

Combined
Quantum Mechanical / Molecular Mechanical
Calculations on Cytochrome P450_{cam}

In a u g u r a l – D i s s e r t a t i o n
zur
Erlangung des Doktorgrades der
Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von
Jan Claasen Curd Schöneboom
aus Waibstadt

2003

Gedruckt mit der Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Heinrich-Heine-Universität Düsseldorf

Referent: Univ.-Prof. Dr. Walter Thiel

Korreferentin: Univ.-Prof. Dr. Christel Marian

Tag der mündlichen Prüfung: 25. Juli 2003

Zusammenfassung

Die vorliegende Arbeit befasst sich mit theoretischen Untersuchungen zur Cytochrom P450 katalysierten Hydroxylierung von C–H-Bindungen. Zur Beschreibung von Häm-Enzym, Substrat und Solvens wird eine kombinierte quantenmechanische/molekülmechanische (QM/MM) Methode verwendet. Im Rahmen dieses Ansatzes wird die elektronische Struktur im reaktiven Zentrum des Enzyms (40–84 Atome) durch Dichtefunktionaltheorie erfasst. Die gesamte Protein- und Solvensumgebung (*ca.* 24000 Atome) wird in den Simulationen durch ein klassisches Kraftfeld repräsentiert.

Zunächst wird die elektronische und geometrische Struktur des katalytisch aktiven Oxo-Eisen(IV) Komplexes $[\text{Fe}^{\text{IV}}(\text{O})(\text{porph}^+)(\text{SR})]$ (SR = Cystein-357, porph = Protoporphyrin IX) in P450_{cam} (*Compound I*) beschrieben, welcher sich bislang dem experimentellen Nachweis entzieht. Es wird dargelegt, wie sich die explizite Berücksichtigung der Proteinumgebung durch das QM/MM-Verfahren im Vergleich zu vereinfachten Modellen des Komplexes in der Gasphase auswirkt. So wird im Enzym ein Porphyrin- π -Radikalkation ($\text{A}_{2\text{u}}$ -Zustand) stabilisiert, während in der Gasphase hauptsächlich ein Schwefel-zentriertes Radikal vorliegt. Wasserstoffbrücken in der Umgebung von Cystein-357 begünstigen Ladungslokalisation am koordinierenden Schwefelatom, wodurch die Fe–S Bindung im Vergleich zu dem isolierten Komplex in der Gasphase verkürzt und verstärkt wird.

Das berechnete Reaktionsprofil für den Abstraktions-Rekombinations- (“*Rebound*”-) Mechanismus der C–H Hydroxylierung läßt darauf schließen, dass die Reaktion in zwei Spin-Zuständen (Dublett und Quartett) stattfindet, wie zuvor aufgrund von einfacheren Modellrechnungen vorgeschlagen wurde (“Zwei-Zustands-Reaktivität”). Während die Umwandlung auf der Dublett-Potenzialfläche asynchron, jedoch effektiv konzertiert abläuft, ist die Reaktion im Quartett-Zustand zweistufig, wobei intermediär ein Substrat-Radikal und ein Hydroxo-Eisen-Komplex vorliegen. Vergleichende Rechnungen in der Gasphase zeigen, dass die polarisierende Wirkung der Proteinumgebung die relative Stabilität von Spin-Zuständen und Redox-Elektromeren beeinflusst.

Rechnungen an der Substrat-freien Form von P450_{cam} sagen die Multiplizität des Grundzustandes korrekt voraus und befinden sich in guter Übereinstimmung mit experimentell bekannten Bindungsabständen und ESR-Hyperfeinkopplungskonstanten der Ligandenatome. Ferner geben die Untersuchungen Aufschluss über die Faktoren, welche die Bindungseigenschaften des axialen Wasserliganden im Enzym bestimmen. Weitere Dichtefunktional-basierte Untersuchungen betreffen die spektroskopischen Eigenschaften von Intermediaten im Enzym und entsprechenden Modellkomplexen. Der Vergleich von Rechnungen an $[\text{FeO}(\text{TPP})]^+$ (TPP = *meso*-tetraphenylporphyrin) und experimentellen Daten seiner Derivate zeigt, dass Heisenberg-Austauschkopplungskonstanten, Mößbauer-Isomerieverschiebungen und Quadrupolaufspaltungen in guter Genauigkeit erhalten werden. Entsprechende Vorraussagen über die spektroskopischen Parameter der hochvalenten Oxo-Eisen(IV)-Zwischenstufe (*Compound I*) in P450_{cam} werden vorgestellt, um deren experimentellen Nachweis zu erleichtern.

Abstract

The present work is concerned with theoretical studies on the cytochrome P450-catalyzed hydroxylation of C–H bonds. To describe the heme-enzyme, substrate, and solvent, a combined quantum mechanical/molecular mechanical (QM/MM) approach is adopted. Density functional theory is employed to treat the electronic structure of the active site (40–84 atoms), while the protein and solvent environment (*ca.* 24000 atoms) is approximated at the classical force field level.

The calculations characterize the electronic and geometric features of the elusive active oxidant of P450_{cam} (compound I), *i.e.* the oxoferryl species $[\text{Fe}^{\text{IV}}(\text{O})(\text{porph}^+)(\text{SR})]$ (SR = cysteine-357, porph = protoporphyrin IX). We analyze how the explicit consideration of the protein environment by the QM/MM treatment influences the results with respect to simplified compound I models in the gas phase. We find that a porphyrin π -cation radical (A_{2u} state) is stabilized in the enzyme, while the gas phase models are mainly sulfur-centered radicals. A hydrogen bonding network around the proximal cysteine-357 favors localization of charge density at the coordinating sulfur atom, which shortens and strengthens the Fe–S bond with respect to the gas phase complex.

The calculated energy profile of the “rebound” mechanism of C–H hydroxylation indicates that the reaction takes place in two spin-states (doublet and quartet), as has been suggested earlier on the basis of calculations on simpler models (“*two-state-reactivity*”). While the reaction on the doublet potential energy surface is non-synchronous, yet effectively concerted, the quartet pathway is truly stepwise, including formation of a distinct intermediate substrate radical and a hydroxo-iron complex. Comparative calculations in the gas phase demonstrate that the polarizing effect of the enzyme environment affects the relative stability of spin states and redox electromers.

Calculations on the substrate-free resting form of P450_{cam} correctly predict the ground state multiplicity and are in good accord with experimentally known bond lengths and EPR hyperfine coupling constants on ligand atoms. Furthermore, these investigations offer insights into the factors that govern the binding properties of the axial water ligand in the enzyme environment.

Additional density functional-based calculations address the spectroscopic parameters of intermediates in the protein environment and corresponding model complexes. The comparison of computational results on the complex $[\text{FeO}(\text{TPP})]^+$ (TPP = *meso*-tetraphenylporphyrin) and experimental data of its derivatives shows that Heisenberg exchange coupling constants, Mößbauer isomer shifts, and quadrupole splittings are obtained with satisfactory accuracy. Corresponding predictions for the spectroscopic parameters of the high-valent oxoferryl intermediate (compound I) in P450_{cam} are presented, to facilitate its spectroscopic detection.

Für Irina

ACKNOWLEDGEMENT

I would like to express my deep gratitude to Prof. Dr. Walter Thiel, for giving me the opportunity to conduct a thesis on a fascinating and stimulating area of research, for his confidence that admitted me to work independently, for his amicable and frank attitude, his accessibility for discussions and his support.

I am indebted to Prof. C. Marian for kindly accepting to act as co-referee for this thesis.

It is a pleasure to thank my colleagues that worked with me on the P450 project: Shimrit Cohen (Hebrew University, Jerusalem), Dr. Hai Lin (Mülheim) and I had a very close, friendly, and fruitful collaboration. I am also grateful to Dr. Nathalie Reuter and Dr. Frank Terstegen for sharing their skills in molecular simulation techniques with me. Prof. S. Shaik is thanked for his stimulating interest and advice.

I am most grateful to Dr. Frank Neese, who patiently introduced me to the field of theoretical EPR and Mößbauer spectroscopy, and always found time (regardless of whether he actually had it) for discussions. Dr. Sebastian Sinnecker and Dr. Eckhard Bill are gratefully acknowledged for giving me insights into calculation and interpretation of spectra.

Special thanks are due to my colleagues of the Thiel group for the good atmosphere and the many interesting and stimulating discussions. Especially, I would like to thank Michael Bühl, Holger Hermann, Hans Martin Senn, and Stephan Thiel for proof-reading parts of this thesis and discussing it with me. A big thank you goes to Axel Koslowski who generously shared his comprehensive skills and knowledge regarding “practical” theoretical chemistry, and for the continuous supply of cookies.

I thank the computer department of the MPI Mülheim, especially Horst Lenk, for help in solving hardware and software problems.

I am obliged to the Fonds der Chemischen Industrie for a Kekulé scholarship and the German Israeli Foundation for financial support regarding travel expenses.

My wife Irina and my parents Sigrid and Klaus have made this work possible through their never ending support and affection. Thank you!

Contents

1	Introduction	1
1.1	General Introductory Remarks and Overview	1
1.2	Cytochrome P450	4
1.2.1	The Catalytic Cycle of P450 Hydrocarbon Hydroxylation	4
1.2.2	Formation of Compound I	5
1.2.3	The Electronic Structure of Compound I	8
1.2.4	The Mechanism of Hydroxylation	12
1.2.5	The Resting State	17
1.3	Theoretical Modeling of Enzymatic Reactions	18
1.3.1	Scope and Limitations of Computational Approaches	18
1.3.2	The QM/MM Approach	21
2	Theoretical Background	27
2.1	Density Functional Theory	27
2.1.1	Foundations of DFT and the Kohn–Sham Formalism	28
2.1.2	Non-Dynamical Correlation and Single-Determinant Nature of DFT	29
2.2	QM/MM Methods	32
2.2.1	The Quantum Mechanical/ Molecular Mechanical Hamiltonian	32

2.2.2	Link Atoms	33
2.2.3	Charge Shifting	34
3	Theoretical Characterization of Compound I	37
3.1	Motivation and Background	37
3.2	Preparation of the System	39
3.3	Computational Methodology	43
3.3.1	QM Regions and Basis Sets	43
3.3.2	Optimized Regions	44
3.3.3	QM/MM Scheme	44
3.3.4	Performance Considerations	46
3.4	Results	48
3.4.1	QM/MM Calibration Study (Snapshot 29)	48
3.4.2	Comparison of Different Snapshots	57
3.5	Discussion	59
3.6	Conclusions	61
4	Investigation of the Rebound Mechanism	63
4.1	Motivation and Background	63
4.2	Computational Methodology	64
4.2.1	QM Regions and Basis Sets	64
4.2.2	Optimized Region	65
4.2.3	QM/MM Scheme	66
4.2.4	QM Calibration Study	66
4.2.5	Strategy	67

4.3	Results	68
4.3.1	Hydrogen Abstraction	68
4.3.2	The Hydroxo Intermediate	79
4.3.3	The Rebound Step	86
4.3.4	The Product Complex	92
4.3.5	Product Release	99
4.4	Discussion	103
4.5	Conclusions	107
5	The Resting State	109
5.1	Motivation and Background	109
5.2	Preparation of the System	112
5.3	Computational Methodology	114
5.3.1	QM Regions and Basis Sets	114
5.3.2	Optimized Regions	115
5.3.3	QM/MM Scheme	115
5.3.4	Strategy	115
5.4	Results	116
5.4.1	Influence of Exact Exchange	116
5.4.2	Influence of the Protein Environment	118
5.5	Discussion	126
5.6	Conclusions	128
6	Calculation of Spectroscopic Parameters	131
6.1	Motivation and Background	131

6.2	Theoretical Background	133
6.2.1	The Spin Hamiltonian	133
6.2.2	Magnetic Perturbing Operators	134
6.2.3	Formulation of Spin Hamiltonian Parameters	136
6.2.4	Calculation of Second Order Contributions by a Double Perturbation Approach	140
6.2.5	Practical Aspects of the Computations	141
6.2.6	Calculation of Mößbauer Parameters	142
6.3	Computational Details	144
6.3.1	Geometries	144
6.3.2	Spectroscopic Parameters	144
6.3.3	Spin Projection	146
6.4	Results and Discussion	148
6.4.1	Model Complexes	148
6.4.2	The Resting State	152
6.4.3	Compound I	154
6.5	Conclusion	159
7	Summary	161
	Bibliography	165
A	Miscellaneous	189
A.1	List of abbreviations	189
A.2	Convergence criteria	190
B	Force Field	191

Chapter 1

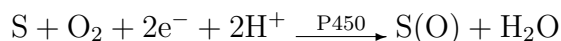
Introduction

1.1 General Introductory Remarks and Overview

Cytochrome P450 enzymes have now been recognized for more than 35 years as one of the most sophisticated and yet common oxidation systems in nature [1]. These ubiquitous enzymes have been isolated from numerous mammalian tissues such as liver, kidney, lung, intestine, and adrenal cortex, as well as from insects, plants, yeasts and bacteria. They constitute a superfamily of more than 500 isozymes cloned and sequenced. The various members of the P450 family exhibit different substrate selectivities which allows them to perform specific bioregulatory functions. Thus, P450 enzymes participate in the metabolism of very diverse compounds, such as steroids, fatty acids and alkaloids. For example, they are involved in every step of steroid hormone biosynthesis. Moreover, P450 enzymes play a crucial role in detoxification of xenobiotics, *e.g.* drugs, by rendering lipophilic compounds water soluble and thus excretable. These enzymes can be isolated, *e.g.*, from liver microsomes. Ironically, in addition to their very beneficial roles in metabolism, biosynthesis, and detoxification, P450 enzymes have been strongly implicated as the activator of many chemical carcinogens, by producing highly reactive products from certain substrates.

From a chemist's point of view, P450 enzymes are remarkable catalysts – they function as monooxygenases, incorporating selectively one atom from dioxygen at physiological conditions into a wide variety of mostly hydrophobic substrates. This reaction, if uncatalyzed, requires high temperatures and proceeds nonspecifically. P450 enzymes mediate an impressive repertoire of reactions, most notably hydroxylations, epoxidations, N-, S-, and O-dealkylations, N-oxidations, sulfoxidations and dehalogenations. Alkane hydroxylation and olefin epoxidation have attracted particular interest, because of the desire to understand the details of biological oxygen activation and transfer, and also because of the potential application of these principles to organic synthesis and to large-scale process chemistry.

Most P450 enzymes (except for one or two that use peroxides) require electrons to reduce molecular oxygen to a state formally equivalent to that of H_2O_2 . As an example, P450_{cam} is reduced *via* the redox proteins putidaredoxin and putidaredoxin reductase, transferring electrons ultimately provided by nicotinamide adenine dinucleotide (NADH) [2]. In contrast, $\text{P450}_{\text{BM-3}}$ interacts directly with a covalently bound reductase containing flavin adenine dinucleotide (FAD) and flavin mononucleotide (FMN), similar to the situation in microsomal P450 enzymes. Hence, in very general terms the overall reaction catalyzed by P450 can be written as



where S symbolizes the substrate, and the protons and electrons are provided by a biological transport system.

The architecture of the active site of P450 enzymes is relatively simple. Its central feature is the heme group, shown in figure 1.1. The heterocyclic ring system of heme is a porphyrin derivative, which consists of four pyrrole rings linked by methene bridges. The porphyrin in heme, with its particular arrangement of four methyl, two propionate, and two vinyl substituents, is known as protoporphyrin IX.

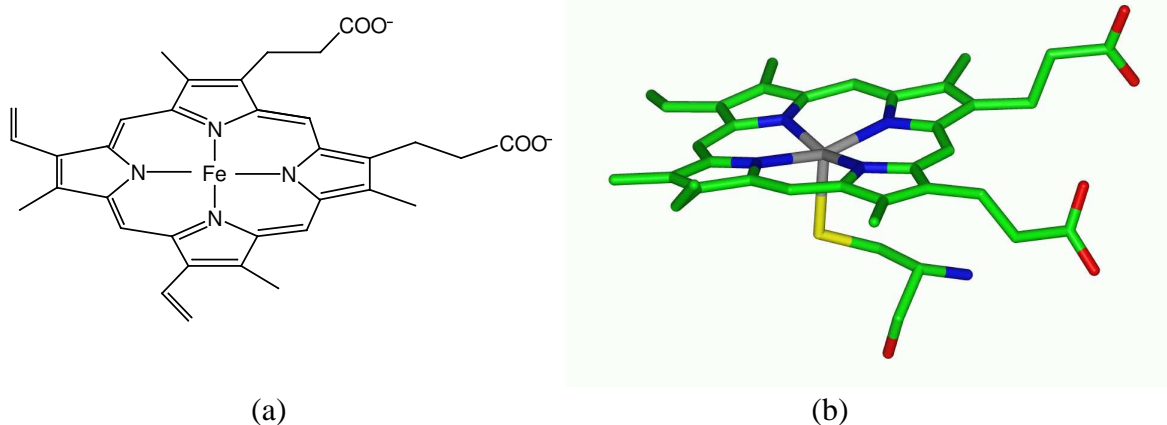


Figure 1.1: The heme prosthetic group of P450 enzymes, an iron-protoporphyrin IX complex. (a) Schematic representation. (b) Stick representation displaying the heme group and the cysteinate ligand (carbon: green, nitrogen: blue, oxygen: red, sulfur: yellow, iron: gray).

This moiety is bound to the protein *via* hydrogen bonds of the two propionate side chains and covalently through a thiolate ligand coordinating to the heme iron. The latter is part of a cysteine in a highly conserved area of the protein, and is placed at the proximal face of the porphyrin. The opposite, distal site is the region where dioxygen and the substrate are bound and transformed. The thiolate ligand is responsible for the characteristic red-shifted Soret absorption maximum at ≈ 450

nm upon CO binding to the reduced ferrous state of the protein to give the $\text{Fe}^{\text{II}}\text{-CO}$ complex. This feature, distinguished from the corresponding maximum at ≈ 420 nm in hemoproteins with the more common histidine ligand [3], has been the reason for the name of these enzymes (*Pigment 450*). The local environment for oxygen binding and activation consists of mostly hydrophobic protein residues and a single highly conserved threonine that has been identified as being essential for catalytic function in most P450 enzymes.

A unique property of the active site is the significant electronic delocalization between the iron center, porphyrin macrocycle, and the axial ligand, which is due to the covalent interactions between these components. This allows for charge transfer and spin density redistribution between metal and ligands. Such effects play a key role in the formation and stabilization of the various heme species present in the catalytic cycle.

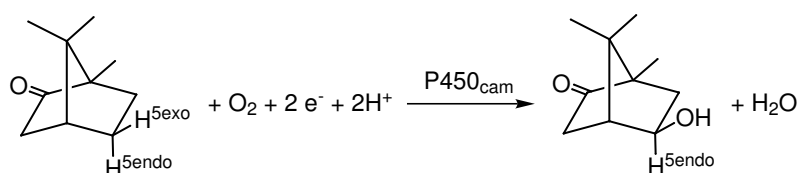


Figure 1.2: Cytochrome P450_{cam} catalyzes the stereoselective hydroxylation of camphor at the 5-*exo* position.

P450_{cam} has *de facto* served as the prototype for the entire P450 family. Using camphor as its only carbon source, the common soil bacterium *Pseudomonas putida* employs P450_{cam} to catalyze the stereospecific hydroxylation at the 5-*exo* position as a first step in a cascade of energy supplying reactions. This soluble enzyme with a molecular mass of *ca.* 45000 Dalton was the first P450 enzyme to be purified and structurally characterized by X-ray diffraction [4]. As P450_{cam} has been intensely studied over the years by numerous biochemical and biophysical techniques [2], a large part of the catalytic cycle is by now well understood. Additionally, synthetic metalloporphyrins have been extensively investigated as models of the active site of P450 and have contributed significantly to a deeper understanding of the properties and reactivity of the many intermediates of the enzymes themselves (see, *e.g.*, [5]). However, many central aspects of the catalytic function are still not resolved experimentally, due to the elusive character of reactive intermediates and the occasional difficulties to interpret spectroscopic or other “indirect” experimental data (*vide infra*).

Theoretical techniques have recently been used successfully to address some of the complex questions associated with P450 catalysis. By providing detailed insight into molecular properties and processes, computational methods have become a powerful and complementary approach to experiment. For example, they allow to study the role of changes in oxidation and spin state of the heme, the influence of the distal ligand on electronic structure and reactivity, or the spectroscopic features of catalytic intermediates [6]. Yet, theoretical methods encountered their own difficulties,

which are mainly associated with the size of the system that has to be considered and the complex electronic structure of the central iron complex. In the present work, the combined quantum mechanical/molecular mechanical (QM/MM) approach is utilized to model unresolved aspects of the P450_{cam}-catalyzed hydroxylation of camphor. In QM/MM calculations, a confined region that is crucial for the chemical properties of a system is treated at a high level of theory (by quantum mechanics – QM), while the environment is taken into account explicitly by a classical force field (molecular mechanics – MM). By treating the entire system with this combined approach, accurate predictions about the active site are possible, while realistically accounting for the influence of the environment on the computed properties at a manageable cost.

This thesis is organized as follows: Section 1.2 presents an overview of the aspects of P450 catalysis relevant to the present work. Section 1.3 discusses the scope and limitations of modern theoretical methods to address questions of biochemical relevance. In chapter 2, the theoretical concepts that underly the computational scheme employed in this work are briefly reviewed. Specifically, section 2.1 deals with aspects of density functional theory that are important in the context of open-shell systems. The QM/MM model employed in the present calculations is presented in section 2.2. In chapters 3 to 6, the results of the QM/MM calculations are presented and discussed. A summary of the results can be found in chapter 7.

1.2 Cytochrome P450

1.2.1 The Catalytic Cycle of P450 Hydrocarbon Hydroxylation

P450-mediated catalytic hydroxylation of non-activated carbon–hydrogen bonds – proceeding at mild conditions and often highly stereoselective – has fascinated and challenged researchers for more than three decades. The field is highly interdisciplinary and involves essentially all areas of chemistry, such as (without claiming completeness) analytical, biochemical, pharmaceutical, physical, synthetic, and theoretical chemistry. Not surprisingly, essentially all existing suitable analytic techniques have been applied to study the intermediates and their transformations associated with the mechanism. Hence, most of the features of the catalytic cycle have been agreed on. The important steps involve (see figure 1.3 on page 6)

1. binding of substrate (RH) to give the enzyme substrate complex **2**,
2. reduction of the substrate-bound complex from the ferric (**2**) to the ferrous (**3**) state by an iron-sulfur or flavoprotein partner;
3. binding of molecular oxygen yielding the ferrous P450-dioxygen complex **4**;

4. reduction by a second electron to give a peroxoiron(III) complex **5**;
5. first protonation at the distal oxygen to give the ferric hydroperoxy intermediate **6**;
6. second protonation and cleavage (heterolysis) of the O–O bond to form a water molecule and the (putative) reactive iron(IV)-oxo species **7**;
7. oxygen atom transfer to the bound substrate, presumably *via* an intermediate **8**, forming the product complex **9**;
8. product dissociation and re-formation of the resting state, a ferric aqua heme complex (**1**).

Extensive experimental studies on cytochrome P450_{cam} and other isozymes have firmly established the ferric (**2**), ferrous (**3**), ferrous dioxy (**4**), anionic ferric peroxy (**5**), and ferric hydroperoxy (**6**) complexes as intermediates in the catalytic cycle of P450 enzymes [7]. However, after injection of the second electron, very rapid reaction steps occur, and despite numerous efforts it has been impossible so far to detect any intermediate in the normal catalytic cycle of P450 enzymes between the ferric hydroperoxy intermediate **6** and the product complex **9**. Based on studies on model compounds and biological analogy with related peroxidase enzymes, the identity of the active oxidant is now generally assumed to be compound I (Cpd I), an oxoiron(IV) porphyrin radical cation (**7**). The consensus mechanism for hydroxylation of hydrocarbons by Cpd I, depicted in figure 1.3, is known as the rebound mechanism and is largely attributed to the work of Groves and coworkers [1]. It involves hydrogen abstraction from the substrate (R–H) followed by rapid recombination (rebound step) of the metal-bound hydroxo radical with the intermediate alkyl (R·) radical. The rebound mechanism is consistent with early stereochemical, regiochemical, and allylic scrambling results observed in the oxidation of norbornane, camphor, and cyclohexene by cytochrome P450 [8,9], yet it has been challenged recently [10,11] by studies that aimed at measuring the rate of the rebound step and the lifetime of the intermediate radicals (*vide infra*).

Current research focuses on the steps that lead to formation of Cpd I, the characterization of this highly elusive species, and the potential involvement of other species in oxidation reactions effected by P450. In the following, some key observations made over the past years are summarized.

1.2.2 Formation of Compound I

Schlichting *et al.* [12] have structurally characterized the ferrous dioxy, ferric hydroperoxy, and product complexes of P450_{cam} using low-temperature crystallographic techniques. They also obtained evidence for a species between the ferric hydroperoxy and product complexes that was tentatively identified as the ferryl species **7**, but this attribution remains uncertain. Electron para-

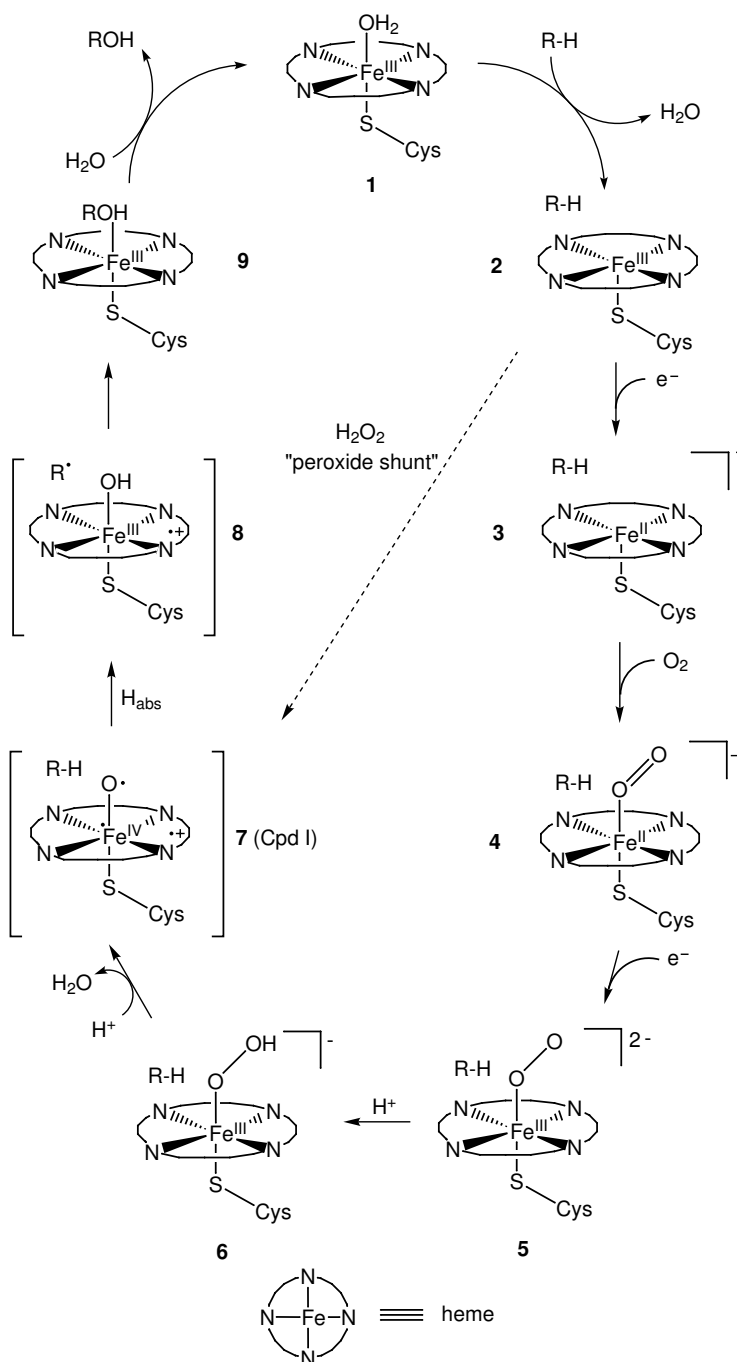


Figure 1.3: The catalytic mechanism of P450 mediated hydroxylation.

magnetic resonance (EPR) and electron nuclear double resonance (ENDOR) studies [13] of the cryogenic reduction of the ferrous dioxy complex by γ -radiation identified the ferric peroxy anion and showed that its protonation at low temperatures yields the ferric hydroperoxy intermediate **6**. Upon warming, this intermediate converted directly into the product complex without detectable accumulation of a ferryl species. Analogous observations were made in another study, where radiolytic reduction of the P450_{cam} ferrous dioxy intermediate at cryogenic temperatures led to the ferric hydroperoxy complex that was identified on the basis of its UV spectrum. Warming of the sample resulted in direct hydroxylation of the bound camphor without any spectral evidence for the intermediacy of Cpd I [14]. The conclusion from these and numerous preceding studies is that Cpd I is present in the catalytic cycle of P450 but eludes detection due to its high reactivity.

Although there is little doubt about the involvement of the ferryl porphyrin radical cation **7** in most of the oxidations supported by cytochrome P450, other intermediates have been proposed to participate in the oxidation of specific classes of substrates. Among these, the most prominent ones are the ferric peroxy anion (**5**) and the ferric hydroperoxy complex (**6**).

The ferric peroxy anion is a strong nucleophile. Convincing evidence has been put forward that it undergoes oxidizing reactions with highly electrophilic substrates such as aldehydes [7, 15]. This type of reaction plays a central role in sterol biosynthetic reactions mediated, *e.g.*, by CYP19 and CYP17. Due to its nucleophilic character, an involvement of this species in hydrocarbon hydroxylation, however, can be ruled out.

The ferric hydroperoxy intermediate has been proposed to participate as an electrophile in the oxidation of electron-rich double bonds, heteroatoms, and even C–H bonds. The potential involvement of this species is suggested by studies of a model metalloporphyrin system. These have demonstrated that the ferryl species is responsible for olefin epoxidation in protic solvents, whereas in aprotic solvents the ferric hydroperoxide acts as the oxidant [16]. Furthermore, it is known that the Fe^{III}–OOH species serves as electrophilic oxidant in the oxidation reaction catalyzed by heme oxygenase. In this special enzyme, the heme serves both as prosthetic group and as substrate. Experimental investigations indicate that the reaction involves electrophilic addition of the terminal oxygen of the ferric hydroperoxide complex to one of the *meso*-carbon atoms of the same heme group [17]. On the other hand, the evidence for the involvement of the ferric hydroperoxide in hydrocarbon hydroxylation remains circumstantial and unconvincing [7]. A further discussion of the potential role of the ferric hydroperoxy complex in hydroxylation reactions is postponed to section 1.2.4.

The route to Cpd I formation has been investigated in P450 enzymes [18] and in model systems [19–21]. The results suggest that heterolytic cleavage of the O–O bond is facilitated by hydrogen bonding of the ferric hydroperoxide, either directly or *via* an intervening water molecule, to a highly conserved threonine residue (Thr252 in P450_{cam}). This residue, in combination with Asp251 and other residues, has been implicated in P450_{cam} as a member of a proton transfer net-

work that promotes the O–O cleavage to give the ferryl species [22]. Molecular dynamics simulations for the reduced ferrous dioxygen species of P450_{eryF} have revealed two dynamically stable hydrogen-bond networks to the distal oxygen atom of the dioxygen ligand [23]. Further DFT studies on this system have demonstrated that the proton transfer from water to the distal oxygen should be rapid and facile [24]. Recent model DFT studies [25] on a proton shuttle similar to that assumed in P450_{cam} predict that the protonation of the ferric hydroperoxide and formation of Cpd I is exothermic and proceeds on two spin state surfaces, *i.e.* in the doublet and quartet state. Moreover, a “push” effect of the thiolate ligand in P450 [18, 26] was proposed to favor heterolytic cleavage, leading to Cpd I formation, over the competing homolytic cleavage. Hirobe and coworkers have demonstrated that for a model porphyrin with a coordinated thiolate tail O–O heterolytic bond cleavage is exclusively observed even in nonpolar aprotic solvents [20, 21]. This indicates that the heterolytic cleavage leading to Cpd I will also take place in the hydrophobic active site environment of P450 enzymes. Consistent with these findings mutation studies on the Cys357His mutant of P450_{cam} demonstrated that the thiolate ligand is crucial for catalytic function [27, 28]. Recent computations supported the pivotal role of the sulfur ligand and revealed the energetic factors that contribute to the “push” effect [29].

The “peroxide shunt” mechanism (see figure 1.3) leads directly from the ferric substrate-bound complex to formation of the ferryl species (Cpd I) by providing H₂O₂ or another small peroxide as a co-substrate. In many, but not all, P450 enzymes it is thus possible to circumvent the step-wise oxygen activation, but the catalytic turnover achieved in this way can yield different product distributions as compared to natural conditions [1]. Yet, peroxides have been used over the years to create and characterize directly the hydroxylating agent of cytochrome P450 enzymes *via* the “peroxide shunt” mechanism. In the most recent of such studies, the reaction of CYP119 from thermophilic bacteria with *meta*-chloroperbenzoic acid produced an intermediate which exhibited UV/VIS spectral features that are consistent with a ferryl porphyrin radical cation [30]. Due to its thermostability, this enzyme is believed to have a more rigid active site structure than its mesophilic counterparts, which could allow for a detectable amount of the active species to build up. The use of peroxides to generate and investigate the active oxidant of the enzyme, however, has to be regarded with caution because it is unclear if such species are identical to those present under normal turnover conditions. One obvious difference is that reactions with peroxides produce detectable protein radicals [31, 32], in contrast to the situation in the native reaction.

1.2.3 The Electronic Structure of Compound I

Both experimental [33] and theoretical [34–36] investigations have shown that the nature of the proximal ligand as well as the proximal/distal binding site environments determine the chemical nature of the high-valent ferryl intermediate, Cpd I. As an example consider the enzyme chloroper-

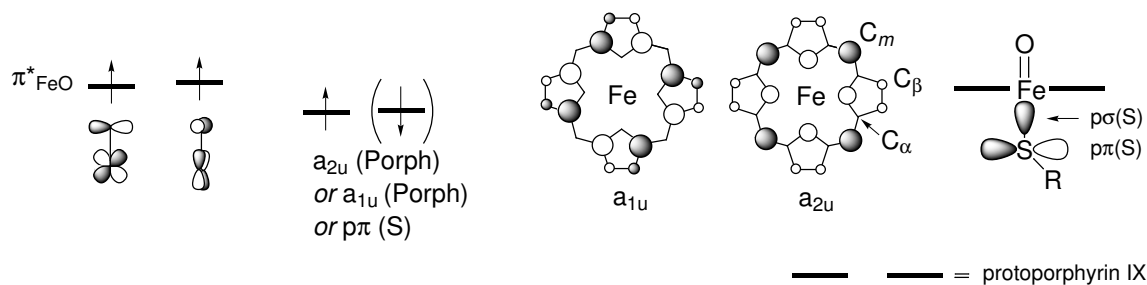


Figure 1.4: Electronic structure of compound I – key orbitals and occupations. The labels on the porphyrin carbon atoms denote the α , β , and $meso$ -positions respectively.

oxidase (CPO) which shares the linkage of the heme to a proximal cysteine with P450 enzymes. Yet like horseradish peroxidase and other peroxidases (where the linkage is provided by a histidine) CPO carries out one- and two-electron peroxidations. In addition to peroxidase activity, CPO also performs P450-typical reactions such as benzylic hydroxylation [37]. The variable properties and reactivity of Cpd I depending on the proximal ligand and the architecture of the binding pocket (in different enzymes) has led to the notion that it behaves as a “*chameleon*” species which adapts to the specific environment of a given enzyme [36].

The electronic nature of Cpd I of P450 enzymes has been studied theoretically at the DFT level employing different truncated model representations of the active species [34, 35, 38–41]. DFT studies on Cpd I model systems [38, 39] using a methyl mercaptate (SMe) ligand or a cysteinate anion devoid of internal hydrogen bonding [41] predict a sulfur-centered radical to be the electronic ground state. Characteristically, these studies yield a long Fe–S bond length of ≈ 2.7 Å and *ca.* 80 % of the unpaired spin density on the proximal sulfur. In contrast to these findings, DFT calculations by Shaik and coworkers [42] as well as by Harris and Loew [43] indicate a predominant porphyrin π cation radical nature of Cpd I.

The conflicting information deduced from different theoretical studies can be rationalized by an analysis of the key orbitals of Cpd I. These are depicted in figure 1.4. Cpd I is a triradicaloid with two unpaired electrons occupying d-p π^* orbitals of the FeO moiety with parallel spin. The third unpaired electron can be located in a π orbital of the porphyrin ring, either of the a_{2u} or a_{1u} variety¹, giving rise to a porphyrin π cation radical. In contrast, a sulfur centered radical results from a singly occupied sulfur based π orbital. All DFT calculations reported so far agree that the lowest porphyrin π cation radical state is predominantly of A_{2u} character. However, as noted above, some studies assign the ground state to a sulfur radical while others predict dominant porphyrin radical character. According to a recent INDO/S-CI treatment the multiconfigurational ground state wavefunction is predominantly of the A_{2u} -type, but contains significant contributions

¹ For simplicity, the labels of the irreducible representations of the idealized D_{4h} -symmetric system are used throughout.

from sulfur radical and A_{1u} configurations [40].

Shaik and coworkers have presented a thorough investigation of the nature of the different Cpd I forms [34, 44, 45] by means of calculations on model compounds. They find that the computed electronic structure of truncated models of Cpd I markedly depends on the representation of the proximal sulfur ligand L. The energy difference between the sulfur and the lowest porphyrin π radical states for $L = ^-\text{SCH}_3$ is less than 1 kcal/mol whereas for a more realistic representation with $L = \text{cysteinate}$ the state splitting is 7–8 kcal/mol with the ground state corresponding to the porphyrin π cation radical. Interestingly, the simplest model using $L = ^-\text{SH}$ gives sulfur radical–porphyrin π cation radical energy splittings of 5–6 kcal/mol in good agreement with the full cysteine model. These findings demonstrate that the $L = ^-\text{SCH}_3$ approximation to cysteine yields artificial results, probably due to the strong electron-donating capability of the methyl group which leads to an overestimation of the stability of sulfur radical states.



Figure 1.5: The redox mesomorphism underlying the P450 Cpd I state mixing. (a) Schematic resonance structures. (b) A plot of the mixed $a_{2u} + \sigma(S)$ orbital as obtained from MO calculations.

It is important to note that even in the states of predominant A_{2u} character, the $p\sigma(S)$ orbital at sulfur mixes with the a_{2u} -type porphyrin orbital. The resulting Cpd I species is in fact best described as a mixture of the porphyrin π cation radical and a sulfur radical. This feature is generally termed “redox mesomorphism” and is schematically depicted in figure 1.5. In molecular orbital (MO) calculations, the unpaired spin density is distributed over the atoms with large contributions to the porphyrin a_{2u} orbital (*i.e.* the *meso*-carbons and pyrrole nitrogens) and the sulfur atom. An intriguing consequence of this mixing is the sensitivity of the electronic structure of Cpd I to environmental effects. Specifically, Ogliaro *et al.* demonstrated that additional hydrogen bond donors to the sulfur and a polarizing medium (represented by a continuum solvent model) effectively increase the charge density on the sulfur atom and thus enhance the stabilization of the porphyrin π cation radical. The more pronounced these effects, the larger (smaller) is the fraction of the unpaired spin density on the porphyrin (sulfur) ligand. In line with the resonance structures in figure 1.5, a stabilization of negative charge on the sulfur atom increases the donor capabilities of the thiolate ligand. This has consequences for the Fe–S bond: the calculated Fe–S bond length is shortened with respect to the isolated system in the gas phase and the Fe–S bond energy is estimated to increase by *ca.* 9 kcal/mol [34, 44].

As discussed above (figure 1.4), Cpd I exhibits a $\text{Fe}^{\text{IV}}\text{O}$ moiety with total spin $S = 1$ and a third unpaired electron with spin $S' = 1/2$ delocalized over the ligands. The antiferromagnetic coupling of the two spin systems gives rise to a low-spin doublet state ($S = 1/2$), whereas ferromagnetic coupling yields a high-spin quartet state ($S = 3/2$). The resulting two lowest electronic states, ^2A and ^4A , are nearly degenerate; DFT calculations predict relative energies on the order of a few tenths of a kcal/mol [35]. The precise ordering of these states is very sensitive to the basis set as well as to environmental effects such as medium polarization [35, 45] and to different models of the proximal ligand. Basis set improvement and inclusion of medium polarization lead to a slight preference of the antiferromagnetic doublet state, which becomes lower in energy than the quartet state by *ca.* 0.1 kcal/mol [45]. However, this state ordering may be system-dependent and may possibly change with different environmental factors, such as bound substrate or hydration and H-bonding to the proximal sulfur. A clear-cut conclusion on the basis of the model computations is thus not possible.

As noted above, the elusive character of Cpd I of P450_{cam} has so far prevented the spectroscopic determination of the nature of its electronic ground state. EPR data are available from the more stable Cpd I of chloroperoxidase (CPO) [46]. In CPO, which also contains a cysteine sulfur ligand bound to the heme iron, the interaction between the ferryl $S=1$ and the porphyrin $S'=1/2$ spin is antiferromagnetic and moderately strong ($J = -35 \text{ cm}^{-1}$). Resonance Raman spectra of this species show an upshift of some characteristic frequencies that argue for a $^2\text{A}_{1\text{u}}$ rather than a $^2\text{A}_{2\text{u}}$ ground state [47]. In Cpd I of horseradish peroxidase (HRP-I), in which the axial ligand is a histidine, the ground state is weakly antiferromagnetic [48]. Based on frequency shifts of marker bands in the resonance Raman spectra of HRP-I, Paeng and Kincaid assigned the electronic ground state to be $\text{A}_{2\text{u}}$ [49], in agreement with earlier EPR and Mössbauer studies [48, 50]. In contrast, another study reported frequency shifts analogous to porphyrin π -cation radicals, which are $\text{A}_{1\text{u}}$ [51]. Thus, the identity of the electronic ground state of this species remains uncertain.

Several spectroscopic studies on Cpd I model compounds have appeared. *Meso*-tetraaryl porphyrin complexes such as $[\text{Fe}^{\text{IV}}\text{O}(\text{TMP})^+]$ (TMP: tetramesityl porphyrin) and its derivatives have been investigated by means of UV/VIS, NMR, EPR, Mössbauer, and EXAFS spectroscopy (for reviews, see [52, 53]). These systems typically exhibit a strong ferromagnetic coupling between the oxo-ferryl and the porphyrin π -cation radical spin. The electronic ground state of most such complexes has been assigned as $\text{A}_{2\text{u}}$. Electron-withdrawing substituents at the *meso*-positions lower the energy of the $\text{a}_{2\text{u}}$ orbital relative to the $\text{a}_{1\text{u}}$ orbital. In the case of pentafluorophenyl-substituted porphyrins the state ordering is thus reversed, and the $\text{A}_{1\text{u}}$ state is the ground state [52, 53].

1.2.4 The Mechanism of Hydroxylation

The consensus mechanism for hydrocarbon hydroxylation by cytochrome P450 is the hydrogen abstraction–oxygen rebound pathway [1]. As suggested by this mechanism, Cpd I abstracts a hydrogen atom from a carbon atom of the substrate to produce an iron-bound hydroxyl radical and a carbon radical. These two species then recombine in the rebound step to produce the oxidized product. This scenario, with its discrete radical intermediate, nicely explains many of the early experimental observations on P450 hydroxylations, such as the partial scrambling of stereochemistry [54], allylic rearrangements [9], and the high intrinsic isotope effects [55]. Stereochemical scrambling has also been observed in P450_{cam}, which can remove either an *exo* or an *endo* hydrogen but delivers the oxygen solely to produce the *exo*-hydroxy isomer [54, 56]. The intrinsic kinetic isotope effects (KIE) k_H/k_D measured for the oxygen insertion into a C–H bond are very large (*e.g.*, hydroxylation of tetra-*exo*-deuterated norbornane: 11.5 ± 1.0 [8]; benzylic hydroxylation of [1,1-D]-1,3-diphenylpropane: 11 [55]). These large KIEs are consistent with a hydrogen abstraction reaction, characterized by a transition state in which the C–H bond is essentially half broken in a linear arrangement $[C \cdots H \cdots O]$.

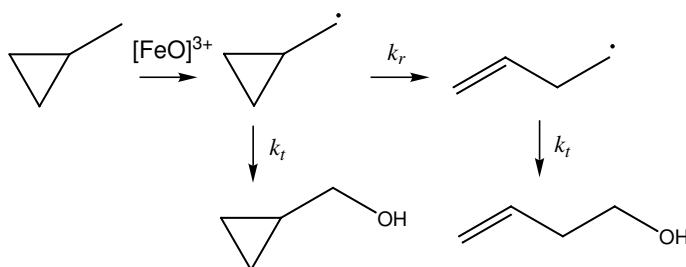


Figure 1.6: The principle of radical clock probes.

So-called “radical clocks” have been employed to probe the rate of the oxygen rebound step and the lifetime of the proposed radical intermediate. The common feature of radical clocks is a cyclopropyl ring either containing the carbon atom that becomes a radical center in the course of the reaction or being vicinal to that center. The ratio of products derived from the unrearranged *vs.* the rearranged radical is thus determined by the relative magnitudes of the rate constants for radical quenching k_t and rearrangement k_r (see figure 1.6). Early studies on the oxidation of bicyclo[2.1.0]pentane (**10**, see figure 1.7) produced a 7:1 mixture of unrearranged *vs.* rearranged alcohols and allowed the rebound rate to be estimated as $1.4 \times 10^{10} \text{ s}^{-1}$ [7, 57, 58].

Subsequently, more sophisticated radical probes have been designed [10] with higher and precisely known rearrangement rate constants. These probes are expected to give a higher portion of rearranged products and thus a more accurate determination of the lifetime of the radical intermediate. However, with increasing rearrangement rates of the radical probes, the apparent recombination rates of the rebound step also increase, resulting in unreasonably fast reactions and

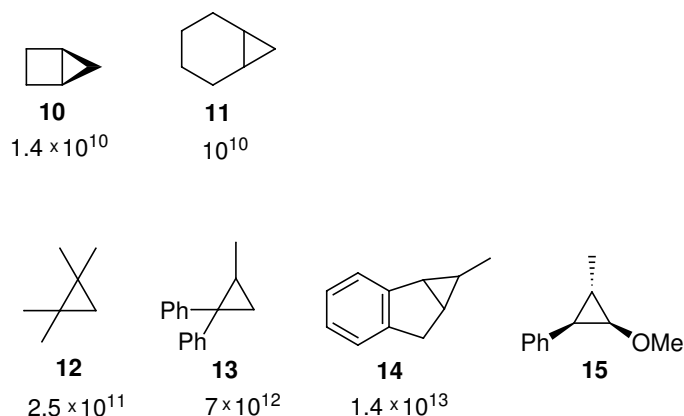


Figure 1.7: Different radical clocks. The values below correspond to the apparent rates of the oxygen rebound step [s^{-1}] deduced from the resulting ratio of rearranged vs. unrearranged products.

short lifetimes. For example, the use of highly substituted cyclopropanes (*e.g.*, **12**, **13** in figure 1.7) leads to values of $2\text{--}7 \times 10^{12} s^{-1}$ [59]. Noting that a reaction that is entirely controlled by encounter proceeds with rate of approximately $6 \times 10^{12} s^{-1}$ at $37^\circ C$, these rebound rates seem impossibly fast. Initially, these results were rationalized by the notion that the protein sterically inhibits the rearrangement reaction, thus leading to erroneously small values for the rearrangement rate k_r . This explanation, however, had to be discarded when the structurally rigid radical clock **14** (expected not to be influenced by steric constraints) also yielded an incredibly fast rebound rate of $1.4 \times 10^{13} s^{-1}$ [60]. However, two recent independent studies of the oxidation of norcaradiene (**11**) by different P450 enzymes yielded comparably slow radical recombination rates ($\approx 10^{10} s^{-1}$ in ref. [61]), in good agreement with results obtained earlier for the oxidation of **10** [61, 62]. In summary, some of the simple cyclopropylmethyl compounds and some bicyclic systems lead to reasonable lifetimes for a radical intermediate, while all of the phenylcyclopropylcarbinyl probes suggest impossibly short lifetimes. These discrepancies preclude firm conclusions regarding the mechanism of hydroxylation, and call for further experimental investigation.

A concerted nonsynchronous oxygen insertion mechanism has been proposed by Newcomb *et al.* [63] to explain the results of radical clock experiments, which suggest that radicals formed during the P450 hydroxylation reaction mediated by Cpd I are not true intermediates. This mechanism is sketched in figure 1.8. In this scenario, the C–H bond of the substrate is assumed to be attacked by the oxo-ligand in a side-on fashion. The intermediate radicals are components of a reacting ensemble (or transition structure) with lifetimes on the order of femtoseconds. To account for the rearrangement products found in various studies, *e.g.*, allylic rearrangements or stereochemical scrambling, the reaction is predicted to proceed through a bifurcated transition state where one reaction channel leads to the formation of the hydroxylated product and the other, minor pathway leads into a radical manifold with subsequent rearrangement. However, as noted by Ortiz de

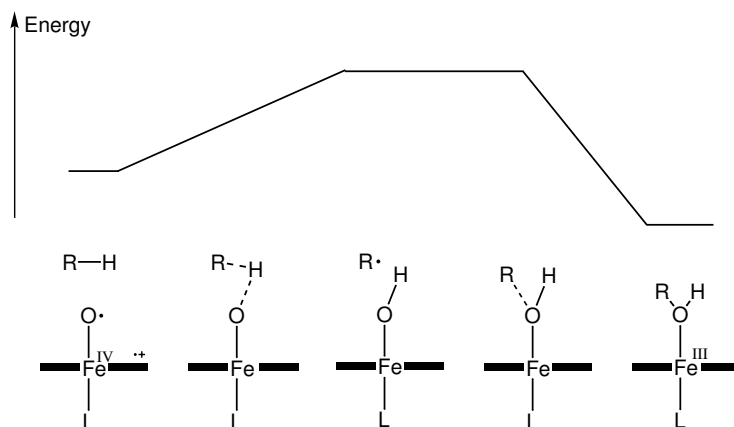


Figure 1.8: The concerted nonsynchronous oxygen insertion mechanism suggested in ref. [63]. For explanations see text.

Montellano [7], a weakness of this scenario is the variety of products that must be explained by electronically and structurally different bifurcated transition states.

A mechanism based on computational results has recently been postulated by Shaik [64–68] that combines features of both the stepwise oxygen rebound and the concerted insertion mechanism (see figure 1.9). The reactive species is predicted to be the ferryl species Cpd I, which exists in two nearly degenerate electronic states, a doublet and a quartet spin state [34]. Both species can abstract hydrogen from non-activated C–H bonds *via* essentially identical transition states. This accounts for the observed isotope effects as well as the parallels in hydrogen abstraction reactivity between P450 enzymes and the *tert*-butoxy radical [69, 70]. These transition states lead to intermediates in which the alkyl radical is coordinated to the iron-bound hydroxyl species. These intermediates exist either in the doublet or quartet spin state, which are close in energy. Corresponding DFT calculations predict that the doublet species collapses in an essentially barrierless process to the hydroxylated product, *i.e.* the overall reaction on the doublet surface proceeds in a non-synchronous, yet effectively concerted, fashion with no true intermediate. The process is barrierless because during recombination the half-filled porphyrin π orbital accepts one “excess” electron and the thiolate ligand increases its interaction with the iron atom resulting in a net energy gain. In contrast, the rebound step on the quartet spin surface exhibits a significant energy barrier according to the calculations. Therefore, this process gives rise to intermediate substrate radicals which have a sufficiently long lifetime to undergo the characteristic rearrangement reactions. The calculated barrier on this pathway is thought to be due to the three unpaired spins that are required to maintain the quartet state. This results in the occupation of an energetically high lying iron d_{z^2} (σ^*) orbital in the course of the reaction which causes a loss of binding energy in the S–Fe–O moiety. The concept of *Two State Reactivity* (TSR), *i.e.* the involvement of two electronic states in the reactions of a system, was recently proposed to be an important feature in the understanding of organometallic systems in general [67, 72]. Subsequent DFT studies by other groups concerning

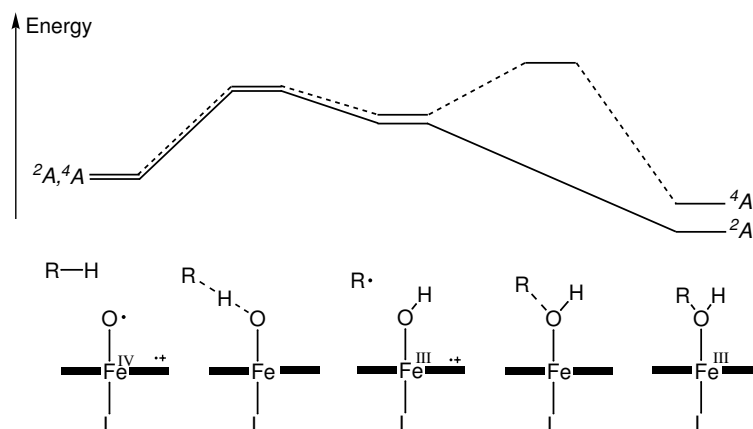


Figure 1.9: The two-state reactivity scenario proposed by Shaik *et al.* [71]. For explanations see text.

ethane and camphor hydroxylation [25,73–77] and methane hydroxylation [78] confirmed that the hydroxylation mechanism is typified by TSR.

The use of specially tailored radical clocks such as **15** has lead to the observation of small amounts of products that apparently originate in rearrangement of a cationic intermediate. Newcomb *et al.* have proposed that a second oxidant is involved in P450 C–H bond hydroxylation that produces corresponding cationic intermediates. A species that has been linked to such cationic hydroxylating activity is the iron hydroperoxy complex [10, 11, 79–81]. The assumed mode of action of this intermediate is depicted in figure 1.10. Insertion of “HO⁺” from the hydroperoxide into a C–H bond would yield a protonated alcohol, which in turn could lead to carbocationic rearrangement products. It should be noted, however, that the absolute amount of cationic rearrangement products is quite small, thus the corresponding pathway does not seem to play a crucial role in normal oxidation reactions.

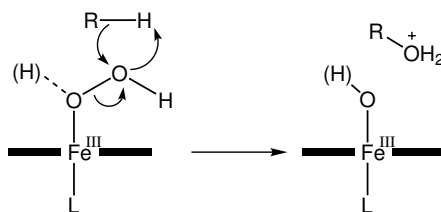


Figure 1.10: The cationic HO⁺ insertion mechanism. For explanations see text.

Arguments against the involvement of the iron hydroperoxy species as hydroxylating agent come from spectroscopic and theoretical investigations. EPR and ENDOR results on cryoreduced P450_{cam} strongly suggest that the ferric hydroperoxy species plays no role in the normal catalytic mechanism of C–H hydroxylation effected by this enzyme [13]. In these experiments,

the hydroperoxy complex was generated under cryogenic conditions, and detected by EPR and ENDOR spectroscopy. Upon warming to ≈ 200 K, it was quantitatively converted into 5-*exo*-hydroxycamphor, the normal product of P450_{cam} hydroxylation. The initial product observed after warming had the hydroxyl group of the oxidized camphor coordinated to the heme iron. This is in agreement with hydroxylation by a ferryl oxene species. Instead, the cationic oxidation mechanism (figure 1.10) suggests that the initially formed hydroxycamphor, which contains the distal oxygen of the hydroperoxy moiety, is initially detached from the heme iron. Assuming this scenario, the experimental finding would imply that hydroxycamphor had displaced the hydroxide or water ligand that is left bound to the heme iron. Yet, such a displacement reaction appears to be implausible at 200 K. Moreover, the ENDOR studies in D₂O suggest that the proton whose bond to the C-5 carbon of the substrate is broken during hydroxylation is trapped in the product complex as the hydroxyl proton. This is consistent with a ferryl-mediated oxidation, whereas a hydroperoxy mechanism implies that the hydroxyl proton could derive from solvent as well. Computational studies at the DFT level of theory have led to the conclusion that the Fe^{III}-OOH moiety is a poor oxidant as compared to the ferryl oxo species and is very unlikely to participate in the P450 catalytic cycle [82]. This conclusion is based on three key results of the computations: (i) The conversion of PorFe^{III}-OOH₂⁺ into the ferryl oxene species is a barrierless process [43, 82]. (ii) The ferric peroxy anion is basic and a good nucleophile. (iii) The PorFe^{III}-OOH moiety is a poor oxidant and is therefore unlikely to be involved in the oxidation of C-H bonds.

Very recently, Jin *et al.* studied the Thr252Ala mutant of P450_{cam}, that lacks the proton shuttle delivering the second proton to form Cpd I [83]. It was shown that the mutant enzyme catalyzes the epoxidation of camphene, albeit with a much smaller efficiency than the wild type enzyme. This indicates that in the absence of Cpd I, the iron hydroperoxy species acts as the oxidant. However, the mutant enzyme exhibited no activity for camphor hydroxylation.

An alternative explanation for the occurrence of cationic intermediates invokes the rebound mechanism. Here, electron transfer from the carbon radical produced by the initial H-abstraction could compete with the oxygen rebound step, thus yielding a cationic intermediate. Direct electron transfer from a substrate to a P450 enzyme as part of a hydrocarbon hydroxylation has been observed in P450 catalysis [84].

In summary, the available results from experiment and theory support the TSR scenario for hydrocarbon hydroxylation. The formation of a radical intermediate on the quartet surface explains many of the characteristics of the reaction, such as stereochemical randomization and structural rearrangements. The presence of two pathways allows the degree of such reactions to vary both with the substrate and the enzyme. The small amounts of cationic rearrangements could result from electron transfer from the radical intermediate to the iron-hydroxyl complex.

1.2.5 The Resting State

The resting form of the substrate-free cytochrome P450 isozymes with a known crystal structure is the ferric aquo heme complex **1** in figure 1.3 [1]. From X-ray diffraction studies on cytochrome P450_{cam} it is known that the sixth ligand is either a water molecule or a hydroxide ion, with a Fe–O distance of 2.28 Å [85]. The identity of a water molecule as the sixth ligand in the resting form of P450_{cam} was confirmed by ENDOR and electron spin-echo envelope modulation (ESEEM) spectroscopy techniques in ¹⁷O-enriched water [86, 87]. The ESEEM experiment gave a direct measure of the anisotropic hyperfine interaction for the protons of the water ligand. Both the ENDOR and ESEEM spectral features could be satisfactorily simulated by assuming two magnetically equivalent protons and a water in an upright conformation (perpendicular to the heme plane).

The ferric iron atom in P450_{cam} equilibrates between the low-spin doublet ($S = 1/2$) and the high-spin sextet ($S = 5/2$) state. The low-spin state is favored in the absence of a substrate, and the iron atom is hexacoordinated. In the presence of the substrate camphor, the spin equilibrium shifts toward the high-spin form and the sixth ligand is displaced, resulting in a pentacoordinated iron (see figure 1.3). In addition to the substrate-induced shift of the spin equilibrium, an increase in the redox potential from about –300 mV to about –173 mV accompanies the change from the low-spin to the high-spin configuration. This shift in redox potential has mechanistic significance. The physiological reductant of P450_{cam}, putidaredoxin, exhibits a redox potential near –196 mV so that reduction of the substrate-bound ferric high-spin P450_{cam} is thermodynamically favorable, while reduction of the ferric low-spin resting form is not [88].

The finding of a Fe^{III} low-spin resting state in P450 enzymes is unexpected because other heme enzymes, such as horseradish peroxidase, cytochrome *c* peroxidase and metmyoglobin are typified by either a sextet or quartet ground state, although with subtle variations of populations of different spin states. The ground states of two known model compounds with water ligands, a ferric diaqua porphyrin complex [89] and an aqua thiolate heme complex [90], are also high-spin. The origin of the preferred low-spin configuration has been studied intensely by means of calculations [6]. An early study [91] on the [Fe^{III}(protoporphyrin IX)(SMe)(H₂O)] model system employing the semi-empirical restricted open-shell INDO/S method determined a sextet ground state, however, with a small energy separation of 4 kcal/mol to the doublet and 6 kcal/mol to the quartet state. When the electrostatic interactions of the complex with the enzyme environment were taken into account by adding the surrounding atoms as point charges to the one-electron Hamiltonian, the doublet state became the ground state in agreement with experimental results. These findings suggest that the protein, by controlling the spin state, plays an important and previously unidentified role in determining the active-site redox potential. In subsequent studies the origin of the low-spin state found for the resting form has been reexamined by DFT calculations [6, 42, 92, 93]. These

calculations reveal that the relative state energies depend on the functionals and basis sets used, however, they generally favor the low-spin doublet state as the ground state without invoking the electric field of the enzyme environment. Scherlis *et al.* [94] performed DFT single point calculations using geometries derived from X-ray data and fixed protein point charges in the QM treatment. In accord with previous DFT studies, the doublet state was found to be the ground state. However, in contrast to the results of the abovementioned INDO/S study [91], the inclusion of the polarizing effect of the enzyme environment was found to stabilize the quartet state, thus reducing the doublet-quartet gap.

Taken together, the results from previous theoretical studies indicate that the relative energies of the low-spin and high-spin states are small and depend not only on the computational approach but also on environmental factors. A definite conclusion regarding the ground state multiplicity on the basis of computations has not yet been reached.

Another open question concerns the conformation of the water ligand. While the ESEEM experiment suggests an upright conformation (*vide supra*), the model computations favor a “tilted” conformation, where the water protons form hydrogen bonds with the porphyrin nitrogen atoms. This disagreement is probably due to the neglect of suitable hydrogen bond acceptors in the model computations, which are present in the protein pocket.

1.3 Theoretical Modeling of Enzymatic Reactions

1.3.1 Scope and Limitations of Computational Approaches

Much effort in contemporary biochemical research is devoted to a detailed understanding of the factors that govern enzyme function. Enzymes can provide rate accelerations of the order of 10^7 or more over the rates in aqueous solution [95, 96]. Which factors are responsible for the high efficiency of enzyme catalysis?

Since the three-dimensional structure of biological macromolecules was made accessible by X-ray crystallography [97], much progress has been made in exploiting innovative experimental techniques for elucidating protein structures and dynamics. For instance, recent advances have been brought about by nuclear magnetic resonance (NMR) [98] and mass spectroscopy. These developments have recently been acknowledged through the Nobel Prize in Chemistry 2002 for “the development of methods for identification and structure analyses of biological macromolecules”, received by Kurt Wüthrich (nuclear magnetic resonance), John B. Fenn (electron spray ionisation, mass spectroscopy) and Koichi Tanaka (soft laser desorption, mass spectroscopy). Other spectroscopic techniques that are especially useful for the study of metalloenzymes are EPR, ENDOR, and Mößbauer spectroscopy. Biochemical research has provided insight into protein function by

numerous methods to identify intermediates along a catalytic pathway and to elucidate the role of catalytic residues, *e.g.*, through site-directed mutagenesis or chemical modifications. Kinetic measurements are used to obtain information about the energy barriers associated with the transformations leading from substrate to product [95]. Genomics now opens the way to the identification and synthesis of a plethora of proteins. Theoretical investigations have contributed significantly to the understanding of enzyme catalysis at the microscopic level. Empirical force field methods, typically used in conjunction with molecular dynamics (MD) or Monte Carlo (MC) techniques, have helped to unravel the dynamics of biomolecular systems in condensed phases, the role of intra- and intermolecular interactions, and their impact on catalytic activity [99, 100].

