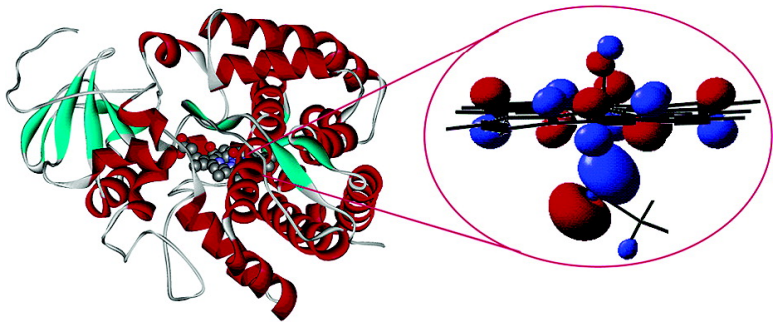
Article

Electronic Structure of Compound I in Human Isoforms of Cytochrome P450 from QM/MM Modeling

Christine M. Bathelt, Jolanta Zurek, Adrian J. Mulholland, and Jeremy N. Harvey

J. Am. Chem. Soc., **2005**, 127 (37), 12900-12908• DOI: 10.1021/ja0520924 • Publication Date (Web): 26 August 2005



J|A|C|S
ARTICLES
Published on Web 08/26/2005

Electronic Structure of Compound I in Human Isoforms of Cytochrome P450 from QM/MM Modeling

Christine M. Bathelt, Jolanta Zurek, Adrian J. Mulholland,* and Jeremy N. Harvey*

Contribution from the Centre for Computational Chemistry, School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS United Kingdom

Received April 1, 2005; E-mail: Jeremy.harvey@bris.ac.uk

Abstract: Human cytochromes P450 play a vital role in drug metabolism. The key step in substrate oxidation involves hydrogen atom abstraction or C=C bond addition by the oxygen atom of the Compound I intermediate. The latter has three unpaired electrons, two on the Fe−O center and one shared between the porphyrin ring and the proximal cysteinyl sulfur atom. Changes in its electronic structure have been suggested to affect reactivity. The electronic and geometric structure of Compound I in three important human subfamilies of cytochrome P450 (P450, 2C, 2B, and 3A) that are major contributors to drug metabolism is characterized here using combined quantum mechanical/molecular mechanical (QM/MM) calculations at the B3LYP:CHARMM27 level. Compound I is remarkably similar in all isoforms, with the third unpaired electron located mainly on the porphyrin ring, and this prediction is not very sensitive to details of the QM/MM methodology, such as the DFT functional, the basis set, or the size of the QM region. The presence of substrate also has no effect. The main source of variability in spin density on the cysteinyl sulfur (from 26 to 50%) is the details of the system setup, such as the starting protein geometry used for QM/MM minimization. This conformational effect is larger than the differences between human isoforms, which are therefore not distinguishable on electronic grounds, so it is unlikely that observed large differences in substrate selectivity can be explained to a large extent in these terms.

Introduction

The cytochromes P450 form a large family of heme-containing enzymes, which are present in a range of species, including plants, fungi, bacteria, insects, and mammals. They perform vital bioregulatory functions, such as detoxification of xenobiotics and biosynthesis of endogenous compounds. A variety of monooxygenase reactions are catalyzed, such as hydroxylations, epoxidations, heteroatom oxidation, and N- and O-dealkylation.[1] There is a particular interest in the human cytochromes P450, as it is thought that more than 90% of drug oxidations in man are mediated by the cytochrome P450 enzyme family.[2−4] It would therefore be of great value to be able to understand and predict the reactivity and selectivity properties of the cytochromes P450. However, these properties are known from experiment to be highly complex, as these enzymes are not restricted to a particular substrate: unlike most enzymes, cytochrome P450 isoforms are capable of oxidizing a structurally diverse set of molecules. Substrates can typically undergo oxidation catalyzed by several different P450 isoforms, with widely varying selectivity patterns. The isoforms 2C9 and 3A4, for example, catalyze the oxidation of the drug compound diclofenac on two different aromatic rings, each with high regioselectivity.[5] Indeed, different isoforms can even catalyze different reaction types, such as aliphatic and aromatic hydroxylation and epoxidation with the same substrate. Coumarin, for example, is hydroxylated by the 3A4 isoform, while cytochromes P450 of the 2E and 1A subfamily catalyze its epoxidation.[6]

These differences in reactivity could in principle have several causes. Perhaps the most obvious is simply that different substrates position themselves differently upon binding in the active site, due to steric constraints[7] and electrostatic interactions in the enzyme. For this reason, considerable effort has been invested in predicting substrate docking geometries.[8,9] However, more "chemical" effects must also play an important role in determining the selectivity. Most oxidation reactions with the P450 enzymes seem to involve the same heme-based oxidizing species, referred to as "Compound I". However, it has been suggested that other oxidizing species might play a role in some reactions. There is currently extensive debate as to how significant the role of such alternative oxidants might be, but the evidence suggests that they would not alone be able to explain the remarkable diversity in reactivity.[10] Equally, the immediate environment of the heme group could also lead to

(1) Guengerich, F. P. *Chem. Res. Toxicol.* **2001**, *14*, 611−650.
(2) Anzenbacher, P.; Anzenbacherova, E. *Cell. Mol. Life Sci.* **2001**, *58*, 737−747.
(3) Lewis, D. F.; Dickins, M. *Drug. Discovery Today* **2002**, *7*, 918−925.
(4) Lewis, D. F. *Pharmacogenomics* **2003**, *4*, 387−395.
(5) Mancy, A.; Antignac, M.; Minoletti, C.; Dijols, S.; Mouries, V.; Duong, N. T.; Battioni, P.; Dansette, P. M.; Mansuy, D. *Biochemistry* **1999**, *38*, 14264−14270.
(6) Born, S. L.; Caudill, D.; Fliter, K. L.; Purdon, M. P. *Drug Metab. Dispos.* **2002**, *30*, 483−487.
(7) Bathelt, C.; Schmid, R. D.; Pleiss, J. *J. Mol. Model.* **2002**, *8*, 327.
(8) Kirton, S. B.; Murray, C. W.; Verdonk, M. L.; Taylor, R. D. *Proteins* **2005**, *58*, 836−844.
(9) Park, J. Y.; Harris, D. *J. Med. Chem.* **2003**, *46*, 1645−1660.
(10) Newcomb, M.; Hollenberg, P. F.; Coon, M. J. *Arch. Biochem. Biophys.* **2003**, *409*, 72−79.

10.1021/ja0520924 CCC: $30.25

different reactivity patterns in different isoforms and with different substrates.[11] For example, the protein environment could "tune" the reactivity of the Compound I intermediate by changing its electronic structure.

