J|A|C|S
ARTICLES
Published on Web 06/14/2002

The Elusive Oxidant Species of Cytochrome P450 Enzymes: Characterization by Combined Quantum Mechanical/Molecular Mechanical (QM/MM) Calculations

Jan C. Schöneboom,[†] Hai Lin,[†] Nathalie Reuter,[†] Walter Thiel,*[,†] Shimrit Cohen,[‡] François Ogliaro,[‡] and Sason Shaik*[,‡]

Contribution from the Max-Planck-Institut für Kohlenforschung, Mülheim an der Ruhr, Germany, and Department of Organic Chemistry and the Lise Meitner Center for Computational Quantum Chemistry, The Hebrew University, 91904 Jerusalem, Israel

Received March 22, 2002

Abstract: The primary oxidant of cytochrome P450 enzymes, Compound I, is hard to detect experimentally; in the case of cytochrome $P450_{cam}$, this intermediate does not accumulate in solution during the catalytic cycle even at temperatures as low as 200 K (ref 4). Theory can play an important role in characterizing such elusive species. We present here combined quantum mechanical/molecular mechanical (QM/MM) calculations of Compound I of cytochrome $P450_{cam}$ in the full enzyme environment as well as density functional studies of the isolated QM region. The calculations assign the ground state of the species, quantify the effect of polarization and hydrogen bonding on its properties, and show that the protein environment and its specific hydrogen bonding to the cysteinate ligand are crucial for sustaining the Fe−S bond and for preventing the full oxidation of the sulfur.

Introduction

Cytochrome P450 enzymes constitute a superfamily of monooxygenases that appear in all bioorganisms and perform vital bioregulatory functions such as detoxification of xenobiotics as well as biosynthesis of sex hormones, muscle-relaxing, antiflammatory, and antihypertensive compounds.[1] These enzymes present an exciting agenda for mechanistic research that is aimed at a detailed understanding of all elementary steps in the respective catalytic cycles.[1−7] Much of the current debate focuses on the key intermediates and transformations shown in Figure 1.

It is commonly accepted that the active oxidant is Compound I (Cpd I)[8,9] (**1** in Figure 1). Decades of ingenious experimental work have provided a rich gallery of model systems,[1] but a definitive characterization of this putative species for P450 itself proved to be difficult because of the fast reaction steps occurring after O_2 binding. Using cryogenic X-ray diffraction and trapping techniques, Schlichting et al.[7] obtained electron density data that would be consistent with the formation of an oxyferryl species in cytochrome $P450_{cam}$ (CYP101). While the restraints of the

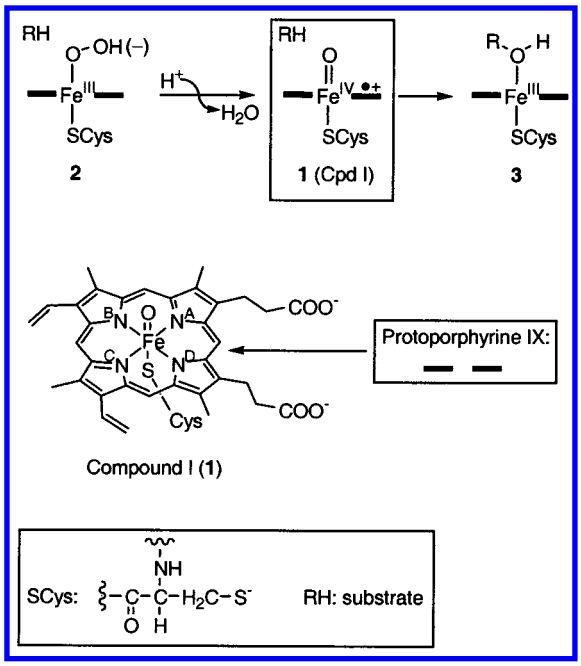


Figure 1. Key steps in the catalytic cycle of cytochrome P450. Cpd I (**1**) is drawn explicitly below the reaction scheme.

* To whom correspondence should be addressed. W.T.: e-mail, thiel@mpi-muelheim.mpg.de; phone, +49-208-306-2150; fax, +49-208-306-2996. S.S.: e-mail, sason@yfaat.ch.huji.ac.il; phone, +972-2-6585909; fax, +972-26584680.

[†] Max-Planck-Institut für Kohlenforschung.

[‡] The Hebrew University.

(1) Ortiz de Montellano, P. R., Ed. *Cytochrome P450: Structure, Mechanisms, and Biochemistry*, 2nd ed.; Plenum Press: New York, 1995; Vol. 2.
(2) Mueller, E. J.; Lioda, P. J.; Sligar, S. G. In *Cytochrome P450: Structure, Mechanisms, and Biochemistry*, 2nd ed.; Ortiz de Montellano, P. R., Ed.; Plenum Press: New York, 1995; Vol. 2, Chapter 3.
(3) Sono, M.; Roach, M. P.; Coulter, E. D.; Dawson, J. H. *Chem. Rev.* **1996**, *96*, 2841.
(4) Davydov, R.; Makris, T. M.; Kofman, V.; Werst, D. E.; Sligar, S. G.; Hoffman, B. M. *J. Am. Chem. Soc.* **2001**, *123*, 1403.
(5) Meunier, B.; Bernadou, J. *Struct. Bonding* **2000**, *97*, 1.
(6) Groves, J. T.; Han, Y. Z. In *Cytochrome P450: Structure, Mechanisms, and Biochemistry*, 2nd ed.; Ortiz de Montellano, P. R., Ed.; Plenum Press: New York, 1995; Vol. 2; Chapter 1.
(7) Schlichting, I.; Berendzen, J.; Chu, K.; Stock, A. M.; Maves, S. A.; Benson, D. A.; Sweet, R. M.; Ringe, D.; Petsko, G. A.; Sligar, S. G. *Science* **2000**, *287*, 1615.
(8) Groves, J. T.; Watanabe, Y. *J. Am. Chem. Soc.* **1988**, *110*, 8443.
(9) Groves, J. T.; McMurry, T. J. *Rev. Port. Quim.* **1985**, *27*, 102.

10.1021/ja026279w CCC: $22.00

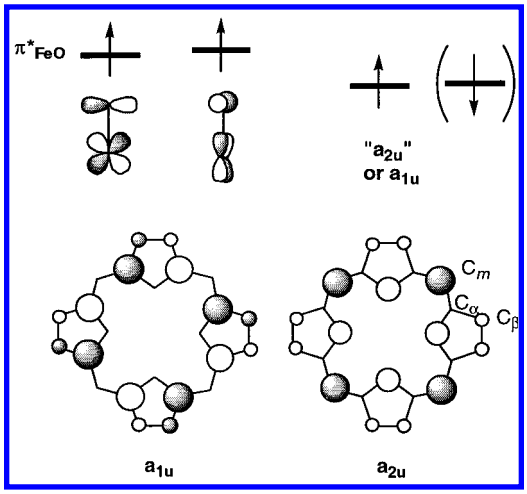


Figure 2. General description of the triradicaloid states of Cpd I. Two electrons occupy π^* (FeO) orbitals. The third one is located in an "a_{2u}" or a_{1u} orbitals of the porphyrin. "a_{2u}" involves a sulfur ligand contribution. The pure a_{1u} and a_{2u} MOs are shown above. The positions C_α, C_β, and C_m are indicated on the a_{2u} orbital.

crystal lattice and the use of an unnatural electron source (X-ray radiolysis) might, in principle, allow for the accumulation of Cpd I under these special conditions, concern about this structure arises because a mixture of different species is likely to be present, and other interpretations remain possible. Reports of a tentative assignment of the species using rapid scan absorption spectroscopy[10] were questioned recently by Davydov et al.[4] whose ENDOR and EPR spectroscopic studies showed that in a frozen solution, Cpd I of CYP101 does not accumulate at a detection temperature of 200 K (nor even below this temperature). The product directly obtained was the iron(III)-hydroxy-camphor complex (**3** in Figure 1), the C-5 hydrogen of camphor being trapped in the hydroxyl group of hydroxy-camphor according to product inventory analysis.[4] These facts were interpreted as being consistent with monooxygenation of camphor by Cpd I (**1**). In a recent study, Kellner et al.[11] measured the kinetics of formation and decay of Cpd I in the thermostable cytochrome P450 CYP119 from *Sulfolobus solfataricus*, which is thought to have a more rigid active-site structure and thus different relative reaction rates than its mesophilic counterpart. The ferryl-oxo-(π) porphyrin cation radical was identified by its characteristic spectral features.[11] In summary, the available experimental evidence indicates that the key species in the catalytic mechanism of the isozyme CYP101 exists but eludes detection because of the special kinetic scenario of the mechanism.