Despite the tremendous progress that has been made, it is safe to say that at present there is still no quantitative understanding of “how enzymes work” [101]. This has been partly attributed to the difficulties of applying “rigorous quantum mechanical and thermodynamic analyses” to the simulation of “active sites in huge molecules surrounded by water” [101, 102]. To gain insight into enzyme function, theoretical methodologies are needed that allow us to identify transient intermediates and transition states (treating electronic and structural changes) along a catalytic pathway, while accounting for the surrounding environment in a realistic manner. Quantum mechanical (QM) calculations are suitable to address such questions, and they can – in principle – provide results within “chemical accuracy” (≈ 1 kcal/mol). But the time needed to accomplish a conventional QM calculation scales formally as $O(N^3) - O(N^7)$, depending on the level of theoretical rigour. Here, N is the number of basis functions, a measure for the size of the system considered (see table 1.1). While improvements both in algorithms and computer power will make such calculations more feasible in the future, these methods are currently not applicable to systems as large as enzymes or solute/solvent systems. Another challenge to computer simulations of macromolecules is the great conformational flexibility of such systems. They are usually characterized by many configurations with similar energy, and the only meaningful energetic quantity is the free energy of the system. The statistical mechanical calculation of the free energy or associated quantities usually requires the consideration of millions of different configurations, which is currently beyond the scope of QM methods. Therefore, QM calculations are often carried out on truncated models of the full system (such as the most essential atoms in the active site of a protein) in the gas phase or a continuum solvent environment. While this approach has proven to give useful insights in many cases, it suffers from the fact that a lot of important interactions, say of a substrate with the surrounding enzyme, are completely neglected. This approach can give qualitative information about the intrinsic reactivity of an enzyme, yet it cannot explain enzyme specificity or the role of microsolvation in an enzyme environment. In fact, enzyme catalysis is best described by comparing corresponding reactions in (aqueous) solution to those in the specific protein environment. By comparing the corresponding reaction profiles, the important elements of the protein responsible for the rate acceleration can be identified.

Empirical force field (molecular mechanics, MM) methods which employ a classical representation of molecular systems scale formally as $O(N^2)$ (see table 1.1), where N now corresponds to the number of particles of a system. They are in general computationally much cheaper than quantum mechanical approaches². Yet, the standard MM schemes are by construction not able to describe processes that involve electronic rearrangements, such as charge transfer, electronic excitation, and the breaking or making of covalent bonds. In the cases where analytical potentials are fitted to reproduce *ab initio* QM results, the parameterization is time-consuming and the transferability of the resulting force field questionable. Especially problematic in this regard are metalloenzymes which contain transition metals in the active site. The interplay between the protein environment and the often complex and variable electronic structure of the embedded transition metal (in terms of oxidation state, multiplicity, orbital occupancy, *etc.*) can in general not be captured by one single force field parameter set.

Table 1.1: Different computational methods, typical size of considered systems and formal scaling behavior (N equals the number of basis functions in the case of QM methods and the number of particles in the case of force field methods).

Method	No. of atoms	Scaling
<i>Ab initio</i> , electron correlation	1 – 10	$\approx N^5 \dots N^7$
<i>Ab initio</i> , SCF	1 – 100	N^4
Density functional methods	10 – >100	N^4
Semi-empirical methods	10 – >1000	N^3
Empirical force fields	10 – >100000	N^2
Linear scaling QM methods	10 – \approx 10000	N
Combined QM/MM methods	10 – >100000	n.a.

The need to represent a large system as a whole in an effective and practical fashion and at the same time to account for changes in electronic structure has led to the development of new methodologies, which can essentially be grouped into two categories. First, there are linear scaling QM methods which pursue the idea to reformulate the algorithms used in the QM calculation in an intelligent manner, such that the computational effort increases linearly with system size, *i.e.* with a scaling behaviour of $O(N)$. This is achieved by exploiting the locality of the interactions in molecular systems and the associated sparsity of the large matrices that have to be treated in QM calculations. Reviews on this subject can be found, *e.g.*, in refs. [103–106]. The second and to date more widely used approach for the modeling of reactions in large and complex systems are combined quantum mechanical/molecular mechanical (QM/MM) schemes, which will be presented in more detail in the following section.

² The prefactor c in the formal scaling relation $c \times N^x$ is also very small in the MM case.

1.3.2 The QM/MM Approach

The QM/MM method was introduced by Levitt and Warshel as early as 1976 [107] and was widely adopted in the 1990's [108–112]. The concept behind QM/MM methods is consistent with basic ideas from enzymology, specifically the notion that in most enzymes the reaction is confined to a small number of atoms in the active site. In this *subsystem* of atoms, the electronic rearrangements associated with chemical reactivity take place. QM/MM schemes use a QM method to treat the electronic degrees of freedom explicitly for the reactive subsystem, while the remaining part of the system is approximated by a classical force field. The coupling of QM and MM schemes provides a potential energy function that can describe the reactivity of a wide range of large and complex systems without the need to reparameterize potential functions for every new reaction. These arguments carry over to reactions of molecules in a surrounding solvent [110] or adsorbed on a solid phase [113–115].

A number of different implementations of this concept has been reported (for a recent review see ref. [116]). Specialized approaches and variations of the general theme include the quantum-classical molecular dynamics method of McCammon *et al.* [117, 118], the ONIOM method of Morokuma *et al.* [119, 120] and the effective fragment potential (EFP) method of Gordon *et al.* [121, 122]. A related method, although not QM/MM in a strict sense, is the empirical valence bond (EVB) technique of Warshel and co-workers [123, 124].

The following discussion will give a brief survey over the general concepts of QM/MM approaches. Section 2.2 presents a more detailed description of the scheme that was adopted in the present work.

First, the *entire* system S has to be divided into an electronically important *inner* region I that is treated quantum mechanically, and an *outer* part O that is described by a classical molecular mechanics scheme (see figure 1.11). The main challenges in QM/MM methods are to define

1. a suitable representation of the interactions between both subsystems and
2. a consistent treatment of the boundary region between QM and MM subsystems.

Energy expressions

The master equation for the total energy of the entire QM/MM system can be defined in two ways.

Additive QM/MM scheme. This approach treats the energies of the QM and MM subsystems E_{QM}^{I} and E_{MM}^{O} as complementary. The energy of the entire system $E_{\text{QM/MM}}^{\text{S}}$ is given by summation of both contributions and the addition of a suitable term $E_{\text{QM-MM}}^{\text{I,O}}$ that describes the interactions

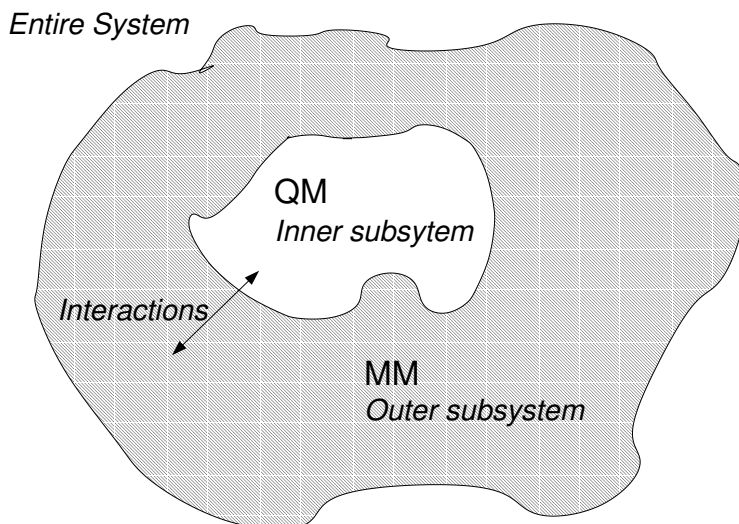


Figure 1.11: Terminology for regions used in QM/MM schemes.

between the QM and MM subsystems, as expressed in equation 1.1.

$$E_{\text{QM/MM}}^{\text{S}} = E_{\text{QM}}^{\text{I}} + E_{\text{MM}}^{\text{O}} + E_{\text{QM-MM}}^{\text{I,O}} \quad (1.1)$$

This method has been widely adopted in biomolecular QM/MM simulations. It is straightforward to implement and is applicable to a wide range of problems. Difficulties arise for the treatment of the boundary region if covalent bonds connect QM and MM regions, since in that case a rigorous definition of the interaction energy $E_{\text{QM-MM}}^{\text{I,O}}$ is not possible. However, several approximate approaches to treat such systems have been developed (*vide infra*).

Subtractive QM/MM scheme. In this scheme, three calculations are performed: The energy of the entire system is calculated on the lower level of theory (MM), whereas the energy of the inner region is calculated both at the lower and higher (QM) level of theory. The QM/MM energy of the entire system can then be written as

$$E_{\text{QM/MM}}^{\text{S}} = E_{\text{QM}}^{\text{I}} + E_{\text{MM}}^{\text{S}} - E_{\text{MM}}^{\text{I}} \quad (1.2)$$

Most applications of this approach have concerned organometallic and zeolite chemistry. The main advantage is that no explicit coupling term $E_{\text{QM-MM}}^{\text{I,O}}$ is required, because all interactions between the two regions are consistently treated by the force field (in the term E_{MM}^{S}). The subtraction ($E_{\text{MM}}^{\text{S}} - E_{\text{MM}}^{\text{I}}$) ensures that the MM energy terms for the inner part exactly cancel, provided that the QM region is large enough. This makes the approach particularly attractive for systems with covalent linkages between QM and MM subsystems. The force field is not required to be able to describe the energetics of the inner region, *e.g.*, in a reaction. However, the force field must be able to calculate the interactions between both regions at all points along a reaction coordinate, which requires, *e.g.*, parameters for atomic charges on all atoms. This is problematic if the elec-

tronic situation in the reacting region varies in the course of the reaction, which makes this scheme therefore not generally applicable.

QM–MM interactions

In the context of additive QM/MM schemes, the interactions between atoms in the QM and MM region that give rise to the term $E_{\text{QM-MM}}^{\text{I,O}}$ have to be considered explicitly. In the simple case where covalent bonds between the two regions are not present, one is faced only with the non-bonded interactions. These comprise short-range (van-der-Waals) and electrostatic terms.

The van-der-Waals interactions are generally calculated at the force field level, which requires, however, the availability of suitable parameters also for the atoms in the QM region. In solvation studies, where such interactions are expected to be significant, a re-fitting of the corresponding parameters has been found to improve the results [125, 126].

The electrostatic energy terms between QM and MM centers can be treated in various ways, using a well-defined hierarchy of coupling models [127, 128]:

Model A: Mechanical embedding. The electrostatic QM–MM interactions are treated at the force field level.

Model B: Electronic embedding. The fixed point charges of the MM region enter the calculation of the QM part, resulting in polarization of the QM electron density. The QM/MM Coulomb energy is evaluated from the MM atomic point charges and the electrostatic potential generated by the QM region at the position of the respective MM centers.

Model C: Model B including polarization of the MM part, by using a polarizable force field.

Model D: Model C with a self-consistent treatment of QM and MM polarization.

In most studies of biomolecular system, model B (electronic embedding) has been adopted and proven to give satisfactory results [110, 112]. Care has to be taken, however, when applying this approach, since charge transfer between the QM and MM region is impossible, and therefore the QM region must be chosen large enough to incorporate the essential fluctuations of electron density. It should be stressed that appropriate correction terms to the QM–MM interactions have to be introduced regardless of the chosen model if covalent bonds are cut by the QM–MM boundary.

Boundary region

The ideal situation for defining QM and MM subsystems is met when there is a natural division of the entire system into weakly interacting regions, *e.g.*, a solute (QM part) and a solvent (MM

part). Generally however, this division may involve the cutting of covalent bonds, which causes a major perturbation of the QM electronic structure.

A widely used approach to deal with this situation is to introduce link atoms [108, 109], *i.e.* an additional QM atom is placed along each QM–MM bond in order to saturate the free valencies of the frontier QM atoms. Link atoms are treated explicitly in the QM calculation and serve to mimic the bond at the QM–MM boundary. An obvious requirement is that interactions with the MM region are not calculated for the link atoms, since they are not part of the physical system. While van-der-Waals terms including link atoms can simply be omitted in the MM calculation, it is quite involved to define a suitable correction for the electrostatic QM–MM interactions for all embedding models apart from model A (*vide supra*). A number of different correction schemes has been designed and critically reviewed [102, 128–130]. Generally, the link atom approach is regarded to be reliable, provided that the frontier bond is placed sufficiently far away from the reactive region. A second issue concerns the nature of the bond that is being cut: preferable are non-polar unconjugated single bonds. Usually hydrogen atoms are used as link atoms. It has been suggested that by making an appropriate choice for the link atom, the nature (in terms of strength, polarity) of the frontier bond may be tuned [116].

To circumvent the introduction of additional (unphysical) atoms, alternative approaches to treat the QM–MM boundary have been devised. In semi-empirical treatments of the QM subsystem, methods based on hybrid orbitals [107, 131, 132] or specially parameterized pseudoatoms [133] have been introduced. Although these schemes are methodologically more elegant than link atom schemes, their performance has not yet been shown to be consistently better [129, 130]. In QM/MM schemes that employ *ab initio* QM methods, hybrid orbital [134–137] and pseudoatom approaches [138] have been implemented. It is yet unclear whether and to which extent these are superior to the newer link atom methods. Murphy *et al.* [137] have demonstrated that their frozen hybrid orbital method can reproduce conformational energetics quite accurately (≈ 5 kJ/mol), however, the formalism of the method is rather complex and it has to be reparametrized every time a given QM–MM bond is placed in a different environment.

Choice of QM and MM method

QM method. The QM calculation usually is the computationally most demanding part of a QM/MM energy or gradient evaluation. Obviously, sophisticated QM methods, like *ab initio* electron correlation approaches, make the QM/MM treatment more accurate, yet the associated computational cost can impose severe constraints on the size of the QM subsystem or the intended application (*e.g.*, the statistical mechanical averaging of properties). It has to be stressed, however, that the overall accuracy of a QM/MM calculation is dependent on various factors, such

as the MM model, the QM–MM interaction model, and the method of sampling conformational space. Therefore the accuracy of the QM/MM calculation can only be systematically improved by equally improving all components involved, and not by only choosing a higher-level QM method – in that case it will only become more time consuming.

Semi-empirical, density functional, and *ab initio* QM methods can be and have been used in QM/MM studies [139]. The choice will depend on the complexity of the electronic situation in the reactive subsystem and on the desired accuracy of the treatment. For instance, standard semi-empirical methods are not suitable for the treatment of open-shell transition metal systems, but they are computationally cheap enough to allow for a sampling of phase space by MD or MC techniques [110]. Calculations to determine the potential of mean force (PMF) [140] of a reaction are usually carried out with semi-empirical treatments of the QM region. Higher-level QM methods are typically used in QM/MM geometry optimizations. Although statistical averaging is neglected, this can be a powerful approach to elucidate the factors underlying enzyme catalysis. However, some insight into the (potentially high-dimensional) system is required in this case to avoid artefacts due to conformational complexity. A promising development are recent QM/MM schemes that use the Car–Parrinello MD method to treat the QM subsystem on the DFT level of theory and allow for a simultaneous dynamic propagation of the entire system [141–143].

MM method. The choice of the MM model depends on the nature of the system under study. Biomolecules are normally represented by valence force fields specifically designed for this purpose, *e.g.*, AMBER [144, 145], CHARMM [146], GROMOS [147], or OPLS [148]. More general valence force fields like CFF [149] or MM3 [150] are usually employed for studies in the area of organometallic or organic chemistry. These valence force fields are constructed from bonding terms (bond stretches, angle bends, *etc.*) and non-bonded terms (van-der-Waals and electrostatic terms). In contrast, ionic force fields are based on electrostatic and van-der-Waals terms exclusively. They are used in studies of ionic solids, *e.g.*, zeolites and are exemplified by the GULP force field [151]. For the use of these schemes in QM/MM calculations, a re-fitting of the force field atomic charges and van-der-Waals parameters is often required.

Chapter 2

Theoretical Background

2.1 Density Functional Theory

As mentioned briefly in section 1.3.2, the quantum mechanical method in a QM/MM calculation has to fulfill certain requirements. First, the method should be capable of adequately describing the properties (in terms of geometric and electronic structure) and the reactivity (in terms of energetics of the processes involved) of the QM part. At the same time, it has to be computationally efficient. This is crucial because the essential part of the QM subsystem has to be considered together with a sufficiently large fraction of the environment in order to minimize boundary effects due to the QM-MM partitioning. In the present work, the QM subsystem consists of a transition metal complex (an iron-porphyrin), which is characterized by an open-shell electronic structure and a conjugated π -system. The delocalized nature of the electron distribution requires the inclusion of 40 to 83 atoms into the QM calculation. The method of choice for the present purposes is density functional theory (DFT), which has proven to be successful in describing complex electronic situations in many cases at a moderate computational cost. However, it is well known that DFT with present-day functionals often gives unreliable results in cases where strong non-dynamical correlation effects are important, as met, *e.g.*, in bond-breaking processes or in fractionally occupied degenerate electronic configurations (*e.g.*, di-, triradicals). Since such situations are important in the context of the present work, this chapter will give a condensed discussion of the basic physical concepts of DFT with a focus on the description of systems with strong non-dynamical correlation. Detailed expositions of the general theory and the performance of DFT methods are available in the literature (*e.g.*, refs [152–154]) and will not be repeated here.

2.1.1 Foundations of DFT and the Kohn–Sham Formalism

Density functional theory is founded on the first Hohenberg–Kohn theorem [155] which states that the exact ground state energy E_0 of a nondegenerate system is a unique functional of the electron density ρ_0 only. In other words, the electron density uniquely determines the Hamiltonian operator and therefore all properties of a given system. Since the electron density only depends on three spatial variables regardless of the size of the system, its use as the basic variable in the energy expression means a tremendous reduction of complexity in comparison to wave function approaches. A wave function depends on $4N$ variables, where N is the number of electrons – three spatial and one spin variable per electron (see, *e.g.*, ref. [156]). Consequently, the exact non-relativistic ground state energy within the Born–Oppenheimer approximation can be written as a functional of the ground state density ρ_0 :

$$E_0[\rho_0] = T[\rho_0] + J[\rho_0] + E_{\text{ncl}}[\rho_0] + E_{\text{eN}}[\rho_0] + E_{\text{NN}}. \quad (2.1)$$

Here, T is the electronic kinetic energy, J the classical Coulomb interaction of the electrons, E_{eN} the electron-nuclear attraction, E_{NN} nuclear-nuclear repulsion, and E_{ncl} the *non-classical* contribution to the electron-electron interaction containing all effects due to exchange and Coulomb correlation. The nuclear-nuclear repulsion term E_{NN} is a constant within the Born–Oppenheimer approximation and will be dropped in the following discussion.

Direct functionals of the kinetic energy $T[\rho]$ perform poorly for molecular systems, in fact they often do not even predict chemical bonding. This deficiency was overcome with the introduction of the Kohn–Sham (KS) formalism in 1965 [157]. The KS concept features the introduction of a fictitious reference system of *non-interacting electrons* built from the Kohn–Sham orbitals φ_i . The exact wave function of such a non-interacting system can be taken as a single Slater determinant Φ_s , giving rise to an associated density ρ_s . For this system the kinetic energy T_s can be exactly expressed as

$$T_s = -\frac{1}{2} \sum_i^N \langle \varphi_i | \nabla^2 | \varphi_i \rangle. \quad (2.2)$$

The reference system is defined by a Hamiltonian which contains an effective local potential V_s :

$$\hat{H}_s = -\frac{1}{2} \sum_i^N \nabla_i^2 + \sum_i^N V_s(\mathbf{r}_i) \quad (2.3)$$

The central assumption in the Kohn–Sham formalism is that the effective potential V_s can be chosen such that it yields *via* the resulting orbitals φ_i the same ground state density as the density of the real, interacting system ρ_0 .

$$\rho_s = \sum_i^N |\varphi_i|^2 = \rho_0 \quad (2.4)$$

Thus the largest fraction of the kinetic energy of the real system is captured by the kinetic energy expression T_s of the fictitious reference system. The remainder is merged together with the non-classical energy contributions of electron-electron interactions E_{ncl} into an *exchange-correlation* energy E_{XC} , which is defined by “inverting” equation 2.1:

$$E_{\text{XC}}[\rho] \equiv (T[\rho] - T_s[\rho]) + (E_{\text{ee}}[\rho] - J[\rho]) \quad (2.5)$$

This yields the total energy expression for the real system:

$$E[\rho] = T_s[\rho] + J[\rho] + E_{\text{XC}}[\rho] + E_{\text{eN}}[\rho] \quad (2.6)$$

The Kohn–Sham orbitals are determined by applying the variational principle and minimizing this energy expression by varying the orbitals φ_i under the usual orthonormality constraint $\langle \varphi_i | \varphi_j \rangle = \delta_{ij}$. The result are the Kohn–Sham equations, a set of N one-particle eigenvalue equations.

$$\left(-\frac{1}{2}\nabla^2 + \left[\int \frac{\rho(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_2 + V_{\text{XC}}(\mathbf{r}_1) - \sum_A^M \frac{Z_A}{r_{1A}} \right] \right) \varphi_i \quad (2.7)$$

$$= \left(-\frac{1}{2}\nabla^2 + V_{\text{eff}}(\mathbf{r}_1) \right) \varphi_i = \epsilon_i \varphi_i \quad (2.8)$$

In this equation, V_{XC} is the potential associated with the exchange-correlation energy E_{XC} . In fact, this term is the great unknown in DFT, and the quest for good approximate formulations of this functional is an active area of research [152]. The expression in square brackets, *i.e.* V_{eff} equals the Kohn–Sham potential V_s .

$$V_s = V_{\text{eff}} = \int \frac{\rho(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_2 + V_{\text{XC}}(\mathbf{r}_1) - \sum_A^M \frac{Z_A}{r_{1A}} \quad (2.9)$$

Provided that the explicit forms of all the potentials contributing to V_s are known, the Kohn–Sham orbitals may be determined, which in turn define the density of the reference system and hence the ground state density of the real system. Inserting this density into equation 2.6 gives the ground state energy. In reality, the exact form of the exchange-correlation functional is not known, and one has to resort to approximations to this functional.

2.1.2 Non-Dynamical Correlation and Single-Determinant Nature of DFT

There is no rigorous theoretical definition of the term non-dynamical correlation. It is generally regarded to be relevant when a wave function has to be represented by a linear combination of a few Slater determinants in order to give a qualitatively correct description of the corresponding system. Non-dynamical correlation occurs, *e.g.*, in bond-breaking processes, avoided crossing situations, excited states, and fractionally occupied degenerate electronic configurations. The prototypical

example is the bond-breaking in H_2 . The principles of quantum mechanics require that even in the dissociation limit (a supermolecule of two non-interacting hydrogen atoms) the wave function retains the $D_{\infty h}$ symmetry of the molecule. Furthermore, the spin density has to be zero everywhere in space and the relative energy of the system in the dissociation limit must equal twice the energy of the isolated hydrogen atom, a feature termed size-consistency. On the basis of the Hohenberg–Kohn theorem, DFT should be able to describe this process since the energy depends on the ground state density only (eq. 2.1). Yet all present-day functionals in spin-restricted calculations predict an energy of the system in the dissociation limit which is much too high. While this is a particular example, the problem of DFT to describe non-dynamical correlation correctly is a rather general issue and has gained much attention [158]. This deficiency is often attributed to the single-determinant nature of the KS approach. Yet the single-determinant representation of the reference wavefunction is generally not the cause of error. By using an essentially exact Kohn–Sham potential V_s for H_2 , Gritsenko and Baerends demonstrated that in fact a single Slater determinant for the non-interacting reference system is sufficient to describe this molecule even at significantly stretched bond distance, where non-dynamical correlation effects are large [159]. These authors numerically determined the potential V_s for H_2 by constructing Kohn–Sham orbitals from a nearly exact charge density, which was taken from high-level *ab initio* calculations, *i.e.* full configuration interaction with large one-electron basis sets. The resulting wave function of the non-interacting system as obtained from the KS orbitals is obviously a bad representation for the real wave function. However, the correct potential energy curve for the H_2 dissociation can be obtained using a single-determinant ansatz for the KS orbitals – provided that the exact functional is known. Such cases are termed *non-interacting pure state V_s representable*. In these cases the reason for the wrong description of non-dynamical correlation is not the single determinant nature of DFT, but the approximate density functionals that are currently in use. Schipper *et al.* [160] later went on to demonstrate that for some systems (*e.g.*, for the $^1\Sigma_g^+$ state of the C_2 molecule) it is in fact not possible to fit the electron density to a single Slater determinant. In other words, for these systems the non-interacting ground state density cannot be represented by a single KS determinant. It was only possible to fit to a determinant with (nearly) degenerate orbitals with fractional occupation numbers (FONs). To be physically founded, these fractional occupation numbers must follow from ensemble averaging. Thus, an ensemble of a small number of accidentally degenerate KS determinants is required to represent the ground state density. Such systems are termed *non-interacting ensemble V_s representable*.

Unrestricted Kohn–Sham (UKS)

As noted above, the KS formalism is in principle – with the exact functional being available – suitable for any kind of atom or molecule, regardless of closed-shell or open-shell character. Even if a system with an odd number of electrons is considered, the basic variable remains the total

density which can be constructed from the individual α and β spin densities, $\rho = \rho_\alpha + \rho_\beta$. In reality, the current approximate functionals cannot account for open-shell problems in a realistic manner, and an unrestricted ansatz for the reference determinant Φ_s is commonly used. To this end, spin-density functionals for exchange and correlation are employed that explicitly depend on the α and β spin densities. This unrestricted Kohn–Sham (UKS) approach can capture more of the essential physics in open-shell systems than the spin-independent functionals. However, the use of UKS has some drawbacks, which can be best understood by considering again the example of the H_2 dissociation. In contrast to the restricted KS method, UKS gives a qualitatively correct potential energy curve, including the asymptotic region with the relative energy of the dissociation limit being equal to twice the energy of the isolated atom. However, the dissociation leads to localization of spin density with opposite sign on the nuclei – the spin density mimics the atomic densities. Thus, the spatial symmetry of the total wave function (charge density) is lower than the symmetry of the system. This unphysical phenomenon is termed (spatial) symmetry breaking and the resulting wave functions are often termed broken symmetry solutions. Additionally, spin unrestricted wave functions are not eigenfunctions of the \hat{S}^2 operator anymore, and breaking of the spin symmetry may occur in addition to the spatial symmetry breaking. The deviation of the expectation value $\langle \hat{S}^2 \rangle$ from the correct value $S(S+1)$ is often taken as a measure for the “contamination” of the wave function with wave functions of other multiplicity. From a fundamental point of view, it is important to note that the KS determinant in DFT does not represent the true wave function of the real system, but only the wave function of the fictitious non-interacting reference system. It is often stressed that the true wave function is not available in DFT. Thus, the deviation of $\langle \hat{S}^2 \rangle$ is not necessarily a probe for the quality of the UKS energies. This is in contrast to the situation in the unrestricted Hartree–Fock (UHF) formalism, where the unrestricted Slater determinant is actually meant to represent the true wave function. In fact, there are opinions which state that KS determinants for open-shell systems which are not spin-contaminated are actually wrong. In practice, however, the performance of the UKS approach depends on the system under consideration, and some severe failures have been noted [161, 162]. In summary, the broken symmetry solutions obtained by the UKS scheme provide in many cases with strong non-dynamical correlation effects surprisingly good results, at the price of unphysical spin densities. Nevertheless the performance of UKS should be evaluated against reliable reference data, and its applicability remains a case-by-case decision.

2.2 QM/MM Methods

2.2.1 The Quantum Mechanical/ Molecular Mechanical Hamiltonian

The present calculations employ an additive QM/MM scheme. The effective Hamiltonian of the entire system $\hat{H}_{\text{QM/MM}}^{\text{S}}$ is expressed as a sum of three terms, which represent the quantum mechanical subsystem ($\hat{H}_{\text{QM}}^{\text{I}}$), the classical partition ($\hat{H}_{\text{MM}}^{\text{O}}$) and the interactions between the two subsystems ($\hat{H}_{\text{QM-MM}}^{\text{I,O}}$). It depends on the positions of the nuclei of inner (\mathbf{R}^{I}) and outer (\mathbf{R}^{O}) regions, as well as on the positions of the electrons in the QM region, \mathbf{r} . (Sums over electron positions are denoted by index i , those over positions of nuclei in the inner region by index A , and those over positions of nuclei in the outer region by index J .)

$$\hat{H}_{\text{QM/MM}}^{\text{S}}(\mathbf{r}, \mathbf{R}^{\text{I}}, \mathbf{R}^{\text{O}}) = \hat{H}_{\text{QM}}^{\text{I}}(\mathbf{r}, \mathbf{R}^{\text{I}}) + \hat{H}_{\text{MM}}^{\text{O}}(\mathbf{R}^{\text{O}}) + \hat{H}_{\text{QM-MM}}^{\text{I,O}}(\mathbf{r}, \mathbf{R}^{\text{I}}, \mathbf{R}^{\text{O}}) \quad (2.10)$$

The total energy of the entire system is given by the expectation value of the effective hamiltonian $\hat{H}_{\text{QM/MM}}^{\text{S}}$.

$$E_{\text{QM/MM}}^{\text{S}} = \langle \Psi(\mathbf{r}, \mathbf{R}^{\text{I}}, \mathbf{R}^{\text{O}}) | \hat{H}_{\text{QM}}^{\text{I}}(\mathbf{r}, \mathbf{R}^{\text{I}}) + \hat{H}_{\text{QM-MM}}^{\text{I,O}}(\mathbf{r}, \mathbf{R}^{\text{I}}, \mathbf{R}^{\text{O}}) | \Psi(\mathbf{r}, \mathbf{R}^{\text{I}}, \mathbf{R}^{\text{O}}) \rangle + E_{\text{MM}}^{\text{O}}. \quad (2.11)$$

The pure MM term can be removed from the integral, since it is independent of the electronic coordinates. It is evaluated by a force field calculation on the outer region O . In the present work, the electronic embedding scheme is adopted. This is appropriate when the environment contains strongly polar or charged groups that influence the electron density distribution of the QM subsystem. The interaction between the QM and MM regions then is given by

$$\hat{H}_{\text{QM-MM}}^{\text{I,O}} = - \sum_{i,J} \frac{q_J}{|r_{iJ}|} + \sum_{A,J} \frac{q_J Z_A}{|R_{AJ}|} + \sum_{A,J} \left(\frac{A_{AJ}}{R_{AJ}^{12}} - \frac{B_{AJ}}{R_{AJ}^6} \right). \quad (2.12)$$

Here, q_J is the atomic point charge on the MM atom J , r_{iJ} is the distance of electron i to MM atom J , Z_A is the core charge of the QM atom A , R_{AJ} is the distance of QM atom A to MM atom J , and A_{AJ} and B_{AJ} are the Lennard–Jones parameters for QM atom A interacting with MM atom J . The first two terms in equation 2.12 describe the electrostatic interaction of the QM region with the MM atoms. The MM atomic charges enter the one-electron hamiltonian \hat{h} of the QM calculation. Thus, the QM electronic structure can respond to its environment by polarization of the electron density through the surrounding protein/solvent.

$$\hat{h}_{\mu\nu}^S = \hat{h}_{\mu\nu}^I - \sum_J^O q_J V_{\mu\nu}^J \quad (2.13)$$

The resulting Coulomb energy can be expressed as the sum over the individual contributions that are given by the product of the charge of MM atom J and the electrostatic potential Φ^J generated by the QM region at the position of J .

$$E^{\text{coul}}(A, J) = \sum_J^O q_J \Phi^J \quad (2.14)$$

$$\Phi^J = - \sum_{\mu,\nu}^I P_{\mu\nu} V_{\mu\nu}^J + \sum_A^I Z_A V^{AJ} \quad (2.15)$$

$P_{\mu\nu}$ are the density matrix elements of the QM calculation. $V_{\mu\nu}$ and V_{AJ} are the nuclear attraction integrals and Coulomb repulsion terms. In the case of density functional theory or general *ab initio* methods, they are straightforwardly defined as

$$V_{\mu\nu}^J = \langle \mu(1) | r_{1J}^{-1} | \nu(1) \rangle \quad (2.16)$$

$$V^{AJ} = R_{AJ}^{-1}. \quad (2.17)$$

The last term in equation 2.12 represents the van-der-Waals interaction between QM and MM atoms. These energy contributions are calculated by the force field. This requires the specification of van-der-Waals parameters also for the atoms of the inner region. The last two terms of equation 2.12 are added to the total energy once the electronic energy has been determined by a self-consistent-field (SCF) procedure and do not affect the electron distribution directly, but do affect the geometry of the total system through the computed gradients (and thus indirectly also the electronic structure).

2.2.2 Link Atoms

In the present work, link atoms are used to saturate free valencies at the QM border. An extra nuclear center, in this case always a hydrogen atom, is introduced together with basis functions and electrons required to form a covalent bond to the QM region that serves to mimic the bond to the MM region.

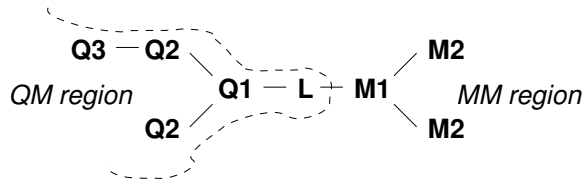


Figure 2.1: Labeling of the atoms at the QM–MM border region.

The coordinates of the link atoms \mathbf{R}_L are expressed as a function of the real atom coordinates \mathbf{R}_X . In this way, it is possible to eliminate them from the set of coordinates that is used in geometry optimizations or molecular dynamics runs. In the present implementation, the link atoms are constrained to lie on the axis defined by the QM and MM atoms at the border Q1–M1 (for labeling of the atoms see figure 2.1). The distance of the bond Q1–L is kept fixed. Because the link atoms experience non-zero forces, it is necessary to add a term to the real atom forces to account for the changes in the link atom positions resulting from movement of the real atoms. When working in cartesian coordinates, chain rule differentiation can be used to establish the contribution to the forces on the real atoms.

$$E_{\text{QM/MM}}^S = E(\mathbf{R}_X, \mathbf{R}_L(\mathbf{R}_X)) = \tilde{E}(\mathbf{R}_X) \quad (2.18)$$

$$\frac{\partial \tilde{E}}{\partial R_{X,r}} = \frac{\partial E}{\partial R_{X,r}} + \sum_{L,s} \frac{\partial E}{\partial R_{L,s}} \frac{\partial R_{L,s}}{\partial R_{X,r}} \quad (2.19)$$

Here, r, s refer to spatial components (x,y,z). The derivatives $\frac{\partial R_{L,s}}{\partial R_{X,r}}$ form a 3×3 matrix describing the coupling of the link atom and real atom motions as a function of the constraint term. There is a term for each atom which appears in the definition of the position of the link atom. In the present case, there is a correction term for the atoms at each end of the bond Q1–M1.

Due to the introduction of link atoms, some adjustment to the force field contribution is required to avoid double counting of energy terms. For example, the MM term for the angle bend Q2–Q1–M1 is effectively supplemented by the bending potential for the group Q2–Q1–L, which is part of the QM calculation. The restoring forces acting on the link atom are transferred to atom M1 as is obvious from equation 2.19. In the present scheme, the force field potential energy term of the respective angle bend is deleted, since the link atom bending potential replaces it. Likewise, dihedral angles of the type Q3–Q2–Q1–M1 are deleted.

2.2.3 Charge Shifting

The close contacts between link atoms and the frontier atoms M1 which they are replacing would lead to unphysical distortions of the QM electron density due to electrostatic attraction or repulsion. In the present scheme this is avoided by selectively deleting the charge on atom M1 from the

list of point charges visible to the QM calculation. Preservation of the total charge of the MM system is achieved by distributing (*shifting*) the charge of atom M1 equally over the next connected atoms M2. This charge shift introduces an unphysical dipole moment to the MM subsystem. To compensate this, dipole moments of opposite sign are placed on the M2 centers, aligned along the M2–M1 axis. The resulting matrix of MM charges exhibits the same total charge and dipole moment as the reference MM system, however, the charges close to the link atoms are removed [163].

Chapter 3

Theoretical Characterization of Compound I

3.1 Motivation and Background

As detailed in section 1.2, the available experimental evidence indicates that compound I (Cpd I) – the key species in P450_{cam} mediated hydroxylation of C–H bonds – exists under turnover conditions [13] but so far eludes detection due to the special kinetic scenario of the mechanism (see figure 3.1). Schlichting *et al.* [12] tentatively assigned X-ray data obtained from low-temperature crystallographic techniques to the ferryl species **7**, however this attribution remains uncertain. The authors themselves note that this structure suffers from contamination with other species in the crystal. Moreover, consecutive ENDOR studies at cryogenic conditions [13] did not provide any evidence for the accumulation of Cpd I. Hence, there is no reliable information on the geometric features of this species. Spectroscopic studies of Cpd I in related enzymes CPO and HRP could not resolve uncertainties about the electronic nature of Cpd I [164], *i.e.* whether the A_{2u} or the A_{1u} state is preferred. This situation invites theory as a viable method that can characterize this elusive species. Yet, theory encounters its own difficulties, associated with the size of Cpd I, and with the need to go beyond isolated molecule calculations. Density functional (DFT) studies [6, 34, 35, 39–41, 44, 45, 73, 165] of different simplified models with pure and hybrid functionals depend on the representation of the cysteinate ligand and the protoporphyrin IX. Most strikingly, the various truncated cysteine-ligand models generated conflicting information regarding the Fe–S bond lengths and the distribution of the unpaired electron density that typifies the ground state of the species. These discrepancies underscore the fact that calculations using truncated and gas phase models cannot resolve the dilemma of Cpd I, and one must resort to more realistic models which involve the electric field and structural constraints of the protein pocket.

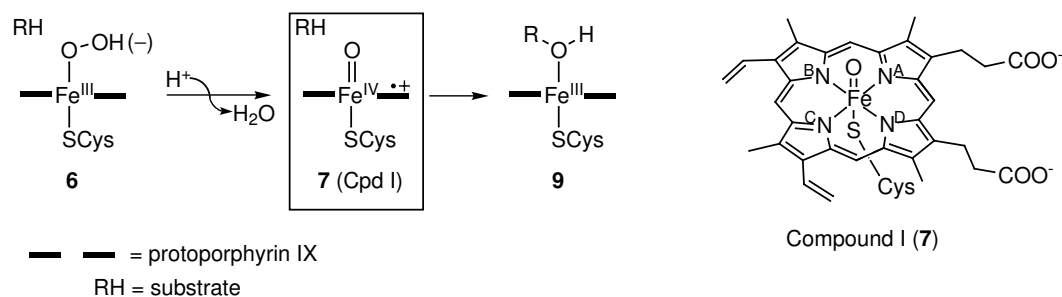


Figure 3.1: Key steps in the catalytic cycle of cytochrome P450. Cpd I (**7**) is drawn next to the reaction scheme.

Structural studies of P450 enzymes show that the sulfur atom of the proximal cysteinate in the native species interacts through three $\text{NH} \cdots \text{S}^-$ hydrogen bonds with Leu358, Gly359, and Gln360 (corresponding to the numbering system of P450_{cam}), and is exposed to the positive ends of dipoles in the protein pocket [12, 166, 167]. Mutation studies prove that the Fe–S bond is stabilized only if the protein pocket is sufficiently polar and supplies $\text{NH} \cdots \text{S}^-$ hydrogen bonding to the cysteinate sulfur [168], and that loss of a single $\text{NH} \cdots \text{S}^-$ hydrogen bond can have a significant effect on structure and function of the enzyme [169]. These features are therefore critical for a reliable representation of Cpd I. Previous theoretical studies that mimicked medium polarization and hydrogen bonding to the thiolate ligand in an approximate fashion indicate that such factors indeed strongly influence the computed properties of Cpd I in two model systems, one with porphyrin and HS^- and the other with octamethyl porphyrin and a cysteinate anion [34, 44]. Hence, the simplicity of the models previously used leaves doubts whether small models can indeed represent the real species. Clearly a definitive assignment of the electronic state and geometry of Cpd I, for P450, requires a realistic modeling of Cpd I within its protein environment.

To accomplish this task, we have carried out QM/MM calculations on Cpd I of the enzyme cytochrome P450_{cam}. In this computational approach the active species of the enzyme feels the electric field and hydrogen-bonding environment of the specific protein pocket, as well as the steric constraints exerted on the bonding of the heme to the cysteinate side chain. Such a calculation can on the one hand define a suitable truncated model, which ought to be used in future mechanistic studies. On the other hand, and more importantly, these calculations can form a useful guide for an eventual experimental characterization of Cpd I of Cytochrome P450_{cam}.

Since this study is aimed at a reliable representation of environmental effects on Cpd I, considerable efforts were made to realistically model the surrounding protein and solvent. The initial geometry of the enzyme was derived from X-ray data, which inherently provides only incomplete information, *e.g.*, positions of hydrogen atoms are not available, and crystal packing effects are present. Therefore a series of preparatory force field calculations was carried out with the aim to determine representative structures of the complete enzyme in aqueous solution. These geome-

tries then served as the starting points for QM/MM investigations. Details of these calculations are given in section 3.2.

In order to theoretically characterize Cpd I in the protein environment, geometry optimizations of an extended part of the enzyme around the active site were performed. Considering the high conformational complexity of this system, the present QM/MM calculations in principle should sample configurational space by MD or MC techniques. Due to the high accuracy required for the QM treatment this was not feasible in the present study. However, since the focus is on the electronic structure of Cpd I alone, geometry optimizations can give valuable insights into “static” properties, such as the influence of the polarizing and steric effect of the protein matrix. To assess the impact of different enzyme conformations on the electronic structure of the chromophore, we separately investigated four structures generated from an MD trajectory obtained in the preparatory force field calculations by QM/MM geometry optimizations.

A central aspect of this study is to investigate the role of the apoprotein in determining the electronic nature of the active iron-oxo-porphyrin species Cpd I. To this end, the results of the QM/MM computations for the full system are compared to corresponding results of the isolated QM systems *in vacuo*. This involves both single-point energy evaluations at the geometry of the QM system in the enzyme environment and complete geometry optimizations. The comparison of QM gas phase calculations with the QM/MM results for the complete system allow us to test the validity of the various model calculations used currently in the literature.

3.2 Preparation of the System

Initial Geometry

Initial geometry data was obtained from the X-ray structure with PDB code 1DZ9 (Brookhaven Protein Database) reported by Schlichting *et al.* and tentatively assigned to the Cpd I species of P450_{cam} [12]. Only one monomer from the asymmetric unit was used, corresponding to the protein chain A and solvent chain Z of the PDB structure. The first ten amino acid residues are unresolved and were omitted in all calculations, as well as a molecule of the TRIS buffer present in the structure. A potassium ion that occupies a special ion-binding site of the enzyme was included in the model.

Protonation State and Solvation

Hydrogen positions were constructed with a built-in procedure of CHARMM (*hbuild*) at pH 7. Because this approach cannot fully account for the variable environment of all the amino acid residues in the protein, a few corrections to the automatically determined protonation state had to be made. The according modifications were done following Lounnas and Wade [170] who assigned protons on the basis of Poisson-Boltzmann-calculations: (i) Glu366 was protonated at the carboxylate side chain. (ii) The basic residues His80, His270, His308, His353, and His355 were protonated at both nitrogens. These five doubly protonated histidines form salt-links to Asp77, Glu269, Glu73, Glu76 and one of the propionic side chains of the heme, which are considered important for the structure and function of the enzyme. All other histidines are protonated at either the δ - or ϵ -nitrogens upon visual inspection of their environment. The resulting total electric charge of the system is -10 e. The positions of the hydrogen atoms were optimized using CHARMM (2400 optimization steps for hydrogen atoms of the crystallographic water only, followed by 2400 optimization steps for all hydrogen atoms).

A water layer of 16\AA thickness was constructed around the enzyme using the InsightII software [171]. To adapt the water layer to the enzyme surface, the inner 8\AA of the solvent were equilibrated by MD (3 ps at 300 K) while keeping the remainder of the system fixed. Since the water density in the solvent was too low after the MD, more water molecules were added to the solvent layer and this procedure was repeated until the water density remained constant (3 times). The complete system generated in this manner consists of 24393 atoms, including 16956 atoms of solvent.

Preparatory Force Field Calculations and Snapshots

To relax the structure, a series of classical energy minimizations and MD runs were performed. These preparatory force field calculations employed the CHARMM22 force field [146] and were carried out with the CHARMM program [172], version 27a2. Details about non-standard force field parameters are given in appendix B. Throughout these calculations, the coordinates of the entire heme-unit and the coordinating Cys357 as well as the outer 8\AA of the solvent layer were kept fixed.

First, a geometry optimization was carried out to remove artificially close contacts between atoms. During the minimization, the atom positions of the enzyme were initially constrained with a harmonic potential using a force constant of $100\text{ kcal}/(\text{mol}\text{\AA}^2)$ for the backbone and $50\text{ kcal}/(\text{mol}\text{\AA}^2)$ for the side chains, respectively. These constraints were scaled down by a factor of 0.65 every 60 optimization steps. The final harmonic constraint force constants were $0.32\text{E-}05\text{ kcal}/(\text{mol}\text{\AA}^2)$ (backbone) and $1.64\text{E-}06\text{ kcal}/(\text{mol}\text{\AA}^2)$ (side chains). The final root mean square value of the cartesian gradient (GRMS) was $0.04\text{ kcal}/(\text{mol}\text{\AA})$.

Subsequently an MD run was performed, with 15 ps heating dynamics and 200 ps equilibration dynamics at 300 K. In the MD simulation, an integration step of 1 fs was used and bonds to hydrogen atoms were constrained by the SHAKE algorithm [173]. After initial velocity rescaling events during the simulation period from 0 – 20 ps, the total energy of the system was conserved.

In the X-ray structure [12], the camphor molecule is held in the protein pocket *via* a hydrogen bond from a tyrosine hydroxyl group (Tyr96) to its carbonyl oxygen. Experimental studies indicate that different substrates have considerable conformational flexibility in the pocket [174]. In accord with these findings, the present MD simulations show that camphor can rotate around this O-H \cdots H hydrogen bond. While a tumbling of the substrate in the pocket is thus not unexpected, the trajectory leads to conformations of camphor which are unproductive in terms of the hydroxylation reaction and are increasingly different from the X-ray data. It was therefore decided to select snapshot structures at the beginning of the time evolution of the MD as input for the QM/MM investigations, which are still close to the experimental structure. Three such structures were extracted from the equilibration trajectory after 29, 40, and 50 ps simulation time respectively. Each of the corresponding geometries was optimized by 4000 steps of ABNR minimization. These structures will be referred to as snapshots 29, 40 and 50.

The crystal structure features a hydrogen bond between the amide NH₂ group of the Gln360 side chain and the backbone carbonyl oxygen of the axial Cys357. Mutation studies have shown that this hydrogen bond has an impact on the redox potential of the heme in P450_{cam} [175]. This interaction is, however, lacking in the discussed snapshots. To probe if the electron-withdrawing effect of this particular hydrogen bond has an effect on QM/MM calculated properties, we have generated an additional structure by manipulating the conformation of the Gln360 side chain in snapshot 40, which has the conformation closest to the X-ray data, to form this hydrogen bond. This structure is denoted snapshot 40 manipulated (man) in the following. By analogy to the above-mentioned snapshots, the system was then also minimized with 4000 ABNR optimization steps.

The sensitivity of the optimized geometries with respect to the choice of the force field charges on the heme was tested [176] by a series of MM geometry optimizations where these charges were varied all the way up to the limit given by the oxidation state formalism (*i.e.*, $-2e$ on the porphyrin macrocycle, $-1e$ on the thiolate, etc.). The resulting optimized geometries exhibit minor differences, which indicates that the pocket around the heme is rather rigid and not significantly disturbed by different choices for the atomic charges on the heme atoms.

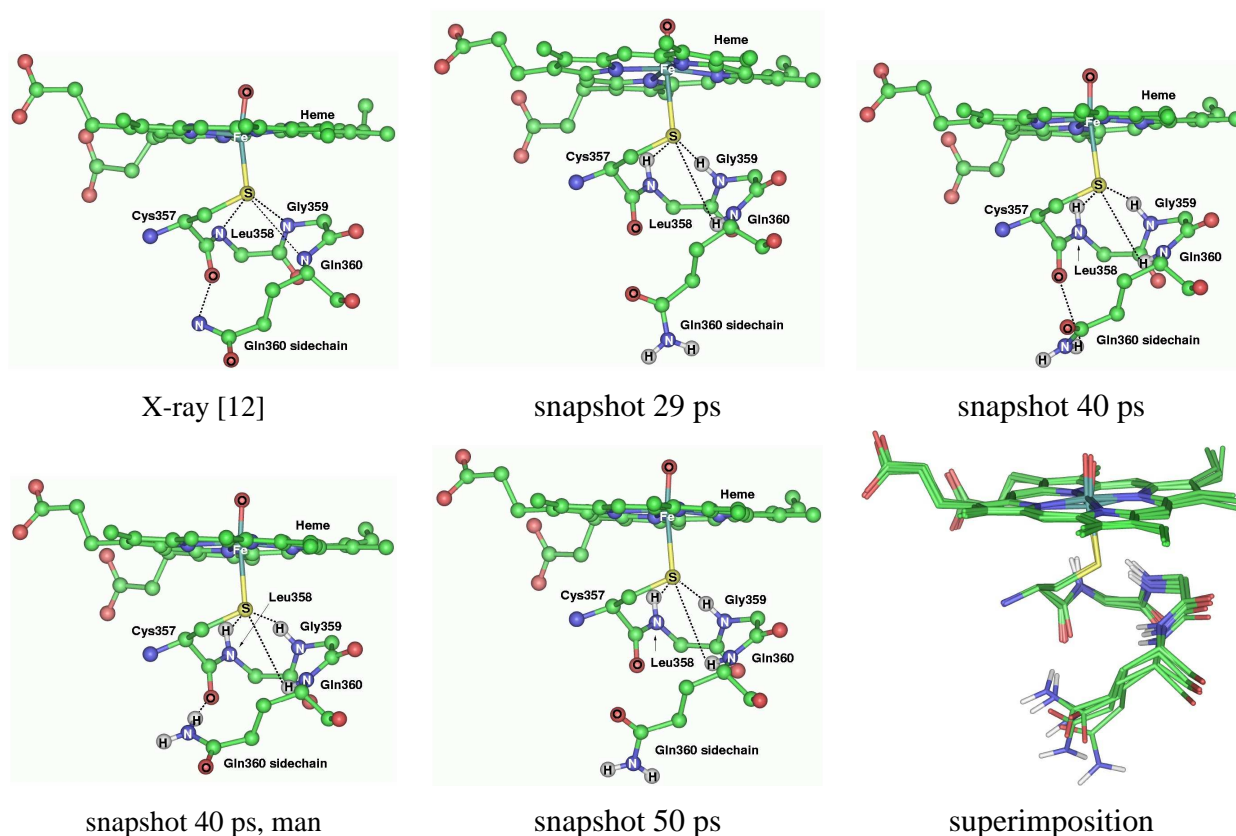


Figure 3.2: Hydrogen bonding situation of Cys357 in P450_{cam}: experimental X-ray structure and the calculated structures (snapshots 29, 40, 50 and snapshot 40 with added hydrogen bond between the side chain amide group of Gln360 and backbone carbonyl oxygen of Cys357). Only the hydrogen atoms of the backbone peptide NH groups and the side chain amide group are shown. The right picture in the second row shows a superimposition of all the structures. Color code: white, hydrogen; green, carbon; dark blue, nitrogen; red, oxygen; yellow, sulfur; light blue, iron.

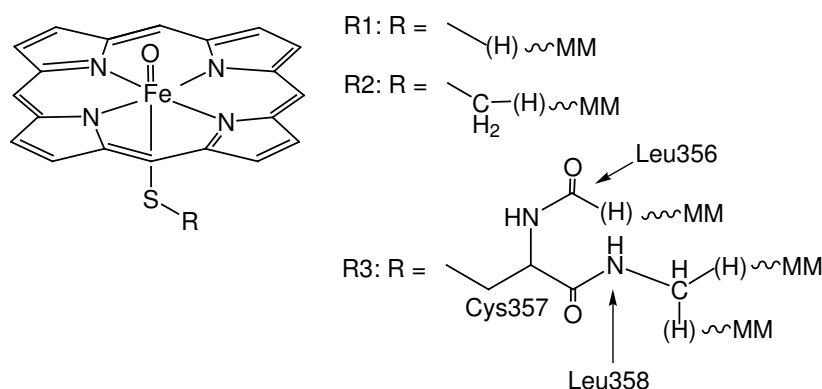


Figure 3.3: Choice of QM regions R1–R3.

3.3 Computational Methodology

3.3.1 QM Regions and Basis Sets

Three different QM regions were considered, all of which include the iron-oxo-porphyrin subunit (without side chains of the protoporphyrin IX) but differ in the extension to the axial cysteine. The smallest choice R1 comprises only the coordinating sulfur atom, replacing the bond of the sulfur to C_β of the cysteine with a bond to a hydrogen link atom at fixed distance. This choice might appear somewhat naive at first sight, considering that the QM–MM border is very close to the region of interest and does not cut through a carbon-carbon single bond as is usually recommended. However, a S–C bond is actually not significantly different in terms of polarity as compared to a S–H bond (Allred-Rochow electronegativity values are H: 2.2, C: 2.5, S: 2.4). Furthermore, the SH ligated Cpd I model has been important in gas phase model calculations, where it proved to give results in good agreement with calculations employing the full cysteine. This model thus appeared to be a natural starting point for the QM/MM investigations. The fixed bond distance S–H to the hydrogen link atom of 1.37988 Å was obtained from DFT optimized gas phase structures for this system.

The next bigger region R2 includes the C_βH₂ group of the cysteine ligand. The corresponding gas phase model [FeO(porphyrin)(SMe)] in some cases gave artificial results concerning the stability of sulfur-centered radicals due to the strong electron-donating character of the methyl group. Hence, one aspect of the QM/MM investigations was to test if the inclusion of explicit polarization by the enzyme environment could cure this effect.

The largest QM subsystem R3 incorporates the full Cys357, the CO group of Leu356, and the NH-C_α unit of Leu358 and can be expected to give results that are essentially devoid of QM–MM border effects. This model system was never investigated before by QM calculations. For QM regions R2 and R3, the bond distances C–H to the hydrogen link atoms are 1.090 Å, which is the default value implemented in the ChemShell QM/MM program.

Three basis sets were used; the iron atom is always assigned the LACVP small-core effective core potential and the associated basis set, which is of double- ζ quality [177]. Basis B1 employs a 6-31G gaussian-type basis set on all other atoms. Basis B2 is augmented on the six atoms directly coordinating to iron (four pyrrole nitrogens, the oxo atom and the sulfur of the axial cysteine) by a set of polarization functions (6-31G*), while basis B3 also includes a set of diffuse functions (6-31+G*) on these atoms [178–181].

3.3.2 Optimized Regions

The standard optimized region considered in the QM/MM calculations comprises 544 atoms and involves the following residues that are grouped around the active site:

Enzyme: Tyr75, Pro100, Thr101, Glu108, Val119, Phe163, Leu244, Leu245, Gly248, Gly249, Thr252, Val253, Phe256, Leu294, Val295, Asp297, Gly298, Arg299, Glu322, Thr349, Phe350, His355, Leu356, Cys357, Leu358, Gly359, Ala363, heme

Camphor

Water: Wat61B, Wat92B, Wat253B, Wat255B, Wat265B, Wat322B, Wat323B, Wat324B, Wat325B

Additional geometry optimizations employing smaller and larger optimized regions served to test the validity of this choice. The biggest optimized region considered in these studies consists of 1900 atoms and was defined by including all residues (including solvent) that have atoms within a distance of 9 Å around any atom of the heme or Cys357. The use of larger optimized regions had no significant effect on relative energies of spin states or the electronic structure of Cpd I as compared to those obtained for the standard region. However, with the larger optimized region, the conformational complexity related to the many degrees of freedom in some cases led to local minima, which had to be dismissed since they differed in conformations that are not related to Cpd I, but to the MM environment (*e.g.*, the hydrogen-bonding network around the A-propionate chain). To avoid corresponding artifacts, it was decided to restrict the optimizations to the smaller standard region.

3.3.3 QM/MM Scheme

QM Method

As discussed in section 1.2.3 Cpd I is a triradicaloid species, and non-dynamical correlation effects are relevant to the description of the doublet (low spin) state of this system. *Ab initio* QM

methods for treating non-dynamical correlation such as CASSCF are, however, computationally too demanding for the present purposes. Density functional theory in the unrestricted Kohn-Sham (UKS) formalism is computationally more economical and has proved to be successful in the description of Cpd I gas phase model compounds [35]. Hence, density functional theory was adopted as the QM method for the present calculations.

The low spin state of Cpd I is modeled by a broken-symmetry UKS determinant which exhibits significant spin contamination (typically $1.6 < \langle \hat{S}^2 \rangle < 1.8$). A correction scheme for the corresponding energy of such determinants has been suggested by Yamaguchi *et al.* [182] that involves computation of the energy of the next higher spin state. In the present case, the corrected and uncorrected energies, however, differ by 0.2 kcal/mol at most, because the LS doublet and quartet state are virtually degenerate. This is due to the weak interactions between the $S = 1$ system on the FeO subunit and the $S = 1/2$ spin on the ligands, as will be further discussed below. In view of the limited significance of such small corrections, plain UKS energies for the LS state are reported.

The DFT calculations in the present section employ the gradient-corrected exchange expressions B88 [183] and B3 [184] due to Becke in combination with the LYP correlation density functional [185].

QM-MM coupling

The QM-MM van-der-Waals interactions were represented by standard CHARMM22 force field terms, including the parameters for the heme group. The electrostatic QM-MM interactions were treated according to the electronic embedding scheme [127], *i.e.* the fixed MM charges were included into the one-electron Hamiltonian of the QM calculation, and the QM/MM electrostatic interactions were evaluated from the QM electrostatic potential and the MM atomic charges (for details see section 2.2.1). To treat the QM-MM boundary, hydrogen link atoms in conjunction with the charge shift model are employed [163]. This approach was recently applied and validated against others in a theoretical calibration study on triosephosphate isomerase [186].

The cutoff for the inclusion of MM charges into the QM one-electron hamiltonian is a technical parameter in QM/MM calculations that requires some attention. The evaluation of the QM/MM gradient involves QM gradient contributions for every MM center that is included into the one-electron hamiltonian. As a consequence, the computational effort for the QM/MM gradient evaluation increases with the number of considered point charges. However, since electrostatic forces are long ranged it is desirable to include as many charges as possible into the QM treatment. Therefore, the issue of different QM-MM cutoffs was investigated in a series of QM/MM calculations which monitored the computational cost and variations in the computed properties as a

function of the cutoff distance. No cutoffs were introduced for pure nonbonding MM terms.

The QM programs employed in the QM/MM as well as in the pure QM calculations were TURBOMOLE [187, 188] and GAMESS-UK [189]. All QM/MM calculations were performed with the ChemShell package [190]. The CHARMM22 force field [146] was used with the DL-POLY [191] program to treat the molecular mechanics part of the system. Geometry minimizations were done with built-in optimizers of ChemShell: gas phase geometries of model compounds R1 and R2 were optimized with NEWOPT [190] while all other minimizations employed the HDLC optimizer [192], applying the default convergence criteria (see appendix A.2).

3.3.4 Performance Considerations

QM performance

Table 3.1: CPU times t_{CPU} [min] for one QM/MM energy and gradient evaluation^a using different DFT implementations and functionals for the QM calculation, determined on an ALPHA EV67, 667 MHz CPU.

program	functional	t_{CPU}
GAMESS-UK	B3LYP	138
TURBOMOLE	B3LYP	68
TURBOMOLE	RI-BLYP ^b	50

^a QM region R1, Basis B1 (286 basis functions), cutoff for QM–MM electrostatic interactions 50 Å. ^b Calculation employed the resolution of the identity (RI) approximation.

For QM calculations we tested TURBOMOLE and GAMESS-UK. In QM/MM geometry optimizations (where in our case the QM calculations take more than 95% of the computer time), TURBOMOLE turned out to be faster typically by a factor of two and was therefore employed for most of the calculations. Results of corresponding benchmark calculations are listed in table 3.1.

The use of the resolution of the identity (RI) approximation allows for an efficient Coulomb integral evaluation and further speeds up the QM calculations. However, the use of RI is discarded in the present study for two reasons: (i) As can be seen from a comparison between RI-BLYP and B3LYP (table 3.1), the speed-up resulting from the use of RI is only moderate, presumably because of a suboptimal choice of the auxiliary basis set ¹. (ii) At present, RI cannot be used in

¹ Since no specifically tailored auxiliary basis sets are available for the basis sets used in the present calculations, a rather large auxiliary basis set (corresponding to basis TZVP) was employed to preclude concerns about the accuracy of the Coulomb fit.

conjunction with the hybrid B3LYP density functional, which has been shown to be superior to GGA functionals in the description of the reaction $\text{Fe}=\text{O}^+ + \text{H}_2 \rightarrow \text{Fe}^+ + \text{H}_2\text{O}$ [193], a simple model of P450 mediated hydroxylation of C–H bonds. It is thus desirable to use B3LYP for the present investigations. This functional has also been employed in most of the previous gas phase model studies on Cpd I.

QM–MM Cutoff for Electrostatic Interactions

The effect of the QM–MM cutoff for electrostatic interactions (*i.e.* how many point charges are included into the one-electron Hamiltonian of the QM calculation) was investigated by means of TURBOMOLE/CHARMM calculations with cutoffs of 55 a.u. (*ca.* 14500 point charges), 66 a.u. (*ca.* 20500 point charges) and 99 a.u. (all point charges, *ca.* 24300). The results in table 3.2 show small effects of the cutoff on the relative and absolute QM energies. Spin densities on other atomic positions (not shown in the table) are not affected significantly.

Table 3.2: Influence of different cutoffs for the QM–MM electrostatic interactions on important QM/MM calculated properties, given for the ^4A (^2A) states^a.

property	QM–MM cutoff / no. of point charges		
	55 a.u. / ≈ 14500	66 a.u. / ≈ 20500	99 a.u. / ≈ 24300 (all)
$R_{\text{Fe-S}}$	2.573 (2.586)	2.573 (2.592)	2.571 (2.592)
$\rho(\text{S})^b$	0.252 (-0.293)	0.252 (-0.294)	0.252 (-0.296)
$\rho_{\text{gr}}(\text{porph})^c$	0.734 (-0.837)	0.734 (-0.837)	0.734 (-0.835)
$\Delta E^{\text{QM}}(^4\text{A}_{2\text{u}} - ^2\text{A}_{2\text{u}})$	-0.50 kcal/mol	-0.37 kcal/mol	-0.25 kcal/mol

^a QM region R1, basis B1, B3LYP for the QM calculation. ^b Unpaired spin density on sulfur atom. ^c Unpaired group spin density on porphyrin macrocycle.

On the other hand, the DFT energy/gradient implementation in TURBOMOLE can treat large numbers of external point charges very efficiently, hence the use of larger cutoffs does not increase the computational effort significantly. For example, for QM model R1 and basis B1 (B3LYP, 286 basis functions), typical CPU times for one QM/MM energy and gradient evaluation with TURBOMOLE on a single node range from 62 to 68 minutes for a cutoff of 55 a.u. (≈ 14500 point charges) and from 75 to 85 minutes if all point charges (≈ 24300) are taken into account. Therefore, all data reported in the following sections were obtained using a QM–MM cutoff of 99 a.u. in QM/MM geometry optimizations, *i.e.* all point charges were included in the QM treatment.

3.4 Results

Section 3.4.1 presents a QM/MM calibration study that probed the effect of size of the QM region (R1–R3) and basis set (B1–B3) on the QM/MM computed properties. Snapshot 29 was chosen for this initial study.

In section 3.4.2, we present a comparison of QM/MM results obtained for the three snapshots (29, 40, 50) and the manipulated structure (40 man) to assess the influence of different enzyme conformations on the electronic properties of the active species. The latter calculations employed the R3 region and B3 basis exclusively.

3.4.1 QM/MM Calibration Study (Snapshot 29)

Much as in previous gas-phase model calculations, there are two close-lying spin states of Cpd I, the high-spin (quartet) and the low-spin (doublet) state, that correspond to ferromagnetic and antiferromagnetic coupling of electrons in the triradicaloid species. For the time being, these states are labeled as 4A and 2A . Their precise identity will be assigned at the end of this section.

Geometries

Table 3.3 on page 49 summarizes the optimized key geometric parameters around the iron center of the Cpd I species in the 4A and 2A states, for the different models of the QM region (R1 - R3) and different basis sets (B1 - B3). For comparison we show also the results for the isolated gas-phase molecules. The species has C_1 symmetry with four different Fe–N distances, which exhibit small differences of the order of 0.01-0.02 Å. These bond lengths show little dependence on the choice of R, but are slightly expanded (0.01-0.03 Å) in the enzyme pocket vs. the gas phase. The calculated Fe–O bond distances are almost identical for different R but are slightly (\approx 0.01 Å) shorter in the isolated molecule than in the enzyme environment. The $^4A/^2A$ pairs have virtually the same Fe–O and Fe–N bond distances, due to the fact that the different spin states arise from the rather weak coupling of the FeO electrons to the third unpaired electron. The use of a polarized basis causes a contraction of the Fe–O bond length (by 0.02-0.03 Å) and a slight elongation of the Fe–N distances (by less than 0.01 Å), whereas diffuse functions have no significant effect on these geometrical parameters. In fact, all calculations done so far find these bonds to be least sensitive to the choice of model or basis set.