The common structural characteristic in the active site of all cytochromes P450 is a thiolate-ligated heme group, which activates molecular oxygen.[12,13] The Compound I active species of cytochrome P450 has not been observed experimentally due to its high reactivity, which means that it does not build up significantly during the catalytic cycle. However, there is substantial indirect evidence, from in situ EPR and ENDOR studies[14] during the oxidation of camphor catalyzed by the bacterial isoform P450cam, that a hydroxylating species analogous to the more stable and well-characterized Compound I of chloroperoxidase (CPO) is formed. Recent studies[15,16] also showed the formation of a species with an optical absorption spectrum identical to that of CPO-Compound I when the substrate-free ferric resting form of cytochrome P450 was reacted with *m*-chloroperoxybenzoic acid. Spectroscopic data[17−19] have identified CPO-Compound I as an iron−oxo porphyrin π-cation radical with a doublet ground state. In this species, the iron is oxidized to an oxyferryl (Fe(IV)) triplet spin state, and the porphyrin ring is oxidized to a π-cation radical. The distribution of the radical spin density on the porphyrin system has been a matter of debate among experimentalists.[20] The unpaired electron on the porphyrin ligand can potentially reside in a porphyrin π orbital with either a_{2u} or a_{1u} character. Computational studies[11,21−27] on models of Compound I predict an A_{2u} electronic ground state, where the two unpaired electrons on the iron−oxo moiety couple to an unpaired electron in the a_{2u} orbital of the porphyrin ring. The unpaired electron on the porphyrin ring can be spin-up or spin-down. Thus, coupling with the triplet electrons on the iron−oxo moiety can lead to the overall doublet and quartet spin states $^2A_{2u}$ and $^4A_{2u}$, respectively, which were found to lie very close in energy. States in which the unpaired porphyrin electron is located in the a_{1u} orbital, or in which it is partly delocalized on the proximal cysteinyl sulfur atom, were also found to lie fairly close in energy.

There is evidence from computational work[11,22,23,25−27] that the electronic character of Compound I varies in response to hydrogen bonding and polarity of the environment. In the gas phase, Compound I shows significant sulfur radical character, due to a large admixture of a sulfur lone pair p orbital with the porphyrin a_{2u} orbital. Upon including an approximate treatment of the hydrogen bonding and polar environment of the thiolate ligand within the enzyme,[11,25] this mixing is much reduced, and the unpaired electron is predominantly located on the porphyrin ring. This change in electronic character was found to influence significantly the calculated barrier heights along pathways for hydroxylation and epoxidation of propene, and thereby the predicted selectivity.[11] It was also shown[28] that the application of an external electric field along the S−Fe−O axis leads to changes in the calculated spin distribution between the thiolate and the porphyrin ligand, and thereby also to changes in the predicted regioselectivity in oxidation reactions. This has led to Compound I being designated a "chemical chameleon", which changes its electronic structure depending on the details of the protein environment.[11,21] These observations suggest that the differing protein environment in different P450 isoforms could lead to changes in the electronic structure of the ground state of the Compound I active oxidant, and hence to the observed differences in selectivity among these catalysts.

There is also experimental evidence for the sensitivity of the electronic structure of Compound I to the enzyme environment. EPR studies[17] of CPO-Compound I confirm that one unpaired electron is not located on the porphyrin ring alone but partly delocalized onto the thiolate ligand. The role of hydrogen bonding of backbone N−H groups of the protein to the sulfur of the coordinating thiolate ligand was investigated by mutation studies[29,30] on P450cam. They show that removing hydrogen bonds to the sulfur alters the reduction potential of the heme iron. Likewise, studies on synthetic model systems[31] found that the ratio between hydroxylation and epoxidation changed with the number of hydrogen bonds formed to the sulfur, and suggest that the electronic structure of Compound I is influenced by the hydrogen bonding network in the enzyme. Also, related Compound I species are known for a variety of other heme proteins, such as cytochrome *c* peroxidase, ascorbate peroxidase, and horseradish peroxidase, and the electronic structure of these species also varies quite significantly despite a coordination environment to the heme group that is rather similar.[32]

So far, computational studies of the effect of the enzyme environment on the electronic structure and geometry of Compound I have been restricted to a single cytochrome P450 isoform, the bacterial P450cam. Combined quantum mechanical/molecular mechanical (QM/MM) calculations on this enzyme[21,33,34] predict that, in the protein, spin density is spread over both the porphyrin and the thiolate ligand but is located

(11) de Visser, S. P.; Ogliaro, F.; Sharma, S.; Shaik, S. *Angew. Chem., Int. Ed.* **2002**, *41*, 1947−1951.
(12) Loew, G. H.; Harris, D. L. *Chem. Rev.* **2000**, *100*, 407−420.
(13) Sono, M.; Roach, M. P.; Coulter, E. D.; Dawson, J. H. *Chem. Rev.* **1996**, *96*, 2841−2887.
(14) Davydov, R.; Makris, T. M.; Kofman, V.; Werst, D. E.; Sligar, S. G.; Hoffman, B. M. *J. Am. Chem. Soc.* **2001**, *123*, 1403−1415.
(15) Egawa, T.; Shimada, H.; Ishimura, Y. *Biochem. Biophys. Res. Commun.* **1994**, *201*, 1464−1469.
(16) Kellner, D. G.; Hung, S. C.; Weiss, K. E.; Sligar, S. G. *J. Biol. Chem.* **2002**, *277*, 9641−9644.
(17) Weiss, R.; Mandon, D.; Wolter, T.; Trautwein, A. X.; Muether, M.; Bill, E.; Gold, A.; Jayaray, K.; Terner, J. *J. Biol. Inorg. Chem.* **1996**, *1*, 377−383.
(18) Hosten, C. M.; Sullivan, A. M.; Palaniappan, V.; Fitzgerald, M. M.; Terner, J. *J. Biol. Chem.* **1994**, *269*, 13966−13978.
(19) Rutter, R.; Hager, L. P. *J. Biol. Chem.* **1982**, *257*, 7958−7961.
(20) Terner, J.; Gold, A.; Weiss, R.; Mandon, D.; Trautwein, A. X. *J. Porphyrins Phthalocyanines* **2001**, *5*, 357−364.
(21) Schöneboom, J. C.; Lin, H.; Reuter, N.; Thiel, W.; Cohen, S.; Ogliaro, F.; Shaik, S. *J. Am. Chem. Soc.* **2002**, *124*, 8142−8151.
(22) Harris, D.; Loew, G.; Waskell, L. *J. Inorg. Biochem.* **2001**, *83*, 309−318.
(23) Harris, D. L.; Loew, G. H. *J. Am. Chem. Soc.* **1998**, *120*, 8941−8948.
(24) Harris, N.; Cohen, S.; Filatov, M.; Ogliaro, F.; Shaik, S. *Angew. Chem., Int. Ed.* **2000**, *39*, 2003−2007.
(25) Ogliaro, F.; Cohen, S.; Filatov, M.; de Visser, S. P.; Shaik, S. *J. Am. Chem. Soc.* **2000**, *122*, 12892−12893.
(26) Ogliaro, F.; Cohen, S.; Filatov, M.; Harris, D.; Shaik, S. *Angew. Chem., Int. Ed.* **2000**, *39*, 3851−3855.
(27) Ogliaro, F.; de Visser, S. P.; Groves, J. T.; Shaik, S. *Angew. Chem., Int. Ed.* **2001**, *40*, 2874−2878.
(28) Shaik, S.; de Visser, S. P.; Kumar, D. *J. Am. Chem. Soc.* **2004**, *126*, 11746−11749.
(29) Yoshioka, S.; Takahashi, S.; Ishimori, K.; Morishima, I. *J. Inorg. Biochem.* **2000**, *81*, 141−151.
(30) Yoshizawa, K.; Tosha, T.; Ishimori, K.; Hori, H.; Morishima, I. *J. Am. Chem. Soc.* **2002**, *124*, 14571−14579.
(31) Suzuki, N.; Higuchi, T.; Urano, Y.; Kikuchi, K.; Uekusa, H.; Ohashi, Y.; Uchida, T.; Kitagawa, T.; Nagano, T. *J. Am. Chem. Soc.* **1999**, *121*, 11571−11572.
(32) Raven, E. L. *Nat. Prod. Rep.* **2003**, *20*, 367−381.
(33) Schoneboom, J. C.; Cohen, S.; Lin, H.; Shaik, S.; Thiel, W. *J. Am. Chem. Soc.* **2004**, *126*, 4017−4034.
(34) Schoneboom, J. C.; Neese, F.; Thiel, W. *J. Am. Chem. Soc.* **2005**, *127*, 5840−5853.

