Insights into Biomolecular Structure from Theoretical Spectroscopy

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Christoph Riplinger

aus

Lebach

Bonn 2011

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der
Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter: Prof. Dr. Frank Neese

2. Gutachter: Prof. Dr. Thomas Bredow

Tag der Promotion: 19. Oktober 2011

Erscheinungsjahr: 2011

IN DER DISSERTATION EINGEBUNDEN:

Zusammenfassung

Lebenslauf

Abstract

The first part of the present PhD thesis is concerned with the reaction mechanism of Cytochrome P450 NO reductase. Two important aspects of this enzyme are studied. 1) The structural and electronic nature of the key intermediate *I* and 2) the unusual direct reduction of the active site by NADH.
In the computational study the inclusion of the entire enzyme structure is important. For that purpose, an interface for quantum mechanics/molecular mechanics (QM/MM) calculations with the program packages ORCA and GROMACS is implemented. A QM/MM study is carried out using density functional theory (DFT) to describe the protein's active site at quantum mechanical level (44–142 atoms) and a classical molecular mechanics force field to describe the remaining parts of the system (about 50000 atoms).
Calculations on two experimentally well characterized intermediates of the reaction cycle are performed. Structural properties, Mössbauer parameters and IR calculations are in good agreement with the experimental data.
The third intermediate of the reaction, the key intermediate *I*, can be observed only under limited NO concentration and is thus experimentally not well characterized. The structural and spectroscopic properties of six possible candidates for *I* are analyzed in detail. The calculated properties exclude four candidates, leaving two for additional analysis: the singlet diradicals Fe(III)-$NHO^{\bullet -}$ and Fe(III)-$NHOH^{\bullet}$.
In further QM/MM calculations two different reaction pathways, connecting all possible intermediates, are modeled. In order to perform these calculations the geometry optimization methodology in ORCA is extended by algorithms for constrained and transition state (TS) geometry optimization. Additionally, several approaches to exact Hessian matrices that are used in TS optimization are developed and implemented.
The calculated kinetic and thermodynamic properties for both pathways are used for the kinetic simulation of the entire reaction course. The kinetic simulation helps to unravel the nature of intermediate *I*. It is identified as the singlet diradical species Fe(III)-$NHOH^{\bullet}$.
For the QM/MM calculations of the reduction of heme-bound nitrosyl by the cofactor NADH the kinetic isotope effect as well as the calculated activation barrier are in agreement with the experimental data. The results reveal that the reduction is carried out in form of a hydride ion transfer. Additionally, they give insight into the factors that determine NADH binding in the active site and explain the loss of catalytic activity observed in site-directed mutagenesis experiments.
The calculations on the formation of the N-N bond show that the mechanism behind the bond formation is spin-recoupling, which is only possible due to the diradical character of intermediate *I*.
In a second part of the thesis a more general question is adressed. Namely, the validity of the point-dipole approximation, which is used in the EPR spectroscopy applications PELDOR (pulsed electron-electron double resonance) and DQC (double quantum coherence). Both techniques are used to assign distances in proteins and other macromolecules. The magnetic dipolar coupling between spin centers in macromolecules is measured experimentally in order to determine their distance and thus get important insight into the conformation of the entire molecule. For distance determination, a simple point-dipole approximation is frequently used. In the present thesis, DFT studies on a variety of model systems are performed to study the limitations of this approximation.
The calculations show that delocalization of the spin density can lead to large errors in the obtained interspin distances. Even small delocalization leads to errors in the range of 1-2 Å.

The deviations become dramatic for systems with large delocalization of the spin density, e.g. unsaturated linkers between the spin centers or aromatic nitroxide systems. For aromatic nitroxide systems there can even be a dependence on the orientation of the spin-carrying fragments.

Zusammenfassung

Der erste Teil der vorliegenden Arbeit beschäftigt sich mit theoretischen Untersuchungen des Reaktionsmechanismus der Cytochrom P450 NO Reduktase und konzentriert sich hierbei auf zwei wichtige Aspekte des Enzyms: 1) Den Aufbau und die elektronischen Eigenschaften des Schlüsselintermediates *I* der Reaktion und 2) die Reduktion des Substrats durch NADH, die ungewöhnlicherweise direkt im aktiven Zentrum des Enzyms stattfindet.
Für die Studie ist die Berücksichtigung der gesamten Enzymstruktur von Bedeutung. Zu diesem Zweck wird eine Schnittstelle für Quantenmechanik/Molekularmechanik (QM/MM) Rechnungen mit den Programmpaketen ORCA und GROMACS implementiert. Es wird hiermit eine QM/MM Studie durchgeführt, in welcher das Aktivzentrum mit Dichtefunktionaltheorie (DFT) auf quantenmechanischem Niveau (44–142 Atome) und das restliche System (ca. 50000 Atome) mit einem klassischen Molekularmechanik Kraftfeld beschrieben wird.
Die enzymatisch katalysierte Reaktion verläuft über drei Reaktionsintermediate. Zunächst werden Berechungen an den beiden experimentell gut charakterisierten Reaktionsintermediaten durchgeführt. Die Ergebnisse stimmen gut mit den entsprechenden experimentellen Strukturparametern und spektroskopischen Daten von Mössbauer- und IR-Spektroskopie überein.
Das Schlüsselintermediat *I* der Reaktion kann nur bei niedrigen NO-Konzentrationen beobachtet werden und ist daher experimentell schlecht zugänglich. Sechs mögliche Kandidaten für Intermediat *I* werden auf ihre strukturellen und spektroskopischen Eigenschaften hin untersucht. Vier der Kandidaten können durch Vergleich der berechneten mit den entsprechenden experimentellen Daten ausgeschlossen werden. Die beiden übrigen Kandidaten, die Singulett Diradikale Fe(III)-$NHO^{\bullet -}$ und Fe(III)-$NHOH^{\bullet}$, werden für weitere Untersuchungen ausgewählt.
Im weiteren Teil der Arbeit werden zwei verschiedene Reaktionswege, die alle selektierten Intermediate miteinander verbinden, modelliert. Um diese Rechnungen durchzuführen, werden zunächst Methoden zur Optimierung unter Zwangsbedingungen und zur Übergangszustandsbestimmung in ORCA implementiert. Darüber hinaus werden verschiedene Verfahren zur Bestimmung von exakten und genäherten zweiten Ableitungen entwickelt, die zur Optimierung von Übergangszuständen notwendig sind.
Mit den berechneten Parametern der beiden Reaktionswege wird eine kinetische Simulation des Reaktionsverlaufs durchgeführt. Diese gibt Aufschluss über die Struktur des Schlüsselintermediates. Intermediat *I* kann als Singulett Diradikal Fe(III)-$NHOH^{\bullet}$ identifiziert werden.
Die QM/MM Ergebnisse eines zentralen Schrittes der Reaktion, der Reduktion des Hämgebundenen Nitrosyls durch den Kofaktor NADH, stimmen sowohl mit dem gemessenen kinetischen Isotopeneffekt, als auch mit der Aktivierungsenthalpie gut überein. Die Rechnungen zeigen, dass der Reduktionsschritt in Form eines Hydridionentransfers abläuft. Darüber hinaus tragen die Ergebnisse zu einem besseren Verständnis der NADH-Bindung im aktiven Zentrum bei und können die experimentellen Resultate von ortsspezifischen Mutationsexperimenten, die teils drastischen Aktivitätsverlust aufweisen, erklären.
Für die weiteren Teile der Reaktion zeigen die QM/MM-Berechnungen, dass die Knüpfung der N-N-Bindung über den Mechanismus der Spin-Paarung abläuft. Dies wird erst durch den Diradikal-Charakter des Intermediates ermöglicht.
Der letzte Teil der Arbeit befasst sich mit der Gültigkeit einer zentralen Näherung, die in der Elektronenspinresonanz (EPR)-Spektroskopie verwendet wird. Die EPR-Methoden PELDOR (gepulste Elektron-Elektron Doppelresonanz) und DQC (Doppelquantenkohärenz) werden verwendet, um Abstände in Proteinen und anderen Makromolekülen zu bestimmen. Hierbei wird die magnetische Dipol-Kopplung zwischen Spinzentren in den Makromolekülen gemessen. Aus der Grösse der Dipol-Kopplung wird der Abstand zwischen den Spinzentren

berechnet, was Aufschluss über die Konformation des Moleküls gibt. Für die Berechnung des Abstands aus der gemessenen magnetischen Dipol-Kopplung wird eine einfache Punkt-Dipol-Näherung verwendet. In der vorliegenden Arbeit werden DFT-Rechnungen an verschiedenen Modellsystemen durchgeführt, um die Grenzen der Punkt-Dipol-Näherung zu untersuchen. Die Berechnungen zeigen, dass eine Delokalisierung der Spindichte zu grossen Abweichungen in den erhaltenen Spin-Spin Abständen führen kann. Selbst kleine Delokalisierungen können Fehler in der Grössenordnung von 1–2 Å verursachen. Die Fehler werden noch weitaus grösser bei Systemen mit stark delokalisierter Spindichte, wie z.B bei aromatischen Nitroxidsystemen oder ungesättigten Verbindungsmolekülen zwischen den Spinzentren. Für aromatische Nitroxidsysteme kann sogar die Orientierung der Spinzentren die Abstände verfälschen.

List of Publications

Publications related to the thesis

- ***Chapter 3***

  C. Riplinger, E. Bill, A. Daiber, V. Ullrich, H. Shoun and F. Neese. Reaction Intermediates in Cytochrome P450 NO Reductase: a Combined Spectroscopic and QM/MM Study. To be submitted to *J. Am. Chem. Soc.*
  In this work the intermediates of the reaction catalyzed by P450 NO reductase were studied in a combined experimental (applying a variety of spectroscopic techniques, e.g. Mössbauer and EPR spectroscopy) and computational (QM/MM calculation of the intermediate structures in the solvated enzyme and their electronic and spectroscopic properties) approach. My contribution to this project was the entire computational study.

- ***Chapter 4***

  C. Riplinger and F. Neese. The reaction mechanism of Cytochrome P450 NO Reductase: A Detailed Quantum Mechanics/Molecular Mechanics Study. *ChemPhysChem,* accpeted, 2011.
  In this work the entire reaction mechanism of P450 NO reductase was studied computationally using QM/MM calculations of the solvated enzymatic system. I carried out the entire study.

- ***Chapter 5***

  C. Riplinger, J.P.Y. Kao, G.M. Rosen, V. Kathirvelu, G.R. Eaton, S.S. Eaton, A. Kutateladze and F. Neese. Interaction of radical pairs through-bond and through-space: Scope and limitations of the point-dipole approximation in electron paramagnetic resonance spectroscopy. *J. Am. Chem. Soc.,* 131 (29), 10092–10106, 2009.
  In this work the validity of the point-dipole approximation, which is used in EPR spectroscopy to evaluate interspin distances from zero-field splitting (ZFS) parameters, was analyzed experimentally (X-ray crystal structure analysis and EPR measurements) and computationally (DFT calculations of the ZFS parameters for various model systems). I carried out the entire quantum chemical study.

Publications by the author not related to the thesis

1. Q. Scheifele, *C. Riplinger*, F. Neese, H. Weihe, A.L. Barra, F. Juranyi, A. Podlesnyak and P.L.W. Tregenna-Piggott. Spectroscopic and theoretical study of a mononuclear manganese (III) complex exhibiting a tetragonally compressed geometry. *Inorg. Chem.,* 47 (2), 439–447, 2008.
   In this work experimental (inelastic neutron scattering and high-field electron paramagnetic resonance (EPR) measurements) and computational (ligand field, density

functional theory (DFT) and complete active space self-consistent field ab initio calculations) approaches were combined to study the electronic and spectroscopic properties of a tetragonally compressed Manganese(III)-complex. My contribution to this work was the development and implementation of the geometry optimization methodology that was used to provide suitable molecular structures for the calculation of electronic and spectroscopic properties.

2. E.L. Klein, A.V. Astashkin, D. Ganyushin, *C. Riplinger*, K. Johnson-Winters, F. Neese and J.H. Enemark. Direct detection and characterization of chloride in the active site of the low-ph form of sulfite oxidase using electron spin echo envelope modulation spectroscopy, isotopic labeling, and density functional theory calculations. *Inorg. Chem.,* 48 (11), 4743–4752, 2009.
   In this work the position of chloride in the active site of sulfite oxidase was analyzed in a combined experimental (electron spin echo envelope modulation spectroscopy of the enzyme with different isotope labeling) and computational (DFT calculations of EPR parameters for different active site cluster models) study. I was concerned with the development and implementation of the geometry optimization methodology for constrained cluster calculations and with the optimization of active site models.

3. M. Sundararajan, *C. Riplinger*, M. Orio, F. Wennmohs, and F. Neese. Spectroscopic properties of protein-bound cofactors: Calculation by combined quantum mechanical/ molecular mechanical (QM/MM) approaches. *Encyc. Inorg. Chem.,* 2009.
   In this work the calculation of spectroscopic properties of enzymatic systems using the quantum mechanics/molecular mechanics (QM/MM) approach was reviewed. In this project I was involved with the implementation of the QM/MM interface of the quantum mechanical program package ORCA to the molecular mechanics program package GROMACS.

4. M. Radoul, M. Sundararajan, A. Potapov, *C. Riplinger*, F. Neese and D. Goldfarb. Revisiting the nitrosyl complex of myoglobin by high-field pulse epr spectroscopy and quantum mechanical calculations. *Phys. Chem. Chem. Phys.,* 12 (26), 7276–7289, 2010.
   In this work NO binding to reduced myoglobine was analyzed using a combined experimental (high field pulsed EPR spectroscopy) and computational (QM/MM calculation of EPR parameters) approach. My contribution to this work was the implementation of the geometry optimization methodology which was used to perform the QM/MM calculations.

5. T. Argirević, *C. Riplinger*, J. Stubbe, F. Neese and M. Bennati. A combined high-field ENDOR/DFT approach for proton-coupled electron transfer investigation in E.coli ribonucleotide reductase. To be submitted to *P. Natl. Acad. Sci. USA*.
   In this work the hydrogen bonding network in ribonucleotide reductase was studied with a combined experimental (high-field electron-nuclear double resonance spectroscopy) and computational (DFT calculations of EPR parameters of enzyme cluster models) approach. In this project I carried out all quantum chemical calculations.

6. S. Ye, *C. Riplinger*, A. Hansen, C. Krebs, J.M. Bollinger and F. Neese. Electronic Structure Analysis of the Oxygen Activation Mechanism by the Fe(II)- and α-Ketoglutarate-

Dependent Dioxygenases. To be submitted to *Chem. Eur. J.*
In this work the mechanism of α-ketoglutarate dependent dioxygenase was studied computationally using DFT and high-level ab initio calculations and using an active site cluster model. I was involved with the development and implementation of the geometry optimization methodology used in this work and with the calculation and analysis of the kinetic simulation.

Acknowledgements

It is a pleasure to thank those who contributed to my PhD thesis over the last four years. There are various academics from whom I had the opportunity to learn. They all guided me to some extent during my PhD period:

First and foremost, I want to express my deep gratitude to my supervisor Prof. Neese, who introduced me with patience into the fields of bioinorganic chemistry, theoretical chemistry and theoretical spectroscopy and who taught me his perspective on many topics related to electronic structure and beyond. He gave me the opportunity to work on a variety of projects. It was especially interesting to witness the growth of our research group first-hand.

I want to thank Profs. Sandra and Gareth Eaton from Denver who were very patient in sharing with me their huge knowledge on EPR spectroscopy. I also had the pleasure to get an excellent guided tour through their large collection of EPR spectrometers. I would like to thank Prof. Hai Lin, also from Denver, who introduced me to the field of QM/MM. My thanks go to Dr. Gerrit Groenhof from Göttingen, who advised me in setting up the interface between ORCA and GROMACS. I had a nice and pleasant visit in Göttingen and had a lot of fun getting to know his young family. My sincere thanks go to Jun.-Prof. Johannes Kästner from Stuttgart, who supported me several times by giving advice on how one should actually set up a QM/MM calculation and on 'defining' a way through the jungle. I also wish to thank Prof. Max Holthausen, with whom I had some inspiring discussions. He helped me in finding some new ideas on Hessian-related topics.

Many other people in the institute in Bonn contributed to a pleasant atmosphere over the last years. My warm thanks go to Claudia Kronz, for running the institute so smoothly. I had a really great time sharing an office with Christoph (Reimann) in the first years. Christoph, Simone, Kanthen and Barbara were joyful company in general, or in our small QC seminar. I wish to thank Shengfa, FrankW, Maylis, Taras, Robert, and many others for interesting and helpful discussions and expertise.

Vera, Robert, Barbara, Kanthen, Dima, Rebecca, Martin, Christian and Noel diligently proof-read parts of the thesis. Simone and Christoph provided the LaTeX template and their expertise. Ingo improved the layout last-minute. I am grateful for all your help in making this document reasonably comprehensible and visually bearable.

My warm thanks go to my family and friends who supported me during that time. Christian and Renate, thanks for being great flatmates in the first three years of the thesis. Martin, it was a lot of fun sharing travel stories and going on some exciting trips. Thanks to Panthea, Thomas and the Ska band for unforgettable parties in the tower. Christian, thanks for endless discussions and our conquest of the Alps. Thanks to all from the Bonn university rowing team who made me get up early and were nice company on the Rhine and in the boathouse. My very special thanks go to Barbara, who has been an inspiring and cheerful companion.

Finally I want to thank the SFB624 for financial support.

Contents

List of Tables

List of Figures

Nomenclature

$\pi^{\star}_{ip}$	$\pi^{\star}$ in-plane
$\pi^{\star}_{oop}$	$\pi^{\star}$ out-of-plane
Arg	Arginine
Asp	Aspartate
BS	Broken Symmetry
cg	Conjugate Gradient
DEER	Double Electron-Electron Resonance
DNA	Deoxyribonucleic Acid
DQC	Double Quantum Coherence
EE	Electrostatic Embedding
EF	Eigenvector Following
EPR	Electron Paramagnetic Resonance
EXAFS	Extended X-Ray Absorption Fine Structure
Glu	Glutamate
KIE	Kinetic Isotope Effect
LA	Link Atom
LJ	Lennard-Jones
Lys	Lysine
MCD	Magnetic Circular Dichroism
ME	Mechanical Embedding
MM	Molecular Dynamics
MM	Molecular Mechanics
MO	Molecular Orbital
NAAD	Nicotinic Acid Adenine Dinucleotide
NADH/NAD^{+}	Nicotinamide Adenine Dinucleotide (reduced/oxidized form)
NMR	Nuclear Magnetic Resonance
NPA	Natural Population Analysis
NR	Newton-Raphson
P450nor	P450 NO Reductase

PBC	Periodic Boundary Condition
PE	Polarized Embedding
PELDOR	Pulsed Electron-Electron Double Resonance
PES	Potential Energy Surface
PHVA	Partial Hessian Vibrational Analysis
QM	Quantum Mechanics
QM/MM	Quantum Mechanics/Molecular Mechanics
QN	Quasi-Newton
RFO	Rational Function Optimization
rR	Resonance Raman
sd	Steepest Descent
Ser	Serine
SOC	Spin-Orbit Coupling
SOMO	Singly Occupied Molecular Orbital
SS	Spin-Spin
Thr	Threonine
TS	Transition State
UV	Ultraviolet
VDW	Van Der Waals
ZFS	Zero-Field Splitting
ZPE	Zero Point Energy

1. Introduction

Proteins are biological macromolecules with a large diversity in structure and functionality. Together with nucleic acids they form the chemical basis of life. Enzymes are proteins that catalyze chemical reactions.[1] All important processes in biological cells are catalyzed by enzymes, which accelerate reactions by several orders of magnitude. The most prominent example in terms of rate acceleration is urease, an enzyme that speeds up the hydrolysis of urea by a factor of $\approx 10^{14}$, compared to the reaction in aqueous solution.[2]
A large number of enzymes contain metal ions. These metalloenzymes are often involved in processes where reduction or oxidation takes place.[3] Photosystem II, the protein complex which is of central importance for life on earth, converting carbon dioxide to dioxygen with the aid of solar energy, is probably the most important example of a multi-site metalloenzyme.[4]
Understanding how enzymes work on a mechanistic, structural and electronic level is of great importance in contemporary biochemical research. For this purpose it is necessary to gain knowledge about the intermediates of a reaction. Many intermediates in reactions that are catalyzed by metalloenzymes have very short life times. Spectroscopy is often the only source for structural and electronic analysis.[5,6]
In the last decades a variety of experimental and computational techniques, which are used to get insight into protein function, have been developed. An overview of these techniques is given in Section 1.1.1.
The QM/MM (quantum mechanics/molecular mechanics) method, which is introduced in Section 1.1.2 and in more detail in Section 2.2, was used to study the reaction mechanism of the enzyme Cytochrome P450 NO reductase (P450nor). Information on the state of knowledge on P450nor is given in Section 1.2. The results of the study on P450nor obtained in this work are presented in Chapters 3 and 4.
Electron paramagnetic resonance (EPR) spectroscopy is used in biochemical research to gain structural information about proteins.[7,8] The method is introduced in Section 1.1.3, while the theoretical background is given in Section 2.3. The validity of approximations that are used for distance determination in proteins were analyzed as part of this thesis. The results are presented in Chapter 5.

1.1. Studying Enzymatic Reactions

1.1.1. Overview of Experimental and Computational Approaches

The techniques to analyze protein function are manifold. Since the mid-20th century the development in biochemical methods has advanced rapidly. Large progress in *molecular biology* made it possible to study the composition of proteins and to manipulate their primary sequence different from the native one.[1,9] The determination of three-dimensional protein structures started in 1958 with the first *X-ray structure* of myoglobine[10] and made it possible to *link protein function and protein structure*. Since that time, about 66000 protein structures (as of early 2011) have been deposited in the protein data bank (PDB, a repository for the three-dimensional structural data of large proteins and nucleic acids).[11,12] From the late

1980's on, nuclear magnetic resonance (NMR) has also been applied in three-dimensional structure determinations of proteins.[13] However, the vast majority of protein structures is obtained using X-ray analysis.[a] There also exist a plethora of other spectroscopic techniques, which do not reveal the entire protein structure, but are particularly useful for analyzing the *immediate environment of metals in metalloenzymes*: For example, electron paramagnetic resonance[14] (EPR), resonance Raman[15–17] (rR), magnetic circular dichroism[15] (MCD) and Mössbauer spectroscopy.[18,19] Applying such methods is often the only possibility to gain information on short-lived intermediates.

However, knowledge of the protein structure does not explain the entire functionality of the protein. There is no understanding of an enzyme's function without knowledge about its reaction mechanism. The intermediates of a whole enzymatic catalytic cycle can nowadays be examined by applying a variety of biochemical techniques. Catalytic residues can be identified by *site-directed mutagenesis* or chemical modification.[20] Reaction rate constants for different reaction steps can be determined with a variety of kinetic experiments (e.g. stopped flow or quenched flow techniques for short-lived intermediates).[21–24]

From the computational chemistry point of view, many methods have been developed over the last decades, which help in understanding protein function at an atomistic or even electronic level. Classical mechanics (also called molecular mechanics, MM) is applied to study the *dynamics of proteins* in aqueous solution using *molecular dynamics (MD) techniques*.[25,26] Herewith, the importance of specific intra- and intermolecular interactions and their role in enzymatic activity can be revealed. In order to study catalytic pathways, classical mechanics is not sufficient, and more accurate methods are necessary.

With *quantum chemical* (also called quantum mechanical, QM) methods all intermediates (also transient ones) and transition states can be modeled. Quantum chemical calculations provide relative energies and reaction barriers at chemical accuracy (depending on the method applied).[26,27] Not only structural, but also electronic changes that are present in the course of a reaction can be investigated. The problem in applying QM methods is their scaling behavior. Depending on the level of theory, the formal scaling is of the order of $O(N^3)$ to $O(N^7)$ (where N is the number of basis functions, which is proportional to the system size).[28] With these methods the systems that can be studied may contain 1–10 (for highly correlated methods) up to 200-300 atoms (for density functional methods). These methods are thus much too slow for the modeling of an enzyme in an explicit and realistic solvent.

Instead of modeling the entire protein, small truncated systems, representing the active site of the protein, are used in *cluster calculations*.[29] Here, the electrostatic effect of the surrounding protein is usually represented by a continuum solvation model. This approach mimics the protein's influence in an isotropic manner, but neglects specific interactions of the truncated active site model with the directly adjacent protein residues. Nevertheless, many useful results can be obtained with this approach, not necessarily quantitatively, but qualitatively.

Molecular mechanics methods with empirically parametrized force fields show a much better scaling behavior than *QM methods*. They formally scale as $O(N^2)$ and are therefore suitable for an atomistic simulation of systems containing several millions of atoms.[4] The important drawback of these methods, with respect to the simulation of enzyme function, is that they are not able to describe chemical processes which are based on electronic rearrangements, such as bond breaking and bond formation, charge transfer and electronic excitations.[28] Es-

[a] In early 2011, 88% of all protein structures in the protein data bank were revealed with X-ray analysis, whereas only 12% were revealed with NMR (early 2011 from PDB statistics[12]).

pecially metalloenzymes are hard to handle within the *MM approach*. The interactions in MM methods have to be parametrized, and parameters for the interaction between ligands and each individual metal type, which plays a role in biology, have to be provided.[30] Unfortunately these parameters are usually not available.
A combination of both techniques, *QM and MM methods*, led to the development of the nowadays widely spread *QM/MM* approach. Its principles are presented in the following section.

1.1.2. QM/MM Approach

The idea of combining QM and MM methods in order to combine their advantages was introduced in 1976 by Warshel and Levitt.[31] Their ideas follow the basic concept of enzymology, namely, that an enzymatic reaction is effectively catalyzed in a very small part of the protein, by only few functional groups (residues, cofactors and metals) in the *active site*. The other, much larger parts of the protein are not necessarily important for the actual enzymatic reaction. They instead can play a role in binding processes to other proteins, membranes and second messengers and in the stability of the overall protein structure (which again ensures the stability of the active site). Following this concept, the whole enzymatic system (protein plus cofactors plus explicit solvent) is divided into *different subsystems* in the QM/MM approach. The smaller subsystem, which is important for the catalytic reaction, is treated by an appropriate QM method. The remaining parts of the protein are approximated by a classical MM force field. This approach can give much insight into enzymatically catalyzed reactions. Instead of simulating reactions in gas phase, the whole protein structure, including its steric and electrostatic influence, can be taken into account. Thus, the influence of the enzyme on reactants and especially transition states can be studied.[32] Furthermore, point mutations can be introduced *in silico*, which allows to test the influence of a mutation on enzymatic reactivity.[33]
The QM/MM approach results in a potential energy function which is applicable on a wide range of systems. It is used frequently for systems with 10000 to 40000 atoms, where the smaller QM subsystem has typically 50–100 atoms.[34] The concept was initially developed for proteins, but it can also be used in other areas of application. Studies of DNA, reactions in an explicit solvent, reactions on the surface of solid states, etc. have been performed in the past.[35–37]
The QM/MM approach, its basic schemes, equations and a workflow for a standard QM/MM calculation are described extensively in Section 2.2. In Chapters 3 and 4 the QM/MM approach is used to study the reaction catalyzed by Cytochrome P450 NO reductase.

1.1.3. Distance Measurement

Electron paramagnetic resonance spectroscopy can be used to study the structural properties of proteins or protein complexes. Pulsed electron-electron double resonance (PELDOR) and double quantum coherence (DQC) are two important and rapidly developing applications of EPR. With PELDOR or DQC, the magnetic dipolar coupling between two unpaired electrons can be used to *precisely assign distances* between spin centers of up to 60 or 70 Å.[7,8,38–41] The spin centers can either be attached to the protein[b] (*spin labels*), or they are already part of it (e.g. metal ions, transient or stable amino acid radicals, cofactor radicals, etc.).

[b] The method can of course be used with any other macromolecular system, either biomolecular or synthetic, provided that the spin centers are present.

Spin labels are frequently nitroxyl radicals, and data analysis for distance determination is typically based on a *point-dipole approximation*, in which it is assumed that the unpaired electron is localized on the nitrogen-oxygen moiety. However, it is also known that, at shorter distances or in delocalized systems, the use of the point-dipole approximation is inadequate, and the relation between the zero-field splitting observed in EPR spectra to intercenter distances is more complicated.[42]

If one wants to deduce distances with EPR with high accuracy, it is necessary to evaluate the *limitations* of the widely used point-dipole approximation. For this purpose, quantum chemical calculations are invaluable. Fortunately, concomitant with the progress in experimental EPR spectroscopy, there has been considerable progress in the quantum chemical calculation of EPR parameters.[43] Reasonably accurate values can now be calculated for organic radicals based on density functional theory (DFT), for hyperfine couplings,[44–48] g-tensors,[45] quadrupole couplings,[45,49] and more recently also for the two parts of the zero-field splitting.[50–55] It is particularly the zero-field splitting that is of importance for distance measurements. The biquadratic zero-field splitting (ZFS) D-tensor contains a first order contribution from the direct magnetic dipole-magnetic dipole spin-spin (SS) interaction, and a second-order contribution from the spin-orbit coupling (SOC). While the general equations have been known for a long time,[56,57] their realization in a DFT framework has been less than obvious and it is only recently that general, efficient programs became available that allow the full treatment of the ZFS in organic[51,52,54] as well as inorganic[50,53] systems.

In this thesis, the quantum chemical theory of the spin-spin ZFS and important approximations to it (amongst them the point-dipole approximation) are presented in Section 2.3. In Chapter 5, the limitations of the point-dipole approximation are analyzed using a dinitroxyl model system. The origin of these limitations and their implications for distance measurement in larger biomolecules are discussed in detail.

1.2. Cytochrome P450nor

The major part of this thesis is devoted to the study of the reaction mechanism of the enzyme Cytochrome P450 NO reductase (P450nor). In this section, different aspects of P450nor are introduced: its biological context, structural characteristics and the current knowledge about its reaction mechanism. In Chapter 3 and 4 the intermediates and the mechanism of the reaction catalyzed by P450nor are investigated with the help of QM/MM results.

1.2.1. Biological Meaning

Nitrate Reduction

Catalyzing the reduction of NO, P450nor plays an important role in fungal dissimilatory denitrification (see Figure 1.1).[58] Here, the fungus uses nitrate reduction as additional energy source for *anaerobic respiration*. The stepwise reduction of nitrate to dinitrogen is catalyzed by different enzymes (as illustrated in Figure 1.1).

Nitric oxide is an intermediate of nitrate reduction. A prolonged increase in NO concentra-

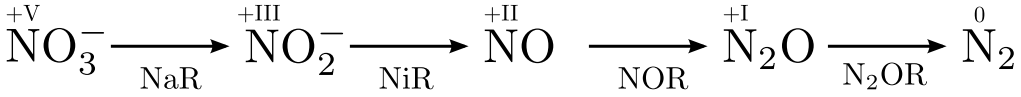


Figure 1.1. – Denitrification pathway for the reduction of nitrate to dinitrogen. The catalyzing enzymes are indicated below the arrows (NaR: Nitrate Reductase, NiR: Nitrite Reductase, NOR: NO Reductase, N_2OR: N_2O Reductase).[58] Roman numerals indicate the oxidation state of nitrogen.

tion is toxic.[c,61–63] Thus, nitric oxide, which is produced in high amounts in denitrification, has to be converted fast and efficiently.

Cytochrome P450 Superfamily

P450 NO reductase is a member of the Cytochrome P450 superfamily. The P450 superfamily obtained its name from the unusual absorption band at 450 nm of the carbon monoxide bound enzyme.[64] The first member of the superfamily was discovered more than 50 years ago.[65] Since that time, several thousands of distinct P450 members were found in all different branches of life. They form a diverse group, not only in terms of their biological function, but also in terms of *substrate selectivity*, and they are involved in a *large number of different reactions*.[66] Among these reactions are essential processes like the detoxification of xenobiotics, drug metabolism, biosynthesis of hormones, bioregulatory functions, etc. In most of these reactions, P450 enzymes are acting as *monooxygenases*, inserting one oxygen atom from dioxygen into nonpolar groups (e.g. aliphatic C-H groups).[5,66] Without a catalyst, such a reaction needs very high temperature and proceeds highly unselectively.[67] This makes P450 enzymes exciting catalysts. Much scientific interest arose in the last decades, not only out of pure interest in understanding their mechanistic details (as e.g. $\approx 70\%$ of the drugs taken orally are metabolized by P450 enzymes[5]), but also due to the potential transfer of knowledge to applications in catalytic chemistry.

[c] NO is poisonous, although it can also act as a second messenger. As a signaling molecule, it occurs at much lower concentration levels and is synthesized by different enzymes in different cellular compartments.[59–62]

P450 NO Reductase

P450NO reductase was first described in 1991 as an enzyme from *Fusarium oxysporum*, a fungus which lives in soils and can live under aerobic as well as anaerobic conditions.[68] P450nor is a 46 kD monomeric protein with a high molar activity of 72000.[69] Under *anaerobic conditions* it catalyzes the *reduction of NO to N_2O* using NADH as electron donor:

$$2\ \mathrm{NO} + \mathrm{NADH} + \mathrm{H^+} \rightarrow \mathrm{N_2O} + \mathrm{H_2O} + \mathrm{NAD^+}\ .$$

Interestingly P450nor is a reductase, whereas most P450 enzymes work as monooxygenases. For this reason it has an outstanding role in the P450 family.

1.2.2. Structural Characteristics

All P450 enzymes have the same prosthetic heme group in common (see Figure 1.2).[66] Heme consists of a protoporphyrin IX which is coordinating an iron ion. Protoporphyrin IX is a large heterocyclic organic ring containing four pyrrole rings, which are linked by methene bridges. The large ring has eight substituents, two propionate, two vinyl and four methyl groups.[1] The heme group is bound in the active site via proximal coordination to a cysteine ligand

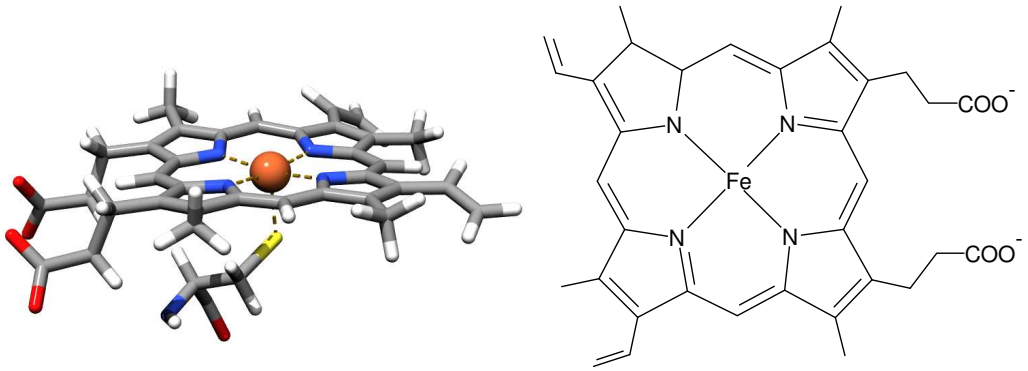


Figure 1.2. – Common heme prosthetic group from the active site of P450 enzymes (left figure made from crystal structure, PDB code 1CL6[70]).

of the protein and via salt bridges between propionate groups and residues from the protein. The *cysteine residue*, which is acting as a fifth ligand on the proximal heme side, is a *specific characteristics* of P450 enzymes, as other hemoproteins usually have a histidine ligand at this position.[71] The sixth coordination site of iron is on the distal heme side, opposite to cysteine. Water or incoming ligands are bound and the reaction takes place at this position. The first X-ray crystal structure of P450nor was obtained in 1997 and showed that the enzyme exhibits a large heme distal pocket, which connects the sixth coordination site to the protein surface.[72] This is visualized in Figure 1.3, where the channel reaching from the protein surface to the active site is shown. The discovery of the large heme distal pocket gave a hint at the mechanism, how NADH (nicotinamide adenine dinucleotide) might reduce the enzyme (see next paragraph).

NADH Binding

Nicotinamide Adenine Dinucleotide Before coming to NADH-binding in P450nor, the structural and biochemical properties of nicotinamide adenine dinucleotide are presented in

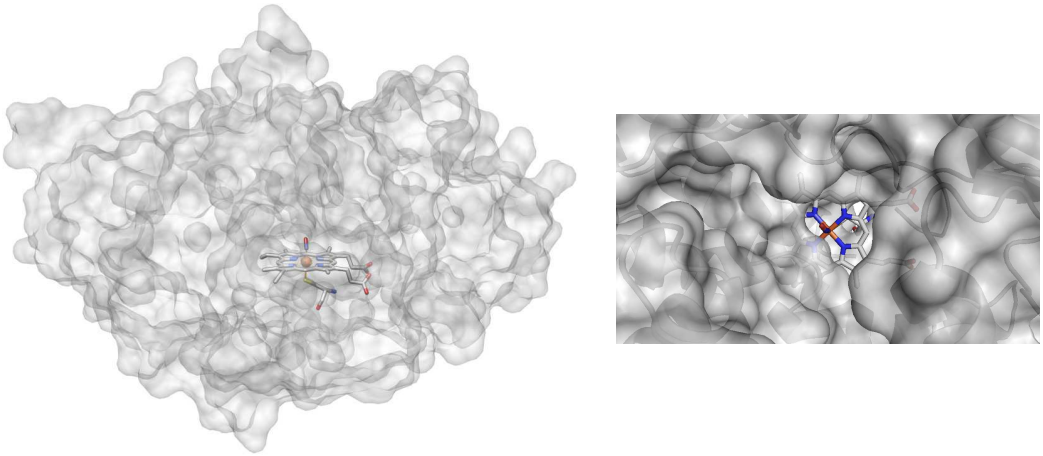


Figure 1.3. – Crystal structure of Cytochrome P450 NO reductase (PDB code 1CL6[70]) showing the location of the heme group and the heme distal pocket. The protein is shown in surface representation and the heme group with the NO ligand in the ball-and-stick representation. On the left side, the entire protein is shown. On the right side, a view from the protein exterior in the direction of the active site is given, showing the heme distal pocket and channel reaching from the protein surface to the heme group. Note that overlapping residues cover a large part of the heme distal pocket.

more detail.[73]
The structure of NADH is drafted in Figure 1.4. NADH consists of two nucleotides (a ribose covalently bound to adenine — lower nucleotide — and a ribose bound to nicotinamide — upper nucleotide) which are connected by two phosphate groups.[1] In Figure 1.5 the structure of nicotinic acide adenine dinucleotide (NAAD) is shown, which is a dinucleotide similar to NADH that possesses a carboxyl group where NADH has an amide group. NAAD is used in crystal structure analysis of P450nor and will be mentioned in the next paragraph.[74]
NADH plays a central role in the metabolism of all types of living cells. Here, it occurs in many different redox reactions, where it accepts or donates electrons in form of a hydride ion.[1] In the metabolizing cell, NADH and NAD^+ are continuously converted into one another as illustrated in Figure 1.6. The redox midpoint potential of the redox pair NAD^+/NADH is relatively low (-320 mV at pH 7), which makes NADH a strong reducing agent.[1]
NADH and NAD^+ show high UV absorbance at different wavelengths (NAD^+ absorbs at 260 nm, whereas NADH absorbs at 260 and 340 nm), which makes it relatively easy to measure their ratio in a kinetic experiment.[1]

NADH in P450nor Usually, bacterial NO reductases use electron carriers in order to transfer the electrons from the reducing cofactor to the heme center.[58,75] P450nor is an exception, it only uses NADH, without other cofactors. Thus, for some time the question was open how the electrons are transferred from NADH to the active site. An additional aspect of NADH oxidation is that it is usually oxidized in two steps with the help of one-electron carriers. Due to the lack of one-electron carriers it was not clear whether, in the course of the reaction, NADH is oxidized in one step by two electrons, or twice by one electron.
The large heme distal pocket was suspected to be a *potential NADH-binding site*. A unique accumulation of positively charged amino acids (Lys and Arg) is present in the upper part of the heme distal pocket, providing a binding site for the phosphate groups of NADH.

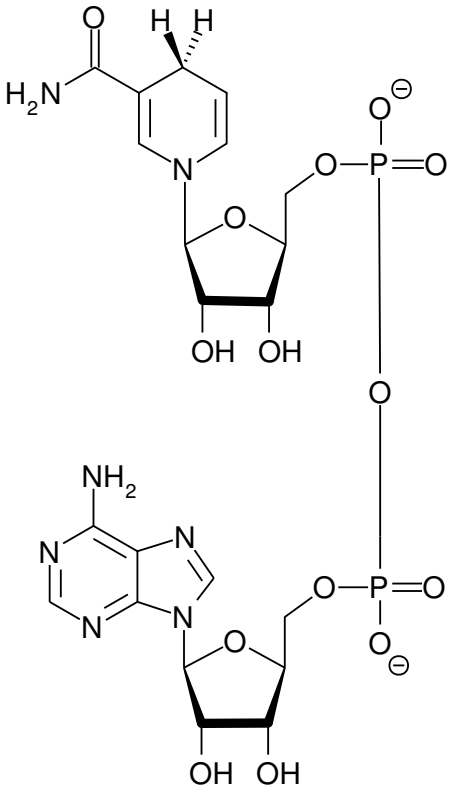


Figure 1.4. – Structure and composition of nicotinamide adenine dinucleotide (NADH).

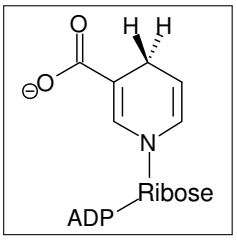


Figure 1.5. – Structure of nicotinic acid adenine dinucleotide (NAAD).

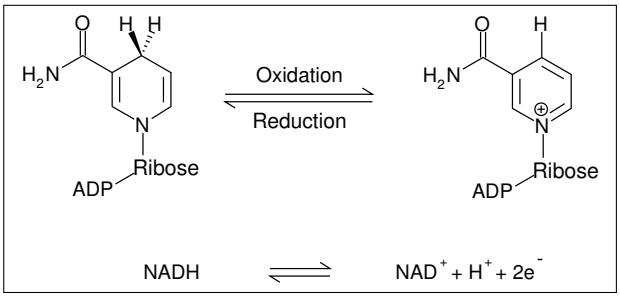


Figure 1.6. – NADH is shown in its two states, the reduced form (NADH) on the left side and the oxidized form (NAD^+) on the right side.

Mutation studies supported this hypothesis by showing that these residues are important for NADH-binding, and that especially Arg_{64} and Arg_{174} in this region are vital for NADH-binding.[76,77] Later, Daiber and coworkers showed by kinetic isotope effect (KIE) studies that the reduction of the NO-bound species is stereoselective with respect to NADH.[78] The question on NADH-binding could finally be answered in 2004 when the crystal structure of P450nor with a bound cofactor became available, providing much experimental information on NADH-binding — unfortunately, not NADH itself, but only the slightly different NAAD (see again Figure 1.5) could be cocrystallized with the protein.[74] This crystal structure shows that the cofactor binds directly in the distal heme pocket, with the hydride transfer group being in direct proximity to the sixth coordination site of iron (see Figure 1.7). Arg_{64} and Arg_{174} can be seen to interact strongly via salt bridges with the pyrophosphate moiety at the center of the cofactor, while the other positively charged residues form hydrogen bonds with other parts of the dinucleotide via water molecules. These two studies (on the KIE and on the cofactor-bound crystal structure) gave enough experimental evidence to confirm the hypothesis of direct reduction by NADH.

Interestingly, two negatively charged residues (Glu_{71} and Asp_{88}) were also found in the vicinity of the positively charged cluster. It turned out that they are acting like guards for Arg_{64}.[79,80] Without the cofactor they form salt bridges with Arg_{64}, whereas after dinucleotide binding the Arg_{64}-Asp_{88} salt bridge breaks. This interaction destabilizes NADH-binding in P450nor, which is found to be important for the mechanism of *rapid NADH release after reduction*. It is concluded that *P450nor has a rather low affinity for NADH*, which is crucial

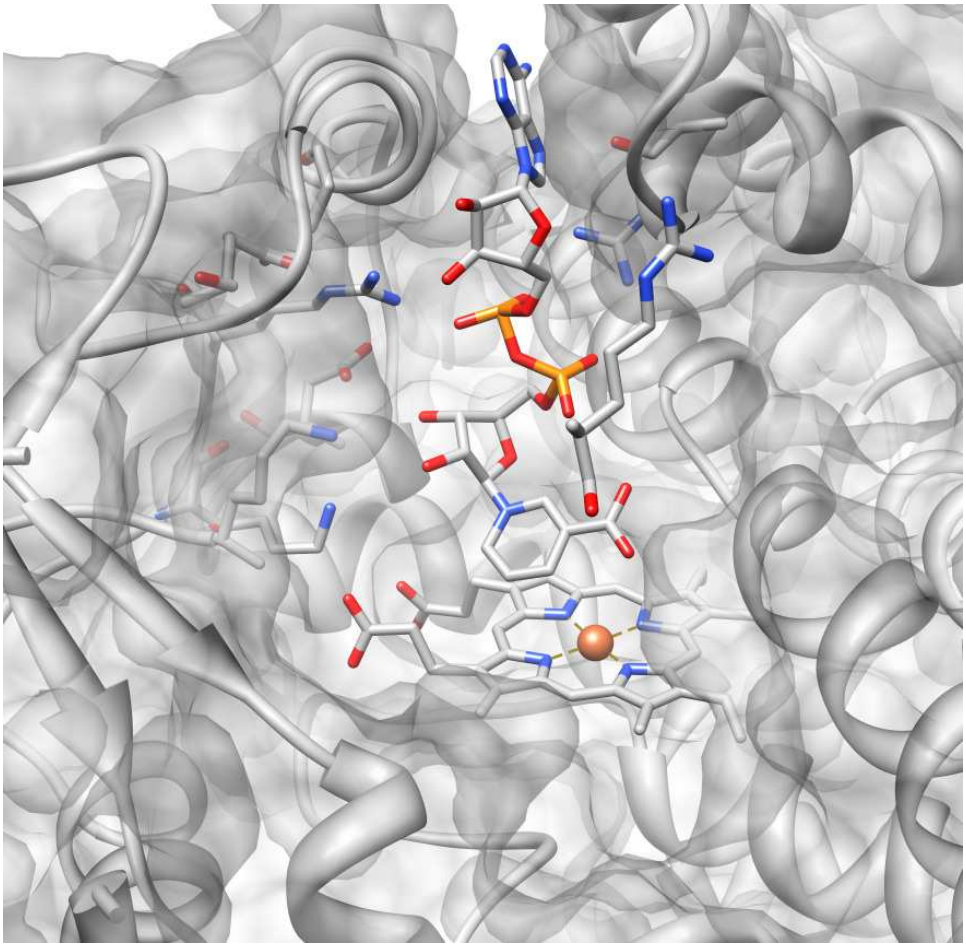


Figure 1.7. – Crystal structure of Cytochrome P450 NO reductase (PDB code 1XQD[74]) cocrystallized with NAAD, showing the cofactor bound in the heme distal pocket. The protein is shown in surface and cartoon representation; the heme group, NAAD and residues that are binding to the cofactor are shown in ball-and-stick representation.

for the NAD^+ release after the reduction.
Thr_{243}, a residue which is highly conserved in P450 enzymes, is well-known for providing protons for the monooxygenase reaction in usual P450 enzymes.[81] In P450nor, Thr_{243} is essential for the reduction step.[81–83] In the crystal structure with NAAD, Thr_{243} interacts with the carboxyl group of the nicotinic acid, leading to the assumption that the conserved Thr_{243} is important for the recognition of the nicotinamide group in NADH-binding.

Proton Transfer Network

Protonation is an important step in the P450nor reaction cycle. In the X-ray structure, a pathway for proton delivery, which reaches from the protein surface to the reaction site, could be revealed.[72] This hydrogen bond network goes via the hydroxyl group of Ser_{286} and the carboxyl group of Asp_{393}. Mutants of Ser_{286} and Asp_{393} show a dramatic decrease of enzyme activity, confirming their role for the reaction mechanism.[70,84] When the cofactor is present in the heme-distal pocket, the original hydrogen bonding network is disrupted and rearranges, connecting the reaction site of iron with the protein surface.[74] Asp_{393} and Ser_{286}

are still involved in the newly formed hydrogen bonding network. However here, Ser_{286} plays two different roles:

- it participates in the hydrogen bonding network, and
- it is also involved in the recognition of the cofactor.

1.2.3. Catalytic Cycle

The reaction catalyzed by P450nor was first discovered in 1991. In the following years, it was revealed that three reaction intermediates can be observed in the P450nor reaction cycle. They can be distinguished by different Soret bands in their absorption spectra, and represent:[69]

- the resting state (**1**),
- the NO-bound state (**2**) and
- an intermediate *I*.

A mechanism which connects the three intermediates is shown in Figure 1.8.[d] The first two species are experimentally well accessible, so that a large amount of experimental data exists.[70,72,83–87] Consequently, there is also a consensus about their properties.[88–90]
The third intermediate is not easily accessible to experimental structural analysis due to its *short lifetime*, and thus no exhaustive knowledge about its composition can be obtained. Also the spectroscopic data about intermediate *I* are quite limited. This lack of data on the experimental side is mirrored by a variety of structures proposed from both experimental and computational studies.[78,89,91,92]

Resting State

The resting state of the enzyme carries a low-spin ferric iron with a water molecule bound at its distal coordination site.[86,87] EPR experiments showed that part of the enzyme also exist in the ferric high spin state, in which iron is five-coordinate and does not carry a water molecule.[85]

Binding of Nitric Oxide

The ferric NO-bound state ($\{FeNO\}^6$ according to the Enemark-Feltham notation[e]) is formed through replacement of the water molecule by an NO molecule. The formation of the NO-bound state is relatively fast with a binding rate of $k = 26 \times 10^6$ M^{-1} s^{-1}.[69] The high binding rate is made possible by the large heme distal pocket.
The crystal structures and, more importantly, EXAFS data show that the Fe-NO unit is bent with an Fe-N-O angle of $\approx 160°$.[70,86] For quite some time it was speculated whether this bending is a steric effect of the protein active site, or an electronic effect. With the help of model complexes it could be shown that the bent Fe-NO unit is an intrinsic feature of ferric heme-nitrosyl species coordinated with cysteine.[96] It could be shown computationally that the Fe-N-O bending is a σ-trans effect of the axial thiolate ligand onto the NO ligand.[90] Electronically, a σ-backbond from the S-Fe σ-bond into the NO-$\sigma^\star$ orbital is formed, and this slightly favors the tilted structure.

[d] Only one scenario, the mechanism via the doubly protonated intermediate *I*, is shown.

[e] For metal-nitrosyl complexes $\{M(NO)_x\}^n$ the Enemark-Feltham notation gives the number *n* of electrons in the metal-d and the NO-$\pi^\star$ orbitals, without assigning a formal oxidation state on M or NO.[93–95]

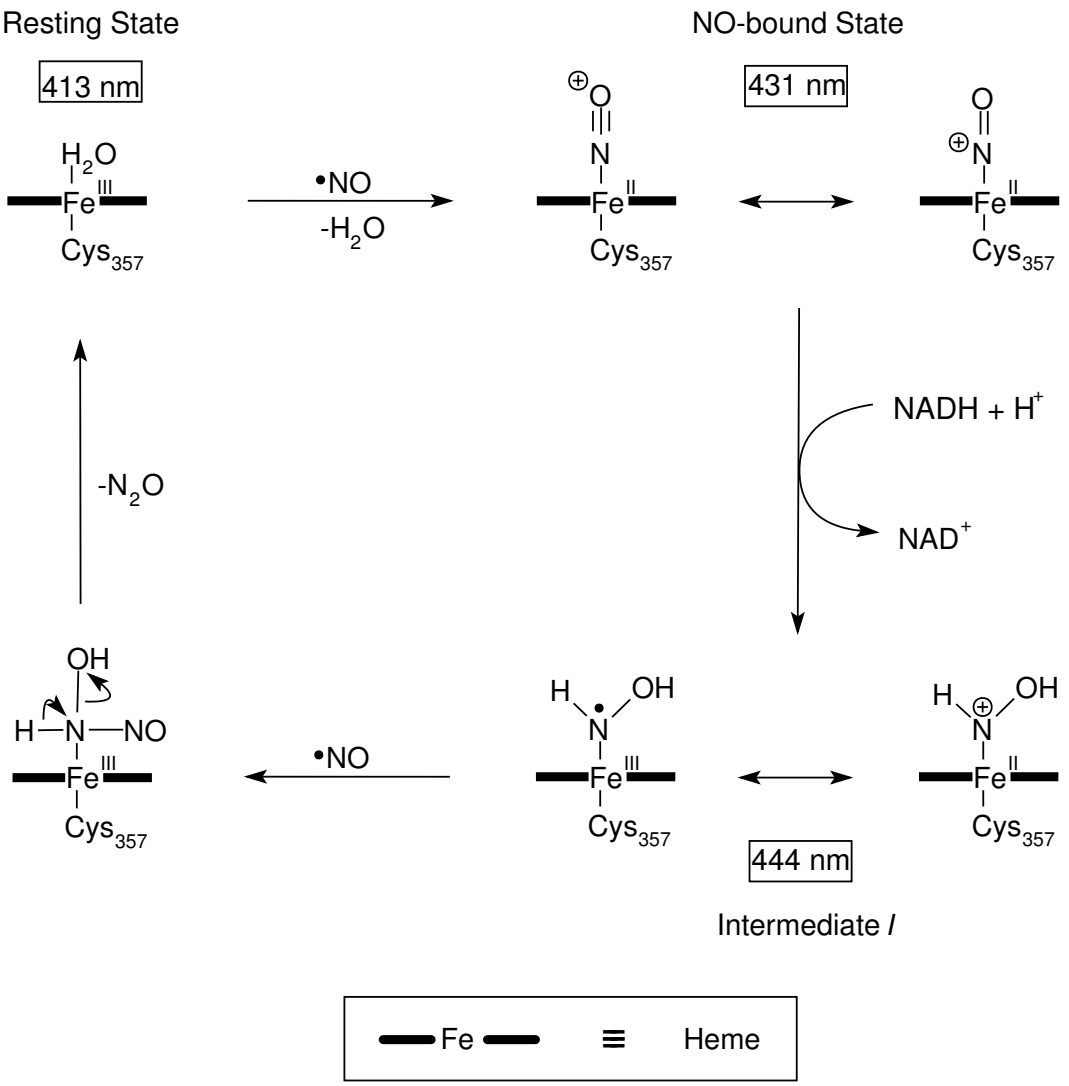


Figure 1.8. – Reaction pathway for the reduction of NO to N_2O as proposed by Daiber et al.[89]

Formation of Intermediate *I*

Intermediate *I* is generated via the reduction of the NO-bound state by NADH. The reduction by NADH is the rate-limiting step in the P450nor reaction cycle, with an experimental rate constant of $k = 0.9 \times 10^6$ M^{-1} s^{-1} (which corresponds to a barrier height of 8.8 kcal/mol at 10°C).[69] Intermediate *I* is seen as the *key intermediate* of the P450nor reaction cycle. It is not always observed, except only under certain reaction conditions.[69,78,89] When NADH is added to the NO-bound enzyme, two different experimental observations are made depending on the prevailing NO concentration [NO] (compared to the enzyme concentration [E]):

- At *very low NO concentrations* ([NO] $\approx$ [E]) $\rightarrow$ *I can be observed* (besides **1** and **2**).
- At *high NO concentrations* ([NO]>>[E]) $\rightarrow$ *I is not observed* (only **1** and **2** are detected.

