P450 Enzymes: Their Structure, Reactivity, and Selectivity—Modeled by QM/MM Calculations

Sason Shaik,*,† Shimrit Cohen,† Yong Wang,† Hui Chen,† Devesh Kumar,‡,§ and Walter Thiel*,‡

Institute of Chemistry and the Lise-Meitner-Minerva Center for Computational Quantum Chemistry, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel, and Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, D-45470 Mülheim an der Ruhr, Germany

Contents

* E-mail: S.S., sason@yfaat.ch.huji.ac.il; W.T., thiel@mpi-muelheim.mpg.de. Further contact information for S.S.: phone, +972-2-6585909; fax, +972-2-6584680/6585345.
† The Hebrew University of Jerusalem.
‡ Max-Planck-Institut für Kohlenforschung.
§ Current address: Molecular Modeling Group, Indian Institute of Chemical Technology, Hyderabad 500 607, India.

1. Introduction

The introduction of oxygen into biochemical processes has brought about an evolutionary leap in the history of life, whereby many organisms have evolved to use oxygen as part of the life sustaining process and at the same time invented means to lower its toxic effects.[1] One of these means

Sason Shaik was born in Iraq. His family immigrated to Israel in the Exodus of the Iraqi Jewry. He received his B.Sc. and M.Sc. degrees in chemistry from Bar-Ilan University and his Ph.D. from the University of Washington under Nicholaos D. Epiotis. In 1978/9 he spent a Post Doctoral year with Roald Hoffmann at Cornell University. In 1980 he started his first academic position as a Lecturer at Ben-Gurion University, where he became a Professor in 1988. He subsequently moved to the Hebrew University, where he is currently the director of The Lise Meitner-Minerva Center for Computational Quantum Chemistry. Among the awards he received are the Lise-Meitner-Alexander von Humboldt Senior Award in 1996−1999, the 2001 Israel Chemical Society Prize, the 2001 Kolthoff Award, and the 2006 Schrödinger Medal. He is an AAAS fellow and a WATOC fellow. His research interests are in the use of quantum chemistry, and in particular of valence bond theory, to develop paradigms which can pattern data and lead to the generation and solution of new problems. He started his P450 research in 1998 and has been fascinated ever since!

Shimrit Cohen was born in Ashdod, Israel, in 1978. She received her B.Sc. in chemistry from the Hebrew University in 2000. She joined the group of Prof. Sason Shaik while she was an undergraduate student, and she is currently completing her Ph.D. in quantum chemistry. In 2005 she received the Wolf scholarship for excellent Ph.D. students. Her research involves applications of QM/MM methodology for understanding the mechanism and catalysis of cytochrome P450.

is the enzyme cytochrome P450 (CYP), which is a heme protein that was scientifically discovered by Omura and Sato a few decades ago.[2] Within a short time, it has become clear that this is a superfamily of isoforms present in a diverse array of organisms and performing biosyntheses of hormones and detoxifications of *xenobiotics* by utilizing oxygen and two reducing equivalents, and thereby oxidizing a variety of organic compounds.[3−21]

Scheme 1 shows the great diversity in the chemical reactivity of P450s. The most common reaction involves a single oxygen insertion into an organic molecule, such as C—H hydroxylation, sulfoxidation, C≡C epoxidation, etc.; hence, the name "monooxygenase" is given to this enzyme. However, as can be seen in the scheme, the enzyme performs

Yong Wang was born in Anhui, China, in 1975. He received his B.Sc. degree in theoretical physics at Anhui University in 1999, his M.Sc. degree in atomic and molecular physics at Henan Normal University in 2002, and his Ph.D degree (supervised by Prof. Keli Han) in physical chemistry at Dalian Institute of Chemical Physics, CAS, in 2007. During his Ph.D. research, he has been fortunate to cooperate with Prof. Shaik since 2005. He then joined Shaik's group as a postdoctoral fellow. He was awarded by the Hebrew University the 2008−2009 Lady-Davis Postdoctoral Fellowship. His current research interests are focused on the mechanisms of biotransformations catalyzed by P450, by means of QM, QM/MM, and MD methods.
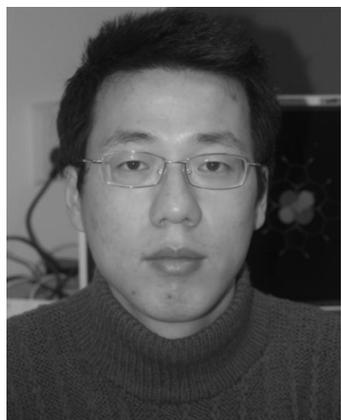
Hui Chen was born in China in 1978. He received his B.Sc. (2001) degree in chemistry from Nanjing University (China) and his Ph.D. (2006) degree in physical chemistry from Nanjing University, under the supervision of Prof. Shuhua Li at the Institute of Theoretical and Computational Chemistry. Since then he has been working with Prof. Sason Shaik at the Hebrew University as a postdoctoral fellow until now. He is the recipient of a Golda Meir Fellowship (2007−2009) of the Hebrew University and won the Lise Meitner Prize of the Lise-Meitner-Minerva Center (2007) for computational quantum chemistry. His main research interests are theoretical studies of reaction mechanisms and chemical bonding in chemistry and biochemistry.

a variety of other transformations, such as desaturation, oxidative dehalogenation, reductive dehalogenation, deformylation, N-dealkylation, O-dealkylation, C—C coupling, and so on. The immense biological impact of this enzyme may be appreciated from the fact that more than 70% of the drugs taken orally by humans are metabolized by the P450 isozymes CYP3A4 and CYP2D6.[21]

Due to this chemical versatility and the biological importance of P450, the area has become a target of intense research both practical and fundamental. An exciting branch of the basic research is related to the structure and mechanism of the enzyme, and attempts *to understand how it performs its reactions so fast and so efficiently*, albeit not in all the isoforms. Experimental mechanistic studies have uncovered important modes of reaction, such as the rebound mechanism

Devesh Kumar was born in India in 1965. He received his B.Sc. from the L. N. Mithla University and his M.Sc. and Ph.D. in Physics from the D. D. U. Gorakhpur University. He worked at the Centre for Liquid Crystal Research and Education, Nagarjuna University, as a research associate (July 2001−March 2002) before joining Prof. Sason Shaik at the Hebrew University of Jerusalem as a postdoc in May 2002. Subsequently, he joined Professor Walter Thiel at the Max-Planck-Institut für Kohlenforschung in Mülheim an der Ruhr, Germany, as a postdoc (June 2006−March 2009). At present he is Ramanujan Fellow at the Indian Institute of Chemical Technology, Hyderabad. His main research interests are QM and QM/MM studies of conformations and interactions of mesogens and enzyme catalysis.
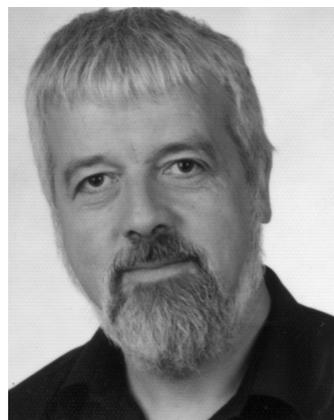
Walter Thiel was born in Treysa, Germany, in 1949. He studied chemistry at the University of Marburg and received his Ph.D. degree in 1973. After two postdoctoral years at the University of Texas at Austin with Michael Dewar, he returned to Marburg and obtained the Habilitation in 1981. He was Professor at the University of Wuppertal from 1983 to 1992, Visiting Professor at the University of California at Berkeley in 1987, and Full Professor of Chemistry at the University of Zürich from 1992 to 1999. He is now Director at the Max-Planck-Institut für Kohlenforschung in Mülheim an der Ruhr, and affiliated with the University of Düsseldorf as an Honorary Professor. He received the Heisenberg Fellowship in 1982, the Alfried-Krupp-Award in 1988, and the WATOC Schrödinger Medal in 2002. His research area is theoretical and computational chemistry, both method development and applications.

in C—H hydroxylation,[3,8,22] but at the same time have raised many questions and interesting puzzles.[18,23,24]

It is very difficult to prove a mechanism and to characterize unstable intermediates with very short lifetimes, such as the active species, Cpd I, of the enzyme shown at the top of Scheme 1, and this is where theory comes into play and offers a complementary tool to aid experiment. Indeed, the allure of P450 has attracted many theoretical studies[25−32] designed to understand these fundamental aspects of P450. Most of the initial studies were done on model iron-porphine systems using primarily density functional theory (DFT) as the quantum mechanical (QM) tool.[25,26,33−52] These DFT calcula-

Scheme 1. Chemical Transformations Achieved by P450s, Shown along with the Putative Active Species, Compound I (Cpd I)
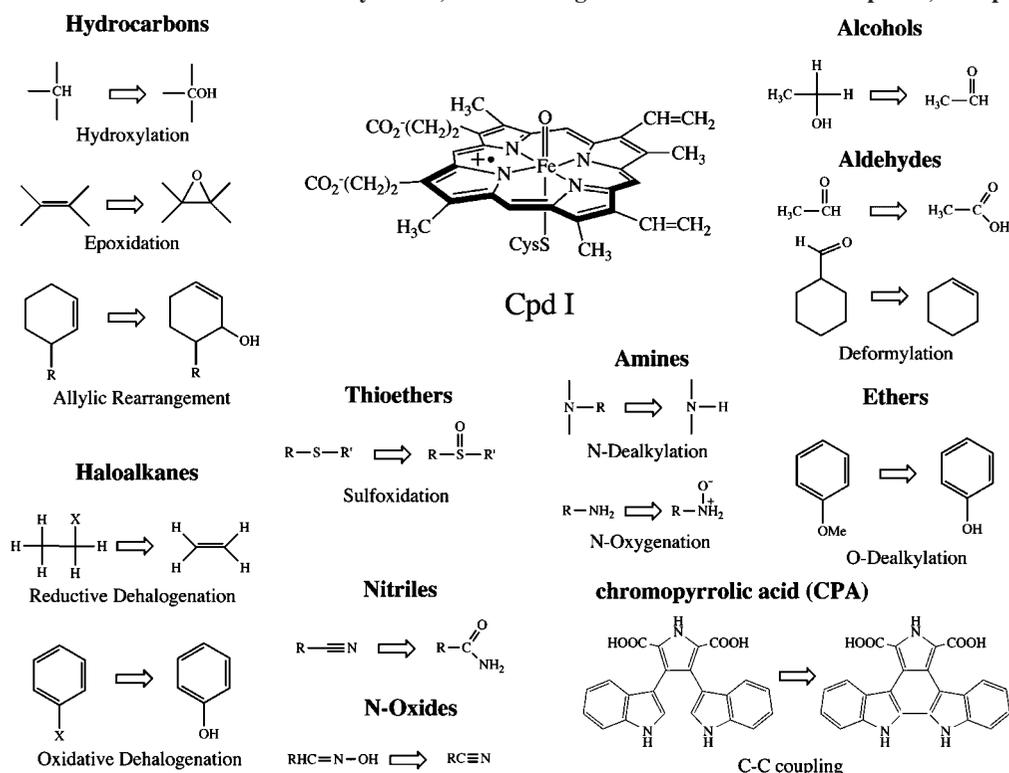
tions have contributed a great deal to the understanding of the electronic structure of the active species of the enzyme, the iron-oxo porphyrin species shown in Scheme 1, and to its reactivity patterns, including its high sensitivity to the environment, such as the bulk polarity or the hydrogen bonding to the axial cysteine ligand.

It was however recognized that part of the versatility of P450s must be attributed to the fact that while all the isoforms share the same active species, they differ by the identity of their proteins. Therefore, a more suitable level to study these isoforms is the hybrid QM/MM (quantum mechanical/molecular mechanical) method (see section 2), which treats the active species within its native protein environment. The first QM(DFT)/MM study[53] of the active species in 2002 revealed the utility of this method. It was followed by an increasing number of QM/MM studies of all the species in the catalytic cycle in CYP101, known also as P450$_{cam}$,[54] the active species of human isoforms,[55,56] and some of the reaction mechanisms initiated by this active species with a variety of organic molecules.[57−61] In-silico mutants were designed,[62,63] and in one case, the design of a selenocysteine mutant[62] even preceded its experimental preparation.[64,65] Species and reactivity patterns of related heme enzymes, such as chloroperoxidase (CPO),[66,67] nitric oxide synthase (NOS),[68,69] and horseradish peroxidase (HRP),[70−74] were also investigated. Some of these studies were done in collaboration with experimentalists[58,66,67] or were used by experimentalists as benchmarks,[75] thus demonstrating the usefulness of QM-(DFT)/MM as a companion to experiment. Recent advances in computing hardware and software have allowed large-scale calculations, e.g., on the mechanism of electron transfer from the reductase to the ferric-heme in CYP101,[76] as well as sophisticated QM/MM calculations with correlated ab initio QM methods using MR-CI (multireference configuration interaction)[77,78] and CASSCF (complete-active-space self-consistent field)[79] approaches. Over the past few years, there has been an increasing number of QM/MM studies

dedicated to P450 that reveal many exciting new insights about the role of the protein in shaping the structures and reactivity patterns of these enzymes. All these studies have had a considerable impact on the field and led to an ever-growing surge of theoretical activity. It was therefore deemed timely to review the theoretical results provided in recent years by QM/MM methods, *especially on the issues of P450 reactivity and selectivity*. The main goal is to project the new insight gained on the bonding and reactivity of P450s, including studies of novel mechanisms whereby P450 functions as a peroxidase and participates in proton-coupled electron transfer (PCET).[58,80]

Some of the material covered in previous reviews[28−32,81−83] is included here too in order to create a unified survey. We start with a brief summary of the available QM/MM methods (section 2), and then we describe the various species in the catalytic cycle (section 3), the principal mechanisms by which the enzyme catalyzes oxygenation reactions (section 4), and predictive models of mechanisms and reactivity (section 5).

2. Theoretical Methods

Combined QM/MM theory has emerged as the method of choice for modeling local electronic events in large biomolecular systems. The basic idea is to describe the active site (where chemical reactions or electronic excitations occur) by quantum mechanics, as accurately as needed, while capturing the effects of the biomolecular environment by molecular mechanics, i.e., at the classical force field level. The QM/MM concept was introduced, as early as 1976, by Warshel and Levitt, who presented a semiempirical QM/MM treatment for a chemical reaction in lysozyme.[84] The QM/MM approach has become ever more popular over the past two decades, and numerous reviews have addressed both the development of the QM/MM methodology and its application in the biomolecular area. Here we cite only a

Scheme 2. Different Model Systems and Their Abbreviations

(Por, SH)   (Por, SMe)   (Por, ext-SCys)   (PPIX, SCys)
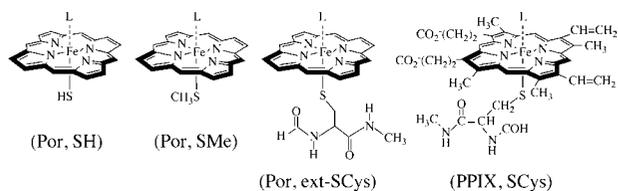
few of these[85−97] and refer to two recent reviews[93,97] for an up-to-date coverage of the field with a comprehensive survey of the literature (755 and 627 references, respectively). In this section, we address theoretical issues relevant for QM/MM studies of enzymes (section 2.1), discuss some practical issues that arise during such QM/MM calculations (section 2.2), and introduce notations and conventions for the following sections (section 2.3).

2.1. Methodological Issues in QM/MM Studies

In composite theoretical approaches such as QM/MM, a guiding principle is to retain the formalism of the methods that are being combined (as far as possible) and to adopt well-defined rules for their coupling. In the following, we address the methodological choices that need to be made, both in general and specifically in the case of QM/MM studies on cytochrome P450 enzymes.

2.1.1. QM/MM Partitioning

Inherent to the QM/MM concept is the division of the system into the inner QM region (active site) that is treated quantum-mechanically and the outer MM region (protein environment) that is described by a force field. In the boundary region at the interface, the standard QM and MM procedures may be modified in some way (e.g., by adding link atoms; see below). The choice of the QM region is usually made by consideration of the chemical problem: chemical arguments normally suggest a minimum-size QM region which can then be enlarged to check the sensitivity of the QM/MM results with regard to such an extension. In the case of cytochrome P450, the minimum-size QM region will contain the heme (iron-porphine without side chains) and the axial ligands (with SH being the smallest model to represent the cysteinate ligand; see Scheme 2). Enlarged QM regions may include the heme side chains, larger representations of the cysteinate ligand, selected residues of the protein environment, and water molecules close to the active site, depending on the problem under study. The actual choice requires care because the QM/MM system needs to be electronically balanced and properly screened with regard to electrostatic interactions.

Standard QM/MM applications employ a fixed QM/MM partitioning which is defined once and for all at the outset. It is possible to allow the boundary to move during QM/MM simulations (adaptive partitioning).[98] In another generalization of the QM/MM approach, one may go from two-layer to three-layer approaches, for example by using a continuum solvation model as a third layer[99,100] or by adopting an ONIOM-type scheme with two different inner QM layers and an outer MM layer.[101] Such more involved QM/MM treatments have not yet been applied to P450 enzymes.

2.1.2. Choice of QM Method

The selection follows the same criteria as in pure QM studies (accuracy and reliability versus computational effort). In the case of metalloproteins such as P450 enzymes, DFT is the natural choice which has been used in the vast majority of QM-only model studies[32] and also in most recent QM/MM work. There is a myriad of density functionals available, ranging from simple local functionals (LDA, local density approximation) via gradient-corrected functionals (GGA, generalized gradient approximation), meta-GGA functionals (with second derivative terms), and hybrid functionals (with admixture of HF exchange) to more advanced forms. The various functionals with their shorthand notations have been described amply, and the reader may consult a recent review for a brief survey.[102] Some of these functionals are derived from first principles and are thus parameter-free, while others involve a semiempirical calibration against theoretical or experimental reference data. In practice, gradient-corrected functionals such as BP86 are normally considered suitable for metalloenzymes, but the preferred choices are hybrid functionals such as B3LYP. Generally speaking, B3LYP is known to perform well for a number of properties,[103] even though the more recent, highly parametrized functionals such as those from the M06 series[102] show smaller overall deviations in various standard benchmarks, especially with respect to relative energies and reaction barriers. On the other hand, it has been established that B3LYP reproduces relative spin state energies in heme systems quite well, apparently since it has the "right" admixture of Hartree−Fock exchange, which is an essential prerequisite for realistic studies on P450 enzymes. Many other successful B3LYP applications have been reported early on in the P450 area,[32] and it is thus no surprise that B3LYP has become by far the most popular QM component in DFT/MM studies of heme enzymes. Nevertheless, in view of several documented shortcomings of B3LYP, one must remain critical and refrain from expecting too much accuracy, in particular with regard to the absolute values of barriers. While B3LYP is dominant at present,[103] we anticipate an increasing use of more advanced density functionals and of correlated ab initio methods in future QM/MM work.

Ab initio QM methods can in principle provide a convergent path to the correct solution of the nonrelativistic Schrödinger equation and are thus indispensable for accurate theoretical work. It has recently been demonstrated in the case of p-hydroxybenzoate hydroxylase (PHBH) and chorismate mutase that QM/MM calculations can reproduce the barriers of enzymatic reactions to within 1−2 kcal mol^{-1} when applying high-level correlated ab initio QM treatments (i.e., local coupled cluster methods).[104,105] These single-reference examples are however easier to handle than the multireference situations that are commonly encountered in heme enzymes. Ab initio MR-CI/MM and CASSCF/MM calculations have been reported for some P450 intermediates and analogous heme enzymes recently,[77−79] but entire P450 reactions have not yet been treated at this level because of technical limitations that prevent the use of a balanced active space along the complete reaction path. The ab initio MR-CI/MM results for relative spin state energies[77,78] are quite similar to those from B3LYP/MM and thus provide further justification for the latter approach. Such ab initio QM/MM calculations are expected to gain importance for validating crucial DFT/MM predictions.

2.1.3. Choice of MM Method

MM force field methods express the total energy of a molecular system as a sum of bonded energy terms, electrostatic terms, and van der Waals interactions. The bonded energy terms normally consist of harmonic stretching and bending potentials and periodic torsion potentials. The nonbonded electrostatic interactions are commonly treated as Coulomb interactions between fixed point charges. The van der Waals interactions are normally represented by a Lennard-Jones potential with attractive $1/r^6$ and repulsive $1/r^{12}$ terms. Established force fields of this kind are available for proteins (e.g., CHARMM, AMBER, GROMOS, and OPLS). These standard force fields are generally considered suitable for biomolecular applications, although their relative merits have not yet been studied systematically at the QM/MM level. In QM/MM work on P450 enzymes, CHARMM has been most widely used as the MM component.

The use of fixed MM point charges in the standard protein force fields neglects polarization effects. The logical next step toward higher accuracy should thus be the transition to polarizable force fields, which are currently developed in the biomolecular simulation community using various classical models (e.g., induced dipoles, fluctuating charges, or charge-on-spring models). The QM/MM formalism has been adapted to handle such polarizable force fields,[93,106] and one may expect corresponding QM/MM applications once these new force fields are widely accepted. In the meantime, essential polarization effects in residues close to the active site may be taken into account in QM/MM studies by a suitable extension of the QM region.

2.1.4. Subtractive versus Additive Schemes

Subtractive QM/MM schemes are interpolation procedures. The QM/MM energy is obtained from three separate calculations: the MM energy of the entire system is corrected by adding the QM energy of the QM region and subtracting the MM energy of the QM region. In this approach, the QM/MM interactions are handled entirely at the MM level. This may be problematic for the electrostatic interactions, which will then typically be computed from fixed atomic charges in the QM and MM regions (neglecting the polarization of the QM region by the MM charges).

In additive QM/MM schemes, the QM/MM energy is the sum of the QM energy of the QM region, the MM energy of the MM region, and the QM/MM coupling terms. The latter normally include bonded terms across the QM/MM boundary and nonbonded van-der-Waals terms (both generally handled at the MM level) as well as electrostatic interaction terms which are modeled explicitly and can thus be described realistically using QM-based treatments (allowing for polarization of the QM region by the MM charges). This is probably the reason why most of the published QM/MM studies on P450 enzymes employ an additive scheme.

2.1.5. Electrostatic QM/MM Interactions

There are several strategies for handling these interactions ranging from simple mechanical to electronic and polarized embedding.[107] In the case of the polar protein environment, an electronic embedding is generally considered necessary. Thus, the electrostatic interactions are incorporated into the QM calculations by including the MM charges into the QM Hamiltonian, so that the QM subsystem is polarized in response to the electric field of the protein environment. The resulting QM density should be much closer to reality than that from a gas-phase model calculation. Electronic embedding is the standard choice in current QM/MM studies of enzymes.

The most refined embedding scheme is polarized embedding where the back-polarization of the MM region by the QM region is included as well, preferably in a fully self-consistent manner. This approach has been applied only rarely up to now (and not at all for P450 enzymes). It is expected to become more popular when more QM/MM studies with polarized force fields appear, because polarized embedding is the natural coupling scheme in this case.

2.1.6. Boundary Treatments

It is often unavoidable that the QM/MM boundary cuts through a covalent bond. The resulting dangling bond must be capped to satisfy the valency at the frontier, and in the case of electronic or polarized embedding, one must prevent overpolarization of the QM density by the charges close to the cut. To cope with these problems, there are several options. The most common approach is to introduce a link atom, typically hydrogen,[108,109] but it is also possible to use specifically parametrized boundary atoms[110] or pseudo-bonds[111] or specifically tailored pseudopotentials[112] or frozen orbitals.[113−115] All of these options seem to work well if proper care is exercised, especially with regard to the treatment of electrostatics at the QM/MM boundary.[116−118] In practice, the link atom approach is used most widely in QM/MM studies of P450 enzymes, often in combination with a charge shift scheme at the boundary.[118,119] It is common to fix the link atom in the bond being cut, at some well-defined distance from the QM frontier atom, and to redistribute the forces acting on it to the two neighbor atoms,[120] which effectively removes the artificial degrees of freedom of the link atom.

2.1.7. QM/MM Geometry Optimization

Special techniques are available to optimize the structures of large biomolecules with thousands of atoms at the QM/MM level that exploit the partitioning into a QM region, where energy and gradient evaluation are computationally expensive, and an MM region, where these calculations are almost for free. These techniques include a linear-scaling fragment-based divide-and-conquer optimizer[121] and microiterative strategies with alternating geometry relaxation in the core region (containing the QM region) and the environment.[122] Their combined use allows the efficient optimization of minima and transition states at the QM/MM level, also in the case of electronic or polarized embedding.[123] However, QM/MM geometry optimizations of the stationary points along a single reaction path are of limited significance, since there are many closely related paths in large biomolecules because of their conformational flexibility. It is thus advisable to determine at least several representative transition states with their corresponding minima in order to assess the influence of the conformational diversity of the environment. This strategy is often followed in DFT/MM studies of P450 enzymes.

2.1.8. QM/MM Molecular Dynamics

The vast configuration space that is accessible to large molecules can be sampled using molecular dynamics (MD) or related techniques. Straightforward QM/MM MD calculations are computationally demanding, however, and routinely affordable only at the semiempirical QM/MM level. For higher-level QM/MM approaches, approximate techniques are required that reduce the computational cost, e.g., by exploiting the QM/MM partitioning. One strategy is to avoid the expensive direct sampling of the QM region while fully sampling the MM configuration. An early example of this approach[124] keeps the QM region fixed and uses ESP (electrostatic potential)-derived charges for the QM atoms to evaluate the electrostatic QM/MM interactions during the MD run, in the context of free energy perturbation theory.[124,125] Using this procedure, it has recently been shown that the effects of sampling on the computed DFT/MM reaction profile are minor in the case of hydrogen abstraction in cytochrome P450$_{cam}$.[126]

2.1.9. QM/MM Energy versus Free Energy Calculations

In principle, theoretical studies should aim at free energy calculations, since chemical thermodynamics and kinetics are governed by free energy differences. These can be determined by statistical mechanics through sampling procedures (e.g., using thermodynamic integration, umbrella sampling, or free energy perturbation theory). Recent advances in this area have been summarized in an excellent review.[96] It is straightforward to perform free energy calculations at the semiempirical QM/MM level using any of these techniques (as demonstrated in several recent applications[125−128]), but the necessary sampling tends to become expensive with ab initio or DFT QM components. This is the reason why QM/MM free energy studies on P450 enzymes are still very rare. Judging from the available evidence, however, the differences between QM/MM energy and free energy profiles in enzymatic reactions tend to be small as long as local chemical events are investigated (e.g., hydrogen abstraction in P450$_{cam}$, OH transfer in PHBH, and nucleophilic substitution in fluorinase).[126] For reactions of this kind, one can thus expect to gain valuable mechanistic information by the less demanding DFT/MM geometry optimization studies that are commonly performed for P450 enzymes.

2.2. Practical Issues in QM/MM Studies

QM/MM methods have matured over the past two decades, but canonical "black-box" procedures have not yet been established. It would thus seem worthwhile to address briefly some of the practical problems and choices that are encountered in QM/MM work.

2.2.1. QM/MM Software

QM/MM calculations require efficient programs with wide-ranging functionality. Several of the standard QM and MM packages offer QM/MM capabilities as an add-on. The alternative is to interface independent external QM and MM programs to a central control module that supplies the QM/MM coupling as well as routines for standard tasks such as geometry optimization and molecular dynamics. In the long run, the latter modular approach is expected to be more flexible, since it can provide access to more QM/MM functionality (e.g., facilitating the combination of many different QM and MM methods). An example for such a modular implementation is the ChemShell package.[118,129]

2.2.2. QM/MM Setup

Realistic starting structures for QM/MM work will normally be derived from experiment (typically from an X-ray crystal structure). Small modifications of experimental structures are common in the setup phase, e.g., replacement of an inhibitor by substrate or substitution of selected residues to generate a specific mutant. The available structural information from experiment is generally incomplete and often not error-free. It is thus usually checked and processed employing the protocols developed by the classical simulation community. This includes, for example, adding missing hydrogen atoms, adding water molecules in "empty" spots, assigning the protonation states of titrable residues (e.g., His, Asp, Glu), and checking the orientation of residues in ambiguous cases (e.g., His, Asn, Gln). The correct choice of protonation states can often be crucial in QM/MM work; the assignments may be guided by empirical rules, structural considerations (e.g., via visual inspection of local hydrogen-bonding networks), empirical algorithms for pK_a predictions (e.g., PROPKA[130]), and explicit pK_a calculations (e.g., of Poisson−Boltzmann type). A related issue concerns the overall neutrality of the entire system, which may be achieved by suitable placement of counterions or by neutralization of charged residues at the surface of the protein.

After these initial setup choices are made, the enzyme is put into a water box and relaxed by a series of constrained energy minimizations and MD runs at the MM level (which may necessitate the derivation of missing force field parameters for the "nonstandard" parts of the system). After equilibration, the system is subjected to a classical production run from which several (5−10) snapshots are taken as starting geometries for the QM/MM work (often supplemented by an "initial" snapshot that resembles the available X-ray structure as much as possible). The starting structures generated in this manner for a P450 enzyme in a droplet of water typically contain around 20,000−25,000 atoms.

It is obvious that the setup phase requires a lot of work prior to the actual QM/MM calculations, and it should be emphasized that errors and wrong choices during setup can normally not be recovered at a later stage. These issues have been discussed in more detail in a recent review.[93] For the specific case of cytochrome P450$_{cam}$, there is a systematic QM/MM investigation of how the barrier to hydrogen abstraction from camphor is affected by the choices made during setup.[131] In this context, it should be noted that early QM/MM work on different P450$_{cam}$ intermediates and on different steps of the catalytic cycle employed separate setups that had been generated independently, which raises the question whether it is possible to handle the whole cycle with the same QM/MM setup. This is indeed the case: following a standard protocol that starts from the X-ray structure of the ferrous dioxygen complex and that involves careful assignment of protonation states and hydration, one can generate a common setup that gives QM/MM results for P450$_{cam}$ which are generally consistent with those obtained previously with individual setups.[54] This confirms that setup issues can be resolved in a consistent manner if proper care is exercised.

2.2.3. Accuracy of QM/MM Results

Given the large number of choices in QM/MM work, it is very difficult to converge the QM/MM results with regard to all computational options. Typical QM/MM studies on P450 enzymes may employ DFT/MM calculations with a standard protein force field such as CHARMM, electronic embedding, and a link atom boundary treatment. In practice, the sensitivity of the QM/MM results with regard to the chosen force field, embedding scheme, and boundary treatment is normally not checked (even though the results will depend on these choices). On the other hand, since the QM treatment influences the outcome rather directly, it is quite common to use different basis sets and different density functionals, at least in single-point energy calculations, to assess the reliability and robustness of the QM/MM results within DFT. Systematic convergence studies with high-level ab initio QM components are still rare, however,[104,105] and the most that has been achieved up to now for P450$_{cam}$ are single-point ab initio MR-CI/MM calculations to validate relative spin state energies for selected intermediates.[77,78] Finally, it is advisable to check the convergence of the results also with regard to the QM/MM approach itself. This can be done by QM/MM calculations with QM regions of increasing size, since extension of the QM region should decrease the influence of the QM/MM coupling and MM force field terms and, thus, make the QM/MM treatment more realistic overall. Such checks are often performed in QM/MM studies on P450 enzymes.

2.2.4. QM/MM Geometry Optimization

For systems with 20,000−30,000 atoms, it is usually sufficient to allow only around 1000 atoms to move during optimization (i.e., the active site and its environment) while the outer part remains fixed at the initially prepared snapshot geometry. It is essential in QM/MM work on reaction profiles to retain the same conformation of the optimized active region in order to ensure a smooth and continuous reaction path. If this requirement is not satisfied (e.g., by the flip of a distant hydrogen bond or some other remote conformational change), the QM/MM results from geometry optimization become spurious.[131] Experience shows that such problems can normally be avoided with an optimized active region up to about 1000 atoms (as typically used in QM/MM studies on P450 enzymes).

2.3. Conventions and Notations

It is obvious from the preceding discussion that there are not yet "canonical" black-box procedures for QM/MM methods and that one must decide on a large number of choices in QM/MM studies. Therefore, QM/MM calculations cannot be characterized by a short label, in contrast to QM studies on small molecules, where it is often sufficient to quote the acronyms for the QM method and basis set. A full specification of the chosen QM/MM options is impractical in the context of a review, and in this regard, the readers will thus normally be referred to the original literature in the following sections.

In spite of this general caveat, it is still possible to outline a "consensus" QM/MM approach that has emerged in the computational P450 community over the past years. This approach employs DFT(B3LYP)/MM calculations with an established protein force field (mostly CHARMM, sometimes OPLS or AMBER), electronic embedding, and a link atom boundary treatment. Open-shell systems are normally described by an unrestricted B3LYP treatment; this is implied when we use the label B3LYP while a restricted open-shell treatment will be denoted by ROB3LYP (analogous conventions apply for other functionals). The QM/MM setup usually starts from a crystal structure and involves initial classical MD calculations for system preparation. The actual QM/MM studies normally employ geometry optimization of the active site and its surroundings. Readers should envision this kind of QM/MM approach when we review QM/MM studies in the following sections.

There are two factors that affect the outcome of QM/MM calculations on P450 enzymes rather directly and that will therefore often be mentioned in the following sections, namely the QM region and the basis set for the QM calculation. The QM region needs to be adapted to the problem under investigation, but there are a few standard choices that will be introduced in section 3.1 (see Scheme 2, which also gives the corresponding notation). The commonly used basis sets show a greater diversity. Many of them include Pople-type basis sets[132] on the nonmetal atoms, such as 6-31G, 6-31G*, etc. Based on the experience with QM model calculations, the earliest QM/MM work on P450[53] employed a basis set (B1) of double ζ (DZ) quality that describes iron by an effective core potential and the LACVP basis[133] and the other atoms by the 6-31G basis. The B1 basis is often applied for geometry optimization, while subsequent single-point QM/MM calculations normally use larger basis sets to determine more reliable energies. Early examples for such larger sets are the B2 basis, which differs from B1 by including polarization functions (6-31G*) on the six ligands to iron (four pyrrole nitrogen atoms, the oxo atom, and the sulfur atom of the coordinating cysteine), and the B2W basis,[134] which differs from B2 by describing iron by an all-electron Wachters basis with an additional diffuse d function and a set of f polarization functions.[135−137] More recent QM/MM work often employs more extended basis sets for single-point calculations, typically of triple ζ (TZ) quality with one or more sets of polarization (P) functions. One such example is the B′ basis for P450$_{cam}$,[54,78] composed of the Wachters all-electron basis for iron, polarized triple ζ valence basis sets[138] for the six atoms coordinated to iron (TZVP), the C^5 atom of camphor (TZVP), the H^{5exo} atom of camphor (TZVPP), and the smaller SV basis[139] for the remaining atoms. Special basis sets are needed for special purposes, e.g., for the computation of Mössbauer parameters, which requires a flexible uncontracted basis in the core region of iron.[77] It is clearly impractical to specify in this review all the basis sets that have been used in recent QM and QM/MM studies on P450 enzymes. In the following sections, we shall refer to the labels introduced above (B1, B2, B2W, B′) when deemed appropriate, using a notation of the type DFT(B3LYP/B1)/MM. We shall explicitly specify other basis sets when necessary, but we shall also quote results from the literature in more general terms (e.g., those obtained with the largest basis set) and refer the reader to the original paper for technical details. As a general caveat, we note that the label B2 has been used in the literature also for larger basis sets than the one introduced above; in the following sections, such usage will be indicated either explicitly or by reference to the original paper.

2.4. Coverage and Sources of Computational Results

Most of the older theoretical literature on P450 up to the year 2000 can be found in the reviews of Loew and Harris[28] and Loew.[140] Studies that have appeared between 2000 and 2004 and during part of 2005 were summarized in our 2005 review.[32] These studies include complete QM(DFT)-only modeling of the catalytic cycle of P450 by Rydberg et al.,[141] Shaik et al.,[31,142,143] and Hata et al.[144] In addition, the 2005 *Chemical Review* article covers extensively the two-state reactivity (TSR) concept in P450 enzymes[25,30,37,83,145−147] and discusses a variety of mechanistic studies, using QM(DFT)-only methods, e.g., of C−H hydroxylation,[39−43,45,146−153] double bond epoxidation,[35−37,44,149,150] benzene and arene hydroxylation,[33,34,38] and sulfoxidation.[154] These studies will be mentioned in passing but will not be covered thoroughly in the present review.

The 2005 *Chemical Review* article included only a few QM(DFT)/MM studies.[53,59,134,155−157] However, since that review has been published, there have appeared a variety of QM(DFT)/MM studies, as well as single-point QM(CI)/MM calculations. Many of these studies concern the entire catalytic cycle of CYP101 and of selected species in human isoforms, e.g., CYP3A4.[54−56,77,78,158−162] Some of these studies describe analogous species in the selenocysteine mutant, the Mn mutant, the D251N and L358P mutants of CYP101, as well as the CPO and NOS enzymes.[62,63,66−69,160,161] In addition, QM(DFT)/MM has also been used to study the mechanisms of C−H hydroxylation, C=C epoxidation, aromatic hydroxylation, and C−C coupling.[57−61,131,163,164] Mechanistic studies since 2005 using QM-only methods addressed dehalogenation of aromatic compounds,[165] N-dealkylation in *N,N*-dimethylanilines,[166,167] ethanol oxidation,[168] and sulfoxidation.[169] QM-only methods were also used for computational investigations of C−H hydroxylation, aromatic hydroxylation, and heteroatom oxygenation,[170−173] as well as for valence bond modeling of the mechanism and reactivity patterns in C−H hydroxylation.[174] These QM/MM and QM-only studies will form the core of the present review.

3. The Catalytic Cycle of P450 Enzymes

To understand P450 reactivity, one must first understand the catalytic cycle of CYP enzymes, since, in many of the isoforms of this enzyme, the rate-determining event is not substrate oxidation; this poses obvious difficulties to mechanistic investigations of these reactions. It is established by now that CYP enzymes operate by means of the generic catalytic cycle drawn in Figure 1,[175,176] where, except for **1**, the heme is depicted by two bold horizontal lines, and the cysteinate proximal ligand is abbreviated as CysS. The resting state (**1**) of the enzyme generally involves a ferric-protoporphyrin complex, coordinated to the proximal cysteine residue and to a water molecule at the distal side; in some isoforms, such as CYP2D6, the water ligand is missing.[177] This hexacoordinated Fe^{III} complex (**1**) exists largely in a low-spin doublet state, and as long as this is the case, the enzyme will remain inactive.[175,178] When the substrate (for example, an alkane, AlkH) enters, it displaces the water molecule, leaving a pentacoordinated ferric-porphyrin (**2**), and changes the spin state to a sextet. Due to the loss of the water ligand, the ferric complex (**2**) is a slightly better electron acceptor than the resting state (by approximately 130−300 mV), and this is sufficient to trigger an electron
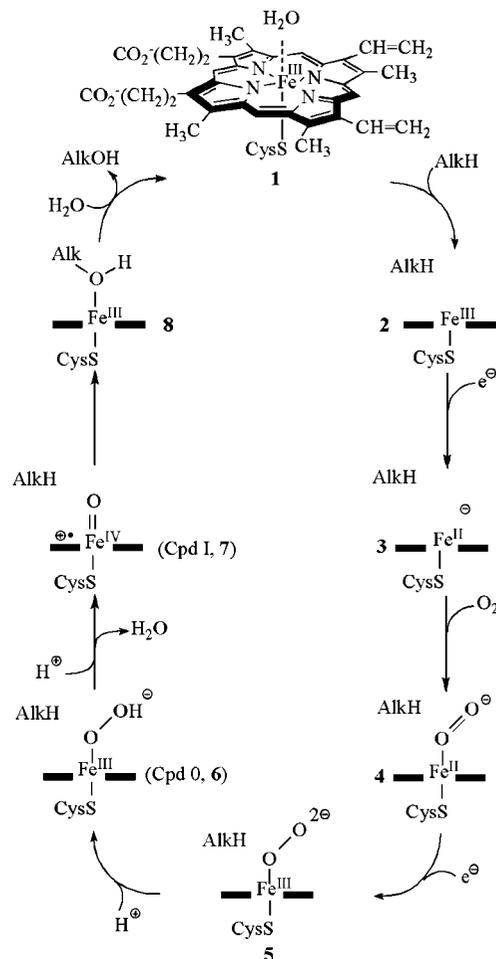
Figure 1. Schematic representation of the catalytic cycle of P450. Reprinted with permission from ref 32. Copyright 2005 American Chemical Society.

transfer from a reductase protein (which operates at 200 mV), leading to a high-spin ferrous complex (**3**), which is a good dioxygen binder. The binding of molecular oxygen yields the oxy-ferrous complex (**4**), which has a singlet spin state and is a good electron acceptor. This, in turn, triggers a second reduction of the system to generate the ferric-peroxo anion species (**5**). This second reduction is in some cases, albeit not always,[179] the rate determining step in the catalytic cycle.[180] Since the reduced dioxygen-complex (**5**) is a good base, it gets quickly protonated to form the ferric-hydroperoxide species (**6**) that is called Compound 0 (Cpd 0). Thereafter, Cpd 0 abstracts another proton, and a water molecule splits off to generate the high-valent iron-oxo species (**7**), so-called Compound I (Cpd I). In the common case, Cpd I monooxygenates the substrate; for example, the alkane in Figure 1 is converted to an alcohol (**8**), which exits the pocket, thereby allowing water molecules to re-enter and restore the resting state (**1**).

Note that a single water molecule is used to gate the cycle in Figure 1; however, the enzyme also requires a protonation machinery to come full cycle. The protonation machinery is believed to involve an acid−alcohol pair (in the case of CYP101, the Asp_{251} and Thr_{252} residues shown in Figure 2).[175,178,181] It is thought that these residues, and perhaps also Glu_{366} that is not shown in the drawing,[182] shuttle the protons to the dioxygen moiety in **4**−**6**, via water molecules.[182] Indeed, mutation of Asp_{251} reduces the rate of camphor hydroxylation, while mutation of Thr_{252} can virtually stop
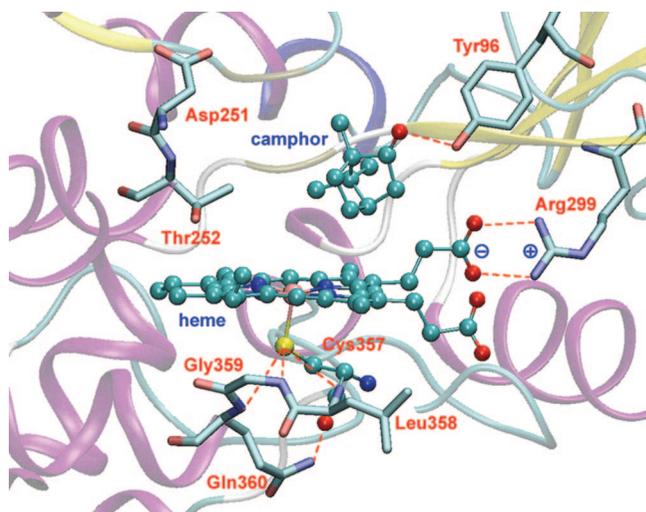
Figure 2. Close-up of the active site of CYP101 as taken from the X-ray structure (pdb code: 1DZ8) of Schlichting et al.[178] with some essential groups highlighted.

the process:[175,178,181] the T252A mutant performs only C=C epoxidation, presumably via Cpd 0.[183] Oddly, the double mutant, with both Asp251 and Thr252 mutated to residues that do not shuttle protons, shows significant camphor hydroxylation activity,[184] thereby implying that proton shuttle can occur through water networks with no need for the alcohol–acid pair of residues. The role of water is like a two-edged sword: On the one hand, water molecules are important for protonation,[182] but on the other hand, too much water leads to formation of hydrogen peroxide and other uncoupling products that waste the oxygen and the reducing equivalents.[185−187] Thus, the enzyme has to function in a middle road: having some water molecules and at the same time denying them free access to the neighborhood of the Fe-OO moiety, in order to be able to activate bonds and to produce oxidized products efficiently. Although some isoforms, especially the human ones,[21] are promiscuous and have large pockets, which can accommodate many water molecules[188] and more than one substrate molecule,[189] nonetheless, the control of water is influenced in part by the tightness of substrate binding by the enzyme.[190,191] For example, in the case of CYP101 (Figure 2), this is achieved by a substrate-binding domain made from Tyr96, which forms a hydrogen bond (H-bond) with the carbonyl group of camphor and holds the substrate steady in place just above the FeO moiety. In addition to this feature, as shown in Figure 2, the sulfur atom of the cysteinate ligand interacts with the three amidic groups of Gln360, Gly359, and Leu358 (only one of these has the geometry of a classical H-bond), while the carbonyl group of the cysteine interacts with Gln360.[178,192] These interactions appear to be essential for the stability and activity of the enzyme.[193,194] These issues have all been studied by QM/MM and will be addressed in the next subsections that summarize the theoretical studies of the different species in the catalytic cycle (structures **1**−**8**).

Concerning reactivity, there is uncertainty about the details of the cycle starting from **6** and onward back to **1**; Cpd I is elusive in the native cycle, and its mechanism of formation is still putative. The reason for its elusiveness is thought to be its fast formation and fast consumption, but this is not yet clear. Furthermore, the range of reactivity patterns of Cpd I is still not fully understood. Thus, while the normal activity involves monooxygenation as in C−H hydroxylation (Scheme 1), the enzyme seems to be much more versatile.

For example, in the bacterial enzyme CYP245A1 (P450 StaP),[195] the enzyme seems to function more like cytochrome *c* peroxidase,[196] while, in CYP3A4, the enzyme often requires a few substrate molecules to operate efficiently.[189] As we shall show, QM/MM theory provides very useful insight into some of these major questions by considering the properties of Cpd I and its reactivity patterns.

3.1. Models of the Enzyme

Since the complete enzyme cannot be described by QM(DFT), one usual practice among the workers in the field is either to truncate the active species and calculate the model system by QM only or to embed it in the protein and use QM/MM. Generally, the heme is treated as porphine by removing all side chains of the protoporphyrin IX (PPIX) macrocycle. In some cases, these side chains were included in the calculation.[59,68,131,155,157,163] The cysteinate ligand is modeled either as thiolate (SH^-), as methyl mercaptide (SCH_3^-), or as cysteinate anion ($SCys^-$), and in some cases as an extended cysteinate (ext-SCys), augmented by the peptide bonds of the adjacent Leu358 and Leu356 residues. DFT(B3LYP)/MM calculations[53] showed that gas-phase DFT(B3LYP) calculations using SH^- give generally better agreement with the DFT/MM results than SCH_3^-, or even the full cysteinate ligand. Many of the calculations, especially those on reaction mechanisms, used the simplest model with porphine and SH^- as the thiolate ligand, or porphine and SCH_3^-. The most recurring model systems and their notations are described above in Scheme 2. There are however, quite a few other models, which incorporate porphyrin substituents, and specific key residues in the active site; these will be mentioned specifically wherever deemed necessary.

3.2. Oxidation States and Key Orbitals of Species in the Catalytic Cycle

Since the species in the cycle (Figure 1) possess a dense orbital manifold that is only partially filled, this gives rise to many closely lying spin and electromeric states, which differ in the oxidation states on the iron and the ligands.[25] It is, therefore, important to establish a common language by beginning with the oxidation state formalism of these P450 species and then describing a generic orbital picture, which describes at least most of the species in the cycle. More orbital diagrams will be presented as we proceed along with the discussion of various specific issues and reactivity patterns.

The oxidation state of iron is given as Roman numerals near the iron atom in the species in Figure 1; note that the total charge of the species is indicated in a circle near the species. We must remember that the oxidation state formalism is a heuristic device, which assumes that all metal−ligand bonds are "ionic". It turns out that this assumption is useful to keep track of the "number of d-electrons" that remain on the metal ion in redox processes; however, the oxidation number is not identical to and may be even remote from the charge distribution in the molecule (or from the real integrated d-population altogether). In this formalism, the porphyrin ligand has an oxidation number of 2−, while the cysteine ligand is 1−. Any molecular ligand that has already an electronic octet like water is counted as zero (0), while, for any other ligand, which does not possess an electronic octet, the assigned oxidation number is determined by the number of electrons that this ligand "takes from iron" to fill
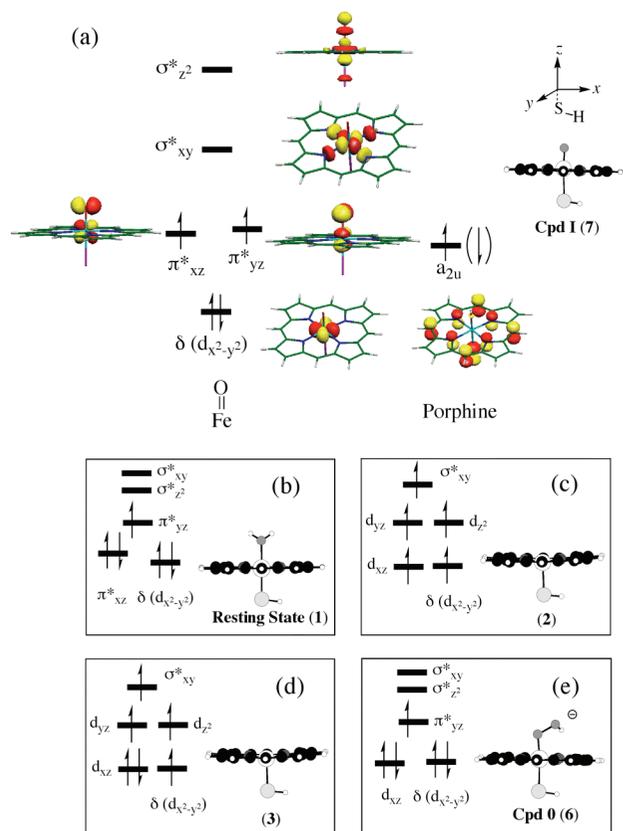
Figure 3. (a) Orbitals and their occupations in Cpd I (**7**). (b–e) Orbital occupations for the resting state (**1**), the pentacoordinated complexes (**2** and **3**), and Cpd 0 (**6**).

its octet; for example, an oxo ligand is 2−, while an OOH ligand is 1−. Thus, for all the neutral complexes, with and without a water ligand, the oxidation number of iron will be 3+, which is labeled in Figure 1 as Fe^{III}. This corresponds to the resting state (**1**) and the pentacoordinated complex (**2**). After accepting one electron from the reductase, the pentacoordinate complex (**3**) becomes negatively charged and iron is in the ferrous state Fe^{II}. This complex (**4**) has an O_2 molecule that possesses an octet, and hence it can be formally written as $Fe^{II}O_2$. It is thus referred to as oxy-ferrous; as we shall see later, the bonding features of **4** are more appropriately represented by $Fe^{III}O_2^-$.[79] In the reduced complex (**5**) with an additional electron, it is customary to assign the dioxygen ligand oxidation number of 2−, and hence, **5** is a ferric-peroxy species, $Fe^{III}O_2^{2-}$, with a total charge of 2−. After protonation, Cpd 0 (**6**) is a ferric-hydroperoxide, Fe^{III}−OOH, with a total charge of 1−. In Cpd I (**7**), which is neutral, the formal oxidation state of the entire heme is 5+, and this is assigned partly to the iron as Fe^{IV} and partly to the porphyrin ring, which becomes a π-cation-radical $Por^{\bullet+}$. This assignment is not part of the rules of the oxidation-state formalism but is based on knowledge of the electronic structure (see Figure 3). Finally, the product complex (**8**) is a neutral Fe^{III} species, and as such, Cpd I has two-oxidation equivalents above those of the resting state and of the product complex.

The oxidation states in Figure 1 provide some rationale for comprehending the electronic structure of the P450 species. The orbital labels and ground state configurations are shown in Figure 3 for **1–3**, **6**, and **7**, whereas, for **4** and **5**, the bonding features are more complex and will be described separately. Thus, each species in the figure has five d-type orbitals, which differ from species to species

depending on the interaction of the d-orbitals with the ligand orbitals. In addition, a porphyrin-based orbital is drawn for **7** alongside the d-type orbitals, in Figure 3a. The total oxidation state on iron and porphyrin determines the occupancy of these orbitals, while the interaction strength with the ligand will determine the spread of these orbitals and hence the propensity for high-spin ground states.

In Figure 3a we show the orbitals and electronic configuration of Cpd I (**7**), on the left-hand side the five metal d-type orbitals, and near them the high-lying mixed porphyrin-thiolate orbital, so-called a_{2u}. The sulfur−porphyrin mixing in a_{2u} is significant, since thiolate is a powerful electron donor,[197,198] but with weaker donor ligands, such as imidazole,[74,199] this mixing is negligible and a_{2u} is a pure porphyrin orbital. As we shall see, the singly occupied a_{2u} orbital is sensitive to the donor capability of the thiolate, and in fact to any good donor that may participate in the electronic structure of Cpd I.[199]

The d-block in Figure 3a is split into the characteristic three-below-two sets, which are nascent from the purely octahedral t_{2g} and e_g sets with an additional splitting of t_{2g} into one-below-two. The lowest three orbitals are, in increasing energy order, the nonbonding $\delta(d_{x^2-y^2})$ orbital and the two π^*_{FeO} (π^*_{xz} and π^*_{yz}) orbitals; the latter two involve antibonding interactions between the metal $3d_{xz,yz}$ and the oxygen $2p_{x,y}$ atomic orbitals. The two high-lying virtual orbitals of Cpd I are σ^* orbitals; one (σ^*_{xy}) describes the Fe−N antibonding interactions and the other ($\sigma^*_{z^2}$) those along the O−Fe−S axis; the ordering of the σ^* orbitals is dominated by the length of the Fe−O bond, and when it is short as in Cpd I, $\sigma^*_{z^2}$ rises above σ^*_{xy}. With the $Por^{\bullet+}Fe^{IV}$ oxidation state, there are four electrons in the d-block, and the porphyrin-based orbital is singly occupied. This occupation gives rise to a pair of almost degenerate spin states, one quartet ($S = 3/2$) with all single electrons being up-spin, and one doublet state with a spin flip in the porphyrin-based orbital. Some other spin states ($S = 5/2$ and $3/2$) are also relatively low lying and involve 1e-promotion from δ to σ^*_{xy}.[77,200]

In those species where the Fe−O bond is longer than that in Cpd I (**1**, **6**) or where the complex loses its sixth ligand altogether (**2**, **3**), two orbitals exhibit weaker antibonding interactions; these are $\sigma^*_{z^2}$ and one of the π^* orbitals, which undergo significant stabilization, as indicated in Figures 3b–d. In **2** and **3**, both orbitals become nonbonding d orbitals, d_{z^2} and d_{xz}, while, in **1** and **6**, the $\sigma^*_{z^2}$ orbital gets stabilized and descends below σ^*_{xy}, whereas the π^*_{xz} orbital is almost a nonbonding d_{xz} orbital. The weakening of the ligand field shrinks the spacing of the d-orbitals, thus giving rise to low lying high-spin states as ground states.

3.3. The Resting State (1) and the Pentacoordinate Intermediates (2, 3)

3.3.1. The Resting State

The resting state (**1**) is the most thoroughly studied species. According to electron spin echo envelope modulation (ES-EEM) spectroscopy, it has a doublet ground state.[201] Other experimental observations show spin equilibrium between the doublet ground state and a sextet state[202−204] and indicate thereby that these two states are closely lying within 2 kcal mol^{-1} or less. QM-only calculations do not reproduce this trend[32] and exhibit particular sensitivity to the representation of the cysteine ligand. Thus, in some studies the ground state

(a)

45(49)[51]

2.141(2.475)[2.467]

2.269(2.491)[2.418]

r_{Fe-N} = 2.038(2.037)[2.095]

(b)

31(39)[38]
42(54)[50]

2.028(2.283)[2.257]
2.056(2.351)[2.295]

2.355(2.558)[2.501]
2.349(2.525)[2.482]

r_{Fe-N} = 2.030(2.031)[2.084]
r_{Fe-N} = 2.028(2.027)[2.080]

DFT(B3LYP)/MM; 2S+1 = 2 (4) [6]

^{2}**1** (^{4}**1**) [^{6}**1**]
E_{rel} = 0.0 (2.4) [3.3]

^{2}**1** (^{4}**1**) [^{6}**1**]
E_{rel} = 0.0 (0.4) [4.0]
E_{rel} = 0.0 (1.1) [4.6]
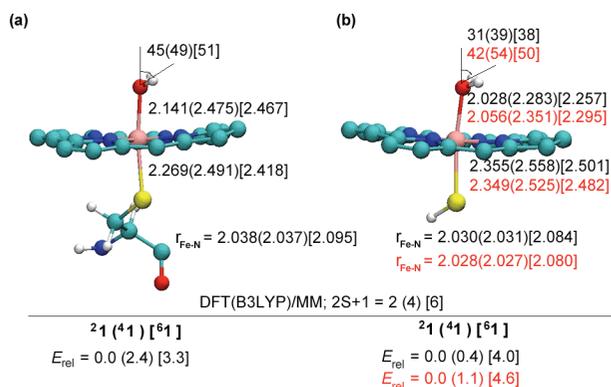
Figure 4. Bond lengths (in Å), angles (in degree), and relative energies (in kcal mol^{-1}) of the spin states of the resting state (**1**). The spin states are indicated by left-hand superscripts. r_{Fe-N} is the average distance of the four Fe–N bonds. (a) DFT(B3LYP/B2)/ MM results (Por,ext-SCys model).[156] (b) DFT(B3LYP/B1)/MM results (Por,SH model) from ref 156 (first line) and ref 54 (second line) which employ slightly different QM/MM setups; the relative energies are average values from six snapshots, respectively.

is the sextet state,[141] and in others, the Fe–O bond lengths of the quartet and sextet states are virtually dissociated.[156]

The two most recent DFT/MM studies[54,156] investigated the resting state thoroughly, in order to establish the most reliable computational levels and procedures. The earlier one[156] compared two different QM regions for the thiolate ligand (SH and ext-SCys), two different protonation states (for Asp$_{297}$), three different density functionals (BLYP, B3LYP, and BHLYP), two different basis sets, and six different snapshots as starting points for QM/MM optimizations. The relative energies of the spin states were strongly affected by the choice of the functional, consistent with previous computational experience[205,206] on other iron complexes: BLYP overestimates the stability of the low-spin states, BHLYP artificially favors high-spin states, and only B3LYP gives realistic multiplet splittings. The results were less dependent on the other choices (such as protonation state or snapshot). The more recent study[54] used DFT(B3LYP)/ MM only on the small QM model (SH) but performed the calculations of the entire cycle in a common setup that enables comparison of the various QM/MM energies.