Table 3.3: Computed bond distances r [Å] around the iron center of Cpd I (in the ${}^4A/{}^2A$ states) and angle $\theta_{\text{Fe-S-X}}$ [deg] for different choices of the QM region (R1–R3) and basis sets (B1–B3)^a.

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

^a R1–R3 refer to the QM regions defined in figure 3.3. Atoms $\text{N}^{\text{A-D}}$ are defined in figure 3.1. ^b B3LYP/CHARMM QM/MM geometry optimization. ^c B3LYP QM optimization of the isolated QM systems.

The computed Fe–S bond lengths in the isolated molecule exhibit strong variations of more than 0.1 Å. For any one of the basis sets, model R1 gives the shortest (*e.g.*, 2.587/2.605 Å), and R3 the longest (*e.g.*, 2.712/2.722 Å) Fe–S bond. In the protein environment, these variations are much less pronounced. Shifting the QM/MM border away from the sulfur atom (*i.e.*, moving from R1 to R3) causes an elongation of the Fe–S bond by 0.02–0.03 Å, while increasing the basis set slightly shortens the bond length, such that the calculated bond lengths vary from a minimum of 2.560/2.589 Å (⁴A/²A) to a maximum of 2.602/2.629 Å in the QM/MM computations. The Fe–S bond within the protein pocket is significantly shorter than in the naked system, by as much as 0.1 Å in the case of the sulfur ligand R3 in combination with the basis B1. These results agree well with the previous conclusions of Ogliaro *et al.* [34, 44, 45], and highlight the stabilizing impact of the hydrogen bonding and electric field of the protein on the Fe–S bond, as deduced from experiment [169, 175, 194]. Generally, the Fe–S bond is computed to be 0.02–0.03 Å longer in the ²A than in the ⁴A state [34, 35, 44, 45, 73]. The angle $\phi_{\text{Fe-S-X}}$ (X being hydrogen or carbon, depending on the choice of the QM part) is almost the same in all QM/MM models, owing to the fact that the geometry is largely determined by the enzyme environment. In the naked system, there is considerable conformational freedom and the angle strongly depends on the nature of the sulfur ligand, ranging from around 100° in the case of R1 to 115° in the case of R3. It is likely that the valence angle at the sulfur atom as well as the Fe–S distance influence the ability of the thiolato ligand to bind to Fe, and hence the electronic structure of Cpd I as a whole.

Relative Energies

To appreciate the influence of the protein environment on the QM system in terms of energetic stabilization, we compare in table 3.4 the energies of Cpd I (the QM system) in three situations defined by the subscripts p – protein and g – gas phase: (i) $S_{\text{p,p}}$ – the system at the geometry and in the electric field due to the protein environment; (ii) $S_{\text{p,g}}$ – the system at its protein geometry but in the gas phase, without the electric field of the protein; and (iii) $S_{\text{g,g}}$ – the naked system optimized in the gas phase. Taking as common reference point the total QM energy of $S_{\text{p,p}}$, the following energy differences are reported in the table:

$$\Delta E_{\text{vert}} = E(S_{\text{p,g}}) - E(S_{\text{p,p}}) \quad (3.1)$$

$$\Delta E_{\text{ad}} = E(S_{\text{g,g}}) - E(S_{\text{p,p}}) \quad (3.2)$$

Thus, ΔE_{vert} corresponds to the “vertical” stabilization energy of Cpd I provided by the protein electric field, at the geometry that the species adopts in the pocket. On the other hand, ΔE_{ad} represents the “adiabatic” stabilization energy, which includes the energy change ΔE_{relax} arising from geometry relaxation in the gas phase. Figure 3.4 shows the relations between these quantities.

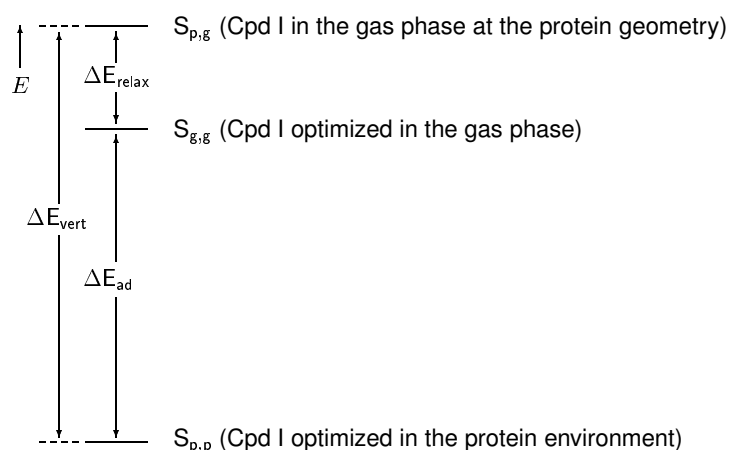


Figure 3.4: The QM stabilization energies of Cpd I in the protein environment (as defined in eqs. 3.1 and 3.2). The first subscript of each system (S) specifies the geometry, and the second one specifies the environment. ΔE_{relax} corresponds to the geometric relaxation of the gas phase species relative to the protein.

Table 3.4: Influence of the protein environment on QM energies of the Cpd I species (common reference point is the QM energy of Cpd I in the protein pocket ($S_{p,p}$ in eq. 3.1)) ^{a,b,c}.

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

^a R1–R3 refer to the QM model Cpd I systems in figure 3.3. ^b ΔE_{vert} (eq. 3.1, figure 3.4) is the vertical stabilization energy of the naked system in its protein geometry relative to the reference. ^c ΔE_{ad} (eq. 3.2, figure 3.4) is the adiabatic stabilization energy of the naked system in the gas phase geometry relative to the reference.

Table 3.5: Relative state energies $E(^4A) - E(^2A)$ [kcal/mol] for different models of Cpd I (defined by R1–R3) at the QM/MM and QM levels^a.

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

^a Energy differences as defined by eq. 3.3: a negative (positive) value implies a 4A (2A) ground state. The triple of entries corresponds to basis B1/B2/B3. ^b QM/MM energy difference. ^c QM contribution to the QM/MM energy difference. ^d Energy difference of the naked system at the protein geometry. ^e Energy difference of the naked system at the relaxed gas phase geometry.

As expected, the polarization by the electric field causes a substantial stabilization of Cpd I (approximately 109, 100 and 150 kcal/mol for Cpd I systems with substituents R1, R2 and R3, respectively), which is virtually identical for both spin states. Compared to this large electrostatic effect, the energy change due to the relaxation in the naked molecule is smaller but nevertheless substantial (9-10 kcal/mol for R1 and R2, 20 kcal/mol for R3), indicating that Cpd I is “strained” in the enzyme. It is clearly an advantage of combined QM/MM schemes that they are able to account for these effects of electrostatic stabilization and steric strain simultaneously. The larger relaxation of Cpd I model R3 is apparently due to the ability of this species to adopt a gas phase conformation with internal hydrogen bonding between the backbone NH group of Cys357 and the carbonyl oxygen of the same residue. This interaction is not feasible in the enzyme matrix. This concurs with the notion that the protein pocket is not a malleable cavity, but rather more a rigid matrix that provides a hydrogen-bonding template.

The energy differences between the quartet and doublet spin states

$$\Delta E = E(^4A) - E(^2A) \quad (3.3)$$

are listed in table 3.5 for different QM regions (R1 - R3) and basis sets (B1 - B3). The first entry gives the difference $\Delta E^{QM/MM}$ for the full systems, while the other three entries correspond to differences in the QM energies of the QM systems considered above ($S_{p,p}; S_{p,g}; S_{g,g}$). Generally speaking, all values in table 3.5 are rather small (typically less than 0.2 kcal/mol in absolute value) emphasizing again that the two spin states are close in energy regardless of the chosen computational approach. Enlargement of the basis set from B1 to B3 tends to favor the doublet state (entry 1) but there is apparently no clear trend when extending the QM region from R1 to R3 (entry 1) or when passing from the full system to the model systems considered (entries 2 - 4).

Our best calculations with the large basis B3 invariably predict that the ground state is 2A in the enzyme (entry 1), in accord with the experimental assignment of the related Cpd I of the CPO enzyme [46]. Qualitatively similar results are obtained for the relaxed gas phase species (entry 4), indicating that the apoprotein has no major influence on the relative energies of the two spin states. Using the data for model R3 and basis B3 in entry 1, the QM/MM energy difference between the two states is *ca.* 11 cm^{-1} which may be compared with the experimental value of 37 cm^{-1} for Cpd I in CPO [38,46]. The computed adiabatic energy differences in table 3.5 come from separate geometry optimizations of the two states, and their precision is therefore limited by the convergence criteria of these optimizations. To assess the intrinsic ordering of the two states, we have also computed the vertical energy differences at the optimized doublet geometry and obtain values of $-0.01/ +0.05/ +0.08\text{ kcal/mol}$ for basis B1/B2/B3 and model R3 (*i.e.* $-4/ +18/ +27\text{ cm}^{-1}$).

Spin densities

Table 3.6 summarizes the calculated net spin densities of the two states of Cpd I. For each Cpd I model (R1–R3) and basis set (B1–B3) three data are given which refer to the full system (QM/MM), the isolated Cpd I system at the geometry in the protein, and the relaxed gas phase species (QM). The FeO unit has a spin density of approximately 2.0, corresponding to two unpaired electrons in a triplet situation, as illustrated in figure 1.4 in section 1.2.3. This spin density is neither affected by the Cpd I model, nor the polarizing effect of the enzyme environment or the basis set. In contrast, the spin densities on the sulfur atom and the porphyrin moiety show a remarkable dependence on all the parameters. For the isolated gas phase compounds, the spin density on sulfur changes in the order $R2 > R3 > R1$, while the change on the porphyrin is precisely the opposite, $R1 > R3 > R2$. As discussed previously [45], the substituent effect on the sulfur density reflects the electron donating capability of the group R attached to sulfur, R1 (R2) being the worst (best) electron donor. Consequently, R1 (R2) underestimates (overestimates) the sulfur radical character with respect to the full cysteine model compound (R3). The spin density on the porphyrin macrocycle depends on the substituent in a complementary manner.

The greatest changes in these spin densities occur as Cpd I is transferred from the gas phase to the protein environment. When this happens, Cpd I is transformed from a dominant sulfur radical species (in all the models R1–R3) to a dominant porphyrin radical type (see table 3.6 and figure 3.5). While the R1 model still slightly underestimates the sulfur radical character in the enzyme environment, R2 and R3 give quite similar results, especially for the larger basis sets. The R2 spin density on sulfur is usually even slightly lower than for R3, which gives rise to an order $R1 < R2 < R3$. Thus, the polarizing effect of the enzyme environment cancels the artificially strong electron pushing effect of the methyl substituent in R2 that is observed in the gas phase

Table 3.6: Computed net spin densities of the 4A and 2A states for models R1–R3 and basis sets B1–B3 ^a

		4A state		
	basis	R1 (SH)	R2 (SMe)	R3 (Cys)
FeO	B1	2.022/ 2.023/ 2.025	2.019/ 2.024/ 2.023	2.018/ 2.026/ 2.007
	B2	2.021/ - / -	2.018/ - / -	2.018/ - / -
	B3	2.041/ 2.047/ 2.037	2.042/ 2.049/ 2.040	2.046/ 2.063/ 2.026
S _{Cys357}	B1	0.252/ 0.467/ 0.566	0.323/ 0.664/ 0.742	0.331/ 0.574/ 0.586
	B2	0.236/ - / -	0.312/ - / -	0.313/ - / -
	B3	0.201/ 0.408/ 0.526	0.282/ 0.645/ 0.719	0.283/ 0.539/ 0.535
Porph	B1	0.734/ 0.528/ 0.431	0.658/ 0.313/ 0.239	0.652/ 0.399/ 0.409
	B2	0.749/ - / -	0.666/ - / -	0.665/ - / -
	B3	0.763/ 0.557/ 0.452	0.671/ 0.296/ 0.229	0.667/ 0.385/ 0.429
		2A state		
		R1 (SH)	R2 (SMe)	R3 (Cys)
FeO	B1	2.121/ 2.115/ 2.094	2.119/ 2.104/ 2.091	2.122/ 2.108/ 2.101
	B2	2.119/ - / -	2.116/ - / -	2.120/ - / -
	B3	2.177/ 2.154/ 2.146	2.172/ 2.145/ 2.141	2.173/ 2.139/ 2.149
S _{cys357}	B1	-0.296/ -0.516/ -0.618	-0.371/ -0.730/ -0.782	-0.381/ -0.641/ -0.627
	B2	-0.285/ - / -	-0.353/ - / -	-0.360/ - / -
	B3	-0.270/ -0.477/ -0.600	-0.349/ -0.728/ -0.777	-0.352/ -0.620/ -0.593
Porph	B1	-0.835/ -0.618/ -0.500	-0.750/ -0.377/ -0.312	-0.744/ -0.466/ -0.475
	B2	-0.841/ - / -	-0.759/ - / -	-0.756/ - / -
	B3	-0.915/ -0.691/ -0.563	-0.820/ -0.407/ -0.350	-0.818/ -0.506/ -0.545

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

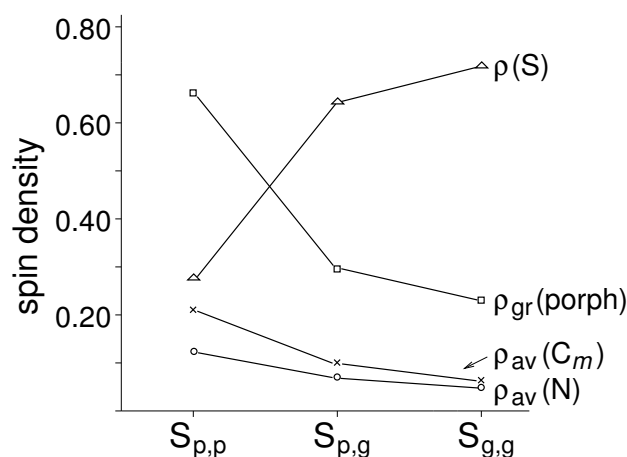


Figure 3.5: Calculated spin density (QM region R2, basis B3, 4A state) for the $S_{p,p}$, $S_{p,g}$, and $S_{g,g}$ case: individual spin density on sulfur $\rho(S)$, group spin density on porphyrin $\rho_{gr}(\text{porph})$, averaged spin densities on *meso*-carbon atoms $\rho_{av}(C_m)$, and averaged spin densities on pyrrole nitrogens $\rho_{av}(N)$.

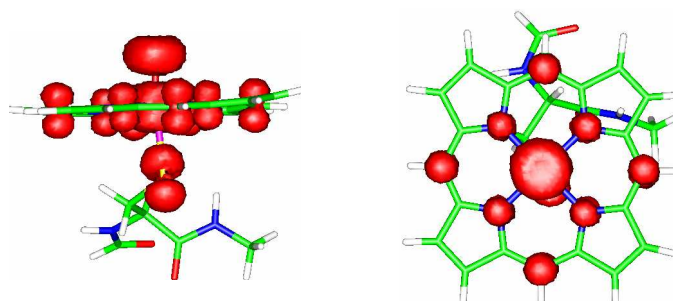


Figure 3.6: Plot of the α unpaired spin density (in red) from QM/MM calculations (B3LYP/CHARMM, QM region R3, basis B3) in the 4A state. Color code: white, hydrogen; green, carbon; dark blue, nitrogen; red, oxygen; yellow, sulfur.

calculations. These results on one hand demonstrate the ability of QM/MM methods to converge a calculated property by increasing the size of the QM region stepwise. More important, however, is that the pitfall of overstabilization of the sulfur radical state due to the lack of a polarizing medium (as in isolated molecule calculations) is avoided. These results are consistent with small model calculations using added NH_3 molecules to mimic $NH-S$ bonding in the protein pocket, and a continuum solvation model to mimic the polarization by the electric field [34,44].

Table 3.7 lists the spin density distribution on the various positions of the porphyrin. An analogous graphic representation for QM region R2 is presented in figure 3.5. It is obvious that as Cpd I is moved from the gas phase to the protein environment, the spin densities increase by a factor of two or even more on the *meso* carbon (C_m) and significantly also on the pyrrole nitrogen atoms. This indicates that the porphyrin spin density is located in an orbital of a_{2u} character (as drawn in figure

Table 3.7: Spin density distribution within the porphyrin ring for models R1 - R3 and basis sets B1- B3 ^a.

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

^a The triplet of numbers in each field refers to the QM/MM-optimized structure / the naked system in its protein geometry / the naked system in its gas phase geometry. See figure 3.7 for the definition of the relevant atoms.

1.4, section 1.2.3). The plot of the unpaired spin density, shown in figure 3.6, clearly confirms that the unpaired spin density is nascent from single occupation of the $\pi^*(\text{FeO})$ and $a_{2u}(\text{porph})$ orbitals.

3.4.2 Comparison of Different Snapshots

The influence of different conformations of the enzyme environment on the calculated electronic properties of Cpd I can be assessed by comparing the results of QM/MM geometry optimizations that start from different snapshots of the initial MD trajectory (see figure 3.2). We have selected three such snapshots (after 29, 40 and 50 ps) and generated an additional structure from the second snapshot by manipulating the side chain of Gln360 such that it forms an H-bond between its NH(amide) group and the carbonyl oxygen of the backbone of Cys357, in order to quantify the effect of this particular H-bond. Table 3.8 presents the most relevant H-bond distances in the QM/MM optimized structures (QM region R3, basis set B3, 2A state) and compares them with the available X-ray data [12] for the heavy atom distances N–X and with estimated distances H–X (H-positions optimized at the CHARMM force field level while keeping the heavy atoms fixed at the X-ray positions). In the case of Leu358 and Gly359, the four calculated N–S distances between the backbone peptide groups and the cysteinate sulfur are quite similar (variations within 0.04 Å and 0.08 Å, respectively), and close to the X-ray values. Hence, the backbone conformation of these two residues seems to be rather rigid. In contrast, the corresponding backbone peptide group of Gln360 exhibits more conformational flexibility since the computed N–S distances to the cysteinate sulfur change by up to 0.22 Å between different snapshots: the H-bond geometry is rather unfavorable in snapshots 29 and 50, and somewhat better in snapshot 40, although the N–S distance still exceeds the X-ray value by 0.26 Å. The side chain of Gln360 clearly does not form an H-bond with the carbonyl oxygen of Cys357 in snapshot 29 and 50, and its orientation is also unfavorable for this purpose in snapshot 40. This H-bond is present in the manipulated structure.

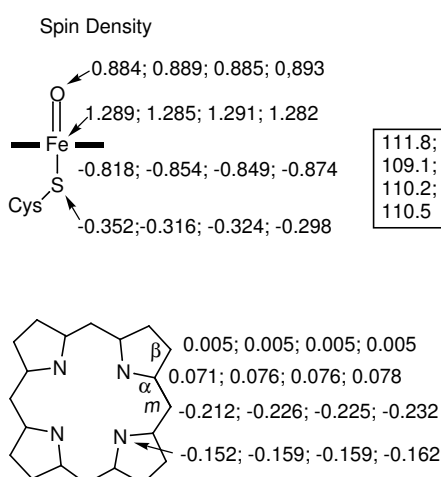
Figure 3.7 shows other selected geometrical parameters and spin densities for the QM/MM optimized structures derived from the different snapshots. The Fe–O and Fe–N distances show only negligible variations, in accordance with their very small sensitivity to the choice of QM region and basis set (for snapshot 29, see above). The Fe–S bond again proves to be more flexible, it ranges from 2.585/2.609 ($^4A/2A$) to 2.547/2.569 Å. The shortest bond is found for snapshot 40 with the manipulated side chain, the presence of this H-bond causes a slight shortening of 0.008 Å with respect to snapshot 40. In line with these findings, the calculated spin densities on the FeO moiety are little affected by the different conformations, whereas those on porphyrin and sulfur show some variation. Thus, the ratio of porphyrin:sulfur unpaired spin density in the quartet changes from 0.67:0.28 in snapshot 29 to 0.72:0.23 in snapshot 40 with the manipulated side chain. A plot of this ratio for the different snapshots is shown in figure 3.8. This shift of spin

Table 3.8: Hydrogen bond distances [Å], for N–X [H–X], where X=S, except for the last column with X=O. The first row gives the X-ray data (ref. [12]) with H-atoms added and optimized at the force field level. The other rows present distances from QM/MM optimizations (model R3, basis B3, 2A state).

	backbone N–H \cdots S(Cys357)			side chain Gln360 amide
	Leu358	Gly359	Gln360	N–H \cdots O=C(Cys357)
X-ray(1DZ9)+H	3.51 [3.58]	3.23 [2.51]	3.31 [3.14]	3.00 [2.02]
snapshot 29	3.46 [3.35]	3.32 [2.42]	3.78 [3.81]	4.97 [5.44]
snapshot 40	3.45 [3.34]	3.40 [2.68]	3.57 [3.47]	3.15 [2.86]
snapshot 40 ^a	3.49 [3.48]	3.33 [2.57]	3.56 [3.43]	2.77 [1.77]
snapshot 50	3.48 [3.43]	3.35 [2.48]	3.78 [3.65]	4.22 [4.50]

^a Side chain of Gln360 manipulated to form H-bond with carbonyl oxygen of Cys357.

a) 2A_2u



b) 4A_2u

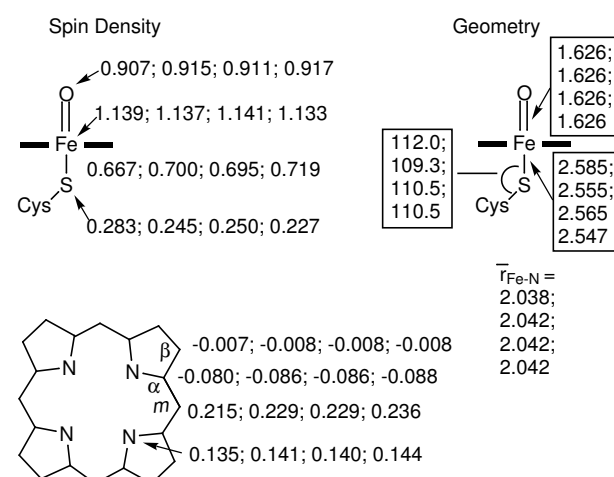


Figure 3.7: Cpd I of P450_{cam} (model R3, basis B3): QM/MM optimized geometries and spin densities for different snapshots. The given values correspond to snapshot 29; snapshot 40; snapshot 50; snapshot 40 with the manipulated Gln360 side chain conformation. a) LS state, b) HS state.

density between sulfur and the macrocycle nicely reflects the differential stabilization of negative charge on sulfur through the variations in the H-bond network. However, compared to the large effect on the spin density that is observed when Cpd I is transferred to the gas phase, these changes appear rather small, and one should stress that the qualitative results - namely that of a dominant porphyrin radical character - remain the same for all snapshots.

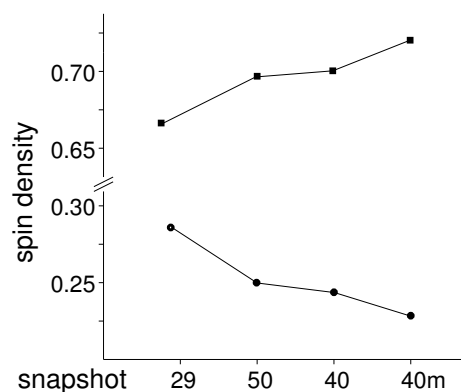


Figure 3.8: QM/MM calculated spin density (B3LYP/CHARMM, QM region R3, basis B3, 4A state) on the porphyrin macrocycle (squares) and on the sulfur atom (circles) for different snapshots.

3.5 Discussion

The present study reveals some key aspects related to the electronic and geometric features of Cpd I and how these features are influenced by the protein environment.

The geometric data in table 3.3 and energetic data (ΔE_{relax} , figure 3.4) in table 3.4 show that Cpd I inside the pocket is strained with respect to its relaxed gas phase geometry. This is especially apparent for the model R3 that uses an extended part of the cysteine loop as a proximal ligand. Thus, while in the gas phase, the small chain curls to form an internal $\text{NH} \cdots \text{O}=\text{C}$ hydrogen bond, within the protein pocket this interaction does not exist, and is replaced by another one from Gln360. These results, as well as other indications from the pure MM calculations, show that the protein pocket is not a malleable cavity, but a rather rigid matrix that provides a hydrogen-bonding template that sustains the Fe–S bond. This is further emphasized by the Fe–S bond shortening (table 3.3) and the increased electron density at sulfur (table 3.6) in the pocket with respect to the gas phase. These results accord well with experimental observations [168, 169, 175] that the hydrogen bonding machinery in the proximal pocket is essential for the stability of the heme-thiolate species.

The spin density data in table 3.6 and figure 3.5 show that the protein environment, by stabilizing charge density localization on the cysteinyl sulfur, transforms Cpd I from a sulfur-centered radical in the gas phase to a porphyrin centered radical cation. Our calculations thus support the experimental assignment of the characterized Cpd I species of CPO [47] and HRP [50, 195], which are porphyrin radical cations. Likewise, the recently probed Cpd I of CYP119 [30] exhibited UV/VIS spectral features that are consistent with a ferryl porphyrin radical cation.

The analysis of the QM/MM calculated spin density distribution on the porphyrin macrocycle, which are large on the nitrogen and C_m positions (see table 3.7 and figure 3.7) allows us to specify

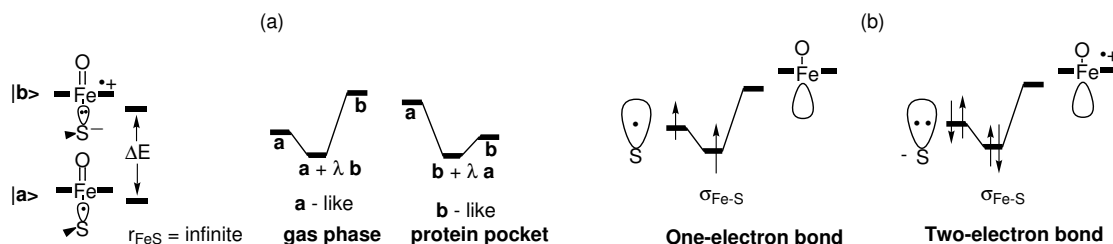


Figure 3.9: The nature of the Fe–S bond in Cpd I. (a) A VB mixing diagram. (b) A complementary orbital mixing diagram of the Fe–S bond orbital.

the ground state of Cpd I as A_{2u} , in agreement with EPR results for HRP [50]. The small amount of spin density on other positions could suggest a small admixture of A_{1u} character, yet this cannot be quantified within the present UKS approach. Consistent with experimental results for Cpd I of HRP [48] and CPO [46], the present calculations predict the interaction between the ferryl $S = 1$ and the porphyrin $S' = 1/2$ spin to be on the order of several wave numbers. A definite conclusion regarding the ground state multiplicity is thus difficult. Our best calculations, however, favour the antiferromagnetic doublet state, consistent with the experimental results obtained for Cpd I in these related enzymes [46, 48]. A more thorough DFT study of this energy difference, employing single point calculations with larger basis sets, confirms this assignment. These results are presented in chapter 6.

Interestingly, the spin density distribution suggests that the simplest model with the QM region R1 gives results in terms of electronic structure that are similar to those obtained with the largest QM region R3 (figure 3.7). This model may thus be used for more demanding calculations on reaction mechanisms.

The interplay between the hydrogen-bonding situation around the thiolate ligand and the resulting changes in the spin density distribution between the sulfur *vs.* the porphyrin is an important feature (see figure 3.7) of Cpd I. As noted in section 1.2.3, the spin density distribution between sulfur and porphyrin is commonly attributed to a “redox-mesomorphism” that can be conceptualized by a simple valence bond (VB) mixing diagram (figure 3.9).

In this description, the species is a resonance hybrid of two forms: $|a\rangle$ involves a thiyl radical and a closed-shell porphyrin, while $|b\rangle$ involves a thiolate anion and a porphyrin radical cation. At infinite Fe–S separation, $|a\rangle$ is considerably lower in energy. However, at the equilibrium geometry the ion pair structure $|b\rangle$ is stabilized and approaches $|a\rangle$ energetically. In the gas phase, $|a\rangle$ is still the lowest structure and the VB mixing leads to a Cpd I species which has a dominant thiyl radical character. As the species is placed in the protein pocket, the electric field and the $\text{NH} \cdots \text{S}$ hydrogen bonding stabilize preferentially the ion-pair form, $|b\rangle$, which becomes the lowest VB structure. The VB mixing then leads to a Cpd I species that has a dominant porphyrin radical-cation character as indeed revealed by the calculations.

The interactions with the protein also strengthen the Fe–S bond. This complementary information becomes more obvious from the orbital mixing diagram in figure 3.9b, which shows the alteration in the nature of the Fe–S bond as the thiol center changes from a radical to an anion. In each case the iron-oxo-porphyrin fragment has an empty hybrid orbital (a mixture of d_{z^2} , p_z , and s orbitals on iron) pointing to the missing coordination site. This hybrid mixes with the p_σ hybrid on sulfur to form the Fe–S bond orbital. In the limiting case of a sulfur radical situation, the bond orbital will contain a single electron, and the Fe–S bond will be a one-electron bond. However, when the sulfur is in an anionic situation, the bond orbital is doubly occupied and the Fe–S bond is a two-electron bond. Thus, the interactions in the protein pocket cause a transformation of the Fe–S linkage from a weak one-electron bond to a stronger two-electron bond. This qualitative change of the Fe–S bond is in line with the calculated shortening and strengthening of this bond [34, 44].

The VB model in figure 3.9a further predicts that the stronger the electric field and the hydrogen bonding, the more dominant will be the porphyrin radical-cation character. Figure 3.7 mirrors this trend beautifully: when going from snapshot 29, in which Gln360 does not form an $\text{NH} \cdots \text{S}$ hydrogen bond, via snapshot 40, in which this hydrogen bond is present to some extent, all the way to the manipulated structure in which Gln360 also forms an $\text{NH} \cdots \text{O}=\text{C}$ hydrogen bond to the cysteine, the corresponding porphyrin spin density in the doublet state changes from -0.818 to -0.874 with an opposite and concomitant change in the cysteinate spin density. This again shows the extraordinary ability of Cpd I to accommodate its electronic structure to the protein environment. These subtle changes predicted by the present calculations are in accord with recent mutation studies [169, 175] concerning the role of the residues in the proximal pocket of P450_{cam} . These investigations revealed small changes in the reduction potential of the heme in the Leu358Pro, Gln360Pro, and Gln360Leu mutants, which suggest that the corresponding hydrogen bonding network serves to stabilize the heme-thiolate coordination and to regulate the redox potential of the heme iron. One may thus expect that cysteinate-Cpd I species in different proteins (with different H-bond patterns) will have different properties. It is yet to be unraveled whether these differences are also expressed in the reactivity of different P450 isozymes and different thiolate enzymes.

3.6 Conclusions

The present study characterizes the geometric and electronic structure of the elusive active oxidant Cpd I of cytochrome P450. It highlights the advantage of performing QM/MM calculations for such a task. The calculations provide information on the species in its specific protein environment, in this case P450_{cam} , without the need to make assumptions about the nature and architecture of this environment. In this respect, QM/MM calculations come as close as possible to being a

faithful partner to structural and spectroscopic determinations by experiment. An added advantage is that these calculations offer important insight into the factors that govern the properties of Cpd I, through comparisons with various gas phase models and through inspection of different conformations within the protein pocket. We find that Cpd I is transformed by the protein environment from a sulfur-centered radical to the $^{4,2}A_{2u}$ porphyrin-centered radical cation (*ca.* 70%). The hydrogen bonding machinery of the protein pocket shortens the Fe–S bond and strengthens it. The best estimates of the Fe–S bond lengths in the $^2A_{2u}/^4A_{2u}$ states are 2.569 Å/ 2.547 Å. A series of calculations on different conformations of the protein pocket shows that the balance of the sulfur/porphyrin radical character and the precise Fe–S bond length are finely tuned by small changes in the architecture of the pocket. Cpd I behaves therefore as a chameleon species that adapts its electronic and structural character to the specific environment.

Chapter 4

Investigation of the Rebound Mechanism

4.1 Motivation and Background

While strong evidence has been found for the role of compound I (Cpd I) as the primary oxidant in P450 catalyzed hydroxylation [7, 13], there is no consensus regarding the mechanistic details of its reaction with hydrocarbons yielding alcohols. As discussed in section 1.2.4, experimental evidence such as partial scrambling of stereochemistry [54], allylic rearrangements [9], and the high intrinsic isotope effects [55], suggest a stepwise hydrogen abstraction–oxygen rebound pathway, depicted in figure 4.1. This mechanism implies the presence of a discrete radical intermediate **8** [1]. On the other hand, the results from radical clock experiments seem to exclude the presence of such an intermediate and have been interpreted in terms of an effectively concerted, non-synchronous oxygen–insertion reaction [63]. A mechanism that is capable to resolve the seemingly contradictory experimental findings has been suggested by Shaik and coworkers on the basis of theoretical considerations and model computations [64–68, 71]. This two-state reactivity (TSR) scenario invokes two spin states, the low spin (doublet) and the high spin (quartet) state. DFT calculations on model compounds indicate that the reaction involves hydrogen abstraction

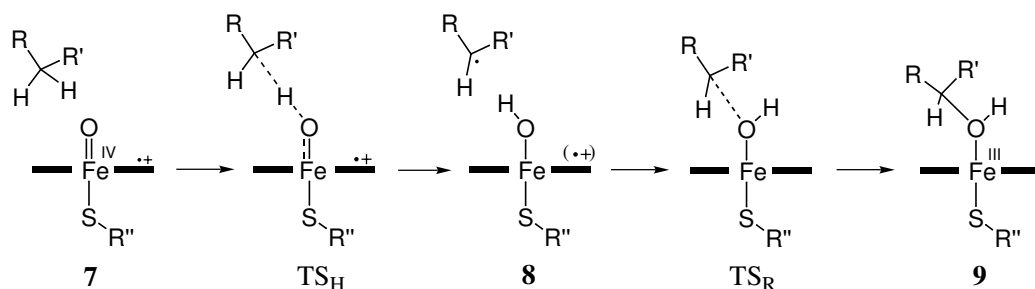


Figure 4.1: The rebound mechanism.

on both spin surfaces *via* essentially identical transition states (TS_H , figure 4.1). The reaction on the doublet surface then proceeds *via* a barrierless recombination to give the product alcohol complex **9** and thus is effectively concerted. In contrast, the quartet pathway involves a significant barrier for rebound and a genuine transition state (TS_R) and thus represents a true two-step reaction that gives rise to a discrete radical intermediate **8**. The TSR scenario for P450 mediated hydroxylation gained support from further DFT calculations on different gas phase model compounds [25, 73–78]. However, these studies cannot account for the specific steric and electrostatic interactions of the substrate with the protein pocket. For example, in P450_{cam}, the substrate camphor is positionally restrained by a hydrogen bond from Tyr96 to its carbonyl oxygen. In contrast, previous investigations in the gas phase assumed that the substrate is mobile and can attack the oxygen in the rebound phase from any conceivable conformation. Furthermore, it was demonstrated that environmental effects, such as hydrogen bonds to the thiolate sulfur and medium polarization, have a distinct impact on computed activation barriers in model calculations [36]. To test the TSR hypothesis, a QM/MM study is required that can realistically account for these effects. In this chapter, the results of combined QM/MM calculations for the complete enzyme, as well as from pure QM calculations on the isolated QM subsystems in the gas phase are reported for the hydroxylation pathway as proposed in the TSR rebound mechanism. These calculations characterize in detail the reaction pathway in the specific protein environment. By calculating the energy profile for both, the LS and the HS reaction, we show that the hydrogen abstraction–oxygen rebound pathway in two spin states (TSR) is in fact operational in P450_{cam} mediated C–H hydroxylation. The comparison to the analogous calculations in the gas phase allows us to interpret factors that are considered to be relevant to the role of the protein environment in enzyme catalysis.

4.2 Computational Methodology

4.2.1 QM Regions and Basis Sets

For QM/MM geometry optimizations, two different QM regions, R1pro and R3cam, were employed in the present study (see figure 4.2). These comprised the following atoms:

R1pro (51 QM atoms): Iron-oxo-porphyrin (without sidechains of the heme), sulfur of Cys357, C^4H , C^5H_2 and C^6H_2 of camphor. In the gas phase calculations, this corresponds to $[\text{FeO}(\text{SH})(\text{porph})] + \text{propane}$.

R3cam (84 QM atoms): Iron-oxo-porphyrin (without sidechains of the heme), Cys357, CO group of Leu346 and $\text{NH-C}_\alpha\text{H}$ unit of Leu348, full camphor.

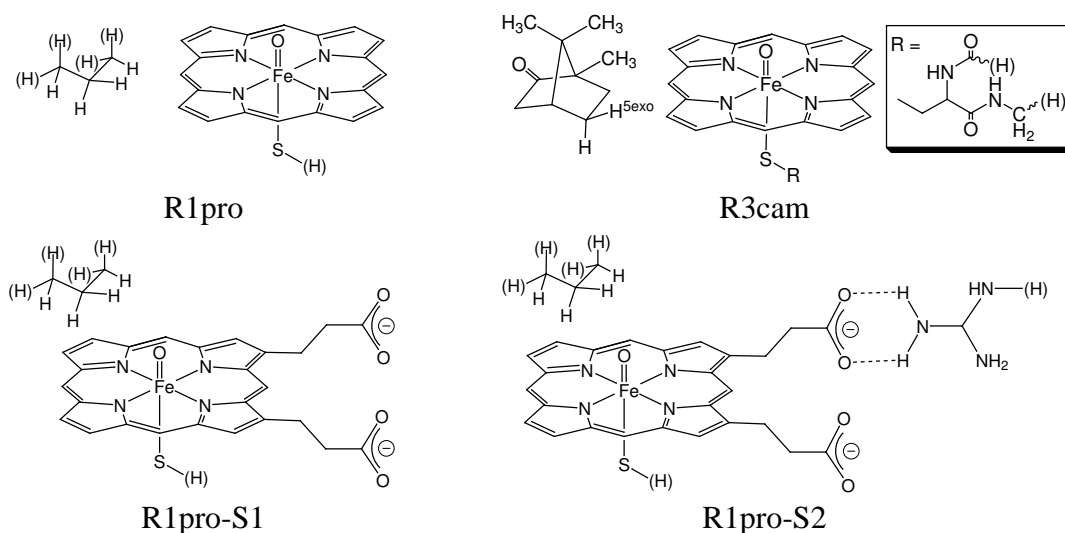


Figure 4.2: Choice of QM regions R1pro, R3cam, R1pro-S1, and R1pro-S2.

Additionally, single point energy calculations were carried out with two QM regions that are derived from R1pro but extend onto the propionate side chains of the heme, and the side chain of the neighbouring Arg299 group. These QM regions are denoted R1pro-S1 and R1pro-S2 and are also depicted in figure 4.2.

The basis sets used in the present investigations, are abbreviated as B1 and B4. Basis B1 is defined in analogy to section 3.3. Basis B4 features the LACVP small-core effective core potential and the associated basis set for iron [177]. The other atoms are assigned the 6-31G gaussian-type basis set, augmented with a set of polarization and diffuse functions ($6-31^+G^*$) on the six atoms coordinated to iron, the C^5 atom ($6-31^+G^*$) and also on the H_{exo}^5 hydrogen atom ($6-31^{++}G^{**}$) of camphor [178–181].

4.2.2 Optimized Region

The QM/MM optimizations included the following set of residues of the active site.

Enzyme: Pro86, Phe87, Tyr96, Pro100, Thr101, Leu244, Gly248, Asp251, Thr252, Val295, Asp297, Arg299, Phe350, His355, Leu356, Cys357, Leu358, Gly359, Gln360, Ile395, Val396, heme

Camphor

Water: Wat61, Wat253, Wat325

This corresponds to a total of 442 optimized atoms.



Figure 4.3: The simplified model used in the QM calibration study.

4.2.3 QM/MM Scheme

The QM/MM coupling scheme adopted in the present study was presented in detail in section 3.3. The QM part was treated with unrestricted Kohn-Sham DFT and the B3LYP density functional [184, 185]. The CHARMM22 force field¹ [146] was employed for the MM calculation. The electrostatic embedding scheme and link atoms were used for the QM–MM coupling. The calculations were done with ChemShell [190], TURBOMOLE [187, 188] and DL-POLY [191] for the QM/MM, QM and MM contributions, respectively.

All optimizations were carried out with the HDLCOPT optimizer [192] applying the default convergence criteria (see appendix A.2). The rational function optimizer with the Powell update (P-RFO) for an explicit hessian [192] was used in transition state optimizations. Unless otherwise indicated, the reaction core² was defined as the FeO subunit and C⁵ and H_{exo}⁵ of camphor, and thus contained 4 atoms.

4.2.4 QM Calibration Study

To probe the intrinsic accuracy of the DFT approach in describing the H-abstraction barrier, we have carried out a calibration study employing different density functionals and basis sets. To this end, we have chosen a simplified molecular model $[\text{FeO}(\text{SH})(\text{NH}_3)_4]^+ + \text{CH}_4$ for the reaction, shown in figure 4.3. This model system includes all essential features of the hydrogen abstraction step, *i.e.*, the oxoferryl moiety, the thiolate ligand and the octahedral coordination geometry of the iron center with four nitrogen donors. The geometries of the model reactive complex **16** and the corresponding transition state **16-TS** were optimized with TURBOMOLE [187, 188] using the B3LYP density functional [184, 185] and the SVP basis [188]. A frequency calculation confirmed that the transition state structure has a single eigenmode with a negative eigenvalue. Single point calculations were done using the popular pure GGA functionals BP86 [183, 196] and BLYP [183, 185]. As alternatives to our standard choice B3LYP we also tested the B97 [197], and PBE0 [198] hybrid density functionals. The B97 functional due to Becke contains ten linear coefficients which

¹ Details about non-standard force field parameters are given in appendix B.

² An explicit hessian is calculated for this subset of atoms.

were optimized by a least-squares fit to energetical data from the G2 set. This functional contains an amount of 20% exact exchange. The PBE0 functional is a parameter-free hybrid functional, where the amount of 25% exact exchange was derived from purely theoretical arguments through a perturbation theory argument. These DFT calculations were carried out with MOLPRO [199]. Additional basis sets employed for the single point energy calculations were the SV, SVP and TZVP sets taken from the TURBOMOLE library [200, 201], the cc-pVDZ and cc-pVQZ basis sets by Dunning *et al.* [202, 203], and a Wachters all electron basis set for iron [204] with an additional diffuse d function [205] and a set of f polarization functions [206] in the contraction [8s7p4d1f] .

4.2.5 Strategy

The QM/MM study involved potential energy surface (PES) scans and full geometry optimizations of minima and saddle points along the reactive pathways. These were started from the Cpd I species in the enzyme environment as described in chapter 3. The initial geometry was obtained from optimizations with QM region R3 and basis B3 of snapshot 40. A second, full study of the reaction pathway including PES scans and optimizations was carried out in the Jerusalem group starting from the Cpd I geometry obtained for snapshot 29 [176]. The consistency of the results obtained in the two independent studies confirmed that the optimized conformations are representative for the reactive pathway under study.

The PES scans were essential to locate regions in the high-dimensional conformational space that were close enough to the stationary points for our geometry optimizations to converge within a reasonable time. The nature of the transition states as saddle points of first order (1 mode with a negative eigenvalue) was confirmed by numerical calculation of the finite-difference-hessian for the set of QM atoms (R1pro/B1).

The computationally demanding PES scans and frequency calculations were performed with the smaller QM region R1pro and the B1 basis set. The fully optimized stationary points on this level of theory were then reoptimized in the enzyme environment with the larger QM region R3cam and basis B4.

In the following, the QM/MM scheme employed in a particular calculation will be denoted by the expression QM region/basis set, *e.g.* R1pro/B1, or R3cam/B4.

To assess the influence of the protein environment on the QM subsystem, gas phase QM calculations employing QM region R1pro and basis B1 were performed (i) as single point energy calculations at the geometry of the system in the enzyme environment, and (ii), as full geometry optimizations of the isolated species in the gas phase. In analogy to chapter 3, the former will be

referred to as $S_{p,g}$ (the system in the gas phase at the protein geometry) and the latter as $S_{g,g}$ (the system optimized in the gas phase).

4.3 Results

4.3.1 Hydrogen Abstraction

Starting from snapshot 40, the PES relevant to the hydrogen abstraction was scanned at the R1pro/B1 level by the following procedure: the H_{exo}^5 atom of camphor was moved in steps of 0.15 Å towards the oxo-atom, while the internal coordinate corresponding to the distance H_{exo}^5-O was constrained and the environment was minimized. This was done separately on the 2A and 4A hypersurface. Figures 4.4a and 4.4b present plots of the energy profiles obtained from the respective scans. The plots show the total QM/MM energy, as well as the individual QM and MM contributions to the latter, as a function of the O–H distance. It is obvious that the QM/MM energy is clearly dominated by the QM contributions, which are due to the chemical bond-breaking and bond-making process in the reactive subsystem. Apparently, the scans yield slightly different conformations and energies at a given step along the O–H distance for the 2A and 4A hypersurface (see figures 4.4a and 4.4b). Since the present approach, however, neglects conformational sampling, the corresponding differences are not related to different energy profiles for the 2A and 4A reaction, but merely reflect the high-dimensionality of the PES. Meaningful with respect to activation energies are only the stationary points that were obtained from structures along the scan by releasing the O–H bond constraint. From these free geometry optimizations we obtained a reactive complex (RC), and the transition state for hydrogen abstraction (TS_H). An analysis of the geometries separately optimized for the 2A and 4A states shows that the structures in the different spin states are in fact very close and correspond to completely analogous conformations of the active site environment.

Table 4.1 summarizes the computed relative energies for the H-abstraction phase of the reaction. We compare the QM/MM computed energies for the full system $E^{QM/MM}(p, p)$, the QM energy contribution to the latter $E^{QM}(p, p)$, the QM energy of the isolated QM subsystem in the gas phase at the protein geometry $E^{QM}(p, g)$, and the QM energy of the isolated QM subsystem optimized in the gas phase $E^{QM}(g, g)$. Selected geometry data of the QM/MM and QM optimized structures are shown in table 4.2. The spin densities, which serve as sensitive reporter of the electronic structure, are listed in table 4.3.

The geometric and electronic structure of the reactive complex obtained in the QM/MM geometry optimizations is very close to the initial structure, snapshot 40. The main difference is a change

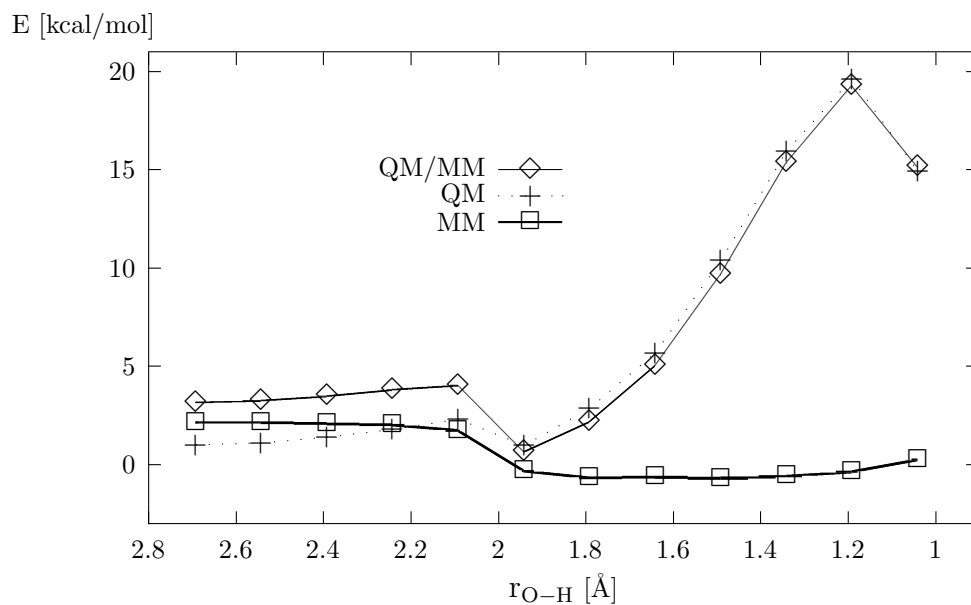
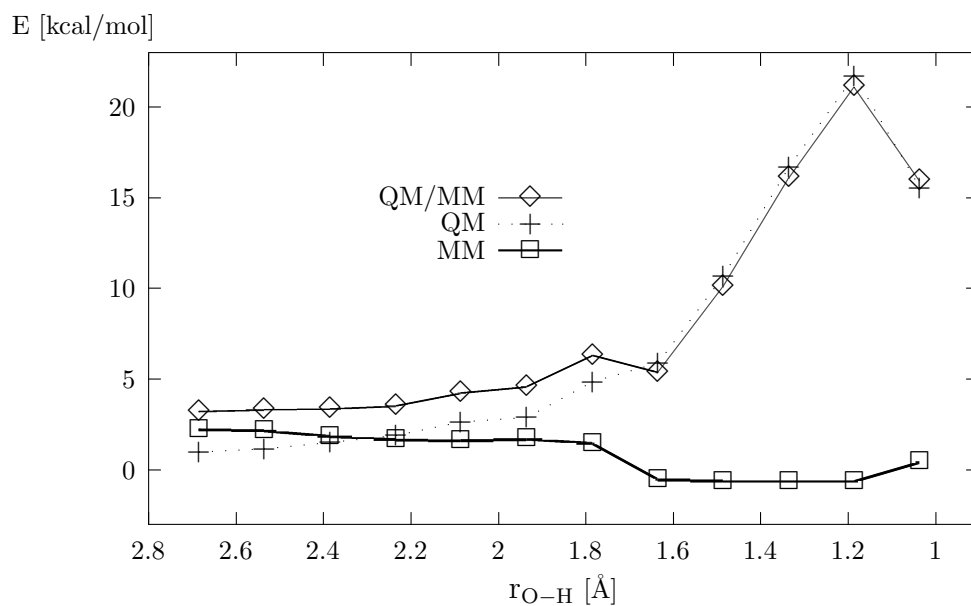
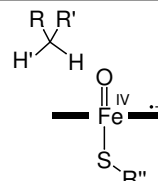
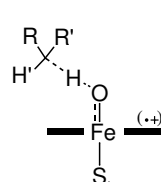
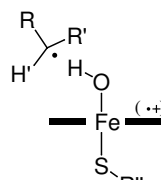
a) 2A stateb) 4A state

Figure 4.4: PES scan of the coordinate $r(O - H_{\text{exo}}^5)$, B3LYP/CHARMM, QM region R1pro, basis set B1.

Table 4.1: Relative energies [kcal/mol] of the stationary points along the H-abstraction reactive pathway (snap 40: QM/MM optimized structure of snapshot 40; RC: reactive complex **7**; TS_H: transition state for H-abstraction; hydroxo intermediate **8**). Reference point is the reactive complex on the ²A surface.

		QM region				
		R1pro/B1		R3cam/B4		
		LS	HS	LS	HS	
7 (snap 40)	$E^{\text{QM/MM}}(\text{p}, \text{p})^a$	3.2	3.1	1.6	1.6	
	$E^{\text{QM}}(\text{p}, \text{p})^b$	1.0	0.9	1.8	1.8	
	$E^{\text{QM}}(\text{p}, \text{g})^c$	- ^e	- ^e	-	-	
	$E^{\text{QM}}(\text{g}, \text{g})^d$	- ^e	- ^e	-	-	
7 (RC)	$E^{\text{QM/MM}}(\text{p}, \text{p})^a$	0	-0.1	0	0.0	
	$E^{\text{QM}}(\text{p}, \text{p})^b$	0	-0.1	0	0.1	
	$E^{\text{QM}}(\text{p}, \text{g})^c$	0	-0.3	-	-	
	$E^{\text{QM}}(\text{g}, \text{g})^d$	0	0.0	-	-	
TS _H	$E^{\text{QM/MM}}(\text{p}, \text{p})^a$	19.5	20.6	21.1	21.8	
	$E^{\text{QM}}(\text{p}, \text{p})^b$	19.8	20.8	22.3	22.9	
	$E^{\text{QM}}(\text{p}, \text{g})^c$	20.4	21.1	-	-	
	$E^{\text{QM}}(\text{g}, \text{g})^d$	19.5	19.5	-	-	
8	$E^{\text{QM/MM}}(\text{p}, \text{p})^a$	14.0	14.9	9.4	10.5	
	$E^{\text{QM}}(\text{p}, \text{p})^b$	13.8	14.6	9.8	10.9	
	$E^{\text{QM}}(\text{p}, \text{g})^c$	13.0	13.6	-	-	
	$E^{\text{QM}}(\text{g}, \text{g})^d$	12.0	12.5	-	-	

^a QM/MM energy difference. ^b QM contribution to the QM/MM energy difference. ^c Energy differences of the isolates species in the gas phase at the protein geometry. ^d Energy differences of the isolated species optimized in the gas phase. ^e Snapshot 40 refers to the protein/solvent environment in the QM/MM calculations (no data given for the gas phase species).

in the conformation of the substrate, camphor, within the protein pocket. The internal geometry parameters of the heme and camphor remain essentially the same, whereas the relative orientation of the two units changes notably. This change accounts for a small energy gain of *ca.* 2–3 kcal/mol. In the reactive complex, the angle $\theta_{\text{O-H-C}}$ is almost perfectly linear, much the same as the geometry in the H-abstraction transition state. The stable conformation representing the reactive complex indicates that the enzyme environment pre-organizes the substrate and Cpd I in the pocket, yielding a geometry that is ideally suited for H-abstraction.

The transition states were calculated at the R1pro/B1 level to have a relative energy of (LS/HS) 19.5/ 20.7 kcal/mol with respect to the reactive complex. The nature of the transition states as saddle points of first order was confirmed by numerical calculation of the hessian and vibrational frequencies of all QM atoms, which gave one mode with a eigenvalues of (LS/HS) 811i cm^{-1} / 561i cm^{-1} . The eigenmode corresponds to the expected motion of the hydrogen atom along the C–H–O axis. The calculations using the larger QM region R3 and basis B4 gave similar, albeit slightly higher activation energies of (LS/HS) 21.1/ 21.8 kcal/mol. We also confirmed by additional minimizations starting from the transition structures that this saddle point truly connects the reactive complex (7) and hydroxo intermediate (8).

In the transition structure the $\text{H}_{\text{exo}}^5\text{--C}_5$ distance is (LS/HS) 1.390/ 1.408 Å for R1pro/B1 and 1.359/1.369 Å for R3cam/B4, while the corresponding distances for $\text{H}_{\text{exo}}^5\text{--O}$ are 1.158/ 1.154 Å and 1.195/ 1.188 Å. Hence, in the calculations employing the larger QM region, the transition state occurs somewhat “earlier” on the reaction coordinate ($r_{\text{C-H}} - r_{\text{O-H}}$). In general, the transition state is closer to the product side, *i.e.* to the hydroxo intermediate 8. Inspection of table 4.3 shows a shift of unpaired spin density mainly from the oxygen atom to C_5 of camphor. The unpaired spin density on iron is also decreased, although to a lesser extent, while the porphyrin retains the π cation radical character. Hence, the system acquires some Fe^{III} character already in the transition state. The reduction of spin density on iron and sulfur causes a concomitant shortening of the Fe–S bond length from (LS/HS) 2.573/ 2.552 Å in the reactive complex to 2.533/ 2.523 Å in the transition state on the R3cam/B4 level. The H_{exo}^5 atom exhibits a small negative amount of spin density in all calculations, consistent with previous model studies [25, 65–68, 73–78].

The activation energies computed for the isolated systems with QM region R1 and basis B1 in the gas phase are very similar to the corresponding QM/MM results: For the isolated species in the gas phase at the protein geometry ($S_{\text{p,g}}$), values of (LS/HS) 20.4/ 21.1 kcal/mol result. Optimizing the system in the gas phase ($S_{\text{g,g}}$) leads to activation energies of (LS/HS) 19.5/ 19.5 kcal/mol, hence the barriers are very close to the QM/MM results. A comparison of the electronic features obtained in the enzyme and gas phase calculations reveals a few subtle changes. (i) In the gas phase, there is significantly more spin density on the sulfur atom (as has already been discussed for Cpd I, see chapter 3). For instance, the ratio of spin density porphyrin:sulfur in the HS reactive complex (gas phase) is 44:56, as compared to 80:20 in the QM/MM calculations. (ii) In the gas

phase transition state, the spin density distribution between iron and porphyrin is different in the doublet and quartet states: In the LS state, the spin density on iron (porphyrin) is 0.963 (-0.437). This species thus essentially retains the porphyrin π cation radical and therefore adopts some Fe^{III} character. In contrast, the HS transition state exhibits a spin density on iron (porphyrin) of 1.269 (0.123), which resembles more a Fe^{IV} state, with minor porphyrin π cation radical character. In the QM/MM calculations both electronic states are of the Fe^{III} -type. It thus seems that the enzyme environment accounts for a stabilization of the porphyrin π cation radical.

By changing the occupation of the orbitals in the KS determinant, we tried to model the Fe^{IV} -type transition state also for the LS case in the gas phase, and for both the QM/MM structures. However, due to the pronounced mixing of the frontier orbitals, the SCF calculation in all cases converged back to the Fe^{III} -type species.

Table 4.2: Selected structural data for the H-abstraction step (snap 40: QM/MM optimized structure of snapshot 40; RC: reactive complex **7**; TS_H : transition state for H-abstraction). Distances in Å, angles in degree.

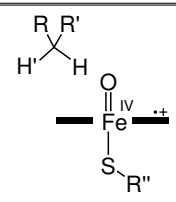
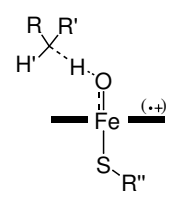
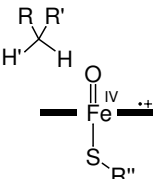
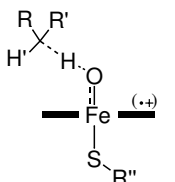
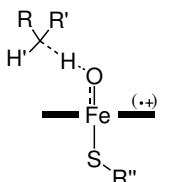
		QM/MM				Gas phase		
		R1pro/B1		R3cam/B4		R1pro/B1		
		LS	HS	LS	HS	LS	HS	
7(snap 40)	r _{Fe–O}	1.654	1.656	1.628	1.628	-	-	
	r _{Fe–S}	2.565	2.546	2.562	2.557	-	-	
	r _{O–H}	2.692	2.684	2.759	2.759	-	-	
	r _{C–H}	1.093	1.093	1.091	1.091	-	-	
	θ _{Fe–O–H}	110.0	110.6	111.2	111.4	-	-	
	θ _{O–H–C}	133.0	133.4	132.9	132.9	-	-	
7(RC)	r _{Fe–O}	1.654	1.656	1.626	1.627	1.650	1.652	
	r _{Fe–S}	2.568	2.545	2.573	2.552	2.602	2.580	
	r _{O–H}	2.145	2.156	2.187	2.190	2.625	2.713	
	r _{C–H}	1.093	1.093	1.091	1.091	1.098	1.098	
	θ _{Fe–O–H}	129.7	129.6	129.7	130.3	135.7	133.3	
	θ _{O–H–C}	174.1	172.2	175.0	175.1	172.1	166.6	
TS_H	r _{Fe–O}	1.794	1.789	1.763	1.761	1.774	1.758	
	r _{Fe–S}	2.533	2.536	2.533	2.523	2.487	2.400	
	r _{O–H}	1.158	1.154	1.195	1.188	1.152	1.181	
	r _{C–H}	1.390	1.408	1.359	1.369	1.371	1.362	
	θ _{Fe–O–H}	120.2	122.0	121.4	122.7	119.3	125.7	
	θ _{O–H–C}	166.4	166.9	167.6	167.8	170.6	171.6	

Table 4.3: Spin densities on selected atoms for the H-abstraction step (snap 40: QM/MM optimized structure of snapshot 40; RC: reactive complex **7**; TS_H: transition state for H-abstraction).

		QM/MM				Gas phase		
		R1pro/B1		R3cam/B4		R1pro/B1		
		LS	HS	LS	HS	LS	HS	
snap 40	$\rho(\text{O})$	0.919	0.963	0.880	0.901	-	-	
	$\rho(\text{Fe})$	1.201	1.062	1.306	1.154	-	-	
	$\rho(\text{S})$	-0.242	0.201	-0.291	0.228	-	-	
	$\rho_{\text{gr}}(\text{por})^a$	-0.887	0.777	-0.892	0.712	-	-	
	$\rho(\text{C})$	0.002	0.003	0.001	0.002	-	-	
	$\rho(\text{H})$	-0.001	-0.001	-0.002	-0.002	-	-	
RC	$\rho(\text{O})$	0.906	0.952	0.877	0.905	0.914	0.972	
	$\rho(\text{Fe})$	1.213	1.071	1.310	1.142	1.182	1.052	
	$\rho(\text{S})$	-0.242	0.200	-0.295	0.225	-0.617	0.558	
	$\rho_{\text{gr}}(\text{por})^a$	-0.889	0.776	-0.891	0.719	-0.503	0.437	
	$\rho(\text{C})$	0.008	0.009	0.006	0.007	0.002	0.002	
	$\rho(\text{H})$	-0.004	-0.00	-0.005	-0.003	-0.001	-0.001	
TS _H	$\rho(\text{O})$	0.504	0.594	0.546	0.612	0.502	0.681	
	$\rho(\text{Fe})$	1.038	0.887	1.157	0.947	0.963	1.269	
	$\rho(\text{S})$	-0.210	0.191	-0.257	0.211	-0.518	0.397	
	$\rho_{\text{gr}}(\text{por})^a$	-0.827	0.761	-0.892	0.717	-0.437	0.123	
	$\rho(\text{C})$	0.560	0.647	0.500	0.553	0.522	0.576	
	$\rho(\text{H})$	-0.057	-0.060	-0.072	-0.066	-0.047	-0.035	

^a $\rho_{\text{gr}}(\text{por})$ denotes the sum of individual atomic spin densities on the atoms of the porphyrin macrocycle.

In the light of the elusive nature of Cpd I and its high reactivity towards substrate hydroxylation, the computed barrier for the H-abstraction step of *ca.* 20 kcal/mol is surprisingly high. Although there is no experimental value, current estimates are in the range of approximately 12 kcal/mol [207]. Furthermore, the present comparative QM/MM and gas phase QM calculations indicate that the height of the barrier is not influenced much by the enzyme environment, which again is unexpected regarding the efficiency of the enzyme catalyzed reaction. In view of this situation, we deemed it necessary to explore factors that could possibly account for an artificially high barrier in our calculations. The following issues have been investigated and will be addressed in detail in the following sections: (i) the QM regions considered in our previous studies did not include the propionate side chains of the heme. However, the enzyme could effect transition state stabilization *via* a charge-transfer mechanism involving these negatively charged groups. This would imply that the electrostatic interactions of the protein salt-bridges with the propionate-carboxylate groups are intensified in the transition state, thus preferentially stabilizing the transition state with respect to the reactant. In effect, this would lower the barrier. (ii) Our computed activation energies do not incorporate thermal effects such as zero-point-energy (ZPE) corrections. However, in the transition state, the energy equivalent of one C–H bond vibration is lost with regard to the equilibrium structures, which is expected to lower the barrier significantly. (iii) Finally, the intrinsic accuracy of the QM method in our computational approach to describe the potential energy of the H-abstraction reaction was investigated by means of calibration studies on a simplified model system.

Another factor that has not been considered in the present work are tunneling effects, which may be expected to effectively lower the observed barrier for H-abstraction.