more on the porphyrin. Another recent computational study[35] on Compound I of P450cam, using slightly different QM/MM methods, reported a delocalization of spin onto the peripheral heme substituents rather than onto the sulfur. This delocalization was proposed to play an important role in defining the reactivity of the intermediate. Even for P450cam, then, there is some uncertainty concerning the electronic structure of this key intermediate and how it affects the oxidation properties.

For a long time, only crystal structures of bacterial cytochromes P450 were available due to the difficulties involved in obtaining crystals of mammalian (e.g., human) isoforms, which are membrane-bound, unlike the soluble bacterial cytochromes. This problem has recently been surmounted, and a small number of crystal structures[36−40] of mammalian P450 isoforms have been resolved crystallographically, namely isoforms of the 2C, 2B, and 3A subfamilies, which are among the major catalysts of human drug metabolism. The general topology of human and bacterial cytochrome P450 structures is similar. However, there are significant differences in structural features involved in substrate binding, membrane binding, and interaction with redox partners.[41] Within the human isoforms, there are also differences in the size and shape of the active sites, and in flexible regions of the enzymes that can adapt their conformation depending on the bound substrate. It is also to be noted that isoforms within P450 family 2 exhibit more than 70% sequence identity, while the sequence identity between P450s of family 2 and family 3 is only 40%.

In previous work,[42,43] we established a structure−reactivity relationship for aromatic hydroxylation in cytochrome P450 based on calculations using a small QM model. The effect of the environment was either ignored or treated using a solvent continuum model. To address the topic of the present study, namely the electronic structure of the Compound I intermediate of the different human isoforms of cytochrome P450, it is however necessary to include an atomistic description of a much larger region of the protein at the atomic level. To do this, we use combined quantum mechanical/molecular mechanical (QM/MM) calculations, in which a small number of atoms, essentially the heme group of the Compound I species, are treated quantum mechanically, while a much larger number of protein atoms are described using molecular mechanics. The QM region interacts with the MM region through polarization by the electric field created by MM point charges, and through the steric restraints associated with the specific enzyme environment. We report here for the first time QM/MM calculations on the electronic and geometric structure of the Compound I active intermediate of the 2C9, 2B4, and 3A4 human cytochrome P450 isoforms (2B4 is isolated from rabbits but is very similar to human 2B6 enzyme). We use the standard B3LYP hybrid density functional in most cases for the QM part of these calculations, as this has been found to give accurate results in many calculations on P450 Compound I[21,22,24,26,27,33,34,42,44,45] species and for many other bioinorganic systems.[46−48] To provide a point of comparison with previous work, we also studied the Compound I species in bacterial P450cam. Additionally, effects of different enzyme conformations and of drug binding were investigated, as well as the influence of the computational setup in terms of QM regions, the use of different QM density functionals and basis sets. This new insight into the properties of the active species Compound I in human cytochromes P450 can help to improve structure−activity relationships and to solve mechanistic questions of drug oxidation.

Computational Details

Initial structures were constructed on the basis of the crystallographical structures obtained from the Protein Data Bank. There are two slightly different X-ray structures of P450 2C9 with the PDB codes 1OG2[36] and 1R9O,[37] which we refer to as P450 2C9•A and P450 2C9•B, respectively. Likewise, there are two structures of P450 3A4 with the PDB codes 1WOE[39] and 1TQN,[40] which we refer to as P450 3A4•A and P450 3A4•B, respectively. For P450 2B4 and P450cam, we use the structures with the PDB codes 1SUO[38] and 1DZ9,[49] respectively. Hydrogen atoms were added according to standard pK_A values and the H-bonding environment for His residues, and the structures were solvated in a sphere of equilibrated water with a radius of 26 Å centered on the heme iron. The added water was then equilibrated and minimized before QM/MM geometry optimization of all atoms within a sphere of 15 or 20 Å radius around the heme iron (or, for the larger QM regions, the meso carbon between the propionate side chains). In some cases, the whole system was equilibrated using molecular dynamics (MD) with an MM force field describing the Compound I intermediate for 200 ps. Snapshots during the equilibration were minimized using MM and then QM/MM. All pure MM and MD calculations used the CHARMM22 force field[50] and the CHARMM package (version c27b2),[51] with additional parameters for the heme group and the substrate molecules diclofenac and ibuprofen as described in the Supporting Information, which also contains full details of the system setup.

In the QM/MM calculations, the QM region (see the Results and Discussion section for a definition of the QM region) was described by density functional theory, using the Jaguar 4.0 software,[52] with the standard B3LYP, B3PW91, and BP86 functionals. Quartet $^4A_{2u}$ states were treated using a restricted open-shell ansatz, while the $^2A_{2u}$ states required the use of an unrestricted approach (unrestricted calculations on the quartet gave near-identical results). For most calculations, the standard Los Alamos effective core potential and the associated double-ζ basis LACVP[53] was used on iron and the 6-31G basis on all other atoms. This combination is referred to as BSI. In some cases, the larger BSII comprising the LACV3P triple-ζ contraction of the LACVP basis, as implemented in Jaguar, on Fe, and the 6-31G(d) basis