Some representative DFT(B3LYP)/MM data for the resting state[54,156] are summarized in Figure 4. In agreement with experiment, the DFT(B3LYP)/MM calculations[54,156] predict a doublet ground state and closely lying quartet and sextet states. As can be seen from the energy differences, the three states are packed within 3–5 kcal mol^{-1} or so, depending on the proximal ligand representation. Comparison of the most recent results[54] to the previous ones[156] in Figure 4b reveals virtually identical features, so that the results are not too dependent on the system setup. A dense spin-state manifold was found also in single-point DFT/MM calculations by Scherlis et al.[207,208] on a fixed geometry of the resting state. Thus, although their doublet-sextet gap is a bit higher, the spin states are still close energetically, in line with the experimental observation of spin equilibrium.[204]

The geometric features in parts a and b of Figure 4 are also similar. In both cases, the optimized geometries show a tilt of the water ligand by 31–54°, because of hydrogen bonding interactions with the other water molecules in the pocket. The calculated tilt in the doublet ground state is in excellent agreement with ESEEM measurements[201] of the distances between the iron and the protons of the bound water ligand (computed 2.646 and 2.620 Å, ESEEM 2.62 Å).

(a)

$\sigma^*_{z^2}$

$p_z(S)$

CysS$^-$

σ_{FeS}

$d_{z^2}(Fe)$

Fe
SCys

(b)

$\sigma^*_{z^2}$

$p_z(S)$

CysS$^-$

σ_{FeS}

$d_{z^2}(Fe)$

Fe
SCys
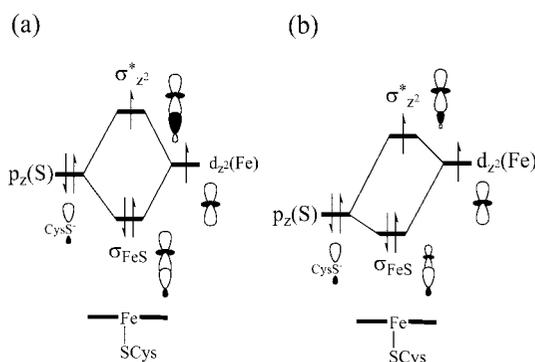
Figure 5. Orbital interactions that generate the Fe–S bond orbital (σ_{FeS}) and the $\sigma^*_{z^2}$ orbital: (a) in the gas phase where the interaction is strong, and (b) in the protein where the interaction is weakened due to stabilization of the $p_z(S)$ orbital.

Furthermore, the predicted Fe–O (2.141 Å) and Fe–S (2.269 Å) distances (Figure 4a) for the ground state are shorter and longer, respectively, than the corresponding gas-phase values and are in reasonable agreement with the X-ray data of 2.28 ± 0.2 Å and 2.25 ± 0.2 Å (pdb file: 1PHC).[209]

The impact of the protein environment is moderate on the doublet ground state but significant on the higher spin states. Thus, in the gas phase, the Fe–O bonds of both quartet and sextet states are almost dissociated (3.6–3.7 Å),[156] but in the protein, this bond is significantly shorter (2.26–2.48 Å). Furthermore, the energy gaps in Figure 4 are smaller than the corresponding gas-phase DFT(B3LYP)-only values; especially affected is the quartet state that is greatly stabilized by the protein. These two trends are connected and are both associated with the effect of the protein on the orbital mixing in Figure 5. Thus,[156] the mixing of the d_{z^2} orbital on iron and the $p_z(S)$ orbital generates a bonding and an antibonding combination; the latter is the $\sigma^*_{z^2}$ orbital of the water complex. In the gas phase (Figure 5a), the $p_z(S)$ orbital is high lying and close to $d_{z^2}(Fe)$, and hence, the mixing is strong, leading to a high-energy $\sigma^*_{z^2}$ orbital with significant S–Fe antibonding character. On the other hand, in the protein (Figure 5b), the hydrogen bonding to the sulfur (see the NH···S interaction above in Figure 2) and the bulk polarity lower the sulfur orbital and weaken its mixing with d_{z^2}, resulting in a lower lying $\sigma^*_{z^2}$ orbital, with less antibonding character. Since $\sigma^*_{z^2}$ is populated in both the quartet and sextet states, the protein stabilizes both states relative to the doublet ground state. Furthermore, the strong Fe–S interaction in the gas phase raises the $\sigma^*_{z^2}$ orbital, and this weakens significantly any potential p(O)–$d_{z^2}(Fe)$ interaction, leading to very long Fe–O bonds in the gas phase and short Fe–S bonds. By contrast, the weaker S–Fe interaction in the protein results in a long Fe–S bond and enables the formation of a short Fe–O bond.

3.3.2. The Pentacoordinated States

The conversion of **1** to **2** is the crucial event that enables the gating of the cycle via reduction of **2** to **3**. Experimental work shows that **2** has a sextet ground state, ($d_{x^2-y^2}^1 d_{xz}^1 d_{yz}^1 \sigma^*_{z^2}^1 \sigma^*_{xy}^1$; see orbitals in Figure 3c), which is in spin equilibrium with a doublet state, in a manner that reflects sensitivity to the protein environment.[204,210] The doublet state, ^{2}**2**, corresponds to $d_{x^2-y^2}^2 d_{xz}^2 \pi^*_{yz}^1$, and the intermediate spin state, ^{4}**2**, to $d_{x^2-y^2}^2 d_{xz}^1 d_{yz}^1 \sigma^*_{z^2}^1$. Much like the pentacoordinated ferric complex, the reduced ferrous complex **3**, obtained after electron uptake by ^{6}**2**, is known
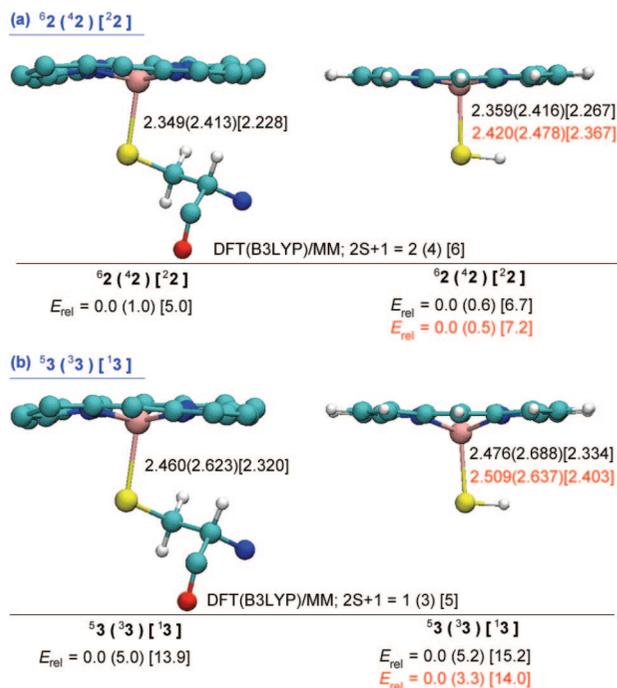
Figure 6. Optimized Fe−S bond lengths (in Å) and relative energies of the various spin states (in kcal mol^{-1}) of the pentaco-ordinated complexes **2** and **3**. The superscripts on the left-hand side refer to the spin multiplicity ($2S + 1$) of the state. (a) DFT(B3LYP)/MM results for the small and large QM subsystems for **2**. (b) DFT(B3LYP)/MM results for the small and large QM subsystems for **3**. The data in parts a and b correspond to different basis sets and setups. They are taken from ref 134 (first line, basis B2W, snapshot 31) and from ref 54 (second line, basis B1 for geometry optimization, basis B' for energies, snapX).

to have a high-spin quintet state, $^5\mathbf{3}$.[203,211] Two alternative electronic configurations are possible, $d_{x^2-y^2}^2 d_{xz}^1 d_{yz}^1 \sigma^*_{z^2}^1 \sigma^*_{xy}^1$ and $d_{x^2-y^2}^1 d_{xz}^2 d_{yz}^1 \sigma^*_{z^2}^1 \sigma^*_{xy}^1$ (see Figure 3); these may be labeled as $^5\mathbf{3}_{xz}$ and $^5\mathbf{3}_{x^2-y^2}$.[143] Closely lying states are the triplet state, $^3\mathbf{3}$, with the $d_{x^2-y^2}^2 d_{xz}^2 d_{yz}^1 \sigma^*_{z^2}^1$ configuration, and a singlet state, $^1\mathbf{3}$, with $d_{x^2-y^2}^2 d_{xz}^2 d_{yz}^2$ occupation.

Early QM(B3LYP)-only studies[134,141,143] for the low-lying spin states of **2** and **3**, with sufficiently large basis sets and different ligand representations, gave remarkably good results. High-spin ground states are found in all these gas-phase model calculations, and the ferrous complex has a $^5\mathbf{3}_{xz}$ ground state, with the $^5\mathbf{3}_{x^2-y^2}$ state lying only ca. 2 kcal mol^{-1} higher.[134,143] The various sets of results show some scatter, of course, but they agree with regard to the trends, and the computed geometries were in reasonable accord with the experimental X-ray data.[178]

The pentacoordinated complexes **2** and **3** have been studied recently by means of DFT(B3LYP)/MM in different QM/MM setups and with large basis sets.[54,134] Key geometric data and spin-state energy gaps are given in Figure 6 for the two extreme ligand representations.

The geometric parameters in Figure 6 are in basic accord with the experimental data.[178] A recent DFT(B3LYP)/MM study of **3** with and without the reductase (putidaredoxin) by Freindorf et al.[212] gave similar structural data and further showed that the Fe−S bond distance gets shorter upon binding of the reductase, to the proximal side of the active site. This bond shortening (by 0.05−0.10 Å) is accompanied by an increase of the stretching frequency of the Fe−S bond in qualitative accord with experimental findings.[213]

The data in Figure 6 show that DFT(B3LYP)/MM also reproduces the precise identity of the ground states of $^6\mathbf{2}$

and $^5\mathbf{3}$ in accord with experimental data. Thus, while the smallest $^2\mathbf{2}-^6\mathbf{2}$ energy difference of 5 kcal mol^{-1} (for the Por,ext-SCys model) is still on the high side for spin equilibrium, this value is sufficiently low to suggest this phenomenon.[214] Furthermore, according to the DFT(B3LYP/B2W)/MM results for the larger QM region,[134] the protein inverts the relative ordering of the quintet electromers of the ferrous complex $^5\mathbf{3}$: a $^5\mathbf{3}_{x^2-y^2}$ ground state is predicted in the enzyme (one electron in the $d_{x^2-y^2}$ orbital; see Figure 3), which is preferred by 0.5 kcal mol^{-1} over the $^5\mathbf{3}_{xz}$ electromer. Indeed, the comparison of calculated[134] Mössbauer parameters with the measured quadrupole splitting and asymmetry parameter[203,211,215] unequivocally assigns the ground state as $^5\mathbf{3}_{x^2-y^2}$.

The experimental values of the reduction potentials of free and putidaredoxin-bound $^6\mathbf{2}$ are −170 and −173 mV, respectively.[180,215] Using the potential of the standard hydrogen electrode (−4.43 V)[216] as reference, this implies that the free energy change ΔG for the reduction $\mathbf{2} + e^- \rightarrow \mathbf{3}$ is −98 kcal mol^{-1}. Approximating ΔG by the energy change ΔE, QM(B3LYP)-only calculations predict the conversion of $^6\mathbf{2}$ into $^5\mathbf{3}$ to be exothermic by −85.5 kcal mol^{-1} with a dielectric constant of $\varepsilon = 5.7$,[143] and by −80.2 kcal mol^{-1} with $\varepsilon = 80$.[141] DFT(B3LYP)/MM calculations[54,134] show a strong dependence on the protonation state of the protein. Thus, when the total charge of the protein is 9−, the reaction energy is endothermic by 53.8 kcal mol^{-1}, while a neutral protein leads to an overall reduction energy of −81.6 kcal mol^{-1} (QM region: Por,SH; basis B1). However, the most recent DFT(B3LYP/B')/MM calculations[54] reveal that this accuracy was obtained at the expense of unrealistically doubly protonating all the His residues; when using more realistic protonation states, of an otherwise neutral protein, the exothermicity of reduction is underestimated by 30−50 kcal mol^{-1}. Free energy calculations on reduction potentials in enzymes[217,218] have emphasized the importance of proper sampling and derived QM/MM-based reorganization energies,[218] which are substantial and will only partly be captured in an approach relying on geometry optimization.

3.3.3. Gating of the Catalytic Cycle

DFT(B3LYP/LACV3P+*) calculations[143] including bulk polarity ($\varepsilon = 5.7$) demonstrated that the reduction $^6\mathbf{2} \rightarrow ^5\mathbf{3}$ is 10.7 kcal mol^{-1} more exothermic than the reduction of the resting state ($\mathbf{1} \rightarrow \mathbf{1}^-$), compared with an experimental value of about 3 kcal mol^{-1}. This result supports the experimentally accepted gating mechanism[180,219] and can be understood based on the orbital picture of the resting state and of **2** and **3**. Thus, as shown in Figure 3b, reducing the resting state adds an electron into a π^* orbital, which exhibits antibonding interactions with the water ligand. By contrast, during the reduction of **2**, an electron is added to a low-lying orbital at iron (d_{xz} in Figure 3c). Of course, electron−electron repulsion of the excess electron with the water ligand in $\mathbf{1}^-$ must also contribute to the resistance of the resting state to accept an electron from the reductase, and as such, **1** is the off-state in the cycle.

Another conclusion from the theoretical study[143] concerns the role of the thiolate ligand in this gating mechanism. Thus, it was demonstrated that upon removal of the thiolate ligand both **1** and **2** could be easily reduced by most reductants. It is therefore the thiolate ligand that controls the gating of the catalytic cycle by making the system a poorer electron

acceptor *so that only the pentacoordinated ferric porphyrin can be reduced by the reductase*.

3.3.4. DFT/MM Study of the Electron Transfer Event

A DFT/MM study of the electron transfer mechanism in the complex of CYP101 and putidaredoxin (PDX) has recently appeared.[76] Starting from crystal structures of the two components, protein–protein docking and refinement procedures were used to generate four CYP101–PDX complexes that were then subjected to a QM/MM electron transfer path search algorithm. In this algorithm, the geometry and atomic charges of the donor (Fe_2S_2 cluster) and the acceptor (heme) are first obtained from QM/MM energy minimizations and held fixed thereafter. The QM region in the following QM/MM calculation includes the entire transfer region between donor and acceptor. A receptor in the QM region is identified by injecting an extra electron and allowing it to localize at a given residue, which is then excluded from the QM region in the subsequent QM/MM calculation. This procedure is iterated until the entire electron transfer path between PDX and the heme has been mapped.

The key findings of this study are the following: (a) The binding of the negative Fe_2S_2 cluster of PDX to CYP101 is maintained by the Arg_{112}^{CYP}–Asp_{38}^{PDX} salt bridge between the two proteins. The minimum Fe–Fe distance between the two redox centers in the different complexes varies between 12.7 Å and 14.4 Å. (b) In all four complexes, Arg_{112}^{CYP} acts as mediator of the electron transfer, either through its carbonyl group or its side chain. On the PDX side, the carbonyl group of Asp_{38}^{PDX} is involved in three cases and Gly_{41}^{PDX} in one case. (c) The so identified electron transfer pathways are *not connected to the propionate side chain* of the heme but rather pass through the Cys_{357} ligand. This unimportance of the 6-propionate side chain is in accord with experimental findings[220,221] that replacement of this side chain by a methyl group has little effect on the consumption of NADH.

3.4. The Oxy-Ferrous (4) and Ferric Peroxo (5) Complexes

The oxy-ferrous (**4**) and ferric peroxo (**5**) complexes are crucial for the propagation of the cycle. The oxy-ferrous species is remarkable, since the two constituents ^{5}**2** and $^3O_2$ have a total of six unpaired spins to begin with, and all these spins couple in **4**, which is known to be a singlet-state species.[222] Its geometry has been determined by X-ray crystallography to be an end-on complex with a bent O_2 moiety (132°) and Fe–O/Fe–S distances of 1.8/2.3 Å, respectively.[178] The ferric-peroxo complex of CYP101 has been detected by EPR/ENDOR (electron paramagnetic resonance/ electron–nuclear double resonance) studies in cryogenic experiments and is known to have a doublet spin state.[176] Its structure has not been determined yet, but its resonance Raman spectrum was recorded in the D251N mutant.[223] A structural investigation of the analogous CPO enzyme revealed spontaneous protonation to Cpd 0 (**6**).[67] However, the structure of the corresponding ferric peroxo complex of myoglobin (Mb) was recently determined in a combined X-ray and DFT(B3LYP)/MM study, and was found to possess an Fe–O bond of 1.85 Å (1.90 Å by DFT/MM).[224]

QM(DFT)-only calculations[31,141,225,226] correctly predict the spin states of the ground states for the two species and the end-on bent geometry but tend to give a somewhat longish Fe–O bond, ca. 0.15 Å longer[141] than the experimental values.[178] Adding the H-bonding protonation machinery to the Fe-OO moiety further elongates the Fe–O distances by 0.05 Å and shortens the Fe–S distances.[226] Thus, the electrostatic interactions of the dioxygen ligand with the water molecules and the peptide groups apply a *trans*-effect; the Fe–S bond gets shorter due to the weakening of the Fe–O bond. This might well be an artifact of B3LYP, which tends to undermine the Fe–O_2 bonding.[103] Indeed, a recent CASPT2 study shows that the only functional that reproduces the binding energy is OLYP.[227]

Both **4** and **5** have recently been investigated by means of DFT(B3LYP)/MM calculations, using different protonation states, different basis sets, and different representations of the proximal ligand.[161,162] The results are shown in Figure 7. Once again, the calculations correctly predict the end-on bent geometry of the Fe–OO moiety but seem to exaggerate the Fe–OO bond length, while the Fe–S bond length is in line with experimental data for **4**.[178] Furthermore, irrespective of the model used in the DFT(B3LYP)/MM calculations, the ground state of **4** is found to be an open-shell singlet state with a very low lying triplet state. An intact ferric peroxo complex **5** is only found in the doublet state, whereas the quartet and sextet states dissociate upon QM/MM optimization into 2(Fe^{II}) and a superoxide ion, $O_2^{\bullet-}$. Reduction of the singlet ground state of **4** will generate **5** in its doublet state. If intersystem crossing to higher spin states of **5** occurs, one would expect dissociation; such uncoupling is indeed observed.[175] Recent resonance Raman data for the Fe–O stretching mode indicate that the Fe–O bond of **4** is weaker than that in **5**,[223] which does not correlate with the relative bond lengths computed at the DFT(B3LYP)/MM level (see Figure 7).

The calculated reduction potential for the conversion **4** → **5** is sensitive to the chosen model, QM region, and protonation state. The DFT(B3LYP)/CHARMM values range between −2.28 and −1.04 V, relative to a standard hydrogen electrode,[161,162] and are comparable to the values obtained by QM(DFT)-only calculations.[141,143,226] While the mechanism of electron transfer was studied,[76] the barrier for this process has not yet been calculated due to the immense complexity of the problem. However, judging from the significant protein reorganization during reduction, one may conclude that the process cannot be facile,[162] thus conforming to experimental inferences that this is the rate controlling step in the CYP101 cycle.[228]

3.4.1. Bonding Features of Oxy-Ferrous (4) and Ferric Peroxo (5) Complexes

As mentioned above, the bonding in **4** and **5** is more complex than the orbital diagrams in Figure 3 suggest. Indeed, the Fe–OO bonding in type-**4** complexes has been a focus of a longstanding debate, ever since Pauling and Coryell[229] measured the magnetic susceptibility of the oxy-ferrous complex of hemoglobin and found it to be diamagnetic. Pauling suggested a model whereby $PorFe^{II}$, in a singlet state, binds $^1O_2(^1\Delta_g)$ in a coordinative dative bond, involving the lone pair on O_2 and the empty d_{z^2} orbital of the ferrous ion. On the other hand, Weiss[230] proposed another model based on spectroscopic and chemical behavior of model oxy complexes, especially their tendency to dissociate to Fe^{III} + O_2^-. In his model, Fe^{II} donates one electron to O_2 and the two radicals couple to a singlet pair. A third model that was

(a) ^{1}**4** (^{3}**4**) [^{5}**4**]

1.354(1.362)[1.338]

1.927(1.968)[2.332]

2.448(2.428)[2.640]

r_{Fe-N} = 2.038(2.036)[2.035]

R1

Thr252

1.362(1.368)[1.362]

1.942(1.989)[2.378]

2.417(2.394)[2.628]

Glu366

r_{Fe-N} = 2.037(2.036)[2.032]

R3

DFT(B3LYP)/MM; 2S+1 = 1 (3) [5]

E_{rel} = 0.0 (3.5) [5.5] E_{rel} = 0.0 (3.0) [4.6]

(b) ^{2}**5**

1.385

1.986

2.580

r_{Fe-N} = 2.044

R1

Thr252

1.394

2.030

2.570

Glu366

r_{Fe-N} = 2.046

R3
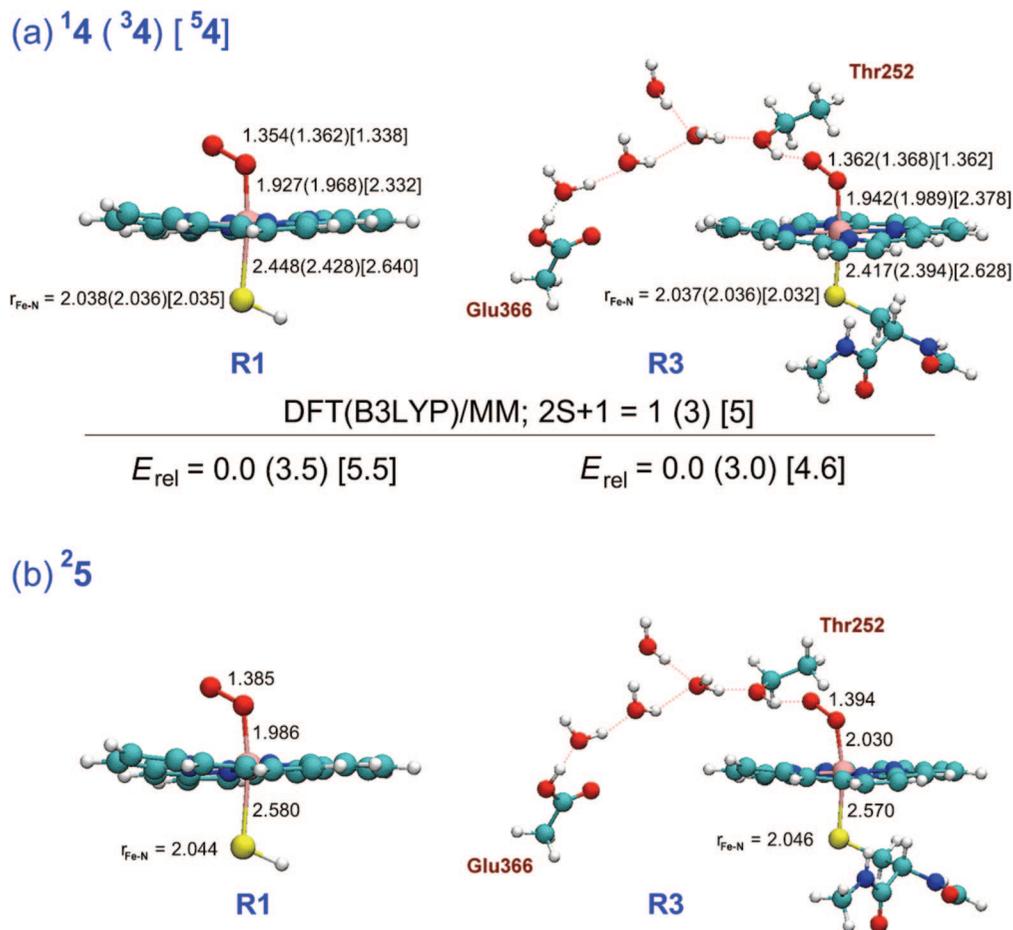
Figure 7. DFT (B3LYP/B1)/MM optimized geometries (Å) of **4** (oxy-ferrous) and ^{2}**5** (ferric peroxo) using the Por,SH (R1) and Por,ext-SCys (R3) models.[161,162] Relative energies for the spin states of **4** are given in kcal mol^{-1}.

initially suggested by McClure[231,232] involves singlet coupling of the triplet PorFe^{II} and $^3O_2$ moieties into a double Fe=O bond.

The DFT(B3LYP) and DFT(B3LYP)/MM calculations for **4** yield an open-shell singlet state configuration, with two unpaired electrons, one in a π_{yz}(FeO) orbital with a dominant iron character and the other one in a π^*_{yz}(FeO) orbital largely on the dioxygen ligand,[161,162] as shown in Figure 8a. As such, the DFT results seem to support the Weiss model for bonding: PorFe^{II} donates one electron to the π^*_{xz}(OO) orbital which becomes doubly occupied and coordinates to Fe via a dative σ-bond, maintained by overlap between the filled 2p AO on oxygen and the vacant d_{z^2} orbital on iron. At the same time, the single electron in the d_{yz} orbital of PorFe^{III} couples with the single electron in the π^*_{yz}(OO) orbital to form a singlet-pair: a weakly coupled π-type bond (described by an open-shell singlet with single electron occupancy in both a π_{yz}(FeO) and a π^*_{yz}(FeO)). The weakness of the bond is further appreciated from the fact that the corresponding triplet state is only 3.5 kcal mol^{-1} higher in energy than the open-shell singlet ground state.

Since open-shell singlet states in DFT have to be considered with some care (the open-shell singlet formulation of DFT involves one determinant and refers to an indefinite spin state), it is advisable to test the DFT results by using correlated ab initio wave function theory. This has been done recently through CASSCF/MM calculations of the oxy-Mb complex which demonstrated indeed that all the properties calculated with DFT/MM and CASSCF/MM were virtually identical and that spin projection of the DFT/MM wave
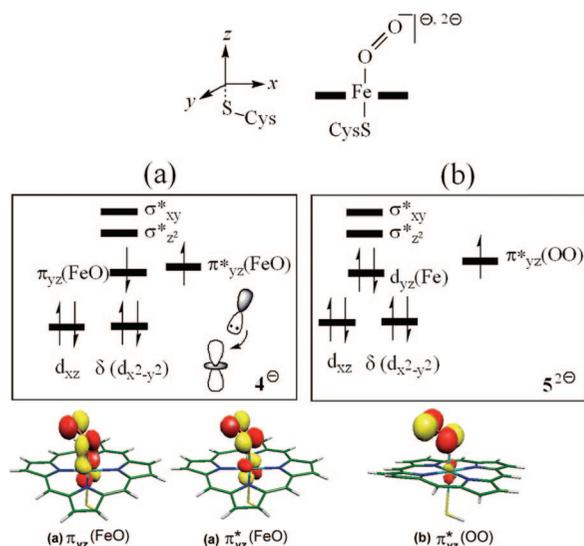
Figure 8. (a) Orbitals and their occupancies in ^{1}**4**. Shown are the singly occupied π-type orbitals and a scheme of the dative σ_{FeO} bond, made by donation from the oxygen filled p orbital to the vacant d_{z^2} iron orbital. (b) Orbitals and their occupancies in ^{2}**5** formed by reduction of **4**. Shown is the singly occupied π^*(OO) orbital.[161,162]

function gives the same singlet-to-triplet energy gap as CASSCF/MM. Furthermore, the CASSCF/MM wave function could be transformed[79] to an open-shell singlet configuration analogous to the DFT result in Figure 8a. In fact, using natural Kohn−Sham orbitals for the DFT(B3LYP)/MM species gives rise to virtually the same electronic
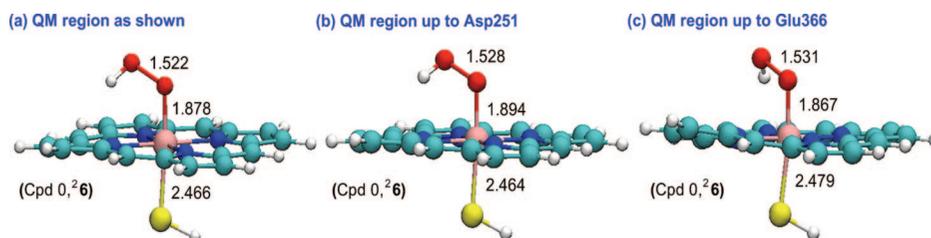
Figure 9. DFT(B3LYP/B1)/MM optimized structures (Å) of Cpd 0 in the doublet ground state, ^{2}**6**:[78,159] (a) QM region as shown; (b) QM region, which includes Thr$_{252}$, W$_{901}$, and Asp$_{251}$; (c) QM region, which includes Thr$_{252}$, W$_{523}$, W$_{566}$, W$_{687}$, W$_{902}$, and Glu$_{366}$.

population as the CASSCF/MM wave function,[79] thus further showing that DFT(B3LYP) captures the chemical essence of the oxy-ferrous complex. Hence, the best description of this species in terms of oxidation states is indeed Fe^{III}–O$_2^-$, with a σ_{FeO} bond and a weakly coupled electron pair in the π^*(OO)-d$_{yz}$ orbitals, precisely as in the Weiss model. The σ_{FeO} bond is formed by electron pair donation from the doubly occupied orbital of the coordinated oxygen atom to the vacant d$_{z^2}$ orbital on Fe, as schematized in Figure 8a. Furthermore,[79] the CASSCF/MM results were found to differ significantly from the gas-phase CASSCF results, highlighting the impact of the protein on the electronic structure. Thus, in the gas phase the oxy-ferrous complex is quite far from the Weiss model, exhibiting more covalency as in the McClure model, and it is the protein and its bulk polarity that drive the structure increasingly more toward the Fe^{III}–O$_2^-$ direction by stabilizing the negative charge on the terminal H-bonded oxygen. This sensitivity to the protein environment suggests that oxy-ferrous complexes of different heme proteins (and model systems) may have different characters of the Weiss-type bonding, thus behaving like an "electronic chameleon".

Once the orbital diagram for **4** in Figure 8a is admitted, the electronic structure of **5** is simple and corresponds to Figure 8b, with one unpaired electron in the π^*_{yz}(OO) orbital. This singly occupied orbital has an antibonding interaction with the d$_{yz}$ orbital, which should be sensitive to H-bonding and bulk polarity effects, even more so than in the oxy-ferrous case. Indeed, a very recent DFT(B3LYP)/MM investigation showed that the ferric-peroxo species of nitric oxide synthase (NOS) forms a very strong H-bond with the positively charged arginine substrate which strongly modifies its electronic structure—the singly occupied orbital is now an almost pure iron d$_{yz}$ orbital while the filled orbital made from the bonding combination of π^*_{yz}(OO) and d$_{yz}$ (labeled as d$_{yz}$(Fe) in Figure 8b) becomes a low lying orbital describing the partially formed O···H bond.[69]

3.5. Cpd 0: The Ferric-Hydroperoxide Complex (6) and its Formation from 5

Cpd 0 (**6**) has been generated as a transient by cryoradiolytic reduction of the oxy-ferrous complex of CYP101 and detected by EPR[176] and more recently also by resonance Raman spectroscopy.[223,233] The structure of Cpd 0 was determined for the analogous CPO enzyme, by means of a combined X-ray crystallography and DFT(B3LYP/B1)/MM study.[67] The corresponding Cpd 0 in a mutant (M69A) of Cyt c552 was also recently identified by EPR and electronic spectroscopy.[234] The EPR studies indicate that the ground state is a $^2\Pi_{yz}$ state, ^{2}**6**, with a single electron in the π^*_{yz} orbital (Figure 3e), and the combined experimental-theoretical study of CPO Cpd 0 also shows that the observed structural parameters fit only those of the doublet spin state.

A doublet ground state ^{2}**6** was in fact anticipated by early DFT(B3LYP)-only calculations[225] and confirmed by subsequent similar work[44,140,142,226] and an early DFT(ROB3LYP)/MM study.[157] A key feature of ^{2}**6** in all DFT-only calculations is the characteristic hydrogen bond between the proton of the distal OH group and one of the nitrogen atoms of the porphyrin ring. However, larger QM models, which include Glu$_{366}$, Asp$_{251}$, Thr$_{252}$, and crystal water molecules (Wat$_{901}$, Wat$_{902}$, Wat$_{903}$), revealed that this internal hydrogen bond can be replaced by an external one where the OH group points upward, donating a hydrogen bond to a water molecule (W$_{901}$) and accepting one from Thr$_{252}$. The same conformations were reported in the DFT(ROB3LYP)/MM study[157] and in more recent DFT(B3LYP)/MM studies.[78,159]

The DFT(B3LYP/B1)/MM geometry of ^{2}**6** for CYP101 is shown in Figure 9 for three different QM regions, i.e., the bare heme moiety (Figure 9a), the heme moiety plus the residues and crystallographic water molecules of the Asp$_{251}$ channel (Figure 9b), and the Glu$_{366}$ channel (Figure 9c). The QM/MM calculations of these three QM regions yield quite similar distances, Fe–O (1.878; 1.894; 1.867 Å), O–O (1.522; 1.528; 1.531 Å), and Fe–S (2.466; 2.464; 2.479 Å), which are in good accord with the experimentally determined parameters for CPO Cpd 0 (1.82, 1.50, and 2.4 Å, respectively).[67] Interestingly, the computed Fe–O bond length in Cpd 0 is shorter than that in the corresponding oxy-ferrous complex (Figure 7), and here this is in accord with the conclusion from resonance Raman spectroscopy that the Fe–O bond is stronger in Cpd 0.[223,233] The three structures in Figure 9 show internal H-bonding between the OH group of the Fe-OOH moiety and one nitrogen atom of the porphyrin ring. However, according to a recent DFT(B3LYP)/MM study of CPO Cpd 0,[66] the conformation wherein the OH group maintains a strong H-bond with Glu$_{183}$ is more stable in CPO, and the barrier between the two conformations is rather small (<7 kcal mol^{-1} with the larger basis set).

Figure 9 does not show the geometries of higher spin states. As already reviewed,[32] the quartet and sextet states are significantly higher in energy than the doublet state. The most recent extensive MR-CI/MM calculations[78] of Cpd 0 for CYP101 give a doublet-quartet separation of 10.6 kcal mol^{-1}, in close agreement with the DFT(B3LYP)/MM results[159] and with earlier studies.[142,235]

3.6. The Protonation Mechanism and Generation of the Ultimate Oxidant, Cpd I (7)

Formation of Cpd I requires two protonations (**5** → **6** → **7**). This involves translocation of the proton from the bulk solvent into the active site and reorientation of the hydrogen-bond network to facilitate the migration of the proton. It is thus important to identify the aqueducts in the enzyme, along with the titratable amino acids that can adopt different protonation states, and thereby facilitate the proton transfer.
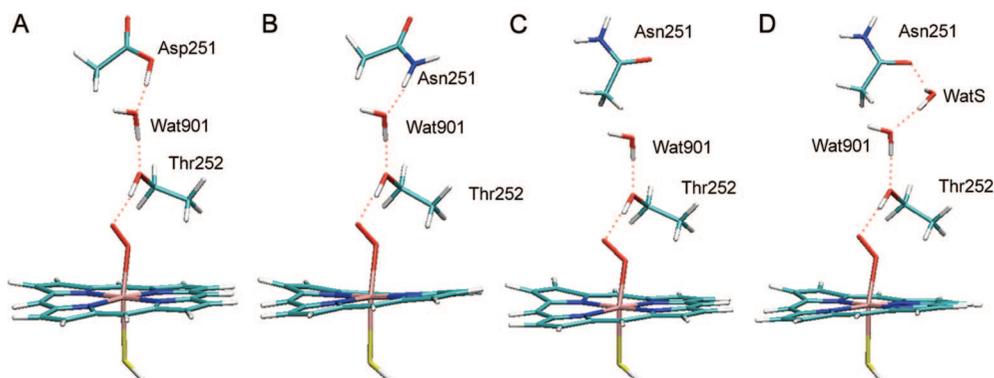
Figure 10. QM-models A–D for DFT/MM studies of the protonation of ^{2}**5** to ^{2}**6**, in the wild type (A) and the D251N mutant of CYP101 with the different Asn$_{251}$ conformations (B–D). WatS is an added water molecule. Reprinted with permission from ref 161. Copyright 2008 American Chemical Society.

As already discussed, the protonation machinery of P450 enzymes is generally thought to involve an acid–alcohol pair and a water network, which shuttles the proton in a Grotthuss type mechanism.[236] In the case of CYP101, possible candidates are Asp$_{251}$-Thr$_{252}$ and/or Glu$_{366}$-Thr$_{252}$, along with active-site water molecules (Wat$_{523}$, Wat$_{566}$, Wat$_{687}$, Wat$_{901}$, and Wat$_{902}$).[175,182,219,237] An alternative path involves Arg$_{299}$, which participates in a salt bridge with the propionate-A side chain (7-propionate) of the heme; it has been postulated to function as a gate that opens wide by a conformational change and thereby allows water molecules in.[238] These many potential channels make the problem very complex. Adding to this complexity is the coupling between reduction and protonation, as indicated by significant kinetic solvent isotope effects (KSIE) that are observed already for the reduction of ^{1}**4**.[182,239–241] The D251N mutant exhibits a very large KSIE of 10,[182] which implies Asp$_{251}$ as a key acid in the protonation process. Indeed, Asp$_{251}$ has access to the bulk water, and it can control the proton shuttle through a "carboxylate switch": deprotonated Asp$_{251}$ blocks the water entry, since it is engaged in salt bridges (Arg$_{186}$; Lys$_{178}$), but upon protonation, it may flip to the inside of the active site and open the channel to water molecules.[178,182] The relevance of Glu$_{366}$ does not seem as compelling, considering the fact that mutation experiments[242] showed little change of activity when replacing Glu$_{366}$ by Met. Moreover, the Glu$_{366}$ residue does not have direct connection with the bulk solvent, such that there is no obvious mechanism of reprotonating this residue after transferring its proton to the heme-oxy moiety.

3.6.1. First Protonation Step: Conversion of (5) into Cpd 0 (6)

Early DFT-only studies of the protonation of ^{2}**5** to form ^{2}**6**[45,235,243,244] provided insight into the mechanism but also showed that protonation cannot be studied with simple model systems; different models yield disparate results for the exothermicity/endothermicity of the process and its barriers. Thus, with a calculated proton affinity of 422 kcal mol^{-1}, ^{2}**5** would even accept a proton from a nearby water molecule,[235] but the energy of the process highly depends on the number of water molecules, as well as a correct choice of the residues near the water molecules.[243,244] A DFT(ROB3LYP)/MM study[157] considered only the process originating from Glu$_{366}$ and reported a small barrier of 4.0 kcal mol^{-1}.

As outlined above, however, the Asp$_{251}$ channel provides the most likely pathway for proton delivery. Therefore, a recent DFT(B3LYP)/MM study focused on this channel in the wild type (WT) and mutant CYP101 enzymes.[161] The

four models A–D that were investigated are shown in Figure 10. According to MD calculations in an earlier study,[159] Asp$_{251}$ can rotate rather easily out of its salt bridge with Arg$_{186}$ (at an energetic cost of less than 4 kcal mol^{-1} for twists up to 40°), and it should thus be dynamically accessible to protonate it from bulk solvent. The wild-type enzyme model A thus employed protonated Asp$_{251}$ as proton source. Three models B–D were used for the D251N mutant: B was generated from A by manually replacing Asp by Asn, C was taken from the MD trajectory of B showing a flip of the Asn side chain within 45 ps to the outside of the pocket (due to repulsion with Arg$_{186}$), while D was generated from C by adding a water molecule, WatS, to the space generated by the flipped Asn.

The mechanism for the WT enzyme is depicted in Figure 11, along with relative energies of the "reactant complex" (**5**$_{RC}$), transition state (**5**$_{TS}$), and "product complex" (**6**$_{PC}$). The barrier for protonation is very small and hardly sensitive to the basis set used, being 2.5 kcal mol^{-1} with the largest basis set (def2-TZVP, 6-311++G(2d,2p)) and 1.2 kcal mol^{-1} with the lowest basis set (B1). The small barrier agrees well with the short lifetime of this species and with the fact that the reduction of the oxy-ferrous complex **4** exhibits KSIE, which means that the formation of **5** and its protonation are almost simultaneous.[239] The generation of Cpd 0, **6**, is seen to be exothermic by 33.6 kcal mol^{-1}, quite close to the result of a large-model DFT-only calculation reviewed before.[32] Indeed, the transition state, **5**$_{TS}$, which involves a synchronous proton shuttle from Asp$_{251}$ all the way to the Fe-OO moiety, is "early", in accord with the highly exothermic reaction and low barrier. Interestingly, **5** has been implicated as an additional oxidant in deformylation reactions (see review of mechanistic aspects[22]). However, with a barrier of only 2.5 kcal mol^{-1} for collapse to Cpd 0, it is unlikely that **5** can be responsible for deformylation in the WT enzyme, unless this barrier is significantly larger, due to conformational changes in the protein.

For the D251N mutant, model B yields a very low barrier (2.2 kcal mol^{-1} with the largest basis set) for proton transfer from the NH$_2$ side chain of Asn$_{252}$ to the FeOO moiety via Wat$_{901}$ and Thr$_{252}$ (see Figure 10), which was attributed to the acidifying effect of the Arg$_{186}$ residue.[161] However, model B is hardly realistic, since classical MD simulations give a very low barrier (<2 kcal mol^{-1}) for the flip of the Asn$_{251}$ side chain that leads to model C. The results for model C are given in Figure 12. Since the flipped Asn$_{251}$ has no good way to deliver a proton, the transfer occurs from Thr$_{252}$, which becomes an anion: the corresponding barrier is
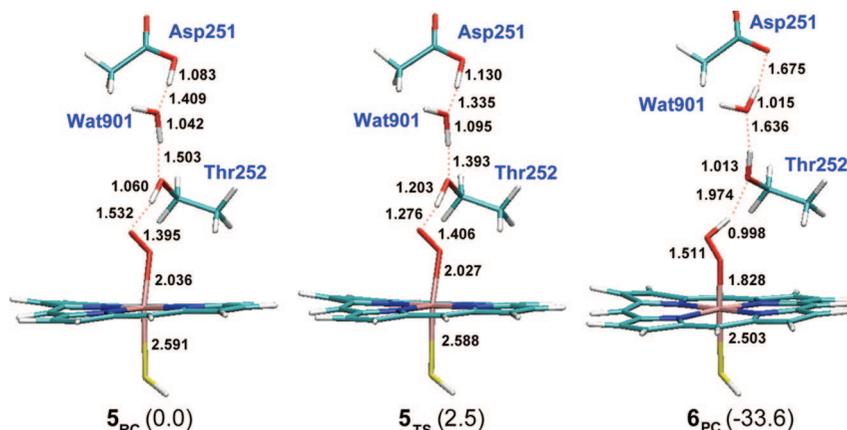
Figure 11. DFT(B3LYP/B1)/MM optimized geometries (Å) and relative energies (kcal mol^{-1}) of 5_{RC}, 5_{TS}, and 6_{PC}, for the proton transfer that leads to the formation of Cpd 0 (**6**) in the WT enzyme (model A in Figure 10).[161] Relative energies refer to the largest basis, def2-TZVP(Fe)/6-311++G(2d,2p) (rest).
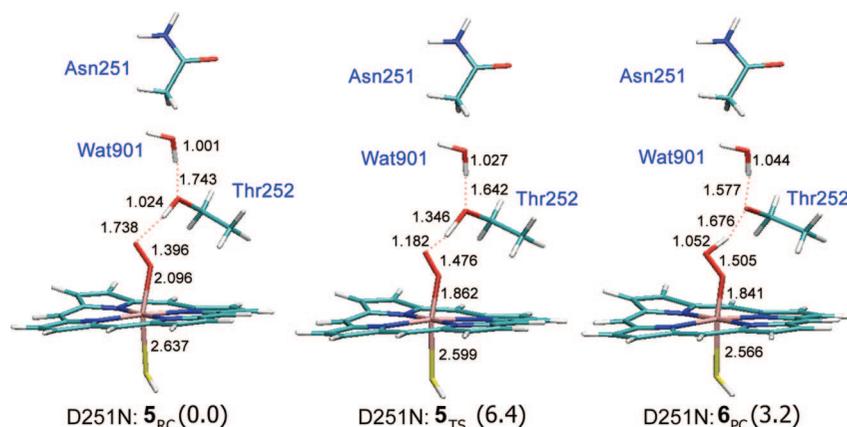
Figure 12. DFT(B3LYP/B1)/MM optimized geometries (Å) and relative energies (kcal mol^{-1}) of 5_{RC}, 5_{TS}, and 6_{PC}, for the proton transfer that leads to the formation of Cpd 0 (**6**) in the D251N mutant (model C in Figure 10).[161] Relative energies refer to the largest basis, def2-TZVP(Fe)/6-311++G(2d,2p) (rest).

significant (6.4 kcal mol^{-1}), the reaction is endothermic, and the transition state is advanced. No facile path for reprotonation of the Thr$_{252}$ anion was found, however, and it may thus well be that an alternative route is actually preferred for model C, namely back-flip to B and subsequent proton transfer in B, which requires an overall activation of ca. 10 kcal mol^{-1}.[161]

A variant of C is D, where an additional water molecule (WatS, Figure 10) can enter the pocket from the bulk and mediate the proton transfer. This process is mildly exothermic (-6.3 kcal mol^{-1} at the highest level), but the barrier remains significant, 7.5 kcal mol^{-1}.[161] It is conceivable that the resulting system with an OH$^-$ anion (from WatS, stabilized by H-bonds) can be reprotonated from bulk solvent without undue activation (not studied). In summary, the DFT(B3LYP)/MM study of several realistic models of the D251N mutant indicates sizable overall barriers for the proton transfer **5** → **6**. Hence, the lifetime of **5** will be significant in the D251N mutant, and the KSIE might be large, especially for model D, where the generated hydroxide anion of WatS will have to be reprotonated from bulk water.

3.6.2. Second Protonation Step: Conversion of Cpd 0 (6) into Cpd I (7)

EPR/ENDOR studies of the T252A mutant of P450$_{cam}$ by Davydov et al.[176] showed that Cpd 0 is still formed but no camphor hydroxylation occurs, thus suggesting that this mutation prevents the protonation of Cpd 0 and the formation

Scheme 3. Schematic Consensus Mechanism of the Conversion of Cpd 0 (6) to Cpd I (7)
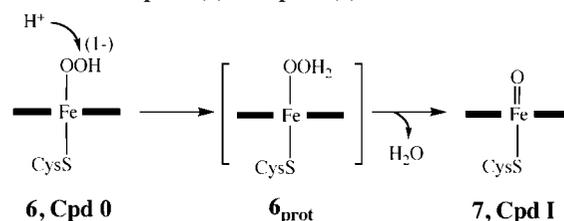
of the ultimate oxidant, presumably Cpd I. The Cpd I formation in P450 enzymes is generally assumed to follow the Poulos–Kraut mechanism that is well established in peroxidases,[245] where the proton is added to the distal oxygen of Cpd 0 followed by water splitting (see Scheme 3). The species labeled as 6_{prot} is not necessarily considered to be stable but perhaps a transition state en route to Cpd I (**7**).

Early attempts to study this second protonation process by means of DFT-only calculations of the isolated molecule suggested that Cpd 0 (**6**) is a very strong base with a proton affinity of 334 kcal mol^{-1},[235] or 330.1 kcal mol^{-1},[142,143] and that it is thus expected to accept a proton and collapse to Cpd I without a barrier.[45,144] A somewhat larger model[45] also gave a highly exothermic reaction and no definite barrier. An early DFT(ROB3LYP)/MM study[157] indicated a highly exothermic process as well, but no precise data appear in the paper to assess the cause of the exothemicity. By contrast, a more recent study of a large model of 96 atoms, which
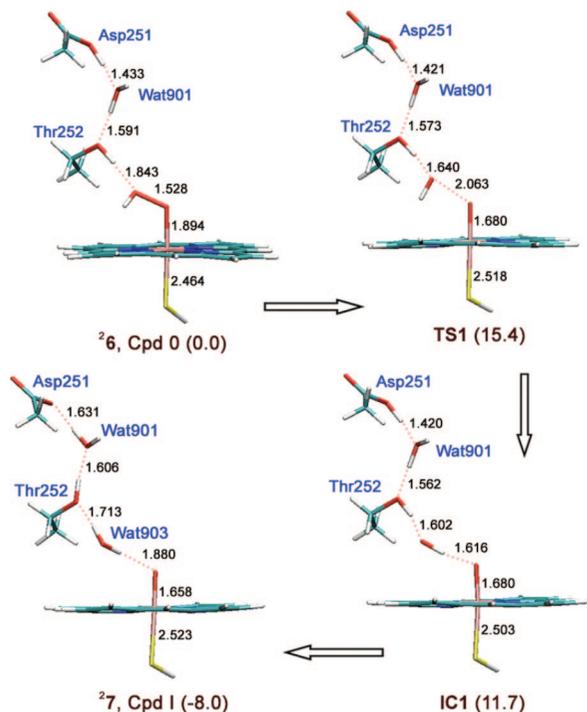
Figure 13. DFT(B3LYP/B1)/MM optimized geometries (Å) of Cpd 0 (^{2}**6**), TS1, IC1, and Cpd I (^{2}**7**). Relative energies (kcal mol^{-1}) refer to the largest basis set employed: LACV3P++**(Fe)/6-311++G** (rest). Reprinted with permission from ref 159. Copyright 2006 American Chemical Society.

included Asp$_{251}$, Thr$_{252}$, Glu$_{366}$, and the water networks connecting the two acids to the FeOOH moiety,[226] suggested for the first time that the process is almost thermoneutral and that **6**$_{prot}$ may be a true intermediate. A significant barrier was found for the formation of **6**$_{prot}$ as well as for its subsequent water splitting step.

This spectrum of results was clearly confusing and motivated a recent DFT(B3LYP)/MM study of CYP101[159] that addressed both the Glu$_{366}$-Thr$_{252}$ and Asp$_{251}$-Thr$_{252}$ channels. The Glu$_{366}$-Thr$_{252}$ channel indeed gave the mechanism shown in Scheme 3 and located in the preceding model study.[226] However, the **6**$_{prot}$ intermediate was found to be unstable at the DFT(B3LYP)/MM level, lying >20 kcal mol^{-1} above Cpd 0 and having a very small barrier (3−4 kcal mol^{-1}) for collapse to Cpd I. The Asp$_{251}$-Thr$_{252}$ channel did not yield this mechanism at all. In any event, with such energetics, this mechanism does not seem to be the likely candidate for the formation of Cpd I.

The second mechanism considered originates from the Asp$_{251}$-Thr$_{252}$ channel, and its key features are shown in Figure 13. It is not driven by initial protonation but rather starts with homolysis of the O−OH bond. As OH departs (see IC1), two events happen simultaneously: (i) an electron from the porphyrin a$_{2u}$ orbital is shifted to the departing OH radical, converting it to OH$^-$, and (ii) simultaneously Thr$_{252}$ protonates the hydroxide anion and gets itself protonated by Wat$_{901}$. Thus, the mechanism involves proton-coupled electron transfer (PCET), which is made possible by the facile proton delivery along the hydrogen bonding network in the Asp$_{251}$-Thr$_{252}$ channel and the "push effect" of the thiolate that promotes the electron shift. The barrier for the O−OH cleavage was calculated to be 15.4 kcal mol^{-1} (at the highest level). The intermediate IC1 is a shallow minimum with standard QM/MM options (using the QM region shown in Figure 13). Inclusion of Arg$_{186}$ into the QM region facilitates

the protonation by acidifying Wat$_{901}$, and as a result, the IC1 minimum disappears in the largest QM/MM model considered. The analogous mechanism in the Glu$_{366}$-Thr$_{252}$ channel gives a similar initial barrier but proceeds in a stepwise manner. It was not pursued further because Glu$_{366}$ is not connected to the bulk water and can thus not easily be reprotonated (see above) which would be necessary in a complete catalytic cycle. In both mechanisms, the formation of Cpd I from Cpd 0 is either almost thermoneutral or slightly exothermic, with a "best" estimated reaction energy of about −8 kcal mol^{-1}. At this stage we cannot say whether the small exothermicity is a general feature or one that is specific to P450$_{cam}$. Nevertheless, the PCET mechanism seems to carry over to other enzymes. Thus, in subsequent QM/MM work, the synchronous PCET mechanism for Cpd I formation has also been found in CPO[66] and in NOS.[69] In all these cases, PCET proceeds by a hybrid homolytic−heterolytic mechanism, and this hybrid nature lowers the O−O cleavage barrier well below that of the pure homolytic mechanism.[66]

What remains to be studied for the wild-type P450$_{cam}$ enzyme is the process for generating the higher spin states of Cpd I, ^{4,6}**7**. Corresponding DFT(B3LYP)/MM calculations for CPO gave large barriers of the order of 20 kcal · mol^{-1}.[66] If this is true also for P450, then we might conclude that the ^{4,6}**7** species are not formed directly but must be generated by intersystem crossing from ^{2}**7**.

3.6.3. Conversion of Cpd 0 (**6**) into Cpd I (**7**) in the T252X Mutants

The mutation of Thr$_{252}$ is known to have an adverse effect on the activity of CYP101.[175] Replacement of Thr$_{252}$ by a small aliphatic residue in mutants such as T252A and T252G leads to an uncoupling of O$_2$ consumption from camphor hydroxylation, since most of the consumed O$_2$ is converted to H$_2$O$_2$ (95% or more). Both hydroxylation and uncoupling are observed in the T252V mutant, to a similar extent, whereas hydroxylation is the dominant channel in the T252S mutant (85%) and the only channel in the wild-type enzyme (100%). The latter trend remains true when the Thr$_{252}$ side-chain OH group is replaced by a methoxyl group (which, however, slows down the reaction significantly). The double mutant D251N/T252A as well as the single mutant D251N also perform hydroxylation, albeit at a reduced rate compared with the wild-type enzyme. These experimental findings clearly indicate that the Thr$_{252}$ residue plays a crucial mechanistic role and influences the ratio of hydroxylation vs uncoupling products. It is also known that some T252X mutants retain partial oxygenation activity. For example, the T252A mutant is still capable of epoxidizing the double bond of 5-methylenylcamphor,[183] which has been interpreted to arise from reactivity of Cpd 0, acting as an oxidant toward the more reactive substrates if Cpd I formation is suppressed (recall the more favorable uncoupling in the T252A mutant, see above). Since the crystal structure of the T252A mutant contains a water molecule near the O$_2$ binding site that is not present in the wild-type enzyme,[184,185] it has been suggested that solvent participation affects the reactivity, in particular the competition between Cpd I formation and uncoupling.

To address these issues, the T252X mutants with X = serine, valine, alanine, and glycine were studied by DFT(B3LYP)/MM calculations.[246] Starting geometries for Cpd 0 were taken from a crystal structure of the wild-type enzyme (pdb: 1DZ8) that had been used in previous QM/
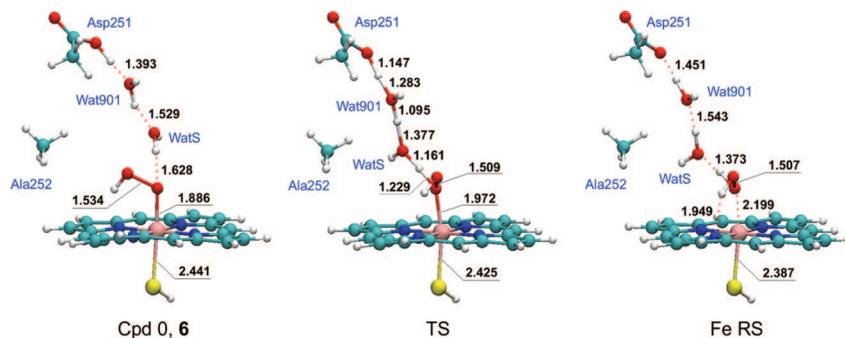
Figure 14. DFT(B3LYP/B1)/MM optimized geometries (Å) for the species in the uncoupling reaction of the T252A mutant in the presence of an extra water molecule, WatS. Reprinted with permission from ref 246. Copyright 2009 American Chemical Society.

MM work, and mutant structures were generated by manually replacing the threonine residue by the desired amino acid X and subsequent relaxation and equilibration. The crystal water molecule Wat_{901} that is located in the X-ray structure, of the wild type (WT) enzyme, between Thr_{252} and Asp_{251} remains in this position during classical 2 ns MD simulations, both in the WT enzyme and in each of the four mutants. By contrast, an additional water molecule, WatS, placed manually into the vicinity of residue 252 quickly escapes from this region during classical MD in the case of the WT enzyme and the T252S mutant, whereas it remains put during the full 2 ns simulation in the case of the T252V, T252A, and T252G mutants. The corresponding substitutions obviously generate enough "empty space" to accommodate an extra water molecule so that the latter three mutants should be modeled with inclusion of WatS.

The conversion of Cpd 0 to Cpd I is computed to proceed by essentially the same mechanism in the WT enzyme and in the four mutants. As discussed above (see section 3.6.2), the first and rate-limiting step involves a homolytic–heterolytic cleavage of the O–O bond followed by a proton-coupled electron transfer. The computed barrier is in the range of 14–17 kcal mol[-1] for all five systems and is not affected much by the presence or absence of WatS in the T252V, T252A, and T252G mutants. The uncoupling reaction (formation of hydrogen peroxide) is found to be a concerted reaction in the WT enzyme, with a rather high barrier of 27 kcal mol[−1], which reflects the large distortions that are necessary to establish a proton delivery path from Asp_{251} via Wat_{901} and Thr_{252} to the proximal (rather than the closer distal) oxygen atom of the FeOOH moiety. The alternative two-step mechanism (Fe–O cleavage first followed by proton transfer) is even less favorable because of the strength of the Fe–O bond. In the mutants, in the absence of WatS, the uncoupling reactions remain concerted and difficult to achieve, with barriers between 23 and 29 kcal mol[−1]. However, in the presence of WatS, a much more favorable hydrogen bond network is established in the Asp_{251} channel for the T252V, T252A, and T252G mutants, with uncoupling barriers of 19, 12, and 12 kcal mol[-1], respectively. The corresponding optimized geometries and reaction profiles are illustrated for the T252A mutant in Figures 14 and 15, respectively.

