending from fit of the energy profiles derived from T-HM all except *, which was derived with F-HM.

	ic6	DIOXY	CPDI
Bonds			
Fe–S	80	80	39
Fe–N	100	98	114
Fe–O		194	572
O–O		372	
Angles			
Fe–N–C	100	142	136
N–Fe–N	137	174	239
S–Fe–N	13	68	48
N–Fe–O		68	65
Fe–O–O		58	
Fe–S–C			39*

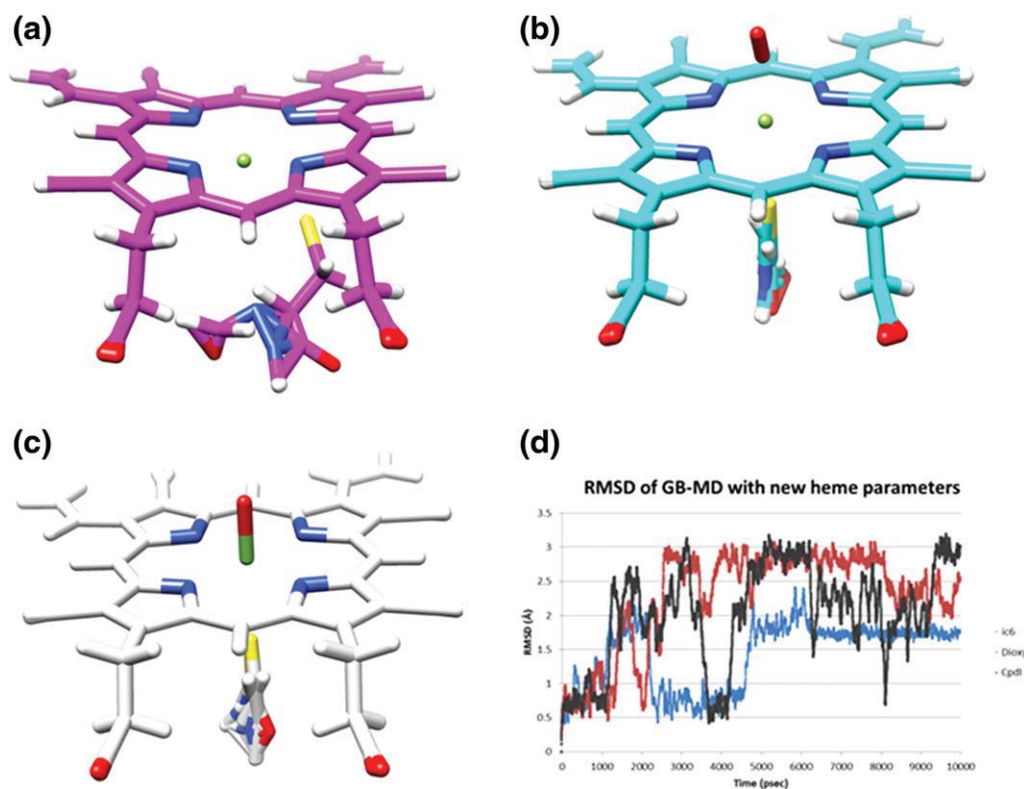


Figure 8. Average structure over 10-ns GB-MD simulation with RED-III derive atomic partial charges and new force constants: a) penta-coordinate ferric high-spin, b) hexa-coordinate dioxygen-bound, c) compound I heme states, and d) RMSD of three GB-MD simulations. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Molecular dynamics of the CYP3A4 protein

MD simulation of the QM parameterized heme states in models of CYP3A4 (PDB ID: 1TQN and 1W0E) was also performed to assess the heme models. Key geometries of the heme for the F-HM structure model with CPDI over the course of full MD simulations of the protein in explicit solvent are shown in Figure 9. The heme Fe—O and Fe—S bond distances and the n—Fe—N, N—Fe—O and N—Fe—S angles appear to remain stable over the course of the ~25-ns MD simulation. The P450 structure appears to reach a steady state based on the RMSD plot 9(C) after 5 ns, and the 2D RMSD indicates that a steady state is reached after 10 ns that differs from that reached during the initial 10 ns. During these small conformational readjustments of the P450 model, the heme geometry maintains key equilibrium geometries with the new heme parameters. Similar steady-state results are observed for the 1W0E structure for CPDI and for both the dioxy and penta-coordinate ferric state with CYP3A4 models 1TQN and 1W0E (data not shown).