Porphyrin Sidechains

The interaction of the protoporphyrin IX propionate side chains with neighbouring positively charged protein residues is shown in figure 4.5. In the previously discussed calculations (*vide supra*) using QM regions R1pro and R3cam, the side chains and their environment were included in the MM calculation. Since the electrostatic force field terms are calculated from non-polarizable atomic charges, this approach cannot account for charge transfer to these groups. Therefore, we have performed single point energy calculations in the enzyme environment and in the gas phase, employing two larger QM regions that extend onto the side chains. The geometries were obtained from QM/MM optimizations at the R1pro/B1 level. The first QM region R1pro-S1 includes both propionate side chains of the heme (see figure 4.2). The second QM region R1pro-S2 additionally incorporates part of the Arg299 side chain. Only the 4A state was investigated.

The calculated QM activation energies ΔE_A , spin densities ρ , and Mulliken atomic charges Q for QM regions R1pro-S1 and R1pro-S2 are collected in table 4.4. In the case of the QM/MM

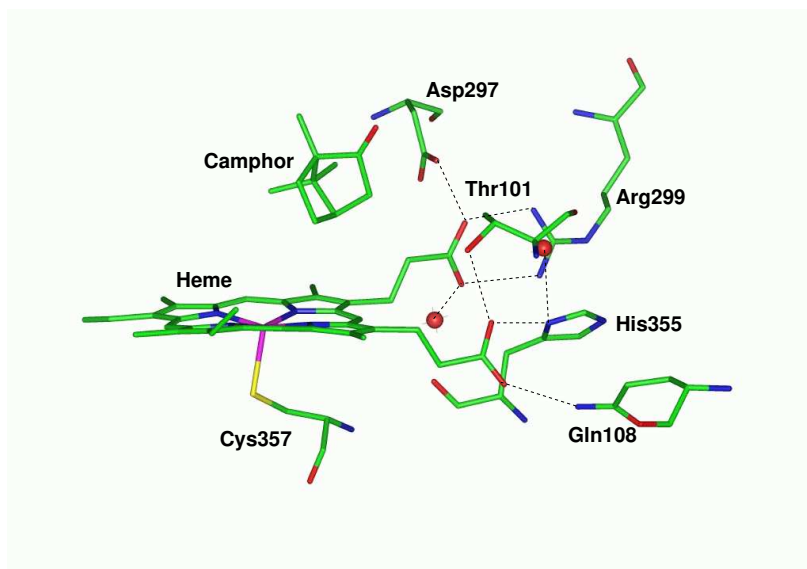


Figure 4.5: The salt bridges to the propionate side chains.

Table 4.4: QM/MM vs. gas phase single point energy calculations on R1pro/B1 optimized geometries including the propionate side chains in the QM region, 4A state, QM activation energies ΔE_A^{QM} (kcal/mol), unpaired spin densities ρ , and Mulliken atomic charges Q [e].

		R1pro-S1		R1pro-S2	
		QM/MM	gas phase	QM/MM	gas phase
RC	ΔE_A^{QM}	20.5	25.1	21.0	24.8
	$\rho(O)$	0.909	0.909	0.953	0.936
	$\rho(Fe)$	1.074	1.164	1.070	1.112
	$\rho(S)$	0.189	0.154	0.209	0.325
	$\rho_{gr}(por)^a$	0.784	0.773	0.767	0.634
	$\rho_{gr}(O_{prop})^b$	0.001	0.902	0.000	0.512
TS _H	$Q_{gr}(O_{prop})^c$	-2.811	-1.927	-2.746	-2.099
	$\rho(O)$	0.597	0.584	0.600	0.600
	$\rho(Fe)$	0.888	0.968	0.888	0.937
	$\rho(S)$	0.198	0.161	0.205	0.304
	$\rho_{gr}(por)^a$	0.751	0.770	0.744	0.630
	$\rho(C)$	0.646	0.621	0.646	0.630
	$\rho_{gr}(O_{prop})^b$	0.000	0.868	0.000	0.482
	$Q_{gr}(O_{prop})^b$	-2.814	-1.944	-2.748	-2.115

^a The sum of individual atomic spin densities on the atoms of the porphyrin macrocycle. ^b The sum of individual atomic spin densities on the oxygen atoms of the propionate side chains. ^c The sum of Mulliken atomic charges of the oxygen atoms of the propionate side chains.

single point calculations, we give the QM contribution to the barrier, because this relative energy reflects the electrostatic energy contributions that we are interested in (recall that in the present approach, the QM–MM electrostatic interactions are included in the QM energy). In the QM/MM calculations, the barriers for H-abstraction are 20.5 and 21.0 kcal/mol for R1pro-S1 and R1pro-S2, respectively. Hence, they are almost identical to the barriers obtained for QM regions R1pro and R3cam (20.9 and 22.8 kcal/mol, respectively, see table 4.1). The differences in QM/MM calculated Mulliken spin densities (Mulliken atomic charges) on the carboxylate oxygens between the reactive complex and transition state are 0.001 (0.003 e) and 0.000 (0.002 e) for R1pro-S1 and R1pro-S2, respectively. This demonstrates that there is no significant charge transfer to the carboxylate groups in the transition state, which could account for a selective stabilization of the transition state with respect to the reactive complex.

On the other hand, the calculated barriers in the gas phase are 25.1 and 24.8 kcal/mol, respectively. The increased spin densities and decreased atomic charges indicate that the localization of electronic charge on the carboxylate groups is unfavourable in the gas phase. For R1pro-S1, where there is no stabilization of the negative charge on the carboxylate group, the total spin density on the propionate oxygen atoms is (RC/TS_H) 0.902/0.868 while the total Mulliken charge is (RC/TS_H) -1.927/-1.944. Hence, there is significantly less electron density on these groups as compared to the QM/MM calculations. In R1pro-S2, the H-bonds provided by the arginine side chain group account for a smaller unpaired spin density of (RC/TS_H) 0.512/0.482 and higher electronic charge of -2.099/-2.115 on this group of atoms. The loss of the polarizing environment in the gas phase calculations indeed increases the energy of the transition state more than that of the reactive complex, resulting in higher calculated barriers. However, this effect is not present in the QM/MM calculations: Qualitatively the same barriers are retrieved for all QM regions R1pro, R3cam, R1pro-S1 and R1pro-S2, *i.e.* regardless of the inclusion of the propionate side chains. Hence, the higher barriers of approximately 25 kcal/mol in the gas phase for the QM systems R1pro-S1 and R1pro-S2 are obviously an artifact of the lacking polarizing environment in these calculations. In summary, the inclusion of the propionate side chains into the QM region does not affect the results obtained at the QM/MM level, but leads to inconsistent results for the isolated gas phase species.

Zero-point vibrational corrections

Vibrational energies of the reactive complex and the transition state TS_H were obtained from frequency calculations of the isolated species in the gas phase using QM region R1 and basis B1, for the ²A as well as the ⁴A state. These calculations predict that the zero-point energy (ZPE) decreases by 4.3 kcal/mol and 4.1 kcal/mol, in the LS and HS states, respectively, when going from the reactive complex to the transition state. This loss of vibrational energy can be rationalized by

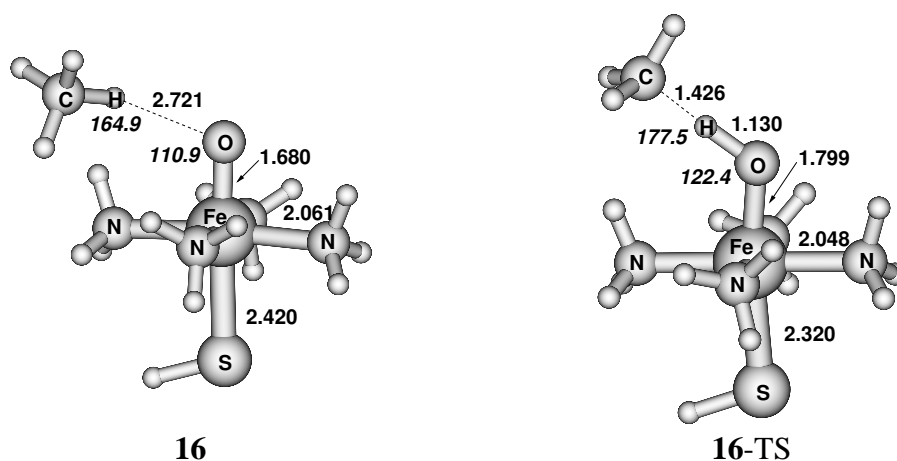


Figure 4.6: Optimized geometries (B3LYP/SV(P)) of the model compound in the reactive complex **16** and the transition state **16-TS**. Distances in Å, angles (oblique numbers) in degree.

the fact that the energy of the C–H vibration of the $C^5 - H_{\text{exo}}^5$ bond in the reactant is missing in the transition state.

Inclusion of zero-point energy leads to barriers of (LS/HS) 15.2/15.4 kcal/mol for the gas phase reaction. It is reasonable to assume that the effect is analogous in the enzyme environment, because it is mainly due to the loss of the C–H vibrational energy within the substrate. Applying the gas phase ZPE correction to the QM/MM calculated activation energies yields estimates for the barriers of (LS/HS) 15.2/16.5 kcal/mol at the R1pro/B1 level and (LS/HS) 16.8/17.7 kcal/mol at the R3cam/B4 level.

QM calibration study

The QM calibration study served to explore the intrinsic accuracy of the DFT method in describing the H-abstraction barrier and employed the simplified molecular model depicted in figure 4.3. A closely related nonheme $Fe^{IV}=O$ complex has recently been characterized by X-ray crystallography [208]. This complex, $[Fe(O)(TMC)(NCCH_3)]^{2+}$ (TMC = 1,4,5,11-tetramethyl-1,4,8,11-tetraazacyclotetradecane), features a Fe–O bond length of 1.646(3) Å and four equatorial amino-ligands with an average Fe–N bond distance of 2.091 Å, in close agreement with the optimized geometry calculated for **16**. The electronic ground state of the cationic ferryl complex is a triplet ($S = 1$) with two unpaired electrons occupying d-p π^* orbitals, and hence isoelectronic to the ferryl oxene moiety in Cpd I. The B3LYP/SV(P) optimized geometries and selected geometric parameters of the model reactive complex **16** and transition state **16-TS** are shown in figure 6.2.

Subsequent to the full geometry optimizations, single point calculations using different density functionals and basis sets were performed. The results are summarized in table 4.5. The B3LYP

calculated barriers with basis sets B1 and B4 are 22.7 and 19.5 kcal/mol, respectively. They are close to the barriers obtained for the full systems (R1pro/B1, $S_{g,g}$: 19.5 kcal/mol both for 2A and 4A state). This underscores the fact that the chosen simple model is realistic.

Applying different density functionals with the same large basis set (AE1, 508 basis functions) yields activation energies ranging from 25.2 kcal/mol (BLYP) to 15.1 kcal/mol (PBE0). The pure GGA functionals BP86 and BLYP give the highest barriers of 22.4 and 25.2 kcal/mol, respectively. The admixture of exact exchange apparently reduces the barrier. Hence, the B3LYP and B97 functionals which both contain 20% exact exchange predict barriers of 19.5 and 20.1 kcal/mol, respectively. The smallest barrier is obtained with the PBE0 functional, containing 25% exact exchange.

Table 4.5: Activation energies [kcal/mol] for H-abstraction in the model system (see figure 4.3) for different density functionals and basis sets.

Different density functionals, basis AE1 ^a					
BP86	BLYP	B3LYP	B97	PBE0	
22.4	25.2	19.5	20.1	15.1	
Different basis sets, B3LYP density functional					
B1	B4	SV	SVP	TZVP	AE2 ^b
22.7	19.5	25.4	21.7	21.4	19.3

^a AE1: Fe Wachters all electron [8s7p4d1f] + cc-pVTZ on others (508 basis functions).

^b AE2: Fe:Wachters[8s7p4d1f], others:cc-pVQZ without g functions (892 basis functions).

The barriers calculated with the B3LYP density functional and different basis sets exhibit a clear trend. The simple double- ζ quality basis sets give the highest barriers (SV: 25.4 kcal/mol, B1: 22.7). The addition of polarization and/or diffuse functions reduces the relative energy by 3-4 kcal/mol (SVP: 21.7 kcal/mol, B4: 19.5 kcal/mol), whereas an increased flexibility in the valence region has only a minor effect (TZVP: 21.4 kcal/mol). The largest basis set AE2 which employs a Wachters basis for iron with additional f-polarization functions and a quadruple- ζ basis set with polarization up to f-functions on the other atoms (892 basis functions total) gives a barrier of 19.3 kcal/mol. One may thus conclude that basis B4, that is also used in the studies on the full system, provides essentially converged B3LYP results with respect to basis set size.

Conclusion

The results presented in the preceding sections lead to the following conclusions: (i) The QM/MM calculated barrier for hydrogen abstraction is not affected significantly by an extension of the QM regions onto the propionate side chains of the heme. The results obtained with QM regions R1pro

and R3cam are thus reliable. Larger QM regions including the propionate side chains give inconsistent results if only the isolated gas phase systems are considered. (ii) Accounting for ZPE corrections obtained at the R1pro/B1 level for the isolated gas phase species lowers the calculated barriers by *ca.* 4 kcal/mol. This large effect can be understood on the basis of the vibrational energy of one C–H bond that is being lost in the transition state. (iii) The calculated barrier is very sensitive to the choice of the density functional: we observe variations of *ca.* 10 kcal/mol. The basis set B4 provides B3LYP results close to the basis set limit.

Tunneling effects may be expected to further lower the effective barrier. For example, Truhlar, Gao and coworkers applied variational transition state theory with explicit treatment of quantum mechanical tunneling to the proton-abstraction catalyzed by enolase and methylamine dehydrogenase, and the hydride-transfer reactions catalyzed by liver alcohol dehydrogenase and xylose isomerase [209]. They found that tunneling reduces the barriers by 0.3, 2.4, 0.8, and 1.1 kcal/mol, respectively. Since the H-abstraction studied in the present work represents a comparable situation, one may expect a reduction of the observed barrier for H-activation on the order of 0–3 kcal/mol through tunneling effects.

4.3.2 The Hydroxo Intermediate

QM/MM geometry optimizations at the R1pro/B1 level that were started from the PES scans described in section 4.3.1 at short distances $r(\text{H}_{\text{exo}}^5\text{--O})$ lead to the intermediate species **8**. This system has been identified in previous gas phase DFT calculations [25, 64–68, 73–78, 210] as an iron-hydroxo complex and the loosely coordinated substrate radical, nascent from the hydrogen abstraction.

The intermediate cluster **8** is characterized by four MOs with similar energy, hosting a total of five electrons. This electronic situation is schematically depicted in figure 4.7. Two of these MOs consist of the two d_{xz} and d_{yz} orbitals of the iron atom, mixing with the oxygen p_x and p_y orbitals. The third orbital is the mixed porphyrin- a_{2u} and $p_\sigma(\text{S})$ orbital discussed in chapter 3, while the fourth orbital is a p-type orbital mainly localized at the C^5 atom of camphor. This orbital is denoted $\phi(\text{C}^5)$ in the following. As a consequence of this near-degeneracy of orbitals, there are several low-lying electronic states. Essentially, two electronic configurations, shown in figure 4.7, are relevant to the representation of the ground state. In the first, one of the three unpaired electrons is mainly localized at the iron center (and delocalized on the coordinating oxygen), another one, in the $a_{2u}/p_\sigma(\text{S})$ orbital, is delocalized over the porphyrin and thiolate ligand, and the third one occupies the $\phi(\text{C}^5)$ orbital. This electronic structure corresponds to iron being in the formal oxidation state III (see figure 4.7). As discussed previously [210], a second electromer with a low energy exists that has two unpaired electrons in d-orbitals of iron and the third one in $\phi(\text{C}^5)$.

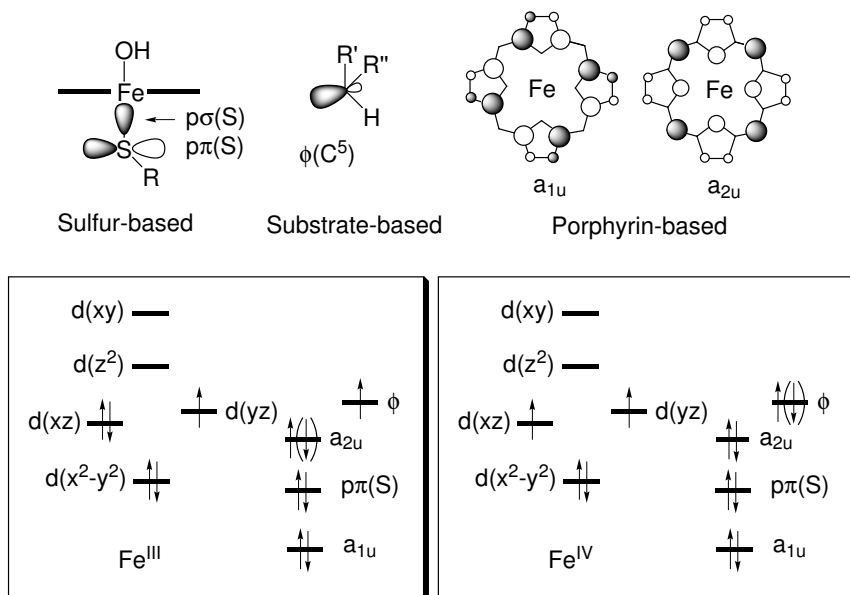


Figure 4.7: The hydroxo intermediate **8**: orbitals and occupations. The electron spin symbol in parantheses denotes the corresponding LS configuration.

In this species, the formal oxidation state of iron is IV and there is only residual unpaired spin density on porphyrin and sulfur ligands (see figure 4.7).

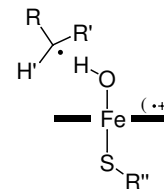
Another (UKS broken symmetry) determinant with a different electronic configuration also gives rise to an energetically low-lying Fe^{III} electromer with doublet multiplicity. This determinant has a spin-up (α) electron in the porphyrin a_{2u}-type orbital and a spin-down (β) electron in the ϕ -orbital centered on C⁵ of camphor. We have performed QM/MM optimizations on this species using QM region R3 and basis B4 and found it to be higher in energy by 0.36 kcal/mol as compared to the Fe^{III} LS species discussed above and shown in figure 4.7. The optimized geometries of the two systems are virtually identical. Because the configuration displayed in figure 4.7 gives the lower energy it is apparently better suited to describe the Fe^{III} LS ground state and all calculations reported here on this state employed this configuration.

We have also performed geometry optimizations of the sextet ($S = 5/2$) state on the R1pro/B1 level. The optimized species is higher in energy by 6–7 kcal/mol as compared to the two energetically close-lying doublet and quartet states and thus is not considered further at this stage.

The Fe^{III} and Fe^{IV} electromers were optimized in their LS and HS states respectively. The resulting relative energies are given in table 4.6. Relevant geometry data are listed in table 4.7, while unpaired spin densities are summarized in table 4.8. Again, we compare the QM/MM results at the R1pro/B1 level to the analogous isolated system in the gas phase. To qualitatively explore the nature of the lowest electronically excited states in each species, we have carried out time-dependent DFT (TDDFT) calculations for QM region R1 and basis B1. The TDDFT excitations were calculated both for the system in the enzyme environment ($S_{p,p}$) and the isolated gas phase

Table 4.6: Relative energies [kcal/mol] of electromers of the hydroxo intermediate **8**.

		R1pro/B1		R3cam/B4	
		LS	HS	LS	HS
Fe ^{III}	$E^{\text{QM/MM}}(\text{p}, \text{p})^a$	0	0.86	0	1.08
	$E^{\text{QM}}(\text{p}, \text{p})^b$	0	0.81	0	1.09
	$E^{\text{QM}}(\text{p}, \text{g})^c$	0	0.56	-	-
	$E^{\text{QM}}(\text{g}, \text{g})^d$	0	0.89	-	-
Fe ^{IV}	$E^{\text{QM/MM}}(\text{p}, \text{p})^a$	1.12	1.96	2.55	3.28
	$E^{\text{QM}}(\text{p}, \text{p})^b$	1.29	1.55	3.07	3.43
	$E^{\text{QM}}(\text{p}, \text{g})^c$	-2.28	-1.73	-	-
	$E^{\text{QM}}(\text{g}, \text{g})^d$	-2.54	-2.08	-	-



^a QM/MM energy difference. ^b QM contribution to the QM/MM energy difference. ^c Energy differences of the isolates species in the gas phase at the protein geometry. ^d Energy differences of the isolated species optimized in the gas phase.

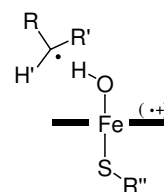
species ($S_{\text{g,g}}$). The $S_{\text{p,p}}$ TDDFT calculations are based on geometries optimized at the QM/MM (R1pro/B1) level in the protein environment. The electron densities that enter the TDDFT calculations have been determined in the presence of the polarizing point charge field representing the surrounding enzyme. The gas phase excitations ($S_{\text{g,g}}$) are calculated at the geometry fully optimized in the gas phase. Apart from yielding information about the excited states, these calculations also served to confirm the stability of the UKS determinants used to describe the ground state. Table 4.9 and 4.10 list the calculated excitations below 1 eV in the HS and LS states, respectively.

The relative energies presented in table 4.6 reveal that the Fe^{III} and Fe^{IV} electromers are close in energy. Generally, the LS states are slightly preferred over the analogous HS states. QM/MM calculations at the R1pro/B1 level predict the Fe^{III} state to be the ground state, with the Fe^{IV} state lying only (LS/HS) 1.12/1.10 kcal/mol higher in energy. When the larger QM region R3cam is used, the gap is somewhat increased (2.55/2.20 kcal/mol). The dominant contributions to this small energy differences come from the QM calculation (R1pro/B1: 1.29/0.75 kcal/mol; R3cam/B4: 3.07/2.34). Interestingly, the ordering of the states is reversed in the gas phase: Without polarizing point charge environment ($S_{\text{p,g}}$), the Fe^{IV} electromer is more stable than the Fe^{III} species by (LS/HS) 2.28/2.29 kcal/mol. As is obvious from the data in table 4.6, geometry optimization in the gas phase ($S_{\text{g,g}}$) has only a minor effect on the relative energies of Fe^{III} vs. the Fe^{IV} state. It is thus the polarizing effect of the enzyme environment that stabilizes the Fe^{III} electromer. This finding is in accord with DFT model calculations by Filatov *et al.* [210], who demonstrated that a polarizing continuum environment stabilizes the Fe^{III} porphyrin radical cation state. This effect

was attributed to the larger dipole moment of the latter as compared to the Fe^{IV} species.

Table 4.7: Selected structural data for hydroxo intermediate **8**. Distances in Å, angles in degree.

		QM/MM				Gas phase	
		R1pro/B1		R3cam/B4		R1pro/B1	
		LS	HS	LS	HS	LS	HS
Fe^{III}	$r_{\text{Fe}-\text{O}}$	1.823	1.818	1.821	1.812	1.820	1.813
	$r_{\text{Fe}-\text{S}}$	2.497	2.495	2.433	2.456	2.474	2.445
	$r_{\text{O}-\text{H}}$	0.989	0.990	0.971	0.973	0.986	0.985
	$r_{\text{C}-\text{O}}$	2.980	3.028	3.117	3.154	3.190	3.310
	$\theta_{\text{Fe}-\text{O}-\text{H}}$	111.5	112.3	109.0	111.0	111.8	112.2
	$\Theta_{\text{C}-\text{O}-\text{Fe}-\text{H}}$	-15.1	-12.4	-27.6	-14.0	-17.3	-10.5
Fe^{IV}	$r_{\text{Fe}-\text{O}}$	1.784	1.793	1.781	1.787	1.796	1.801
	$r_{\text{Fe}-\text{S}}$	2.454	2.445	2.414	2.405	2.366	2.357
	$r_{\text{O}-\text{H}}$	1.000	0.994	0.982	0.978	0.995	0.990
	$r_{\text{C}-\text{O}}$	2.955	3.001	3.065	3.110	3.041	3.008
	$\theta_{\text{Fe}-\text{O}-\text{H}}$	117.8	116.5	114.8	113.6	117.7	117.0
	$\Theta_{\text{C}-\text{O}-\text{Fe}-\text{H}}$	-11.6	-11.8	-12.2	-9.6	-4.1	-11.8



The differences in the electronic structures of the Fe^{III} vs. the Fe^{IV} state are also manifest in the respective optimized structures. Most notably, the bond distances to the axial ligands are decreased in the Fe^{IV} electromer. This effect may be rationalized by the smaller Pauli repulsion of the electrons occupying the Fe–ligand σ bonds with the d_{xz}/d_{yz} electrons [210].

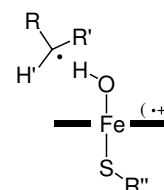
The dihedral angle $\text{C}^5\text{--O--Fe--H}$ determines the relative position of the C^5 atom with respect to the plane spanned by the atoms O--Fe--H . This parameter is important since it indicates from which position the carbon radical will attack the oxygen atom in the rebound phase. This dihedral angle is found to be negative, ranging from -10 to -28 degrees in the QM/MM optimized structures. Hence, there is a clear preference for one conformation, indicating that attack in the rebound step will occur from one side of the O--Fe--H plane only.

The unpaired spin densities summarized in table 4.8 conform to the electronic structures discussed above: The Fe^{III} (Fe^{IV}) species is typified by a spin density equivalent of approximately one (two) unpaired electron(s) on the iron center. The results obtained for the isolated species in the gas phase show little influence of environmental effects for the Fe^{IV} case, while the Fe^{III} electromer exhibits the expected shift of spin density from the porphyrin ring to the sulfur atom.

The TDDFT calculations reveal several interesting features. Generally, the canonical orbitals obtained from the HS calculation mix less than those of the LS species and the excitations are

Table 4.8: Spin densities on selected atoms for the hydroxo intermediate **8**.

		QM/MM				Gas phase	
		R1pro/B1		R3cam/B4		R1pro/B1	
		LS	HS	LS	HS	LS	HS
Fe ^{III}	$\rho(\text{O})$	0.135	0.219	0.085	0.175	0.094	0.244
	$\rho(\text{Fe})$	1.020	0.877	1.120	0.905	1.005	0.922
	$\rho(\text{S})$	-0.169	0.151	-0.154	0.142	-0.478	0.443
	$\rho_{\text{gr}}(\text{por})^a$	-0.881	0.811	-0.896	0.810	-0.586	0.437
	$\rho(\text{C})$	0.955	1.012	0.876	0.976	0.994	1.034
	$\rho(\text{H})$	0.001	0.001	-0.006	0.004	0.006	-0.003
Fe ^{IV}	$\rho(\text{O})$	0.215	0.275	0.236	0.250	0.260	0.320
	$\rho(\text{Fe})$	1.933	1.981	2.042	2.122	1.830	1.803
	$\rho(\text{S})$	-0.117	-0.088	-0.140	-0.116	-0.023	0.035
	$\rho_{\text{gr}}(\text{por})^a$	-0.119	-0.123	-0.213	-0.217	-0.134	-0.122
	$\rho(\text{C})$	-0.950	0.982	-0.928	0.952	-0.963	0.993
	$\rho(\text{H})$	-0.021	0.029	-0.005	0.017	-0.017	0.018



^a The sum of individual atomic spin densities on the atoms of the porphyrin macrocycle.

thus easier to interpret in terms of MOs in the ⁴A case. Table 4.9 lists the excitations in the ⁴A species below 1 eV together with an analysis of the corresponding transitions. The HS Fe^{III} species exhibits a number of excitations with comparably low energy into the unoccupied β - a_{2u} spin-orbital. The first excited state (0.45 eV) corresponds to the Π sulfur radical, while the A_{1u} porphyrin cation radical state is obtained at 0.55 eV. The $d_{xz} \rightarrow a_{2u}$ transition gives rise to the Fe^{IV} species and dominates the transitions at 0.85 and 0.99 eV. In the gas phase, the Π sulfur radical state has a slightly lower excitation energy (0.39 eV), reflecting the lacking stabilization of charge on the sulfur atom through the polarizing enzyme environment. In contrast, the A_{1u} state is calculated at a higher energy in the gas phase (0.82 eV). As discussed above, the global minimum in the gas phase corresponds to the Fe^{IV} state. In line with these findings, excitations giving rise to the Fe^{IV} species ($d_{xz} \rightarrow a_{2u}$) are calculated at significantly lower energy in the gas phase, *i.e.* at 0.58 and 0.78 eV.

The lowest excitation calculated for the HS Fe^{IV} species is obtained at (enzyme/gas phase) 0.72/-0.93 eV and is dominated by the $d_{x^2-y^2} \rightarrow d_{xz}$ transition. The next higher excited state (0.94/0.99 eV) is characterized by the $a_{2u} \rightarrow d_{xz}$ excitation and corresponds to the Fe^{III} state. Regarding these low excitation energies, interconversion of the Fe^{III} and Fe^{IV} states is likely under ambient conditions. Moreover, one may expect some population of Fe^{III} Π sulfur radical and A_{1u} porphyrin cation radical states.

Table 4.9: TDDFT excitation energies (QM region R1, B3LYP, basis B1) for the hydroxo intermediate in the HS (4A) state. Geometries are optimized at the R1pro/B1 level. Comparison of QM/MM and gas phase calculations.

species	no.	E[eV]		excitation	c^2
QM/MM					
Fe ^{III}	1a	0.45	β	$p_\pi \rightarrow a_{2u}$	87
	2a	0.55	β	$a_{1u} \rightarrow a_{2u}$	88
	3a	0.59	β	$d_{xz} \rightarrow \pi^*(\text{porph})$	28
			β	$d_{xz} \rightarrow d_{yz}$	21
	4a	0.80	β	$d_{x^2-y^2} \rightarrow \pi^*(\text{porph})$	24
			β	$d_{x^2-y^2} \rightarrow d_{yz}$	18
			β	$d_{xz} \rightarrow a_{2u}$	18
	5a	0.85	β	$d_{xz} \rightarrow a_{2u}$	67
	6a	0.99	β	$d_{xz} \rightarrow a_{2u}$	76
Fe ^{IV}	1a	0.72	β	$d_{x^2-y^2} \rightarrow d_{xz}$	51
			β	$d_{x^2-y^2} \rightarrow d_{xz}$	31
	2a	0.94	β	$a_{2u} \rightarrow d_{xz}$	88
	3a	0.98	α	$a_{2u} \rightarrow d_{z^2}$	83
Gas phase					
Fe ^{III}	1a	0.39	β	$p_\pi \rightarrow a_{2u}$	82
	2a	0.58	β	$d_{xz} \rightarrow a_{2u}$	45
			β	$d_{xz} \rightarrow d_{yz}$	20
	3a	0.78	β	$d_{xz} \rightarrow a_{2u}$	38
			β	$d_{xz} \rightarrow d_{yz}$	21
	4a	0.82	β	$a_{1u} \rightarrow a_{2u}$	99
	5a	0.91	β	$d_{x^2-y^2} \rightarrow d_{yz}$	43
Fe ^{IV}	1a	0.93	β	$d_{x^2-y^2} \rightarrow d_{xz}$	37
			β	$p_\pi \rightarrow d_{xz}$	19
	2a	0.99	β	$a_{2u} \rightarrow d_{xz}$	47
			β	$p_\pi \rightarrow d_{xz}$	15

Table 4.10: TDDFT excitation energies (QM region R1, B3LYP, basis B1) for the hydroxo intermediate in the LS (2A) state. Geometries are optimized at the R1pro/B1 level. Comparison of QM/MM and gas phase calculations.

species	no.	E[eV]		excitation	c^2
QM/MM					
Fe^{III}	1a	0.47	α	$d_{xz}, \phi \rightarrow a_{2u}$	50
			α	$a_{1u} \rightarrow a_{2u}$	29
			α	$p_{\pi} \rightarrow a_{2u}$	19
	2a	0.53	α	$d_{xz}, \phi \rightarrow a_{2u}$	47
			α	$p_{\pi} \rightarrow a_{2u}$	31
			α	$a_{1u} \rightarrow a_{2u}$	20
	3a	0.57	α	$a_{1u} \rightarrow a_{2u}$	50
			α	$p_{\pi} \rightarrow a_{2u}$	49
	4a	0.73	β	$d_{xz} \rightarrow d_{yz}$	54
	5a	0.91	β	$d_{x^2-y^2} \rightarrow d_{yz}$	54
Fe^{IV}	1a	0.82	β	$d_{x^2-y^2} \rightarrow d_{z^2}$	75
	2a	0.95	β	$a_{2u} \rightarrow d_{z^2}$	91
Gas phase					
Fe^{III}	1a	0.38	α	$p_{\pi} \rightarrow a_{2u}$	77
	2a	0.43	α	$\phi, d_{xz} \rightarrow a_{2u}$	90
	3a	0.74	β	$d_{xz} \rightarrow d_{yz}$	62
	4a	0.82	α	$a_{1u} \rightarrow a_{2u}$	99
	5a	0.94	β	$d_{x^2-y^2} \rightarrow d_{yz}$	53
Fe^{IV}	1a	0.98	β	$d_{x^2-y^2} \rightarrow d_{xz}$	59

The TDDFT calculations on the LS states are summarized in table 4.10. Qualitatively, the results are very similar to the HS case discussed above. For example, the lowest excitations of the Fe^{III} species are dominated by transitions into the empty $\alpha\text{-a}_{2\text{u}}$ spin-orbital. In the enzyme environment, however, the first transition is dominated by a $\text{d}_{xz} \rightarrow \text{a}_{2\text{u}}$ transition, corresponding to a Fe^{IV} -type excited state, albeit with another occupation pattern than the one shown in figure 4.7. This transition also dominates the 2a excitation, but there is also a significant contribution from the $\text{p}_{\pi} \rightarrow \text{a}_{2\text{u}}$ excitation, giving rise to the sulfur II radical state. The latter transition clearly dominates the lowest excited state in the gas phase. This emphasizes again the stabilization of electronic charge on the sulfur atom in the enzyme environment.

The present TDDFT calculations on one hand serve to confirm that the considered configurations correspond to the electronic ground state of the discussed electromers. This is not a trivial issue, since in the present system, with high local symmetry and several possible occupations, the present approach might have converged to unstable solutions of the Kohn–Sham equations. On the other hand, and more importantly, the calculations demonstrate that several electronic states exist below 1 eV. This is due to the presence of nearly degenerate orbitals on both the ligands and the iron atom. The calculated excitation energies are not regarded as being quantitatively accurate, but they do indicate that several electronic states may be populated at ambient conditions and contribute to the reactivity of the system. The consequences of such a multi-state-scenario are yet to be explored.

4.3.3 The Rebound Step

In the second phase of the rebound mechanism – the rebound step – the camphor radical and the hydroxo complex **8** recombine to give the product alcohol complex **9**. Two internal coordinates are relevant to the PES of this process: the dihedral angle $\Theta(\text{C}^5\text{--Fe--O--H}^5)$ and the distance $r(\text{C}^5\text{--O})$. The true reaction coordinate that connects **8** and **9** is a combination of these two internal coordinates (of the form $c_1r + c_2\Theta$). Since a two dimensional scan is too expensive for the present level of theory, the PES of the HS and LS state for snapshot 40 have been scanned following both internal coordinates separately, while fully optimizing all other degrees of freedom. The results of the scan along $r(\text{C}^5\text{--O})$ for the LS state are shown in figure 4.8a. After the two fragments have approached each other to a distance of $r(\text{C}^5\text{--O}) = 2.67 \text{ \AA}$, the hydroxyl group flips around the Fe–O axis such that a free coordination site at the oxygen atom is present for the substrate radical to attack. Beyond this point the two fragments collapse in an exothermic fashion to give the alcohol complex **9**. The highest QM/MM energy obtained from the scan corresponds to a barrier of 1.7 kcal/mol with respect to the hydroxo intermediate. We have performed frequency calculations and transition state searches starting from this conformation. However, the vibrational analysis gave no mode with a negative eigenvalue and the transition state searches failed to locate a stationary point. It is thus likely that the small barrier obtained from the scan is due to a “hysteresis” effect

of the scan procedure.

We could not carry out a PES scan of the dihedral angle $\Theta(\text{C}^5\text{-Fe-O-H}^5)$ on the LS surface³. However, when this dihedral was slightly increased manually from -15 to -45 degree, free geometry optimizations directly led to the product complex. This indicates that the process is barrierless once the OH group has slightly rotated and the oxygen atom becomes accessible for attack by the substrate radical. Hence, consistent with previous gas phase studies, all our results indicate that this process has a small – if any – barrier.

Table 4.11: PES scan of the coordinate $r(\text{C}^5 - \text{O})$ on the ^4A surface, B3LYP/CHARMM, QM region R1pro, basis set B1. Distance in Å, dihedral angle in degree; relative QM/MM energy with QM and MM contributions in kcal/mol^a.

$r(\text{C}^5\text{-O})$	$\Theta(\text{C}^5\text{-Fe-O-H}^5)$	$\Delta\text{E}(\text{QM/MM})$	$\Delta\text{E}(\text{QM})$	$\Delta\text{E}(\text{MM})$
3.028	-12.4	0.0	0.0	0.0
2.591	-55.6	6.0	6.1	-0.1
2.517	-56.2	7.5	7.9	-0.4
2.442	-60.8	9.1	9.4	-0.2
2.368	-106.2	-11.9	-16.6	4.7
2.292	-107.0	-15.5	-20.2	4.7
2.217	-108.0	-19.2	-23.9	4.7
2.142	-108.6	-23.3	-28.0	4.7
2.067	-108.3	-27.7	-32.3	4.6
1.992	-108.9	-32.2	-36.9	4.7

^a Reference point is the hydroxo complex (**8**), Fe^{III} electrophile, ^4A state.

The scans of the rebound step on the HS surface proved to be complex. Starting from the hydroxo intermediate **8**, the scan along the distance $r(\text{C}^5\text{-O})$ does not lead to the product, but back to the reactants, *i.e.* to Cpd I and camphor. This is because the unconstrained dihedral angle $\Theta(\text{C}^5\text{-Fe-O-H}^5)$ is still close to the initial configuration (as in the hydroxo intermediate **8**), such that the scan along the distance $r(\text{C}^5\text{-O})$ leads to an approach of the carbon-centered radical and the hydrogen atom of the hydroxo group. In effect, this scan crosses the transition state for H-abstraction (TS_H) and reaches the initial reactants. On the other hand, following the dihedral angle $\Theta(\text{C}^5\text{-Fe-O-H}^5)$ simply gave the two dissociated fragments (the camphoryl radical and the Fe^{III} hydroxo complex), with a relative energy approximately 4 kcal/mol above the intermediate **8**. None of the two internal coordinates alone thus leads to the product, and a linear combination of the two seems to be required to describe the desired process. To circumvent expensive two-dimensional scans, one of

³Due to technical limitations of the HDLC optimizer it was not possible to constrain the internal coordinate corresponding to the dihedral angle $\Theta(\text{C}^5\text{-Fe-O-H}^5)$ at all points along the scan.

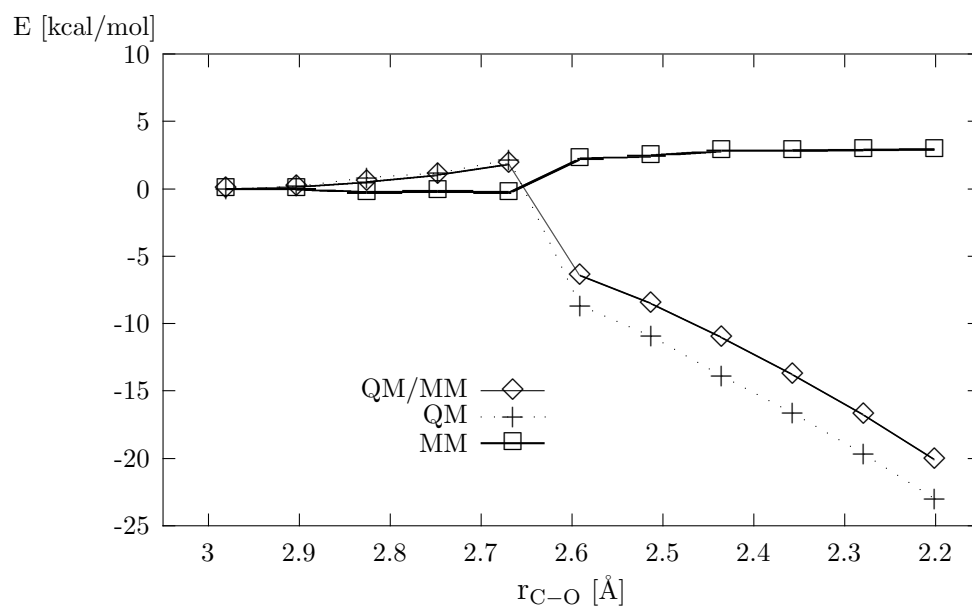
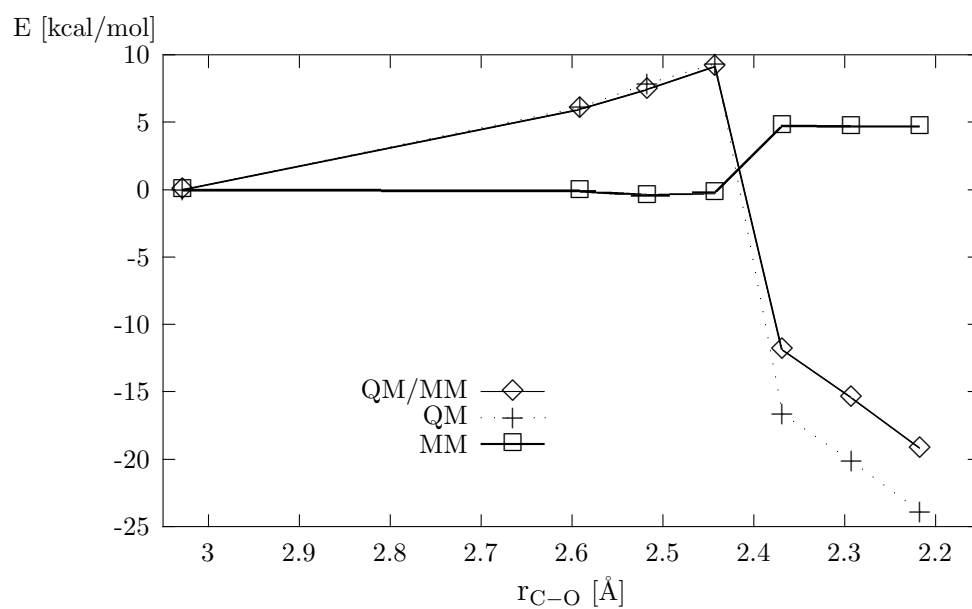
a) 2A stateb) 4A state

Figure 4.8: PES scan of the coordinate $r(C^5 - O)$, B3LYP/CHARMM, QM region R1pro, basis set B1 (see text).

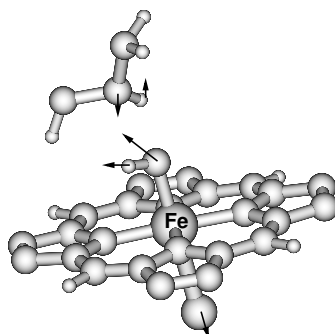


Figure 4.9: Transition mode for the rebound step on the 4A surface (TS_R) from B3LYP/CHARMM calculations (QM region R1pro, basis B1).

the steps of the scan along the dihedral coordinate, in which the hydroxo group has already rotated and a free site on oxygen for the attack of the camphoryl is present, was chosen as the starting point for a subsequent scan of the $r(C^5-O)$ -coordinate. The resulting energy profile of this scan is shown in figure 4.8b. Table 4.11 lists the QM/MM energy with QM and MM contributions, and the dihedral angle $\Theta(C^5-Fe-O-H^5)$ as a function of the constraint distance $r(C^5-O)$. The scan was started from the structure with a distance $r(C^5-O) = 2.591$ Å and $\Theta(C^5-Fe-O-H^5) = -55.6$ degrees, which is 6 kcal/mol higher in energy than the hydroxo intermediate **8**. The point with the highest energy occurs at $r(C^5-O) = 2.442$ Å and $\Theta(C^5-Fe-O-H^5) = -60.8$ degrees and is 9 kcal/mol above **8**. Beyond that point the energy drops sharply and the dihedral angle $\Theta(C^5-Fe-O-H^5)$ remains constant at approximately -107 degrees. Free geometry optimizations starting from such conformations yield the product complex **9**. Again all relative QM/MM energies are clearly dominated by the QM contributions (see figures 4.8a and 4.8b).

Table 4.12: Selected structural data for the rebound transition state (TS_R), QM region R1pro, basis B1. Distances in Å, angles in degree.

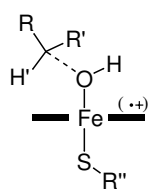
		QM/MM ^a	Gas phase ^b	
TS_R	r_{C-O}	2.401	2.863	
	r_{Fe-O}	1.917	1.852	
	r_{Fe-S}	2.607	2.406	
	r_{O-H}	0.981	0.984	
	θ_{Fe-O-H}	105.3	111.5	
	$\Theta_{C-O-Fe-H}$	-103.3	-171.5	

^a B3LYP/CHARMM optimization in the enzyme environment. ^b B3LYP optimization in the gas phase.

Starting optimizations from the highest energy point of the scan (figure 4.11b), we have successfully located the transition state for rebound on the HS (4A) surface. This species was fully

optimized using QM region R1 and basis set B1. The calculated transition mode with a negative eigenvalue of $92i\text{ cm}^{-1}$ corresponds to a coupled motion of the O atom approaching C^5 , a pyramidalization at the carbon radical center C^5 , and a rotation of the OH group around the Fe–O axis. A plot of this mode is shown in figure 4.9. The QM/MM barrier at the R1pro/B1 level is 9.9 kcal/mol with respect to the corresponding hydroxo intermediate **8** (Fe^{III} electromer, HS state), and is clearly dominated by the QM contributions (9.08 kcal/mol). The transition state thus lies 24.9 kcal/mol (QM/MM) above the reactive complex (**7**) and hence is the point of maximum potential energy in the overall HS rebound pathway. This has obvious consequences for the stability of the intermediate radicals on the HS surface *vs.* those on the LS surface, as shall be discussed in detail at the end of this chapter.

Table 4.13: Spin densities on selected atoms for the rebound transition state (TS_R), QM region R1pro, basis B1.

		QM/MM ^a	Gas phase ^b	
TS_R	$\rho(\text{O})$	-0.088	0.146	
	$\rho(\text{Fe})$	1.433	2.092	
	$\rho(\text{S})$	0.222	0.022	
	$\rho_{\text{gr}}(\text{por})^c$	0.689	-0.135	
	$\rho(\text{C})$	0.803	0.924	
	$\rho(\text{H})$	0.009	0.009	

^a B3LYP/CHARMM optimization in the enzyme environment. ^b B3LYP optimization in the gas phase. ^c $\rho_{\text{gr}}(\text{por})$ denotes the sum of individual atomic spin densities on the atoms of the porphyrin macrocycle.

The relevant structural features of TS_R are summarized in table 4.12, while table 4.13 gives unpaired spin densities. It is interesting to note that the bonds to the axial ligands are significantly elongated in TS_R as compared to **8**; the Fe–O and Fe–S bond lengths in these species are 1.917 \AA *vs.* 1.818 \AA (Fe–O) and 2.607 \AA *vs.* 2.495 \AA (Fe–S), respectively. This elongation is indicative of the occupation of the iron d_{z^2} orbital, which makes the dominant contribution to the antibonding σ^* orbital lying along the O–Fe–S axis. As can be seen from the spin densities in table 4.13, the transition state is of Fe^{III} nature. Hence, this species is electronically equivalent to the lowest hydroxo complex and there is a smooth electronic transition between these species along the lowest energy reactive pathway.

In contrast to the relatively large rebound barrier obtained in the protein environment, single point calculations at the protein geometry in the gas phase ($\text{S}_{\text{p,g}}$) yield a significantly smaller value of 3.63 kcal/mol with respect to the lowest HS hydroxo species **8**. This shows that there is a large electronic effect of the protein environment on the transition state, and consequently, that the potential energy surface might be fundamentally different in the gas phase. Indeed, all attempts to

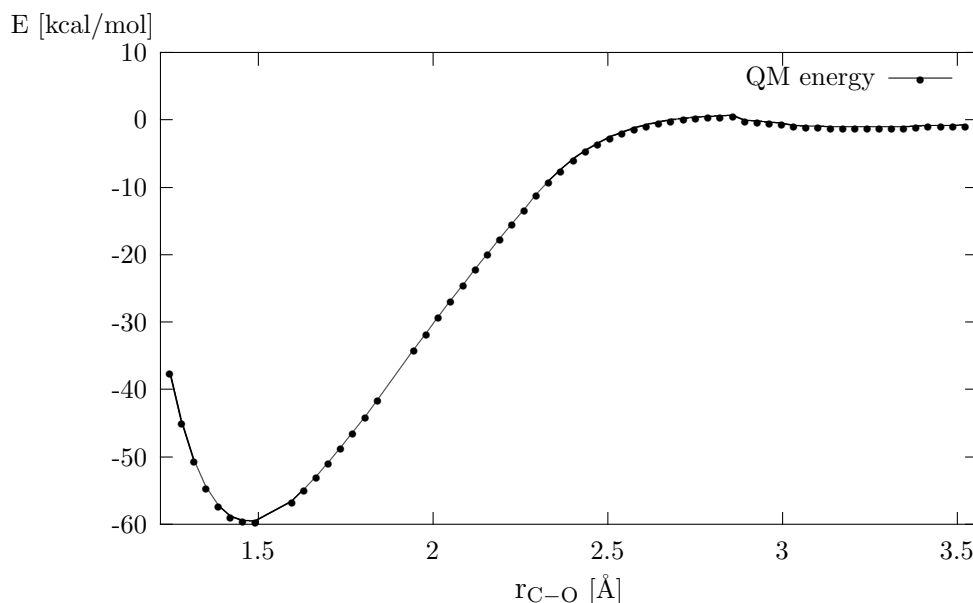


Figure 4.10: PES scan of the coordinate $r(\text{C}^5 - \text{O})$ in the gas phase, B3LYP, QM region R1pro basis set B1. Reference point is the energy at $r(\text{C}^5 - \text{O}) = 3.7 \text{ \AA}$.

directly optimize the transition state in the gas phase starting from this geometry failed, indicating that the isolated QM system is rather far away from the transition state region. Previous DFT studies on different model compounds have characterized the rebound transition state in the gas phase [25, 65, 211]. In the most recent of these studies [25], Kamachi and Yoshizawa investigated a gas phase model including full camphor and a methylthiolate ligated iron-porphyrin. They found a barrier of 3.3 kcal/mol for this process on the HS surface, and an optimized C–O bond length of 2.555 Å. Because this is rather different from the value of 2.401 Å obtained in the present QM/MM calculations, we have carried out a potential energy scan in the gas phase at the B3LYP level with basis set B1. To this end, the distance $r(\text{C}^5\text{--O})$ was constrained at values ranging from 3.7 Å to 1.2 Å, while all other internal degrees of freedom were fully relaxed. The resulting energy profile is shown in figure 4.10. In contrast to previous gas phase studies, the present calculations do not predict a barrier in the range $2.2 \text{ \AA} < r(\text{C}^5\text{--O}) < 2.6 \text{ \AA}$. Instead, consecutive transition state searches located a saddle point at $r(\text{C}^5\text{--O}) = 2.863 \text{ \AA}$, with a low imaginary frequency ($40i \text{ cm}^{-1}$), indicative of a very flat potential energy surface at this point. This transition state is 3.72 kcal/mol above the Fe^{IV} hydroxo intermediate **8**, discussed above. However, by further increasing the distance $r(\text{C}^5\text{--O})$ from this point, we arrived at another conformation of the hydroxo species, which is only 1.68 kcal/mol below the transition state. This conformation is typified by a dihedral angle $\Theta_{\text{C--O--Fe--H}}$ of nearly 180 degrees, which shows that the relative orientation of the attacking carbon radical and the hydroxo group at the iron-complex is very different from that found in the gas phase for the hydroxo intermediate ($\Theta_{\text{C--O--Fe--H}} = -12$ degrees), as well as from that of the TS structure optimized in the enzyme environment by QM/MM calculations ($\Theta_{\text{C--O--Fe--H}} = -103$

degrees). Similar observations were made by Ogliaro *et al.* [65] in gas phase model calculations using methane as substrate for hydroxylation: the transition state for rebound found in this study does not directly link the product with the energetically lowest hydroxo intermediate, but with an energetically less favourable conformation (“rebound cluster”), in which the dihedral angle $\Theta_{\text{C-O-Fe-H}}$ is 180 degrees.

The transition state found in the present study in the gas phase is qualitatively different from the species obtained in the enzyme environment by QM/MM optimizations, which confirms that the protein exerts a strong influence on the conformation and energetics along the rebound pathway in our calculations. The transition structure in the enzyme is more compact, with a distance $r(\text{C}^5\text{-O})$ of 2.401 Å, as opposed to 2.863 Å in the gas phase. Importantly, the electronic structures are also different, as can be seen from the spin densities in table 4.13: the transition state is a Fe^{IV} species in the gas phase and a Fe^{III} species in the protein environment. In summary, the gas phase calculations at the R1pro/B1 level (using propane as the substrate) give a transition structure that is energetically, electronically, and structurally different from the QM/MM result, which emphasizes again the advantage of going beyond isolated molecule calculations. As a note of caution, it is stressed that we cannot strictly rule out the presence of another rebound transition state in the gas phase, which might be more similar to the QM/MM situation, but was not found by the procedure outlined above. Future investigations will have to employ two-dimensional PES scans, in order to provide a definitive characterization of the possible rebound pathways in the gas phase.

Further QM/MM studies could address the significance of the Fe^{IV} electromer to the rebound process in the enzyme environment. Although the corresponding reactant state (Fe^{IV} hydroxo intermediate **8**) was found to lie 1 – 3 kcal/mol higher in energy than the Fe^{III} state (see table 4.6) and is consequently much less populated, the overall barrier on the Fe^{IV} potential energy surface could still be lower, thereby giving rise to a competitive rebound pathway.

4.3.4 The Product Complex

(Collaboration of Dr. Hai Lin and J.C.S who contributed equally to this work.)

The product complex obtained in the catalytic transformation contains 5-*exo*-hydroxycamphor coordinated to the heme iron atom (**9**). X-ray [212] and EPR [13, 213] data for this system are available. UV-VIS and EPR measurements [212, 213], indicate that a mixture of LS and HS species is present, with the LS species being predominant. According to detailed low temperature ENDOR studies at 200 K [13] product formation occurs through three distinct hydroxycamphor-bound product conformations. The authors suggest that the immediate product after hydroxylation involves a more or less normal Fe–O distance (≈ 2 Å) and that the heme pocket then relaxes in two detectable steps to accommodate the longer Fe–O bond of 2.67 Å, that is found in the crystal

structure [212]. The long Fe–O bond in the equilibrium state is thought to be a compromise between the favourable Fe–O interactions on one hand, and the unfavourable desolvation of the polar hydroxy group in the nonpolar protein pocket on the other hand.

We have performed geometry optimizations both for the full system (QM/MM) and the isolated species (QM) starting from different structures:

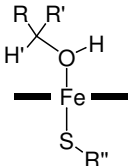
1. from the optimized geometries obtained in the scan of the rebound step of the LS state (as discussed above) at the R1pro/B1 and the R3cam/B4 level.
2. from the X-ray structure reported by Poulos *et al.* [212] at the R1pro/B1 and the R3cam/B1 level.

Starting from the Rebound Scan

In line with the experimental findings [13], free geometry optimizations at the R1pro/B1 level starting from the last steps of the rebound scan on the LS surface lead to a structure with a relatively short Fe–O distance of 2.26 Å, that is identified as the product hydroxycamphor complex. This species is denoted **9-LS1**. On the HS surface, the Fe–O interaction is repulsive, because an unpaired electron occupies a σ^* molecular orbital of the associated bond. Hence, the conformation with the short Fe–O distance **9-LS1** is unstable on the HS surface and geometry optimizations for this state lead to another minimum conformation **9-HS**, which exhibits a significantly longer Fe–O distance of 2.843 Å. In this species, the Fe–O bond is essentially broken, leaving a penta-coordinated iron center. Camphor is still loosely attached to the heme *via* a hydrogen bond from the hydroxy proton to one of the pyrrole nitrogens of the porphyrine macrocycle. Starting from this geometry, we have located a corresponding minimum **9-LS2** on the LS surface with similar geometrical features, most notably also with a broken Fe–O bond (the Fe–O distance is 2.819 Å). Of these three structures, the HS minimum **9-HS** has the lowest relative QM/MM energy; it is approximately 5 kcal/mol more stable than the two LS species that are nearly isoenergetic (see table 4.14). It is well known that the substrate complex (camphor·P450_{cam}) with a pentacoordinated iron atom also has a HS ground state, whereas the LS state is preferred in hexacoordinated iron systems as in the resting state, an aquaferriheme complex.

We have re-optimized the structures obtained at the R1pro/B1 level with the larger QM region R3cam and basis set B4. These relative energies are also given in table 4.14 for the complexes **9-HS**, **9-LS1**, and **9-LS2** and include relative QM/MM energies, QM contributions to the QM/MM energies, QM energies in the gas phase at the protein geometry, and QM energies after optimization in the gas phase. The key geometric features of these systems are summarized in table 4.15, while spin densities are compared in table 4.16.

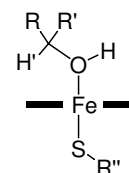
Table 4.14: Relative energies [kcal/mol] of product complexes **9** obtained from the PES scans^a.

		R1pro/B1	R3cam/B4	
9-HS	$E^{\text{QM/MM}}(\text{p}, \text{p})$ ^b	0	0	
	$E^{\text{QM}}(\text{p}, \text{p})$ ^c	0	0	
	$E^{\text{QM}}(\text{p}, \text{g})$ ^d	0	-	
	$E^{\text{QM}}(\text{g}, \text{g})$ ^e	0	-	
9-LS1	$E^{\text{QM/MM}}(\text{p}, \text{p})$ ^b	4.87	6.38 ^f	
	$E^{\text{QM}}(\text{p}, \text{p})$ ^c	4.96	7.92 ^f	
	$E^{\text{QM}}(\text{p}, \text{g})$ ^d	-0.57	-	
	$E^{\text{QM}}(\text{g}, \text{g})$ ^e	-5.16	-	
9-LS2	$E^{\text{QM/MM}}(\text{p}, \text{p})$ ^b	5.18	2.65	
	$E^{\text{QM}}(\text{p}, \text{p})$ ^c	8.11	5.13	
	$E^{\text{QM}}(\text{p}, \text{g})$ ^d	-	-	
	$E^{\text{QM}}(\text{g}, \text{g})$ ^e	-	-	

^a **9-HS**: the product complex in the HS (⁴A) state; **9-LS1**: the product complex in the LS (²A) state as obtained when starting optimizations from the rebound PES scans; **9-LS2**: the product complex in the LS state as obtained when starting optimizations from the HS minimum (**9-HS**). ^b QM/MM energy difference. ^c QM contribution to the QM/MM energy difference. ^d Energy differences of the isolated species in the gas phase at the protein geometry. ^e Energy differences of the isolated species optimized in the gas phase. ^f Constrained minimization, Fe–O distance fixed to 2.262 Å.

Table 4.15: Selected structural data for the product complexes **9** obtained from the PES scans^a. Distances in Å, angles in degree.

		QM/MM		Gas phase	X-ray ^b
		R1pro/B1	R3cam/B4	R1pro/B1	
9-HS	r _{Fe–O}	2.843	3.198	2.450	2.68
	r _{Fe–S}	2.525	2.467	2.489	2.11
	r _{O–H}	0.977	0.965	0.979	
	r _{C–O}	1.482	1.433	1.475	
	θ _{Fe–O–H}	83.7	92.0	106.2	
	D _{C–O–Fe–H}	-113.7	-122.8	-151.0	
	r _{O:Tyr96–O:Cam}	2.73	2.68	-	2.72
9-LS1	r _{Fe–O}	2.262	2.262 ^c	2.086	2.68
	r _{Fe–S}	2.392	2.310 ^c	2.297	2.11
	r _{O–H}	0.976	0.967 ^c	0.976	
	r _{C–O}	1.499	1.461 ^c	1.482	
	θ _{Fe–O–H}	97.5	104.0 ^c	112.5	
	D _{C–O–Fe–H}	-130.4	-137.3 ^c	-162.8	
	r _{O:Tyr96–O:Cam}	2.83	2.81 ^c	-	2.72
9-LS2	r _{Fe–O}	2.819	3.211	-	2.68
	r _{Fe–S}	2.371	2.277	-	2.11
	r _{O–H}	0.981	0.965	-	
	r _{C–O}	1.484	1.434	-	
	θ _{Fe–O–H}	68.1	91.2	-	
	D _{C–O–Fe–H}	-96.7	-122.0	-	
	r _{O:Tyr96–O:Cam}	2.74	2.68	-	2.72

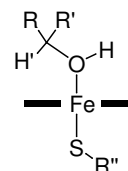


^a **9-HS**: the product complex in the HS (⁴A) state; **9-LS1**: the product complex in the LS (²A) state as obtained when starting optimizations from the rebound PES scans; **9-LS2**: the product complex in the LS state as obtained when starting optimizations from the HS minimum (**9-HS**). ^b Taken from ref. [212], spin state uncertain. ^c Constrained minimization, Fe–O distance fixed to 2.262 Å.

Regarding the two LS species **9**-LS1 and **9**-LS2, it is instructive to analyze the contributions to the total QM/MM energy (see table 4.14). In the case of the QM system R1pro, the loss of the Fe–O binding interaction in **9**-LS2 accounts for a destabilization of ≈ 3 kcal/mol relative to **9**-LS1 in the QM contributions. This is, however, compensated by a stabilization of approximately the same magnitude due to the MM terms, which could be a result of more favourable interactions of camphor with the pocket. Candidates for these energy contributions are dispersive interactions of the methyl groups with residues Phe87, Val295, Ile395 and Val396 and the hydrogen bond of the camphor carbonyl group with Tyr96. In fact, the distance R(O–O) between the camphor carbonyl oxygen and the hydroxo oxygen of Tyr96 is significantly shortened (by ≈ 0.1 Å) in **9**-LS2 (see also table 4.15).

Table 4.16: Spin densities on selected atoms for the product complex **9** obtained from the PES scans^a.

		QM/MM		Gas phase
		R1pro/B1	R3cam/B4	R1pro/B1
9 -HS	$\rho(\text{O})$	0.010	-0.018	0.016
	$\rho(\text{Fe})$	2.674	2.740	2.578
	$\rho(\text{S})$	0.320	0.350	0.447
	$\rho_{\text{gr}}(\text{por})^b$	0.001	-0.080	-0.035
9 -LS1	$\rho(\text{O})$	-0.003	-0.014 ^c	0.006
	$\rho(\text{Fe})$	1.254	1.297 ^c	1.102
	$\rho(\text{S})$	-0.158	-0.136 ^c	-0.017
	$\rho_{\text{gr}}(\text{por})^b$	-0.102	-0.148 ^c	-0.091
9 -LS2	$\rho(\text{O})$	-0.001	-0.007	-
	$\rho(\text{Fe})$	1.377	1.403	-
	$\rho(\text{S})$	-0.201	-0.167	-
	$\rho_{\text{gr}}(\text{por})^b$	-0.185	-0.232	-



^a **9**-HS: the product complex in the HS (⁴A) state; **9**-LS1: the product complex in the LS (²A) state as obtained when starting optimizations from the rebound PES scans; **9**-LS2: the product complex in the LS state as obtained when starting optimizations from the HS minimum (**9**-HS). ^b The sum of individual atomic spin densities on the atoms of the porphyrin macrocycle. ^c Constrained minimization, Fe–O distance fixed to 2.262 Å.

The QM/MM calculations with the larger QM region also predict the global minimum to be the quartet species, **9**-HS. However, the stability of the dissociated LS complex, **9**-LS2, is enhanced in comparison to the R1pro/B1 calculations. The relative QM/MM energies of this complex with respect to **9**-HS is +3 kcal/mol (R1pro/B1: +5 kcal/mol). In contrast, the other doublet species, **9**-LS1, is found to be unstable: Geometry optimizations of this system lead to a minimum with a broken Fe–O bond, corresponding to **9**-LS2. It was only possible to compute **9**-LS1 by imposing

a distance constraint on the Fe–O bond (2.26 Å, taken from the R1pro/B1 optimization). This species is calculated to lie +6 kcal/mol above **9**-HS (R1pro/B1: +5 kcal/mol).