In summary, the DFT(B3LYP)/MM calculations for the T252X mutants agree with the experimental findings and offer a consistent mechanistic scenario. In the WT enzyme, Cpd I formation is favored, since its rate-limiting barrier is 13 kcal mol[-1] lower than that for uncoupling. The difference is reduced to 7 kcal mol[-1] in the T252S mutant. In the case of the other three mutants, an extra water molecule enters
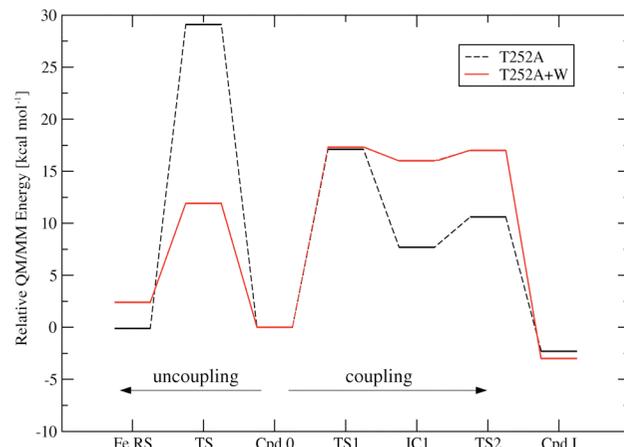
Figure 15. DFT(B3LYP/TZVP)/MM energy profile for coupling and uncoupling reactions with and without an extra water molecule in the T252A mutant. Reprinted with permission from ref 246. Copyright 2009 American Chemical Society.

the active site and lowers the activation energy for uncoupling significantly. With the additional water molecule, Cpd I formation and uncoupling have similar barriers in the T252V mutant, and uncoupling becomes favored in the T252A and T252G mutants. These DFT(B3LYP)/MM thus confirm the hypothesis that proton delivery by solvent water is responsible for the uncoupling reaction in some mutants.

3.6.4. Conversion of Cpd 0 (6) into Cpd I (7) in the D251N Mutant

As mentioned above, it is known experimentally[182] that the D251N mutant does not undergo uncoupling but performs hydroxylation of camphor, albeit at a reduced rate (lower by a factor of 30 compared with wild-type P450$_{cam}$). This suggests that the mutant forms Cpd I, but more slowly than the wild-type enzyme. DFT(B3LYP)/MM calculations[247] on the D251N mutant confirm that the barriers for uncoupling are indeed always significantly higher than those for coupling, in agreement with experiment. As in the case of the first proton transfer that leads to Cpd 0 (see section 3.6.1), three models of the D251N active site were used to study the second proton transfer that converts Cpd 0 into Cpd I: the mutated Asn$_{251}$ residue can be directed toward the heme (model I, no flip) or away from the heme (model II, flip), and in the latter case, the reorientation of the Asn$_{251}$ side chain generates enough empty space to hold an extra water molecule (model III, flip with WatS). Classical MD simulations show that model I rearranges spontaneously to model II within less than 400 ps and that the extra water molecule of model III remains stable in the active site during 2 ns of

MD simulation. For all three models, Cpd I formation in the D251N mutant follows the same basic mechanism as in the wild-type case (see section 3.6.2), with initial O−O cleavage followed by proton transfer. The barrier for the initial step is similar in all D251N models (around 14 kcal/mol), but the proton transfer is most facile in model III (not feasible in model II, significantly higher barriers in model I). The proton transfer in model III generates a hydroxide anion (at Wat$_{901}$) that is not reprotonated easily by neighboring residues (Arg$_{186}$); however, proton delivery from bulk solvent seems possible by a Grotthuss mechanism via a water network that remains intact during 2 ns of classical MD simulation. In summary, the DFT(B3LYP)/MM calculations show that Cpd I formation in the D251N mutant is feasible only when extra water molecules enter the active site (model III). They reproduce the experimentally observed preference of the mutant for coupling, and the lower hydroxylation activity can be tentatively ascribed to a less facile proton transfer (since the initial O−O cleavage has similar barriers in the wild-type enzyme and in the D251N mutant).

3.7. Cpd I (7): The Ultimate Active Species

As mentioned above, Cpd I is elusive in the native working cycle. X-ray structural evidence for the existence of Cpd I[178] seemed initially very convincing but has subsequently been cast in doubt,[176] although not definitely ruled out. The presence of Cpd I as a transient species was inferred by cryogenic and electronic absorption studies[184,248−250] and from cryogenic EPR/ENDOR spectroscopy of the alcohol product of camphor hydroxylation starting from deuterated camphor at the C^5−H position.[176] Attempts to generate Cpd I using oxygen surrogates such as iodosylbenzene or peracids led usually to the one-electron reduced species, so-called Compound II (Cpd II), and a Tyr radical.[251−254] More recent experiments[250,255] showed how one can optimize the formation of Cpd I, before the heme bleaches due to destruction by the peracid, and observe the characteristic chemistry of Cpd I. Strong but indirect support for the participation of Cpd I was provided by the analogous reactivity of synthetic Cpd I species and P450 toward the same substrates (similar product distribution and stereochemical scrambling).[22] A recent study by Dowers et al.[256] used a surrogate oxygen atom donor to generate the N-oxide derivative of *N,N*-dimethyl amines and demonstrated that this leads to hydroxylation of the methyl group of the *N,N*-dimethyl amines, thereby strongly implying that the only competent oxidant in the reaction was Cpd I. Nevertheless, the elusiveness of Cpd I, in the working cycle, prompted suggestions for alternative oxidant species, such as Cpd 0,[23,183,257] Cpd II,[258−260] and recently also the perferryl PorFe(V)=O species.[261,262] However, recent experiments, which directly compared the reactivity of Cpd 0 and Cpd I in enzymes and in model systems showed Cpd 0 to be far less reactive than Cpd I, so that in all likelihood the latter is the only oxidant in a working P450 enzyme.[234,263−265] According to computational work (described in more detail below), the same applies to the reactivity of Cpd II,[164] while the perferryl proposal seems at present untenable because of the high energy of this species.[164,266] Furthermore, in a more recent study of Newcomb et al.,[262] Cpd I of the thermophilic enzyme CYP119 was generated by laser flash photolysis (LFP) from Cpd II and was allowed to react with a variety of substrates, thus providing rate constants for substrate oxidations by a P450 Cpd I. The so obtained rate constants were generally

2−3 orders of magnitude larger than the rate constants for the reactions of model Cpd I species with the same set of substrates.[267] While the reaction rate for CYP119 Cpd I, generated in the native working cycle, is still not known, these results already show that a P450 Cpd I generated by LFP is a competent oxidant. As such, the weight of evidence seems to be in favor of Cpd I being the primary if not the sole oxidant species in the P450 cycle.

Because of the uncertainty regarding Cpd I, and because of its status as the primary oxidant of P450, theoretical studies have tended to focus primarily on this species, using QM-only models[28,46−50,141,197,198,266,268−271] and DFT/MM studies of many P450 isoforms[53,55,56,77,78,157,160] and other thiolate enzymes such as CPO[66] and NOS,[68,69] as well as other heme enzymes.[74,272,273] All these DFT-based studies as well as the available correlated ab initio studies[77,78,274] agree on one feature, namely, that the species has two low-lying spin states, doublet and quartet, which are almost degenerate, as already discussed by reference to Figure 3a.

Early DFT-only studies of Cpd I models in vacuum led to controversial results about the electronic ground state.[28,46−50,53,141,197,198,266,268−271] In particular, the results were very sensitive to the choice of the axial ligand. Still, all calculations agreed that Cpd I possesses three singly occupied orbitals, two of which are the π^*_{xz} and π^*_{yz} orbitals (Figure 3a). However, the nature of the third singly occupied orbital appeared to depend on the chosen thiolate model and on the interactions of the sulfur atom with the environment.[28,46−50,53,141,197,198,266,268−271] In cases where the cysteinate was modeled by SMe[46,141,271] or SCys without internal hydrogen bonding,[49] the third unpaired electron resides in a lone-pair orbital on the sulfur, leading to a very large spin density located on the sulfur ligand and hardly any on the porphyrin ring. By contrast, the choice of SH$^-$ or cysteinate with internal hydrogen bonding leads to $^{4,2}A_{2u}$ electronic states where the a$_{2u}$ porphyrin orbital is singly occupied ($\pi^*_{xz}{}^1\pi^*_{yz}{}^1a_{2u}{}^1$), with spin densities distributed evenly between porphyrin and sulfur.[29,50,198] Furthermore, the Fe−S distance was found to depend on the nature of the ligand model used for cysteine and to vary as a result of environmental perturbations, which were modeled by NH···S hydrogen bonding and a dielectric constant to mimic the effect of the electric field of the protein.[50,197,266] The environmental effects on the geometry and spin densities of Cpd I are depicted in Figure 16, which shows results of Ogliaro et al.[50] and Shaik et al.[26] for the system (Por,SH) and those of Rydberg et al.[141] for the system (Por,SMe).

It is seen from Figure 16 that, generally, the FeO moiety bears a spin density (ρ) of around 2, indicative of singly occupied π^*_{xz} and π^*_{yz} orbitals (see Figure 3a). The third spin is distributed on the sulfur and porphyrin groups, in a manner sensitive to the environment and ligand representation. Thus, in the gas phase and without any external interactions, the species has a predominant thiolate radical character; for example, for $^4A_{2u}$ 54% (Figure 16a) or 72% (Figure 16b) of the spin density is on the thiolate ligand. External interactions with a dielectric medium ($\varepsilon = 5.7$) or NH···S hydrogen bonding (Figure 16a) shift spin density from sulfur to porphyrin; in the case of the Por,SH model, the species acquires a predominant porphyrin radical cationic character (74−89% in Figure 16a).[26,50] Further inspection of Figure 16 reveals that, while the Fe=O and Fe−N bonds are not sensitive to the environment, the Fe−S distance is significantly affected by the external perturbation. Thus,
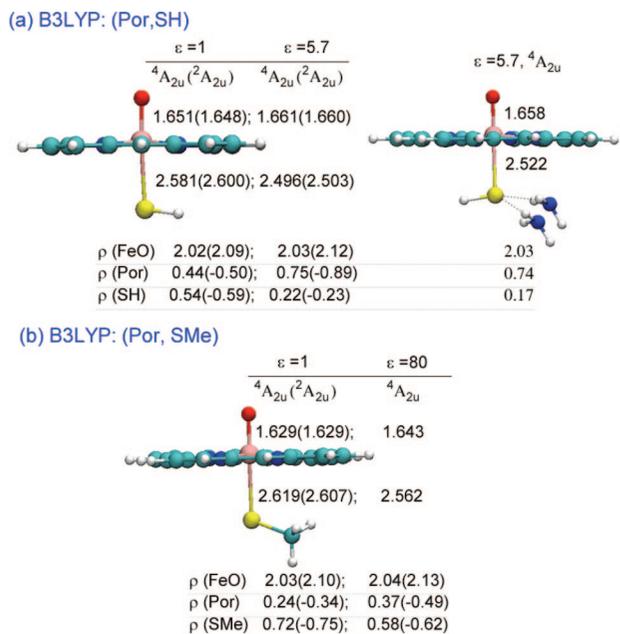
Figure 16. Environmental effects on geometrical parameters (Å) and group spin densities ρ of Cpd I: (a) Model calculations[26,50] with different dielectric constants ε and with hydrogen bonding (NH···S); (b) model calculations[141] with different dielectric constants ε.

relative to the gas phase, the Fe−S distance is shortened, in some cases by almost 0.1 Å.[50,197] In view of this sensitivity of the electronic structure and Fe−S bond length of Cpd I, Ogliaro et al.[197] characterized Cpd I *as a chameleon species* that can change its nature depending on external conditions.[50,197] Later it was shown that an external electric field (EEF) oriented along the S−Fe−O axis can induce these changes or create an exclusive sulfur radical by simply switching the direction of the EEF, thereby further supporting the chameleon nature of the species.[275]

These geometric and electronic changes exerted by the environment, noted initially in DFT model studies that accounted for the environment,[50,197] were confirmed by the DFT(B3LYP)/MM study of Cpd I in the native protein.[53] The first DFT(B3LYP)/MM calculation on Cpd I of CYP101 was published in 2002.[53] It employed three QM regions with different cysteine parts (labeled in Scheme 2 as SH, SMe, and ext-SCys), a few snapshots from equilibrated MD simulations, and three different basis sets. Representative DFT(B3LYP)/MM results obtained for the 29 ps snapshot and the largest basis set are shown in Figure 17a−c and compared with the corresponding gas-phase data. It is apparent that for all QM regions the protein shortens the Fe−S bond and converts the species from a predominantly sulfur radical to a porphyrin cation radical. Thus, irrespective of the ligand model, all the species in the protein environment look very similar to one another. This picture did not change with subsequent studies of CYP101, which all gave Fe=O/Fe−S distances of approximately 1.64/2.55 Å and an electronic structure that is typified by a porphyrin cation radical with some thiolate spin density ≤ 0.23.[56,77,78,160] As shown in Figure 17d, these results are in general agreement with recent EXAFS (extended X-ray absorption fine structure) structural data[276] and ENDOR spin density data[75] for CPO Cpd I.

Another impact of the protein environment can be seen from Table 1, which summarizes Mössbauer spectroscopic
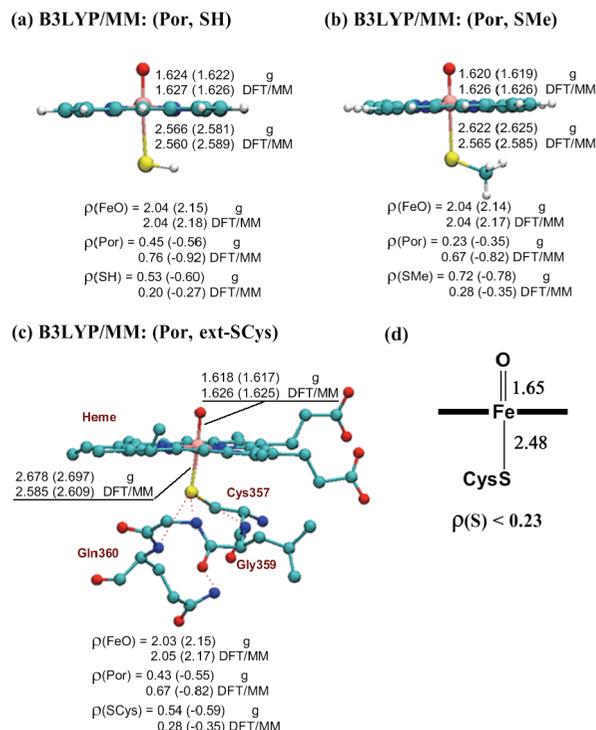
Figure 17. (a−c) Gas-phase and DFT(B3LYP/B1)/MM optimized structures (Å) and group spin density data for three different QM regions.[53] Data refer to $^4A_{2u}(^2A_{2u})$ states. (d) Experimental EXAFS geometry[276] (Å) and EPR-ENDOR spin densities[75] for the analogous CPO Cpd I.

Table 1. DFT(B3LYP)/MM and DFT(B3LYP)-Only Computed[a] Mössbauer Properties—Isomer Shift δ, Quadrupole Splitting ΔE_Q, and Asymmetry Parameter η—of Cpd I Species of CYPs and Analogous Thiolate Enzymes, Including Available Experimental Data for CPO

Cpd I species	δ (mm/s)	ΔE_Q (mm/s)	η	source
CYP101-enzyme	0.10	0.58	0.19	ref 160
CYP101-enzyme	0.13	0.67	0.09	ref 77
CYP101-model	0.09	1.34	0.06	ref 77
CYP3A4	0.11	1.36	0.17	ref 55
Se-CYP101	0.11	1.19	0.08	ref 62
CPO	0.11	0.73	0.32	ref 160
NOS	0.08	1.00	0.21	ref 160
CPO expt	0.15	1.02	n/a	ref 422

[a] DFT-only data refer to gas-phase calculations for a CYP101 model system.

parameters of Cpd I in the gas phase and in the protein. The first two entries correspond to the Por,ext-SCys model of CYP101,[77] while other entries refer to the Por,SH model of CYP3A4[55] and of CYP101, its L358P mutant, and the analogous CPO and NOS enzymes.[160] The largest difference between the protein and the gas phase (see the first three entries) is seen in the quadrupole splitting parameter. However, some changes are also apparent in the isomer shift and asymmetry parameters. The changes of the quadrupole splitting and isomer shift parameters reflect the reduced charge density on the iron center in the protein environment, in line with the increased porphyrin cation radical character, while the larger asymmetry parameter reflects the greater off-planar doming distortion of the porphyrin moiety from planarity in the protein. Similar differences were noted for other spectral properties, such as zero field splitting parameters and EPR/ENDOR parameters.[77] Interestingly, the enzymes that are closer than others to the gas-phase environment are CYP3A4 (4th entry) and NOS (7th entry).
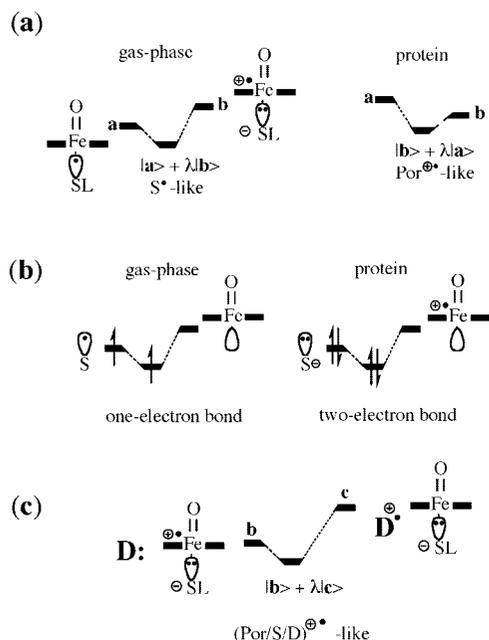
Figure 18. Valence bond modeling of the influence of the protein environment on the electronic structure of Cpd I: (a) Valence bond mixing of the contributing structures. (b) The Fe−S bond orbital, its occupancy, and the type of the Fe−S bond, depending of the environment. (c) The valence bond mixing situation in which an additional electron donor, **D:**, contributes to the electronic structure and participates in the cation-radical manifold.[50,53,198,199] Reprinted with permission from ref 32. Copyright 2005 American Chemical Society.

In the case of NOS vs the CYP and CPO enzymes, this difference was found[160] to correlate with the differences in the orientation of the EEF vector of the protein vis-à-vis Cpd I in a manner analogous to the model study of the EEF effect;[275] thus, whereas, in most enzymes, the EEF vector is aligned approximately with the S−Fe−O axis, in the case of NOS it lies in the porphyrin plane. This finding highlights the importance of the protein field in shaping the properties of Cpd I.

The protein-field effect also has energetic consequences that were reviewed before.[32] Thus, the range of QM energetic stabilization (within the QM/MM energy) of Cpd I by the protein relative to the gas-phase species, for the three ligand models (Por,SH; Por,SMe; Por,ext-SCys), is of the order of $100-150$ kcal mol^{-1}. Recalling that the QM energy component in QM/MM includes the effect of the MM charges, these stabilization energies reflect the field effect exerted by the protein on Cpd I, due to the interaction of the partial charges and H-bonding with the active species. Interestingly, the same quantity in CYP3A4 was calculated to be significantly lower (85 kcal mol^{-1}), in accord with the similarity of the corresponding Mössbauer data to the gas phase noted above in Table 1.

3.7.1. A Model for the Chameleon Behavior of Cpd I

The importance of the protein field effect, and its effect on the properties of Cpd I, can be rationalized most simply using the valence bond (VB) models shown in Figure 18.[31,50,53,198,199] The main idea in this model is that the iron-oxo porphyrin cation-radical moiety is a powerful electron acceptor, and therefore, any good electron donor can donate electrons into the porphyrin hole and participate in the electronic structure as part of the cation-radical moiety. In the Cpd I species itself, the good electron donor is the thiolate

ligand, and therefore, we can view Cpd I as a composite of the two resonance structures in Figure 18a, one in which the porphyrin is closed-shell and the thiolate is a radical (state $|\mathbf{a}\rangle$) while the other (state $|\mathbf{b}\rangle$) is an ion-pair with a cation radical on the porphyrin and a thiolate anion. In the gas-phase $|\mathbf{a}\rangle$ is below $|\mathbf{b}\rangle$ and the energy separation depends mainly on the choice of the axial ligand, e.g., thiolate or methyl mercaptide, etc. However, a dielectric medium and/or hydrogen bonding interactions will stabilize the ion-pair state $|\mathbf{b}\rangle$, to the extent that it moves below $|\mathbf{a}\rangle$, as evident from the large energetic protein-field effect discussed above. In this case, the mixed state becomes more $|\mathbf{b}\rangle$-like, having more spin density on the porphyrin. This is indeed what the above calculations show in Figures 16 and 17. An alternative way to understand the change in Figure 18a is by noting that the singly occupied "a_{2u}" orbital of Cpd I changes its character from a sulfur-dominated orbital in the gas phase to a porphyrin cation radical in the protein or in the condensed phase.[26,77,197] Generally speaking, the electronic structure in this picture is expected to show sensitivity to nearby additives, which are polar, such as water molecules[50] or charged residues.[160] Moreover, as shown in Figure 18b, the Fe−S bond changes from a weak and long one-electron bond in the gas phase to a stronger and shorter two-electron bond in the protein. Therefore, Cpd I is a mixed-valent state and as such, its Fe−S distance and electronic properties depend on the hydrogen-bonding network and polarity of the surrounding enzyme.

However, as shown in Figure 18c, thiolate is not the only donor in the protein environment. There may be other electron donors, labeled generally as D: (with an electron pair in a donor orbital). These donors may be other protein residues or the substrate itself.[199] In such an event, we can think of an additional resonance structure, where the external donor molecule/moiety donates one electron to Cpd I, leading to D$^{\bullet+}$/Cpd II. The mixing of the two structures now creates a new hybrid state, where the cation radical moiety will be spread over the sulfur, the porphyrin, and the D moieties, i.e., (S/Por/D)$^{\bullet+}$. Alternatively, if the donor moiety is a powerful one and there is no delocalization path that allows mixing, the *electronic* structure will tend to localize as D$^{\bullet+}$/Cpd II, altogether. Such an electronic situation is known among peroxidases, where cytochrome *c* peroxidase (C*c*P) involves a Cpd II and Trp$^{\bullet+}$, whereas ascorbate peroxidase (APX) has a normal Cpd I, due to subtle changes in the proximal side of the active site.[196,272] By contrast, the chameleon-like behavior with delocalization of the cation radical species over a few moieties has recently been observed computationally in NOS, where Cpd I, the arginine substrate, and the cofactor H$_4$B seem to share the cation radical.[69] Moreover, this phenomenon has recently been postulated to occur in the P450 isoform called P450 StaP (CYP245A1),[195] which has been confirmed by DFT(B3LYP)/MM computations.[58]

The DFT(B3LYP)/MM spin density distributions and singly occupied orbitals of the complex of P450 StaP Cpd I with the substrate chromopyrrolic acid (CPA) are shown in Figure 19. It is seen that, in all three situations considered, the cation radical is distributed over Por, sulfur, and CPA. CPA participates through the indolic moiety, while the Por/sulfur moieties participate through the normal a_{2u} orbital. When the His$_{250}$ residue is protonated in its distal side, in Figure 19a, the spin density on CPA is small (0.23), and when it is protonated at the proximal side, in Figure 19b,
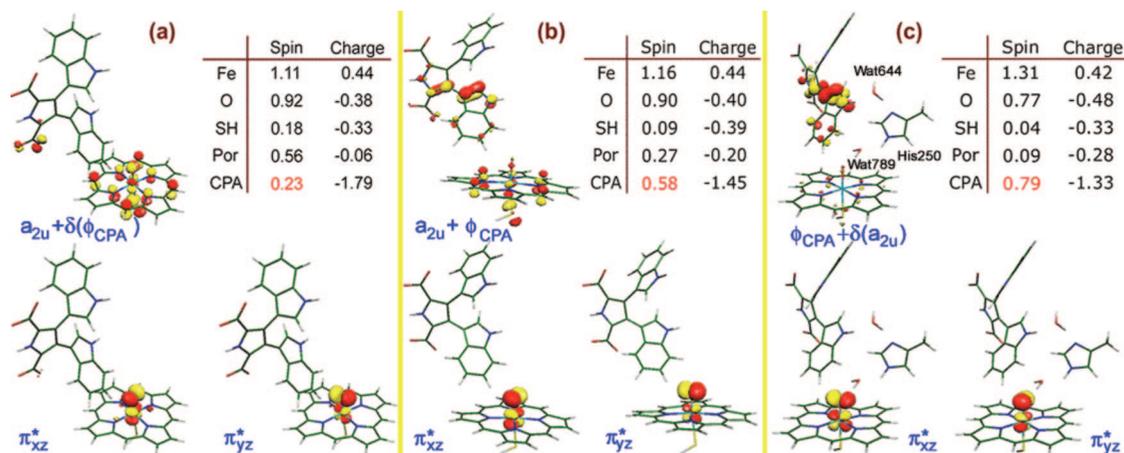
Figure 19. Electronic structures and spin densities of Cpd I of P450 StaP with the substrate, chromopyrrolic acid (CPA) in three situations: (a) His_{250} protonated at the N atom distal to Cpd I and Wat_{789} absent; (b) His_{250} protonated at the N atom proximal to Cpd I and Wat_{789} absent; (c) His_{250} protonated at the N atom proximal to Cpd I and Wat_{789} present. Note that the third unpaired electron is initially (a) almost completely in the a_{2u} porphyrin orbital and finally (c) almost completely on the indole group of CPA. Reprinted with permission from ref 58. Copyright 2008 American Chemical Society.

the CPA spin density rises to 0.58 and the singly occupied orbital acquires significant indolic character. Finally, in Figure 19c, when Wat_{789} (the water molecule that splits off during the formation of Cpd I; see above) is added, the CPA spin density rises to 0.79 and the singly occupied orbital looks virtually indolic. This delocalization of the cation radical species over to CPA is determined by the fact that the protein causes double deprotonation of CPA and holds it by an extensive array of H-bonds. However, in the dianionic state, CPA becomes a powerful enough donor to participate in the cation-radical species, and at the same time, the hybrid $(Por/S/CPA)^{\bullet+}$ character arises due to the delocalization path that involves $CPA\text{-}Wat_{644}\text{-}His_{250}\text{-}Wat_{789}$ moieties and the interaction of the CPA carboxylate with the propionic acid side chain on the porphyrin. Thus, while the chameleon property varies with protein fluctuation, it is still apparent in all conformations.

At the same time, we must recall that not every interaction will make a difference. For example, recent calculations[160] on the L358P mutant of CYP101,[194] where the NH···S interaction due to the NH group of Leu_{358} is severed by mutating it to Pro, showed that the removal of the H-bonding-like interaction has hardly any effect on the geometry and spin density of Cpd I. Similarly, in CYP3A4, the analogous NH···S interaction with the Met_{445} residue was shown to be insignificant.[55] However, the other NH···S interactions (in CYP101 with Gly_{459} and Gln_{360}, and in CYP3A4 with Gly_{444} and Ile_{443}) are important[55] and are responsible for the shifting of the spin density from sulfur to porphyrin, as shown in Figures 16 and 17.

Another case in point is the early DFT(ROB3LYP)/MM study of Cpd I of CYP101,[157] which reported considerable spin density on the propionate side chain A (propionates 7). However, in this case, extensive test calculations using DFT(B3LYP)/MM[55,59,131,163,277] did not find any significant spin density at the propionate groups whenever these are properly screened by the neighboring arginines, *both in the gas phase and in the protein*. This screening makes the propionates poor electron donors, and therefore, they do not participate in the electronic structure, as initially proposed.[157] However, it may well be that, under the conformational fluctuations of the protein, Cpd I may acquire transitory spin density on the propionate side chains[74] and provide a delocalization path as found in NOS, where the H_4B cofactor

properties $^2A_{2u}/^4A_{2u}$

CYP-isozyme	$r_{Fe\text{-}S}$ $^2A_{2u}/^4A_{2u}$	ρ_S $^2A_{2u}/^4A_{2u}$
2C9(X)	2.54–2.57/2.54–2.57	-(0.37–0.39)/0.34–0.37
2C9(MD)	2.63/–	-0.45/–
2B4(X)	2.51/2.52	-0.33/0.32
2B4(MD)	2.62/–	-0.41/–
3A4(X)$_1$	2.53/2.53	-0.39/0.37
3A4(X)$_2$	2.55/2.56	-0.47/0.45
3A4(MD)	–/2.62	–/0.50
101(X)	2.53/2.54	-0.20/0.20
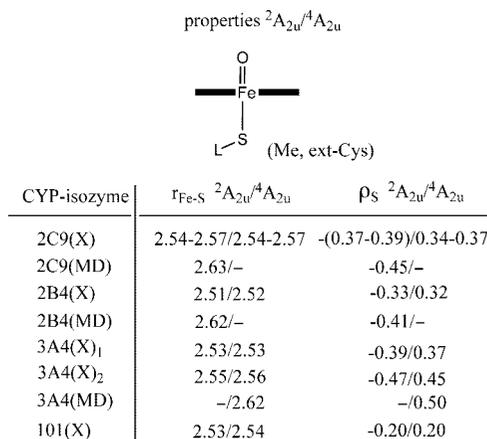
Figure 20. DFT(B3LYP)/MM optimized Fe−S distances (Å) and sulfur spin densities for the doublet and quartet spin states of Cpd I for different CYP isozymes. X and MD denote the snapshots derived from the X-ray structure and from the preparatory MD simulations, respectively. Reprinted with permission from ref 56. Copyright 2005 American Chemical Society.

acquires spin density (partially as $H_4B^{\bullet+}$) due to the interaction with the propionate side chain.[69]

3.7.2. Cpd I Species in Human P450 Isoforms

An extensive DFT(B3LYP)/MM study of Cpd I in the human isoforms CYP2C9, 2B4, and 3A4, with comparison to CYP101,[56] involved the Por,SMe and Por,ext-SCys models and several other variants, with residues in the distal and proximal sides as well as on the porphyrin ring. The starting geometries for QM/MM optimization were mostly taken from an X-ray structure (snapX), and in some cases from snapshots of an MD run (snapMD). The results are shown in Figure 20; there are changes from one isozyme to the other, but they are not larger than the changes due to fluctuations within a given isozyme. The authors concluded, therefore, that while the electronic structure of Cpd I, in the human isoforms tested, is somewhat dependent on the environment, the differences are not larger than those found for a given isoform in different snapshots. This is clearly true for the Cpd I species of the limited human isoform group used in this study but not valid for the entire range of Cpd I structures investigated since then (see below).
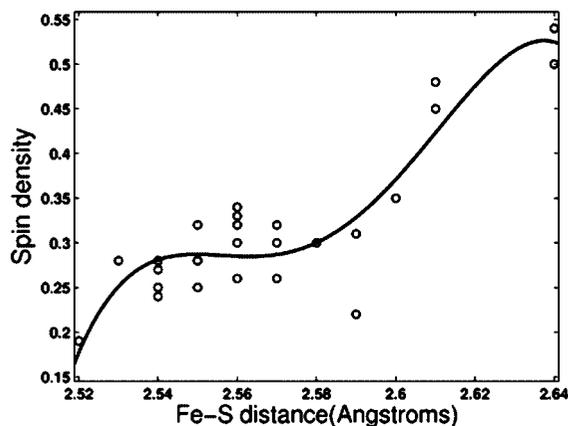
Figure 21. The variation of spin densities on sulfur as a function of the Fe−S distance in CYP3A4 Cpd I in 28 DFT(B3LYP)/MM structures optimized along the MD trajectory. Reprinted with permission from ref 55. Copyright 2007 American Chemical Society.

Figure 22. Three models used for studying the effect of cooperative binding (CB) on Cpd I of CYP3A4 with the drug molecule called diazepam (Dia). From left to right: (a) Dia-far, Cpd I with one diazepam after 5.9 ns of MD; the diazepam is far away. (b) Dia-prox, Diazepam docked with the methylene group undergoing activation near the oxo group of Cpd I. (c) Dia_CB, Cpd I with two diazepam molecules; the farthest one is labeled as "the effector". Reprinted with permission from ref 55. Copyright 2007 American Chemical Society.

3.7.3. Effect of Substrate Binding on Cpd I of CYP3A4

An extensive DFT(B3LYP)/MM study[55] focused on CYP3A4 in order to understand the observation of cooperative binding, whereby substrate oxidation follows faster kinetics when two or more molecules of substrate are admitted into the active site, which is large in this isoform. Seven substrate-free models of Cpd I were used, with different thiolate ligand representations and with propionate substituents on the porphyrin ring and the corresponding arginines that form salt bridges, as well as distal and proximal residues (in the latter cases those substituents that form NH···S interactions with the proximal Cys$_{442}$ ligand). Two starting structures (PDB: 1TQN and 1W0E) were chosen and subjected to a 0.5 ns MD simulation. Optimizations starting from the initial X-ray structures produced DFT(B3LYP)/MM geometries and spin densities similar to those obtained in the above study of Bathelt et al.[56] Changes occurred, however,[55] in the human isoforms tested, when the optimizations started from the snapshots along the MD trajectory. The results are summarized in Figure 21 for 28 different QM/MM optimized structures. It is seen that as long as the Fe−S bond is shorter than 2.61 Å, the spin density on sulfur remains approximately at 0.25, whereas, for longer Fe−S bonds, the spin density reaches values larger than 0.50; the maximum occurs for the structures optimized after 0.5 ns MD. Thus, within 0.5 ns, the substrate-free active species tends to elongate its Fe−S bond and to localize the radical on the sulfur; both aspects may signal instability of the species (possibly by sulfur-based radical abstraction of a hydrogen atom from neighboring residues).

In a second set of calculations, Fishelovitch et al.[55] interrogated the effect of the substrate by studying the three different models shown in Figure 22. The substrate is a drug molecule called diazepam (Dia), known to exhibit cooperative binding with two molecules inside the pocket. As shown in Figure 22, the first model (Dia_far) includes a single unrestricted diazepam molecule that was QM/MM optimized starting from the 5.5 ns snapshot from the MD simulation trajectory;[189] here, the diazepam molecule drifted away from Cpd I, reaching a distance of 9 Å. In the second model (Dia_prox), one diazepam molecule was docked and constrained to the same position it has when the other diazepam molecule is present, but in the absence of the second molecule (the effector, see below). And finally, in the third model (Dia_CB),

there are two diazepam molecules, simulating cooperative binding (CB); one close to Cpd I is called the "substrate"; the other one, far away, is called "the effector".

All three models were subjected to MM optimization and 0.5 ns MD prior to QM/MM optimization; the results are depicted in Figure 23. Figure 23a and b shows the situation of Cpd I in the former model (Dia_far) where the substrate was allowed to drift away. It is seen that the Fe−S bond here is long, and this elongation is attended by spin density localization on the sulfur. Moreover, the oxo group of Cpd I is H-bonded over short distances to nearby protein residues, T309 (Thr$_{309}$) and A305 (Ala$_{305}$) (Figure 23a and b). Thus, when the substrate is not in the proximity of the active site, Cpd I will be surrounded by abstractable H-bonds from the protein residues and will be consumed before it can react with the substrate. This is reminiscent of attempts to generate Cpd I by shunting the catalytic cycle with peracid, which generally results in Cpd II and a protein radical.[253] The long Fe−S bond and the high S-radical content further aggravate the situation, since sulfur itself may abstract hydrogen and convert P450 to the inactive form known as P420.[7]

By contrast, the situation in Figure 23c is very different: in the presence of the substrate, the Fe−S bond remains short even after 0.5 ns, the sulfur spin density is well below 0.5, and the only C−H bond close to Cpd I is the one of the diazepam molecule; all other H-bonds from the adjacent residues are either displaced or elongated. Adding the effector diazepam to the calculations has no noticeable effects on the geometry or electronic structure of Cpd I, other than what we just mentioned. As such, the role of the effector is simply to push the substrate mechanically and hold it close to Cpd I. Therefore, when the substrate is close to Cpd I, the only abstractable hydrogen will be the one from the substrate. Thus, the effect of cooperative binding in a wide-mouthed isoform such as CYP3A4 is to optimize substrate oxidation by preventing the destabilization of Cpd I (via Fe−S elongation) and by ensuring that no other abstractable hydrogens are in the vicinity of the active species. A detailed analysis showed that this robustness of Cpd I was associated also with the stabilization of the Fe−S bond by the two NH···S interactions (from Ile$_{443}$ and Gly$_{444}$) with the proximal ligand.

Interestingly, a substrate effect on the lifetime of the oxidizing species was noted experimentally in CYP101.[176,241,278] Thus, in the substrate-free case, the last dioxygen-containing species (Cpd 0) has an extremely short lifetime.[278] When
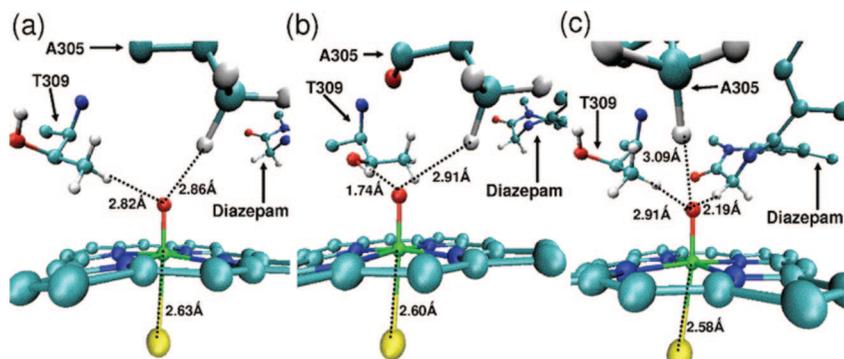
Figure 23. DFT(B3LYP/B1)/MM optimized geometries of the three models in Figure 22: Parts a and b are two snapshots of Dia_far; note that the closest hydrogens to Cpd I are those in the CH group of A305, and the OH or CH groups of T309 (A is Ala; T is Thr). (c) Snapshot from Dia_prox; note that diazepam displaces the interacting H-moieties of the A305 and T309 residues, and its own C−H bond is now closest to Cpd I with an O···H distance of 2.19 Å. Note also the short Fe−S bond in part c compared with those in parts a and b. Reprinted with permission from ref 55. Copyright 2007 American Chemical Society.
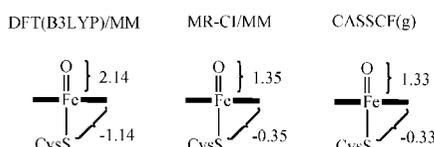
Figure 24. Spin densities obtained from DFT(B3LYP)/MM, MR-CI/MM, and CASSCF(g) studies of the doublet state of CYP101 Cpd I.[77,78,274]

camphor was added, the lifetime of the oxidant species increased 20-fold, while using a bulkier substrate increased the lifetime 80-fold. As such, the substrate stabilizes Cpd I, allowing it to live longer near the substrate so it can optimize substrate oxidation. In the absence of the substrate, Cpd I gets consumed all too quickly by the protein residues. In CYP101, with its tight binding pocket, one molecule of substrate is sufficient to achieve the effect, but in CYP3A4, with a very large pocket, the same effect requires cooperative binding.

3.7.4. Description of Cpd I by Highly Correlated QM Levels

Three studies of Cpd I by highly correlated levels have appeared between 2005 and 2008.[77,78,274] All three studies used initial CASSCF wave functions of varying sizes, which were augmented subsequently by accounting for dynamic correlation; in the first two by MR-CI methods[77,78] and in the third one by CASPT2.[274] Furthermore, the first two studies included the effect of the protein, hence MR-CI/MM, while the third one employed a gas-phase model, hence CASPT2(g). The three studies agree that Cpd I has two virtually degenerate doublet and quartet states, and while the MR-CI/MM results show conclusively a doublet ground state,[77,78] the CASPT2(g) results are less decisive, reflecting probably the lack of the protein environment, as was reasoned before from model studies.[198]

In both sets of studies, the quartet state properties closely resemble the DFT(B3LYP)/MM results. However, the doublet state comes out quite different. Figure 24 shows the spin density distribution from DFT(B3LYP)/MM, MR-CI/MM, and CASPT2(g) calculations for the doublet state. The difference is apparent: while the DFT method yields two spins on the FeO moiety, one with spin-up and one spin-down on the combined Por/S moiety, the correlated calculations give a spin density of 1.33−1.35 on FeO and a total of −(0.33−0.35) on the Por/S moieties.

Scheme 4. (a) Spin Arrangements of the π^*_{xz}, π^*_{yz}, and a_{2u} Orbitals in the Three Determinants (D_1−D_3) That Are Required To Obtain a Spin-Adapted Doublet-State Wave Function $^2\Phi_{111}$ and (b) Covalent Configuration $^2\Phi_{111}$ and the "Ionic" Ones, $^2\Phi_{210}$ and $^2\Phi_{012}$, Which Contribute to the Doublet $^2A_{2u}$ State
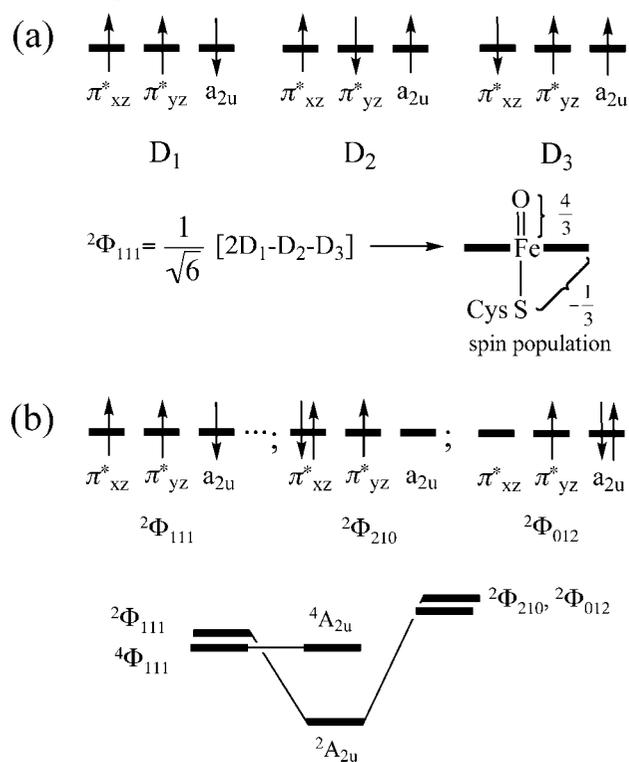
One of the problems of DFT is the use of a single Kohn−Sham determinant to describe the doublet state. This leads to a broken symmetry solution, since the single determinant is not a definitive spin state and does not properly describe the open-shell singlet coupling of the a_{2u} and π^* electrons; therefore, the spin distribution of the $^2A_{2u}$ state is incorrectly evaluated by DFT. However, this can be easily remedied by constructing from the DFT determinant a spin-adapted wave function that describes properly the doublet spin, as shown in Scheme 4a. Thus, since the electron in the a_{2u} orbital is singlet-coupled to each one of the electrons in the π^* orbitals, the corresponding wave function will involve a combination of spin arrangements as shown in Scheme 4a. Therefore, if we order the orbitals as $\pi^*_{xz}\pi^*_{yz}a_{2u}$, the spin arrangements will be $D_1 = |\uparrow\uparrow\downarrow\rangle$, $D_2 =$

$|\uparrow\downarrow\uparrow\rangle$, and $D_3 = |\downarrow\uparrow\uparrow\rangle$, with the former having a double coefficient and hence a quadrupled weight compared with the latter two. Since the spin of the a_{2u} electron has a spin down with a relative weight of 4 and a spin up with a relative weight of 2, the sum is a spin down weight of 2. When we normalize the wave function, the spin distribution in the doublet states should be $^2/_3$ in each of the π^*_{xz} and π^*_{yz} orbitals and $-^1/_3$ in a_{2u}. This distribution in Scheme 4 is close to the MR-CI/MM and CASPT2(g) distributions, and we may conclude that spin adaptation is sufficient to retrieve these high level results. As such, the $^2A_{2u}$ state is described at the correlated level as a triradicaloid with singly occupied π^*_{xz}, π^*_{yz}, and a_{2u} orbitals.

The recent MR-CI/MM study[78] emphasizes that the wave function of the $^2A_{2u}$ state has a multireference character, more so than the $^4A_{2u}$ state, which is dominated by a single configuration analogous to the DFT/MM Kohn–Sham determinant. As shown in Scheme 4b, in addition to the triradicaloid configuration, $^2\Phi_{111}$, with a single electron in each one of the occupied π^*_{xz}, π^*_{yz}, and a_{2u} orbitals, the $^2A_{2u}$ state wave function has contributions from $^2\Phi_{210}$ and $^2\Phi_{012}$. The latter two are obtained from $^2\Phi_{111}$ by back-and-forth single electron shifts between the π^*_{xz} and a_{2u} orbitals. This configuration mixing is analogous to the covalent-ionic mixing that stabilizes singlet diradicals in valence bond (VB) theory.[279] By analogy, if we refer to $^2\Phi_{111}$ as a covalent electronic structure, then the other two configurations are simply the corresponding ionic ones. This "covalent-ionic" mixing stabilizes the $^2A_{2u}$ state below the $^4A_{2u}$ state, as shown in Scheme 4b. Interestingly, these configurations were not found in the CASPT2(g) study,[274] which reported a close analogy of the $^2A_{2u}$ state to the corresponding DFT description, presumably just the triradicaloid structure $^2\Phi_{111}$. The mixing found in the MR-CI/MM study, of course, cannot be very stabilizing, since the two states remain very close (within <1 kcal mol^{-1}) to each other. Nevertheless, the presence of the "ionic configurations" is sufficient to lower $^2A_{2u}$ below $^4A_{2u}$ and may reflect the effect of the protein on the electronic structure, discussed in model studies.[198]

3.7.5. Low-Lying Excited States of Cpd I

Here we restrict ourselves to excited states that may have some impact on the reactivity of Cpd I, while excited states that are useful for spectroscopic identification (e.g., the split Soret band) of Cpd I species have been discussed by Harris et al.[268] The $^{4,2}A_{1u}$ excited states have been considered either as ground or low-lying excited states in Cpd I and other metalloporphyrin species.[280] DFT and DFT/MM calculations usually predict these states of P450 Cpd I to lie 7–10 kcal mol^{-1} higher than the corresponding $^{4,2}A_{2u}$ pair.[78,266] However, as noted recently for HRP Cpd I,[74] the distortion of the porphyrin ring by the protein causes mixing of these states, such that the ground states have A_{2u}–A_{1u} mixed character. Furthermore, a recent treatment[280] showed how to analyze a variety of metalloporphyrin systems and obtain the A_{2u}–A_{1u} contents in the ground states. Having said that, the state with dominant $^{4,2}A_{1u}$ character is still predicted by DFT to lie well above the $^{4,2}A_{2u}$ ground states. However, a recent MR-CI/MM treatment[78] gave a rather small $^2A_{1u}$–$^2A_{2u}$ gap of only 1.9 kcal mol^{-1}, a feature which merits further investigation.

The other pair of states, which are interesting from the point of view of chemical reactivity,[200] are the pentaradicaloid $^{6,4}A_{2u}$ states, which arise from electron promotion from the
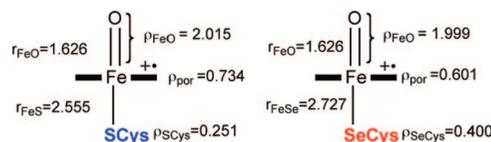
Figure 25. Key geometric features (Å) and spin densities obtained in the quartet state of the Cpd I species of the wild type CYP101 (S-Cpd I) and the selenocysteine mutant (Se-Cpd I) from DFT(B3LYP)/MM calculations.[62]

δ to the σ^*_{xy} orbital in Figure 3a. Early DFT(B3LYP) calculations[266] predicted this state to lie only 14 kcal mol^{-1} higher than the $^{4,2}A_{2u}$ ground states. Subsequent DFT/MM calculations[277] verified that these states are indeed low lying, ca. 11 kcal mol^{-1} above the ground state. The major factor that stabilizes these pentaradicaloid states is the increased number of d–d exchange interactions generated by the excitation. This stabilization may lower the transition states nascent from these states sufficiently to make them relevant for the reactivity of Cpd I.

3.7.6. Se-Cpd I for an in-Silico Designed Mutant of P450

The recently discovered 21st amino acid is selenocysteine, in which selenium replaces the sulfur atom. It was therefore interesting to find out whether an in-silico DFT(B3LYP)/MM designed mutant would be stable and what kind of properties the corresponding Se-Cpd I would have compared with the wild-type (WT) S-Cpd I.[57,277] The in-silico mutation was done by simply replacing the sulfur by Se in the DFT(B3LYP)/MM calculated[53,156] resting state, **1**, and in Cpd I of the WT enzyme, taking a few snapshots from the MD trajectory and optimizing them by DFT(B3LYP)/MM.[57,277] The resting state Se-**1** was found to be stable and sufficiently close to WT-**1** (having a doublet ground state), with an Fe–Se distance close to that in known organometallic Fe/Se complexes. The geometric and spin density features of Se-Cpd I are shown in Figure 25 alongside the S-Cpd I analogue. It is seen that the major difference is the greater radical character of Se-Cpd I, which is almost identical to the gas-phase spin density in S-Cpd I. Indeed, the calculated Mössbauer parameters of Se-Cpd I are quite close to the gas-phase parameters of S-Cpd I.[57] Thus, the increased donor ability of CysSe vis-à-vis CysS causes a greater mixing of the Cys-Se•/Cpd II state into the Cys-Se:$^-$/Cpd I state in terms of the VB mixing model in Figure 18a.

Once again we see how the powerful electron acceptor property of the Por•$^+$FeO moiety results in a chameleonic behavior of the Cpd I species. In the meantime, the mutant was made experimentally using different techniques.[64,65]

3.8. Concluding Remarks on the Catalytic Cycle of P450

The cycle of P450 has many subtle factors that influence the function of the enzyme. As discussed before,[143] and further argued above based on DFT(B3LYP)/MM calculations, the gating of the cycle by a single water molecule is enabled by the negative charge and electron donor property of the thiolate (cysteinate) ligand, i.e., "push"-effect. In the absence of axial thiolate, the resting state would have been

reduced too, while in its presence there is a narrow window of redox potential that applies only to the pentacoordinate intermediate **2** and the oxy-ferrous species **4**. Further into the catalytic cycle, the "push"-effect increases the proton basicity of the species **5** and Cpd 0 (**6**), thereby affecting their efficient protonation, while at the same time preventing the reduction of these species. In the last step of the catalytic cycle, the "push"-effect generally prevents reduction of Cpd I so *that it preferentially reacts by hydrogen abstraction or bond formation*.[143,281] However, as the example of P450 StaP above shows, this is not a firm rule, since in some cases the substrate is a good enough donor to partly reduce Cpd I and to participate in the electronic structure of this species. This brings us to the chameleon nature of Cpd I, which originates in the electron acceptor capability of the iron-oxo porphyrin cation radical moiety that tends to accept electron density from its internal donor, the cysteine, as well as from any sufficiently good donor around, be this protein residue or a substrate. The potential implications of the chameleon nature may be far reaching but have not yet been fully explored.

There are still unresolved issues with regard to other steps in the catalytic cycle, for example, concerning the role of the reductase, the participation of individual water molecules, and the mechanism of the proton transfers, which may be different in different mutants.[282] What is also missing is a way to assess the cycle as a whole based on the integration of its many features. Preliminary MD studies by Fishelovitch et al.[283] suggest that the reductase plays a role in regulating the cycle, not only by providing the electrons needed for reduction but also by the conformational changes it induces in the protein such that formerly closed water pathways[178] can open and assist the function of the enzyme. The MD simulations in CYP3A4 show that the arginine residue that forms the salt bridge with propionate 7 (propionate A) is flexible and functions as a gate. When the substrate (e.g., testosterone) is in the pocket, the gate is closed. However, as soon as the reductase is attached to the heme from the proximal side close to the cysteine ligand, it causes the Arg residue to rotate and open the gate. When the gate opens up, water molecules flow into the pocket and form a water chain that connects with the acid−alcohol pair (Glu and Thr).[283] Some support for the specific status of the salt bridge to propionate 7 was recently demonstrated by the work of Hayashi et al., who showed that severing the propionate substituent has an adverse effect on the functionality of the enzyme and its efficiency of electron uptake.[221] If these conclusions are general for the P450 family, then the catalytic cycle is completely regulated by the attachment of the reductase to the heme. Thus, upon substrate binding, water molecules including the aqua ligand leave the site,[178] thereby making the heme a better electron acceptor and causing an attachment of the electron donor reductase. Once the reductase is attached, it causes both reduction of the heme and a timely entrance of water molecules that can shuttle protons required for the O−O bond cleavage and the formation of Cpd I. In this manner, P450 enzymes possess a functional catalytic cycle that drives off water molecules, when these are not necessary, and lets them in when they are required. Quite an ingeneous cycle!

4. QM/MM and QM-Only Studies of P450 Reactivity

4.1. Coverage of Reactivity and Mechanistic Issues

Let us start with issues that have been reviewed before and will thus receive only minor coverage herein. Then we shall address new reactivity issues that will be discussed here at some length.

One of the key features of reactivity derives from the nearly degenerate spin states of Cpd I, which brings about a hybrid reactivity pattern, so-called two-state reactivity (TSR), that affects product distributions, kinetic isotope effects (KIEs), and other reactivity aspects through the interplay of the two reactive spin states of Cpd I.[30,145,284] This has been reviewed amply, by some of us[25,26,31,32,83] as well as by others,[18,22,285−287] focusing on how the interplay of the two spin states in TSR elucidates the controversial radical-clock lifetimes, the rearrangement patterns, and other features in C−H hydroxylation. We shall therefore not cover these topics as such, although the notion of TSR and multistate reactivity (MSR) will be mentioned throughout the following discussions.

Another important issue that has been prominent in mechanistic discussions is the putative presence of a second oxidant species in the cycle of P450, which until recently was assumed to be Cpd 0.[24] The evidence for the participation of Cpd 0 in substrate oxidation arose from the observation that mutants in which the proton relay that converts Cpd 0 to Cpd I is disrupted by site-directed mutagenesis[23,183,257] exhibit different reactivity patterns compared with the WT enzyme. For example, the T252A mutant of P450$_{cam}$ does not hydroxylate camphor but is capable of epoxidizing the double bond of 5-methylenylcamphor, albeit more sluggishly than the WT enzyme.[183] Since the catalytic cycle of T252A is thought to terminate at Cpd 0, this necessarily means that Cpd 0 should be involved in the epoxidation of an activated double bond but that its reactivity is smaller than that of Cpd I. Curiously, the double mutant of P450$_{cam}$, T252A/D251N, in which the protonation machinery has been disrupted by mutations of both Thr$_{252}$ and Asp$_{251}$, shows partial activity in camphor hydroxylation,[184] which indicates that Cpd I might after all be present even in these mutants. Moreover, a mechanistic study[256] reports that oxidation reactions, thought before to involve Cpd 0, are caused in fact by Cpd I. As we already mentioned, there are now direct experimental data showing that Cpd 0 is orders of magnitudes less reactive than Cpd I, and that in WT P450 it should be silent in the presence of Cpd I.[234,263−265] Thus, while the reactivity of Cpd 0 in the "absence" of Cpd I seems to be generally accepted, the very recent QM/MM calculations of the T252X mutants (section 3.6.3) show that even in these mutants Cpd I is formed, albeit somewhat more slowly than in the WT enzyme, and might be the actual oxidant (it will possess a smaller steady-state concentration due to increased uncoupling). Be this as it may, the simultaneous reactivity of Cpd I and Cpd 0 in the WT enzyme is hardly considered to be viable anymore, with a few exceptions.[288,289] Since this Cpd I/Cpd 0 conundrum has been reviewed by some of us before,[26] it will not be covered here.

Another but similar issue is the recent suggestion that the only oxidant in P450 is the perferryl PorFe(V)=O species and not at all Cpd I, since model Cpd I species react much slower than the assumed reaction rate in P450$_{cam}$.[261,267] As discussed above (section 3.7), however, both QM-only and

QM/MM calculations show that the perferryl species lies too high in energy to react and that the electromeric conversion of $PorFe(V)=O$ to $Por^{•+}Fe(IV)=O$ should be virtually barrier free. In fact, the more recent study of Newcomb et al.[262] actually shows that an in situ generated CYP119 Cpd I, by means of LFP,[267,290] is more reactive than model Cpd I species in bimolecular reactions and is, hence, a viable oxidant.

Two comments might be in order here: (a) The use of bimolecular reaction rates vs rates in the enzyme is an insecure grounds for deducing mechanistic conclusions on the enzymatic reaction: A factor that slows down bimolecular reactions is the entropic contribution due to the loss of translational and rotational degrees of freedom in the transition state, which raises the free energy barrier, by up to 12 kcal mol^{-1}, where these degrees of freedom are completely lost, or less if they are only partially lost (this loss enters the $T\Delta S^{\ddagger}$ term that raises the free energy barrier).[26,291] This factor does not play a role in P450 enzymes, because here, the loss of degrees of freedom in the formation of an enzyme–substrate reaction complex is compensated by an increase of the degrees of freedom as the water molecules leave the pocket upon substrate entrance; another factor is the enthalpy of substrate binding (e.g., in CYP101, there is a $Tyr_{96}\cdots O=C$ hydrogen bond with camphor).[59,180] The enzymatic reaction then starts from a "complex" that has already lost two rotational and all translational degrees of freedom, and hence, the free-energy barrier for the reaction from this complex should be lower than that of the bimolecular models by a quantity that can reach up to a maximum of 12 kcal mol^{-1} if the translational and rotational degrees of freedom are completely lost. Thus, the complexation effect accounts for a large part, if not all, of the difference observed between the rate constants of the bimolecular reactions vis-à-vis the enzymatic process. (b) Deducing mechanistic conclusions by comparing reaction rates for an LFP generated Cpd I of CYP119[262] to the reaction rate in P450$_{cam}$ estimated by extrapolation from the cryogenic temperature[176] is also insecure, since one has to assume a certain rate constant for these cryogenic experiments and to extrapolate this assumed value from a very low temperature to ambient temperature, which is a questionable procedure. An ultimate resolution of this mechanistic conundrum will come by comparing the reaction rates of the LFP generated Cpd I with the native reaction of the same enzyme toward a given substrate.