All known model and enzymatic Cpd I species are triradicaloids,[1,3,5,6,8,9,12] shown in Figure 2, that exhibit two closely lying quartet and doublet spin states. These arise from the weak coupling of a triplet pair on the FeO moiety with a third electron, which in most cases resides in a porphyrin-based orbital of either a_{2u} or a_{1u} character. The electronic states of Cpd I are generally classified according to the identity of the latter orbital as either $^{4,2}A_{2u}$ or $^{4,2}A_{1u}$. Many of the known Cpd I species are "green", belonging to the A_{2u} type. Experimental observations of Cpd I of chloro peroxidase (CPO), a cysteinate enzyme analogous to P450,[12a,13−18] suggest that it has a doublet ground state with a quartet state lying higher by ca. 37 cm^{-1}, but there does not seem to be a consensus on the identity of the state. Resonance Raman studies by Hosten et al.[16] indicate that CPO(I) is the "gray" A_{1u} state, with an unpaired electron in the a_{1u} orbital of porphyrin. In contrast, EPR studies by Rutter et al.[17] assign it as an A_{2u} state. Mössbauer and EPR studies indicate a state of a mixed A_{1u}/A_{2u} character with some delocalization of the orbital on the sulfur atom of the cysteinate ligand.[12a] Thus, the experimental characterization leads to uncertainty about the electronic nature of the species. More so, nothing is known about the geometry of Cpd I other than the X-ray structure by Schlichting et al.,[7] which, as mentioned above, probably suffers from contamination with other species in the crystal.

This situation invites theory as a viable method that can characterize this elusive species. Yet, theory encountered its own difficulties, associated with the size of Cpd I, and with the need to go beyond isolated molecule calculations.[19−23] Density functional (DFT) studies of different simplified models with pure and hybrid functionals depend on the representation of the cysteinate ligand and the protoporphyrin IX. Thus, the cysteinate ligand, which is part of a polypeptide chain, was modeled by HS^-, CH_3S^-, and by the anion of the cysteine amino acid. Similarly, the native protoporphyrin IX was modeled by a pristine porphine, by octamethyl porphyrin, as well as by the native ligand itself. While the porphyrin substituents did not seem to make much of a difference, the various truncated cysteine-ligand models generated conflicting information regarding the Fe−S bond lengths and the distribution of the unpaired electron density that typifies the ground state of the species. Thus, the Fe−S bond depends on the model ligand and ranges from ca. 2.72 to 2.37 Å.[19−23] Similarly, calculations using CH_3S^- found a ground state that was neither A_{1u} nor A_{2u}, but rather a π-sulfur state ($^{4,2}\Pi_S$) in which the third electron of the triradicaloid was to a large extent (>80%) located in the π-type lone pair of the sulfur.[18,20,22] In other cases, the unpaired electron was reported to be located more on the porphyrin, in an a_{2u}-type mixed porphine-sulfur orbital.[19−21] These differences were shown to reflect the donor capabilities of the different ligand models vis-à-vis the porphyrin iron-oxo moiety.[21a] Two studies

(10) Egawa, T.; Shimada, H.; Ishimura, Y. *Biochem. Biophys. Res. Commun.* **1994**, *201*, 1464.
(11) Kellner, D. G.; Hung, S.-C.; Weiss, K. E.; Sligar, S. G. *J. Biol. Chem.* **2002**, *277*, 9641.
(12) (a) Weiss, R.; Mandon, D.; Wolter, T.; Trautwein, A. X.; Müther, M.; Bill, E.; Gold, A.; Jayaraj, K.; Terner, J. *J. Biol. Inorg. Chem.* **1996**, *1*, 377. (b) Watanabe, Y.; Fujii, H. *Struct. Bonding* **2000**, *97*, 61. (c) Mandon, D.; Weiss, R.; Jayaraj, K.; Gold, A.; Terner, J.; Bill, E.; Trautwein, A. X. *Inorg. Chem.* **1992**, *31*, 4404.
(13) Placic, M. M.; Rutter, R.; Araiso, T.; Hager, L. P.; Dunford, H. B. *Biochem. Biophys. Res. Commun.* **1980**, *94*, 1123.
(14) Schulz, C. E.; Rutter, R.; Sage, J. T.; Debrunner, P. G.; Hager, L. P. *Biochemistry* **1984**, *23*, 4743.
(15) Rutter, R.; Hager, L. P.; Dhonau, H.; Hendrich, M.; Valentine, M.; Debrunner, P. *Biochemistry* **1984**, *23*, 6809.
(16) Hosten, C. M.; Sullivan, A. M.; Palaniappan, V.; Fitzgerald, M. M.; Terner, J. *J. Biol. Chem.* **1994**, *269*, 13966.
(17) Rutter, R.; Valentine, M.; Hendrich, M. P.; Hager, L. P.; Debrunner, P. G. *Biochemistry* **1983**, *22*, 4769.
(18) Antony, J.; Grodzicki, M.; Trautwein, A. X. *J. Phys. Chem. A* **1997**, *101*, 2692.
(19) For general reviews, see: (a) Loew, G. H.; Harris, D. L. *Chem. Rev.* **2000**, *100*, 407. (b) Harris, D. L. *Curr. Opin. Chem. Biol.* **2001**, *5*, 724. (c) Harris, D. L.; Loew, G.; Waskell, L. *J. Inorg. Biochem.* **2001**, *83*, 309. (d) Harris, D. L.; Loew, G. H. *J. Porphyrins Phthalocyanines* **2001**, *5*, 334.
(20) Yoshizawa, K.; Kagawa, Y.; Shiota, Y. *J. Phys. Chem. B* **2000**, *104*, 12365.
(21) (a) Ogliaro, F.; Cohen, S.; Filatov, M.; Harris, N.; Shaik, S. *Angew. Chem., Int. Ed.* **2000**, *39*, 3851. (b) Ogliaro, F.; Cohen, S.; de Visser, S. P.; Shaik, S. *J. Am. Chem. Soc.* **2000**, *122*, 12892. (c) Ogliaro, F.; de Visser, S. P.; Cohen, S.; Kaneti, J.; Shaik, S. *ChemBioChem* **2001**, *2*, 848.
(22) (a) Green, M. T. *J. Am. Chem. Soc.* **1999**, *121*, 7939−7940. (b) Green, M. T. *J. Am. Chem. Soc.* **2000**, *122*, 9495.
(23) Ohta, T.; Matsuura, K.; Yoshizawa, K.; Morishima, I. *J. Inorg. Biochem.* **2000**, *82*, 141.

of the cysteinate model[21a,c,23] ligand led to different spin densities on the sulfur (65% vs 87%) and different ground-state character ($^2A_{2u}$ vs $^2\Pi_S$). This discrepancy originates from the different conformations of the cysteinate, one with internal hydrogen bonds and the other without them, which greatly affects its donor capability.[24] While all of these differences are by now well understood, they nevertheless underscore the fact that calculations using truncated and gas-phase models cannot resolve the dilemma of Cpd I, and one must resort to more realistic models which involve the electric field and structural constraints of the protein pocket.

Structural studies of P450 enzymes show that the sulfur atom of the coordinating cysteinate in the native species is coordinated by three NH- - -S^- hydrogen bonds with Leu_{358}, Gly_{359}, and Gln_{360} (corresponding to the numbering system of $P450_{cam}$), and is exposed to the positive ends of dipoles in the protein pocket.[7,25,26] Mutation studies proved that the Fe−S bond is stabilized only if the protein pocket is sufficiently polar and supplies NH- - -S^- hydrogen bonding to the cysteinate sulfur.[27] Another mutation study showed that severing of a single NH- - -S^- hydrogen bond can have a significant effect on all aspects of the P450 enzyme.[28] These features are therefore critical for a reliable representation of Cpd I. Preliminary theoretical studies by some of us indicated that medium polarization and hydrogen bonding to the thiolate ligand exert a dramatic effect on the properties of Cpd I in two model systems, one with porphyrin and HS^- and the other with octamethyl porphyrin and a cysteinate anion.[21b,c] Thus, it was found that by taking into account hydrogen bonding, the ground state of Cpd I changes from a dominant sulfur radical type to a dominant porphyrin radical type. Medium polarization further increases this tendency but mostly serves to shorten the Fe−S bond, by as much as 0.1 Å. It was postulated that Cpd I is a *chameleon state that adapts its electronic structure and geometry to the protein environment into which it has to accommodate*.[21b,c] Nevertheless, the simplicity of the models used to represent Cpd I naturally leaves doubts whether small models can indeed represent the real species. Clearly, a definitive assignment of the electronic state and geometry of Cpd I, for P450, requires a realistic modeling of Cpd I within its protein environment.

At present, combined quantum mechanical/molecular mechanical calculations (QM/MM)[29] are among the most realistic methods. In these calculations, a suitable truncated Cpd I species is treated with a QM method, while the protein environment is treated with MM. This paper presents such QM/MM calculations on Cpd I of the enzyme cytochrome $P450_{cam}$ that hydroxylates C(5)−H of camphor.[2,3] In this computational approach, the active species of the enzyme feels the electric field and hydrogen-bonding environment of the *specific* protein pocket, as well as the steric constraints exerted on the bonding of the heme to the cysteinate side chain. Such a calculation can, on one hand, define a suitable truncated model, which ought to be used in future mechanistic studies. On the other hand, and more importantly, these calculations can form a useful guide for an eventual experimental characterization of Cpd I of cytochrome $P450_{cam}$.