The energy data presented in table 4.14 show that the differences between R1pro/B1 and R3cam/B4 results are mainly due to changes in the QM energy contributions: The QM energy of **9**-LS1 relative to the global quartet minimum is +8 kcal/mol (R1pro/B1: +5 kcal/mol), while **9**-LS2 is +5 kcal/mol higher in energy (R1pro/B1: +8 kcal/mol). As mentioned above, there are stabilizing interactions between camphor and neighboring protein residues in the pocket, which are intensified in the dissociated product complex. In the case of QM system R1pro, which incorporates only a small part of camphor, this is mirrored by the MM contributions, which stabilize **9**-LS2 relative to **9**-LS1. On the other hand, the larger QM region R3cam includes the whole camphor molecule, thus the net energy gain upon dissociation is now apparent from the QM contributions. This QM stabilization energy apparently over-compensates the loss of Fe–O binding energy, thus the dissociated product is now clearly preferred. We attribute this stabilization mainly to electrostatic interactions between camphor and MM atoms of the environment. The most prominent interaction of this type is certainly the Tyr96–camphor hydrogen bond; indeed the shortest distances $r_{\text{O:Tyr96}-\text{O:Cam}}$ of 2.68 Å are found in the dissociated complexes at the R3cam/B4 level. More detailed studies on the product dissociation will be presented in the following section.

A comparison of the QM/MM results with the QM results obtained at the R1pro/B1 level for the isolated gas phase system reveals that the enzyme environment has a distinct effect on the HS and LS product species. In accord with previous DFT model studies [65], the relative energies presented in table 4.14 show that the LS complex **9**-LS1 is the lowest minimum in the gas phase. It is calculated to be more stable than the HS minimum by 5 kcal/mol. Hence, the ordering of spin states is reversed in the gas phase as compared to the situation in the enzyme environment. One explanation for this finding is offered by the geometry data given in table 4.15. The optimized Fe–O bond distances in the gas phase of (**9**-HS/**9**-LS1) 2.450/2.086 Å are significantly shorter than in the protein (2.843/2.262 Å) so that the repulsive Fe–O interactions (present only in the ^4A state) cause a significant destabilization of the HS with respect to the LS species in the gas phase. Another factor becomes obvious upon inspection of the calculated relative energies of the species in the gas phase at the geometry optimized in the protein environment ($S_{\text{p,g}}$): in this situation the LS and HS states are approximately isoenergetic, with the LS state 0.57 kcal/mol below the HS state. This demonstrates that the enzyme environment stabilizes the HS species through electronic polarization of the QM part, by as much as 5.5 kcal/mol.

Starting from the X-ray conformation

The QM/MM calculations presented above indicate that the global minimum of the product complex **9** corresponds to the HS structure **9**-HS. However, experimental studies [212, 213] suggest

that **9** is predominantly in the LS state. Due to the high dimensionality of the PES in the enzyme environment, we cannot strictly rule out the existence of another, global LS minimum. To check the state ordering obtained from the QM/MM calculations discussed above, we replaced⁴ the coordinates of the camphor molecule and the oxo atom in the simulated structure (snap 40) with the coordinates of 5-*exo*-hydroxocamphor taken from the crystal structure (PDB code: 1NOO) [212]. The relative orientation of the substrate with respect to the porphyrin plane changes with the replacement, however the hydrogen-bond between the substrate and the Tyr96 residue is retained. The resulting structure of the product complex was optimized in both gas phase (QM) and protein/solvent environment (QM/MM). Only the B1 basis set was employed in combination with QM regions R1pro and R3cam.

Table 4.17 lists the results obtained from geometry optimizations with QM region R1pro. The LS minimum found in this study exhibits a Fe–O distance of 2.212 Å in the enzyme environment, in close agreement with the structure **9**-LS1 ($r_{\text{Fe-O}} = 2.262$ Å). Likewise, the HS minimum is characterized by a long Fe–O distance of 2.753 Å, indicative of an essentially broken Fe–O bond. The relative energies at the QM/MM level predict the quartet minimum to lie 4 kcal/mol below the doublet species, again in agreement with the energy separation of 5 kcal/mol discussed above. This order is changed if the isolated gas phase species are considered: single point calculations in the gas phase at the protein geometry ($S_{\text{p,g}}$) favour the LS by 0.8 kcal/mol over the HS structure. Optimizing the systems in the gas phase leads to a gap of 3 kcal/mol with the LS state being the ground state. Qualitatively the same results are obtained from geometry optimizations employing the larger QM region R3cam [214].

Table 4.17: Selected structural data and relative energies of the product complexes **9** obtained from optimizations starting from the X-ray conformation^a. Distances in Å, angles in degree.

QM/MM optimized geometries ($S_{\text{p,p}}$)			Relative energies		
Parameter	HS	LS	Energy	HS	LS
$r_{\text{Fe-O}}$	2.753	2.212	$E^{\text{QM/MM}}(\text{p}, \text{p})^a$	0	3.57
$r_{\text{Fe-S}}$	2.527	2.388	$E^{\text{QM}}(\text{p}, \text{p})^b$	0	4.31
$r_{\text{C-O}}$	1.483	1.506	$E^{\text{QM}}(\text{p}, \text{g})^c$	0	-0.83
$\theta_{\text{Fe-O-H}}$	86.9	99.7	$E^{\text{QM}}(\text{g}, \text{g})^d$	0	-3.20

^a QM region R1pro, basis set B1; initial coordinates of heme and camphor taken from the X-ray structure with PDB code 1NOO [212].

⁴ The structures were aligned using a least squares fit, taking five atoms in the porphyrine (Fe, NA, NB, NC, and ND) as reference.

Conclusion

In summary, the QM/MM calculations on the product complex **9** indicate that the system reaches the product state on the LS surface after the nearly barrierless rebound step. The initial product state corresponds to a species with a relatively short Fe–O distance (**9**-LS1). This minimum is, however, not a thermodynamic sink: another structure on the LS surface is isoenergetic or lower in energy (**9**-LS2), where the hydroxycamphor is detached from the heme, leaving a pentacoordinated iron. The loss of the Fe–O binding energy is compensated by more favourable interactions of hydroxycamphor with the pocket, *e.g.*, a stronger hydrogen bond between Tyr96 and the carbonyl group of the substrate. As expected for a pentacoordinated iron system, the HS state is energetically lower at this conformation (**9**-HS). Moreover, as shown by the comparative gas phase and QM/MM calculations, the polarizing effect of the enzyme environment stabilizes the HS with respect to the LS state by polarization of the electron density in the QM calculation. The QM/MM optimizations did not locate a minimum on the LS surface that has a lower energy than the dissociated product complex **9**-HS, even when starting from a conformation derived from the experimental structure of the product complex [212].

4.3.5 Product Release

(Collaboration of Dr. Hai Lin and J.C.S who contributed equally to this work.)

The intriguing differences in the geometries and relative stabilities of the product complex **9** depending on spin state and environment (gas phase vs. enzyme) prompted us to further investigate this issue. Specifically, the following points were addressed:

1. The QM/MM calculations presented in section 4.3.4 indicate that the global minimum of **9** corresponds to the 4A state at a long distance $r(\text{Fe}-\text{O})$. The doublet state exhibits a minimum with a comparably short Fe–O bond. At what distance $r(\text{Fe}-\text{O})$ is the LS (doublet) state the ground state, and where is the crossing point of the two spin surfaces?
2. What is the energy cost associated with dissociation of the Fe–O bond, yielding the free hydroxylated product?
3. At large Fe–O distances, the sextet state (6A) is expected to be comparably low in energy. What is the relative energy of the sextet state as compared to the doublet and quartet states?

To answer these questions, PES scans of the internal coordinate $r(\text{Fe}-\text{O})$ were carried out for complex **9** both in the enzyme and the gas phase. These were done separately in the 2A , 4A , and 6A states, respectively. The calculations employed QM region R1 and basis set B1, exclusively.

Table 4.18 summarizes the relative QM/MM energies $E(\text{QM/MM})$, and the QM contributions to

the latter, E(QM), obtained from the scans in the enzyme environment. The potential energy profiles for the gas phase scans are given in table 4.19. Plots of the energy profiles obtained from the scans are depicted in figures 4.11 and 4.12. The reference point in each case is the energy of the minimum in the 2A state. In analogy to section 4.3.4, the starting geometry for the scan was obtained by replacing the oxygen atom and camphor atoms in the QM/MM optimized structure (snapshot 40) by the corresponding atoms from the X-ray structure [212] after a superimposition of these two structures. Based on the scanned energy profiles, the Fe–O distances at the crossing points (entries 2 for 2A , 4A and 6 for 2A , 6A) were estimated and constrained geometry optimizations at these estimated distances were carried out to confirm the degeneracy of the spin states.

Table 4.18: QM/MM PES scan of the coordinate $r(\text{Fe–O})$ in the 2A , 4A , and 6A states in the enzyme environment. Distance in Å; QM/MM energy E(QM/MM) and QM contributions E(QM) in kcal/mol.

entry	$r_{\text{Fe–O}}$	E(QM/MM)			E(QM)		
		2A	4A	6A	2A	4A	6A
1	2.084	0.42	1.90	5.45	-0.88	-2.24	3.96
2	2.155 ^a	0.12	0.04	-	-0.66	-3.90	-
3	2.235	0.06	-1.41	2.38	-0.49	-4.37	1.97
4	2.249 ^b	0.00	-	-	0.00	-	-
5	2.385	0.48	-3.11	0.98	0.77	-5.66	1.11
6	2.445 ^c	0.72	-	0.70	1.24	-	1.06
7	2.535	1.57	-3.74	0.84	2.60	-5.48	1.68
8	2.685	1.90	-3.97	0.54	3.26	-5.31	1.86
9	2.753 ^d	-	-3.97	-	-	-5.32	-
10	2.832 ^e	-	-	0.08	-	-	1.54
11	2.835	2.01	-3.89	0.39	3.90	-4.98	1.81
12	2.984	2.05	-3.83	0.31	4.17	-4.79	1.97
13	3.134	2.45	-3.58	0.22	4.44	-4.30	2.24

^a The crossing point of 2A and 4A QM/MM energy surfaces. ^b From an unconstrained optimization in the 2A state.

^c The crossing point of 2A and 6A QM/MM energy surfaces. ^d From an unconstrained optimization in the 4A state.

^e From an unconstrained optimization in the 6A state.

The minimum structures in the respective spin states were obtained by unconstrained geometry optimizations from the points of lowest energy along the PES scan. Consistent with the studies discussed in section 4.3.4, we locate a doublet ground state in the gas phase, with the lowest quartet and sextet minimum being 3 and 5 kcal/mol higher in energy, respectively. In contrast, the QM/MM optimizations predict a quartet ground state minimum that is *ca.* 4 kcal/mol lower than the doublet equilibrium structure. The $S = 5/2$ state minimum resulting from QM/MM geometry

optimization has a similar energy as the doublet species, the energy difference being only 0.1 kcal/mol. The sextet equilibrium structure is typified by a long Fe–O distance of 2.974 Å in the QM/MM, and 2.712 Å in the gas phase QM optimized geometries, respectively.

Table 4.19: QM PES scan of the coordinate $r(\text{Fe–O})$ in the ^2A , ^4A , and ^6A states in the gas phase. Distance in Å; QM energy $E(\text{QM})$ in kcal/mol.

entry	$r_{\text{Fe–O}}$	$E(\text{QM})$		
		^2A	^4A	^6A
1	1.935	1.47	15.13	15.22
2	2.085	0.00	8.83	9.46
3	2.235	1.61	6.00	8.54
4	2.385	2.41	5.90	7.55
5	2.535	3.56	5.61	7.57
6	2.685	4.45	5.27	7.22
7	2.835	5.32	5.34	7.12
8	2.984	5.82	5.41	7.07
9	3.134	6.33	5.48	7.00

In the gas phase scan, the sextet state has the highest energy over the whole range of distances $r(\text{Fe–O})$ considered (1.9–3.1 Å). The doublet state is the ground state for $r(\text{Fe–O}) < 2.9$ Å. At short distances $r(\text{Fe–O})$, the quartet and sextet states are higher in energy, by as much as 9 kcal/mol at the doublet equilibrium Fe–O distance. For $r(\text{Fe–O})$ close to the experimental value (2.7 Å), the $S = 1/2$ state is still 1 and 3 kcal/mol lower than the $S = 3/2$ and $S = 5/2$ states, respectively. The crossing of the doublet and quartet surfaces occurs at a distance $r(\text{Fe–O})$ of approximately 2.905 Å (0.19 kcal/mol difference in energy).

The potential energy profiles change significantly upon going from the isolated gas phase species to the QM/MM system. Most notably, the protein environment stabilizes the ^4A state considerably. Hence, this state becomes the ground state for $r(\text{Fe–O}) > 2.16$ Å and it is significantly (4–6 kcal/mol) lower in energy than the ^2A state for $r(\text{Fe–O}) > 2.4$ Å. Due to the stabilization of the ^4A state, the crossing of doublet and quartet surfaces occurs at a much shorter distance $r(\text{Fe–O})$ of 2.16 Å – as compared to 2.9 Å in the gas phase. The crossing point of the ^2A and ^6A states is obtained at $r(\text{Fe–O}) = 2.45$ Å, while this crossing is not observed in the gas phase calculations in the range scanned for $r(\text{Fe–O})$. It is interesting to note that the potential energy minimum of the ^2A state, entry 4 in table 4.18, is close to the crossing point of doublet and quartet surfaces, both in energetic ($\Delta E < \approx 0.2$ kcal/mol) and geometric ($\Delta r(\text{Fe–O}) \approx 0.06$ Å) terms. The energy of the doublet state increases from the equilibrium structure to 2 kcal/mol at $r(\text{Fe–O}) = 3.134$ Å, hence the dissociation of the hydroxylated substrate is facile. Consistent with the observations discussed

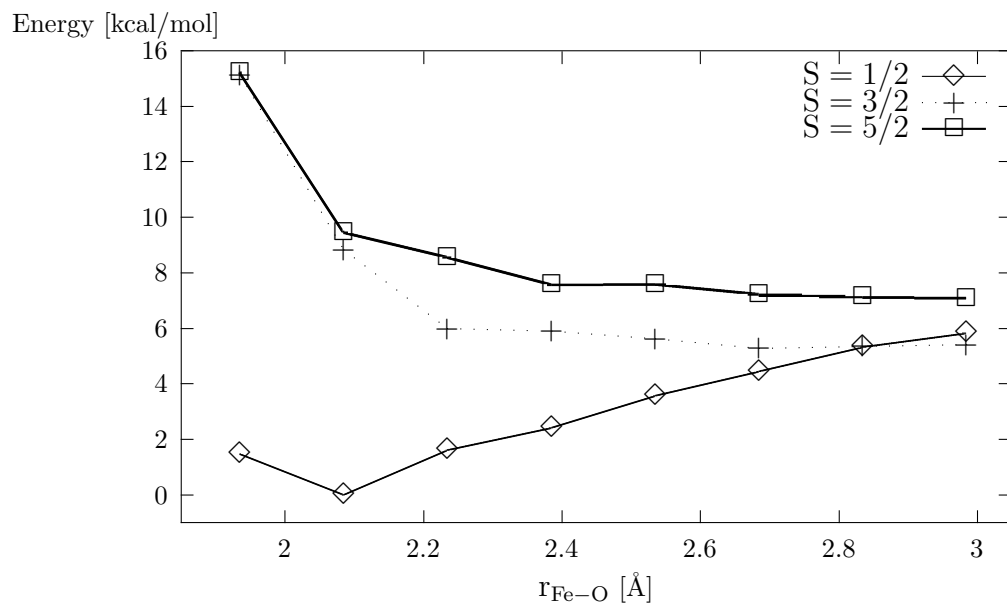


Figure 4.11: QM PES scan of the coordinate $r(\text{Fe-O})$ of the product complex **9** in different spin states in the gas phase (QM region R1pro, basis B1).

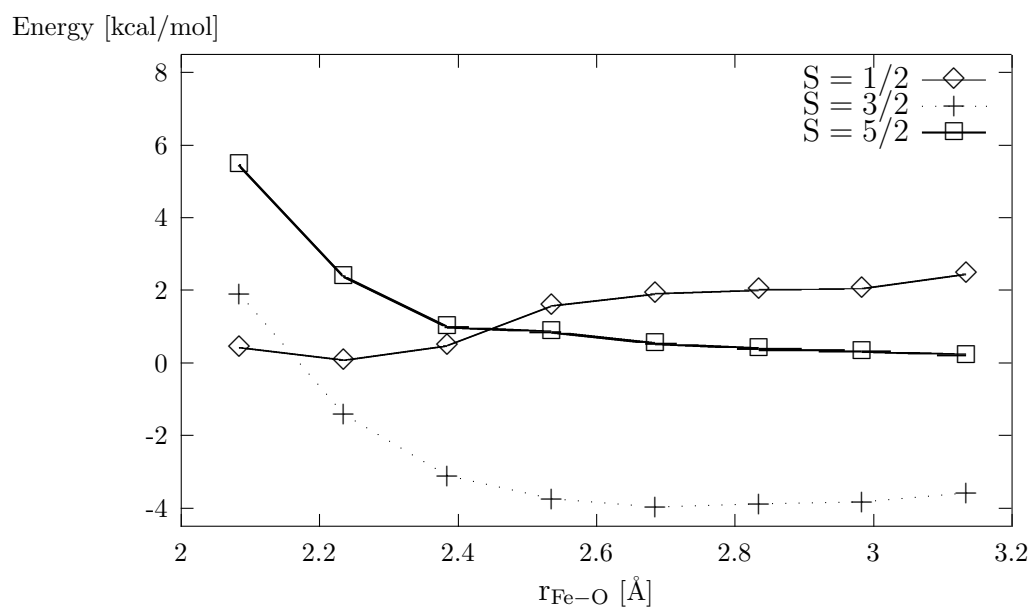


Figure 4.12: QM/MM PES scan of the coordinate $r(\text{Fe-O})$ of the product complex **9** in different spin states in the enzyme environment (QM region R1pro, basis B1).

in section 4.3.4, the QM contributions to the QM/MM energy account for the higher energy of the dissociation products (4 kcal/mol), which is mainly attributed to the loss of Fe–O binding energy. However, the MM terms that reflect the interactions of the substrate with the protein pocket partly compensate for this destabilizing effect.

4.4 Discussion

The QM/MM (R1pro/B1) calculated overall energy profile of P450_{cam} mediated camphor hydroxylation *via* the hydrogen-abstraction, oxygen-rebound mechanism is shown in figure 4.13. The computed barrier for hydrogen abstraction is approximately equally high in the LS (doublet) and HS (quartet) states, with a slightly lower barrier for the LS state. After this initial step, the energy profiles bifurcate. On the LS surface, the rebound step is predicted to proceed with a very low barrier (< 2 kcal/mol), whereas the HS process exhibits a significant barrier and a genuine transition state. These findings are consistent with the involvement of a two-state reactivity (TSR) scenario as proposed by Shaik *et al.* on the basis of DFT computations on model compounds [64–68]. Hence, the HS reaction is truly stepwise with a finite lifetime of the substrate radical intermediate which is loosely bound to the iron-hydroxo complex (**8**). In contrast, the LS mechanism is effectively concerted with an ultrashort lifetime for the intermediate species **8**. A recent DFT study on camphor hydroxylation in the gas phase identified a rebound transition state also for the LS rebound process with a barrier of 0.7 kcal/mol [25]. This small activation energy has been attributed to the rotational barrier of the camphor molecule to a conformation which is best suited for the rebound reaction. In contrast, the steric constraints in the protein pocket prevent such a rotation of the substrate in the present QM/MM calculations. Instead, the OH group of the iron-hydroxo complex has to rotate around the Fe–O axis, such that the C–O interaction becomes possible.

As discussed previously [65, 71], a TSR scenario, as met in the present case, is capable to explain the results of radical clock experiments by Newcomb and coworkers [10, 11]. As detailed in section 1.2.4, these measurements seem to exclude the presence of a discrete radical intermediate and have been interpreted in terms of a nonsynchronous concerted oxygen insertion pathway [63]. The lifetime of the radical intermediate was derived from the inverse of the rate constant of radical rearrangement (k_r) and the ration $[R/U]$ of rearranged (**R**) to unrearranged (**U**) alcohol products (see section 1.2.4), assuming single-state reactivity (SSR). However, in a two-state reactivity scenario, the rearranged product **R** is formed only on the HS surface, while the unrearranged product is formed on the LS, and, to less extent, also on the HS surface. Hence, the ratio $[R/U]$ is associated with the relative yields of the HS *vs.* the LS reaction, and not with the radical lifetime as such [65, 71]. Presupposing a simple TSR scenario, the following expression may be derived [65] for the ratio of the real lifetime of the radical intermediate τ_{real} and the apparent lifetimes τ_{app}

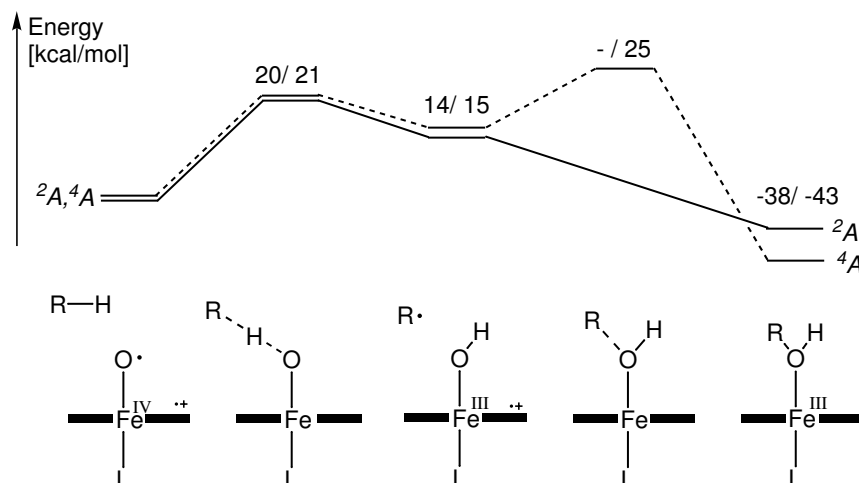


Figure 4.13: The rebound mechanism – B3LYP/CHARMM (R1pro/B1) computed energy profile of the overall reaction.

obtained from experiments assuming SSR:

$$\tau_{\text{real}}/\tau_{\text{app}} = \{[U/R](1 + F)\}/\{[U]/[R] - F\}; \quad F = [LS]/[HS]. \quad (4.1)$$

Here, the quantity F is the relative yield of the LS and the HS reactions. It is thus clear that the real lifetime of the radicals on the HS surface may be longer than the apparent lifetime, depending on F . In the present QM/MM calculations, the bond activation barrier (TS_H) is consistently lower in the LS than in the HS state, and thus, F will be larger than unity. Values for F of the order of 10 were reported in the case of allylic hydroxylation when the effect of the protein electric field and hydrogen bonding were taken into account [36,211]. It is reasonable to assume that the situation is similar for the probes employed in the radical clock experiments. As a consequence, the apparent radical lifetimes will be unrealistically short compared with the real lifetimes. Yet, future QM/MM investigations will have to deal with the radical clock systems themselves, in order to model the actual experimental situation and hence make realistic predictions.

In contrast to the present results, recent DFT studies in the gas phase using CH_3S^- as a model of the proximal ligand and camphor as the substrate [25], predict the barrier for H-abstraction to be lower in the HS state than in the LS state. It thus seems that the CH_3S^- misrepresents the actual situation in the enzyme environment. This finding, much the same as the wrong assignment of a $^2\Pi$ sulfur radical ground state [39,41], is probably due to the strong electron donor property of the CH_3S^- ligand.

The QM/MM barriers for hydrogen abstraction calculated in the present study are 20–22 kcal/mol (see table 4.1), depending on QM region and basis set. From the R1pro/B1 calculations of the isolated QM system in the gas phase we obtained activation energies of 20 kcal/mol. These values are essentially close to the activation energies reported in previous DFT studies in the gas phase for different Cpd I models and similar substrates R (LS/HS: R = CH_4 , 27/27 kcal/mol [65];

$R = C_2H_6$, 16/19 kcal/mol [77]; $R = \text{camphor}$, 20/18 kcal/mol [25]). In close agreement with a previous study on methane hydroxylation [65], we find that inclusion of ZPE reduces the barrier by *ca.* 4 kcal/mol. This effect is due to the loss of the C–H vibrational energy of the bond that is cleaved in the TS_H species. Quantum mechanical tunneling may be expected to further reduce the barrier by up to 3 kcal/mol [209].

Taking into account all these effects and the intrinsic accuracy of the DFT approach, the calculated barriers for hydrogen abstraction may be extrapolated to 13 (± 5) kcal/mol. This value is reasonable regarding the experimental findings of Davydov *et al.*, who could not detect Cpd I in the presence of substrate even at temperatures below 200 K.

Since the barriers are obviously not significantly changed in the gas phase as compared to the enzyme, one might wonder what role the protein environment plays in the present enzymatic reaction. Another feature discussed in previous theoretical studies in the gas phase is a significant entropic contribution to the free energy barrier of H-abstraction. Originating in the loss of translational and rotational degrees of freedom of the substrate in the transition state, this effect accounts for an increase of the barrier of *ca.* 10 kcal/mol in the case of methane hydroxylation [65]. On the other hand, in enzyme catalysis, substrate binding is largely entropically driven, by expulsion of water molecules from the binding pocket. Hence, one role of the protein environment in P450 mediated hydroxylation certainly is to absorb the entropic cost of establishing the transition state into the preceeding step of substrate binding, thus effectively lowering the barrier for H-abstraction. A clear indication for this effect is the stable conformation representing the reactive complex, where the substrate and Cpd I are pre-organized in the pocket at a geometry with a linear arrangement C–H–O, which is ideally suited for H-abstraction.

Apart from lowering the entropic barrier for H-activation, a second important role of the protein environment in the overall reaction is to control the access of water to the active site. It is known that uncoupling of reducing equivalents is enhanced if water has uncontrolled access to the pocket, if the substrate spends much of its time in the active site at unproductive distances or orientations, or if the substrate is difficult to oxidize [2, 56, 215]. The third (and most intriguing) effect of the enzyme environment is the high regio- and stereoselectivity of substrate hydroxylation, which is probably due to selective constraints and preferences of certain modes of substrate movement within the active site.

The hydroxo intermediate **8** exhibits a number of low-lying electronic states (Fe^{III} , Fe^{IV} , HS, LS) that are probably all populated at ambient temperature. The present study demonstrates that the Fe^{III} variety is stabilized by polarization and represents the ground state in the enzyme environment, while the isolated gas phase species are of the Fe^{IV} -type. Recent model studies on camphor hydroxylation found the LS electromer to have Fe^{III} -character, while the HS species was predicted to have the Fe^{IV} occupation [25]. All these states can participate in the rebound reaction.

The consequences of such a multi-state-scenario in terms of the kinetics of the reaction will be a subject of future investigations. As discussed in section 4.3.3, we find a significant barrier for the rebound process on the HS surface (9.85 kcal/mol, R1pro/B1). Another factor revealed by the QM/MM investigations is that the rebound process in the enzyme occurs by OH rotation around the Fe–O bond. Similar observations were made in a QM/MM study on the methane monooxygenase enzyme [216]. This is contrast to model studies in the gas phase, where the transition mode is characterized by rotation of the alkyl group about the Fe–O bond [25, 65, 77, 211]. Clearly, this finding is a manifestation of the selective constraints of the protein pocket with respect to substrate motion. Interestingly, experimental results show that both *exo*- and *endo*-hydrogens are abstracted in camphor hydroxylation, but only the *exo*-product is formed [56]. Hence, the directional effect of the pocket in the rebound step is crucial for the observed stereoselectivity.

The initial LS product alcohol complex formed after the nearly barrierless rebound step is characterized by a relatively short Fe–O bond. This species can undergo Fe–O dissociation and intersystem crossing to give the quartet species. The resulting quartet minimum structure, where 5-*exo*-hydroxocamphor is dissociated from the iron center, is stabilized in the enzyme environment and is calculated at the QM/MM level to lie below all the LS species considered up to now. The factors that cause the stabilization of the quartet (and to a smaller extent the sextet) state with respect to the doublet state in the enzyme environment have been studied in the closely related ferric aquo heme complex **1**, *i.e.*, the resting form of the enzyme. These results are presented and discussed in detail in chapter 5. In short, this study shows that the polarizing effect of the protein/solvent environment causes a stabilization of electronic charge on the sulfur atom, which effectively lowers the energy of the Fe–S σ^* orbital. Since this orbital has to accomodate an unpaired electron in the HS states but not in the LS state, the former are stabilized significantly – in the ferric aqua complex this accounts for a shift of *ca.* 8 kcal/mol with respect to the gas phase.

The work on the resting state **1** also indicates (see chapter 5) that the B3LYP functional, due to the admixture of exact exchange, tends to overestimate the stability of the HS states [217, 218]. Exploratory studies using the BLYP functional (without exact exchange) indeed show that the relative state energies in **9** are changed by several kcal/mol with regard to the B3LYP results, leading to a preference for the LS state. Hence, definite conclusions regarding the multiplicity of the ground state of **9**, and the crossing point of the doublet and quartet surfaces, may be hard to reach at the DFT level. This issue requires further studies on the basis of *ab initio* methods, which can offer a balanced description of exchange and correlation effects. It should be stressed, however, that the qualitative (TSR) picture is not changed by these findings, since the product release is not the rate limiting step of the overall reaction.

During the course of this thesis, QM/MM studies using the same methodology as presented here were carried out in our group [214] with the aim to characterize a concerted oxygen-insertion

pathway [63] for camphor hydroxylation. Specifically, these studies employed constrained geometry optimizations at the R1pro/B1 level. The results indicate that the process has high barriers of approximately 48 and 61 kcal/mol for the LS and HS states, respectively. Moreover, when the number of geometry constraints was reduced, the structures along the path were found to be unstable and optimizations led to geometries not related to the reaction. This study thus indicates that no competitive pathway for the concerted oxygen-insertion pathway is viable in the enzyme environment. This is in accord with previous gas phase model studies, which reached the same conclusion [219, 220].

4.5 Conclusions

On the basis of combined QM/MM calculations, the present study provides a complete picture of the hydrogen-abstraction, oxygen-rebound pathway, by which P450_{cam} mediates the selective hydroxylation of C–H bonds. Our calculations show that this reaction is typified by TSR, as has been suggested earlier on the basis of model computations. As such, this mechanism provides a satisfactory explanation for seemingly contradictory experimental findings, which on one hand suggested the presence of radical intermediates, but on the other hand yielded unreasonably short lifetimes for these intermediates.

Comparative QM calculations of model complexes in the gas phase allowed us to investigate the influence of the protein environment on the catalytic process. We studied a potential mechanism of electrostatic transition state stabilization, that implies the selective stabilization of the transition state for hydrogen abstraction by interactions of the propionate side chains of the heme with neighbouring positively charged protein residues. Our results demonstrate that there is no charge transfer to the propionate groups, and hence no differential stabilization of the transition state.

The results from a QM calibration study show that the height of the barrier for hydrogen abstraction largely depends on the chosen density functional, hence quantitative predictions for this activation energy are currently beyond the scope of DFT methods. However, by estimating the effect of zero point energies and tunneling, we arrive at barriers that seem reasonable with regard to experiments that could not detect Cpd I even at low temperatures (200 K). It is suggested that the protein promotes the reaction by absorbing the entropic cost that is required to establish the transition state for hydrogen abstraction (*ca.* 10 kcal/mol) into the process of substrate binding, which is driven by the expulsion of water molecules from the active site and the desolvation of the hydrophobic substrate.

Our calculations furthermore indicate that substrate motion is hindered in the active site, mainly by a hydrogen bond of the camphor carbonyl oxygen with Tyr96. The rebound transition state that is thought to control the stereoselectivity of camphor hydroxylation involves OH rotation, in

contrast to substrate rotation, as was found in previous model studies.

The comparative QM/MM and QM work further established that the protein environment has an effect on the stability of redox electromers (*i.e.* Fe^{III} vs. Fe^{IV} states, porphyrin cation radical vs. sulfur radical states) and the relative energies of different spin states. Thus, it was shown that the stability of the quartet state in the product complex is largely increased by the polarizing enzyme environment.

Chapter 5

The Resting State

5.1 Motivation and Background

The ferric resting state **1** is one of the most thoroughly studied intermediates in the catalytic cycle of P450 enzymes. As discussed in section 1.2.5, the presence of an axial water ligand in this species was verified by X-ray diffraction [85], and ENDOR/ESEEM studies [86, 87]. The active site environment with the cluster of six water ligands is depicted in figure 5.1.

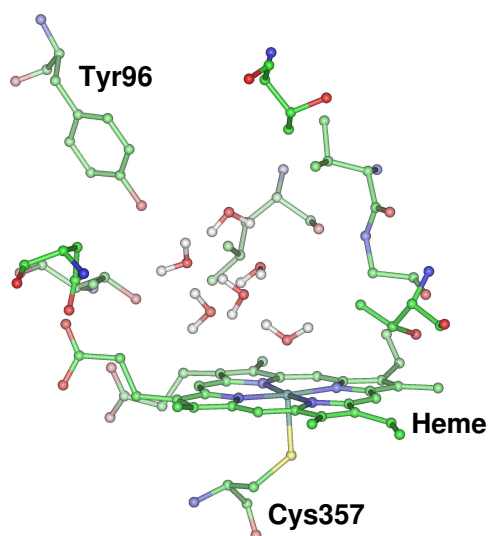


Figure 5.1: Active site of the ferric resting form (**1**) of P450_{cam}. The six water molecules filling the pocket are shown.

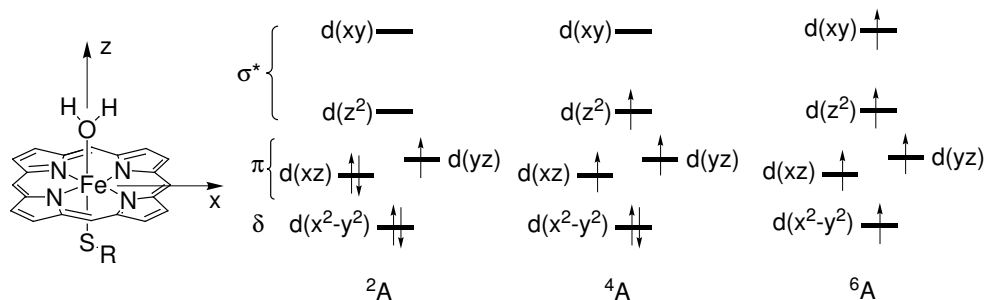


Figure 5.2: Electronic configurations of the ferric ion in **1**.

The ESEEM measurement of the resting form of cytochrome P450_{cam} [87] determined the hyperfine interactions with the water protons of the axial ligand, and allowed an estimation for the distance of these protons to the iron atom of 2.62 Å. In the X-ray structure, the Fe–O distance refined to 2.28 Å, and the Fe–S distance to 2.25 Å [85]. However, since the structure was determined only at 2.20 Å resolution, the experimental uncertainty of these bond lengths is at least ± 0.2 Å. No significant differences in the conformation of the protein backbone or the side chain atoms were found between the substrate-free [85] and substrate-bound structures [12], apart from a slight repositioning of the Phe87 side chain. Because the X-ray data available for the substrate-bound complex exhibits a significantly better resolution (1.6 Å) [12], we have chosen to build the model for our simulations based on this structure. To model the resting state, the camphor molecule was removed, and replaced by the cluster of six water molecules from the substrate-free structure [85]. In the latter, only the iron-linked aqua ligand could be resolved as an isolated sphere of electron density. The remaining five solvent molecules in the substrate pocket occupy a large lobe of unresolved electron density. Hence, the cluster of water molecules resulting from the refinement procedure represents only one of many energetically similar configurations that co-exist. This is probably due to the absence of hydrophilic protein groups in the active site capable of introducing a strong preference to any particular configuration. Therefore, we performed MM MD simulations to determine a set of representative configurations of the water cluster and the protein environment. These structures were separately investigated by QM/MM optimizations, in order to assess the effect of different conformations on the resulting computed properties.

Experimentally, the ferric complex **1** has a doublet ground state [88]. Figure 5.2 displays the corresponding orbital occupation, and the alternative quartet and sextet configurations. Recent calculations on model compounds [6, 221] show that even qualitative predictions regarding the multiplicity of the ground state, *i.e.* the energy differences between the HS (quartet and sextet) and LS states, of the ferric complex **1** are a challenge to theoretical approaches. The results from restricted open-shell Hartree-Fock (ROHF) as well as from density functional calculations with and without exact exchange show that the stability of the HS state markedly depends on the amount of exact exchange in the theory employed. The ROHF calculations strongly favor the sextet state.

On the other hand, the “pure” density functional BLYP (no exact exchange) clearly predicts a doublet ground state. Hybrid functionals using the B3 exchange expression due to Becke (20% exact exchange) give smaller energy separations with a slight preference for the doublet state.

The underlying cause of this effect has recently been systematically investigated by Hess and coworkers [217,218]. These authors found a large dependence of the high spin – low spin splitting in Fe^{II} transition metal complexes on the amount of exact exchange included in the computational method. For complexes with a sulfur-rich first coordination sphere, different density functionals yield results which vary by up to 1 eV. These variations were tracked back to the systematic overestimation of the stability of HS states by HF theory, where Fermi correlation is explicitly included through the exchange terms, while Coulomb correlation is not. With density functionals, the high spin – low spin splitting in transition metal complexes was shown to linearly depend on the coefficient of exact (HF) exchange admixture. While the pure density functionals BP86 or BLYP are strongly biased towards LS states, the popular hybrid functional B3LYP was found to overestimate the stability of HS states. Based on these observations and by reference to experimental data, it was suggested to reduce the 20% admixture of HF exchange in the B3 functional to 15%. A corresponding implementation of a density functional with 15% HF exchange, dubbed B3LYP*, was demonstrated to be capable of correctly predicting the ground state multiplicity of a range of iron-sulfur complexes, while retaining the overall good accuracy in applications to “standard” systems [218].

In the light of these findings, it was of interest to investigate the influence of the amount of the exact exchange admixture in the density functional on the computed doublet-quartet and doublet-sextet gaps of the iron-aqua complex **1** in a comprehensive manner. This will help us to test the validity of the approaches used in model studies in the literature [6,42,92,93], and to estimate the intrinsic accuracy of the computational method used in the present study.

Another aspect of the present work is to study the influence of the protein environment on the central iron complex. Previous theoretical studies did not consider, *e.g.*, the specific interactions of the axial ligand with other water molecules or protein residues in the active site. The only previous DFT/MM study, by Scherlis *et al.*, involved only single-point energy calculations at the X-ray geometry of P450_{BM3} [94], which does not provide a realistic measure for the relative stabilities of the different spin states. The present study involves full QM/MM geometry optimizations of the species in the enzyme environment for all spin states, and hence, realistic, “adiabatic” energy differences. The optimized geometry of the ground state will additionally serve to compute EPR parameters that can directly be compared to reliable experimental data (see chapter 6).

5.2 Preparation of the System

Initial Geometry

Initial coordinates were obtained from the crystal structure of the ferric substrate complex (PDB code 1DZ4, Brookhaven Protein Database) by Schlichting *et al.* [12]. No ordered water molecules in the active site are present in this species. Only one monomer from the crystallographic dimer was used, which corresponds to the A chain of the PDB structure, where the first 10 residues are unresolved. These 10 residues were omitted in the calculations. 310 associated crystal water molecules (chain Z) were included in the model. The coordinates of camphor and the TRIS buffer molecule were deleted. The potassium ion in the binding site is present in the simulations.

The positions of 5 water molecules in the active site were taken from X-ray data by Poulos *et al.* [85] (PDB code 1PHC, Brookhaven Protein Database). The relative position of this water cluster was determined by aligning the backbone atoms of the two crystal structures by a least squares fit. Another water molecule was added as the sixth ligand to iron. ESEEM studies indicate the presence of such a ligand, with the plane spanned by the water molecule perpendicular to the porphyrin plane [86, 87]. The initial geometry of the aqua-heme complex was determined by reference to gas phase QM calculations at the DFT level (B3LYP/LACVP,6-31G).

The protonation state of the protein and the positions of hydrogen atoms were derived in analogy to the procedure detailed in section 3.2. The enzyme was solvated with a layer of TIP3P water molecules of 16 Å thickness, leading to a system of 28497 atoms in total, with 21117 atoms in the solvent. The net charge is $-10e$.

Force Field Parameters

The force field parameters for the axial water ligand were taken from the standard TIP3P parameter set (oxygen: type OT, charge $-0.834e$; hydrogen: type HT, charge $0.417e$). The sulfur atom of the proximal cysteine was assigned an atomic charge of $-0.07e$. All other parameters were used as obtained from the CHARMM22 library [146].

Preparatory Force Field Calculations

To relax the structure, a series of classical force field calculations was carried out using the CHARMM program [172], version 27a2. In these simulations, the coordinates of the entire heme unit, the oxygen atoms of the 5 water molecules in the active site, the proximal cysteine (Cys357),

and the outer 8 Å of the solvent were kept fixed. The details of these simulations are completely analogous to the procedure described in section 3.2, and are shortly summarized as follows:

1. An initial geometry optimization included 3600 steps of ABNR minimization to remove close contacts. Heavy atom positions of the enzyme were assigned harmonic constraints [100 kcal/(mol Å²) for the backbone and 50 kcal/(mol Å²) for the side chains] that were scaled down by a factor of 0.65 every 80 steps. The final GRMS was 0.00073 kcal/(mol Å²).
2. MD (NVT): 15 ps heating dynamics (0 → 300 K), 200 ps equilibration dynamics (300K), timestep 1 fs, SHAKE constraints [173] for bonds to hydrogen atoms.
3. Several snapshot structures obtained from the equilibration trajectory were optimized by 5000 steps of ABNR minimization, respectively. The final GRMS was always below 0.003 kcal/(mol Å²).

Table 5.1: MM energy minimization of selected snapshots of the equilibration trajectory (coordinates of heme, Cys357, and outer 8 Å of solvent layer fixed): Maximum and average RMS deviation [Å] of atomic positions with respect to the X-ray structure (1DZ4 [12]) after alignment of the backbone atoms, potential energy [kcal/mol], and root mean square of the gradient (GRMS) [kcal/(mol Å²)].

snapshot	RMS deviation from X-ray		Potential Energy ^a	GRMS
	Maximum	Average		
0 ps ^{b,c}	0.2371	0.1677	-62811.26210	0.00073
50 ps ^b	0.5354	0.3786	-66447.91348	0.00000
100 ps	0.6234	0.4408	-66907.16674	0.00260
150 ps ^b	0.6470	0.4575	-66945.82037	0.00001
175 ps ^b	0.5872	0.4152	-66805.36913	0.00075
190 ps	0.5655	0.3998	-66723.80612	0.00011
195 ps ^b	0.5643	0.3990	-67784.90956	0.00190
200 ps ^b	0.5768	0.4079	-67470.07154	0.00001

^a CHARMM22 force field potential energy [kcal/mol]. ^b Structure was selected for QM/MM investigations. ^c The structure after an initial MM optimization, before the MD run.

Table 5.1 summarizes the results of the energy minimizations of the snapshot structures taken from the MD trajectory. The first structure (0 ps) corresponds to the MM optimized geometry before the MD. We measured the root mean square (RMS) deviation of the backbone atom positions in the minimized structures with respect to the X-ray geometry (1DZ4 [12]). Generally, the deviation in aligned position is relatively small in all structures. Obviously, “snapshot 0 ps” exhibits the smallest deviations of (maximum/average) 0.24/0.17 Å, whereas all structures extracted from the MD

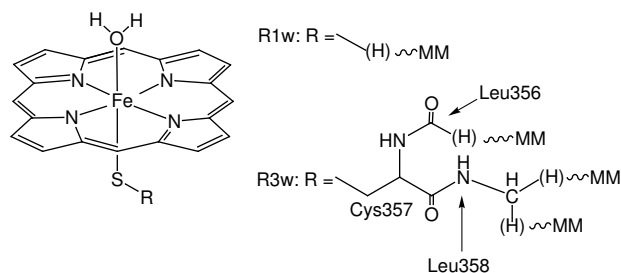


Figure 5.3: Choice of QM regions R1w and R3w.

show rather similar values, varying from 0.65/0.46 Å (snapshot 150 ps) to 0.54/0.38 Å (snapshot 50 ps). Five of the energy minimized snapshots were selected at random for the QM/MM studies described below. These are the structures after 0 ps, 50 ps, 150 ps, 175 ps, 195ps, and 200 ps simulation time, respectively. Note that snapshot 195 exhibits the lowest potential energy of all structures.

5.3 Computational Methodology

5.3.1 QM Regions and Basis Sets

In the QM/MM calculations, two different QM regions were employed, which are defined in analogy to section 3.3 as R1w and R3w (see figure 5.3). These QM subsystems comprise the following set of atoms:

R1w (42 QM atoms): Iron-oxo-porphyrin (without sidechains of the heme), sulfur of Cys357, and the axial water ligand. In the gas phase calculations, this corresponds to [FeO(SH)(porph)(H₂O)].

R3w (59 QM atoms): Iron-oxo-porphyrin (without sidechains of the heme), Cys357, CO group of Leu346, NH-C_αH unit of Leu348, and the axial water ligand.

We have used the basis sets defined in sections 3.3 and 4.2 that are abbreviated as B1 and B4. Basis B1 consists of the LACVP small-core effective core potential with the associated basis set for iron [177], and the 6-31G basis set for all other atoms. Basis B4 additionally contains a set of polarization and diffuse functions (6-31⁺⁺G*) on the six atoms coordinated to iron and a set of polarization and diffuse functions on the two protons of the axial water ligand (6-31⁺⁺G**) [178–181].

5.3.2 Optimized Regions

The QM/MM optimizations included the following set of residues of the active site.

Enzyme: Tyr75, Phe87, Tyr96, Pro100, Thr101, Gln108, Arg112, Val119, Leu244, Leu245, Gly249, Thr252, Val253, Leu294, Val295, Asp297, Arg299, Gln322, Thr349, Phe350, Gly351, Hsp355, Leu356, Cys357, Leu358, Gly359, Gln360, Leu362, Ala363, Ile367, heme

Water: 6 water molecules within the binding pocket, Wat63, Wat257.

This corresponds to a total of 586 optimized atoms.

5.3.3 QM/MM Scheme

The QM part was treated with unrestricted Kohn-Sham DFT. We have employed different density functionals for the exchange energy to assess the influence of the amount of Hartree-Fock (HF) exchange on the relative energies of the spin states (doublet, quartet, sextet) of the central iron(III)-aqua complex (see above). Geometry optimizations were performed with the B88 [183] (no HF exchange) and B3 [184] (20% HF exchange) functionals. Additional single-point calculations were done with the Becke half-and-half functional [222] (50% HF exchange). The correlation energy was calculated from the LYP functional [185] in all cases. The CHARMM22 force field [146] was used to represent the MM part of the system. The electrostatic embedding scheme and link atoms were used for the QM-MM coupling. This QM/MM coupling scheme has been presented in detail in section 3.3. All calculations were done with ChemShell [190], TURBOMOLE [187,188] and DL-POLY [191] for the QM/MM, QM and MM contributions, respectively. Geometry optimizations were carried out with the HDLCOPT optimizer [192] applying the default convergence criteria (see appendix A.2).

5.3.4 Strategy

To assess the influence of different protein conformations on the computed properties, separate QM/MM geometry optimizations were performed for the 6 structures denoted snapshot 0, 50, 150, 175, 195, and 200 that were obtained from the MM calculations described above. All these structures were optimized in the doublet, quartet, and sextet states, respectively. Due to the high conformational complexity of the system, separate geometry optimizations for individual spin states often lead to different conformations that are not related to the characteristics of the central

iron complex but to the protein/solvent environment. In order to avoid corresponding artifacts, we first performed QM/MM geometry optimizations for the ^2A state, using the B3LYP functional for the QM calculation. These geometries then served as the starting point for geometry optimizations of the ^4A and ^6A states, using the B3LYP functional. Similarly, the B3LYP geometry of the ^2A state was first reoptimized with the BLYP functional. The resulting structure represented the starting point for corresponding optimizations of the higher spin states. We also report the vertical relative energies of the ^4A and ^6A states at the geometry optimized for the ^2A species. These extensive QM/MM studies were carried out with the smaller QM region R1w and basis B1 (results are denoted R1w/B1 in the following). To compare the geometry and electronic structure of the iron complex to the corresponding situation in the gas phase, QM subsystem R1w was also fully optimized as isolated species.

Snapshot 195, which exhibits the lowest absolute potential energy of all selected structures, was studied by QM/MM calculations using the larger QM region R3w and basis B4. Geometry optimizations for this system were done in analogy to the procedure detailed above, using the B3LYP and BLYP functionals. The corresponding results are referred to as R3w/B4 in the following.

5.4 Results

5.4.1 Influence of Exact Exchange

To demonstrate the influence of the amount of HF exchange on the calculated relative energies of different spin states, we have summarized the relative single-point energies obtained with the B88 [183], B3 [184], and BH [222] exchange expressions in table 5.2. The LYP functional [185] was used for the computation of the correlation energy in all cases. The geometries were always optimized with the B3LYP functional for the doublet, quartet, and sextet states respectively. We have chosen to report BLYP and BHLYP single-point energies on B3LYP optimized geometries, because, at this stage, we are primarily interested in the effect of the HF exchange on electronic energies. Relative energies from optimized structures, however, include geometry relaxation effects, which sometimes give misleading results. For example, the quartet-doublet gap from separate geometry optimizations in the gas phase using the B3LYP and BLYP functionals, respectively, is 5 and 4 kcal/mol (see below, tables 5.3 and 5.4). However, the optimized geometries of the ^4A state are very different: while the B3LYP optimized structure features a weak Fe–O interaction, BLYP predicts the axial water to be completely unbound. The close agreement of the calculated quartet-doublet gaps is thus coincidental, since the two functionals predict qualitatively different equilibrium structures for the ^4A state in the gas phase. Hence, this comparison cannot give a measure for the intrinsic preference of a given functional for the HS or LS situation.

Table 5.2: Relative single-point energies [kcal/mol] of the 4A state (E^{Q-D}) and the 6A state (E^{S-D}) with respect to the 2A state. Calculated using different exchange density functionals.

	QM/MM, snap 195 ^a		QM, gas phase ^b	
	$E^{(Q-D)}$	$E^{(S-D)}$	$E^{(Q-D)}$	$E^{(S-D)}$
B-LYP	3.9	19.8	10.9	21.8
B3-LYP	-2.4	4.5	5.5	6.9
BH-LYP	-10.8	-18.1	-0.8	-13.5

^a QM/MM single-point energy in the enzyme environment (only QM energies contribute), geometries from B3LYP/CHARMM22 (R1w/B1) optimization. ^b QM single-point energy in the gas phase, geometries from B3LYP (R1w/B1) optimization.

Table 5.2 shows the relative energies of the quartet ($E^{(Q-D)}$) and sextet ($E^{(S-D)}$) state with respect to the doublet state. The first two entries correspond to QM/MM single-point energies including explicit polarization, from structures optimized at the R1w/B1, B3LYP/CHARMM22 level of snapshot 195. The third and fourth entry are the corresponding relative energies for the isolated species optimized in the gas phase. As expected, the BLYP functional exhibits a clear preference for the doublet state, and gives the ordering of the spin states as doublet below quartet below sextet. In going to the B3LYP results, the quartet and sextet states are lowered by 5–6 and 14–15 kcal/mol, respectively. Increasing the fraction of exact exchange from 20% (B3) to 50% (BH) further lowers these states by 6–8 and 20–23 kcal/mol. Here, the ordering of the spin states is reversed, being sextet below quartet below doublet. The lowering of the HS states with increasing fraction of HF exchange is approximately equally pronounced in the QM/MM and in the gas phase system. A second finding is that the HS states, especially the 4A state, are apparently stabilized in the enzyme environment as compared to the gas phase. This is obvious from the relative energies $E^{(Q-D)}$, which are smaller for the QM/MM calculation as compared to the gas phase results. A more detailed discussion of this effect will be presented in the following sections.

Hence, similar to the results of others for Fe^{II} -sulfur complexes [217] and model compounds of the resting form of P450 (**1**) [6], we find large variations of the relative stabilities of the HS vs. the LS states, depending on the amount of HF exchange in the density functional. As noted above, Hess and coworkers demonstrated [217, 218] that pure density functionals like BLYP are strongly biased towards LS states, whereas B3LYP overestimates the stability of HS states. Very good agreement with experimental data was obtained, when the amount of 20% exact exchange in the B3LYP functional was reduced to 15% [217]. Instead of choosing this pragmatic approach, we apply the standard B3LYP functional in the following study, which is apparently closest to the ideal value of admixture of HF exchange. This seems reasonable if we aim at qualitative trends regarding the influence of the protein environment, rather than at accurate values for the LS-HS separation. Additionally, we report results obtained with the BLYP functional. These might serve

as an “upper bound” for estimating the effect of reducing the amount of exact exchange in the density functional.

5.4.2 Influence of the Protein Environment

Relative Energies

Table 5.3 summarizes the relative energies of the 4A and 6A states with respect to the 2A state, computed with the B3LYP functional for the QM part. The analogous data obtained with the BLYP functional are given in table 5.4. The “adiabatic” energy differences refer to optimized equilibrium structures of the respective spin states, whereas “vertical” energy differences are calculated at the equilibrium geometry of the 2A state. The relative QM/MM energies are a sum of the contributions from the QM and the MM calculation. The QM contributions are explicitly given in the tables. In the case of the “vertical” energies, the MM contributions to the QM/MM energy are constant and hence do not contribute to the HS-LS splittings.

Entries 1–6 in tables 5.3 and 5.4 show the influence of different protein conformations on the QM/MM (R1w/B1) calculated state energy differences. Generally, these fluctuations in all relative energies are rather small, varying by less than 2 kcal/mol. If only the QM contributions are considered, the B3LYP calculations predict a quartet ground state, which is favored over the doublet state by only 1–3 kcal/mol. The sextet state is significantly higher in energy in all calculations. As expected, the BLYP functional favors the doublet state, which now becomes the ground state, with the quartet state lying 4–5 kcal/mol higher in energy.

By adding the MM contributions to the relative QM energies, one arrives at relative QM/MM energies, which are given in the first two columns of tables 5.3 and 5.4. The inclusion of MM terms causes a stabilization of the doublet state with respect to the quartet state in all snapshots by up to 3 kcal/mol. This effect is much less pronounced for the sextet state. As a consequence, the QM/MM calculated doublet-quartet gaps obtained with the B3LYP functional for the QM calculation in all snapshots vary from -1 to $+1$ kcal/mol, while the corresponding BLYP results range from 6–8 kcal/mol.

A notable exception from this rule is found in the BLYP/MM optimized structure of snapshot 40 (entry 6 in table 5.4): Here, the MM contributions to the doublet-quartet gap apparently favor the quartet state. The QM energy difference of 4 kcal/mol is thus nearly canceled by the MM terms, giving rise to a QM/MM energy difference of only 0.39 kcal/mol. The reason for this large negative MM contribution was investigated by an analysis of the 2A and 4 optimized structures, which shows that two of the six water molecules in the active site have different conformations in the two structures, giving rise to qualitatively different H-bond networks. This underscores again

the problem of the high dimensionality of the potential energy surface encountered in such calculations. Since our approach does not involve sampling of conformational space, we are restricted to the analysis of local potential minima that exhibit qualitatively the same conformations in the protein/solvent environment, *e.g.*, the same numbers of H-bonds. In snapshot 40, this requirement is not met, as the protein/solvent environment is inconsistent in the two compared structures. As a result, the MM contributions to the relative QM/MM energy give an artificial stabilization of one structure, which leads to an erroneously small energy gap.

Consistent with the results of Harris and Loew [91], the present QM/MM study indicates that inclusion of the protein environment has an influence on the relative stabilities of the spin states: it increases the doublet-quartet gap, provided that the protein/solvent environment exhibits consistent conformations. A comparison of the doublet and quartet QM/MM optimized geometries indicates that the stabilization is due to subtle variations in the conformations of the MM environment, and cannot be traced back to individual atomic positions. As a note of caution, it has to be stressed again that the optimizations of the quartet and sextet state started from the doublet equilibrium geometries, and one is dealing with local potential energy minima only. A definitive answer to the question of the effect of protein conformations on the stability of spin states requires a (QM/MM) MD study, including extensive sampling of conformational space.

Entry 7 in tables 5.3 and 5.4 shows the relative energies obtained from QM/MM calculations with the larger QM region R3w and basis B4 for snapshot 195. The use of the extended QM region favors the doublet state, by increasing the B3LYP (BLYP) QM contributions to the doublet-quartet gap from -2.4 (4.6) to 0.1 (6.1) kcal/mol, respectively. Using the B3LYP/CHARMM22 (BLYP/CHARMM22) QM/MM energies, the doublet state is thus consistently predicted as the ground state, with the quartet state 2.4 (8.5) and the sextet state 3.3 (18.3) kcal/mol higher in energy.

The relative doublet-quartet and doublet-sextet energies of the isolated QM subsystem R1w in the gas phase are given in entry 8 of tables 5.3 and 5.4. In accord with previous DFT model studies [6,42,92,93], we obtain a doublet ground state with both functionals employed. At the B3LYP/B1 (BLYP/B1) level of theory, the quartet minimum is calculated to be 5.5 (4.2) kcal/mol higher in energy, while the sextet minimum is calculated at a relative energy of 6.9 (20.5) kcal/mol. As noted above, the low doublet-quartet gap obtained from the BLYP as compared to the B3LYP calculation is rationalized by comparing the respective optimized structures: in the BLYP geometry the axial water ligand is not coordinated to the iron, while the B3LYP geometry features a weak Fe–O bond. Apparently, the BLYP functional underestimates the strength of the Fe–O interaction in the gas phase. BLYP single-point calculations on B3LYP optimized geometries give a doublet-quartet gap of 10.9 kcal/mol. This demonstrates again the preference of the LS state by the pure density functional. Moreover, the agreement of the doublet-quartet gap obtained from the BLYP

Table 5.3: Relative energies [kcal/mol] of the 4A state (E^{Q-D}) and the 6A state (E^{S-D}) with respect to the 2A state. Calculated using the B3LYP functional.

entry	snapshot	adiabatic ^a				vertical ^b	
		QM/MM ^c		QM ^d		QM ^d	
		E^{Q-D}	E^{S-D}	E^{Q-D}	E^{S-D}	E^{Q-D}	E^{S-D}
1	0	0.73	4.95	-2.32	4.00	8.27	14.99
2	50	1.06	4.32	-1.99	4.15	8.51	14.75
3	150	-0.29	3.53	-2.38	3.81	7.15	13.35
4	175	0.73	3.65	-1.01	4.78	7.94	13.89
5	195	0.57	4.01	-2.41	4.50	7.98	13.94
6	200	-0.62	3.78	-1.62	4.15	7.53	13.80
7	195,R3w/B4	2.37	3.30	0.12	4.13	10.42	14.50
8	gas phase						
	R1w/B1	-	-	5.48	6.92	13.26	19.23

^a Relative energies from optimized geometries. ^b Relative energies from single-point calculations at the 2A optimized geometry. ^c Relative QM/MM energy including MM contributions. ^d Relative QM energy.

and B3LYP optimizations is shown to be coincidental, since the energies relate to two qualitatively different geometries.

Interestingly, the doublet-quartet gap is dramatically reduced when going from the gas phase (entry 8 in table 5.3) to the snapshots optimized in the enzyme environment (entries 1-6 in table 5.3). For example, E^{Q-D} is decreased from 5.5 to -2.4 kcal/mol in the case of snapshot 195. A similar, albeit much smaller effect is apparent for the sextet state. The reason for this significant stabilization of the quartet state will be addressed below.

The B3LYP (BLYP) “vertical” energies for the doublet-quartet gap are 7-9 (14-15) kcal/mol, and 14-15 (30-31) kcal/mol for the doublet-sextet gap, respectively. This should allow for an equilibration of the doublet state with the higher spin states at ambient temperature, as is indeed observed experimentally.

Optimized geometries

Key features of the optimized geometries from the QM/MM calculations in the enzyme environment as well as the isolated gas phase systems are summarized in tables 5.5 and 5.6.

The differences in the electronic structures of the various spin states are manifest in their optimized

Table 5.4: Relative energies [kcal/mol] of the 4A state (E^{Q-D}) and the 6A state (E^{S-D}) with respect to the 2A state. Calculated using the BLYP functional.

entry	snapshot	adiabatic ^a				vertical ^b	
		QM/MM ^c		QM ^d		QM ^d	
		E^{Q-D}	E^{S-D}	E^{Q-D}	E^{S-D}	E^{Q-D}	E^{S-D}
1	0	7.04	20.01	4.12	18.78	14.96	31.23
2	50	7.58	19.82	4.46	18.89	15.46	31.08
3	150	5.98	18.48	3.85	18.41	13.87	29.57
4	175	6.86	18.86	5.03	19.29	14.79	30.22
5	195	6.91	18.95	4.58	19.24	14.71	30.21
6	200	(0.39) ^e	18.05	4.05	19.30	14.28	30.01
7	195,R3w/B4	8.45	18.27	6.07	18.63	17.59	31.32
8	gas phase						
	R1w/B1	-	-	4.16 ^f	20.51	27.64	35.31

^a Relative energies from optimized geometries. ^b Relative energies from single-point calculations at the 2A optimized geometry. ^c Relative QM/MM energy including MM contributions. ^d Relative QM energy. ^e Different H-bond network in the active site. ^f Axial water molecule is not coordinated to iron.

geometries. The doublet state exhibits the shortest bonds of the six ligands to the central iron atom. In contrast, the quartet state is typified by significantly elongated bonds to the axial ligands. This is due to the occupation of the iron d_{z^2} orbital, which contributes to the antibonding σ^* orbital along the S–Fe–O axis. In the sextet state, an unpaired electron additionally occupies the iron d_{xy} orbital. This orbital mixes with sp^2 orbitals of the pyrrole nitrogens to give a strongly antibonding Fe–N σ^* MO. As a consequence, the Fe–N bonds are elongated by *ca.* 0.05 Å and the central iron atom is displaced from the heme plane towards the proximal face of the heme.

The improper dihedral angle $\Theta_{Fe-O-H-H}$ and the torsion angle $\Theta_{O-H-H-Fe}$ are a measure for the “tilting” of the water ligand with respect to the heme plane. A value of 180° (0°) for $\Theta_{Fe-O-H-H}$ ($\Theta_{O-H-H-Fe}$) indicates a perfectly upright conformation, with the plane spanned by the three water atoms perpendicular to the heme plane. As obvious from tables 5.5 and 5.6, the QM/MM optimizations always predict a slight deviation from the perfectly upright geometry, even in snapshot 0, which is optimized starting from the X-ray conformation. This deviation is smallest for the doublet species and slightly more pronounced for the quartet and sextet equilibrium structures. In general, there is a correlation between the “tilting” of the water ligand and the Fe–O bond length: the shorter the Fe–O bond length, the weaker the “tilting” of the water ligand.

As compared to the differences in geometries of the different spin states, the variations arising from different snapshots are relatively small. For example, the Fe–O bond lengths vary by up to

(doublet/quartet/sextet) 0.032/0.029/0.041 Å in different snapshots. This is in line with the small differences in the calculated relative energies of the different spin states.

In contrast, there are notable differences between the geometries optimized in the enzyme environment, and those for the isolated QM subsystem R1w in the gas phase. This is especially obvious from the Fe–O bond lengths, which are shortened by (doublet/quartet/sextet) 0.082/0.274/0.242 Å in snapshot 195 as compared to the isolated species. In contrast, the Fe–S bond is slightly elongated in the enzyme environment. In snapshot 195, this bond is longer than in the gas phase species by (doublet/quartet/sextet) 0.066/0.088/0.086 Å. Moreover, the water ligand is strongly tilted in the quartet and sextet gas phase minimum geometries. In these structures, the water protons form hydrogen-bonds with the pyrrole nitrogen atoms of the porphyrin. These interactions are not present within the protein environment, where they are replaced with hydrogen bonds to the neighbouring water molecule occupying the pocket. Hence, our calculations indicate that the interaction of the axial water ligand with the other water molecules present in the active site favors the upright conformation of this ligand, and shortens the Fe–O bond. This effect is apparent in all spin states, but it is most pronounced in the quartet and sextet species.