Discussion

Accurate description of the electronic ground state of different P450 heme species has been clearly shown to require full QM/MM treatments of the enzyme system based on experimental structures. Yet smaller and more tractable gas-phase models—despite displaying model system specific limitations—have been shown to provide useful insights into reac-

tion mechanisms.^[11] The smaller porphyrine–thiolate model system, such as the T-HM, has been especially useful in this manner and has proven to be surprisingly accurate producing electronic spin densities in good agreement with the results from QM/MM treatment of the full enzyme system.^[51] As QM gas-phase calculations using truncated model systems have proven useful for the development of transferable parameters for classical MD force fields such as for AMBER^[23]/GAFF,^[36] simpler and less computationally expensive truncated models of the heme system has been used for earlier studies for MM parameter development for the resting heme state of the P450 catalytic cycle.

Here, we have tested the use of two model systems, T-HM and F-HM, for the development of MM parameters. The results of the geometry optimizations show that the T-HM more closely matches key axial geometries with available X-ray structures, but lacks some of the experimentally observed asymmetry. Both the F-HM and T-HM are closer in agreement on “in plane” heme geometries with each other than available X-ray results. This is in agreement with other published work and raises the question of what effect these differences may produce in the charge and spin.

Our calculations show that the charges do vary between the two systems, and that charge transfer is observed, in agreement with previously published work,^[51] but that these differences are smaller than that observed elsewhere when charges are derived using RESP with RED-III. Furthermore, the changes in atomic partial charges because of the charge transfer due

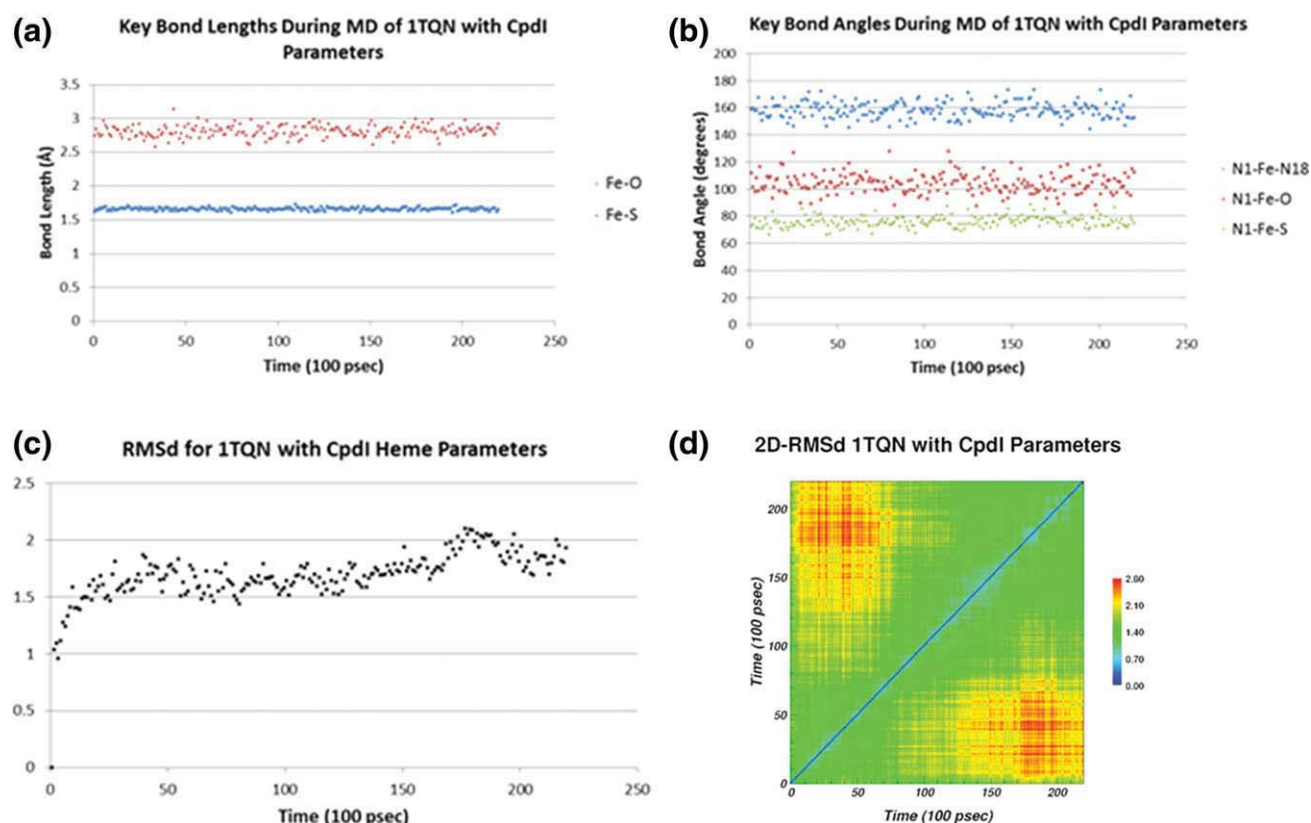


Figure 9. a) Key heme bond distances, b) key heme bond angles, c) 2D RMSD of C α , and d) RMSD of MD simulation of 3A4 (PDB: 1TQN) with the parameters for compound I (Cpdl). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