Computational Methodology

A. Setup of the System. To prepare suitable initial structures for the QM/MM calculations, we started from the experimental X-ray structure for Cpd I reported by Schlichting et al.[7,30] and built a complete model of the solvated enzyme by adding missing hydrogen atoms and solvent water. The system then was relaxed by performing pure force field energy minimizations and molecular dynamics (MD) simulations, using the CHARMM force field[31] as implemented in the CHARMM program,[32] during which the coordinates of the entire heme unit and the coordinating Cys_{357} were kept fixed. The complete system consists of 24 394 atoms, including 16 956 atoms in the solvent. Details of these force field calculations are given in the Supporting Information.

B. Snapshots. In principle, our QM/MM calculations should sample the configurational space of the system by means of a statistical method like MD or Monte Carlo. Because of the high computational effort caused by the required accuracy of the QM treatment, this was not feasible in the present investigation. However, because we focus on the electronic structure of Cpd I alone, it should be sufficient to perform only geometry optimizations for this species. To assess the impact of different enzyme conformations on the electronic structure of the chromophore, we separately investigated three snapshots (after 29, 40, and 50 ps of equilibration) from an MD trajectory obtained in the preparatory force field calculations by means of QM/MM geometry optimizations.

An important structural feature that was discussed by Poulos et al.[25,26] is the hydrogen-bonding situation of the sulfur of Cys_{357} with the backbone NH groups of amino acid residues that form the cysteine loop, Leu_{358}, Gly_{359}, and Gln_{360}. The conformation of these hydrogen bonds in the different snapshots is shown in Figure 3. All snapshots nicely reproduce the hydrogen-bonding situation with the first two residues, but less so with Gln_{360} where snapshot 40 is closest to the experimental structures.[7,25,26]

Another hydrogen bond found in the crystal structure[33] is that between the amide NH_2 group of the Gln_{360} side chain and the backbone carbonyl oxygen of Cys_{357}. This interaction is not present in snapshot 29, and is far less than optimal in snapshots 40 and 50 as compared with the crystal structures.[7,25,26] To probe if the electron-withdrawing effect of this particular hydrogen bond has an impact on the electronic nature of the proximal sulfur ligand, we have manipulated the conformation of this side chain in snapshot 40, which has the conformation closest to the X-ray data, to form this respective hydrogen bond. Snapshot 29 was chosen for QM/MM calibration studies (vide infra), while the comparison of QM/MM results for the three snapshots and the manipulated structure will be used to assess the influence of different enzyme conformations on the electronic properties of the active species.

(24) See method section in: Ogliaro, F.; de Visser, S. P.; Cohen, S.; Sharma, P. K.; Shaik, S. *J. Am. Chem. Soc.* **2002**, *124*, 2806.
(25) Poulos, T. L. *J. Biol. Inorg. Chem.* **1996**, *1*, 356.
(26) Poulos, T. L.; Vickery, J. C.; Li, H. In *Cytochrome P450: Structure, Mechanisms, and Biochemistry*, 2nd ed.; Ortiz de Montellano, P. R., Ed.; Plenum Press: New York, 1995; Vol. 2, Chapter 4.
(27) Sigman, J. A.; Pond, A. E.; Dawson, J. H.; Lu, Y. *Biochemistry* **1999**, *38*, 11122.
(28) Yoshioka, S.; Takahashi, S.; Ishimori, K.; Morishima, I. *J. Inorg. Biochem.* **2000**, *81*, 141.
(29) See, for example: (a) Gao, J. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH Publishers: New York, 1996; Vol. 7, p 119. (b) Gao, J., Ed. *Combined Quantum Mechanical and Molecular Mechanical Methods*; ACS Symposium Series 712; American Chemical Society: Washington, DC, 1998. (c) Monard, G.; Merz, K. M. *Acc. Chem. Res.* **1999**, *32*, 904.
(30) Brookhaven Protein Database, PDB code 1DZ9.
(31) CHARMM22 force field: MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586.
(32) Brooks, B. R.; Burccoleri, R. E.; Olafson, B. D.; States, D. J.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
(33) Yoshioka, S.; Takahashi, S.; Hori, H.; Ishimori, K.; Morishima, I. *Eur. J. Biochem.* **2001**, *268*, 252.

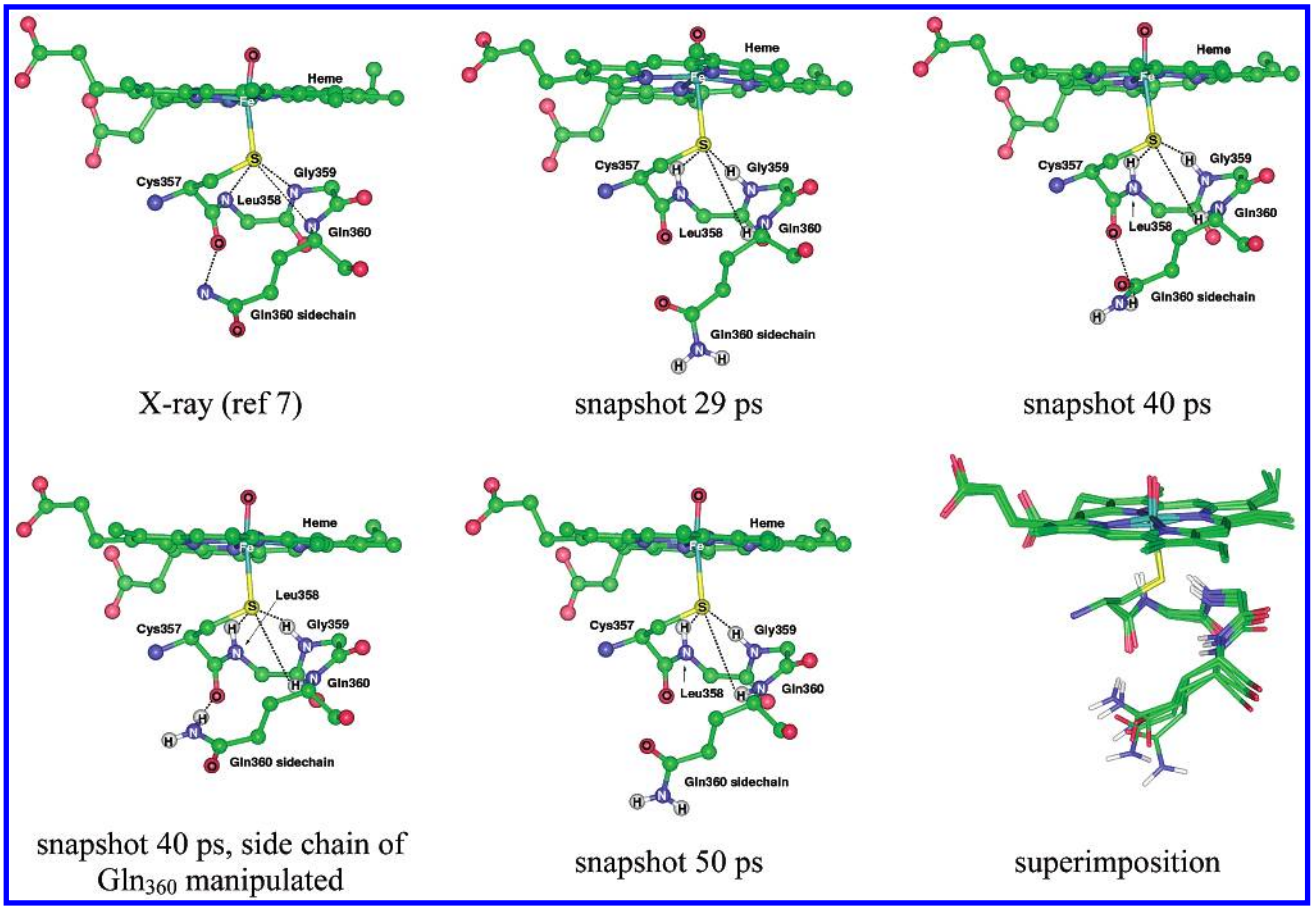


Figure 3. Hydrogen-bonding situations of Cys$_{357}$ in P450$_{cam}$: experimental X-ray structure and the calculated structures (snapshots 29, 40, 50 and snapshot 40 with added hydrogen bond between the side-chain amide group of Gln$_{360}$ and backbone carbonyl oxygen of Cys$_{357}$). Only the hydrogens of the backbone peptide NH groups and the side-chain amide group are shown. The right picture in the second row shows a superimposition of all the structures. For calculated bond distances, see Table 6. Color code: white, hydrogen; green, carbon; dark blue, nitrogen; red, oxygen; yellow, sulfur; light blue, iron.

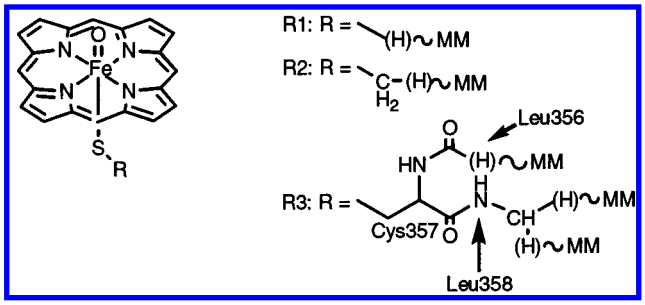


Figure 4. Choice of the QM regions R1−R3.

