Selected Aspects of Enzymatic Catalytic Activity Studied by Theoretical Methods

and

Implementation of the Analytic Second Derivatives of Hartree-Fock and Hybrid Density Functional Energies

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Dmytro Bykov

aus

Kiew

Bonn 2013

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

1. Gutachter: Prof. Dr. Frank Neese

2. Gutachter: Prof. Dr. Stefan Grimme

Tag der Promotion: 20.12.2013

Erscheinungsjahr: 2014

IN DER DISSERTATION EINGEBUNDEN:

Zusammenfassung

Моїй родині

ABSTRACT

This thesis deals with application of modern theoretical methods for studying enzymatic reactivity as well as with the efficient implementation of second analytical energy derivatives on the self-consistent field (SCF) theory level. The enzymatic mechanisms are investigated in the framework of electronic structure theory with special accent put on kinetics, proton-coupled electron transfer and theoretical support of EPR and Mössbauer experiments. The theory part of the thesis includes conventional implementation of SCF second derivatives into ORCA set of programs combined with efficient approximations like "resolution of identity" and chain of spheres. The results of present work are divided into six chapters 3.1-3.6, each dealing with different aspects of the aforementioned subjects. In the following the content of each chapters will be briefly outlined.

The chapters 3.1-3.3 are dedicated to cytochrome *c* nitrite reductase (C*c*NiR) enzyme mechanism. C*c*NiR is a homodimeric enzyme, containing five covalently attached *c*-type hemes per subunit. Four of the heme-irons are bis-histidine ligated, while the fifth, the active site of the protein, has an unusual lysine coordination and calcium site nearby. A fascinating feature of this enzyme is that the full six-electron reduction of the nitrite to ammonia is achieved without release of any detectable reaction intermediate.

In chapter 3.1 the possible role of second-sphere active site amino acids as proton donors is investigated by taking different possible protonation states and geometrical conformations into account. It was found that the most probable proton donor is His_{277}, whose spatial orientation and fine-tuned acidity lead to energetically feasible, low-barrier protonation reactions. However, substrate protonation may also be accomplished by Arg_{114}. The calculated barriers for the Arg_{114} pathway are only slightly higher than the experimentally determined value of 15.2 kcal/mol for the rate-limiting step.

Chapter 3.2 presents the study of the active site reactivation with protons and electrons modeled by the series of reaction intermediates based on nitrogen monoxide (Fe(II)-NO^+, Fe(II)-$NO^\bullet$, Fe(II)-NO^- and Fe(II)-HNO). The activation barriers for the various proton and electron transfer steps were estimated in the framework of Marcus theory. Using the obtained barriers, the kinetics of the reduction process was simulated. It was found that the complex reactivation process can be accomplished in two possible ways: either through two consecutive proton coupled electron transfers (PCET) or in the form of three consecutive elementary steps involving reduction, PCET and protonation. Kinetic simulations revealed the recharging through two PCETs to be a means of overcoming the predicted deep energetic minimum that is calculated to occur at the stage of the Fe(II)-$NO^\bullet$ intermediate.

In chapter 3.3 the second half cycle of the nitrite reduction catalyzed by C*c*NiR was considered in details. In total 3 electrons and 4 protons must be provided to reach the final product ammonia starting form HNO intermediate. The first event in this half cycle is the reduction of the HNO intermediate accomplished by PCET reaction. The isomeric intermediates $HNOH^\bullet$ and $H_2NO^\bullet$ are formed. Both intermediates are active and are readily transformed into hydroxylamine most likely through intramolecular proton transfer either from Arg_{114} or His_{277}. The protonated side chain then

provides its proton to initiate a heterolytic cleavage reaction of the N-O bond. As a result the H_2N^+ intermediate is formed. The latter readily picks up an electron forming $H_2N^{+\bullet}$, which in turn reacts with Tyr_{218}. Intramolecular reaction with Tyr_{218} in the final step of the nitrite reduction process leads directly to the ammonia final product. The product dissociation was found to proceed through the change of spin state, which was also observed in resonance Raman investigation of Martins (Martins, G., et. al. (2010), *J Phys Chem B 114*, 5563)

Chapter 3.4 is concerned with cd_1 nitrite reductase (NIR) enzyme. NIR is a key enzyme in the denitrification process that reduces nitrite to nitric oxide (NO). There are three residues at the "distal" side of the active site heme (Tyr_{10}, His_{327} and His_{369}) and in this work the focus was set on the identification and characterization of possible H-bonds they can form with the NO, thereby affecting the stability of the complex. It was shown that the NO in the nitrosyl d_1-heme complex of cd_1 NIR forms H-bonds with Tyr_{10} and His_{369} whereas the second conserved histidine, His_{327}, appears to be less involved in NO H-bonding. Moreover, it was shown that the H-bonding network within the active site is dynamic and that a change in the protonation state of one of the residues does affect the strength and position of the H-bonds formed by the others. In the Y10F mutant His_{369} is closer to the NO, whereas mutation of both distal histidines displaces Tyr_{10} removing its H-bond. The implications of the H-bonding network found in terms of the complex stability and catalysis are discussed.