(35) Guallar, V.; Baik, M. H.; Lippard, S. J.; Friesner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 6998−7002.
(36) Williams, P. A.; Cosme, J.; Ward, A.; Angove, H. C.; Matak Vinkovic, D.; Jhoti, H. *Nature* **2003**, *424*, 464−468.
(37) Wester, M. R.; Yano, J. K.; Schoch, G. A.; Yang, C.; Griffin, K. J.; Stout, C. D.; Johnson, E. F. *J. Biol. Chem.* **2004**, *279*, 35630−35637.
(38) Scott, E. E.; White, M. A.; He, Y. A.; Johnson, E. F.; Stout, C. D.; Halpert, J. R. *J. Biol. Chem.* **2004**, *279*, 27294−27301.
(39) Williams, P. A.; Cosme, J.; Vinkovic, D. M.; Ward, A.; Angove, H. C.; Day, P. J.; Vonrhein, C.; Tickle, I. J.; Jhoti, H. *Science* **2004**, *305*, 683−686.
(40) Yano, J. K.; Wester, M. R.; Schoch, G. A.; Griffin, K. J.; Stout, C. D.; Johnson, E. F. *J. Biol. Chem.* **2004**, *279*, 38091−38094.
(41) Williams, P. A.; Cosme, J.; Sridhar, V.; Johnson, E. F.; McRee, D. E. *J. Inorg. Biochem.* **2000**, *81*, 183−190.
(42) Bathelt, C. M.; Ridder, L.; Mulholland, A. J.; Harvey, J. N. *Org. Biomol. Chem.* **2004**, *2*, 2998−3005.
(43) Bathelt, C. M.; Ridder, L.; Mulholland, A. J.; Harvey, J. N. *J. Am. Chem. Soc.* **2003**, *125*, 15004−15005.
(44) Harris, D.; Loew, G.; Waskell, L. *J. Am. Chem. Soc.* **1998**, *120*, 4308−4318.
(45) de Visser, S. P.; Ogliaro, F.; Sharma, P. K.; Shaik, S. *J. Am. Chem. Soc.* **2002**, *124*, 11809−11826.
(46) Siegbahn, P. E.; Blomberg, M. R. *Chem. Rev.* **2000**, *100*, 421−438.
(47) Siegbahn, P. E. *Q. Rev. Biophys.* **2003**, *36*, 91−145.
(48) Harvey, J. N. *J. Am. Chem. Soc.* **2000**, *122*, 12401−12402.
(49) Schlichting, I.; Berendzen, J.; Chu, K.; Stock, A. M.; Maves, S. A.; Benson, D. E.; Sweet, R. M.; Ringe, D.; Petsko, G. A.; Sligar, S. G. *Science* **2000**, *287*, 1615−1622.
(50) MacKerell, A. D.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(51) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(52) *Jaguar*, 4.2 ed.: Schrodinger, Inc.: Portland, OR, 1991−2002.
(53) Hay, J. P.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 299−308.

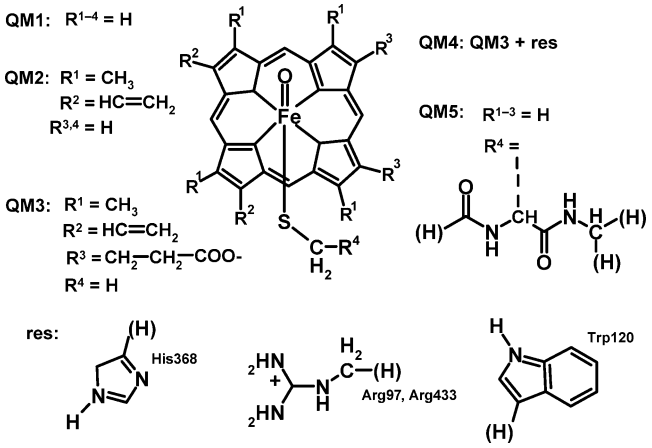


Figure 1. QM regions QM1−QM5; atoms in parentheses indicate H link atoms.

on all other atoms, was also used. The TINKER MM code with the CHARMM27 all-atom force field[54] was used for the treatment of the MM part of the system. Potential energy terms were computed for atoms within a sphere of 15 or 20 Å centered on the heme iron for substrate-free and substrate-bound enzymes, respectively. Input and output from the Jaguar and Tinker codes were coupled using our own set of routines (QoMMMa[55]), which was also used to optimize the geometry of the QM subset of the atoms; MM atoms were optimized at each QM step within the Tinker code. Polarization effects on the QM atoms are accounted for by including the fixed MM charges in the QM Hamiltonian. The steric QM/MM interactions are described using van der Waals radii for the QM atoms that are taken from the CHARMM27 force-field values for similar atoms. No electrostatic or nonbonding cutoff was applied in the QM/MM calculations. The valences of the covalent bonds at the QM/MM boundary were satisfied using hydrogen link atoms.[56] Like the other QM atoms, these additional hydrogen capping atoms were exposed to the electric field of the MM point charges. However, MM charges on the atom replaced by the link atom, as well as those on a few atoms directly bonded to this atom, were set to zero to avoid nonphysical effects. The positions of the bonds replaced by link atoms were chosen such that no artificial charge transfer between the MM and QM regions was introduced by the link atom scheme.

Results and Discussion

P450 2C9. In the first part of this work, we examine how various physical effects and computational parameters combine to affect the electronic structure of the Compound I intermediate in cytochrome P450. For this work, we chose the human isoform 2C9, as a representative of one of the important drug-metabolizing classes of cytochromes P450. The starting structure for these QM/MM geometry optimizations is one of the available crystal structures (2C9•A), which was prepared as described in the Computational Details.

We begin by examining the effect of varying the size and nature of the QM region on the predicted electronic structure for the substrate-free 2C9 Compound I system. Five different QM regions were used, which differ in the substitution of the porphyrin ring, the inclusion of protein side chains, and the extension to the coordinating cysteine (Figure 1). The simplest model, QM1, comprises the iron−oxo−porphyrin subunit without any side chains and the S−CH_2 group of the cysteine (43 atoms). This QM region has been used successfully in QM/MM calculations on P450cam[21] and in pure QM studies.[23,57−59] QM2 additionally contains the methyl and vinyl substituents of the porphyrin (63 atoms), whereas QM3 includes all porphyrin substituents (79 atoms), and QM4 additionally includes parts of the protein side chains that are involved in salt bridges with the propionate substituents on the porphyrin ring (131 atoms), that is, the aromatic rings of His368 and Trp120 and the guanidinium groups of Arg97 and Arg433. Finally, QM5 consists of the unsubstituted iron−oxo−porphyrin subunit as in QM1 but includes the complete Cys435, the CO group of Ile434, and the NHCH group of Val436 (Figure 1).

By analogy with the electronic structure proposed for the related chloroperoxidase Compound I species,[17] the electronic ground state of the cytochrome P450 Compound I is generally recognized to be a triradicaloid and near-degenerate pair of spin states, labeled $^2A_{2u}$ and $^4A_{2u}$. In these states, two unpaired electrons occupy the π^* orbitals of the ferryl group with a triplet spin coupling. The third unpaired electron, which can either be spin-up or spin-down, resides in an orbital that is a mixture of the a_{2u} orbital of the porphyrin ring (hence the label) and the p_σ orbital of the sulfur. This is confirmed by our calculations using QM region QM1: starting with a variety of different initial guesses, the $^2A_{2u}$ and $^4A_{2u}$ pair of states are found to lie very close to each other, and lower than any other state. Apart from the spin coupling, these two states are also found to resemble each other very strongly in terms of their geometry, their electronic structure, and their energy. The geometry of these states is also similar to that found in previous work on Compound I in cytochrome P450cam[21,33] and on gas-phase models,[22,26,27] with the iron atom lying in the porphyrin plane and with an Fe−O bond length of 1.65 Å. The singly occupied orbitals derived from a QM/MM calculation with QM1 on the $^4A_{2u}$ state of the Compound I species of the 2C9 isoform are shown in Figure 2.