The geometries optimized with the BLYP functional exhibit analogous trends to the B3LYP structures discussed above. The Fe–O bond is predicted by this functional to be slightly longer, *e.g.*, we obtain a bond length of (doublet/quartet/sextet) 2.071/2.356/2.329 Å in snapshot 195 as compared to the B3LYP values of 2.032/2.280/2.251 Å. As noted above, the greatest difference between the two density functionals is observed for the isolated gas phase species. Here, the BLYP optimization predicts a structure for the quartet and sextet state where the Fe–O bond is completely broken (distances in quartet: 3.074 Å, sextet: 3.210 Å). The water molecule is loosely coordinated *via* hydrogen bonds to the pyrrole nitrogens of the porphyrin. As discussed above, it is well known that pure density functionals like BP86 or BLYP overestimate the stability of LS states. This effect apparently carries over to an underestimation of the iron-ligand bond strengths in the HS states.

Our best estimates from the B3LYP/CHARMM22 QM/MM geometry optimizations employing the large QM region R3w yield Fe–O distances of 2.141/2.475/2.467 Å in the doublet/quartet/sextet states, respectively. The computed Fe–S bond lengths at this level of theory are 2.269/2.491/2.418 Å. Hence, the optimized doublet geometries agree reasonably well with the experimental values of 2.28 Å for the Fe–O and 2.25 Å for the Fe–S bond lengths (estimated uncertainty ± 0.2 Å, see section 5.1). The Fe–H distances to the protons of the axial water ligand are 2.646 and 2.620 Å, in perfect agreement with the result of 2.62 Å obtained from the ESEEM experiment [87].

In summary, the QM/MM calculations employing the large QM region R3w and basis B4 in conjunction with B3LYP/CHARMM22 optimizations lead to a geometry for the doublet species in good agreement with the experimental data. A comparison of the geometries from the R1w/B1 QM/MM study with corresponding optimizations in the gas phase shows that the enzyme environ-

Table 5.5: Selected structural data (distances r in Å, angles θ , dihedral angles Θ and improper dihedral angles Θ' in degree). Calculations employed the B3LYP functional ^{a,b}.

snapshot		$r_{\text{Fe-N}}^{\text{av}}$	$r_{\text{Fe-O}}$	$r_{\text{Fe-S}}$	$\theta_{\text{Fe-O-H}}^{\text{av}}$	$\Theta_{\text{Fe-O-H-H}}$	$\Theta'_{\text{O-H-H-Fe}}$
0	D	2.028	2.011	2.348	115.2	136.1	-31.6
	Q	2.028	2.268	2.545	109.5	120.8	-44.7
	S	2.080	2.235	2.493	110.7	123.6	-42.3
50	D	2.030	2.025	2.349	114.0	132.7	-33.9
	Q	2.031	2.277	2.543	110.4	121.9	-42.8
	S	2.084	2.278	2.484	107.0	118.0	-48.5
150	D	2.032	2.023	2.365	119.0	146.7	-22.2
	Q	2.031	2.272	2.581	118.6	137.9	-26.6
	S	2.083	2.242	2.519	118.9	140.7	-25.4
175	D	2.031	2.043	2.353	114.4	133.0	-33.5
	Q	2.032	2.297	2.554	112.0	123.9	-40.0
	S	2.086	2.276	2.496	112.0	125.2	-39.7
195	D	2.031	2.032	2.356	115.8	135.9	-30.5
	Q	2.031	2.280	2.567	114.2	127.5	-36.0
	S	2.085	2.251	2.503	115.4	131.4	-33.2
200	D	2.030	2.035	2.358	115.4	134.9	-31.3
	Q	2.030	2.293	2.572	115.9	130.0	-32.9
	S	2.083	2.262	2.510	114.1	128.4	-35.9
195,R3w/B4	D	2.038	2.141	2.269	109.5	118.8	-44.9
	Q	2.037	2.475	2.491	108.1	113.1	-48.7
	S	2.095	2.467	2.418	106.9	113.4	-50.6
gas phase (R1w/B1)	D	2.019	2.114	2.290	108.1	118.3	-46.0
	Q	2.021	2.554	2.479	94.3	95.8	-70.6
	S	2.086	2.493	2.417	94.7	96.1	-69.6

^a av denotes an averaged value. ^b D: ^2A (doublet) state; Q: ^4A (quartet) state; S: ^6A (sextet) state.

Table 5.6: Selected structural data (distances r in Å, angles θ , dihedral angles Θ and improper dihedral angles Θ' in degree). Calculations employed the BLYP functional ^{a,b}.

snapshot		$r_{\text{Fe-N}}^{\text{av}}$	$r_{\text{Fe-O}}$	$r_{\text{Fe-S}}$	$\theta_{\text{Fe-O-H}}^{\text{av}}$	$\Theta_{\text{Fe-O-H-H}}$	$\Theta_{\text{O-H-H-Fe}}$
0	D	2.040	2.057	2.338	111.0	125.0	-40.8
	Q	2.041	2.351	2.553	104.5	112.5	-53.4
	S	2.097	2.312	2.505	106.1	115.2	-50.6
50	D	2.043	2.068	2.339	109.4	122.3	-43.3
	Q	2.043	2.411	2.541	99.2	106.1	-61.8
	S	2.101	2.353	2.490	101.8	110.2	-57.4
150	D	2.043	2.065	2.358	115.1	132.6	-32.8
	Q	2.044	2.349	2.588	112.9	122.6	-39.2
	S	2.100	2.318	2.531	113.0	124.4	-38.7
175	D	2.043	2.083	2.340	111.4	125.1	-40.0
	Q	2.045	2.383	2.565	106.6	113.9	-50.1
	S	2.102	2.354	2.508	107.2	115.8	-48.9
195	D	2.043	2.071	2.348	112.6	127.0	-37.7
	Q	2.045	2.356	2.571	110.4	118.7	-43.6
	S	2.101	2.329	2.517	110.0	119.5	-44.1
200	D	2.042	2.077	2.349	111.6	125.0	-39.7
	Q	2.041	2.449	2.590	105.9	114.3	-51.7
	S	2.099	2.352	2.524	110.1	119.0	-44.1
195,R3w/B4	D	2.050	2.191	2.239	107.1	114.8	-48.9
	Q	2.051	2.605	2.490	102.6	105.7	-57.7
	S	2.113	2.611	2.421	100.9	105.9	-60.4
gas phase (R1w/B1)	D	2.030	2.159	2.291	103.8	110.9	-53.3
	Q ^c	2.032	3.074	2.477	66.1	56.4	-122.8
	S ^c	2.109	3.210	2.416	62.3	47.7	-131.8

^a av denotes an averaged value. ^b D: ^2A (doublet) state; Q: ^4A (quartet) state; S: ^6A (sextet) state. ^c Axial water molecule is not coordinated to iron.

ment shortens the Fe–O bond, and favors the upright conformation of the axial water ligand. This is especially pronounced in the HS (quartet and sextet) states. In contrast, the Fe–S bond is slightly elongated under the influence of the surrounding protein. Different snapshot structures representing different protein conformations show only small geometrical variations in the coordination sphere of the iron complex.

Spin Densities

The unpaired spin densities from the unrestricted B3LYP/CHARMM calculations on snapshot 195 and from B3LYP calculations on the isolated gas phase system are presented in table 5.7. The other snapshots optimized at the R1w/B1 QM/MM level show only minor variations in computed spin densities, and are therefore not addressed in the following discussion. Likewise, the corresponding results obtained with the BLYP functional show analogous trends, and are therefore not presented here.

The spin density distribution of the doublet state shows that the unpaired electron is predominantly localized on the iron atom. This is expected for the single occupation of the d_{yz} orbital, which has weak interactions with ligand orbitals. In contrast, the quartet state exhibits a significant amount of positive spin density on the sulfur atom, due to the mixing of the iron d_{z^2} with the sulfur p_z orbital. In the sextet state, only 4 equivalents of the 5 unpaired electrons are attributed to the iron center. This is due to the strong interaction of the d_{xz} with sp^2 orbitals at the pyrrole nitrogen atoms. These four nitrogen atoms each carry an average spin density of 0.1 electron equivalents in the sextet state (see last column in table 5.7).

A comparison of the QM/MM results with the gas phase calculations of the isolated R1w/B1 system reveals a redistribution of spin density between the iron and sulfur atom in the HS states: Going from the enzyme environment to the gas phase, the sulfur atom acquires more spin density (quartet: 0.271→0.459; sextet: 0.323→0.491), with a concomitant reduction of spin density on iron (quartet: 2.739→2.557; sextet: 4.067→3.987). This finding indicates that the protein environment, by polarization and the presence of H-bond donors in the proximal pocket, effects a stabilization of electron density on the sulfur atom. As discussed in chapter 3, a similar effect has been observed in the influence of the protein on the ratio of sulfur *vs.* porphyrin radical character of compound I of P450 (7).

Interestingly, the unpaired spin density on the oxygen atom of the axial water ligand is relatively small in all calculations. This indicates a rather weak mixing of the iron d_{z^2} and the oxygen p_z orbitals that yields the singly occupied antibonding σ^* Fe–O orbital.

Table 5.7: Spin densities on selected atoms. Calculations employed the B3LYP functional ^{a,b}.

snapshot		$\rho(\text{Fe})$	$\rho(\text{O})$	$\rho(\text{S})$	$\rho_{\text{gr}}(\text{por})$	$\rho_{\text{av}}(\text{N})$
195, R1w/B1	D	1.131	0.002	-0.058	-0.077	-0.027
	Q	2.739	0.036	0.271	-0.044	-0.036
	S	4.067	0.041	0.323	0.574	0.114
195, R3w/B4	D	1.151	-0.011	-0.029	-0.113	-0.033
	Q	2.812	-0.013	0.329	-0.132	-0.055
	S	4.181	-0.008	0.367	0.458	0.087
gas phase (R1w/B1)	D	1.090	0.001	-0.005	-0.085	-0.023
	Q	2.557	0.021	0.459	-0.025	-0.022
	S	3.987	0.028	0.491	0.507	0.110

^a ρ_{av} denotes an averaged value, ρ_{gr} a sum of individual values for a group of atoms. ^b D: ²A (doublet) state; Q: ⁴A (quartet) state; S: ⁶A (sextet) state.

5.5 Discussion

The relative energies presented in tables 5.3 and 5.4 show that the ground state of the ferric resting state of P450 has doublet multiplicity, in agreement with the experimental assignment [88]. The best estimates from B3LYP/CHARMM energies with QM region R3w and basis B4 predict the quartet state at 2.4 kcal/mol, and the sextet state at 3.3 kcal/mol relative to the doublet. Since the B3LYP functional is known to underestimate the stability of the LS state due to the admixture of exact exchange (see above), the true energy differences are expected to be even slightly larger. Analogous calculations with the BLYP functional give an upper bound for the doublet-quartet (doublet-sextet) gap of 8.5 (18.3) kcal/mol.

The optimized geometries of the doublet ground state obtained from B3LYP/CHARMM calculations yield a Fe–O distance of 2.141 Å and a Fe–S bond length of 2.269 Å. This agrees well with the X-ray data of 2.28 (± 0.2) Å for the Fe–O and 2.25 (± 0.2) Å for the Fe–S bond lengths. The computed Fe–H distances to the protons of the axial water ligand are 2.646 and 2.620 Å, in excellent agreement with the result of 2.62 Å obtained from the ESEEM experiment [87]. This supports the reliability of the QM/MM approach in predicting the geometrical features of the iron complex in the enzyme environment.

An analysis of the QM and MM energy contributions to the relative QM/MM energies shows that the MM terms favor the doublet state by up to 3 kcal/mol over the quartet state. A similar, but smaller effect is also found for the sextet state. However, it was not possible to unambiguously relate this stabilization to individual contributions from the MM environment. It would seem desir-

able to confirm this finding by QM/MM MD or MC calculations with appropriate conformational sampling.

The present QM/MM study reveals some features that add to the understanding of the influence of the protein environment on the computed properties of the Fe^{III} -aqua complex **1**. These may be summarized as follows: (i) Comparing the relative QM/MM spin state energies to those obtained from QM optimizations of the isolated gas phase species indicates that the enzyme environment provides a significant stabilization of the quartet state, and, to a smaller extent, of the sextet state. (ii) The geometries optimized at the QM/MM level show a preference for the conformation with an essentially upright axial water ligand, although this ligand is always found to be slightly tilted with respect to the heme plane. (iii) There are significant differences in the conformation of the axial water in the HS (quartet and sextet) states between QM/MM and analogous gas phase optimized structures: The protein/solvent environment favors a significant shortening of the Fe–O bond and an upright axial water conformation. In contrast, the isolated species have long Fe–O distances and the axial water ligand nearly parallel to the heme plane, since the water protons form hydrogen-bonds with the pyrrole nitrogens of the porphyrin. (iv) The Fe–S bond in all spin states is slightly elongated in the enzyme environment as compared to the gas phase. (v) In the HS states the sulfur atom exhibits significantly more unpaired spin density in the gas phase than in the enzyme environment. This indicates that the enzyme environment favors localization of electron density on sulfur.

The largest influence of the protein/solvent environment on energies, geometries, and spin densities is found for the quartet state. Particularly notable is the stabilization of the quartet relative to the doublet state (which also occurs in the product complex **9**, see section 4.3).

Figure 5.4 shows a MO diagram that rationalizes these findings in a simple manner. Decisive for the relative stability of the quartet vs. the doublet configuration is (in first approximation) the relative energy of the d_{z^2} orbital with respect to the d_{xz} orbital. If this energy difference is large, spin pairing (the doublet state) will be preferred over occupation of the higher d_{z^2} orbital (the quartet state) which gives rise to a larger exchange stabilization. However, the d_{z^2} orbital mixes with the p_z orbital on sulfur and also with a corresponding orbital on the oxygen of the axial water ligand, yielding a bonding σ and an antibonding σ^* orbital (shown in figure 5.4). The polarizing effect of the enzyme and the presence of H-bond donors in the Cys357 loop stabilize electronic charge on the sulfur atom, as has already been shown for the compound I species **7** in chapter 3.4. In the present situation, this effectively lowers the orbital energy of the (doubly occupied) sulfur p_z orbital. As a consequence, the interaction of this orbital with the d_{z^2} orbital becomes weaker (see figure 5.4), and therefore, the splitting of the bonding σ and an antibonding σ^* orbital is smaller, as compared to the gas phase. Hence, the destabilization $\Delta\epsilon$ of the σ^* orbital due to the mixing is smaller and its occupation (and thus the quartet state) becomes more favorable, which is mirrored by the smaller calculated doublet-quartet gap in the enzyme environment. In the present case, this

accounts for a stabilization of the quartet state by *ca.* 8 kcal/mol. These results accord well with the DFT/MM single-point calculations by Scherlis *et al.* [94], who found a decrease of the vertical doublet-quartet gap by 7 kcal/mol when the polarizing effect of the protein was included into the QM treatment (DFT, BP86 functional) *via* an electrostatic interaction model similar to the one used here.

To underscore this qualitative explanation we present in table 5.8 the following properties computed in the enzyme and in the gas phase: (i) The relative energy of the occupied α -spin Kohn-Sham orbitals in the quartet state corresponding to d_{z^2} (σ^*) and d_{xz} orbital: This gap is increased from 0.848 eV in the enzyme to 1.057 eV in the gas phase. This indicates that the polarizing environment predominantly lowers the orbital energy of the σ^* orbital. (ii) The total Mulliken population on the sulfur: The electron population is decreased from 16.452 e in the enzyme environment to 16.202 e in the gas phase. This demonstrates that the enzyme environment effects localization of electron density on the sulfur atom. Because the σ^* orbital is unoccupied in the doublet state, the corresponding populations in this state are smaller (16.240 and 16.023). Hence, the quartet state will preferentially benefit from this effect.

Based on these findings, one may also rationalize the differences in the Fe–S and Fe–O bond lengths between the enzyme and the gas phase. The argument is similar to the concept of the well-known *trans*-effect: two different donor ligands to a transition metal in *trans* orientation “compete” in forming a bond with the same orbital of the metal (in this case the d_{z^2} orbital). Generally, the better σ -donor forms the stronger bond, while the bond to the other ligand is more labile as compared to the situation with identical ligands (“*trans*-effect”). In the present situation, the thiolate sulfur is clearly the stronger donor. However, the interaction of the sulfur p_z orbital with the d_{z^2} orbital is weakened by the enzyme environment, because the energetic separation between the two orbitals becomes larger (see figure 5.4). As a consequence, the interaction of the d_{z^2} orbital with the “competing” water ligand becomes stronger in the enzyme environment. In effect, the Fe–S bond is elongated, and the Fe–O bond is shortened in the protein as compared to the gas phase.

5.6 Conclusions

This study presents a comprehensive investigation of the energetic, geometric, and electronic factors that typify the resting form of P450_{cam} (**1**). In the first part of this work, we have probed the dependence of the results on the chosen density functional. We find that the computed relative energies of the doublet, quartet, and sextet states are very sensitive to the admixture of exact exchange. Consistent with the recent systematic work of Hess *et al.* [217, 218], we conclude that pure density functionals such as BP86 or BLYP overestimate the stability of the LS state. Results

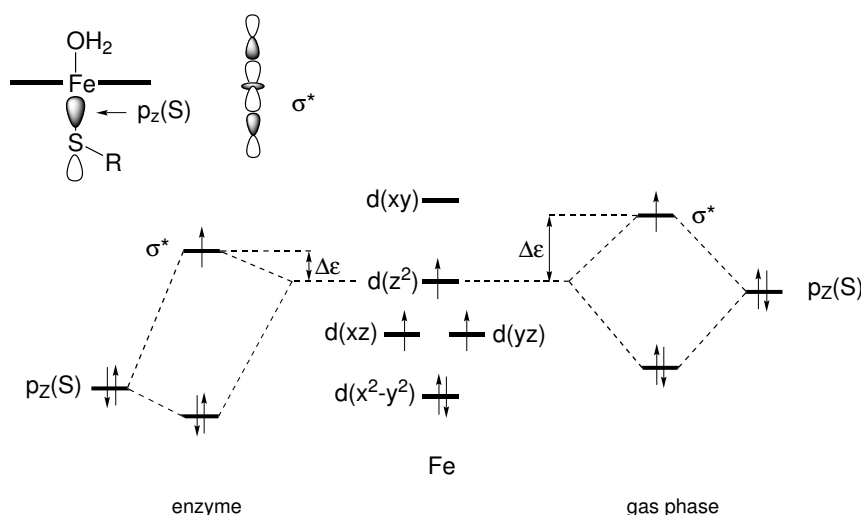


Figure 5.4: MO mixing diagram, rationalizing the stabilization of the quartet (4A) state by the polarizing effect of the enzyme (shown on the left) as compared to the gas phase (shown on the right).

Table 5.8: Resting state of P450 (**1**), quartet (4A) state: Orbital energy differences $\Delta\epsilon$ of the α -spin Kohn-Sham orbitals d_{z^2} and d_{xz} [eV], Mulliken total electron population $P_e(S)$ [e] on the proximal sulfur atom, and relative energies with respect to the doublet state [kcal/mol] in the enzyme and in the gas phase ^a.

	$S_{p,p}$	$S_{g,g}$
$\Delta\epsilon(d_{z^2} - d_{xz})$	0.848	1.057
$P_e(S)$	16.452	16.202
$E^{Q-D} (QM)$	-2.41	5.48

^a $S_{p,p}$: The system optimized at the QM/MM level (B3LYP/CHARMM, QM region R1w, basis set B1), in the enzyme environment; $S_{g,g}$: the isolated QM subsystem optimized in the gas phase (B3LYP, QM region R1w, basis set B1).

from previous studies employing such functionals thus have to be interpreted with care.

Using the B3LYP functional in conjunction with a QM/MM coupling scheme, our results for the full system indicate a doublet ground state of **1**, in accord with experimental results. Since the B3LYP functional is known to slightly underestimate the stability of the LS state [6,217,218,221], the calculated doublet-quartet and doublet-sextet gaps of 2.4 kcal/mol and 3.3 kcal/mol may be regarded as a lower bound to the true values. Analogous calculations with the BLYP functional allow us to estimate an upper bound for the doublet-quartet (doublet-sextet) gap of 8.5 (18.3) kcal/mol.

Comparing the QM/MM results to analogous results for the isolated QM subsystem in the gas phase reveals the effects of the protein environment on the central iron-aqua complex, which may

be summarized as follows. (i) The presence of other water molecules in the active site provides a hydrogen-bonding network that favors the upright conformation (perpendicular to the heme plane) of the axial water ligand. (ii) The polarization exerted by the enzyme environment stabilizes electron density on the proximal sulfur, which effectively lowers the orbital energy of the antibonding Fe–S σ^* orbital. As a consequence, the HS states, which have an unpaired electron in this orbital, are significantly stabilized in the enzyme as compared to the gas phase. This effect is most prominent in the quartet state, which is lowered by *ca.* 8 kcal/mol. (iii) Likewise, the electronic polarization by the protein weakens the σ -donor capabilities of the sulfur ligand. This is obvious from the optimized Fe–S bond lengths, which are elongated by *ca.* 0.1 Å in the enzyme environment. Due to the *trans*-effect, this causes a concomitant shortening of the Fe–O bond to the axial water ligand.

It is known that substrate binding induces the displacement of the axial water ligand, which in turn is linked to the reversal of the stability of the HS and LS states. The associated change in the reduction potential of the heme triggers the catalytic mechanism of substrate hydroxylation. Our QM/MM calculations reveal how the protein/solvent environment modulates the conformation and binding properties of the axial water ligand by electronic polarization as well as the presence of other water molecules in the active site. The relative energies of the HS states *vs.* the LS state are also affected by these factors. This highlights the advantage of the combined QM/MM approach, that allows us to simultaneously model the electronic factors at the heme site and the steric and electrostatic influence of the environment.

Chapter 6

Calculation of Spectroscopic Parameters

6.1 Motivation and Background

The characterization of short-lived intermediates in the catalytic cycles of cytochrome P450, CPO, HRP, catalases and other heme enzymes has become possible in the past decades by virtue of rapid developments in spectroscopic techniques [164,223]. EPR, ENDOR, and Mößbauer spectroscopy, particularly in combination with rapid freeze quench techniques, have been most successful in this regard. For example, EPR, ^1H - and ^{14}N -ENDOR studies [13] on P450_{cam} have recently identified the ferric peroxy anion and the ferric hydroperoxy intermediate of this enzyme (see section 1.2).

A definitive characterization of the primary oxidant of P450_{cam}, compound I (Cpd I), has not been possible to date [17]. If the ongoing efforts to detect this species should succeed, an interpretation of the experimental spectra may be difficult, *e.g.*, due to the presence of impurities. The analysis of experimental data can be significantly facilitated by theoretical predictions for the corresponding spin Hamiltonian parameters. Therefore, the goal of the present work is the calculation of such parameters in the enzyme environment, in order to assist experimentalists in an ultimate detection of this species. Specifically, this study is concerned with the prediction of exchange coupling constants, hyperfine and nuclear quadrupole coupling tensors, the ^{57}Fe Mößbauer isomer shift, as well as *g*-values for Cpd I of P450_{cam}. These parameters are reported for both the LS (doublet) and HS (quartet) state of Cpd I. We also compare the results obtained from calculations including the enzyme environment with those for the isolated system in the gas phase. In doing so, we can assess the validity of calculations that rely on gas phase model representations of the real system.

To test the accuracy of our approach, we have studied the oxoferryl tetraphenylporphyrin π cation radical complex **17** (see figure 6.1). Derivatives of this complex, which differ from **17** by substituents at the phenyl groups, have been characterized spectroscopically as models of the native

Cpd I species in P450, CPO, or HRP. We compare our results to the tetramesitylporphyrin (TMP) derivative, and the tetrakis(2,6-dichlorophenyl)porphyrin (TDCPP) derivative of **17**, where results from detailed EPR and Mößbauer studies are available [224]. The preparation of these complexes involves oxidation of a precursor in methylene chloride/methanol by oxidation with *meta*-chloroperbenzoic acid. Hence, the presence of a sixth ligand to the iron atom, presumably an oxygen donor, is likely. Thus, we additionally considered complex **18** (see figure 6.1) with an axial water ligand. Other effects due to the presence of a solvent, however, were neglected.

The resting state of P450_{cam}, *i.e.*, the iron aqua complex **1** (figure 6.1), has been studied by EPR, ENDOR and ESEEM techniques [86,87,225,226]. Because the corresponding spectroscopic data derive from the native species in the enzyme, QM/MM calculations allow for a direct comparison with these experimental values. Thus, we have performed calculations on this species, using the QM/MM optimized structure of **1** discussed in chapter 5.

The adequate description of the antiferromagnetically coupled spin systems, in this case the LS state of Cpd I, requires special theoretical methods. In the present work, the broken symmetry density functional approach (BS-DFT) [227–229] is adopted. Hence, the multi-determinant electronic state with unpaired α - and β -electrons delocalized over both the ferryl and the porphyrin subunit is described by a single spin-unrestricted determinant Kohn-Sham wave function with unpaired α - and β -electrons localized at the two individual sites. As a consequence, energies and spin densities of the broken symmetry state differ from those of the multi-determinant antiferromagnetic state, but its exchange coupling constants, hyperfine coupling constants and g-tensors can be estimated from mapping procedures with spin Hamiltonians and spin projection methods [230,231]. As shown in numerous studies, the BS-DFT approach can provide reasonable geometries, energies and spectroscopic parameters. This has been demonstrated, *e.g.*, by the calculation of g- and hyperfine tensors for 2-Fe Ferredoxin model systems [232].

Kuramochi *et al.* reported a DFT study on model peroxidase compounds I and II including predictions of isotropic and dipolar hyperfine coupling constants and Mößbauer quadrupole splittings [233]. However, to the best of our knowledge, this is the first comprehensive theoretical study on spin Hamiltonian parameters for Cpd I of P450 enzymes.

This chapter is organized as follows. In section 6.2, we give an overview of the basic theoretical concepts that make the connection between the experimentally determined spectroscopic parameters and electronic structure calculations. Computational details of our approach are given in section 6.3, while the results are presented and discussed in section 6.4. A summary of the present study can be found in section 6.5.

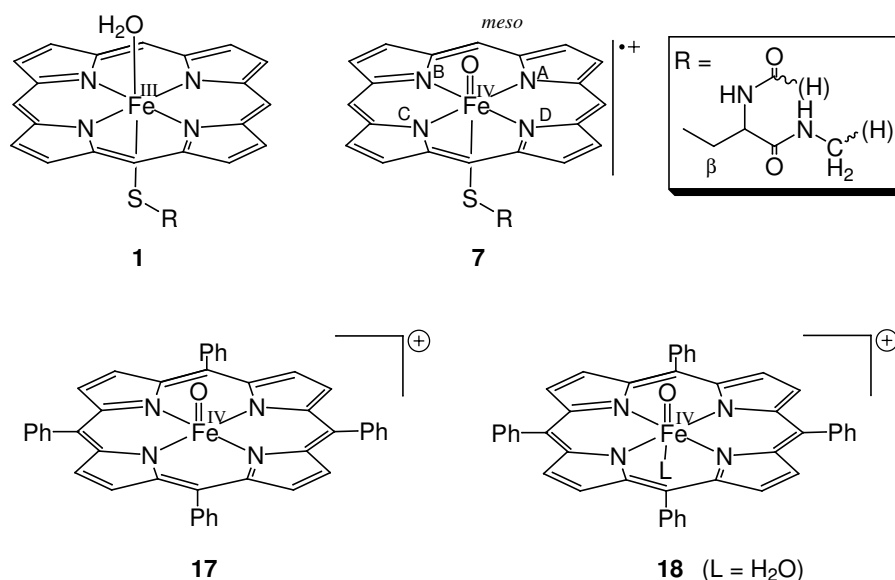


Figure 6.1: Compounds for theoretical EPR and Mößbauer studies: ferric aqua complex **1**, compound **I** (**7**), and the ferryl *meso*-tetraphenyl-porphyrin complexes **17** and **18**.

6.2 Theoretical Background

6.2.1 The Spin Hamiltonian

In order to transform the immediate results of an EPR measurement into a simple and unified language, the corresponding spectra are commonly interpreted in terms of a phenomenological spin Hamiltonian (SH) \hat{H}_{spin} . The spin Hamiltonian itself does not refer explicitly to the geometric or electronic structure of the compound under investigation, instead the spectral features of a molecule are expressed in terms of a few free parameters – **D**, **g**, **A** – that are extracted from the raw experimental data by fitting procedures [234].

$$\hat{H}_{\text{spin}} = \hat{\mathbf{S}}\mathbf{D}\hat{\mathbf{S}} + \beta\mathbf{B}\mathbf{g}\hat{\mathbf{S}} + \sum_A \hat{\mathbf{S}}\mathbf{A}^{(A)}\hat{\mathbf{I}}^{(A)} - \beta_N g_N^{(A)}\hat{\mathbf{I}}^{(A)}\mathbf{B} \quad (6.1)$$

Here, β is the Bohr magneton (1.3316×10^{-4} MHz T⁻¹), $\hat{\mathbf{S}}$ is the operator for total electron spin, $\hat{\mathbf{I}}^{(A)}$ is the operator for the nuclear spin of nucleus *A*, β_N is the nuclear magneton (7.2521×10^{-8} MHz T⁻¹) and $g_N^{(A)}$ the nuclear *g*-value¹. The first term in equation 6.1 describes the zero-field splitting, which arises from interelectronic spin–spin interactions, the second term represents the interaction of the net electron spin magnetic moment with the external magnetic field **B** (the

¹ The nuclear *g*-value is taken to be scalar in equation 6.1, with a value corresponding to that of the free atom. This is not true in reality, because in molecules deviations from this value occur, which give rise to the phenomenon of “chemical shielding”. This effect is of fundamental importance in nuclear magnetic resonance (NMR) spectroscopy, it is, however, negligible in the context of EPR experiments.

electronic Zeeman effect), and the third term describes the magnetic interactions between the electron magnetic moment and the nuclear spin magnetic moments. The last term refers to the interaction of the nuclear spin magnetic moments with the external magnetic field and is of no concern in the following discussion.

The SH parameters that are of interest in the present work are the elements of the 3×3 matrices \mathbf{g} and $\mathbf{A}^{(A)}$. Many empirical or semi-empirical relationships exist that explain trends in the measured SH parameters for different classes of substances [235, 236]. Quantum mechanical calculations on the other hand can provide a quantitative relation between the experimental parameters and the geometric and electronic structure of a given compound, by predicting the SH parameters from first principles. The standard Born-Oppenheimer (BO) Hamiltonian does not include the terms that are necessary to describe the lifting of the degeneracy of the magnetic sublevels and thus to simulate magnetic spectra. By assuming that the interactions between the magnetic moments of the particles themselves and the particles and the external magnetic field are small, they may be introduced as perturbing operators to the standard BO Hamiltonian. In the following, the necessary perturbing operators are introduced in an *ad hoc* fashion, since their derivation is beyond the scope of this thesis.

6.2.2 Magnetic Perturbing Operators

The formulations of the magnetic perturbing operators given in this section are based on ref. [237] and references therein.

The interaction between the external magnetic field \mathbf{B} and the magnetic moment caused by the orbital motion of the electrons is described by the orbital Zeeman operator \hat{H}_{LB} ².

$$\hat{H}_{\text{LB}} = \frac{\alpha}{2} \mathbf{B} \hat{\mathbf{L}}, \quad (6.2)$$

where α is the fine structure constant ($\alpha = c^{-1}$ in atomic units) and $\hat{\mathbf{L}}$ is the electron angular momentum operator.

The operators necessary to describe interactions including electron spin can be derived from the Dirac equation and its generalization to many particle systems (in the Breit-Pauli approximation) [234]. The spin Zeeman operator \hat{H}_{SB} describes the coupling of the electron spin magnetic moment to the external magnetic field.

$$\hat{H}_{\text{SB}} = \frac{\alpha}{2} g_e \mathbf{B} \hat{\mathbf{S}} \quad (6.3)$$

² The expression for this operator can be derived by substituting the electron momenta \hat{p}_i in the standard Born-Oppenheimer Hamiltonian by the expression $\hat{\pi}_i = \hat{p}_i + \alpha A_i$, with α being the fine structure constant and A_i the vector potential for an external magnetic field.

Here, g_e is the free-electron g-value (2.002319...). The fully relativistic treatment introduces a kinetic energy correction to the spin Zeeman energy (the relativistic mass correction operator) that is given by the expression

$$\hat{H}_{\text{SB}}^{\text{RMC}} = \frac{\alpha^3}{2} g_e \sum_i \nabla_i^2 \mathbf{B} \hat{\mathbf{s}}_i, \quad (6.4)$$

where $\hat{\mathbf{s}}_i$ is the spin operator of the i th electron.

The largest and most important perturbation comes from the spin-orbit coupling (SOC) operator. It describes the coupling of the electron spin magnetic moment to the orbital magnetic moment of the same electron as well as of the other electrons. In the Breit-Pauli approximation this operator is given by eqs. 6.5 to 6.9. It consists of a one-electron (eq. 6.6) and a two-electron (eq. 6.7) part.

$$\hat{H}_{\text{SO}} = \hat{H}_{\text{SO}}^{(1)} + \hat{H}_{\text{SO}}^{(2)} \quad (6.5)$$

$$\hat{H}_{\text{SO}}^{(1)} = \frac{\alpha^2}{2} \sum_A \sum_i \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}_i|^3} \hat{\mathbf{l}}_i^A \hat{\mathbf{s}}_i \quad (6.6)$$

$$\hat{H}_{\text{SO}}^{(2)} = -\frac{\alpha^2}{2} \sum_i \hat{\mathbf{s}}_i \sum_{j \neq i} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|^3} \{\hat{\mathbf{l}}_i^j + 2\hat{\mathbf{l}}_j^i\} \quad (6.7)$$

$$\hat{\mathbf{l}}_i^A = (\mathbf{r}_i - \mathbf{R}_A) \times \hat{\mathbf{p}}_i \quad (6.8)$$

$$\hat{\mathbf{l}}_i^j = (\mathbf{r}_i - \mathbf{r}_j) \times \hat{\mathbf{p}}_i \quad (6.9)$$

Here, \mathbf{r}_i (\mathbf{R}_A) is the position vector of the i th electron (A th nucleus), and the symbol $\hat{\mathbf{l}}_i^A$ ($\hat{\mathbf{l}}_i^j$) denotes the angular momentum of electron i relative to nucleus A (electron j). The two-electron part (eq. 6.7) of the spin-orbit coupling operator is difficult to evaluate, and it is therefore commonly approximated by an effective one-electron operator. This approximation is rationalized by the finding that for atoms a one-electron formulation using an effective spin-orbit coupling constant ζ' [238] leads to good agreement with experiment for 2p and 3d elements. In the case of molecules, it is assumed that the two electron part provides a “screening” of the nuclear charge, which can be incorporated into $\hat{H}_{\text{SO}}^{(1)}$ through an effective nuclear charge. Because of the r^{-3} dependence (eq. 6.6), $\hat{H}_{\text{SO}}^{(1)}$ is of rather local nature, and it thus seems reasonable to consider only one-center matrix elements of this operator. The effective operator may thus be written as

$$\hat{H}_{\text{SO}} = \sum_A \sum_i \xi(r_{iA}) \hat{\mathbf{l}}_i^A \hat{\mathbf{s}}_i \quad (6.10)$$

$$\xi(r_{iA}) = \frac{\alpha^2}{2} \frac{Z_{\text{eff}}^A}{r_{iA}^3}, \quad (6.11)$$

where $r_{iA} = |\mathbf{r}_{iA}| = |\mathbf{r}_i - \mathbf{R}_A|$. The functions $\xi(r_{iA})$ contain effective nuclear charges Z_{eff}^A , which are semi-empirical parameters that are determined with reference to experiment or higher level calculations. For *ab initio* calculations a set of parameters Z_{eff}^A has been proposed by Koseki *et al.* [239–241]. Alternative approximations to treat spin-orbit coupling are clearly possible and

include, *e.g.* the “molecular potential” approach or the use of effective core potentials [237]. An accurate method to efficiently approximate the full Breit-Pauli treatment of spin-orbit coupling is the “atomic mean field” method introduced by Hess, Marian *et al.* [242]. In this approach, the two-electron interaction of the spin-orbit coupling operator (eq. 6.7) is approximated by a simpler pseudo-single particle operator, similar to the treatment of electron-electron repulsion in Hartree–Fock theory.

In the present formulation (eqs. 6.10, 6.8) the momentum $\hat{\mathbf{p}}_i$ has been used instead of the gauge invariant momentum $\hat{\boldsymbol{\pi}}_i$. Therefore, an additional gauge correction for the effective SOC operator has to be introduced. This correction term can be written as

$$\hat{H}_{SO}^{GC} = \frac{\alpha}{2} \sum_{i,A} \xi(r_{iA}) \{ (\hat{\mathbf{s}}_i \mathbf{B})(\mathbf{r}_{iA} \mathbf{r}_i) - (\hat{\mathbf{s}}_i \mathbf{r}_i)(\mathbf{B} \mathbf{r}_{iA}) \}. \quad (6.12)$$

For the interaction of nuclear and electronic magnetic momenta, three terms have to be considered. These are the (classical) magnetic dipole-dipole interaction of electron and nuclear spin \hat{H}_{SI}^d , the well-known Fermi contact interaction \hat{H}_{SI}^c and the coupling of nuclear magnetic moments with the orbital magnetic moment of electrons \hat{H}_{LI} . These terms read

$$\hat{H}_{SI}^d = \frac{\alpha}{2} g_e \beta_N \sum_A g_N^{(A)} \sum_i \frac{\hat{\mathbf{s}}_i \hat{\mathbf{I}}^{(A)}}{r_{iA}^3} - 3 \frac{(\hat{\mathbf{s}}_i \mathbf{r}_{iA})(\hat{\mathbf{I}}^{(A)} \mathbf{r}_{iA})}{r_{iA}^5} \quad (6.13)$$

$$\hat{H}_{SI}^c = \frac{\alpha}{2} \frac{8\pi}{3} g_e \beta_N \sum_A g_N^{(A)} \sum_i \hat{\mathbf{s}}_i \hat{\mathbf{I}}^{(A)} \delta(r_{iA}) \quad (6.14)$$

$$\hat{H}_{LI} = \frac{\alpha}{2} \beta_N \sum_A g_N^{(A)} \sum_i \frac{\hat{\mathbf{I}}_i^A \hat{\mathbf{I}}^{(A)}}{r_{iA}^3} \quad (6.15)$$

6.2.3 Formulation of Spin Hamiltonian Parameters

The operators introduced in the previous section describe the physics behind the empirical parameters \mathbf{g} and \mathbf{A} that are employed in the spin Hamiltonian concept. In order to make the connection between the spin Hamiltonian parameters and electronic structure theory, one may use a perturbation theory formulation [243],

$$\hat{H} = \hat{H}_0 + \hat{H}_1 \quad (6.16)$$

where the various magnetic operators are added as perturbing operators (\hat{H}_1) to the standard Born–Oppenheimer Hamiltonian (\hat{H}_0). An elegant route to the formulation of the SH parameters in quantum mechanical terms is the “effective Hamiltonian” treatment suggested by McWeeny [243]. The eigenstates of the unperturbed system are represented by a set of wave functions of the general form $|aSM\rangle$, where S is the total spin quantum number and M the projection onto the z-axis. The index a is a compound label that denotes the eigenstate of the standard Born–Oppenheimer

Hamiltonian. It also incorporates implicitly the spatial irreducible representation Γ and its components M_Γ . Thus, there are $\dim(\Gamma)(2S + 1)$ degenerate states with energy E_a .

The secular equations arising from the variation principle can then be transformed to the following expression [237, 243]

$$\begin{aligned} \langle 0SM | \hat{H} | 0SM' \rangle &= E_0 \delta_{MM'} + \langle 0SM | \hat{H}_1 | 0SM' \rangle \\ &- \sum_{a>0, S', M''} \Delta_a^{-1} \langle 0SM | \hat{H}_1 | aS' M'' \rangle \langle aS' M'' | \hat{H}_1 | 0SM' \rangle, \end{aligned} \quad (6.17)$$

which is analogous to the result of second-order Rayleigh-Schrödinger perturbation theory. Here, $\Delta_a = E_a - E_0$, the energy difference between the a th excited state and the ground state.

The various matrix elements arising from this treatment, with the perturbing operators discussed above (see section 6.2.2), may be grouped such that they match one by one the matrix elements of the spin Hamiltonian (eq. 6.1) in the same basis of states, and they can thus be identified with the elements of the matrices **D**, **g** and **A** [237]. The many electron wave functions $|aSM\rangle$ are normally expressed in terms of molecular orbitals ψ . Since the molecular orbitals are in turn linear combinations of atomic orbitals φ (LCAO ansatz), most of the matrix elements may be reduced to integrals of atomic orbitals over the perturbing operators. In this form, the resulting matrix elements can be implemented in combination with any semi-empirical, density functional, or *ab initio* method. Many of these matrix elements are however difficult to evaluate, therefore specialized approaches have been developed, particularly in the field of DFT (*vide infra*).

The g-tensor. By definition, only terms that are bilinear in **B** and **S** contribute to the matrix **g** (see eq. 6.1). Three contributions arise in first order perturbation theory. The first one comes from the spin Zeeman interaction (eq. 6.3), which simply yields the free electron *g*-value. The other two terms originate from the relativistic mass correction (eq. 6.4) and the gauge correction to the spin-orbit coupling operator (eq. 6.12), respectively. These three contributions are computationally easy to evaluate, however, they are generally small and tend to cancel each other. The explicit terms read:

$$g_{rs}^{\text{SB}} = g_e \delta_{rs} \quad (6.18)$$

$$\Delta g^{\text{RMC}} = -\frac{\alpha^2}{S} \sum_{\mu, \nu} P_{\mu\nu}^{\alpha-\beta} \langle \varphi_\mu | \hat{T} | \varphi_\nu \rangle \quad (6.19)$$

$$\Delta g_{rs}^{\text{GC}} = \frac{1}{2S} \sum_{\mu, \nu} P_{\mu\nu}^{\alpha-\beta} \langle \varphi_\mu | \sum_A \xi(r_A) \{ \mathbf{r}_A \mathbf{r}_O - r_{A,r} r_{O,s} \} | \varphi_\nu \rangle \quad (6.20)$$

In general, indices p, r, s denote spatial components (x, y, z). Here, α is the fine structure constant, $P_{\mu\nu}^{\alpha-\beta}$ is the spin density matrix, $\hat{T} = -1/2\nabla^2$ is the kinetic energy operator and $\{\varphi\}$ is the basis set. In equation 6.20, \mathbf{r}_A is the position vector of the electron relative to nucleus A , \mathbf{r}_O the position vector relative to the gauge origin and $\xi(r_A)$ is the screening function as defined in equation 6.11.

The fourth contribution to the g-tensor emerges in second order perturbation theory as a cross term between the orbital Zeeman operator \hat{H}_{LB} (OZ, eq. 6.2) and the SOC operator \hat{H}_{SO} (eq. 6.11). In almost all cases it accounts for the dominant contribution. Inserting $\hat{H}_{\text{LB}} + \hat{H}_{\text{SO}}$ as perturbing operators into the perturbation theory formulation, eq. 6.17, and dropping all terms that are not bilinear in **B** and **S** leads to the expression

$$\begin{aligned} \Delta g_{rs}^{\text{OZ/SOC}} = & \frac{1}{2S} \sum_{i(\text{doubly})} \sum_{o(\text{singly})} \Delta_{i;o}^{-1} (L_{2r}^{io} L_{1s}^{io} + L_{1r}^{io} L_{2s}^{io}) \\ & - \frac{1}{2S} \sum_{a(\text{empty})} \sum_{o(\text{singly})} \Delta_{o;a}^{-1} (L_{2r}^{oa} L_{1s}^{oa} + L_{1r}^{oa} L_{2s}^{oa}) \end{aligned} \quad (6.21)$$

where

$$L_{1p}^{ij} = \text{Im} \langle \psi_i | \sum_A \xi(r_A) l_p^A | \psi_j \rangle \quad (6.22)$$

$$L_{2p}^{ij} = \text{Im} \langle \psi_i | l_p | \psi_j \rangle. \quad (6.23)$$

Here, ψ_i are molecular orbitals. The symbol $\Delta_{i;j}$ represents the energy difference between two configurational state functions (CSF)³ that differ in the occupation of molecular orbitals i and j . It is important to note that these energy denominators are not simply given by orbital energy differences but contain corrections in the form of two-electron integrals. In equation 6.21, the orbital index i is used for doubly occupied, o for singly occupied, and a for unoccupied orbitals. Because the orbital angular momentum operator is diagonal in electron spin, the sum over excited states contains only states of the same spin as the ground state. It can be readily shown that only excitations from doubly to singly occupied orbitals and from singly occupied to empty orbitals contribute, since in other cases either the SOC matrix element with the ground state, or the orbital Zeeman matrix element vanishes.

In summary, the g matrix elements may be written as a sum of four contributions [234, 243]:

$$g_{rs} = g_e \delta_{rs} + \Delta g_{rs}^{\text{RMC}} \delta_{rs} + \Delta g_{rs}^{\text{GC}} + \Delta g_{rs}^{\text{OZ/SOC}} \quad (6.24)$$

The terms $\Delta g_{rs}^{\text{RMC}}$ and $\Delta g_{rs}^{\text{GC}}$ are readily calculated from the ground state spin density. In contrast, the term $\Delta g_{rs}^{\text{OZ/SOC}}$ (eq. 6.21) requires the explicit generation of a sufficient number of excited state wave functions and associated energies, which may be achieved by configuration interaction (CI) calculations. For instance, a recent implementation of a highly accurate yet efficient MR-CI approach to g-tensors that uses the atomic mean-field method to spin-orbit coupling is described in ref. [244].

The A-tensor. The relevant expressions for the calculation of hyperfine interactions contain terms proportional to the product of the electron spin operator and the nuclear spin operator. First order

³ The energies of the CSFs can be chosen as the expectation value of the BO operator.

contributions arise from the (classical) anisotropic magnetic dipole–dipole interaction of electron and nuclear spin \hat{H}_{SI}^d (eq. 6.13) and from the well-known isotropic Fermi contact interaction \hat{H}_{SI}^c (eq. 6.14).

The isotropic Fermi contribution reads:

$$\begin{aligned} A^{(A;\text{iso})} &= \frac{8\pi}{3} \frac{1}{2S} \frac{\alpha}{2} g_e \beta_N g_N^{(A)} \sum_{\mu,\nu} P_{\mu\nu}^{\alpha-\beta} \langle \varphi_\mu | \delta(\mathbf{r} - \mathbf{R}_A) | \varphi_\nu \rangle \\ &= \frac{2\pi\alpha}{3S} g_e \beta_N g_N^{(A)} \rho^{\alpha-\beta}(\mathbf{R}_A) \end{aligned} \quad (6.25)$$

This shows that the Fermi term is directly proportional to the spin density $\rho^{\alpha-\beta}$ at the considered nucleus.

The dipolar contribution can be written as:

$$A_{rs}^{(A;d)} = \frac{\alpha}{4S} g_e \beta_N g_N^{(A)} \sum_{\mu,\nu} P_{\mu\nu}^{\alpha-\beta} \langle \varphi_\mu | r_A^{-5} \{ \delta_{rs} r_A^2 - 3r_{A,r} r_{A,s} \} | \varphi_\nu \rangle \quad (6.26)$$

Hence, the dipolar HFC tensor can be calculated from the spin density matrix and the field gradient integrals.

The remaining contribution arises in second order as a cross term between the spin-orbit coupling operator \hat{H}_{SO} (eq. 6.11) and \hat{H}_{LI} (eq. 6.15), which describes the coupling of nuclear magnetic moments with the orbital magnetic moment of electrons⁴. By applying the perturbation treatment discussed above (eq. 6.17) one finds

$$\begin{aligned} A_{rs}^{(A;\text{SOC})} &= \frac{1}{2S} \frac{\alpha}{2} \beta_N g_N^{(A)} \{ \sum_{i(\text{doubly})} \sum_{o(\text{singly})} \Delta_{i;o}^{-1} (L_{3r}^{io} L_{1s}^{io} + L_{1r}^{io} L_{3s}^{io}) \\ &\quad - \sum_{o(\text{singly})} \sum_{a(\text{empty})} \Delta_{o;a}^{-1} (L_{3r}^{ao} L_{1s}^{ao} + L_{1r}^{ao} L_{3s}^{ao}) \} \end{aligned} \quad (6.27)$$

where

$$L_{3p}^{ij} = \text{Im} \langle \psi_i | l_p^A r_A^{-3} | \psi_j \rangle. \quad (6.28)$$

For molecules containing only light nuclei, it is a reasonable approximation to neglect this term, which, however, can make a significant contribution in the case of transition metals [245–248].

⁴ There is another spin-orbit contribution to hyperfine interactions, which arises as a cross term between the spin-orbit coupling operator (eq. 6.11) and the magnetic-dipole-dipole operator (eq. 6.13). However, this term does not reduce to the form $\hat{\mathbf{S}}\mathbf{A}\hat{\mathbf{I}}$, and hence it should not be included into the spin-Hamiltonian parameter \mathbf{A} . For a discussion see ref. [245].

6.2.4 Calculation of Second Order Contributions by a Double Perturbation Approach

Using the sum-over-states formulation discussed above, the second order contributions to the spin-Hamiltonian parameters g and A may be calculated quite accurately, *e.g.*, by a CI treatment. However, this method is computationally very challenging and currently not applicable to larger molecules. A more practical approach in connection with DFT is the use of coupled perturbed SCF theory. In general, a second order property may be expressed as a (mixed) second derivative of the total energy with respect to one or two perturbations. Thus, the contribution to the g -tensor is defined as the derivative with respect to the magnetic moment μ of the electron and the external magnetic field B , while the contribution to the A -tensor is obtained by differentiation with respect to the electron spin and the nuclear spin:

$$\Delta g_{rs}^{\text{OZ/SOC}} = \frac{1}{\beta S} \frac{\partial^2 E}{\partial B_r \partial \mu_s} \quad (6.29)$$

$$A_{rs}^{A;\text{SOC}} = \frac{\partial^2 E}{\partial S_r \partial I_s^{(A)}} \quad (6.30)$$

These derivatives may be calculated by double perturbation theory [243]. In the following, we give a very brief outline of this procedure for the case of hybrid DFT in the unrestricted formalism. Detailed expositions are available in the literature [237, 243, 249]. Coupled perturbed SCF theory may be employed to calculate the first order wave function from the linear response of the Kohn–Sham ground state to one of the two possible perturbations⁵. The perturbation is provided by the orbital Zeeman operator (\hat{H}_{LB} , eq. 6.2) in the case of the g -tensor [249–253], and by the nucleus-orbit interaction (\hat{H}_{LI} , eq. 6.15) in the case of the A -tensor [245]. The result are the first order wave function coefficients $U^{\sigma;r}$. The symbol σ denotes the spin of either the α or β spin orbitals. Having obtained these coefficients the second order properties are calculated as:

$$\begin{aligned} \Delta g_{rs}^{\text{OZ/SOC}} = & -\frac{1}{\beta S} \sum_{\sigma=\alpha,\beta} (-1)^{\delta_{\sigma\beta}} \sum_{i,a \in \sigma} \text{Im}(U_{ia}^{\sigma;r}) \\ & \times \text{Im} \left(\langle \psi_i^{\sigma(0)} | \sum_A \xi(r_A) l_s^A | \psi_a^{\sigma(0)} \rangle \right) \end{aligned} \quad (6.31)$$

and

$$\begin{aligned} A_{rs}^{(A);\text{SOC}} = & -\frac{1}{S} \frac{\alpha}{2} \beta_N g_N^{(A)} \sum_{\sigma=\alpha,\beta} (-1)^{\delta_{\sigma\beta}} \\ & \times \sum_{i,a} U_{ia}^{\sigma;r} \text{Im} \langle \psi_i^{\sigma(0)} | \sum_A \xi(r_A) l_s^A | \psi_a^{\sigma(0)} \rangle \end{aligned} \quad (6.32)$$

Here, labels i, j refer to occupied and a, b to unoccupied orbitals.

⁵ The interchange theorem proves that the order in which the two perturbations are applied is immaterial [243].

In this way, (infinite) summations over excited states, as in eqs. 6.21 and 6.27, are completely avoided. In practice, the solution of the coupled perturbed SCF equations is the computational bottleneck in the evaluation of $\Delta g_{rs}^{\text{OZ/SOC}}$ and $A_{rs}^{\text{A;SOC}}$, which requires approximately the same time as the SCF calculation itself.

We note that another approach to spin-orbit contributions is to include spin-orbit coupling directly into the electronic structure calculation *via* a relativistic DFT method [254, 255]. In this case, first order perturbation theory is sufficient to calculate all magnetic parameters. This approach is superior to double perturbation methods in the case of heavier transition metals, where relativistic effects such as spin-orbit coupling are important. Studies using the relativistic zeroth-order regular approximation (ZORA) DFT method have highlighted the importance of the SOC contributions to metal hyperfine interactions [246–248]. Unfortunately the ZORA approach is currently available only in spin-restricted form and for a single unpaired electron.

6.2.5 Practical Aspects of the Computations

g-tensor. In practical calculations which necessarily employ finite basis sets, the orbital angular momentum operator introduces an unphysical gauge dependence of the g-tensor. As a consequence, predictions for the g-tensor may change with different choices for the gauge origin. This problem is well-known in the calculation of magnetic properties, such as NMR chemical shifts, magnetizabilities and g-tensors. Of these three properties, the g-tensor is least sensitive with respect to the chosen gauge origin [237]. Thus, if a good choice for the gauge origin is made, the deviations from the exact result are in most cases negligible. This is referred to as the “common gauge” approach. Luzanov *et al.* have shown that the best choice for a common gauge origin is the center of electronic charge, which leads to g-values that effectively approximate the fully gauge invariant results [256], particularly when using larger basis sets [237, 256]; triple- ζ quality is usually sufficient in practice. A more rigorous treatment is to introduce magnetic field dependent phase factors into the electronic wave function. The two most popular approaches of this kind are the GIAO (Gauge Including Atomic Orbitals) method, that uses London atomic orbitals [257], and the IGLO (Individual Gauge for Localized Orbitals) [258] method, where the phase dependence is introduced into the molecular orbitals, which have to be localized according to established criteria.

A-Tensor. The isotropic HFC contribution is directly proportional to the spin density at the position of the nucleus. Even if the formally singly occupied orbitals do not carry significant s character, isotropic hyperfine coupling interactions can arise from subtle mechanisms of spin polarization [259]. A realistic description of spin polarization effects by the theoretical model is thus extremely important. Therefore, basis functions with high flexibility in the core region are required, which means that inner s functions should at least be decontracted. Barone and co-

workers have specifically designed basis sets for this purpose, denoted EPR-II and EPR-III [260]. For the prediction of metal hyperfine coupling constants, standard Gaussian type basis functions additionally need to be enhanced with primitive Gaussians with very large exponents in order to mimic the cusp behavior of the true wave function near the nuclei. Since spin polarization is crucial, spin-restricted wave functions cannot be used for the prediction of metal hyperfine interactions. For the treatment of large molecules, unrestricted DFT calculations are usually the method of choice, although pure LDA and GGA functionals tend to underestimate spin polarization. Thus, Munzarova and Kaupp have shown by comparison to high-level *ab initio* calculations, that present-day density functionals tend to underestimate the core level spin polarization which is crucial for obtaining accurate spin densities at the metal nucleus [261]. Hybrid functionals benefit from some error compensation, since the Hartree–Fock method overestimates spin polarization and the admixture of exact exchange in hybrid functionals apparently adjusts the isotropic coupling constants in the right direction. Care has to be taken, however, when spin polarization is caused by significant spin contamination, because the DFT results then tend to become unreliable [261].

For the prediction of dipolar hyperfine coupling constants, standard basis sets usually are sufficient, as long as they are flexible enough, preferably of triple- ζ quality.

6.2.6 Calculation of Mößbauer Parameters

The Mößbauer effect consists of the recoilless emission and resonant absorption of γ -radiation due to associated changes in the state of the atomic nuclei. In biology, this effect is especially useful to study iron containing systems, where the $I = 1/2 \rightarrow I = 3/2$ transition in the ^{57}Fe isotope is probed [262]. Of particular interest is the possibility to combine Mößbauer spectroscopy with rapid freeze quench techniques to study short lived intermediates. The energy of the transition is dependent on the immediate environment of the absorbing atom. The absorption wave length with respect to a given standard (*i.e.*, iron foil) is called the isomer shift δ . The quadrupole splitting ΔE_Q is due to the fact that the $M_I = \pm 1/2, \pm 3/2$ components of the $I = 3/2$ excited state are split in energy by the electric quadrupole interaction. The physical origin of this effect is the interaction of the nuclear quadrupole moment with the electric field gradient at the nucleus. Both quantities, the isomer shift and the quadrupole splitting, give valuable information about the spin and valence state of the metal and about the nature of the metal–ligand bonds.

The theoretical prediction of Mößbauer properties requires an accurate description of the electron density around the nucleus. Density functional theory studies [263–266] have demonstrated that the method is capable of predicting these properties in very good agreement with experiment [267]. Similar to the calculation of isotropic hyperfine coupling constants, basis sets with enhanced flexibility in the core region seem to be required [266]. The isomer shift δ is proportional

to the electron density at the nucleus $\rho(0)$ according to equation 6.33:

$$\delta = \frac{4}{5}\pi Z S(Z) e_0^2 R^2 \left(\frac{\delta R}{R} \right) [\rho_0^A(0) - \rho_0^S(0)] \quad (6.33)$$

Here, Z is the nuclear charge of the Mößbauer absorbing atom, $S(Z)$ is a relativistic correction factor, e_0 is the elementary charge, R is the average value of the radii of the nucleus in the ground and excited state, and δR is the difference of the two radii. $\rho_0^A(0)$ is the non-relativistic electron density at the nucleus of the absorbing atom and $\rho_0^S(0)$ is the same quantity for a given standard. Because all quantities except for the electron density of the absorber are constant for a given isotope, one may simplify this expression to equation 6.34:

$$\delta = a[\rho_0^A(0) - b] \quad (6.34)$$

The theoretical approach to isomer shifts has been to construct a linear correlation plot according to equation 6.34 between electron densities at the nucleus calculated from DFT or MO theory and experimentally measured Mößbauer data [264–266]. Given the parameters a and b , predictions for other iron-containing systems can be made. The slope and intercept of the linear correlation line depend on the density functional and basis set used, and as such are not universal. A correlation for the popular B3LYP functional and standard Gaussian-type basis sets has been reported by Neese [266].

To calculate the quadrupole splitting, the electric field gradient V at the position of the iron nucleus needs to be evaluated. This quantity is relatively easy to compute and enters the expression of the quadrupole splitting as

$$\Delta E_Q = \frac{1}{2} e_0 Q V_{zz} \left(1 + \frac{\eta^2}{3} \right)^{1/2}, \quad (6.35)$$

where Q is the quadrupole moment of the excited state, and the components of the electric field gradient are labeled according to the convention

$$|V_{zz}| > |V_{yy}| > |V_{xx}| \quad (6.36)$$

with the asymmetry parameter

$$\eta = (V_{xx} - V_{yy})/V_{zz}. \quad (6.37)$$

An uncertainty concerns the value for the quadrupole moment of ^{57}Fe . Values in the literature vary from ≈ 0.1 barn⁶ to ≈ 0.3 barn, however, 0.15 barn is most commonly adopted and gives results in good accord with experiment. Systematic DFT studies using the B3LYP density functional of quadrupole splittings were done by Oldfield *et al.* for a variety of heme proteins [268–270]. A study of compound I-type intermediates in heme proteins demonstrated that the calculated

⁶ The dimension of the electric quadrupole moment is charge times area, however, commonly the electronic charge is not explicitly given.

Mössbauer parameters for the $^4A_{1u}$ and $^4A_{2u}$ π -cation radical states are similar and in good agreement with experiment [233]. These studies indicate that DFT is a reliable tool for the prediction of quadrupole splittings.

6.3 Computational Details

6.3.1 Geometries

Compound I. The QM region used in the Cpd I study is shown in figure 6.1. The protoporphyrin IX in heme was represented by a porphyrin, while the axial sulfur ligand was taken as the full Cys357 residue, the CO group of Leu356, and the NH-C $_{\alpha}$ unit of Leu358.

Chapter 3 gives details about the geometry optimization of this system in the enzyme environment as well as in the gas phase (QM region R3). We have used the B3LYP/CHARMM22 geometries optimized separately for the LS and HS state of snapshot 40 with basis set B3. The analogous isolated species were optimized in the gas phase on the B3LYP level with basis set B3.

Model compounds. Complexes **17** and **18** were fully optimized without symmetry constraints. These calculations were entirely done with TURBOMOLE [187, 188] and employed the B3LYP functional together with the SV(P) basis set [200] (784 basis functions total). The quartet and the doublet spin state were separately optimized with tight convergence criteria. In agreement with the experimental assignment [224], we found the quartet state to be the ground state, being more stable than the doublet by *ca.* 0.1 kcal/mol (*vide infra*).

Resting State. The ferric aqua complex considered in the present study is shown in figure 6.1. The corresponding QM/MM geometry optimizations were described in chapter 5. Here we use the geometry of snapshot 195 optimized at the B3LYP/CHARMM22 level, employing the same QM system (R3w) with basis set B4. Only the ground state, *i.e.* the doublet state, was considered.

6.3.2 Spectroscopic Parameters

^{57}Fe , ^{14}N , ^{17}O and ^1H isotropic hyperfine coupling constants (HFCs) A^{iso} and the traceless anisotropic HFCs A^{d} are calculated directly from Fermi contact terms (eq. 6.25) and classical dipolar contributions (eq. 6.26), respectively. The principal values of the total hyperfine coupling tensor A^{tot} are obtained from addition of these two contributions. Furthermore, the ^{57}Fe spin-orbit contribution of the hyperfine interaction was calculated, using double-perturbation theory (eq. 6.32) [245]. Its isotropic contribution was added to the ^{57}Fe isotropic Fermi term A^{iso} , while its

anisotropic pseudo-contact contribution was added to the anisotropic ^{57}Fe tensor A^d . Mößbauer quadrupole splittings ΔE_Q (eq. 6.35) and the asymmetry parameter η (eq. 6.37) were calculated from the electric field gradient components V_{rs} . A nuclear quadrupole moment $Q(^{57}\text{Fe})$ of 0.15 barn was employed. The Mößbauer isomer shift was obtained from the electron density at the iron nucleus ρ_0 with the correlation (eq. 6.34) given by Neese [266] for the B3LYP functional. The g-values were calculated employing double-perturbation theory (eq. 6.31) [249] in conjunction with a parameterized one-electron spin-orbit operator [240, 241]. This approach covers spin-orbit coupling, spin polarization, relativistic mass and diamagnetic gauge correction. The common gauge origin was chosen as the center of electronic charge. The sensitivity of the computed g-tensor with respect to the chosen gauge origin was probed by a series of calculations that used the center of mass, the center of electronic charge, and the iron atom as the gauge origin. The resulting deviations of the individual g values (g_1, g_2, g_3) from the respective averages ($\bar{g}_1, \bar{g}_2, \bar{g}_3$) were found to be less than 4 %.

The standard definition of the spin Hamiltonian for the exchange interaction between two magnetic centers with site spins S_1 and S_2 is given by the relation

$$\hat{H} = -2JS_1S_2 \quad (6.38)$$

where J is the Heisenberg exchange coupling constant. In the broken symmetry formalism, J may be calculated from the energy differences between high spin and broken symmetry states, using the general expression [271]

$$J = -\frac{E_{\text{HS}} - E_{\text{BS}}}{4S_1S_2}. \quad (6.39)$$

Alternatively, Yamaguchi [272] suggested the following expression for the exchange coupling constant:

$$J = -\frac{E_{\text{HS}} - E_{\text{BS}}}{\langle S_{\text{HS}}^2 \rangle - \langle S_{\text{BS}}^2 \rangle}. \quad (6.40)$$

The spin expectation values are obtained from the unrestricted DFT calculation. We compare the results obtained from both formulas. The magnitude of J is a measure for the strength of the exchange interactions and its sign indicates whether exchange coupling is ferromagnetic ($J > 0$) or antiferromagnetic ($J < 0$). The energies E_{HS} and E_{BS} were obtained from single point energies on the geometries optimized for the HS and BS states, respectively, *i.e.* J is based on adiabatic energy differences and is thus directly related to the experimental values.