Based on the above considerations, Cpd I emerges as the most important if not the sole oxidant of P450s. Therefore, although we shall present some DFT(B3LYP)/MM results that compare the reactivity of Cpd I to other possible oxidants such as Cpd II, etc., the discussion in the following sections will focus on the reactivity patterns of Cpd I. Accordingly, we shall cover DFT/MM results on C–H hydroxylation,[59,61,131,163,164,277] C=C epoxidation,[277] aromatic hydroxylation,[60] and C–C coupling.[58,80] In addition, we shall describe DFT-only results for N-dealkylation,[166,167] ethanol oxidation,[168] hydroxylation of arenes,[172] dehalogenation of polyhalogenated benzene derivatives,[165] and heteroatom (N,S) oxidations.[170,171,292,293] In the following section (section 5), we shall discuss attempts to correlate P450 barriers,[170,172] as well as the recent valence bond approach to the modeling of the barriers and mechanisms for C–H hydroxylations.[174]
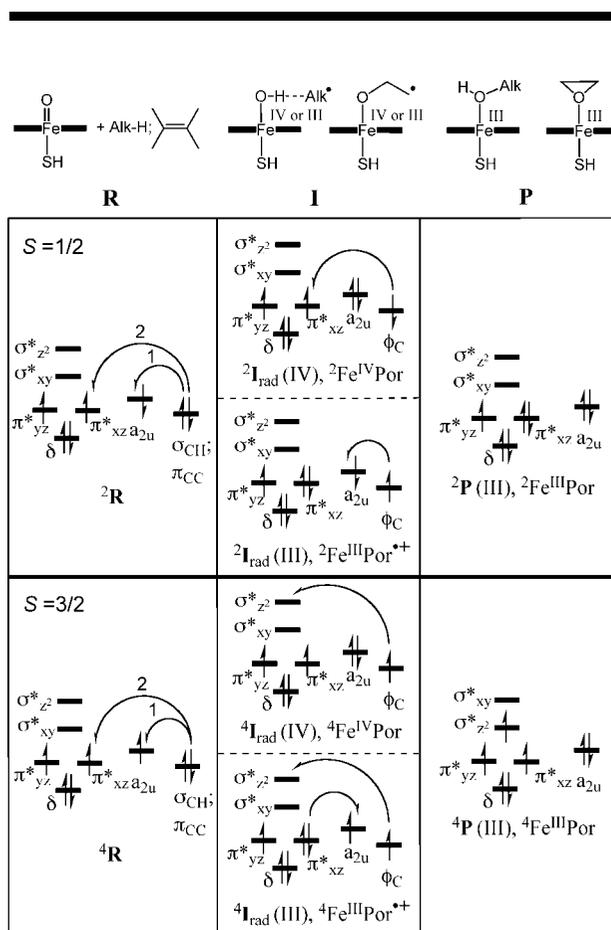
Figure 26. Orbital occupancy evolution diagrams during C–H hydroxylation and/or C=C epoxidation by Cpd I, in the doublet ($S = {}^{1}/_{2}$) and quartet ($S = {}^{3}/_{2}$) states. Reprinted with permission from ref 25. Copyright 2007 American Chemical Society.

4.2. Reactivity of Cpd I: Counting Electrons, Spin States, and Electromeric Situations

As we argued at the outset, the "oxidation state formalism" has been used fruitfully in inorganic chemistry, as a heuristic device for following the electronic reorganization and for assigning the "right number of electrons" to the d-block orbitals of the metal. Here we use it in Figure 26 in terms of orbital occupancy evolution during C–H hydroxylation and/or C=C epoxidation.

The reactant states, $^{2,4}\mathbf{R}$, involve the two spin states ($S = {}^{1}/_{2}$, ${}^{3}/_{2}$) of Cpd I, having the orbital picture already shown in Figure 3a, and the substrate, which is represented here by a doubly occupied orbital, σ_{CH} or π_{CC}. The heme of Cpd I effectively contains an Fe^{IV} center and a hole in the porphyrin (i.e., Por$^{•+}$) and thus has two oxidation equivalents more than the product state complexes $^{2,4}\mathbf{P}$ that have an Fe^{III} center and a closed-shell porphyrin. Thus, any reaction of Cpd I will ultimately involve two formal electron "transfer" events from the substrate undergoing oxidation to the heme. Starting from the left-hand side, Figure 26 shows two of these electron shift events, one to the a_{2u} orbital and the second to the π^{*}_{xz}(FeO) orbital, in the quartet and doublet spin states. The first electron shift generates an Fe^{IV}-type intermediate with a closed-shell porphyrin, while the second shift generates an Fe^{III}-type intermediate with a radical cationic situation on the porphyrin ring. In both cases, the substrate retains a radical center with a singly occupied orbital labeled ϕ_C. In addition, each of these two electromers will come in two

spin varieties, so we might expect involvement of at least four radical intermediate states during the reactions.[83] During the second bond formation step, an additional electron is "transferred" from the substrate to fill either the a_{2u} orbital or one of the d-orbitals of the iron. Filling the π_{xz}^* orbital generates the doublet-spin ferric product, whereas the filling of the $\sigma_{z^2}^*$ orbital leads to the quartet-spin ferric product. Of course, in synchronous reactions, both "electron transfer" events will transpire in one step (see later the sulfoxidation mechanism).

It is apparent already from Figure 26 that there are other possible states, which are obtained by redistribution of the electrons in this dense orbital manifold. For example, shifting electrons from the ϕ_C orbital of the radical intermediate center into one of the d-orbitals will generate a variety of states in which the substrate is a cationic species, as described in studies of the clock substrates.[147,152,294] Additionally, there are low-lying sextet and quartet states (section 3.7.5), which may generate intermediates of sufficiently low energy that can participate in the reaction.[200,277] As such, the reactivity of Cpd I will involve at least two spin states, and possibly many electromeric states, hence TSR and MSR.

4.3. DFT/MM and DFT-Only Studies of C−H Hydroxylation by Cpd I of P450

Many DFT-only studies were performed to elucidate the mechanism of C−H hydroxylation and to resolve the controversy around the radical clock data.[45,146−153,295] The more recent studies involve DFT/MM modeling of camphor hydroxylation by P450$_{cam}$.[59,61,131,155,157,163,164,277,296] The nature of the iron-hydroxo intermediate during the reaction was recently studied by means of MR-CI/MM calculations.[78] The DFT and DFT/MM studies cover more than 10 different substrates, which has led to a general and unified mechanistic picture that involves TSR and MSR, as reviewed before.[32]

4.3.1. Camphor Hydroxylation by CYP101 and the Controversial Role of Spin Density on the Propionate Side Chains

An early DFT(ROB3LYP)/MM(OPLS) study[155] of camphor hydroxylation by CYP101 yielded a surprisingly low barrier of 11.7 kcal mol^{-1} for the quartet state reaction, as shown in Figure 27. The low barrier which could have potentially explained the elusiveness of Cpd I was ascribed to a drift of spin density to the propionate side chain (propionate-A); thus, the propionate spin density, which starts as 0.2 at the reactant complex, diminished to 0.10 at the transition state. It was argued[155,157] that the diminishment of the propionate spin density at $^4\mathbf{TS}_H$ causes an increase of the propionate negative charge, which is stabilized by the salt bridge to the neighboring arginine, thereby lowering the barrier and stabilizing the intermediate $^4\mathbf{I}$. However, subsequent studies did not find support for this attractive mechanism.

The first comprehensive DFT(B3LYP)/MM study of camphor hydroxylation by CYP101[59,277,296] employed four different snapshots after MD equilibration, using two different models of the QM subsystem (Por,SH and Por,ext-SCys), and two different basis sets. The results are depicted in Figure 28, where it is seen that the reaction coordinate exhibits a hydrogen abstraction phase (bond activation) and a reorientation of the alkyl moiety attended by a rebound process that generates the ferric-alcohol complexes. The bond activation phase involves the quartet high-spin (HS) and
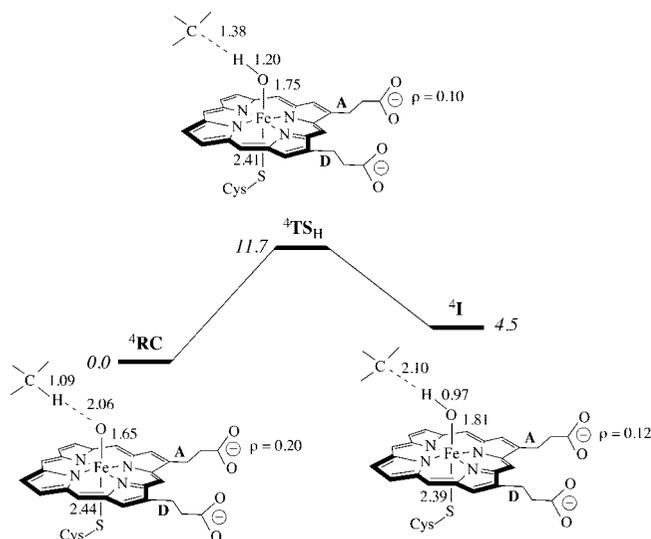
Figure 27. DFT(ROB3LYP)/MM results[155,157] for geometries (Å) and spin densities (ρ values) during H-atom abstraction from camphor by CYP101. Energies in kcal mol^{-1}.

doublet low-spin (LS) bond activation transition states ($^{4,2}\mathbf{TS}_H$), which originate in the two spin states of Cpd I. The LS species, $^2\mathbf{TS}_H$, lies slightly below the HS one; generally, the two bond activation transition states are close in energy to within 2 kcal mol^{-1} or less. This phase generates the clusters $\mathbf{I}$, which involve an alkyl radical (CAM$^•$) weakly coordinated to the iron-hydroxo species; the $^{2,4}\mathbf{I}$ intermediates appear in the Fe^{III}Por$^{•+}$ and Fe^{IV}Por varieties discussed above (Figure 26), but in the figure we show generically the two lowest $^{4,2}\mathbf{I}$ states (for a discussion of all the states, see later). Subsequently, the rebound on the HS manifold encounters a significant barrier, while from the LS manifold the barrier is negligible, ca. <0.3 kcal mol^{-1}. The scenario is extremely similar to the findings in the corresponding gas-phase process,[59,277,296] as well as in all the gas-phase reactions studied thereafter.[32,174] Thus, this study confirms the TSR picture that emerged from gas-phase model calculations using the small model representation, Por,SH.[146]

However, comparison of the DFT(B3LYP)/MM barriers to those in Figure 27 shows that while the transition state geometries are similar, the barriers in the DFT(B3LYP)/MM study are significantly higher. Since the DFT(ROB3LYP)/MM calculations (Figure 27) included also the propionate side chains, while the DFT(B3LYP)/MM calculations (Figure 28) did not, it was deemed necessary to test the putative appearance of the spin density on the propionate, by including these side chains in the DFT(B3LYP)/MM calculations and performing single-point calculations on the so optimized structures.[59,277,296] These tests show that as long as the propionates are properly screened by Arg$_{299}$, these side chains have zero spin density. Furthermore, the presence or absence of propionate spin density does not lower the barriers in these calculations, and DFT(B3LYP)/MM geometry optimization did not change these conclusions.[277]

Since then, three other comprehensive studies were carried out to benchmark the barrier for the camphor hydroxylation reaction and to test the propionate spin density hypothesis.[61,131,163] Thus, Zurek et al.[163] used DFT(B3LYP)/MM, different setups of the system, different MD procedures, and different QM models. The results are shown in Figure 29, and it is apparent that, despite the differences in the two DFT(B3LYP)/MM approaches, the results are very similar to those of Schöne-
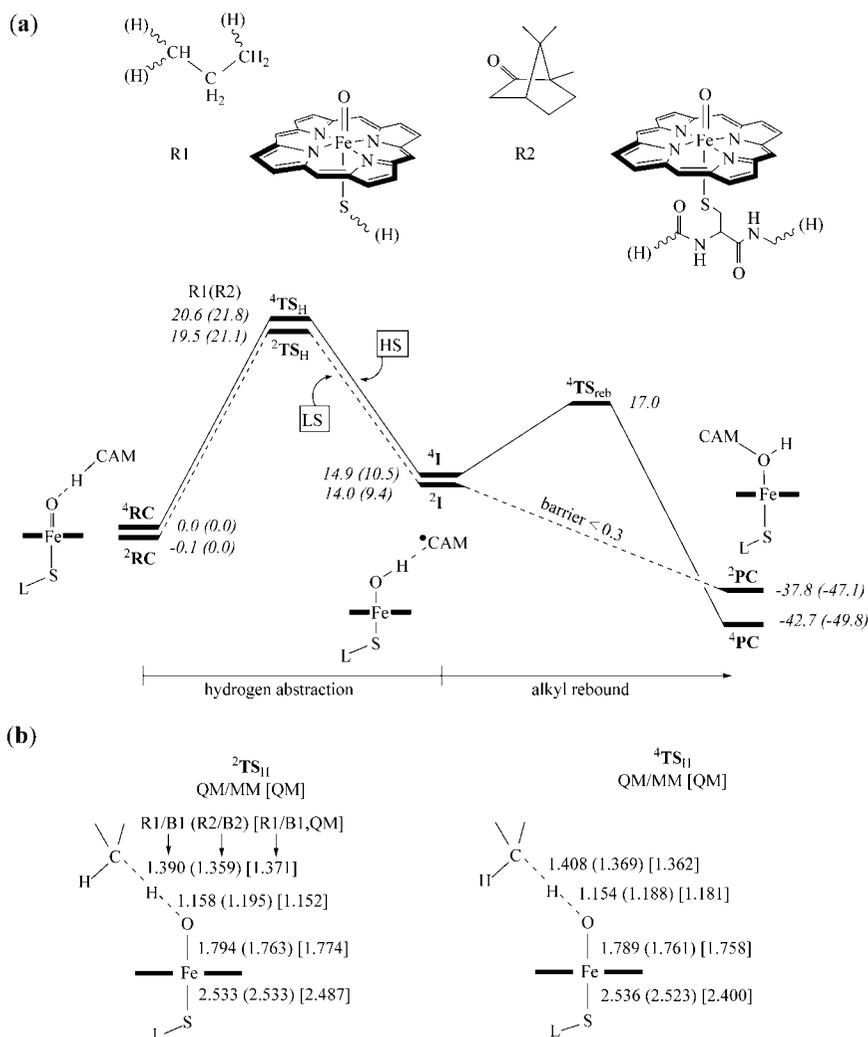
Figure 28. Representative DFT(B3LYP/B2)/MM energy profiles for camphor (CAM-H) hydroxylation calculated for the equilibrated 40 ps snapshot. The QM regions are labeled as R1 (Por,SH; truncated CAM-H) and R2 (Por,ext-SCys; full CAM-H). Data correspond to R1 (R2), respectively. Energies (a) in kcal mol^{-1}, distances (b) in Å. Reprinted with permission from ref 59. Copyright 2004 American Chemical Society.

boom et al.[59,296] in Figure 28, in the sense that the two studies lead to barriers of 15−20 kcal mol^{-1} depending on models and basis sets. The largest barriers were obtained for QM region R2 when the Asp$_{297}$ residue adjacent to propionate-A is deprotonated. Zurek et al.[163] carefully studied the role of propionate spin density, which appears with deprotonated Asp$_{297}$ and then actually raises the barrier. Furthermore, the observed short O···O distance of 2.4−2.8 Å between the oxygen atoms of Asp$_{297}$ and propionate-A[178] could not be reproduced with deprotonated Asp$_{297}$, in which case both QM/MM optimizations and MD runs result in O···O distances of up to 4 Å. By contrast, with protonated Asp$_{297}$ the O···O distance is stable at 2.7 Å.

The source of the differences between the published DFT(B3LYP)/MM and DFT(ROB3LYP)/MM results[59,155,157] for the hydrogen abstraction barrier in the HS state was traced through systematic calculations[61,131] that employed a variety of QM models ranging from 51 atoms for the simplest one (the Por,SH model in Scheme 2) up to 170 atoms (with full camphor, ext-SCys ligand, all porphyrin substituents, and their salt bridges, as well as a crystal water molecule, Wat$_{903}$). It was found that (i) the two procedures gave essentially the same barriers *as long as one used the same geometries*, (ii) no spin density appeared on the propionate side chains *as long as these groups were properly screened*, (iii) *the*

barriers were raised somewhat whenever spin density appeared on the propionate side chains, and (iv) reoptimization of the DFT(ROB3LYP)/MM **RC** and **TS**$_H$ species,[155,157] *with a consistent protein environment,* gave a barrier about 5 kcal mol^{-1} higher than reported originally.[155,157] The latter point (iv) serves as a reminder that QM/MM optimizations require great technical care to avoid artifacts and that it is actually easier to determine technically reliable barriers with the smallest QM-model (Por,SH), which poses less technical difficulties during optimization. Overall, these systematic comparisons have resolved the noted discrepancies and have shown that the different DFT/MM approaches yield compatible results when applied in a consistent manner.

4.3.2. Camphor Hydroxylation by CYP101—How Does Wat$_{903}$ Catalyze the Reactions?

Another important finding of these systematic studies[61,131] is shown in Figure 30, which depicts the ^{4}**RC** and ^{4}**TS**$_H$ species of a representative model, along with the HS barrier and the charges on the oxygen atom of the FeO bond. There is a nearby water molecule, labeled as Wat$_{903}$, which is seen in the X-ray structure (1DZ9)[178] and presumably represents the water molecule liberated during the formation of Cpd I from Cpd 0 (see Figure 1). During
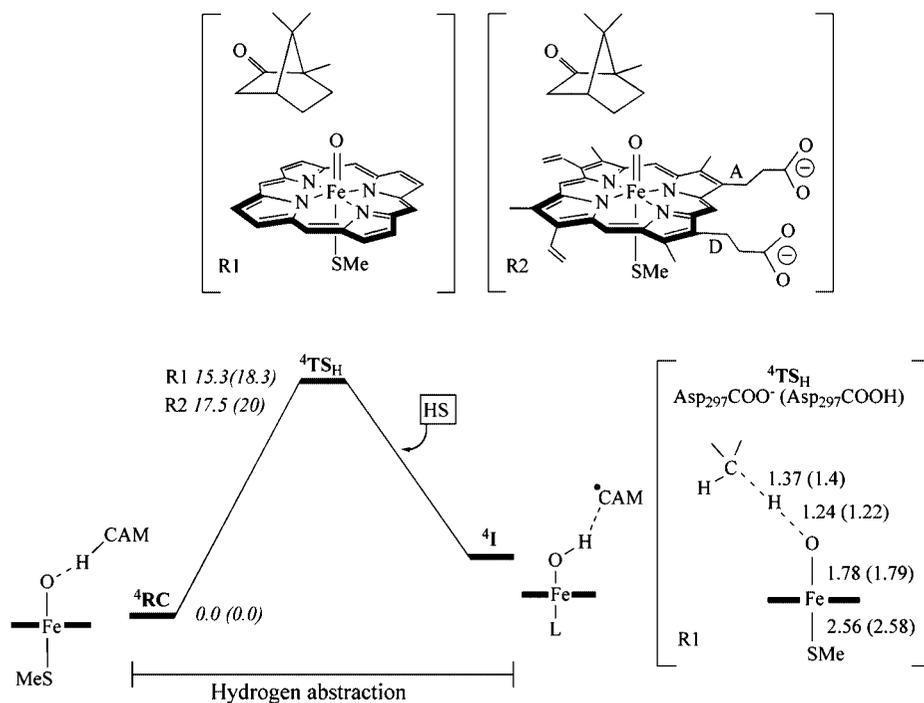
Figure 29. DFT(B3LYP/B2)/MM energy profiles for camphor hydroxylation in the high-spin state of Cpd I, for two models, R1 (Por,SMe; full CAM-H) and R2 (ext-Por,SMe with heme side chains; full CAM-H). Relative energies (kcal mol^{-1}) and distances (Å) are given for setups with deprotonated Asp$_{297}$ (protonated Asp$_{297}$).[163]
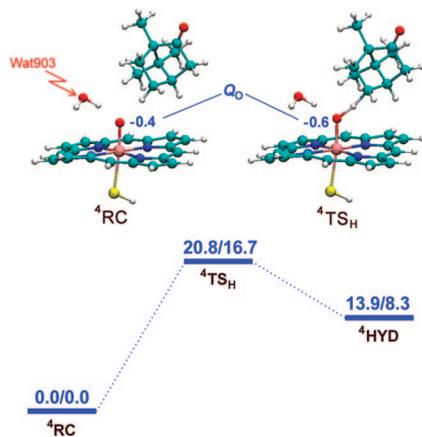
Figure 30. DFT(B3LYP/B1)/MM relative energies (kcal mol^{-1}) of the stationary points for hydrogen abstraction from camphor by quartet Cpd I, in the absence/presence of Wat$_{903}$. Shown are the ^{4}**RC** and ^{4}**TS**$_H$ species and the Mulliken charge of the oxo ligand in the presence of Wat$_{903}$.[61,131]

the preparatory MD runs in our original QM/MM work,[296] Wat$_{903}$ moved to the cavity around camphor and thus did not directly participate in the reaction.[59] By contrast, the QM/MM calculations of Guallar and Friesner[157] started from the X-ray structure and followed the QM/MM reaction path without prior MD runs, with Wat$_{903}$ remaining near the oxo group. The presence of this water molecule lowers the hydrogen abstraction barrier; when it is included near the oxo group in our original setup, the barrier goes down by 4 kcal mol^{-1} [61,131] and the endothermicity of the intermediate, ^{4}**I**, is reduced by 6 kcal mol^{-1}. Virtually the same energy changes are obtained whether the Wat$_{903}$ molecule is part of the QM region or the MM region, which suggests that the effect of Wat$_{903}$ is primarily electrostatic. The same effect is observed in isolated gas-phase models, with and without the Wat$_{903}$ molecule; in each case, the water molecule lowers the barrier and the endothermicity of ^{4}**I**. Furthermore, simple electrostatic calculations using the Mulliken charges reproduce the barrier lowering of the Wat$_{903}$ molecule. As can be seen from Figure 30, the favorable interaction of the Wat$_{903}$ molecule with ^{4}**TS**$_H$ arises because the negative charge on its oxo group increases compared with the reactant complex, ^{4}**RC**, due to the charge transfer ($\sim$0.3−0.4e) that occurs from camphor to the Cpd I moiety in the transition state. This autocatalytic effect of the Wat$_{903}$ molecule is intriguing; since Wat$_{903}$ is liberated during the conversion of Cpd 0 to Cpd I, it catalyzes the consumption of its own creation.

A subsequent DFT(B3LYP)/MM study[164] further explored the role of Wat$_{903}$ on the barriers as well as on the TSR feature of C−H hydroxylation. The study used three different QM models, a few snapshots, and three different basis sets. Figure 31 shows representative results for the R1 model (Por,SH) using the B1 basis set. The results confirm that the presence of Wat$_{903}$ does not affect the TSR feature; at the bond activation transition state, ^{2}**TS**$_H$ is still below ^{4}**TS**$_H$, the LS rebound is still almost barrier free, while the HS rebound has a significant barrier. Additionally, Figure 31 shows that the barriers without Wat$_{903}$ are 4 kcal mol^{-1} higher than those calculated in its presence, and similarly, the intermediates ^{4,2}**I** are stabilized by 6 kcal mol^{-1} or so.

Therefore, a water molecule that is present near the oxo group of Cpd I will play a biological role as a catalyst. These results imply that the C−H hydroxylation barrier will be susceptible to additives present near the reaction center (water molecules, ions, positive residues, etc), as suggested in the past on the basis of DFT-only model calculations.[150,275,292]

4.3.3. Camphor Hydroxylation by CYP101—More States than Just Two

As already discussed above, there are more states that can be involved in the C−H hydroxylation (as well as other reactions) by Cpd I: (a) the pentaradicaloid states,[200] which
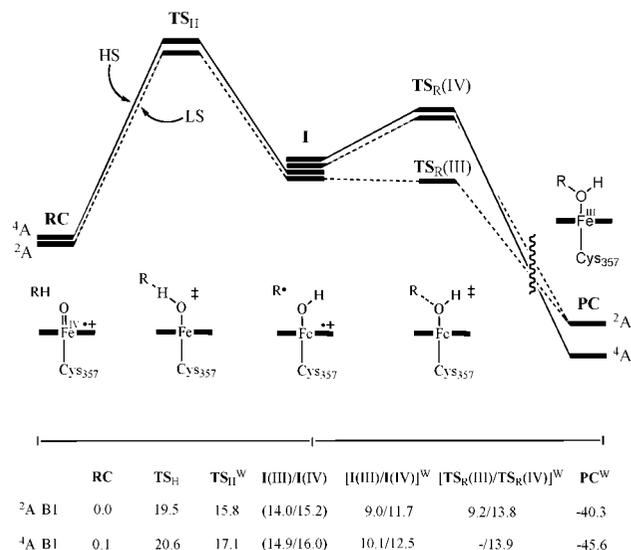
Figure 31. DFT(B3LYP/B1)/MM energy profiles for camphor hydroxylation using the R1 model (Por,SH; truncated-CAM-H) with and without Wat$_{903}$. The relative energies (kcal mol^{-1}) are given below the drawing. Species with W$_{903}$ are labeled with a superscripted "W", e.g., TS_H^W, etc. Reprinted with permission from ref 164. Copyright 2007 American Chemical Society.

	RC	TS$_H$	TS$_{II}^W$	I(III)/I(IV)	[I(III)/I(IV)]^W	[TS$_R$(III)/TS$_R$(IV)]^W	PC^W
^{2}A B1	0.0	19.5	15.8	(14.0/15.2)	9.0/11.7	9.2/13.8	-40.3
^{4}A B1	0.1	20.6	17.1	(14.9/16.0)	10.1/12.5	-/13.9	-45.6

originate by promotion of an electron from the δ to the σ^*_{xy} orbital (see Figure 3a), and (b) the two electromeric varieties of the intermediates, ^{4,2}I(IV) and ^{4,2}I(III).[32,59,83] In fact, as previously suggested,[32,59,83] all these states may contribute to reactivity and product distribution during the oxidation by Cpd I. The role of these states was examined by Altun et al.[78,164] using DFT(B3LYP)/MM calculations for all these extra states, and MR-CI/MM calculations for the intermediate states ^{4,2}I(IV) and ^{4,2}I(III) states.

Figure 32 shows the energy profiles of the pentaradicaloid states and the orbital occupancy evolution diagram. While, at the reaction onset, these states are 12 kcal mol^{-1} above the ground states, the gap decreases to less than 5 kcal mol^{-1} at the TS_H stage and further goes down to 4 kcal mol^{-1} at the intermediate state **I** (see Figure 32a), quite similar to a DFT-only study done before.[200] The reason for the low energy of these states at the intermediate stage can be understood from the orbital occupation diagrams shown in Figure 32b. It is seen that going from the reactant state to the intermediate, ^{6}**I**, the spin-up electron that is shifted from the substrate orbital (σ_{CH}) to the $\sigma^*_{z^2}$ d-orbital on iron adds four new d-d exchange interactions (which can contribute as much as 10−20 kcal mol^{-1} per interaction), and this exchange stabilization overrides the orbital energy gap and lowers the energy of the intermediate, which is now a heptaradicaloid. The reduction of the energy gap between the ^{6}**I** and the lower ^{2,4}**I** states comes from these increased exchange interactions, and the small gaps imply that all these states may be accessible for rebound.

The energy spacing of all the intermediates is shown in detail in Figure 33, for the Por,SH model (R1) using two basis sets, B1 and B2W. With B1, the lowest states are the ^{2,4}I(III) electromers, and the entire energy spread is 4.9 kcal mol^{-1}. In the larger B2W basis set, the four ^{2,4}I(III) and ^{2,4}I(IV) states are condensed into 1.5 kcal mol^{-1}.[164] ^{6}**I** was not computed at this level, but most likely it will go down as well.[200] As shown in Figure 33, MR-CI/MM calculations give rise to similar results, even though the wave functions of all these species show some multireference character.[78] The match between MR-CI/MM and DFT(B3LYP)/MM
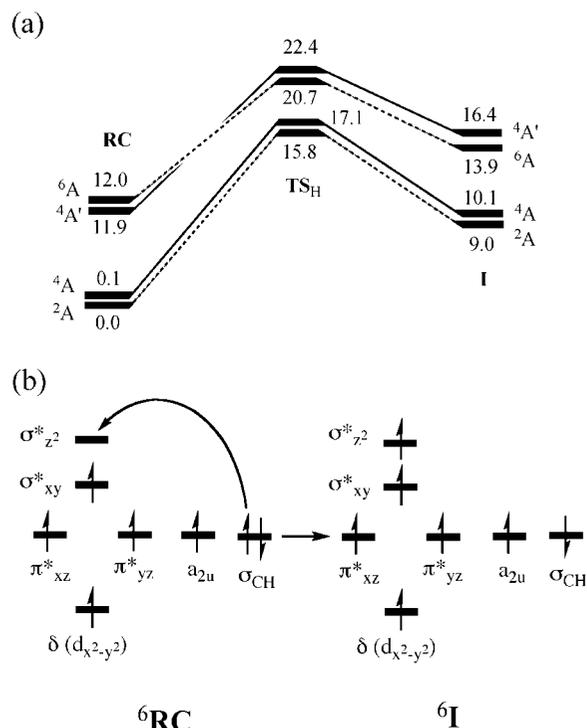
Figure 32. (a) DFT(B3LYP/B1)/MM energy profiles (kcal mol^{-1}) for hydrogen abstraction from camphor by the triradicaloid (doublet and quartet) and pentaradicaloid (sextet and quartet) of Cpd I. (b) Orbital occupancy evolution diagram showing the transformation from ^{6}**RC** to ^{6}**I**. Reprinted with permission from ref 164. Copyright 2007 American Chemical Society.
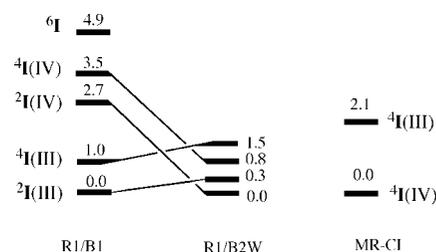
Figure 33. (a) DFT(B3LYP)/MM relative energies (kcal mol^{-1}) of the various intermediate states during camphor hydroxylation, ^{2,4,6}**I**, for R1 (Por,SH; truncated-Por) with two basis sets, B1 and B2W.[164] Shown are also some corresponding MR-CI/MM results with basis B'.[78]

shows once again that the latter method is reliable for P450 species. Finally, the small energy spacing of all the states in the intermediate stage argues for MSR, where the various intermediates exhibit complex dynamics due to their different rebound barriers and possible interconversions.

4.3.4. Camphor Hydroxylation by CYP101—The Product Release Step

According to detailed low temperature electron nuclear double resonance (ENDOR) studies at 200 K, product formation occurs through three distinct hydroxycamphor-bound conformations.[176] The spectroscopic results suggest that the immediate product after hydroxylation corresponds to a nonequilibrium state, involving a more or less normal Fe−O distance ($\approx$2 Å) and that the heme pocket then relaxes in two detectable steps to accommodate the anomalously long Fe−O bond of 2.67 Å, found in the crystal structure.[297]

The product release was reviewed before[32] and will thus only be summarized here briefly. Thus, DFT(B3LYP/B1)/MM calculations[158] on the ferric complex of hydroxy
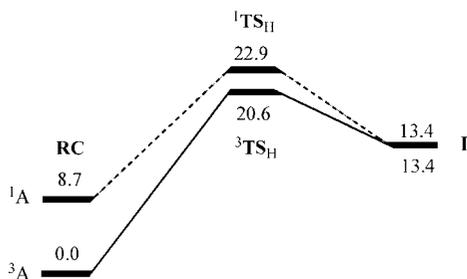
Figure 34. DFT(B3LYP/B1)/MM energy profiles (kcal mol^{-1}) for hydrogen abstraction from camphor by Cpd II in the presence of Wat$_{903}$.[164]

Snap-1, B1 [Mean, B1] (Snap-1, B2W//B1)

Figure 35. DFT(B3LYP)/MM results for the allylic hydroxylation of cyclohexene, using the R1 model (Por,SH; C_6H_{10}). Each data line contains relative energies for Snap-1, B1 [Mean, B1] (Snap-1, B2W), where the second datum refers to the mean value of seven snapshots (see text).[57,277]

camphor in its different spin states indicate a doublet minimum at an Fe−O distance of ca. 2.2 Å, and a flat, barrierless potential for the dissociation of the Fe−O bond. Comparisons with analogous calculations on the isolated QM system in the gas phase show that inclusion of the protein/ solvent environment lowers the activation energy for bond dissociation in the doublet state because of interactions within the binding pocket and accounts for a significant stabilization of the quartet and sextet states. The DFT(B3LYP)/MM results support the interpretation of the low-temperature ENDOR data[176] that the initially observed nonequilibrium species with a short Fe−O distance is a hexacoordinated product complex in the doublet state. The other two observed species with longer Fe−O distances could well correspond to pentacoordinated quartet and doublet minima with a relaxed protein environment, but a conclusive assignment is not possible in view of the flatness of the computed potential curves.[158]

4.3.5. Camphor Hydroxylation by CYP101—Can Cpd II Compete with Cpd I?

As mentioned above, all attempts to generate Cpd I by shunting the cycle with, e.g., peracids resulted in Cpd II and a radical on a protein residue. To recall, Cpd II is the 1-electron reduced form of Cpd I, with a filled a_{2u} orbital and a triplet ground state, ^3A (see orbitals in Figures 3a and 26). Theoretical considerations of reactivity factors during C−H hydroxylation[174] suggest that Cpd II should be less reactive than Cpd I for two major reasons that are associated with the much poorer electron acceptor capability of Cpd II compared with Cpd I; one is the diminished charge transfer from camphor to the iron-oxo moiety in the transition state, which denies the transition states of Cpd II the stabilization contributed by the polar character of the TS, and the other is the lesser stability of the H-abstracted PorFe(III)−OH/ CAM• intermediates. Experimentally it has been shown indeed that synthetic models of Cpd II are able to carry out C−H hydroxylation as well as other oxidation reactions, albeit more slowly than the corresponding Cpd I species.[258] In order to gauge the relative reactivities, Altun et al.[164] explored by means of DFT(B3LYP)/MM the reactivity of Cpd II toward camphor hydroxylation in CYP101, and the energy profiles are depicted in Figure 34. It is seen that, with the same QM model, R1 (Por,SH; truncated-CAM-H), and with the same basis set (B1), the barrier for H-abstraction by Cpd II is 5 kcal mol^{-1} higher than the corresponding one for Cpd I in Figure 31. This sluggish reactivity of Cpd II is in accord with known experimental data.[258] Interestingly, the reaction of Cpd II involves TSR too, since the singlet state of the species (with spin-up and spin-down electrons in the π* orbitals in Figures 3a and 26) leads to a transition state
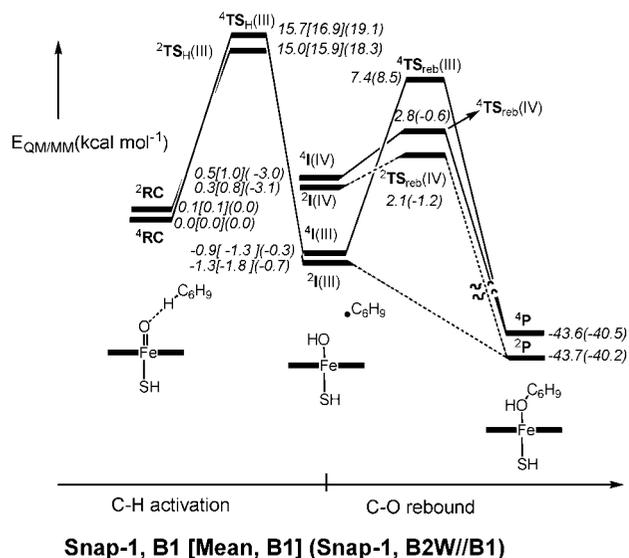
not much higher than ^3TS_H and a degenerate $^{1,3}I$ intermediate state. While the rebound was not studied, it is clear that the two states, which connect to different electronic structures of the Fe^{II} product complex, will possess different rebound barriers. Additionally, it was found that the H-abstraction barrier of Cpd II is not affected much by the local environment of the propionate side chains (e.g., by the protonation state of Asp$_{297}$).[164]

4.3.6. Allylic Hydroxylation of Cyclohexene by CYP101

The allylic hydroxylation of cyclohexene by CYP101 was studied by means of DFT(B3LYP)/MM using the simplest QM model R1 (Por,SH) with Wat$_{903}$ (Snap-X) and without it, for altogether eight different snapshots from the MD trajectory.[57,277] The energy profile for the C−H hydroxylation is depicted in Figure 35, for a representative snapshot (Snap-1) and reveals the same biphasic behavior, common to all C−H hydroxylations studied so far, with a bond activation phase and a rebound phase. The relative energies in Figure 35 are given for three different situations (i.e., values for snap-1 and basis B1, mean values for the seven snapshots and basis B1, values for snap-1 and basis B2W). As in camphor hydroxylation, here too, the four electromeric and spin states participate in the rebound. The highest rebound barrier is again[277,296] found for the $^4I(III)$ intermediate, while the $^2I(III)$ state possesses a barrier free rebound. Unlike in gas-phase models, the $^{2,4}I(IV)$ states both exhibit small barriers to rebound, perhaps due to steric constraints. But the overall picture is similar as in the camphor hydroxylation case; namely, the HS intermediates will possess longer lifetimes than the LS ones.

The role of the pentaradicaloid state of Cpd I ($^6A_{2u}$) was tested for one of the snapshots (Snap-6), and the relative spacing of the transition states, $^{2,4,6}TS_H$, for allylic hydroxylation is shown in Figure 36. It is seen that the energy gap between the lowest of the two transition states, $^{2,4}TS_H$, nascent from the ground triradicaloid state of Cpd I, and the sextet state, ^6TS_H, species is only 2.0 kcal mol^{-1} at the highest level. Thus, as in camphor hydroxylation, here too DFT(B3LYP)/MM predicts that the pentaradicaloid states
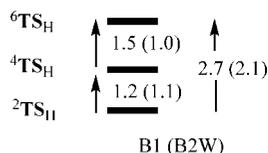
Figure 36. DFT(B3LYP)/MM $^{2,4,6}TS_H$ relative energies (kcal mol^{-1}) for the allylic hydroxylation of cyclohexene, using the R1 model (Por,SH; C_6H_{10}) and the B1 (B2W) basis sets.[277]
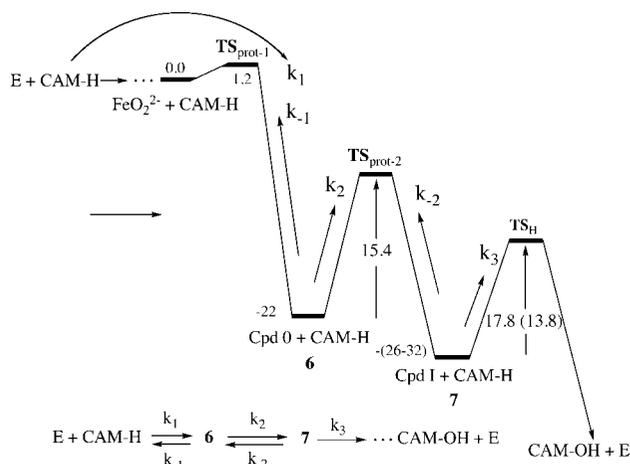
Figure 37. Mechanistic scheme for camphor (CAM-H) hydroxylation by Cpd I (**7**), starting from the enzyme−camphor complex (E+CAM-H) and going via Cpd 0 and Cpd I to the product complex. The relative energies (kcal mol^{-1}) are taken from DFT(B3LYP)/MM calculations.[61,164] The H-abstraction barrier during hydroxylation is given without and with ZPE correction (4 kcal mol^{-1}, as determined from gas-phase calculations[59]). At the bottom is the corresponding kinetic scheme with rate constants, and the steady-state condition for the concentration of Cpd I.

of Cpd I may play a role in C−H hydroxylation reactivity, thus making this a multistate scenario, MSR.

Finally, the H-abstraction barriers in the presence of Wat$_{903}$ (Snap-X) are lower than the Snap-1 barriers for the B1 (B2W) basis by 2.5 (1.7) kcal mol^{-1} in the LS state and by 1.5 (0.5) kcal mol^{-1} in the HS state.[277] Thus, in principle, Wat$_{903}$ still catalyzes the H-abstraction process, but only to a minor extent. Moreover, for B2W, there is one snapshot that has equal barriers with and without Wat$_{903}$, and therefore the status of the Wat$_{903}$ effect here is not clear without further sampling. As we shall see later, however, the effect on the C=C epoxidation barriers is clear-cut and rate-enhancing. It might thus be that catalysis by Wat$_{903}$ is substrate and reaction dependent.

4.3.7. Why is Cpd I Elusive in CYP101?

As mentioned before, Cpd I of P450$_{cam}$ continues to be an elusive species, which has not yet been detected experimentally in the native working cycle. Under steady-state conditions, the detectability of Cpd I will depend on a few rate constants that involve this species as well as Cpd 0. By now, the relevant energetics are available from DFT(B3LYP)/MM calculations, and we may therefore attempt to tackle this issue better than we did in the past.[32] Figure 37 shows a schematic DFT(B3LYP)/MM potential energy profile for the process from the enzyme camphor complex ES at the peroxo stage (FeO$_2^{2-}$) through Cpd 0 and Cpd I up to the hydroxycamphor product. The corresponding kinetic scheme is included below the energy profile. Assuming steady state

conditions (applied to the concentration of Cpd I), the ratio of Cpd 0 to Cpd I is given by the following expression:

$$d[Cpd\ I]/dt = 0 \rightarrow [Cpd\ 0]/[Cpd\ I] = (k_3 + k_{-2})/k_2 \quad (1)$$

Here k_3 is the rate constant for camphor hydroxylation, k_2 for the formation of Cpd I, and k_{-2} for the reverse process, going back from Cpd I to Cpd 0.

Using the Eyring expression, these rate constants are given as exponentials of the corresponding free energies of activation. While we do not have these free energies, computations of hydrogen abstraction in P450$_{cam}$ indicate a negligible entropic effect.[126] Thus, the computed DFT(B3LYP)/MM barriers with ZPE correction are expected to be good approximations to the corresponding free energies. These barriers and relative energies are presented in Figure 37 for the simplest model R1 (Por,SH) with the largest basis set available for the two processes.[61,159,164] The best estimate of the barrier for O−O bond heterolysis is 15.4 kcal mol^{-1} (LACV3P++**/6-311++G**), while the reaction exothermicity is in the range of 4−10 kcal mol^{-1}.[159] The experimental activation energy for O−O cleavage in the FeOOR complex of CYP119, obtained by shunting the reaction with m-chloroperbenzoic acid, is 14.1 kcal mol^{-1};[250] a value which is close to the computational datum and is slightly lower, as would be expected when the departing moiety is m-chlorobenzoic acid. The LS H-abstraction barrier from camphor at the highest level is 17.8 kcal mol^{-1} (for the large QM region labeled as R2s).[164] One should not forget, however, that zero-point energy (ZPE) correction reduces this latter value by about 4 kcal mol^{-1} because the contribution from one C−H stretching mode is lost at the transition state.[32] This will lower the H-abstraction barrier down to about 13.8 kcal mol^{-1} ($k_3 \sim 10^3$ s^{-1}); at the same time, the ZPE correction for the O−O cleavage should be significantly smaller. Taking the ZPE correction and the range of calculated exothermicity into account, the reverse barrier for going back to Cpd 0 is 6.4−10.4 kcal mol^{-1} higher than the barrier for H-abstraction. This means that Cpd I will be consumed (k_3) $10^4−10^7$ faster than it can revert to Cpd 0 (with k_{-2}). Hence, eq 1 can be simplified to:

$$[Cpd\ 0]/[Cpd\ I] = k_3/k_2 = \exp[(\Delta E^{\ddagger}_2 - \Delta E^{\ddagger}_3)/RT] \quad (2)$$

While the limited accuracy of our computational data suggests some caution, it is still interesting to describe the scenario that follows from the mechanistic scheme with this data. Thus, if we ignore the tunneling contribution to k_3, the ratio of k_3/k_2 at room temperature will be >10, and at 200 K, which is the cryogenic temperature used to probe the various intermediates in the cycle,[176] this ratio is 55:1, already reaching the detection limit of the experiment. Moreover, if tunneling is important during H-abstraction, this could further lower the effective barrier, maybe by as much as 3 kcal mol^{-1},[298] which would then lead to a value of ca. 10.8 kcal mol^{-1}. In this case, the ratio of concentration becomes 40,000:1, which is clearly way beyond the detection limits. In fact, even a small tunneling correlation that cuts only 1 kcal mol^{-1} from the barrier would place the ratio at 600:1, again outside the detection limit. One should note, however, that the measured kinetic isotope effects (KIEs) do not provide conclusive evidence for tunneling in P450$_{cam}$: the KIE values for camphor hydroxylation (ca. 4−5)[299] are lower

than those for norbornane hydroxylation (ca. 11)[300] and thus less indicative of tunneling. Be this as it may, the scenario in Figure 37 seems already sufficient to rationalize the elusiveness of Cpd I. Furthermore, in the case of the T252X mutants (section 3.6.3), where the barrier for the conversion of Cpd 0 to Cpd I can rise up to 17 kcal mol^{-1}, the likelihood of probing Cpd I is even smaller than in the WT enzyme. One can also include in the steady state derivation of eq 1 other reactions which consume Cpd I, e.g., the reduction of Cpd I to the resting state (the oxidase reaction)[180] or its autoxidation etc. In such a case, eq 1 will include in the numerator a sum of all the rate constants corresponding to the consuming reactions. For the WT enzyme, all these rate constants are smaller than k_3 (the rate constant for oxidation) and hence eq 2 is still valid. Of course, this scenario predicts that Cpd I can be observed in cases where k_3 is made very small, e.g., by using a perfluorinated derivative of camphor, or in cases where k_2 is large, e.g., by the design of an appropriate mutant. The experiment with 5,5-difluorocamphor led to exclusive hydroxylation at C-9.[12] To the best of our knowledge, experiments with perfluorinated substrates have not been carried out.

4.4. DFT/MM Studies of C=C Epoxidation by Cpd I of P450

Alkene epoxidation by P450 and other Cpd I reagents has many of the puzzling mechanistic features that have been discussed for C–H hydroxylation.[8,301] Several experimental observations, such as occasional *cis/trans* isomerization,[302] sensitivity of the reaction to the donor capability of the olefin,[303] production of aldehydes,[302,304] and formation of heme-alkylated products ("suicidal complexes"),[10,18] indicate that the mechanism involves intermediates which are very likely of different nature. Indeed, the proposed mechanistic schemes[8,301] suggest the incursion of radicals, cations, and radical cationic intermediates. Theory was used to shed light on these mechanistic puzzles.

DFT(B3LYP)-only calculations have been carried out on the epoxidation of ethene,[36,305] propene,[150] cyclohexene,[277] styrene,[306] and 5-methylenylcamphor.[26,307] Most of these studies have been reviewed already[32] and will not be discussed here in detail; we shall only mention the essential features. The reaction profile is very similar to the one found for C–H hydroxylation; the mechanism follows a nonsynchronous MSR with competing HS and LS pathways and a few electromeric varieties, as shown in Scheme 5. An initial C=C bond activation step proceeds via nonsynchronous transition states, to the iron alkoxy radical species, $^{4,2}\text{I}_{rad}$ (i.e., PorFeO-CR'H-CRH$^{\bullet}$), which are of the I_{rad}(III) and I_{rad}(IV) varieties. In addition, there are cationic situations, I_{cat} (e.g., PorFeO-CR'H-CRH^{+}), that are generated by electron transfer from the carbon radical moiety into the heme. In all the cases, the LS species, be they radicals or cations, collapse in a barrier-free manner to the epoxide with conserved stereochemistry of the alkene, while all the HS state electromers give rise to epoxide with scrambled stereochemistry or to aldehyde and suicidal byproduct (^{4}AC, ^{4}SC). The choice depends on the barriers for epoxide-ring closure; thus, $^{4}\text{I}_{rad}$(IV) with the lowest ring closure barrier gives epoxides with scrambled stereochemistry, while $^{4}\text{I}_{rad}$(III) with the highest ring closure barrier crosses (and/or mixes) with the HS cationic states and leads to the aldehyde complex and the suicidal complex.

Scheme 5. DFT-Only Computed Mechanistic Schemes of Alkene Epoxidation by Cpd I and of Byproduct Formation in the High-Spin (HS) and Low-Spin (LS) Manifolds
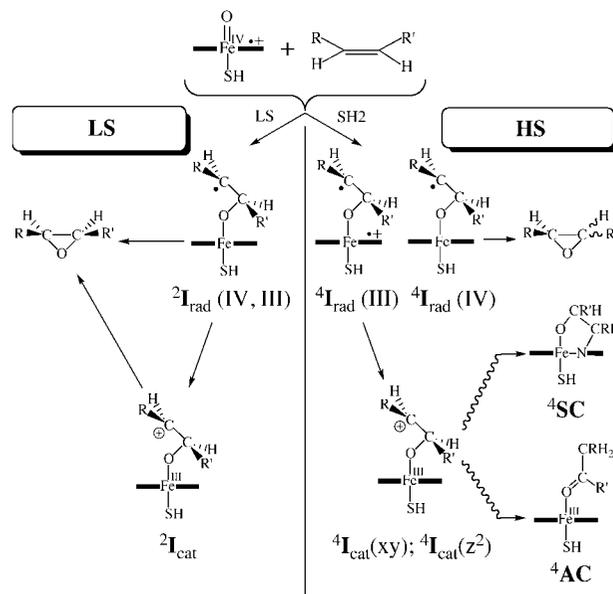
To complement the picture, two DFT(B3LYP)/MM studies of C=C epoxidation have been carried out by the Jerusalem group, using propene and cyclohexene as substrates.[57,277] The two substrates were chosen because they have been used amply in P450 chemistry.[194,308] The focus of the study was on the mechanistic details of epoxidation and the effects of the protein on these mechanisms; no attempt was made to explore the side product formation, as was done in the DFT(B3LYP)-only studies. In the following we discuss only the reactions of cyclohexene, since propene exhibits essentially similar trends.

Figure 38a shows the energy profiles in the gas phase.[57,277] The mechanism displays MSR whereby different electromers and spinomers participate in two phases, C=C bond activation followed by ring closure. The gas-phase profile reveals rather small bond activation barriers, which increase by 2–3 kcal mol^{-1} with the larger basis set (B2W). ZPE corrections are negligible, and we can assume that this will also be the case in the protein. As expected, the ^{2,4}TS$_C$ transition states are essentially of Fe^{IV} character and lead to ^{2,4}I(IV) intermediates. Using TDDFT, ^{4}TS$_C$(III) was found to lie ca. 15 kcal mol^{-1} above the ^{4}TS$_C$(IV) species. The ring closure exhibits barrier-free LS processes and a moderate barrier for the ^{4}I(IV) species. The ^{4}I(III) species was located and had a much higher barrier to ring closure, as expected from results for other model systems.[26,37,306,307]

The situation within the protein is shown in Figure 38b, which depicts the DFT(B3LYP)/MM energy profile for one of the snapshots (Snap-5, others are similar) selected from the MD trajectory, as well as the mean relative energies for all snapshots. Since cyclohexene is a symmetric molecule, one might expect similar bond activation barriers for an attack on either carbon atom of the double bond C$^{\alpha}$=C$^{\beta}$. However, in the protein environment, these two carbon atoms are no longer equivalent, and therefore, the reacting carbon atom in each snapshot was selected as the one having the shortest C···O distance in the initial complex.

Inspection of Figure 38b shows that the lowest-energy transition states are ^{2}TS$_C$(III) and ^{4}TS$_C$, with the latter having mixed Fe^{IV} and Fe^{III} character (as judged from the spin density). However, in other snapshots (Snaps-X, 1, 2, and
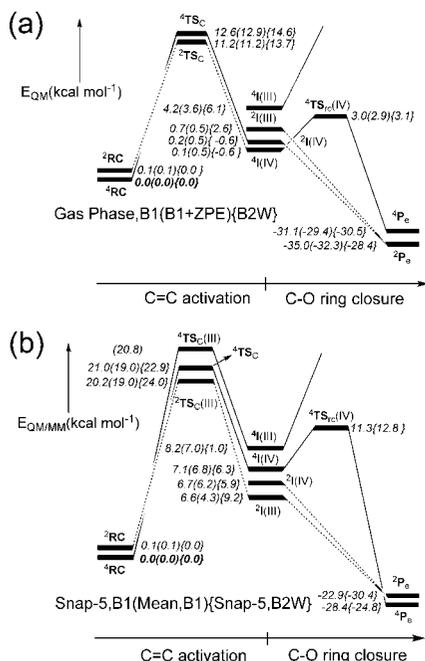
Figure 38. (a) Gas-phase DFT(B3LYP)-only energy profiles during cyclohexene epoxidation, using the R1 model (Por,SH; C_6H_{10}); energy data are given for B1 (B1+ZPE) {B2W}. (b) DFT(B3LYP)/MM results for cyclohexene epoxidation using the R1 model (Por,SH; C_6H_{10}). Each data line contains relative energies for Snap-5,B1 (Mean,B1) {Snap-5,B2W}, where the second datum refers to the mean value of seven snapshots (see text).[57,277]

6), $^4\mathbf{TS}_C$ is of Fe^{III} type; this underscores the fact that the electromeric states tend to be very close in the protein. TDDFT calculations were performed on all transition states and confirmed that the quartet transition states are the ground state in each snapshot. Similar to the gas phase, in the protein too, the ring-closure phase involves the two electromers in two spin-state varieties; the LS species exhibit a barrier-free process, while the HS ones have significant barriers, and the longest lived intermediate is the $^4\mathbf{I}(III)$ species. This confirms the conclusions derived from DFT(B3LYP)-only studies[306] that the stereorandomization (*cis* and *trans*) of the epoxide is due to $^4\mathbf{I}(IV)$ and $^4\mathbf{I}(III)$, while the suicidal and aldehyde complexes will be nascent from the $^4\mathbf{I}(III)$ type intermediate (see Scheme 5).

The bond activation barriers in the enzyme (Figure 38b) are seen to be significantly higher than those in the gas phase (Figure 38a), and this is reflected in the corresponding TS energies. The transition state geometries for the bond activation phase are depicted in Figure 39 for both the gas phase and the protein. In all these structures, the cyclohexene assumes an upright position; the parallel position was not explored because the corresponding initial $^{2,4}\mathbf{RC}$ complexes were higher in energy. Figure 39 displays structures typical for epoxidation TSs,[37,150] showing elongation of the C=C and Fe—O bonds and shortening of the Fe—S bond. The gas-phase species have earlier "reactant like" structures compared with the in-protein species, i.e., longer C···O distances, shorter C=C distances, and longer Fe—S distances. This shift of the transition state by the protein is no doubt the origin of the increased barriers within the protein. As discussed before,[150] the electromeric nature of the species is apparent by comparing $^4\mathbf{TS}_C$ and $^4\mathbf{TS}_C(III)$; the latter has shorter C···O bonds and longer Fe—O and Fe—S bonds, caused by the electron transfer from the $\pi_{C=C}$ orbital to the antibonding π^*_{xz} orbital of the Fe=O moiety (see Figure 26).

The corresponding snapshot-averaged geometrical properties of $^4\mathbf{TS}_C$ and $^4\mathbf{TS}_C(III)$ are similar to those of Snap-5.

Finally, we note that, as in the camphor hydroxylation study,[61] the lowest barriers in the protein are again found for Snap-X, which includes the Wat$_{903}$ molecule. This barrier is invariant to whether Wat$_{903}$ is part of the QM or MM region, which indicates that the influence of Wat$_{903}$ is largely electrostatic, as noted before. For the mean barrier over all snapshots, the energy-lowering effect of Wat$_{903}$ amounts to 4.0 (3.5) kcal mol^{-1} for B1 (B2W), thus once again highlighting the biological role of water as a catalytic species. This effect is somewhat less prominent in the allylic hydroxylation case (see above), suggesting that the catalytic role of water is reaction and substrate dependent.

4.5. DFT/MM and DFT-Only Studies of Aromatic Hydroxylation by Cpd I of P450

Arene hydroxylation by cytochrome P450 enzymes still poses tantalizing questions, concerning both the reaction mechanism and the toxicity of some of the metabolites.[10,12] The reaction yields typically three types of products, phenol, cyclohexenone, and arene oxide, which are shown in Scheme 6 using the reaction of benzene as an example.

The P450 mediated conversion of arene to phenol (**P**) has traditionally been viewed as a means of detoxification and excretion of the toxic arene by the biosystem. By contrast, arene oxide (**AO**) is carcinogenic and mutagenic through its interaction with, e.g., DNA and other cellular material.[309−312] The relationship between the two products became intriguing, when mechanistic investigations[311,312] led to the conclusion that the arene oxide is an intermediate in this reaction and the phenol is one of its byproducts. Further studies indicated that arene oxide is an obligatory intermediate in arene hydroxylation.[313,314] However, in the course of time, new evidence appeared that questioned these conclusions and served to propose alternative pathways proceeding through radical and/or cationic Meisenheimer intermediates.[10,12,315,316] However, based on the detailed analysis of Ortiz de Montellano,[10] it is fair to say that as yet there is no consensus mechanism that is unequivocally established or that can rule out all others. Scheme 7 summarizes the main mechanistic hypotheses for arene hydroxylation by Cpd I.