Under the latter reaction conditions, *I* decomposes so quickly that it doesn't accumulate.
The *structural and electronic nature* of intermediate *I* has been discussed for quite a long time.[58,69,78] Several DFT cluster model studies were carried out to find a satisfying answer.[88,91,97,98] It is consensus nowadays that the intermediate is either of two possibilities:

- The singly protonated form of the reduced NO-bound state (formally [Fe-NHO]$^{2+}$, **3**).
- The respective doubly protonated form (formally [Fe-NHOH]$^{3+}$, **4**).

Note, that in Figure 1.8 only the latter scenario is shown.
The most convincing computational study about intermediate *I* was performed by Lehnert et al.,[91] who used vibrational spectroscopy and pK_A-values in order to correlate their modeled structures to experiment. They claimed that the doubly protonated form [Fe-NHOH]$^{3+}$ is 8.6 kcal/mol more stable than the singly protonated form, and that electronically it corresponds to a Fe(IV)-NHOH$^-$ species.[f]

Decomposition

Intermediate *I* finally decomposes to yield nitrous oxide and water. The decomposition is initiated by the addition of another NO molecule to intermediate *I* in the active site.
Experimentally, a convincing reaction rate for the decomposition process has not yet been obtained. All experimental kinetic data on the last reaction step were obtained under limited NO concentrations, and thus under the condition that almost all NO is bound as intermediate *I*. The rate of decomposition observed in this manner thus depends highly on the *back reaction* of intermediate *I* to the resting state and free NO. It is a known fact that decomposition is faster than intermediate *I* formation,[89] and thus at least an upper barrier for the activation enthalpy of the decomposition reaction rate is known (8.8 kcal/mol, see previous paragraph).
Computationally the decomposition reaction was modeled in two DFT studies.[91,98] They found barriers of 26 and 35 kcal/mol, which are far too high compared to the here given upper barrier.

[f] The electronic state is, however, not in agreement with experimental Mössbauer data, obtained for intermediate *I* later (see also Chapter 3) and Figure B.1.

2. Methods

As a part of this thesis, three implementation projects were carried out. These projects are related to the fields of geometry optimization, QM/MM methodology and theoretical EPR spectroscopy.

Geometry optimization The existing geometry optimization methods in the QM program package ORCA[99] were extended. Transition state (TS) optimization and constrained optimization were implemented following algorithms described in the literature.[100,101] Additionally, two approximations to the exact Hessian matrix, needed for TS optimization, were developed and implemented. The methodologies and their background are described in Section 2.1. The implemented methods were used in the calculations that are described in Chapters 3 to 5.

QM/MM methodology The QM/MM interface of the MM program package GROMACS[102] was extended in order to be applicable with ORCA. Additionally, QM-MM VDW interaction terms were implemented in the energy and gradient calculation part of the ORCA code. These terms were implemented in ORCA in order to make an optimization (performed by ORCA) possible, in which the QM system is interacting with surrounding MM atoms that are represented by their atomic point charges and Lennard-Jones[a] coefficients. This allows the ORCA user to perform

- a QM/MM optimization (with ORCA/GROMACS) applying a microiterative optimization scheme or
- use the ORCA optimizer (without GROMACS) for a QM/MM optimization, in which the MM atoms are kept fixed.

The latter is of importance, since TS optimization methods are not implemented in GROMACS, but in ORCA.
A QM/MM study was performed on Cytochrome P450nor, using the implemented QM/MM methodology and geometry optimization methods. The results of the study are summarized in Chapters 3 and 4.

Approximations to Spin-Spin Zero-Field Splitting Finally, in Section 2.3, the theoretical derivation of different approximations to the spin-spin zero-field splitting is given. The equations were implemented in a development version of ORCA.
Results obtained with the approximate treatment are compared to those results obtained with the exact treatment in Chapter 5. This enables us to study the limitations of different approximations that are used in the experimental EPR community.

[a] Lennard-Jones coefficients for the description of the QM-MM VDW interaction.

2.1. Geometry Optimization

Probably the most important step in quantum chemistry applications is the optimization of molecular structures. Geometry optimization usually has to be done before other properties of a chemical system can be calculated. Optimized geometries are of importance for the calculation of thermodynamic and kinetic properties of a chemical reaction, relative energies of isomers, spectroscopic properties and many more. In the field of geometry optimization, different minimization schemes can be used. The fastest and most accepted way to optimize systems with a few to several hundreds of atoms (which is the target group of molecules for quantum chemical methods) is to use second-derivative optimization algorithms. With such methods usually up to several dozens of optimization steps are sufficient to find an equilibrium geometry. The overall convergence of the optimization is improved by avoiding too large steps, and it is much accelerated by performing the geometry steps in internal coordinates (instead of the usually used Cartesian coordinates).
An overview of a typical geometry optimization is illustrated in the flowchart in Figure 2.1. In this section, we first describe the properties of the energy landscape of a molecular structure, before coming to different aspects of geometry optimization. In summary, the following topics are discussed:

- Second-derivative minimization algorithms,
- step size control,
- transition state optimization,
- transformation between coordinate systems,
- initial Hessian and Hessian update,
- constrained optimization.

2.1.1. Potential Energy Surface

To understand the concept of geometry optimization, it is helpful to first have a look at the properties of the potential energy surface (PES) of a molecular structure (see also the book chapter by Hratchian and Schlegel[103]). A PES is a multi-dimensional hypersurface, in which the energy of a molecular system is a function of the coordinates of its N nuclei (where N is the number of atoms in the system).
Important points on a hypersurface are *stationary points*, i.e. points with zero gradient (see equation (2.2)) in all directions. For a chemical system, two classes of stationary points are relevant: minima and first order saddle points. They can be distinguished by the *curvature of the PES* at the stationary point. The curvature at any point on the PES can be determined by calculating the Hessian matrix. The latter is the second derivative of the energy with respect to all $3\times N$ coordinates (for its definition see equation (2.2)). Calculating its eigenvectors and eigenvalues yields the principal directions of the system with corresponding pure curvature. A positive (negative) eigenvalue of the Hessian matrix means that the PES is curved to the left (right) in the direction of the respective eigenvector.
A *minimum* is defined as a stationary point with a positive definite Hessian matrix (i.e. all eigenvalues are positive). A *saddle point* on the other hand has at least one negative Hessian eigenvalue. The number of negative eigenvalues defines the order of the saddle point. If

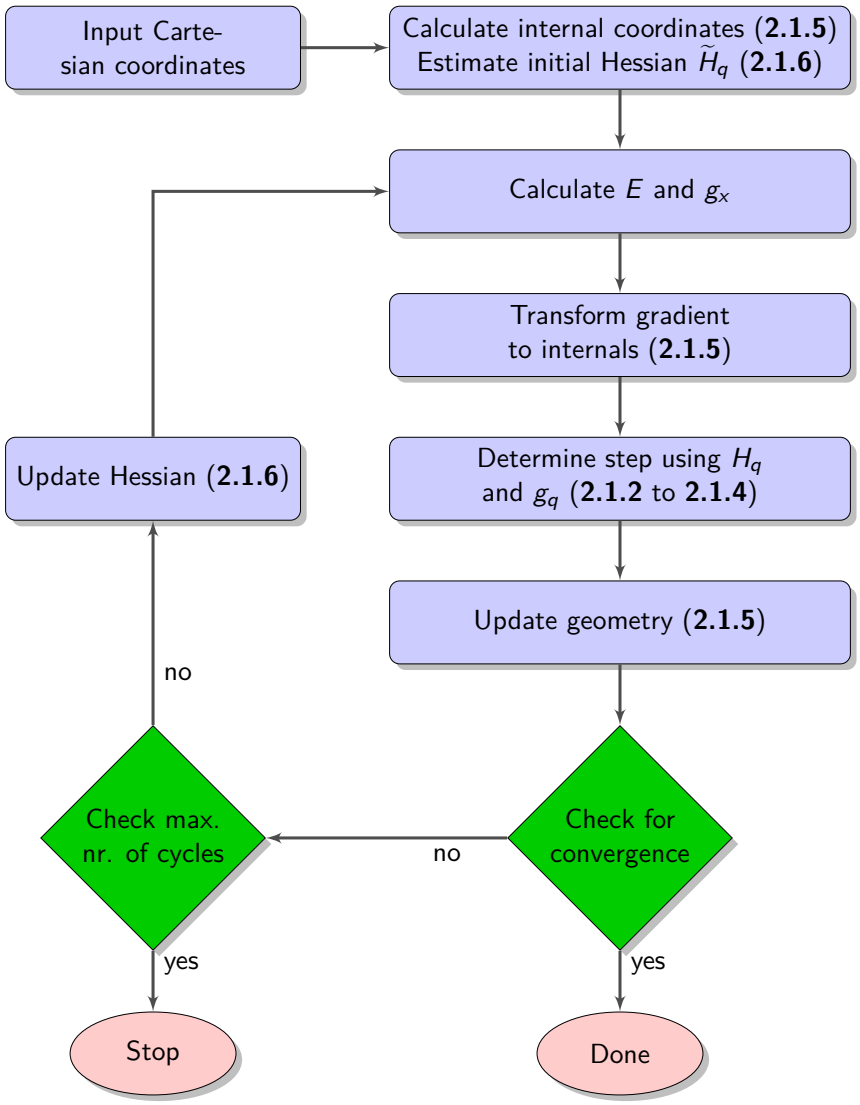


Figure 2.1. – Flowchart of a typical geometry optimization. For the important steps references are given to the respective sections.

the point has one (two, etc.) negative Hessian eigenvalues (and all other eigenvalues are positive), then the respective stationary point is a first- (second-, etc.) order saddle point. It will become clear below that only first-order saddle points are chemically relevant.
A PES can be fully displayed in internal coordinates only for two-atomic molecules. For larger molecules one usually chooses one or two degrees of freedom of interest (here denoted x and y) and draws the energy of the system as a function of x and y. For each point on the thus displayed PES the energy of the system is usually minimal with respect to all other degrees of freedom.
In Figure 2.2 a simplified, two-dimensional PES is drawn. Let us assume that the PES shows two competing reactions of an educt A to two different products B and C (i.e. the reactions $A \rightarrow B$ and $A \rightarrow C$) via two different transition states $AB^{\ddagger}$ and $AC^{\ddagger}$ and that the two chosen internal coordinates are two different bond lengths of the molecular structure. Three local minima are shown in Figure 2.2, corresponding to the three stable intermediates A, B and C. At these points, the PES has left curvature in both displayed directions. That

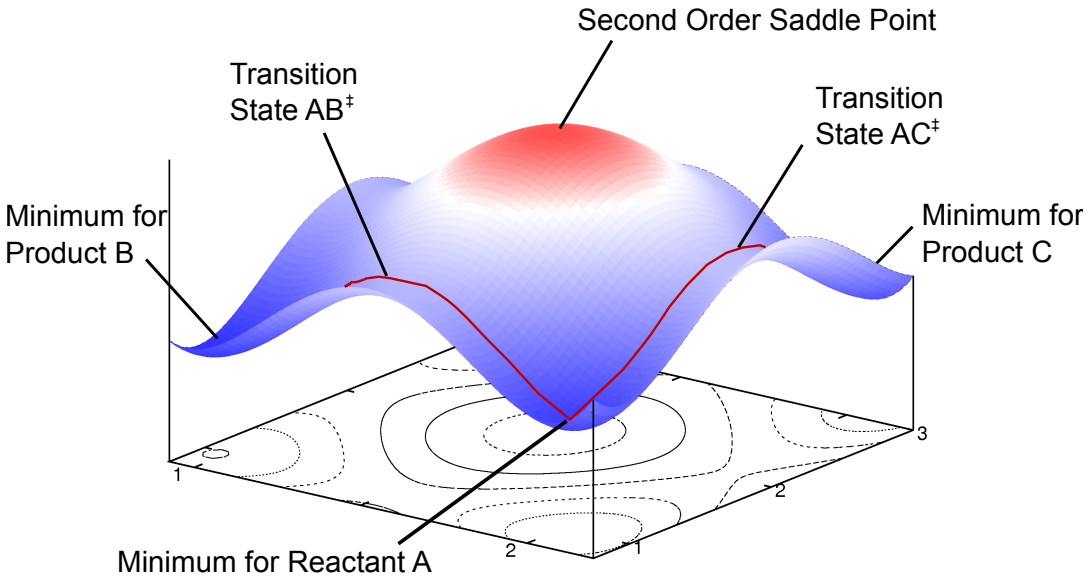


Figure 2.2. – Illustration of important stationary points on a Potential Energy Surface.

is, moving away from a minimum automatically leads to increasing energy. In the chemical reaction $A \rightarrow B$, the system is climbing upwards from A to the first order saddle point $AB^{\ddagger}$ (red curve in Figure 2.2).[b] After the saddle point is reached, the system continues its way downhill to the minimum B. $AB^{\ddagger}$ thus corresponds to the TS of the chemical reaction $A \rightarrow B$. From this example it becomes clear that only first-order saddle points are important for chemical reactions. Such a first-order saddle point has to be traversed in order to come from one minimum to the other. The barrier to reach it determines the *minimum energy that has to be spent for the reaction to take place*. What can also be drawn from Figure 2.2 is the meaning of the Hessian eigenmode corresponding to the negative eigenvalue of the TS (also called TS mode or TS eigenvector). The TS eigenvector points in the direction of the reaction from A to B. With respect to all other directions, the saddle point is left curved (but only one of them is shown here). The path $A \rightarrow C$ can be similarly described as that of $A \rightarrow B$.
Higher order saddle points (e.g. the second-order saddle point in Figure 2.2) are not of interest for the study of chemical reactions.
Whether one ($A \rightarrow B$) or the other ($A \rightarrow C$) reaction takes place depends on the relative stability of the intermediates A, B and C, on the barrier heights from A to $AB^{\ddagger}$ and $AC^{\ddagger}$ and on the reaction conditions applied (thermodynamic or kinetic control).

2.1.2. Second-Derivative Methods

Having introduced the properties of a PES, we can now turn our attention to methods that are used to find stationary points on the PES.
The energy at any arbitrary point $\mathbf{p}$ on the PES (which corresponds to a specific molecular

[b] This is an illustrative approach to describe the chemical reaction $A \rightarrow B$, neglecting dynamic effects, etc.

configuration) can be formulated by the Taylor expansion around another point $\mathbf{k}$,

$$E(\mathbf{p}) = E(\mathbf{k} + \Delta\mathbf{s}) = E(\mathbf{k}) + \mathbf{g}_k^{\mathsf{T}}\Delta\mathbf{s} + \frac{1}{2}\Delta\mathbf{s}^{\mathsf{T}}\mathbf{H}_k\Delta\mathbf{s} + \dots \ , \tag{2.1}$$

with the gradient vector

$$\mathbf{g} = \begin{pmatrix} \frac{\partial E}{\partial x_1} \\ \frac{\partial E}{\partial x_2} \\ \vdots \\ \frac{\partial E}{\partial x_n} \end{pmatrix} \ , \tag{2.2}$$

the second-derivative matrix of the energy with respect to the nuclear coordinates

$$\mathbf{H} = \begin{pmatrix} \frac{\partial^2 E}{\partial x_1^2} & \frac{\partial^2 E}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 E}{\partial x_1 \partial x_n} \\ \frac{\partial^2 E}{\partial x_2 \partial x_1} & \frac{\partial^2 E}{\partial x_2^2} & \cdots & \frac{\partial^2 E}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 E}{\partial x_n \partial x_1} & \frac{\partial^2 E}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 E}{\partial x_n^2} \end{pmatrix} \ , \tag{2.3}$$

and $\Delta\mathbf{s}$ being the displacement vector from $\mathbf{k}$ to $\mathbf{p}$.

Quadratic Approximation

Experience shows that the local properties around stationary points on a PES can approximately be modeled by a quadratic hypersurface (*quadratic approximation*). Within this approximation, all terms in the Taylor expansion of third order or higher are neglected,

$$E(\mathbf{k} + \Delta\mathbf{s}) \approx E^{QA}(\mathbf{k} + \Delta\mathbf{s}) = E(\mathbf{k}) + \mathbf{g}_k^{\mathsf{T}}\Delta\mathbf{s} + \frac{1}{2}\Delta\mathbf{s}^{\mathsf{T}}\mathbf{H}_k\Delta\mathbf{s} \ . \tag{2.4}$$

Equation (2.4) gives a good approximation to the energy at any point near a stationary point. Differentiating this equation with respect to $\Delta\mathbf{s}$ results in the relation:

$$\frac{\mathrm{d}E^{QA}(\mathbf{k} + \Delta\mathbf{s})}{\mathrm{d}\Delta\mathbf{s}} = \mathbf{g}_k + \mathbf{H}_k\Delta\mathbf{s} \ . \tag{2.5}$$

We are looking for a step $\Delta\mathbf{s}$ which brings us to a stationary point. At a stationary point the gradient vanishes, and the left side of equation (2.5) becomes zero (*stationary condition*). Reordering the obtained equation finally leads to the *Newton-Raphson (NR) step*:

$$\Delta\mathbf{s}^{NR} = -\mathbf{H}_k^{-1}\mathbf{g}_k \ , \tag{2.6}$$

for which information about the first and second derivatives is needed.
In the ideal case of a quadratic surface, and with exact knowledge of the gradient $\mathbf{g}_k$ and the Hessian $\mathbf{H}_k$, the stationary point could be reached within one single Newton-Raphson step $\Delta\mathbf{s}^{NR}$. However, in practice, the PES is not exactly quadratic and thus more than one geometry step is needed to reach the stationary point.

Quasi-Newton Step

Unfortunately, the calculation of the exact Hessian matrix $\mathbf{H}_k$ is much more expensive than an energy or gradient calculation. It is thus very advantageous not to calculate $\mathbf{H}_k$ explicitly, but to use an approximate Hessian matrix $\widetilde{\mathbf{H}}_k$ instead. This then defines the *Quasi-Newton (QN) step*:

$$\Delta\mathbf{s} = -\widetilde{\mathbf{H}}_k^{-1}\mathbf{g}_k \ . \tag{2.7}$$

The approximate Hessian (there exist different methods to form an approximate Hessian, see Section 2.1.6) is usually constructed only once at the beginning of the geometry optimization. In the course of the optimization several geometry steps are taken and the approximate Hessian can be improved by using Hessian update schemes (see Section 2.1.6). This approach slightly increases the number of geometry optimization steps needed to converge to a stationary point, but on the other hand a significant amount of time is saved in the determination of the Hessian matrix.

2.1.3. Step Size Control

The quadratic approximation is only valid for a small region around the stationary point. If a step is too large, it can end up outside of the quadratic region and thus deteriorate the convergence of the optimization procedure. Thus, the step size has to be controlled.
The simplest method to control the step size is to scale down the length of each component s_i and the average length of $\Delta\mathbf{s}$, if they are too large.
A more elaborate approach is the rational function optimization (RFO), which is quite often used for step size control.

Rational Function Optimization

In the RFO approach,[104,105] the second order Taylor approximation of equation (2.4) is expanded with a scaling factor in the following way:

$$E^{QA}(\mathbf{k}+\Delta\mathbf{s}) = E(\mathbf{k}) + \frac{\mathbf{g}_k^\mathsf{T}\Delta\mathbf{s} + \frac{1}{2}\Delta\mathbf{s}^\mathsf{T}\mathbf{H}_k\Delta\mathbf{s}}{1+\Delta\mathbf{s}^\mathsf{T}\mathbf{S}\Delta\mathbf{s}} \ . \tag{2.8}$$

The scaling factor $(1+\Delta\mathbf{s}^\mathsf{T}\mathbf{S}\Delta\mathbf{s})$ is step size dependent. Equation (2.8) can be rewritten as:

$$\epsilon = E^{QA}(\mathbf{k}+\Delta\mathbf{s}) - E(\mathbf{k}) = \frac{\frac{1}{2}\begin{pmatrix} \Delta\mathbf{s}^\mathsf{T} & 1 \end{pmatrix}\begin{pmatrix} \mathbf{H}_k & \mathbf{g}_k \\ \mathbf{g}_k^\mathsf{T} & 0 \end{pmatrix}\begin{pmatrix} \Delta\mathbf{s} \\ 1 \end{pmatrix}}{\begin{pmatrix} \Delta\mathbf{s}^\mathsf{T} & 1 \end{pmatrix}\begin{pmatrix} \mathbf{S} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \Delta\mathbf{s} \\ 1 \end{pmatrix}} \ , \tag{2.9}$$

with the $(N+1)\times(N+1)$ *augmented Hessian matrix* $\mathbf{H}_{aug} = \begin{pmatrix} \mathbf{H}_k & \mathbf{g}_k \\ \mathbf{g}_k^\mathsf{T} & 0 \end{pmatrix}$, and $\mathbf{S}$ is usually taken to be the unit matrix.[c] Transforming the equation further yields the relation:

$$\begin{pmatrix} \mathbf{H}_k & \mathbf{g}_k \\ \mathbf{g}_k^\mathsf{T} & 0 \end{pmatrix}\begin{pmatrix} \Delta\mathbf{s} \\ 1 \end{pmatrix} = \lambda\begin{pmatrix} \mathbf{S} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \Delta\mathbf{s} \\ 1 \end{pmatrix} \ , \tag{2.10}$$

[c] If the denominator of (2.9) is set to 1, the equation results in the quadratic model equation (2.4).

with $\lambda = 2 \cdot \epsilon$. Setting the scaling matrix $\mathbf{S}$ to $\mathbf{1}$, a simple eigenvalue equation is obtained:

$$\begin{pmatrix} \mathbf{H}_k & \mathbf{g}_k \\ \mathbf{g}_k^\mathsf{T} & 0 \end{pmatrix} \begin{pmatrix} \Delta\mathbf{s} \\ 1 \end{pmatrix} = \lambda \begin{pmatrix} \Delta\mathbf{s} \\ 1 \end{pmatrix} \ . \tag{2.11}$$

Diagonalization of the augmented Hessian matrix $\mathbf{H}_{aug}$ then yields the eigenvectors $\begin{pmatrix} \Delta\mathbf{s} \\ 1 \end{pmatrix}$ and corresponding eigenvalues λ. The step $\Delta\mathbf{s}$ can then be evaluated from the eigenvector with the lowest eigenvalue, λ_0. As the optimization converges, λ_0 approaches zero.

2.1.4. Transition State Optimization

TS optimization is used to find first-order saddle points on a PES. It searches for a point on the PES which is *maximal with respect to one eigenvector of the Hessian and minimal with respect to all other eigenvectors*. A very efficient method for this purpose is the partitioned RFO approach. Here, a distinct eigenmode, which should be maximized, *can be chosen* at the beginning of the procedure. The approach, including its implementation, is described in the following. A flowchart for a typical TS optimization is drawn in Figure 2.3.

Eigenvector Following

The Newton-Raphson step (equation (2.6)) can be reformulated in terms of the eigenspace of the Hessian matrix:

$$\Delta\mathbf{s} = -\mathbf{H}^{-1}\mathbf{g} = -\sum_i \frac{\mathbf{v}_i^\mathsf{T}\mathbf{g} \cdot \mathbf{v}_i}{b_i} = -\sum_i \frac{f_i \cdot \mathbf{v}_i}{b_i} \ , \tag{2.12}$$

with $\mathbf{v}_i$ and b_i being the eigenvectors and eigenvalues of the Hessian matrix $\mathbf{H}$ and

$$f_i = \mathbf{v}_i^\mathsf{T}\mathbf{g} \ , \tag{2.13}$$

being the component of $\mathbf{g}$ along the Hessian eigenvector $\mathbf{v}_i$. This equation forms the basis of the well-established *eigenvector following* (EF) method.[100,105]
In the limit of the local quadratic approximation, the NR step $\Delta\mathbf{s}$ leads to the nearest stationary point on the PES. The orientation of the NR step in each direction $\mathbf{v}_i$ depends on the respective eigenvalue b_i. It is directed *along* the gradient for each $\mathbf{v}_i$ with negative eigenvalue b_i (minimization), and *opposite* to the direction of $\mathbf{v}_i$ with positive eigenvalue b_i (maximization). Thus the stationary point where the NR step ends up can be either a minimum, or a saddle point, depending on the curvature of the PES (and the eigenvalue composition of the Hessian) at the actual starting point $\mathbf{k}$.
This behavior implies that only for points $\mathbf{k}$ with exactly one negative Hessian eigenvalue (i.e. for $\mathbf{k}$ in the realm of the TS) the algorithm can find a first order saddle point. If the Hessian has a different composition, the NR step is not leading in direction of the desired region, but to a minimum or a higher order saddle point. This means that, when applying the NR method, the starting point has to be very close to the requested stationary point. This is a severe drawback of the NR step.

Hessian Eigenvalue Shift

A solution to this problem is a simple modification of the NR step. A constant λ is subtracted from each b_i in the denominator of equation (2.12) .[106]

$$\Delta \mathbf{s} = -\sum_i \frac{f_i \cdot \mathbf{v}_i}{b_i - \lambda} \ , \tag{2.14}$$

This modification guides the step into the direction of a stationary point with the desired characteristics. The shift λ of the Hessian eigenvalues b_i can also be interpreted as changing the eigenvalue structure of the Hessian. Depending on the shift, the number of negative and positive *effective Hessian eigenvalues* $(b_i - \lambda)$ can be influenced.

Partitioned Rational Function Optimization

Banerjee et al. showed that equation (2.14) can also be obtained within the RFO approach (see Section 2.1.3).[105] Starting from equation (2.14), they went a step further and partitioned the TS optimization problem into two separate optimization problems. Here, the energy is

1. maximized along one Hessian eigenmode and
2. minimized along the remaining eigenmodes.

The full $(N+1) \times (N+1)$ augmented Hessian matrix (for a system with N degrees of freedom) is thus divided into two smaller *partitioned RFO (P-RFO) matrices*, namely a 2×2 matrix (as given in the eigenvalue equation 2.15),

$$\begin{pmatrix} b_k & f_k \\ f_k & 0 \end{pmatrix} \begin{pmatrix} \Delta s_k \\ 1 \end{pmatrix} = \lambda_p \begin{pmatrix} \Delta s_k \\ 1 \end{pmatrix} \ , \tag{2.15}$$

for *maximization* along the k-th mode, and a $N \times N$ matrix (as given in the eigenvalue equation 2.16),

$$\begin{pmatrix} \lambda_1 & & & & & & f_1 \\ & \ddots & & & 0 & & \vdots \\ & & \lambda_{k-1} & & & & f_{k-1} \\ & & & \lambda_{k-1} & & & f_{k+1} \\ & 0 & & & \ddots & & \vdots \\ & & & & & \lambda_n & f_n \\ f_1 & \dots & f_{k-1} & f_{k+1} & \dots & f_n & 0 \end{pmatrix} \begin{pmatrix} \Delta s_1 \\ \vdots \\ \Delta s_{k-1} \\ \Delta s_{k+1} \\ \vdots \\ \Delta s_n \\ 1 \end{pmatrix} = \lambda_n \begin{pmatrix} \Delta s_1 \\ \vdots \\ \Delta s_{k-1} \\ \Delta s_{k+1} \\ \vdots \\ \Delta s_n \\ 1 \end{pmatrix} \ , \tag{2.16}$$

for *minimization* along all other modes.
In theory, λ_p is obtained as the largest eigenvalue of the 2×2 augmented Hessian in equation (2.15) and λ_n as the lowest eigenvalue of the $N \times N$ augmented Hessian in equation (2.16). In practice, λ_p and λ_n are not determined by diagonalization of the two reduced augmented Hessian matrices. For the determination of λ_p, equation (2.15) can be converted to

$$\lambda_p = \frac{1}{2} b_k + \frac{1}{2} \sqrt{b_k^2 + 4 f_k^2} \ , \tag{2.17}$$

and λ_p can thus directly be calculated from the eigenvalue b_k and gradient f_k of the TS vector. Note here that λ_p is always positive and ensures that the effective Hessian eigenvalue $(b_k - \lambda_p)$ is negative.
For the determination of λ_n equation (2.16) can be converted to

$$\lambda_n = \sum_{i \neq k} \frac{f_i^2}{\lambda_n - b_i} \ . \tag{2.18}$$

From this relation, λ_n cannot be calculated directly, but has to be solved iteratively. Some care has to be taken when determining the solution to λ_n. In order to yield valid effective Hessian eigenvalues for minimization, it has to fulfill the following conditions:

1. $\lambda_n < 0$ and
2. $\lambda_n < b_i$, for $i \neq k$.

These conditions ensure that all effective Hessian eigenvalue $(b_i - \lambda_n)$ (for $i \neq k$) are positive. If a solution, fulfilling the above conditions, is not found, then the eigenvalue composition of the actual Hessian matrix can be modified:

- all eigenvalues are shifted upwards by a constant (e.g. smallest b_i $(i \neq k)$ $+10^{-4}$) or
- all negative, non-TS mode eigenvalues ($b_i < 0$, for $i \neq k$) are multiplied by a factor of -1.

Usually a valid λ_n can be determined after one of the above modifications.
After having obtained the shift parameters λ_n and λ_p, the total step $\Delta\mathbf{s}$ can be calculated:

$$\Delta\mathbf{s} = -\sum_{i=1}^{n} \Delta\mathbf{s}_i \ , \text{ with } \begin{cases} \Delta\mathbf{s}_i = \dfrac{f_i \cdot \mathbf{v}_i}{b_i - \lambda_p} \ , \text{ if } i = k \ . \\ \\ \Delta\mathbf{s}_i = \dfrac{f_i \cdot \mathbf{v}_i}{b_i - \lambda_n} \ , \text{ if } i \neq k \ . \end{cases} \tag{2.19}$$

Mode Following

The mode l, which has to be maximized, is usually chosen to be the mode with the lowest eigenvalue ($l = 0$). But this is not necessarily the case. The described algorithm can also be used to maximize along a different mode ($l > 0$). In principle any mode can be chosen as TS mode at the beginning of the TS optimization and the algorithm attempts to *follow the same mode* throughout the calculation. Here, the problem arises that the Hessian eigenvectors change with every optimization step. In order to find the mode in the actual iteration j which corresponds to the TS mode $\mathbf{v}_{TS}$ from the previous optimization step $j-1$, the overlap ov between all Hessian eigenvectors $\mathbf{v}_i$ (from the actual iteration) and $\mathbf{v}_{TS}$ (from the previous iteration) is calculated and $\mathbf{v}_i$ with highest overlap is chosen.

$$ov_i = \mathbf{v}_{TS,\ j-1} \cdot \mathbf{v}_{i,\ j} \tag{2.20}$$

In the course of the TS optimization, the eigenvalue corresponding to $\mathbf{v}_{TS}$ ideally becomes the lowest one.
If this is the case, i.e. if $l = 0$, the mode following algorithm is switched off and the mode with lowest eigenvalue is followed in the remainder of the TS optimization.

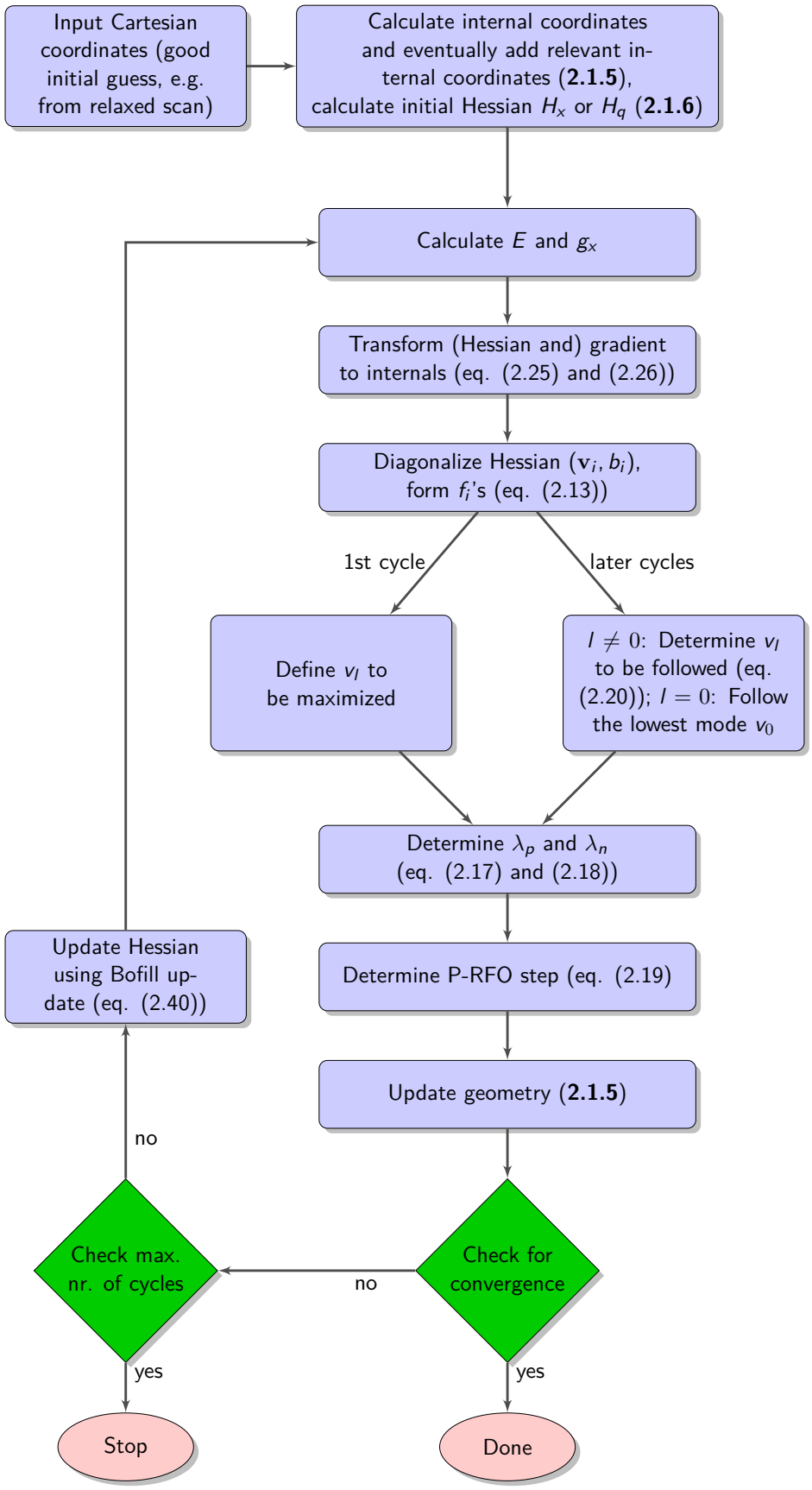


Figure 2.3. – Flowchart of a typical TS optimization. References to the respective equations and sections (bold) are given.

2.1.5. Transformation between Coordinate Systems

In quantum chemical calculations, the energy and the gradient are usually determined in Cartesian coordinates. The geometry optimization step is, on the other hand, most efficiently carried out in an *internal coordinate system*. Here, the quadratic approximation (basic assumption of all second-derivative methods, see Section 2.1.2) is best fulfilled.[103,107] An internal coordinate system consists of a set of bond lengths, bond angles and dihedral angles of the molecular system. There are several types of internal coordinate sets, and we are focusing on the most widely used for small to medium sized molecules: the set of *redundant internal coordinates*. Such a set includes, as the name already suggests, a higher number of coordinates than the necessary $3N - 6(5)$ degrees of freedom (usually about $3 - 4 \times N$ internal coordinates, for a system with N atoms).[d]
In this section, first the standard build-up of a redundant internal coordinate set is described, before the necessary transformations between the Cartesian coordinate system and the internal coordinate system are discussed.

Setup of Internal Coordinates

A redundant internal coordinate set consists of bond lengths, bond angles and dihedral angles. Their definitions are summarized in Figure 2.4. The construction of such a set (as

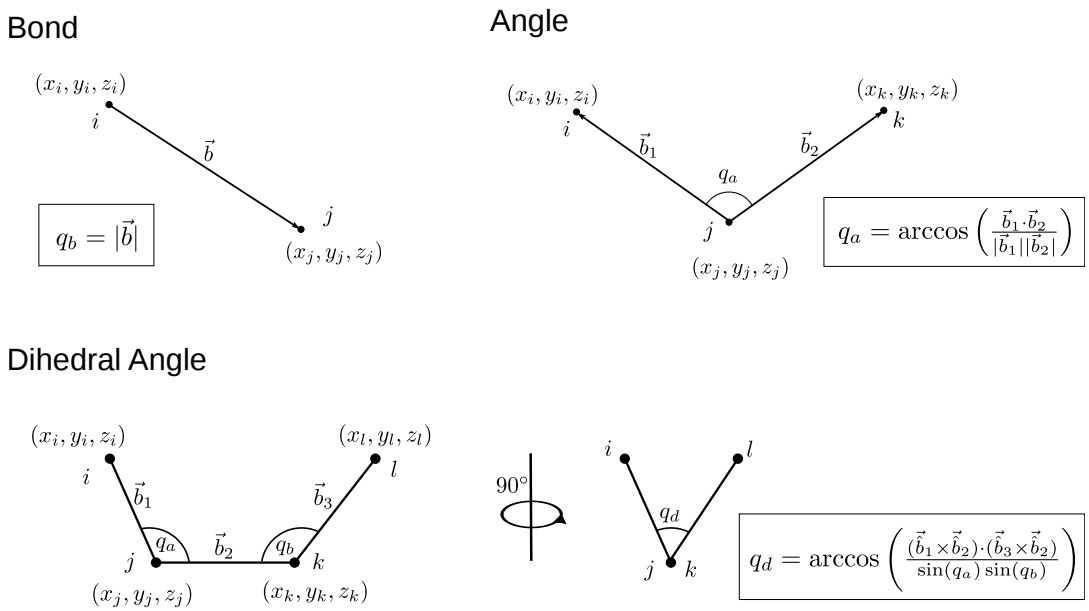


Figure 2.4. – Definition of the internal coordinates: bond lengths, bond angles and dihedral angles. The equations to determine the values of the internal coordinates from the Cartesian coordinates of the atoms are given (q_b, q_a and q_d).

e.g. implemented in ORCA) consists of three consecutive steps:

Bonds In the first step, *regular bonds* are assigned to all atom pairs i and j, whose interatomic distance is smaller than 1.3 times the sum of tabulated covalent radii of atoms i and j.

[d] One has to remove these redundancies before the geometry step is determined, as will be discussed later.

Since this procedure might result in an incomplete set of bonds with disconnected fragments (e.g. for a system with different, separated molecules), it is checked after the generation of the regular bonds whether there are still fragments in the overall system that are not connected. If two or more such fragments are found, the first two fragments (here named A and B) are connected with an *interfragmental bond* between the closest atom pair between A and B. Additional interfragmental bonds are added, if there are other close[e] atom pairs between A and B. This procedure is repeated until all fragments are connected.
The definition of all bond and dihedral angles in the following two steps is based on the thus obtained set of regular and interfragmental bonds.

Bond Angles In the second step of the setup, a bond angle is assigned to all *atom triplets out of pairs of bonds*, which have exactly one atom in common (e.g. as atom j in bonds $\vec{b_1}$ and $\vec{b_2}$, shown in Figure 2.4).
If a bond angle (q_a) is larger than $175°$, it is categorized as a *linear bond angle*.[f] In this case not only one, but two angle coordinates have to be assigned. Hereto, the positions of the three atoms are projected onto two orthogonal planes, as illustrated in Figure 2.5, and the two angles are determined as the angles of the projected vectors.

Dihedral Angles In the third step of the coordinate setup, a *dihedral angle* is assigned to *atom quadruples of directly neighboring bond angles*, which share two atoms in the following manner (see also Figure 2.4): Let us assume, one bond angle is defined by the atom triplet $i-j-k$, and another bond angle is defined by the atom triplet $j-k-l$. In this case, a dihedral angle is assigned to the atom set $i-j-k-l$.[g]

Add/Remove Internal Coordinates The automatically generated set of redundant internal coordinates can be modified. It is convenient (e.g. for TS optimizations, where the internal coordinate setup might not recognize a bond between two atoms belonging to the TS mode) to manually modify the above, automatically built set of redundant internal coordinates. Bonds, bond angles and dihedral angles can be added (removed) to (from) the automatically generated set.

Transformation between different Coordinate Sets

Once the internal coordinate set is built, the values for the respective coordinates can be easily determined from the Cartesian coordinates by using the equations given in Figure 2.4. The transformation is much more elaborated for the *reverse direction* (internal to Cartesian coordinates). The gradient and Hessian can be transformed in both directions in a relatively straightforward manner. These transformations have been described in the literature, and here we are following the descriptions from Peng et al.[101] (another excellent publication was written by Bakken and Helgaker[108]). Before describing how the transformations are carried out, it is necessary to introduce *Wilson's* $\mathbf{B}$ *matrix* and the mathematical concept of the *generalized inverse*.

[e] Close here means that if the closest atom pair between A and B has a distance of $d_{0,AB}$, then all atom pairs between A and B with a distance of up to $min(1.3 \times d_{0,AB}, d_{0,AB} + 2\ \text{Å})$ are selected as interfragmental bonds.

[f] Linear bond angles require special attention, since their existence can lead to problems with regard to Wilson's B matrix.[108]

[g] Note that e.g. the pair of bond angles i-j-k and i-j-l would not result in a dihedral angle.

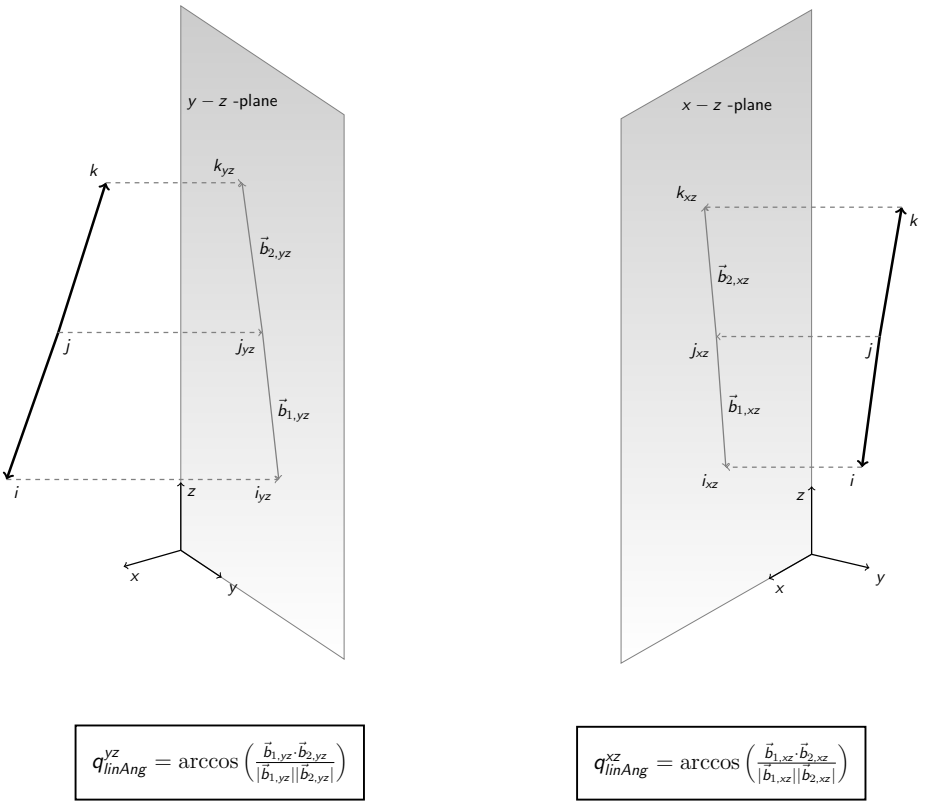


Figure 2.5. – Definition of a *linear angle*. The two planes, on which the atomic positions are projected ($x-y$, $x-z$ or $y-z$ plane), are chosen depending on the position of i, j and k in the Cartesian space. The two planes, on which the vector $(\vec{i}-\vec{k})$ has highest magnitude, are used for the definition of a linear angle. In this figure, these are the $x-z$ and $y-z$ planes.

Generalized Inverse Only for quadratic, nonsingular matrices $\mathbf{A}$ there exists a well-defined inverse matrix $\mathbf{A}^{-1}$. For other types of matrices with linear dependent rows or columns, the usual inverse $\mathbf{A}^{-1}$ is not defined. In such cases, the *generalized inverse* $\mathbf{A}^{-}$ is needed. $\mathbf{A}^{-}$ can also be derived for rank-deficient matrices. Without going into further detail, it is shortly described here at a technical level, how the generalized inverse is generated.[101]

The generalized inverse $\mathbf{A}^{-}$ is obtained by *diagonalization* of $\mathbf{A}$ and *inverting only its nonzero eigenvalues* Λ.

$$\mathbf{V}^{\mathsf{T}}\mathbf{A}\mathbf{V} = \begin{bmatrix}\Lambda & 0\\ 0 & 0\end{bmatrix} \;;\; \mathbf{A}^{-} = \mathbf{V}\begin{bmatrix}\Lambda^{-1} & 0\\ 0 & 0\end{bmatrix}\mathbf{V}^{\mathsf{T}} \;, \tag{2.21}$$

where $\mathbf{V}$ is a matrix containing the eigenvectors of $\mathbf{A}$.[h]

Wilson's $\mathbf{B}$ matrix A relation between displacements in Cartesian coordinates and in redundant internal coordinates is given by Wilson's $\mathbf{B}$ matrix.[i,108] The elements of $\mathbf{B}$ are

[h] For a nonsingular matrix $\mathbf{A}$, the generalized inverse reduces to the regular inverse $\mathbf{A}^{-1}$.

[i] Wilson's $\mathbf{B}$ matrix has its origins in vibrational spectroscopy.

defined as:

$$B_{ij} = \frac{\partial q_i}{\partial x_j} \ , \tag{2.22}$$

with q_i as internal coordinates (bonds, angles and dihedrals) and x_j as Cartesian coordinates (x, y and z coordinate) of the molecular system. $\mathbf{B}$ is a *mxn* matrix, where m is the number of redundant internal coordinates and n is the number of Cartesian coordinates.
Small displacements in Cartesian coordinates can herewith be transformed to small displacements in internal coordinates by simply applying the $\mathbf{B}$ matrix as follows:

$$\partial q = \mathbf{B} \partial x \ . \tag{2.23}$$

The transformation of $\partial q \rightarrow \partial x$ can also be done with Wilson's $\mathbf{B}$ matrix, but here, much more effort has to be invested, as will be described later. The gradient and Hessian can be easily transformed in both directions. Before this can be done, the $\mathbf{G}$ matrix has to be defined, which is the product of the $\mathbf{B}$ matrix and its transpose $\mathbf{B}^\mathsf{T}$,

$$\mathbf{G} = \mathbf{B}\mathbf{B}^\mathsf{T} \ . \tag{2.24}$$

For the following transformations, the generalized inverse $\mathbf{G}^-$ is needed, which is generated from $\mathbf{G}$ as described above.

Transformation of Gradient and Hessian With the help of Wilson's $\mathbf{B}$ matrix and the generalized inverse $\mathbf{G}^-$ the gradient and Hessian can now be transformed from Cartesian to internal coordinates. For the gradient, the relation is

$$\mathbf{g}_q = \mathbf{G}^- \mathbf{B} \mathbf{g}_x \ , \tag{2.25}$$

where $\mathbf{g}_q$ and $\mathbf{g}_x$ are the gradients in internal and Cartesian coordinates, respectively. The internal gradient $\mathbf{g}_q$ contains redundancies (since the internal coordinate set is redundant). Before the geometry step $\Delta\mathbf{s}$ can be calculated, the redundancies have to be removed with the help of a projection operator (vide infra).
From the Hessian matrix, the relation is as follows:

$$\mathbf{H}_q = \mathbf{G}^- \mathbf{B} (\mathbf{H}_x - \mathbf{K}) \mathbf{B}^\mathsf{T} \mathbf{G}^- \ , \tag{2.26}$$

where the matrix $\mathbf{K}$ is the product of the internal gradient and the first derivative of $\mathbf{B}$,[j]

$$\mathbf{K} = \mathbf{g}_q \cdot \mathbf{B}' \ ; \quad K_{jk} = \sum_i [g_q]_i \ B'_{ijk} = \sum_i [g_q]_i \ \frac{\partial^2 q_i}{\partial x_j \ \partial x_k} \ . \tag{2.27}$$

In some cases, it is also necessary to do the reverse transformation, from internal to Cartesian coordinates.[k] The transformation is done by applying the following equations:

$$\begin{aligned} \mathbf{g}_x &= \mathbf{B}^\mathsf{T} \mathbf{g}_q \ , \\ \mathbf{H}_x &= \mathbf{B}^\mathsf{T} \mathbf{H}_q \mathbf{B} + \mathbf{K} \ . \end{aligned} \tag{2.28}$$

[j] It should be noted that the matrix $\mathbf{K}$ is usually set to zero.[109]

[k] The exact Hessian is needed e.g. in a TS optimization. It is calculated in Cartesian space, but used in internal coordinate space.

Projection of Gradient and Hessian After the transformation, both the internal gradient as well as the internal Hessian contain redundant entries. It has to be ensured that the gradient and Hessian act only on the nonredundant part of the internal coordinate space. This can be realized with the help of a projection matrix $\mathbf{P}$, which is constructed from the $\mathbf{G}$ matrix and its generalized inverse $\mathbf{G}^{-}$:

$$\mathbf{P} = \mathbf{G}\mathbf{G}^{-} = \mathbf{G}^{-}\mathbf{G} \ . \tag{2.29}$$

The gradient and the Hessian can then be projected by simply applying $\mathbf{P}$,

$$\begin{aligned} \mathbf{g}_{q,P} &= \mathbf{P}\mathbf{g}_q \ , \\ \mathbf{H}_{q,P} &= \mathbf{P}\mathbf{H}_q\mathbf{P} \ , \end{aligned} \tag{2.30}$$

ensuring that the redundant entries are zero. Having done all this, the geometry step $\Delta\mathbf{s}$ can be determined.

Calculation of the Step in Internal Coordinates With the projected internal gradient and Hessian, the geometry step can be calculated using any of the methods described in Section 2.1.2 – 2.1.4. Applying e.g. the QN step from equation (2.7), we obtain the step

$$\Delta\mathbf{s}_q = -\widetilde{\mathbf{H}}_{q,P}^{-1}\ \mathbf{g}_{q,P} \ . \tag{2.31}$$

bringing us from the actual point in the direction of the stationary point.

Transformation of the Step from Internal to Cartesian Coordinates The energy and gradient of the new geometry can be calculated *in Cartesian space only*. When acting *in the internal coordinate space*, the new geometry $\mathbf{q}_{k+1}$ can be obtained by simply adding the determined geometry step $\Delta\mathbf{s}_q$ to the old geometry $\mathbf{q}_k$. That is, the new geometry $\mathbf{q}_{k+1}$ would have to be converted to Cartesian coordinates ($\mathbf{q}_{k+1} \rightarrow \mathbf{x}_{k+1}$). As already mentioned above, neither the new geometry $\mathbf{q}_{k+1}$ nor the step $\Delta\mathbf{s}_q$ can be transformed to Cartesian coordinates *in a direct manner*. There exists *no unique transformation* from the internal coordinate system to the Cartesian coordinate system.[l]
Instead, the transformation has to be done *iteratively*. A scheme for this transformation is shown in the flowchart in Figure 2.6 and can be described as follows:
A first estimate of the displacement in Cartesian coordinates $\Delta\mathbf{s}_x$ is calculated from $\Delta\mathbf{s}_q$,

$$\Delta\mathbf{s}_x = \mathbf{B}^{\mathsf{T}}\mathbf{G}^{-1}\Delta\mathbf{s}_q \ . \tag{2.32}$$

With this estimate, a *first guess* of the new Cartesian coordinates $\mathbf{x}_1$ can be calculated:

$$\mathbf{x}_1 = \mathbf{x}_{old} + \Delta\mathbf{s}_x \ . \tag{2.33}$$

The quality of this guess $\mathbf{x}_1$ in Cartesian coordinates is now evaluated by calculating the respective internal coordinates $\mathbf{q}_1$ of the guess structure $\mathbf{x}_1$[m] and comparing the *requested step* $\Delta\mathbf{s}_q$ with the *actual step* $(\mathbf{q}_1 - \mathbf{q}_0)$,

$$\Delta\Delta\mathbf{s}_{q,i} = \Delta\mathbf{s}_q - (\mathbf{q}_i - \mathbf{q}_0) \ , \tag{2.34}$$

[l] The reverse conversion from Cartesian to internal coordinates is simple, see equations in Figure 2.4.
[m] This can be done by simply calculating the distances, angles and dihedrals for $\mathbf{x}_1$.

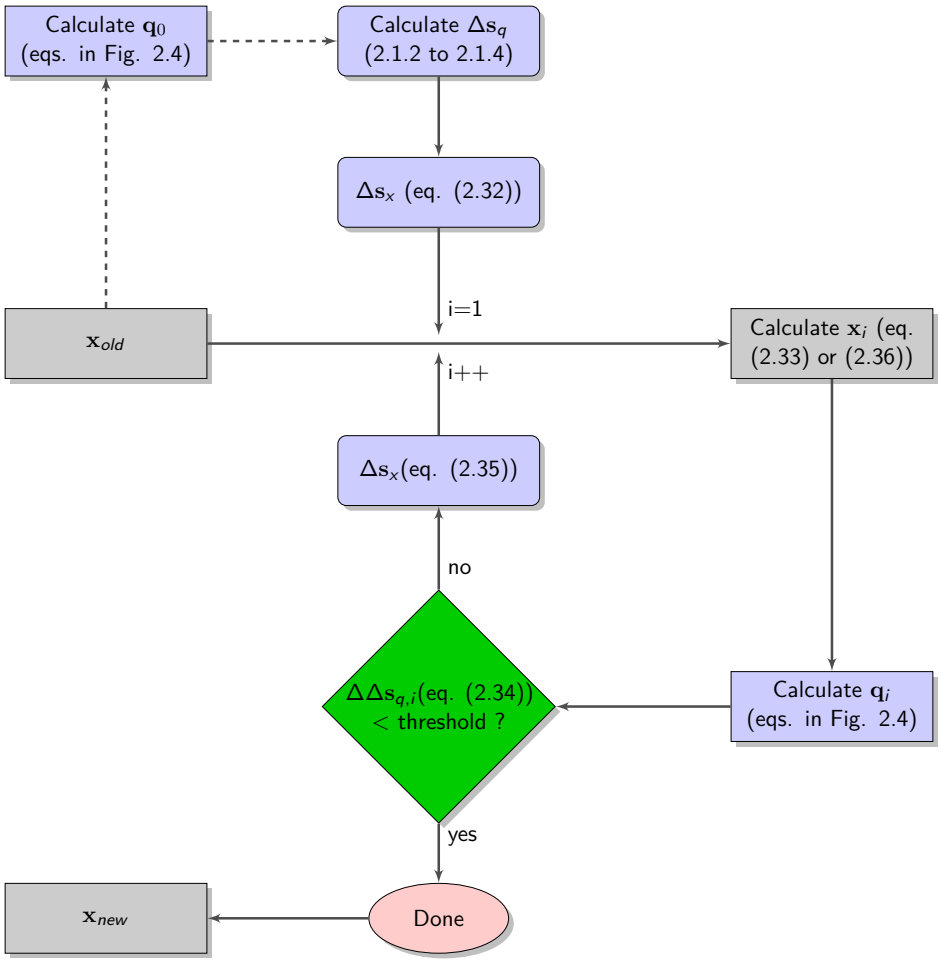


Figure 2.6. – Flowchart for Backtransformation. References to the respective equations and sections (bold) are given. $\mathbf{x}_0$ is the geometry (in Cartesian space) before the step is taken, $\Delta\mathbf{s}_q$ is the determined step which has to be taken.

with $i = 1$ for the first iteration. The difference vector $\Delta\Delta\mathbf{s}_{q,i}$ gives the difference with respect to the requested step, which still needs to be transformed. This is put into the next transformation

$$\Delta\mathbf{s}_{x,i+1} = \mathbf{B}^{\mathsf{T}}\mathbf{G}^{-}\Delta\Delta\mathbf{s}_{q,i} \ , \tag{2.35}$$

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \Delta\mathbf{s}_{x,i+1} \ . \tag{2.36}$$

This procedure is repeated until $\Delta\Delta\mathbf{s}_{q,i}$ converges to zero. In most cases, the backtransformation succeeds after a couple of iteration steps. $\mathbf{x}_{new}$ is the 'new' geometry which corresponds to $\mathbf{q}_k + \Delta\mathbf{s}_q$, and which is used for the next energy and gradient calculation.
If the requested step size $\Delta\mathbf{s}_q$ is large, problems might occur in the convergence of the iterative backtransformation. The problems may be due to the fact that usually the $\mathbf{B}$ matrix is calculated only once for the initial geometry $\mathbf{x}_0$, but used for all subsequent iteration steps in the backtransformation. In principle, $\mathbf{B}$ is not valid for these later occuring $\mathbf{x}_i, i \neq 0$. Recalculation of the $\mathbf{B}$ matrix for new $\mathbf{x}_i$ usually solves the problem.
In rare cases, where also the recalculation of $\mathbf{B}$ does not improve convergence, the (conservative) convergence threshold for $\Delta\Delta\mathbf{s}_{q,i}$ should be decreased.
In the remaining problematic cases, where $\Delta\Delta\mathbf{s}_{q,i}$ diverges in the course of the backtrans-

formation, it is advisable to simply use the very first guess $\mathbf{x}_1$ obtained from equation (2.33).