to “artificial” unpaired electron-spin density on the propionate oxygen atoms are difficult to correlate with changes in atomic partial charges of the entire system. We believe that this is inherent to differences between the model systems, as it is also observed to occur with the closed shell DIOXY state.

As no clear difference or superiority is observed in partial atomic charges derived from either model system, we empirically tested both the T-HM and F-HM-derived charges and previously published charges with a model system that had been experimentally confirmed in our lab. Our results indicate that with AMBER-based docking software that is sensitive to changes in geometries and charges derived from different spin states for the heme, despite the issues with charge transfer from the F-HM produce improved docking results.

Therefore, we developed and tested a set of MM parameters for/with the new AMBER ff99SBildn force field using the T-HM for the force constants and the F-HM charges. These models appear to be stable and produce reasonable results both in GB-MD and with MD of the full enzyme system in explicit solvent.

It remains to be seen how these parameters perform in simulations with other P450 enzymes, substrates, and ligands and which parameters will need additional refinement. However, on the basis of our current experience in MM applications, we recommend these B3LYP/LACVP-derived parameters and RESP-derived charges which are provided in the Supporting Information for the various P450-catalytic cycle heme states in improving P450 MM-based simulations.

Acknowledgments

Thomas E. Cheatham, III acknowledges financial support from NIH GM079383, computer time from the Center for High Performance Computing at Utah and from the NSF TeraGrid MCA01S027. Garold S. Yost acknowledges financial support from NIH GM0742249.

- W. F. van Gunsteren, D. Bakowies, R. Baron, I. Chandrasekhar, M. Christen, X. Daura, P. Gee, D. P. Geerke, A. Glatli, P. H. Hunenberger, M. A. Kastenholz, C. Oostenbrink, M. Schenk, D. Trzysiak, N. F. van der Vegt, H. B. Yu, *Angew Chem Int Ed* 2006, 45, 4064.
- L. Afzelius, C. H. Arnby, A. Broo, L. Carlsson, C. Isaksson, U. Jurva, B. Kjellander, K. Kolmodin, K. Nilsson, F. Raubacher, L. Weidolf, *Drug Metabol Rev* 2007, 39, 61.
- H. Sun, D. O. Scott, *Chem Biol Drug Design* 2010, 75, 3.
- (a) Q. Wang, J. R. Halpert, *Drug Metab Dispos* 2002, 30, 86; (b) S. Ekins, *Biochem Soc Trans* 2003, 31(Pt 3), 611; (c) F. P. Guengerich, *Chem Res Toxicol* 2008, 21, 70.
- M. Otyepka, J. Skopalik, E. Anzenbacherova, P. Anzenbacher, *Biochim Biophys Acta* 2007, 1770, 376.
- (a) E. E. Scott, J. R. Halpert, *Trends Biochem Sci* 2005, 30, 5; (b) E. E. Scott, Y. A. He, M. R. Wester, M. A. White, C. C. Chin, J. R. Halpert, E. F. Johnson, C. D. Stout, *Proc Natl Acad Sci USA* 2003, 100, 13196; (c) E. E. Scott, Y. Q. He, J. R. Halpert, *Chem Res Toxicol* 2002, 15, 1407; (d) E. E. Scott, M. Spatzeneberger, J. R. Halpert, *Arch Biochem Biophys* 2001, 395, 57; (e) E. E. Scott, M. A. White, Y. A. He, E. F. Johnson, C. D. Stout, J. R. Halpert, *J Biol Chem* 2004, 279, 27294.
- (a) D. E. Bikiel, L. Boechi, L. Capece, A. Crespo, P. M. De Biase, S. Di Lella, M. C. Gonzalez Lebrero, M. A. Marti, A. D. Nadra, L. L. Perissinotti, D. A. Scherlis, D. A. Estrin, *Phys Chem Chem Phys* 2006, 8, 5611; (b) A. Dey, Y. Jiang, P. Ortiz de Montellano, K. O. Hodgson, B. Hedman, E. I. Solomon, *J Am Chem Soc* 2009, 131, 7869; (c) P. R. Ortiz de