A universal feature of arene hydroxylation[10,12,311,312] is the migration of the substituent from the site of hydroxylation to the adjacent carbon (so-called "NIH-shift"). The substituent migration results in a partial skeletal retention of the original hydrogen of the activated C−H bond. This observation leaves in Scheme 7 those hypotheses that account for the NIH-shift. For clarity we labeled in the scheme the migrating hydrogen atom with D to identify the position of this atom in all intermediates.[12] In mechanism a, everything is channeled through the arene oxide,[312] which is postulated to be the primary product of the enzymatic reaction. In a subsequent nonenzymatic step, the epoxide ring opens to the zwitterionic species, which can further react in two different manners: A hydrogen/deuterium exchange will produce phenol (**P**, **P′**) whereas a 1,2-hydride (deuteride) shift to the carbocationic center (NIH-shift) will give rise to cyclohexenone (**K**), which can, in turn, enolize to the phenol.[10,12,311,312] Mounting experimental evidence, of various kinds, has revealed, however, that phenol and ketone formation is unlikely to proceed through arene oxide.[315,316] Hence, in the alternative mechanistic hypothesis b, all products are generated from the tetrahedral intermediate σ-complexes; the
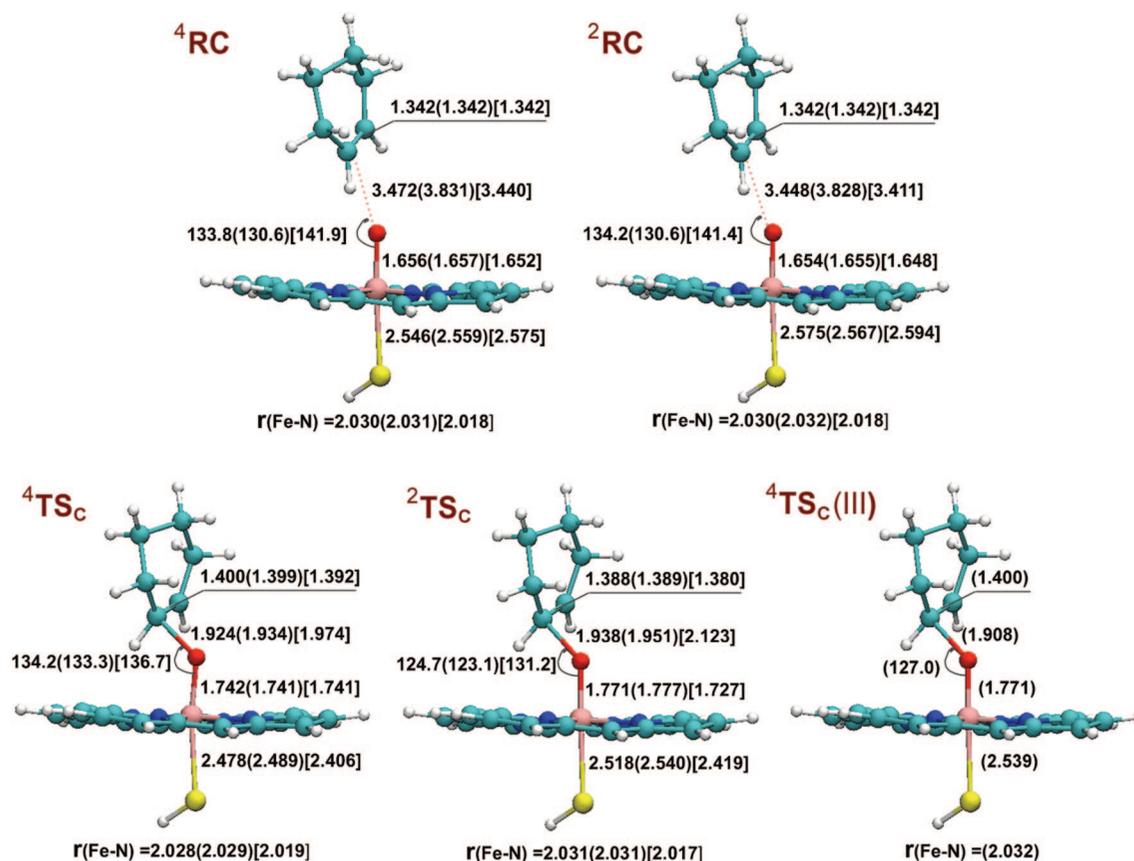
Figure 39. DFT(B3LYP/B1)-only and DFT(B3LYP/B1)/MM optimized key distances (Å) and angles (degree) of the stationary points during cyclohexene epoxidation (see Figure 38). Each data line reports three values, Snap-5 (mean) [gas-phase]. The mean values for the eight snapshots studied are always found to be rather similar to those for Snap-5 (see text).[57,277]

Scheme 6. Observed Products during Arene Hydroxylation by P450
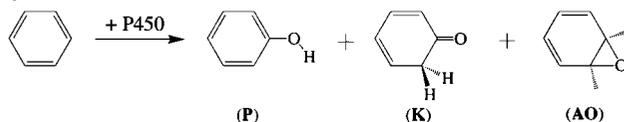
phenol and ketone are formed directly from the intermediate cationic σ-complex (**σ-C**$^+$) (Scheme 7),[315−317] while arene oxide formation is suggested[316] to transpire from the radical σ-complex intermediate (**σ-C**$^\bullet$). Parts of scheme b have gained indirect support from kinetic isotope effect measurements,[316] regiochemical studies,[316] frontier orbital theoretical arguments, and local density Xα type calculations that ruled out concerted epoxidation as the initial activation step of substituted benzene by Cpd I.[318,319] A related mechanism c that involves an initial electron transfer,[6,10,12,320] followed by collapse to the σ-complexes, was suggested too, but the existence of a free (diffusive) radical cation of the arene has not been demonstrated yet.[10,12,315] The role of charge transfer from the arene to the iron porphyrin has nevertheless been inferred from kinetic isotope effects, e.g., for the *meta* hydroxylation of chlorobenzene.[315] Clearly, the catalysis of arene hydroxylation by P450 offers a rich and complex mechanistic puzzle. Despite the support accumulated thus far for the σ-complex based mechanism b, there still exist counterarguments and compelling evidence for the intermediacy of benzene oxide, as in mechanism a.[10,12] Furthermore, the postulated pathways that lead to arene oxide vis-à-vis phenol/ketone products are still under debate. Is the arene oxide intermediary tenable? Are there really both radicalar

and carbocationic σ-complexes?[315] And if they exist, how precisely do they rearrange to yield ketone and phenol products?

Initial DFT(B3LYP)-only calculations, by de Visser and Shaik,[38] and subsequently by Bathelt et al.,[33,34] addressed the reaction mechanism and concluded, in agreement with the mechanisms proposed by Korzekwa et al.[315] and Rietjens et al.,[316] that the process involves initial electrophilic attack on the π-system of benzene, to form Meisenheimer complexes of hybrid character, i.e., σ-C$^+$ with a substantial admixture of σ-C$^{\bullet}$[38] (see mechanism b in Scheme 7). Incorporation of bulk polarity and amidic H-bonds to the thiolate[38] was found to increase the cationic nature of the intermediate, raising the positive charge on the benzene moiety to >0.5. These studies[33,34,38] also found that the lowest pathway occurs preferentially via the doublet spin state. Bathelt et al.[34] demonstrated that the same LS mechanism operates for an extended series of arenes, and they derived useful structure−reactivity correlations. There was also consensus[33,34,38] that the σ-C$^+$ Meisenheimer complex leads to all the products in Scheme 6, and de Visser and Shaik[38] further demonstrated that the epoxide product (**AO**) is a dead-end product that does not transform to phenol (**P**) or ketone (**K**). Many of these features were confirmed in the most recent DFT-only study of 17 arene hydroxylation reactions.[172]

Theory also revealed some novel aspects of the mechanism, which are shown in Scheme 8. Thus, de Visser and Shaik[38] found that, starting from the hybrid ^{2}σ-C$^+$/^{2}σ-C$^\bullet$ Meisenheimer complex, the oxygenated benzene moiety transfers a proton from the *ipso* carbon to one of the nitrogen atoms of the porphyrin ring to form the protonated porphyrin intermediate, ^{2}**I**$_{NH}$ (path a in Scheme 8). This is possible since

Scheme 7. Alternative Mechanistic Hypotheses for Arene Oxidation by P450 (Reprinted with Permission from ref 38. Copyright 2003 American Chemical Society)
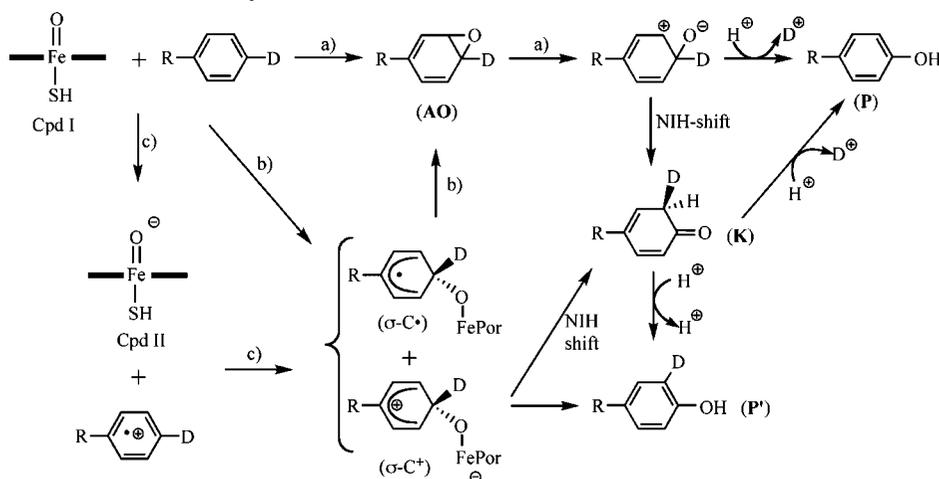
Scheme 8. Theoretically Derived Mechanisms of Phenol (P) and Ketone (K) Formation from the Hybrid Meisenheimer Complexes in Side-on and Face-on Conformations[33,34,38]
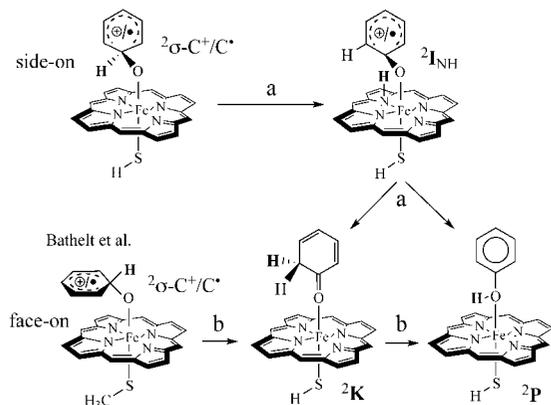
The original "H" that migrates during the processes is indicated in bold.

Figure 40. DFT(B3LYP/B1)-only and DFT(B3LYP/B1)/MM optimized transition state structures and barriers (in kcal mol^{-1}), for the (a) side-on and (b) face-on conformations. Reprinted with permission from ref 60 using new data from the author (J. N. Harvey). Copyright 2007 American Chemical Society.

the nitrogens in the porphyrin are basic. The intermediate then reprotonates the oxo group to form phenol (^{2}**P**) and the *ortho* carbon to produce the ketone (^{2}**K**). In contrast to this proton-shuttle mechanism, Bathelt et al.[33,34] found that the Meisenheimer $^2\sigma$-C$^+$/C$^{\bullet}$ intermediate undergoes a 1,2-H shift (path b in Scheme 8) from the *ipso* to the *ortho* carbon and generates ^{2}**K**, which can subsequently undergo further rearrangement to the phenol, ^{2}**P**.

Thus, while the two mechanisms in Scheme 8 are different, they both account for the NIH shift, whereby the original "H" from the arene ends on either the ketone or phenol products. The reason for the two different mechanisms is obvious from Scheme 8, where it is seen that, in the study of de Visser and Shaik,[38] the phenyl plane is perpendicular to the porphyrin (side-on) and the *ipso* C—H bond is in direct contact with one of the nitrogen atoms of the porphyrin, thus leading to an instantaneous proton transfer and reshuttle mechanism as described in mechanism a in Scheme 8. By contrast, in the initial study of Bathelt et al.[33] the phenyl ring is parallel to the porphyrin plane (face-on), and the *ipso* C—H bond points away, thus being prone to skeletal rearrangement by 1,2 hydride shifts, as in mechanism b in Scheme 8. In a subsequent study, Bathelt et al.[34] confirmed the proton-shuttle mechanism a for the NIH shift, and the

results of the two groups converged. Indeed, in the protein pocket both conformations may exist and contribute to the NIH shift.

4.5.1. DFT/MM Studies of Benzene Hydroxylation by CYP2C9

The mechanism was recently reinvestigated by Bathelt et al.[60] using DFT(B3LYP)/MM calculations for benzene hydroxylation by CYP2C9, starting from an X-ray structure (pdb: 1OG2), which lacked a substrate and hence could be used to accommodate benzene without conformational bias due to the presence of another substrate. After a short MD equilibration, two snapshots were selected for each of the benzene conformations of Scheme 8, the face-on and the side-on structures. The calculations were restricted to the LS state.

Figure 40 depicts transition state geometries for the two bond activation transition states and provides additional data as well as comparison with the gas-phase structures of Bathelt et al.[60] The calculations show that the two conformations, face-on and side-on, are competitive inside the protein, as postulated before.[32] The computed barriers of 18.1–21.0 kcal mol^{-1} are in good accord with experimental data in CYP2E1,[321] obtained by conversion of the rate constants to free energy barriers. The barriers in the protein are slightly larger than those in the gas phase, and the corresponding
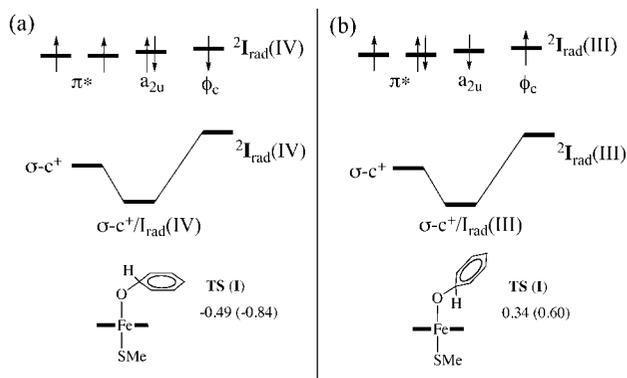
Figure 41. VB mixing diagrams showing the cationic/radicalar hybrid character, as well as the generation of (a) negative and (b) positive spin density on the phenyl moiety in the radical. DFT(B3LYP)/MM computed spin densities[60] in the transition state (**TS**) and intermediate (**I**) are shown beneath the mixing diagrams.

TSs occur somewhat earlier (see the C···O distances). It was noted[60] that the Meisenheimer complexes are destabilized in the protein relative to the gas phase, and hence, this behavior follows the Hammond principle.

Another interesting feature is the electronic structure of the transition states and the Meisenheimer hybrid $^2\sigma$-C$^+$/C$^\bullet$ intermediate. Thus, while, in both the DFT-only and DFT/MM calculations, the transition states and intermediates have a hybrid cationic/radicalar nature, with spin density as well as charge (+0.3 to +0.57) residing on the benzene moiety, the computed spin density was positive in both face-on species, while it was positive in one and negative in the other snapshot of the side-on species. This hybrid nature and the sign of spin densities on the phenyl moiety can be understood by reference to the orbital occupancy diagrams for the Meisenheimer intermediates in Figure 41. Thus, in Figure 41a, we show a mixing of the cationic state with a radicalar state, corresponding to ^{2}**I**(IV). We recall from Figure 26 that this radicalar state arises from a shift of a spin-up electron to the a$_{2u}$ orbital of the heme, thus resulting in a negative spin density on the benzene moiety (see the ϕ_C orbital); the mixing of this radicalar state with the cationic state will lead to some negative spin density on the benzene moiety in the hybrid state. This is precisely the spin density computed in the transition state and the intermediate Meisenheimer complex, shown underneath the mixing diagram. By comparison, in Figure 41b, we show a similar mixing but now with the ^{2}**I**(III) electromer, which by reference to Figure 26 arises by shifting a spin-down electron to one of the π*(FeO) orbitals of the heme and, hence, generating now a positive spin density on the benzene moiety, as shown underneath the mixing diagram. Clearly, therefore, benzene hydroxylation may involve MSR as well with different electromeric states, which may be addressed in future studies.

Finally, the DFT(B3LYP)/MM calculations support the proton-shuttle mechanism with an N-protonated intermediate (^{2}**I**$_{NH}$) as a viable pathway for phenol production. Figure 42 shows in a schematic manner the various barriers nascent from the collapse of the side-on (SO) and face-on (FO) Meisenheimer $^2\sigma$-C$^+$/C$^\bullet$ intermediates to the various products. Considering these barriers, it seems that the side-on complex leads preferentially to epoxide and to phenol, the latter via the N-protonated porphyrin species (the proton shuttle mechanism in Scheme 8). The ketone is generated from both face-on and side-on pathways with approximately the same barriers.
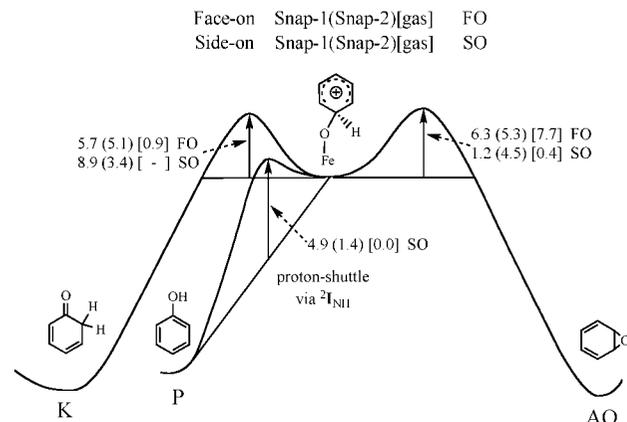
Figure 42. DFT(B3LYP)/MM and DFT(B3LYP)-only computed barriers (kcal mol^{-1}) for the collapse of the face-on (FO) and side-on (SO) Meisenheimer complexes to the three products of arene oxidation by P450.[60] The energy profiles are featureless; note, however, that the SO path proceeds to the phenol product via ^{2}**I**$_{NH}$.

Scheme 9. Products Formed during Oxidative P450 Dehalogenation of Polyhalogenated Benzene[165]
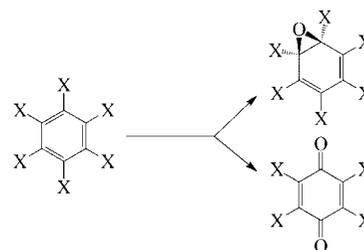
In conclusion, the DFT(B3LYP)/MM study of Bathelt et al.[60] recovers the basic features of the DFT(B3LYP)-only studies,[33,34,38] while accounting in detail for the role of the protein.

4.5.2. DFT-Only Studies of Oxidative Dehalogenation of Perhalogenated Benzenes by P450

Polyhalogenated aromatics are persistent hazardous environmental pollutants.[322,323] While partially halogenated arenes are degradable by some organisms,[324−327] more extensively halogenated species are degraded at a very slow rate.[328] P450 enzymes have the potential to oxidatively metabolize halogenated benzenes, so much so, that a considerable effort has been put into engineering of bacterial P450s with enhanced dehalogenation kinetics.[329−335] P450-catalyzed dehalogenation generally produces 1,4-diquinones and arene oxides, as shown in Scheme 9. The mechanism of formation of these and other species during P450 dehalogenation has been a topic of extensive debate.[336]

In order to elucidate the mechanism of dehalogenation and to identify potential primary products, Hackett et al.[165] performed DFT-only calculations of a few polyhalogenated benzene derivatives. The authors used the PBE functional with the resolution-of-identity approximation to expedite geometry optimization, hence RI-PBE, and complemented this by single-point energy calculations using B3LYP. Despite the different geometry of Cpd I compared with B3LYP, the triradicaloid nature of Cpd I is apparent even at the RI-PBE level, so that the mechanisms deduced by this study are most likely reliable.

The reaction of Cpd I with hexachlorobenzene (HCB) is shown in Figure 43. It is seen that the reaction is concerted,
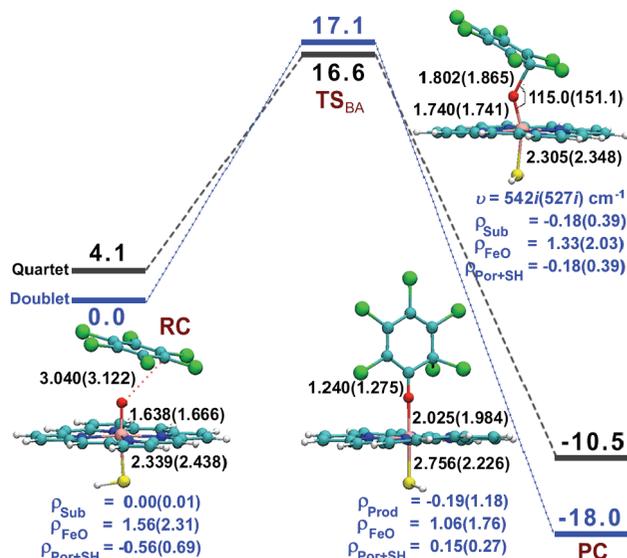
Figure 43. DFT-only calculations of the first step in the oxidative dehalogenation of hexachlorobenzene by P450 Cpd I. Numerical results for bond lengths (Å), angles (degree), and spin densities ρ refer to the doublet (quartet) state. The energy profiles (kcal mol^{-1}) come from RI-PBE/TZVPP calculations at RI-PBE/TZVP,SV(P) optimized geometries. Reprinted with permission from ref 165. Copyright 2007 American Chemical Society.

involving π-attack of Cpd I on the hexachlorobenzene, with concomitant 1,2-Cl shift. The transition states are virtually degenerate for the doublet and quartet spin states, but in the product complex, the doublet state is significantly lower than the quartet at the RI-PBE level, which tends to overstabilize low-spin states. The final organic product in the reaction is 2,2,3,4,5,6-hexachloro-3,5-cyclohexadienone.

Even though the Meisenheimer intermediates are not formed in this concerted process, inspection of the spin density (ρ) data for the transition states (^{2,4}**TS**$_{BA}$) reveals a similar electronic structure as during benzene hydroxylation (see above). The negative spin density on the HCB moiety of ^{2}**TS**$_{BA}$ indicates an Fe^{IV} type species, and the small charge development on the same moiety reveals a primary radical nature, and not cationic or hybrid ones, as in the benzene hydroxylation case.

The concerted nature of the transition states depends however on the halogen, and in the case of hexafluorobenzene (HFB), shown in Figure 44, there are two well separated TSs: a bond activation **TS**$_{BA}$ species, going to a Meisenheimer-type intermediate, $\mathbf{I}_{\sigma\text{-}C^{\bullet}}$, followed by a 1,2-F-shift **TS**$_{shift}$ species. The final product is a 2,2,3,4,5,6-hexafluoro-3,5-cyclohexadienone. The similarity to the benzene hydroxylation is apparent; the spin density on the HFB moiety in the $^{2}\mathbf{I}_{\sigma\text{-}C^{\bullet}}$ intermediate is large and negative (-0.50), while the charge is small, thus indicating a mostly radicalar mechanism nascent from the Fe^{IV} electromer (see Figure 41a). Interestingly, the bond activation barrier for the reaction of Cpd I with HFB is substantially lower (5.3 (RI-PBE) and 13.8 kcal mol^{-1} (B3LYP)) than that with HCB. While Hackett et al.[165] did not account for this difference, we propose tentatively that the effect is associated with the well-known fluoro effect of stabilizing radical species.[337]

2,2,3,4,5,6-Hexahalo-3,5-cyclohexadienones, which are the primary products in Figures 43 and 44, have not been observed experimentally but could well be key intermediates for the formation of tetrahaloquinones during P450 oxidation. The experimentally observed product from hexachloroben-

zene is pentachlorophenol. Hackett et al.[165] suggested that this product is obtained from 2,2,3,4,5,6-hexacholoro-3,5-cyclohexadienones by electron capture from the reductase followed by a loss of chloride ion, another reduction, and formation of pentachlorophenol as shown in Scheme 10. Alternatively, the reduction can arise from nonenzymatic soluble biological factors in aqueous solution. Hackett et al. have verified that the 2,2,3,4,5,6-hexahalo-3,5-cyclohexadienones have sufficiently large electron affinities to be easily reducible.

Once the pentachlorophenol is formed, another Cpd I-mediated oxidation by π-attack would lead, in a mechanism analogous to the one shown in Figure 43, to 4-hydroxy-2,2,3,5,6-pentachloro-3,5-hexadienone (and the corresponding *meta*- and *ortho*-attack products), as sketched in Scheme 11. At physiological pH, deprotonation of 4-hydroxy-2,2,3,5,6-pentachloro-3,5-hexadienone and Cl^{-} expulsion will predominate and yield the tetrachloroquinone. This mechanism is general, only in that, in the perfluoro case, the F^{-} expulsion requires incorporation of bulk solvation in the calculations.

With mixed chloro$-$fluoro derivatives, the calculations reproduce the experimentally observed preferential oxidation at the fluorinated sites. The calculations predict a barrier that is substantially smaller, by 4$-$6 kcal mol^{-1}, for the π-attack on the fluorinated site compared with a chlorinated site. Once again, this intriguing fact may reflect effects of radical stabilization by fluorine.[337] However, confirmation of this effect will require a close analysis of the spin density distribution on fluorine vs chlorine to assess the role of hyperconjugative interactions between the developing radical in the benzenic moiety and the COF vs COCl moieties in the **TS**$_{BA}$ species.

In summary, the study of Hackett et al.[165] on oxidative dehalogenation and the above studies on benzene hydroxylation demonstrate that activation of aromatic substrates generally proceeds by π-attack on the aromatic ring. The low barriers for oxidative defluorination suggest that P450s can be very effective in detoxification of fluorinated aromatics. Of course, without further DFT/MM studies or at least DFT-only studies augmented with bulk polarity effects, this conclusion should be considered tentative.

4.6. DFT-Only Studies of *N*-Dealkylation of *N,N*-Dimethylaniline Derivatives

The mechanism of *N*-demethylation of *N,N*-dimethylanilines (DMAs) by cytochrome P450 is a highly debated topic in mechanistic bioinorganic chemistry. Scheme 12 describes the two rival mechanistic hypotheses. Thus, one mechanism labeled as (i) in Scheme 12 suggests that the reaction starts by the usual H-atom abstraction (HAT) reaction, followed by rebound to form the carbinolamine, which decomposes subsequently in a nonenzymatic reaction to give rise to formaldehyde and a dealkylated amine.[338–342] The alternative mechanistic hypothesis, labeled as (ii), starts with a single electron transfer (SET) from the amine to Cpd I,[343–354] followed by proton transfer to the Cpd II species and proceeding with a rebound to form the carbinolamine, which then decomposes as outlined above.

The HAT-first mechanism has been supported by kinetic isotope effect (KIE) profiles, which showed that the KIEs measured for P450 oxidation of a series of *para*-substituted *N,N*-dimethylanilines (*p*-X-DMAs), by a few isoforms, correlate linearly with the KIEs measured for the same *p*-X-
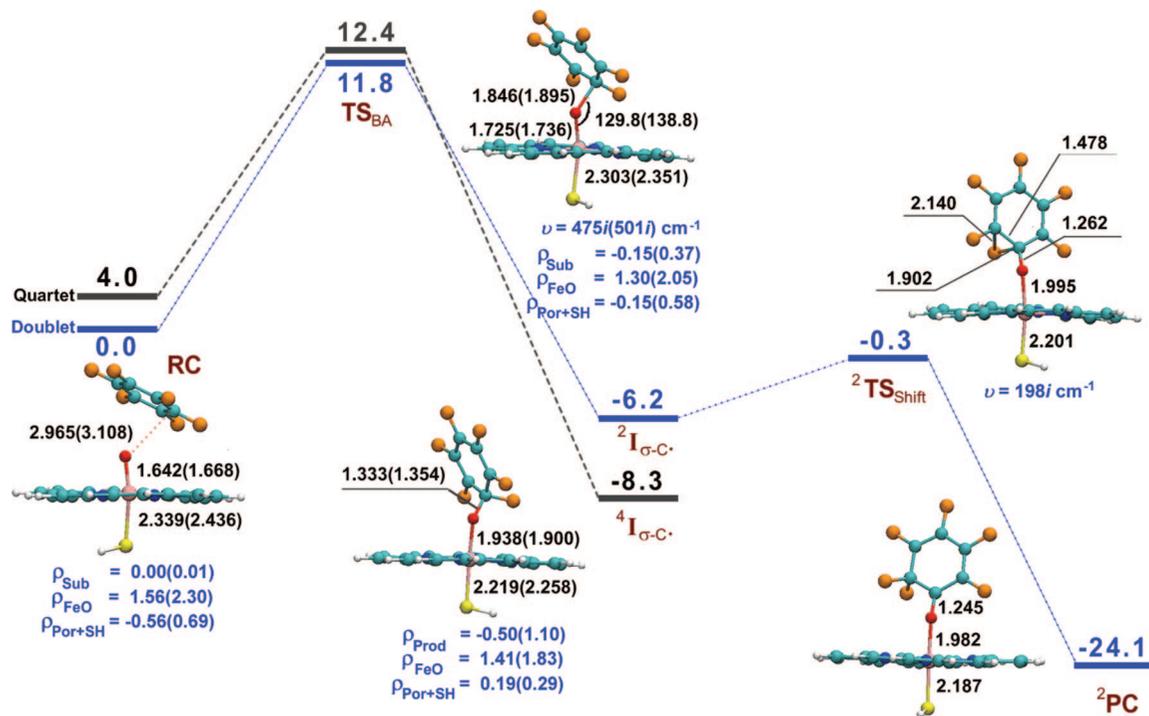
Figure 44. DFT-only calculations of the first step in the oxidative dehalogenation of hexafluorobenzene by P450 Cpd I. Numerical results for bond lengths (Å), angles (degree), and spin densities ρ refer to the doublet (quartet) state. The energy profiles (kcal mol^{-1}) come from RI-PBE/TZVPP calculations at RI-PBE/TZVP,SV(P) optimized geometries. Reprinted with permission from ref 165. Copyright 2007 American Chemical Society.

Scheme 10. Proposed Mechanism for the Formation of the Ferric-Phenolate Complex by Electron Transfers from the Reductase, Followed by Cl$^-$ Expulsion[165]
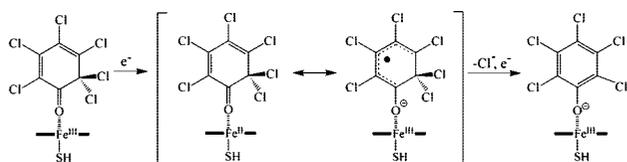
Scheme 11. Proposed Mechanism for the Formation of the Tetrachloroquinone Product, Starting with Cpd I Oxidative Dehalogenation of Pentachlorophenol and Followed by Nonenzymatic Elimination of HCl[165]

Scheme 12. Mechanistic Hypotheses for *N*-Dealkylation of Amines: Mechanism (i) Involves Hydrogen Atom Transfer (HAT) First, While Mechanism (ii) Involves Single Electron Transfer (SET) First[a]
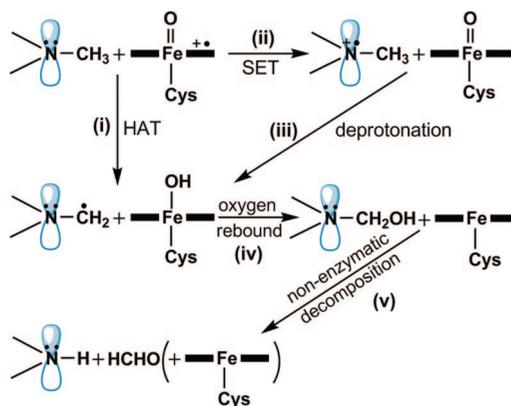
[a] The two mechanisms converge at the carbinolamine decomposition, which is proposed to occur nonenzymatically. Reprinted with permission from ref 167. Copyright 2007 American Chemical Society.

DMAs series in a real HAT reaction with the *t*BuO$^\bullet$ radical.[338–341] In addition, use of clock substrates[355–359] and other experimental criteria[342,360–363] seemed to rule out amine cation radical formation. The SET-first mechanism has been supported by the facts that (a) the KIE values are low and seem atypical of HAT reactions[343,344,346,352,353] and (b) the reaction rates for N-demethylation of *p*-X-DMAs correlate nicely with Hammett substituent parameters (which are

negative)[345,347–351,354,364–367] and with the redox potential $E_{1/2}$ values of the *p*-X-DMAs.[345,347–351,354,364–367]

Since all these pieces of evidence are more or less circumstantial, it was essential to study the *N*-dealkylation by theoretical means. This was done using DFT(B3LYP)-only calculations of the reactions of Cpd I with four *para*-(H, Cl, CN, NO$_2$) substituted DMAs used in the experimental research.[167] An initial study of the unsubstituted DMA itself with Cpd I[166] showed that the SET-first hypothesis, at least for this substrate, is less likely, since the electron-transferred DMA$^{\bullet+}$/Cpd II species are significantly higher in energy than the DMA/Cpd I pair, both at the cluster geometry and at the corresponding TS geometry (without and with inclusion of bulk polarity effects). The second study[167] showed also a
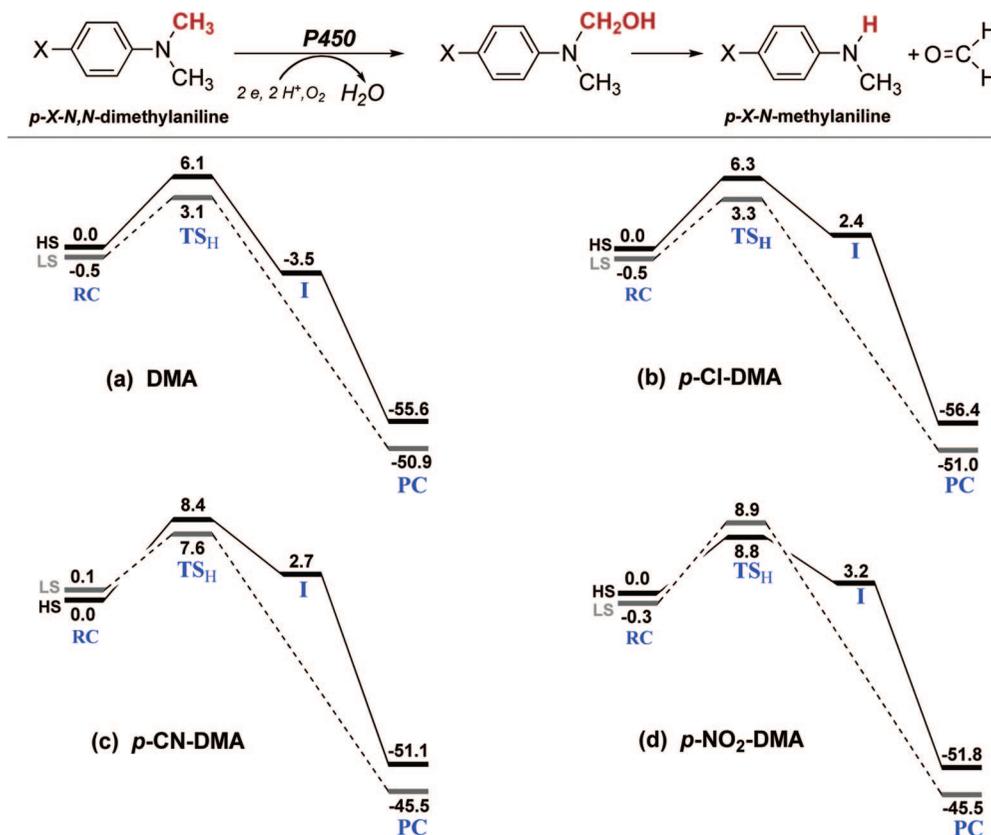
Figure 45. DFT(B3LYP)-only energy profiles (kcal mol^{-1}) for C−H hydroxylation of p-X-DMA (X = H, Cl, CN, NO$_2$). Geometries were optimized with the B1 basis, while the relative energies were determined using a larger basis, LACV3P+(Fe)/6-31+G*(rest), with inclusion of ZPE, bulk polarity (ε = 5.6), and NH···S corrections. Reprinted with permission from ref 167. Copyright 2007, American Chemical Society.

HAT mechanism. In addition, the calculated KIE$_{LS}$ for the low-spin (LS) HAT reaction of the pristine DMA (X = H) agrees closely with experiment. These results then led to the conclusion that all the reactions proceed by HAT-first, and it remained to understand the various structure−reactivity correlations and low KIEs that implied SET-first, using the HAT-first mechanism.

Figure 45 shows the energy profiles for the reactions of Cpd I with the *para*-(H, Cl, CN, NO$_2$) substituted DMAs, at the highest level, with the LACV3P+*(Fe)/6-311+G*(rest) basis set, including bulk polarity (ε = 5.7) and two NH···S amidic H-bonds to the thiolate ligand.[167] Both the LS and HS profiles exhibit a transition state for HAT, ^{2,4}TS$_H$; there are no intermediates, and the "^{4}I" feature in the figure is a shoulder and not a real intermediate. It is seen that the energy difference between the ^{2}TS$_H$ and ^{4}TS$_H$ species depends on the *para* substituent: the gap is largest for p-H and p-Cl, ca. 3.0 kcal mol^{-1}, and smaller in the other cases, 0.8 kcal mol^{-1} for p-CN and 0.1 kcal mol^{-1} for p-NO$_2$. As a whole, there is a switch from a single-state mechanism (SSM) in the p-H and p-Cl case, which will proceed preferentially via the LS state, to TSR in the p-CN and p-NO$_2$ case.

The structures of ^{2,4}TS$_H$ are displayed in Figure 46. They exhibit some interesting trends with regard to the progress along the H-transfer coordinate, C···H···O. In the LS species, ^{2}TS$_H$, the pristine case (p-H) has the shortest C−H and longest O−H bond and is hence the "earliest". As the substituent becomes more electron withdrawing, from p-Cl to p-NO$_2$, the transition state occurs progressively later. In the HS species, ^{4}TS$_H$, the transition states are more advanced along the H-transfer coordinate but exhibit a smaller amount of variation in the series, compared with the LS species.

The calculated KIE values are collected in Table 2, together with the experimental KIE values. The KIE$_{LS}$ values are smaller than the corresponding KIE$_{HS}$ values, and they exhibit a variation consistent with the Melander−Westheimer postulate that the KIE value reflects the position of the TS along the H-transfer coordinate; the earlier the position, the smaller is the KIE—much the same as in the experimental data. The small values for p-H and p-Cl reflect the early nature of the corresponding TSs, as reasoned by Dinnocenzo and co-workers.[338−341] By contrast, the KIE$_{HS}$ values are almost constant, in line with the smaller variation in geometry of these species (Figure 46), and the larger values manifest the more "central" position of the respective TSs, precisely as expected from the Melander−Westheimer postulate. Furthermore, the absolute magnitudes of the KIEs correlate with the reaction barriers and with the extent of the C−H elongation in the TS$_H$ species. It is apparent that the KIE values and the KIE profile are entirely congruent with the HAT nature of the reaction.

Figure 47 shows plots of computed KIE$_{HS,LS}$ values against experimental values, KIE$_{expt}$. Obviously, the KIE$_{HS}$ values do not correlate with the experimental data, while the KIE$_{LS}$ values do so quite well (see Figure 47a and b). Finally, Figure 47c presents the correlation of a blended KIE with the experimental values. The blending coefficient was determined by fitting the KIE differences in the series to the experimental ones.[167] As can be seen, the computational data match experiment perfectly, such that, in the case of p-H and p-Cl with a large energetic preference for the LS state, the reactivity is exclusively LS, and as the electron withdrawing power of the substituent increases and the LS preference decreases, there is a shift to TSR. This blending fits also
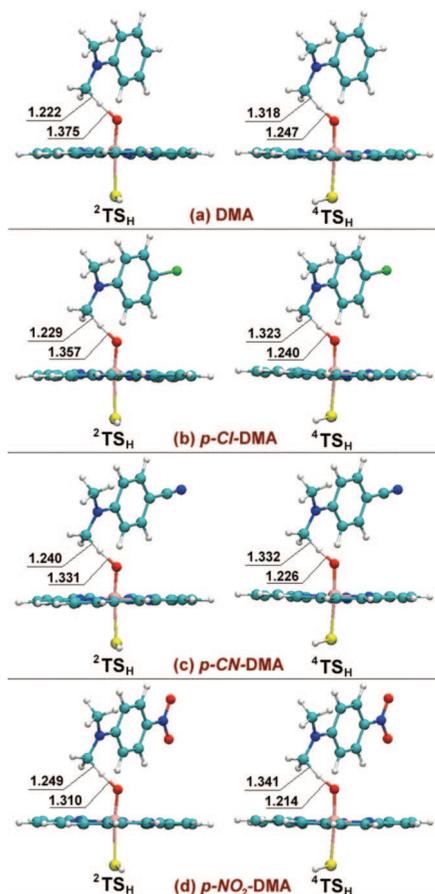
Figure 46. DFT(B3LYP/B1) optimized transition states (Å) for C−H hydroxylation of p-X-DMA (X = H, Cl, CN, NO$_2$). Reprinted with permission from ref 167. Copyright 2007 American Chemical Society.

Table 2. DFT(B3LYP/B1)-Only Computed^a KIE Values for p-X-DMA, along with Experimental Values

	DMA	p-Cl-DMA	p-CN-DMA	p-NO$_2$-DMA
KIE$_{HS}$	5.5	5.4	5.5	5.6
KIE$_{LS}$	3.6	3.8	4.2	4.4
KIE$_{exp}$	2.6	2.8	3.6	4.0

^{a} From ref 167.

nicely with the change in the computed $^4\text{TS}_H-^2\text{TS}_H$ energy gaps in Figure 45, where initially the LS TS is significantly lower and finally becomes virtually degenerate with the HS TS for the electron withdrawing substituents.

Since the KIEs fit a HAT-first mechanism, then, what about the observed Hammett correlations and the correlation with the redox potentials of the DMAs, discussed above? Figure 48a shows the computed barriers for the LS states plotted against the Hammett substituent parameters. The correlation is linear and has a negative slope, in perfect agreement with the experimentally observed negative Hammett slope value. Figure 48b indicates that the LS barriers for HAT correlate nicely with the redox potential $E_{1/2}$ of the p-X-DMAs. Clearly, therefore, the computed barriers for a HAT-first mechanism reproduce the experimental correlations and demonstrate thereby that these correlations cannot serve as a proof for a SET-first mechanism.

Having seen that all the experimental trends are reproduced by the DFT(B3LYP) calculations using a HAT-first mechanism, the question to ask is: what is the mechanistic information of the experimentally observed correlations?
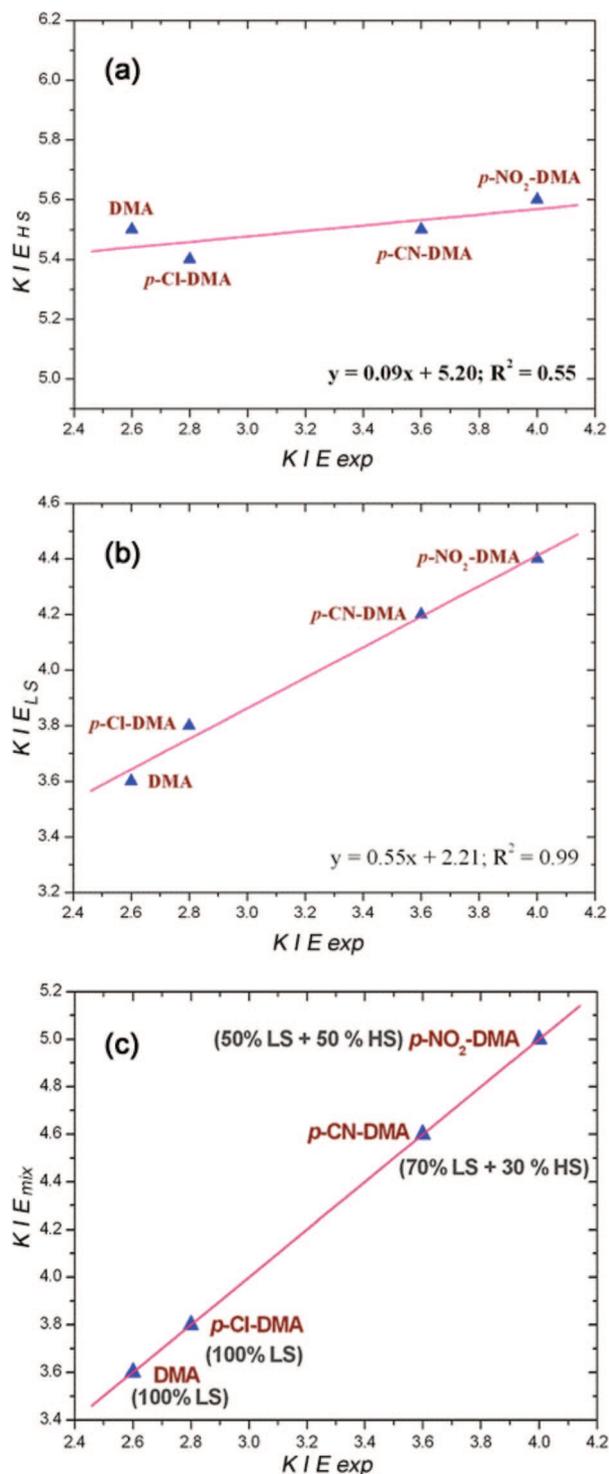
Figure 47. DFT(B3LYP/B1) calculated KIE$_{HS,LS}$ values and experimental data, KIE$_{exp}$, for C−H hydroxylation of p-X-DMA (X = H, Cl, CN, NO$_2$): (a) KIE$_{HS}$ plotted against KIE$_{exp}$, (b) KIE$_{LS}$ plotted against KIE$_{exp}$, and (c) KIE$_{mixed}$ plotted against KIE$_{exp}$. Reprinted with permission from ref 167. Copyright 2007 American Chemical Society.

Some clue is given by Figure 49, which shows linear correlations of the spin density, $\rho(N)$, on the nitrogen adjacent to the CH$_3$ group being activated, for the HS and LS TS$_H$ species, with the Hammett substituent constants. The nitrogen atom accumulates spin density during the reaction because its lone pair can conjugate with the nascent radical on the carbon as the C−H bond is being cleaved and stabilize it. Thus, as the *para* substituent becomes increasingly more electron withdrawing, the nitrogen lone pair is "sucked" by
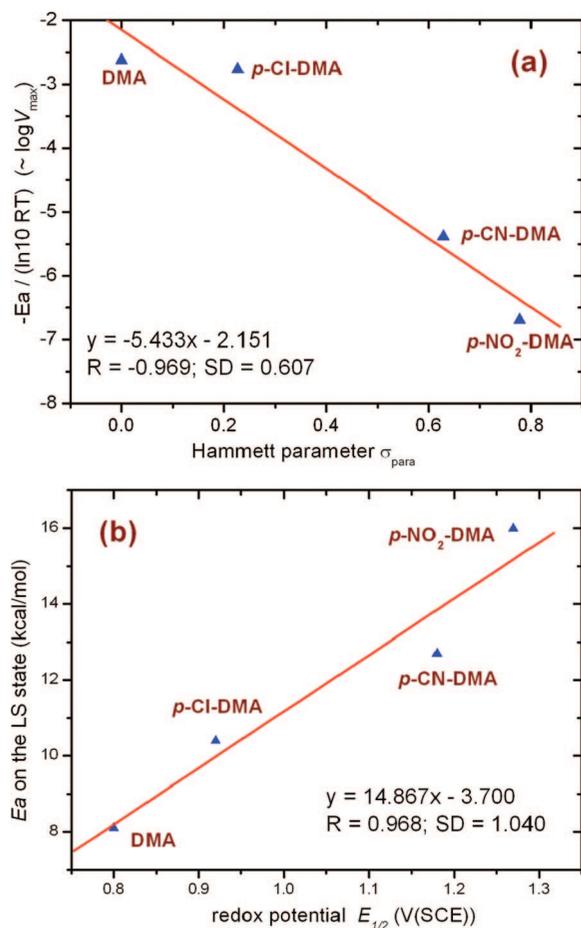
Figure 48. Plots of calculated LS barriers for *p*-X-DMAs against (a) the Hammett substituent parameter of *p*-X and (b) the redox $E_{1/2}$ values of *p*-X-DMAs. Reprinted with permission from ref 167. Copyright 2007 American Chemical Society.

the substituent and the spin density on N decreases. Hence, the *para* substituent X communicates with the TS by changing the accessibility of the nitrogen lone pair to interact with the activated C—H bond, as well as by transferring charge to the Cpd I moiety. The less electron withdrawing the substituent, the more accessible is the lone pair, and the larger is the stabilization of the nascent radical, since more charge will be transferred from the aryl group to the C---H---O(FePor) moiety in the TS. This is the main reason for the good correlation of the barriers with the Hammett substituent constants and with the redox potential of the DMA. All these correlations reflect the "polar" character of the process; it is not an SET process but an HAT process that involves "partial charge transfer character". Indeed, all the cases exhibit a positive charge accumulation on the *p*-X-DMA moiety in the TS.

This stabilization mechanism endowed by the nitrogen lone pair is shown in Scheme 13 using three resonance VB structures (**R**1−**R**3) which represent the formation of a new covalent O—H bond (**R**1), the three-electron resonance stabilization of the $CH_2^{\bullet}$ moiety by the nitrogen lone pair (**R**2), and polar contributions to the TS (**R**3). As the *para* substituent X becomes more electron withdrawing, the resonance structures **R**2 and **R**3 become less important and the TS stabilization diminishes. The correlation with the Hammett substituent constants thus reflects the attenuation of the conjugation between the nitrogen lone-pair and the cleaving C—H bond in the TS by the ring substituent. One can witness in fact this weakening of the conjugation due to

the electron withdrawal of X = CN and NO_2, by inspecting the orbitals of the $\mathbf{TS_H}$ species.[167] The impaired delocalization causes in turn destabilization of the $\mathbf{TS_H}$ species for the more electron withdrawing substituents and requires a higher degree of C—H bond cleavage, and hence also a larger KIE value.

4.6.1. KIE is a Reliable Probe of Spin State Selectivity in Oxidation of DMAs

The correlation of computed and experimental KIE values in Figure 47 as well as the results of the preliminary study for the pristine DMA[166] show that in fact KIE can serve as a reliable probe of the spin state selectivity of Cpd I. Thus, with the pristine DMA and the *p*-Cl-DMA, where the H-abstraction $^2\mathbf{TS_H}$ has a significantly lower energy than the corresponding HS species, the observed low KIE values reflect the LS-only reactivity. As the substituent becomes increasingly more electron withdrawing, the electronic advantage of the LS mechanism diminishes, causing $^2\mathbf{TS_H}$ and $^4\mathbf{TS_H}$ to approach one another in energy, and as a result the observed KIE becomes a blended mixture of KIE_{LS} and KIE_{HS}. The use of KIE as a probe for spin state reactivity has been demonstrated before for the gas-phase activation of norbornane by FeO^+, where the exo-CH and endo-CH bonds exhibited very different KIEs that turned out to correspond to the LS and HS process, respectively.[368] This idea found use in other C—H bond activation processes as well, e.g., in the observation of KIE jumps during the oxidation of DMA by the native Cpd I vis-à-vis Cpd I prepared using the PhI=O oxygen surrogate.[292] As shown in that study, the use of PhI=O leads preferentially to the HS Cpd I species, and when the oxidation barrier is as small as in DMA, the latter gets oxidized by the HS spin state before spin interconversion can occur.

4.6.2. Decomposition of the Carbinolaniline Initial Product

After methyl hydroxylation of DMA, the so formed carbinolaniline decomposes as shown in Scheme 14. The conventional wisdom is that this process occurs in a nonenzymatic manner after the carbinolaniline is released and most likely leaves the active site. The DFT(B3LYP) calculations[167] confirm this supposition; thus, the binding energies of the carbinolaniline to the heme are very small (8.5 kcal mol^{-1} for the doublet ferric complex), whereas the decomposition barriers are much larger ($\geq$30.3 kcal mol^{-1} for the doublet state complex). Furthermore, as shown in Figure 50, the decomposition reaction is assisted by water molecules, which are abundant on the protein surface. Thus, the barrier for decomposition is 29.6 kcal mol^{-1} without water assistance and drops to 14.8 kcal mol^{-1} after the addition of a single water molecule. The reason for this large effect is clear from the structure of the decomposition transition state, $\mathbf{TS_{decomp}}$, in Figure 50. Thus, in the absence of a water molecule, the hydroxyl H migrates to the nitrogen atom via a strained four-membered cyclic structure, while in the presence of a water molecule the water mediates the H-migration via a six-membered cyclic TS which is strain-free. We have not pursued further whether more water molecules will reduce the barrier even more.
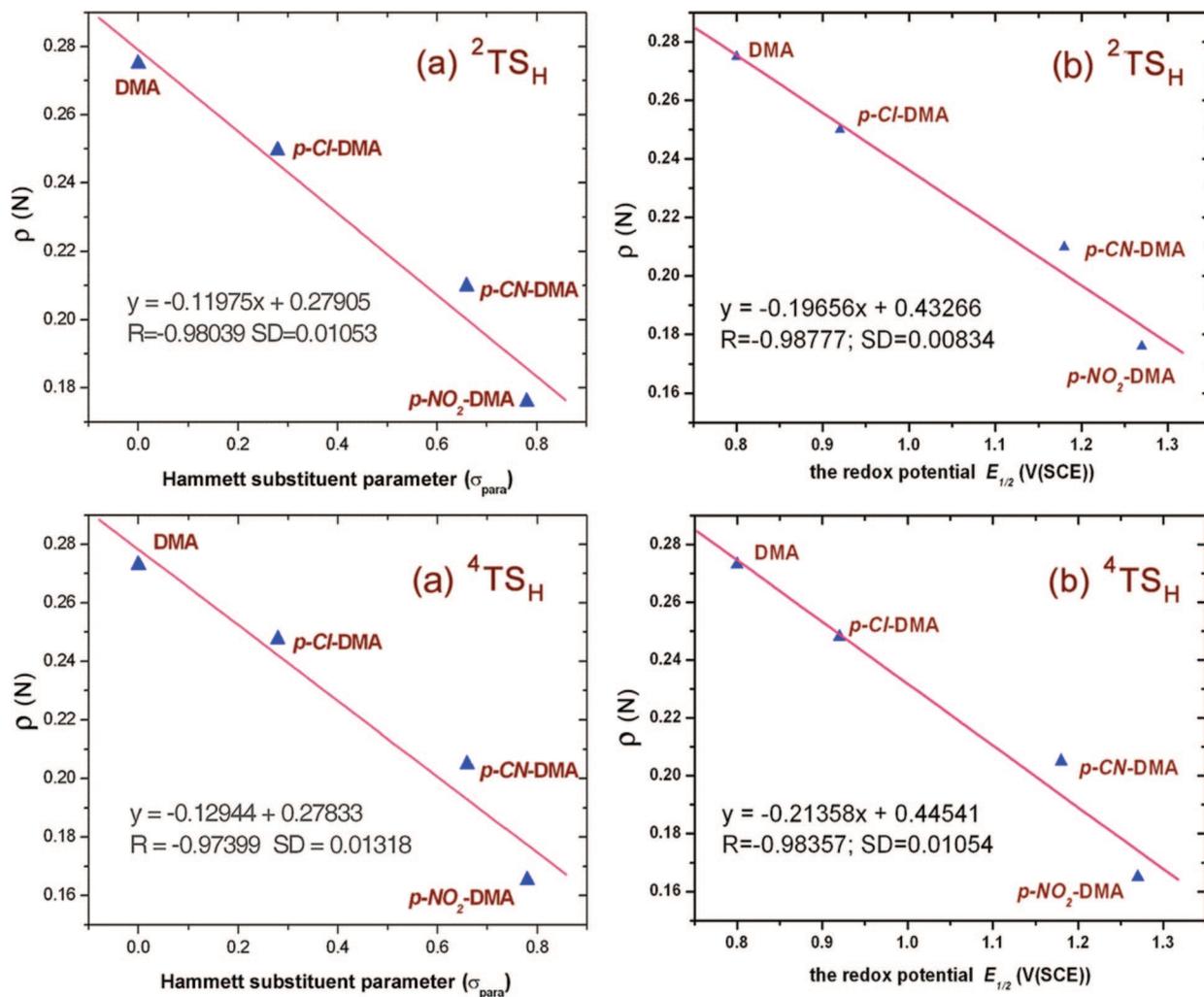
Figure 49. Plots of calculated spin density, $\rho(N)$, on the nitrogen atom adjacent to the activated C−H bond of ^2TS_H and ^4TS_H, against the (a) Hammett substituent parameter of p-X, and (b) redox $E_{1/2}$ values of p-X-DMAs. Reprinted with permission from ref 167. Copyright 2007 American Chemical Society.

Scheme 13. Three of the Resonance Structures That Contribute to the Stabilization of TS_H for C−H Hydroxylation in p-X-DMA[a]
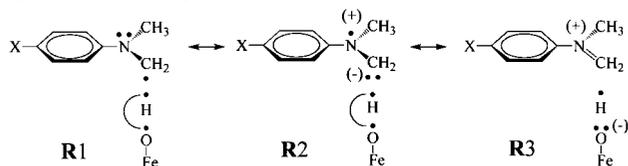
[a] The curved lines connecting the electrons signify spin-paired electrons (covalent bonding).

Scheme 14. Decomposition Reaction of Carbinolanilines
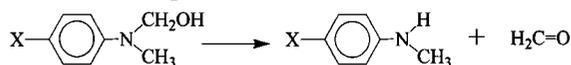
4.6.3. Summary of the Mechanistic Findings on Dealkylation of DMAs by Cpd I

The DFT(B3LYP) calculations resolve mechanistic controversies, offer a consistent mechanistic view, and reveal the following features. (a) The reaction pathways involve C−H hydroxylation by Cpd I followed by a nonenzymatic carbinolamine decomposition, which is assisted by water molecules. (b) C−H hydroxylation is initiated by a hydrogen atom transfer (HAT) step that possesses a "polar" character. As such, the HAT energy barriers correlate with the $E_{1/2}$

values and the HOMO energies of the DMAs. (c) The series exhibits a switch from spin-selective reactivity for DMA and p-Cl-DMA to two-state reactivity, with low- and high-spin states, for p-CN-DMA and p-NO$_2$-DMA. (d) The computed kinetic isotope effect profiles for these scenarios match the experimentally determined profiles. Furthermore, KIEs and TS structures vary in a manner predicted by the Melander−Westheimer postulate: as the substituent becomes more electron withdrawing, the TS is shifted to a later position along the H-transfer coordinate and the corresponding KIE increases. (e) Theory shows that a KIE can serve as a reliable probe of the spin-state selectivity of Cpd I.