2.1.6. Initial Hessian

In second-derivative based optimization methods (see Sections 2.1.2, 2.1.3 and 2.1.4), the Hessian matrix is needed for the determination of the geometry step $\Delta\mathbf{s}$.
Unfortunately, it is quite expensive to calculate the Hessian exactly. To circumvent the exact calculation of the force constants, it is a commonly used technique to use an approximate Hessian at the beginning of the calculation, and to use an updated Hessian at subsequent optimization steps.

Initial Hessian for Minimization

For minimizations, only the positive Hessian eigenvalues are needed, regardless of the curvature of the PES at the actual configuration. Initial Hessians for minimization usually consist of empirically estimated force constants. At this point, an advantage of internal coordinates comes into play. Approximate Hessians can easily be built up in internal coordinates, since experimental force constants are well known for bond stretching, bending and dihedral torsion e.g. from infrared spectroscopy. Such a *model force-field* Hessian $\widetilde{\mathbf{H}}_q^M$ is constructed as a diagonal Hessian (see equation (2.37)),

$$\widetilde{\mathbf{H}}_q^M = \begin{pmatrix} f_{b,1} & & & & & & & & 0 \\ & \ddots & & & & & & & \\ & & f_{b,m} & & & & & & \\ & & & f_{a,1} & & & & & \\ & & & & \ddots & & & & \\ & & & & & f_{a,k} & & & \\ & & & & & & f_{d,1} & & \\ & & & & & & & \ddots & \\ 0 & & & & & & & & f_{d,k} \end{pmatrix}, \tag{2.37}$$

with $f_b/f_a/f_d$ being the estimated force constant for the m bonds/k angles/l dihedrals. The off-diagonal force constants are zero, which means that cross-coupling between different internal coordinates is neglected. Several empirical formulas for approximate nuclear harmonic force constants exist (e.g. Almlöf,[110] Lindh[111] or Schlegel[109]). They usually depend on the type of internal coordinate (bond length, bond angle, dihedral angle), on the types of atoms involved, and on the value of the internal coordinate.

Initial Hessian for TS Optimization

In TS optimization, the most crucial point is the initial guess structure. The TS optimization should start in the proximity of the TS, i.e. in a region where the PES shows already some curvature to the right with respect to the TS vector.[n] A simple model Hessian based on empirical force constants contains only positive harmonic force constants (also at near-TS molecular configurations). With such an eigenvalue composition even the mode following

[n] This can usually be realized by performing a relaxed surface scan beforehand.

method of the P-RFO algorithm (see Section 2.1.4) is usually not able to succeed in finding the TS. A better Hessian is needed.
The most straightforward way for TS optimization is to *calculate the exact Hessian*. Diagonalizing this Hessian yields the TS eigenvector with its negative eigenvalue, and the exact positive force constants for all other eigenmodes (depending on the 'quality' of the initial guess structure).[o] The calculation of the exact Hessian is very time consuming, and normally it is not necessary to know the exact Hessian for the *whole* molecular system. Most often, a TS eigenvector shows some locality, i.e. it represents the coupling between only a few atoms out of the whole molecular system. One can take advantage of this in two different ways, described in the following.

Modification of Diagonal Hessian Values In the internal coordinate space most TS modes can be described by only a few coordinates. To illustrate this, consider the example of a hydrogen atom transfer process. Here, a hydrogen atom (H) is transferred from a donor (D) to an acceptor (A). The majority of the changes in the course of the reaction occur in the two bond coordinates D-H and A-H (the bond distance of D-H is increasing, while the bond distance of A-H is decreasing). The TS vector, which corresponds to this reaction, is more or less fully described by those two bond coordinates.
The most simple approach to construct a Hessian matrix, which incorporates the negative force constant for the selected internal coordinates, is the *modification of selected diagonal entries in an approximate Hessian matrix*. First the model force-field Hessian $\widetilde{\mathbf{H}}_q^M$ is built up in the usual way. Then the selected diagonal entries are overwritten, either by a user-defined value (good values here are -0.3 Hartree/bohr^2 for bonds and -0.1 Hartree/bohr^2 for bond angles and dihedrals), or by a more exact value which is calculated from a difference gradient calculation of two neighboring molecular structures from a relaxed surface scan. This procedure can be applied for all types of internal coordinates.

Hybrid Hessian When the TS mode is not localized on one or two internal coordinates, another approach is available: the *hybrid Hessian*. The basic Hessian matrix is again the model force-field Hessian. In this approach, the modification of the Hessian is not done for selected internal coordinates, but for a selected *range of atoms*.
For the construction of the hybrid Hessian $\widetilde{\mathbf{H}}^H$ the beforehand generated $\widetilde{\mathbf{H}}_q^M$ matrix is first transformed from internal to Cartesian coordinate space ($\widetilde{\mathbf{H}}_q^M \rightarrow \widetilde{\mathbf{H}}_x^M$, see equation (2.28)). Then, the exact Hessian values are calculated for the selected atoms (i.e. for the respective rows and columns of the selected atoms) in Cartesian space. The hybrid Hessian is finally built up in the following manner:

$$\widetilde{\mathbf{H}}_x^H = \left(\begin{array}{ccc|ccc} H_x^M(1,1) & \dots & H_x^M(1,i) & H_x^E(1,j) & \dots & H_x^E(1,n) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ H_x^M(i,1) & \dots & H_x^M(i,i) & H_x^E(i,j) & \dots & H_x^E(i,n) \\ \hline H_x^E(j,1) & \dots & H_x^E(j,i) & H_x^E(j,j) & \dots & H_x^E(j,n) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ H_x^E(n,1) & \dots & H_x^E(n,i) & H_x^E(n,j) & \dots & H_x^E(n,n) \end{array}\right) . \qquad (2.38)$$

[o] Usually some other eigenmodes also show (low) negative eigenvalues.

For the purpose of displaying it is assumed that the selected range of atoms (E for exact treatment, atoms 1 to i) is arranged as a block after all other atoms (M for model force-field treatment, atoms j to n). Note that the – in internal coordinates diagonal – model force-field Hessian has non-zero off-diagonal elements in the Cartesian representation.
We can identify four blocks in the hybrid Hessian matrix:

$$\widetilde{\mathbf{H}}_x^H = \left(\begin{array}{c|c} \mathbf{H}_x^M & \mathbf{H}_x^E \\ \hline \mathbf{H}_x^E & \mathbf{H}_x^E \end{array} \right) ,$$

The coupling of the i M-atoms is approximated at the model force-field level and the coupling of the $n - i$ E-atoms is evaluated at the exact level. Furthermore, the off-diagonal blocks $D - E$ and $E - D$ are treated exactly. Usually, the M group is much larger than the E group. If the Hessian is calculated by numerical differentiation, the calculation time is directly proportional to the number of E atoms. Thus, applying this approach can save considerable time compared to a full exact Hessian calculation.
After the construction of the hybrid Hessian $\widetilde{\mathbf{H}}_x^H$ in Cartesian coordinate space, it is transformed back to the internal coordinate space ($\widetilde{\mathbf{H}}_x^H \rightarrow \widetilde{\mathbf{H}}_q^H$), where the P-RFO step is taken (see section 2.1.5).

In rare, complicated cases, where these two approximate Hessians do not lead to the desired TS structure, either the initial structure is too bad and a new one has to be generated, or a full exact Hessian is needed.

Hessian Update

The curvature of the PES is different for each molecular configuration of atoms in the same molecule. After an optimization step, the old Hessian is not valid any more for the new geometry (neither the exact, nor the approximate ones). The Hessian at a new configuration k can be *updated* from the Hessian at the old configuration $k-1$ by including the information on the actual step $\Delta\mathbf{s}$ and on the difference gradient between last and actual step ($\Delta\mathbf{g}_k = \mathbf{g}_k - \mathbf{g}_{k-1}$).

- If the initial Hessian was exactly calculated (e.g. at the beginning of a TS optimization), the Hessian update ensures that the Hessian maintains its quality (provided that the steps are not too large → Section 2.1.3).

- If the force constants were estimated, then the Hessian update improves the quality of the Hessian with each optimization step, since the *estimated* curvature of the Hessian from the first step is updated with information about the *real* curvature of the PES.

There exist several Hessian update schemes, and they all have different advantages. The two update schemes from Broyden, Fletcher, Goldfarb and Shanno (BFGS) and from Bofill are mentioned here.

- The BFGS update[112] (given in equation (2.39)) is an update scheme which guarantees a positive definite Hessian. It is thus perfectly suited for minimizations.

$$\mathbf{H}_k^{\text{BFGS}} = \mathbf{H}_{k-1} + \frac{\Delta\mathbf{g}_k \Delta\mathbf{g}_k^\mathsf{T}}{\Delta\mathbf{g}_k^\mathsf{T} \Delta\mathbf{x}_k} - \frac{\mathbf{H}_{k-1} \Delta\mathbf{x}_k \Delta\mathbf{x}_k^\mathsf{T} \mathbf{H}_{k-1}}{\Delta\mathbf{x}_k^\mathsf{T} \mathbf{H}_{k-1} \Delta\mathbf{x}_k} \tag{2.39}$$

- The Bofill update[113] (given in equation (2.40)) does, on the other hand, not enforce a positive definite matrix. It is thus applicable for TS optimizations.

$$\mathbf{H}_k^{\text{Bofill}} = (1-\phi_k)\mathbf{H}_k^{\text{MS}} + \phi_k \mathbf{H}_k^{\text{PSB}} \tag{2.40}$$
$$\phi_k = 1 - \frac{(\Delta\mathbf{x}_k^{\mathsf{T}}\mathbf{F}_k)^2}{(\Delta\mathbf{x}_k^{\mathsf{T}}\Delta\mathbf{x}_k)(\mathbf{F}_k^{\mathsf{T}}\mathbf{F}_k)} \;;\quad \mathbf{F}_k = \Delta\mathbf{g}_k - \mathbf{H}_{k-1}\Delta\mathbf{x}_k \;.$$

$$\mathbf{H}_k^{\text{MS}} = \mathbf{H}_{k-1} + \frac{(\Delta\mathbf{g}_k - \mathbf{H}_{k-1}\Delta\mathbf{x}_k)(\Delta\mathbf{g}_k - \mathbf{H}_{k-1}\Delta\mathbf{x}_k)^t}{(\Delta\mathbf{g}_k - \mathbf{H}_{k-1}\Delta\mathbf{x}_k)^t\Delta\mathbf{x}_k} \;,$$

$$\mathbf{H}_k^{\text{PSB}} = \mathbf{H}_{k-1} + \frac{(\Delta\mathbf{g}_k - \mathbf{H}_{k-1}\Delta\mathbf{x}_k)\Delta\mathbf{x}_k^{\mathsf{T}} + \Delta\mathbf{x}_(\Delta\mathbf{g}_k - \mathbf{H}_{k-1}\Delta\mathbf{x}_k)^{\mathsf{T}}}{\Delta\mathbf{x}_k^{\mathsf{T}}\Delta\mathbf{x}_k} - \frac{(\Delta\mathbf{x}_k^{\mathsf{T}}\Delta\mathbf{g}_k - \Delta\mathbf{x}_k^{\mathsf{T}}\mathbf{H}_{k-1}\Delta\mathbf{x}_k)\Delta\mathbf{x}_k\Delta\mathbf{x}_k^{\mathsf{T}}}{(\Delta\mathbf{x}_k^{\mathsf{T}}\Delta\mathbf{x}_k)^2}$$

2.1.7. Constrained Optimization

A constrained optimization is an optimization with one or several internal coordinates fixed to a certain value. All types of internal coordinates (as well as Cartesian coordinates) can be constrained. This method is useful for a variety of applications: Exploration of a PES, the search for a good starting structure for TS optimization, refinement of crystal structures (by keeping the positions of the heteroatoms constrained and optimizing only the hydrogen positions), etc.
The constraints have to be included in two steps of the geometry optimization scheme:

1. in the projection of the internal gradient and Hessian and

2. after the update of the Cartesian geometry.

Projection

The gradient and Hessian are projected before the step $\Delta\mathbf{s}$ is determined (see equation (2.30)). Constraints can be inserted by simply modifying the projector (equation (2.29)). The 'constrained' projector $\mathbf{P}'$ is derived from the 'normal' projector $\mathbf{P}$ in the following manner:

$$\mathbf{P}' = \mathbf{P} - \mathbf{PC}(\mathbf{CPC})^{-1}\mathbf{CP} \;. \tag{2.41}$$

Here, the matrix $\mathbf{C}$ contains ones on the diagonal for constrained coordinates and zeros elsewhere. This modified projector is then applied on the gradient and the Hessian matrix:

$$\tilde{\mathbf{g}}_{q,P} = \mathbf{P}'\mathbf{g}_q \;, \tag{2.42}$$
$$\widetilde{\mathbf{H}}_{q,P} = \mathbf{P}'\mathbf{H}\mathbf{P}' + \alpha(\mathbf{1} - \mathbf{P}') \;, \tag{2.43}$$

where α is a large constant (e.g. in our implementation in ORCA: 1000). The large constant ensures that the matrix elements of the Hessian, belonging to constrained coordinates, are set to high values. The projection of the gradient removes the gradient contribution for the constrained internal coordinates, which prevents displacements in the respective coordinates.

Geometry Update Step

Sometimes it is not sufficient to modify the gradient and Hessian. Small changes in the constrained coordinates might be introduced in the course of the backtransformation. To ensure that the constraints are still present after the backtransformation and before the next energy and gradient calculation, an *additional backtransformation* is performed, which again imposes the constraints. The obtained configuration x_{new} (from equation (2.36)) is therefore first transformed to internal space (q_{new}). The deviation to the requested imposed constraints is calculated as:

$$\Delta\Delta\mathbf{s}_q^C = \mathbf{C}(\mathbf{q}_{new} - \mathbf{q}_C) \ , \tag{2.44}$$

where $\mathbf{q}_C$ is a vector with the constraint values for the respective coordinates. Multiplication by $\mathbf{C}$ ensures that only the deviations of the constrained coordinates are included in $\Delta\Delta\mathbf{s}_q^C$. This displacement vector $\Delta\Delta\mathbf{s}_q^C$ is transformed to the Cartesian coordinate space as already described in equations (2.32) to (2.36), while it is always ensured (by applying the matrix $\mathbf{C}$) that solely the constraints, and no other coordinates, are reimposed:

$$\Delta\mathbf{s}_x^C = \mathbf{B}^\mathsf{T}\mathbf{G}^-\Delta\Delta\mathbf{s}_{q,k}^C \ , \tag{2.45}$$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta\mathbf{s}_x^C \ . \tag{2.46}$$

2.2. QM/MM Methodology

The QM/MM approach was already introduced in Section 1.1.2. It combines the accurate description of the protein's active site by quantum chemical methods with the efficient description of the remaining part of the system by classical mechanics. Since its first implementation in 1976 the method has been developed further and used in a wide range of applications.[35, 114–116]
In this section, different aspects of the QM/MM methodology are discussed in detail. First, the theoretical background of the method is presented, starting with the division into separate subsystems to different coupling schemes for the determination of the QM/MM energy, including the basic equations. Excellent reviews on this topic exist,[34, 114, 116, 117] and we follow in this section to a large part the descriptions by Senn and Thiel.[117]
The theoretical part is followed by a description of the interface between the MM and QM program packages GROMACS and ORCA, which was implemented as a part of this thesis. Finally, a detailed workflow for a standard protein QM/MM calculation is given.

2.2.1. Subsystems

The entire system $\mathbb{S}$ is divided into the *smaller inner subsystem* $\mathbb{I}$ (or *QM subsystem*), which is handled at the higher level of theory, and the *larger outer subsystem* $\mathbb{O}$ (or *MM subsystem*), which is described at the lower level of theory. (The division of different subsystems is illustrated in Figure 2.7.) The boundary region between both is only loosely defined, and it represents here the junction between inner and outer subsystem. An ideal division is met,

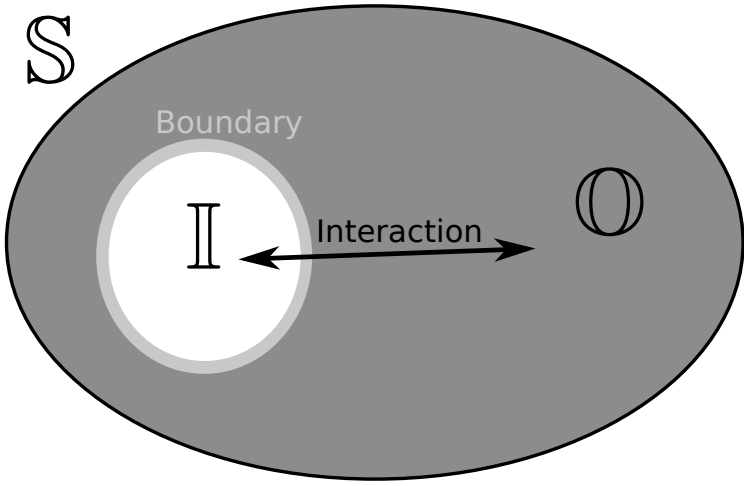


Figure 2.7. – Illustration of the division of the total system ($\mathbb{S}$) into inner ($\mathbb{I}$) and outer ($\mathbb{O}$) subsystems.

when inner and outer subsystem are distinct molecules or groups of molecules (e.g. a solute within an explicit solvent). In this case, the boundary is quasi non-existent. In enzymatic QM/MM systems the situation is far from ideal. Both subsystems are usually connected by one or several covalent bonds. Here, dividing the system means cutting off one or several covalent bonds, which dramatically perturbs the electronic structure of the QM subsystem (e.g. a homolytic cut of a non-polar C-C bond would generate a carbon radical with totally different properties).
In order to circumvent this problem, there exist several techniques.[117] The most widely

spread technique is the *link atom (LA) approach*.[a] Here, usually non-polar bonds, like C-C bonds (in which the C atoms have sp^3 hybridization), are cleaved. The free valencies of the cleaved bond are saturated by a hydrogen atom on the side of the QM subsystem. These so-called link atoms (also called $\mathbb{L}$ in the following) are added to the primary subsystem, but do not belong to the entire system. They saturate the free valencies, and their energy and gradient are distributed onto the atoms of the cleaved bonds. (More information on link atoms is given in Section 2.2.4.)
Usually, the subsystems are statically defined before the calculation. But it should be mentioned that there exist also QM/MM schemes with a dynamic division. In such schemes atoms belonging to the outer subsystem can, in the course of the simulation, be shifted to the primary subsystem and vice versa ('adaptive QM/MM').[116]

2.2.2. QM/MM Schemes

The QM/MM potential energy function is not simply the sum of the energies of inner and outer subsystem, but it also contains coupling terms between both subsystems. There exist two different definitions for the QM/MM energy, the subtractive and additive QM/MM schemes.

Subtractive QM/MM Scheme

The major equation of the subtractive QM/MM scheme (with the link atom approach) is:

$$E_{QM/MM}(\mathbb{S}) = E_{MM}(\mathbb{S}) + E_{QM}(\mathbb{I}+\mathbb{L}) - E_{MM}(\mathbb{I}+\mathbb{L}) \ , \tag{2.47}$$

with the subscript indicating the level of theory at which the energy calculation is performed. $\mathbb{I}+\mathbb{L}$ designates the 'capped inner subsytem' (inner subsystem saturated with link atoms). If no covalent bonds connect the inner and outer subsystem, then $\mathbb{I}+\mathbb{L}$ reduces to $\mathbb{I}$. Using the subtractive QM/MM scheme, three calculations must be carried out:

- An MM calculation on the entire system $\mathbb{S}$,
- a QM calculation on the capped inner subsystem $\mathbb{I}+\mathbb{L}$ and
- an MM calculation on the capped inner subsystem $\mathbb{I}+\mathbb{L}$.

The subtractive scheme can be regarded as an MM calculation, from which the inner region is cut out and is instead treated at a higher level of theory. This approach is simple in terms of implementation, since only pure QM and MM calculations, but no extra QM-MM coupling calculations, are needed.
A subtractive scheme is handled correctly only in the mechanical embedding scheme (*vide infra*), and suffers the drawbacks of this method.

Additive QM/MM Scheme

The basic equation for the additive scheme is given here:

$$E_{QM/MM}(\mathbb{S}) = E_{MM}(\mathbb{O}) + E_{QM}(\mathbb{I}+\mathbb{L}) + E_{QM-MM}(\mathbb{I},\mathbb{O}) \ . \tag{2.48}$$

The three calculations, which have to be carried out here, are:

[a] Other approaches are: Localized-orbital schemes, boundary-atom schemes, etc.

- An MM calculation on the outer subsystem $\mathbb{O}$,
- a QM calculation on the capped inner subsystem $\mathbb{I} + \mathbb{L}$ and
- a calculation of the interaction between the two subsystems $\mathbb{I}$ and $\mathbb{O}$.

In the additive scheme, the MM calculation is only carried out for the outer subsystem. The third term, $E_{QM-MM}(\mathbb{I}, \mathbb{O})$, sums up the *coupling terms between inner and outer subsystem*. The different coupling terms can be calculated at different levels, which is described in the following section.[b]

2.2.3. QM-MM Interaction

The QM-MM interaction term usually includes bonded and nonbonded interaction contributions:

$$\begin{aligned} E_{QM-MM}(\mathbb{I}, \mathbb{O}) &= E^{bonded}_{QM-MM} + E^{nonbonded}_{QM-MM} \\ &= E^{bonded}_{QM-MM} + E^{el}_{QM-MM} + E^{VDW}_{QM-MM} \ . \end{aligned} \tag{2.49}$$

The bonded terms E^{bonded}_{QM-MM} and the van der Waals interaction E^{VDW}_{QM-MM} between inner and outer subsystem are usually treated at the MM level. The electrostatic coupling E^{el}_{QM-MM} can be handled at *different levels of sophistication*, which gives rise to three different embedding schemes. In this subsection, the three different embedding schemes for the electrostatic QM-MM interaction are presented, before the other terms (E^{bonded}_{QM-MM} and E^{VDW}_{QM-MM}) are discussed.

Electrostatic QM-MM Interaction

The different schemes of the electrostatic QM-MM interaction are characterized by the *extent of mutual polarization*. One distinguishes between:

- Mechanical embedding,
- electrostatic embedding and
- polarized embedding.

Mechanical Embedding In mechanical embedding (ME) the QM-MM electrostatic interaction is treated at the same level as the MM-MM electrostatics. This is typically an interaction between rigid atomic point charges (see Figure A.1 in the Appendix). Thus, fixed point charges have to be assigned not only to the MM subsystem, but also to the atoms of the QM subsystem. This is the source of a number of problems:

- In the course of a chemical process (e.g. a reaction) the charge distribution of the QM subsystem can change, which cannot be represented well by fixed atomic charges. This problem can be circumvented by updating the atomic charges during the calculation, although this has again the drawback that discontinuities are introduced into the potential energy function of the entire system.

[b] Usually no extra calculation is carried out for the QM-MM interaction energy $E_{QM-MM}(\mathbb{I}, \mathbb{O})$, but it can instead be included in one or both of the two other calculations.

- The charge density of the active site is 'in reality' influenced by the charges of the whole protein. This polarization effect is not taken into account in the ME scheme.
- The force field atomic charges and other force field parameters are interdependent and are optimzed to yield a balanced description of the entire system. They are not designed as a *true* charge distribution.

In general, the ME scheme should not be used if the electron density of the inner subsystem is significantly polarized by the charge distribution of the outer subsystem.

Electrostatic Embedding The major drawbacks of the ME scheme can be eliminated using the electrostatic embedding (EE) scheme. Here, the QM-MM electrostatic interaction is not calculated at the MM level, but at the QM level. Therefore, the fixed atomic point charges of the outer subsystem are included in the QM calculation by adding terms to the one-electron Hamiltonian describing the interaction between the MM point charges and the QM nuclei or the QM electrons, respectively:

$$\widetilde{H}^{el}_{QM-MM} = -\sum_{i}^{N_e} \sum_{M\in\mathbb{O}} \frac{q_M}{|\mathbf{r}_i - \mathbf{R}_M|} + \sum_{\alpha\in\mathbb{I}+\mathbb{L}} \sum_{M\in\mathbb{O}} \frac{q_M Z_\alpha}{\mathbf{R}_\alpha - \mathbf{R}_M} \ , \tag{2.50}$$

where q_M are the MM point charges, Z_α the nuclear charges, N_e is the number of electrons, M runs over all MM point charges and α over all QM nuclei, including the link atoms.
In such a scheme, the QM atoms do not need to carry fixed point charges for the calculation of the electrostatic interaction at the MM level. Most importantly, the electron density of the QM subsystem adapts automatically to the charge distribution of the outer subsystem. The polarization of the QM subsytem thus depends not only on the QM atoms themselves, but also on the spatial distribution and magnitude of the MM point charges.
There are still small problems with the EE scheme:

- The problem of *overpolarization* at the boundary, when link atoms come too close to MM point charges (see Section 2.2.4).
- MM atomic point charges (which are assigned to the MM subsystem and act on the QM subsystem) are not designed to mimic a *real* charge distribution perturbing the QM subsystem. Nevertheless, EE is shown to yield reasonable results regarding the polarization effect of the surrounding environment.

EE is the most popular QM/MM approach used nowadays.[117] It allows to treat the electronic structure of the active site correctly while incorporating the long-range electrostatic effect of the protein and its solvent, concomitantly with the steric constraints induced by the protein.

Polarized Embedding Taking the extent of mutual polarization one step further leads to the polarized embedding (PE) scheme.[116] Here, not only the QM density is polarized by the charge distribution of the outer subsystem, but also the MM charge distribution of the outer subsystem is polarized by the QM density. This procedure can be accomplished by using a polarizable force field. Two approaches exist here:

- In the simpler one, the polarizable MM region is polarized by the QM charge density.
- In the more sophisticated approach, QM and MM charge distributions are mutually polarized in a self-consistent manner.

PE is not widely spread due to its technical complexity, higher computational cost and the fact that only few polarizable force fields are currently available.

Other QM-MM Interactions

The other QM-MM interactions, bonded and VDW, are always handled at the MM level.

QM-MM van der Waals Interaction E^{VDW}_{QM-MM} is evaluated at the MM level and is usually described by a Lennard-Jones type potential. This means that for the atoms of the QM subsystem suitable parameters have to be provided. Here, similar problems arise as within the ME scheme for the fixed atomic charges of the QM atoms:

- QM atoms can change their character in the course of a reaction, and consequently previously assigned LJ coefficients are not suitable any more.
- There might be no atom types and thus no standard parameter set available for non-standard functional groups[c] in the inner subsystem.

Due to the short-range nature of the VDW interaction (see also Figure A.1 in the Appendix), these problems can easily be smoothed out. Atoms which are very close to the boundary contribute most to the QM-MM VDW interaction, whereas atoms in the second or third shell below the boundary contribute only marginally. That is, if the chosen QM region is large enough, the LJ coefficients of the nonstandard or chemically changing atoms in the *inner part* of the QM region have no influence on the outcome of the calculation.

QM-MM Bonded Interactions The bonded QM-MM interactions (bond stretching, angle bending and torsional; see also Figure A.1) are only calculated for atoms in the boundary region. Again, the same arguments against the MM parameters (vide supra) prevail. But, also in this case, the solution is pragmatic. The QM-MM bonded interactions occur strictly only for atoms in the boundary region. If the QM region is large enough, these interactions do not influence the behavior in the critical inner part of the QM region.

2.2.4. Link Atoms

Link atoms are used to saturate free valencies (of the frontier QM atoms) that arise, when QM and MM subsystems are separated. Link atoms are used only in the QM calculation, but not in the MM calculation.

Location of the Link Atom

The placement of a link atom in the boundary region is depicted in Figure 2.8. The division between QM and MM subsystem runs between the atoms $Q1$ and $M1$. That is, the bond between $Q1$ and $M1$ is cleaved. The link atom L saturates the free valencies of the $Q1$ atom. There exist various ways to define the position of the link atom, and only the most common approach is described here:[116]

The link atom is placed on the line connecting $Q1$ and $M1$. Its exact placement on the

[c] Nonstandard groups in contrast to the usual *standard groups* (i.e. amino acids) found in proteins.

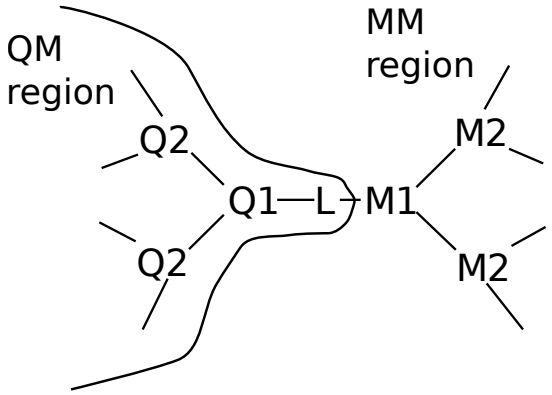


Figure 2.8. – Position of the link atom in the boundary region between the inner and outer subsystem. The bond between atoms $Q1$ and $M1$ is cleaved and saturated by the link atom L. QM atoms directly bound to $Q1$ are denoted $Q2$, MM atoms directly bound to $M1$ are denoted $M2$.

line is defined by the ratio of the 'standard' bond lengths[d] d_0 of $Q1$ and hydrogen and of $Q1$ and $M1$:

$$d(Q1, L) = \frac{d_0(Q1, H)}{d_0(Q1, M1)} \cdot d(Q1, M1) \ , \tag{2.51}$$

where $d(Q1, M1)$ is the actual bond length of $Q1$ and $M1$. Most often, the distance of $Q1$ and $M1$ is constrained to $d_0(Q1, M1)$, and consequently the distance of the link atom to $Q1$ is equal to $d_0(Q1, H)$.

Overpolarization

The above definition for the $Q1$-L distance ensures that the $Q1$ atom does not have a much different character than in the 'real' system. At the same time, this approach places the link atom in very close proximity to the frontier MM atom. Let us illustrate this with a simple example: If we assume a standard bond length of 1.09 Å for a C-H bond and 1.53 Å for an aliphatic C-C bond, then it easily follows that the L-$M1$ distance is only about 0.44 Å. This short distance leads to the so-called *overpolarization problem*, when applying the electrostatic embedding scheme. Here, the whole QM subsystem (and thus also the link atom) is surrounded by the field of the MM point charges (to which $M1$ belongs). Depending on the magnitude of the $M1$ charge and on the size of the basis set,[e] electron density accumulates at the site of $M1$, which is an undesired phenomenon.

Charge Shifting There are several ways to get rid of the overpolarization problem.[119] The simplest approach is to just delete the charge of $M1$. A much better way is to delete the charge on $M1$ and distribute it equally onto the $M2$ atoms. This has the big advantage that it preserves the charge of the overall system.

[d] With respect to the applied force field.

[e] Large basis sets with polarization and diffuse functions are more prone to overpolarization than small basis sets.[118]

2.2.5. QM/MM Interface for ORCA and GROMACS

As part of this thesis an interface of GROMACS to several QM programs, which was already present in the GROMACS source code, was extended to be operative with ORCA. Using this interface, QM/MM calculations (QM/MM optimization or QM/MM MD simulations) can be performed with GROMACS in conjunction with ORCA.
Here, GROMACS is the 'main driver' of the calculation[f], and it requests, in each optimization or MD step, the necessary energy and gradient calculation from ORCA.
In Figure 2.9, the flowchart of an ORCA/GROMACS QM/MM geometry optimization is outlined. The user starts the QM/MM calculation by providing GROMACS with

- general information on the calculation,
- the coordinates and the topology of the system and
- the specific input for the QM calculation (filename ending <.ORCAINFO>).

GROMACS then prepares the optimization step by

- calculating the interactions within the MM subsystem itself,
- calculating the (MM-level) interactions between QM and MM subsystem and
- writing input files for ORCA and starting the QM calculation.

With respect to the QM calculation, there are two possibilities (in the framework of electrostatic embedding):

Conventional optimization The input files for ORCA contain information about the type of calculation[g] (see ORCAINFO-file), the QM atoms (coordinates, atomic numbers) and the MM atoms (coordinates and atomic charges). ORCA performs an energy and gradient calculation.

Microiterative optimization additionally, the input files contain the Lennard Jones coefficients of the QM and MM atoms. ORCA performs an optimization of the QM subsystem (with frozen boundary atoms, i.e. constrained $Q1$ and L atoms) *within the electrostatic and 'Lennard Jones' field of the MM subsystem*. Finally, ORCA performs an energy and gradient calculation.

The collection of data by GROMACS depends on the type of optimization:

- For conventional optimization, GROMACS collects all energies and forces and performs a geometry step.
- For microiterative optimization GROMACS collects all energies, forces *and the new geometry of the QM subsystem* and performs a geometry step.

The microiterative approach has the advantages that

- the faster optimization algorithms in ORCA can be used and

[f] Main driver means that data handling for the entire system and the optimization or simulation steps are carried out by GROMACS.
[g] Method, basis set, etc.

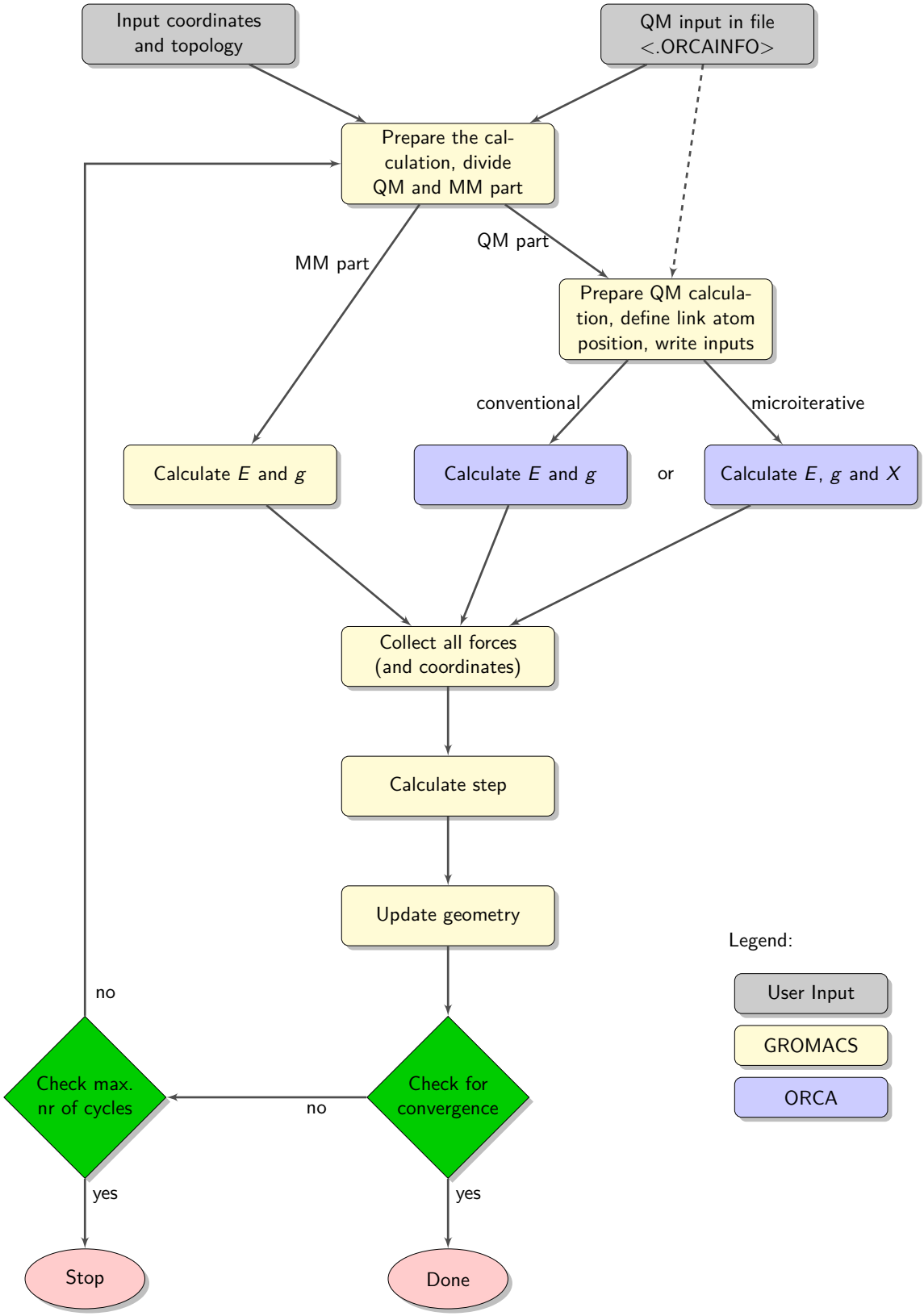


Figure 2.9. – Flowchart of a QM/MM optimization using the interface between GROMACS and ORCA. (E: energy, g: forces.)

- that TS optimization can be performed (which is only possible with ORCA, and not with GROMACS).

Furthermore, it allows the use of different convergence thresholds for the two subsystems (e.g. loose convergence thresholds for the MM atoms and tight thresholds for the QM atoms). After the step performed by GROMACS, the program checks for convergence and proceeds accordingly.

2.2.6. Standard Workflow for a QM/MM Calculation

Performing a QM/MM calculation requires multiple steps of preparation. The crystal structure has to be checked for errors, MM parameters have to be provided, the system has to be solvated, etc. In this section, a workflow is given describing the different steps of a QM/MM calculation, starting from the initial structure preparation to the final property calculations on the optimized geometries. This workflow, which was also followed when working on the projects presented in Chapters 3 and 4, is valid for a QM/MM calculation within the electrostatic embedding scheme using GROMACS and ORCA with the above described interface. The boundary QM atoms are saturated with link atoms and overpolarization is prevented by using the charge distribution scheme.
The steps of the workflow are outlined in Figure 2.10 and are described, in a quite technical way, in the following.

(1) Initial Coordinates

The initial coordinates of a protein are taken from its PDB-file. These coordinates usually come from an X-ray structure analysis. The crystal structure can contain cofactors and mutated residues. It might also contain irregularities, if it was not derived with enough care from experiment. Thus, first the crystal structure should be checked for anomalies (for this purpose there exist several programs, e.g. PROCHECK[120]). Then, residues which do not have the correct residue type[h] can be mutated *in silico*. Other, non-proteic parts of the structure can also be changed at this point:

- Cofactors can be modified, removed or added (e.g. by 'protein-ligand docking'),
- reactants can be introduced or exchanged and
- non-native ions or solvent molecules, which are present for the purpose of crystallization, can be removed.

(2) Force Field and Topologies

In a QM/MM calculation the interactions described at the MM level are calculated using the parameters from a classical force field (see Appendix A). The *topology* of a system collects the information on the actual *atom types* and *connectivities* of the molecule. The topology of the *entire system* must be available during a QM/MM calculation, although the accuracy of the force-field parameters of the atoms in the *inner part* of the QM subsystem is of minor importance (see also Section 2.2.3).
Usually, the topology can be generated with the GROMACS tool 'pdb2gmx':

[h] This might be the case if one wants to study the wild type protein structure, but the crystallized protein is a mutant protein; or if the wild type protein structure is available, but the interest of the calculation lies in a mutant structure.

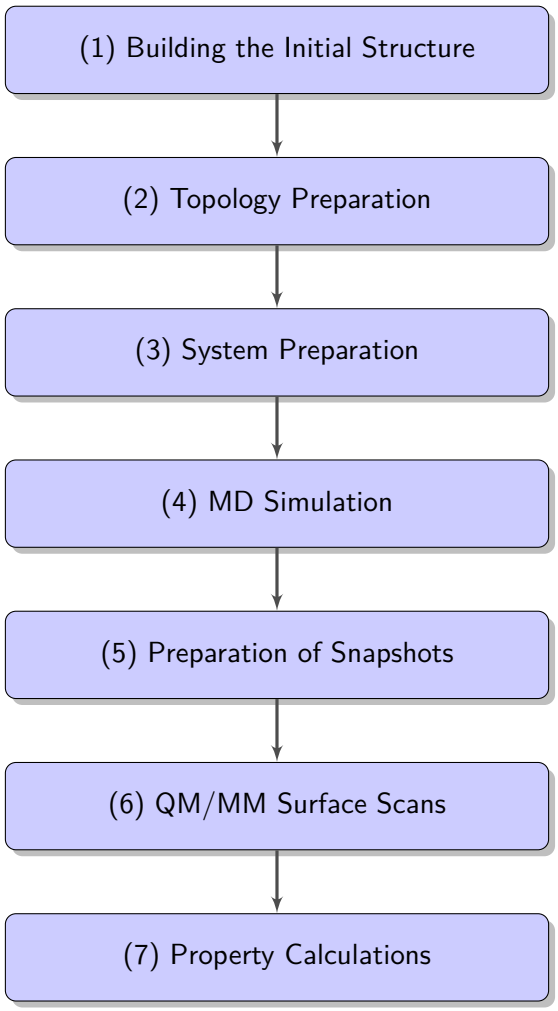


Figure 2.10. – Flowchart showing the different steps necessary for a QM/MM Calculation.

```
pdb2gmx -f inputProt.pdb -p topol.top -o outputProt.g96
```

Here <inputProt.pdb> is the file containing the crystal structure, in <topol.top> the topology of the system is stored and <outputProt.g96> contains the structural data in GROMACS format.
'Non-standard' residues of the system[i] cannot be processed with 'pdb2gmx'. Their topology has to be generated in another way. Here, e.g. the program 'PRODRG'[121] can be used to generate the atom types and connectivities (see example in Figure 2.11). Since PRODRG only roughly estimates atomic point charges, a more accurate description of the non-standard MM point charges should be obtained. The CHELPG charges[122] can be easily calculated with ORCA.[99]
Crystal structures normally come without hydrogen positions. Another highly useful feature

[i]Non-standard residues, i.e. non-protein parts, may belong to the QM subsystem, but also to the MM subsystem.

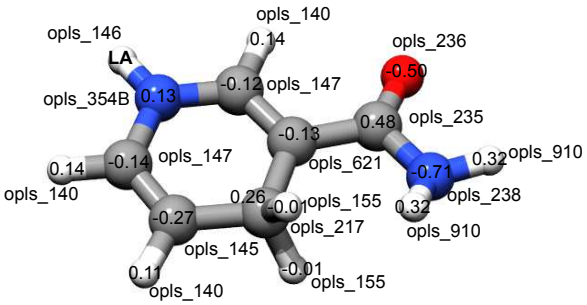


Figure 2.11. – Example for MM parameters of a 'non-standard' group: Nicotinamide. The OPLS-AA atom types (as generated by PRODRG) and atomic point charges (as generated by ORCA) are given for all atoms.

of 'pdb2gmx' and PRODRG is that they add hydrogen atoms to the crystal structure (see Figure 2.12). The hydrogen positions are chosen according to standard bond lengths, angles and dihedrals, and this procedure might thus introduce high repulsive forces into the system. This drawback is taken care of at a later stage of the preparation.

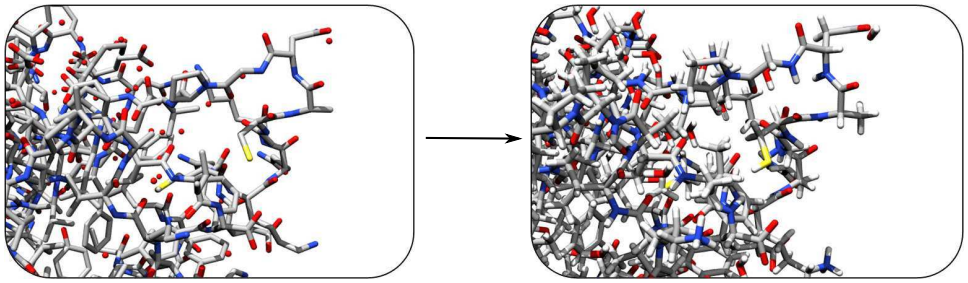


Figure 2.12. – Example showing the addition of hydrogen atoms to an X-ray crystal structure of a protein (as done by pdb2gmx).

(3) System Preparation

Apart from 'pdb2gmx', several other tools are provided by GROMACS for the preparation of the system. Before the protein can be solvated it has to be put into a 'virtual box':[j]

```
editconf -f outputProt.g96 -o outputProt_Box.g96 -d 1.0
```

The dimensions of the box are chosen such that the protein is covered by a water layer of at least 1 nm ('-d 1.0') in each direction.
When the box is present, the system can be solvated with water molecules (see also Figure 2.13):

[j] The virtual box is necessary for the periodic boundary conditions.

```
genbox -cp outputProt_Box.g96 -cs spc216.gro -p topol.top
-o outputProt_Box_Sol.g96
```

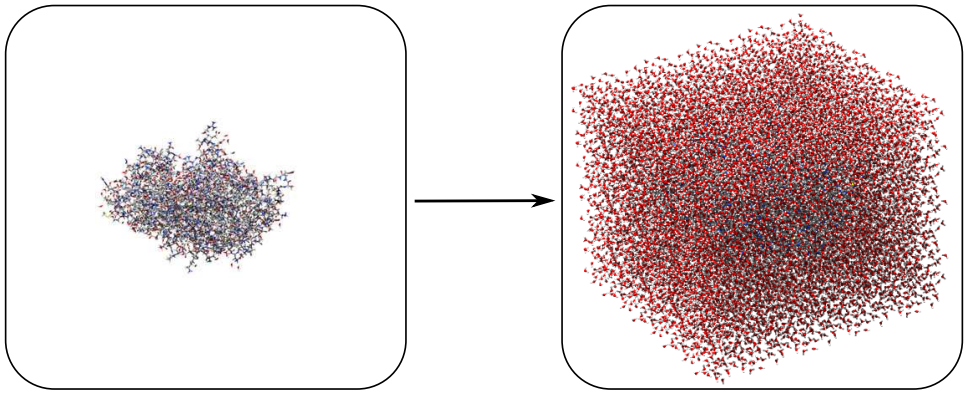


Figure 2.13. – Example showing the solvation of a protein with water molecules (as done by genbox).

Here, <spc216.g96> carries information about the solvent model. The solvation algorithm fills the box with water molecules and removes those water molecules which collide with any atom of the already existing system.
At this point of the preparation, the overall charge of the system is usually not zero. Since the system should be neutral for the MD and QM/MM calculation, it has to be neutralized. For this purpose, some solvent water molecules (not crystal waters) are replaced by sodium or chloride ions using the GROMACS program 'genion':

```
genion -np 10 -pname NA -o outputProt_Box_Sol.g96
-s outputProt_Box_Sol.tpr -norandom -n index.ndx
```

In this example, 10 solvent molecules are replaced by sodium ions ('NA') at positions with most favorable electrostatic potential ('-norandom'). The <index.ndx> (containing information about the different groups of atoms, e.g. solvent atoms, heteroatoms, etc.) and <outputProt_Box_Sol.tpr> files have to be generated beforehand.

(4) MD Simulation

Up to this point, hydrogen atoms and water molecules were introduced without taking care of repulsive interactions with other atoms. Thus, high repulsive forces might prevail in the system. Before the MD simulation can be started, the system has to be relaxed and equilibrated by classical energy minimizations and classical MD runs using periodic boundary conditions (PBC).
During these classical minimization and MD simulation runs, the non-standard residues can be frozen[k] (depending on the 'quality' of their topology) in order to prevent unphysical

[k] Freezing an atom means keeping its Cartesian coordinates constrained.

structural deformation.
For the following minimization and MD runs, the following command must be executed:

```
mdrun -v -deffnm outputProt_Box_Sol_em.tpr
-c outputProt_Box_Sol.g96
```

Here, <outputProt_Box_Sol_em.tpr> is the file containing all necessary information about the energy minimization run (it has to be generated beforehand and includes all calculation parameters, coordinates, topologies and index groups), <outputProt_Box_Sol.g96> contains the coordinates of the last minimization step. In an MD run, other files, containing the trajectory and log-files about energies, forces, etc., are additionally generated.
The step numbers given below for minimization and heating may serve as a rough guide. Of course the system has to be checked for convergence or equilibrium after each run.

Minimization First, constrained energy minimizations are carried out subsequently in order to remove artificially introduced close contacts:

1. 200 steps of steepest descent (sd) minimization only for hydrogen atoms followed by
2. 200 steps of sd minimization for the solvent molecules and finally
3. 1000 steps of conjugate gradient (cg) minimization for all atoms.

After these minimizations the system should be relatively free of strains.

Heating the System For further equilibration the system is

4. heated up to 200K with position restraints (pr) in an MD run for 5000 steps (5ps). Here, harmonic restraints of 1000 kJ/(mol·nm^2) are applied on the Cartesian positions of all heteroatoms.
5. Thereafter, the system is equilibrated without position restraints at 300K for 50ps.

MD Simulation The final NVT[l] sampling run is performed for 450ps. In this run, the system is equilibrated and snapshots can be taken from the classical MD run as starting structures for the subsequent QM/MM optimization (see Figure 2.14).

(5) Preparation of Snapshots

Before starting the QM/MM calculations, the QM region and the active region have to be defined.

QM Region The atoms that are chosen for the QM subsystem are added to the index file. Additionally, link atoms are introduced and their positions are defined. The $Q1$-$M1$ distance is constrained to their standard bond length (according to the topology). The topology is modified by distributing the charges of the $M1$ atom onto the $M2$ atoms.[123]

[l] NVT: the number of particles (N), the volume (V) and the temperature (T) of the system are kept constant.

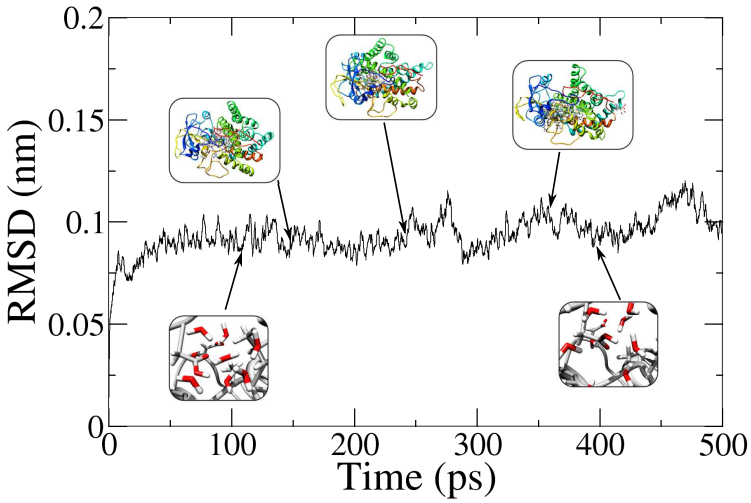


Figure 2.14. – Example of an MD simulation run. The RMSD of the backbone atom position with respect to the initial configuration is displayed. The first 50 ps belong to the equilibration phase, which is followed by the productive sampling run. Snapshots can be taken at random positions for the QM/MM optimization.

Active Region The *active region* is the set of (QM and MM) atoms whose positions are relaxed in a QM/MM optimization (see Figure 2.15). It is advisable to choose an active region which is not too large.[124] All atoms that belong to residues, which have atoms within a distance of 4 Å around any of the QM atoms, are included in the active region. This usually sums up to about 1000 active atoms.

Preoptimization The snapshots are first preoptimized classically for 300 steps with the cg minimizer (only the active region, again with frozen non-standard residues). After this, the QM/MM calculations can start.

(6) QM/MM Calculations

For the QM/MM calculations, the PBC are switched off and neither a Coulomb cutoff, nor a VDW cutoff are used. That is, the QM part is polarized by the charges of the entire protein and of all solvent molecules.
Usually a QM/MM surface scan is performed first, before stationary points (minima and transition states) are further optimized.

QM/MM Surface Scan For a QM/MM surface scan (e.g. when modeling a reaction mechanism), successive distance constrained QM/MM minimizations are carried out using a step size of 0.1 to 0.2 Å in order to obtain a smooth PES. The minimization is driven by GROMACS with the cg minimizer with a maximum force of 100 kJ/(mol·nm) as convergence criterion.

QM/MM Minimization Intermediates are taken at suitable starting points from the surface scan. They are first minimized with the GROMACS cg minimizer within a conventional optimization without constraints (same convergence criteria as above) to reach a sufficient

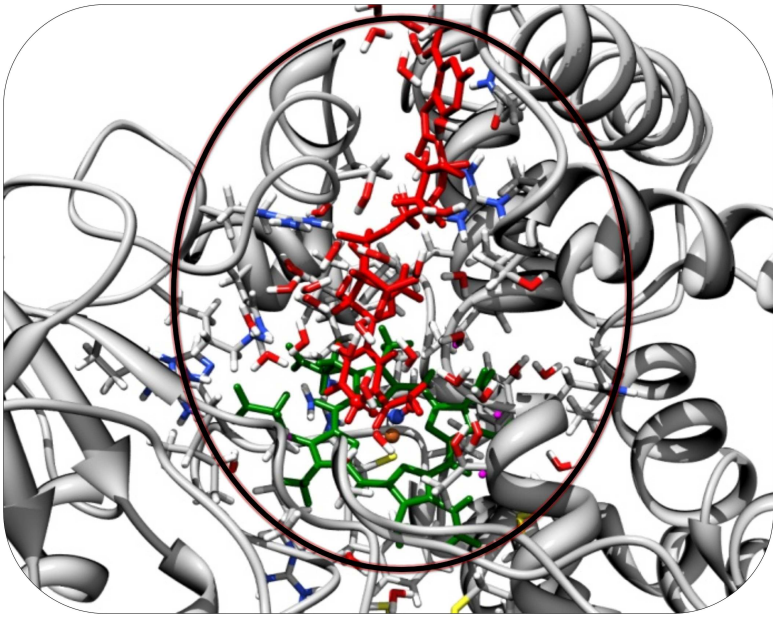


Figure 2.15. – Example of an active region. P450nor is shown with bound NADH cofactor (heme as green sticks, NADH as red sticks). The ellipse sketches the active region. Only the positions of atoms within the ellipse are optimized, all atoms outside this region are fixed to their Cartesian position.

optimization for the MM charges. Only the QM atoms are then further optimized within an ORCA QM/MM optimization (microiterative optimization with a single GROMACS step).