C. QM/MM Calculations. The QM/MM calculations started from the geometries of the snapshots prepared in the setup phase. Three different QM regions were employed (see Figure 4), all of which include the iron-oxo-porphyrin subunit (without side chains) but differ in the extension to the coordinating cysteine residue. The smallest choice for the sulfur ligand, denoted R1, comprises only the sulfur atom, the next bigger region R2 includes the $C_\beta H_2$ group, and the biggest QM subsystem R3 contains the full Cys$_{357}$ residue, the CO group of Leu$_{356}$, and the NH-C_αH unit of Leu$_{358}$. The covalent bonds cut at the QM/MM border are saturated by hydrogen link atoms, leading to QM subsystems of 40 (R1), 43 (R2), and 57 (R3) atoms. Models R1 and R2 were investigated before[19−23] as naked molecules by QM calculations. These calculations were also performed with a full cysteinate ligand with porphine as well as octamethyl-porphyrin macrocycles.[21a,c] Model system R3 was never investigated before by QM calculations. Thus, the QM and QM/MM results of the present models, R1−R3, will enable us to resolve the difficulties mentioned in the Introduction, and at the same time test the validity of the various model calculations used currently in the literature.

QM Methods. Because Cpd I is a triradicaloid (Figure 2), the QM method of choice should be able to incorporate static and dynamic electron correlation effects. Ab initio multireference methods such as CASSCF or CASPT2 are too demanding for the actual size of the system that has to be considered. Density functional theory in the unrestricted Kohn−Sham formalism (UKS)[34] is computationally more economical and has proven to be successful in the description of Cpd I gas-phase model compounds.[19b] Thus, in line with the model studies described above, the QM method employed in the present calculations is the gradient-corrected B3LYP hybrid density functional.[35] Three basis sets were used; the iron atom is always described by a small-core effective core potential and the associated LACVP basis of a double ζ quality.[36] Basis B1 employs a 6-31G basis on all other atoms. Basis B2 is augmented on the six ligands to iron (four pyrrole nitrogens, the oxo atom, and the sulfur of the coordinating cysteine) with a set of polarization functions (6-31G*), while basis B3 includes also diffuse functions (6-31+G*) on these atoms.[37] The influence of the size of

(34) Low spin (LS) UKS determinants in this work usually suffer from spin contamination because of mixing with higher spin states. The corresponding energy of such determinants can be corrected with a scheme (Yamaguchi, K.; Jensen, F.; Dorigo, A.; Houk, K. N. *Chem. Phys. Lett.* **1988**, *149*, 537) that incorporates the energy of the next higher spin state. Presently, the uncorrected and corrected energies typically differ by less than 0.2 kcal/mol, because the LS and next higher spin state are virtually degenerate. This is because of the weak interaction of the two spins on the FeO unit with the third spin on the ligands. We therefore report uncorrected energies for the LS state.

(35) (a) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098. (b) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785. (c) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(36) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 299.

Table 1. Computed Bond Distances *r* [Å] around the Iron Center of Cpd I (in the ^{4}A/^{2}A States) and Angles θ_{Fe-S-X} [deg] for Different Choices of the QM Region (R1−R3) and Basis Sets (B1−B3)[a]

			QM/MM[b]			gas phase[c]		
entry	variable	basis	R1 (SH)	R2 (SMe)	R3 (cys)	R1 (SH)	R2 (SMe)	R3 (cys)
1	r_{Fe-O}	B1	1.655/1.653	1.654/1.652	1.653/1.651	1.651/1.648	1.646/1.645	1.644/1.643
		B2	1.630/1.628	1.629/1.628	1.628/1.627			
		B3	1.627/1.626	1.626/1.626	1.626/1.625	1.624/1.622	1.620/1.619	1.618/1.617
2	r_{Fe-S}	B1	2.571/2.592	2.576/2.602	2.602/2.629	2.587/2.605	2.655/2.66	2.712/2.722
		B2	2.559/2.591	2.569/2.588	2.590/2.614			
		B3	2.560/2.589	2.565/2.585	2.585/2.609	2.566/2.581	2.622/2.625	2.678/2.697
3	r_{Fe-NA}	B1	2.027/2.026	2.027/2.025	2.025/2.024	2.019/2.018	2.018/2.018	2.009/2.010
		B2	2.031/2.028	2.029/2.027	2.028/2.026			
		B3	2.032/2.030	2.032/2.030	2.031/2.029	2.026/2.024	2.026/2.024	2.017/2.016
4	r_{Fe-NB}	B1	2.038/2.037	2.037/2.035	2.035/2.034	2.016/2.011	2.014/2.011	2.017/2.013
		B2	2.045/2.043	2.044/2.043	2.042/2.041			
		B3	2.049/2.050	2.048/2.047	2.046/2.045	2.024/2.019	2.022/2.019	2.026/2.020
5	r_{Fe-NC}	B1	2.030/2.032	2.028/2.030	2.029/2.031	2.018/2.019	2.019/2.019	2.019/2.019
		B2	2.038/2.041	2.035/2.038	2.037/2.039			
		B3	2.042/2.045	2.038/2.042	2.039/2.043	2.027/2.030	2.029/2.030	2.029/2.030
6	r_{Fe-ND}	B1	2.025/2.027	2.027/2.029	2.027/2.029	2.021/2.026	2.021/2.026	2.015/2.018
		B2	2.029/2.031	2.031/2.033	2.032/2.034			
		B3	2.030/2.031	2.033/2.035	2.034/2.037	2.027/2.034	2.031/2.036	2.022/2.030
7	θ_{Fe-S-X}	B1	111.8/111.6	112.0/111.4	112.2/112.0	98.4/98.4	109.8/109.7	115.7/114.9
		B2	111.5/111.6	112.5/112.2	112.4/112.2			
		B3	111.5/111.2	112.1/111.9	112.0/111.8	97.5/97.5	109.3/109.4	115.5/115.1

[a] R1−R3 refer to the QM model Cpd I systems in Figure 4. Atoms N^{A-D} are defined in Figure 1. [b] B3LYP/CHARMM QM/MM geometry optimization, 544 optimized atoms. [c] B3LYP QM optimization of the isolated QM systems.

the QM region and the basis set was investigated for snapshot 29 only. Snapshots 40, 50, and the manipulated structure of snapshot 40 were studied with the R3 region and B3 basis exclusively.

QM/MM Methods. An electronic embedding scheme[38] was applied; that is, the fixed MM charges were included into the one-electron Hamiltonian of the QM calculation, and the QM/MM electrostatic interactions were evaluated from the QM electrostatic potential and the MM atomic charges. No cutoffs were introduced for the nonbonding MM and QM/MM interactions. To treat the QM/MM boundary, we used hydrogen link atoms[39] with the charge shift model.[40] The approach outlined above was recently applied and validated against others in a theoretical calibration study on triosephosphate isomerase.[41]

Geometry Optimizations. The standard optimized region treated in the QM/MM calculations comprises 544 atoms. We have also tested smaller and larger optimized regions, reaching up to 1900 atoms (defined by including all residues in a sphere of 9 Å around the heme), but did not observe a significant influence on the results as compared to those of the standard region. With the larger optimized region, the conformational complexity related to the many degrees of freedom in some cases led to local minima, which had to be dismissed because they differed in conformations that are not related to Cpd I, but to the MM environment (e.g., the hydrogen-bonding network around the A-propionato chain). To avoid corresponding artifacts, it was decided to restrict the optimizations to the smaller standard region. Further details are given in the Supporting Information.

The main goal of the present paper is to clarify the role of the apoprotein in determining the electronic nature of the active iron-oxo-porphyrin species (Cpd I). To do so, we compare our QM/MM results for the full system to corresponding computations of the isolated QM systems in vacuo, which involve both single-point energy evaluations at the geometry of the QM system in the enzyme environment and complete geometry optimizations. These in vacuo calculations used only the B1 and B3 basis sets as defined above.

The QM programs employed in the QM/MM as well as in the pure QM calculations were TURBOMOLE[42a] and GAMESS-UK.[42b] All QM/MM calculations were performed with the ChemShell package.[43] The CHARMM22 force field[31] run through the DL-POLY[44] program was used for the treatment of the molecular mechanics (MM) part of the system. Geometry minimizations were done with built-in optimizers of ChemShell; gas-phase geometries of model compounds R1 and R2 were optimized with NEWOPT,[43] while all other minimizations employed the HDLC optimizer,[45] applying the default convergence criteria[45] (see Supporting Information).

Results

A. Snapshot 29. As in all gas-phase model calculations,[19−23] there are two closely lying spin states of Cpd I, the high-spin quartet and the low-spin doublet, that correspond to ferromagnetic and antiferromagnetic coupling of electrons in the triradicaloid species. For the time being, we label these states as ^{4}A and ^{2}A. Their precise identity will be assigned at the end.

Geometries. Table 1 summarizes the optimized key geometric parameters around the iron center of the Cpd I species in the ^{4}A and ^{2}A states, for the different models of the QM region (R1−R3) and different basis sets (B1−B3). For comparison, we show also the results for the isolated gas-phase molecules.