4.7. DFT-Only Studies of Sulfur and Nitrogen Oxidation

4.7.1. Sulfur Oxidation

An early DFT(B3LYP) investigation of sulfoxidation of dimethyl sulfide (DMS) from 2003[52] has already been reviewed.[32] In this initial work, it was found that the sulfoxidation of DMS involves a single step that occurs preferentially from the HS quartet state. The LS TS turned out to be higher in energy and distorted as though performing porphyrin oxidation. Subsequently, two independent studies showed that there is indeed a lower-energy TS[169,171,293] that

(a) without water assistance
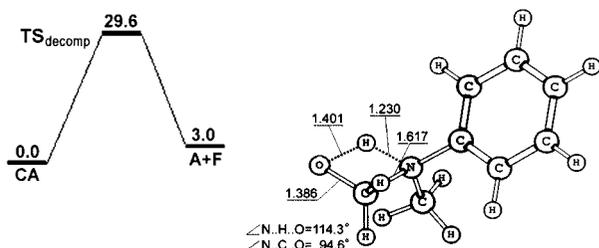
(b) with water assistance
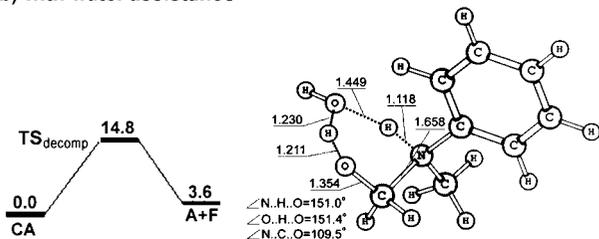
Figure 50. DFT(B3LYP) calculated transition states (distances in Å) and energy profiles (kcal mol^{-1}) for the carbinolaniline (CA) decomposition reaction to aniline (A) and formaldehyde (F): (a) without water assistance, and (b) with water assistance. Relative energies were determined using a large basis, LACV3P+(Fe)/6-31+G*(rest), and including ZPE, bulk polarity (ε = 5.6), and NH···S corrections.[167]
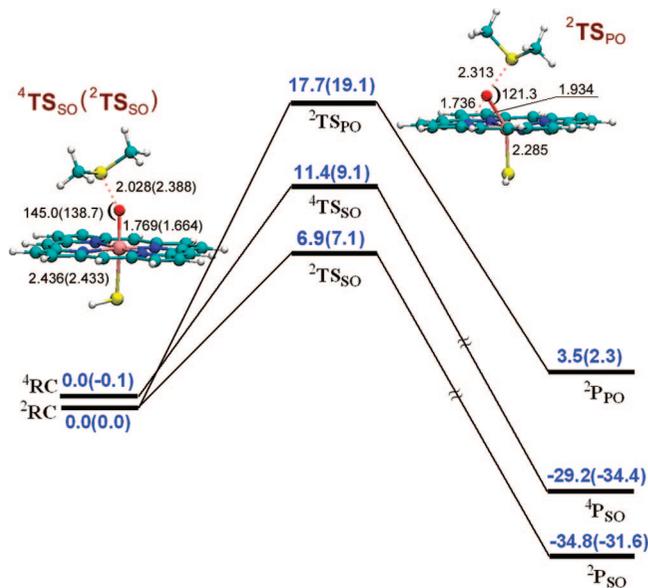
Figure 51. DFT(B3LYP/LACVP**) optimized transition states during sulfoxidation of dimethyl sulfide by Cpd I in the doublet and quartet state (^{2,4}TS$_{SO}$) and corresponding energy profiles. The ^{2}TS$_{PO}$ species corresponds to porphyrin self-oxygenation by Cpd I. The relative energies (kcal mol^{-1}) in each line come from single-point calculations with the LACV3P+(Fe)/6-311++G**(rest) basis (in parentheses: with inclusion of ZPE, bulk solvation (ε = 5.7), and NH···S hydrogen bonding corrections). Reproduced with permission from ref 169. Copyright 2007 Wiley-VCH Verlage GmbH & Co KGaA.

was missed in the original study because of the extreme flatness of the energy surface.

Figure 51 shows the results obtained by Li et al.[169,293] using DFT(B3LYP/LACV3P++**) and including ZPE corrections and the effects of bulk polarity and NH···S hydrogen
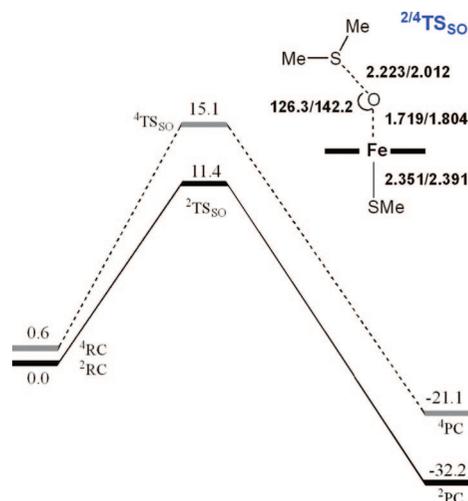
Figure 52. DFT(B3LYP/DZP) transition states for the sulfoxidation of dimethyl sulfide by Cpd I in the doublet and quartet states, and corresponding energy profiles (kcal mol^{-1}). Relative energies from single-point calculations using a large basis set with double polarization and diffuse functions. Reprinted with permission from ref 171. Copyright 2008 American Chemical Society.

Scheme 15. Different Electronic Reorganization Events Leading to ^{2}TS$_{SO}$ and ^{4}TS$_{SO}$ during Sulfoxidation by Cpd I
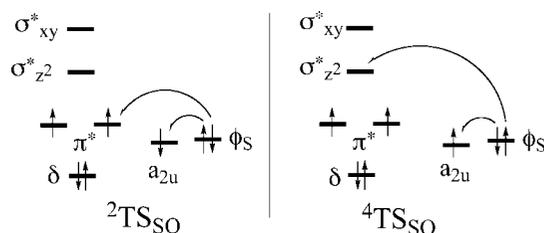
bonding. It is seen that the lowest-energy **TS**$_{SO}$ species is the doublet with a barrier of merely 7.1 kcal mol^{-1}, somewhat higher than the bond activation barrier in DMA but still very small, showing that sulfoxidation is facile. About 2.0 kcal mol^{-1} higher is a quartet ^{4}**TS**$_{SO}$ species. When neglecting bulk polarity, H-bonding, and ZPE corrections, the ^{4}**TS**$_{SO}$ species lies 4.5 kcal mol^{-1} above ^{2}**TS**$_{SO}$. Still considerably higher in energy is the distorted LS transition state, labeled as ^{2}**TS**$_{PO}$, which was wrongly assigned in 2003 as a sulfoxidation TS.[52] The label ^{2}**TS**$_{PO}$ signifies that this is a transition state for porphyrin oxidation, leading to N−O porphyrin adducts found experimentally.[16]

Figure 52 shows the DFT(B3LYP)-only results of Olsen and Ryde[171] using a large basis set. While the numerical values of the barriers and the key geometric features of the ^{2,4}**TS**$_{SO}$ species are different from those in Figure 51, the trends are similar, showing preference for the LS pathways.

In retrospect, the preference for the LS process should have been clear all along, for the same reasons that have been discussed for the HS and LS rebound processes during C−H hydroxylation. Thus, as shown in the orbital occupancy diagrams in Scheme 15, a LS state can be generated by shifting an electron from the sulfur lone-pair orbital (ϕ_S) to one of the low-lying π^* orbitals of the oxo-heme. By contrast, in the HS process, one electron must be shifted to the high lying $\sigma^*_{z^2}$ orbital, which raises the energy of the ^{4}**TS**$_{SO}$ species, compared with ^{2}**TS**$_{SO}$.

4.7.2. Nitrogen Oxidation

DFT(B3LYP) studies of oxo-transfer to aminic nitrogen have been reported by two groups,[171,292,369] using DMA and trimethylamine (TMA). They arrive at a mechanism similar to the one discussed above for sulfoxidation: the reaction is a synchronous oxo-transfer from Cpd I to the nitrogen atom, and it occurs preferentially on the LS surface, with the HS mechanism having barriers that are higher by 7 kcal mol^{-1} for DMA[292,369] and by 4−5 kcal mol^{-1} for TMA.[171,369]

Interestingly, the reverse reaction whereby the N−O compound transfers an oxygen to the heme to form Cpd I has been used by Dowers et al.[256] to study the mechanism of N-dealkylation of DMA derivatives by Cpd I. The feasibility of the oxo-transfer from the N-oxide derivative of DMA to the heme has been addressed in two recent studies,[292,369] which found a barrier of the order of 17−18 kcal mol^{-1} (or 13.6 kcal mol^{-1} after including bulk polarity and NH···S H-bonding effects). The complete energy profile[369] is shown in Figure 53 for only the LS state. It is seen that if one starts from the LS ferric complex of the N-oxo-DMA ($^2\mathbf{P}_{NO}$), the oxo-transfer to the heme is the rate controlling step, followed by a very fast step of C−H hydroxylation of one of the CH$_3$ groups of DMA. In such a scenario, Cpd I is not observable, but one might think about using the same technique with the N-oxide of hexafluorinated p-NO$_2$-DMA-F$_6$, in which case the subsequent reaction of Cpd I will be slow, while the oxo-transfer reaction will become competitive or faster.

4.8. A DFT-Only Study of Ethanol Oxidation by P450

Lieber and DeCarli have suggested already in the late 1960s that the microsomal CYP2E1 is responsible for alcohol dehydrogenation activity in the liver.[370−372] This suggestion was followed by a surge of studies of toxicology aspects, such as acidosis and fatty liver disease, associated with
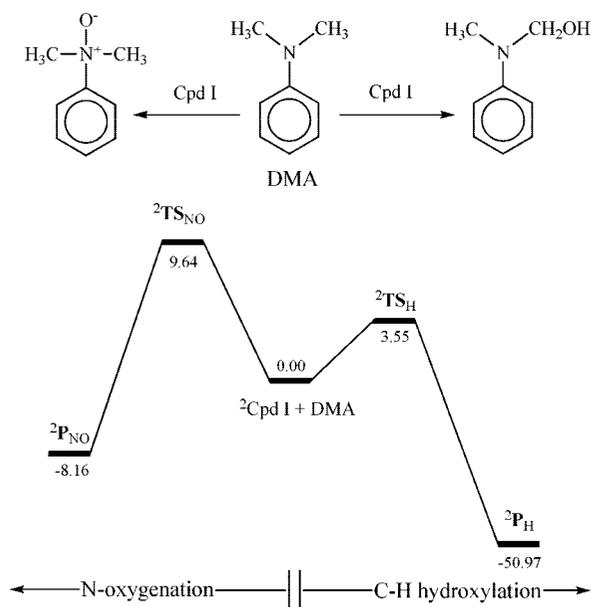
Figure 53. DFT(B3LYP/B1) optimized transition states and energy profiles for C−H hydroxylation and N-oxidation of N,N-dimethylaniline (DMA) by Cpd I in the LS state. Relative energies from single-point calculations using the LACV3P+(Fe)/6-311+G*(rest) basis, and including ZPE, bulk polarity ($\varepsilon = 5.7$), and NH···S H-bonding corrections. Reproduced with permission from ref 369. Copyright 2009 Wiley-VCH Verlage GmbH & Co KGaA.

Scheme 16. Three Mechanistic Hypotheses for Ethanol Dehydrogenation by CYP2E1 (Reproduced with Permission from Ref 168. Copyright 2007 Wiley-VCH Verlage GmbH & Co KGaA)
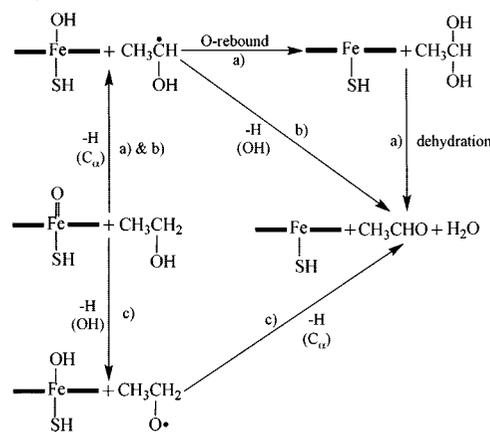
CYP2E1 malfunction, and of the mechanism of alcohol dehydrogenation.[373−383] Three mechanisms have been proposed and are shown in Scheme 16. The consensus mechanism is the gem-diol mechanism a, in which the rate-determining step is H-abstraction from the α-carbon atom of the alcohol, followed by radical rebound to form the diol, which then undergoes dehydration and produces acetaldehyde.[374−376,378−382] The second mechanistic hypothesis b starts with the same H-abstraction step from C$_\alpha$ as in mechanism a but follows with a second H-abstraction from the radical and is hence called the dual hydrogen abstraction (DHA) mechanism.[373,377,379,383] Finally, the third hypothesis c, so-called reverse dual hydrogen abstraction (R-DHA), involves an initial H-abstraction from the hydroxyl group of the alcohol, followed by a second H-abstraction from the α-carbon atom. This is the one discovered by computational means, using B3LYP(LACVP**/6-31G**) and corrections to account for basis set (LACV3P+*/6-311+G*) and bulk polarity effects.[168]

The computational study explored all three mechanistic hypotheses. Figure 54 shows the DHA and gem-diol mechanisms, which both start with an initial H-abstraction step (Figure 54a). As usual, this step involves TSR, leading to the iron-hydroxo/ethanolyl radical complexes, $^{2,4}\mathbf{I}$. These $^{2,4}\mathbf{I}$ complexes are not stable intermediates but appear as shoulders on the potential energy surface (the reason for the missing rebound barrier is the low ionization potential of the radical; see discussion later); and collapse in a barrier-free fashion to the products of mechanisms a and b, as shown in Figure 54b for $^4\mathbf{I}$. Thus, a rotation of the iron-hydroxo OH group, on the left side of Figure 54b, leads directly to the gem-diol product, while combining this rotation with the enol radical rotation, as on the right side of Figure 54b, gives rise to a second and barrier-free H-abstraction that generates the acetaldehyde complex.

The H-abstraction barrier in Figure 54a is affected by bulk polarity, by as much as 1.5−4.1 kcal mol^{-1} when the bulk polarity is taken to be that of ethanol; the barriers are 13.2/ 14.4 kcal mol^{-1} for the HS/LS states.[168]

Figure 55 shows the energy profile calculated for the R-DHA mechanism in the gas phase and with bulk polarity effects taken into account. It is seen that the mechanism starts in a regular TSR scenario, leading to an iron-hydroxo/ ethanoyl radical complex, $^{2,4}\mathbf{I}$. This complex then undergoes a second H-abstraction from the α-methylene group of the ethanoyl radical. In this step, the LS state has a tiny barrier,

(a)

(b)
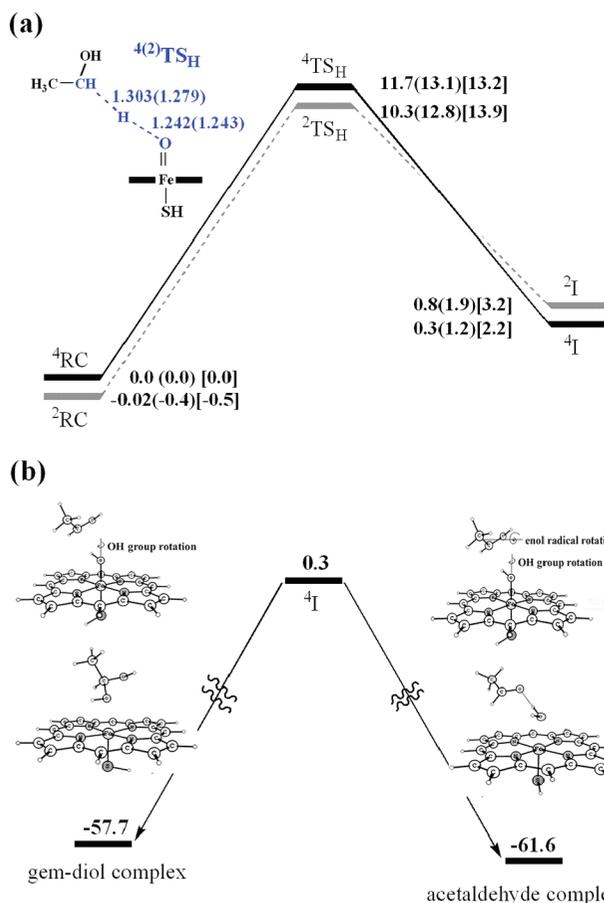
Figure 54. DFT(B3LYP/LACVP**) optimized transition states and energy profiles (kcal mol^{-1}) for ethanol dehydrogenation by Cpd I in the DHA and gem-diol mechanisms. (a) H-abstraction step and (b) collapse of ^{4}I to the gem-diol or acetaldehyde complexes. In part a, the relative energies refer to the LACV3P+(Fe)/6-311+G*(rest) basis, labeled as B. In part a, the three relative energy values in each line correspond to B+ZPE (B+ZPE; ε = 5.62) [B+ZPE; ε = 24.55]. Reproduced with permission from ref 168. Copyright 2007 Wiley-VCH Verlage GmbH & Co KGaA.
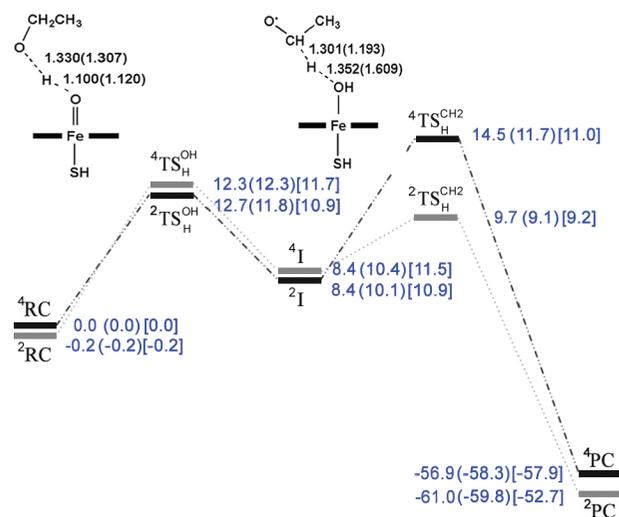
Figure 55. DFT(B3LYP/LACVP**) optimized transition states and energy profiles (kcal mol^{-1}) for ethanol dehydrogenation by Cpd I in the R-DHA mechanism. The three relative energy values in each line correspond to B+ZPE (B+ZPE; ε = 5.62) [B+ZPE; ε = 24.55], where B denotes the LACV3P+(Fe)/6-311+G*(rest) basis. Reproduced with permission from ref 168. Copyright 2007 Wiley-VCH Verlage GmbH & Co KGaA.

Scheme 17. Orbital Occupancy Evolution during the Second H-Abstraction Step of the R-DHA Mechanism in Figure 55^a
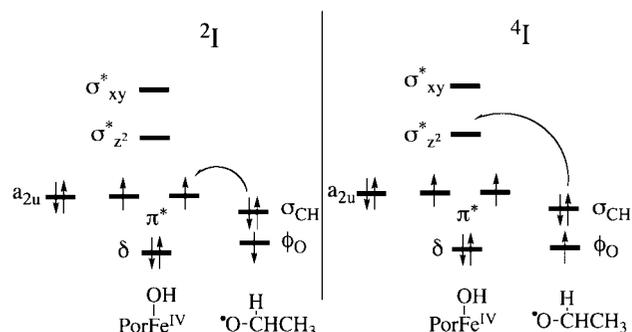
^{a} Note that the electron shift to a$_{2u}$ occurs en-route from ^{2,4}TS$_H^{OH}$ down to the intermediate complexes.

while the HS state has a large barrier. The reason for this spin-selective second H-abstraction can be understood from the orbital occupancy evolution diagrams in Scheme 17. Thus, the H-abstraction involves a one-electron shift from the σ_{CH} orbital of the methylene group to the iron hydroxo complex; in the LS case, the electron is shifted to the low lying π^* orbital, while in the HS case, the shift occurs to the high lying $\sigma^*_{z^2}$ orbital (that is 18 kcal mol^{-1} higher than the π^* orbital).

Assuming a fast spin crossover from quartet to doublet in the second H-abstraction event (see reasoning in the original paper[168]), the rate controlling step during the R-DHA mechanism in Figure 55 is the first H-abstraction from the hydroxyl group of ethanol. Comparison of Figures 54a and 55 reveals that, in the gas phase, the consensus mechanisms (DHA and gem-diol) possesses a slightly smaller barrier than the R-DHA mechanism. However, with bulk polarity included, the reverse is true; the barriers for R-DHA are 11.7/11.1 vis-à-vis 13.2/14.4 kcal mol^{-1} (for HS/LS, respectively). Thus, the R-DHA mechanism competes with the traditional DHA or gem-diol mechanism in a weakly polar environment (ε = 5.62), which reflects the hydrophobic heme pocket of CYP2E1. However, under stronger polarity (ε = 24.55), corresponding to that of ethanol, the R-DHA mechanism should prevail. This might explain the adaptive response of CYP2E1 to high ethanol concentration in the blood, resulting in the acceleration of ethanol dehydrogenation.[370−372] The influence of ethanol concentration on the CYP2E1 pocket has in fact been observed by changes in the Fe−CO bond of the ferrous−carbon monoxide complex, which are in line with the expected modulation of the polarity of the pocket.[384] A reasonable scenario was proposed whereby the increased ethanol concentration in the blood modifies the pocket polarity and induces a mechanistic shift from a competitive gem-Diol-DHA/R-DHA situation to a dominant R-DHA mechanism, which offers a faster alcohol dehydrogenation path and thereby accounts for the observed acceleration of ethanol dehydrogenation.[370−372]

4.9. DFT/MM Study of the C−C Bond Forming Reactivity of P450

P450 StaP (CYP245A1) catalyzes the formation of staurosporine (STA) from chromopyrrolic acid (CPA) as depicted in Scheme 18.[195,385] Staurosporines are natural products that possess strong inhibitory activity on protein kinase, which makes them therapeutically important anticancer agents. The staurosporine indolocarbazole cores are derived from L-

Scheme 18. Conversion of L-Trp to Staurosporine (STA) via CPA[a]
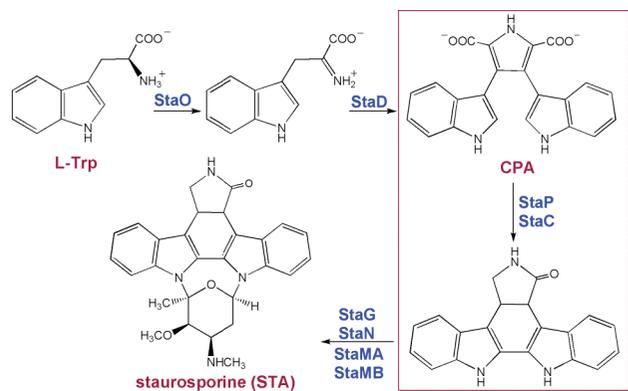
[a] The oxidation of CPA by P450 StaP in the frame involves C−C formation and loss of two H atoms. Acronyms on the arrows refer to the enzymes that catalyze the specific steps. Reprinted with permission from ref 80. Copyright 2009 American Chemical Society.

tryptophan (L-Trp) via a series of oxidative transformations catalyzed by a series of enzymes that are indicated on the arrows in Scheme 18. The most remarkable transformation is the fusion of the two indole rings of CPA by cytochrome P450 StaP and its coenzyme StaC, to generate a six-ring indolocarbazole scaffold that undergoes subsequent N-glycosylation, followed by methyl transfer steps leading in the end to STA. How does an enzyme that regularly monooxygenates organic compounds perform this remarkable transformation that involves C−C bond formation and a loss of two H atoms? This transformation becomes all the more puzzling in view of the fact that CPA is held steadfast in the crystal structure by an array of H-bonds quite distant from Cpd I, as shown in Figure 56. As such, there is no chance of moving the CPA closer to enable Cpd I to abstract hydrogens or affect CPA directly in any other manner. The action must be from a distance and/or mediated by water molecules and active residues.

The electronic structure of Cpd I of P450 StaP was already discussed above in Figure 19, based on DFT(B3LYP)/MM calculations.[58,80] In this Cpd I species, the cation-radical moiety is distributed over Por, sulfur, and CPA in an orbital that has a mixed character, made from a_{2u} for the Por/S ligand and ϕ_{CPA} that is localized in the proximal indolic moiety of CPA. Furthermore, the DFT(B3LYP)/MM calculations[58] show that when the His₂₅₀ residue is protonated in the proximal side to the FeO moiety of Cpd I, there is an H-bond triad consisting of Wat₇₈₉ (the water molecule that is liberated during the formation of Cpd I), Wat₆₄₄, and His₂₅₀ that connects the indolic NH of CPA to the oxo group of Cpd I. As shown in Figure 57, this provides a nice Grotthuss-type proton transfer pathway[236] for proton delivery from NH to the oxo group, which occurs simultaneously to an electron transfer from CPA to Cpd I. This proton-coupled electron transfer (PCET) event results in a net H-atom transfer from the indole NH to Cpd I, without bringing CPA close to Cpd I. In this mechanism P450 StaP functions much like a peroxidase.[58,71] Three elements join to provide this remarkable functionality: (i) the protein environment deprotonates the two carboxylic acid groups of CPA and immobilizes the substrate via an array of H-bonds that claw the carboxylate groups (see Figure 56), (ii) the so created CPA²⁻ dianion is an excellent electron donor, while Cpd I is a powerful electron acceptor, which together promote electron transfer

from a distance, and (iii) the Wat₆₄₄-His₂₅₀-Wat₇₈₉ triad enables us to shuttle protons from the CPA to the Cpd I species.

However, the PCET step in Figure 57 just initiates the mechanism of CPA oxidation, which was recently elucidated by means of extensive DFT(B3LYP)/MM calculations.[80] The entire computed mechanism is shown in Scheme 19. It involves 10 steps labeled as **s1−s10**. In addition to the PCET step (**s1**), there is another key element, a bond-formation coupled electron transfer (BFCET) step (**s4**) that involves the second electron transfer from the substrate to the heme along with the C−C bond formation. There are also four proton transfer (PT) events; in the first one, PT1 (**s3**), the proton from the distal indolic N−H group of CPA, is transferred to the proximal N, while thereafter (**s6**) this proton is relayed to the iron hydroxo to form the ferric-aqua complex of the resting state. Steps PT3 and PT4 (**s8** and **s10**) rearrange the protons from the C−H groups of the HC−CH moiety of the indole back to the indolic bare nitrogen atoms, thus completing the transformation and coming full cycle.

However, recalling that the CPA substrate is immobilized, none of these events could have happened without the water relocation steps (**s2**, **s5**, **s7**, and **s9**) in Scheme 19. As an example, **s2** is shown in Figure 58. Here, after the PCET event, **s1** in Scheme 19, Wat₇₈₉ and Wat₆₄₄ migrate to establish together with His₂₅₀ a new triad that can relay the proton from the distal indolic N−H group to the proximal bare N atom. Similarly, the hydrogen shifts from the carbons to the N positions in the indole (**s8** and **s10**) are mediated by water molecules that have to migrate initially to the correct positions to effect these rearrangements. Without the water molecules, these H-rearrangements have large barriers, which become smaller than 12 kcal mol⁻¹ in their presence.

Indeed, the water molecules are the "working-class bees" of this complex mechanism. By moving around, they tend to "nurture" the immobilized CPA, allowing it to undergo a most complex transformation. *They act as biocatalysts.*

4.10. DFT-Only and DFT/MM Calculations of Regioselectivity in the Reactions of Cpd I

An important function of an enzyme is its ability to carry out regioselective reactions of substrates that have two functionalities. P450 enzymes fulfill this role quite well: they may perform C−H hydroxylations in the presence of double bonds, or sulfur, or phenyl groups, and vice versa (oxidation of these groups in the presence of C−H bonds). A few of these regiochemical problems have been addressed by theory and are discussed below.

4.10.1. Competitive Heteroatom Oxidation vs C−H Hydroxylation

For substrates such as N,N-dimethylaniline (DMA), trimethylamine (TMA), and dimethylsulfide (DMS), there is a possibility for competitive C−H hydroxylation of the α-methyl group and oxo-transfer to the heteroatom. Interestingly, amines such as DMA and TMA prefer C−H hydroxylation (eventually leading to N-demethylation; see section 4.6), while sulfur containing compounds such as DMS prefer sulfoxidation.[18,287,386] The former issue has been addressed by DFT(B3LYP)-only calculations by Li et al.[369] and Cho et al.,[292] who showed that C−H hydroxylation is preferred over oxo-transfer to nitrogen, by as much as 3 kcal mol⁻¹ for TMA and 6 kcal mol⁻¹ for DMA. Rydberg et al.[171]
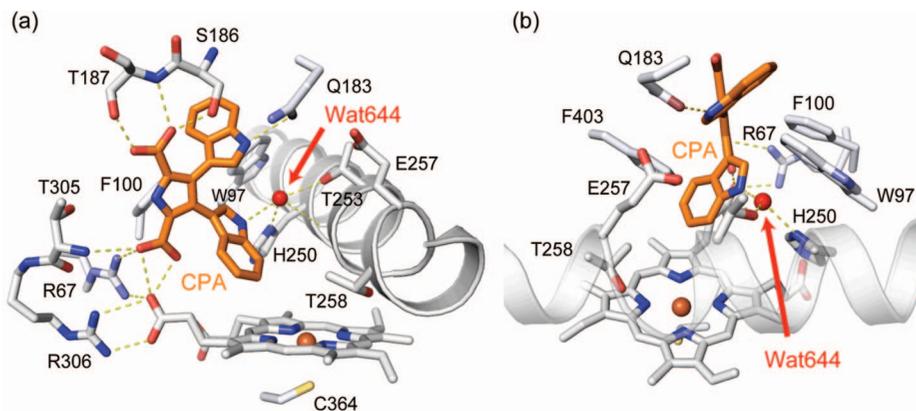
Figure 56. X-ray structure showing the binding of chromopyrrolic acid (CPA) by P450 StaP via an array of H-bonds, which are indicated by dotted lines. (a) and (b) represent two different perspectives; (b) shows His$_{250}$ being H-bonded to CPA via a water molecule.[195] Reprinted with permission from ref 80. Copyright 2009 American Chemical Society.
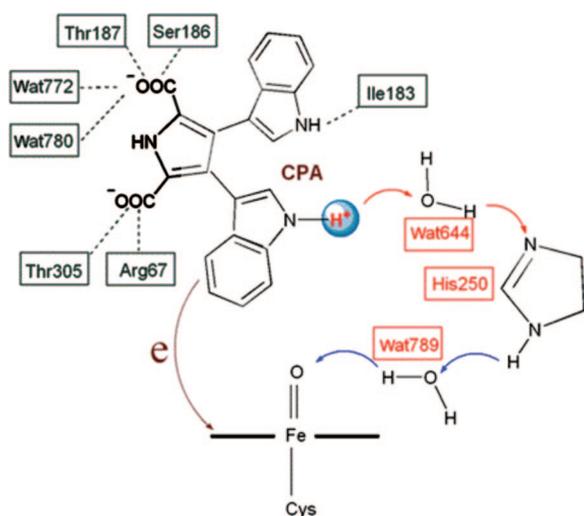
Figure 57. Schematic representation of the proton-coupled electron transfer (PCET) event involving a proton relay from the indole group of CPA to the iron-oxo moiety of Cpd I, via the Wat$_{644}$-His$_{250}$-Wat$_{789}$ triad, with concomitant electron transfer from CPA to Cpd I. Reprinted with permission from ref 58. Copyright 2008 American Chemical Society.

studied both types of competition for TMA and DMS as substrates, using DFT(B3LYP)-only calculations, and they reported a preference of 4.4 kcal mol^{-1} for C–H hydroxylation in TMA and of 4 kcal mol^{-1} for oxo transfer to sulfur in DMS. The slightly different values of the two groups reflect differences in basis sets and in the incorporation or neglect of bulk polarity corrections. Notwithstanding these minor differences, these results show that DFT(B3LYP)-only calculations are capable of addressing the regiochemical problem in heteroatom containing compounds.

4.10.2. DFT-Only Studies of the Competition between Epoxidation and Allylic Hydroxylation in Propene and Cyclohexene by Cpd I

The competition of C–H hydroxylation and C=C epoxidation in olefins that possess allylic hydrogens, such as cyclohexene and propene in Scheme 20, is a fundamental problem in P450 selectivity. Experimentally, with liver microsomal P450, propene gives products that are exclusively related to C=C activation,[8,387] while cyclohexene yields a roughly 1:1 ratio of the two products with a variety of P450s.[194,308] Other olefins give different values, generally

with preference to C=C activation, but not always.[257,388] Interestingly, using synthetic Cpd I reagents, one can get with cyclohexene a variable C–H/C=C activation ratio that ranges from the typical 1:1 ratio of the enzymatic reaction up to 100% C–H activation, depending on conditions such as solvent polarity and small amounts of water in the solvent.[389,390] Thus, the model studies show that a single oxidant (Cpd I) may yield different C–H/C=C regioselectivity ratios under different conditions of polarity, water content, axial ligand, and so on.

Early DFT(B3LYP) studies with propene as a model substrate by de Visser et al.[149,150] addressed the C–H/C=C competition. In the gas phase, Cpd I was found to be nonselective, yielding C–H/C=C ratios close to 1. However, inclusion of the effects of bulk polarity and hydrogen bonding to the sulfur ligand led to a clear preference for C–H activation, by as much as 7 kcal mol^{-1}, thus predicting exclusive allylic hydroxylation. For cyclohexene, DFT(B3LYP) calculations show a preference for C–H hydroxylation already in the gas phase, which is augmented by polarity.[57] Another recent study of propene oxidation indicates[275] that an electric field alone can achieve selectivity in this reaction and that, by simply reversing the field direction along the S–Fe–O-olefin axis, one could control the selectivity at will to give C=C or C–H activation. This is an exciting result, and studies of synthetic models seem to suggest that it is also a viable one.[389,390]

4.10.3. DFT/MM Studies of the Competition between Epoxidation and Allylic Hydroxylation in Propene and Cyclohexene by Cpd I

The results of the model studies,[149,150] which show a preference for C–H hydroxylation for both propene and cyclohexene, are in discord with the experimental results in the enzyme. Clearly, the enzymatic regioselection issue must be studied by means of QM/MM calculations within the native protein. However, before discussing the results of such calculations, it is important to clarify the conditions whereby experimental selectivity can be studied in P450 enzymes. Experimentally, the rate controlling step in these enzymes is not at all substrate oxidation but rather the second electron transfer step[180] in the catalytic cycle (see Figure 1), at least for quite a few of the isozymes, albeit not for all.[391] Thus, a regioselectivity that reflects the relative barriers for C=C and C–H activations, under experimental conditions, requires a rapid equilibrium between the two reactant complexes of

Scheme 19. DFT(B3LYP)/MM-Elucidated Mechanism of CPA Oxidation by P450 StaP Involving 10 Steps, Labeled as s1−s10 (Reprinted with Permission from Ref 80. Copyright 2009 American Chemical Society)
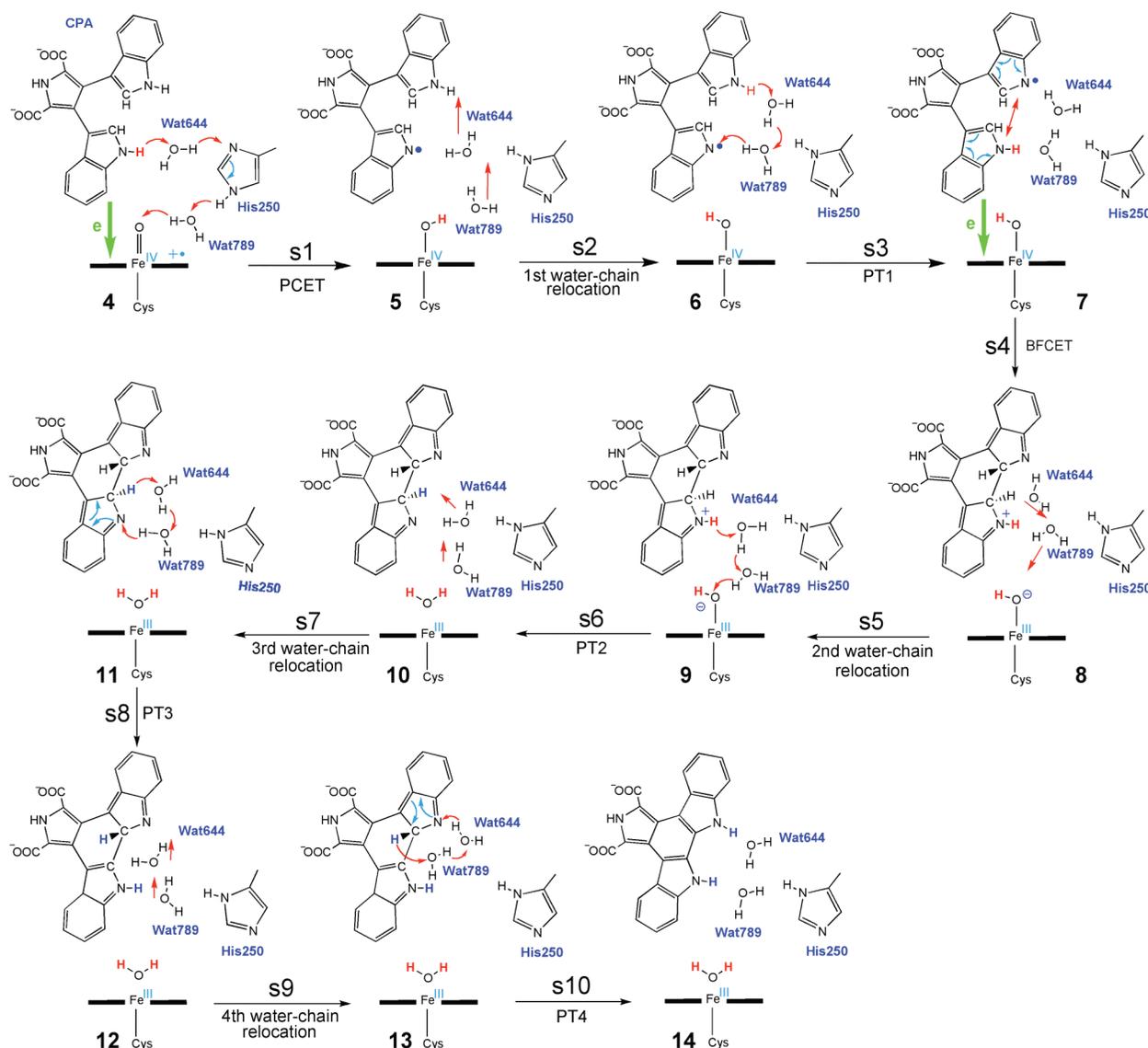
Cpd I with the C=C and C−H moieties. In such a case, the C−H/C=C ratio is determined by the relative energies of the corresponding bond activation TSs. Otherwise, the measured regioselection ratio will not reflect the actual oxidation process. Therefore, any QM/MM calculation of regioselectivity must ascertain this condition of free tumbling of the substrate. Under these conditions, the relative barriers will be determined solely by the difference in TS energies, while the relative energies of the reactant complexes do not matter.

The C−H/C=C regioselection problem was tackled by Cohen et al.[57,277] using DFT(B3LYP)/MM calculations for both propene and cyclohexene in CYP101. Initial MD simulations showed that, unlike camphor, which is bound by Tyr96 and Val295, both cyclohexene and propene tumble freely within the pocket and have no favored conformation of binding by the residues of the enzyme. This means that, in principle, there should be a rapid equilibrium between the reactant complexes of Cpd I and the two regiochemical sites, i.e. Cpd I/C−H and Cpd I/C=C. Under these conditions, the reactant complexes do not influence the regioselectivity, and *the relevant comparison is between the QM/MM energies of the respective transition states*.

The results of the DFT(B3LYP)/MM calculations[57] for two different basis sets were very similar for the two olefins. In both cases, the protein environment was shown to prefer the C−H activation by as much as 7−8 kcal mol^{-1} for cyclohexene and 3−5 kcal mol^{-1} for propene. These results are clearly still in discord with experimental reports on the two reactions in P450 enzymes. To tackle this disparity, the original paper[57] suggested a model based on the kinetics of the full catalytic cycle. The model seemed successful in predicting the C−H/C=C ratio for the two olefins and others. However, in retrospect, it was recognized that the suggested model described two disconnected cycles, with no free tumbling between the reactant complexes, and as such was inappropriate for addressing this problem, since the MD simulations clearly showed free tumbling. Evidently, DFT(B3LYP)/MM calculations are still incapable of treating this delicate problem. At the time of writing of this review, this problem remains unresolved. Another pending issue is the fact that the results from synthetic model studies of C=C/C−H competition[389,390] are different from those in the enzyme, at least with respect to the regioselectivity of cyclohexene oxidation.
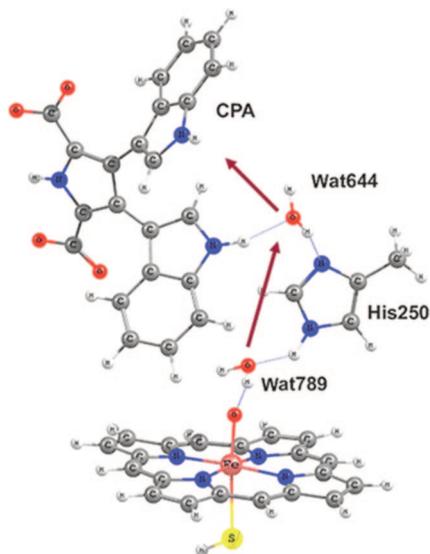
Figure 58. DFT(B3LYP/LACVP) optimized product of the PCET event (**s1** in Scheme 19) with arrows indicating the movements of Wat$_{644}$ and Wat$_{789}$ in the water relocation step (**s2** in Scheme 19). Reprinted with permission from ref 80. Copyright 2009 American Chemical Society.

Scheme 20. Regiochemical Pathways of Cyclohexene and Propene Epoxidation, as Representatives of the Competition of C=C Epoxidation vis-à-vis Allylic C−H Hydroxylation (Reprinted with Permission from Ref 57. Copyright 2006 American Chemical Society)
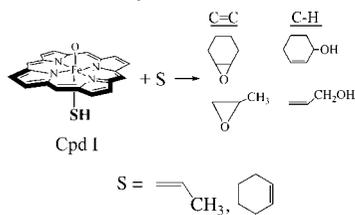
4.10.4. DFT-Only Studies of the Metabolic Switching from Benzylic C−H Hydroxylation to Phenyl Hydroxylation in Toluene by Cpd I

Metabolic switching is the name given to the changes in regioselectivity induced by replacement of hydrogen by deuterium.[317,357,391−394] This is exemplified in the case of toluene in Scheme 21. Thus, while the pristine toluene gives mostly benzyl alcohol, deuteration of the methyl group results in "metabolic switching" to phenyl hydroxylation products.[317,392,394] This of course depends on the magnitude of the KIE during benzylic H-atom abstraction, as well as on bulk polarity and environmental effects (H-bonds to sulfur, etc.) that might differentiate the radicalar mechanism for H-abstraction from the polar π-attack mechanism of the phenyl activation (see section 4.5). Indeed, DFT(B3LYP)-only calculations[395] show that, in the pristine toluene, the H-abstraction from the benzylic position has a free energy barrier which is 5.6 kcal mol^{-1} lower than that of phenyl activation, whereas, for deuterated toluene and with the inclusion of bulk polarity and NH⋯S H-bonding effects, the C−H hydroxylation preference went down by a significant amount to only 0.4 kcal mol^{-1}. Thus, while these results do not accurately reproduce the experimental data, due to the limited accuracy of DFT, they still give the correct trend, namely, that the combination of deuteration and bulk polarity effects tends to favor phenyl over C−H hydroxylation, as also observed in model systems.[389]

Scheme 21. Regiochemical Pathways of Toluene Oxidation by Cpd I, Showing the Metabolic Switching upon Methyl Deuteration (Reproduced with Permission from Ref 395. Copyright 2007 Wiley-VCH Verlage GmbH & Co KGaA)
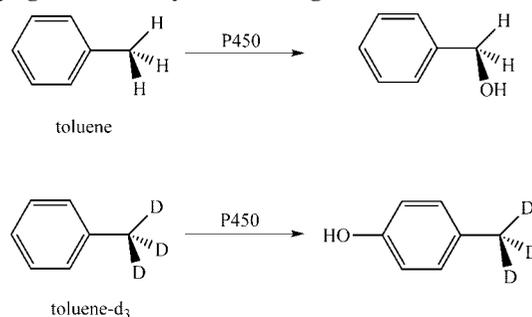
	o-cresol	*m*-cresol *p*-cresol	benzyl alcohol
toluene	14.8	16.2	69.0
toluene-d$_3$	34.7	41.0	24.3
toluene-d$_8$	35.2	45.2	19.6

A recent DFT(B3LYP)-only study by Lin et al.[173] shows that gas-phase calculations can predict the regioselectivity trend in testosterone hydroxylation by CYP3A4. The regioselectivity follows the bond dissociation energies of the various C−H bonds undergoing hydroxylation, much as demonstrated previously by de Visser et al.[153]

4.11. Miscellaneous Reactions: Carbocation Intermediates

The incursion of carbocations during P450 bond activation has been reviewed before.[32,152] It involves an electron shift from the radical center, initially formed by bond activation, to the iron-hydroxo or iron-alkoxy complex. This was shown to be the main mechanism of the dehydrogenase-oxidase activity often observed in P450 enzymes.[380,396,397] The same mechanism was found recently by DFT(B3LYP)-only calculations to be responsible for the last of the aromatization steps in the catalytic conversion of androgens to estrogen products by CYP19 aromatase.[381]

5. Predictive Models for P450 Reactivity Patterns

Predictive models for P450 reactivity are sought for two main reasons. One reason is pragmatic, since finding correlative relationships for P450 reactivity should be extremely useful for fast predictions of drug metabolism.[170,172,398] Another reason is fundamental, trying to achieve understanding and predictive ability from first principles. In reviewing this activity, we shall focus mainly on models that relate reactivity to a few fundamental properties of the molecules and are based on an underlying physical model behind the correlation. We do not consider schemes such as QSAR (quantitative structure−activity relation) that use a multitude of molecular descriptors, and we refer readers interested in such schemes to recent reviews in the field of drug metabolism.[399,400]

5.1. Predictive Pattern of C−H Hydroxylation Barriers Based on Bond Dissociation Energies

As reviewed previously,[32] there have been a number of studies that attempt to correlate hydrogen abstraction barriers

with the corresponding bond dissociation energies (BDE$_{CH}$),[401-403] which were reasonably successful for P450 hydroxylation.[404,405] Notably, de Visser et al.[153] used DFT(B3LYP) calculations with two different basis sets (B1 and LACV3P+/6-311++G**) and studied the barriers for C−H hydroxylation by Cpd I for 10 different substrates, ranging from methane through camphor to toluene and phenyl ethane. The use of ZPE corrected BDE$_{CH}$ values led to very good correlation with ZPE corrected barriers:[153]

$$\Delta E(ZPE)^{\ddagger} = 0.5887 BDE_{CH} - 39.418 \ (kcal \ mol^{-1}) \tag{3}$$

More recently,[174] the same authors took a different set of reactions and used a different basis set and different software to run the DFT(B3LYP)-only calculations (GAUSSIAN instead of JAGUAR, which have small differences). Interestingly, while the quality of the correlation was less good than the one found in the past, the equation for the correlation remained essentially the same:

$$\Delta E(ZPE)^{\ddagger} = 0.5849 BDE_{CH} - 39.625 \ (kcal \ mol^{-1}) \tag{4}$$

The utility of this correlation can be illustrated for the case of camphor hydroxylation by P450$_{cam}$. The computed low-spin $\Delta E(ZPE)^{\ddagger}$ barrier (LACV3P+), relative to the gas-phase reactant complex, is 15.1 kcal mol^{-1}, and the corresponding $\Delta H^{\ddagger}$ value is within 0.4−0.6 kcal mol^{-1}.[277] A similar $\Delta E(ZPE)^{\ddagger}$ barrier is obtained from DFT(B3LYP)/MM calculations[59] that do not include the barrier lowering effect of Wat$_{903}$.[61] The computed BDE$_{CH}$ value (6-311++G**) for camphor is 93.7 kcal mol^{-1}, close to the experimental estimate of 98 ± 4 kcal mol^{-1}.[281] Using the computed BDE$_{CH}$ value of 93.7 kcal mol^{-1} in eq 4 yields an estimate of 15.2 kcal mol^{-1} for the barrier, almost identical to the actually computed value of 15.1 kcal mol^{-1} (see above). Furthermore, a single turnover kinetics datum for camphor hydroxylation[406] by P450$_{cam}$ yields an upper limit of the free energy barrier as $\Delta G^{\ddagger} < 15$ kcal mol^{-1} ($k > 200$ s^{-1}). The estimated DFT(B3LYP)/MM barrier in Figure 37, with the Wat$_{903}$ effect and ZPE correction included, is 13.8 kcal mol^{-1} and is not too far from these correlation-based estimates. This match supports the utility of the correlations given in eqs 3 and 4.

Recently, Olsen et al.[170] calculated the C−H abstraction barriers for a larger set of 24 molecules with sp^3 and sp^2 hybridizations, as well as primary, secondary and tertiary carbons. They found good correlations of the HS barriers with BDE$_{CH}$ values. They further showed that one could use computationally cheaper methods and obtain reasonable correlations. Based on these correlations, the authors were able to derive qualitative bracketing rules for the various C−H abstraction barriers (excluding methane), depending on the chemical environment of the carbon atom: (i) Primary sp^3 hybridized carbons, ∼17−18 kcal mol^{-1}, (ii) secondary and tertiary sp^3 hybridized carbons, 14−15 kcal mol^{-1}, (iii) carbons adjacent to aromatic and C=C moieties (e.g., allylic), 12−13 kcal mol^{-1}, (iv) carbons with a heteroatom O or S in an α-position, 11−12 kcal mol^{-1}, and (v) carbons with an N heteroatom in the α-position, 7 kcal mol^{-1}. These qualitative rules were used subsequently to predict the regiochemistry in the metabolism of progesterone and dextromethorphan by CYP3A4, once again attesting to the utility of these correlations. The authors noted that even the high barriers in these series correspond to a rate constant of 0.1 s^{-1} and thereby illustrate the high reactivity of Cpd I. Furthermore, for metabolic purposes, these high rates reduce the predictive task to locating the C−H sites that are most reactive and sterically most accessible to the reactive species. The second factor, in the authors' opinion, depends on the P450 isozyme and can be studied by means of docking and MD simulations. Similar ideas were shown to apply for 17 cases of arene hydroxylation by Cpd I.[172]

5.2. Valence Bond Modeling of Reactivity Patterns in C−H Hydroxylation

Among the most successful chapters in theoretical chemistry have been those that involved constructions of quantum chemical models for conceptualizing the reactivity and mechanism of complex organic reactions. Such theories, like the Woodward−Hoffmann rules[407-411] and Fukui's frontier orbital theory,[412] have formed an extremely useful interface between experiment and numerical calculations. As such, these models enable chemists to design better experiments and to benefit from the interplay between theoretical calculations and experiments. The field of P450 certainly involves a great deal of complexity on the experimental side, but it is equally complex on the computational side. Here too we need simpler models that can bridge the two disciplines and provide a conceptual framework for understanding the P450 reactivity patterns from a few first principles. Some progress along these lines has been made in recent years, and these attempts are described in this subsection.[26,83,174]

During the studies of correlation of reactivity with BDE$_{CH}$,[170,174] it was noted that using in situ bond strengths, which do not allow for structural relaxation of the radical after bond dissociation, generally gave a better correlation than the experimental BDE$_{CH}$ values. Moreover, in C−H activation by nonheme iron-oxo reagents, where the correlation with the BDE$_{CH}$ values was successful,[260] one could find an equally good correlation with the ionization potential of the alkane (IP$_{AlkH}$).[413] This means that the BDE$_{CH}$ and IP$_{AlkH}$ values are internally correlated, and the correlation with BDE$_{CH}$ is not necessarily fundamental, albeit very useful. Indeed, a correlation by itself cannot imply a causal relationship, and this recognition formed an incentive to seek a more fundamental principle than the Bell−Evans−Polanyi (BEP) one that is responsible for these correlations. Such a model should deal not only with C−H activation but also with the rebound step and its spin-state selectivity, and it should be sufficiently general to deal with other reactions such as C=C epoxidation, heteroatom oxidation, and so on. This principle was derived using the valence bond (VB) diagram model of reactivity, which explains the origins of barriers and of the incursion of reaction intermediates.[402,414,415]

5.2.1. Valence Bond Diagrams for C−H Hydroxylation

As already noted, there are two electromeric intermediates that may be involved during the reactions: these are the Por$^{\bullet+}$Fe^{III}OH and PorFe^{IV}OH species. In addition, there are two spin states, doublet and quartet, which lead to distinct reaction profiles. In order to economize the discussion, we shall start by modeling the mechanisms for the two electromeric states, without specifying spin states, and then comment on the spin-state effect mainly by reference to the rebound step.[26,83]
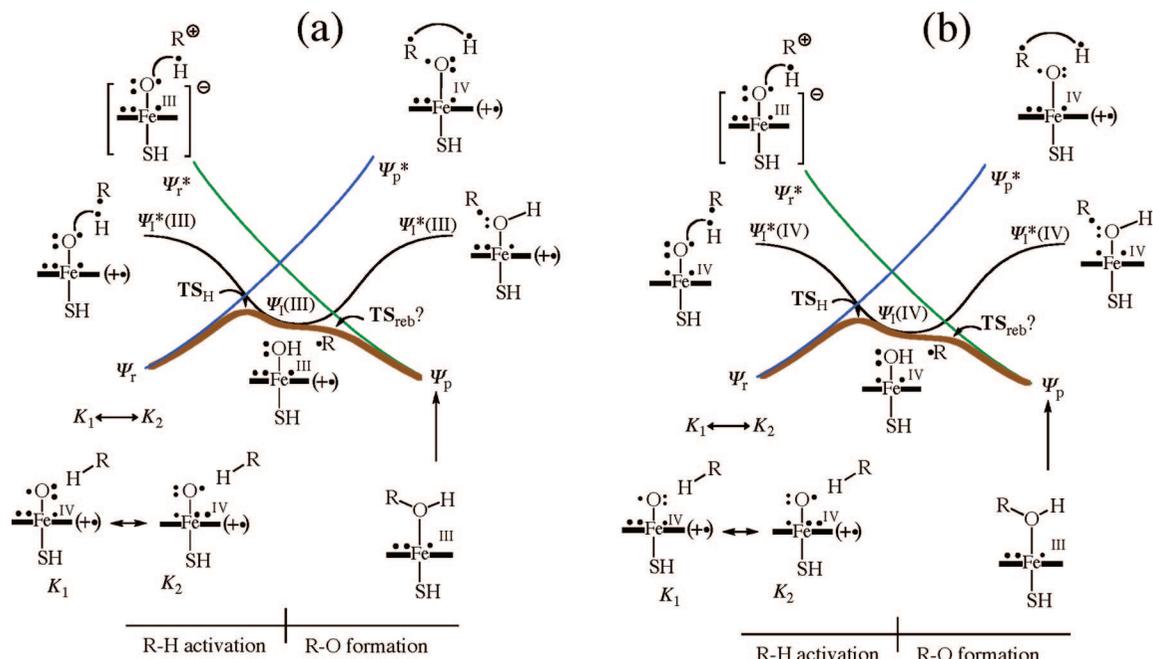
Figure 59. Valence bond (VB) diagrams describing the mechanistic scenarios for C−H hydroxylation by Cpd I, for the two electromeric pathways, without specifying a spin state. In each case, an intermediate curve cuts through the high barrier of the reactants' and products' curves. In part a, the intermediate curve corresponds to the Por$^{•+}$Fe(III)OH electromer, and in part b, it corresponds to the PorFe(IV)OH electromer. The VB mixing and avoided crossing creates the reaction profile, which is shown in both cases by the bold curve. The question mark alongside **TS**$_{reb}$ signifies that the existence of this species depends on the spin-state variety. The various species are illustrated in more detail in Scheme 22. Reprinted with permission from ref 174. Copyright 2008 American Chemical Society.

Figure 59 shows the VB diagrams for the hydroxylation reactions involving the two electromeric intermediates; the alkane is symbolized as R−H, and the porphyrin ring by the bold lines flanking the iron. The two diagrams involve two major curves that correspond to reactant and product states shown in blue and green, respectively. A third curve (in black) that cuts through the high energy ridge of the two major curves provides a low energy pathway for the transformation.[402,416] Looking at the wave functions of reactants and products, labeled as Ψ_r and Ψ_p at the lower two ends of each diagram, it is seen that the overall process of conversion of the alkane, R−H, to an alcohol involves also a two-electron reduction of iron-oxo porphyrin, which changes from Por$^{•+}$Fe^{IV}=O to PorFe^{III}(alcohol). This however is mediated by the intermediate curve that describes an H-abstraction process, whereby the electron of the abstracted H$^•$ reduces either the oxidation state on iron to Fe^{III} (Figure 59a) or the oxidation state of the porphyrin from a cation-radical to a closed-shell species (Figure 59b). This step is followed by radical rebound, where an electron shift from the alkyl radical (R$^•$) to the iron-porphyrin moiety generates the alcohol and the ferric porphyrin, which is twice reduced vis-à-vis the reactant state. *This three-curve VB diagram is a typical case, where an intermediate state internally facilitates the otherwise more difficult transformation of Ψ_r directly to Ψ_p.*[402,416]

To facilitate understanding of the state correlation in the VB diagram, we have to translate the VB representation of the various anchor states to the molecular orbital picture and the language of oxidation state formalism, used throughout this review. The two major curves describe the transformation of reactants to products and are anchored in two ground states, Ψ_r and Ψ_p, for the reactants and products, and the two corresponding excited states, Ψ_r* and Ψ_p*. Each state-curve represents a single bonding scheme: Thus Ψ_r describes the bonding pattern in the reactants, and so does Ψ_p*, only

that the latter state represents the reactant-type bonding in the geometry of the products. Similarly, Ψ_r* involves the product-type bonding but in the geometry of the reactants. As a result of these bonding features, as we move along the reaction coordinate, by stretching the C−H bond and shrinking the O···H and R···O distances and thereby changing the geometry from reactants to products, the Ψ_r state connects to Ψ_p* and, at the same time, Ψ_r* connects to Ψ_p. Let us now inspect these four states in terms of electronic structure cartoons in detail.