ORCA QM/MM Optimization In this optimization, which is fully run by ORCA, only the QM atoms are optimized, whereas the MM atoms are present only with their point charges and Lennard Jones (LJ) coefficients.[m] That is, the calculation takes into account the nonbonded interactions between the QM and MM parts. In such an optimization, the QM/MM boundary is frozen.[n] In the ORCA QM/MM optimization the ORCA default optimization convergence thresholds are used (RMS gradient: 5 kJ/(mol·nm) and MAX gradient: 15 kJ/(mol/nm)).

ORCA QM/MM TS Optimization As an initial structure for a TS optimization, the structure with the highest energy from the QM/MM surface scan is chosen. Here, the MM environment is again present with its point charges and LJ coefficients, but frozen at its minimized configuration (from the surface scan). Only the QM atoms (except the boundary atoms) are allowed to move (see also above, paragraph QM/MM Optimization). The TS is optimized using the P-RFO algorithm (see Section 2.1.4) with an approximate Hessian as initial Hessian (see Section 2.1.6).

(7) Property Calculations

Having obtained the QM/MM optimized structures, additional properties can be calculated by simply using the files obtained from the ORCA optimization containing QM Cartesian

[m] MM atomic charges are only present for the QM-MM interaction here.
[n] Link atoms and QM atoms connected to the link atoms.

coordinates, MM Cartesian coordinates, MM point charges and MM LJ coefficients (e.g. single point energy calculations using higher level electronic structure methods, frequencies, spectroscopic parameters, etc.).

QM/MM Frequency Analysis Due to its technical peculiarity, it is shortly presented at this point how QM/MM frequencies are calculated. A QM/MM frequency calculation can be carried out with the partial Hessian vibrational analysis[125] (PHVA) as implemented in ORCA. This is useful in order to

- verify the nature of minima and transition states and
- calculate enthalpic corrections (ZPE) for the intermediates.

In this analysis, the Hessian matrix is calculated only for the QM system, and the MM environment is represented only by its point charges and LJ coefficients ($\rightarrow$ QM-MM electrostatic and QM-MM VDW interactions). In the mass-weighted Hessian matrix, which is diagonalized in order to obtain the vibrational frequencies, the atomic weight of the boundary atoms is set to a very large value. This approach yields frequencies for the QM subsystem, which is interacting with the outer MM subsystem. It results at the same time in significant savings of computer time compared to the Hessian calculation of the entire QM/MM system.

2.3. Spin-Spin Zero-Field Splitting

In Section 1.1.3 an overview of experimental distance measurement in macromolecules with EPR spectroscopy was given. As already described there the simple *point-dipole approximation* to the ZFS is a popular approximation that is used in the experimental EPR community to determine the distance between two spin centers from the experimentally derived ZFS parameter.
For organic spin labels the spin-spin ZFS is the main contribution to the observed ZFS. It can be accurately calculated using quantum chemical methods. In this section, the quantum chemical treatment of the spin-spin ZFS, its different contributions and their relation to the approximations used in the experimentalist's community are discussed.
These equations (exact as well as approximate spin-spin ZFS) will be used in Chapter 5, where the scope and limitations of the point-dipole approximation are analyzed in detail.

2.3.1. Quantum Chemical Treatment

The spin-spin contribution to the ZFS parameter $D^{(SS)}$ dominates the ZFS of organic triplets and diradicals. As discussed recently,[54] within a DFT framework, the tensor components of $D^{(SS)}$ may be calculated from the equation of McWeeny and Mizuno:[126]

$$D_{KL}^{(SS)} = -\frac{g_e^2}{16}\frac{\alpha^2}{S(2S-1)}\sum_{\mu\nu}\sum_{\kappa\tau} P_{\mu\nu}^{\alpha-\beta} P_{\kappa\tau}^{\alpha-\beta}\left[\left\langle \mu\nu \left| g^{KL} \right| \kappa\tau \right\rangle - \left\langle \mu\kappa \left| g^{KL} \right| \nu\tau \right\rangle\right] , \quad (2.52)$$

where g_e is the free electron g-value (2.002319...), α is the fine structure constant ($\approx 1/137$ in atomic units), the indices μ, ν, κ, τ refer to basis functions, $P_{\mu\nu}^{\alpha-\beta}$ is an element of the spin density matrix and $g^{KL} = r_{12}^{-5}\left(3r_{12,K}r_{12,L} - \delta_{KL}r_{12}^2\right)$ is the operator for the electron-electron magnetic dipole-dipole interaction. In analogy to Hartree Fock theory we can distinguish between two different contributions to $D_{KL}^{(SS)}$. The $P_{\mu\nu}^{\alpha-\beta}P_{\kappa\tau}^{\alpha-\beta}\left\langle \mu\nu \left| g^{KL} \right| \kappa\tau \right\rangle$ part of equation (2.52) is a 'Coulomb' contribution, while the $P_{\mu\nu}^{\alpha-\beta}P_{\kappa\tau}^{\alpha-\beta}\left\langle \mu\kappa \left| g^{KL} \right| \nu\tau \right\rangle$ term is an 'exchange' contribution that arises from the antisymmetry requirement of the N-electron molecular wavefunction. Thus, even the direct dipolar spin-spin interaction contains an exchange contribution that is of fundamentally different origin than the isotropic Heisenberg type exchange interaction, which is frequently used in the modeling of interacting spins.[127,128] The Heisenberg isotropic exchange interaction is responsible for the splitting of the energy levels into singlet- and triplet states of the interacting radical pair. The exchange contribution to the spin-spin interaction in equation (2.52) on the other hand does not preserve singlet and triplet spin symmetry and represents a quantum mechanical correction to the magnetic dipole-dipole interaction.

2.3.2. Coulomb versus Exchange Contributions

In the following paragraph, we will concentrate on a system with two unpaired electrons in the singly occupied molecular orbitals (SOMOs) i and j. If we expand equation (2.52) in the

MO basis we get:

$$
\begin{aligned}
D_{KL}^{(SS)} = & -\frac{g_e^2}{16}\frac{\alpha^2}{S(2S-1)} \times \\
& [\left\langle ii\left|g^{KL}\right|ii\right\rangle + \left\langle ii\left|g^{KL}\right|jj\right\rangle + \left\langle jj\left|g^{KL}\right|ii\right\rangle + \left\langle jj\left|g^{KL}\right|jj\right\rangle \\
& -\left\langle ii\left|g^{KL}\right|ii\right\rangle - \left\langle ij\left|g^{KL}\right|ij\right\rangle - \left\langle ji\left|g^{KL}\right|ji\right\rangle - \left\langle jj\left|g^{KL}\right|jj\right\rangle] \\
= & -\frac{g_e^2}{8}\frac{\alpha^2}{S(2S-1)} \times \left[\left\langle ii\left|g^{KL}\right|jj\right\rangle - \left\langle ij\left|g^{KL}\right|ij\right\rangle\right] \ .
\end{aligned}
\tag{2.53}
$$

Four terms arise that represent a self-interaction of the unpaired electrons. These terms cancel out each other. Thus, one is left with only two terms that represent the Coulomb and exchange contributions to the spin-spin interaction. Unfortunately, the division of the total spin-spin interaction into these two contributions is not unique. This is readily seen upon performing a unitary transformation of the SOMOs ($|i\theta\rangle = |i\rangle\cos\theta + |j\rangle\sin\theta$, $|j\theta\rangle = -|i\rangle\sin\theta + |j\rangle\cos\theta$, arbitrary rotation angle θ). Such a transformation leaves the sum of Coulomb and exchange terms invariant, but changes their relative magnitudes. Thus, since the SOMOs of a spin system with $S > 1/2$ are only specified up to an arbitrary unitary transformation, one has to specify the representation that one wishes to work with. For interpretation purposes, we argue that transforming the SOMOs to a local representation (i_L, j_L) is most revealing. In this case, the two interactions smoothly go into 'classical' limits upon separating the spin-carrying fragments. If the localized orbitals are 'compressed' to δ-functions (say, at the center of gravity of i_L and j_L given by $R_{i_L} = \langle i_L|r|i_L\rangle$ and $R_{j_L} = \langle j_L|r|j_L\rangle$), the Coulomb part of the spin-spin interaction simply becomes:

$$
D_{KL}^{PD,J} \approx -\frac{g_e^2}{8}\frac{\alpha^2}{S(2S-1)} \times R_{i_L j_L}^{-5}\left[3R_{i_L j_L,K}R_{i_L j_L,L} - \delta_{KL}R_{i_L j_L}^2\right] \ . \tag{2.54}
$$

This is clearly a point-dipole type equation with the effective intercenter distance vector $R_{i_L j_L} = R_{i_L} - R_{j_L}$. Under the plausible assumption that the orbitals (i_L, j_L) decay exponentially at large distances from their respective centroids, the exchange contribution should also decay very rapidly. A Mulliken type approximation would suggest that the leading term goes as $\exp(-\kappa R_{i_L j_L})\, R_{i_L j_L}^{-5}$ (with constant κ) — that is, much faster than the Coulomb interaction.

2.3.3. Multicenter Contributions

A different way to partition the spin-spin interaction is to recognize that the basis functions are tied to parent atomic centers. Hence, one obtains 1- to 4-center contributions to the spin-spin interaction as in equation (2.55) and Table 2.1:

$$
\begin{aligned}
D_{KL}^{(SS)} = & -\frac{g_e^2}{16}\frac{\alpha^2}{S(2S-1)} \times \sum_{\mu\in A}\sum_{\nu\in B}\sum_{\kappa\in C}\sum_{\tau\in D} P_{\mu_A\nu_B}^{\alpha-\beta}P_{\kappa_C\tau_D}^{\alpha-\beta}[\left\langle \mu_A\nu_B\left|g^{KL}\right|\kappa_C\tau_D\right\rangle \\
& -\left\langle \mu_A\kappa_B\left|g^{KL}\right|\nu_C\tau_D\right\rangle] \ .
\end{aligned}
\tag{2.55}
$$

The various interaction types have a transparent physical interpretation: The (AA|AA) 1-center contributions represent the interactions of local magnetic dipoles that arise through a certain amount of triplet character at a given atom A. The contributions (AA|BB) represent the quasi-classical point-dipole interactions, while (AB|AB) exchange and (AA|AB) hybrid

Table 2.1. – Definition of the 1- to 4-center- contributions to $D^{(SS)}$ in equation (2.55).

ZFS-Contribution	Definition
1-center	A=B=C=D
2-center-Coulomb	A=B, C=D, A≠C
2-center-Exchange	A=C, B=D, A≠B
2-center-Hybrid	A=B=C≠D or A=B=D≠C or A=C=D≠B or B=C=D≠A
3-center-Coulomb	A=B, A≠C≠D or C=D, A≠B≠C
3-center-Exchange	A=C, A≠B≠D or B=D, B≠A≠C or A=D, A≠B≠C or B=C, B≠A≠D
4-center	A≠B≠C≠D

type 2-center integrals provide quantum mechanical corrections that depend primarily on the overlap of spin-carrying orbitals on centers 'A' and 'B'. The hybrid type (AA|AB) integral may also be interpreted as the dipole-dipole interaction of unpaired electrons in the bond A-B with unpaired electrons localized on center A. Likewise, the 3-center contribution (AA|BC) represents the interaction of unpaired electrons in the bond B-C with unpaired electrons on A. Finally, the 4-center contributions (AB|CD) are interpreted as 'distant', and the 3-center exchange integrals (AB|AC) as 'adjacent' bond-bond interactions.

2.3.4. Distributed Point-Dipole Model

The popular distributed point-dipole model[129] is obtained upon contracting the individual basis functions (instead of the entire localized SOMOs) to δ-functions. Assuming the spin density matrix to be diagonal in this basis and neglecting the exchange interaction, one then obtains:

$$D_{KL}^{DPD} \approx -\frac{g_e^2}{8}\frac{\alpha^2}{S(2S-1)} \times \sum_{AB} P_A^{\alpha-\beta} P_B^{\alpha-\beta} R_{AB}^{-5} \left[3R_{AB,K}R_{AB,L} - \delta_{KL}R_{AB}^2\right] , \qquad (2.56)$$

with a sum over all nuclei A and B, and $P_A^{\alpha-\beta} = \sum_{\mu\in A} P_{\mu\mu}^{\alpha-\beta}$ being the 'gross' spin population on atom A. Equation (2.56) describes the interaction of collections of point-dipoles centered at atomic positions where each atom pair is weighted by the product of the spin populations that reside on these atoms. Models of this type have been frequently used in EPR investigations,[42] where the spin populations have either been estimated from measured hyperfine couplings or from quantum chemical calculations. It should be noted that the justification of these approaches does not involve the *net* spin populations as output by quantum chemi-

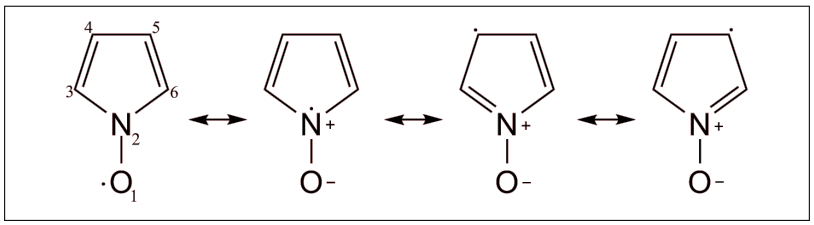


Figure 2.16. – Scheme showing the dominant resonance structures of the nitroxyl ring.

cal program packages within various schemes (Mulliken, Löwdin, natural population analysis (NPA[130])), but the *gross* atomic spin populations. The differences between the two sets of numbers are not negligible. As an illustrative example, in Table 2.2 the

- *net* Mulliken spin population,
- *gross* atomic Mulliken spin population and
- spin population obtained with a natural population analysis

are given for the model system depicted in Figure 2.16. The calculated spin distribution

Table 2.2. – *Net* and *gross* atomic Mulliken spin population obtained with an NPA for the truncated dinitroxyl model system (orthogonal configuration, numbering as in Figure 2.16 and 5.2).

atom	*Net* Mulliken spin population	*Gross* atomic Mulliken spin population	NPA spin population
O_1	0.50	0.46	0.48
N_2	0.19	0.21	0.21
C_3	0.03	0.03	0.05
C_4	0.12	0.14	0.10
C_5	0.13	0.10	0.10
C_6	0.03	0.06	0.05

suggests that the dominant resonance structures are the ones shown in Figure 2.16. This description emphasizes the delocalization of the spin onto the ring system of the nitroxide — an important factor in the analysis that follows below.

2.3.5. Effective Point-Dipole Model

If the distance between two spin-carrying fragments is large enough, it may be possible to reduce equation (2.56) to a single term, where R_{AB} refers to an 'effective' distance. Quite frequently, the origins of the spin-distributions are simply fixed in an *ad hoc* manner based on chemical intuition (for nitroxides commonly the center of the N-O bond is chosen).[131] A perhaps somewhat more rational approach is the following: First calculate the centers of

gravity (COG) of the spin populations on fragments 'F1' and 'F2' as:

$$R^{(F1)} = \sum_{A \in F1} \bar{P}_A^{\alpha-\beta} R_A \ , \tag{2.57}$$

$$R^{(F2)} = \sum_{A \in F2} \bar{P}_B^{\alpha-\beta} R_B \ , \tag{2.58}$$

and then obtain the ZFS D-value as:

$$D_{KL}^{COG} \approx -\frac{g_e^2}{8} \frac{\alpha^2}{S\,(2S-1)} \times R_{12}^{-5} \left[3R_{12,K} R_{12,L} - \delta_{KL} R_{12}^2 \right] \ , \tag{2.59}$$

where $R_{12} = R^{(F2)} - R^{(F1)}$ is the distance vector between the centers of gravity of fragments F1 and F2.
All the approximations and their limitations will be explored in Chapter 5.

3. Characterization of the Intermediates of P450nor

3.1. Introduction

As already presented in Section 1.2, the first two species of the P450nor reaction mechnanism are experimentally readily accessible, and a considerable amount of experimental data exist.[70,72,83–87] There is thus consensus about their properties in the scientific community.[88–90] The third intermediate of the reaction, to which a key role in the mechanism is assigned, is intermediate *I*. This has always been a central focus in studies on P450nor. Due to its short lifetime it is not easy to access experimentally. In particular, a crystal structure is not available and hence spectroscopy has to be relied upon for providing structural information. However, thus far spectroscopic data about intermediate *I* have been quite limited. In the absence of sufficient experimental constraints, a rather wide variety of structures were proposed in various experimental and theoretical studies.[78,88,89,91,92]
The aim of this chapter is to analyze and elucidate the geometric and electronic structures of all three intermediates in P450nor:

- The resting state (**1**),
- the NO-bound state (**2**) and
- the key intermediate *I*.

In order to obtain a complete picture, QM/MM calculations were carried out to study the intermediate structures in the protein environment. The selected intermediates that are obtained in this chapter will be analyzed further by calculating the *entire reaction path* in the protein environment in Chapter 4.
For all intermediates their spectroscopic properties are determined and compared with the available experimental data.[a] Furthermore, they are characterized in a detailed analysis of their structural and electronic properties.
For intermediate *I* three different ligands are modeled in the active site (only iron and the sixth ligand are given here):

(3) $[FeNHO]^{2+}$,

(4) $[FeNHOH]^{3+}$ and

(5) $[FeN]^{3+}$,

yielding in total six different electronic structures (two electromers for each ligand). The spectroscopic, structural and electronic properties of all intermediates (**1** to **5**) are analyzed in the following.

[a] Mössbauer spectra for all three intermediates, which were measured by E. Bill, but are not yet published, can be used for the characterization of the intermediates. The spectra are shown in the Appendix in Figure B.1.

3.2. Computational Details

The workflow presented in Section 2.2.6 was applied in the calculations of this chapter. Details for the study are given below.

3.2.1. Preparation of the Protein Structure and MD Simulation

The calculations were started from the experimental X-ray structure of P450nor (PDB code: 1CL6).[70] The protonation states of the residues were assigned with PROPKA 2.0 (Glu, Asp) and the standard algorithm of GROMACS (other residues).[102,132,133] As MM parameter set the OPLS-AA force field was used.[134–136] The topologies for the non-standard residues (cysteinate-ligand to iron, heme-group and the nitrogenous intermediates) were generated using PRODRG2.5[121] and CHELPG charges[122] derived with ORCA[99] (BP86[137–139] with SV(P)[140–142] basis set). The protein was put in a rhomboedric box (protein dimensions plus 2 nm in each direction) and solvated with water molecules. In order to neutralize the system, solvent water molecules (not crystal waters) were replaced by sodium ions using the GROMACS program genion[102] resulting in an overall neutral charge.
The thus prepared system was relaxed and equilibrated by classical energy minimizations and classical MD runs with periodic boundary conditions. During the classical minimization and MD simulation the Cartesian coordinates of the iron ion and its ligands (heme, Cys_{352} and NHO ligand) were kept fixed.

3.2.2. QM and QM/MM Calculations on Intermediates

The QM/MM calculations started from the snapshots taken from the MD simulation (at (A) 250ps, (B) 350ps and (C) 450ps). The different ligands (given as L1 to L5 in Figure 3.1) were modeled at the sixth coordination site of heme. Depending on the ligand charge, the total system charge was adapted by deleting sodium ions to maintain overall neutrality. Four different QM regions were used (see Figure 3.1), all of which consist of the iron-ligand-porphyrin unit without (**R1**) and with (**R2**) side chains and additionally with important residues around the carboxyl groups (**R3**) and around the Cys_{352} residue (**R4**), respectively. All residues that have atoms within a distance of 4 Å of the heme-group, Cys_{352} and the respective sixth ligand were optimized, resulting in approximately 800 optimized atoms. QM/MM calculations were carried out with an electrostatic embedding scheme using the QM/MM interface between GROMACS[102,143] and ORCA. Periodic boundary conditions were switched off and no cutoffs were used for the nonbonded QM-MM interactions. For the bonds connecting the QM/MM boundary the hydrogen link atom approach[143] was used in conjunction with the charge shift model.[123] Gas phase calculations (D) were started from the obtained QM/MM structures, without using any constraints, modeling the protein environment through a dielectric continuum using COSMO[144] (conductor-like screening model) with a dielectric constant of $\epsilon = 4$.[145]

3.2.3. Minimization Scheme

The snapshots were first minimized with QM region **R2** with the GROMACS conjugate gradient minimizer to reach a sufficient optimization for the MM charges (threshold: maximum force: 100 kJ/(mol·nm)). The different QM regions were further optimized with ORCA including the nonbonding QM-MM interactions, with the positions of the link atoms and their directly adjacent atoms being constrained.

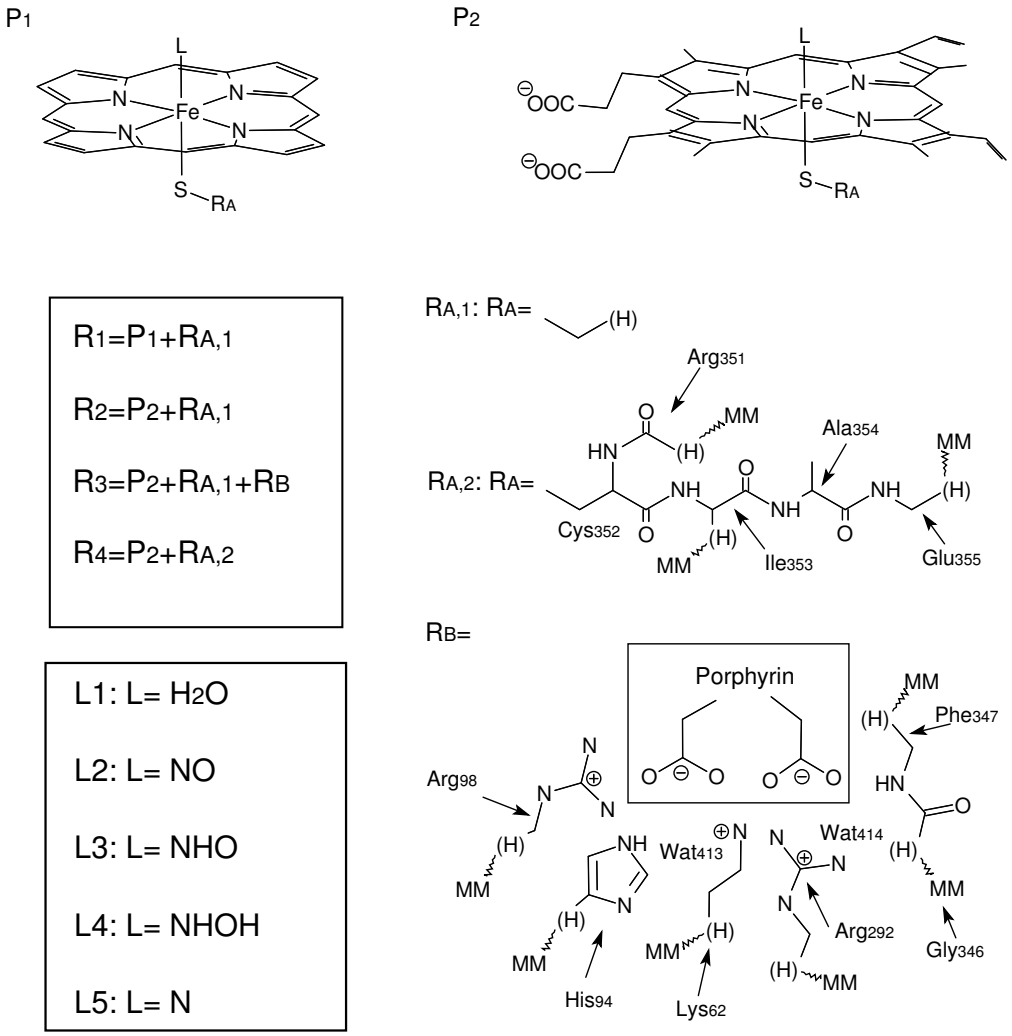


Figure 3.1. – QM systems **R1** to **R4** and the ligands **L1** to **L5** which characterize the different states of the protein.

3.2.4. QM Method

The QM/MM optimization calculations were performed with unrestricted RIJCOSX[146]-B3LYP[137,147,148] using the def2-SVP basis set in accordance with the def2-SVP/J auxiliary basis set (**B1**).[b,140,149,150] A second, larger basis set was used for optimization to test the basis set convergence. This basis set (**B2**) comprises of def2-TZVPP[149–151] (with def2-TZVPP/J as auxiliary basis set) for iron and the first ligand sphere atoms including the full sixth ligand, and of **B1** for the remaining atoms. Single-point calculations were carried out with RIJCOSX-B3LYP*[152,153] using the def2-TZVPP basis set in accordance with def2-TZVPP/J as auxiliary basis set for all atoms (**B3**). The Mössbauer parameters were calculated using B3LYP* in conjunction with the CP(PPP) basis set[154] for iron and TZVP[151] for the remaining atoms (**B4**). In all calculations ZORA was used to account for relativistic effects,[155–157] the dispersion correction implemented as VDW06 in ORCA was included to account for dispersion effects within the QM part[158] and the MM point charges of the protein

[b] During the course of the project it turned out that the use of GGA functionals is insufficient, but that these have to be replaced by more accurate hybrid functionals.

environment were included in the one-electron Hamiltonian of the QM-calculation[159] (see also equation (2.50)).
In the frequency calculations the same level of theory as for the optimization was used. The partial Hessian vibrational analysis[125] (see page 49) as implemented in ORCA was used here in order to circumvent the inclusion of MM bonding interactions at the QM/MM boundary. For the localization of molecular orbitals the Pipek-Mezey algorithm[160] was used.

Calibration of the Density Functional

Goodrich et al. pointed out that the bonding situation in heme-nitrosyls is a challenge for DFT calculations.[161] They showed that BP86 overestimates and B3LYP underestimates the Fe-NO bond in ferrous heme-nitrosyls. Olah et al. performed a comparative study using DFT and correlated ab initio calculations for NO bonding to ferric and ferrous nitrosyls.[162] Their data set is used in this work for calibration (see Table 3.1). It turned out that B3LYP* reproduces the NO bonding energies best. Thus B3LYP* was chosen for the calculation of the single point energies (vide supra).

Table 3.1. – NO bonding energy to heme-thiolate. Model structures and reference energies are taken from Olah et al.[162] Single point energies are calculated using basis set **B3**. Relative energies are given in kcal/mol.

Density Functional	Model 1[a]	FeP(SH)[b]
B2PLYP	-44.5	-57.1
B3LYP	6.5	-5.9
B3LYP*	-7.3	-19.8
PBE	-42.8	-55.0
TPSSh	-19.7	-32.1
Reference value	-9.3 (-4.5)[c]	$\approx$ -15 – -20[d]

[a] Small model mimicking a heme-thiolate NO complex. Heme represented by amidines and thiolate by SH^--ligand.
[b] Bigger model for hemethiolate as used in Olah et al.[162]
[c] Relative CCSD(T) energy extrapolated to the infinite basis set limit (calculated with cc-pV8Z/cc-pVDZ for Fe/amidines). In parentheses the BSSE corrected energy is given.
[d] As given in Olah et al.[162]

3.3. MD Simulation Results

After the preparation of the protein structure 450ps of productive sampling were carried out (see Figure 3.2). The plots show that the system is equilibrated. For the subsequent QM/MM calculations three snapshots were used (A – C), taken at 250, 350 and 450 ps.

3.4. QM/MM Results

The structures of the different P450nor intermediates were optimized starting from the three different snapshots (A – C). Their results are compared with the results that were obtained

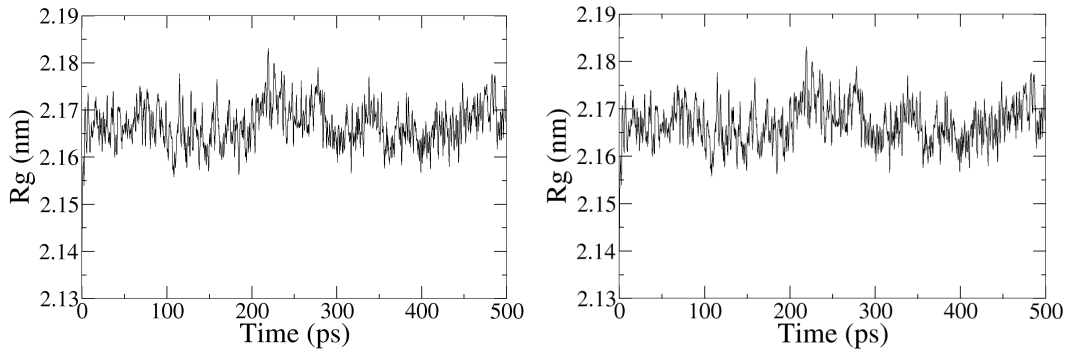


Figure 3.2. – The gyration radius of the protein (left) and the root-mean-square deviation (rmsd) of the backbone atoms (right) with respect to the initial configuration of 1CL6 are plotted versus the MD simulation time.

with a cluster model (D). This comparison allows us to examine the effect of the protein environment.
Five different models are studied, differing by the ligand in the axial position. In the different snapshots the sixth ligand position is occupied with: **1** a water molecule in the resting state and **2** NO in the NO-bound state. In the case of intermediate *I* the mono- (**3**, NHO^-) and diprotonated (**4**, NHOH) alternatives are used. An additional ligand **5**, which is obtained by formally removing one water molecule from **4**, is also used. In the course of the study it emerged that **2** – **5** exist in two different electronic states each. These two states are a closed shell and a broken symmetry[c] (BS) state. Thus, a total of nine different species is analyzed and denoted in the following as:

(1) Fe(III)-H_2O,

(2a) closed shell $\{FeNO\}^6$,

(2b) BS(2,2) $\{FeNO\}^6$,

(3a) closed shell $\{FeNHO\}^8$,

(3b) BS(1,1) $\{FeNHO\}^8$,

(4a) closed shell $\{FeNHOH\}^8$,

(4b) BS(1,1) $\{FeNHOH\}^8$,

(5a) closed shell $\{FeN\}^8$ and

(5b) BS(2,2) $\{FeN\}^8$.

It will become clear in Section 3.4.5 that the calculated Mössbauer parameters for both heme-nitrido complexes (**5a** and **5b**) are far off the experimental values. Thus, the analysis of the structural and electronic properties will be focused on the other ligands.

[c]The broken symmetry state BS(m,n) refers to a state with m unpaired spin-up electrons on fragment A and n unpaired spin-down electrons on fragment B.

3.4.1. Geometric Structures

Convergence of Basis-Set and QM-Region Size Before analyzing the results for the different models, the validity of the used basis sets and QM regions is discussed. **1** and **2** were optimized in snapshot B with two basis sets, **B1** and **B2**. Basis **B1** is sufficient for the characterization of the iron-nitrosyl species (as can be seen by comparing the results for **B1** and **B2** in Table 3.3). For the resting state Table 3.2 shows that there is still some deviation in Fe-O bond length upon going from **B1** to **B2**. The relatively weak bond between iron and the water ligand is only insufficiently described by a basis of size **B1**. But on the other hand the spectroscopic properties (see e.g. Table 3.8) are hardly influenced by these deviations in the bond distance to a weak axial ligand. In the interest of efficiency, the efforts were thus concentrated on the structures obtained with basis **B1**.
Regarding the QM regions, **1** and **2** were again optimized in snapshot B with four different QM regions, **R1** to **R4**. The important structural features of the first ligand sphere do virtually not change for all four QM regions. The following analysis will thus be based on the results obtained with QM regions **R2** and **R1**.

Resting State

In the resting state the water molecule only binds weakly to the iron. The iron-oxygen distance thus varies significantly over the different snapshots and the cluster model, as shown in Table 3.2. In the snapshots, the Fe-O bond varies by more than 10 pm, depending on

Table 3.2. – Computed and experimental structural parameters for the Fe(III)-H_2O species for different QM regions in snapshots A – C and the cluster model D. Distances are given in Ångstrom.

Snapshot	QM region	Fe-O	Fe-S
A	**R2**	2.337	2.176
B	**R2**	2.214	2.185
	R2[a]	2.296	2.180
	R3	2.230	2.185
	R4	2.241	2.187
	R1	2.224	2.190
C	**R2**	2.168	2.190
D	**R2**	2.162	2.188
P450nor	Exp. data[86]	2.44 ± 0.02	2.26 ± 0.02

[a]Basis set **B2** was used

hydrogen bonding interactions with surrounding water molecules in the distal heme pocket. Without these competing hydrogen bonds (as evident in the cluster calculation), the Fe-O distance is even shorter. Compared to the experimental value, the Fe-O bond length is slightly underestimated in the QM/MM calculations, but it can be seen that the explicit modeling of the surrounding in the active site and the larger basis set improve the computed value significantly.

NO-bound State

Two different electromers exist for the NO-bound form: a closed shell (**2a**) and a broken symmetry BS(2,2) (**2b**) solution. Table 3.3 summarizes the important structural features of these species in the different calculations. The most striking difference between the two

Table 3.3. – Computed structural parameters for **2a**/**2b** (in the closed shell-/ BS(2,2)-state) for different QM regions in snapshots A – C and in the cluster model D. Distances are given in Ångstrom, bond and dihedral angles are given in degrees.

Snapshot	QM region	Fe-N	N-O	Fe-N-O	Fe-S	C-S-N-O
A	**R2**	1.649/1.756	1.161/1.173	157/152	2.307/2.281	159/160
	R2[a]	1.656/1.760	1.155/1.166	159/154	2.309/2.283	148/129
B	**R2**	1.641/1.713	1.148/1.157	171/171	2.313/2.291	178/172
	R2[b]	1.641/1.712	1.143/1.154	171/171	2.316/2.295	177/171
	R2[a]	1.659/1.754	1.155/1.165	158/155	2.309/2.280	-127/-122
	R3	1.641/1.713	1.147/1.157	171/171	2.313/2.294	-179/162
	R4	1.640/1.712	1.147/1.157	171/170	2.312/2.297	169/169
	R1	1.645/1.713	1.148/1.156	168/170	2.302/2.281	-138/-145
C	**R2**	1.648/1.754	1.162/1.173	158/154	2.305/2.277	147/145
D	**R2**	1.657/1.757	1.146/1.157	163/153	2.295/2.251	-177/-88
	R1	1.663/1.738	1.142/1.151	162/157	2.285/2.257	88/89
P450nor	Exp.data[86]	1.66 ± 0.02	–/–	160 - 180	2.26 ± 0.02	–/–
P450cam	Exp.data[86]	1.76 ± 0.02	–/–	160 - 180	2.26 ± 0.02	–/–

[a] All solvent molecules were deleted and the QM region was reoptimized.
[b] Basis set **B2** was used.

electromers is the Fe-N bond length with a deviation of up to 10 pm (see also Section 3.4.2). But this is not the only exciting observation on the Fe-N bond length. What is much more surprising is that for **2b** the Fe-N bond length of the cluster model (D) is reproduced in two of the snapshot structures (A and C), but in the other snapshot (B) a much shorter bond distance is observed. The origin for this difference lies in:

- either the differing protein conformations at different MD simulation times or
- different hydrogen bonding networks in the snapshots due to the water molecules which are present in the large heme distal pocket.

This question can be answered by removing the water molecules from two snapshot structures (A and B) and reoptimizing QM region **R2** again. The results (given under **R2**[a] in Table 3.3) indicate that the protein structure itself influences the Fe-N bond length only marginally. Thus, the difference is due to the different water networks in the heme distal pocket.
The computed Fe-N-O angle varies from 152° to 171° for the different models. At the same time, the orientation of NO with respect to the heme plane (C-S-N-O dihedral angle) varies in a range of about 180°. These results are not unexpected, since the potential energy surface for Fe-N-O bending is flat over a wide range (150° – 180°) and NO can freely rotate above the heme plane (as can be seen in Figures 3.4 and 3.3, compare also similar results for ferric and ferrous heme-nitrosyls with an axial imidazole ligand[163,164]).

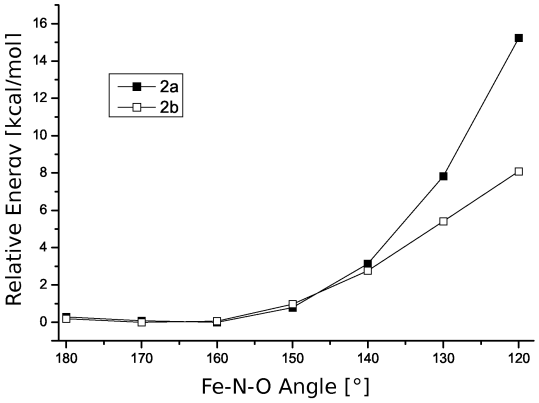


Figure 3.3. – Relative energy for Fe-N-O bending in **2**. Geometries obtained with relaxed surface scan for QM region **R1** (model D) with basis **B1**. Relative energies obtained with B3LYP*/**B3**. COSMO was applied for all calculations.

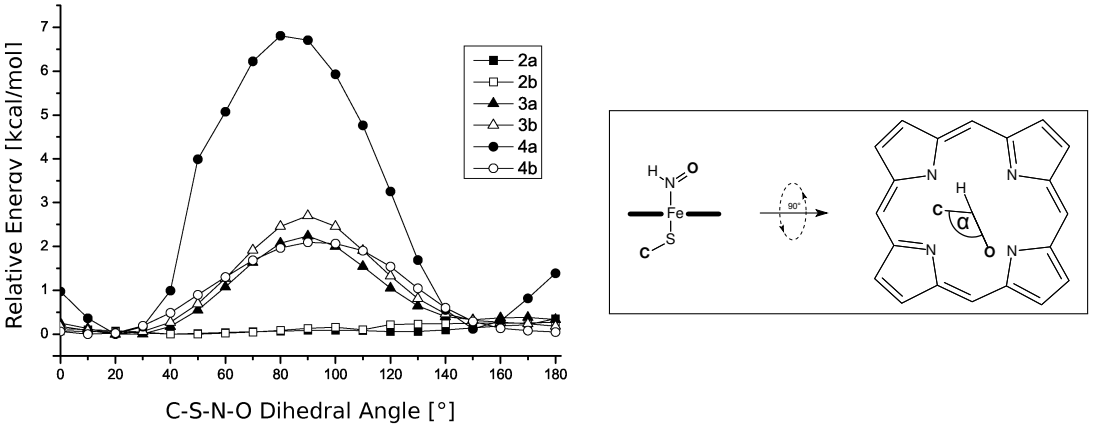


Figure 3.4. – Barrier for the rotation of the ligand above the heme plane for **2**, **3** and **4** (left figure). Definition of α (dihedral angle C-S-N-O) in the right figure. Geometries obtained with relaxed surface scans for QM region **R1** (model D) with basis **B1**. Relative energies obtained with B3LYP*/**B3**. COSMO was applied for all calculations.

The N-O bond length in both electromers corresponds approximately to the bond length of free NO (at the same level of theory: NO^+ 1.062 Å, NO 1.148 Å, NO^- 1.257 Å), where the N-O bond length in the BS species has a slight tendency towards NO^-. Compared to the available experimental EXAFS data of NO-binding to P450nor,[86] the closed shell species fits the experimental data better than the BS species.

Intermediate *I*

The structural data for the optimized geometries of **3** – **5** are summarized in Table 3.4. Two electromers exist for all ligands: a closed shell (**3a**, **4a**, **5a**) and a broken symmetry (**3b**, **4b**, **5b**) state. As mentioned above, we concentrate on the results obtained for **3a**, **3b**, **4a** and **4b**.

In comparison to the NO-bound species, two more electrons are present in intermediate *I*. Thus, the NHO(H) ligand is now bent (Fe-N-O angle: 124 – 134°). It was expected that

Table 3.4. – Computed structural parameters for **3a**/**3b**, **4a**/**4b** (in the closed shell/BS(1,1)-state) and **5a**/**5b** (in the closed shell/BS(2,2)-state) for QM region **R2** in snapshots A – C and in the cluster model D. Distances are given in Ångstrom, bond and dihedral angles are given in degrees.

Ligand	Snapshot	Fe-N	N-O	Fe-N-O	Fe-S	C-S-N-O
NHO	A	1.793/1.896	1.264/1.281	134/131	2.329/2.296	175/173
(**3a**/**3b**)	B	1.793/1.886	1.254/1.275	134/131	2.384/2.348	-130/-133
	C	1.783/1.887	1.266/ 1.282	133/131	2.335/2.306	-168/-169
	D	1.796/1.857	1.235/1.258	132/130	2.354/2.333	179/179
NHOH	A	1.796/1.962	1.361/1.337	126/128	2.241/2.216	167/-174
(**4a**/**4b**)	B	1.768/1.930	1.376/1.343	126/126	2.257/2.223	-150/-151
	C	1.786/1.984	1.371/1.342	128/128	2.236/2.219	178/178
	D	1.799/1.947	1.367/1.331	124/126	2.234/2.221	-179/180
N (**5a**/**5b**)	B	1.521/1.645	–/–	–/–	2.400/2.355	167/168

the bent nature of the ligand would result in a larger barrier to rotation of the ligand above the heme plane. In order to study this, relaxed surface scans were performed for the rotation of the NHO and NHOH ligand in the cluster model D (see Figure 3.4). In all four cases rotation barriers are present. Species **4a** shows a high barrier of 6.8 kcal/mol. The other three species have relatively low barriers (2.1 to 2.7 kcal/mol). The existence of intrinsic rotation barriers goes hand in hand with a much smaller variation of the C-S-N-O dihedral angle in the QM/MM results, as compared to the NO-bound state **2** (compare Table 3.4). There are differences between the closed shell and BS structures for **3** and **4**:

- In **3a** the N-O bond length is, with 1.254 to 1.266 Å, in between the N-O bond length of free NHO and NHO^- (1.200 and 1.322 Å). The N-O bond in **3b** is slightly longer (by 2 pm), indicating a larger resemblance to the NHO^- ligand. The difference in the Fe-N bond length is, again, more pronounced and elongated by about 10 pm upon going from the closed shell to the BS structure.
- In **4a** the N-O bond is slightly longer than that of free NHOH ($NHOH^+$ 1.215 Å, NHOH 1.352 Å, $NHOH^-$ 1.555 Å), whereas the N-O bond of **4b** is slightly shorter than that of free NHOH. The Fe-N bond is elongated by 16 pm in **4b** compared to **4a**.

Comparing **3** and **4**, similar Fe-N distances are observed for **3a** and **4a**, but a larger deviation for **3b** and **4b**. The most important difference between **3** and **4** is the Fe-S bond length (deviation of ≈ 10 pm). The reason for this difference is that, in **3**, more electrons are pushed into the iron moiety, thus preventing stronger electron donation from the sulfur atom.

3.4.2. Bonding

For the analysis of the leading bonding interactions the occupied KS orbitals were localized. For the broken symmetry species, it was necessary to exclude the magnetically coupled orbitals (unrestricted corresponding orbitals, UCOs) from the localization procedure. The thus obtained localized molecular orbitals (LMOs) are discussed in the following. First the $\{FeNO\}^6$ (**2**) species is described before the changes of the bonding pattern upon reduction to $\{FeNHO(H)\}^8$ (**3**, **4**) are discussed.

NO-Bound State

At the beginning of the project the expected result was a closed shell structure for **2** (**2a**). Surprisingly, the BS(2,2)-state **2b** was also found. This is in contrast to the results obtained by previous studies on P450nor.[88] The reason for this is that, so far, GGA-functionals were used, which only yield one SCF solution, the closed shell solution. In this study more accurate hybrid functionals are used, which yield two electronic states for $\{FeNO\}^6$.[d] The LMOs for **2a** and **2b** are displayed in Figure 3.5.

The closed shell species has one non-interacting d_{xy} orbital and two Fe 3d orbitals (d_{xz}

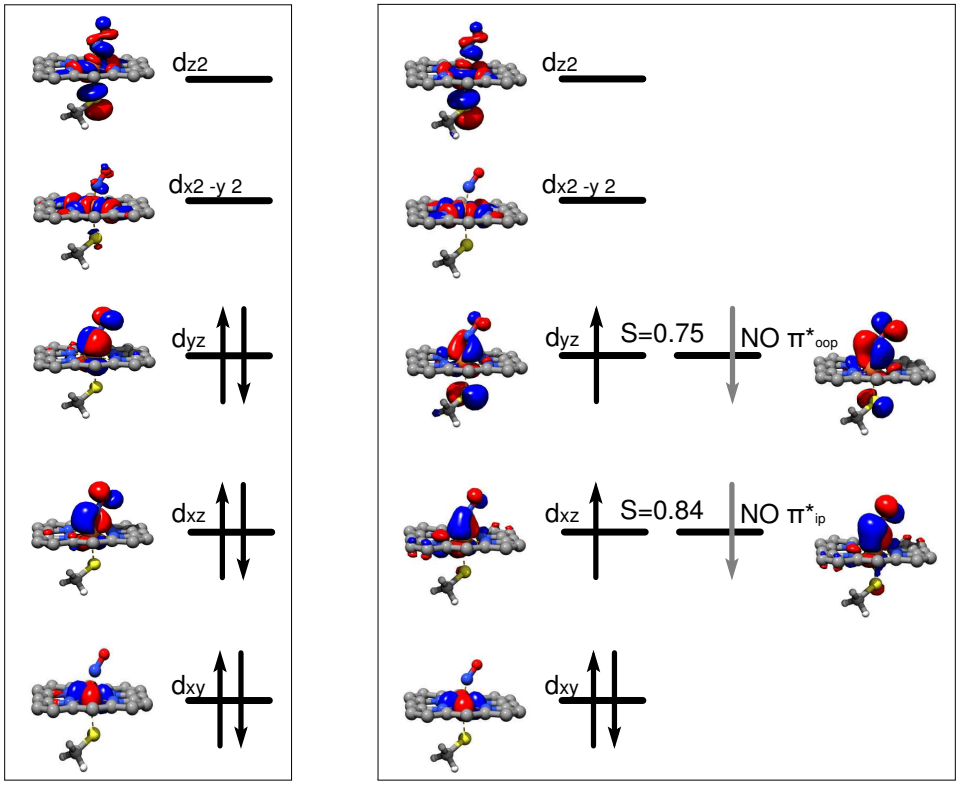


Figure 3.5. – Schematic MO diagram for **2a** (left) and **2b** (right). Spin-coupled pairs are represented by unrestricted corresponding orbitals with S being their spatial overlap. All other orbitals are localized canonical (**2a**), quasi-restricted (**2b**) and virtual orbitals. Note that these orbitals do not have well defined orbital energies.

and d_{yz}) with strong π-backbonding to the two unoccupied NO-$\pi^\star$ orbitals (both $\approx 70\%$ Fe and 30% NO-$\pi^\star$ character). This backbonding explains why the N-O bond length in the closed shell form is similar to the N-O bond length of the free NO radical (although from the occupation pattern **2a** is rather a Fe(II)-NO^+ species). Switching to the BS form, the d_{xz} and d_{yz} orbitals become polarized and give two new pairs of magnetically coupled orbitals. Here, the β spin is mainly distributed over the NO-$\pi^\star$ orbitals, and the α spin is mainly located on the Fe d_{xz} and d_{yz} orbitals. From the Mulliken analysis a spin population of $+1.14$ on Fe and -1.00 on the NO ligand is obtained. Thus, this state is best formulated as a singlet diradical involving Fe(III) and $NO^\bullet$.

[d] It has to be mentioned that this has already been observed before: Two degenerate electronic states were already found by Praneeth et al. for a six-coordinate ferric heme-nitrosyl complex with an imidazole trans ligand.[165]

Intermediate *I*

As in the case of the NO-bound state, intermediate *I* was also studied before,[88,91] but here only the closed shell solutions (**3a** and **4a**) were considered. This study is the first to obtain BS solutions for intermediate *I*. In fact, the BS solutions **3b** and **4b** will turn out to be energetically favored over the closed shell solutions **3a** and **4a** and their existence will have implications on the reaction mechanism (see Sections 3.4.4 and 4.4.3).
The LMOs of **3a**/**3b** and **4a**/**4b** are displayed in Figures 3.6 and 3.7. The two additional

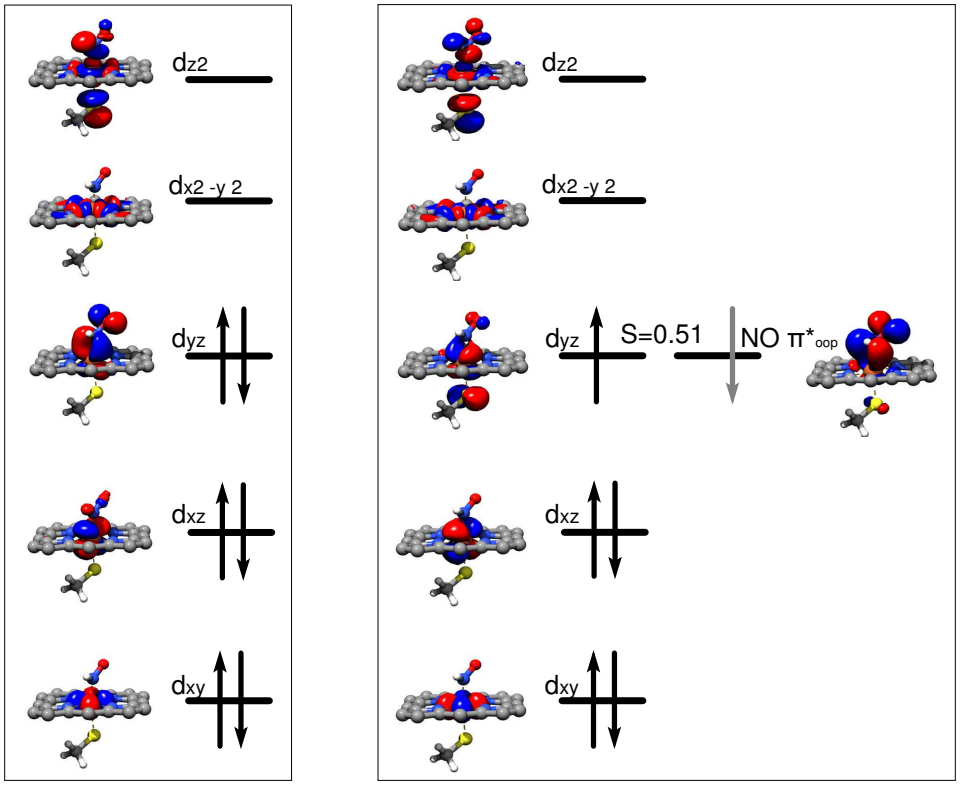


Figure 3.6. – Schematic MO diagram for **3a** (left) and **3b** (right). The spin-coupled pair is represented by unrestricted corresponding orbitals with S being their spatial overlap. All other orbitals are localized canonical (**3a**), quasi-restricted (**3b**) and virtual orbitals. Note that these orbitals do not have well defined orbital energies.

electrons in **3** and **4** (compared to **2**) occupy the newly formed N-H σ-bond. Due to the N-H σ-bond, the NO-$\pi^\star_{ip}$ orbital does not exist anymore. Back donation is thus only possible into the NO-$\pi^\star_{oop}$ orbital. All four species have two non-interacting d_{xy} and d_{xz} orbitals in common. The difference between these four species is found in the distribution of the two remaining electrons of the Fe 3d / NO-$\pi^\star$ moiety. In the BS forms (**3b** and **4b**) only one pair of magnetically coupled orbitals is present. This pair is similar to the d_{yz} / NO-$\pi^\star_{oop}$ pair in **2b**, but shows smaller UCO overlap and thus less π-backbonding character. Moreover, the UCO overlap is decreasing from **3b** to **4b**, i.e. the π-backbonding character is decreasing. This causes a lengthening of the Fe-N bond, which can be observed when going from **2b** via **3b** to **4b** (see again Section 3.4.1).
The closed shell species **3a** and **4a** differ significantly from each other. Whereas the Fe d_{yz} orbital of **3a** shows 65% Fe and 35% NO character, the composition of the corresponding orbital in **4a** is reversed (≈ 35% Fe and 65% NO character). That is in **4a** these two electrons

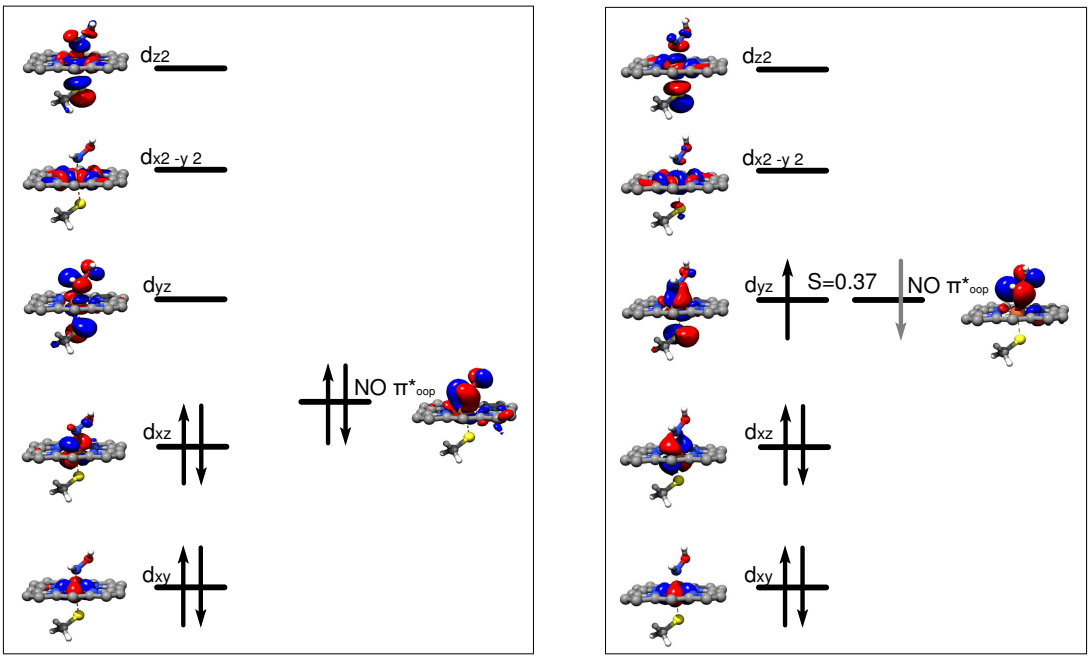


Figure 3.7. – Schematic MO diagram for **4a** (left) and **4b** (right). The spin-coupled pair is represented by unrestricted corresponding orbitals with S being their spatial overlap. All other orbitals are localized canonical (**4a**), quasi-restricted (**4b**) and virtual orbitals. Note that these orbitals do not have well defined orbital energies.

are mainly located on the NO ligand, and only two of the Fe 3d orbitals are occupied. Thus, **4a** is best described as a Fe(IV) species.
In summary, the putative intermediate *I* species can be described as:

(3a) low-spin Fe(II)-NHO,

(4a) low-spin Fe(IV)-$NHOH^-$,

(3b) $NHO^{\bullet -}$ antiferromagnetically coupled to low-spin Fe(III) and

(4b) $NHOH^{\bullet}$ antiferromagnetically coupled to low-spin Fe(III).