The most notable aspect of the electronic structure is the spatial localization of the third unpaired electron. In our calculations on the isolated QM region QM1 in the gas phase,[42] we found extensive mixing between the porphyrin a_{2u} orbital and the sulfur bonding pair, with more than half the unpaired spin density corresponding to this singly occupied orbital located on the sulfur (e.g., in the $^2A_{2u}$ state, the sulfur bears excess spin-down electrons leading to a spin population (ρ_S) of −0.57). We find a much smaller spin density (ca. 1/3 of an unpaired electron, see Table 1) on sulfur in our QM/MM calculations including the protein environment. This effect has been noticed before for various models of P450cam,[11,21,25] and can be attributed to electrostatic and hydrogen bonding interactions that lower the energy of the sulfur-based orbitals, and hence decrease the mixing with the porphyrin orbital. As has been noted before, the optimum Fe−S bond length is somewhat sensitive to the balance of charge and spin between the thiolate and porphyrin ligands: upon including the protein environment, the Fe−S bond is strengthened and shortens, from 2.63 Å in gas-phase[42] calculations on model QM1 to values between 2.52 and 2.57 Å in B3LYP QM/MM calculations. This bond length, together with

(54) MacKerell, A. D.; Banavali, N.; Foloppe, N. *Biopolymers* **2000**, *56*, 257−265.
(55) Harvey, J. N. *Faraday Discuss.* **2004**, *127*, 165−177.
(56) Reuter, N.; Dejaegere, A.; Maigret, B.; Karplus, M. *J. Phys. Chem. A* **2000**, *104*, 1720−1735.
(57) Kamachi, T.; Yoshizawa, K. *J. Am. Chem. Soc.* **2003**, *125*, 4652−4661.
(58) Hata, M.; Hirano, Y.; Hoshino, T.; Tsuda, M. *J. Am. Chem. Soc.* **2001**, *123*, 6410−6416.
(59) Yoshizawa, K.; Kamachi, T.; Shiota, Y. *J. Am. Chem. Soc.* **2001**, *123*, 9806−9816.

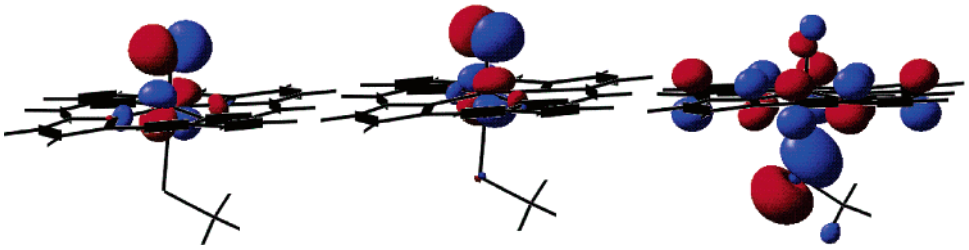


Figure 2. Frontier orbitals in QM/MM optimized P450 2C9 Compound I: (left) two π^* orbitals of Fe−O; (right) porphyrin "a2u" orbital with considerable contributions of sulfur p orbital.

Table 1. Spin Densities[a] and Fe−S Distances [Å] in QM/MM Optimized Compound I of 2C9 for Doublet and Quartet States ($^2A_{2u}/^4A_{2u}$)[b]

	DFT	ρ_S	d(Fe–S)
QM1	B3LYP	−0.39/0.37	2.55/2.55
QM2	B3LYP	−0.38/0.34	2.54/2.54
	BP86	−0.38/0.34	2.54/2.50
	B3PW91	−0.38/0.36	2.50/2.50
QM3	B3LYP	−0.39/0.37	2.55/2.55
	BP86	−0.39/−	2.55
	B3PW91	−0.39/−	2.55
QM4	B3LYP	−0.39/0.36	2.54/2.54
QM5	B3LYP	−0.37/0.35	2.57/2.57
	BP86	−0.33/−	2.44
	B3PW91	−0.37/−	2.52

[a] Negative sign refers to β spin density. [b] Values are given for calculations using different QM regions and different DFT funtionals.

the spin density on sulfur, is reported below for all our QM/MM optimizations as it gives indirect information on the electronic structure of Compound I.

The small QM region (QM1) used here has been used successfully as a model of Compound I in many QM-only studies[23,57−59] and also in QM/MM calculations on P450cam.[21] There has, however, been considerable debate[33,35] as to the importance of including substituents on the heme ring. As the discussion on this topic has previously focused only on P450cam, it is unclear as yet whether small QM regions can be used in QM/MM calculations on human isoforms. The results obtained using QM regions QM2−QM5 are summarized in Table 1. The difference between QM1 and QM2 is fairly small, with the latter including methyl and vinyl side chains on the porphyrin ring. These groups could in principle have an electronic effect due to their electron-donating and resonance character, respectively. However, as in previous work, we found no effect from these groups in our calculations.

The next QM region, QM3, also includes the negatively charged propionate side chains. Again, no change in electronic structure is observed with this larger region, and in particular, no unpaired electronic density is found on the terminal oxygen atoms, which are found instead to have a charge and geometry typical of a deprotonated carboxylate group. When carrying out gas-phase QM-only calculations on QM region QM3, we did observe significant spin density (0.34 unpaired electrons) on both carboxylate groups, but this disappeared upon including the MM environment. This occurred even when the converged gas-phase orbitals (with spin density on oxygen) were used to construct the initial guess for the QM/MM calculation. The lack of spin density on the propionate side chains is in agreement with other calculations on Compound I of P450cam,[33] but differs from results obtained in another recent study.[35]

QM region QM3 has an overall charge of −2, and important salt bridges and hydrogen bonds with amino acid side chains

Table 2. Spin Densities[a] and Fe−S Distances [Å] in QM/MM Optimized P450 2C9 Compound I Doublet and Quartet States ($^2A_{2u}/^4A_{2u}$), with His368 in the Doubly Protonated Form with Different QM Regions of Figure 1

	ρ_S	d(Fe–S)
QM1	−0.34/0.32	2.52/2.52
QM3	−0.35	2.54
QM4	−0.35	2.53

[a] Negative sign refers to β spin density.

in the vicinity of the propionate groups are treated across the QM/MM boundary. To further test the screening of the carboxylate charges by the protein environment, QM region QM4 was constructed. This enlarged core includes the basic residues surrounding the propionate side chains, namely, in the 2C9 structure two Arg residues, one His, and one Trp, which all form hydrogen bonds to the carboxylate groups. With this very large QM region also, no significant change in electronic structure, and no charge transfer from the porphyrin ring to the carboxylate groups, was observed. The results in Table 1 refer to a protonation state in which the hydrogen bonding His was neutral, which was the charge assigned based on the environment of this group. To further probe the effect of the hydrogen bonding environment on the electronic structure of Compound I, the calculations were repeated using a positively charged doubly protonated His side chain, using QM regions QM1, QM3, and QM4. All these calculations gave the same electronic structure for the porphyrin ring (and the propionate side chains) as was observed using the neutral His group (see Table 2).

The final QM region, QM5, was designed to probe the effect of the way in which the important network of hydrogen bonds near the proximal cysteine group is treated. QM region QM5 treats all the cysteine atoms quantum mechanically. As for the other changes to the QM region, this does not lead to a significant change in the electronic structure of Compound I. This means that the truncation of the part of the cysteine residue treated in the QM region to a methylthiolate group is a good approximation.