(37) (a) Ditchfield, R.; Hehre, W. J.; Pople, J. A. *J. Chem. Phys.* **1971**, *54*, 724. (b) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257. (c) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213. (d) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. v. R. *J. Comput. Chem.* **1983**, *4*, 294.
(38) Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996**, *100*, 10580.
(39) Antes, I.; Thiel, W. In *Hybrid Quantum Mechanical and Molecular Mechanical Methods*; Gao, J., Ed.; ACS Symposium Series 712; American Chemical Society: Washington, DC, 1998; pp 50−65.
(40) de Vries, A. H.; Sherwood, P.; Collins, S. J.; Rigby, A. M.; Rigutto, M.; Kramer, G. J. *J. Phys. Chem. B* **1999**, *103*, 6133.
(41) Lennartz, C.; Schäfer, A.; Terstegen, F.; Thiel, W. *J. Phys. Chem. B* **2002**, *106*, 1758.
(42) (a) Turbomole: Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165. (b) GAMESS-UK is a package of ab initio programs written by M. F. Guest, J. H. van Lenthe, J. Kendrick, K. Schoffel, and P. Sherwood, with contributions from R. D. Amos, R. J. Buenker, H. J. J. van Dam, M. Dupuis, N. C. Handy, I. H. Hillier, P. J. Knowles, V. Bonacic-Koutecky, W. von Niessen, R. J. Harrison, A. P. Rendell, V. R. Saunders, A. J. Stone, D. J. Tozer, and A. H. de Vries. The package is derived from the original GAMESS code of M. Dupuis, D. Spangler, and J. Wendoloski. NRCC, Software Catalog, Vol. 1, Program No. QG01 (GAMESS), 1980.
(43) ChemShell is a modular QM/MM program based on the TCL interpreter and developed in the European QUASI project under the coordination of P. Sherwood. See: http://www.cse.clrc.ac.uk/Activity/ChemShell.
(44) Smith, W.; Forester, T. R. *J. Mol. Graphics* **1996**, *14*, 136.
(45) Billeter, S. R.; Turner, A. J.; Thiel, W. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2177.

The species has C_1 symmetry with four different Fe−N distances, which exhibit small differences on the order of 0.01−0.02 Å. These bond lengths show little dependence on the identity of R, but are slightly expanded (0.01−0.03 Å) in the enzyme pocket vis-à-vis the gas phase. The calculated Fe−O bond distances are almost identical for different R but are slightly (∼0.01 Å) shorter in the isolated molecule than in the enzyme environment. The ^{4}A/^{2}A pairs have virtually the same Fe−O and Fe−N bond distances, because of the fact that the different spin states arise from the rather weak coupling of the FeO electrons to the third unpaired electron. The use of a polarized basis causes a contraction of the Fe−O bond length (by 0.02−0.03 Å) and a slight elongation of the Fe−N distances (by less than 0.01 Å), whereas diffuse functions have no significant effect on these geometrical parameters. In fact, all calculations done so far find these bonds to be least sensitive to the choice of model or basis set.

The computed Fe−S bond lengths in the isolated molecule exhibit strong variations of more than 0.1 Å (as discussed in the Introduction). For any one of the basis sets, model R1 gives the shortest (e.g., 2.587/2.605 Å), and R3 the longest (e.g., 2.712/2.722 Å), Fe−S bond length. In the protein environment, these variations are much less pronounced. Shifting the QM/MM border away from the sulfur atom (i.e., moving from R1 to R3) causes an elongation of the Fe−S bond by 0.02−0.03 Å, while increasing the basis set slightly shortens the bond length, such that the calculated bond lengths vary from a minimum of 2.560/2.589 Å (^{4}A/^{2}A) to a maximum of 2.602/2.629 Å in the QM/MM computations. The Fe−S bond within the protein pocket is significantly shorter than it is in the naked system, by as much as 0.1 Å in the case of the sulfur ligand R3 in combination with the basis B1. These results accord well with the previous conclusions of Ogliaro et al.[21] and highlight the stabilizing impact of the hydrogen bonding and electric field of the protein on the Fe−S bond, as deduced from experiment.[25−28,46] Generally, the Fe−S bond is 0.02−0.03 Å longer in the ^{2}A than it is in the ^{4}A state.[19−23]

The angle θ_{Fe-S-X} (X being hydrogen or carbon, depending on the choice of the QM part) is almost the same in all QM/MM models, because the geometry is largely determined by the enzyme environment. In the naked system, there is considerable conformational freedom, and the angle strongly depends on the nature of the sulfur ligand, ranging from around 100° in the case of R1 to 115° in the case of R3. It is likely that the valence angle at the sulfur atom as well as the Fe−S distance influence the ability of the thiolate ligand to bind to Fe and, hence, the electronic structure of Cpd I as a whole.

Relative Energies. To appreciate the influence of the protein environment on the QM system in terms of energetic stabilization, we compare in Table 2 the energies of Cpd I (the QM system) in three situations defined by the subscripts p, protein, and g, gas phase: (i) $S_{p,p}$ − the system in the geometry and electric field corresponding to the protein environment; (ii) $S_{p,g}$ − the system at its protein geometry but in the gas phase, without the electric field of the protein; and (iii) $S_{g,g}$ − the naked system optimized in the gas phase.

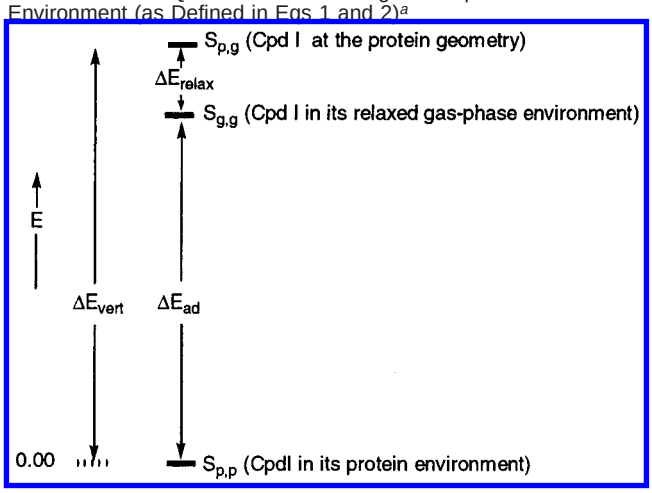
(46) Hydrogen bonding to sulfur shortens the Fe−S bond in model complexes by ca. 0.028 Å. See: Suzuki, N.; Higuchi, T.; Urano, Y.; Kikuchi, K.; Uekusa, H.; Ohashi, Y.; Uchida, T.; Kitagawa, T.; Nagano, T. *J. Am. Chem. Soc.* **1999**, *121*, 11571.

Table 2. Influence of the Protein Environment on QM Energies of the Cpd I Species (Common Reference Point Is the QM Energy of Cpd I in the Protein Pocket ($S_{p,p}$ in Eq 1))[a,b,c]

		$\Delta E_{vert}/\Delta E_{ad}$		
		R1 (SH)	R2 (SMe)	R3 (Cys)
^{4}A	B1	+106.8/+97.9	+99.5/+90.0	+150.1/+128.4
	B3	+108.7/+99.5	+100.7/+90.9	+150.6/+129.7
^{2}A	B1	+106.7/+97.7	+99.5/+90.0	+150.3/+128.1
	B3	+109.0/+99.7	+100.7/+90.9	+150.6/+129.6

[a] R1−R3 refer to the QM model Cpd I systems in Figure 4. [b] ΔE_{vert} (eq 1, Scheme 1) is the vertical stabilization energy of the naked system in its protein geometry relative to the reference. [c] ΔE_{ad} (eq 2, Scheme 1) is the adiabatic stabilization energy of the naked system in the gas-phase geometry relative to the reference.

Scheme 1. The QM Stabilization Energies of Cpd I in the Protein Environment (as Defined in Eqs 1 and 2)[a]



[a] The first subscript of each system (S) specifies the geometry, and the second one specifies the environment. ΔE_{relax} corresponds to the geometric relaxation of the gas-phase species relative to the protein.

Taking as a common reference point the total QM energy of $S_{p,p}$, we reported the following energy differences in the table:

$$\Delta E_{vert} = E(S_{p,g}) - E(S_{p,p}) \quad (1)$$

$$\Delta E_{ad} = E(S_{g,g}) - E(S_{p,p}) \quad (2)$$

Thus, ΔE_{vert} corresponds to the "vertical" stabilization energy of Cpd I provided by the protein electric field, at the geometry that the species adopts in the pocket. On the other hand, ΔE_{ad} represents the "adiabatic" stabilization energy, which includes the energy change ΔE_{relax} arising from geometry relaxation in the gas phase. Scheme 1 shows the relations between these quantities.