Starting with the reactants state, Ψ_r (Figure 59), we depict in Scheme 22 the VB-MO representations of Ψ_r for ^{4}Cpd I and R−H. In the VB formulation (Scheme 22a), Cpd I has six π-electrons, distributed in two perpendicular planes, across the Fe−O moiety, and coupled to produce a triplet spin state.[417] This state is described by a resonance mixture of the two forms (K_1 and K_2) that lead to formation of two three-electron bonds in the two perpendicular planes; the line connecting the Fe and O atoms in Figure 59 and in Scheme 22 represents the σ_{Fe-O} bond. The porphyrin is open-shell, as indicated by a cation−radical sign. In the MO formulation (Scheme 22b) these six π-electrons on the FeO moiety occupy two bonding π_{FeO} orbitals and two singly occupied antibonding π^*_{FeO} orbitals. The third unpaired electron of Cpd I resides on the porphyrin in an a_{2u} type orbital, and its spin is either ferromagnetically or antiferromagnetically coupled to the triplet FeO pair, thereby yielding two spin states: overall quartet or doublet. Note that the doubly occupied δ orbital on iron and two virtual antibonding orbitals (σ^*_{xy} and $\sigma^*_{z^2}$) are not shown explicitly in Scheme 22. The alkane R−H is represented in VB diagrams by the corresponding Lewis structure and in MO representations by the bond orbital σ_{CH} of the R−H bond undergoing activation.

The excited state Ψ_r* in Figure 59 is a charge transfer state in which one electron is transferred from the R−H

Scheme 22. (a) Valence Bond (VB) Scheme Describing the FeO Bonding in Cpd I in Terms of the Two Resonance Structures, K_1 and K_2, That Account for π-Bonding. (b) Corresponding MO Representation Showing That the Six π-Electrons Occupy the π and π* Manifold of Perpendicular Orbitals, While the Cation-Radical State Is Described by the a_{2u} Orbital[a]
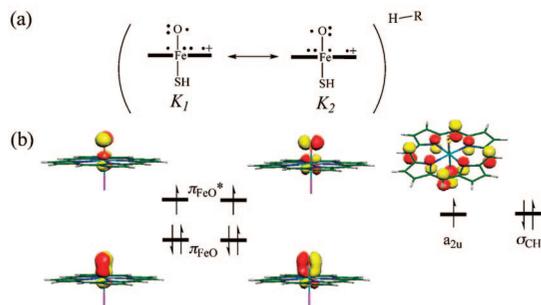
[a] In part a, the heavy dots represent the π-electrons in two perpendicular planes, while the line connecting Fe–O represents the Lewis σ_{FeO} bond. The (•+) sign near the porphyrin represents the cation-radical state of this moiety. The alkane is represented by the R–H bond. In part b, the alkane is represented by the corresponding σ_{CH} orbital. Reprinted with permission from ref 174. Copyright 2008 American Chemical Society.

alkane to fill the "hole" in the a_{2u} orbital of the porphyrin radical cation of Cpd I, while at the same time the H• moiety is spin-paired to the odd electron on the oxo group. This spin pairing[418] is shown in Figure 59a by the curved line connecting the single electrons on H• and on the oxo group of Cpd I; the RH moiety in this state is a cation radical. With these bonding changes, the oxidation state of iron in Ψ_r* is reduced from Fe^{IV} to Fe^{III}. The assignment of oxidation state for Ψ_r* is based on the electron counting formalism explained in section 3.2. Thus, in Ψ_r*, the oxo is spin-paired to H•, and as an "OH" group its oxidation state becomes HO^{1-} so that iron changes to oxidation state Fe^{III}. As such, by linking the H• atom to the iron-oxo, the oxo group accepts formally one electron from the H-atom, and hence, it "releases" back one of the two electrons it took from Fe in Cpd I. This in addition to the actual charge transfer from R–H in Ψ_r* makes this state "twice reduced" relative to the reactant state, Ψ_r. As such, along the reaction coordinate, which involves R•••H bond elongation, and O•••H and O•••R shortening, the Ψ_r* state gets stabilized by forming O–H and O–R bonds and becomes the ground state of the ferric-alcohol product, Ψ_p, which involves three electrons in iron dπ orbitals (drawn as heavy dots on iron). Thus, Ψ_r* is the electronic image state of Ψ_p and, hence, the two states correlate along the reaction coordinate. The same applies in fact to the relationship between the states Ψ_r and Ψ_p*, with the latter being the electronic image of the former, with spin pairing between R and H, but with the geometry of the product. As such, the two states correlate with one another along the reaction coordinate.

Let us turn now to the intermediate curves, which are labeled as Ψ_I*(III) in Figure 59a and Ψ_I*(IV) in Figure 59b. In the VB representation, the C–H bond in R–H is unpaired into a triplet spin,[402,416,419] and the H• atom is spin-paired to the odd electron on the oxo group. As already discussed above, this spin pairing changes the oxidation state of the OH moiety to −1 and "releases" one electron to the heme. In the case of Ψ_I*(III), this electron is given to iron that is reduced to Fe^{III}, while, in Ψ_I*(IV), the electron shifts from the iron into the porphyrin hole, while the iron keeps two unpaired electrons in the dπ orbitals, thereby remaining in

oxidation state Fe^{IV}. In each case, the entire intermediate curve corresponds to a single VB structure with energy changes reflecting the geometric changes along the reaction coordinate. Thus, at the intermediate stage, it corresponds to an H-abstraction intermediate (for example, **I** in Figure 35), composed of an iron-hydroxo/R• species, where the iron-porphyrin is in two electromeric states ($Por^{•+}Fe^{III}OH$ or $PorFe^{IV}OH$). Similarly, at the product geometry, this state involves a charge transfer from porphyrin to the R–O linkage in Figure 59a or from the iron d-electrons to the same linkage in Figure 59b.

The final energy profile for the mechanism in Figure 59 involves VB mixing of the three curves. The mixing that is shown by the bold-brown lines leads to an avoided crossing and generates the corresponding transition states; initially for the H-abstraction step, TS_H, and subsequently for the radical rebound via TS_{reb}. The latter transition state exists only on the quartet energy profile (see above in section 4.3, and discussion below), as has already been rationalized using VB diagrams.[26,83] To denote the spin-state condition on the existence of this particular barrier, a question mark is attached to the symbol TS_{reb} in Figure 59.

The TS_H species is expected to have a polar character, with some charge transfer, Q_{CT}, from the alkane moiety to the iron-oxo species, due to the secondary VB mixing of the Ψ_r* curve into TS_H. Indeed, the many transition states calculated so far possess Q_{CT} values between 0.32 and 0.49.[174] All other qualitative features, which are predicted by the diagram, are apparent in the computational results of P450 hydroxylation, summarized in papers and reviews.[26,30,32,83] Furthermore, it is seen from Figure 59 that *the intermediate curve catalyzes the two-electron oxidation process by cutting through the high barrier*, which would have been encountered otherwise due to the crossing of the main curves that emanate from reactants and products, Ψ_r and Ψ_p. Therefore, in addition to revealing the source of the polar effect in TS_H, the VB diagram accounts nicely and simply for the findings, both experimentally[420] and computationally,[32] of a stepwise rather than a concerted process. Additionally, the model predicts that there should exist a correlation between the degrees of C–H bond breaking and O–H bond formation in TS_H, on the one hand, and the corresponding reaction barriers on the other hand. Such correlations have been observed before (see section 5.1),[170,174] and they reflect the mechanism of an avoided crossing that leads to the transition states: Thus, the crossing of the reactant (Ψ_r) and intermediate (Ψ_I*(III) or (Ψ_I*(IV)) curves, to establish TS_H, is caused by C–H stretching and O–H approach, and hence the barrier correlates with the progress of these geometric features compared with the reactants state.

Let us next discuss the spin-state effect on the rebound process.[26,83] The VB diagrams for the rebound step on the two spin states of the $PorFe^{IV}OH/R•$ electromer are shown in Figure 60. They start from the center of Figure 59b, where the intermediate curve Ψ_I(IV) has to cross the Ψ_r* curve that descends toward the product state Ψ_p. To clarify the argument, we show small orbital occupation diagrams in Figure 60 describing the three upper orbitals of the d-block. Figure 60 shows the situation for the quartet spin state on the left and for the doublet spin curve on the right. At the intermediate geometry, the Ψ_r* state includes an electron transfer from the radical moiety R• to the iron-hydroxo complex. To maintain the quartet spin state, the electron is transferred into the high lying $\sigma^*_{z^2}$ orbital, while, for the
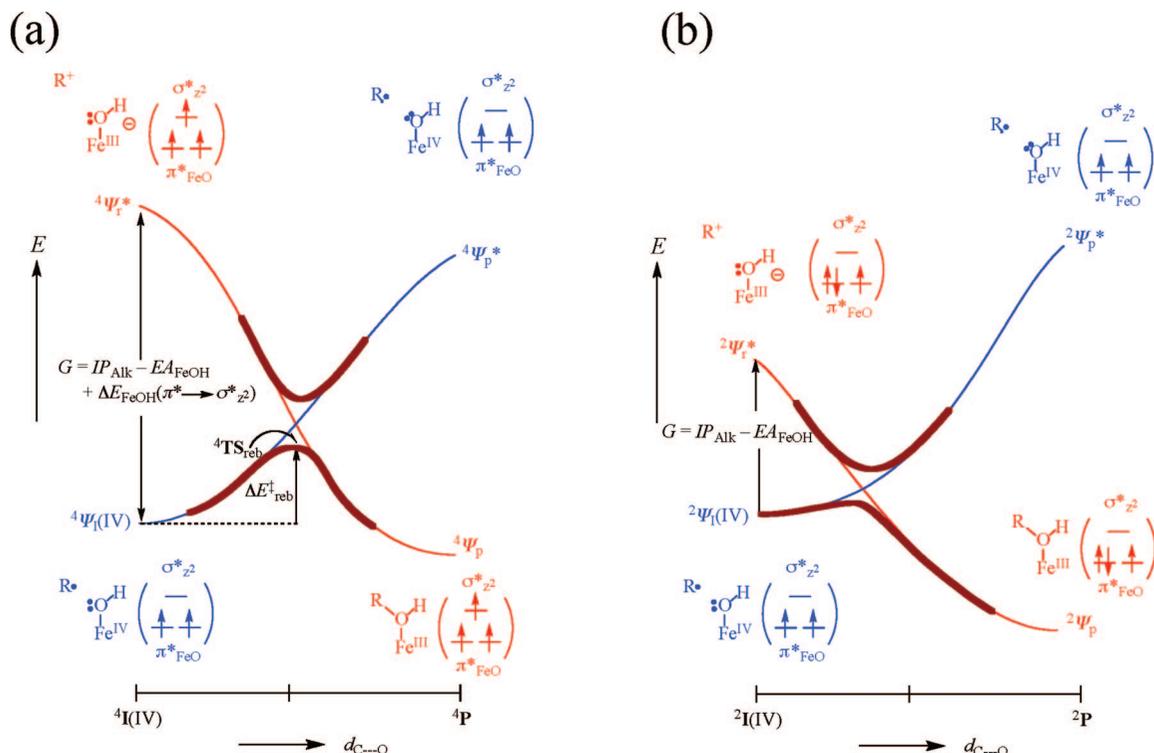
Figure 60. VB diagrams showing the energy profiles of the rebound process for the PorFe(IV)FeOH/R$^{\bullet}$ intermediates in the quartet (a) and doublet (b) spin states. It is seen that the promotion gap, G, for the quartet rebound is larger than that for the doublet by the promotion energy of an electron from the π^* orbital to the $\sigma^*_{z^2}$ orbital, labeled in part a as $\Delta E_{FeOH}(\pi^* \rightarrow \sigma^*)$. As a result of the larger gap, the quartet state will generally exhibit a barrier for rebound, while the doublet state will rebound in a barrier-free manner. These scenarios are shown by the bold energy profile that reflects the VB mixing of the two curves. Reprinted with permission from ref 174. Copyright 2008 American Chemical Society.

doublet state, it goes into a lower lying π^* orbital. The height of the crossing point depends on the initial gap between the two curves: the larger this energy gap, the higher the crossing point. Consequently, in the quartet spin state, where the gap is large[26,83] the VB mixing will generate a barrier for rebound. On the other hand, in the doublet spin state, the gap is much smaller (by 26 kcal mol^{-1}) so that the avoided crossing yields a barrier-free radical rebound. This trend has been observed so far in all computational studies of this reaction.[26,32,83]

The reason why the quartet rebound barrier is significantly higher for the $^4\Psi_I(III)$ electromer compared with the $^4\Psi_I(IV)$ electromer has been analyzed before in VB terms.[26,83] The chief reason is the amount of mixing between the VB-state curves (the avoided crossing energy) at the crossing point of Figure 59. Thus, in the case shown in Figure 60, the $^4\Psi_I(IV)$ and $^4\Psi_r^*$ states differ by a single electron transfer from the radical to the $\sigma^*_{z^2}$ orbital, whereas, in the case of the $^4I(III)$ rebound, the $^4\Psi_I(III)$ and $^4\Psi_r^*$ states differ by two electron transfers, one from the radical to $\sigma^*_{z^2}$ and the other from the filled π^*_{FeO} orbital to the singly occupied a$_{2u}$ orbital. Therefore, the avoided crossing energy for the $^4I(IV)$ rebound is large, since the two states mix strongly, while for the $^4I(III)$ rebound the two states mix weakly and the avoided crossing energy is small;[418] the result is a small barrier for the $^4I(IV)$ rebound and a significantly larger one for the $^4I(III)$ rebound. It is apparent from the various C$-$H hydroxylation (and C$=$C epoxidation) cases in this review that these trends are manifest in all the computations, showing that the $^4I(III)$ intermediate is the longest-lived species during the reaction.

All in all, the VB diagram model provides a good rationale for the computational and experimental findings on various aspects of P450 hydroxylation.
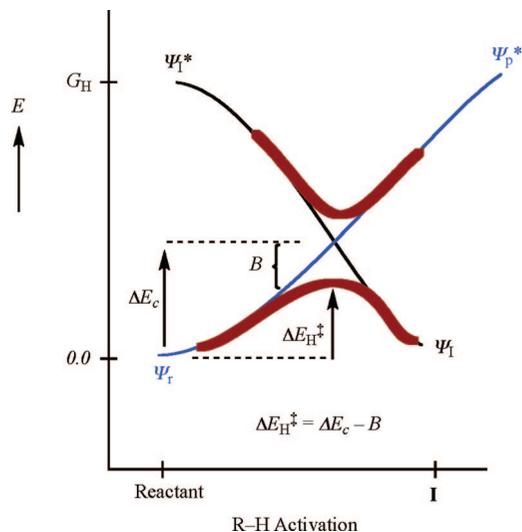
Figure 61. A VB state-correlation diagram (VBSCD) showing the formation of the barrier for the H-abstraction step of alkane hydroxylation by P450, without specifying the spin state. G_H is the promotion energy gap, B is the resonance energy of the transition state due to VB mixing, and $\Delta E_H^{\ddagger}$ is the resulting barrier. Reprinted with permission from ref 174. Copyright 2008 American Chemical Society.

5.2.2. Quantitative Aspects of Valence Bond Modeling for C$-$H Hydroxylation

Let us now turn to a more quantitative aspect of the modeling. Figure 61 shows the VB diagram for the first step in the reaction mechanism, i.e., C$-$H bond activation. The two curves (Ψ_r and Ψ_I^*) are initially separated by an energy gap G_H (the so-called promotion gap[402,416]) and cross one

another along the reaction coordinate to generate a transition state (**TS**$_H$) through VB mixing and avoidance of the crossing.

The barrier for the reaction, $\Delta E_H^{\ddagger}$, can be expressed as the height of the crossing point, ΔE_c, minus the avoided crossing or resonance energy (B) of **TS**$_H$ due to the VB mixing:

$$\Delta E_H^{\ddagger} = \Delta E_c - B \qquad (5)$$

The height of the crossing point in turn can be expressed as a fraction (f) of the promotion gap (G_H):

$$\Delta E_c = f G_H \qquad (6)$$

Combining these two equations, an expression for the barrier heights is derived:

$$\Delta E_H^{\ddagger} = f G_H - B \qquad (7)$$

This effective equation includes the BEP effect of the reaction energy difference (between Ψ_r and Ψ_1^* in Figure 61) directly into the f factor, and although there exist explicit expressions which depend on all the variables in the diagram,[402,416,419] eq 7 is the most convenient one to use.

Based on previous modeling of the barrier for H-abstraction processes, it has been established[419] that f is approximately 0.3, while B is given by eq 8:

$$B \approx {}^1\!/_2 (BDE_W) \qquad (8)$$

Here BDE$_W$ stands for the bond dissociation energy of the weakest bond among the two that are formed and broken during the reaction: either the C−H or FeO−H bond. Thus, B can be estimated from the raw data of BDEs.

Finally, the promotion gap G_H, which is related to the triplet unpairing of the C−H bond,[174,415] was shown to be well approximated by eq 9, as twice the in situ bond strength, D_{CH}, of the activated C−H bond:

$$G_H = 2 D_{CH} = 2(BDE_{CH} + RE_{Alk\bullet}) \qquad (9)$$

The in situ bond strength is given as the sum of the BDE and the reorganization energy of the alkyl radical, RE$_{Alk\bullet}$, from its equilibrium geometry to its geometry in the alkane. The only missing quantity is RE$_{Alk\bullet}$, which can be calculated easily for a given radical. Thus, in principle, once we define the various alkane substrates, all the quantities necessary for calculating the barriers for C−H activation from raw fundamental data are available for the VB scheme outlined above.

This model was applied to a series of alkanes, depicted in Scheme 23, which are the same substrates used in eq 4 to find a correlation between barriers and BDE$_{CH}$ properties. The computed BDE$_{CH}$ and RE$_{Alk\bullet}$ values are collected in Table 3, while the corresponding barriers are given in Table 4.[174] It is seen that the BDE$_{CH}$ values and the associate barriers range over 20 and 17 kcal mol^{-1}, respectively, and hence these substrates form a good test set to attempt a quantitative evaluation of the barrier from raw data.

To this end, one could use eqs 6−9 to calculate the barriers or to correlate the barriers, e.g., with D_{CH}. At the outset, it should be clear from the above discussion that neither f nor B can be constant for all the alkanes. Nevertheless, it would be interesting to ascertain whether the proposed values of f

Scheme 23. Alkane Substrates, 1−10, Used for the Application of the VB Model To Predict H-Abstraction Barriers
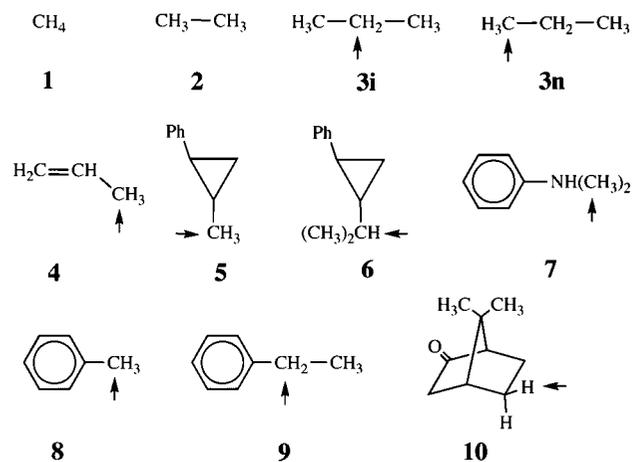
~ 0.3 and B (eq 8) will emerge from barriers by applying eq 7 in conjunction with eq 9. There are two ways to do so: Either one can use the equations and find the f values that reproduce the calculated barriers, or one can apply the equations with $f = 0.3$ for all the reactions and find the corresponding B values that are required to reproduce the barriers. This was done and is discussed here only for the barriers at the highest level (DFT(B3LYP/B2)).

In the first method of modeling, the B values based on eq 8 are all ≤ 44.45 kcal mol^{-1}, with the smallest value being 41.4 kcal mol^{-1}. Using these values and G_H according to eq 9, one obtains an average value of $f_{av} = 0.285 \pm 0.015$. Thus, while f is not constant, its values cluster around the theoretical value of 0.3. The largest deviations are found for substrates **1** and **7**. Using $f_{av} = 0.285$ to calculate barriers, the resulting values are within 2.4 kcal mol^{-1} of the B3LYP/B2 computed ones.

In the second method of modeling, the computed B values cluster around an average $B_{av} = 46.39$ kcal mol^{-1} with a scatter of ± 2.24 kcal mol^{-1}. The largest deviations are for the substrates **1** and **7**, for which the so found B values are 3.97 kcal mol^{-1} smaller and 4.76 kcal mol^{-1} larger than the average quantity. Thus, once again, while the B values are not constant, they still cluster around an average value with a moderate scatter, as suggested by eq 8.

It is clear therefore that the two ways of modeling the barriers lead to f and B values that are not too far from each other. Therefore, we calculated all the barriers obtained with eq 7 using $f = 0.3$ and $B_{av} = 46.39$ kcal mol^{-1}. A plot of the resulting barriers against the computed ones gives a correlation coefficient $r \approx 0.9$. Table 5 displays the DFT and VB-derived barriers. The correlation coefficient between the two sets is modest, but considering the fact that only raw data were used, the correlation is good. We may therefore conclude that the VB models can predict barrier heights of P450-catalyzed reactions in the gas phase quite well.

The VB-derived barriers are close to the DFT-computed ones, with an average deviation of 1.53 kcal mol^{-1}. The greatest deviations (Table 5) are found for methane (**1**), where the VB equation underestimates the barrier by 4.0 kcal mol^{-1}, and N,N-dimethylaniline (**7**), for which the barrier is overestimated by 4.7 kcal mol^{-1}. Without these outliers, the VB equation predicts the computed barriers to within 0.90 kcal mol^{-1} (rms). The two outliers have clear physical origins, which are explained below.

Table 3. Calculated BDE$_{CH}$, BDE$_{OH}$, and RE$_{Alk}$. Data for the Alkanes 1−10 in Scheme 23 and for the PorFeO−H Species^a

	BDE$_{CH}$				RE$_{Alk}$.^{b}	
	B1		B2		ΔE	
	ΔE	ΔE + ZPE	ΔE	ΔE + ZPE	B1	B2
methane (**1**)	112.45	102.77	111.29	101.62	7.42	6.79
ethane (**2**)	107.51	97.70	106.67	96.87	7.45	7.00
isopropane (**3i**)	103.39	93.55	102.87	93.03	7.56	7.21
n-propane (**3n**)	107.98	98.37	107.09	97.49	7.23	6.87
propene (**4**)	91.48	82.89	91.39	82.80	17.69	16.69
trans-methylphenylcyclopropane (**5**)	103.66	94.32	102.96	93.62	9.00	8.30
trans-isopropylphenylcyclopropane (**6**)	95.92	86.56	95.95	86.60	9.04	8.60
N,N-dimethylaniline (**7**)	95.72	86.99	94.81	86.08	8.14	7.94
toluene (**8**)	94.45	86.08	94.16	85.80	13.15	12.27
phenylethane (**9**)	91.30	82.44	91.31	82.45	19.32	18.38
camphor (**10**)	104.08	94.58	103.39	93.89	7.90	7.61
PorFeO−H	88.59^c	81.81	95.69	88.91		

^{a} All data in kcal mol^{-1} and calculated with basis set B1 or B2.[174] ZPE data with basis set B1. B1 is LACVP(Fe)/6-31G(rest); B2 is LACV3P+(Fe)/6-311+G*(rest). ^{b} These are absolute magnitudes. ^{c} BDE$_{FeO-H}$.

Table 4. DFT(B3LYP)-Only Barriers, $\Delta E^{\ddagger}$ and $(\Delta E + ZPE)^{\ddagger}$, Relative to Separated Reactants, ^{4,2}Cpd I + Alk-H, for TS$_H$^{a,b}

		B1		B2	
		$\Delta E^{\ddagger}$	$(\Delta E + ZPE)^{\ddagger}$	$\Delta E^{\ddagger}$	$(\Delta E + ZPE)^{\ddagger}$
HS Data					
1	methane	26.77	22.81	26.87	22.91
2	ethane	21.32	17.71	21.02	17.41
3i	isopropane	19.05	15.39	19.46	15.80
3n	*n*-propane	21.47	17.84	21.12	17.48
4	propene	14.68	11.70	15.93	12.95
5	*trans*-methyl phenylcyclopropane	18.28	14.73	18.09	14.53
6	*trans*-isopropylphenylcyclopropane	16.03	12.46	17.03	13.46
7	*N,N*-dimethylaniline	7.66	5.04	8.17	5.54
8	toluene	14.81	11.79	15.45	12.43
9	phenylethane	14.57	11.19	15.93	12.55
10	camphor	17.96	14.21	18.27	14.51
LS Data					
1	methane	25.87	22.24	25.97	22.34
2	ethane	20.23	16.65	18.88	15.30
3i	isopropane	17.68	14.17	17.36	13.85
3n	*n*-propane	20.36	16.70	18.85	15.19
4	propene	15.15	12.13	15.85	12.82
5	*trans*-methyl phenylcyclopropane	17.31	13.90	16.54	13.13
6	*trans*-isopropylphenylcyclopropane	15.05	11.64	15.68	12.27
7	*N,N*-dimethylaniline	6.65	4.53	7.11	4.99
8	toluene	14.78	11.73	15.10	12.05
9	phenylethane	14.21	10.69	15.00	11.47
10	camphor	18.39	14.02	20.25	15.88

^{a} All energies are in kcal mol^{-1}. Basis sets B1 and B2 are defined in Table 3.[174] ^{b} HS, quartet (high-spin); LS, doublet (low-spin).

Table 5. B3LYP/B2//B1 Barriers (kcal mol^{-1}), $\Delta E^{\ddagger}$$_{(DFT)}$, and Barriers Modeled by VB Theory, $\Delta E^{\ddagger}$$_{(VB)}$

alkane	$\Delta E^{\ddagger}$$_{(DFT)}$^{a}	$\Delta E^{\ddagger}$$_{(VB)}$^{b}
1	22.63	18.66
2	16.35	15.93
3i	14.82	13.75
3n	16.34	16.23
4	12.89	13.31
5	13.83	14.76
6	12.86	10.73
7	5.27	10.02
8	12.24	12.45
9	12.01	14.10
10	15.20	14.51

^{a} These are averaged HS and LS barriers.[174] ^{b} The barriers were calculated using eq 6: $\Delta E^{\ddagger}(VB) = 0.3G_H - B_{AV}$; $B_{AV} = 46.39$ kcal mol^{-1}.

5.2.2.1. Deviations for Methane.

As shown by the spin density distribution,[174] the ^{2,4}TS$_H$ transition states for methane hydroxylation are the only ones in the computed series with a clear ^{2,4}TS$_H$(III) character, having a porphyrin cation radical moiety; all other substrates exhibit ^{2,4}TS$_H$(IV) species. Referring back to Figure 59, we can see that while G_H is virtually the same in both cases, the two states Ψ_I^*(IV) and Ψ_I^*(III), which cross Ψ_r and generate the corresponding intermediate, are different. In Figure 59b, the state, Ψ_I^*(IV), that crosses Ψ_r down to the PorFe^{IV}OH/R$^{\bullet}$ intermediate, has an odd electron that is localized on the oxo group. By contrast, Ψ_I^*(III) in Figure 59a has two resonance structures that delocalize the odd electron on the oxo group between two planes. Therefore, the bond coupling interaction of the spin paired moieties, H$^{\bullet}$−$^{\bullet}$O, in Ψ_I^*(IV) is strong, while in Ψ_I^*(III) the same interaction is weak. The result is that Ψ_I^*(IV) establishes a low energy crossing point (smaller f in eq 6), while Ψ_I^*(III) establishes a higher crossing point (larger f in eq 6). Consequently, the barrier for the Fe^{III} electromer will be generally higher than the one for the Fe^{IV} electromer in the gas phase. Furthermore, this higher barrier is expected to be attended by TS$_H$(III) species that lie later
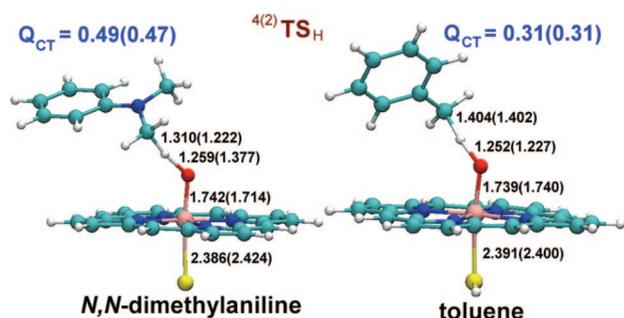
Figure 62. DFT(B3LYP/LAVCP) transition state structures (Å) and degree of charge transfer from the alkane to Cpd I, Q_{CT}, for *N,N*-dimethylaniline vs toluene. Q_{CT} is determined from NBO charges computed with the LACV3P+(Fe)/6-311+G*(rest) basis.[174] The charges are given in the quartet (doublet) states, respectively.

on the H-transfer coordinate, having greater extents of elongations of the R−H and Fe−O bonds than the corresponding TS_H(IV) species. Coming back to methane, we note that since the strength of the C−H bond in methane is the largest in the series, it requires extensive bond stretching to reach the **TS**. And indeed, the computed C−H bond lengths in the ^2TS_H (^4TS_H) species of methane are highly stretched, being 1.489 (1.499) Å, respectively. With such a stretched C−H bond, the transition states will always be of TS_H(III)-type.[174] This in turn means that a larger f value is required in the VB model to reproduce more accurately the corresponding barrier for methane hydroxylation, and this is the reason why the use of $f = 0.3$ (in Table 5) underestimates the barrier; an f value of 0.318 would have given a perfect barrier.

We note that, notwithstanding methane, in the gas phase we seldom expect to find the TS_H(III) species being lower than the TS_H(IV) species. However, in a polarizing medium such as the protein, the Fe^{III} electromeric states are stabilized compared with the Fe^{IV} states and become the lower energy species. Indeed, the **TS** geometries in the protein show larger extents of C···H stretching and O···O formation than the corresponding gas-phase species. Camphor is a case in point, where the QM/MM optimized transition states are of the $^{4,2}TS_H$(III) variety; their C−H bonds are longer compared with those of the corresponding gas-phase species.[59]

5.2.2.2. Deviations for *N,N*-Dimethylaniline−Polar Effects in P450 Hydroxylation. We recall that the DFT computed barrier for C−H bond activation of *N,N*-dimethylaniline (DMA, **7**) is very small and deviates from the VB barriers of the other substrates. In fact, looking at Tables 3 and 5, it is seen that the BDE_{CH} value of **7** is much like that in toluene, **8**, but the corresponding barrier is 7−8 kcal mol^{-1} lower! The reason is a strong polar effect in the corresponding TS_H species, which was discussed above by reference to Figure 59 and which in the specific case of DMA is apparent in Scheme 13.

Indeed, Figure 62 shows side by side the TS_H species of toluene and DMA along with the corresponding Q_{CT} values (calculated with the NBO method), i.e., the charge transferred from the alkane to Cpd I in the transition state. It is seen that the Q_{CT} value for DMA is close to 0.5 and is significantly larger than the one for toluene. Such a large Q_{CT} value means a very strong mixing of the radical cation state into the TS_H species (see Ψ_r^* in Figure 59). As already discussed in Scheme 13, this mixing is enhanced by the conjugation between the lone pair and the empty orbital on carbon in the N−CH$_2^+$ moiety.
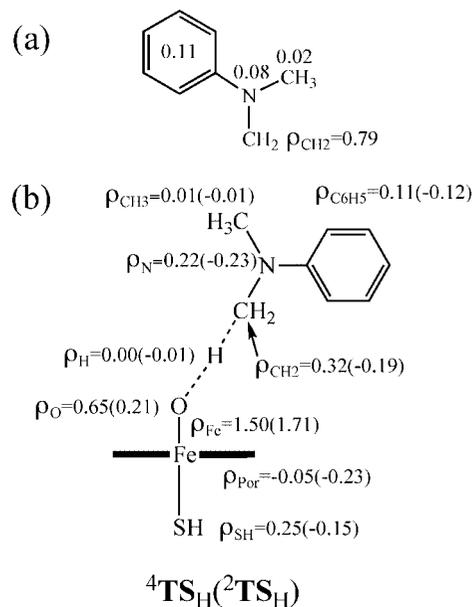
Figure 63. DFT(B3LYP/LACV3P+/6-311+G*) computed spin densities for the *N*-methylene-*N*-methylaniline radical and the TS_H species for hydrogen abstraction by Cpd I.[174] The spin densities are given in the quartet (doublet) states, respectively.

The extensive lone pair conjugation in the TS_H species can be deduced also by inspecting the amount of radical development at the nitrogen atom in Figure 63. It can be seen that the nitrogen develops spin density (ρ) in the TS_H species to an extent that is even larger than that in the isolated radical species, the *N*-methylene-*N*-methylaniline radical. This enhanced conjugation increases the resonance energy of the transition state[166] and thereby lowers the barrier. Thus, the deviation of **7** in Table 5 is caused by the B factor that gets larger (ca. 51 kcal mol^{-1} vs 46.39 kcal mol^{-1}) due to the enhanced conjugation of the nitrogen lone pair in the transition state. Indeed, as discussed above in the *p*-X-DMA series,[167] the H-abstraction barriers exhibit a Hammett correlation that reflects the *para* substituent effect on the ability of the nitrogen to conjugate with the stretched C−H bond in TS_H. In fact, in the absence of this conjugation for X = NO$_2$,[167] the barrier for H-abstraction is close to the one calculated here for **7**, using the standard B value of 46.39 kcal mol^{-1} for the series.

5.2.3. Comments on the VB Modeling of the Rebound Step

A quantitative modeling of the rebound has not been attempted yet. However, the utility of the model was demonstrated in its ability to predict the observed rearrangement patterns, notably in the series of radical clock substrates studied by Newcomb[26,30,32,83] but not only in this series. Thus, the VB model of the rebound in Figure 60 was used[83] to predict and rationalize the variation of the rebound barrier with the change of the proximal axial ligand from thiolate to imidazole and to PhSO$_3$, and for the change of the metal from Fe to Ru. As can be seen from Figure 60, the promotion gap for the HS rebound depends also on the energy difference between the low lying π^*_{FeO} orbital and the higher lying $\sigma^*_{z^2}$ orbital. The latter orbital is affected by the antibonding interaction of the metal with the axial ligand. Therefore, in heme complexes, the $\sigma^*_{z^2}$ orbital is less antibonding and lower in energy, thus lowering the rebound

barrier, while, in the Ru-Cpd I, the $\sigma^*_{z^2}$ orbital lies very high, since Ru is a much stronger binder of ligand than Fe, and hence the rebound barrier is much larger. Another prediction of the model in Figure 60 is that different P450 isozymes should exhibit different amounts of rearrangement for a given substrate as a result of the stronger or weaker binding strength between iron and the cysteine ligand of the particular isozyme. There is ample evidence in the literature that the amount of rearrangement for a given probe substrate depends on the identity of the P450 isozyme, and while this does not prove the correctness of the prediction, still the prediction is clearly testable by a combination of QM/MM calculations and experimental means.

5.3. Summary of Valence Bond Modeling of P450 Reactivity

All in all, the VB diagram model provides a good rationale for the computational and experimental findings on various aspects of P450 hydroxylation. VB modeling of the H-abstraction barrier shows, according to eqs 6−9, that the promotion gap depends on BDE$_{CH}$, which can be used to estimate barriers quickly from experimental BDE values, without having to compute the RE$_{Alk}$. quantities. In fact, eq 3 has the form of the barrier equation, eq 6, with $f = 0.293$ and $B = 39.6$ kcal mol^{-1}, except that the promotion gap will here simply be 2 BDE$_{CH}$. In any event, the barrier equation shows the usefulness of BDE$_{CH}$, but the correlation with these quantities has little to do with the original BEP correlation; it is a correlation with the promotion gap that leads to the barrier.

While this has not yet been done, the diagrams in Figures 59 and 60 can be used with small modifications (using the π-charge transfer state for Ψ_r^* and a triplet excited state for the intermediate curve) to discuss the reactivity and mechanistic aspects of C=C epoxidation. The mechanism of arene hydroxylation can also find a rationale in the VB diagram of Figure 59, with the modification that the excited state Ψ_r^* is a charge transfer state involving an electron transfer from the π-system of the arene to Cpd I and is thus likely to be sufficiently low-lying to lead to a transition state with a hybrid radicalar-cationic character in the arene, as discussed above (section 4.5) in conjunction with the corresponding DFT(B3LYP) and DFT(B3LYP)/MM calculations.[33,34,38,60] Similarly, heteroatom oxidation, which involves a concerted mechanism dominated by the LS state, can be accommodated with a two-state scenario, involving the reactants curve and the curve that starts at the charge transfer state, Ψ_r^*, which will be low-lying for sulfides and amines. As such, the VB method can provide a unified framework for the reactivity of P450 enzymes.

6. Summary of Reactivity Patterns

Theory shows that Cpd I is a versatile reagent with uncanny solutions for complex reactions. On the one hand, it is a two-state triradicaloid reagent (or even a four-state reagent if one counts the pentaradicaloid states in section 3.7.5), and therefore, it participates in radical-type reactions that involve bond activation processes on at least two spin states, with state-specific chemistry, known as TSR or MSR.[26,32,146] At the same time, Cpd I is highly electron deficient, having an effective overall oxidation state of V. As such, it participates also in electrophilic reactions,

electron transfer reactions, and bond-reorganization coupled electron transfer reactions, such as the proton-coupled electron transfer (PCET) mechanism of P450 StaP.[58] Taken together, Cpd I is a chameleon species that adapts its electronic structure to the protein environment, to additives such as water molecules,[61,131] ions, etc., and to the substrate with which it reacts.[58] These features are revealed by both QM-only and QM/MM calculations. Theory shows that kinetic isotope effect (KIE) measurements can probe TSR and the reactive spin states of Cpd I, at least for C−H hydroxylation.[166,167] In addition, VB modeling provides an understanding of the mechanism and reactivity patterns. This has been demonstrated for C−H hydroxylation,[174] but the principles of this model can clearly be applied to other reactions as well.

An important aspect that is revealed by the QM/MM calculations is the biological role of small molecules that may affect the reactivity of the enzyme. Chief among these is the water molecule. Placed in strategic positions inside the enzyme cavity[58,61,131] or coming from the surface or the deep of the protein,[73] water molecules act as promoters of otherwise difficult transformations by lowering reaction barriers[58,61,131] or by bridging the distance separation between the immobilized substrate and Cpd I reactants.[58]

7. Concluding Remarks

With current computational capabilities, QM(DFT)-only and QM(DFT)/MM calculations can be used to investigate models of cytochrome P450 systems, with typically around 50−100 and up to about 200 atoms (including heme and substrate, plus a few neighboring residues and water molecules) in the QM region and with thousands of atoms (typically around 25000) in the MM region. It has become apparent by extensive validation studies in several groups that the current technology is good enough for meaningful investigations of metalloenzymes such as cytochrome P450, when using B3LYP as QM(DFT) component, with medium-size basis sets for geometry optimization and larger basis sets for single-point energy evaluations; realistic QM/MM calculations employ electronic embedding and established protein force fields as MM component. Higher accuracies are expected by QM/MM methods with high-level correlated ab initio QM methods (e.g., MR-CI). Currently, only a few such calculations have been published, on Cpd I and other intermediates[77,78] in the catalytic cycle of P450$_{cam}$ and related enzymes.[79] This limited set of calculations has given results that are rather similar to those from DFT(B3LYP)/MM, but of course more tests are required to properly assess the quantitative accuracy of DFT/MM methods in this area.

The QM/MM methods have advanced to the point that the complete catalytic cycle of a metalloenzyme can now be studied. In fact, QM/MM calculations have been carried out for all intermediates of cytochrome P450$_{cam}$ and for most of the reactions in the catalytic cycle, including the mechanism of electron transfer,[76] the formation of Cpd 0 and of Cpd I in the wild-type and various mutant enzymes,[159,161,246] and the various reactions catalyzed by the active species, Cpd I, for example in bond activation (C−H, C=C, aryl, etc), oxo-transfer (sulfoxidation, N-oxidation, etc.), and C−C bond coupling reactions. Furthermore, they have addressed the in-silico design of new mutants, such as the selenocysteine mutant of P450,[61] which has already received attention

by the experimental community.[64] Coupled with MD simulations, it is now possible to probe the role of the reductase and to pinpoint the factors that may regulate and make for a functional catalytic cycle.[283]

There remain open-ended issues, of course, that will require further improvements in the theoretical tools. The DFT(B3LYP)/MM data in Figure 37 seem to suggest a solution for one of the puzzling open problems, i.e., the inability to observe the active species Cpd I in P450$_{cam}$. The proposed explanation depends on the difference between the barriers for formation of Cpd I from Cpd 0 and for H-abstraction from camphor. However, since a question mark still looms over the accuracy of DFT barriers, these two reactions should be investigated by more accurate ab initio QM(MR-CI)/MM calculations, which will hopefully provide reliable barriers for Figure 37 and allow us to draw more definitive conclusions on the elusive nature of Cpd I in P450$_{cam}$.

Another key aspect is the role of additives on the reactivity of P450 enzymes. DFT(B3LYP)/MM calculations have so far demonstrated nicely that Cpd I is a chameleon species and that various additives and different substrates can dramatically affect its reactivity patterns (see, for example, the role of Wat$_{903}$ in P450$_{cam}$[61,131] and of Wat$_{789}$ in P450 StaP[58]). The potential of the latter concept is far-reaching and awaits exploration by further theoretical and experimental studies.

The issue of selectivity is still a sore point, since it is still hard to understand, e.g., the regioselectivity of C—H vs C=C activation or of N-demethylation vs sulfoxidation. As pointed out above, the DFT(B3LYP)/MM results for the C—H/C=C competition in propene and cyclohexene oxidation do not reproduce the experimental data.[57] This might require QM(MR-CI)/MM calculations to ascertain whether the problem is associated with the inaccuracy of B3LYP barriers or with the lack of sufficient conformational sampling.

The DFT(B3LYP) and DFT(B3LYP)/MM calculations invariably find two or more low-lying spin states, which contribute to reactivity and product formation (in a state-specific manner). It has indeed been shown that TSR and MSR scenarios can rationalize many puzzling experimental observations on P450 reactivity.[18,32,285,286] The actual kinetics of these processes will depend on the probability of spin crossover. There is a clear need for more quantitative theoretical studies of spin crossover in order to better assess the role of TSR and MSR in P450 chemistry. There is a further need for studying spin-state perturbations, e.g., by external magnetic and/or electric fields, that may affect the stereospecificity and selectivity of P450 enzymes.[25] This requires development of DFT(MR-CI)/MM methods to handle such external fields, as well as experimental studies of these phenomena.

Dynamical aspects in P450 have generally only been addressed at the classical force-field level up to now, and it would certainly be of interest to study them also at higher QM/MM levels, for example with regard to the biological role of water molecules in the various proton-relay mechanisms during Cpd 0 and Cpd I formation and in the reactions of Cpd I (e.g., with regard to the catalytic effect of Wat$_{903}$ in P450$_{cam}$ reactions[61,131] and the mechanistic role of the two water molecules during aryl—aryl coupling by P450 StaP[58,80]). In addition, there is a need for a thorough thermodynamical analysis of the whole catalytic cycle in P450 enzymes that includes the reduction steps and their dependence on the protonation state of the enzyme.

Finally, a treatment of plant P450s is still missing. Some of these P450 enzymes are interesting because they do not use exogenous oxygen, but rather activate hydroperoxides of polyunsaturated fatty acids to catalyze several remarkable reactions.[421] Investigation of these mechanisms may provide more insight into the uncanny ability of P450 enzymes to find solutions for performing complex transformations.

8. Abbreviations

AC	aldehyde complex
Alk	alkyl
AlkH	alkane
AO	arene oxide
APX	ascorbate peroxidase
Arg	arginine
Asn	asparagine
Asp	aspartic acid
B2W	a basis set consisting of Wachters basis (with polarization and diffuse f) on iron, and 6-31G* on the atoms of the first coordination shell
B3LYP	a hybrid functional used in DFT consisting of 20% Hartree−Fock and 80% Becke88 exchange combined with the Lee−Yang−Parr correlation functional
BDE	bond dissociation energy
BEP	Bell−Evans−Polanyi
BFCET	bond-formation coupled electron transfer
BHLYP	a hybrid functional used in DFT, consisting of half Hartree−Fock and half Becke88 exchange combined with the Lee−Yang−Parr correlation functional
BLYP	a pure gradient functional used in DFT, consisting of Becke88 exchange combined with the Lee−Yang−Parr correlation functional
BO	ferric benzene oxide complex
BP86	a pure gradient corrected functional used in DFT consisting of Becke88 exchange and Perdew86 correlation functional
CAM	camphor
CASSCF	complete active space self-consistent field method
CASPT2	multiconfigurational second-order perturbation theory
CB	cooperative binding
CcP	cytochrome c peroxidase
CCSD(T)	coupled cluster method with single and double excitations, and a perturbative treatment of triple excitations
CI	configuration interaction
CPA	chromopyrrolic acid
Cpd 0	ferric hydroperoxide species in heme protein
Cpd I	the active species of heme enzymes, Compound I, involving Fe(IV)O and a porphyrin radical cation
Cpd II	one-electron reduced species of Cpd I
CPO	chloroperoxidase
CYP	cytochrome P450
Cys	cysteine
D251N	a mutant of P450$_{cam}$ in which Asp$_{251}$ is replaced by asparagine (N)
def2-TZVP	a triple ζ quality basis set of Ahlrichs and co-workers (TURBOMOLE notation)
DFT	density functional theory
DFT/MM	a density functional/molecular mechanical method, a variant of QM/MM
DHA	dual hydrogen transfer
Dia	diazepam

DMA	*N,N*-dimethylaniline
DMS	dimethyl sulfide
DNA	deoxyribonucleic acid
DZ	double ζ basis set
DZP	double ζ basis set with polarization
EC	ferric epoxide complex
EEF	external electric field
EPR/ENDOR	electron paramagnetic resonance/electron–nuclear double resonance
ESEEM	electron spin echo envelope modulation spectroscopy
ESP	electrostatic potential
EXAFS	extended X-ray absorption fine structure
ext-SCys	extended cysteinate
FO	face-on
fs	femtosecond
GGA	generalized gradient approximation
Gln	glutamine
Glu	glutamic acid
Gly	glycine
H_4B	tetrahydrobiopterin
HAT	H-atom abstraction
HCB	hexachlorobenzene
HF	Hartree–Fock
HFB	hexafluorobenzene
His	histidine
HO	heme oxygenase
HOMO	highest occupied molecular orbital
HRP	horseradish peroxidase
HS	high spin
I	intermediate
Ile	isoleucine
IP	ionization potential
K	ketone
KIE	kinetic isotope effect
KSIE	kinetic solvent isotope effect
L358P	a P450 mutant created by replacement of leucine 358 (Leu_{358}) by proline
LACV3P++**	a triple ζ basis set augmented with diffuse and polarization functions describing the valence orbitals on the transition metal (the core is described by an effective core potential), combined with 6-311++G** for other atoms
LACV3P+*	a triple ζ basis set augmented with diffuse and polarization functions describing the valence orbitals on the transition metal (the core is described by an effective core potential), combined with 6-311+G* for other atoms
LACVP	a double ζ basis set describing the valence orbitals on the transition metal (the core is described by an effective core potential), combined with 6-31G on all other atoms
LDA	local density approximation
Leu	leucine
LFP	laser flash photolysis
LS	low spin
Lys	lysine
M06	M06 suite of density functionals of Truhlar and co-workers
Mb	myoglobin
MD	molecular dynamics
MM	molecular mechanics
MR-CI	multireference configuration interaction
MSR	multistate reactivity
NADH	reduced nicotinamide adenine dinucleotide
NIH shift	a chemical rearrangement where a hydrogen atom on an aromatic ring undergoes an intramolecular migration, primarily during a hydroxylation relation
NOS	nitric oxide synthase
ns	nanosecond

OLYP	a pure gradient functional used in DFT, consisting of Handy's OPTX modification of Becke's exchange combined with the Lee–Yang–Parr correlation functional
P	products (in benzene hydroxylation, P also refers to phenol)
P450	cytochrome P450 enzymes
PBE	Perdew, Burke, and Ernzerhof exchange-correlation functional
PCET	proton-coupled electron transfer
PDX	putidaredoxin
PHBH	*p*-hydroxybenzoate hydroxylase
pK_a	$pK_a = -\log K_a$, K_a is the dissociation constant of an acid
Por	porphyrin or porphine
PPIX	protoporphyrin IX, the native porphyrin for heme enzymes
ps	picosecond
PT	proton transfer
Q_{CT}	charge transfer
QM	quantum mechanics
QM/MM	a hybrid method that couples a quantum mechanical (QM) method and a molecular mechanical (MM) method. The QM method can be DFT, hence QM(DFT)/MM.
QSAR	quantitative structure–activity relationship
R	reactant state
RC	reactant complex
RE	reorganization energy
RI	resolution of identity
ROB3LYP	restricted open-shell B3LYP
SC	suicidal complex
SCys	cysteinate
SET	single electron transfer
SMe	methyl mercaptide
SO	side-on
SSM	single-state mechanism
STA	staurosporine
SV(P)	split-valence basis set with valence polarization of Ahlrichs and co-workers
T252A	a P450 mutant created by replacement of threonine 252 (Thr_{252}) by alanine
T252G	a P450 mutant created by replacement of threonine 252 (Thr_{252}) by glycine
T252S	a P450 mutant created by replacement of threonine 252 (Thr_{252}) by serine
T252V	a P450 mutant created by replacement of threonine 252 (Thr_{252}) by valine
T252X	a P450 mutant created by replacement of threonine 252 (Thr_{252}) by amino acid X
Thr	threonine
TMA	trimethylamine
Trp	tryptophan
TS	transition state
TS_E	epoxidation transition state
TS_H	hydroxylation transition state
TSR	two-state reactivity
Tyr	tyrosine
TZ	triple ζ basis set
TZVP	a valence triple ζ with valence polarization basis set of Ahlrichs and co-workers
TZVPP	a valence triple ζ with valence multiple polarization basis set of Ahlrichs and co-workers, including 2p, 1d polarization functions for H, and 2d, 1f polarization functions for heavy atoms
UB3LYP	unrestricted B3LYP
VB	valence bond
Wat	water molecule
WT	wild type (native enzyme)
ZPE	zero point energy

9. Acknowledgments

10. Note Added in Proof

A recent paper used DFT/MM calculations for the mechanism of Cpd I formation in CYP51. The authors found a heterolytic mechanism like the one described in ref. 70 for HRP, alongside an energetically close hybrid homolytic-heterolytic mechanism like the one presented here. See: Sen, K.; Hackett, J. C. *J. Phys. Chem. B* **2009**, *113*, 8170-8182.

11. References

(1) Raymond, J.; Segrè, D. *Science* **2006**, *311*, 1764–1767.
(2) Omura, T.; Sato, R. *J. Biol. Chem.* **1962**, *237*, 1375–1376.
(3) Groves, J. T. *J. Chem. Educ.* **1985**, *62*, 928–931.
(4) Dawson, J. H.; Sono, M. *Chem. Rev.* **1987**, *87*, 1255–1276.
(5) Ortiz de Montellano, P. R. *Trends Pharmacol. Sci.* **1989**, *10*, 354–359.
(6) Guengerich, F. P.; MacDonald, T. L. *FASEB J.* **1990**, *4*, 2453–2459.
(7) *Cytochrome P450: Structure, Mechanism, and Biochemistry*, 2nd ed.; Ortiz de Montellano, P. R., Ed.; Plenum Press: New York, 1995.
(8) Groves, J. T.; Han, Y.-Z. In *Cytochrome P450: Structure, Mechanism, and Biochemistry*, 2nd ed.; Ortiz de Montellano, P. R., Ed.; Plenum Press: New York, 1995; Chapter 1, pp 3−48.
(9) Guengerich, F. P. In *Cytochrome P450: Structure, Mechanism, and Biochemistry*, 2nd ed.; Ortiz de Montellano, P. R., Ed.; Plenum Press: New York, 1995; Chapter 14, pp 473−535.
(10) Ortiz de Montellano, P. R. In *Cytochrome P450: Structure, Mechanism, and Biochemistry*, 2nd ed.; Ortiz de Montellano, P. R., Ed.; Plenum Press: New York, 1995; Chapter 8, pp 245−303.
(11) Poulos, T. L.; Cupp-Vickery, J.; Li, H. In *Cytochrome P450: Structure, Mechanism, and Biochemistry*, 2nd ed.; Ortiz de Montellano, P. R., Ed.; Plenum Press: New York, 1995; Chapter 4, pp 125−150.
(12) Sono, M.; Roach, M. P.; Coulter, E. D.; Dawson, J. H. *Chem. Rev.* **1996**, *96*, 2841–2887.
(13) Woggon, W.-D. *Top. Curr. Chem.* **1996**, *184*, 39–96.
(14) Guengerich, F. P. *Annu. Rev. Pharmacol. Toxicol.* **1999**, *39*, 1–17.
(15) Sligar, S. G. *Essays Biochem.* **1999**, *34*, 71–83.
(16) Watanabe, Y. *J. Biol. Inorg. Chem.* **2001**, *6*, 846–856.
(17) Lewis, D. F. V. *Guide to Cytochromes P450*; Taylor and Francis: New York, 2001.
(18) Ortiz de Montellano, P. R.; de Voss, J. J. *Nat. Prod. Rep.* **2002**, *19*, 477–493.
(19) *Cytochrome P450: Structure, Mechanism, and Biochemistry*, 3rd ed.; Ortiz de Montellano, P. R., Ed.; Kluwer Academic/Plenum Publishers: New York, 2005.
(20) *The Ubiquitous Role of Cytochrome P450 Proteins In Metal Ions in Life Sciences*; Sigel, A., Sigel, H., Sigel, R. K. O., Eds.; John Wiley & Sons Ltd: Chichester, England, 2007; Vol. 3.
(21) Guengerich, F. P. In *Cytochrome P450: Structure, Mechanism, and Biochemistry*, 3rd ed.; Ortiz de Montellano, P. R., Ed.; Kluwer Academic/Plenum Publishers: New York, 2005; Chapter 10, pp 377−530.
(22) Groves, J. In *Cytochrome P450: Structure, Mechanism, and Biochemistry*, 3rd ed.; Ortiz de Montellano, P. R., Ed.; Kluwer Academic/Plenum Publishers: New York, 2005; Chapter 1, pp 1−43.
(23) Newcomb, M.; Toy, P. H. *Acc. Chem. Res.* **2000**, *33*, 449–455.
(24) Newcomb, M.; Shen, R.; Choi, S.-Y.; Toy, P. H.; Hollenberg, P. F.; Vaz, A. D. N.; Coon, M. J. *J. Am. Chem. Soc.* **2000**, *122*, 2677–2686.
(25) Shaik, S.; Hirao, H.; Kumar, D. *Acc. Chem. Res.* **2007**, *40*, 532–542.
(26) Shaik, S.; Hirao, H.; Kumar, D. *Nat. Prod. Rep.* **2007**, *24*, 533–552.
(27) Yamamoto, S.; Kashiwagi, H. *Chem. Phys. Lett.* **1989**, *161*, 85–89.
(28) Loew, G. H.; Harris, D. L. *Chem. Rev.* **2000**, *100*, 407–419.
(29) Harris, D. L. *Curr. Opin. Chem. Biol.* **2001**, *5*, 724–735.
(30) Shaik, S.; de Visser, S. P.; Ogliaro, F.; Schwarz, H.; Schröder, D. *Curr. Opin. Chem. Biol.* **2002**, *6*, 556–567.
(31) Shaik, S.; de Visser, S. P. In *Cytochrome P450: Structure, Mechanism, and Biochemistry*, 3rd ed.; Ortiz de Montellano, P. R., Ed.; Kluwer Academic/Plenum Publishers: New York, 2005; Chapter 2, pp 45−85.
(32) Shaik, S.; Kumar, D.; de Visser, S. P.; Altun, A.; Thiel, W. *Chem. Rev.* **2005**, *105*, 2279–2328.
(33) Bathelt, C. M.; Ridder, L.; Mulholland, A. J.; Harvey, J. N. *J. Am. Chem. Soc.* **2003**, *125*, 15004–15005.
(34) Bathelt, C. M.; Ridder, L.; Mulholland, A. J.; Harvey, J. N. *Org. Biomol. Chem.* **2004**, *2*, 2998–3005.
(35) de Visser, S. P.; Ogliaro, F.; Shaik, S. *Angew. Chem., Int. Ed.* **2001**, *40*, 2871–2874.
(36) de Visser, S. P.; Ogliaro, F.; Shaik, S. *Chem. Commun.* **2001**, 2322–2323.
(37) de Visser, S. P.; Ogliaro, F.; Harris, N.; Shaik, S. *J. Am. Chem. Soc.* **2001**, *123*, 3037–3047.
(38) de Visser, S. P.; Shaik, S. *J. Am. Chem. Soc.* **2003**, *125*, 7413–7424.
(39) Yoshizawa, K.; Ohta, T.; Eda, M.; Yamabe, T. *Bull. Chem. Soc. Jpn.* **2000**, *73*, 401–407.
(40) Yoshizawa, K.; Shiota, Y.; Kagawa, Y. *Bull. Chem. Soc. Jpn.* **2000**, *73*, 2669–2673.
(41) Yoshizawa, K.; Kagawa, Y.; Shiota, Y. *J. Phys. Chem. B* **2000**, *104*, 12365–12370.
(42) Yoshizawa, K.; Kamachi, T.; Shiota, Y. *J. Am. Chem. Soc.* **2001**, *123*, 9806–9816.
(43) Yoshizawa, K. *Coord. Chem. Rev.* **2002**, *226*, 251–259.
(44) Kamachi, T.; Shiota, Y.; Ohta, T.; Yoshizawa, K. *Bull. Chem. Soc. Jpn.* **2003**, *76*, 721–732.
(45) Kamachi, T.; Yoshizawa, K. *J. Am. Chem. Soc.* **2003**, *125*, 4652–4661.
(46) Green, M. T. *J. Am. Chem. Soc.* **1999**, *121*, 7939–7940.
(47) Green, M. T. *J. Am. Chem. Soc.* **2000**, *122*, 9495–9499.
(48) Green, M. T. *J. Am. Chem. Soc.* **2001**, *123*, 9218–9219.
(49) Ohta, T.; Matsuura, K.; Yoshizawa, K.; Morishima, I. *J. Inorg. Biochem.* **2000**, *82*, 141–152.
(50) Ogliaro, F.; Cohen, S.; de Visser, S. P.; Shaik, S. *J. Am. Chem. Soc.* **2000**, *122*, 12892–12893.
(51) de Visser, S. P.; Ogliaro, F.; Gross, Z.; Shaik, S. *Chem.—Eur. J.* **2001**, *7*, 4954–4960.
(52) Sharma, P. K.; de Visser, S. P.; Ogliaro, F.; Shaik, S. *J. Am. Chem. Soc.* **2003**, *125*, 2291–2300.
(53) Schöneboom, J. C.; Lin, H.; Reuter, N.; Thiel, W.; Cohen, S.; Ogliaro, F.; Shaik, S. *J. Am. Chem. Soc.* **2002**, *124*, 8142–8151.
(54) Zheng, J. J.; Altun, A.; Thiel, W. *J. Comput. Chem.* **2007**, *28*, 2147–2158.
(55) Fishelovitch, D.; Hazan, C.; Hirao, H.; Wolfson, H. J.; Nussinov, R.; Shaik, S. *J. Phys. Chem. B* **2007**, *111*, 13822–13832.
(56) Bathelt, C. M.; Zurek, J.; Mulholland, A. J.; Harvey, J. N. *J. Am. Chem. Soc.* **2005**, *127*, 12900–12908.
(57) Cohen, S.; Kozuch, S.; Hazan, C.; Shaik, S. *J. Am. Chem. Soc.* **2006**, *128*, 11028–11029.
(58) Wang, Y.; Hirao, H.; Chen, H.; Onaka, H.; Nagano, S.; Shaik, S. *J. Am. Chem. Soc.* **2008**, *130*, 7170–7171.
(59) Schöneboom, J. C.; Cohen, S.; Lin, H.; Shaik, S.; Thiel, W. *J. Am. Chem. Soc.* **2004**, *126*, 4017–4034.
(60) Bathelt, C. M.; Mulholland, A. J.; Harvey, J. N. *J. Phys. Chem. A* **2008**, *112*, 13149–13156.
(61) Altun, A.; Guallar, V.; Friesner, R. A.; Shaik, S.; Thiel, W. *J. Am. Chem. Soc.* **2006**, *128*, 3924–3925.
(62) Cohen, S.; Kumar, D.; Shaik, S. *J. Am. Chem. Soc.* **2006**, *128*, 2649–2653.
(63) Hirao, H.; Cho, K.-B.; Shaik, S. *J. Biol. Inorg. Chem.* **2008**, *13*, 521–530.
(64) (a) Jiang, Y.; Ortiz de Montellano, P. R. *Inorg. Chem.* **2008**, *47*, 3480–3482. (b) Jiang, Y.; Sivaramakrishnan, S.; Hayashi, T.; Cohen, S.; Moënne-Loccoz, P.; Shaik, S.; Ortiz de Montellano, P. R. *Angew. Chem., Int. Ed.* **2009**, *48*, 7193–7195.
(65) Aldag, I.; Gromov, I.; García-Rubio, I.; von Koenig, K.; Schlichting, I.; Jaun, B.; Hilvert, D. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 5481–5486.
(66) Chen, H.; Hirao, H.; Derat, E.; Schlichting, I.; Shaik, S. *J. Phys. Chem. B* **2008**, *112*, 9490–9500.
(67) Kühnel, K.; Derat, E.; Terner, J.; Shaik, S.; Schlichting, I. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 99–104.
(68) Cho, K.-B.; Derat, E.; Shaik, S. *J. Am. Chem. Soc.* **2007**, *129*, 3182–3188.
(69) Cho, K.-B.; Carvajal, M. A.; Shaik, S. *J. Phys. Chem. B* **2009**, *113*, 336–346.
(70) Derat, E.; Shaik, S. *J. Phys. Chem. B* **2006**, *110*, 10526–10533.
(71) Derat, E.; Shaik, S. *J. Am. Chem. Soc.* **2006**, *128*, 13940–13949.
(72) Derat, E.; Shaik, S. *J. Am. Chem. Soc.* **2006**, *128*, 8185–8198.
(73) Derat, E.; Shaik, S.; Rovira, C.; Vidossich, P.; Alfonso-Prieto, M. *J. Am. Chem. Soc.* **2007**, *129*, 6346–6347.