To summarize these results, we have found no significant changes in electronic structure of the Compound I species in P450 2C9 upon varying the size of the QM region. No unpaired electron density is observed on the propionate side chains in the large QM3 and QM4 systems, which were designed to probe this possibility. The smallest QM region, QM1, therefore seems to be adequate for describing the electronic structure of Compound I in human isoforms of cytochrome P450.

We have also tested other aspects of the computational procedure. First, we explored the effect of changing the DFT functional. This leads to significant changes in electronic structure for a gas-phase treatment of QM1. B3LYP gives a

Table 3. Spin Densities[a] and Fe−S Distances [Å] in QM/MM Optimized 2C9 Compound I for QM Regions QM1 and QM2 Using Different Basis Sets (BSI/BSII)

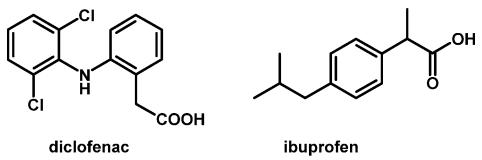
QM region	ρ_S	d(Fe–S)
QM1	−0.39/−0.37	2.55/2.55
QM2	−0.38/−0.36	2.54/2.56

[a] Negative sign refers to β spin density.

large spin density on sulfur, a long Fe−S bond, and a low Fe−S homolytic bond dissociation energy to yield methylthioxy radical (3.3 kcal/mol). The B3PW91 functional, which is similar to B3LYP except for the correlation functional, gives a very slightly larger gas-phase bond dissociation energy of 4.3 kcal/mol, whereas the "pure" BP86 functional gives a significantly larger value of 9.9 kcal/mol. This last value suggests that bonding in this system is described differently using "pure" gradient-corrected density functionals than with the hybrid B3LYP. To examine whether this has an impact on the description of the electronic structure of Compound I, QM/MM geometry optimizations were carried out using BP86 and B3PW91 for 2C9. No major changes in the electronic structure were observed. The spin densities on sulfur are nearly identical to the B3LYP values, and the Fe−S bond length is either identical or shows a slight decrease in some cases (Table 1). When using the QM region QM5, which includes the whole cysteine residue, a slightly more distinct shortening of the Fe−S bond from 2.57 to 2.44 Å, and a decrease in sulfur spin density from 0.37 to 0.33, is observed with BP86. However, this difference is not systematic and the magnitude of the change in spin density and bond length is similar to that obtained using different starting geometries (vide infra), so it is probably not meaningful.

We have also examined the effect of the basis set. Due to computational expense, the small BSI was used for most of our QM/MM calculations. Similar-sized basis sets have been shown to give good results in previous computational studies of Compound I.[21,33,45] To test that this is the case here also, two QM/MM geometry optimizations of P450 2C9 Compound I were carried out with the larger basis set BSII. The results for QM regions QM1 and QM2 were found to be nearly identical to those obtained with the smaller basis set (Table 3). As basis set effects were shown to be negligible in these two cases, it is unlikely that the results for the other P450 isoforms studied here will depend strongly on the basis set.

It has been suggested that typical MM van der Waals parameters are not suitable to describe nonbonded interactions between MM and QM atoms at the QM/MM boundary.[60,61] This is because these parameters were optimized to describe nonbonded interactions in the absence of polarization effects, which are however present in part in the QM/MM calculations. Other studies have found that optimization of these nonbonded parameters leads to only negligible effects on energetic observables.[62] Nevertheless, we tested the effect of increasing by 5% the van der Waals radii on the S atom, the propionate oxygens, the ferryl oxygen, and the porphyrin ring in both the QM1 and QM3 QM/MM models for 2C9. Geometry optimization with these modified parameters leads as expected to larger hydrogen-bond distances (e.g., the Val436 backbone N−H to sulfur distance increases from 2.43 to 2.53 Å), but to no significant change in either the spin density on sulfur (−0.394 and −0.393 with standard and modified parameters, respectively) or the optimized Fe−S distance (2.546 and 2.540 Å, respectively).



diclofenac ibuprofen

Figure 3. 2C9 substrates diclofenac and ibuprofen used in calculations of the electronic structure of Compound I in a drug complexed structure.

Previous work on Compound I in P450cam[63] suggests that the details of the system setup may play a role in defining its electronic structure. The previous calculations were all carried out using a solvated cytochrome P450 2C9 system with the atomic coordinates derived from X-ray crystallographical structure 2C9•A. To test the effect of the system setup, we also carried out a new QM/MM minimization of Compound I for the same 2C9 enzyme, but starting from a geometry obtained after equilibration for 200 ps at 310 K of all the atoms in a 25 Å sphere around iron. This equilibration does not lead to any significant change in the tertiary structure of the protein, but does lead to a large number of small conformational changes.[64] After QM/MM minimization, the final structure derived from this equilibration has a somewhat different electronic structure from that of the unequilibrated 2C9•A system: the spin density on sulfur has increased from −0.38 to −0.45 and the Fe−S bond length has increased from 2.54 to 2.63 Å. Similar changes in electronic structure were observed upon QM/MM optimization starting from different snapshots along an MD pathway for P450cam.[21]

Most of the calculations were carried out on Compound I in the absence of a substrate, with only MM water molecules in the active site. However, during the real catalytic cycle, a substrate will be present, and this may affect the electronic structure of Compound I, even in the prereactive complex. To explore this effect, we have carried out QM/MM geometry optimizations on complexes with two typical substrates, diclofenac and ibuprofen. For the diclofenac (Figure 3) complex of 2C9•A, the dichloroaminophenyl group of diclofenac and the atoms contained in QM1 were treated quantum mechanically (57 QM atoms). In the ibuprofen (Figure 3) complex of 2C9•B, the QM region also consisted of the QM1 atoms and the isopropyl group of ibuprofen (54 QM atoms). This way the interaction between the substrate and Compound I could be accounted for quantum mechanically, which is likely to be more accurate than considering the effect of the substrates by means of force-field charges on the MM atoms.

The substrates were equilibrated in the active site by MM MD simulations, and a structure was chosen, where the substrates were in a position favorable for reaction with the ferryl oxygen. Although the spin densities on sulfur vary between the two complexes and compared to the substrate-free structure, the observed changes are fairly modest (Table 4). From inspecting the final structures, it appears unlikely that these changes are due to an electronic effect of the substrate on the

(60) Murphy, R. B.; Philipp, D. M.; Friesner, R. A. *J. Comput. Chem.* **2000**, *21*, 1442−1457.
(61) Freindorf, M.; Gao, J. L. *J. Comput. Chem.* **1996**, *17*, 386−395.
(62) Riccardi, D.; Li, G. H.; Cui, Q. *J. Phys. Chem. B* **2004**, *108*, 6467−6478.
(63) Shaik, S.; Kumar, D.; de Visser, S. P.; Altun, A.; Thiel, W. *Chem. Rev.* **2005**, *105*, 2279−2328.
(64) Elber, R.; Karplus, M. *Science* **1987**, *235*, 318−321.