As expected, the polarization by the electric field causes a substantial stabilization of Cpd I (approximately 109, 100, and 150 kcal/mol for Cpd I systems with substituents R1, R2, and R3, respectively), which is virtually identical for both spin states. As compared to this large electrostatic effect, the energy change because of the relaxation in the naked molecule is smaller but nevertheless substantial (9−10 kcal/mol for R1 and R2, 20 kcal/mol for R3), indicating that Cpd I is "strained" in the enzyme. It is clearly an advantage of combined QM/MM schemes that they are able to account for these effects of electrostatic stabilization and steric strain simultaneously. The larger relaxation of Cpd I model R3 is apparently because of the ability of this species to adopt a gas-phase conformation with internal hydrogen bonding between the backbone NH group of Cys$_{357}$

Table 3. Relative $^4A-^2A$ State Energies [kcal/mol] for Different Models of Cpd I (Defined by R1−R3) at the QM/MM and QM Levels[a]

		QM region		
entry		R1 (SH)	R2 (SMe)	R3 (Cys)
1	$\Delta E^{QM/MM}(p,p)$[b]	−0.18/−0.14/+0.06	−0.10/−0.01/+0.04	−0.04/+0.04/+0.03
2	$\Delta E^{QM}(p,p)$[c]	−0.25/+0.07/+0.21	+0.06/+0.02/+0.08	+0.04/+0.11/+0.04
3	$\Delta E^{QM}(p,g)$[d]	−0.15/−/−0.10	+0.06/−/+0.09	−0.13/−/+0.06
4	$\Delta E^{QM}(g,g)$[e]	−0.07/−/+0.03	+0.05/−/+0.14	+0.37/−/+0.08

[a] Energy differences as defined by eq 3; a negative (positive) value implies an ^4A (^2A) ground state. The triplet of entries corresponds to basis B1/B2/B3. [b] QM/MM energy difference. [c] QM contribution to the QM/MM energy difference. [d] Energy difference of the naked system at the protein geometry. [e] Energy difference of the naked system at the relaxed gas-phase geometry.

Table 4. Computed Net Spin Densities of the ^4A and ^2A States for Models R1−R3 and Basis Sets B1−B3[a]

		^4A state			^2A state		
		R1 (SH)	R2 (SMe)	R3 (Cys)	R1 (SH)	R2 (SMe)	R3 (Cys)
FeO	B1	2.022/2.023/2.025	2.019/2.024/2.023	2.018/2.026/2.007	2.121/2.115/2.094	2.119/2.104/2.091	2.122/2.108/2.101
	B2	2.021/−/−	2.018/−/−	2.018/−/−	2.119/−/−	2.116/−/−	2.120/−/−
	B3	2.041/2.047/2.037	2.042/2.049/2.040	2.046/2.063/2.026	2.177/2.154/2.146	2.172/2.145/2.141	2.173/2.139/2.149
S_{cys357}	B1	0.252/0.467/0.566	0.323/0.664/0.742	0.331/0.574/0.586	−0.296/−0.516/−0.618	−0.371/−0.730/−0.782	−0.381/−0.641/−0.627
	B2	0.236/−/−	0.312/−/−	0.313/−/−	−0.285/−/−	−0.353/−/−	−0.360/−/−
	B3	0.201/0.408/0.526	0.282/0.645/0.719	0.283/0.539/0.535	−0.270/−0.477/−0.600	−0.349/−0.728/−0.777	−0.352/−0.620/−0.593
porph	B1	0.734/0.528/0.431	0.658/0.313/0.239	0.652/0.399/0.409	−0.835/−0.618/−0.500	−0.750/−0.377/−0.312	−0.744/−0.466/−0.475
	B2	0.749/−/−	0.666/−/−	0.665/−/−	−0.841/−/−	−0.759/−/−	−0.756/−/−
	B3	0.763/0.557/0.452	0.671/0.296/0.229	0.667/0.385/0.429	−0.915/−0.691/−0.563	−0.820/−0.407/−0.350	−0.818/−0.506/−0.545

[a] The triplet of numbers in each field refers to the QM/MM-optimized structure/the naked system in its protein geometry/the naked system in its gas-phase geometry.

and the carbonyl oxygen of the same residue. This interaction is not feasible in the enzyme matrix. This concurs with the notion that the protein pocket is not a malleable cavity, but rather more a rigid matrix that provides a hydrogen-bonding template.

The energy differences between the quartet and doublet spin states

$$\Delta E = E(^4\mathrm{A}) - E(^2\mathrm{A}) \qquad (3)$$

are listed in Table 3 for different QM regions (R1−R3) and basis sets (B1−B3). The first entry gives the difference $\Delta E^{QM/MM}$ for the full systems, while the other three entries correspond to differences in the QM energies of the QM systems considered above ($S_{p,p}$; $S_{p,g}$; $S_{g,g}$). Generally speaking, all values in Table 3 are rather small (typically less than 0.2 kcal/mol in absolute value), emphasizing again that the two spin states are close in energy regardless of the chosen computational approach. Enlargement of the basis set from B1 to B3 tends to favor the doublet state (entry 1), but there is apparently no clear trend when extending the QM region from R1 to R3 (entry 1) or when passing from the full system to the model systems considered (entries 2−4).

Our best calculations with the large basis B3 invariably predict that the ground state is ^2A in the enzyme (entry 1), in accord with the experimental assignment of the related Cpd I of the CPO enzyme.[12a,13−18] Qualitatively similar results are obtained for the relaxed gas-phase species (entry 4), indicating that the apoprotein has no major influence on the relative energies of the two spin states. Using the data for model R3 and basis B3 in entry 1, the QM/MM energy difference between the two states is ca. 11 cm^{-1} which may be compared with the experimental value of 37 cm^{-1} for Cpd I in CPO.[15,18] The computed adiabatic energy differences in Table 3 come from separate geometry optimizations of the two states, and their precision is therefore limited by the convergence criteria of these optimizations. To assess the intrinsic ordering of the two states, we have also computed the vertical energy differences at the optimized doublet geometry and have obtained values of −0.01/+0.05/+0.08 kcal/mol for basis B1/B2/B3 and model R3 (i.e., −4/+18/+27 cm^{-1}).

Spin Densities. Table 4 summarizes the calculated net spin densities of the two states of Cpd I. For each Cpd I model (R1−R3) and basis set (B1−B3), three data are given which refer to the full system (QM/MM), the isolated Cpd I system at the geometry in the protein, and the relaxed gas-phase species (QM). The FeO unit has a spin density of approximately 2.0 and thus corresponds to two unpaired electrons in a triplet situation, as illustrated in Figure 2. This spin density is affected neither by the Cpd I model nor by the polarizing effect of the enzyme environment or the basis set. In contrast, the spin densities on the sulfur atom and the porphyrin moiety show a remarkable dependence on all of the parameters. For the isolated gas-phase compounds, the spin density on sulfur changes in the order R2 > R3 > R1, while the change on the porphyrin is precisely the opposite, R1 > R3 > R2. As discussed previously,[21a] the substituent effect on the sulfur density reflects the electron-donating capability of the group R attached to sulfur, R1 (R2) being the worst (best) electron donor. Consequently, R1 (R2) underestimates (overestimates) the sulfur radical character with respect to the full cysteine model compound (R3). The spin density on the porphyrin macrocycle depends on the substituent in a complementary manner. The greatest changes in these spin densities occur as Cpd I is transferred from the gas phase to the protein environment. When this happens, Cpd I is transformed from a dominant sulfur radical species (in all of the models R1−R3) to a dominant porphyrin radical type. While the R1 model still slightly underestimates the sulfur radical character in the enzyme environment, R2 and R3 give quite similar results, especially for the larger basis sets. The R2 spin density on sulfur is usually even slightly lower than for R3, which gives rise to an order R1 < R2 < R3. Thus, the polarizing effect of the enzyme environment cancels the artificially strong electron pushing effect of the methyl substituent in R2 that is

Table 5. Spin Density Distribution within the Porphyrin Ring for Models R1−R3 and Basis Sets B1−B3[a]

		$^4A_{2u}$			$^2A_{2u}$		
		R1 (SH)	R2 (SMe)	R3 (Cys)	R1 (SH)	R2 (SMe)	R3 (Cys)
C_α	B1	−0.087/−0.069/−0.054	−0.079/−0.043/−0.031	−0.078/−0.054/−0.049	0.076/0.055/0.041	0.068/0.029/0.020	0.068/0.039/0.039
	B2	−0.087/−/−	−0.077/−/−	−0.078/−/−	0.075/−/−	0.067/−/−	0.067/−/−
	B3	−0.089/−0.072/−0.056	−0.079/−0.042/−0.029	−0.080/−0.054/−0.052	0.078/0.061/0.044	0.070/0.029/0.020	0.071/0.041/0.044
C_β	B1	−0.008/−0.007/−0.005	−0.007/−0.004/−0.002	−0.007/−0.005/−0.005	0.005/0.004/0.003	0.005/0.001/0.000	0.005/0.003/0.003
	B2	−0.008/−/−	−0.007/−/−	−0.007/−/−	0.005/−/−	0.005/−/−	0.005/−/−
	B3	−0.008/−0.007/−0.005	−0.007/−0.004/−0.002	−0.007/−0.005/−0.005	0.005/0.005/0.003	0.005/0.001/0.000	0.005/0.003/0.003
C_m	B1	0.229/0.167/0.123	0.205/0.099/0.065	0.202/0.127/0.117	−0.226/−0.161/−0.116	−0.202/−0.090/−0.063	−0.200/−0.117/−0.114
	B2	0.236/−/−	0.209/−/−	0.209/−/−	−0.230/−/−	−0.207/−/−	−0.206/−/−
	B3	0.243/0.183/0.133	0.214/0.101/0.066	0.215/0.133/0.131	−0.238/−0.178/−0.127	−0.212/−0.092/−0.065	−0.212/−0.125/−0.129
N	B1	0.153/0.122/0.107	0.139/0.077/0.063	0.139/0.096/0.098	−0.154/−0.119/−0.101	−0.139/−0.068/−0.058	−0.139/−0.088/−0.092
	B2	0.149/−/−	0.134/−/−	0.135/−/−	−0.149/−/−	−0.135/−/−	−0.136/−/−
	B3	0.150/0.122/0.107	0.134/0.067/0.057	0.135/0.086/0.096	−0.166/−0.133/−0.113	−0.151/−0.074/−0.066	−0.152/−0.095/−0.106

[a] The triplet of numbers in each field refers to the QM/MM-optimized structure/the naked system in its protein geometry/the naked system in its gas-phase geometry. See Figures 1 and 2 for the definitions of the relevant atoms.