- Montellano, *Chem Rev* 2010, 110, 932; (d) P. R. Ortiz de Montellano, J. J. De Voss, *Cytochrome P450*, 3rd ed.; Kluwer Academic/Plenum Publishers: New York, NY, USA, 2005, p. 183.
- [8] I. Schlichting, J. Berendzen, K. Chu, A. M. Stock, S. A. Maves, D. E. Benson, R. M. Sweet, D. Ringe, G. A. Petsko, S. G. Sligar, *Science* 2000, 287, 1615.
- [9] P. R. Ortiz de Montellano, R. Paul, Eds. *Cytochrome P450, Structure, Mechanism and Biochemistry*, 3rd ed.; Kluwer Academic/Plenum Publishers: New York, 2003.
- [10] (a) S. Shaik, D. Kumar, S. P. de Visser, A. Altun, W. Thiel, *Chem Rev* 2005, 105, 2279; (b) D. Harris, G. Loew, L. Waskell, *J Inorg Biochem* 2001, 83, 309; (c) C. M. Bathelt, L. Ridder, A. J. Mulholland, J. N. Harvey, *J Am Chem Soc* 2003, 125, 15004.
- [11] S. Shaik, S. Cohen, Y. Wang, H. Chen, D. Kumar, W. Thiel, *Chem Rev* 2010, 110, 949.
- [12] G. D. Szklarz, J. R. Halpert, *J Comput-Aided Mol Des* 1997, 11, 265.
- [13] F. Autenrieth, E. Tajkhorshid, J. Baudry, Z. Luthey-Schulten, *J Comput Chem* 2004, 25, 1613.
- [14] (a) A. Oda, N. Yamaotsu, S. Hirono, *J Comput Chem* 2005, 26, 818; (b) P. Rydberg, L. Olsen, P.-O. Norrby, U. Ryde, *J Chem Theory Comput* 2007, 3, 1765.
- [15] (a) E. H. Lee, J. Hsin, M. Sotomayor, G. Comellas, K. Schulten, *Structure* 2009, 17, 1295; (b) R. Lavery, K. Zakrzewska, D. Beveridge, T. C. Bishop, D. A. Case, T. Cheatham, III, S. Dixit, B. Jayaram, F. Lankas, C. Laughton, J. H. Maddocks, A. Michon, R. Osman, M. Orozco, A. Perez, T. Singh, N. Spackova, J. Sponer, *Nucleic Acids Res* 2010, 38, 299; (c) R. O. Dror, M. O. Jensen, D. W. Borhani, D. E. Shaw, *J Gen Physiol* 2010, 135, 555.
- [16] (a) J. Kottalam, D. A. Case, *J Am Chem Soc* 1988, 110, 7690; (b) D. A. Case, J. A. McCammon, *Annals N Y Acad Sci* 1986, 482, 222.
- [17] S. C. Hoops, K. W. Anderson, K. M. Merz, *J Am Chem Soc* 1991, 113, 8262.
- [18] C. D. Moore, C. A. Reilly, G. S. Yost, *Biochemistry* 2010, 49, 4466.
- [19] M. Ekroos, T. Sjogren, *Proc Natl Acad Sci USA* 2006, 103, 13682.
- [20] P. A. Williams, J. Cosme, D. M. Vinkovic, A. Ward, H. C. Angove, P. J. Day, C. Vonnrhein, I. J. Tickle, H. Jhoti, *Science* 2004, 305, 683.
- [21] F. Y. Dupradeau, A. Pigache, T. Zaffran, C. Savineau, R. Lelong, N. Grivel, D. Lelong, W. Rosanski, P. Cieplak, *Phys Chem Chem Phys* 2010, 12, 7821.
- [22] G. M. Morris, D. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, A. J. Olson, *J Comput Chem* 1998, 19, 1639.
- [23] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J Am Chem Soc* 1995, 117, 5179.
- [24] C. I. Bayly, P. Cieplak, W. Cornell, P. Kollman, *J Phys Chem* 1993, 97, 10269.
- [25] (a) P. C. Hariharan, J. A. Pople, *Theor Chim Acta* 1973, 28, 213; (b) M. M. Francl, W. J. Pietro, W. J. Hehre, J. S. Binkley, M. S. Gordon, D. J. Defrees, J. A. Pople, *J Chem Phys* 1982, 77, 3654.
- [26] (a) P. J. Hay, W. R. Wadt, *J Chem Phys* 1985, 82, 299; (b) T. H. Dunning, Jr., P. J. Hay, *Methods of Electronic Structure Theory*; Plenum: New York, 1977.