It has to be stressed that **3a** and **4a** have already been studied before,[88,91] but not **3b** and **4b**.
The electronic character of **3b** and **4b** is of obvious importance for the catalytic behavior of P450nor. In Chapter 4 it will be shown shown that the N-N bond formation is a spin-recoupling reaction which requires intermediate *I* to have radical character.

Fe-S Backbonding

One noticeable electronic property of all BS species is the involvement of the S p_y orbital in the hole stabilization (see Figures 3.5, 3.6 and 3.7) The Fe d_{yz} UCO has about 10–15% sulfur character for **2b**, **3b** and **4b**. This is in contrast to the closed shell species, which have essentially no sulfur contribution. This difference can partly be rationalized by the geometric differences. Due to the longer Fe-N bonds the electron density on iron is lower, which leads to more electron donation from sulfur to iron and thus to a more covalent character of the

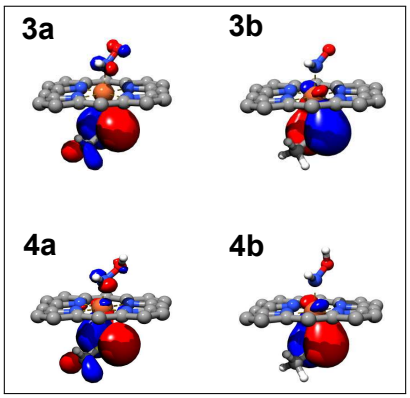


Figure 3.8. – Doubly occupied S p_y orbital for **3a**, **3b**, **4a** and **4b**. The orbitals are localized canonical (**3a**, **4a**) and localized quasi-restricted orbitals (**3b**, **4b**).

Fe-S π-bond (see Figure 3.8) for **3b** and **4b**. That is, the hole that is created at the iron center attracts electron density from the soft thiolate ligand via a Fe-S π-bond.

3.4.3. Spin Distribution

Table 3.5 summarizes the calculated spin populations of the three BS species. The spin

Table 3.5. – Computed spin population (Mulliken population analysis) for Fe, S and the N(H)O(H) ligand in **2b**, **3b** and **4b** in QM region **R1** at different geometries and with and without the protein environment. Given are the data for the enzymatic structure A with point charges/A without point charges/cluster model D (without COSMO).

	2b (NO)	**3b** (NHO)	**4b** (NHOH)
Ligand	-0.97/-0.80/-0.73	-0.83/-0.66/-0.63	-0.89/-0.90/-0.90
Fe	1.07/0.86/0.72	0.81/0.68/0.67	0.81/0.76/0.78
S	0.01/0.04/0.08	0.10/0.09/0.07	0.18/0.23/0.20

population of the singlet BS species is, of course, not measurable. As discussed elsewhere, it is an artifact of the BS method that has to be taken into account when computing spin dependent properties.[166] Here, the spin population in the singlet case signifies the antiferromagnetic coupling of the nitrosyl ligand to the Fe moiety in the single-configuration DFT framework, its strength being correlated to the diradical character of the species. It is interesting to compare the results obtained for the calculations with and without the protein matrix.

- For **2b** and **3b** the protein environment significantly increases the spin polarization between the ligand and iron-sulfur moiety (by about 25%). The protein seems to lower the diradical character.

- The diradical character of **4b** on the other hand is not influenced by the protein environment, but it is as high as that of **3b** in the protein matrix.

3.4.4. Energetics

In order to get insight into the stability of the different electronic states, the energy differences between the closed shell and the BS species were calculated for **2**, **3** and **4**.

NO-Bound State

For the $\{FeNO\}^6$ form, the two electromers are found to be very similar in energy. In Table 3.6 the energy differences for the snapshots and the cluster model are given for QM region **R1** to **R4** (if available). In comparison to the calculation with the protein environment

Table 3.6. – Computed energy differences calculated with basis set **B3** for **2a** and **2b** for different QM regions (**R1** to **R4**) in snapshots A – C and in the cluster model D. Given are the QM/MM energy differences for snapshots A – C and QM energy differences for cluster model D. The energy differences are given for **R1**/**R2**/**R3**/**R4**. In parentheses the energy differences without surrounding point charges (COSMO) for A – C (D) are given. The sign '–' indicates that the respective calculation was not performed. Positive (negative) values imply a BS(2,2) (closed shell) ground state.

Snapshot	**2** (NO)
A	+0.31 (-0.50)[a]/+0.46/–/–
B	+0.02 (-0.06)[a]/+0.08/-0.01/-0.08
C	–/+0.43/–/–
D	-0.38 (-0.03)[b]/-0.75/–/–

[a]Calculated at QM/MM geometry, but without point charges.
[b]Energy difference calculated without COSMO.

the calculation without point charges and the cluster model calculation slightly prefer the closed shell state. The same is found when enlarging the QM region to **R3** and **R4**, but these effects are only in the range of less than 1 kcal/mol.
The results for the closely lying closed shell and broken symmetry form are consistent with calculations by Praneeth et al., who studied a ferric heme-nitrosyl, but with an imidazole instead of a thiolate ligand.[165] They found a Fe(II)-NO^+ form to be only ≈ 1.0 kcal/mol more stable than a BS Fe(III)-$NO^\bullet$ form.

Intermediate *I*

The results for the putative intermediate *I* forms are summarized in Table 3.7. Compared to **2**, the ligands **3** and **4** show much larger energy differences between the two electromers. In the protein matrix, the BS species are about 4.5 to 5.5 kcal/mol more stable than the closed shell species. What is surprising at first sight is the different effect of the protein on **3** and **4**: **3a** is stabilized in the electrostatic field of the protein (energy difference increases by 2 to 3 kcal/mol compared to calculations without point charges and in the cluster model), whereas for **4a** only a small stabilization of the closed shell state by the protein environment is observed. This behavior can be rationalized when regarding the formal net charge of the respective species. **3** is best described as Fe(II)-NHO (Fe(III)-$NHO^{\bullet-}$) in the closed shell (broken symmetry) state and **4** is described as Fe(II)-$NHOH^+$ (Fe(IV)-$NHOH^{\bullet-}$) in the

Table 3.7. – Computed energy differences calculated with B3LYP* and basis set **B3** for **3a**/**3b**, **4a**/**4b** and **5a**/**5b** for QM regions **R1** and **R2** in snapshots A – C and in the cluster model D. Given are the QM/MM energy differences for snapshots A – C and QM energy differences for cluster model D. The energy differences are given for **R1**/**R2**. In parentheses the energy differences without surrounding point charges (COSMO) for A – C (D) are given. The sign '–' indicates that the respective calculation was not performed. Positive (negative) values imply a BS (closed shell) ground state.

Snapshot	**3** (NHO)	**4** (NHOH)	**5** (N)
A	+5.40 (+2.02)[a]/+5.55	+4.83(+4.01)[a]/+4.49	–/–
B	+4.53 (+1.93)[a]/+4.78	+4.30(+4.04)[a]/+4.82	-0.48/–
C	+5.76/–	+5.13/–	–/–
D	+2.50 (+2.29)[b]/+2.95	+5.79(+4.22)[b]/+5.73	–/–

[a]Calculated at QM/MM geometry, but without point charges.
[b]Energy difference calculated without COSMO.

closed shell (broken symmetry) state. This means, for **3a** there is a larger charge separation between the iron and the NHO ligand than for **3b**, whereas for **4** there is no such difference. The protein naturally stabilizes the charge separation and thus increases the energy difference for **3a** and **3b**.
The ligand **5** shows a similar behavior as the NO-bound state. Both electromers, the closed shell and the BS state, are nearly isoenergetic.

3.4.5. Spectroscopy

Mössbauer

The Mössbauer parameters were calculated for all models using QM region **R2**. For the different intermediates in the protein structure (A – C) the Mössbauer properties were additionally calculated at the protein geometry, but without the point charges, which allows to assess the effect of the protein electrostatic field on the spectroscopic properties. The calculated values for the resting state are summarized in Table 3.8. Those for the NO-bound state are found in Table 3.9 and the values for all six putative intermediate *I* models, **3a**/**3b**, **4a**/**4b** and **5a**/**5b**, are found in Table 3.10.

The observed Mössbauer parameters for the resting state are well reproduced by the QM/MM calculations. The effect of bringing the water bound heme into the protein increases the quadrupole splitting (in absolute values) by almost 1 mm/s and brings it to good agreement with the experimental value. This effect does not simply come from the different Fe-O(H_2O) distances, but essentially arises from the polarization of the QM region by the point charges. This mirrors the known steep dependence of the Mössbauer quadrupole splitting on the counter ion in low molecular weight complexes.[49] Since these are relatively small effects with large consequences, it is evident that the isomer shift is the more readily interpretable quantity.
For models **2** to **4** the effect of the protein polarization is even much more pronounced and is also consistent for all six models (see Table 3.9 and 3.10). The calculated isomer shift is lowered by up to 0.4 mm/s compared to the values without the point charges and at cluster

Table 3.8. – Computed Mössbauer parameters determined with B3LYP* and basis set **B4** for the resting state **1** at QM region **R2** in snapshots A – C and in the cluster model D. Given are the isomer shift δ and quadrupole splitting ΔE_Q.

	Snapshot	δ [mm s^{-1}]	ΔE_Q [mm s^{-1}]
Fe(III)-Wat	A	0.31	-3.0
	B	0.31 (0.31)[a]	-2.9 (-2.9)[a]
	B[b]	0.32	-2.7
	C	0.31	-2.8
	D	0.31	-2.1
Expt. value[c]		0.32	-2.96

[a]Computed at the geometry optimized with **B2**.
[b]Computed at QM/MM geometry, but without point charges.
[c]For experimental Mössbauer spectra see Appendix, Figure B.1.

Table 3.9. – Computed Mössbauer parameters for the NO-bound state (**2a**, **2b**). The parameters were determined with B3LYP* and basis set **B4** at QM region **R2** in snapshots A – C and in the cluster model D. Given are the isomer shift δ and quadrupole splitting ΔE_Q.

	Snapshot	δ [mm s^{-1}]	ΔE_Q [mm s^{-1}]
Closed shell	A	0.04	1.0
	B	0.04 (0.06)[a]	0.8 (0.9)[a]
	B[b]	0.06	0.6
	C	0.03	0.9
	D	0.06	0.7
BS	A	0.11	1.6
	B	0.10 (0.13)[a]	1.2 (1.4)[a]
	B[b]	0.14	0.9
	C	0.11	1.5
	D	0.13	1.1
Expt. value[c]		0.15	1.31

[a]Computed at QM/MM geometry optimized without solvent molecules.
[b]Computed at QM/MM geometry, but without point charges.
[c]For experimental Mössbauer spectra see Appendix, Figure B.1.

Table 3.10. – Computed Mössbauer parameters for intermediate *I* (**3a**, **3b**, **4a**, **4b**, **5a**, **5b**). The parameters were determined with B3LYP* and basis set **B4** at QM region **R2** in snapshots A – C and in the cluster model D. Given are the isomer shift δ and quadrupole splitting ΔE_Q.

		3 (NHO)		**4** (NHOH)		**5** (N)	
	Snapshot	δ [mm s^{-1}]	ΔE_Q [mm s^{-1}]	δ [mm s^{-1}]	ΔE_Q [mm s^{-1}]	δ [mm s^{-1}]	ΔE_Q [mm s^{-1}]
Closed	A	0.22	-2.1	0.12	-2.9	–	–
shell	B	0.23	-1.6	0.08	-2.9	-0.35	-0.8
	B[a]	0.26	1.3	0.12	-2.4	–	–
	C	0.20	-2.1	0.09	-3.1	–	–
	D	0.25	1.1	0.10	-2.1	–	–
BS	A	0.26	-2.3	0.25	-2.6	–	–
	B	0.26	-1.6	0.23	-2.5	-0.15	-1.0
	B[a]	0.30	-1.2	0.24	-2.1	–	–
	C	0.25	-2.3	0.26	-2.7	–	–
	D	0.27	-1.1	0.24	-1.8	–	–
Expt. value[b]		0.24	1.95	0.24	1.95	0.24	1.95

[a] Computed at QM/MM geometry, but without point charges.
[b] For experimental Mössbauer spectra see Appendix, Figure B.1.

model geometry, respectively. The absolute values of the quadrupole splitting are increased when switching from cluster model geometry to QM/MM geometry and additionally upon proceeding to electronic embedding at protein geometry.
For the two NO-bound states, **2a** and **2b**, both Mössbauer calculations agree with the experimental data.
The Mössbauer parameters for ligand **5** were calculated only for snapshot B. The calculated Mössbauer data are far off from the experimental value and thus closed shell (**5a**) and BS form (**5b**) can both be ruled out as putative intermediate *I*.
Surprisingly, comparing the calculated parameters for **3** and **4**, three of the four species have similar calculated isomer shifts. Model **4a**, shown to be the only Fe(IV) species (see Section 3.4.2), is the only one that deviates significantly from the rest of this set. Its computed isomer shift is 0.15 to 0.20 mm/s lower than that of **3a**, **3b** and **4b**, and thus does not fit to the experimental data of intermediate *I*. For the quadrupole splitting the behavior is different. Here, **3a** and **3b** are similar on the one hand and **4a** and **4b** are similar on the other hand. But, **3a**, **3b** and **4b** are still in the usual error range of such calculations. It is interesting that these four different electronic structures give only two different isomer shift values in the Mössbauer property calculations. Quite typically, the variations in isomer shifts are dominated by bond distance changes.[154,167,168] However, this appears not to be the case here. A correlation between the electron density in the iron 3d shell and the isomer shift was already shown earlier.[154,167] The more electron density is taken out of the iron 3d shell, the higher is the electron density at the nucleus and thus the higher is the isomer shift. Hence for **3a**/**3b** and **4a**/**4b** the electron density in the Fe 3d orbitals is analyzed,

since these orbitals play a crucial role in shielding the core electrons. The iron contributions to those LMOs where the ligands are interacting with the Fe 3d shell were calculated. The values are given in Table 3.11. The crucial difference between **4a**, which has the much lower

Table 3.11. – Computed electron population on Fe for important iron-ligand orbitals in **3a**, **3b**, **4a** and **4b**. Contributions derived from Löwdin reduced orbital population of spin-coupled orbitals and LMOs. Given are the data for snapshot A and QM region **R2**. Σ sums up the different contributions.

		Fe d_{xz}	S p_y	Fe d_{yz}	NO $\pi^{\star}_{oop}$	Σ
NHO	Closed shell (**3a**)	1.98	0.06	1.29		3.33
	BS (**3b**)	2.00	0.32	0.89[a]	0.09[a]	3.30
NHOH	Closed shell (**4a**)	1.98	0.12	0.74		2.84
	BS (**4b**)	2.00	0.46	0.83[a]	0.02[a]	3.31

[a]Unrestricted corresponding orbital.

isomer shift, and the other three species is readily seen. It has about 0.5 electrons less in its Fe 3d shell. All other species have different contributions to the Fe 3d population, but in net effect they have very similar 3d populations. This is surprising especially for **3a** and **3b** which have quite different electronic structures. What is important here is that the loss of electron density in the Fe d_{yz} orbital is compensated for by high electron donation from the S p_y orbital.

Infrared Spectroscopy

The N-O stretching frequencies for the NO-bound state were calculated using QM region **R2**. Unfortunately, the respective experimental information for intermediate *I* is not available. The calculated values for **2a** and **2b** are summarized in Table 3.12. The calculated values for both electromers cannot be used to distinguish between the two electromers, but a clear trend can be seen when going from **2a** to **2b**. In the QM/MM calculations the N-O stretching frequency for **2b** is constantly about 50–60 cm^{-1} lower than that for **2a** (whereas the respective shift is ≈ 100 cm^{-1} without the protein). This shift corresponds to the shift which is observed experimentally when comparing NO-bound P450nor and NO-bound P450cam, which supports the identification of NO-bound P450nor as the closed shell species **2a** (compare to the results on the Fe-N bond length in Table 3.3).

3.5. Discussion and Conclusion

This chapter was aimed at obtaining a complete description of the reaction intermediates of P450 NO reductase. It was shown that there are two possible electromers for the NO-bound state. Six species were studied as putative intermediate *I* forms.

Through comparison between theory and experiment, the NO-bound species could be identified as the closed shell form **2a** of NO^+ bound to Fe(II) with characteristically strong back-donation from iron to the nitrosyl ligand. Here, the calculated Mössbauer parameters and Fe-N bond length agree well with the respective experimental results. It is tempting to speculate that the BS(2,2) $\{FeNO\}^6$ species **2b**, which is isoenergetic to the closed shell

Table 3.12. – Computed IR frequencies for the NO-bound state (**2a**/**2b**). The parameters were determined with B3LYP and basis set **B1** at QM region **R2** in snapshots A – C and in the cluster model D. For A – C the electrostatic and LJ interaction with the protein is included. For D the frequencies are calculated using COSMO with $\epsilon = 4.0$. Given are the harmonic vibrational frequencies scaled with an appropriate scale factor.[a]

Snapshot	ν(N-O)	ν(^{15}N-O)
A	1839 / 1777	1803 / 1743
B[b]	1922 / 1872	1884 / 1835
C	1832 / 1772	1796 / 1783
D	1898 / 1800	1861 / 1766
Exp. P450nor	1851	1820
Exp. P450cam	1806	1775

[a]The scaling factor for the determination of the frequencies was taken from[169] for B3LYP and a similar basis set.

[b]The frequencies for snapshot B deviate from those of the other snapshots. This is due to the same reasons as for the structural deviation, see Section 3.4.1.

form, might play a role in the NO-bound form of P450cam.[86] Although the respective QM/MM calculations with P450cam were not carried out, the cluster model calculations show that the Fe-N bond length in **2b** is 10 pm longer than in **2a**, which is also the case when comparing the EXAFS Fe-N bond length of NO-bound P450nor (1.66 Å) with that of NO-bound P450cam (1.76 Å). Also the N-O IR stretching frequencies ν(N-O) show a trend supporting this hypothesis: The experimental ν(N-O) in NO-bound P450cam is 45 cm^{-1} lower than in NO-bound P450nor.[86] On the computational side a shift of 100 cm^{-1} is found for the cluster model ν(N-O) when going from **2a** to **2b**, and this shift is becoming smaller (50–60 cm^{-1}) upon comparing the respective QM/MM results for P450nor.

Fewer experimental data are available for intermediate *I* in the P450nor reaction cycle. The properties of six putative intermediates were calculated and studied in detail, the closed shell and BS forms of singly and doubly protonated $\{$FeNO$\}^8$, named **3a**/**3b** and **4a**/**4b** in the above analysis, and the closed shell and BS (dehydrated ligand **4**) $\{$FeN$\}^8$, **5a** and **5b**. With the aid of the experimentally obtained Mössbauer parameters, **4a**, **5a** and **5b** could be excluded out of the set, although **4a** was proposed to represent intermediate *I* previously.[91] Apart from this result, **3b** is favored energetically over **3a** by about 5 kcal/mol, which leaves **3b** and **4b** as putative intermediate *I* species. Both forms are Fe(III) species, differing only in the protonation state of the antiferromagnetically coupled ligand: Fe(III)-NHO$^{\bullet-}$ (**3b**) or Fe(III)-NHOH$^{\bullet}$ (**4b**). The diradical character of these two proposed species is of importance for the mechanism of P450nor. The decomposition reaction of intermediate *I* can easily be initiated by a spin-recoupling radical reaction of **3b** or **4b** with a second NO molecule, as will be discussed in the next chapter. The question, which of both species (**3b** or **4b**) represents intermediate *I*, will also be investigated in Chapter 4.

4. Characterization of the Reaction Kinetics of P450nor

4.1. Introduction

In the previous chapter the intermediates of the P450nor reaction cycle were studied. Still, several aspects of the mechanism remain open.

- The reduction of Fe(III)-NO by NADH was shown experimentally to be the rate limiting step of the reaction cycle.[89] The cofactor reduces the active site without additional electron carriers, and two experimental studies could show that the reduction is done by NADH directly in the active site.[74,78] From the computational side neither the binding of NADH to the active site, nor the direct hydride transfer to Fe(III)-NO has yet been studied.

- Several residues in P450nor are experimentally known to be essential for the catalytic mechanism (see Section 1.2).[70,84] To date there exists no theoretical study which gives a complete picture of the role of these residues in the enzymatic reaction.

- Intermediate *I*, which is produced in the reduction step, is considered as the key intermediate in the reaction mechanism. Under the commonly employed higher NO concentrations ([NO]=1.0-2.5 mM, compared to the enzyme concentration of [E]=5.4 nM), intermediate *I* does not accumulate in measurable quantities.[69] However, at low NO concentrations ([NO] almost equal to [E]) it can be spectroscopically detected. Hence, it must represent a structure that occurs in the reaction cycle prior to N-N bond formation. In the previous chapter two putative intermediate *I* species could be selected: Fe(III)-NHO$^{\bullet-}$ (**3b**) and Fe(III)-NHOH$^{\bullet}$. Nevertheless, conclusive evidence for the nature of *I* has not yet been obtained.

Thus, despite accumulating experimental and theoretical data, the complete reaction cycle of P450nor has not yet been established.
In this chapter QM/MM calculations are again used. The presumed intermediates are connected through reaction paths. Three snapshots are used as starting point for the reduction by NADH and two different QM regions are used, where the larger one includes all residues for cofactor binding near the heme reaction site. For protonation and decomposition one snapshot with a sufficiently large QM region is investigated.

- The first aim of this chapter is to obtain the *first complete theoretical picture of the reduction of NO-bound P450nor in a 'realistic' enzyme environment*. The dynamics of NADH in the protein and the structural influences of the protein on the reduction step are studied. With the help of DFT calculations the mechanism of the electron transfer is investigated. Since rate constants and the kinetic isotope effect (KIE) for the reduction step are experimentally known, these data can be used to validate the theoretical results.

- The results of the MD simulation and QM/MM calculations are correlated with site-directed mutagenesis experiments that shed light on the role of individual amino acids.

- The entire reaction course of the P450nor reaction cycle is simulated using the calculated rate constants. Using this approach, the reaction course can be simulated under varying reaction conditions. Herewith, it is aimed to reveal the structural and electronic nature of intermediate *I* and to explain the mechanism behind the behavior under varying NO concentrations.

4.2. Computational Details

In the calculations presented in this chapter the QM/MM workflow described in Section 2.2.6 was applied.[a] Specific details of this study are given in the following.

4.2.1. Preparation of the Protein Structure and MD Simulation

For the kinetic study two different enzyme structures were used (see also Section 1.2.2). The first one, **P1** (PDB code: 1XQD[74]), has NAAD bound in its active site. The cofactor was chemically modified by replacing the carboxyl group with an amide group. **P1** could not be used for the modeling of the entire reaction cycle. Due to the presence of the cofactor there is not enough space in the heme distal pocket for the attack of another NO molecule and for the rearrangements of the reactant.[b] Thus, a second enzyme structure without NADH, **P2** (PDB code: 1CL6[70]), was used for the other reaction steps (**P2** is the system also used in Chapter 3). The protonation states of the residues were assigned using the PROPKA 2.0 program (Glu, Asp) together with the default algorithm of GROMACS (other residues).[102,132,133] Asp_{393}, which is essential for the proton transfer in P450nor[70,72,74,84] and which has a relatively high pK_A value of 6 in P450nor (according to PROPKA), was protonated. It can thus function as proton donor for the reaction.
For both enzymatic systems the same procedure as in the previous chapter was applied. The OPLS-AA force field was used.[134–136] The topologies of the non-standard residues (cysteinate-ligand to iron, heme-group, NADH and the nitrogenous ligand) were generated using PRODRG2.5.[121] Atomic MM charges were derived as CHELPG charges[122] with ORCA[146] (BP86[137–139]/SV(P)[140–142]). The protein was put in a rhomboedric box (protein dimensions plus 2 nm in each direction) and solvated with water molecules. In order to neutralize the system, solvent water molecules (not crystal waters) were replaced by sodium ions using the genion algorithm,[102] in order to arrive at an overall neutral charge. The resulting parameters of this procedure are summarized in Table 4.1.
The prepared system was geometry relaxed and equilibrated through classical energy minimization and classical MD runs with periodic boundary conditions (see also Section 2.2.6). During the classical minimization and MD simulations, the Cartesian coordinates of the iron ion and its ligands (heme, Cys_{352} and NHO ligand) were kept fixed.
In **P2**, a second NO molecule had to be introduced for the QM/MM calculations. This was done by replacing a water molecule in the active site by NO, about 3.6 Ångstrom away from NHO.

[a] The computational methodology used in this chapter is mostly similar to that in the previous chapter, but important differences exist.

[b] It was also shown experimentally that after reduction NADH first has to leave the active site, before the reaction can proceed further.[80]

Table 4.1. – System properties of the prepared systems **P1** and **P2**.

	P1	**P2**
Box dimensions [nm × nm × nm]	$8.4 \times 6.9 \times 7.9$	$7.5 \times 8.2 \times 8.2$
Nr. of solvent molecules	14199	12742
Nr. of sodium ions	8	7
Nr. of active atoms	995/988/1025 (snapshot A/B/C)	827
Nr. of QM atoms	61/85 (**R1**/**R2**)	76 (**R3**)

4.2.2. QM and QM/MM Calculations

The QM/MM calculations started from three snapshots of **P1** (snapshot A/B/C taken at 150/250/350 ps) and one snapshot of **P2** (taken at 350 ps). In the reaction course different ligands occupy the sixth coordination site of heme. The most important ones are depicted in Figure 4.1.[c] Depending on the ligand charge, the total system charge was adapted by deleting sodium ions in order to maintain overall neutrality. Four different QM regions were used (see Figure 4.1). For **P1**, QM region **R1** consists of the heme-thiolate-NO unit together with the nicotinamide moiety. QM region **R2** additionally incorporates Ser_{286} (modeled as methanol), Thr_{243} (modeled as ethanol) and the peptide bond between Ala_{239} and Gly_{240}. For **P2**, QM region **R3** was used, which consists of the heme-thiolate-NHO unit, Ser_{286} (modeled as methanol), Asp_{393} (modeled as acetic acid) and six water molecules. QM system **R4** comes with the heme-thiolate-NO unit and one water molecule. For **P1**, all residues that have atoms within a distance of 4 Å of the heme-group, Cys_{352} the NO ligand and NADH, were optimized. The active region for **P2** includes all residues that have atoms within 4 Å of **R3** (see also Table 4.1).
QM/MM calculations were carried out with an electrostatic embedding scheme using the QM/MM interface between GROMACS and ORCA.[102,143] Periodic boundary conditions were switched off and no cutoffs were used for the nonbonded QM-MM interactions. For the bonds connecting the QM/MM boundary the hydrogen link atom approach,[143] applying the charge shift model,[123] was used.
Gas phase calculations on **R4** were carried out modeling the protein environment through a dielectric continuum using COSMO[144] with a dielectric constant of $\epsilon = 4$.[145]

4.2.3. Reaction Path Scans

The QM/MM optimizations were performed according to the standard workflow described in Section 2.2.6. Reaction path scans with QM regions **R1**/**R3** (in systems **P1**/**P2**) were performed using the GROMACS conjugate gradient (cg) minimizer. Here, successive distance constrained QM/MM minimizations were carried out using a step size of 0.1–0.2 Ångstrom, in order to obtain a smooth reaction path. Stationary points for **R1**, **R2** and **R3** (minima

[c] Here the same nomenclature as in Chapter 3 is used.

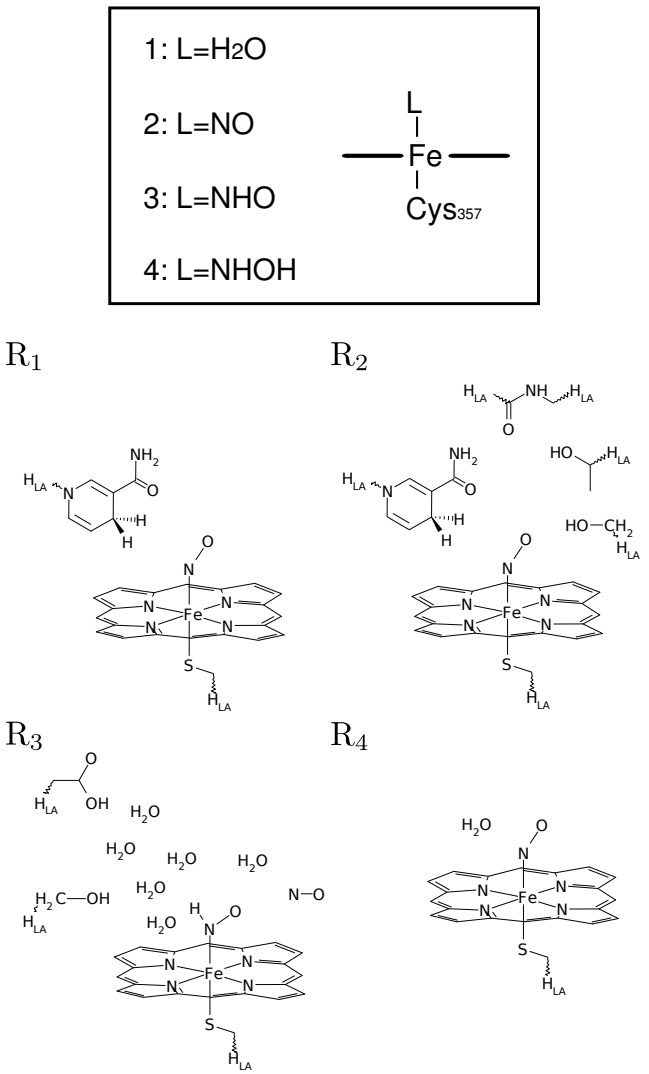


Figure 4.1. – QM systems **R1** to **R4** and ligands **L1** to **L4** which characterize the different states of the protein.

and transition states) were further optimized using the ORCA optimizer.[d]

4.2.4. QM method

The same QM methodology was used as in the work described in the previous chapter. The QM/MM optimizations were performed with unrestricted RIJCOSX[146]-B3LYP[137,147,148] using the def2-SVP basis set together with the def2-SVP/J auxiliary basis set (**B1**).[140,149,150] Single-point calculations were carried out with RIJCOSX-B3LYP*[152,153] using the def2-TZVPP basis set together with def2-TZVPP/J as auxiliary basis set for all atoms (**B2**). In all calculations ZORA was used to account for relativistic effects.[155–157] Grimme's 2006 semiempirical dispersion correction was included to account for dispersion effects within the QM part.[158] The MM point charges that represent the protein environment were included

[d]For **P1**, QM region **R1** turned out to be too small to obtain reliable relative energies. Thus, for **P1** we will focus on the results obtained with **R2**.

into the one-electron Hamiltonian of the QM-calculation.
For the frequency calculations, the partial Hessian vibrational analysis[125] was used at the same level of theory as for the optimization. From these frequency calculations the zero point energies (ZPE) for the relative energies were derived.

4.2.5. Kinetic Isotope Effect

The kinetic isotope effect (KIE) of the NADH reduction step was calculated by applying the following equation:

$$KIE = e^{(E_{a,D}-E_{a,H})/RT} , \tag{4.1}$$

with $E_{a,H}$ and $E_{a,D}$ being the respective activation enthalpies for hydrogen and deuterium, R being the gas constant and T the absolute temperature.

4.2.6. Kinetic Simulation of the Reaction Course

For the kinetic simulation of the entire P450nor reaction course the program Tenua,[170] which is based on the program KINSIM,[171] was used.

4.3. MD Simulation Results

Even though the MD simulations are only used as a preparation step before the QM/MM calculations, some important conclusions can be drawn from the MD runs. Two systems are studied, one with NADH (**P1**) and one without (**P2**).

4.3.1. P450nor with bound cofactor (P1)

After having prepared the system as described above, it was fully equilibrated after 50ps and 350ps of productive NVT sampling were performed (see Figure 4.2). Despite the structural

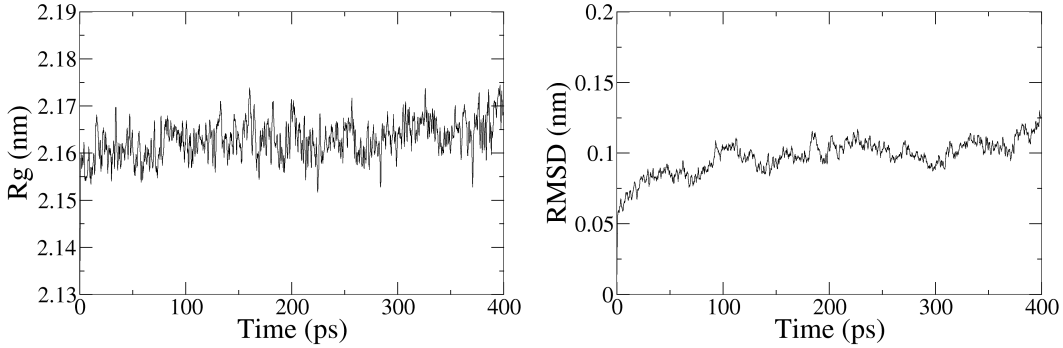


Figure 4.2. – The gyration radius (Rg) of the protein (left) and the root-mean-square deviation (RMSD) of the backbone atoms (right) with respect to the initial configuration of **P1** are plotted versus the MD simulation time.

and chemical modifications compared to the crystal structure of 1XQD (NAAD → NADH), NADH stays bound in the active site. The orientation of the NADH nicotinamide group with respect to the ligand should play a crucial role in the reduction step. Visual inspection of the

trajectory shows that NADH stays in the active site near the NO ligand, stabilized by the surrounding amino acids. The resulting distance distribution of NADH to the NO ligand can be seen in Figure 4.3. The hydrogen atom, which is carrying the reducing electrons, shows in average a distance of 2.7 to 3.7 Å to N(NO). Taking a closer look at the interaction of

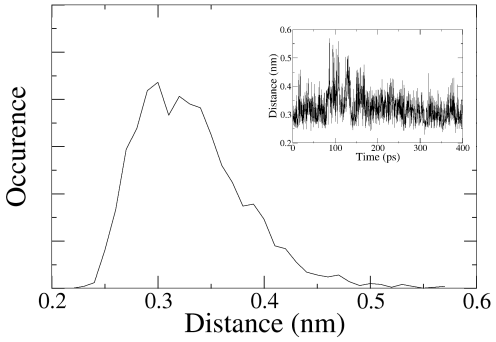


Figure 4.3. – The relative distribution of the distance between H(NADH) and N(NO) during the MD simulation of **P1** (from 50ps to 400ps) is displayed. The inlet shows the distance plotted versus the time domain for the whole run.

nicotinamide in the heme distal pocket, it is found that the only stabilization comes from hydrogen bonds to the amide group of nicotinamide. In order to reveal the contributions to the stabilization, the interaction strengths of the important hydrogen bonds are analyzed in Figure 4.4. The amino group switches between two different types of hydrogen bonds. The competing hydrogen bond acceptors are:

1. the phosphate group of NADH (intramolecular hydrogen bond, Figure 4.4, C) and
2. both the side chain of Thr_{243} (Figure 4.4, D) and the backbone carbonyl group of Ala_{239}. (Figure 4.4, B)

The intramolecular (C) and intermolecular (B and D) hydrogen bonds alternate. The histograms for these distances show thus a broadening between 3–6 Å. *The carbonyl group of nicotinamide forms a constant hydrogen bond with the OH-group of* Ser_{286} (Figure 4.4, A) and thus plays the most important role in the placement of the hydride donor near the NO ligand.

Conclusions It is known from experiment that Ser_{286} and Thr_{243} are impotant for the P450nor reaction.[70,72,74,81–84] From the MD results it emerges that Ser_{286} and Thr_{243} both form stable hydrogen bonds with NADH in the active site and thus pull it close to the NO ligand. A further crucial interaction exists that cannot be deduced from mutation studies, namely an interaction of NADH with the backbone of Ala_{239}.[e]

4.3.2. P450nor without cofactor (P2)

After the preparation of the NHO-bound P450nor 450ps of productive sampling were performed (see also Section 3.3).

[e]With mutagenesis studies only the relevance of side chain specific interactions can be probed, as the backbone is not altered.

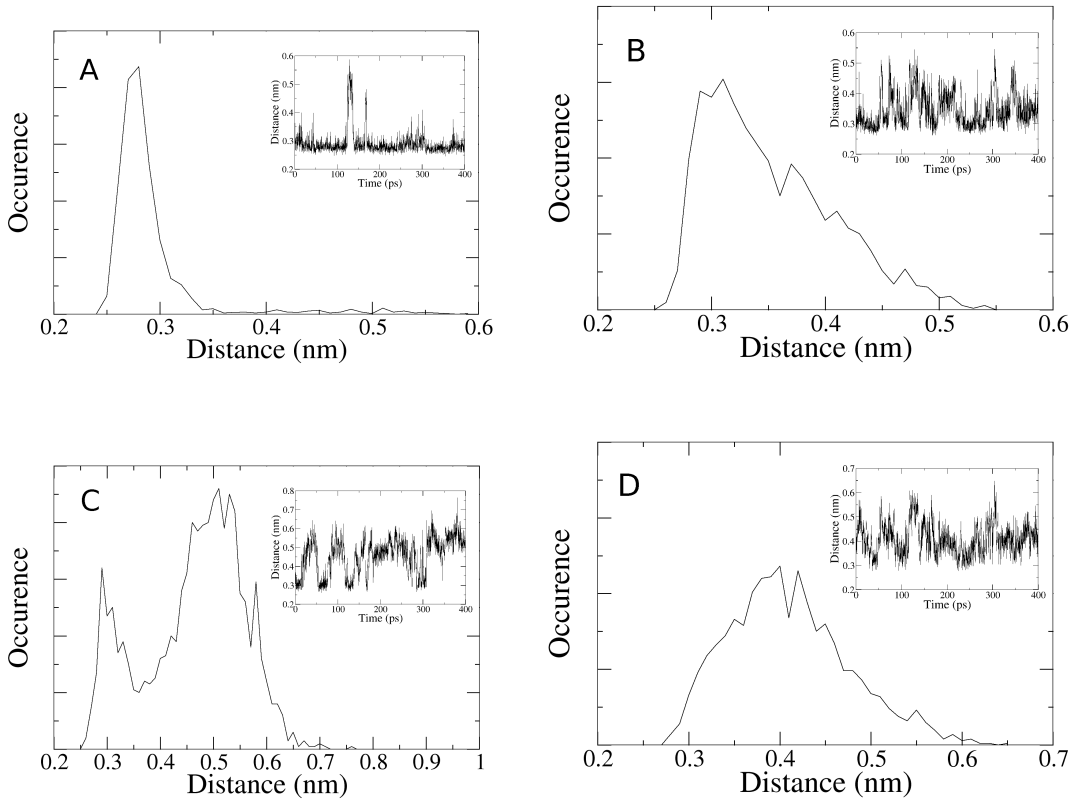


Figure 4.4. – Histograms showing the relative distribution of the hydrogen bond lengths for the important hydrogen bonds of NADH in the active site during the MD simulation from 50ps to 400ps. The inlets show the respective hydrogen bond length versus the time domain. A: Distance between Ser_{286}-OG and NADH-amide O. B: Distance between Ala_{253}-(C=)O and NADH-amide N. C: Distance between NADH-O_{12} (phosphate) and NADH-amide N. D: Distance between Thr_{239}-OG and NADH-amide N.

The MD simulation trajectory for **P2** shows that the water molecules do not escape from the active site. Their number stays relatively constant (visual inspection). During the MD simulation Ser_{286} and Asp_{393}, which belong to the proton transfer chain,[70,72,74,84] form hydrogen bonds to water molecules throughout the simulation: Ser_{286} has one or more hydrogen bonds in about 95% of the simulation time. Here, e.g. crystal water WAT_{478} is hydrogen bonded in 90% of the simulation time.
In the course of the MD run the carboxyl group of Asp_{393} switches many times between two conformations. In about half of the time it is hydrogen bonded to the backbone-NH group of Ala_{287}, in the remainder it points towards the protein exterior (see Figure 4.5). The flexibility of the carboxyl group probably plays a role in the protonation mechanism. After being deprotonated, it can easily flip towards the bulk solvent and be reprotonated in order to facilitate the next protonation step.

Conclusions Experimentally it has been inferred that proton donation happens via Asp_{393}. A possible mechanism for this function involves the high flexibility of Asp_{393}, which allows it to adopt two different positions, one oriented versus the protein exterior and one versus

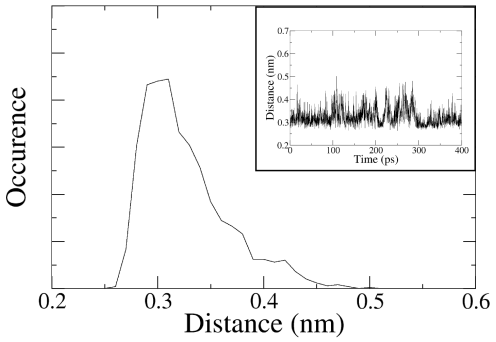


Figure 4.5. – The relative distribution of the hydrogen bond distance of the carbonyl group of Asp_{393} and the NH-group of Ala_{287} during the MD simulation from 50ps to 400ps. The inlet shows the distance plotted versus the time domain for the whole run, including equilibration.

the active site. This situation hints towards a 'paddle' mechanism that enables the protein to supply the active site with protons (see Figure 4.6).

4.4. QM/MM Results

In this section, the steps connecting the three key intermediates (see also Figure 1.8) are studied. First, NO binds to the enzyme's resting state (**1** → **2**). This is followed by hydride transfer (**2** → **3**) and, possibly, a consecutive protonation step (**3** → **4**) that occurs only for the double protonated variant of intermediate *I*. Finally, the product decomposition takes place that is initiated by the reaction of intermediate *I* with a second NO molecule.

4.4.1. NO Binding to Heme (1 → 2)

Unfortunately, it was not possible to find a TS in the QM/MM calculations for the ligand exchange from water-bound to the NO-bound state. This is probably due to the many interactions that NO and water undergo with the surrounding water molecules. Instead, a two-dimensional relaxed surface scan was carried out (scanning both the Fe-O(H_2O) and the Fe-N(NO) distances) in the absence of the protein matrix (see Figure 4.7). In this calculation, a pseudo-TS with a very low imaginary frequency (-100 cm^{-1}) was found, thus indicating an exceedingly flat potential barrier for this reaction step. This TS is denoted by **TS**(**1**,**2**).
TS(**1**,**2**) decomposes into the NO-bound state **2** and a water molecule. As already shown in the previous chapter, **2** can exist in two electromeric states, a closed shell (**2a**) and a broken symmetry, BS(2,2) state **2b**. For the further elucidation of the reaction process the kinetic and thermodynamic properties of the NO binding process are important.

- The activation enthalpy of the ligand exchange (**1** → **TS**(**1**,**2**)) is 5.4 kcal/mol.
- The reaction enthalpy ((**1** → **2a**)/(**1** → **2b**)) is -4.3/-4.1 kcal/mol.

The calculated value for the activation enthalpy $\Delta H^{\ddagger}$ is in good agreement with the experimentally derived value for the NO-bound state formation, which is 2.6 x 10^7 M^{-1} s^{-1} ($\Delta H^{\ddagger} = 6.9$ kcal/mol at 10°C).[69]

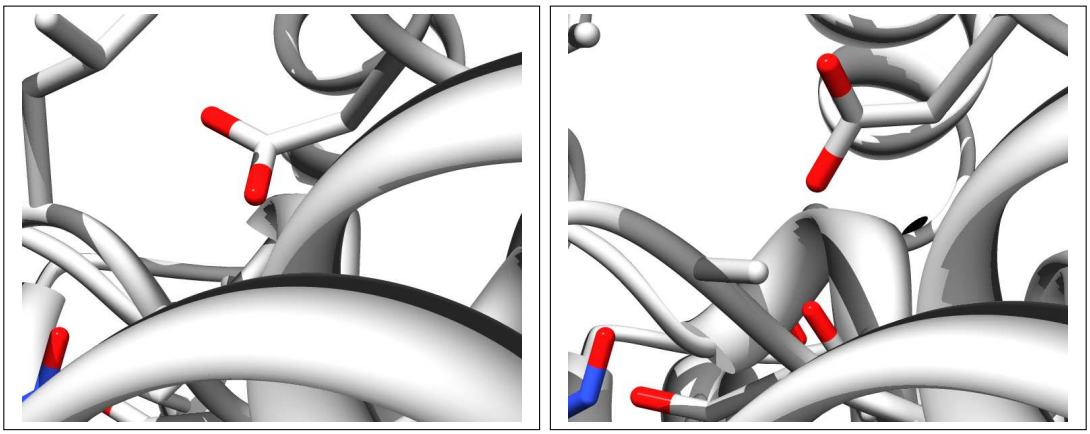


Figure 4.6. – Asp_{393} in two different conformations, one pointing towards the active site (left figure, 160ps) and one pointing towards the exterior (right figure, 260ps). The protein is shown in cartoon representation, Asp_{393} and surrounding residues are shown in stick representation (for the purpose of clarity the hydrogen atoms are not shown).

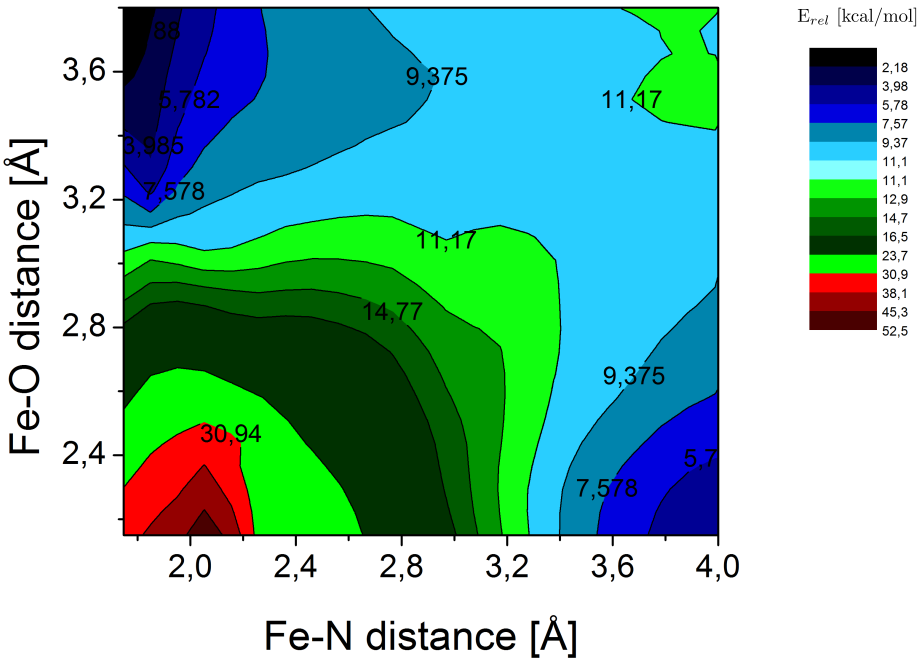


Figure 4.7. – Water-to-ligand exchange in P450nor. Relative energies (in kcal/mol) for a two-dimensional relaxed surface scan of Fe-N(NO) and Fe-O(H_2O) are displayed. The scan was performed with QM region **R4** and basis set **B1**, followed by single point calculations using B3LYP*/**B1**. COSMO ($\epsilon = 4$) was used to model the protein environment.

4.4.2. Hydride Transfer by NADH (2 → 3)

P450nor is one rare example of a reductase, where NADH reacts directly at the protein's active site (see Section 1.2.2). Therefore, this reaction step is studied in depth in an attempt to elucidate how the electrons and the proton are being transferred to the substrate.[f] Possible mechanisms are

- reduction followed by proton transfer,
- protonation followed by reduction and
- concomitant proton and electron transfer in form of a hydride ion.

QM/MM surface scans were performed on three different snapshots of **P1** (at 150ps, 250ps and 350ps, called snapshot A, B and C) with subsequent optimizations to connect the NO-bound state **2** with the reduced state **3** via the corresponding transition state **TS**(**2**,**3**).[g]
As shown in the previous chapter, the product of this reaction step can exist in two electronic states, a closed shell (**3a**) and a BS(1,1) (**3b**) species. Thus, the reaction (**2** → **3**) was calculated on two different PESs, (**2a** → **3a**) and (**2b** → **3b**), via TSs **TS**(**2**,**3**)**a** and **TS**(**2**,**3**)**b**.

Structural Properties

The structural properties of the central Fe-NO unit of the intermediates and the TS are summarized in Figure 4.8.[h] The properties of the Fe-NO unit of the educt and product were already described in detail in Section 3.4.2. However, since here the reduction of the NO-bound state is modeled *including the actual reducing agent*, namely NADH, the special properties of NADH binding to the active site for the educt, TS and product deserve to be highlighted.
Nicotinamide is found in proximity to the NO ligand in the educt state with its ring plane facing the NO ligand. The hydrogen, which is to be transferred to NO, is pointing directly towards the NO ligand, ready for the nucleophilic attack. Moving further to the TSs of the reaction step, **TS**(**2**,**3**)**a** and **TS**(**2**,**3**)**b**, the hydrogen is located in between the carbon and the nitrogen atom, where **TS**(**2**,**3**)**a** is more product-like and **TS**(**2**,**3**)**b** is more educt-like. In the product state, **3a** and **3b**, the nicotinamide ring plane is still facing the ligand, and has a similar distance to the iron as in the educt state. One noticeable feature is that the Fe-N bond in the TSs adopts already the length of the respective product state.
In Section 4.3 the stability of the hydrogen bonds between the amide group and the active site surrounding was already pointed out. Importantly, the two hydrogen bonding partners Thr_{243} and Ser_{286} are *the only hydrophilic residues in the heme distal pocket*. They are located near the reaction site and their interaction with the nicotinamide ring is such that

1. NADH comes close enough to the reaction site and

[f] In the calculations it was also tried to first protonate the NO ligand and subsequently reduce NHO^+ by NADH, which should be even more favorable, since the NHO^+ ligand is much more electrophilic than the NO ligand. However, the attempts to optimize a state with protonated NO ligand failed. The attached proton always spontaneously transferred back to the proton donor Asp_{393} in the course of the optimization.

[g] Optimizations were done with QM regions **R1** and **R2**. **R1** was found to be too small to deliver reliable results (see Table B.1), thus only the results for **R2** are presented here.

[h] Here only the values for snapshot B are shown. Snapshot A and C show very similar properties (differences < 0.02 Å for Fe-N and N-O bond distances and $< 4^\circ$ for the Fe-N-O bond angle), see also Table B.2 in the Appendix.

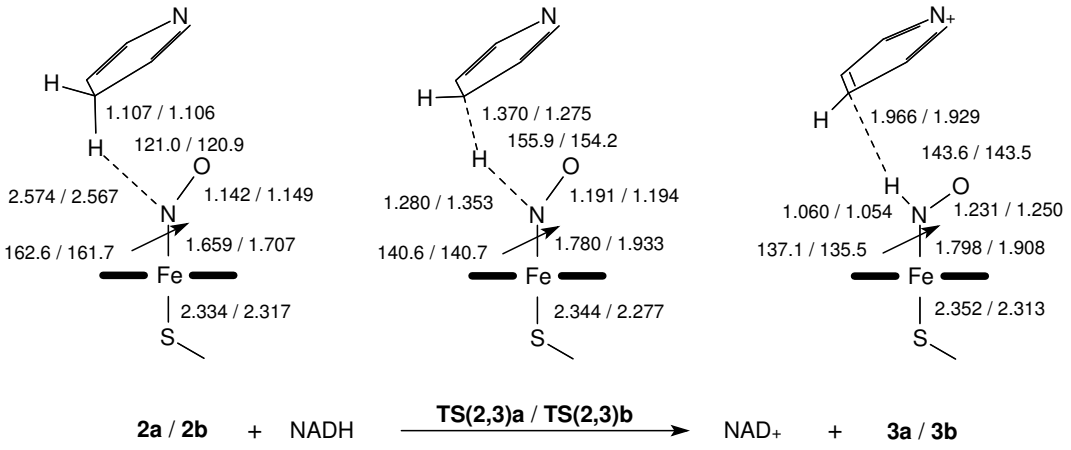


Figure 4.8. – Important structural data for the educt (left), transition state (middle) and product (right) of the hydride transfer reaction. The structural parameters refer to closed shell / BS form obtained for snapshot B and QM region **R2** in system **P1**. Bond distances are given in Ångstrom, angles are given in degree.

2. the nicotinamide ring points with pro-R hydrogen towards the nitrogen.

Remarkably, the hydrogen bonds are not only responsible to bring NADH near the NO ligand, but also to *stabilize the TS of the hydrogen transfer step*. The key parameters for the three hydrogen bonds (the two mentioned above and an additional hydrogen bond to the protein backbone) in the TS of all three snapshots are given in Figure 4.9 and Table 4.2.

The hydrogen bonds are present in all snapshots, but their strength varies from snapshot to

Table 4.2. – Relevant structural data for the transition state stabilization of the hydride transfer reaction for all three snapshots. The structural data refer to the QM/MM optimization of QM-region **R2**. Distances are given in Ångstrom. An overview of the interactions in the active site is given in Figure 4.9. The following abbreviations are used for the hydrogen bond distances: Ser-HB: O(NADH-C=O) - H(Ser_{283}-OH), Thr-HB: H_a(NADH-NH_2) - O(Thr_{243}-OH) and Ala-HB: H_b(NADH-NH_2) - O(Ala_{239}-C=O).

Distance	Electromer	
	TS(2,3)a	**TS(2,3)b**
N-H	1.27/1.28/1.30	1.33/1.35/1.38
C-H	1.38/1.37/1.35	1.29/1.28/1.26
Ser-HB	1.81/1.77/1.71	1.77/1.77/1.71
Thr-HB	2.66/2.38/1.70	2.79/2.41/1.71
Ala-HB	2.71/1.71/1.74	2.71/1.73/1.76

snapshot. Inspection of the protein structures shows that this variation is *depending on the initial snapshot structures* and consequently depends on the strain of the protein backbone, reorienting the different hydrogen bonding partners relative to each other. Overall, the *most stable hydrogen bond is the one to* Ser_{286}.