Table 6. Hydrogen Bond Distances [Å] for N−X [H−X], Where X = S, Except for the Last Column with X = O (First Row Gives the X-ray Data (Ref 7) with H-Atoms Added and Optimized at the Force Field Level; Other Rows Present Distances from QM/MM Optimizations (Model R3, Basis B3, ^{2}A State))

	backbone N–H- - -S(Cys_{357})			side-chain Gln_{360} amide
	Leu_{358}	Gly_{359}	Gln_{360}	N–H- - -O═C(Cys_{357})
X-ray(1DZ9) + H	3.51 [3.58]	3.23 [2.51]	3.31 [3.14]	3.00 [2.02]
snapshot 29	3.46 [3.35]	3.32 [2.42]	3.78 [3.81]	4.97 [5.44]
snapshot 40	3.45 [3.34]	3.40 [2.68]	3.57 [3.47]	3.15 [2.86]
snapshot 40[a]	3.49 [3.48]	3.33 [2.57]	3.56 [3.43]	2.77 [1.77]
snapshot 50	3.48 [3.43]	3.35 [2.48]	3.78 [3.65]	4.22 [4.50]

[a] Side chain of Gln_{360} manipulated to form an H-bond with the carbonyl oxygen of Cys_{357}.

observed in the gas-phase calculations. These results demonstrate the ability of QM/MM methods to converge a calculated property by increasing the size of the QM region stepwise. More important, however, is that the pitfall of overstabilization of the sulfur radical state because of the lack of a polarizing medium (as in isolated molecule calculations) is avoided.

These results are consistent with small model calculations using added NH_3 molecules to mimic NH- - -S bonding in the protein pocket and a continuum solvation model to mimic the polarization by the electric field.[21b,c]

Table 5 displays the spin density distribution on the various positions of the porphyrin (see Figure 2). It is obvious that as Cpd I is moved from the gas phase to the protein environment, the spin densities increase by a factor of 2 or even more on the meso carbon (C_m) and significantly also on the pyrrole nitrogen ($N_{pyrrole}$). This is a clear indication that the porphyrin spin density is located in an orbital that has a_{2u} character (as drawn above in Figure 2).

B. Comparison of Different Snapshots. The influence of different conformations of the enzyme environment on the calculated electronic properties of Cpd I can be assessed by comparing the results of QM/MM geometry optimizations that start from different snapshots of the initial MD trajectory (see Figure 3). We have selected three such snapshots (after 29, 40, and 50 ps) and generated an additional structure from the second snapshot by manipulating the side chain of Gln_{360} such that it forms an H-bond between its NH(amide) group and the carbonyl oxygen of the backbone of Cys_{357}, to quantify the effect of this particular H-bond.[47] Table 6 presents the most relevant H-bond distances in the QM/MM optimized structures (QM region R3, basis set B3, ^{2}A state) and compares them with the available X-ray data[7] for the heavy atom distances N−X and with estimated distances H−X (H-positions optimized at the CHARMM force field level while keeping the heavy atoms fixed at the X-ray positions). In the case of Leu_{358} and Gly_{359}, the four calculated N−S distances between the backbone peptide groups and the cysteinate sulfur are quite similar (variations within 0.04 and 0.08 Å, respectively) and close to the X-ray values. Hence, the backbone conformation of these two residues seems to be rather rigid. In contrast, the corresponding backbone peptide group of Gln_{360} exhibits more conformational flexibility because the computed N−S distances to the cysteinate sulfur change by up to 0.22 Å between different snapshots; the H-bond geometry is rather unfavorable in snapshots 29 and 50, and somewhat better in snapshot 40, although the N−S distance still exceeds the X-ray value by 0.26 Å. The side chain of Gln_{360} clearly does not form an H-bond with the carbonyl oxygen of Cys_{357} in snapshots 29 and 50, and its orientation is also unfavorable for this purpose in snapshot 40. This H-bond is present in the manipulated structure.

Figure 5 shows other selected geometrical parameters and spin densities for the QM/MM optimized structures derived from the different snapshots. The Fe−O and Fe−N distances show only negligible variations, in accordance with their very small sensitivity to the choice of QM region and basis set (for snapshot 29, see above). The Fe−S bond again proves to be more flexible; it ranges from 2.585/2.609 (^{4}A/^{2}A) to 2.547/2.569 Å. The shortest bond is found for snapshot 40 with the manipulated side chain; the presence of this H-bond causes a slight shortening of 0.008 Å with respect to snapshot 40. In line with these findings, the calculated spin densities on the FeO moiety are little affected by the different conformations, whereas those on porphyrin and sulfur show some variation. Thus, the ratio of

(47) Note that the H-bond interactions of Cys_{357} with Gly_{359} and Gln_{360} are calculated between QM and MM atoms, and thus the physical picture is that of a constant (unpolarizable) charge that interacts with the polarizable QM electron density. Thus, although we can quantify the effect of polarization of the QM subsystem, the importance of individual H-bond interactions of this kind should not be overestimated.

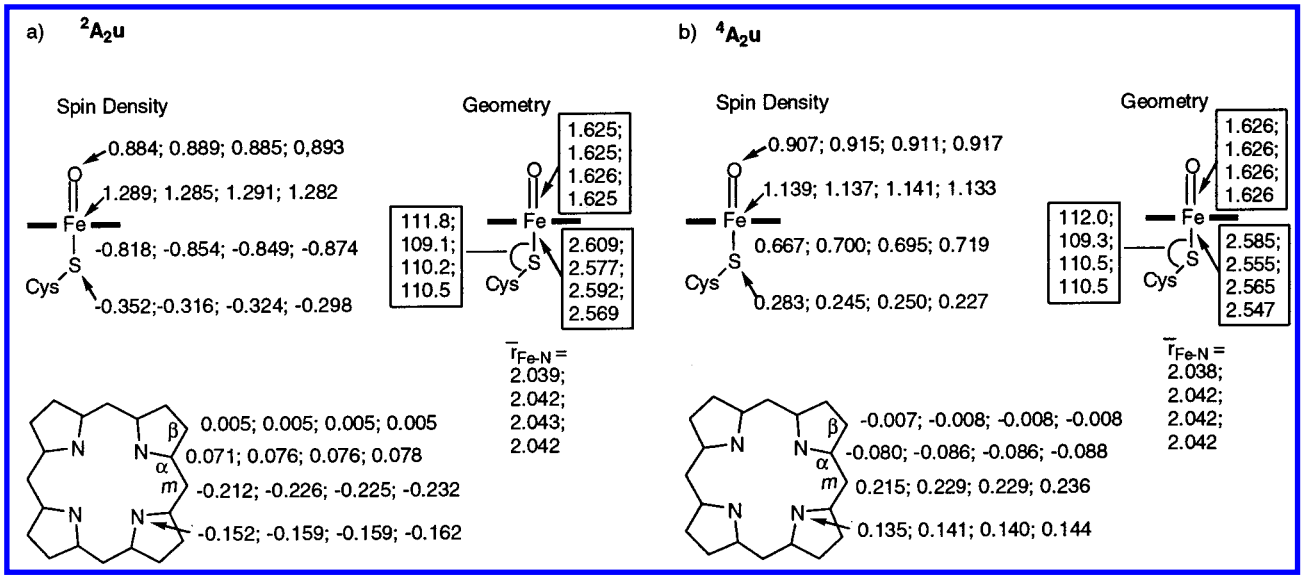


Figure 5. Cpd I of $P450_{cam}$ (model R3, basis B3): QM/MM optimized geometries and spin densities for different snapshots. The given values correspond to snapshots 29, 40, 50 and snapshot 40 with the manipulated Gln_{360} side-chain conformation.

porphyrin:sulfur unpaired spin density in the quartet changes from 0.67:0.28 in snapshot 29 to 0.72:0.23 in snapshot 40 with the manipulated side chain. This shift of spin density between sulfur and the macrocycle reflects the differential stabilization of negative charge on sulfur through the different H-bond networks. However, as compared to the large effect on the spin density that is observed when Cpd I is transferred to the gas phase, these changes appear rather small, and one should stress that the qualitative result − that of a dominant porphyrin radical character − remains the same for all of the snapshots.