In the calculation of the hyperfine tensor and the Mößbauer properties, the triply polarized (17s-11p5d1f) \rightarrow [17s7p3d1f] core properties basis set CP(PPP) for iron [266] was used. The other atoms were assigned the Ahlrichs-SV(P) [200] basis set, however, the inner s-functions were left completely uncontracted. For the iron atom, an enhanced integration grid was used and the overall integration accuracy was increased using the `SpecialGridIntAcc=7` keyword. The

g-tensors as well as the exchange coupling constants J were obtained with the TZVP basis [201] at all atoms. In the case of Cpd I, the calculated hyperfine- and g-tensors of the broken-symmetry doublet state were corrected by spin projection as described below⁷.

All calculations were carried out with the ORCA 2.2 program package [273] and employed the B3LYP density functional. This functional has been shown to give better results for transition metal compounds than GGA functionals, which is thought to be due to a better description of metal-ligand covalency [267].

QM/MM single point calculations were done according to the electrostatic embedding scheme (see chapter 3 for details), *i.e.* the outer part of the system was represented by point-charges that entered the one-electron Hamiltonian of the QM calculation.

6.3.3 Spin Projection

Cpd I (7) exhibits two weakly coupled spin systems. One is the $\text{Fe}^{\text{IV}}\text{O}$ moiety with a formal site spin of $S_1 = 1$. The other consists of the porphyrin and cysteinate ligands with formal site spin $S_2 = 1/2$. These systems interact by ferromagnetic or antiferromagnetic coupling, yielding total spins of $S = 3/2$ for the HS state and $S = 1/2$ for the broken symmetry (BS) state. The spin projection coefficients c_1 , c_2 for the exchange coupled system result from spin algebra and may be written as

$$c_1 = \frac{S(S+1) + S_1(S_1+1) - S_2(S_2+1)}{2S(S+1)} \quad (6.41)$$

$$c_2 = \frac{S(S+1) - S_1(S_1+1) + S_2(S_2+1)}{2S(S+1)}. \quad (6.42)$$

In the present case, the spin coupling model results in projection coefficients of $c_1 = 4/3$ and $c_2 = -1/3$. The theoretical foundation of the definition of such coefficients is the Wigner-Eckart theorem associated with the vector model. This model states that two spin vectors S_n are vectorially coupled to give a resultant total spin whose magnitude can be any of the integrally spaced values between $|S_1 - S_2|$ and $|S_1 + S_2|$. The quantity c_n gives the projection of the “local” spin onto the total spin of the system [230, 231, 274].

Correction of hyperfine coupling constants. The broken-symmetry calculated isotropic and anisotropic hyperfine coupling constants a^n corresponding to an atom on site n of the antiferromagnetic doublet system were converted to intrinsic site hyperfine coupling constants $a^{n,\text{site}}$ by [275, 276]

$$a^{n,\text{site}} = \frac{a^n}{\pm 2S_n}. \quad (6.43)$$

⁷ In contrast to the spin density, the electron density of the antiferromagnetic state is well described by the broken symmetry formalism, thus no spin projection had to be applied in the calculation of the Mößbauer parameters.

S_n is the site spin of either the $\text{Fe}^{\text{IV}}\text{O}$ ($S_1 = 1$) subunit or the ligands ($S_2 = 1/2$). The positive sign is applied for all atoms on site 1 (here: $\text{Fe}^{\text{IV}}\text{O}$), the negative sign for those of site 2 (ligands). The final coupling constants A that are related to the experimental values are then calculated employing the spin projection coefficients c_n ⁸:

$$A = c_n a^{n, \text{site}} \quad (6.44)$$

Correction of g-tensors. The concept of spin projection requires the calculation of site values also for the g-tensor. For that purpose, the g site values were extracted from the g-tensors of both the HS and BS states [277]. The HS g-tensor can be decomposed into contributions C_1 and C_2 from sites 1 and 2. Neglecting electronic relaxation, the HS and BS states differ only by a spin flip of the magnetic electron on site 2. Therefore one can assume that the g-tensor of the BS state is made up by the same contributions C_1 and C_2 , differing only in the sign of C_2 due to the rotation of the spin vector on that site:

$$g_{\text{HS}} = \frac{1}{S_{\text{HS}}}(C_1 + C_2) \quad (6.45)$$

$$g_{\text{BS}} = \frac{1}{S_{\text{BS}}}(C_1 - C_2) \quad (6.46)$$

The same assumption underlies the calculation of J-values with the BS-DFT approach. Inserting the g-tensor site values

$$g_1 = \frac{1}{S_1}C_1 \quad (6.47)$$

$$g_2 = \frac{1}{S_2}C_2 \quad (6.48)$$

in eqs. 6.45 and 6.46 yields with $S_{\text{HS}} = S_1 + S_2$ and $S_{\text{BS}} = |S_1 - S_2|$ the following equations for the calculation of the site values from the g-tensors of the HS and BS states:

$$g_1 = \frac{1}{2S_1}[g_{\text{HS}}S_{\text{HS}} + g_{\text{BS}}S_{\text{BS}}] \quad (6.49)$$

$$g_2 = \frac{1}{2S_2}[g_{\text{HS}}S_{\text{HS}} - g_{\text{BS}}S_{\text{BS}}] \quad (6.50)$$

The individual matrix elements of the molecular g-tensors of the antiferromagnetic state were then calculated from the spin projection relation

$$g = 4/3g_1 - 1/3g_2. \quad (6.51)$$

The matrix product $\mathbf{g}^T \mathbf{g}$ was subsequently diagonalized to yield the final g-tensor, *i.e.* the principal g values and the orientation.

⁸ This relation is derived under the assumption that the coupling constant J is large and the zero field splitting D is small with respect to the hyperfine coupling constants.

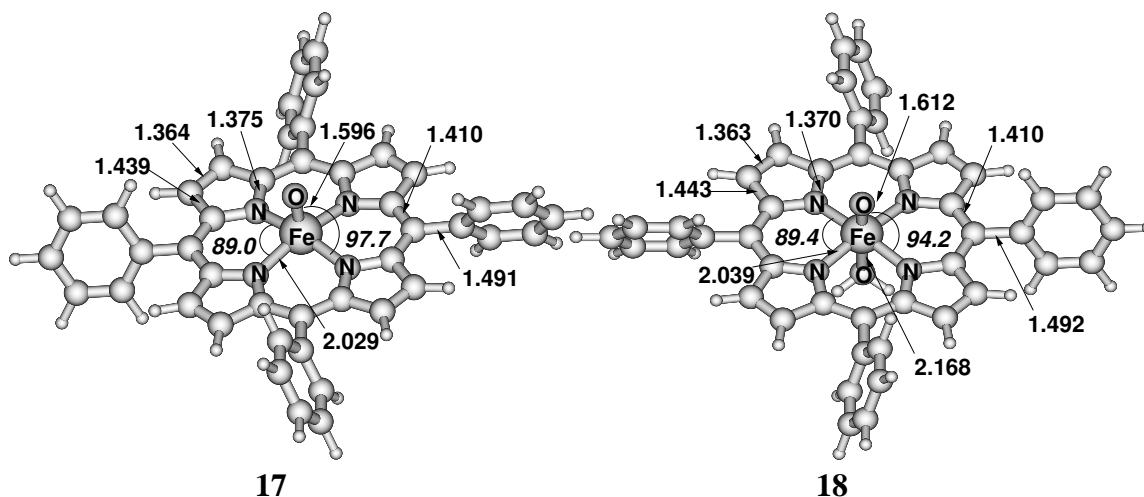


Figure 6.2: Optimized geometries of the iron-oxo *meso*-tetraphenyl model compounds **17** and **18**. Distances in Å, angles (oblique numbers) in degree.

6.4 Results and Discussion

6.4.1 Model Complexes

The geometries of the oxoferryl complexes **17** and **18** optimized in the HS state are displayed in figure 6.2. The corresponding LS geometries are virtually identical. Both systems exhibit short Fe–O distances, indicative of partial double bond character. The calculated Fe–O bond length in **18** (1.612 Å) is slightly larger than in **17** (1.596 Å) and closer to the experimental value of 1.64 Å determined by EXAFS studies on $[\text{Fe}^{\text{IV}}\text{O}(\text{TMP}^+)]$ [53, 278]. The phenyl substituents at the *meso*-carbon atoms are rotated by 25° with respect to the perfectly perpendicular arrangement towards the porphyrin plane. This is apparently a compromise between repulsive interactions with the porphyrin β -hydrogen atoms and the energy gain due to the partial overlap of the π -systems. In **18**, the bond to the other axial ligand, the water molecule, is relatively short (2.17 Å).

In agreement with the experimental assignment [53, 224], both complexes are porphyrin cation radicals, with one unpaired electron in the porphyrin a_{2u} orbital⁹. The spin density of the porphyrin π system delocalizes partly into the adjacent phenyl rings, which underscores that a consideration of these groups is important for the description of the electronic structure of the complexes.

Table 6.1 lists the Mößbauer parameters calculated for models **17** and **18** and compares them with the experimental results on the complexes $[\text{FeO}(\text{TDCPP})]^+$ and $[\text{FeO}(\text{TMP})]^+$ [224]. The quadrupole splitting ΔE_Q is a measure for the deviation from cubic coordination symmetry around the iron atom. Expectedly, the five coordinated species shows the largest splitting of 2.21 mm s^{-1} ,

⁹ Symmetry labels refer to idealized D_{4h} symmetry of the porphyrin.

as opposed to the experimental values of 1.62 and 1.48 mm s⁻¹. This strongly indicates that a sixth ligand to the iron is present in the experiments. Indeed, the water ligated model **18** shows the smallest splitting of 1.16 mm s⁻¹. Hence, the coordination sphere in the true systems is less symmetric than in **18**, which agrees with the notion that the ligand is either methanolate or *meta*-chlorobenzoate [164]. The calculated small Mößbauer isomer shifts $\delta_{57\text{Fe}}$ (0.08 and 0.11 mm s⁻¹) are typical for the HS Fe^{IV} species and accord very well with the experimental data. In contrast, HS Fe^{III} porphyrin complexes exhibit larger isomer shifts, *e.g.*, $\delta_{57\text{Fe}} = 0.46$ mm s⁻¹ for the complex [Fe^{III}(Br)(TMP)] and 0.42 mm s⁻¹ for the corresponding porphyrin π cation radical [Fe^{III}(Br)(TMP^{•+})] [164].

Table 6.1: Mößbauer isomer shift $\delta_{57\text{Fe}}$ [mm s⁻¹], quadrupole splitting ΔE_Q [mm s⁻¹], and asymmetry parameter η for complexes **17** and **18**.

	compound	ΔE_Q [mm s ⁻¹]	η	$\delta_{57\text{Fe}}$ [mm s ⁻¹]
theory	17	2.21	0.0	0.08
theory	18	1.16	0.0	0.11
exp. [224]	[FeO(TDCPP)] ⁺	1.48	0	0.06
exp. [224]	[FeO(TMP)] ⁺	1.62	0	0.08

Up to now, spectral simulations of measured Mößbauer and EPR spectra of compound I analogues have not been possible with a unique parameter set [279]. Instead, a whole parameter space is available due to the fact that several of the parameters are covariant, *i.e.* local *g*-values and zero-field-splitting *D* for iron(IV), and the anisotropy contribution to the exchange interaction J_d . In the experimental study [224], the *g*-tensor of the ferryl iron was related to the zero-field splitting *D* using a perturbation formula derived from a ligand-field treatment by Oosterhuis and Lang [280], assuming a one-electron spin-orbit coupling constant of $\zeta = 500$ cm⁻¹. For the unpaired spin on the ligands, an isotropic *g*-tensor with principal values of 2 was assumed. Moreover, in these complexes, the exchange coupling *J* is of similar magnitude as the zero-field splitting *D*. The competition of zero-field and exchange interaction leads to a mixing of spin states and to a strong dependence of effective *g*-factors on $|J|$ and *D*. Using the abovementioned assumptions, it was possible to simulate the overall EPR features, which provided a sensitive estimate for *D* and *J*. However, due to the covariance of parameters, there is essentially no independent information on the *g*-values. In fact, it is important to note that *D* may contain substantial contributions from excited states with lower or higher multiplicities than the ground state, whereas only excited states of the same multiplicity as the ground state contribute to the *g*-shift [237]. Thus, the assumption of proportionality between *D* and *g* is questionable.

The experimentally determined values for *J* and *g* are compared to the calculated numbers in

Table 6.2: Principal g-values and Heisenberg exchange coupling constants J [cm^{-1}] for complexes **17** and **18**.

	compound	g	$J(1)^a$ [cm^{-1}]	$J(2)^b$ [cm^{-1}]
theory	17	2.010, 2.010, 2.006	+56	+55
theory	18	2.010, 2.010, 2.006	+26	+26
exp. [224]	$[\text{Fe}^{\text{IV}}\text{O}(\text{TDCPP})]^+$	2.20, 2.24, 1.99 ^c		+38
exp. [224]	$[\text{Fe}^{\text{IV}}\text{O}(\text{TMP})]^+$	2.21, 2.23, 1.99 ^c		+43

^a Calculated according to equation 6.39. ^b Calculated according to equation 6.40. ^c Electronic g values of ferryl iron were related to D and E/D via a ligand field model [280], see text.

table 6.2. The values obtained from the different expressions for J (eqs. 6.39 and 6.40) are virtually identical. Most notably, our calculations correctly predict a ferromagnetically coupled spin system with positive sign of J . The magnitude of the calculated Heisenberg exchange coupling constants for compounds **17** and **18** agrees very well with moderately strong coupling found in the experiment.

The calculated g-values indicate an isotropic g-tensor with principle values of *ca.* 2.01. This is in contrast to the fitted g-tensor of the complex $[\text{Fe}^{\text{IV}}\text{O}(\text{TDCPP})]^+$ ($[\text{Fe}^{\text{IV}}\text{O}(\text{TMP})]^+$), which exhibits a rhombic pattern with $g = 2.20, 2.24, 1.99$ (2.21, 2.23, 1.99). As noted above, a definitive experimental determination of g-values has not been possible, hence this poor agreement is not unexpected. However, our calculated g-shifts appear rather small. To check this further, we have calculated the relative energies of the lowest excited states in the simplified oxoferryl model system $[\text{Fe}^{\text{IV}}\text{O}(\text{NH}_3)_4(\text{SH})]^+$ (see section 4.2.4) by means of TDDFT (B3LYP/TZVP) and INDO/S-CI calculations. This system contains the central $\text{Fe}^{\text{IV}}\text{O}$ moiety with spin $S = 1$, which is mainly responsible for the g-shifts in the *meso*-tetraaryl porphyrin substituted parent compounds. The lowest excited states with spin $S = 1$ are found more than 2 eV above the ground state. As shown in equation 6.21, these relative state energies appear in the denominator of the $g_{rs}^{\text{OZ/SOC}}$ term, which makes the dominating contribution to the g-tensor. Hence, one may rationalize the small g-shifts¹⁰ by the weak interaction of the electronic ground state with the excited states of equal spin. Another factor that influences the g-shift is the metal-ligand covalency. Our calculations show that the d-orbitals on iron strongly mix with ligand orbitals. The unrestricted natural orbitals (UNOs) with occupation numbers of approximately 1 electron are a reasonable approximation to the “magnetic” orbitals, *i.e.* those orbitals which make the dominant contributions to the interactions with the external magnetic field. The UNOs that correspond to the FeO d-p π^* orbitals contain only *ca.* 50-60% contributions from iron d-orbitals. This large covalency causes a concomitant reduction of the spin-orbit coupling interactions with respect to the free iron ion. Integrals of molecular orbitals over the effective spin-orbit coupling function ξ and the orbital

¹⁰ The g-shift is the deviation from the free-electron g-value g_e (2.002319).

angular momentum operator directly contribute to $g_{rs}^{OZ/SOC}$ (see eq. 6.21). This effect thus leads, likewise, to a decrease in the calculated g-shifts.

On the basis of these considerations the calculated isotropic g-tensor with $g \approx 2.01$ for **17** and **18** does not seem unrealistic. It would therefore be desirable to re-evaluate the spin Hamiltonian simulations of the spectral data [224] using our computed g-values. On the other hand, one should keep in mind that the precise calculation of the g-tensor for such complex systems remains a challenge to present-day theoretical methods, and one should not expect quantitative accuracy.

Table 6.3: ^{57}Fe Hyperfine coupling constants [MHz] for complexes **17** and **18**.

	compound	A^{iso}	$A_{x,y}^{\text{d}}$	A_z^{d}
theory	17	-10.0	-18.5	36.9
theory	18	-9.7	-16.8	33.1
exp. [224]	$[\text{Fe}^{\text{IV}}\text{O}(\text{TDCPP})]^+$	-18.3	-8.5	17.0
exp. [224]	$[\text{Fe}^{\text{IV}}\text{O}(\text{TMP})]^+$	-18.3	-9.2	18.4

The ^{57}Fe hyperfine coupling tensor determined experimentally shows an axial pattern with an isotropic coupling constant of -18 MHz, and anisotropic components of -9 MHz ($A_{x,y}^{\text{d}}$) and 17-18 MHz (A_z^{d}). The experimental result is compared to the corresponding calculated values in table 6.3. The calculations correctly predict the pattern and signs of the tensor. However, the agreement is qualitative rather than quantitative: the isotropic HFCs are underestimated, while the anisotropic HFCs are overestimated by a factor of approximately 2. The underlying cause for the too positive Fermi contact terms is a significant underestimation of the core level spin polarization by DFT, as has been analyzed in detail by Munzarova and Kaupp [261]. This deficiency of DFT seems to plague the predictions of isotropic metal HFCs in general, see also ref. [245].

We note that the spin-orbit contribution to the ^{57}Fe A-tensor (eq. 6.32) accounts for a small adjustment of only 1-2 MHz to the total A-tensor. These small values are consistent with the calculated small g-shifts, considering the similarities in the expressions of the spin-orbit contribution to both quantities (eqs. 6.31 and 6.32).

The calculated hyperfine interactions with the pyrrole nitrogen atoms and the oxygen atoms directly coordinated to iron in the complexes **17** and **18** are given in table 6.4. These parameters have not been determined experimentally for the complexes $[\text{Fe}^{\text{IV}}\text{O}(\text{TDCPP})]^+$ and $[\text{Fe}^{\text{IV}}\text{O}(\text{TMP})]^+$. However, a comparison of these values with those obtained for the native Cpd I species in the enzyme environment might be instructive, giving a measure for the impact of environmental factors (polarization, cysteinate ligation) on the computed ligand HFCs.

The data in table 6.4 show that the oxo ligand exhibits an axial splitting with a prominent coupling constant of (**17/18**) 121/132 MHz for the anisotropic A_z^{d} component. In contrast, the pyrrole

nitrogens show small splittings. Interestingly, model complex **18** shows two sets of similar, but distinguishable resonances, which reflect the deviation from 4-fold symmetry of the porphyrin caused by the axial water ligand.

Table 6.4: Calculated ^{17}O and ^{14}N Hyperfine coupling constants [MHz] for complexes **17** and **18**.

compound	atom	A^{iso}	A_x^{d}	A_y^{d}	A_z^{d}
17	oxo O	-11.2	-60.4	-60.4	120.8
	pyrrole N	2.3	-4.0	-1.7	5.7
18	oxo O	-12.3	-67.7	-63.8	131.5
	aqua O	-3.1	-3.1	0.9	2.2
	pyrrole N ⁽¹⁾	2.1	-5.1	-3.0	8.1
	pyrrole N ⁽²⁾	2.1	-5.4	-3.5	9.0

6.4.2 The Resting State

The results from ^1H -ENDOR measurements of the resting form of P450_{cam} [225, 226], together with X-ray data [85], have resolved a long-standing controversy over the identity of the second axial ligand to the heme iron in **1** [3]. These studies have confirmed the presence of either a water or a hydroxide as the sixth ligand. From recent ESEEM measurements it was concluded that the ligand is a water molecule [86, 87], with two strongly exchangeable protons at a distance of 2.62 Å from the heme iron; the presence of a hydroxide ion was excluded.

The hyperfine coupling to the pyrrole nitrogen atoms, as determined by ^{14}N -ENDOR studies, provides a good measure of the active site environment of heme enzymes, and also for the spin density and electronic structure of the central iron atom. For example, Fann *et al.* have demonstrated how protein-induced perturbations can alter these hyperfine couplings between the ferric resting form of CPO and P450 enzymes, which share the same metal coordination environment [226]. While aquomet-CPO shows distinct resonances for all four pyrrole nitrogens, the ^{14}N resonances in P450_{cam} are less resolved, indicating that the active site environment is more symmetric in the latter species. The experimental ^{14}N -ENDOR spectrum of **1** shows a broad absorption from *ca.* 5.0–6.4 MHz, with an average value of 5.7 MHz.

The calculated ^1H HFCs for the water protons are compared against experimental data in table 6.5. The spectral features could be simulated by assuming two magnetically equivalent protons, with an isotropic HFC of 1.5–2.0 MHz and anisotropic components of 4.2–4.5 MHz. The comparison to the calculated values (table 6.5) shows nearly perfect agreement of the anisotropic part of the

Table 6.5: ^1H Hyperfine coupling constants [MHz] calculated for **1**^a.

	theory				experiment ^b			
	A^{iso}	A_x^{d}	A_y^{d}	A_z^{d}	A^{iso}	A_x^{d}	A_y^{d}	A_z^{d}
H ⁽¹⁾	2.0	4.7	5.0	9.4	1.5–2.0	4.2–4.5	4.2–4.5	8.4–9.0
H ⁽²⁾	0.4	4.6	4.4	9.0	1.5–2.0	4.2–4.5	4.2–4.5	8.4–9.0

^a The sign of the HFCs was not determined. ^b ref. [87].

A-tensor. For the isotropic HFCs, we obtained two distinct resonances of 0.4 and 2.0 MHz. This is related to the different spin densities on the two protons, one of which, H(1), participates in a hydrogen bond to a neighbouring water molecule in the active site, while the other one is not. Obviously, the assumption of two equivalent protons is not transferable to our calculations, which considered only one single conformation, while the experimental values derive from an ensemble of conformations. Hence, one might expect that on average both protons are involved in hydrogen-bond interactions, considering the fact that the theoretical prediction for the isotropic HFC on H(1) virtually coincides with the experimental value of 1.5–2.0 MHz.

The calculated HFCs for the four pyrrole nitrogens are listed in table 6.6. One can see that the pattern falls almost exactly into the experimentally determined range of 5.0–6.4 MHz. The average value for the calculated total HFCs is 5.9 MHz, in nearly perfect agreement with the experimental average of 5.7 MHz. A roughly isotropic distribution around 5.8–5.9 MHz has also been obtained by Davydov *et al.* for other ferric heme intermediates in the P450_{cam} catalytic cycle, namely the iron(III) peroxo-complex **5**, hydroperoxy complex **6**, and product complex **9** (see section 1.2.1). In contrast, the HFCs calculated for the pyrrole nitrogens in the ferryl model complex **17** (**18**) are $A_x^{\text{tot}} = -2$, $A_y^{\text{tot}} = +1$, $A_z^{\text{tot}} = +8$ MHz ($A_x^{\text{tot}} = -3$, $A_y^{\text{tot}} = -1$, $A_z^{\text{tot}} = +10$ MHz). These differences in the signals could allow for an experimental distinction between the high-valent ferryl compound and its ferric precursors in the catalytic cycle.

Table 6.6: ^{14}N Hyperfine coupling constants [MHz] on pyrrole nitrogens^a

	theory				experiment ^b
	A^{iso}	A_x^{tot}	A_y^{tot}	A_z^{tot}	A
N ^A	-6.3	-6.8	-6.2	-6.0	±5.0–6.4
N ^B	-6.3	-6.8	-6.1	-5.9	±5.0–6.4
N ^C	-5.9	-6.7	-5.8	-5.2	±5.0–6.4
N ^D	-5.5	-6.2	-5.5	-4.7	±5.0–6.4

^a For explanation of the labels N^A–N^D refer to figure 6.1. ^b ref. [226].

The ^1H HFCs calculated for the *meso*-protons displayed an isotropic distribution in the range of 0–2 MHz, consistent with the experimentally observed small HFCs in **1** (≈ 1 MHz) [87,225].

The *g*-values of the resting form **1** were calculated as $g_1 = 2.016$, $g_2 = 2.151$, $g_3 = 2.173$. The experimental EPR spectrum of **1** is a typical powder pattern of a $S = 1/2$ spin system with a non-axial *g*-tensor where $g_1 = 1.91$, $g_2 = 2.26$, $g_3 = 2.45$ [213]. The poor agreement of the calculated and experimental values is due to the near-degenerate LS Fe^{III} ground state, which gives rise to very large spin-orbit contributions to the *g*-shifts. Furthermore, the observed large negative shift in g_1 indicates that perturbations of higher order need to be taken into account [281], which are not included in our approach. The method of choice to address this problem would be a relativistic two-component DFT approach [282] in spin-polarized form, which is, however, not yet available. Alternatively, one might use a parameterized model (based, *e.g.*, on a ligand-field argument [237]) and determine the necessary input quantities (excitation energies, orbital energy differences) from DFT calculations.

In conclusion, the results presented in the preceding two sections demonstrate that realistic predictions for the Mößbauer quadrupole splitting and isomer shift, the exchange coupling constant J , and the ligand hyperfine coupling interactions are possible on the basis of DFT calculations. The calculated metal HFCs of oxoferryl model compounds are in qualitative agreement with the experimental data, yet the isotropic contributions are underestimated, while the anisotropic part is overestimated by a factor of 2. Due to the lack of reliable reference data, we could not assess the accuracy of the *g*-values calculated for Fe^{IV} Cpd I model complexes (**17**, **18**). However, the limitations of the present approach to *g*-tensors become obvious when spin-orbit effects become very large, as shown by the calculated *g*-shifts of the ferric aqua complex **1**.

6.4.3 Compound I

The calculated Mößbauer data for Cpd I of P450_{cam} are collected in table 6.7. We compare the parameters from the QM/MM treatment in the enzyme environment ($S_{\text{p,p}}$) to those of isolated system in the gas phase ($S_{\text{g,g}}$). The gas phase results are very similar to those obtained for the Cpd I model complexes **17** and **18** (see table 6.1). Comparing only the equivalent HS states, the quadrupole splitting (isomer shift) is 1.34 (0.09) mm s^{-1} as compared to 1.48 (0.06) mm s^{-1} in **17** and 1.62 (0.08) mm s^{-1} in **18**. The spectral features of Cpd I are significantly changed in the enzyme environment: ΔE_Q is reduced to 0.64 mm s^{-1} and $\delta_{57\text{Fe}}$ is increased to 0.13 mm s^{-1} . The HS and LS states exhibit nearly identical Mößbauer parameters, since these depend only on the total electron density, which is essentially the same in both spin states.

Calculated Heisenberg exchange coupling constants for the exchange coupled system are given in table 6.8, together with *g*-values. As described in section 3.4, QM/MM calculations with the

Table 6.7: Calculated Mößbauer isomer shift $\delta_{57\text{Fe}}$ [mm s⁻¹], quadrupole splitting ΔE_Q [mm s⁻¹], and asymmetry parameter η for Cpd I of P450_{cam}^a.

Spin state	system	ΔE_Q [mm s ⁻¹]	η	$\delta_{57\text{Fe}}$ [mm s ⁻¹]
² A	S _{p,p}	0.67	0.09	0.13
	S _{g,g}	1.34	0.06	0.09
⁴ A	S _{p,p}	0.64	0.10	0.13
	S _{g,g}	1.33	0.03	0.09

^a S_{p,p} refers to the system in the enzyme environment, S_{g,g} to the system in the gas phase.

same QM region as employed here (R3) and the basis set B4 predict an antiferromagnetic doublet ground state, being more stable than the quartet state by 11 cm⁻¹. This assignment is confirmed by the present single point calculations which employ a larger basis set (TZVP) and lead to similar values, J being -16 cm⁻¹. In the gas phase calculations, we obtained a coupling constant of $J = -27$ cm⁻¹. This value may be compared to the experimental datum of -37 cm⁻¹ for Cpd I in CPO [38, 46]. Likewise, J is also negative, but smaller in compound I of HRP [48]. Our calculations thus correctly reproduce the intriguing experimental finding that the sign of J is reversed in the native Cpd I species as opposed to Cpd I model complexes, like the ones discussed above, which all show ferromagnetic coupling [164]. Weiss *et al.* suggested [164] that this effect is a result of the delocalization of one oxidizing equivalent not only over the porphyrin, but also over the proximal axial ligand. Specifically, interactions of orbitals on sulfur and the oxoferryl moiety are thought to provide the overlap of magnetic orbitals, that would give rise to the observed antiferromagnetic coupling. This qualitative notion agrees very well with the redox mesomorphism in Cpd I, discussed in section 3.4.

The g-shifts computed for Cpd I are rather small, similar to the situation in the model complexes discussed above. However, in going to the cysteinate ligated system, there is a slight tendency towards a non-axial pattern. Such a qualitative difference to the situation in the model systems, albeit small, should be considered in the analysis of corresponding spectra. No significant changes in the g-values are observed between the isolated gas phase system (S_{g,g}) and the system in the enzyme environment (S_{p,p}).

Table 6.9 summarizes the calculated hyperfine coupling constants on the ferryl atom, as well as on the oxo and pyrrole nitrogen ligands. Comparing again the HS data with those for the synthetic compounds **17** and **18** (tables 6.3 and 6.4), it is obvious that the overall features of the hyperfine tensor are very similar in the native Cpd I species and in these complexes. For example, the HFCs of the central iron atom are the same to within 6 MHz. Similarly, the oxo ligand exhibits the same axial splitting with a large anisotropic A_z^d component of *ca.* 120 MHz, compared to

Table 6.8: Calculated principal g-values and Heisenberg exchange coupling constant J [cm^{-1}] for Cpd I of P450_{cam}^a.

State	system	g	J(1) ^b [cm^{-1}]	J(2) ^c [cm^{-1}]
² A	S _{p,p}	2.016, 2.009, 1.982 (2.020, 2.012, 1.918) ^d	-16	-16
	S _{g,g}	2.015, 2.008, 1.980 (2.020, 2.007, 1.916) ^d	-27	-27
⁴ A	S _{p,p}	2.045, 2.012, 2.006	-	-
	S _{g,g}	2.045, 2.013, 2.006	-	-

^a S_{p,p} refers to the system in the enzyme environment, S_{g,g} to the system in the gas phase. ^b Calculated according to equation 6.39. ^c Calculated according to equation 6.40. ^d Spin projected (BS uncorrected) values.

(17/18) 121/132 MHz in the model compounds. The pyrrole nitrogen atoms exhibit a somewhat broader spread of the resonances, which is probably due to the more unsymmetric environment of the cysteinate ligated system, but the absolute values are generally quite small. In the protein environment, the spin density in p orbitals at the pyrrole nitrogen positions (which contribute to the a_{2u} orbital, see chapter 3.4), is increased. As a consequence, the dipolar contributions to the ¹⁴N HFCs are also increased by a factor up to 2 with respect to the gas phase, so that the ¹⁴N HFCs are generally larger in the enzyme environment.

Comparing the theoretical results for the doublet ground state with those for the quartet state (table 6.9) the metal HFCs exhibit somewhat larger isotropic and smaller anisotropic contributions, while the ligand HFCs are generally smaller. Qualitatively, the resonance patterns are however very similar.

The largest hydrogen HFCs of Cpd I are listed in table 6.10. A notable interaction is found with the β -hydrogen atoms on the cysteinate ligand, one of which shows isotropic coupling constants of (S_{pp}/S_{gg}) -8/-17 MHz and 6/16 MHz in the LS and HS state, respectively. The other significant couplings arise from the *meso*-hydrogen atoms, which have resonances in the range of $\pm(2-6)$ MHz and $\pm(5-13)$ MHz in the LS and HS states, respectively. These HFCs are apparently sensitive to environmental factors, since they are decreased in the gas phase calculations. Similar to the situation for the ¹⁴N HFCs, this is due to a concentration of spin density in the p-orbitals at the *meso*-carbon atoms, which contribute to the singly-occupied a_{2u} orbital.

Hence, analogous to the calculated Mößbauer data, the inclusion of the steric and electronic influence of the protein environment leads to non-negligible effects on the spectral features.

Table 6.9: Calculated ^{57}Fe , ^{17}O , and ^{14}N hyperfine coupling constants [MHz] in compound I for the LS and the HS states^{a,b}.

		$^2\text{A state}^c$		$^4\text{A state}$	
		$S_{\text{p,p}}$	$S_{\text{g,g}}$	$S_{\text{p,p}}$	$S_{\text{g,g}}$
Fe	A^{iso}	-16.0 (-24.0)	-16.9 (-25.3)	-10.0	-10.1
	A_x^{d}	-13.0 (-19.5)	-12.3 (-18.5)	-17.1	-16.7
	A_y^{d}	-12.2 (-18.3)	-12.0 (-18.0)	-16.1	-16.2
	A_z^{d}	25.2 (37.8)	24.4 (36.6)	33.2	32.9
O^d	A^{iso}	-21.5 (-32.2)	-21.7 (-32.5)	-12.5	-11.7
	A_x^{d}	-43.9 (-65.9)	-43.5 (-65.3)	-61.5	-61.3
	A_y^{d}	-40.6 (-60.9)	-41.3 (-61.9)	-55.3	-58.7
	A_z^{d}	84.5 (126.8)	84.8 (127.2)	116.8	119.9
N^{A}	A^{iso}	-1.8 (-5.3)	-0.6 (-1.9)	2.4	1.9
	A_x^{d}	-3.2 (-9.7)	-1.9 (-5.8)	-5.9	-4.0
	A_y^{d}	1.3 (4.0)	0.6 (1.9)	-3.9	-1.6
	A_z^{d}	1.9 (5.7)	1.3 (3.9)	9.7	5.6
N^{B}	A^{iso}	-2.1 (-6.3)	-0.3 (-1.0)	1.8	1.2
	A_x^{d}	-3.6 (-10.8)	-1.6 (-4.9)	-6.3	-3.4
	A_y^{d}	1.5 (4.5)	0.5 (1.4)	-4.5	-1.3
	A_z^{d}	2.1 (6.3)	1.2 (3.5)	10.9	4.7
N^{C}	A^{iso}	-1.9 (-5.6)	-0.5 (-1.6)	3.0	2.4
	A_x^{d}	-3.6 (-10.7)	-2.1 (-6.4)	-6.6	-4.3
	A_y^{d}	1.5 (4.4)	0.7 (2.1)	-4.3	-1.8
	A_z^{d}	2.1 (6.3)	1.4 (4.3)	10.8	6.1
N^{D}	A^{iso}	-2.7 (-8.2)	-3.6 (-10.8)	2.4	5.0
	A_x^{d}	-4.0 (-11.9)	-4.2 (-12.6)	-7.0	-7.5
	A_y^{d}	1.7 (5.2)	1.8 (5.4)	-5.0	-5.0
	A_z^{d}	2.3 (6.8)	2.4 (7.2)	12.1	12.5

^a $S_{\text{p,p}}$ refers to the system in the enzyme environment, $S_{\text{g,g}}$ to the system in the gas phase. ^b For explanation of the labels $\text{N}^{\text{A}}\text{--}\text{N}^{\text{D}}$ see figure 6.1. ^c Spin projected (BS uncorrected) values. ^d The oxo ligand.

Table 6.10: Calculated ^1H hyperfine coupling constants [MHz] Compound I for the LS and the HS state.^a

		$^2\text{A state}^b$		$^4\text{A state}$	
		$S_{\text{p,p}}$	$S_{\text{g,g}}$	$S_{\text{p,p}}$	$S_{\text{g,g}}$
H_1^β	A^{iso}	-7.6 (-22.9)	-17.0 (-50.9)	6.4	16.4
	A_x^{d}	-0.6 (-1.7)	-1.1 (-3.3)	-3.2	-4.3
	A_y^{d}	0.0 (0.0)	-0.0 (-0.1)	-1.8	-2.3
	A_z^{d}	0.6 (1.7)	1.1 (3.4)	5.0	6.5
H_2^β	A^{iso}	-2.7 (-8.0)	-6.8 (-20.3)	1.9	6.1
	A_x^{d}	-1.1 (-3.3)	-1.7 (-5.2)	-4.5	-5.0
	A_y^{d}	-0.5 (-1.4)	-0.0 (-0.0)	-2.9	-2.1
	A_z^{d}	1.5 (4.6)	1.7 (5.2)	7.4	7.1
H_1^{meso}	A^{iso}	3.4 (10.1)	1.9 (5.6)	-3.4	-1.9
	A_x^{d}	-1.6 (-4.8)	-0.9 (-2.6)	-8.4	-5.7
	A_y^{d}	-0.1 (-0.4)	0.1 (0.2)	-2.0	-1.9
	A_z^{d}	1.7 (5.1)	0.8 (2.5)	10.3	7.6
H_2^{meso}	A^{iso}	3.6 (10.8)	1.9 (5.6)	-3.6	-1.8
	A_x^{d}	-1.7 (-5.1)	-0.7 (-2.2)	-8.9	-5.5
	A_y^{d}	-0.2 (-0.5)	-0.0 (-0.1)	-2.0	-1.8
	A_z^{d}	1.9 (5.6)	0.8 (2.3)	10.8	7.2
H_3^{meso}	A^{iso}	3.6 (10.9)	2.0 (6.1)	-3.7	-2.0
	A_x^{d}	-1.8 (-5.4)	-0.9 (-2.6)	-9.3	-5.9
	A_y^{d}	-0.1 (-0.4)	-0.0 (-0.1)	-2.1	-1.8
	A_z^{d}	2.0 (5.9)	0.9 (2.7)	11.4	7.7
H_4^{meso}	A^{iso}	3.2 (9.6)	1.8 (5.3)	-3.2	-1.8
	A_x^{d}	-1.4 (-4.3)	-0.8 (-2.4)	-8.3	-5.4
	A_y^{d}	-0.2 (-0.6)	0.1 (0.3)	-2.0	-1.7
	A_z^{d}	1.6 (4.9)	0.7 (2.2)	10.2	7.2

^a $S_{\text{p,p}}$ refers to the system in the enzyme environment, $S_{\text{g,g}}$ to the system in the gas phase. ^b Spin projected (BS uncorrected) values.

6.5 Conclusion

In this chapter, DFT-based theoretical methods were employed to calculate the Mößbauer-parameters, hyperfine interactions, exchange coupling constants, and g-tensors of iron-porphyrins with oxygen donor ligands, which are related to key intermediates in the catalytic cycle of P450 enzymes. To describe the antiferromagnetic doublet state of the exchange-coupled ferryl porphyrin cation radicals, we employed the broken-symmetry (BS) DFT approach [227–229].

In the first part of the study, QM calculations were carried out for synthetic oxoferryl tetraphenylporphyrin complexes, which have been characterized spectroscopically. Our calculations demonstrate that the present theoretical approach gives reliable predictions for Mößbauer quadrupole splittings and isomer shifts, as well as for Heisenberg exchange coupling constants and ligand HFCs. The metal HFCs were found to be in qualitative agreement with experiment. A conclusive experimental determination of g-values of these complexes has not yet been possible, hence we could not relate the calculated g-shifts to reference data. To make the connection to the experimental work, theoretical predictions regarding the D-tensor of these complexes are required, which may be obtained in future studies from a sum-over-states CI approach.

The second part of the present investigation addressed the calculation of ligand HFCs in the resting form of P450_{cam} (**1**) by means of a QM/MM treatment. The calculations are in excellent agreement with ENDOR and ESEEM results that determined ¹H HFCs couplings, as well as with ¹⁴N HFCs on the pyrrole nitrogens. The g-shifts are significantly underestimated by our calculations. This is due to the near-degenerate electronic ground state of **1**, which gives rise to large spin-orbit effects.

On the basis of these results, theoretical predictions were made for the spin Hamiltonian parameters of Cpd I, the primary oxidant of P450_{cam}. These calculations predict a weak antiferromagnetic coupling with $J = -16 \text{ cm}^{-1}$, equivalent to a doublet ground state. The calculated Mößbauer quadrupole splitting is 0.67 mm s^{-1} and the isomer shift is 0.13 mm s^{-1} . The calculated hyperfine interactions on the pyrrole nitrogens yield isotropic HFCs of $\pm(2\text{--}3) \text{ MHz}$ and anisotropic HFCs of 1–4 MHz. The most prominent resonances are due to ¹⁷O coupling at the oxo ligand ($A^{\text{iso}} = -22 \text{ MHz}$, $A_x^{\text{d}} = -44 \text{ MHz}$, $A_y^{\text{d}} = -41 \text{ MHz}$, $A_z^{\text{d}} = 85 \text{ MHz}$). The computed g-values remain close to 2.0 and show a non-axial pattern, however the predicted values may not be too reliable. A comparison to analogous calculations on isolated species in the gas phase demonstrates that the inclusion of steric and electrostatic effects due to the protein/solvent environment leads to non-negligible effects on Mößbauer parameters and hyperfine coupling constants, particularly if the spin densities at these atomic sites are prone to polarization effects.

While these results already give useful information on spectroscopic properties of the compounds under study, the present theoretical methodology can certainly be improved. This is especially true with regard to quantitative predictions for the g-tensor and metal hyperfine coupling constants. For

instance, the semi-empirical spin-orbit coupling operator should be replaced by a theoretically more founded treatment. In the present case, this is crucial for the near-degenerate Fe^{III} ground state of **1**. Moreover, the errors introduced by the density functional approach itself are probably quite large, due to the fact that in the design of such functionals spin-dependent properties are not considered explicitly. The development of new functionals, which take properties like the g-tensor or HFCs into account, therefore seems desirable. Ultimately, *ab initio* electron correlation approaches will be required in order to make reliable predictions for systems that are electronically as complex as the ones discussed here.

Chapter 7

Summary

Cytochrome P450 enzymes function as monooxygenases, incorporating selectively one atom from dioxygen at physiological conditions into a wide variety of mostly hydrophobic substrates – this reaction, if uncatalyzed, would require high temperatures and would proceed nonspecifically. The remarkable catalytic potency of P450, in combination with the high degree of regio- and stereo-selectivity, has attracted considerable interest. However, due to the elusive character of some of the reactive intermediates, central aspects of the catalytic cycle are still not resolved experimentally. This thesis presents results from combined quantum mechanical/molecular mechanical (QM/MM) calculations, which are aimed at an understanding of the factors that govern P450 function by providing insights into processes at the atomistic level. In our investigations, we have dealt with key intermediates in the catalytic hydroxylation of C–H bonds, mediated by the bacterial enzyme P450_{cam}, which is well characterized by biochemical and biophysical techniques. The QM/MM calculations provide information on these species and their transformations in the specific protein environment, by treating this environment explicitly at the force field level. Density functional theory (DFT) was employed to describe the electronic structure of the active site.

The present study characterizes the geometric and electronic structure of the elusive active oxidant of cytochrome P450, an oxoferryl complex [FeO(porph)(SR)] (SR = cysteinate linkage to the protein, porph = protoporphyrin IX) commonly named compound I. By comparisons with various gas phase models and through inspection of different conformations within the protein pocket, these calculations offer important insight into the factors that govern the properties of compound I. Compound I is an exchange coupled triradicaloid, with an $S = 1$ system on the FeO moiety and a single unpaired electron ($S = 1/2$) on the ligands. This situation gives rise to two nearly degenerate electronic states; ferromagnetic coupling yields a quartet state, whereas antiferromagnetic coupling leads to a doublet state. We find that the unpaired electron on the ligands is mainly localized in a porphyrin π orbital (of the a_{2u} variety in idealized D_{4h} symmetry; *ca.*

70%), in contrast to the situation in gas phase models, which are mainly sulfur-centered radicals. This transformation is mainly achieved by a hydrogen-bonding network in the proximal protein pocket, that stabilizes electron density on the sulfur and thereby strengthens Fe–S coordination. The best estimates of the Fe–S bond lengths in the $^2A_{2u}/^4A_{2u}$ states are 2.569 Å/ 2.547 Å. A series of calculations on different conformations of the protein pocket shows that the balance of the sulfur/porphyrin radical character and the precise Fe–S bond length are finely tuned by small changes in the architecture of the pocket. Cpd I behaves therefore as a “chameleon” species that adapts its electronic and structural character to the specific enzyme environment.

Mechanistic studies concern the hydrogen-abstraction, oxygen-rebound pathway for C–H hydroxylation. This mechanism *via* radical intermediates is currently under dispute, because radical-clock experiments have arrived at lifetimes of these radicals which are much too short to correspond to true intermediates. Our detailed calculations support a two-state-reactivity (TSR) scenario, that has been suggested earlier on the basis of model computations. Thus, the reaction proceeds on two spin state surfaces, the doublet and the quartet state. While the reaction on the quartet surface is truly stepwise with a distinct radical intermediate, the doublet reaction is non-synchronous, yet effectively concerted. As such, this mechanism provides a satisfactory explanation for the seemingly contradictory experimental findings.

Comparative QM calculations of model complexes in the gas phase allow us to investigate the role of the protein environment in the C–H hydroxylation reaction. This establishes that the protein environment influences the stability of redox electromers (*i.e.* Fe^{III} vs. Fe^{IV} states, porphyrin cation radical vs. sulfur radical states) and the relative energies of different spin states. Moreover, our calculations indicate that substrate motion is hindered in the active site, mainly by a hydrogen bond of the camphor carbonyl oxygen with a tyrosine hydroxy-group of the protein. These constraints have consequences for the C–O bond formation step, which is known to control the stereochemistry of the product.

We have also considered a mechanism of electrostatic TS stabilization that could account for a lowering of the barrier in the initial C–H activation step. This mechanism involves enhanced interactions of the negatively charged heme propionate side chains with neighbouring positively charged protein residues in the transition state. Our calculations clearly demonstrate that there is no charge transfer to the propionate groups in the transition state, which could lead to such a lowering of the barrier. Instead, it is suggested that the protein promotes the reaction by absorbing the entropic cost required to establish the TS conformation into the process of substrate binding, which is driven by the expulsion of water molecules from the active site and the desolvation of the hydrophobic substrate.

Another part of the present work addresses the resting form of P450_{cam}, the ferric aqua heme complex $[Fe^{III}(\text{porph})(\text{SR})(\text{H}_2\text{O})]$. It is known that displacement of the axial water ligand upon substrate binding changes the ground state multiplicity from low spin to high spin, which modifies the

reduction potential of the heme and triggers the catalytic cycle. Our QM/MM calculations reveal how the protein/solvent environment modulates the conformation and binding properties of the axial water ligand by electronic polarization as well as the presence of other water molecules in the active site. In agreement with experimental results, our QM/MM calculations indicate a doublet ground state. The presence of other water molecules in the active site provides a hydrogen-bonding network that favours the upright conformation (perpendicular to the heme plane) of the axial water ligand. Electronic polarization by the enzyme environment causes a stabilization of electron density on the proximal sulfur atom. This effect apparently lowers the orbital energy of the doubly occupied sulfur p orbital, which mixes with orbitals on iron to give a bonding/antibonding combination of Fe–S σ orbitals. As a consequence, the mixing is less pronounced, effectively lowering the orbital energy of the antibonding Fe–S σ^* orbital. Thus, the high spin states, with an unpaired electron in this orbital, are significantly stabilized in the enzyme as compared to the gas phase. This effect is most prominent in the quartet state, which is lowered by *ca.* 8 kcal/mol with respect to the doublet state. Likewise, the Fe–S bond is longer by 0.1 Å in the enzyme environment as compared to the gas phase. Due to the *trans*-effect, this causes a concomitant strengthening of the bond to the other axial ligand, the water molecule, which is manifest in a shortening of the Fe–O bond length.

DFT-based theoretical methods have been employed to calculate the spin-Hamiltonian parameters of iron-porphyrins with oxygen donor ligands, which represent key intermediates in the catalytic cycle of P450 enzymes.

For the purpose of validation, calculations were carried out on synthetic oxoferryl tetraphenylporphyrin complexes $[\text{FeO}(\text{TPP})]^+$ (TPP = *meso*-tetraphenylporphyrin), derivatives of which have been characterized spectroscopically. These calculations demonstrate that the present theoretical approach gives realistic predictions for Mößbauer quadrupole splittings and isomer shifts, as well as for Heisenberg exchange coupling constants and ligand hyperfine coupling constants. QM/MM calculations of ligand hyperfine coupling constants in the resting form of P450_{cam} are in excellent agreement with ENDOR and ESEEM results for ^1H resonances from the distal water ligand, as well as with ^{14}N coupling constants on the pyrrole nitrogens, while the g-shifts are significantly underestimated. Theoretical predictions were made for Mößbauer parameters, hyperfine interactions, exchange coupling constants, and g-tensors of compound I, the primary oxidant of P450_{cam}, to guide experimentalists in the detection of this species. The antiferromagnetic doublet state of the exchange-coupled system was described by the broken-symmetry (BS) DFT approach and spin projection. These calculations predict a weak antiferromagnetic coupling, equivalent to a doublet ground state. A comparison to analogous calculations on isolated species in the gas phase shows that the inclusion of steric and electrostatic effects due to the protein/solvent environment significantly influences the computed Mößbauer parameters and ligand hyperfine couplings in compound I of P450_{cam}.

Bibliography

- [1] P. R. Ortiz de Montellano (ed.). *Cytochrome P450. Structure, Mechanism, and Biochemistry*. Plenum Press, New York, 2nd edition, 1995.
- [2] E. J. Mueller, P. J. Loida, and S. G. Sligar. Twenty-five years of P450_{cam} research: Mechanistic insights into oxygenase catalysis. Ref. [1], pp. 83-124.
- [3] J. H. Dawson and M. Sono. Cytochrome P-450 and chloroperoxidase: Thiolate-ligated heme enzymes. Spectroscopic determination of their active site structures and mechanistic implications of thiolate ligation. *Chem. Rev.*, 87:1255–1276, 1987.
- [4] T. L. Poulos, B. C. Finzel, I. C. Gunsalus, G. C. Wagner, and J. Kraut. The 2.6 Å crystal structure of *Pseudomonas putida* cytochrome P-450. *J. Biol. Chem.*, 260:16122–16130, 1985.
- [5] H. Patzelt and W.-D. Woggon. The first observation of O₂ cleavage by a P-450 enzyme model in the presence of a thiolate ligand. *Helv. Chim. Acta*, 75:523–530, 1992.
- [6] G. H. Loew and D. L. Harris. Role of the heme active site and protein environment in structure, spectra, and function of the cytochrome P450s. *Chem. Rev.*, 100:407–419, 2000.
- [7] P. R. Ortiz de Montellano and J. J. De Voss. Oxidizing species in the mechanism of cytochrome P450. *Nat. Prod. Rep.*, 19:1–18, 2002.
- [8] J. T. Groves, G. A. McClusky, R. E. White, and M. J. Coon. Aliphatic hydroxylation by highly purified liver microsomal cytochrome P-450 - Evidence for a carbon radical intermediate. *Biochem. Biophys. Res. Commun.*, 81:154–160, 1978.
- [9] J. T. Groves and D. V. Subramanian. Evidence for radical intermediates in allylic hydroxylation by cytochrome P-450. *J. Am. Chem. Soc.*, 106:2177–2181, 1984.
- [10] M. Newcomb and P. H. Toy. Hypersensitive radical probes and the mechanism of cytochrome P450 catalyzed hydroxylation reactions. *Acc. Chem. Res.*, 33:449–455, 2000.

- [11] M. Newcomb, P. F. Hollenberg, and M. J. Coon. Multiple mechanisms and multiple oxidants in P450-catalyzed hydroxylations. *Arch. Biochem. Biophys.*, 409:72–79, 2003.
- [12] I. Schlichting, J. Berendzen, K. Chu, A. M. Stock, S. A. Maves, D. E. Benson, R. M. Sweet, D. Ringe, G. A. Petsko, and S. G. Sligar. The catalytic pathway of cytochrome P450cam at atomic resolution. *Science*, 287:1615–1622, 2000.
- [13] R. Davydov, T. M. Makris, V. Kofman, D. E. Werst, S. G. Sligar, and B. M. Hoffman. Hydroxylation of camphor by reduced oxy-cytochrome P450cam: Mechanistic implications of EPR and ENDOR studies of catalytic intermediates in native and mutant enzymes. *J. Am. Chem. Soc.*, 123:1403–1415, 2001.
- [14] I. G. Denisov, T. M. Makris, and S. G. Sligar. Cryotrapped reaction intermediates of cytochrome P450 studied by radiolytic reduction with phosphorus-32. *J. Biol. Chem.*, 276:11648–11652, 2001.
- [15] M. Akhtar, V. C. O. Njar, and J. N. Wright. Mechanistic studies on aromatase and related C–C bond cleavage P-450 enzymes. *J. Steroid. Biochem.*, 44:375–387, 1993.
- [16] W. O. Nam, M. H. Lim, H. J. Lee, and C. Kim. Evidence for the participation of two distinct reactive intermediates in iron(III) porphyrin complex-catalyzed epoxidation reactions. *J. Am. Chem. Soc.*, 122:6641–6647, 2000.
- [17] R. Davydov, V. Kofman, H. Fujii, T. Yoshida, M. Ikeda-Saito, and B. M. Hoffman. Catalytic mechanism of heme oxygenase through EPR and ENDOR of cryoreduced oxy-heme oxygenase and its Asp140 mutants. *J. Am. Chem. Soc.*, 124:1798–1808, 2002.
- [18] K. Yamaguchi, Y. Watanabe, and I. Morishima. Push effect on the heterolytic O–O bond cleavage of peroxo(III) porphyrin adducts. *Inorg. Chem.*, 31:156–157, 1992.
- [19] J. T. Groves and Y. Watanabe. Reactive iron porphyrin derivatives related to the catalytic cycle of cytochrome P450 and peroxidases. Studies of the mechanism of oxygen activation. *J. Am. Chem. Soc.*, 110:8443–8452, 1988.
- [20] T. Higuchi, S. Uzu, and M. Hirobe. Synthesis of a highly stable iron porphyrin coordinated by alkylthiolate anion as a model for cytochrome P-450 and its catalytic activity in oxygen-oxygen bond cleavage. *J. Am. Chem. Soc.*, 112:7051–7053, 1990.
- [21] T. Higuchi, K. Shimada, N. Maruyama, and M. Hirobe. Heterolytic oxygen-oxygen bond cleavage of peroxy acid and effective alkane hydroxylation in hydrophobic solvent mediated by an iron porphyrin coordinated by thiolate anion as a model for cytochrome P-450. *J. Am. Chem. Soc.*, 115:7551–7552, 1993.

- [22] M. Vidakovic, S. Sligar, H. Li, and T. L. Poulos. Understanding the role of the essential Asp251 in cytochrome P450cam using site-directed mutagenesis, crystallography, and kinetic solvent isotope effect. *Biochemistry*, 37:9211–9219, 1998.
- [23] D. L. Harris and G. H. Loew. Investigation of the proton-assisted pathway to formation of the catalytically active ferryl species of P450s by molecular dynamics studies of P450eryF. *J. Am. Chem. Soc.*, 118:6377–6387, 1996.
- [24] D. L. Harris. Oxidation and electronic state dependence of proton transfer in the enzymatic cycle of cytochrome P450eryF. *J. Inorg. Biochem.*, 91:568–585, 2002.
- [25] T. Kamachi and K. Yoshizawa. A theoretical study on the mechanism of camphor hydroxylation by compound I of cytochrome P450. *J. Am. Chem. Soc.*, 125:4652–4661, 2003.
- [26] J. H. Dawson, R. H. Holm, J. R. Trudell, G. Barth, R. E. Linder, E. Bunnenberg, C. Djerassi, and S. C. Tang. Magnetic circular dichroism studies. 43. Oxidized cytochrome P-450. Magnetic circular dichroism evidence for thiolate ligation in the substrate bound form. Implications for the catalytic mechanism. *J. Am. Chem. Soc.*, 98:3707–3709, 1976.
- [27] K. Auclair, P. Moenne-Loccoz, and P. R. Ortiz de Montellano. Roles of the proximal heme thiolate ligand in cytochrome P450(cam). *J. Am. Chem. Soc.*, 123:4877–4885, 2001.
- [28] S. Yoshioka, S. Takahashi, H. Hori, K. Ishimori, and I. Morishima. Proximal cysteine residue is essential for the enzymatic activities of cytochrome P450(cam). *Eur. J. Biochem.*, 268:252–259, 2001.
- [29] F. Ogliaro, S. P. de Visser, and S. Shaik. The 'push' effect of the thiolate ligand in cytochrome P450: a theoretical gauging. *J. Inorg. Biochem.*, 91:554–567, 2002.
- [30] D. G. Kellner, S.-C. Nung, K. E. Weiss, and S. G. Sligar. Kinetic characterization of compound I formation in the thermostable cytochrome P450 CYP119. *J. Biol. Chem.*, 277:9641–9644, 2002.
- [31] V. Schünemann, C. Jung, A. X. Trautwein, D. Mandon, and R. Weiss. Intermediates in the reaction of substrate-free cytochrome P450(cam) with peroxy acetic acid. *FEBS Lett.*, 179:149–154, 2000.
- [32] V. Schünemann, C. Jung, J. Turner, A. X. Trautwein, and R. Weiss. Spectroscopic studies of peroxyacetic acid reaction intermediates of cytochrome P450cam and chloroperoxidase. *J. Inorg. Biochem.*, 91:586–596, 2002.
- [33] X. Yi, M. Mroczko, K. M. Manoj, X. Wang, and L. P. Hager. Replacement of the proximal heme thiolate ligand in chloroperoxidase with a histidine residue. *Proc. Natl. Acad. Sci. USA*, 96:12412–12417, 1999.

- [34] F. Ogliaro, S. P. de Visser, S. Cohen, J. Kaneti, and S. Shaik. The experimentally elusive oxidant of cytochrome P450: A theoretical “trapping” defining more closely the “real” species. *ChemBioChem*, 2:848–851, 2001.
- [35] D. L. Harris. High-valent intermediates of heme proteins and model compounds. *Curr. Opin. Chem. Biol.*, 5:724–734, 2001.
- [36] S. P. de Visser, F. Ogliaro, P. K. Sharma, and S. Shaik. Hydrogen bonding modulates the selectivity of enzymatic oxidations by P450: Chameleon oxidant behavior by compound I. *Angew. Chem. Int. Ed.*, 41:1947–1951, 2002.
- [37] A. Zaks and D. R. Dodds. Choroperoxidase-catalyzed asymmetric oxidations: Substrate specificity and mechanistic study. *J. Am. Chem. Soc.*, 117:10419–10424, 1995.
- [38] J. Antony, M. Grodzicki, and A. X. Trautwein. Local density functional study of oxo-iron(IV) porphyrin complexes and their one-electron oxidized derivatives: axial ligand effects. *J. Phys. Chem.*, 101:2692–2701, 1997.
- [39] M. T. Green. Evidence for sulfur-based radicals in thiolate compound I intermediates. *J. Am. Chem. Soc.*, 121:7939–7940, 1999.
- [40] D. Harris, G. H. Loew, and L. Waskell. Calculation of the electronic structure and spectra of model cytochrome P450 compound I. *J. Inorg. Biochem.*, 83:309–318, 2001.
- [41] T. Ohta, K. Matsuura, K. Yoshizawa, and I. Morishima. The electronic and vibrational structures of iron-oxo porphyrin with a methoxide or cysteinate axial ligand. *J. Inorg. Biochem.*, 82:141–152, 2000.
- [42] M. Filatov, N. Harris, and S. Shaik. A theoretical study of electronic factors affecting hydroxylation by model ferryl complexes of cytochrome P-450 and horseradish peroxidase. *J. Chem. Soc. Perkin Trans.*, 2:399–410, 1999.
- [43] D. L. Harris and G. H. Loew. Theoretical investigation of the proton assisted pathway to formation of cytochrome P450 compound I. *J. Am. Chem. Soc.*, 120:8941–8948, 1998.
- [44] F. Ogliaro, S. Cohen, S. P. de Visser, and S. Shaik. Medium polarization and hydrogen bonding effects on compound I of cytochrome P450: What kind of radical is it really? *J. Am. Chem. Soc.*, 122:12892–12893, 2000.
- [45] F. Ogliaro, S. Cohen, M. Filatov, N. Harris, and S. Shaik. The high-valent compound of cytochrome P450: the nature of the Fe–S bond and the role of the thiolate ligand as an internal electron donor. *Angew. Chem. Int. Ed.*, 39:3851–3855, 2000.

- [46] R. Rutter, L. P. Hager, H. Dhonau, M. Hendrich, M. Valentine, and P. Debrunner. Chloroperoxidase compound I: electron paramagnetic resonance and Mössbauer studies. *Biochemistry*, 23:6809–6816, 1984.
- [47] C. M. Hosten, A. M. Sullivan, V. Palaniappan, M. M. Fitzgerald, and J. Turner. Resonance Raman spectroscopy of the catalytic intermediates and derivatives of chloroperoxidase from *Caldariomyces fumago*. *J. Biol. Chem.*, 269:13966–13978, 1994.
- [48] C. E. Schultz, R. Rutter, J. T. Sage, P. G. Debrunner, and L. P. Hager. Mössbauer and electron-paramagnetic resonance studies of horseradish-peroxidase and its catalytic intermediates. *Biochemistry*, 23:4743–4754, 1984.
- [49] K.-J. Paeng and J. R. Kincaid. The resonance Raman spectrum of horseradish-peroxidase compound I. *J. Am. Chem. Soc.*, 110:7913–7915, 1988.
- [50] R. Rutter, M. Valentine, M. P. Hendrich, L. P. Hager, and P. G. Debrunner. Chemical nature of the porphyrin- π cation radical in horseradish-peroxidase compound I. *Biochemistry*, 22:4769–4774, 1983.
- [51] V. Palaniappan and J. Turner. Resonance Raman spectroscopy of horseradish-peroxidase derivatives and intermediates with excitation in the near ultraviolet. *J. Biol. Chem.*, 264:16046–16053, 1989.
- [52] Y. Watanabe and H. Fujii. Characterization of high-valent oxo-metalloporphyrins. *Struc. Bond.*, 97:61–89, 2000.
- [53] H. Fujii. Electronic structure and reactivity of high-valent oxo iron porphyrins. *Coord. Chem. Rev.*, 226:51–60, 2002.
- [54] M. H. Gelb, D. C. Heimbrook, P. Malkonen, and S. G. Sligar. Stereochemistry and deuterium-isotope effects in camphor hydroxylation by the cytochrome P450cam monooxygenase system. *Biochemistry*, 21:370–377, 1982.
- [55] L. M. Hjelmeland, L. Aronow, and J. Trudell. Intramolecular determination of primary kinetic isotope effects in hydroxylations catalyzed by cytochrome P-450. *Biochem. Biophys. Res. Commun.*, 76:541–549, 1977.
- [56] W. M. Atkins and S. G. Sligar. Metabolic switching in cytochrome P-450cam. Deuterium-isotope effects on regiospecificity and the monooxygenase oxidase ratio. *J. Am. Chem. Soc.*, 109:3754–3760, 1987.
- [57] P. R. Ortiz de Montellano and R. A. Stearns. Timing of the radical recombination step in cytochrome P-450 with ring-strained probes. *J. Am. Chem. Soc.*, 109:3415–3420, 1987.