Table 4. Spin Densities[a] and Fe−S Distances [Å] in QM/MM Optimized Substrate-Free 2C9 Compound I and in Substrate-Bound 2C9 Compound I[b]

substrate	ρ_S	d(Fe–S)
none ($^2A_{2u}/^4A_{2u}$)−2C9•A	−0.39/0.37	2.55/2.55
diclofenac ($^2A_{2u}$)−2C9•A	−0.32	2.50
ibuprofen ($^4A_{2u}$)−2C9•B	0.44	2.58

[a] Negative sign refers to β spin density. [b] Values for the doublet and quartet A_{2u} states for diclofenac and ibuprofen complexes, respectively, are shown.

Table 5. Spin Densities[a] and Fe−S Distances [Å] in QM/MM Optimized Diclofenac Complexes of 2C9 Compound I for Doublet A_{2u} States[b]

orientation	2C9 structure	ρ_S	d(Fe–S)
A1	2C9•A	−0.36	2.54
A2		−0.32	2.50
A3		−0.39	2.52
A4		−0.42	2.53
A5		−0.26	2.49
B1	2C9•B	−0.36	2.53
B2		−0.39	2.53
B3		−0.33	2.53

[a] Negative sign refers to β spin density. [b] Values for complexes using different initial 2C9 X-ray structures and different orientations of the diclofenac molecule in the active site are shown.

Compound I center, as there are no close contacts or observable structural deformations arising due to the presence of the substrate in what is a fairly large binding site. It is much more likely that the small variation in electronic structure is simply due to the changes in enzyme conformation arising from the fact that the systems are set up and extensively equilibrated independently of each other—the effect of system setup and equilibration has already been noted above. This interpretation is supported by the fact that the calculated electronic structure varies to an extent similar to that in calculations carried out on the 2C9•A Compound I in the presence of diclofenac, but using starting structures derived from docking the substrate in a slightly different way before equilibration. These results are shown in Table 5, together with the associated results for a diclofenac−2C9 complex derived from the other crystal structure (2C9•B).

In summary, then, our QM/MM calculations on Compound I of cytochrome P450 2C9 show that the electronic structure of this key intermediate is independent of the choice of QM region, the DFT functional used, and the presence or absence of a substrate in the active site. The small differences observed from one calculation to another are due to slight differences in system setup and cannot be attributed to any chemically meaningful cause.

Compound I of Cytochrome P450 Isoforms. The work discussed above has shown that QM/MM geometry optimization using the small QM region QM1, the medium-sized basis set BSI, and the standard B3LYP DFT functional gives a consistent account of the electronic structure of the Compound I intermediate for one typical human cytochrome P450 isoform, i.e., 2C9. We have therefore used the same procedure to probe the electronic structure of this intermediate in a variety of other substrate-free isoforms, to examine whether the different reactivities of these different enzymes can be attributed to different electronic structures of the oxidizing species. We have examined enzymes from two further cytochrome P450 subfamilies, 2B4 and 3A4, and also compared our results with those obtained for the bacterial P450cam. For the mammalian isoforms, the effect of varying the enzyme conformation was also investigated. Thus, Compound I of P450 2B4 was optimized at the QM/MM level, starting from both the crystal structure itself and a geometry obtained after equilibration using an MD simulation (2B4 MD). For the 3A4 isoform, two different crystal structures were available, 3A4•A and 3A4•B, and the electronic structure in both forms was studied. A third conformation of 3A4 was obtained by equilibrating the crystal structure 3A4•B in an MD simulation (3A4•B MD).

Table 6. Spin Densities[a] and Fe−S Distances [Å] in Compound I of P450 Isoforms 2B4 and 3A4 and for Bacterial P450cam, for Doublet and Quartet States: $^2A_{2u}/^4A_{2u}$[b]

P450 structure	ρ_S	d(Fe–S)
2B4	−0.33/0.32	2.51/2.52
2B4 MD	−0.41	2.62
3A4•A	−0.39/0.37	2.53/2.53
3A4•B	−0.47/0.45	2.55/2.56
3A4•B MD	−/0.50	−/2.62
cam	−0.20/0.20	2.53/2.54

[a] Negative sign refers to β spin density. [b] QM/MM optimized geometries for 3A4 and 2B4 are based on different crystal structures and on structures equilibrated in MD simulations.

In all cases, an electronic structure very similar to that found in the 2C9 form, i.e., with A_{2u} character, was obtained. As already noted above, some changes in the unpaired electron distribution between the porphyrin ring and the proximal cysteine sulfur are observed when varying the conformation of the enzyme. This is true for both P450 subfamilies (Table 6). Compound I in the equilibrated structure of P450 2B4 shows more delocalization of spin density onto the sulfur atom, and therefore also a longer Fe−S bond, than Compound I in the 2B4 crystal structure. A similar observation was made for the 3A4 isoform. The spin densities on sulfur differ when comparing the Compound I intermediates derived from the two X-ray structures. Likewise, the Compound I characterized using the structure derived from MD equilibration shows a slightly different spin density distribution. It is also to be noted that, compared to these human isoforms, we found Compound I in bacterial P450cam to have a somewhat lower spin density on the thiolate sulfur, but a comparable Fe−S bond length. Our results for this system are similar to those already reported in the literature by others using the same QM region.[21]

These results show that the electronic structure of Compound I is somewhat dependent on its environment. However, the differences obtained from one isoform to another are not significantly larger than those derived for the same isoform but using a different system setup, as discussed above in the case of 2C9. We note also that the observed electronic structure, as measured by the spin density on sulfur, at the beginning of our QM/MM minimizations (i.e., at the geometry of the crystal structure, or at an MM optimized structure) often differs from that of the final optimized structure by an amount similar to that found from one enzyme to another or from one system setup to another. This again suggests that, as the protein, solvent, and substrate atoms of a given cytochrome P450 Compound I move around due to thermal motion, the electrostatic effects of the environment on the porphyrin ring and the iron cysteinate ligand (and perhaps also the strain exerted on the Fe−S bond) fluctuate slightly, thereby causing small changes in electronic structure. To fully probe the magnitude of these changes, long

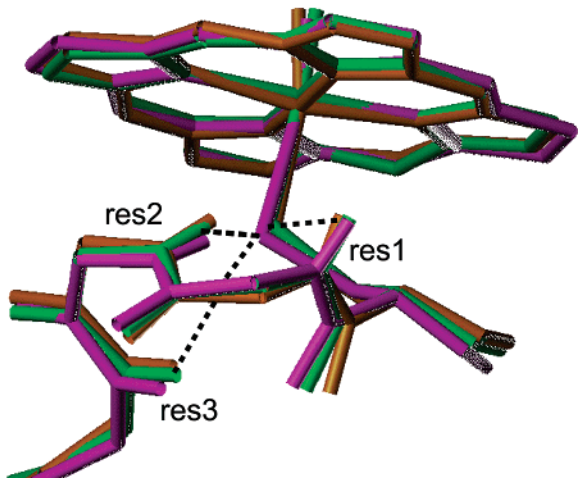


Figure 4. Superimposition of residues res1−3 surrounding the coordinating cysteine in the crystal structures of P450cam, P450 2C9, and P450 2B4. H bonds to the cysteine sulfur are indicated by dotted lines.