- [27] A. Altun, D. Kumar, F. Neese, W. Thiel, *J Phys Chem A* 2008, 112, 12904.
- [28] M. Radoń, E. Broclawik, *J Chem Theory Comput* 2007, 3, 728.
- [29] J. K. Yano, M. R. Wester, G. A. Schoch, K. J. Griffin, C. D. Stout, E. F. Johnson, *J Biol Chem* 2004, 279, 38091.
- [30] P. Rowland, F. E. Blaney, M. G. Smyth, J. J. Jones, V. R. Leydon, A. K. Oxbrow, C. J. Lewis, M. G. Tennant, S. Modi, D. S. Eggleston, R. J. Chenery, A. M. Bridges, *J Biol Chem* 2006, 281, 7614.
- [31] Y. Zhao, M. A. White, B. K. Muralidhara, L. Sun, J. R. Halpert, C. D. Stout, *J Biol Chem* 2006, 281, 5973.
- [32] (a) R. Seeger, J. A. Pople, *J Chem Phys* 1977, 66, 3045; (b) R. Bauernschmitt, R. Ahlrichs, *J Chem Phys* 1996, 104, 9047.
- [33] M. Swart, A. R. Groenhof, A. W. Ehlers, K. Lammertsma, *J Phys Chem A* 2004, 108, 5479.
- [34] C. D. Moore, K. Shahrokh, S. F. Sontum, T. E. Cheatham, G. S. Yost, *Biochemistry* 2010, 49, 9011.
- [35] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, *Proteins* 2006, 65, 712.
- [36] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J Comput Chem* 2004, 25, 1157.
- [37] (a) R. D. Hancock, *Progress in Inorganic Chemistry*, Vol. 37; Wiley, 1991; Hoboken, NJ, USA, p. 187; (b) K. M. Merz, *J Am Chem Soc* 1991, 113, 406.
- [38] M. B. Peters, Y. Yang, B. Wang, L. S. Füsti-Molnár, M. N. Weaver, K. M. Merz, *J Chem Theory Comput* 2010, 6, 2935.
- [39] G. D. Hawkins, C. J. Cramer, D. G. Truhlar, *J Phys Chem* 1996, 100, 19824.
- [40] D. A. Case, T. E. Cheatham, III, T. Darden, H. Gohlke, R. Luo, K. M. Merz, Jr., A. Onufriev, C. Simmerling, B. Wang, R. J. Woods, *J Comput Chem* 2005, 26, 1668.
- [41] (a) C. Simmerling, B. Strockbine, A. E. Roitberg, *J Am Chem Soc* 2002, 124, 11258; (b) K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, D. E. Shaw, *Proteins* 2010, 78, 1950.
- [42] I. S. Joung, T. E. Cheatham, III, *J Phys Chem B* 2008, 112, 9020.
- [43] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J Chem Phys* 1983, 79, 926.
- [44] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, *J Chem Phys* 1995, 103, 8577.
- [45] J.-P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J Comput Phys* 1977, 23, 327.
- [46] R. W. Pastor, B. R. Brooks, A. Szabo, *Mol Phys* 1988, 65, 1409.
- [47] (a) A. D. Becke, *J Chem Phys* 1993, 98, 5648; (b) C. Lee, W. Yang, R. G. Parr, *Phys Rev B* 1988, 37, 785; (c) S. H. Vosko, L. Wilk, M. Nusair, *Can J Phys* 1980, 58, 1200; (d) P. J. Stephens, F. J. Delvin, C. F. Chabalowski, M. J. Frisch, *J Phys Chem* 1994, 98, 11623.
- [48] B. Meunier, S. P. de Visser, S. Shaik, *Chem Rev* 2004, 104, 3947.
- [49] R. Davydov, T. M. Makris, V. Kofman, D. E. Werst, S. G. Sligar, B. M. Hoffman, *J Am Chem Soc* 2001, 123, 1403.
- [50] T. M. Makris, K. von Koenig, I. Schlichting, S. G. Sligar, *J Inorg Biochem* 2006, 100, 507.
- [51] J. C. Schöneboom, S. Cohen, H. Lin, S. Shaik, W. Thiel, *J Am Chem Soc* 2004, 126, 4017.

Received: 25 May 2011

Revised: 28 July 2011

Accepted: 30 July 2011

Published online on 14 October 2011