- [58] V. W. Bowry, J. Lusztyk, and K. U. Ingold. Calibration of the bicyclo[2.1.0]pent-2-yl radical ring opening and an oxygen rebound rate constant for cytochrome P-450. *J. Am. Chem. Soc.*, 111:1927–1928, 1989.
- [59] J. K. Atkinson and K. U. Ingold. Cytochrome P450 hydroxylation of hydrocarbons – variation in the rate of oxygen rebound using cyclopropyl radical clocks including 2 new ultrafast probes. *Biochemistry*, 32:9209–9214, 1993.
- [60] M. Newcomb, M.-H. Le Tadic, D. A. Putt, and P. F. Hollenberg. An incredibly fast apparent oxygen rebound rate constant for hydrocarbon hydroxylation by cytochrome P450 enzymes. *J. Am. Chem. Soc.*, 117:3312–3313, 1995.
- [61] K. Auclair, Z. Hu, D. M. Little, P. R. Ortiz de Montellano, and J. T. Groves. Revisiting the mechanism of P450 enzymes with the radical clocks norcarane and spiro[2,5]octane. *J. Am. Chem. Soc.*, 124:6020–6027, 2002.
- [62] M. Newcomb, R. Shen, Y. Lu, M. J. Coon, P. F. Hollenberg, D. A. Kopp, and S. J. Lippard. Evaluation of norcarane as a probe for radicals in cytochrome P450 and soluble methane monooxygenase-catalyzed hydroxylation reactions. *J. Am. Chem. Soc.*, 124:6879–6868, 2002.
- [63] M. Newcomb, M.-H. Le Tadic-Biadatti, D. L. Chestney, E. S. Roberts, and P. F. Hollenberg. A nonsynchronous concerted mechanism for cytochrome P450 catalyzed hydroxylation. *J. Am. Chem. Soc.*, 117:12085–12091, 1995.
- [64] S. Shaik, M. Filatov, D. Schröder, and H. Schwarz. Electronic structure makes a difference: Cytochrome P-450 mediated hydroxylations of hydrocarbons as a two-state reactivity paradigm. *Chem. Eur. J.*, 4:193–199, 1998.
- [65] F. Ogliaro, N. Harris, S. Cohen, M. Filatov, S. P de Visser, and S. Shaik. A model “rebound” mechanism of hydroxylation by cytochrome P450: Stepwise and effectively concerted pathways, and their reactivity patterns. *J. Am. Chem. Soc.*, 122:8977–8989, 2000.
- [66] N. Harris, S. Cohen, M. Filatov, F. Ogliaro, and S. Shaik. Two-state reactivity in the rebound step of alkane hydroxylation by cytochrome P-450: Origins of free radicals with finite lifetimes. *Angew. Chem. Int. Ed.*, 39:2003–2007, 2000.
- [67] D. Schröder, S. Shaik, and H. Schwarz. Two-state reactivity as new concept in organometallic chemistry. *Acc. Chem. Res.*, 33:133–145, 2000.
- [68] F. Ogliaro, S. P. de Visser, J. T. Groves, and S. Shaik. Chameleon states: high-valent metal-oxo species of cytochrome P450 and its ruthenium analog. *Angew. Chem. Int. Ed.*, 40:2874–2878, 2001.

- [69] J. I. Manchester, J. P. Dinnocenzo, L. A. Higgins, and J. P. Jones. A new mechanistic probe for cytochrome P450: An application of isotope effect profiles. *J. Am. Chem. Soc.*, 119:5069–5070, 1997.
- [70] S.-Y. Choi, P. E. Eaton, P. F. Hollenberg, K. E. Liu, S. J. Lippard, M. Newcomb, D. A. Putt, S. P. Upadhyaya, and Y. Xiong. Regiochemical variations in reactions of methylcubane with *tert*-butoxyl radical, cytochrome P-450 enzymes, and a methane monooxygenase system. *J. Am. Chem. Soc.*, 118:6547–6555, 1996.
- [71] S. Shaik, S. P. de Visser, F. Ogliaro, H. Schwarz, and D. Schröder. Two-state reactivity mechanisms of hydroxylation and epoxidation by cytochrome P-450 revealed by theory. *Curr. Opin. Chem. Biol.*, 6:556–567, 2002.
- [72] R. Poli and J. N. Harvey. Spin forbidden reactions of transition metal compounds. New ideas and new computational challenges. *Chem. Soc. Rev.*, 32:1–8, 2003.
- [73] K. Yoshizawa, Y. Kagawa, and Y. Shiota. Kinetic isotope effects in a C–H bond dissociation by the iron-oxo species of cytochrome P450. *J. Phys. Chem. B*, 104:12365–12370, 2000.
- [74] K. Yoshizawa. Theoretical study on kinetic isotope effects in the C–H bond activation of alkanes by iron-oxo complexes. *Coord. Chem. Rev.*, 226:251–259, 2002.
- [75] K. Yoshizawa, T. Ohta, M. Eda, and T. Yamabe. Two-step concerted mechanism for the hydrocarbon hydroxylation by cytochrome P450. *Bull. Chem. Soc. Jpn.*, 73:401–407, 2000.
- [76] K. Yoshizawa, Y. Shiota, and Y. Kagawa. Energetics for the oxygen rebound mechanism of alkane hydroxylation by the iron-oxo species of cytochrome P450. *Bull. Chem. Soc. Jpn.*, 73:2669–2673, 2000.
- [77] K. Yoshizawa, T. Kamachi, and Y. Shiota. A theoretical study of the dynamic behavior of alkane hydroxylation by a compound I model of cytochrome P450. *J. Am. Chem. Soc.*, 123:9806–9816, 2001.
- [78] M. Hata, Y. Hirano, T. Hoshino, and M. Tsuda. Monooxygenation mechanism by cytochrome P-450. *J. Am. Chem. Soc.*, 123:6410–6416, 2001.
- [79] P. H. Toy, B. Dhanabalasingam, M. Newcomb, I. H. Hanna, and P. F. Hollenberg. A substituted hypersensitive radical probe for enzyme-catalyzed hydroxylations: Synthesis of racemic and enantiomerically enriched forms and application in a cytochrome P450-catalyzed oxidation. *J. Org. Chem.*, 62:9114–9122, 1997.
- [80] P. H. Toy, M. Newcomb, and P. F. Hollenberg. Hypersensitive mechanistic probe studies of cytochrome P450-catalyzed hydroxylation reactions. Implications for the cationic pathway. *J. Am. Chem. Soc.*, 120:7719–7729, 1998.

- [81] P. H. Toy, M. Newcomb, M. J. Coon, and A. D. N. Vaz. Two distinct electrophilic oxidants effect hydroxylation in cytochrome P-450-catalyzed reactions. *J. Am. Chem. Soc.*, 120:9718–9719, 1998.
- [82] F. Ogliaro, S. P. de Visser, S. Cohen, P. K. Sharma, and S. Shaik. Searching for the second oxidant in the catalytic cycle of cytochrome P450: A theoretical investigation of the iron(III)-hydroperoxo species and its epoxidation pathways. *J. Am. Chem. Soc.*, 124:2806–2817, 2002.
- [83] S. Jin, T. M. Makris, T. A. Bryson, S. G. Sligar, and J. H. Dawson. Epoxidation of olefins by hydroperoxo-ferric cytochrome P450. *J. Am. Chem. Soc.*, 125:3406–3407, 2003.
- [84] R. A. Stearns and P. R. Ortiz de Montellano. Cytochrome P-450 catalyzed oxidation of quadricyclane – evidence for a radical cation intermediate. *J. Am. Chem. Soc.*, 107:4081–4082, 1985.
- [85] T. L. Poulos, B. C. Finzel, and A. J. Howard. Crystal Structure of substrate-free *Pseudomonas putida* cytochrome P-450. *Biochemistry*, 25:5314–5322, 1986.
- [86] H. Thomann, M. Bernardo, D. Goldfarb, P. M. H. Kroneck, and V. Ullrich. Evidence for water binding to the Fe center in cytochrome P450cam obtained by O-17 electron-spin echo envelope modulation spectroscopy. *J. Am. Chem. Soc.*, 117:8243–8251, 1995.
- [87] D. Goldfarb, M. Bernardo, H. Thomann, P. M. H. Kroneck, and V. Ullrich. Study of water binding to low-spin Fe(III) in cytochrome P450 by pulsed ENDOR and four-pulse ESEEM spectroscopies. *J. Am. Chem. Soc.*, 118:2686–2693, 1996.
- [88] S. G. Sligar and I. C. Gunsalus. Thermodynamic model of regulation: modulation of redox equilibria in camphor monooxygenase. *Proc. Natl. Acad. Sci. USA*, 73:1078–1082, 1976.
- [89] W. R. Scheidt, I. A. Cohen, and M. E. Kastner. Structural model for heme in high-spin ferric hemoproteins. Iron atom centering, porphinato core expansion, and molecular stereochemistry of high-spin diaquo(meso-tetraphenylporphinato)iron(III) perchlorate. *Biochemistry*, 18:3546–3552, 1979.
- [90] H. Aissaoui, R. Bachmann, A. Schweiger, and W.-D. Woggon. On the origin of the low-spin character of cytochrome P450(cam) in the resting state. Investigations of enzyme models with pulse EPR and ENDOR spectroscopy. *Angew. Chem. Int. Ed.*, 37:2998–3002, 1998.
- [91] D. Harris and G. Loew. Determinants of the spin state of the resting state of cytochrome P450cam. *J. Am. Chem. Soc.*, 115:8775–8779, 1993.
- [92] M. D. Segall, M. C. Payne, S. W. Ellis, G. T. Tucker, and R. N. Boyes. First principles calculation of the activity of cytochrome P450. *Phys. Rev. E*, 57:4618–4621, 1998.

- [93] M. T. Green. Role of the axial ligand in determining the spin state of resting cytochrome P450. *J. Am. Chem. Soc.*, 120:10772–10773, 1998.
- [94] D. A. Scherlis, M. Marti, P. Ordejon, and D. A. Estrin. Environment effects on chemical reactivity of heme proteins. *Int. J. Quantum Chem.*, 90:1505–1514, 2002.
- [95] A. R. Fersht. *Enzyme Structure and Mechanism*. W. H. Freeman and Co., New York, 1985.
- [96] A. Radzicka and R. Wolfenden. A proficient enzyme. *Science*, 267:90–93, 1995.
- [97] T. L. Blundell and L. N. Johnson. *Protein Crystallography*. Academic Press, New York, 1976.
- [98] K. Wüthrich. *NMR of Proteins and Nucleic Acids*. John Wiley and Sons, New York, 1986.
- [99] J. A. McCammon and S. C. Harvey. *Dynamics of Proteins and Nucleic Acids*. Cambridge University Press, New York, 1987.
- [100] C. L. Brooks III, M. Karplus, and B. M. Pettitt. *Proteins: A Theoretical Perspective of Dynamics, Structure and Thermodynamics*. John Wiley and Sons, New York, 1988.
- [101] D. Blow. So do we understand how enzymes work? *Struct. Fold. Des.*, 8:R77–R81, 2000.
- [102] M. J. Field. Simulating enzyme reactions: Challenges and perspectives. *J. Comput. Chem.*, 23:48–58, 2002.
- [103] S. L. Dixon and K. M. Merz. Divide and conquer for semiempirical MO methods. In P. v. R. Schleyer, editor, *Encyclopedia of Computational Chemistry*, volume 1, pages 761–776. Wiley, Chichester, 1998.
- [104] W. Yang and J. M. Perez-Jorda. Linear scaling methods for electronic structure calculations. In P. v. R. Schleyer, editor, *Encyclopedia of Computational Chemistry*, volume 2, pages 1496–1513. Wiley, Chichester, 1998.
- [105] J. J. P. Stewart. Localized MO SCF methods. In P. v. R. Schleyer, editor, *Encyclopedia of Computational Chemistry*, volume 2, pages 1513–1519. Wiley, Chichester, 1998.
- [106] G. E. Scuseria. Linear scaling density functional calculations with Gaussian orbitals. *J. Phys. Chem. A*, 103:4782–4790, 1999.
- [107] A. Warshel and M. Levitt. Theoretical studies of enzymatic reactions – dielectric, electrostatic, and steric stabilization of carbonium-ion in reaction of lysozyme. *J. Mol. Biol.*, 103:227–249, 1976.

- [108] U. C. Singh and P. A. Kollman. A combined ab-initio quantum-mechanical and molecular mechanical method for carrying out simulations on complex molecular systems. Applications to the $\text{CH}_3\text{Cl} + \text{Cl}$ exchange reaction and gas phase protonation of polyethers. *J. Comput. Chem.*, 7:718–730, 1986.
- [109] M. J. Field, P. A. Bash, and M. Karplus. A combined quantum-mechanical and molecular mechanical potential for molecular dynamics simulations. *J. Comput. Chem.*, 11:700–733, 1990.
- [110] J. Gao. Methods and applications of combined quantum mechanical and molecular mechanical potentials. In K. B. Lipkowitz and D. B. Boyd, editors, *Reviews in Computational Chemistry*, volume 7, pages 119–185. VCH Publishers, Inc., New York, 1996.
- [111] T. Z. Mordasini and W. Thiel. Combined quantum mechanical and molecular mechanical approaches. *Chimia*, 58:288–291, 1998.
- [112] G. Monard and K. M. Merz Jr. Combined quantum mechanical/molecular mechanical methodologies applied to biomolecular systems. *Acc. Chem. Res.*, 32:904–911, 1999.
- [113] P. Sherwood, A. H. de Vries, S. J. Collins, S. P. Greatbanks, N. A. Burton, M. A. Vincent, and I. H. Hillier. Hybrid quantum mechanics/molecular mechanics approaches. *Faraday Discuss.*, 106:79–92, 1997.
- [114] U. Eichler, C. M. Kölmel, and J. Sauer. Combining ab initio techniques with analytical potential functions for structure predictions of large systems: Method and application to crystalline silica polymorphs. *J. Comput. Chem.*, 18:463–477, 1997.
- [115] D. Nachtigallova, P. Nachtigall, M. Sierka, and J. Sauer. Coordination and siting of Cu^+ ions in ZSM-5: A combined quantum mechanics and interatomic potential function study. *Phys. Chem. Chem. Phys.*, 1:2019–2026, 1999.
- [116] P. Sherwood. Hybrid quantum mechanics/ molecular mechanics approaches. In J. Groendorst, editor, *Proceedings of the Winterschool on Modern Methods and Algorithms of Quantum Chemistry*, volume 1, pages 257–277. John von Neumann Institute for Computing, Research Center Jülich, D-52425 Jülich, Germany, 2000.
- [117] P. Bala, P. Grochowski, B. Lesyng, and J. A. McCammon. Quantum-classical molecular dynamics simulations of proton-transfer processes in molecular complexes and in enzymes. *J. Phys. Chem.*, 100:2535–2545, 1996.
- [118] P. Grochowski, B. Lesyng, P. Bala, and J. A. McCammon. Density-functional based parametrization of a valence bond method and its applications in quantum-classical molecular dynamics simulations of enzymatic reactions. *Int. J. Quantum Chem.*, 60:1143–1164, 1996.

- [119] F. Maseras and K. Morokuma. IMOMM – a new integrated ab-initio plus molecular mechanics geometry optimization scheme of equilibrium structures and transition states. *J. Comput. Chem.*, 16:1170–1179, 1995.
- [120] M. Svensson, S. Humbel, R. J. Froese, T. Matsubara, S. Sieber, and K. Morokuma. ONIOM: A multilayered integrated MO + MM method for geometry optimizations and single point energy predictions. A test for Diels-Alder reactions and Pt(P(t-Bu)(3))(2)+H2 oxidative addition. *J. Phys. Chem.*, 100:19357–19363, 1996.
- [121] P. N. Day, J. H. Jensen, M. S. Gordon, S. P. Webb, W. J. Stevens, M. Krauss, D. Garmer, H. Basch, and S. Cohen. An effective fragment method for modeling solvent effects in quantum mechanical calculations. *J. Chem. Phys.*, 105:1968–1986, 1996.
- [122] G. N. Merrill and M. S. Gordon. Study of small water clusters using the effective fragment potential method. *J. Phys. Chem. A*, 102:2650–2657, 1998.
- [123] A. Warshel. *Computer Modelling of Chemical Reactions in Enzymes and Solutions*. John Wiley and Sons, New York, 1991.
- [124] J. Åqvist and A. Warshel. Simulation of enzyme reactions using valence bond force fields and other hybrid quantum-classical approaches. *Chem. Rev.*, 93:2523–2544, 1993.
- [125] Y. Q. Tu and A. Laaksonen. On the effect of Lennard-Jones parameters on the quantum mechanical and molecular mechanical coupling in a hybrid molecular dynamics simulation of liquid water. *J. Chem. Phys.*, 111:7519–7525, 1999.
- [126] J. Gao and X. Xia. A priori evaluation of aqueous polarization effects through Monte Carlo QM-MM simulations. *Science*, 258:631–635, 1992.
- [127] D. Bakowies and W. Thiel. Hybrid models for combined quantum mechanical and molecular mechanical approaches. *J. Phys. Chem.*, 100:10580–10594, 1996.
- [128] I. Antes and W. Thiel. On the treatment of link atoms in hybrid methods. In J. Gao and M. A. Thompson, editors, *Combined Quantum Mechanical and Molecular Mechanical Methods*, volume 712 of *ACS Symposium Series*, pages 50–65. American Chemical Society, Washington, DC, 1998.
- [129] N. Reuter, A. Dejaegere, B. Maigret, and M. Karplus. Frontier bonds in QM/MM methods: A comparison of different approaches. *J. Phys. Chem. A*, 104:1720–1735, 2000.
- [130] M. J. Field, M. Albe, C. Bret, F. Proust-De Martin, and A. Thomas. The Dynamo library for molecular simulations using hybrid quantum mechanical and molecular mechanical potentials. *J. Comput. Chem.*, 21:1088–1100, 2000.

- [131] G. Monard, M. Loos, V. Thery, K. Baka, and J.-L. Rivail. Hybrid classical quantum force field for modeling very large molecules. *Int. J. Quantum Chem.*, 58:153–159, 1996.
- [132] J. Gao, P. Amara, C. Alhambra, and M. J. Field. A generalized hybrid orbital (GHO) method for the treatment of boundary atoms in combined QM/MM calculations. *J. Phys. Chem. A*, 102:4714–4721, 1998.
- [133] I. Antes and W. Thiel. Adjusted connection atoms for combined quantum mechanical and molecular mechanical methods. *J. Phys. Chem. A*, 103:9290–9295, 1999.
- [134] X. Assfeld and J.-L. Rivail. Quantum chemical computations on parts of large molecules: The ab initio local self consistent field method. *Chem. Phys. Lett.*, 263:100–106, 1996.
- [135] D. M. Phillip and R. A. Friesner. Mixed ab initio QM/MM modelling using frozen orbitals and tests with alanine dipeptide and tetrapeptide. *J. Comput. Chem.*, 20:1468–1494, 1999.
- [136] R. B. Murphy, D. M. Philipp, and R. A. Friesner. Frozen orbital QM/MM methods for density functional theory. *Chem. Phys. Lett.*, 321:113–120, 2000.
- [137] R. B. Murphy, D. M. Phillip, and R. A. Friesner. A mixed quantum mechanics/molecular mechanics (QM/MM) methods for large-scale modeling of chemistry in protein environments. *J. Comput. Chem.*, 21:1442–1457, 2000.
- [138] Y. Zhang, T.-S. Lee, and W. Yang. A pseudobond approach to combining quantum mechanical and molecular mechanical methods. *J. Chem. Phys.*, 110:46–54, 1999.
- [139] J. Gao and D. G. Truhlar. Quantum mechanical methods for enzyme kinetics. *Annu. Rev. Phys. Chem.*, 53:467–505, 2002.
- [140] A. R. Leach. *Molecular Modelling: Principles and Applications*. Prentice Hall, Harlow, UK, 2nd edition, 2001.
- [141] T. K. Woo, P. M. Margl, P. E. Blöchl, and T. Ziegler. A combined Car-Parrinello QM/MM implementation for ab-initio molecular dynamics simulations of extended systems: Application to transition metal catalysis. *J. Phys. Chem. B*, 101:7877–7880, 1997.
- [142] A. Laio, J. VandeVondele, and U. Rothlisberger. A Hamiltonian electrostatic coupling scheme for hybrid Car-Parrinello molecular dynamics simulations. *J. Chem. Phys.*, 116:6941–6947, 2002.
- [143] M. Eichinger, P. Tavan, J. Hutter, and M. Parrinello. A hybrid method for solutes in complex solvents: Density functional theory combined with empirical force fields. *J. Chem. Phys.*, 110:10452–10467, 1999.

- [144] S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta, and P. Weiner. A new force-field for molecular mechanical simulation of nucleic-acids and proteins. *J. Am. Chem. Soc.*, 106:765–784, 1984.
- [145] W. D. Cornell, Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. A second generation force field for the simulation of proteins, nucleic acids and organic molecules. *J. Am. Chem. Soc.*, 117:5179–5197, 1995.
- [146] A. D. McKerell Jr., D. Bashford, M. Bellott, R. L. Dunbrack Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczero, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyend, B. Prodhom, W. W. Reiher III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczero, D. Yin, and M. Karplus. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B*, 102:3586, 1998.
- [147] W. F. v. Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. R. P. Scott, and I. G. Tironi. *Biomolecular Simulation: The GROMOS96 manual and user guide*. Biomos, Zürich and Groningen, 1996.
- [148] W. L. Jorgensen and J. Tirado-Rives. The OPLS potential functions for proteins - energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.*, 110:1657–1666, 1988.
- [149] M. J. Hwang, T. P. Stockfish, and A. T. Hagler. Derivation of class II force fields. 2. Derivation and characterization of a class II force field, CFF93, for the alkyl functional group and alkane molecules. *J. Am. Chem. Soc.*, 116:2515–2525, 1994.
- [150] N. L. Allinger, Y. H. Yuh, and J.-H. Lii. Molecular Mechanics. The MM3 force field for hydrocarbons, 1. *J. Am. Chem. Soc.*, 111:8551–8565, 1989.
- [151] J. D. Gale. GULP: A computer program for the symmetry-adapted simulation of solids. *J. Chem. Soc., Faraday Trans.*, 93:629–637, 1997.
- [152] W. Koch and M. C. Holthausen. *A Chemist's Guide to Density Functional Theory*. Wiley-VCH, Weinheim, 2nd edition, 2002.
- [153] R. G. Parr and W. Yang. *Density-Functional Theory of Atoms and Molecules*. Oxford University Press, New York, 1989.
- [154] R. G. Parr and W. Yang. Density-functional theory of the electronic structure of molecules. *Annu. Rev. Phys. Chem.*, 46:701–728, 1995.

- [155] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev. B*, 136:864–871, 1964.
- [156] A. Szabo and N. S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. MacMillan Publishing Co., New York, 1st edition, 1982.
- [157] W. Kohn and L. J. Sham. Self consistent equations including exchange and correlation effects. *Phys. Rev. A*, 140:1133–1138, 1965.
- [158] A. Savin. On degeneracy, near degeneracy and density functional theory. In J. M. Seminario, editor, *Recent Developments and Applications of Modern Density Functional Theory*, pages 327–357. Elsevier, Amsterdam, 1996.
- [159] O. V. Gritsenko and E. J. Baerends. Electron correlation effects on the shape of the Kohn–Sham molecular orbitals. *Theor. Chem. Acc.*, 96:44–50, 1997.
- [160] P. R. T. Schipper, O. V. Gritsenko, and E. J. Baerends. One-determinantal pure state versus ensemble Kohn–Sham solutions in the case of strong electron correlation: CH₂ and C₂. *Theor. Chem. Acc.*, 99:329–343, 1998.
- [161] C. W. Bauschlicher, D. M. Hudgins, and L. J. Allamandola. The infrared spectra of polycyclic aromatic hydrocarbons containing a five-membered ring: Symmetry breaking and the B3LYP functional. *Theor. Chem. Acc.*, 103:154–162, 1999.
- [162] M. Alcami, O. Mo, M. Yanez, and I. L. Cooper. The performance of density-functional theory in challenging cases: Halogen oxides. *J. Chem. Phys.*, 112:6131–6140, 2000.
- [163] A. H. de Vries, P. Sherwood, S. J. Collins, A. M. Rigby, M. Rigutto, and G. J. Kramer. Zeolite structure and reactivity by combined quantum-chemical-classical calculations. *J. Chem. Phys. B*, 103:6133–6141, 1999.
- [164] R. Weiss, D. Mandon, T. Wolter, A. X. Trautwein, M. Mütter, E. Bill, A. Gold, K. Jayaraj, and J. Turner. Delocalization over the heme and the axial ligands of one of the two oxidizing equivalents stored above the ferric state in the peroxidase and catalase compound-I intermediates: Indirect participation of the proximal axial ligand of iron in the oxidation reactions catalyzed by heme-based peroxidases and catalases? *J. Biol. Inorg. Chem.*, 1:377–383, 1996.
- [165] M. T. Green. Imidazole-ligated compound I intermediates: The effects of hydrogen bonding. *J. Am. Chem. Soc.*, 122:9495–9499, 2000.
- [166] T. L. Poulos, J. Cupp-Vickery, and H. Li. Structural studies on prokaryotic cytochromes P450. Ref. [1], pp. 125–150.

- [167] T. L. Poulos. The role of the proximal ligand in heme enzymes. *J. Biol. Inorg. Chem.*, 1:356–359, 1996.
- [168] J. A. Sigman, A. E. Pond, J. H. Dawson, and Y. Lu. Engineering cytochrome c peroxidase into cytochrome P450: A proximal effect on heme-thiolate ligation. *Biochemistry*, 38:11122–11129, 1999.
- [169] S. Yoshioka, S. Takahashi, K. Ishimori, and I. Morishima. Roles of the axial push effect in cytochrome P450cam studied with the site-directed mutagenesis at the heme proximal site. *J. Inorg. Biochem.*, 81:141–151, 2000.
- [170] V. Lounnas and R. C. Wade. Exceptionally stable salt bridges in cytochrome P450cam have functional roles. *Biochemistry*, 36:5402–5417, 1997.
- [171] Accelrys, Inc., San Diego, CA, USA. *Insight II: An integrated modeling environment*, 2000.
- [172] B. R. Brooks, R. E. Burccoleri, B. D. Olafson, D. J. States, and M. Karplus. CHARMM – a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.*, 4:187–217, 1983.
- [173] W. F. van Gunsteren and H. H. C. Berendsen. Algorithms for macromolecular dynamics and constraint dynamics. *Mol. Phys.*, 34:1311–1327, 1977.
- [174] H. Lee, P. R. Ortiz de Montellano, and A. E. McDermott. Deuterium magic angle spinning studies of substrates bound to cytochrome P450. *Biochemistry*, 38:10808–10813, 1999.
- [175] S. Yoshioka, T. Tosha, S. Takahashi, K. Ishimori, H. Hori, and I. Morishima. Roles of the proximal hydrogen bonding network in cytochrome P450cam-catalyzed oxygenation. *J. Am. Chem. Soc.*, 124:14571–14579, 2002.
- [176] S. Cohen and S. Shaik. Unpublished results.
- [177] P. J. Hay and W. R. Wadt. Ab initio effective core potentials for molecular calculations. Potentials for K to Au including the outermost core orbitals. *J. Chem. Phys.*, 82:299–310, 1985.
- [178] R. Ditchfield, W. J. Hehre, and J. A. Pople. Self-consistent molecular-orbital methods. 9. Extended gaussian-type basis for molecular-orbital studies of organic molecules. *J. Chem. Phys.*, 54:724–732, 1971.
- [179] W. J. Hehre, R. Ditchfield, and J. A. Pople. Self-consistent molecular-orbital methods. 12. Further extensions of gaussian-type basis sets for use in molecular-orbital studies of organic molecules. *J. Chem. Phys.*, 56:2257–2261, 1972.

- [180] P. C. Hariharan and J. A. Pople. Influence of polarization functions on molecular-orbital hydrogenation energies. *Theor. Chim. Acta*, 28:213–222, 1973.
- [181] T. Clark, J. Chandrasekhar, G. W. Spitznagel, and P. v. R. Schleyer. Efficient diffuse function augmented basis sets for anion calculations. 3. The 3-21+G basis set for 1st-row elements Li–F. *J. Comput. Chem.*, 4:294–301, 1983.
- [182] K. Yamaguchi, F. Jensen, A. Dorigo, and K. N. Houk. A spin correction procedure for unrestricted Hartree–Fock and Møller–Plesset wavefunctions for singlet diradicals and polyradicals. *Chem. Phys. Lett.*, 149:537–542, 1988.
- [183] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A*, 38:3098–3100, 1988.
- [184] A. D. Becke. Density-functional thermochemistry. 3. The role of exact exchange. *J. Chem. Phys.*, 98:5648–5652, 1993.
- [185] C. Lee, W. Yang, and R. G. Parr. Development of the Colle–Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B*, 37:785–789, 1988.
- [186] C. Lennartz, A. Schäfer, F. Terstegen, and W. Thiel. Enzymatic reactions of triosephosphate isomerase: A theoretical calibration study. *J. Phys. Chem. B*, 106:1758–1767, 2002.
- [187] R. Ahlrichs, M. Bär, M. Häser, H. Horn, and C. Kölmel. Electronic-structure calculations on workstation computers - the program system TURBOMOLE. *Chem. Phys. Lett.*, 162:165–169, 1989.
- [188] R. Ahlrichs, M. Bär, H.-P. Baron, R. Bauernschmitt, S. Böcker, M. Ehrig, K. Eichkorn, S. Elliot, F. Furche, M. Häser, H. Horn, C. Hättig, C. Huber, U. Huniar, M. Kattanneck, A. Köhn, C. Kölmel, M. Kollwitz, K. May, C. Ochsenfeld, H. Öhm, A. Schäfer, U. Schneider, O. Treutler, M. v. Arnim, F. Weigend, P. Weis, and H. Weiss. *TURBOMOLE 5.5 - Program package for ab initio electronic structure calculations*. University of Karlsruhe, 2002.
- [189] GAMESS-UK is a package of ab initio programs written by M. F. Guest, J. H. van Lenthe, J. Kendrick, K. Schoffel, and P. Sherwood, with contributions from R. D. Amos, R. J. Buenker, J. J. van Dam, M. Dupuis, N. C. Handy, I. H. Hillier, P. J. Knowles, V. Bonacic-Koutecky, W. von Niessen, R. J. Harrison, A. P. Rendell, V. R. Saunders, A. J. Stone, D. J. Tozer, and A. H. de Vries. The package is derived from the original GAMESS code of M. Dupuis, D. Spangler, and J. Wendolowski. NRCC, Software Catalog, Vol. 1, Program No. QG01 (GAMESS), 1980.

- [190] ChemShell is a modular QM/MM program based on the TCL interpreter and developed in the European QUASI project under the coordination of P. Sherwood. See: <http://www.cse.clrc.ac.uk/qcg/chemshell/>.
- [191] W. Smith and T. R. Forester. DL-POLY 2.0: A general-purpose parallel molecular dynamics simulation package. *J. Mol. Graphics*, 14:136–141, 1996.
- [192] S. R. Billeter, A. J. Turner, and W. Thiel. Linear scaling geometry optimisation and transition state search in hybrid delocalised internal coordinates. *Phys. Chem. Chem. Phys.*, 2:2177–2186, 2000.
- [193] Ref. [152], pp. 251–259.
- [194] N. Suzuki, T. Higuchi, Y. Urano, K. Kikuchi, H. Uekusa, Y. Ohashi, T. Uchida, T. Kitagawa, and T. Nagano. Novel iron porphyrin-alkanethiolate complex with intramolecular $\text{NH} \cdots \text{S}$ hydrogen bond: Synthesis, spectroscopy, and reactivity. *J. Am. Chem. Soc.*, 121:11571–11572, 1999.
- [195] J. E. Roberts, B. M. Hoffman, R. Rutter, and L. P. Hager. Electron-nuclear double resonance of horseradish peroxidase compound I. Detection of the porphyrine π -cation radical. *J. Biol. Chem.*, 256:2118–2121, 1981.
- [196] J. P. Perdew. Density functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B*, 33:8822–8824, 1986.
- [197] A. D. Becke. Density functional thermochemistry. 5. Systematic optimization of exchange-correlation functionals. *J. Chem. Phys.*, 107:8554–8560, 1997.
- [198] C. Adamo and V. Barone. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.*, 110:6158–6170, 1999.
- [199] R. D. Amos, A. Bernhardsson, A. Berning, P. Celani, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, C. Hampel, G. Hetzer, P. J. Knowles, T. Korona, R. Lindh, A. W. Lloyd, S. J. McNicholas, F. R. Manby, W. Meyer, M. E. Mura, A. Nicklass, P. Palmieri, R. Pitzer, G. Rauhut, M. Schütz, U. Schumann, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, and H.-J. Werner. MOLPRO, a package of ab initio programs designed by H.-J. Werner and P. J. Knowles, version 2002.1, 2002.
- [200] A. Schäfer, H. Horn, and R. Ahlrichs. Fully optimized contracted gaussian basis sets for atoms Li to Kr. *J. Chem. Phys.*, 97:2571–2577, 1992.
- [201] A. Schäfer, C. Huber, and R. Ahlrichs. Fully optimized contracted gaussian basis sets of triple zeta quality for atoms Li to Kr. *J. Chem. Phys.*, 100:5829–5835, 1994.

- [202] T. H. Dunning Jr. Gaussian basis sets for use in correlated molecular calculations. 1. The atoms boron through neon and hydrogen. *J. Chem. Phys.*, 90:1007–1023, 1989.
- [203] D. E. Woon and T. H. Dunning Jr. Gaussian basis sets for use in correlated molecular calculations. 3. The atoms aluminium through argon. *J. Chem. Phys.*, 98:1358–1371, 1993.
- [204] A. J. H. Wachters. Gaussian basis sets for molecular wavefunctions containing third-row atoms. *J. Chem. Phys.*, 52:1033–1042, 1970.
- [205] P. J. Hay. Gaussian basis sets for molecular calculations. Representation of 3d orbitals in transition metal atoms. *J. Chem. Phys.*, 66:4377–4384, 1977.
- [206] C. W. Bauschlicher Jr., S. R. Langhoff, and L. A. Barnes. Theoretical studies of the 1st-row and 2nd-row transition metal methyls and their positive ions. *J. Chem. Phys.*, 91:2399–2411, 1989.
- [207] W.-D. Woggon and S. Shaik, 2003. Personal communication.
- [208] J.-U. Rohde, J.-H. In, M. H. Lim, W. W. Brennessel, M. R. Bukowski, A. Stubna, E. Münck, W. Nam, and L. Que Jr. Crystallographic and spectroscopic characterization of a nonheme Fe(IV)=O complex. *Science*, 299:1037–1038, 2003.
- [209] D. G. Truhlar, J. Gao, C. Alhambra, M. Garcia-Viloca, J. Corchado, M. L. Sanchez, and J. Villa. The incorporation of quantum effects in enzyme kinetics modeling. *Acc. Chem. Res.*, 35:341–349, 2002.
- [210] M. Filatov, N. Harris, and S. Shaik. On the “rebound” mechanism of alkane hydroxylation by cytochrome P450: Electronic structure of the intermediate and the electron transfer character in the rebound step. *Angew. Chem. Int. Ed.*, 38:3510–3512, 1999.
- [211] S. P. de Visser, F. Ogliaro, P. K. Sharma, and S. Shaik. What factors affect the regioselectivity of oxidation by cytochrome P450? A DFT study of allylic hydroxylation and double bond epoxidation in a model reaction. *J. Am. Chem. Soc.*, 124:11809–11826, 2002.
- [212] H. Li, N. Shakunthala, L. M. Havran, J. D. Winkler, and T. L. Poulos. Crystal-structure of cytochrome P450(cam) complexed with its catalytic product, 5-*exo*-hydroxycamphor. *J. Am. Chem. Soc.*, 117:6297–6299, 1995.
- [213] J. D. Lipscomb. Electron-paramagnetic resonance detectable states of cytochrome P-450cam. *Biochemistry*, 19:3590–3599, 1980.
- [214] H. Lin and W. Thiel. Unpublished results.

- [215] M. Imai, H. Shimada, Y. Watanabe, Y. Matsushima-Hibiya, R. Makino, H. Koga, T. Horiuchi, and Y. Ishimura. Uncoupling of the cytochrome P-450_{cam} monooxygenase reaction by a single mutation, threonine-252 to alanine or valine. A possible role of the hydroxy amino-acid in oxygen activation. *Proc. Natl. Acad. Sci. USA*, 86:7823–7827, 1989.
- [216] V. Guallar, B. F. Gherman, W. H. Miller, S. J. Lippard, and R. A. Friesner. Dynamics of alkane hydroxylation at the non-heme diiron center in methane monooxygenase. *J. Am. Chem. Soc.*, 124:3377–3384, 2002.
- [217] M. Reiher, O. Salomon, and B. A. Hess. Reparameterization of hybrid functionals based on energy differences of states of different multiplicity. *Theor. Chem. Acc.*, 107:48–55, 2001.
- [218] O. Salomon, M. Reiher, and B. A. Hess. Assertion and validation of the performance of the B3LYP* functional for the first transition metal row and the G2 test set. *J. Chem. Phys.*, 117:4729–4737, 2002.
- [219] D. Schröder, A. Fiedler, J. Hrusak, and H. Schwarz. Experimental and theoretical studies toward a characterization of conceivable intermediates involved in the gas-phase oxidation of methane by bare FeO⁺. *J. Am. Chem. Soc.*, 114:1215–1222, 1992.
- [220] M. Filatov, W. Reckien, S. D. Peyerimhoff, and S. Shaik. What are the reasons for the kinetic stability of a mixture of H₂ and O₂? *J. Phys. Chem. A*, 104:12014–12020, 2000.
- [221] D. A. Scherlis and D. A. Estrin. Structure and spin-state energetics of an iron porphyrin model: an assessment of theoretical methods. *Int. J. Quantum Chem.*, 87:158–166, 2002.
- [222] A. D. Becke. A new mixing of Hartree–Fock and local density functional theories. *J. Chem. Phys.*, 98:1372–1377, 1993.
- [223] J. Turner, A. Gold, R. Weiss, D. Mandon, and A. X. Trautwein. Symmetry states of metalloporphyrin π -cation radicals, models for peroxidase compound I. *J. Porph. Phtalolycanines*, 5:357–364, 2001.
- [224] D. Mandon, R. Weiss, K. Jayaraj, A. Gold, J. Turner, E. Bill, and A. X. Trautwein. Models for peroxidase compound I: Generation and spectroscopic characterization of new oxoferryl porphyrin π cation radical species. *Inorg. Chem.*, 31:4404–4409, 1992.
- [225] R. LoBrutto, C. P. Scholes, G. C. Wagner, I. C. Gunsalus, and P. G. Debrunner. Electron nuclear double resonance of ferric cytochrome P450_{cam}. *J. Am. Chem. Soc.*, 102:1167–1170, 1980.
- [226] Y.-C. Fann, N. C. Gerber, P. A. Omulski, L. P. Hager, S. G. Sligar, and B. M. Hoffman. EN-DOR determination of heme ligation in chloroperoxidase and comparison with cytochrome P450. *J. Am. Chem. Soc.*, 116:5989–5990, 1994.

- [227] F. Illas, I. D. R. Moreira, C. de Graaf, and V. Barone. Magnetic coupling in biradicals, binuclear complexes and wide-gap insulators: A survey of ab initio wave function and density functional theory approaches. *Theor. Chem. Acc.*, 104:265–272, 2000.
- [228] L. Noodleman and E. R. Davidson. Ligand spin polarization and antiferromagnetic coupling in transition-metal dimers. *Chem. Phys.*, 109:131–143, 1986.
- [229] L. Noodleman. Valence bond description of antiferromagnetic coupling in transition metal dimers. *J. Chem. Phys.*, 74:5737–5743, 1981.
- [230] A. Bencini and D. Gatteschi. *EPR of Exchange Coupled Systems*. Springer, Berlin, 1990.
- [231] M. Zheng, S. V. Khangulov, G. C. Dismuskes, and V. V. Barynin. Electronic structure of dimanganese(II,III) and dimanganese(III,IV) complexes and dimanganese catalase enzyme: A general EPR spectral simulation approach. *Inorg. Chem.*, 33:382–387, 1994.
- [232] L. Noodleman and E. J. Baerends. Electronic structure, magnetic properties, electron-spin-resonance, and optical spectra for 2-Fe ferredoxin models by LCAO- $X\alpha$ valence bond theory. *J. Am. Chem. Soc.*, 2316-2327:106, 1984.
- [233] H. Kuramochi, L. Noodleman, and D. A. Case. Density functional study on the electronic structures of model peroxidase compounds I and II. *J. Am. Chem. Soc.*, 119:11442–11451, 1997.
- [234] J. E. Harriman. *Theoretical Foundations of Electron Spin Resonance*. Academic Press, New York, 1978.
- [235] F. E. Mabbs and D. Collison. *Electron Paramagnetic Resonance of d-Transition Metal Compounds*. Elsevier, Amsterdam, 1992.
- [236] J. B. Pilbrow. *Transition Ion Electron Paramagnetic Resonance*. Clarendon Press, Oxford, 1990.
- [237] F. Neese and E. I. Solomon. Interpretation and calculation of spin-Hamiltonian parameters in transition metal complexes. In J. S. Miller and M. Drillon, editors, *Magnetoscience - From Molecules to Materials*, volume 4, pages 345–466. Wiley, New York, 2002.
- [238] M. Blume, A. J. Freeman, and R. E. Watson. Theory of spin-orbit coupling in atoms. 3. *Phys. Rev. A*, 134:320–323, 1964.
- [239] S. Koseki, M. W. Schmidt, and M. S. Gordon. MCSCF/6-31G(d,p) calculations of one-electron spin-orbit-coupling constants in diatomic molecules. *J. Phys. Chem. A*, 102:10768–10772, 1992.

- [240] S. Koseki, M. S. Gordon, M. W. Schmidt, and N. Matsunaga. Main-group effective nuclear charges for spin-orbit calculations. *J. Phys. Chem.*, 99:12764–12772, 1995.
- [241] S. Koseki, M. W. Schmidt, and M. S. Gordon. Effective nuclear charges for the first-through third-row transition metal elements in spin-orbit calculations. *J. Phys. Chem. A*, 102:10430–10435, 1998.
- [242] B. A. Hess, C. M. Marian, U. Wahlgren, and O. Gropen. A mean-field spin-orbit method applicable to correlated wavefunctions. *Chem. Phys. Lett.*, 251:365–371, 1996.
- [243] R. McWeeny. *Methods of Molecular Quantum Mechanics*. Academic Press, London, 1992.
- [244] S. Brownridge, F. Grein, J. Tatchen, M. Kleinschmidt, and C. M. Marian. Efficient calculation of electronic paramagnetic resonance g-tensors by multireference configuration interaction sum-over-state expansions, using the atomic mean-field spin-orbit method. *J. Chem. Phys.*, 118:9552–9562, 2003.
- [245] F. Neese. Metal and ligand hyperfine couplings in transition metal complexes: The effect of spin-orbit coupling as studied by coupled perturbed Kohn–Sham theory. *J. Chem. Phys.*, 118:3939–3984, 2003.
- [246] P. Belanzoni, E. Van Lenthe, and E. J. Baerends. An evaluation of the density functional approach in the zero order regular approximation for relativistic effects: Magnetic interactions in small metal compounds. *J. Chem. Phys.*, 114:4421–4433, 2001.
- [247] M. Stein, E. Van Lenthe, E. J. Baerends, and W. Lubitz. g- and A-tensor calculations in the zero-order approximation for relativistic effects of Ni complexes $\text{Ni}(\text{mnt})_2^-$ and $\text{Ni}(\text{CO})_3\text{H}$ as model complexes for the active center of [NiFe]-Hydrogenase. *J. Phys. Chem. A*, 105:416–435, 2001.
- [248] M. Stein, E. Van Lenthe, E. J. Baerends, and W. Lubitz. Relativistic DFT calculations of the paramagnetic intermediates of [NiFe] hydrogenase. Implications for the enzymatic mechanism. *J. Am. Chem. Soc.*, 123:5839–5840, 2001.
- [249] F. Neese. Prediction of electron paramagnetic resonance g values using coupled perturbed Hartree–Fock and Kohn–Sham theory. *J. Chem. Phys.*, 115:11080–11096, 2001.
- [250] G. Schreckenbach and T. Ziegler. Calculation of the g-tensor of electron paramagnetic resonance spectroscopy using gauge-including atomic orbitals and density functional theory. *J. Phys. Chem. A*, 101:3388–3399, 1997.
- [251] S. Patchkovskii and T. Ziegler. Calculation of the EPR g-tensor of high-spin radicals with density functional theory. *J. Phys. Chem. A*, 105:5490–5497, 2001.

- [252] O. L. Malkina, V. Vaara, B. Schimmelpfennig, M. Munzarova, V. Malkin, and M. Kaupp. Density functional calculations of electronic g-tensors using spin-orbit pseudopotentials and mean-field all-electron spin-orbit operators. *J. Am. Chem. Soc.*, 122:9206–9218, 2000.
- [253] M. Kaupp, R. Reviakine, O. L. Malkina, A. Arbuznikov, B. Schimmelpfennig, and V. G. Malkin. Calculation of electronic g-tensors for transition metal complexes using hybrid density functionals and atomic mean-field spin-orbit operators. *J. Comput. Chem.*, 23:794–803, 2002.
- [254] E. Van Lenthe, P. E. S. Wormer, and A. van der Avoird. Density functional calculations of molecular g-tensors in the zero-order regular approximation for relativistic effects. *J. Chem. Phys.*, 107:2488–2498, 1997.
- [255] K. M. Neyman, D. I. Ganyushin, A. V. Matveev, and V. A. Nasluzov. Calculation of electronic g-tensors using a relativistic density functional Douglas-Kroll method. *J. Phys. Chem. A*, 106:5022–5030, 2002.
- [256] A. V. Luzanov, E. N. Babich, and V. V. Ivanov. Gauge-invariant calculations of magnetic properties in semiempirical approaches. Application to full-CI π -electron scheme. *J. Mol. Struct.: THEOCHEM*, 311:211–220, 1994.
- [257] R. Ditchfield. Self-consistent perturbation-theory of diamagnetism. 1. Gauge-invariant LCAO method for NMR chemical-shifts. *Mol. Phys.*, 27:789–807, 1974.
- [258] W. Kutzelnigg, U. Fleischer, and M. Schindler. The IGLO-method: Ab-initio calculation and interpretation of NMR chemical shifts and magnetic susceptibilities. In P. Diehl, E. Fluck, H. Günther, R. Kosfield, and J. Seeling, editors, *NMR Basic Principles and Progress*, volume 213, pages 165–262. Springer, Berlin/Heidelberg, 1991.
- [259] M. L. Munzarova, P. Kubacek, and M. Kaupp. Mechanisms of EPR hyperfine coupling in transition metal complexes. *J. Am. Chem. Soc.*, 122:11900–11913, 2000.
- [260] V. Barone. Structure, magnetic properties and reactivities of open-shell species from density functional and self-consistent hybrid methods. In D. P. Chong, editor, *Recent Advances in Density Functional Methods (Part I)*, pages 287–334. World Scientific, Singapore, 1995.
- [261] M. L. Munzarova and M. Kaupp. A critical validation of density functional and coupled-cluster approaches for the calculation of EPR hyperfine coupling constants in transition metal complexes. *J. Phys. Chem. A*, 103:9966–9983, 1999.
- [262] A. X. Trautwein, E. Bill, E. L. Bominaar, and H. Winkler. Iron-containing proteins and related analogs - complementary Mössbauer, EPR, and magnetic susceptibility studies. *Structure and Bonding*, 78:1–95, 1991.

- [263] Y. Zhang, J. Mao, and E. Oldfield. Fe-57 Moessbauer isomer shifts of heme protein model systems: Electronic structure calculations. *J. Am. Chem. Soc.*, 124:7829–7839, 2002.
- [264] T. Lovell, J. Li, T. Liu, and L. Noodleman. FeMo cofactor of nitrogenase: A DFT study of states M^N , M^{OX} , M^R , and M^I . *J. Am. Chem. Soc.*, 123:12392–12410, 2001.
- [265] T. Lovell, W. G. Han, T. Liu, and L. Noodleman. A structural model for the high-valent intermediate Q of methane monooxygenase from broken-symmetry density functional and electrostatic calculations. *J. Am. Chem. Soc.*, 124:5890–5894, 2002.
- [266] F. Neese. Prediction and interpretation of the ^{57}Fe isomer shift in Moessbauer spectra by density functional theory. *Inorg. Chim. Acta*, 337:181–192, 2002.
- [267] F. Neese. Quantum chemical calculation of spectroscopic properties of metalloproteins and model compounds: EPR and Mößbauer properties. *Curr. Op. Chem. Biol.*, 7:125–135, 2003.
- [268] N. Godbout, R. H. Havlin, R. Salzmänn, P. G. Debrunner, and E. Oldfield. Iron-57 NMR chemical shifts and Moessbauer quadrupole splittings in metalloporphyrins, ferrocytochrome c, and myoglobins: a density functional theory investigation. *J. Phys. Chem. A*, 102:2342–2350, 1998.
- [269] N. Godbout, L. K. Sanders, R. Salzmänn, R. H. Havlin, M. Wojdelski, and E. Oldfield. Solid-state NMR, Moessbauer, crystallographic and density functional theory investigation of Fe-O₂ and Fe-O₂-analogue metalloporphyrins and metalloproteins. *J. Am. Chem. Soc.*, 121:3829–3844, 1999.
- [270] R. H. Havlin, N. Godbout, R. Salzmänn, M. Wojdelski, W. Arnold, C. E. Schulz, and E. Oldfield. An experimental and density functional theoretical investigation of iron-57 Moessbauer quadrupole splittings in organometallic and heme-model compounds: Applications to carbonmonoxy-heme protein structure. *J. Am. Chem. Soc.*, 120:3144–3151, 1998.
- [271] J. M. Mouesca, J. L. Chen, L. Noodleman, D. Bashford, and D. A. Case. Density functional Poisson–Boltzmann calculations of redox potentials for iron-sulfur clusters. *J. Am. Chem. Soc.*, 116:11898–11914, 1994.
- [272] T. Soda, Y. Kitawaga, T. Onishi, Y. Takano, Y. Shigeta, H. Nagao, Y. Yoshioka, and K. Yamaguchi. Ab initio computations of effective exchange integrals for H–H, H–He and Mn₂O₂ complex: Comparison of broken-symmetry approaches. *Chem. Phys. Lett.*, 319:223–230, 2000.
- [273] F. Neese. *ORCA – an ab initio, DFT, and semiempirical electronic structure package. Version 2.2, Revision 14*. Max-Planck Institut für Strahlenchemie, Mülheim a. d. Ruhr, Germany, 2002.

- [274] J.-M. Mouesca, L. Noodleman, D. A. Case, and B. Lamotte. Spin densities and spin coupling in iron-sulfur clusters: A new analysis of hyperfine coupling constants. *Inorg. Chem.*, 34:4347–4359, 1995.
- [275] L. Noodleman, J. L. Chen, D. A. Case, C. Giori, G. Rius, J. M. Mouesca, and B. Lamotte. Spin dependent electron delocalization, vibronic and antiferromagnetic couplings in iron-sulfur clusters. In G. N. La Mar, editor, *Nuclear Magnetic Resonance of Paramagnetic Macromolecules*, pages 339–367. Kluwer Academic Publishers, Leiden, 1995.
- [276] K. O. Schäfer, R. Bittl, W. Zwegart, F. Lendzian, G. Haselhorst, T. Weyhermüller, K. Wieghardt, and W. Lubitz. Electronic structure of antiferromagnetically coupled dinuclear manganese ($\text{Mn}^{\text{IV}} \text{Mn}^{\text{III}}$) complexes studied by magnetic resonance techniques. *J. Am. Chem. Soc.*, 120:13104–13120, 1998.
- [277] S. Sinnecker, F. Neese, L. Noodleman, and W. Lubitz. A broken symmetry study on an exchange coupled metal system - EPR parameters of a mixed valence manganese dimer complex from density functional theory. manuscript in preparation.
- [278] J. E. Penner-Hahn, K. S. Eble, T. J. McMurphy, M. Renner, A. L. Balch, J. T. Groves, J. H. Dawson, and K. O. Hodgson. Structural characterization of horseradish-peroxidase using EXAFS spectroscopy – evidence for Fe=O ligation in compound-I and compound-II. *J. Am. Chem. Soc.*, 108:7819–7825, 1986.
- [279] H. Paulsen, M. Müther, M. Grodzicki, A. X. Trautwein, and E. Bill. Complementary Mössbauer, EPR, and theoretical studies of the exchange-coupled porphyrin cation radical system $[(\text{Cl})\text{Fe}(\text{IV})=\text{O}(\text{TMP})]$. *Bull. Soc. Chim. Fr.*, 133:703–710, 1996.
- [280] W. T. Oosterhuis and G. Lang. Magnetic properties of the t_{2g}^4 configuration in low symmetry crystal fields. *J. Chem. Phys.*, 58:4757–4765, 1973.
- [281] F. Neese, J. M. Zaleski, K. E. Loeb-Zaleski, and E. I. Solomon. Electronic structure of activated bleomycin: Oxygen intermediates in heme versus nonheme iron. *J. Am. Chem. Soc.*, 122:11703–11724, 2000.
- [282] E. van Lenthe, A. van der Avoird, W. R. Hagen, and E. J. Reijerse. Density functional calculations of g-tensors of low-spin iron(I) and iron(III) porphyrins. *J. Phys. Chem. A*, 104:2070–2077, 2000.
- [283] K. Kuczera, J. Kuriyan, and M. Karplus. Temperature-dependence of the structure and dynamics of myoglobin. A simulation approach. *J. Mol. Biol.*, 213:351–373, 1990.

Appendix A

Miscellaneous

A.1 List of abbreviations

ABNR	Adopted basis Newton-Raphson algorithm
CASSCF	Complete active space self consistent field
CI	Configuration interaction
Cpd I	Compound I
CPO	Chloroperoxidase
CYP450	Cytochrome P450
DFT	Density functional theory
ENDOR	Electron nuclear double resonance
EPR	Electron paramagnetic resonance
ESEEM	Electron spin-echo envelope modulation
FAD	Flavin adenine dinucleotide, oxidized form
FMN	Flavin mononucleotide
GRMS	Gradient root mean square
HF	Hartree-Fock
HFC	Hyperfine coupling constant
HRP	Horseradish peroxidase
HRP-I	Compound I of horseradish peroxidase
HS	High spin
INDO	Intermediate neglect of differential overlap
KIE	Kinetic isotope effect
LS	Low spin

MC	Monte Carlo
MD	Molecular dynamics
MM	Molecular mechanics
MO	Molecular Orbital
NADH	Nicotinamide adenine dinucleotide, reduced form
NMR	Nuclear magnetic resonance
OZ	Orbital Zeeman
PMF	Potential of mean force
QM	Quantum mechanics
QM/MM	combined quantum mechanics/molecular mechanics
RC	Reactive complex
RMS	Root mean square
SOC	Spin orbit coupling
SOS	Sum over states
SSR	Single state reactivity
TMP	<i>Meso</i> -tetramesityl porphyrin
TS	Transition state
TSR	Two state reactivity
ZFS	Zero field splitting

A.2 Convergence criteria

Table A.1: Convergence criteria for geometry optimizations expressed in atomic units (HDLCOPT: angles in rad).

optimizer	HDLCOPT	NEWOPT
representation of coordinates	hybrid delocalised internal	cartesian
maximum step component	0.0018	0.001
RMS of step vector	0.0012	0.0002
maximum gradient component	0.00045	0.001
RMS of the gradient	0.0003	0.0002

Appendix B

Force Field

The CHARMM22 force field [146] was employed to represent the molecular mechanical part of the system. The force field parameters for the heme group present in the CHARMM22 library refer to a Fe^{II} containing heme [283]. This existing parameter set was used in the present work, however to account for the different electronic situation in Cpd I, the atomic charges on iron, the α -, β -, *meso*-carbons and the pyrrole nitrogens of the protoporphyrin IX, and the sulfur of the coordinating thiolato ligand, were modified with reference to B3LYP/6-31G* derived atomic charges for $[\text{Fe}^{\text{IV}}\text{O}(\text{porphyrin}^+)(\text{SH})]$ (see figure B.1). As detailed in section 3.2, the influence of the modified atomic charges on optimized geometries was carefully checked.

The substrate camphor is not among the predefined residues of the CHARMM22 library. The force field for this compound was derived as follows: (i) Atom types were assigned according to CHARMM conventions which implicitly defines most of the parameters. (ii) Atomic charges were chosen in analogy to CHARMM whenever possible (H, C in CH_2 and CH_3) and were derived from QM calculations otherwise. (iii) Three internal coordinates were represented by special parameters because no corresponding terms were available in the CHARMM force field. It is noted that these terms – two C–C–C angles and one C–C–C–O improper torsion term – could be chosen by minor modifications to existing terms for closely related topological situations. The resulting parameters are listed in table B.1.

The following series of calculations were carried out to validate the new camphor parameter set:

1. MM (CHARMM) geometry optimization in a sphere of water (TIP3P).
2. DFT (B3LYP/6-31G*) geometry optimization *in vacuo*.
3. MM (CHARMM) MD in a water sphere (TIP3P) ¹, 25 ps equilibration (NVT), 100 ps

¹ The spherical “miscellaneous mean field potential” (MMFP) implemented in CHARMM was imposed on the water molecules to constrain the system.

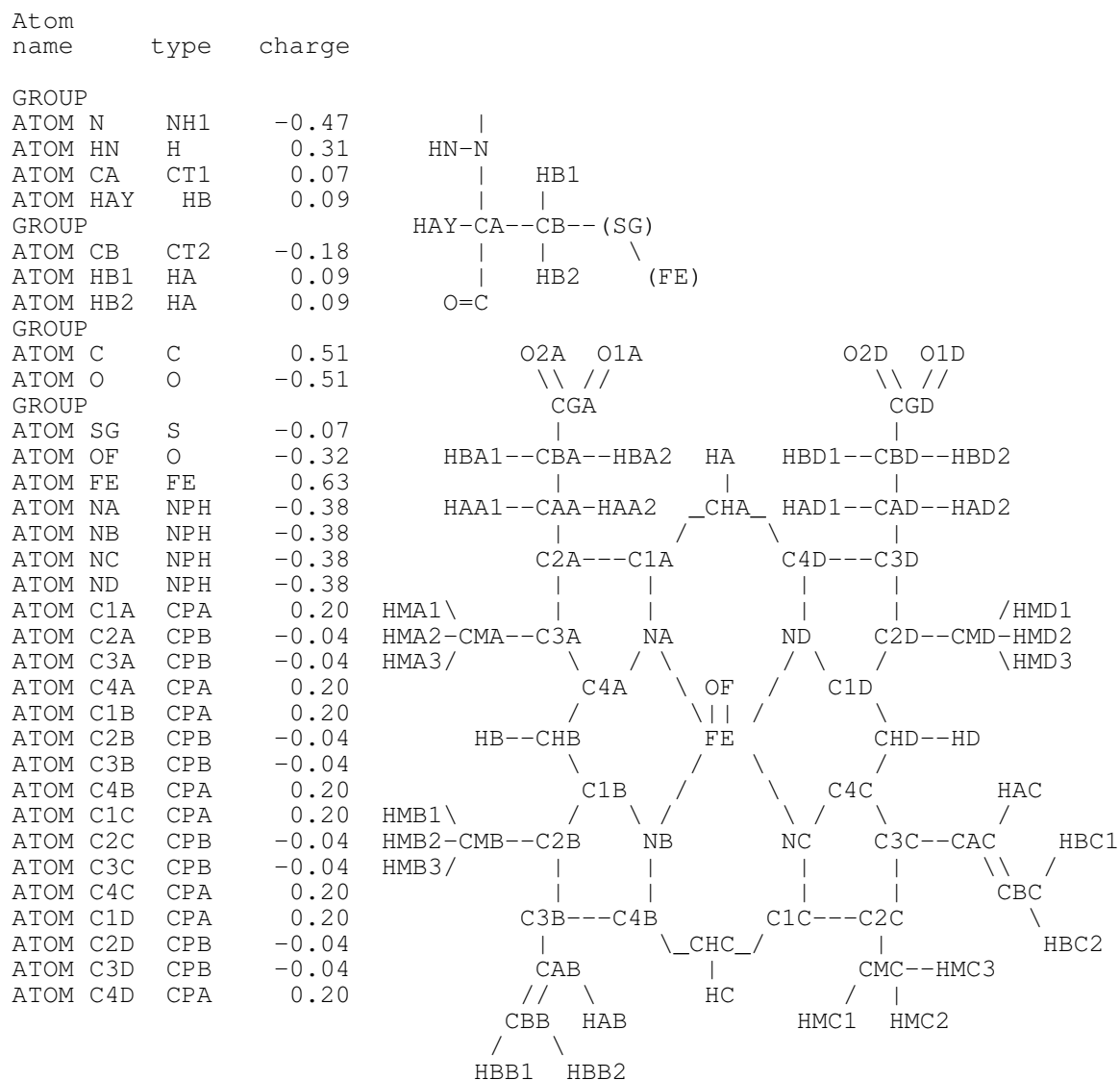
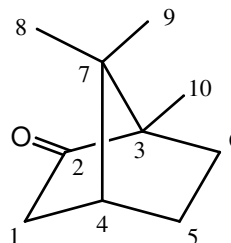


Figure B.1: Force field atom types and charges used for the protoporphyrin IX unit in the pure MM calculations during the setup of the system. All heme parameters were taken from the existing CHARMM22 library except charges on iron, sulfur, oxo-ligand, α -, β -, *meso*-carbons and the pyrrole nitrogen.

Table B.1: Atom types and parameters derived for camphor.

CHARMM22		QM calculated charges		
Name	atom type	charge	HF/3-21G* (Mulliken)	HF/3-21G* (NPA)
O	O	-0.54	-0.576	-0.555
C2	CC	0.64	0.643	0.633
C1	CT2	-0.18	-0.458	-0.559
H1	HA	0.09	0.243	0.262
H2	HA	0.09	0.251	0.266
C3	CT1	-0.07	-0.290	-0.162
C7	CT1	-0.03	-0.131	-0.077
C5	CT2	-0.18	-0.368	-0.462
H5	HA	0.09	0.214	0.240
H6	HA	0.09	0.209	0.238
C6	CT2	-0.18	-0.353	-0.454
H7	HA	0.09	0.222	0.248
H8	HA	0.09	0.221	0.245
C4	CT1	-0.09	-0.280	-0.249
H4	HA	0.09	0.211	0.254
C8	CT3	-0.27	-0.530	-0.660
H9	HA	0.09	0.207	0.232
H10	HA	0.09	0.204	0.234
H11	HA	0.09	0.199	0.232
C9	CT3	-0.27	-0.523	-0.652
H12	HA	0.09	0.204	0.229
H13	HA	0.09	0.197	0.231
H14	HA	0.09	0.200	0.231
C10	CT3	-0.27	-0.530	-0.653
H15	HA	0.09	0.208	0.237
H16	HA	0.09	0.215	0.244
H17	HA	0.09	0.192	0.226



Special parameters:

- (a) angle C3–C2–C1:
 $k = 40 \text{ kcal mol}^{-1} \text{ rad}^{-2}$
 $\theta_e = 118^\circ$
- (b) angle C1–C4–C5:
 $k = 53.35 \text{ kcal mol}^{-1} \text{ rad}^{-2}$
 $\theta_e = 112^\circ$
- (c) improper torsion C2–C1–C3–O:
 $k = 45 \text{ kcal mol}^{-1} \text{ rad}^{-2}$
 $\phi_e = 0^\circ$

production run (NVE).

4. QM/MM (AM1/CHARMM) MD in water sphere, 25 ps equilibration (NVT), 100 ps production run (NVE); AM1 for camphor, force field for surrounding water.
5. MM (CHARMM) MD of the entire Cpd I system (enzyme, substrate, solvent), 200 ps equilibration dynamics (see section 3.2).

The analysis of the resulting geometrical data (see table B.2) shows that the force field parameters yield structures in very good agreement with the theoretical reference data. The improper torsion term C2-C1-C3-O is apparently slightly biased towards positive values in the case of the MD run in a water sphere (run 3, table B.2) however there is no such behavior in the simulation of the full system (run 5, table B.2).

Table B.2: Validation runs: angles (degree) and improper torsion angle (degree) in camphor.

Angle ^a	Run ^b	Value ^c	MD results		
			fluctuation	minimum	maximum
C3-C2-C1	1	106.5	-	-	-
	2	107.6	-	-	-
	3	106.9	2.8	99.4	114.6
	4	106.1	2.3	101.1	112.1
	5	106.9	2.7	99.1	114.8
C1-C4-C5	1	105.6	-	-	-
	2	105.8	-	-	-
	3	105.8	2.8	100.5	113.0
	4	105.4	3.6	93.1	116.3
	5	105.9	3.0	99.3	113.9
C2-C1-C3-O	1	2.3	-	-	-
	2	-0.4	-	-	-
	3	5.5	3.7	-3.0	13.6
	4	-0.3	3.5	-8.9	8.6
	5	-0.4	5.1	-17.6	11.5

^a See table B.1. ^b See text. ^c Optimized values for runs 1-2, MD average values for runs 3-5.

ERKLÄRUNG

Hiermit erkläre ich, Jan Schöneboom, geboren am 12.02.1975 in Waibstadt, dass ich diese Arbeit eigenständig verfasst und ohne unerlaubte Hilfe angefertigt und diese in der vorgelegten oder ähnlichen Form noch keiner anderen Institution eingereicht habe.

Mülheim an der Ruhr, den 30.06.2003

.....