(74) Derat, E.; Cohen, S.; Shaik, S.; Altun, A.; Thiel, W. *J. Am. Chem. Soc.* **2005**, *127*, 13611–13621.

(75) Kim, S. H.; Perera, R.; Hager, L. P.; Dawson, J. H.; Hoffman, B. M. *J. Am. Chem. Soc.* **2006**, *128*, 5598–5599.

(76) Wallrapp, F.; Masone, D.; Guallar, V. *J. Phys. Chem. A* **2008**, *112*, 12989–12994.

(77) Schöneboom, J. C.; Neese, F.; Thiel, W. *J. Am. Chem. Soc.* **2005**, *127*, 5840–5853.

(78) Altun, A.; Kumar, D.; Neese, F.; Thiel, W. *J. Phys. Chem. A* **2008**, *112*, 12904–12910.

(79) Chen, H.; Ikeda-Saito, M.; Shaik, S. *J. Am. Chem. Soc.* **2008**, *130*, 14778–14790.

(80) Wang, Y.; Chen, H.; Makino, M.; Shiro, S.; Nagano, S.; Onaka, H.; Shaik, S. *J. Am. Chem. Soc.* **2009**, *131*, 6748–6762.

(81) Harris, D. L. *Curr. Opin. Drug Discovery Dev.* **2004**, *7*, 43–48.

(82) Meunier, B.; de Visser, S. P.; Shaik, S. *Chem. Rev.* **2004**, *104*, 3947–3980.

(83) Shaik, S.; Cohen, S.; de Visser, S. P.; Sharma, P. K.; Kumar, D.; Kozuch, S.; Ogliaro, F.; Danovich, D. *Eur. J. Inorg. Chem.* **2004**, 207–226.

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

(85) Gao, J. L. In *Reviews in Computational Chemistry*; Lipkowitz, K., Boyd, D., Eds.; VCH: Weinheim, 1996; Vol. 7, pp 119–185.

(86) Sherwood, P. In *Modern Methods and Algorithms of Quantum Chemistry*; NIC Series; Grotendorst, J., Ed.; John von Neumann Institute for Computing: Jülich, 2000; Vol. 1, pp 257–277.

(87) Gao, J. L.; Truhlar, D. G. *Annu. Rev. Phys. Chem.* **2002**, *53*, 467–505.

(88) Field, M. J. *J. Comput. Chem.* **2002**, *23*, 48–58.

(89) Warshel, A. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 425–443.

(90) Monard, G.; Prat-Resina, X.; González-Lafont, A.; Lluch, J. M. *Int. J. Quantum Chem.* **2003**, *93*, 229–244.

(91) Ridder, L.; Mulholland, A. J. *Curr. Top. Med. Chem.* **2003**, *3*, 1241–1256.

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

(93) Senn, H. M.; Thiel, W. *Top. Curr. Chem.* **2007**, *268*, 173–290.

(94) Senn, H. M.; Thiel, W. *Curr. Opin. Chem. Biol.* **2007**, *11*, 182–187.

(95) Lin, H.; Truhlar, D. G. *Theor. Chem. Acc.* **2007**, *117*, 185–199.

(96) Hu, H.; Yang, W. T. *Annu. Rev. Phys. Chem.* **2008**, *59*, 573–601.

(97) Senn, H. M.; Thiel, W. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198–1229.

(98) Heyden, A.; Lin, H.; Truhlar, D. G. *J. Phys. Chem. B* **2007**, *111*, 2231–2241.

(99) Riccardi, D.; Schaefer, P.; Cui, Q. *J. Phys. Chem. B* **2005**, *109*, 17715–17733.

(100) Benighaus, T.; Thiel, W. *J. Chem. Theory Comput.* **2008**, *4*, 1600–1609.

(101) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357–19363.

(102) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215–241.

(103) Siegbahn, P. E. M. *J. Biol. Inorg. Chem.* **2006**, *11*, 695–701.

(104) Claeyssens, F.; Harvey, J. N.; Manby, F. R.; Mata, R. A.; Mulholland, A. J.; Ranaghan, K. E.; Schütz, M.; Thiel, S.; Thiel, W.; Werner, H.-J. *Angew. Chem., Int. Ed.* **2006**, *45*, 6856–6859.

(105) Mata, R. A.; Werner, H.-J.; Thiel, S.; Thiel, W. *J. Chem. Phys.* **2008**, *128*, 025104/1−025104/8.

(106) Geerke, D. P.; Thiel, S.; Thiel, W.; van Gunsteren, W. F. *J. Chem. Theory Comput.* **2007**, *3*, 1499–1509.

(107) Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996**, *100*, 10580–10594.

(108) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718–730.

(109) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700–733.

(110) Antes, I.; Thiel, W. *J. Phys. Chem. A* **1999**, *103*, 9290–9295.

(111) Zhang, Y. K.; Lee, T.-S.; Yang, W. T. *J. Chem. Phys.* **1999**, *110*, 46–54.

(112) Laio, A.; VandeVondele, J.; Rothlisberger, U. *J. Chem. Phys.* **2002**, *116*, 6941–6947.

(113) Théry, V.; Rinaldi, D.; Rivail, J.-L.; Maigret, B.; Ferenczy, G. G. *J. Comput. Chem.* **1994**, *15*, 269–282.

(114) Gao, J. L.; Amara, P.; Alhambra, C.; Field, M. J. *J. Phys. Chem. A* **1998**, *102*, 4714–4721.

(115) Murphy, R. B.; Philipp, D. M.; Friesner, R. A. *J. Comput. Chem.* **2000**, *21*, 1442–1457.

(116) Reuter, N.; Dejaegere, A.; Maigret, B.; Karplus, M. *J. Phys. Chem. A* **2000**, *104*, 1720–1735.

(117) Lennartz, C.; Schäfer, A.; Terstegen, F.; Thiel, W. *J. Phys. Chem. B* **2002**, *106*, 1758–1767.

(118) Sherwood, P.; de Vries, A. H.; Guest, M. F.; Schreckenbach, G.; Catlow, C. R. A.; French, S. A.; Sokol, A. A.; Bromley, S. T.; Thiel, W.; Turner, A. J.; Billeter, S.; Terstegen, F.; Thiel, S.; Kendrick, J.;

Rogers, S. C.; Casci, J.; Watson, M.; King, F.; Karlsen, E.; Sjøvoll, M.; Fahmi, A.; Schäfer, A.; Lennartz, C. *THEOCHEM* **2003**, *632*, 1–28.

(119) Sherwood, P.; de Vries, A. H.; Collins, S. J.; Greatbanks, S. P.; Burton, N. A.; Vincent, M. A.; Hillier, I. H. *Faraday Discuss.* **1997**, *106*, 79–92.

(120) Eichler, U.; Kölmel, C. M.; Sauer, J. *J. Comput. Chem.* **1997**, *18*, 463–477.

(121) Billeter, S. R.; Turner, A. J.; Thiel, W. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2177–2186.

(122) Maseras, F.; Morokuma, K. *J. Comput. Chem.* **1995**, *16*, 1170–1179.

(123) Kästner, J.; Thiel, S.; Senn, H. M.; Sherwood, P.; Thiel, W. *J. Chem. Theory Comput.* **2007**, *3*, 1064–1072.

(124) Zhang, Y. K.; Liu, H. Y.; Yang, W. T. *J. Chem. Phys.* **2000**, *112*, 3483–3492.

(125) Kästner, J.; Senn, H. M.; Thiel, S.; Otte, N.; Thiel, W. *J. Chem. Theory Comput.* **2006**, *2*, 452–461.

(126) Senn, H. M.; Kästner, J.; Breidung, J.; Thiel, W. *Can. J. Chem.* **2009**, *87*, 1322−1337.

(127) Senn, H. M.; Thiel, S.; Thiel, W. *J. Chem. Theory Comput.* **2005**, *1*, 494–505.

(128) Kästner, J.; Thiel, W. *J. Chem. Phys.* **2005**, *123*, 144104/1–144104/5.

(129) http://www.chemshell.org.

(130) Li, H.; Robertson, A. D.; Jensen, J. H. *Proteins* **2005**, *61*, 704–721.

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

(132) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley-Interscience: New York, 1986.

(133) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 299–310.

(134) Altun, A.; Thiel, W. *J. Phys. Chem. B* **2005**, *109*, 1268–1280.

(135) Wachters, A. J. H. *J. Chem. Phys.* **1970**, *52*, 1033–1036.

(136) Hay, P. J. *J. Chem. Phys.* **1977**, *66*, 4377–4384.

(137) Bauschlicher, C. W.; Langhoff, S. R.; Partridge, H.; Barnes, L. A. *J. Chem. Phys.* **1989**, *91*, 2399–2411.

(138) Ahlrichs, R.; May, K. *Phys. Chem. Chem. Phys.* **2000**, *2*, 943–945.

(139) Schäfer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1992**, *97*, 2571–2577.

(140) Loew, G. H. *Int. J. Quantum Chem.* **2000**, *77*, 54–70.

(141) Rydberg, P.; Sigfridsson, E.; Ryde, U. *J. Biol. Inorg. Chem.* **2004**, *9*, 203–223.

(142) Ogliaro, F.; de Visser, S. P.; Cohen, S.; Sharma, P. K.; Shaik, S. *J. Am. Chem. Soc.* **2002**, *124*, 2806–2817.

(143) Ogliaro, F.; de Visser, S. P.; Shaik, S. *J. Inorg. Biochem.* **2002**, *91*, 554–567.

(144) Hata, M.; Hirano, Y.; Hoshino, T.; Nishida, R.; Tsuda, M. *J. Phys. Chem. B* **2004**, *108*, 11189–11195.

(145) Shaik, S.; Filatov, M.; Schröder, D.; Schwarz, H. *Chem.−Eur. J.* **1998**, *4*, 193–199.

(146) Ogliaro, F.; Harris, N.; Cohen, S.; Filatov, M.; de Visser, S. P.; Shaik, S. *J. Am. Chem. Soc.* **2000**, *122*, 8977–8989.

(147) Kumar, D.; de Visser, S. P.; Sharma, P. K.; Cohen, S.; Shaik, S. *J. Am. Chem. Soc.* **2004**, *126*, 1907–1920.

(148) Ogliaro, F.; Filatov, M.; Shaik, S. *Eur. J. Inorg. Chem.* **2000**, 2455–2458.

(149) de Visser, S. P.; Ogliaro, F.; Sharma, P. K.; Shaik, S. *Angew. Chem., Int. Ed.* **2002**, *41*, 1947–1951.

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

(151) Kumar, D.; de Visser, S. P.; Shaik, S. *J. Am. Chem. Soc.* **2003**, *125*, 13024–13025.

(152) Kumar, D.; de Visser, S. P.; Shaik, S. *J. Am. Chem. Soc.* **2004**, *126*, 5072–5073.

(153) de Visser, S. P.; Kumar, D.; Cohen, S.; Shacham, R.; Shaik, S. *J. Am. Chem. Soc.* **2004**, *126*, 8362–8363.

(154) Sharma, P. K.; de Visser, S. P.; Shaik, S. *J. Am. Chem. Soc.* **2003**, *125*, 8698–8699.

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

(156) Schöneboom, J. C.; Thiel, W. *J. Phys. Chem. B* **2004**, *108*, 7468–7478.

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

(158) Lin, H.; Schöneboom, J. C.; Cohen, S.; Shaik, S.; Thiel, W. *J. Phys. Chem. B* **2004**, *108*, 10083–10088.

(159) Zheng, J. J.; Wang, D.; Thiel, W.; Shaik, S. *J. Am. Chem. Soc.* **2006**, *128*, 13204–13215.

(160) Cho, K.-B.; Hirao, H.; Chen, H.; Carvajal, M. A.; Cohen, S.; Derat, E.; Thiel, W.; Shaik, S. *J. Phys. Chem. A* **2008**, *112*, 13128–13138.

(161) Wang, D.; Zheng, J. J.; Shaik, S.; Thiel, W. *J. Phys. Chem. B* **2008**, *112*, 5126–5138.

(162) Wang, D.; Thiel, W. *THEOCHEM* **2009**, *898*, 90–96.

(163) Zurek, J.; Foloppe, N.; Harvey, J. N.; Mulholland, A. J. *Org. Biomol. Chem.* **2006**, *4*, 3931–3937.

(164) Altun, A.; Shaik, S.; Thiel, W. *J. Am. Chem. Soc.* **2007**, *129*, 8978–8987.

(165) Hackett, J. C.; Sanan, T. T.; Hadad, C. M. *Biochemistry* **2007**, *46*, 5924–5940.

(166) Li, C. S.; Wu, W.; Kumar, D.; Shaik, S. *J. Am. Chem. Soc.* **2006**, *128*, 394–395.

(167) Wang, Y.; Kumar, D.; Yang, C. L.; Han, K. L.; Shaik, S. *J. Phys. Chem. B* **2007**, *111*, 7700–7710.

(168) Wang, Y.; Yang, C. L.; Wang, H. M.; Han, K. L.; Shaik, S. *ChemBioChem* **2007**, *8*, 277–281.

(169) Li, C. S.; Zhang, L.; Zhang, C.; Hirao, H.; Wu, W.; Shaik, S. *Angew. Chem., Int. Ed.* **2007**, *46*, 8168–8170.

(170) Olsen, L.; Rydberg, P.; Rod, T. H.; Ryde, U. *J. Med. Chem.* **2006**, *49*, 6489–6499.

(171) Rydberg, P.; Ryde, U.; Olsen, L. *J. Chem. Theory Comput.* **2008**, *4*, 1369–1377.

(172) Rydberg, P.; Ryde, U.; Olsen, L. *J. Phys. Chem. A* **2008**, *112*, 13058–13065.

(173) Zhang, Y.; Morisetti, P.; Kim, J.; Smith, L.; Lin, H. *Theor. Chem. Acc.* **2008**, *121*, 313–319.

(174) Shaik, S.; Kumar, D.; de Visser, S. P. *J. Am. Chem. Soc.* **2008**, *130*, 10128–10140.

(175) Denisov, I. G.; Makris, T. M.; Sligar, S. G.; Schlichting, I. *Chem. Rev.* **2005**, *105*, 2253–2277.

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

(177) Rowland, P.; Blaney, F. E.; Smyth, M. G.; Jones, J. J.; Leydon, V. R.; Oxbrow, A. K.; Lewis, C. J.; Tennant, M. G.; Modi, S.; Eggleston, D. S.; Chenery, R. J.; Bridges, A. M. *J. Biol. Chem.* **2006**, *281*, 7614–7622.

(178) 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.

(179) Guengerich, F. P.; Miller, G. P.; Hanna, I. H.; Sato, H.; Martin, M. V. *J. Biol. Chem.* **2002**, *277*, 33711–33719.

(180) Mueller, E. J.; Loida, P. J.; Sligar, S. G. In *Cytochrome P450: Structure, Mechanism, and Biochemistry*, 2nd ed.; Ortiz de Montellano, P. R., Ed.; Plenum Press: New York, 1995; Chapter 3, pp 83–124.

(181) von Köning, K.; Schlichting, I. In *Metal Ions in Life Sciences*; Sigel, A., Sigel, H., Sigel, R. K. O., Eds.; John Wiley & Sons Ltd: Chichester, England, 2007; Vol. 3; Chapter 8, pp 235–266.

(182) Vidakovic, M.; Sligar, S. G.; Li, H.; Poulos, T. L. *Biochemistry* **1998**, *37*, 9211–9219.

(183) Jin, S.; Makris, T. M.; Bryson, T. A.; Sligar, S. G.; Dawson, J. H. *J. Am. Chem. Soc.* **2003**, *125*, 3406–3407.

(184) Hishiki, T.; Shimada, H.; Nagano, S.; Egawa, T.; Kanamori, Y.; Makino, R.; Park, S.-Y.; Adachi, S.; Shiro, Y.; Ishimura, Y. *J. Biochem. (Tokyo, Jpn.)* **2000**, *128*, 965–974.

(185) Raag, R.; Martinis, S. A.; Sligar, S. G.; Poulos, T. L. *Biochemistry* **1991**, *30*, 11420–11429.

(186) Bell, S. G.; Sowden, R. J.; Wong, L.-L. *Chem. Commun.* **2001**, 635–636.

(187) Jung, C. In *Metal Ions in Life Sciences*; Sigel, A., Sigel, H., Sigel, R. K. O., Eds.; John Wiley & Sons Ltd: Chichester, England, 2007; Vol. 3; Chapter 7, pp 187–234.

(188) Rydberg, P.; Rod, T. H.; Olsen, L.; Ryde, U. *J. Phys. Chem. B* **2007**, *111*, 5445–5457.

(189) Fishelovitch, D.; Hazan, C.; Shaik, S.; Wolfson, H. J.; Nussinov, R. *J. Am. Chem. Soc.* **2007**, *129*, 1602–1611.

(190) Atkins, W. M.; Sligar, S. G. *J. Biol. Chem.* **1988**, *263*, 18842–18849.

(191) Bell, S. G.; Hoskins, N.; Whitehous, C. J. C.; Wang, L.-L. In *Metal Ions in Life Sciences*; Sigel, A., Sigel, H., Sigel, R. K. O., Eds.; John Wiley & Sons Ltd: Chichester, England, 2007; Vol. 3, pp 437–476.

(192) Poulos, T. L. *J. Biol. Inorg. Chem.* **1996**, *1*, 356–359.

(193) 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.

(194) Yoshioka, S.; Takahashi, S.; Ishimori, K.; Morishima, I. *J. Inorg. Biochem.* **2000**, *81*, 141–151.

(195) Makino, M.; Sugimoto, H.; Shiro, Y.; Asamizu, S.; Onaka, H.; Nagano, S. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 11591–11596.

(196) Poulos, T. L. *Nat. Prod. Rep.* **2007**, *24*, 504–510.

(197) Ogliaro, F.; de Visser, S. P.; Cohen, S.; Kaneti, J.; Shaik, S. *ChemBioChem* **2001**, *2*, 848–851.

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

(199) de Visser, S. P.; Shaik, S.; Sharma, P. K.; Kumar, D.; Thiel, W. *J. Am. Chem. Soc.* **2003**, *125*, 15779–15788.

(200) Hirao, H.; Kumar, D.; Thiel, W.; Shaik, S. *J. Am. Chem. Soc.* **2005**, *127*, 13007–13018.

(201) Thomann, H.; Bernardo, M.; Goldfarb, D.; Kroneck, P. M. H.; Ullrich, V. *J. Am. Chem. Soc.* **1995**, *117*, 8243–8251.

(202) Tsai, R.; Yu, C. A.; Gunsalus, I. C.; Peisach, J.; Blumberg, W.; Orme-Johnson, W. H.; Beinert, H. *Proc. Natl. Acad. Sci. U. S. A.* **1970**, *66*, 1157–1163.

(203) Sharrock, M.; Debrunner, P. G.; Schulz, C.; Lipscomb, J. D.; Marshall, V.; Gunsalus, I. C. *Biochim. Biophys. Acta* **1976**, *420*, 8–26.

(204) Sligar, S. G. *Biochemistry* **1976**, *15*, 5399–5406.

(205) Reiher, M.; Salomon, O.; Hess, B. A. *Theor. Chem. Acc.* **2001**, *107*, 48–55.

(206) Salomon, O.; Reiher, M.; Hess, B. A. *J. Chem. Phys.* **2002**, *117*, 4729–4737.

(207) Scherlis, D. A.; Cymeryng, C. B.; Estrin, D. A. *Inorg. Chem.* **2000**, *39*, 2352–2359.

(208) Scherlis, D. A.; Marti, M. A.; Ordejon, P.; Estrin, D. A. *Int. J. Quantum Chem.* **2002**, *90*, 1505–1514.

(209) Poulos, T. L.; Finzel, B. C.; Howard, A. J. *Biochemistry* **1986**, *25*, 5314–5322.

(210) Auclair, K.; Moënne-Loccoz, P.; Ortiz de Montellano, P. R. *J. Am. Chem. Soc.* **2001**, *123*, 4877–4885.

(211) Champion, P. M.; Lipscomb, J. D.; Münck, E.; Debrunner, P.; Gunsalus, I. C. *Biochemistry* **1975**, *14*, 4151–4158.

(212) Freindorf, M.; Shao, Y.; Kong, J.; Furlani, T. R. *J. Inorg. Biochem.* **2008**, *102*, 427–432.

(213) Unno, M.; Christian, J. F.; Benson, D. E.; Gerber, N. C.; Sligar, S. G.; Champion, P. M. *J. Am. Chem. Soc.* **1997**, *119*, 6614–6620.

(214) Lipscomb, J. D. *Biochemistry* **1980**, *19*, 3590–3599.

(215) Debrunner, P. G. In *Iron Porphyrins Part 3: Physical Bioinorganic Chemistry Series*; Lever, A. B. P., Gray, H. B., Eds.; VCH: New York, 1989; Vol. 4; Chapter 2, pp 137–234.

(216) Reiss, H.; Heller, A. *J. Phys. Chem.* **1985**, *89*, 4207–4213.

(217) Olsson, M. H. M.; Hong, G. Y.; Warshel, A. *J. Am. Chem. Soc.* **2003**, *125*, 5025–5039.

(218) Li, G. H.; Zhang, X. D.; Cui, Q. *J. Phys. Chem. B* **2003**, *107*, 8643–8653.

(219) Makris, T. M.; Denisov, I. G.; Schlichting, I.; Sligar, S. In *Cytochrome P450: Structure, Mechanism and Biochemistry*, 3rd ed.; Ortiz de Montellano, P. R., Ed.; Kluwer Academic/Plenum Publishers: New York, 2005; Chapter 5, pp 149–182.

(220) Harada, K.; Sakurai, K.; Ikemura, K.; Ogura, T.; Hirota, S.; Shimada, H.; Hayashi, T. *J. Am. Chem. Soc.* **2008**, *130*, 432–433.

(221) Hayashi, T.; Harada, K.; Sakurai, K.; Shimada, H.; Hirota, S. *J. Am. Chem. Soc.* **2009**, *131*, 1398–1400.

(222) Chottard, G.; Schappacher, M.; Ricard, L.; Weiss, R. *Inorg. Chem.* **1984**, *23*, 4557–4561.

(223) Denisov, I. G.; Mak, P. J.; Makris, T. M.; Sligar, S. G.; Kincaid, J. R. *J. Phys. Chem. A* **2008**, *112*, 13172–13179.

(224) Unno, M.; Chen, H.; Kusama, S.; Shaik, S.; Ikeda-Saito, M. *J. Am. Chem. Soc.* **2007**, *129*, 13394–13395.

(225) Harris, D. L.; Loew, G. H.; Waskell, L. *J. Am. Chem. Soc.* **1998**, *120*, 4308–4318.

(226) Kumar, D.; Hirao, H.; de Visser, S. P.; Zheng, J. J.; Wang, D.; Thiel, W.; Shaik, S. *J. Phys. Chem. B* **2005**, *109*, 19946–19951.

(227) Radoń, M.; Pierloot, K. *J. Phys. Chem. A* **2008**, *112*, 11824–11832.

(228) Benson, D. E.; Suslick, K. S.; Sligar, S. G. *Biochemistry* **1997**, *36*, 5104–5107.

(229) Pauling, L.; Coryell, C. D. *Proc. Natl. Acad. Sci. U.S.A* **1936**, *22*, 210–216.

(230) Weiss, J. J. *Nature* **1964**, *202*, 83–84.

(231) McClure, D. S. *Radiat. Res. Suppl.* **1960**, *2*, 218–242.

(232) Harcourt, R. D. *Chem. Phys. Lett.* **1990**, *167*, 374–377.

(233) Mak, P. J.; Denisov, I. G.; Victoria, D.; Makris, T. M.; Deng, T. J.; Sligar, S. G.; Kincaid, J. R. *J. Am. Chem. Soc.* **2007**, *129*, 6382–6383.

(234) Ichikawa, Y.; Nakajima, H.; Watanabe, Y. *ChemBioChem* **2006**, *7*, 1582–1589.

(235) Harris, D. L.; Loew, G. H. *J. Am. Chem. Soc.* **1998**, *120*, 8941–8948.

(236) de Grotthuss, C. J. T. *Anal. Chim.* **1806**, *58*, 54–73.

(237) Taraphder, S.; Hummer, G. *J. Am. Chem. Soc.* **2003**, *125*, 3931–3940.

(238) Oprea, T. I.; Hummer, G.; García, A. E. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 2133–2138.

(239) Aikens, J.; Sligar, S. G. *J. Am. Chem. Soc.* **1994**, *116*, 1143–1144.

(240) Purdy, M. M.; Koo, L. S.; Ortiz de Montellano, P. R.; Klinman, J. P. *Biochemistry* **2006**, *45*, 15793–15806.

(241) Kim, S.-H.; Yang, T.-C.; Perera, R.; Jin, S.-X.; Bryson, T. A.; Sono, M.; Davydov, R.; Dawson, J. H.; Hoffman, B. M. *Dalton Trans.* **2005**, 3464–3469.

(242) Shimada, H.; Makino, R.; Unno, M.; Horiuchi, T.; Ishimura, Y. In *8th International Conference on Cytochrome P450: Biochemistry, Biophysics and Molecular Biology*, Lisbon, Portugal, 1994; pp 299–306.

(243) Guallar, V.; Harris, D. L.; Batista, V. S.; Miller, W. H. *J. Am. Chem. Soc.* **2002**, *124*, 1430–1437.

(244) Harris, D. L. *J. Inorg. Biochem.* **2002**, *91*, 568–585.

(245) Poulos, T. L. In *The Porphyrin Handbook*; Kadish, K. M., Smith, K. M., Guilard, R., Eds.; Academic Press: San Diego, 2000; Vol. 4, pp 189–218.

(246) Altarsha, M.; Benighaus, T.; Kumar, D.; Thiel, W. *J. Am. Chem. Soc.* **2009**, *131*, 4755–4763.

(247) Altarsha, M.; Wang, D.; Benighaus, T.; Kumar, D.; Thiel, W. *J. Phys. Chem. B* **2009**, *113*, 9577–9588.

(248) Egawa, T.; Shimada, H.; Ishimura, Y. *Biochem. Biophys. Res. Commun.* **1994**, *201*, 1464–1469.

(249) Denisov, I. G.; Makris, T. M.; Sligar, S. G. *J. Biol. Chem.* **2001**, *276*, 11648–11652.

(250) Kellner, D. G.; Hung, S. C.; Weiss, K. E.; Sligar, S. G. *J. Biol. Chem.* **2002**, *277*, 9641–9644.

(251) Jung, C. *J. Mol. Recognit.* **2000**, *13*, 325–351.

(252) Jung, C. *Biochim. Biophys. Acta, Proteins Proteomics* **2002**, *1595*, 309–328.

(253) Schünemann, V.; Jung, C.; Terner, J.; Trautwein, A. X.; Weiss, R. *J. Inorg. Biochem.* **2002**, *91*, 586–596.

(254) Meilleur, F.; Contzen, J.; Myles, D. A. A.; Jung, C. *Biochemistry* **2004**, *43*, 8744–8753.

(255) Spolitak, T.; Dawson, J. H.; Ballou, D. P. *J. Biol. Chem.* **2005**, *280*, 20300–20309.

(256) Dowers, T. S.; Rock, D. A.; Jones, J. P. *J. Am. Chem. Soc.* **2004**, *126*, 8868–8869.

(257) Vaz, A. D. N.; McGinnity, D. F.; Coon, M. J. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 3555–3560.

(258) Nam, W.; Lee, H. J.; Oh, S.-Y.; Kim, C.; Jang, H. G. *J. Inorg. Biochem.* **2000**, *80*, 219–225.

(259) Nam, W.; Park, S.-E.; Lim, I. K.; Lim, M. H.; Hong, J. K.; Kim, J. *J. Am. Chem. Soc.* **2003**, *125*, 14674–14675.

(260) Kaizer, J.; Klinker, E. J.; Oh, N. Y.; Rohde, J.-U.; Song, W. J.; Stubna, A.; Kim, J.; Münck, E.; Nam, W.; Que, L., Jr. *J. Am. Chem. Soc.* **2004**, *126*, 472–473.

(261) Newcomb, M.; Zhang, R.; Chandrasena, R. E.; Halgrimson, J. A.; Horner, J. H.; Makris, T. M.; Sligar, S. G. *J. Am. Chem. Soc.* **2006**, *128*, 4580–4581.

(262) Sheng, X.; Horner, J. H.; Newcomb, M. *J. Am. Chem. Soc.* **2008**, *130*, 13310–13320.

(263) Watanabe, Y.; Nakajima, H.; Ueno, T. *Acc. Chem. Res.* **2007**, *40*, 554–562.

(264) Han, A.-R.; Jeong, Y. J.; Kang, Y.; Lee, J. Y.; Seo, M. S.; Nam, W. *Chem. Commun.* **2008**, 1076–1078.

(265) Franke, A.; Fertinger, C.; van Eldik, R. *Angew. Chem., Int. Ed.* **2008**, *47*, 5238–5242.

(266) Ogliaro, F.; de Visser, S. P.; Groves, J. T.; Shaik, S. *Angew. Chem., Int. Ed.* **2001**, *40*, 2874–2878.

(267) Pan, Z.; Zhang, R.; Newcomb, M. *J. Inorg. Biochem.* **2006**, *100*, 524–532.

(268) Harris, D. L.; Loew, G. H.; Waskell, L. *J. Inorg. Biochem.* **2001**, *83*, 309–318.

(269) Filatov, M.; Harris, N.; Shaik, S. *Angew. Chem., Int. Ed.* **1999**, *38*, 3510–3512.

(270) Filatov, M.; Harris, N.; Shaik, S. *J. Chem. Soc., Perkin Trans. 2* **1999**, 399–410.

(271) Antony, J.; Grodzicki, M.; Trautwein, A. X. *J. Phys. Chem. A* **1997**, *101*, 2692–2701.

(272) Harvey, J. N.; Bathelt, C. M.; Mulholland, A. J. *J. Comput. Chem.* **2006**, *27*, 1352–1362.

(273) Alfonso-Prieto, M.; Borovik, A.; Carpena, X.; Murshudov, G.; Melik-Adamyan, W.; Fita, I.; Rovira, C.; Loewen, P. C. *J. Am. Chem. Soc.* **2007**, *129*, 4193–4205.

(274) Radoń, M.; Broclawik, E. *J. Chem. Theory Comput.* **2007**, *3*, 728–734.

(275) Shaik, S.; de Visser, S. P.; Kumar, D. *J. Am. Chem. Soc.* **2004**, *126*, 11746–11749.

(276) Stone, K. L.; Behan, R. K.; Green, M. T. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 16563–16565.

(277) Cohen, S. Ph.D. Thesis, Hebrew University of Jerusalem, Israel, in preparation.

(278) Davydov, R.; Perera, R.; Jin, S.-X.; Yang, T.-C.; Bryson, T. A.; Sono, M.; Dawson, J. H.; Hoffman, B. M. *J. Am. Chem. Soc.* **2005**, *127*, 1403–1413.

(279) Shaik, S.; Hiberty, P. C. *A Chemist's Guide to Valence Bond Theory*; John Wiley & Sons Inc.: 2007; Chapter 2, pp 26–33.

(280) Hirao, H.; Shaik, S.; Kozlowski, P. M. *J. Phys. Chem. A* **2006**, *110*, 6091–6099.

(281) Green, M.; Dawson, J.; Gray, H. *Science* **2004**, *304*, 1653–1656.

(282) Kimata, Y.; Shimada, H.; Hirose, T.; Ishimura, Y. *Biochem. Biophys. Res. Commun.* **1995**, *208*, 96–102.

(283) Fishelovitch, D. Ph.D. Thesis, Tel-Aviv University, Israel, 2009.

(284) Schröder, D.; Shaik, S.; Schwarz, H. *Acc. Chem. Res.* **2000**, *33*, 139–145.

(285) Cryle, M. J.; Stuthe, J. M.; Ortiz de Montellano, P. R.; de Voss, J. J. *Chem. Commun.* **2004**, 512–513.

(286) He, X.; Ortiz de Montellano, P. R. *J. Biol. Chem.* **2004**, *279*, 39479–39484.

(287) Ortiz de Montellano, P. R.; de Voss, J. J. In *Cytochrome P450: Structure, Mechanism and Biochemistry*, 3rd ed.; Ortiz de Montellano, P. R., Ed.; Kluwer Academic/Plemun Publishers: New York, 2005; Chapter 6, pp 183–245.

(288) Bach, R. D.; Dmitrenko, O. *J. Am. Chem. Soc.* **2006**, *128*, 1474–1488.

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

(290) Zhang, R.; Chandrasena, R. E. P.; Martinez, E.; Horner, J. H.; Newcomb, M. *Org. Lett.* **2005**, *7*, 1193–1195.

(291) Page, M. I.; Jencks, W. P. *Proc. Natl. Acad. Sci. U. S. A.* **1971**, *68*, 1678–1683.

(292) Cho, K.-B.; Moreau, Y.; Kumar, D.; Rock, D. A.; Jones, J. P.; Shaik, S. *Chem.—Eur. J.* **2007**, *13*, 4103–4115.

(293) Li, C. S.; Zhang, L.; Zhang, C.; Hirao, H.; Wu, W.; Shaik, S. *Angew. Chem., Int. Ed.* **2008**, *47*, 8148.

(294) Derat, E.; Kumar, D.; Hirao, H.; Shaik, S. *J. Am. Chem. Soc.* **2006**, *128*, 473–484.

(295) Harris, N.; Cohen, S.; Filatov, M.; Ogliaro, F.; Shaik, S. *Angew. Chem., Int. Ed.* **2000**, *39*, 2003–2007.

(296) Schöneboom, J. C. Ph.D. Thesis, University of Düsseldorf, Germany, 2003.

(297) Li, H. Y.; Narasimhulu, S.; Havran, L. M.; Winkler, J. D.; Poulos, T. L. *J. Am. Chem. Soc.* **1995**, *117*, 6297–6299.

(298) Truhlar, D. G.; Gao, J. L.; Alhambra, C.; Garcia-Viloca, M.; Corchado, J.; Sánchez, M. L.; Villà, J. *Acc. Chem. Res.* **2002**, *35*, 341–349.

(299) Gelb, M. H.; Heimbrook, D. C.; Malkonen, P.; Sligar, S. G. *Biochemistry* **1982**, *21*, 370–377.

(300) Groves, J. T.; McClusky, G. A.; White, R. E.; Coon, M. J. *Biochem. Biophys. Res. Commun.* **1978**, *81*, 154–160.

(301) Guengerich, F. P.; Macdonald, T. L. *Acc. Chem. Res.* **1984**, *17*, 9–16.

(302) Groves, J. T.; Gross, Z.; Stern, M. K. *Inorg. Chem.* **1994**, *33*, 5065–5072.

(303) Groves, J. T.; Nemo, T. E. *J. Am. Chem. Soc.* **1983**, *105*, 5786–5791.

(304) Gross, Z.; Nimri, S.; Barzilay, C. M.; Simkhovich, L. *J. Biol. Inorg. Chem.* **1997**, *2*, 492–506.

(305) de Visser, S. P.; Kumar, D.; Shaik, S. *J. Inorg. Biochem.* **2004**, *98*, 1183–1193.

(306) Kumar, D.; de Visser, S. P.; Shaik, S. *Chem.—Eur. J.* **2005**, *11*, 2825–2835.

(307) Hirao, H.; Kumar, D.; Shaik, S. *J. Inorg. Biochem.* **2006**, *100*, 2054–2068.

(308) White, R. E.; Groves, J. T.; McClusky, G. A. *Acta Biol. Med. Ger.* **1979**, *38*, 475–482.

(309) Boyland, E. *Biochem. Soc. Symp.* **1950**, *5*, 40–54.

(310) Boyland, E.; Sims, P. *Biochem. J.* **1965**, *95*, 780–787.

(311) Jerina, D.; Daly, J.; Witkop, B.; Zaltzman-Nirenberg, P.; Udenfriend, S. *Arch. Biochem. Biophys.* **1968**, *128*, 176–183.

(312) Jerina, D.; Daly, J. *Science* **1974**, *185*, 573–582.

(313) Jerina, D.; Daly, J.; Witkop, B.; Zaltzman-Nirenberg, P.; Udenfriend, S. *J. Am. Chem. Soc.* **1968**, *90*, 6525–6527.

(314) Jerina, D.; Daly, J.; Witkop, B.; Zaltzman-Nirenberg, P.; Udenfriend, S. *Biochemistry* **1970**, *9*, 147–156.

(315) Korzekwa, K. R.; Swinney, D. C.; Trager, W. F. *Biochemistry* **1989**, *28*, 9019–9027.

(316) Rietjens, I.; Soffers, A.; Veeger, C.; Vervoort, J. *Biochemistry* **1993**, *32*, 4801–4812.

(317) Hanzlik, R.; Hogberg, K.; Judson, C. *Biochemistry* **1984**, *23*, 3048–3055.

(318) Zakharieva, O.; Grodzicki, M.; Trautwein, A. X.; Veeger, C.; Rietjens, I. *J. Biol. Inorg. Chem.* **1996**, *1*, 192–204.

(319) Zakharieva, O.; Grodzicki, M.; Trautwein, A.; Veeger, C.; Rietjens, I. *Biophys. Chem.* **1998**, *73*, 189–203.

(320) Burka, L. T.; Plucinski, T. M.; Macdonald, T. L. *Proc. Natl. Acad. Sci. U. S. A.* **1983**, *80*, 6680–6684.

(321) Nedelcheva, V.; Gut, I.; Soucek, P.; Tichavska, B.; Tynkova, L.; Mraz, J.; Guengerich, F. P.; Ingelman-Sundberg, M. *Arch. Toxicol.* **1999**, *73*, 33–40.

(322) Kimbrough, R. D. *Crit. Rev. Toxicol.* **1995**, *25*, 133–163.

(323) Evangelista de Duffard, A. M.; Duffard, R. *Environ. Health Perspect.* **1996**, *104*, 353–360.

(324) Janssen, D. B.; Oppentocht, J. E.; Poelarends, G. J. *Curr. Opin. Biotechnol.* **2001**, *12*, 254–258.

(325) Lee, M. D.; Odom, J. M.; Buchanan, R. J., Jr. *Annu. Rev. Microbiol.* **1998**, *52*, 423–452.

(326) Fetzner, S. *Appl. Microbiol. Biotechnol.* **1998**, *50*, 633–657.

(327) Commandeur, L. C. M.; Parsons, J. R. *Biodegradation* **1990**, *1*, 207–220.

(328) Motosugi, K.; Soda, K. *Cell. Mol. Life Sci.* **1983**, *39*, 1214–1220.

(329) Bell, S. G.; Chen, X.; Xu, F.; Rao, Z.; Wong, L.-L. *Biochem. Soc. Trans.* **2003**, *31*, 558–562.

(330) England, P. A.; Harford-Cross, C. F.; Stevenson, J.-A.; Rouch, D. A.; Wong, L.-L. *FEBS Lett.* **1998**, *424*, 271–274.

(331) Nickerson, D. P.; Harford-Cross, C. F.; Fulcher, S. R.; Wong, L.-L. *FEBS Lett.* **1997**, *405*, 153–156.

(332) Stevenson, J.-A.; Westlake, A. C. G.; Whittock, C.; Wong, L.-L. *J. Am. Chem. Soc.* **1996**, *118*, 12846–12847.

(333) Jones, J. P.; O'Hare, E. J.; Wong, L.-L. *Eur. J. Biochem.* **2001**, *268*, 1460–1467.

(334) Harford-Cross, C. F.; Carmichael, A. B.; Allan, F. K.; England, P. A.; Rouch, D. A.; Wong, L.-L. *Protein Eng.* **2000**, *13*, 121–128.

(335) Chen, X.; Christopher, A.; Jones, J. P.; Bell, S. G.; Guo, Q.; Xu, F.; Rao, Z.; Wong, L.-L. *J. Biol. Chem.* **2002**, *277*, 37519–37526.

(336) Rietjens, I.; den Besten, C.; Hanzlik, R. P.; van Bladeren, P. J. *Chem. Res. Toxicol.* **1997**, *10*, 629–635.

(337) Bernardi, F.; Cherry, W.; Shaik, S.; Epiotis, N. D. *J. Am. Chem. Soc.* **1978**, *100*, 1352–1356.

(338) Dinnocenzo, J. P.; Karki, S. B.; Jones, J. P. *J. Am. Chem. Soc.* **1993**, *115*, 7111–7116.

(339) Karki, S. B.; Dinnocenzo, J. P.; Jones, J. P.; Korzekwa, K. R. *J. Am. Chem. Soc.* **1995**, *117*, 3657–3664.

(340) Karki, S. B.; Dinnocenzo, J. P. *Xenobiotica* **1995**, *25*, 711–724.

(341) Manchester, J. I.; Dinnocenzo, J. P.; Higgins, L.; Jones, J. P. *J. Am. Chem. Soc.* **1997**, *119*, 5069–5070.

(342) Jurva, U.; Bissel, P.; Isin, E. M.; Igarashi, K.; Kuttab, S.; Castagnoli, N. *J. Am. Chem. Soc.* **2005**, *127*, 12368–12377.

(343) Abdel-Monem, M. M. *J. Med. Chem.* **1975**, *18*, 427–430.

(344) Miwa, G.; Garland, W.; Hodshon, B.; Lu, A.; Northrop, D. *J. Biol. Chem.* **1980**, *255*, 6049–6054.

(345) Shono, T.; Toda, T.; Oshino, N. *J. Am. Chem. Soc.* **1982**, *104*, 2639–2641.

(346) Miwa, G.; Walsh, J.; Kedderis, G.; Hollenberg, P. *J. Biol. Chem.* **1983**, *258*, 14445–14449.

(347) Galliani, G.; Rindone, B.; Dagnino, G.; Salmona, M. *Eur. J. Drug Metab. Pharmacokinet.* **1984**, *9*, 289–293.

(348) Burka, L. T.; Guengerich, F. P.; Willard, R. J.; Macdonald, T. L. *J. Am. Chem. Soc.* **1985**, *107*, 2549–2551.

(349) Galliani, G.; Nali, M.; Rindone, B.; Tollari, S.; Rocchetti, M.; Salmona, M. *Xenobiotica* **1986**, *16*, 511–517.

(350) Macdonald, T. L.; Gutheim, W. G.; Martin, R. B.; Guengerich, F. P. *Biochemistry* **1989**, *28*, 2071–2077.

(351) Parker, V. D.; Tilset, M. *J. Am. Chem. Soc.* **1991**, *113*, 8778–8781.

(352) Hollenberg, P. F.; Miwa, G. T.; Walsh, J. S.; Dwyer, L. A.; Rickert, D. E.; Kedderis, G. L. *Drug Metab. Dispos.* **1985**, *13*, 272–275.

(353) Guengerich, F. P.; Yun, C.-H.; Macdonald, T. L. *J. Biol. Chem.* **1996**, *271*, 27321–27329.

(354) Goto, Y.; Watanabe, Y.; Fukuzumi, S.; Jones, J. P.; Dinnocenzo, J. P. *J. Am. Chem. Soc.* **1998**, *120*, 10762–10763.

(355) Shaffer, C. L.; Morton, M. D.; Hanzlik, R. P. *J. Am. Chem. Soc.* **2001**, *123*, 349–350.

(356) Shaffer, C. L.; Morton, M. D.; Hanzlik, R. P. *J. Am. Chem. Soc.* **2001**, *123*, 8502–8508.

(357) Shaffer, C. L.; Harriman, S.; Koen, Y. M.; Hanzlik, R. P. *J. Am. Chem. Soc.* **2002**, *124*, 8268–8274.

(358) Cerny, M. A.; Hanzlik, R. P. *Arch. Biochem. Biophys.* **2005**, *436*, 265–275.

(359) Cerny, M. A.; Hanzlik, R. P. *J. Am. Chem. Soc.* **2006**, *128*, 3346–3354.

(360) Bhakta, M. N.; Wimalasena, K. *J. Am. Chem. Soc.* **2002**, *124*, 1844–1845.

(361) Bhakta, M. N.; Wimalasena, K. *Eur. J. Org. Chem.* **2005**, 4801–4805.

(362) Bhakta, M.; Hollenberg, P. F.; Wimalasena, K. *Chem. Commun.* **2005**, 265–267.

(363) Bhakta, M.; Hollenberg, P. F.; Wimalasena, K. *J. Am. Chem. Soc.* **2005**, *127*, 1376–1377.

(364) Baciocchi, E.; Lanzalunga, O.; Lapi, A.; Manduchi, L. *J. Am. Chem. Soc.* **1998**, *120*, 5783–5787.

(365) Baciocchi, E.; Gerini, M. F.; Lanzalunga, O.; Lapi, A.; Mancinelli, S.; Mencarelli, P. *Chem. Commun.* **2000**, 393–394.

(366) Baciocchi, E.; Gerini, F. M.; Lanzalunga, O.; Lapi, A.; Grazia Lo Piparo, M.; Mancinelli, S. *Eur. J. Org. Chem.* **2001**, 2305–2310.

(367) Baciocchi, E.; Bietti, M.; Gerini, M. F.; Lanzalunga, O. *J. Org. Chem.* **2005**, *70*, 5144–5149.

(368) Harris, N.; Shaik, S.; Schröder, D.; Schwarz, H. *Helv. Chim. Acta* **1999**, *82*, 1784–1797.

(369) Li, C. S.; Wu, W.; Cho, K.-B.; Shaik, S. *Chem.—Eur. J.* **2009**, *15*, 8492–8503.

(370) Lieber, C. S.; Decarli, L. M. *Science* **1968**, *162*, 917–918.

(371) Lieber, C. S. *Physiol. Rev.* **1997**, *77*, 517–544.

(372) Lieber, C. S. *Drug. Metab. Rev.* **2004**, *36*, 511–529.

(373) Cheng, K.; Schenkman, J. *J. Biol. Chem.* **1983**, *258*, 11738–11744.

(374) Ekstrom, G.; Norsten, C.; Cronholm, T.; Ingelmansundberg, M. *Biochemistry* **1987**, *26*, 7348–7354.

(375) Wood, A. W.; Swinney, D. C.; Thomas, P. E.; Ryan, D. E.; Hall, P. F.; Levin, W.; Garland, W. A. *J. Biol. Chem.* **1988**, *263*, 17322–17332.

(376) Vaz, A. D. N.; Coon, M. J. *Biochemistry* **1994**, *33*, 6442–6449.

(377) Bellucci, G.; Chiappe, C.; Pucci, L.; Gervasi, P. G. *Chem. Res. Toxicol.* **1996**, *9*, 871–874.

(378) Bell, L. C.; Guengerich, F. P. *J. Biol. Chem.* **1997**, *272*, 29643–29651.

(379) Bell-Parikh, L. C.; Guengerich, F. P. *J. Biol. Chem.* **1999**, *274*, 23833–23840.

(380) Guengerich, F. P. *Chem. Res. Toxicol.* **2001**, *14*, 611–650.

(381) Hackett, J. C.; Brueggemeier, R. W.; Hadad, C. M. *J. Am. Chem. Soc.* **2005**, *127*, 5224–5237.

(382) Han, J. H.; Yoo, S. K.; Seo, J. S.; Hong, S. J.; Kim, S. K.; Kim, C. *Dalton Trans.* **2005**, 402–406.

(383) Oh, N. Y.; Suh, Y.; Park, M. J.; Seo, M. S.; Kim, J.; Nam, W. *Angew. Chem., Int. Ed.* **2005**, *44*, 4235–4239.

(384) Smith, S. V.; Koley, A. P.; Dai, R.; Robinson, R. C.; Leong, H.; Markowitz, A.; Friedman, F. K. *Biochemistry* **2000**, *39*, 5731–5737.

(385) Howard-Jones, A. R.; Walsh, C. T. *J. Am. Chem. Soc.* **2007**, *129*, 11016–11017.

(386) Seto, Y.; Guengerich, F. P. *J. Biol. Chem.* **1993**, *268*, 9986–9997.

(387) Groves, J. T.; Avarianeisser, G. E.; Fish, K. M.; Imachi, M.; Kuczkowski, R. L. *J. Am. Chem. Soc.* **1986**, *108*, 3837–3838.

(388) Ruettinger, R. T.; Fulco, A. J. *J. Biol. Chem.* **1981**, *256*, 5728–5734.

(389) Song, W.-J.; Ryu, Y.-O.; Song, R.; Nam, W. *J. Biol. Inorg. Chem.* **2005**, *10*, 294–304.

(390) Agarwala, A.; Bandyopadhyay, D. *Chem. Commun.* **2006**, 4823–4825.

(391) Yun, C.-H.; Miller, G. P.; Guengerich, F. P. *Biochemistry* **2001**, *40*, 4521–4530.

(392) Hanzlik, R. P.; Hogberg, K.; Moon, J. B.; Judson, C. M. *J. Am. Chem. Soc.* **1985**, *107*, 7164–7167.

(393) Atkins, W. M.; Sligar, S. G. *J. Am. Chem. Soc.* **1987**, *109*, 3754–3760.

(394) Hanzlik, R. P.; Ling, K.-H. J. *J. Org. Chem.* **1990**, *55*, 3992–3997.

(395) Hazan, C.; Kumar, D.; de Visser, S. P.; Shaik, S. *Eur. J. Inorg. Chem.* **2007**, 2966–2974.

(396) Wand, M. D.; Thompson, J. A. *J. Biol. Chem.* **1986**, *261*, 14049–14056.

(397) Rettie, A. E.; Boberg, M.; Rettenmeier, A. W.; Baillie, T. A. *J. Biol. Chem.* **1988**, *263*, 13733–13738.

(398) Lewis, D. F. V.; Sams, C.; Loizou, G. D. *J. Biochem. Mol. Toxicol.* **2003**, *17*, 47–52.

(399) Bugrim, A.; Nikolskaya, T.; Nikolsky, Y. *Drug Discovery Today* **2004**, *9*, 127–135.

(400) Afzelius, L.; Arnby, C. H.; Broo, A.; Carlsson, L.; Isaksson, C.; Jurva, U.; Kjellander, B.; Kolmodin, K.; Nilsson, K.; Raubacher, F.; Weidolf, L. *Drug. Metab. Rev.* **2007**, *39*, 61–86.

(401) Mayer, J. M. *Acc. Chem. Res.* **1998**, *31*, 441–450.

(402) Shaik, S.; Shurki, A. *Angew. Chem., Int. Ed.* **1999**, *38*, 586–625.

(403) Mayer, J. M. In *Biomimetic Oxidations Catalyzed by Transition Metal Complexes*; Meunier, B., Ed.; Imperial College Press: London, 2000; Chapter 1, pp 1–45.

(404) Korzekwa, K. R.; Jones, J. P.; Gillette, J. R. *J. Am. Chem. Soc.* **1990**, *112*, 7042–7046.

(405) Park, J.-Y.; Harris, D. *J. Med. Chem.* **2003**, *46*, 1645–1660.

(406) Brewer, C. B.; Peterson, J. A. *J. Biol. Chem.* **1988**, *263*, 791–798.

(407) Hoffmann, R.; Woodward, R. B. *J. Am. Chem. Soc.* **1965**, *87*, 2046–2048.

(408) Hoffmann, R.; Woodward, R. B. *J. Am. Chem. Soc.* **1965**, *87*, 4388–4389.

(409) Hoffmann, R.; Woodward, R. B. *J. Am. Chem. Soc.* **1965**, *87*, 4389–4390.

(410) Woodward, R. B.; Hoffmann, R. *J. Am. Chem. Soc.* **1965**, *87*, 395–397.

(411) Woodward, R. B.; Hoffmann, R. *J. Am. Chem. Soc.* **1965**, *87*, 2511–2513.

(412) Fukui, K.; Yonezawa, T.; Shingu, H. *J. Chem. Phys.* **1952**, *20*, 722–725.

(413) Sydora, O. L.; Goldsmith, J. I.; Vaid, T. P.; Miller, A. E.; Wolczanski, P. T.; Abruña, H. D. *Polyhedron* **2004**, *23*, 2841–2856.

(414) Shaik, S. *J. Am. Chem. Soc.* **1981**, *103*, 3692–3701.

(415) Shaik, S.; Hiberty, P. C. *A Chemist's Guide to Valence Bond Theory*; John Wiley & Sons Inc.: Hoboken, NJ, 2007.

(416) Shaik, S.; Hiberty, P. C. In *A Chemist's Guide to Valence Bond Theory*; John Wiley & Sons Inc.: 2007; Chapter 6, pp 116.

(417) Shaik, S.; Danovich, D.; Fiedler, A.; Schröder, D.; Schwarz, H. *Helv. Chim. Acta* **1995**, *78*, 1393–1407.

(418) Shaik, S.; Hiberty, P. C. In *A Chemist's Guide to Valence Bond Theory*; John Wiley & Sons Inc.: 2007; Chapter 3, pp 40−74.

(419) Su, P.; Song, L.; Wu, W.; Hiberty, P. C.; Shaik, S. *J. Am. Chem. Soc.* **2004**, *126*, 13539–13549.

(420) Groves, J. T.; McClusky, G. A. *J. Am. Chem. Soc.* **1976**, *98*, 859–861.

(421) Lee, D.-S.; Nioche, P.; Hamberg, M.; Raman, C. S. *Nature* **2008**, *455*, 363–370.

(422) Rutter, R.; Hager, L. P.; Dhonau, H.; Hendrich, M.; Valentine, M.; Debrunner, P. *Biochemistry* **1984**, *23*, 6809–6816.