Table 7. Residues in P450cam, 2C9, 2B4, and 3A4 Involved in the H-bond Network with the Thiolate Ligand

P450 isoform	Cys/res1/res2/res3
cam	Cys357/Leu358/Gly359/Gln360
2C9	Cys435/Val436/Gly437/Glu438
2B4	Cys436/Leu437/Gly438/Glu439
3A4	Cys442/Ile443/Gly444/Met445

molecular dynamics simulations using the QM/MM method would be required, and this is as yet impossible due to the enormous computational expense that would be associated with such calculations. However, based on the results obtained using different "snapshots" and using different starting crystal structures, these thermal changes appear to be as large as the overall changes in electronic structure between different human isoforms.

It has been proposed that the extent of electron delocalization onto the thiolate ligand and the strength of the Fe−S bond are controlled by the local H-bonding environment of the cysteine.[29,30] We have therefore examined whether notable changes in this hydrogen bonding environment are observed from one isoform to another. In P450cam, the cysteine is surrounded by backbone amide protons of the subsequent three protein residues Leu358, Gly359, and Gln360 (Table 7), which are located at the positive N-terminal end of a helix.[65] This structural feature is in fact very well conserved among the human isoforms studied here (Figure 4). In all enzyme structures used for QM/MM calculations the cysteine sulfur is enclosed within three N−H groups provided by the protein backbone. The equivalent residues in the various isoforms are given in Table 7.

A comparison of the residues around cysteine in the QM/MM optimized structures of P450 2C9, 2B4, and 3A4 with that found for P450cam shows that there are only slight variations in the NH−S distances (Table 8). Among the human isoforms, all NH−S distances in the QM/MM optimized structures are nearly identical. As seen for P450 3A4, this is also independent of the chosen crystal structure and conformation, and thus does not explain the observed variation of spin density on sulfur found in the different calculations. There is, however, a difference between P450cam and all the mammalian cytochromes concerning the length of the hydrogen bond between the third amide proton (H_{res3}) and the cysteine sulfur. This distance is only 2.6 Å in the QM/MM optimized structure of P450cam, while it is increased to 3.3−3.6 Å in the structures of the human isoforms.

Table 8. Distances $d(NH_{resx}-S)$ [Å] between the Amide Proton of Residues 1−3 and the Cysteine Sulfur and Spin Densities on Sulfur in $^{2,4}A_{2u}$ Compound I for QM/MM Optimized (QM Region QM1) Structures of Mammalian Isoforms and P450cam

structure	res1	res2	res3	ρ_S
cam	3.29	2.33	2.59	−0.20
2C9	3.31	2.43	3.44	−0.39
2B4	3.30	2.34	3.31	−0.33
3A4•A	3.38	2.52	3.56	−0.39
3A4•B	3.39	2.54	3.57	−0.47
3A4•B MD[a]	3.49	2.43	3.39	0.50

[a] Structure of Compound I obtained from MD equilibration and subsequent QM/MM optimization.

Mutations in P450cam,[30] which replaced the residue at this position with proline, which does not have an amide proton, have indicated that this amide proton plays a critical role in stabilizing the coordination of the thiolate ligand to the heme iron. By contrast, the first residue (res1) directly adjacent to cysteine was found not to play an important role.[30] Therefore, it might be possible that the more favorable position of the third amide proton in P450cam leads to the significantly lower spin density on sulfur in the electronic structure of its Compound I compared to all the human isoforms.

This suggests that the specific enzyme environment can indeed have an influence on the electronic structure of Compound I, in terms of the delocalization of spin density onto the thiolate ligand. However, the mammalian cytochromes P450 studied here are all very similar, and our calculations do not show any clear differences between the electronic structures of the Compounds I in the various mammalian isoforms.

Conclusions

The electronic and geometric properties of the active species Compound I in the specific enzyme environment of the human cytochrome P450 isoforms were investigated by QM/MM calculations at the B3LYP/CHARMM27 level. One representative of each drug-metabolizing cytochrome P450 subfamily 2B, 2C, and 3A was studied based on the X-ray structures that have recently become available.[36−40] In all cases, QM/MM geometry optimizations predicted the lowest electronic states to be a near-degenerate pair of doublet/quartet A_{2u} states, each having two unpaired electrons on the Fe−O moiety with a third unpaired electron located in the a_{2u} orbital, which is delocalized over the porphyrin and the thiolate ligand. Unlike some recent studies[35,66] on the related Compound I intermediate of P450cam, we did not find any delocalization of the third unpaired electron onto the carboxylate groups of the heme propionate side chains.

Extensive studies on the P450 2C9 isoform revealed that the overall features of the computed electronic structure of Compound I are independent of (i) the inclusion of heme substituents in the QM region, (ii) the truncation of the thiolate ligand in the QM region, (iii) the DFT functional, (iv) the protonation state of the histidine residue His368 close to the heme group, and (v) the presence or absence of a substrate in the active site. This suggests that our computational results should be fairly

(65) Poulos, T. L. *J. Bioinorg. Chem.* **1996**, *1*, 356−359.

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

reliable, as they are not sensitive to the details of how the calculation is set up.

Although the electronic structure of Compound I shows the same general characteristics in all P450 isoforms, there are significant variations in the amount of delocalization of the third unpaired electron onto the axial thiolate ligand, with the spin density on sulfur varying from 0.20 to 0.50. In this sense, our calculations confirm the previously noted[11,21] chameleonic nature of the Compound I intermediate in that the calculated electronic structure is found to vary substantially with relatively small changes in the environment of the heme group. Changes are observed between the different isoforms, between different X-ray structures of a given isoform, and between different conformations obtained from MD simulation for a given starting X-ray structure. However, the changes from one isoform to another are not larger than those obtained in different calculations on the same isoform, so the electronic structures of Compound I in the different human isoforms are not distinguishable. Thus, from the present study, it appears unlikely that the very small differences in the electronic structure of Compound I among the human cytochromes P450 can have a big influence on the specific substrate selectivity of each isoform. It cannot, however, be ruled out that more significant changes occur for the corresponding transition states. Also, there does seem to be a noticeably lower amount of electron delocalization onto the thiolate ligand in the Compound I species of P450cam compared to that obtained for all of the mammalian P450 isoforms. This might be due to a stronger hydrogen bond of an amide proton to the iron-coordinating sulfur in P450cam, which reduces the electron donor capability of the thiolate ligand, so the unpaired electron is more heavily located on the porphyrin ligand.

In summary, the present QM/MM calculations reveal the properties of Compound I in the human cytochromes P450 and address factors that can influence the electronic structure of Compound I. This has not been studied by computational means before and can help investigate the metabolism of drugs in the human cytochromes P450 by an improved understanding of their active species.

Acknowledgment. We acknowledge support from the EPSRC (J.N.H., Advanced Research Fellowship) and the EU (MCInet research and training network).

Supporting Information Available: Complete ref 50, details of system setup, and additional force-field parameters used for the heme group, diclofenac, and ibuprofen. This material is available free of charge via the Internet at http://pubs.acs.org.

JA0520924