Discussion

The present study reveals some key aspects related to the electronic and geometric features of Cpd I and its synergistic response to the protein pocket. These aspects are discussed below.

The geometric data in Table 1 and energetic data (ΔE_{relax}, Scheme 1) in Table 2 show that Cpd I inside the pocket is strained with respect to its relaxed gas-phase geometry. This is especially apparent for the model R3 that uses an extended part of the cysteine loop as a proximal ligand. Thus, while in the gas phase, the small chain curls to form an internal NH- - -O═C hydrogen bond; within the protein pocket, this interaction does not exist and is replaced by another one from Gln_{360}. These results, as well as other indications from the pure MM calculations, show that the protein pocket is not a malleable cavity, but a rather rigid matrix that provides a hydrogen-bonding template that sustains the Fe−S bond. This is further emphasized by the Fe−S bond shortening (Table 1) and the increased electron density at sulfur (Table 4) in the pocket vis-à-vis the gas phase. These results accord well with experimental observations[27,33] that the hydrogen-bonding machinery in the proximal pocket is essential for the stability of the heme-thiolate species.

The spin density data in Table 4 show that the protein pocket is responsible for transforming Cpd I from a sulfur-centered radical, in the gas phase, to a porphyrin-centered radical in the protein pocket. Mimetic Cpd I species with a porphyrin-centered radical cation are "green", while when the porphyrin macrocycle is closed-shell, the corresponding compound is "red".[6,8,9] The characterized Cpd I of CPO is "green",[16] and the UV−vis spectrum of the recently probed Cpd I of CYP119 matches that of CPO.[11] This indicates that the Cpd I species of cysteinate enzymes are also porphyrin-centered radical cation species and that the QM/MM results thus match experiment. Furthermore, the calculated porphyrin spin densities, which are large on the nitrogen and C_m positions (Table 5), enable us to specify the state of Cpd I as an A_{2u} type, in accord with EPR results.[17] Nevertheless, the small spin densities on other positions suggest a small admixture of A_{1u} character.[19c,21c] Finally, the calculations predict that the ground state of the enzyme is the doublet A_{2u} state (Table 3), which is consistent with the experimental assignment of the related Cpd I of CPO.[12a,18]

These general trends are reproduced by small model systems, which contain two external hydrogen bonds with NH_3 molecules and are placed in a solvent with a dielectric constant of 5.7.[21b,c] Interestingly, the spin density distribution suggests that the simplest model which consists of iron-oxo-porphine with HS^- as a proximal ligand is not too far from the full model (Figure 5), and might thus be used for more demanding calculations on reaction mechanisms.

A most interesting feature in Figure 5 is the chameleon behavior of Cpd I.[21c] Small changes in the hydrogen-bonding situation of the thiolate ligand result in changes in the spin density on the sulfur vis-à-vis the porphyrin ring. As shown before,[21b] these changes can be conceptualized with a simple valence bond (VB) mixing model (Figure 6a), which is similar to the resonance-theory picture used by experimentalists in the area.[3,5,12a,48]

In this description, the species is a resonance hybrid of two forms: $|\mathbf{a}\rangle$ involves a thiolyl radical and a closed-shell porphyrin, while $|\mathbf{b}\rangle$ involves a thiolate anion and a porphyrin radical-cation. At infinite Fe−S separation, $|\mathbf{a}\rangle$ is considerably lower in energy. However, at the equilibrium geometry, the ion pair structure $|\mathbf{b}\rangle$ is stabilized and approaches $|\mathbf{a}\rangle$ energetically. In the gas phase, $|\mathbf{a}\rangle$ is still the lowest structure, and the VB

(48) Woggon, W.-D. *Top. Curr. Chem.* **1996**, *184*, 39.

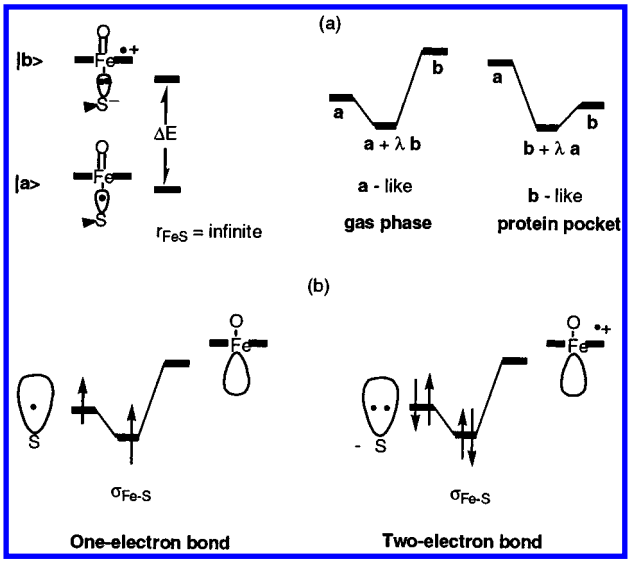


Figure 6. The nature of the Fe−S bond in Cpd I. (a) A VB mixing diagram. (b) A complementary orbital mixing diagram of the Fe−S bond orbital.

mixing leads to a Cpd I species which has a dominant thiolyl radical. As the species is placed in the protein pocket, the electric field and the NH- - -S hydrogen bonding stabilize preferentially the ion-pair form, |**b**>, which becomes the lowest VB structure. The VB mixing then leads to a Cpd I species that has a dominant porphyrin radical-cation character as indeed revealed by the calculations.

The interactions with the protein also strengthen the Fe−S bond. This complementary information becomes more obvious from the orbital mixing diagram in Figure 6b, which shows the alteration in the nature of the Fe−S bond as the thiol center changes from a radical to an anion. In each case, the iron-oxo-porphyrin fragment has an empty hybrid orbital (a mixture of d_{z^2}, p_z, and s orbitals on iron) pointing to the missing coordination site. This hybrid mixes with the p_σ hybrid on sulfur to form the Fe−S bond orbital. When the sulfur is in a radical situation, the bond orbital will contain a single electron, and the Fe−S bond will be a one-electron bond. However, when the sulfur is in an anionic situation, the bond orbital is doubly occupied, and the Fe−S bond is a two-electron bond. Thus, the interactions in the protein pocket cause a transformation of the Fe−S linkage from a weak one-electron bond to a stronger two-electron bond. This qualitative change of the Fe−S bond is in line with the calculated shortening and strengthening of this bond.[21b,c]

The VB model in Figure 6a further predicts that the stronger the electric field and the hydrogen bonding, the more dominant will be the porphyrin radical-cation character. Figure 5 reveals this trend beautifully: when going from snapshot 29, in which Gln_{360} does not form an NH- - -S hydrogen bond, via snapshot 40, in which this hydrogen bond is present to some extent, all the way to the manipulated structure in which Gln_{360} also forms an NH- - -O═C hydrogen bond to the cysteine, the corresponding porphyrin spin density in the doublet state changes from −0.818 to −0.874 with an opposite and concomitant change in the cysteinate spin density. This again shows the extraordinary ability of Cpd I to accommodate its electronic structure to the protein environment. Although there is not yet a direct experimental verification of these subtle changes, the results from site-directed mutageneses in the proximal pocket of $P450_{cam}$[23,33] indicate that replacement of Leu_{358} and Gln_{360} by proline, which does not form hydrogen bonds, exerts a marked influence on all properties of the enzyme. One may thus expect that cysteinate-Cpd I species in different proteins (with different H-bond patterns) will have different properties. It is yet to be unraveled whether these differences are also expressed in the reactivity of different P450 isozymes and different thiolate enzymes.

Conclusion

The present study characterizes the geometric and electronic structure of the elusive active oxidant Cpd I of cytochrome P450. It highlights the advantage of performing QM/MM calculations for such a task. The calculations provide information on the species in its specific protein environment, in this case $P450_{cam}$, without the need to make assumptions about the nature and architecture of this environment. In this respect, QM/MM calculations come as close as possible to being a faithful partner to structural and spectroscopic determinations by experiment. An added advantage is that these calculations offer important insight into the factors that govern the properties of Cpd I, through comparisons with various gas-phase models and through inspection of different conformations within the protein pocket. We find that Cpd I is transformed by the protein environment from a sulfur-centered radical to the "green" $^{4,2}A_{2u}$ porphyrin-centered radical cation (ca 70%). The hydrogen-bonding machinery of the protein pocket shortens the Fe−S bond and strengthens it. The best estimates of the Fe−S bond lengths in the $^2A_{2u}/^4A_{2u}$ states are 2.569 Å/2.547 Å. A series of calculations on different conformations of the protein pocket show that the balance of the sulfur/porphyrin radical character and the precise Fe−S bond length are finely tuned by small changes in the architecture of the pocket. *Cpd I behaves, therefore, as a chameleon species that adapts its electronic and structural character to the specific environment.*

Acknowledgment. This research is sponsored by the Binational German Israeli Foundation (GIF). J.C.S. thanks the Fonds der Chemischen Industrie for a Kekulé scholarship. F.O. thanks the EU for a Marie Curie Fellowship.

Supporting Information Available: Details of the preparatory force field calculations, force field parameters, and technical information on the QM/MM calculations (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA026279W