The effects of the varying hydrogen bond strengths on the TS structure are differences in

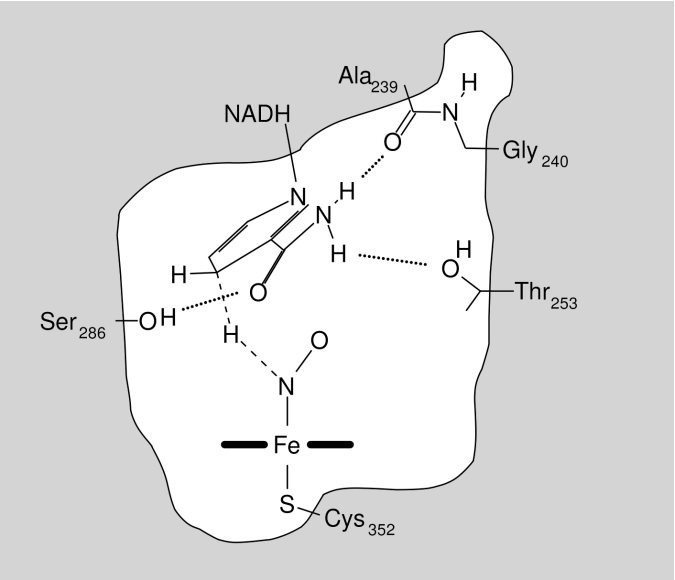


Figure 4.9. – Structural stabilization of NADH in the active site by three different hydrogen bonds of nicotinamide with functional groups from the protein. The grey region indicates the MM subsystem and the white region indicates the QM subsystem **R2**.

the N-H and C-H distances of up to 5 pm. The activation barrier of the hydride transfer is consequently affected (*vide infra*).

Reaction Energetics

The relative QM/MM energies for the intermediates of both PESs are given in Table 4.3.

Table 4.3. – Calculated energies for the intermediates of the hydride transfer reaction (see Figure 4.10) for the three snapshots. The QM/MM energies refer to the structures obtained for snapshot A/B/C with QM region **R2**. Relative energies are given in kcal/mol and refer to the lowest energy of both electronic states of the respective snapshot. The QM/MM energies already include ZPE correction.

	Closed shell	BS
NADH + $\{FeNO\}^6$	0.0/0.0/0.0	0.1/0.1/0.1
TS	13.2/9.3/10.8	10.1/6.5/8.6
NAD^+ + $\{FeNHO\}^8$	9.2/7.3/5.9	5.0/3.0/2.0
Exp. barrier[69]	≈ 8.8	

It can be seen that reaction via **TS**(**2**,**3**)**b** is faster than via **TS**(**2**,**3**)**a**. On both surfaces, the reaction is endergonic, but the energy cost to obtain **3b** is lower than to obtain **3a**.
It can also be seen that the height of the reaction barrier varies with the three snapshots. Taking into account the varying stabilization of the TS in the different snapshots, it can be concluded that the different barriers are due to the different protein conformations in the snapshots.

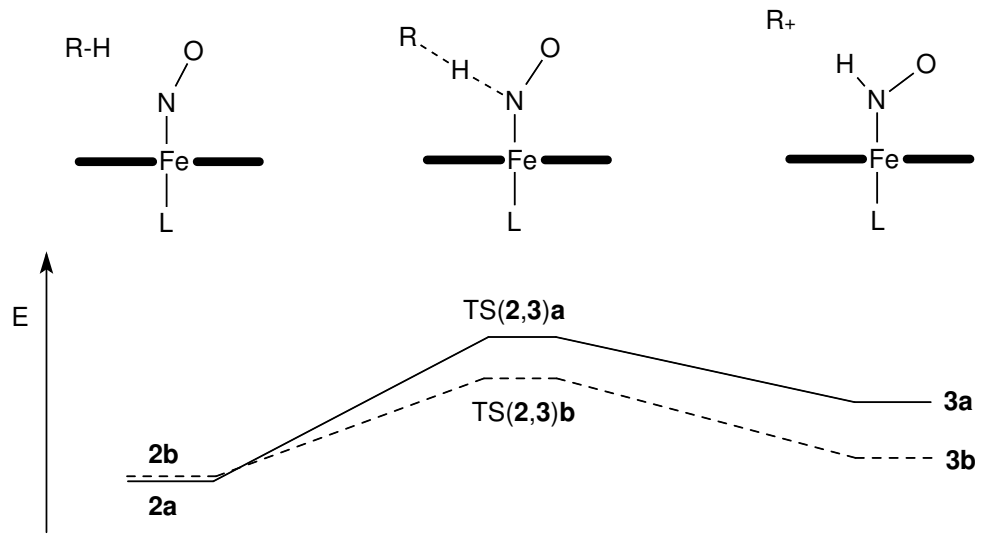


Figure 4.10. – Schematic representation of the hydride transfer reaction. The calculated energies (including ZPE) for the intermediates are summarized in Table 4.3.

Bonding Properties

The unusual direct reduction by NADH in the active site raises the question, how the electrons and the proton are transferred to the NO-bound species. Some insight can be obtained by analyzing the electronic structure of the TSs. The localized orbitals[i] carrying the electrons of the pro-R hydrogen for **TS**(2,3)**a** and **TS**(2,3)**b**, respectively, are visualized in Figure 4.11. In both TS structures, the respective orbital is doubly occupied. The main part of the

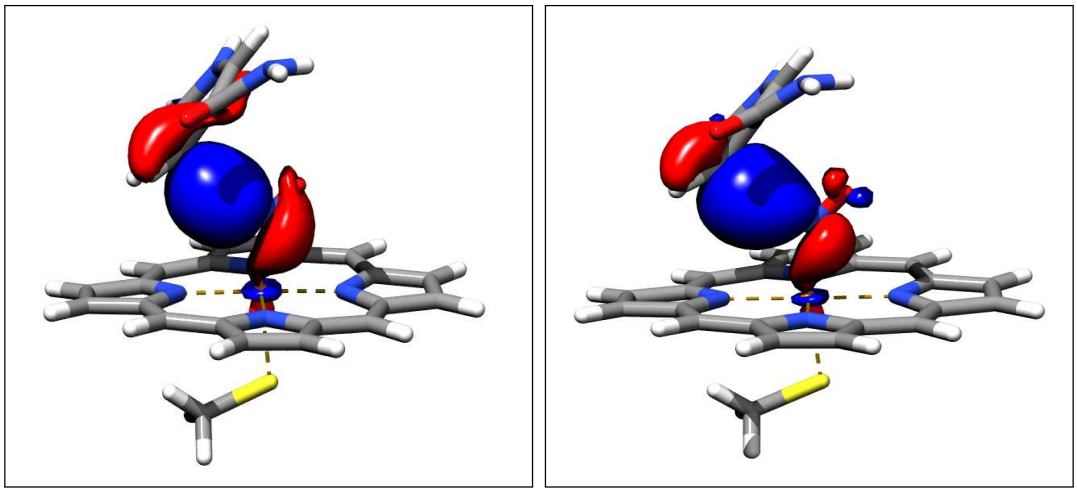


Figure 4.11. – Doubly occupied localized MOs showing the σ_{N-H-C} orbital for **TS**(2,3)**a** (left) and **TS**(2,3)**b** (right).

electron density is found near the proton, in between the nitrogen and the carbon atom.[j] Hence it can be concluded that the *proton and the electrons are transferred concertedly from NADH to the NO ligand in form of a hydride ion.*

[i] The localized orbitals were generated applying the Pipek-Mezey algorithm.[160]

[j] The localized electron density is already a hint that proton tunneling does not play a role in this reaction.

Kinetic Isotope Effect

The kinetic isotope effect is a sensitive probe for the electronic state of a transition state.[172] The KIE for the reaction via **TS**(**2**,**3**)**a** and **TS**(**2**,**3**)**b** was calculated for all three snapshots.[k] Since two PESs are involved, it is possible that the reaction (**2** → **3**) proceeds via a two state reactivity (TSR) mechanism. The KIE was thus calculated not only for the 'One state reactivity' combinations from **2a** via **TS**(**2**,**3**)**a** and from **2b** via **TS**(**2**,**3**)**b**, but also for the TSR combinations from **2a** via **TS**(**2**,**3**)**b** and from **2b** via **TS**(**2**,**3**)**a**.
The results are summarized in Table 4.4. It can be seen that the KIE is only dependent on

Table 4.4. – Calculated kinetic isotope effect for the hydride transfer reaction for all three snapshots at room temperature. The KIE was calculated for the structures obtained with the QM/MM optimization of QM-region **R2** with deuterated pro-R position. The data are given for all four combinations of the reaction from **2a**/**2b** via **TS**(**2**,**3**)**a** and **TS**(**2**,**3**)**b** in the ordering of snapshot A/B/C.

	KIE	
Educt \TS	**TS**(**2**,**3**)**a**	**TS**(**2**,**3**)**b**
2a	3.81/3.97/3.70	3.52/3.42/3.05
2b	3.84/3.99/3.72	3.54/3.44/3.07
Exp. KIE[78]	2.7 ± 0.4	

the electronic state of the TS, and the electronic state of the educt (**2a** or **2b**) does not have an influence hereon.
A comparison to the experimental value shows that the reaction presumably proceeds via **TS**(**2**,**3**)**b**. This result is in accordance with the lower activation barrier that was obtained for this alternative (see again Table 4.3).

4.4.3. Further Reaction Steps: Protonation and Decomposition

In order to complete the reaction, two groups are still missing: one proton and a second NO molecule. For the reaction following the hydride transfer two different pathways are considered. In pathway **M1** the second NO molecule directly binds to **3**, and the necessary protonation takes place in the course of the following decomposition. Pathway **M2** starts with a protonation of **3** (**3** → **4**), before the decomposition is initiated by the second NO molecule binding to **4**. Both pathways yield N_2O and H_2O as end products of the reaction cycle.
Before describing the key steps of both pathways, it is important to note what was observed in the calculations when including the second NO molecule. Species **3**, which exists in two

[k] The experimentally derived KIE is in the semiclassical range[173] ($k_H/k_D < 7$) and thus proton tunneling effects were not analyzed for this process (see also Footnote j).

forms in the hydride transfer step (**3a** and **3b**, see Section 3.4.1) shows only one SCF solution when including the NO radical into the system, namely species **3b**. Efforts to converge to the 'pure doublet' solution failed.
Since **3b** is significantly more stable than **3a** (4-5 kcal/mol lower in energy), **3a** is not investigated further and protonation and decomposition are studied using only species **3b**.

Pathway M1

The QM/MM optimized structures and relative energies obtained for pathway **M1** are shown in Figure 4.13. The starting point for **M1** is the system with species **3** and an additional NO molecule in the heme distal pocket. The N-N bond formation process of NO with **3** was simulated by performing a QM/MM surface scan, in which the N-N distance was continuously shortened.
In the starting structure (A in Figure 4.13) the spin density is distributed as sketched in Figure 4.12. α spin density is found on iron, β spin density is found on the NHO ligand and again α spin density on the second NO molecule. This corresponds to the diradical species **3b** and an additional NO radical, which has opposite spin to the NHO ligand, thus resulting in an overall doublet multiplicity for the whole QM system.
With decreasing N-N distance, the energy decreases successively without showing an activation barrier (A to B). At the same time, the α and β spin density on both the NHO ligand and NO diminishes concertedly, thus indicating a spin-recoupling reaction (see Figure 4.14). In the resulting product B, the ligand has lost its bond to iron. In the following reaction steps the 'free' N_2O_2H undergoes first a tautomerization (B to C), in which one N-O bond is elongated (by 12 pm). This N-O bond is then cleaved (C to D) resulting in N_2O and OH^-. The reaction cycle is completed by a protonation of OH^- by Asp_{393} (D to E). It should

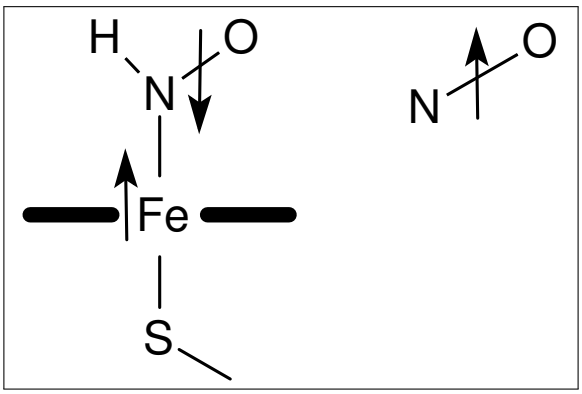


Figure 4.12. – Scheme illustrating the spin distribution before N-N binding in Figure 4.13, A.

be mentioned that, by inclusion of the ZPE corrections, the activation barriers are reduced significantly or even vanish (as for the last protonation step, D to E). This is partly due to the highly delocalized transition state, indicating that the proton is transferred concertedly via several molecules. The same is observed in pathway **M2** described below.

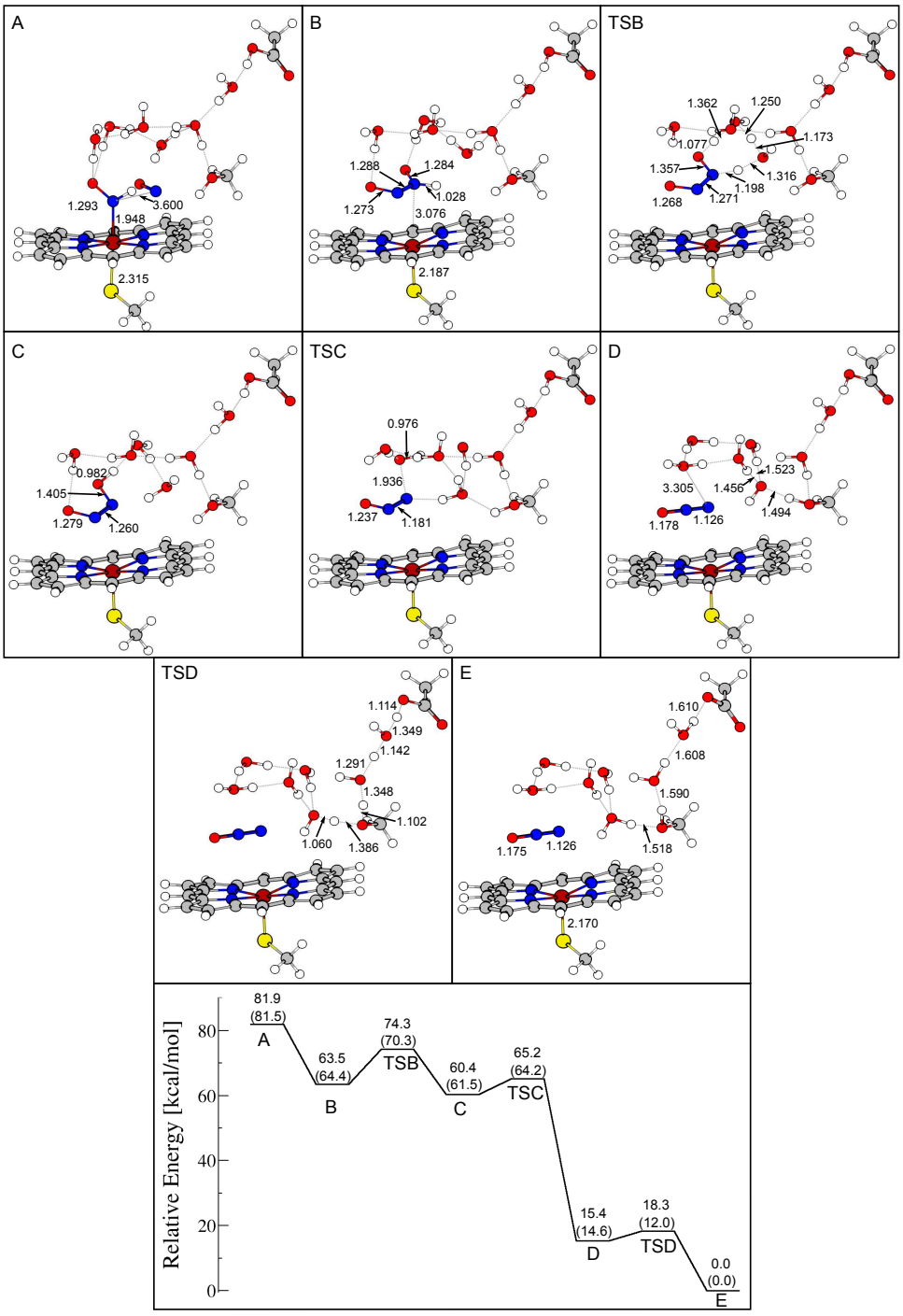


Figure 4.13. – Structures of the reaction intermediates and transition states of the **M1** pathway. Selected bond lengths are given in the respective plots (in Ångstrom). The energy diagram shows the relative energies with respect to the final product. The values are QM/MM energies (QM/MM energies including ZPE correction).

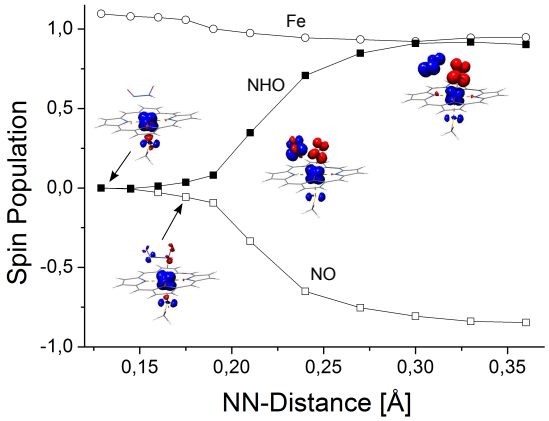


Figure 4.14. – Mulliken spin population distribution on iron, NHO ligand and the second NO molecule in the course of the QM/MM surface scan A to B in **M1**. The spin density for the system at different N-N distances is plotted (α spin density in blue, β spin density in red).

Pathway M2

Figure 4.15 shows the QM/MM optimized structures and the respective QM/MM energy profile for pathway **M2**. This pathway starts with a protonation of **3b** yielding the doubly protonated ligand **4b** (A to C in Figure 4.15).[l] During the protonation, one proton from Asp_{393} is transferred via four water molecules to the NHO ligand. This process has an overall barrier of 5.9 kcal/mol and results in the doubly protonated intermediate **4b**, which is 2.4 kcal/mol more stable than **3b**.

With this finding the present QM/MM results are consistent with previous cluster model calculations[88,91] which found that the doubly protonated species is more stable than the singly protonated form. Note however that the energy differences reported in previous works differ from the energy difference that has been calculated here. This is readily explained by different protonation schemes and the fact that in the previous studies only closed shell forms were investigated, while diradical species were also taken into account in this work.

Continuing from **4b**, the binding of the second NO molecule was modeled in the same way as in mechanism **M1** (C to D). Again, a barrierless, spin-recoupling N-N bond formation is found (see Figure 4.16). The product (D) is also not bound to iron any more. The 'free' $N_2O_2H_2$ decomposes to N_2O and H_2O by cleavage of the (H)N-O(H) bond. This cleavage starts with deprotonation of nitrogen (D to E) and is completed by the protonation of the cleaved OH-group (E to F).

As can be concluded from Figure 4.15, there is no significant barrier for the decomposition of **4b**.[m] This is in agreement with the experimental observations.[69,89]

[l] The observation done for species **3** (see page 89) was also done for **4**. Two electromers, **4a** and **4b**, are found without the additional NO molecule (see also Section 3.4.2), but only one SCF solution is found for the calculation including the second NO radical in the QM system. Since **4b** is about 4-5 kcal/mol more stable than **4a**, the latter is not analyzed further and only **4b** is investigated in the following.

[m] Decomposition of **4** was analyzed in DFT gas phase cluster studies as well,[91,98] where significant barriers were calculated (of 26 kcal/mol and 35 kcal/mol, respectively).

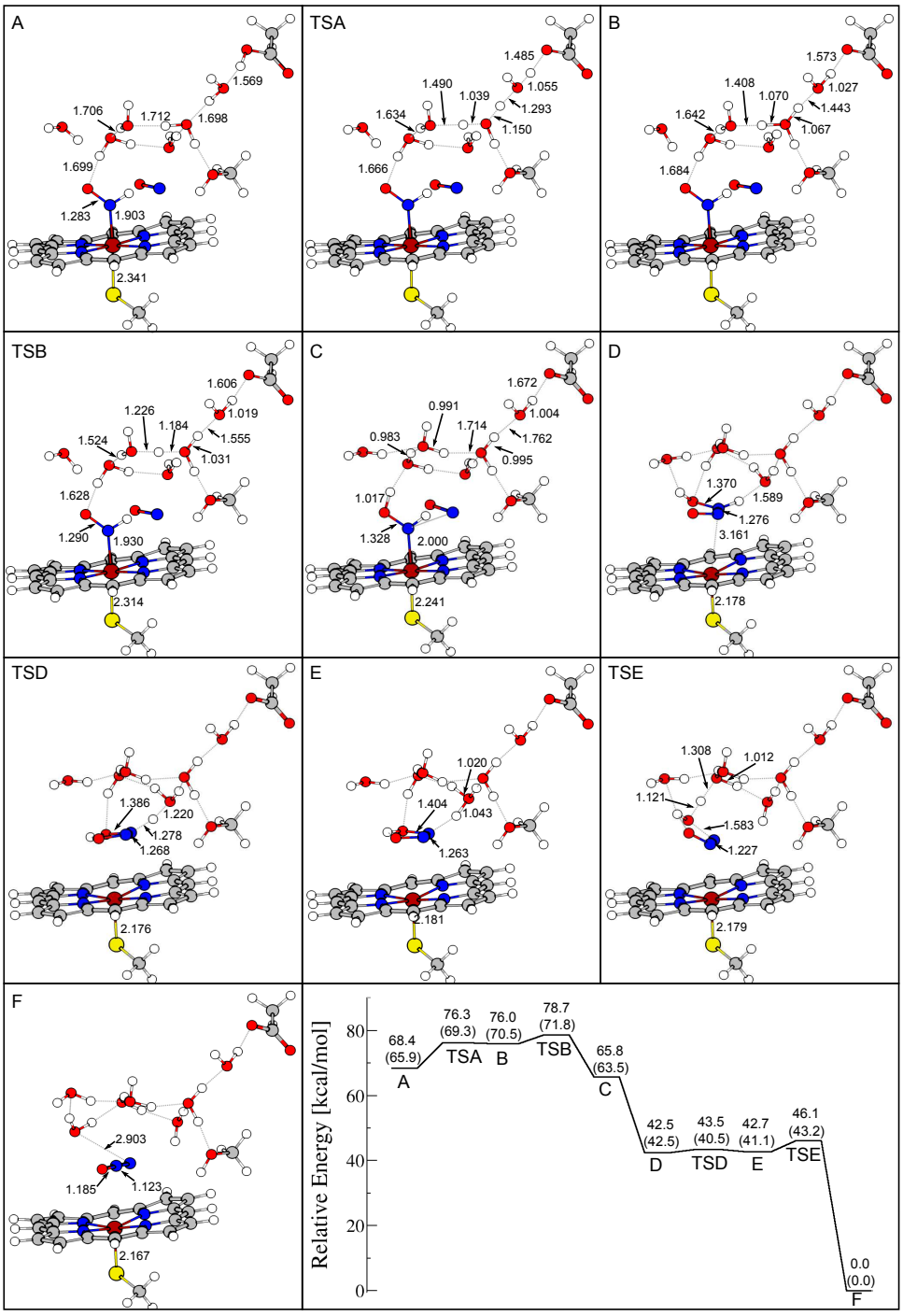


Figure 4.15. – Structures of the reaction intermediates and transition states of the **M2** pathway. Selected bond lengths are given in the respective plots (in Ångstrom). The energy diagram shows the relative energies with respect to the final product. The values are QM/MM energies (QM/MM energies including ZPE correction).

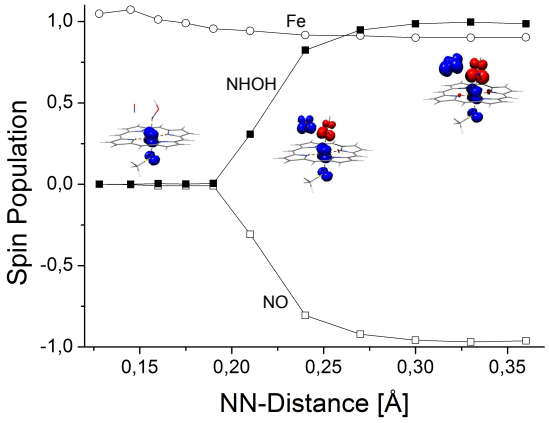


Figure 4.16. – Mulliken spin population distribution on iron, NHOH ligand and the second NO molecule in the course of the QM/MM surface scan C to D in **M2**. The spin density for the system at different N-N distances is plotted (α spin density in blue, β spin density in red).

4.4.4. Kinetic Simulation of the whole Reaction Cycle

In the previous parts of this chapter the kinetic and thermodynamic parameters for the entire reaction cycle were calculated, starting from the resting state **1** via the NO-bound state **2a**/**2b** to the reduced species **3a**/**3b**. From **3b** the mechanism went either directly (pathway **M1**) or via **4b** (pathway **M2**) to the final products.

Table 4.5. – Activation barriers and thermodynamic properties used in the simulation of the reaction cycle. B(**M1**), C(**M1**), C(**M2**) and D(**M2**) represent the states depicted in Figures 4.13 and 4.15.

Elementary step	$\Delta H^{\ddagger}$ [kcal/mol]	ΔH [kcal/mol]
1 + NO $\leftrightarrow$ **2a**[a]	5.4	-4.3
2a $\leftrightarrow$ **3b**[a,b,c]	8.6/6.5	2.0/3.0
3b + NO $\leftrightarrow$ B(**M1**)[d]	2	-17.1
B(**M1**) $\leftrightarrow$ C(**M1**)	5.9	-2.9
C(**M1**) $\leftrightarrow$ **1** + N_2O	2.7	-61.5
3b $\leftrightarrow$ **4b**	5.9	-2.4
4b + NO $\leftrightarrow$ C(**M2**)[d]	2	-21.0
D(**M2**)$\leftrightarrow$ **1** + N_2O	0.7	-42.5

[a] Only the closed shell species **2a** is used.
[b] The activation enthalpy of the reaction via **TS**(**2**,**3**)**b** is used.
[c] Values for snapshot C/B.
[d] Although NO binding has no barrier, a small barrier has to be included for the kinetic simulation due to numerical reasons.

In this section, the thus obtained thermodynamic and kinetic parameters are used to simulate the entire reaction cycle of P450nor. Only the candidates that were selected in Chapter 3 are used here, i.e. the reaction is simulated from **1** to **2a**, via **3b** and via **4b** to the end products. Through comparison with experimental observations, a central question of the P450nor mechanism is addressed: which of the species (**3b** or **4b**) is representing intermediate *I* that can be observed under limiting [NO] conditions?
The parameters used for the simulation are summarized in Table 4.5. In order to simulate the reaction under different [NO] conditions, two different reaction conditions were used:

K1 [NO] is much higher than the enzyme concentration ([NO] = 50 x [E]).

K2 [NO] is set to be equal to the enzyme concentration ([NO] = [E]).

The resulting reaction course for **K1** is shown in Figure 4.17. It can be readily seen that under

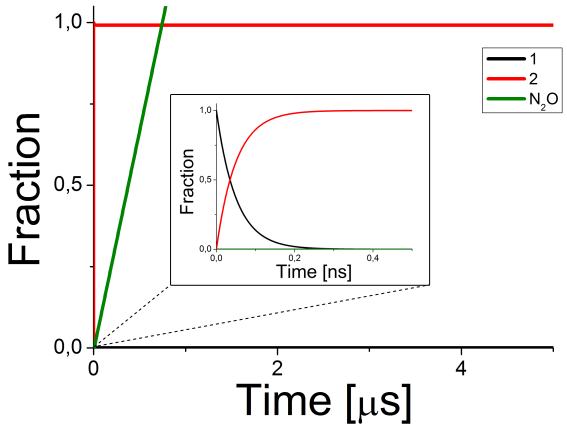


Figure 4.17. – Reaction course of the P450nor reaction cycle at high NO concentration (**K1**) at 10°C using the results of snapshot C for the reduction step.

high NO concentration the only observable intermediate is the NO-bound state **2a**. According to the activation barrier of the slowest reaction step (**2a** → **3b**), about 10^6 molecules N_2O are produced per enzyme molecule and second. With this value, the obtained turnover rate from the simulation is three orders of magnitude higher than the experimentally observed turnover rate (1200 s^{-1}). This discrepancy is probably due to the neglect of transport phenomena in our model simulation. Those should play an important role in such a reaction with a large cofactor (as NADH) and multiple reaction steps. Consequently, the focus here should lie on the qualitative results, which show that only intermediate **2a** is accumulating. Later in the reaction cycle occurring intermediates are transformed too fast to ever accumulate and hence cannot be detected, which is consistent with experiment.
The reaction course under reaction condition **K2** looks quite different and is much more revealing (Figure 4.18). In the early phases of the reaction all NO molecules bind to the enzyme (**1** → **2a**), as in case **K1**. The fraction of **4b** is then increasing at the expense of **2a**, until a ratio of about 2:1 (**4b** versus **2a**) is reached. This ratio stays constant, but both populations (**2a** and **4b**) decline in the course of the reaction, whilst the end product N_2O is released and the resting state **1** re-accumulates. At no time intermediate **3b** is observed in any measurable quantity.
Importantly, the simulation of the reaction course shows that it is species **4b**, and not species

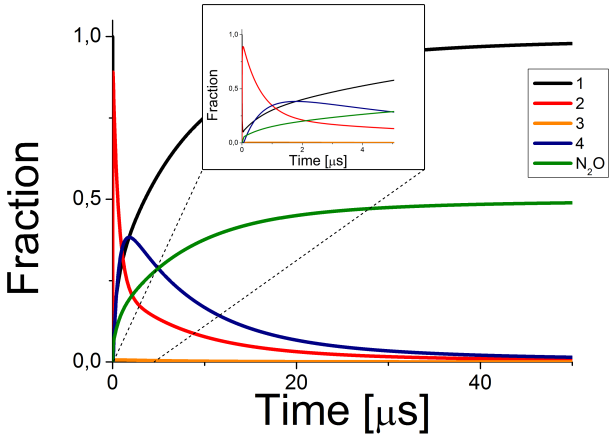


Figure 4.18. – Reaction course of the P450nor reaction cycle at low NO concentration (**K2**) at 10°C using the results of snapshot C for the reduction step.

3b, that is accumulating at low NO concentration. *With this result in hand, intermediate I is revealed to be the doubly protonated species **4b**.*

What has not yet been discussed is the mechanism behind this behavior. Here, the essential questions are: What is happening at low NO concentrations, especially in the region where the ratio of intermediate *I* to NO-bound state (**4b** to **2a**) is constant (see Figure 4.18)? And why does intermediate *I* accumulate at all?
To answer these questions it is important to recall what happens to **4b** in the reaction cycle: A second NO molecule binds to **4b**, which initiates the decomposition to the final products. For this process, free NO must be available. However, under low NO concentrations, the majority of the NO molecules is bound to P450nor, and only very few are in free solution and can react with **4b**. Thus, **4b** is now accumulating and 'waiting' for the next free NO molecule. Under these circumstances there is enough time for **2a**, **3b** and **4b** to reach equilibrium, which is, of course, directly dependent on the relative stabilities of all species. Since **3b** is higher in energy than **2a** and **4b**, the problem can be reduced to the equilibrium between **2a** and **4b**. This equilibrium is, of course, a very sensitive function of the relative stabilities of the species involved. Changing the energy by merely 1.2 kcal/mol changes the equilibrium constant by an order of magnitude. The dependence of the ratio of the two species **2a** and **4b** on their relative stability is drawn in Figure 4.19.
It is seen from Figure 4.19 that for the case ΔE = -0.4 kcal/mol (**2a** $\rightarrow$ **4b**) the ratio of **4b** to **2a** is 2:1 - as obtained in simulation **K2** (see Figure B.2). Changing ΔE to -1 kcal/mol moves this ratio to 6:1 while a change in the opposite direction to +1 kcal/mol leads to a ratio to 1:6 (compare also **K2** for snapshot B, Figure B.2 in the Appendix). Obviously, the DFT results cannot be trusted to this level of accuracy. Hence, the expected ratio cannot be quantitatively predicted with a high degree of confidence.
In simulation **K2**, the ratio of **4b** to **2a** stays constant, but their populations are simultaneously decreasing during the course of the reaction. Nevertheless the reaction proceeds. What is happening there? We know that **4b** needs free NO to react with. If the majority of NO is bound to P450nor, the NO concentration depends on the equilibrium of **1** and **2a**, and thus also on the back reaction from **2a** to **1**. If one active site releases NO in this manner, another active site, which is waiting in intermediate *I* form, can proceed with the reaction

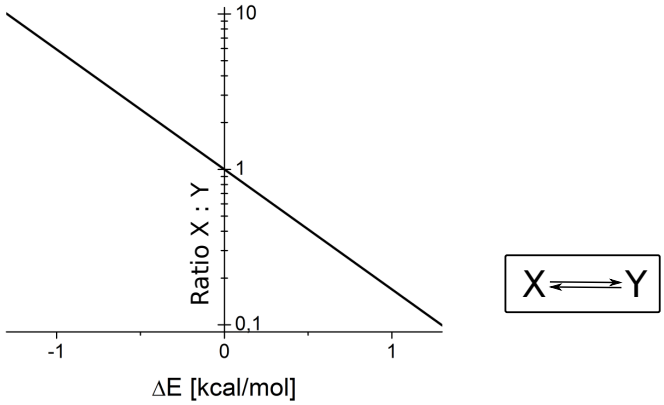


Figure 4.19. – Expected ratio of compounds X to Y for a given reaction enthalpy ΔE

and decompose to N_2O and H_2O. Forward (decomposition of **4b**) and backward reaction (**2a** → **1**) are coupled this way. This means that the rate, at which *I* decomposes, depends also on the rate constant of NO release from **2a** at low NO concentration.[n] On the other hand, at high NO concentration *I* does not accumulate. Here, enough free NO is in solution, and *I can immediately react with another NO radical and decompose* (see **K1**, Figure 4.17).

4.5. Discussion

In this chapter the reaction cycle of P450nor was calculated in the QM/MM framework. Kinetic and thermodynamic constants were derived for all elementary reaction steps and used to simulate the time course of the reaction under the conditions of either excess [NO] or limiting [NO].

The first important part of this chapter is the modeling of the reduction of NO-bound P450nor by NADH. Here, an extensive QM/MM model with the reducing cofactor NADH in the active site of the enzyme molecule was used. The reduction process was modeled on two PESs corresponding to closed shell and diradical states of the educts and products. Reaction rate constants and the kinetic isotope effect of the hydrogen transfer reaction were calculated and compared to experimental data. Comparison between theory and experiment revealed that the reaction presumably proceeds via a diradical transition state. Analysis of the electronic structure of the TS suggests that the reduction step is carried out simultaneously with hydrogen transfer in the sense that it is a hydride ion that is actually transferred, rather than a proton or a hydrogen atom. Both electromers of the educt, NO-bound P450nor, are nearly isoenergetic. Nevertheless, the results in Chapter 3 show that the NO-bound state must be a closed shell species. Since the NHO adduct and the TS is best described by a diradical electronic structure, a rather unusual two-state reactivity scenario arises. *Instead of the crossing of two states of different multiplicity, two singlet states with different electronic structures cross in this scenario*.
The influence of the protein structure on NADH binding and reactivity was studied in detail. The importance of hydrogen bonding of Ser_{286} and Thr_{243} to NADH was demonstrated by

[n] The back reaction (**2a** → **1**) is slower than the reduction of **2a** (**2a** → **3b**), and this is the reason why the equilibrium can be set up at all.

MD simulations and QM/MM calculations. These calculated hydrogen bonds are important for the binding and positioning of nicotinamide next to the NO ligand in the active site. It was shown that the hydrogen bonds significantly influence the activation barrier. Thus, they play an essential role in the *preparation as well as in the realization of the reduction step*. This finding helps to rationalize the experimental results obtained by site-directed mutagenesis studies on P450nor. Several studies demonstrated that substitution of Thr_{243} and Ser_{286} greatly reduces the enzyme activity.[70,81] For Thr_{243} mutations (to Asn, Gly and Ser) a large part of the enzyme activity (30–70%) was retained, whereas all other Thr_{243} mutations greatly reduced the activity.[81] Ser_{286} mutations (to Val and Thr) give a totally different result. Here, the enzyme activity was reduced down to 1–7%.[70] These observations are readily explained with the present theoretical model. The hydrogen bond of nicotinamide to Thr_{243} has a length of 1.7 to 2.4 Å in the catalytically relevant structures, and its strength is less important than that of Ser_{286} (here the calculated hydrogen bond length is between 1.7 and 1.8 Å in all calculated models). Thus, it can be concluded that *the hydrogen bond with Ser_{286} is much more important than that with Thr_{243}*. Indeed, among the Thr_{243} mutations that retain activity, two are polar amino acids of different size (Asn and Ser) and one cannot form a hydrogen bond at all (Gly).

The second important result of this chapter is the identification of intermediate *I* as the doubly protonated species low-spin Fe-NHOH. In contrast to previous theoretical works[88,91] it could be shown that *I* is not the postulated closed shell Fe(IV)-$NHOH^-$ species, but has to be described as the singlet diradical Fe(III)-NHOH, i.e. an NHOH radical antiferromagnetically coupled to a low-spin ferric iron (see also Section 3.4.2). While both, the monoprotonated and diprotonated species occur in the reaction cycle, the kinetic simulations demonstrated that under limiting [NO] only the diprotonated species accumulates and hence is amenable to spectroscopic identification.

The final question addressed in this paper concerns the role of the protein environment in the reaction. It is apparent that the protein has several important functions:

1. Residues in the active site bring NADH and NO together and stabilize the TS (especially Thr_{243} and Ser_{286}).

2. The ferric active site enables the ligand to switch from the closed shell NO-bound state to the diradical TS, which lowers the activation barrier by about 3 kcal/mol. The diradical product, intermediate *I*, can then react in a spin-recoupling manner with the second NO radical.

3. After N-N bond formation, the heme active site releases the ligand and the decomposition occurs essentially spontaneously in the active site pocket. Thus, undesired side reactions with different species in solution are excluded.

5. Spin-Spin Zero-Field Splitting and the Point-Dipole Approximation

5.1. Introduction

Determination of structure and conformation in macromolecular systems, either biomolecular or synthetic, requires determination of distances between specified points in the system. As introduced in Section 1.1.3, most often organic spin labels are attached to the macromolecules, and the magnetic dipolar coupling between them is used to determine distances over a wide range (up to 70 Å). In this area of application PELDOR and DQC are rapidly developing techniques. The probes frequently used are nitroxyl radicals, and data analysis typically is based on a *point-dipole approximation*, in which it is assumed that the unpaired electron is localized on the nitrogen-oxygen moiety.
Distances determined by PELDOR for dinitroxyls with well defined structures have been shown to be in good agreement with expected interspin distances in the range of 20–28 Å.[174,175] However, it is also known that *it is at shorter distances or in delocalized systems inadequate to relate the zero-field splitting observed in EPR spectra to intercenter distances by applying the point-dipole approximation*.[42]
In this chapter, a nonbiological model system with two nitroxyl centers, named diradical [**6**], is studied. Diradical [**6**] was characterized by single-crystal X-ray crystallography as well as by EPR spectroscopy by collaborators, and their results are summarized in Section 5.2. These data show that evaluation of the measured zero-field splitting by the point-dipole approximation gives a distance between the spin centers, which is not in accordance with the distance derived from the X-ray structure.
The equations for exact and approximate spin-spin ZFS were already discussed in Section 2.3. In Section 5.4.3 it is first demonstrated that *explicit quantum chemical calculations* provide a ZFS that agrees very well with experiment. The factors that contribute to the observed ZFS are analyzed in detail and the parts that are missing from the point-dipole treatment are identified (Section 5.4.4). The analysis provides insight into the *scope and limitations of the important point-dipole approximation*. The distance dependence of the zero-field splitting in nitroxyl diradicals is discussed and the 'through-bond' and 'through-space' interactions, a scenario that is also important for biological macromolecules, are studied (Section 5.4.5 to 5.4.7). The predictions are compared with experimental results for diradicals with longer intercenter distances (Section 5.4.8).
The work presented in this chapter is already published in Riplinger et al.[176]

5.2. Summary of the Experimental Results

The experimental work on diradical [**6**] and the analysis of the obtained experimental data was done by our collaborators from Denver and Maryland. In this thesis, only a short summary of the experimental results and analysis is given. For more detailed information see Riplinger et al.[176]

5.2.1. X-Ray Structure

The structure of dinitroxyl [**6**], determined by X-ray crystallography, is shown in Figure 5.1. Key features of the structure are summarized below.

Owing to steric constraints, the two pyrrole rings are far from coplanar. Indeed, the dihedral

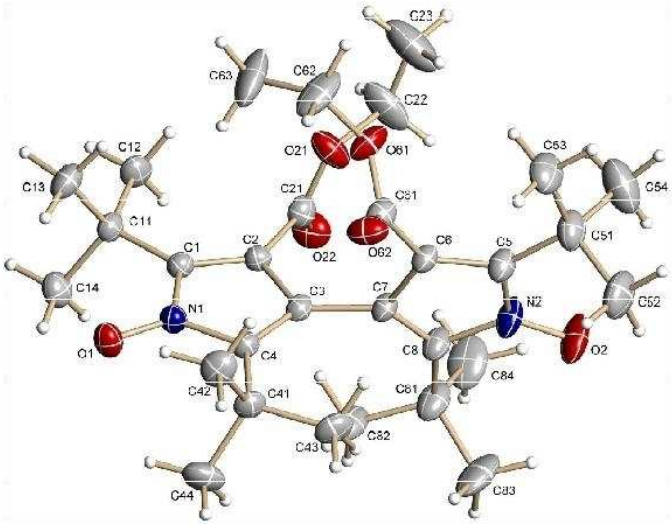


Figure 5.1. – Structure of dinitroxyl [**6**]. Anisotropic atomic displacement ellipsoids for non-hydrogen atoms are shown at the 30% probability level. Hydrogens are displayed with arbitrarily small radii.

angle between the mean planes defined by the two sets of ring atoms (N_1-C_1-C_2-C_3-C_4 and N_2-C_5-C_6-C_7-C_8) is $80.77°$. The C_3-C_7 bond linking between the two rings has a length of 1.476 ± 0.005 Å. The nitrogen-oxygen bond lengths are 1.279 ± 0.004 Å (N_1-O_1) and 1.281 ± 0.004 Å (N_2-O_2).

The distance between the two nitrogen atoms is 5.830 Å and the distance between the midpoints of the two N-O bonds is 7.006 Å. The latter corresponds to the spin-spin distance, if one assumes the unpaired electrons to be located in the middle of the N-O bond.

The structure shows only moderate disorder: one of the two ethoxycarbonyl groups (attached to C_6, Figure 5.1) is found in two orientations, with corresponding site occupancy factors of 0.592 ± 0.017 and 0.408 ± 0.017.

5.2.2. EPR Spectroscopy

Analysis of the X-band EPR spectrum of dinitroxyl [**6**] in glassy decalin at 89 K yielded the parameters summarized in Table 5.1. To match all features of the spectrum it was necessary to assume two conformations. For the dominant conformation the zero-field splitting parameter $D = 211$ G (0.020 cm^{-1}) and $E = 13$ G. For the minor component $D = 244$ G (0.023 cm^{-1}) and $E = 6$ G. If the point-dipole approximation is used, these values of D correspond to *interspin distances of 5.1 and 4.8 Å*, respectively, which are *significantly shorter than the 7.0 Å distance* between the centroids of the N-O bonds found by X-ray crystallography. The structure suggests that extensive delocalization of the unpaired electron might contribute to this discrepancy. Hence this system was selected for detailed computational studies.

Table 5.1. – Parameters used for simulation of spectra of dinitroxyl [**6**] in decalin at 89 K.

	g_{xx}	g_{yy}	g_{zz}	|D|[G]	|E/D|	A_{zz}[G]
Conformation 1[a]	2.0064[b]	2.0052[b]	2.0097[b]	211	0.061	
Conformation 2[a]	2.0064[b]	2.0052[b]	2.0097[b]	244	0.026	
Monoradical	2.0110	2.0059	2.0028			15.7[c]

[a]The populations of Conformation 1 : Conformation 2 are 70 : 30.
[b]The g values for the diradical [**6**] are effective values along the axes of the spin-spin interaction tensor.
[c]For the monoradical the nitrogen hyperfine splitting along the x and y axes is less than linewidth.

5.3. Computational Details

All computations in this chapter were carried out with a development version of the ORCA program package.[177] Geometry optimizations were performed with the BP86[137–139] density functional employing the TZVP[151] basis set.

5.3.1. Model Systems for the Analysis of Coulomb, Exchange and Multicenter Contributions

Ten different sets of calculations were carried out: two sets of calculations (a - b) with the full system as shown in Figure 5.1 and eight sets of calculations (c - j) for different truncated model systems as shown in Figures 5.2 and 5.3. The first five sets of calculations are:

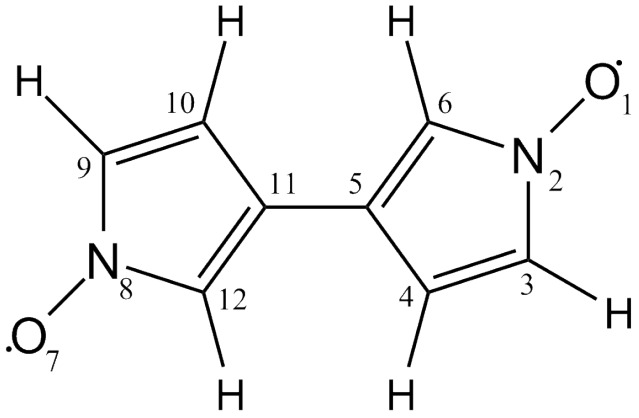


Figure 5.2. – The truncated dinitroxyl model system studied in this work together with its atom numbering scheme.

(a) a structure in which only the positions of the hydrogen atoms were relaxed,

(b) a structure that was fully optimized,

(c) a truncated system (substituents replaced by hydrogen atoms, see Figure 5.2) where only the hydrogen positions were optimized,

(d) a fully optimized truncated system (see Figure 5.2) and

(e) a relaxed surface scan of a full rotation of the two rings for the truncated system. Here, the dihedral angle, defined by the atoms 12, 11, 5 and 6, was fixed at 70 different values.

The model system was constructed in an effort to separate electronic and steric effects that may arise from the bulky substituents of diradical [**6**].

5.3.2. Model Systems for the Analysis of 'Through-Bond' and 'Through-Space' Contributions

To analyze the distance dependence of the ZFS with respect to 'through-bond' and 'through-space' contributions, additional model systems were constructed in which a linker was inserted between the two interacting nitroxyl radical fragments. This part of the study is designed to investigate the effective distance dependence of the ZFS interaction in order to find out under which electronic and geometric circumstances the widely used traditional point-dipole approach to the ZFS provides accurate distances.
The linkers were inserted between the two ring fragments (atoms 5 and 11 in Figure 5.2) yielding the model systems illustrated in Figure 5.3:

(f) fully optimized structures of the ring fragments with an unsaturated allyl linker consisting of consecutive ethenyl-groups ($C_{2n}H_{2n}$, n=1-10),

(g) partially optimized structures of the ring fragments with a saturated alkyl linker consisting of consecutive ethyl-groups ($C_{2n}H_{4n}$, n=1-10) and

(h) partially optimized structures of the ring fragments with a partially saturated linker ($C_{4n+2}H_{4n+4}$, n=0-4).

Furthermore model systems with non-aromatic nitroxyl rings were used:

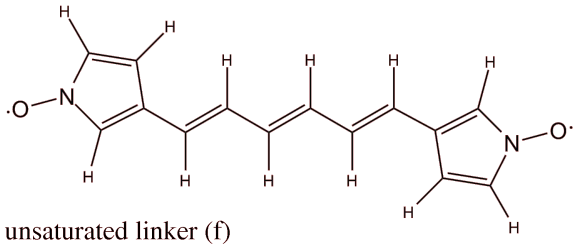
(i) fully optimized structures of the ring fragments with an unsaturated allyl linker and

(j) partially optimized structures of the ring fragments with a saturated linker.

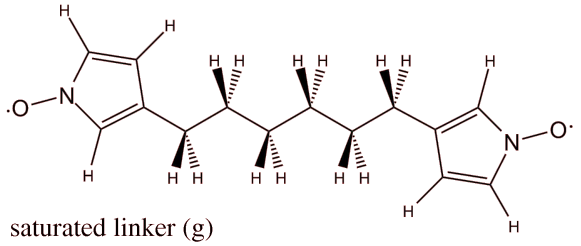
The structures of (g), (h) and (j) were partially constrained. The carbon skeleton of the linker was constrained to be planar (this prevented distortion of the systems with long linkers). With this approach the interspin distances were comparable for different linker types, but similar linker lengths.

5.3.3. Zero-Field Splitting Calculations

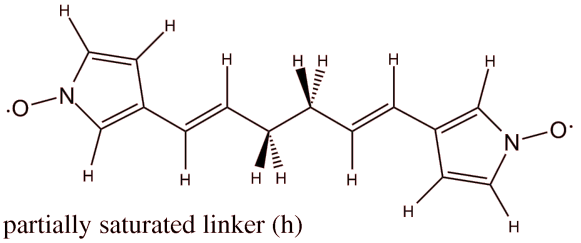
The spin-spin contribution to the ZFS parameter $D^{(SS)}$ was obtained from additional single point calculations. For this purpose, the BP86 density functional was applied in combination with the EPR-II basis set.[178] As discussed previously, the spin-spin term was calculated on the basis of the UNO determinant.[54] The SOMOs were localized according to the Pipek-Mezey criterion[179] as implemented in ORCA. Calculations with the much larger EPR-III basis set[180] led to 2–7% changes in the results, which was deemed insignificant. Hence, for the present purposes the EPR-II basis set was considered to be appropriate and was employed in all calculations reported below. Results obtained with this method will be designated as D^{I} in the following.



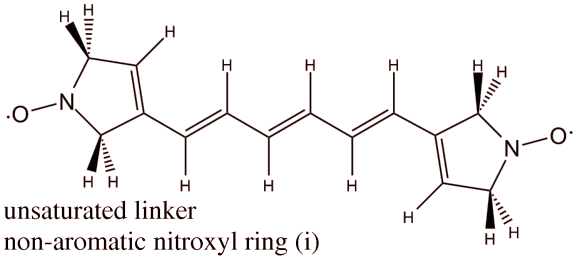
unsaturated linker (f)



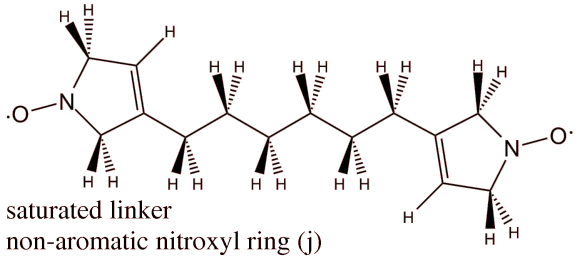
saturated linker (g)



partially saturated linker (h)



unsaturated linker
non-aromatic nitroxyl ring (i)



saturated linker
non-aromatic nitroxyl ring (j)

Figure 5.3. – Model systems with selected linkers.

The spin population needed for the calculation of the COG approximation and for the distributed point-dipole approximation were obtained from a natural population analysis (NPA) of the localized SOMOs of the UNO determinant using the 'gennbo' code developed by Weinhold and coworkers,[130] that is interfaced to ORCA. In these calculations, atoms 1–6 plus the adjacent hydrogen atoms (see Figure 5.2) were attributed to one spin center and atoms 7–12 plus the adjacent hydrogen atoms were attributed to the other spin center. If a linker was present it was cut into two halves and the atoms of each half were attributed to the neighboring ring fragment.
For the calculation of the ZFS within the commonly used point-dipole approximation the two spin centers were located in the middle of the NO group. Results obtained via the COG approximation (D^{COG}), distributed point-dipole approximation (D^{DPD}) or point-dipole approximation (D^{PD}) will be designated collectively as D^{II} in the following.

Table 5.2. – Overview of the different ZFS terms.

Designation		ZFS calculation level	see equation
D^{I}		Explicit quantum chemical calculation	(2.52)
	D^{DPD}	Distributed point-dipole model	(2.56)
D^{II}	D^{COG}	Center of gravity approximation	(2.59)
	D^{PD}	Point-dipole approximation	(2.59)[a]

[a] With $R^{(F1)}$ and $R^{(F2)}$ being the N-O midpoints (for the definition of R see Section 2.3.5).

5.3.4. Heisenberg Isotropic Exchange Coupling Constant

The Heisenberg isotropic exchange coupling constant J was obtained from additional single point calculations. The BP86 density functional used for the calculation of $D^{(SS)}$ would not have been appropriate for the calculation of the exchange coupling constant and thus the hybrid density functional B3LYP[181,182] was chosen for the prediction of J.[183] For this purpose, a broken symmetry calculation with the B3LYP density functional was applied in combination with the TZVP basis set.[151] The Yamaguchi approach[184,185] was used to estimate the Heisenberg exchange coupling constant J. Thus, J-values are based on the spin-Hamiltonian $H_{DvV} = -2J\hat{S}_A\hat{S}_B$.

5.3.5. Analysis of Distance Dependence

For the model systems with linkers the distance dependence of the D-values obtained with the different approximations was analyzed by least-squares fitting of the D-values to the function

$$D = C\left(\frac{1}{x+b}\right)^{a}, \tag{5.1}$$

with x being the length of the linkers and $C = -\frac{g_e^2}{16}\frac{\alpha^2}{S(2S-1)}$ being the prefactor for the dipole-dipole interaction already described in equation (2.52); a and b are fit-parameters.

5.4. Quantum Chemical Results

5.4.1. Geometry Optimizations

Full and constrained optimizations were performed for the complete system (starting from the X-ray structure) as well as for the truncated model system. For the complete system the relaxation of the coordinates of all atoms leads to very limited changes compared to the optimization where only the hydrogen atom positions were relaxed (see Figure 5.4). In

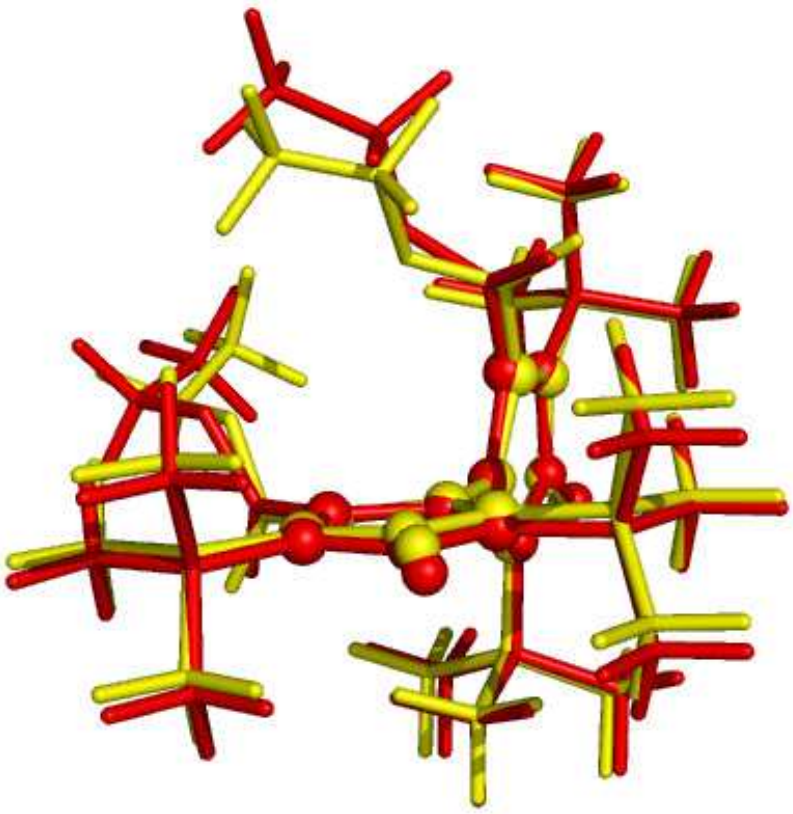


Figure 5.4. – Overlay of the fully optimized structure (BP86/TZVP) in red and of the constrained optimization (with fixed heteroatoms) structure in yellow. The atoms of the two aromatic nitroxyl rings that are used in our model system are shown in the ball and stick representation; the substituents are shown as sticks only.

particular, the relative orientation of the central spin-carrying ring systems changes by only about 10° and even the dangling carboxylate groups rotate by only about 30° relative to the original structure. By contrast, the relaxation of all degrees of freedom in the truncated model system leads to large changes in the geometry. Without the steric constraints provided by the bulky substituents, the two rings are predicted to be coplanar compared to a dihedral angle of 104.6° observed experimentally and 95.3° obtained in the optimization of the complete system.

Given that the dihedral angle between the two ring systems is an important degree of freedom for the truncated systems, the properties of the system are investigated as a function of this angle. This provides essential insight into the electronic communication between the

two subsystems. A relaxed surface scan was therefore performed for the truncated system (Figure 5.5). Compared to the orthogonal configuration, the coplanar structures are stabilized

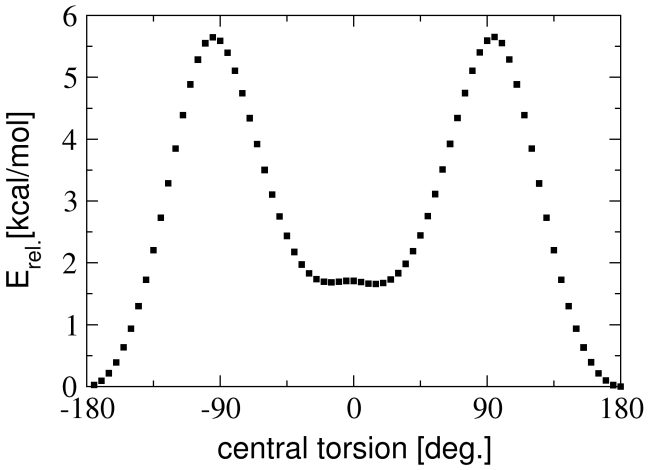


Figure 5.5. – Calculated energy for the truncated model system as a function of the constrained central torsion employing the BP86 density functional with the TZVP basis set.

by 4.0 and 5.7 kcal/mol, respectively.

5.4.2. Electronic Structure

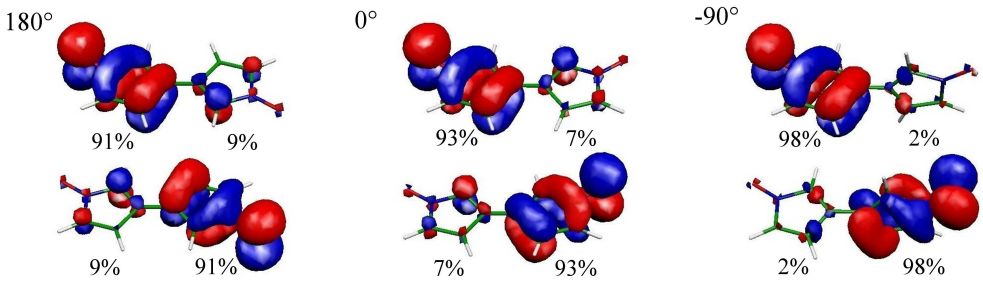


Figure 5.6. – Localized SOMOs at three different configurations: On the left side at a central torsion of $180°$, in the middle a central torsion of $0°$ and on the right side an orthogonal configuration with a central torsion of $-90°$. The population of the SOMOs on the two fragments is indicated as percentages.

The localized SOMOs are shown in Figure 5.6 for three configurations: the two coplanar configurations at $0°$ and $180°$ and a perpendicular configuration at -90°. Not surprisingly, in all configurations the SOMOs are π orbitals that are delocalized over the N-O group and the parent ring system.[186–188] However, in the coplanar configurations even the localized SOMOs have significant contributions not only on the parent ring, but also on the neighboring ring (up to 9%). In the perpendicular configuration the SOMOs show only small orthogonalization tails into the σ-system of the adjacent ring system ($\approx 2\%$).

5.4.3. Zero-Field Splitting: Comparison to Experiment

The spin-spin contribution to the ZFS was calculated for both the fully relaxed and the partially optimized untruncated systems. Both optimized structures give the same D^I of -0.020 cm^{-1} which is in excellent agreement with the experimentally determined value of 0.020 cm^{-1}. Thus, it is concluded that the electronic structure of diradical [**6**] is well represented by the truncated model system. The distance between the N-O midpoints in the experimental structure is 7.006 Å. Assuming the point-dipole approximation to be valid leads to a predicted $|D^{II}|$ of 0.008 cm^{-1}, which represents only about 40% of the experimental (or the quantum chemically calculated) value. As pointed out above, if the point-dipole approximation is employed to interpret the measured D-value and to estimate an intercenter distance, the result is 5.1 Å. *Thus, an error as large as 2* Å *would result from the use of the point-dipole approximation!* Since the full calculation reproduces the experimental value accurately, the reason for this dramatic failure must be of quantum mechanical origin. Hence, it is important to investigate in detail at which stage deviations from the point-dipole behavior occur and under which circumstances one will have to exercise more care.
Interestingly, for the fully relaxed structure of the truncated system with completely coplanar rings the magnitude of D^I increases by as much as 65% to 0.033 cm^{-1}, although the intercenter distance does not change at all. This finding together with the various approximations to the D-value is analyzed extensively below.

5.4.4. Zero-field Splitting Analysis

To obtain insight into the nature of the failure of the point-dipole approximation a series of calculations was carried out. It is instructive to compare the behavior of the individual contributions to D^I along the twisting coordinate. This will clearly show which parts of the D-value arise quantum mechanically and which are of purely classical origin.

Coulomb versus Exchange Contributions

D^I was calculated for the structures obtained with the relaxed surface scan. Its dependence on the orientation of the two rings is shown in Figure 5.7. The magnitude of D^I is minimal for the perpendicular configuration and increases to a maximum value for the coplanar configuration. To analyze this phenomenon D^I is split into an exchange-like contribution (D^I_X) and a Coulomb-like contribution (D^I_J) (Figure 5.7). Three interesting observations are made:

1. D^I_X is zero for the perpendicular configuration,
2. D^I_X is of the same sign as D^I_J and
3. its magnitude in the coplanar configuration is about 20% of D^I.

This $\approx 20\%$ contribution is a quantum mechanical correction to the magnetic dipole interaction and it can obviously not be neglected in an analysis that aims to be quantitative.
The magnitude of D^I_J increases upon approaching the coplanar conformation. A much weaker geometry dependence would be expected if the spins behaved as classical magnetic dipoles situated at the N-O bond midpoint. As will be elaborated below, this finding can be attributed to electron delocalization effects that are maximal in the coplanar orientation.

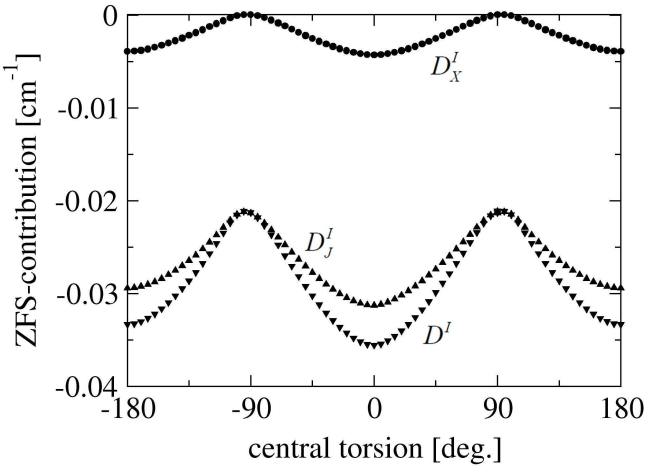


Figure 5.7. – Calculated spin-spin contribution to the ZFS parameter D^I for the dinitroxyl as a function of the constrained central torsion, employing the EPR-II basis set. Results are given for D^I (black filled lower triangles), Coulomb-like contribution (D^I_J, black filled upper triangles) and exchange-like contribution (D^I_X, black filled circles).

Multicenter Contributions

If a point-dipole model is assumed with spins localized on individual centers, it is exclusively the (AA|BB) class of integrals that would contribute to the final D-value. The extent to which this is true can be investigated by studying the 1- to 4-center contributions to the D^I-value evaluated quantum mechanically (Figure 5.8) and defined in Table 2.1. According to the results of these calculations, there are two main contributions to D^I. The largest contribution originates, as expected, from the 2-center-Coulomb integrals (AA|BB) that make up D^I_{2cC}. This contribution follows the same trend as the final D^I-value, but significantly (9–23%) *overestimates* its magnitude. Perhaps unexpectedly, the 3-center-Coulomb integrals (AA|BC), D^I_{3cC}, account for the second largest contribution. D^I_{3cC} also shows the same trend as D^I, but with opposite sign. The opposite sign results from the antibonding character of the SOMOs that reduces the spin density in the N-O bonding region. Hence, this important bond correction compensates for the overestimation of the 'pure' 2-center-Coulomb integral contribution.

Comparison of Point-Dipole Models

In Figure 5.9 D^I, the Coulomb part of the 2-center-Coulomb contribution D^I_{2cCC} and the D^{II}-values obtained at different levels of approximation are plotted. With the approximations made to obtain the 'distributed point-dipole' like equations, D^{DPD} should reproduce the Coulomb part of the 2-center-Coulomb contribution to D^I. For the orthogonal configuration, where the two SOMOs show only very small delocalization tails onto the neighboring spin center, this is indeed the case. The distributed dipole model approximates the Coulomb contribution to D^I to within 2% here, but still has an error of 15% compared to the full D^I-value since it does not include the multicenter and exchange interactions properly. However, compared to an error of 40% observed for the 'naïve' point-dipole model, the distributed dipole approximation is much more successful and reduces the error by more than a factor

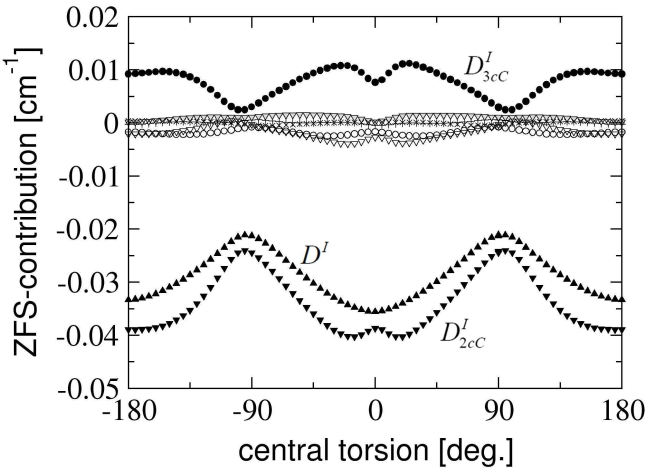


Figure 5.8. – Calculated spin-spin contribution to the ZFS parameter D^I from the 1- to 4-center contributions defined in Table 2.1 for the truncated dinitroxyl as a function of the constrained central torsion, employing the EPR-II basis set. Results are given for D^I (black filled upper triangles), 1-center-Coulomb- (+), 2-center-Coulomb- (D^I_{2cC}, black filled lower triangles), 2-center-Exchange- (grey filled lower triangles), 2-center-Hybrid- (open lower triangles), 3-center-Coulomb- (D^I_{3cC}, black filled circles), 3-center-Exchange- (open circles) and 4-center-integral contribution (x).

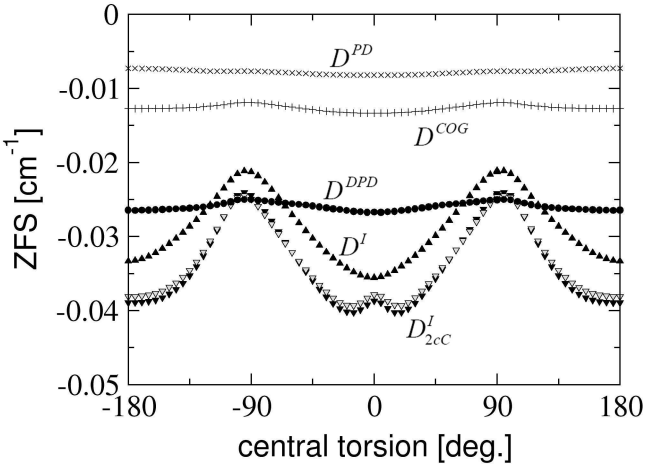


Figure 5.9. – Comparison of D^I and its important contributions calculated with the EPR-II basis set, and D^{II} obtained from spin populations calculated with the TZVP basis set for the truncated dinitroxyl as a function of the constrained central torsion. Results are given for D^I (black filled upper triangles), distributed dipole approximation (D^{DPD}, black filled circles), center of gravity approximation (D^{COG}, +), point-dipole approximation (D^{PD}, x), 2-center-Coulomb (D^I_{2cC}, black filled lower triangles) and the Coulomb part of the 2-center-Coulomb contribution (grey filled lower triangles).

of two.
On top of the 15% error relative to the full D^I-value, there are additional failures of the distributed point-dipole model: foremost, the D^{DPD} has too weak an orientation dependence and thus strongly underestimates the D^I_{2cCC}-value for the coplanar structure. For these configurations, both SOMOs are significantly delocalized over the neighboring fragment (see Figure 5.6). Thus, the magnetic dipole-dipole interaction occurs (on average) at much smaller distances than can be modeled within the 'distributed point-dipole' type approximation where the spin centers are necessarily separated. The observed small orientation-dependence of D^{DPD} is due to the fact that the spin density within the rings increases at the expense of the spin density on the oxygens with increasing planarity of the rings.
The center of gravity approximation yields D^{COG}, which shows only about half of the magnitude of D^{DPD}, but the same orientation dependence. For a highly delocalized system this approximation highly underestimates the contribution of the dipole-dipole interaction of atom pairs at small distances, since it just takes an average distance of the unpaired electrons.
Within the simplest approximation, the point-dipole approximation, D^{PD} is even smaller in magnitude than D^{COG}. Again this arises from the fact that the spin density is delocalized over the ring system. The distance between the two nitroxyl groups is essentially independent of the actual conformation and thus there is an even smaller dependence on the orientation of the two rings than with the two other approximations.
Taken together, the deviations of the naïve point-dipole model (D^{PD}) from the precisely calculated values may be summarized as follows: The point-dipole model, assuming the unpaired electrons in the middle of the N-O bond, strongly underestimates the D-value. Depending on the amount of exchange contribution and of electron delocalization, this simplest possible model accounts for only 22–36% of D^I. The center of gravity approximation, which takes the delocalization of the SOMOs on the ring atoms to some extent into account, results in 38–56% of D^I. The more elaborate introduction of the electron delocalization by the distributed dipole approximation yields a D-value that is 75–118% of D^I. This oscillation around the exact value is due to different errors occurring at different configurations. Comparing the D^{DPD} more rigorously with the Coulomb part of the 2-center-Coulomb contribution, 70–100% of D^I_{2cCC} are reached, depending on the configuration of the two rings and thus on electron delocalization onto the neighboring ring.

5.4.5. 'Through-Bond' versus 'Through-Space' Interactions

The analysis presented above suggests that, at the heart of the failure of the point-dipole approximation for dinitroxyl [**6**] is the close proximity of the two electronic systems and the significant electron delocalization effects, which lead to large quantum mechanical corrections to the point-dipole approximation.
One can therefore ask, at which distance and separation of the two spin systems does the (naïve) point-dipole approximation become accurate? In order to investigate this question, an *in silico* study was carried out in which a linker of increasing length was inserted between the two nitroxyl spin centers. Two types of linkers were investigated:

(a) an unsaturated linker that would allow for through-bond interactions and

(b) a saturated linker that greatly attenuates through-bond interactions.

The (partially optimized) structures and the localized SOMOs for these systems are shown in Figure 5.10 for the inserted unsaturated and saturated C_{10}-linkers. All structures with the allyl linker are planar. The structures with the alkyl linker were optimized such that both

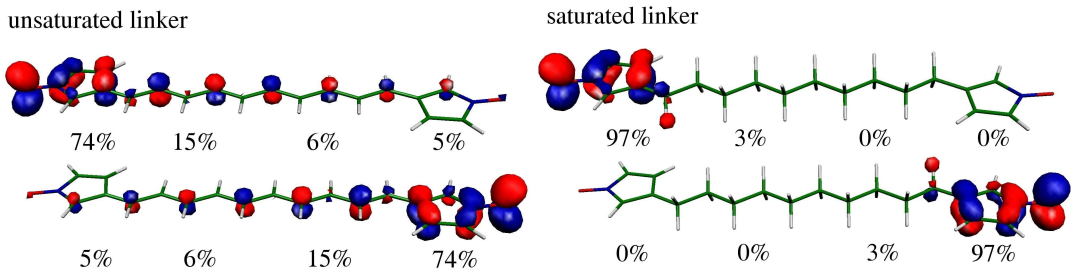


Figure 5.10. – Localized SOMOs for the unsaturated linker — $C_{10}H_{10}$ — on the left side and the saturated linker — $C_{10}H_{20}$ — on the right side. The population of the SOMOs on the fragments is indicated in percentages. It is given separately for the ring fragments and for the linker halves.

rings and the heteroatomic part of the linker are planar. The SOMOs of both systems are of similar shape as in the case without linkers. However, the contour plots show that the unsaturated linker, with its conjugated π-electron system, functions as a kind of conductor and thus even the localized SOMOs extend over the whole linker into the remote ring system. By contrast, the saturated linker cuts the electronic connection between both spin centers. Here, only a slight delocalization into the σ-system of the first few carbon atoms is found. The calculated D-values for the unsaturated linker are shown in Figure 5.11 as a function of intervening linker length. Not surprisingly, the magnitude of all D^{II}-values decreases with

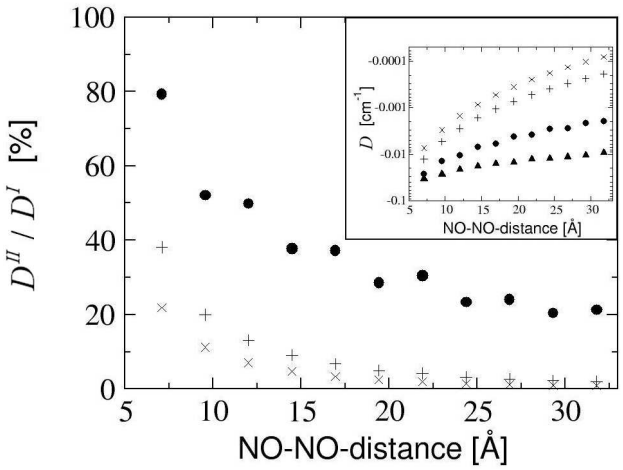


Figure 5.11. – Unsaturated linker. Comparison of D^I and D^{II} as a function of the NO-NO midpoint distance with the unsaturated linker. The D-value is given as percentage of D^I and as absolute values in the inlay diagram. Results are given for D^I (black filled upper triangles), distributed dipole approximation (D^{DPD}, black filled circles), center of gravity approximation (D^{COG}, +) and point-dipole approximation (D^{PD}, x).

increasing separation of the spin centers. However, most importantly, *the absolute D-value decreases much faster with interfragment distance for all point-dipole based schemes than for D^I, which is used as the standard for evaluating other methods.*
Apparently, the least erroneous approximation is the distributed dipole model. Without ad-

ditional linker atoms (which corresponds, at the same time, to an NO-NO bond midpoint distance of about 5.7 Å), it accounts for 80% of D^I. However, this fraction decreases to 50% of D^I with a C_4-linker. The decline of the quality of the distributed dipole approximation is due to the fact that the delocalization of both SOMOs over the neighboring fragment is supported by the unsaturated linker, which cannot be modeled by this approximation, where the spin centers are localized.
The center of gravity approximation accounts for only 40% of D^I without any additional linker atoms. This fraction decreases even faster than that of D^{DPD}. With increasing linker length of the unsaturated linkage and thus increasing delocalization over the linker atoms the contribution to D^I from the dipole-dipole interaction of the spin density on the ring fragments decreases very fast. Nevertheless, the majority of the spin population can still be found on the ring atoms. The interaction at small distances is thus strongly underestimated and D^{COG} is only 10% of D^I with a C_6-linker.
The point-dipole approximation does not account for the conducting bridge at all. Without additional linker atoms D^{PD} yields 20% of D^I. This fraction declines to less than 10% with a C_4-linker.
Thus, no matter what the chain length is, any point-dipole model produces incorrect results if the two spin-carrying fragments are connected via an unsaturated bridge. Effective distances determined for such systems via the point-dipole model will therefore have serious errors. The calculated D-values for the saturated linker (shown in Figure 5.12) are significantly different. Comparing D^I for the saturated linker with those of the unsaturated linker

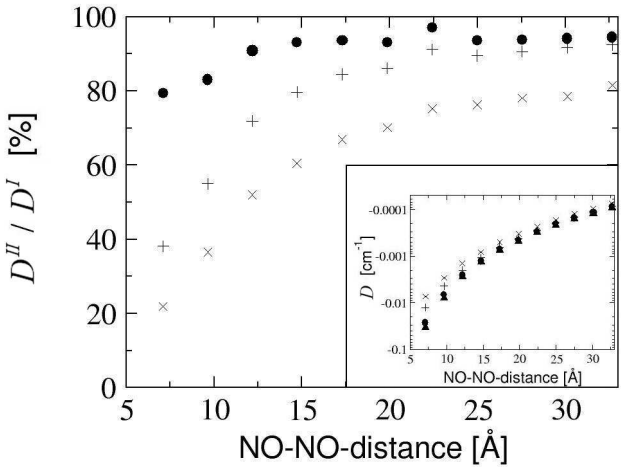


Figure 5.12. – Saturated linker. Comparison of D^I and D^{II} as a function of the NO-NO midpoint distance for the saturated linker. The D-values are given as percentage of D^I and as absolute values in the inset. Results are given for D^I (black filled upper triangles), distributed dipole approximation (D^{DPD}, black filled circles), center of gravity approximation (D^{COG}, +) and point-dipole approximation (D^{PD}, x).

it can be seen that the magnitude of D^I decreases much faster for the saturated bridge (compare Figure 5.11). Thus, the different point-dipole approximations become more and more accurate with increasing spin center distance.
Within the 'distributed point-dipole' like approximation the calculated D^{II}-value D^{DPD} accounts for more than 90% of D^I with a linker of 4 C atoms and reaching a limit of around

94% with a C_8-linker. For a C_4-linker, the fraction of the exchange contribution D_X^I falls below 1% of the total D^I (see also Figure 5.13). Thus, for linkers with more than 4 C atoms,

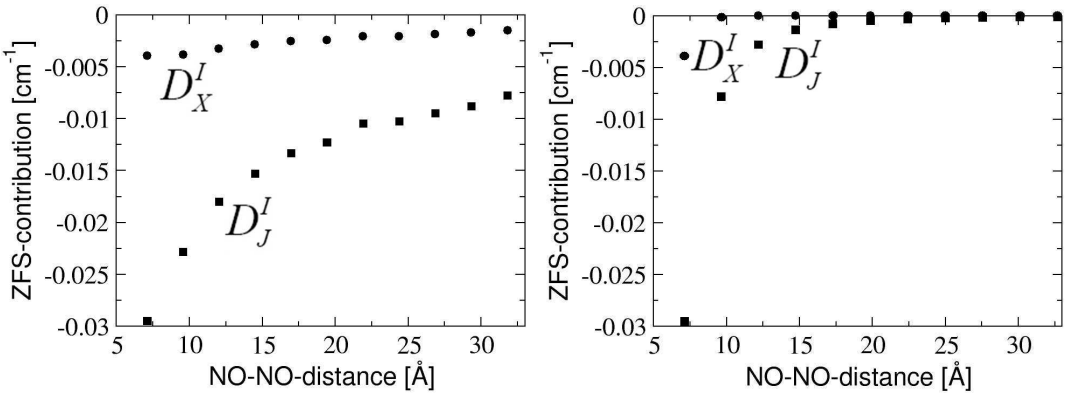


Figure 5.13. – Comparison of the exchange- and the Coulomb-contribution to D^I as a function of the NO-NO midpoint distance with the unsaturated linker (left figure) and with the saturated linker (right figure). Results are given for the exchange- (D_X^I, black filled circles) and for the Coulomb-contribution (D_J^I, black filled squares).

the remaining errors of the distributed dipole approximation are those introduced by the spin population analysis together with the slight errors arising from the 'compression' of the spin density onto the atomic centers.

The center of gravity approximation can account for more than 80% of D^I for a C_6-linker, and for more than 90% for a C_{12}-linker. Since the SOMOs in the systems with saturated linkers are more or less independent of the length of the linker, the centers of gravity of the SOMOs are constantly around 4.4 Å more distant than the length of the linker. With increasing linker length the underestimation of the dipole-dipole interaction of atom pairs at small distances becomes less important due to the $1/R^3$ behavior.

Finally, the simplest approximation, the point-dipole approximation, only very slowly approaches D^I. It reaches 80% of D^I only for a linker with 20 C atoms! The point-dipoles representing the SOMOs in this approximation are always about 5.7 Å further apart than the length of the linker and thus the magnitude of D^{PD} is always below that of D^{COG}.

To study whether there is a difference between the interaction through vacuum and through the saturated linker, the saturated linkers of different lengths were replaced by a hydrogen atom on each ring, the hydrogen positions were relaxed, D^I was calculated and the results were compared to D^I with the saturated, non-conducting bridge. A significant difference of up to 25% of the D-value is found only at small linker lengths. This deviation decreases quickly to only 5% at an NO-NO distance of 20 Å (i.e. with C_{10}-linker, if present). The increase in the D-value is not due to a different type of interaction, but can be traced back to the slight delocalization of the SOMOs into the σ-system of the first few carbon atoms (see also Figure 5.10). Thus, the saturated linker itself behaves essentially identical to vacuum. However, since the small tails of the spin density (caused by hyperconjugation) are likely to be present in many of the real systems, the point-dipole approximation still introduces a recognizable error whenever there is any chemical contact between the two spin-carrying fragments.

5.4.6. Power Laws

To compare the distance dependence of the D-values for the different applied models and the different linkers the data were fitted to a power function. The results are shown in Table 5.3. Here the parameter 'a' measures the decline of the D-values with increasing linker length.

Table 5.3. – Distance Dependence of D for the different approximations. Least squares fit to $D = C \cdot \left(\frac{1}{x+b}\right)^a$.

	Saturated linker		Unsaturated linker	
ZFS calculation model	a	b [Å]	a	b [Å]
D^I	3.1	-2.9	1.5	11.2
Distributed point-dipole	3.1	-2.6	2.1	2.0
Center of gravity	3.0	-1.1	2.8	-0.2
Point-dipole	3.0	0.1	3.0	0.1

Assuming the SOMOs to be contracted to one point, the parameter 'b' can be interpreted as an extrapolation of the location of the real 'contracted' point-dipoles.
For the saturated linker, about the same expected inverse third power law was found for all four methods of calculation, since the delocalization of the SOMOs does not change. The only difference lies in the localization of the 'contracted' point-dipoles (parameter b). As already pointed out, the simple point-dipole and the center of gravity approximation overestimate the distance of the point-dipoles.
For the unsaturated linker, the power laws predicted by different approximations to the D-value vary significantly. The center of gravity and the distributed point-dipole approximations both lead to a significantly faster decay than D^I. The origin of this effect is that both approximations take insufficient account of the spin-delocalization onto the bridge. Hence, for D^I, where the delocalization is exactly treated, the decay is found to be slowest. Strikingly, the D-value through an unsaturated bridge only decays as $R^{-1.5}$ rather than the expected R^{-3}.
Furthermore, the 'contracted' point-dipoles for D^I or the D-value based on the distributed dipole approximation are not even found within the ring systems. Rather, the positive sign of b indicates that they are found outside of the molecule (5.6 Å away from the NO midpoint for D^I). This reflects the fact that the electronically correct effective intercenter distance is much smaller than the one that is read off the geometric structure. In order to disrupt conjugation in the unsaturated tether linkers, consisting of allyl-groups that are separated in the middle by an alkyl-group, were inserted. The partially optimized structure and the localized SOMOs for the partially saturated C_{10}-linker are shown in Figure 5.14. In comparison to the completely unsaturated linker, the SOMOs are delocalized up to the saturated alkyl-group, which 'insulates' between the two conjugated π-electron systems. The calculated D-values for the partially saturated linker are shown in Figure 5.15 as a function of intervening linker length. Compared to the unsaturated linker the distributed dipole model follows D^I rather closely. Only D^{COG} and D^{PD} decrease much faster.
As expected, the distributed dipole model performs well for this system, since the SOMOs are delocalized only up to the saturated center of the system. Nevertheless, a slight decline

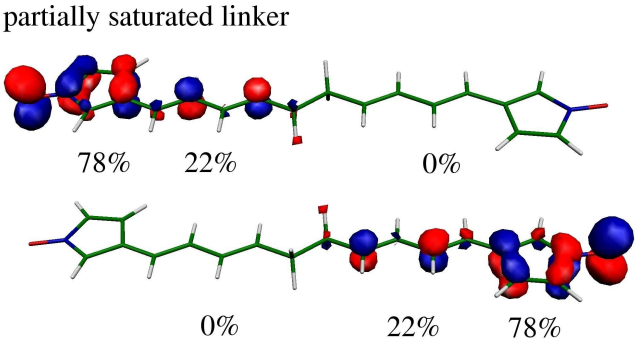


Figure 5.14. – Localized SOMOs for the partially saturated linker - $C_{10}H_{12}$. The population of the SOMOs on the fragments is indicated as percentages. It is given separately for the ring fragments and for the linker halves.

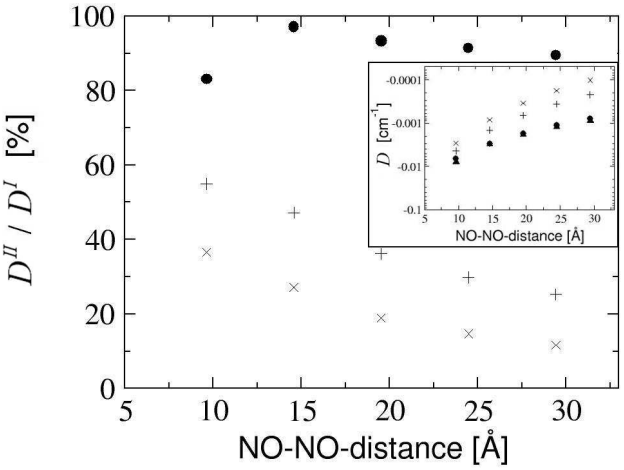


Figure 5.15. – Partially saturated linker. Comparison of D^{I} and D^{II} as a function of the NO-NO midpoint distance with the partially saturated linker. The D-value is given as a percentage of D^{I} and as absolute values in the inset diagram. Results are given for D^{I} (black filled upper triangles), distributed dipole approximation (black filled circles), center of gravity approximation (+) and point-dipole approximation (x).

of the quality of the distributed dipole approximation (evidenced by the ratio D^{DPD}/D^{I}) is observed beyond a linker length of 10 Å. This behavior can be traced back to the growing impact of the errors in the spin population analysis. The center of gravity approximation and the point-dipole approximation perform slightly better here than for the unsaturated linker, but they still greatly underestimate the target D^{I}-value.

5.4.7. Errors in Point-Dipole Based 'Experimental Distances'

In many experimental cases measured D-values are used to determine distances.[38–40, 174, 175] Thus, it is of critical importance to determine the errors in distance that are introduced in the experimental analysis by assuming the point-dipole approximation to be valid. To this end the NO-NO midpoint distances in the molecular structures are compared with the NO-NO midpoint distances that are calculated using the point-dipole approximation and the values of D^{I} (Figure 5.16).

For the saturated linker the derived distances are 1–2 Å smaller than the distances in the

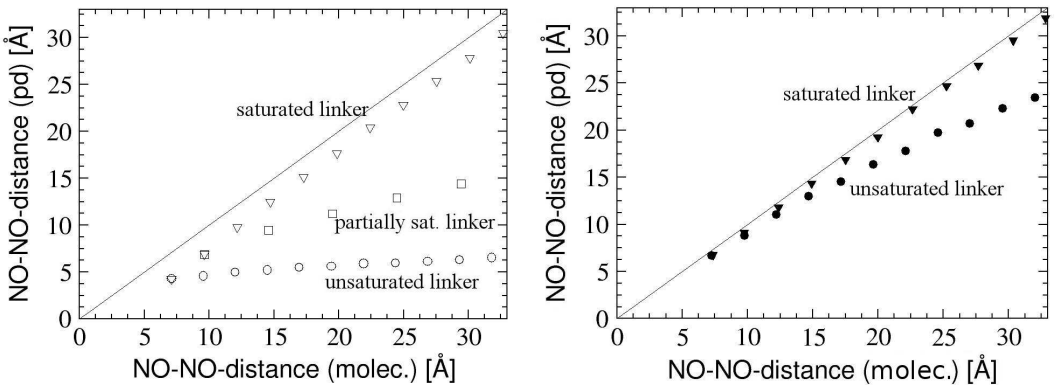


Figure 5.16. – Comparison of the NO-NO midpoint distances in the molecular structures (molec.) and the NO-NO midpoint distances calculated using the point-dipole approximation and the D^{I}-value for the aromatic nitroxyl ring (left figure) and for the non-aromatic nitroxyl ring (right figure). Results are given for the saturated linker (lower triangles), the unsaturated linker (circles) and the partially saturated linker (open squares).

molecular structures. The spin delocalization from the N-O group into the aromatic ring is much less significant for non-aromatic nitroxyl spin labels, since in this case the saturated ring effectively prevents delocalization of the spin onto the bridge. As expected, the error is larger at small linker lengths, since in this regime the neglect of the exchange-contribution to the D-value introduces a non-negligible error in the derived distance. The distances derived for the unsaturated linker increase only slightly with increasing linker length. Here, the approximation of the spin being localized on the nitroxyl group is not valid and introduces substantial error for shorter linkers. The errors in distances derived for the partially saturated linker are in between the extremes for saturated and unsaturated linkages.

For the non-aromatic nitroxyl rings that are used in most spin labeling experiments[189] the SOMOs are much more localized on the nitroxyl group. In this situation the point-dipole approximation becomes significantly better. If the rings are connected by a saturated linker, the error in distance calculated by the point-dipole approximation ranges from 0.5 Å at the

smallest distance to 1 Å at an NO-NO distance of 32 Å. Fitting the data to the power law $D = A \cdot \left(\frac{1}{x+b}\right)^a$, we find a perfect inverse third power law with the parameter b of -0.5. With an N-O bond length of 1.3 Å the 'contracted' point-dipoles can thus be found predominantly on the nitrogen nucleus. For the unsaturated linker the results predicted become worse again. Here, the errors of the predicted distance accumulate up to 8.5 Å for an experimental distance of 32 Å. Although the bonds next to the nitroxyl group are σ-bonds, both SOMOs delocalize over them and a small — but critical for the distance measurements — spin population (about 1%) is found on the unsaturated linker. Due to this effect the power function for the unsaturated linker decays slightly more slowly (as $R^{-2.9}$ with $b \approx 0.3$). In this case, the 'contracted' point-dipoles are located essentially in the middle of the N-O bond.
The bottom line of this analysis is that, whenever there are significant spin-delocalization effects, the point-dipole approximation is not a reliable way to interpret dipolar splittings measured by EPR.

5.4.8. Further Validation through Comparison of Calculated and Experimental Zero-Field Splittings

The study is further validated by examining two diradicals (depicted in Figure 5.17) for which interspin distances were derived from double electron-electron resonance (DEER) measurements by using the point-dipole approximation.[174,175] The thus derived interspin distance

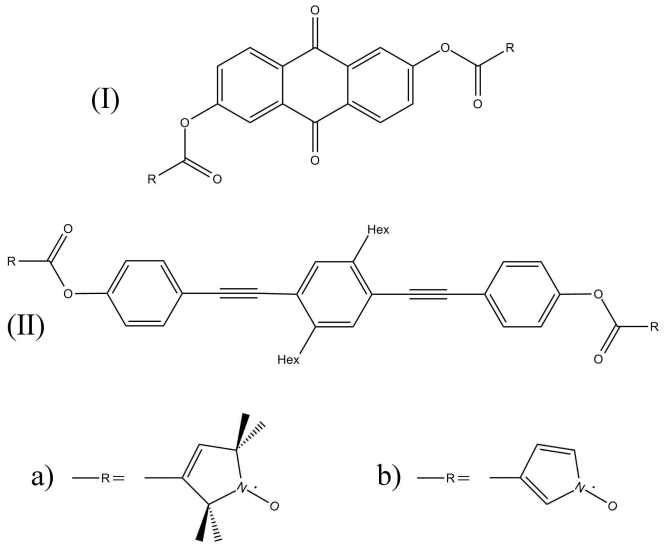


Figure 5.17. – Structural diagram of diradical **7** (I) and diradical **8** (II). These diradicals were studied with non-aromatic oxypyrrolin groups (a) and aromatic pyrroloxyl groups (b).

and the experimental D-value are compared with the calculated values from DFT optimized models. For these DFT calculations, model structures were built for the diradicals. The molecule structures were optimized and D' was calculated without any further approximation as explained in Section 5.3. An interspin distance was then derived from the calculated D'-value. If the point-dipole approximation is valid, this calculation should give the same

distance as the one that is present in the optimized molecular structures.
For both diradicals the interspin distances measured from the optimized structure and calculated from the DFT spin densities are in reasonable agreement (see Table 5.4). Diradical

Table 5.4. – N-N distances for diradicals **7** and **8** measured experimentally from the molecular structure (with nitroxyl ring a) and determined via DFT-calculations from D^I (with nitroxyl rings a and b).

System	Diradical **7**		
	D [10^{-6} cm^{-1}]	Interspin distance [Å]	N-N-distance [Å]
Experimental results[a]	-341 ± 5	19.73 ± 0.14	–
DFT model (a)	-335	19.8	19.11[c]
DFT model (b)	-455	17.9	18.93[c]

System	Diradical **8**		
	D [10^{-6} cm^{-1}]	Interspin distance [Å]	N-N-distance [Å]
Experimental results[a]	-115 ± 6	28.3 ± 0.5	27.84 ± 0.01[b]
DFT model (a)	-109	28.8	28.46[c]
DFT model (b)	-357	19.4	28.29[c]

[a] Available only for systems 7a and 8a.
[b] N-N distance from X-ray diffraction, from ref.[175]
[c] From optimized structure.

7a gives an N-N distance which is 0.7 Å shorter than the point-dipole interspin distance. For diradical **8a** the N-N distance was measured from X-ray diffraction and is in the range of experimental uncertainty of the point-dipole derived interspin distance. The DFT results show a deviation of about 0.3 Å between the N-N distance and the point-dipole derived interspin distance. These results demonstrate that the point-dipole approximation is valid in these systems, but that uncertainties in the derived distance of a few tenth of an Ångstrom must be tolerated.
The computations were then taken one more step forward. An *in silico* experiment was performed in which the spin-carrying non-aromatic oxypyrrolin groups were substituted by aromatic pyrroloxyl groups, in order to see the effect of delocalization. The results of these calculations are also given in Table 5.4.
For diradical **7b**, the interspin distance derived from D^I and the point-dipole approximation is compared to the results obtained for the non-aromatic ring. The change in the inferred distance is about 2 Å, whereas the N-N distance in the molecule only differs by about 0.2 Å. This change in the point-dipole interspin distance is due to the delocalization of the SOMOs in the aromatic ring.
For diradical **8b**, the substitution of the non-aromatic ring by the aromatic nitroxyl ring changes the results even more drastically. The calculated D^I-value increases by a factor of three. The derived point-dipole interspin distance shows a decrease of almost 10 Å (from

28.8 Å down to 19.4 Å), although the N-N distance shrinks only by about 0.2 Å. This dramatic effect is not only due to the delocalization of the SOMOs onto the ring fragment, but also due to a very small amount of delocalization onto the bridging linker (which is a large conjugated system). This small delocalization occurs despite the presence of the ester groups, which separate the conducting bridge from the ring fragments, and is responsible for the drastic effect on the derived interspin distance.
Thus, as soon as the nitroxyl group becomes aromatic, significant delocalization effects take place and the point-dipole derived distances have to be treated with caution.

5.4.9. Separation of the Heisenberg Isotropic Exchange Interaction and the Anisotropic Dipolar Interaction

For most systems where distance measurements are done by DEER, it is assumed that the Heisenberg isotropic exchange coupling constant J is negligible.[39,40] But if J is of the same order of magnitude as the anisotropic dipolar splitting D', problems can arise for the measurement of distances.[190] Under these circumstances the splittings of the energy levels are functions of both D' and J.[a]
One can therefore ask for what situations are J and D' of the same order of magnitude? In order to address this question calculations of the isotropic exchange coupling constant J were carried out for:

(a) the fully optimized structure,

(b) two aromatic nitroxyl rings connected by a saturated linker and

(c) two non-aromatic nitroxyl rings connected by a saturated linker.

For (b) and (c) the J- and D'-values were calculated as a function of increasing linker length. Generally it is assumed that the isotropic exchange interaction decreases exponentially with increasing spin-spin distance following the equation $J = A \cdot \exp(-\beta \cdot R)$, where R is the spin-spin distance.[192,193] For our system thus a least squares fit of the calculated J-values with the NO-NO distance was performed to yield the proportional factor β.
The D-value for the fully relaxed structures was already discussed above and has a magnitude of 0.020 cm^{-1}. The exchange coupling constant J for both fully relaxed structures is 37.52 cm^{-1} and -32.41 cm^{-1}, respectively.[b] That is, the J-value is three orders of magnitude higher than the D-value and thus the EPR signals do not depend on the Heisenberg exchange coupling.
The magnitude of the calculated D'- and J-values for the aromatic nitroxyl ring are shown in the logarithmic plot in Figure 5.18 as a function of intervening linker length. The D-value shows the already discussed R^{-3}-behavior. The J-value apparently shows an exponential dependence on the linker length. Up to the C_4-linker the J-value is much larger than D', i.e. we are in the *strong exchange limit*, the observed EPR signals will be only within the $S = 1$ level manifold. Only for the C_6- and the C_8-linker we are in the *intermediate exchange limit*, D' and J are on the same order of magnitude and the observed EPR signals will depend on both. From the C_{10}-linker on the *weak exchange limit* is reached, the isotropic exchange

[a] Nevertheless the effects of J and D' can in principle be separated by careful analysis of the full 'dipolar' splitting pattern. Exchange interactions through 8–12 bonds in spin-labeled metal complexes are large enough to impact continuous wave EPR line shapes at distances of up to ca. 14 Å.[191]

[b] For antiferromagnetic coupling.

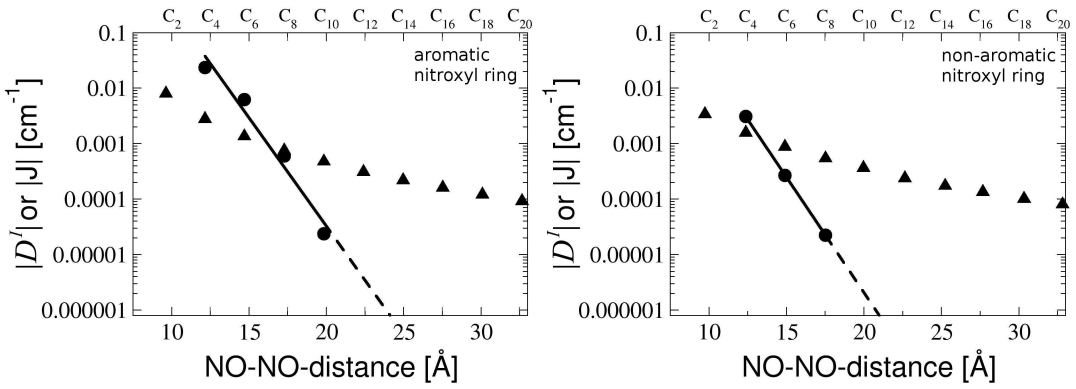


Figure 5.18. – Comparison of the dipolar splitting D^I and the Heisenberg exchange coupling constant J as a function of the NO-NO midpoint distance for the saturated linker with the aromatic nitroxyl ring (left figure) and with the non-aromatic nitroxyl ring (right figure). Results are given for the magnitude of D^I (triangles) and for the magnitude of J (circles). A least squares fit was done for the J-values (black line, extrapolation for higher NO-NO distances as dashed line).

coupling is much smaller than the anisotropic dipolar splitting and the interpretation of the dipolar splitting is predicted to become simple again.

For the saturated linker with the non-aromatic nitroxyl ring the absolute D^I- and J-values are shown in Figure 5.18. The D-values are lower than for the aromatic ring for comparable linker length, but show the same R^{-3}-behavior. As expected, the J-value shows also here an exponential dependence on the linker length. The exchange coupling constant is much lower than with the aromatic rings for comparable linker lengths. This is due to the spins that are mainly localized in the N-O bond for the non-aromatic rings. For the non-aromatic ring D^I and J are on similar scales up to the C_6-linker. Here we are in the intermediate exchange regime. From the C_8-linker on the weak exchange limit is reached, where J is much smaller than D^I and can be assumed to be negligible in the EPR signals. The linker length where J- and D-values are on similar scales, and where thus the splittings of the energy levels are functions of both values, is shorter by about 3 Å for the non-aromatic rings than for the aromatic rings.

Finally, a least squares fit was performed to study the correlation between the isotropic exchange interaction and the NO-NO distance. Here, J was fitted to the equation $J = A \cdot \exp(-\beta \cdot R)$ and β was found to be approximately 1 Å^{-1} for both sets (0.90 for the aromatic nitroxyl ring and 0.95 for the non-aromatic nitroxyl ring). This value may be compared with estimates of $\beta \approx 0.7$ Å^{-1} for the electron transfer matrix element H_{AB} for covalent linkages.[194] Relationships between both quantities have been discussed.[195,196]

5.5. Conclusions

In the present chapter a detailed analysis of the point-dipole approximation for the evaluation of interspin distances was presented. The focus of the investigation has been to understand under which circumstances this central approximation is likely to fail. This is of critical importance in methods for determining interspin distances by electron paramagnetic resonance. The bottom line of the present analysis is the following: delocalization of the spin density

strongly affects the D-value and leads to errors in the interspin distances obtained using the point-dipole model. Distances calculated using the point-dipole approximation usually are underestimated. This is due to delocalization effects that lead to tails of the spin density being closer together than the majority of the spin density. Spins at closer distances are much more strongly weighted by the dipole-dipole interaction operator than at larger distances. Thus even small delocalization effects lead to errors of 1-2 Å in derived distances. These effects become particularly dramatic for aromatic nitroxides or if there are bridging groups between the two spin-carrying fragments that are (partially) unsaturated. In the latter case the D-value not only deviates from the expected R^{-3} behavior but also becomes significantly dependent on the relative orientation of the two spin-carrying fragments. Any type of point-dipole model is bound to fail under such circumstances and derived distances from such models have substantial errors. Thus, whenever the spin label or the bridge that separates the two paramagnetic centers is significantly unsaturated one is well advised to avoid the point-dipole approximation. If there is sufficient structural information available, quantum chemical calculations can be of great help in the analysis of experimental data.

6. Conclusion

In the present Ph.D. thesis three different subjects are adressed.

1. The geometry optimization methodology in ORCA is improved. Constrained optimization, TS optimization and several initial Hessians for TS optimization are developed and implemented. Furthermore, a QM/MM interface between ORCA and GROMACS is established. All implemented techniques are described in Chapter 2. They are applied in the calculations that are presented in Chapters 3 to 5.

2. The reaction mechanism of the enzyme Cytochrome P450nor is studied in a realistic enzyme environment for the first time (Chapter 3 and 4). The study is pursued with the aim to identify the key intermediate *I* of the reaction as well as to analyze the unusual reduction process in detail.

3. The point-dipole approximation, an approximation used in EPR spectroscopy to determine distances from measured ZFS data, is analyzed in detail (Chapter 5). The aim of this second investigation is to predict under which circumstances this central approximation is likely to fail.

P450 NO Reductase

QM/MM calculations are employed on P450 NO reductase. Two different enzymatic systems are used, one including the reducing cofactor NADH (**P1**), and one without the cofactor (**P2**). The systems (**P1**/**P2**) consist of 45487/50459 atoms, containing 13061/14692 solvent molecules, and the used QM regions consist of 44–142 atoms.

Properties of Intermediates In Chapter 3 all three intermediates of the P450nor catalyzed reaction are investigated. For the resting state, a water molecule is bound to the distal heme side (**1**). In the NO-bound state, NO is modeled in the active site (**2**). The calculations show that two different electronic states exist (**2a**, **2b**). For intermediate *I* six species are studied (**3a**/**3b**, **4a**/**4b**, **5a**/**5b**). The structural and electronic properties are calculated and analyzed for all models. Relative energies, Mössbauer shifts and IR frequencies are calculated. Experimental data are used to identify the intermediates.
Through comparison with experimentally derived structural, Mössbauer and IR data the NO-bound species is identified as the closed shell form **2a** of NO^+ bound to low-spin Fe(II). The calculated structural properties and IR frequencies for the BS(2,2) $\{FeNO\}^6$ species **2b** indicate that this form might play a role in the NO-bound state of the related enzyme Cytochrome P450cam. Out of the six putative species which are taken into account for intermediate *I* three (**4a**, **5a** and **5b**) can be excluded due to comparison with experimental Mössbauer data. Apart from the spectroscopic result, **3b** is favored energetically over **3a** by about 5 kcal/mol. This leaves **3b** and **4b** as putative intermediate *I* species. Both forms are Fe(III) species with an antiferromagnetically coupled ligand, differing only in the protonation state: singlet diradical Fe(III)-$NHO^{\bullet -}$ (**3b**) and Fe(III)-$NHOH^{\bullet}$ (**4b**).

Reaction Course and Identification of Intermediate *I* In Chapter 4 further QM/MM calculations are presented. The reaction cycle is analyzed by connecting the putative intermediates (**1**, **2a**, **3b** and **4b**) and the end products N_2O and H_2O through relaxed surface scans in the framework of the QM/MM environment using **P1** and **P2**. Factors influencing NADH binding and reduction by the cofactor are analyzed in detail. The kinetic isotope effect and activation barrier of the reduction step are calculated and compared to the experimental data. The relative energies of all intermediates and transition states of the modeled reaction cycle are used to derive the kinetic and thermodynamic constants for all elementary reaction steps. These are finally used to simulate the time course of the reaction under different experimentally relevant reaction conditions.
The reduction of NO-bound P450nor by NADH is modeled on two PESs corresponding to closed shell and diradical states of the intermediates. The calculated kinetic isotope effect and reaction rate constant for (**2a** $\rightarrow$ **3b**) agree with the respective experimental data. It is thus shown that the reaction proceeds via a diradical transition state. The electronic structure of the transition state suggests that a hydride ion is transferred from NADH to the NO ligand. Besides, the reduction in P450nor is carried out in a rather unusual two-state reactivity mechanism, in which, instead of the crossing of two states of different multiplicity, two singlet states with different electronic structures cross.
In addition, the influence of the protein structure on NADH binding and reactivity is studied extensively. With the results from MD simulations and QM/MM calculations it can be shown that Ser_{286} and Thr_{243} are highly important for NADH binding. Their hydrogen bonds to nicotinamide position the cofactor next to the reaction site. The QM/MM results show that the strength of the hydrogen bonds influences the activation barrier of the reduction step. Both residues are thus essential for the preparation as well as for the realization of the reduction step in P450nor. This theoretical result explains experimental observations obtained from site-directed mutagenesis studies.
Apart from the reduction step, the even more important result in Chapter 4 is the identification of intermediate *I*. The P450nor reaction course is simulated at different NO concentrations ([NO] equal to or much larger than the enzyme concentration). The simulations show that both species, **3b** and **4b**, occur in the reaction cycle, but only the diprotonated form accumulates at low NO concentration, making it amenable to spectroscopic identification. The results from Chapters 3 and 4 identify intermediate *I* as the singlet diradical Fe(III)-NHOH form **4b**, described as $NHOH^{\bullet}$ antiferromagnetically coupled to low-spin ferric iron.

P450nor Reaction Cycle The entire reaction cycle of Cytochrome P450nor, as it is conclusively revealed in this work, is illustrated in Figure 6.1. Starting from the resting state, one molecule of NO binds to the enzyme after the water molecule is dissociated. The incoming NO reduces the ferric iron resulting in NO^+ bound to iron with strong backdonation from iron to the NO ligand. The cofactor NADH is positioned in direct proximity to the reaction site with the aid of the amino acids in the heme distal pocket. The reducing hydrogen of nicotinamide faces the Fe-NO unit. After this preparation NADH can reduce the $\{FeNO\}^6$ moiety by transferring a hydride ion. This reduction is the step with highest activation barrier ($\approx$ 7 kcal/mol) in the whole reaction cycle, as is also observed in experiment. While the NO^+ ligand is reduced, the iron gives back one electron to the ligand resulting in a transient singlet diradical Fe(III)-$NHO^{\bullet -}$. This species has a high proton affinity and thus is readily protonated. The resulting intermediate *I* can be described as an $NHOH^{\bullet}$ ligand that is antiferromagnetically coupled to ferric iron. The second NO molecule binds to intermedi-

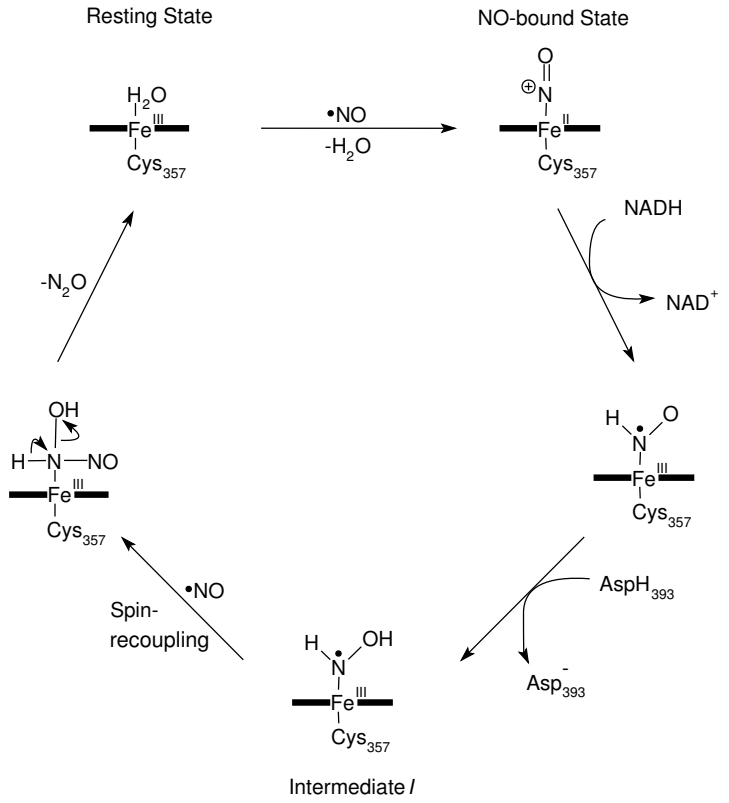


Figure 6.1. – The reaction cycle of the reaction catalyzed by Cytochrome P450nor according to the results obtained in this thesis.

ate *I* in a barrierless, spin-recoupling manner. In this step the N-N bond is formed, and the ligand looses its tight bond with iron. The loosely (via hydrogen bonds) bound 'HNOHNO' tautomerizes and yields the end products N_2O and H_2O.

Point-Dipole Approximation for Spin-Spin Zero-Field Splitting

The last part of the Ph.D. thesis is presented in Chapter 5. Density functional theory is used to calculate the structural and spectroscopic EPR properties of different dinitroxyl systems. The dinitroxyl systems serve as model systems for the study of different contributions to the spin-spin zero-field splitting. It is shown that the exchange-like part of the spin-spin ZFS is low and vanishes under certain circumstances. The biggest contribution to the spin-spin ZFS is shown to be the 2-center-Coulomb-integral contribution. Different approximations (point-dipole, center-of-gravity and distributed point-dipole) to the spin-spin ZFS are used for the determination of interspin distances of dinitroxyls with varying linker length. The comparison to the exactly calculated ZFS shows the performance of the approximations under varying circumstances. It is found that distances calculated using the point-dipole approximation are usually underestimated. The effect is due to delocalization of spin density. This leads to tails of the spin density that are closer together than the majority of the spin density. The spin density at closer distances contributes much stronger to the dipole-dipole interaction than at larger distances. Thus, delocalization of the spin density strongly affects the ZFS-value and leads to errors in the interspin distances that are obtained when using the point-dipole

model. Even for small delocalization these errors sum up to deviations in derived distances of 1–2 Å. In particular, the discrepancy becomes dramatic for aromatic nitroxides or if the linker connecting the nitroxyls is unsaturated or partially unsaturated. In this case, the ZFS-value can also depend on the relative orientation of the two spin-carrying fragments. Here, all point-dipole models fail and the derived distances have substantial errors.

Appendices

A. Molecular Mechanics Force Fields

Molecular mechanics (MM) is a method to describe molecular systems on the basis of classical mechanics. The basic idea is that a molecular system consists of particles, i.e. atoms (all-atom model) or atom groups (united-atom model), and that their interaction can be described by simple interaction terms.
A *force field* for an MM calculation collects

- information on the potential energy function of the system and
- a parameter set that is needed to calculate the interaction energy.

A typical energy expression in an MM force field is given in equation (A.1).

$$E_{MM} = \underbrace{E_{bonds} + E_{angles} + E_{dihedrals}}_{E_{bonded}} + \underbrace{E_{el} + E_{VDW}}_{E_{nonbonded}} \tag{A.1}$$

$$= \underbrace{\sum_{bonds} k_b(d - d_0)^2}_{E_{bonds}} + \underbrace{\sum_{angles} k_\theta(\theta - \theta_0)^2}_{E_{angles}} + \underbrace{\sum_{dihedrals} k_\phi \left[1 + cos\left(n\phi + \delta\right)\right]}_{E_{dihedrals}}$$

$$+ \underbrace{\sum_{\substack{\text{nonbonded} \\ \text{pairs } AB}} \epsilon_{AB} \left[\left(\frac{\sigma_{AB}}{r_{AB}}\right)^{12} - \left(\frac{\sigma_{AB}}{r_{AB}}\right)^{6}\right]}_{E_{el}} + \underbrace{\sum_{\substack{\text{nonbonded} \\ \text{pairs } AB}} \frac{1}{4\pi\epsilon_0} \frac{q_A q_B}{r_{AB}}}_{E_{VDW}} .$$

E_{MM} contains

- *bonded* and
- *nonbonded* interactions.

The *bonded interaction* consists of terms for

- bond stretching (E_{bonds}),
- angle bending (E_{angles}) and
- torsional distortion ($E_{dihedrals}$).

E_{bonds} and E_{angles} depend on the distortion of the actual distance d or bond angle θ from the equilibrium distance d_0 or bond angle θ_0. The interaction term for the dihedral angles $E_{dihedrals}$ is a bit more complicated, including the actual dihedral angle ϕ, the multiplicity of the function n and the phase shift δ.
The *nonbonded interaction* sums up

- the electrostatic interaction E_{el} between charged particles and
- the VDW interaction between the particles, here described by a Lennard-Jones (LJ) potential E_{VDW}.

They are dependent on the distance r_{AB} of the nonbonded atom pair and the 'equilibrium VDW distance' σ_{AB} or the charges (q_A and q_B) of both atom types.

In an all-atom force field one distinguishes between *atom types* (e.g. a carbon atom can

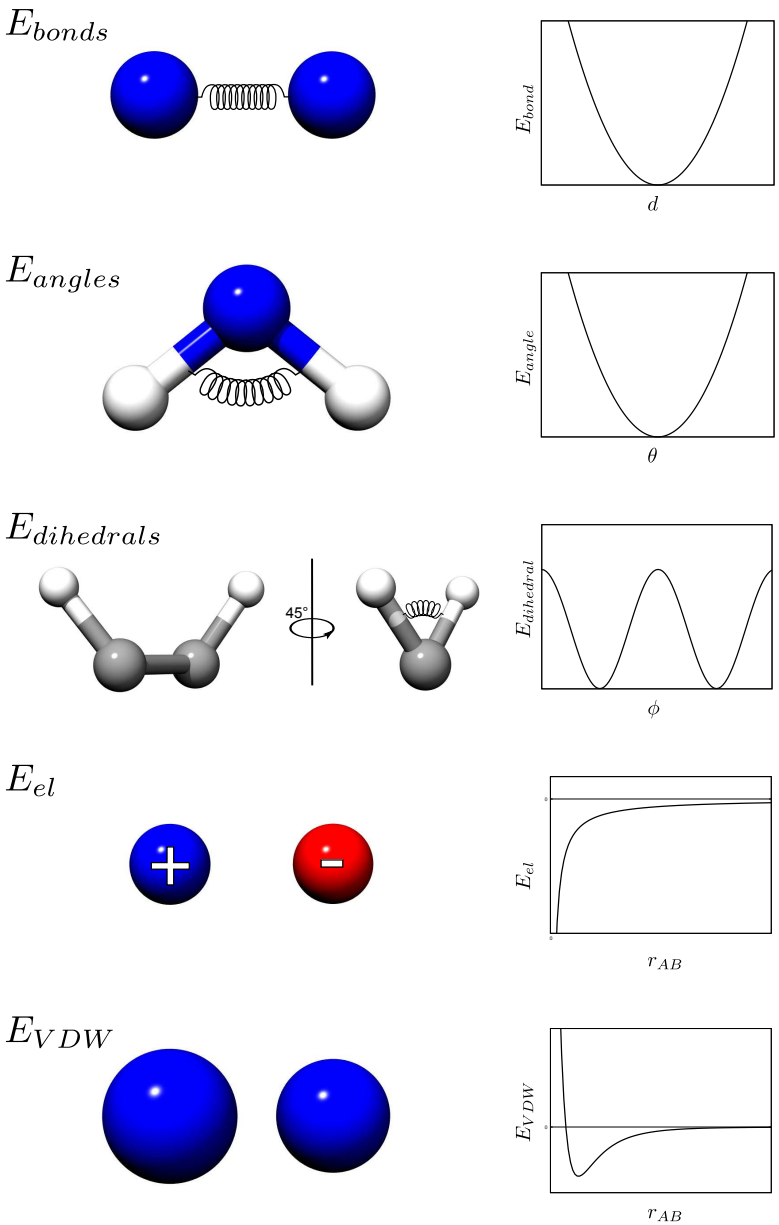


Figure A.1. – Visualization of the definition of interaction terms arising in a typical molecular mechanics force field.

have different atom types, depending on the functional group that it belongs to — a methyl, a methylene group, etc.).

The parameters for the bonded interaction (k_b, d_0, k_θ, θ_0, k_ϕ, n and δ, as given above) are tabulated in the force field. The parameters are different for the pair, triplet or quadruplet of atom types that the respective bond, bond angle or dihedral angle consists of.
Also the parameters for the nonbonded interaction are tabulated in the force field. Here, each atom type is assigned a 'VDW radius' σ and a partial charge (q_A, q_B).
In an MM calculation the *topology* of the molecular system has to be defined. In the topology

- the atom type of each atom and
- the connectivities of all atoms in the system

are defined. The connectivity defines for which subsets of atoms the bonding terms (*rightarrow* E_{bonds}, E_{angles} and $E_{dihedrals}$) are calculated (see also Figure A.1).
The nonbonded interaction is calculated for all atom pairs with the exception of those that are connected by three adjacent bonds or less.

The typical bonded and nonbonded terms are visualized in Figure A.1. It can be seen that E_{bonds} and E_{angles} are simple 'spring forces'. $E_{dihedrals}$ is also a 'spring force', but its shape is somewhat different due to the symmetry present in dihedral torsion.
It can also be seen that there is an equilibrium distance for the LJ interaction, but of course not for the electrostatic interaction. Two oppositely charged particles attract each other, and in the classical force field they are only prevented to unify due to the strong repulsive character ($\propto \frac{1}{r_{AB}^{12}}$) of the LJ interaction at very short distances.

B. Additional Results on P450nor Study

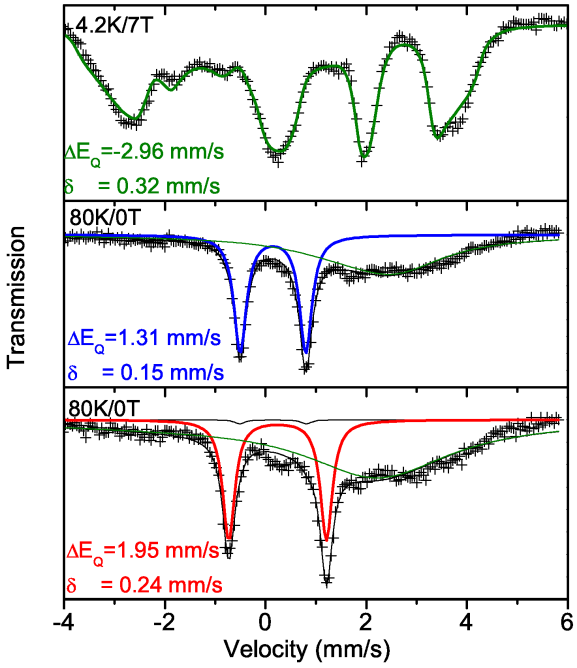


Figure B.1. – Experimental Mössbauer spectra for the three intermediates of the P450nor reaction cycle. Upper figure: Applied field (7 T) Mössbauer spectrum of the resting state at 4.2 K. Mid figure: Zero-field Mössbauer spectrum of the NO-bound state at 80 K. Lower figure: Zero-field Mössbauer spectrum of intermediate *I* at 80 K. The experiments were carried out and the spectra were kindly provided by Eckhard Bill and Frank Neese.

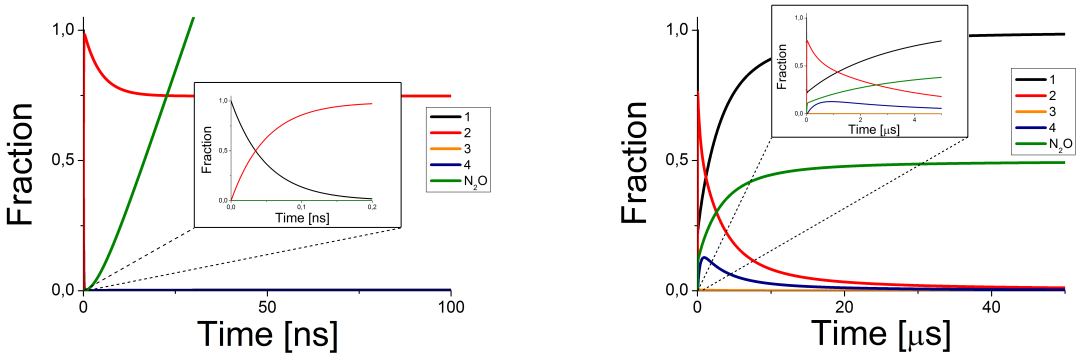


Figure B.2. – Reaction course of the P450nor reaction cycle at high (left, **K1**) and low (right, **K2**) NO concentration at 10°C using the results of snapshot B for the reduction step.

Table B.1. – Calculated energies for the intermediates of the hydride transfer reaction (see Figure 4.10) for the three snapshots. The QM/MM energies refer to the structures obtained for snapshot A/B/C with QM region **R1**. Relative energies are given in kcal/mol and refer to the lowest energy of both electronic states of the respective snapshot. The QM/MM energies already include ZPE correction.

	Closed shell	BS
NADH + $\{FeNO\}^6$	0.3/0.4/0.0	0.0/0.0/0.1
TS	20.3/10.2/13.5	15.7/6.2/10.6
NAD^+ + $\{FeNHO\}^8$	10.6/8.5/6.5	5.5/3.7/1.6

Table B.2. – Selected structural properties for the educt (upper table), transition state (middle table) and product (lower table) of the hydride transfer reaction for snapshots A/B/C. The structural data refer to the QM/MM optimization with QM-region **R2**. Bond distances are given in Ångstrom, angles are given in degree.

Educt		
Distance/Angle	Electromer	
	2a	**2b**
Fe-N(NO)	1.66/1.66/1.65	1.70/1.71/1.70
N-O	1.14/1.14/1.14	1.15/1.15/1.14
Fe-N-O angle	165.5/162.6/166.3	164.7/161.7/165.6

Transition state		
Distance/Angle	Electromer	
	TS(2,3)a	**TS(2,3)b**
Fe-N(NO)	1.78/1.78/1.77	1.92/1.93/1.93
N-O	1.19/1.19/1.19	1.19/1.19/1.19
Fe-N-O angle	139.9/140.6/141.1	139.2/140.7/139.6

Product		
Distance/Angle	Electromer	
	3a	**3b**
Fe-N(NO)	1.80/1.80/1.80	1.92/1.91/1.91
N-O	1.23/1.23/1.23	1.24/1.25/1.25
Fe-N-O angle	133.1/137.1/134.9	129.7/135.5/132.1

Bibliography

[1] A.L. Lehninger, D.L. Nelson, and M.M. Cox. *Lehninger principles of biochemistry.*, volume 1. Wh Freeman, 2005.

[2] W.P. Jencks. *Catalysis in chemistry and enzymology.* Dover Publications, 1987.

[3] S.J. Lippard and J.M. Berg. *Principles of bioinorganic chemistry.* Univ Science Books, 1994.

[4] M. Zink and H. Grubmüller. Mechanical properties of the icosahedral shell of Southern Bean Mosaic virus: a molecular dynamics study. *Biophys. J.*, 96(4):1350–1363, 2009.

[5] S. Shaik, S. Cohen, Y. Wang, H. Chen, D. Kumar, and W. Thiel. P450 Enzymes: Their Structure, Reactivity, and Selectivity — Modeled by QM/MM Calculations. *Chem. Rev.*, 110(2):949–1017, 2009.

[6] A.M. Valentine and S.J. Lippard. Principles of small molecule activation by metalloenzymes as exemplified by the soluble methane monooxygenase from Methylococcus capsulatus. *J. Chem. Soc., Dalton Trans.*, (21):3925–3932, 1997.

[7] H. Hara, T. Tenno, and M. Shirakawa. Distance determination in human ubiquitin by pulsed double electron-electron resonance and double quantum coherence ESR methods. *J. Magn. Reson.*, 184(1):78–84, 2007.

[8] O. Schiemann and T.F. Prisner. Long-range distance determinations in biomacromolecules by EPR spectroscopy. *Q. Rev. Biophys.*, 40(1):1–53, 2007.

[9] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. *Molecular biology of the cell.* 2002.

[10] J.C. Kendrew, G. Bodo, H.M. Dintzis, R.G. Parrish, H. Wyckoff, and D.C. Phillips. A three-dimensional model of the myoglobin molecule obtained by X-ray analysis. *Nature*, 181(4610):662–666, 1958.

[11] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, and P.E. Bourne. The protein data bank. *Nucleic Acids Res.*, 28(1):235, 2000.

[12] Protein Database Website. `http://www.pdb.org`, 2011.

[13] K. Wuethrich. The development of nuclear magnetic resonance spectroscopy as a technique for protein structure determination. *Accounts Chem. Res*, 22(1):36–44, 1989.

[14] B.M. Hoffman. Electron nuclear double resonance (ENDOR) of metalloenzymes. *Accounts Chem. Res.*, 24(6):164–170, 1991.

[15] W. Fu, P.M. Drozdzewski, M.D. Davies, S.G. Sligar, and M.K. Johnson. Resonance Raman and magnetic circular dichroism studies of reduced [2Fe-2S] proteins. *J. Biol. Chem.*, 267(22):15502, 1992.

[16] B.A. Brennan, J.G. Cummings, D.B. Chase, I.M. Turner Jr, and M.J. Nelson. Resonance Raman spectroscopy of nitrile hydratase, a novel iron-sulfur enzyme. *Biochemistry*, 35(31):10068–10077, 1996.

[17] S.D. Garton, J. Hilton, H. Oku, B.R. Crouse, KV Rajagopalan, and M.K. Johnson. Active site structures and catalytic mechanism of Rhodobacter sphaeroides dimethyl sulfoxide reductase as revealed by resonance Raman spectroscopy. *J. Am. Chem. Soc.*, 119(52):12906–12916, 1997.

[18] F. Neese. Quantum chemical calculations of spectroscopic properties of metalloproteins and model compounds: EPR and Mössbauer properties. *Curr. Opin. Chem. Biol.*, 7(1):125–135, 2003.

[19] S.C. Bhatia and N. Ravi. A Mössbauer study of the interaction of chitosan and D-glucosamine with iron and its relevance to other metalloenzymes. *Biomacromolecules*, 4(3):723–727, 2003.

[20] M.R. Seyedsayamdost, T. Argirevic, E.C. Minnihan, J.A. Stubbe, and M. Bennati. Structural Examination of the Transient 3-Aminotyrosyl Radical on the PCET Pathway of E. coli Ribonucleotide Reductase by Multifrequency EPR Spectroscopy. *J. Am. Chem. Soc.*, 131(43):15729–15738, 2009.

[21] B.D. Grant and J.A. Adams. Pre-steady-state kinetic analysis of cAMP-dependent protein kinase using rapid quench flow techniques. *Biochemistry*, 35(6):2022–2029, 1996.

[22] D.B. Northrop and F.B. Simpson. Beyond enzyme kinetics: direct determination of mechanisms by stopped-flow mass spectrometry. *Bioorg. Med. Chem.*, 5(4):641, 1997.

[23] K. Zhang, L. Mao, and R. Cai. Stopped-flow spectrophotometric determination of hydrogen peroxide with hemoglobin as catalyst. *Talanta*, 51(1):179–186, 2000.

[24] J. Ruzicka and T. Guebeli. Principles of stopped-flow sequential injection analysis and its application to the kinetic determination of traces of a proteolytic enzyme. *Anal.Chem.*, 63(17):1680–1685, 1991.

[25] P.J. Steinbach and B.R. Brooks. Protein hydration elucidated by molecular dynamics simulation. *P. Natl. Acad. Sci. USA*, 90(19):9135, 1993.

[26] N.J. Krogan, G. Cagney, H. Yu, G. Zhong, X. Guo, A. Ignatchenko, J. Li, S. Pu, N. Datta, A.P. Tikuisis, et al. Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. *Nature*, 440(7084):637–643, 2006.

[27] F. Claeyssens, J.N. Harvey, F.R. Manby, R.A. Mata, A.J. Mulholland, K.E. Ranaghan, M. Schütz, S. Thiel, W. Thiel, and H.J. Werner. High-Accuracy Computation of Reaction Barriers in Enzymes. *Angew. Chem.*, 118(41):7010–7013, 2006.

[28] F. Jensen. *Introduction to computational chemistry.* Wiley, 2007.

[29] P.E.M. Siegbahn and F. Himo. Recent developments of the quantum chemical cluster approach for modeling enzyme reactions. *J. Biol. Inorg. Chem.*, 14(5):643–651, 2009.

[30] P. Comba and R. Remenyi. Inorganic and bioinorganic molecular mechanics modeling — the problem of the force field parameterization. *Coord. Chem. Rev.*, 238:9–20, 2003.

[31] A. Warshel and M. Levitt. Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.*, 103(2):227–249, 1976.

[32] J.C. Schöneboom, S. Cohen, H. Lin, S. Shaik, and W. Thiel. Quantum Mechanical/Molecular Mechanical Investigation of the Mechanism of C-H Hydroxylation of Camphor by Cytochrome P450cam: Theory Supports a Two-State Rebound Mechanism. *J. Am. Chem. Soc.*, 126(12):4017–4034, 2004.

[33] D. Wang, J. Zheng, S. Shaik, and W. Thiel. Quantum and Molecular Mechanical Study of the First Proton Transfer in the Catalytic Cycle of Cytochrome P450cam and Its Mutant D251N. *J. Phys. Chem. B*, 112(16):5126–5138, 2008.

[34] H.M. Senn and W. Thiel. QM/MM methods for biomolecular systems. *Angew. Chem. Int. Edit.*, 48(7):1198–1229, 2009.

[35] J. Gao. Hybrid quantum and molecular mechanical simulations: an alternative avenue to solvent effects in organic chemistry. *Accounts Chem. Res.*, 29(6):298–305, 1996.

[36] K. Gkionis and J.A. Platts. QM/MM investigation into binding of square-planar platinum complexes to DNA fragments. *J. Biol. Inorg. Chem.*, 14(8):1165–1174, 2009.

[37] S. Difley, L.P. Wang, S. Yeganeh, S.R. Yost, and T.V. Voorhis. Electronic Properties of Disordered Organic Semiconductors via QM/MM Simulations. *Accounts Chem. Res.*, 2010.

[38] P. P. Borbat, J. H. Davis, S. E. Butcher, and J. H. Freed. Measurement of Large Distances in Biomolecules using Double-Quantum Filtered Refocused Electron Spin-Echoes. *J. Am. Chem. Soc.*, 126:7746–7747, 2004.

[39] G. Jeschke and Y. Polyhach. Distance measurements on spin-labeled biomacromolecules by pulsed electron paramagnetic resonance. *Phys. Chem. Chem. Phys.*, 9:1895–1910, 2007.

[40] G. R. Eaton and S. S. Eaton. Measurement of Interspin Distances by EPR. In B.C. Gilbert, editor, *Specialist Periodical Report: Electron Paramagnetic Resonance*, volume 21, pages 59–75. Royal Society of Chemistry, London, 2008.

[41] O. Schiemann, P. Cekan, D. Margraf, T.F. Prisner, and S.T. Sigurdsson. Relative orientation of rigid nitroxides by PELDOR: Beyond distance measurements in nucleic acids. *Angew. Chem. Int. Edit.*, 48(18):3292–3295, 2009.

[42] P. Bertrand, C. More, B. Guigliarelli, A. Fournel, B. Bennett, and B. Howes. Biological polynuclear clusters coupled by magnetic interactions: From the point dipole approximation to a local spin model. *J. Am. Chem. Soc.*, 116:3078–3086, 1994.

[43] F. Neese. First Principle Calculation of EPR Parameters. In B.C. Gilbert, editor, *Specialist Periodical Report: Electron Paramagnetic Resonance*, volume 20, pages 73–95. Royal Society of Chemistry, London, 2007.

[44] 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, 2003.

[45] S. Sinnecker, F. Neese, and W. Lubitz. Dimanganese catalase-spectroscopic parameters from broken-symmetry density functional theory of the superoxidized Mn-III/Mn-IV state. *J. Biol. Inorg. Chem.*, 10(3):231–238, 2005.

[46] S. Kababya, J. Nelson, C. Calle, F. Neese, and D. Goldfarb. Electronic structure of binuclear mixed valence copper azacryptates derived from integrated advanced EPR and DFT calculations. *J. Am. Chem. Soc.*, 128(6):2017–2029, 2006.

[47] M. van Gastel, M. Stein, M. Brecht, O. Schröder, F. Lendzian, R. Bittl, H. Ogata, Y. Higuchi, and W. Lubitz. A single-crystal ENDOR and density functional theory study of the oxidized states of the [NiFe] hydrogenase from Desulfovibrio vulgaris Miyazaki F. *J. Biol. Inorg. Chem.*, 11:41–51, 2006.

[48] S. Kossmann, Kirchner. B., and F. Neese. Performance of modern Density Functional Theory for the prediction of hyperfine structure: meta-GGA and double-hybrid functionals. *Molec. Phys.*, 105(15-16):2049–2071, 2007.

[49] K. Ray, E. Bill, T. Weyhermüller, and K. Wieghardt. Redox-Noninnocence of the S,S'-Coordinated Ligands in Bis(benzene-1,2-dithiolato)iron Complexes. *J. Am. Chem. Soc.*, 127(15):5641–5654, 2005. PMID: 15826204.

[50] M.R. Pederson and S.N. Khanna. Magnetic Anisotropy Barrier for Spin Tunneling in $Mn_{12}O_{12}$ Molecules. *Phys. Rev. B.*, 60(13):9566–9572, 1999.

[51] M. Shoji, K. Koizumi, T. Hamamoto, T. Taniguchi, R. Takeda, Y. Kitagawa, T. Kawakami, M. Okumura, S. Yamanaka, and K. Yamaguchi. A theoretical study of zero-field splitting of organic biradicals. *Polyhedron*, 24(16-17):2708–2715, 2005.

[52] R. Takeda, S. Mitsuo, S. Yamanaka, and K. Yamaguchi. Density functional study of zero-field splitting. *Polyhedron*, 24(16-17):2238–2241, 2005.

[53] F. Neese. Importance of Direct Spin-Spin Coupling and Spin-Flip Excitations for the Zero-Field Splittings of Transition Metal Complexes: A Case Study. *J. Am. Chem. Soc.*, 128:10213–10222, 2006.

[54] S. Sinnecker and F. Neese. Spin-spin contributions to the zero-field splitting tensor in organic triplets, carbenes and biradicals - A density functional and ab initio study. *J. Phys. Chem. A*, 110(44):12267–12275, 2006.

[55] F. Neese. Analytic Derivative Calculation of Zero-Field Splittings. *J. Chem. Phys.*, 127:164112, 2007.

[56] J.E. Harriman. *Theoretical Foundations of Electron Spin Resonance.* Academic press, New York, 1978.

[57] R. McWeeny. *Methods of Molecular Quantum Mechanics.* Academic press, London, 1992.

[58] B.A. Averill. Dissimilatory nitrite and nitric oxide reductases. *Chem. Rev.*, 96(7):2951–2964, 1996.

[59] J. Garthwaite. Glutamate, nitric oxide and cell-cell signalling in the nervous system. *Trends Neurosci.*, 14(2):60–67, 1991.

[60] J. Garthwaite and C.L. Boulton. Nitric oxide signaling in the central nervous system. *Annu. Rev. Physiol.*, 57(1):683–706, 1995.

[61] F.L.M. Ricciardolo. Multiple roles of nitric oxide in the airways. *Thorax*, 58(2):175, 2003.

[62] F.L.M. Ricciardolo, P.J. Sterk, B. Gaston, and G. Folkerts. Nitric oxide in health and disease of the respiratory system. *Physiol. Rev.*, 84(3):731, 2004.

[63] L.J. Ignarro. *Nitric oxide: biology and pathobiology.* Academic Pr, 2000.

[64] P.R.O. De Montellano. *Cytochrome P450: structure, mechanism, and biochemistry.* Springer, 2005.

[65] M. Klingenberg. Pigments of rat liver microsomes. *Arch. Biochem. Biophys.*, 75(2):376–386, 1958.

[66] A. Sigel, H. Sigel, and R.K.O. Sigel. *The ubiquitous roles of cytochrome P450 proteins.* Wiley, 2007.

[67] R.G. Bergman. C–H activation. *Nature*, 446(7134):391–393, 2007.

[68] H. Shoun and T. Tanimoto. Denitrification by the fungus Fusarium oxysporum and involvement of cytochrome P-450 in the respiratory nitrite reduction. *J. Biol. Chem.*, 266(17):11078, 1991.

[69] Y. Shiro, M. Fujii, T. Iizuka, S. Adachi, K. Tsukamoto, K. Nakahara, and H. Shoun. Spectroscopic and kinetic studies on reaction of cytochrome P450nor with nitric oxide. Implication for its nitric oxide reduction mechanism. *J. Biol. Chem.*, 270(4):1617–1623, 1995.

[70] H. Shimizu, E. Obayashi, Y. Gomi, H. Arakawa, S.Y. Park, H. Nakamura, S. Adachi, H. Shoun, and Y. Shiro. Proton delivery in NO reduction by fungal nitric-oxide reductase. *J. Biol. Chem.*, 275(7):4816, 2000.

[71] O. Gotoh, Y. Tagashira, T. Iizuka, and Y. Fulii-Kuriyama. Structural characteristics of cytochrome P-450. Possible location of the heme-binding cysteine in determined amino-acid sequences. *J. Biochem.*, 93(3):807, 1983.

[72] S.Y. Park, H. Shimizu, S. Adachi, A. Nakagawa, I. Tanaka, K. Nakahara, H. Shoun, E. Obayashi, H. Nakamura, T. Lizuka, and Y. Shiro. Crystal structure of nitric oxide reductase from denitrifying fungus Fusarium oxysporum. *Nat. Struct. Mol. Biol.*, 4(10):827–832, 1997.

[73] N. Pollak, C. Dölle, and M. Ziegler. The power to reduce: pyridine nucleotides–small molecules with a multitude of functions. *Biochem. J.*, 402(2):205, 2007.

[74] R. Oshima, S. Fushinobu, F. Su, L. Zhang, N. Takaya, and H. Shoun. Structural evidence for direct hydride transfer from NADH to cytochrome P450nor. *J. Mol. Biol.*, 342(1):207–217, 2004.

[75] W.G. Zumft. Cell biology and molecular basis of denitrification. *Microbiol. Mol. Biol. Rev.*, 61(4):533–616, 1997.

[76] T. Kudo, N. Takaya, S. Y. Park, Y. Shiro, and H. Shoun. A positively charged cluster formed in the heme-distal pocket of cytochrome P450nor is essential for interaction with NADH. *J. Biol. Chem.*, 276(7):5020–5026, 2001.

[77] L. Zhang, T. Kudo, N. Takaya, and H. Shoun. The B'-helix determines cytochrome P450nor specificity for the electron donors NADH and NADPH. *J. Biol. Chem.*, 277(37):33842, 2002.

[78] A. Daiber, T. Nauser, N. Takaya, T. Kudo, P. Weber, C. Hultschig, H. Shoun, and V. Ullrich. Isotope effects and intermediates in the reduction of NO by P450nor. *J. Inorg. Biochem.*, 88(3-4):343–352, 2002.

[79] F. Su, S. Fushinobu, N. Takaya, and H. Shoun. Involvement of a Glu71–Arg64 Couple in the Access Channel for NADH in Cytochrome P450nor. *Biosci. Biotech. Bioch.*, 68(5):1156–1159, 2004.

[80] M. Umemura, F. Su, N. Takaya, Y. Shiro, and H. Shoun. D88A mutant of cytochrome P450nor provides kinetic evidence for direct complex formation with electron donor NADH. *Eur. J. Biochem.*, 271(14):2887–2894, 2004.

[81] N. Okamoto, Y. Imai, H. Shoun, and Y. Shiro. Site-directed mutagenesis of the conserved threonine (Thr243) of the distal helix of fungal cytochrome P450nor. *Biochemistry*, 37(25):8839–8847, 1998.

[82] N. Okamoto, K. Tsuruta, Y. Imai, D. Tomura, and H. Shoun. Fungal P450nor: expression in Escherichia coli and site-directed mutageneses at the putative distal region. *Arch. Biochem. Biophys.*, 337(2):338–344, 1997.

[83] E. Obayashi, H. Shimizu, S.Y. Park, H. Shoun, and Y. Shiro. Mutation effects of a conserved threonine (Thr243) of cytochrome P450nor on its structure and function. *J. Inorg. Biochem.*, 82(1-4):103–111, 2000.

[84] H. Shimizu, S. Park, D. Lee, H. Shoun, and Y. Shiro. Crystal structures of cytochrome P450nor and its mutants (Ser286→Val, Thr) in the ferric resting state at cryogenic temperature: a comparative analysis with monooxygenase cytochrome P450s. *J. Inorg. Biochem.*, 81(3):191–205, 2000.

[85] Y. Shiro, M. Fujii, Y. Isogai, S. Adachi, T. Iizuka, E. Obayashi, R. Makino, K. Nakahara, and H. Shoun. Iron-ligand structure and iron redox property of nitric oxide reductase cytochrome P450nor from Fusarium oxysporum: relevance to its NO reduction activity. *Biochemistry*, 34(28):9052–9058, 1995.

[86] E. Obayashi, K. Tsukamoto, S. Adachi, S. Takahashi, M. Nomura, T. Iizuka, H. Shoun, and Y. Shiro. Unique binding of nitric oxide to ferric nitric oxide reductase from Fusarium oxysporum elucidated with infrared, resonance Raman, and X-ray absorption spectroscopies. *J. Am. Chem. Soc.*, 119(33):7807–7816, 1997.

[87] U.P. Singh, E. Obayashi, S. Takahashi, T. Iizuka, H. Shoun, and Y. Shiro. The effects of heme modification on reactivity, ligand binding properties and iron-coordination structures of cytochrome P450nor. *Biochim. Biophys. Acta*, 1384(1):103–111, 1998.

[88] R. Silaghi-Dumitrescu. Nitric Oxide Reduction by Heme-Thiolate Enzymes (P450nor): A Reevaluation of the Mechanism. *Eur. J. Inorg. Chem.*, 2003(6):1048–1052, 2003.

[89] A. Daiber, H. Shoun, and V. Ullrich. Nitric oxide reductase (P450nor) from Fusarium oxysporum. *J. Inorg. Bioche.*, 99(1):185–193, 2005.

[90] F. Paulat and N. Lehnert. Electronic structure of ferric heme nitrosyl complexes with thiolate coordination. *Inorg. Chem.*, 46(5):1547–1549, 2007.

[91] N. Lehnert, V.K.K. Praneeth, and F. Paulat. Electronic structure of iron(II)-porphyrin nitroxyl complexes: molecular mechanism of fungal nitric oxide reductase (P450nor). *J. Comput. Chem.*, 27(12):1338–1351, 2006.

[92] E. Obayashi, S. Takahashi, and Y. Shiro. Electronic structure of reaction intermediate of cytochrome P450nor in its nitric oxide reduction. *J. Am. Chem. Soc.*, 120(49):12964–12965, 1998.

[93] J.H. Enemark and R.D. Feltham. Principles of structure, bonding, and reactivity for metal nitrosyl complexes. *Coord. Chem. Rev.*, 13(4):339–406, 1974.

[94] R.D. Feltham and J.H. Enemark. *Structures of Metal Nitrosyls.*, pages 155–215. John Wiley & Sons, Inc., 1981.

[95] B.L. Westcott and J.H. Enemark. *Transition metal nitrosyls.*, volume 2, pages 403–450. Wiley: New York, 1999.

[96] N. Xu, D.R. Powell, L. Cheng, and G.B. Richter-Addo. The first structurally characterized nitrosyl heme thiolate model complex. *Chem. Comm.*, 2006(19):2030–2032, 2006.

[97] D.L. Harris. Cytochrome P450nor: a nitric oxide reductase? Structure, spectra, and mechanism. *Int. J. Quantum. Chem.*, 88(1):183–200, 2002.

[98] M.A. Vincent, I.H. Hillier, and J. Ge. How is NN bond formation facilitated by P450 NO reductase? A DFT study. *Chem. Phys. Lett.*, 407(4-6):333–336, 2005.

[99] F. Neese. ORCA, an ab initio, density functional and semiempirical program package. *University of Bonn, Germany*, 2007.

[100] J. Baker. An algorithm for the location of transition states. *J. Comput. Chem.*, 7(4):385–395, 1986.

[101] C. Peng, P.Y. Ayala, H.B. Schlegel, and M.J. Frisch. Using redundant internal coordinates to optimize equilibrium geometries and transition states. *J. Comput. Chem.*, 17(1):49–56, 1996.

[102] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl. Gromacs 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.*, 4(3):435–447, 2008.

[103] H.P. Hratchian and H.B. Schlegel. Finding minima, transition states, and following reaction pathways on ab initio potential energy surfaces. *Theory and applications of computational chemistry: the first forty years*, page 195, 2005.

[104] J. Simons, P. Joergensen, H. Taylor, and J. Ozment. Walking on potential energy surfaces. *J. Phys. Chem.*, 87(15):2745–2753, 1983.

[105] A. Banerjee, N. Adams, J. Simons, and R. Shepard. Search for stationary points on surfaces. *J. Phys. Chem.*, 89(1):52–57, 1985.

[106] C.J. Cerjan and W.H. Miller. On finding transition states. *J. Chem. Phys.*, 75:2800, 1981.

[107] H.B. Schlegel. Exploring potential energy surfaces for chemical reactions: an overview of some practical methods. *J. Comp. Chem.*, 24(12):1514–1527, 2003.

[108] V. Bakken and T. Helgaker. The efficient optimization of molecular geometries using redundant internal coordinates. *J. Chem. Phys.*, 117:9160, 2002.

[109] B.H. Schlegel. Estimating the hessian for gradient-type geometry optimizations. *Theor. Chim. Acta.*, 66(5):333–340, 1984.

[110] T.H. Fischer and J. Almlöf. General methods for geometry and wave function optimization. *J. Phys. Chem.*, 96(24):9768–9774, 1992.

[111] R. Lindh, A. Bernhardsson, G. Karlström, and P. Malmqvist. On the use of a Hessian model function in molecular geometry optimizations. *Chem. Phys. Lett.*, 241(4):423–428, 1995.

[112] R. Fletcher. Unconstrained Optimization. In *Practical Methods of Optimization*, volume 1. Wiley, New York, 1980.

[113] J.M. Bofill. Updated Hessian matrix and the restricted step method for locating transition structures. *J. Comput. Chem.*, 15(1):1–11, 1994.

[114] R.A. Friesner and V. Guallar. Ab initio quantum chemical and mixed quantum mechanics/molecular mechanics (QM/MM) methods for studying enzymatic catalysis. *Annu. Rev. Phys. Chem.*, 56:389–427, 2005.

[115] D. Riccardi, P. Schaefer, Y. Yang, H. Yu, N. Ghosh, X. Prat-Resina, P. König, G. Li, D. Xu, H. Guo, et al. Development of effective quantum mechanical/molecular mechanical (QM/MM) methods for complex biological processes. *J. Phys. Chem. B*, 110(13):6458–6469, 2006.

[116] H. Lin and D.G. Truhlar. QM/MM: what have we learned, where are we, and where do we go from here?. *Theor. Chim. Acta*, 117(2):185–199, 2007.

[117] H. Senn and W. Thiel. QM/MM methods for biological systems. *Top. Curr. Chem.*, pages 173–290, 2007.

[118] B. Wang and D.G. Truhlar. Including Charge Penetration Effects in Molecular Modeling. *J. Chem. Theory Comput.*, 6(11):3330, 2010.

[119] H. Lin and D.G. Truhlar. Redistributed charge and dipole schemes for combined quantum mechanical and molecular mechanical calculations. *J. Phys. Chem. A*, 109(17):3991–4004, 2005.

[120] R.A. Laskowski, M.W. MacArthur, D.S. Moss, and J.M. Thornton. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.*, 26(2):283–291, 1993.

[121] A.W. Schuttelkopf and D.M.F. van Aalten. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr. D*, 60(8):1355–1363, 2004.

[122] C.M. Breneman and K.B. Wiberg. Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *J. Comput. Chem.*, 11(3):361–373, 1990.

[123] 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. Phys. Chem. B*, 103(29):6133–6141, 1999.

[124] J. Zheng, A. Altun, and W. Thiel. Common system setup for the entire catalytic cycle of cytochrome P450cam in quantum mechanical/molecular mechanical studies. *J. Comput. Chem.*, 28(13):2147–2158, 2007.

[125] H. Li and J.H. Jensen. Partial Hessian vibrational analysis: the localization of the molecular vibrational energy and entropy. *Theor. Chim. Acta*, 107(4):211–219, 2002.

[126] R. McWeeny and Y. Mizuno. The density matrix in many electron quantum mechanics II. Separation of space and spin variables; spin coupling problems. *Proc. Roy. Soc. (London)*, A259:554, 1961.

[127] A. Bencini and D. Gatteschi. *EPR of Exchange Coupled Systems.* Springer, Heidelberg, 1990.

[128] O. Kahn. *Molecular Magnetism.* VCH Publishers, New York, 1993.

[129] P. Bertrand, P. Camensuli, C. More, and B. Guigliarelli. A local spin model to describe the magnetic interactions in biological molecules containing [4Fe-4S]$^{+}$ clusters. Applications to Ni-Fe hydrogenases. *J. Am. Chem. Soc.*, 118:1426 – 1434, 1996.

[130] E.D. Glendening, J.K. Badenhoop, A.E. Reed, J.E. Carpenter, J.A. Bohmann, C.M. Morales, and F. Weinhold. NBO., 2001.

[131] G.B. Wells and M.W. Makinen. ENDOR-determined molecular geometries of spin-labeled fluoroanilides in frozen solution. *J. Am. Chem. Soc.*, 110(19):6343–6352, 1988.

[132] H. Li, A.D. Robertson, and J.H. Jensen. Very fast empirical prediction and interpretation of protein pKa values. *Proteins*, 61:704–721, 2005.

[133] D.C. Bas, D.M. Rogers, and J.H. Jensen. Very fast prediction and rationalization of pKa values for protein–ligand complexes. *Proteins*, 73(3):765–783, 2008.

[134] W.L. Jorgensen and J. Tirado-Rives. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.*, 110(6):1657–1666, 1988.

[135] G.A. Kaminski, R.A. Friesner, J. Tirado-Rives, and W.L. Jorgensen. Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J. Phys. Chem. B*, 105(28):6474–6487, 2001.

[136] M. Radoul, M. Sundararajan, A. Potapov, C. Riplinger, F. Neese, and D. Goldfarb. Revisiting the nitrosyl complex of myoglobin by high-field pulse EPR spectroscopy and quantum mechanical calculations. *Phys. Chem. Chem. Phys.*, 12(26):7276–7289, 2010.

[137] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A*, 38(6):3098–3100, 1988.

[138] J.P. Perdew. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B*, 33(12):8822–8824, 1986.

[139] J.P. Perdew. Erratum: Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B*, 34(10):7406, 1986.

[140] A. Schäfer, H. Horn, and R. Ahlrichs. Fully optimized contracted Gaussian basis sets for atoms Li to Kr. *J. Chem. Phys.*, 97:2571, 1992.

[141] K. Eichkorn, O. Treutler, H. Öhm, M. Häser, and R. Ahlrichs. Auxiliary basis sets to approximate Coulomb potentials. *Chem. Phys. Lett.*, 240(4):283 – 289, 1995.

[142] K. Eichkorn, O. Treutler, H. Öhm, M. Haeser, and R. Ahlrichs. Auxiliary basis sets to approximate Coulomb potentials.[Erratum to document cited in CA123: 93649]. *Chem. Phys. Lett.*, 242:652–660, 1995.

[143] D. Bakowies and W. Thiel. Hybrid models for combined quantum mechanical and molecular mechanical approaches. *J. Phys. Chem.*, 100(25):10580–10594, 1996.

[144] A. Klamt and G. Schüürmann. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc. Perk. T. 2*, 1993(5):799–805, 1993.

[145] P.E.M. Siegbahn, M.R.A. Blomberg, and M. Pavlov. A comparison of electron transfer in ribonucleotide reductase and the bacterial photosynthetic reaction center. *Chem. Phys. Lett.*, 292(4-6):421–430, 1998.

[146] F. Neese, F. Wennmohs, A. Hansen, and U. Becker. Efficient, approximate and parallel Hartree-Fock and hybrid DFT calculations. A 'chain-of-spheres' algorithm for the Hartree-Fock exchange. *Chem. Phys.*, 356(1-3):98–109, 2009.

[147] 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(2):785–789, 1988.

[148] A.D. Becke. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.*, 98(1):5648–5652, 1993.

[149] F. Weigend and R. Ahlrichs. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: design and assessment of accuracy. *Phys. Chem. Chem. Phys.*, 7(18):3297–3305, 2005.

[150] D.A. Pantazis, X.Y. Chen, C.R. Landis, and F. Neese. All-Electron Scalar Relativistic Basis Sets for Third-Row Transition Metal Atoms. *J. Chem. Theory Comput.*, 4(6):908–919, 2008.

[151] A. Schäfer, C. Huber, and R. Ahlrichs. Fully Optimized Contracted Gaussian-Basis Sets of Triple Zeta Valence Quality for Atoms Li to Kr. *J. Chem. Phys.*, 100(8):5829–5835, 1994.

[152] M. Reiher, O. Salomon, and B. A. Hess. Reparameterization of hybrid functionals based on energy differences of states of different multiplicity. *Theor. Chim. Acta.*, 107:48–55, 2001.

[153] 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(10):4729–4737, 2002.

[154] F. Neese. Prediction and interpretation of the 57Fe isomer shift in Mössbauer spectra by density functional theory. *Inorg. Chim. Acta*, 337:181–192, 2002.

[155] E. van Lenthe, E.J. Baerends, and J.G. Snijders. Relativistic regular two-component Hamiltonians. *J. Chem. Phys.*, 99(6):4597–4610, 1993.

[156] E. van Lenthe, E.J. Baerends, and J.G. Snijders. Relativistic total energy using regular approximations. *J. Chem. Phys.*, 101(11):9783, 1994.

[157] E. van Lenthe, R. van Leeuwen, E.J. Baerends, and J.G. Snijders. Relativistic regular two-component Hamiltonians. *Int. J. Quantum. Chem.*, 57(3):281–293, 1996.

[158] S. Grimme. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.*, 27(15):1787–1799, 2006.

[159] R.V. Stanton, D.S. Hartsough, and K.M. Merz Jr. An examination of a density functional/molecular mechanical coupled potential. *J. Comput. Chem.*, 16(1):113–128, 1995.

[160] J. Pipek and P.G. Mezey. A fast intrinsic localization procedure applicable for ab initio and semiempirical linear combination of atomic orbital wave functions. *J. Chem. Phys.*, 90:4916, 1989.

[161] L.E. Goodrich, F. Paulat, V.K.K. Praneeth, and N. Lehnert. Electronic Structure of Heme-Nitrosyls and Its Significance for Nitric Oxide Reactivity, Sensing, Transport, and Toxicity in Biological Systems. *Inorg. Chem.*, 49(14):6293–6316, 2010.

[162] J. Olah and J.N. Harvey. NO Bonding to Heme Groups: DFT and Correlated ab Initio Calculations. *J. Phys. Chem. A*, 113(26):7338–7345, 2009.

[163] D.P. Linder and K.R. Rodgers. Fe-NO structure and bonding in six-coordinate $\{FeNO\}^6$ porphyrinates containing imidazole: implications for reactivity of coordinated NO. *Inorg. Chem.*, 44(5):1367–1380, 2005.

[164] W.R. Scheidt, A. Barabanschikov, J.W. Pavlik, N.J. Silvernail, and J.T. Sage. Electronic Structure and Dynamics of Nitrosyl Porphyrins. *Inorg. Chem.*, 49(14):6240–6252, 2010.

[165] V.K.K. Praneeth, F. Paulat, T.C. Berto, S.D.B. George, C. Näther, C.D. Sulok, and N. Lehnert. Electronic Structure of Six-Coordinate Iron (III)- Porphyrin NO Adducts: The Elusive Iron (III)- NO (radical) State and Its Influence on the Properties of These Complexes. *J. Am. Chem. Soc.*, 130(46):15288–15303, 2008.

[166] F. Neese. Definition of corresponding orbitals and the diradical character in broken symmetry DFT calculations on spin coupled systems. *J. Phys. Chem. Solids*, 65(4):781 – 785, 2004. Design, Characterization and Modelling of Molecule-Based Magnetic Materials Proceedings of Symposium K, EMRS Spring Meeting, June 2003, Strasbourg, France.

[167] R.G. Serres, C.A. Grapperhaus, E. Bothe, E. Bill, T. Weyhermüller, F. Neese, and K. Wieghardt. Structural, Spectroscopic, and Computational Study of an Octahedral, Non-Heme $\{Fe\text{- }NO\}^{6-8}$ Series:[Fe (NO)(cyclam-ac)]$^{2+/+/0}$. *J. Am. Chem. Soc.*, 126(16):5138–5153, 2004.

[168] S.C. Bart, K. Chlopek, E. Bill, M.W. Bouwkamp, E. Lobkovsky, F. Neese, K. Wieghardt, and P.J. Chirik. Electronic Structure of Bis(imino)pyridine Iron Dichloride, Monochloride, and Neutral Ligand Complexes: A Combined Structural, Spectroscopic, and Computational Study. *J. Am. Chem. Soc.*, 128(42):13901–13912, 2006. PMID: 17044718.

[169] A.P. Scott and L. Radom. Harmonic vibrational frequencies: an evaluation of Hartree-Fock, Moller- Plesset, quadratic configuration interaction, density functional theory, and semiempirical scale factors. *J. Phys. Chem.*, 100(41):16502–16513, 1996.

[170] D. Wachsstock. Tenua, a chemical kinetics simulation program. Available free online at: `http://bililite.com/tenua`.

[171] B.A. Barshop, R.F. Wrenn, and C. Frieden. Analysis of numerical methods for computer simulation of kinetic processes: development of KINSIM–a flexible, portable system. *Anal. Biochem.*, 130(1):134–145, 1983.

[172] C. Li, W. Wu, D. Kumar, and S. Shaik. Kinetic Isotope Effect is a Sensitive Probe of Spin State Reactivity in C- H Hydroxylation of N, N-Dimethylaniline by Cytochrome P450. *J. Am. Chem. Soc.*, 128(2):394–395, 2006.

[173] M.J. Knapp and J.P. Klinman. Environmentally coupled hydrogen tunneling. *Eur. J. Biochem.*, 269(13):3113–3121, 2002.

[174] R.G. Larsen and D.J. Singel. Double electron-electron resonance spin-echo modulation: spectroscopic measurement of electron spin pair separations in orientationally disordered solids. *J. Chem. Phys.*, 98:5134–5146, 1993.

[175] M. Pannier, S. Veit, A. Godt, G. Jeschke, and H.W. Spiess. Dead-time Free Measurement of Dipole-Dipole Interactions between Electron Spins. *J. Magn. Reson. A*, 142:331–340, 2000.

[176] C. Riplinger, J.P.Y. Kao, G.M. Rosen, V. Kathirvelu, G.R. Eaton, S.S. Eaton, A. Kutateladze, and F. Neese. Interaction of Radical Pairs Through-Bond and Through-Space: Scope and Limitations of the Point- Dipole Approximation in Electron Paramagnetic Resonance Spectroscopy. *J. Am. Chem. Soc.*, 131(29):10092–10106, 2009.

[177] F. Neese. ORCA - an ab initio, Density Functional and Semiempirical Program Package., 2007.

[178] V. Barone. EPR-II. In D.P. Chong, editor, *Recent Advances in Density Functional Methods, Part I.* World Scientific, Singapore, 1996.

[179] J. Pipek and P.G. Mezey. A fast intrinsic localization procedure applicable for ab initio and semiempirical linear combination of atomic orbital wave-functions . *J. Chem. Phys.*, 90:4916–4926, 1989.

[180] N. Rega, M. Cossi, and V. Barone. Development and validation of reliable quantum mechanical approaches for the study of free radicals in solution. *J. Chem. Phys.*, 105(24):11060–11067, 1996.

[181] 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, 1988.

[182] A. D. Becke. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.*, 98(7):5648–5652, 1993.

[183] F. Neese. Prediction of molecular properties and molecular spectroscopy with density functional theory: from fundamental theory to exchange-coupling. *Coord. Chem. Rev.*, 253:526–563, 2008.

[184] K. Yamaguchi, Y. Takahara, and T. Fueno. Yamaguchi formula. In V.H. Smith, editor, *Applied Quantum Chemistry, V*, page 155. Reidel, Dordrecht, 1986.

[185] T. Soda, Y. Kitagawa, T. Onishi, Y. Takano, Y. Shigeta, H. Nagao, Y. Yoshika, and K. Yamaguchi. Yamaguchi formula. *Chem Phys. Lett.*, 319:223, 2000.

[186] Y. Pontillon, A. Caneschi, D. Gatteschi, A. Grand, E. Ressouche, R. Sessoli, and J. Schweizer. Experimental Spin Density in a Purely Organic Free Radical: Visualisation of the Ferromagnetic Exchange Pathway in p-(Methylthio)phenyl Nitronyl Nitroxide, Nit(SMe)Ph. *Chem. Eur. J.*, 5(12):3616–3624, 1999.

[187] U. Schatzschneider and E. Rentschler. Electronic and conformational influences on the spin density distribution in substituted tert-butyl phenyl nitroxides. *J. Mol. Struct. Theochem*, 638:163–168, 2003.

[188] M. Yao, H. Inoue, and N. Yoshioka. Novel aromatic N-oxyl radical based on the benzo[g]quinoline skeleton: synthesis and intermolecular ferromagnetic interaction. *Chem. Phys. Lett.*, 402:11–16, 2005.

[189] C. Altenbach, K.-J. Oh, R. J. Trabanino, K. Hideg, and W.L. Hubbell. Estimation of Inter-Residue Distances in Spin Labeled Proteins at Physiological Temperatures: Experimental Strategies and Practical Limitations. *Biochemistry*, 40:15471 – 15482, 2001.

[190] G. Jeschke and H.W. Spiess. Distance Measurements in Solid-State NMR and EPR Spectroscopy. *Lect. Notes Phys.*, 684:21–63, 2006.

[191] G.R. Eaton and S.S. Eaton. EPR studies of long-range intramolecular electron-electron exchange interaction. *Accts. Chem. Res.*, 21(3):107–13, 1988.

[192] R.E. Coffman and G.R. Buettner. A limit function for long-range ferromagnetic and antiferromagnetic superexchange. *J. Phys. Chem.*, 83:2387–2392, 1979.

[193] G.I. Likhtenshtein. Depth of Immersion of Paramagnetic Centers in Biological Systems. In J.B. Lawrence, editor, *Distance measurements in biological systems by EPR*, volume 19 of *Biol. Mag. Res.* Kluwer Academic, 2000.

[194] C.C. Moser, C.C. Page, R. Farid, and P.L. Dutton. Biological Electron Transfer. *J. Bioenerg. Biomembr.*, 27:263–274, 1995.

[195] D.R. Gamelin, E.L. Bominaar, M.L. Kirk, K. Wieghardt, and E.I. Solomon. Excited-State Contributions to Ground-State Properties of Mixed-Valence Dimers: Spectral and Electronic-Structural Studies of [Fe(OH)(tmtacn)] Related to the [FeS] Active Sites of Plant-Type Ferredoxins. *J. Am. Chem. Soc.*, 118:8085–8097, 1996.

[196] E.L. Bominaar, C. Achim, S.A Borshch, J.-J. Girerd, and E. Münck. Analysis of Exchange Interaction and Electron Delocalization as Intramolecular Determinants of Intermolecular Electron-Transfer Kinetics. *Inorg. Chem.*, 36(17):3689–3701, 1997.