The electronic structure of the [4Fe-3S] cluster in Hydrogenase I (Hase I) is discussed in chapter 3.5. The cluster performs two redox transitions within a very small potential range, forming a super-oxidized state above +200 mV *vs* SHE. Crystallographic data has revealed that this state is stabilized by the coordination of one of the iron atoms to a backbone nitrogen. Thus, the proximal [4Fe-3S] cluster undergoes redox-dependent structural changes to serve multiple purposes beyond classical electron transfer. The field-dependent ^{57}Fe-Mössbauer and EPR data for Hase I is presented, which in conjunction with spectroscopically calibrated DFT calculations reveal the distribution of Fe valences and spin-coupling schemes for the iron-sulfur clusters. The data demonstrate that the electronic structure of the [4Fe-3S] core in its three oxidation states closely resembles that of corresponding conventional [4Fe-4S] cubanes, albeit with distinct differences for some individual iron sites.

The implementation of the SCF energy second derivatives is discussed in chapter 3.6. The second derivatives of electronic energy are the base for the calculation of force constants, harmonic vibration frequencies, infra red (IR) and Raman intensities. To speed up the evaluation of the hessian, in particular the two-electron integrals and their derivatives, the resolution of the identity (RI) and the Chain of Spheres (COS) approximations can be applied. As part of the present work, the RI and COS approximations are introduced at various stages of the molecular Hessian evaluation procedure, e.g., the reference energy calculation, various steps of the coupled-perturbed SCF procedure, and the final integral derivative evaluation. The performance of the approximations and possible errors introduced are discussed in details. The applicability of the Hessian program was also greatly extended by the additional functionality such as effective core potentials (ECP), Van der Waals corrected second derivatives and QM/MM hessian.

ZUSAMMENFASSUNG

Diese Dissertation umfasst sowohl die Anwendung moderner, computergestützter Methoden zur Untersuchung von Enzym-Reaktivitäten als auch eine effiziente Implementierung der analytischen zweiten Ableitungen der SCF-Energie (SCF vom engl. Self Consistent Field). Die enzymatischen Mechanismen werden mit Hilfe der Elektronenstrukturtheorie untersucht, wobei der Schwerpunkt auf der Ermittlung kinetischer Parameter, dem protonengekoppelten Elektronentransfer und der theoretischen Unterstützung von EPR- und Mössbauer-Experimenten liegt. Der theoretische Teil der Arbeit beinhaltet die konventionelle Implementierung der zweiten SCF-Ableitungen unter Berücksichtigung von Näherungen wie „Resolution of the Identity“ und „Chain of Spheres“ in das Programmpaket ORCA. Die Ergebnisse der vorliegenden Arbeit sind in sechs Kapitel unterteilt (3.1–3.6), von welchen jedes einzelne unterschiedliche Aspekte der oben genannten Themen betrachtet. Im Folgenden wird jedes Kapitel kurz zusammengefasst.

Die Kapitel 3.1–3.3 widmen sich dem Mechanismus des Cytochrome *c* Nitrite Reductase (C*c*NiR) Enzyms. C*c*NiR ist ein homodimeres Enzym, das fünf kovalent verknüpfte Häm *c* pro Untereinheit enthält. Vier der Häme sind *bis*-Histidin-koordiniert, während das fünfte, welches das aktive Zentrum des Proteins ist, eine ungewöhnliche Lysin-Koordination aufweist und zudem ein Calcium-Ion in seiner unmittelbaren Umgebung hat. Das besondere Merkmal dieses Enzyms ist, dass die vollständige sechs-Elektronen-Reduktion des Nitrits zu Ammoniak ohne ein detektierbares Intermediat verläuft.

In Kapitel 3.1 wird die Fähigkeit der Aminosäuren aus der zweiten Koordinationssphäre des aktiven Zentrums hinsichtlich möglicher Protonierungen des Edukts untersucht, indem verschiedene Protonierungsmuster und geometrische Konformationen berücksichtigt werden. Die Aminosäure His_{277} wird als vielversprechender Protonendonor identifiziert, da ihre räumliche Orientierung und feinjustierte Azidität eine energetisch günstige Protonierungsreaktionen mit niedriger Energiebarriere ermöglichen. Ein weiterer Kandidat für diese Protonierungsreaktion ist die Aminosäure Arg_{114}. Die berechnete Barriere für den geschwindigkeitsbestimmenden Schritt ist für die Reaktion mit Arg_{114} nur geringfügig höher als der experimentell bestimmte Wert von 15,2 kcal/mol.

Kapitel 3.2 beschreibt die Untersuchung der Reaktivierung des aktiven Zentrums mit Protonen und Elektronen. Die Reaktivierung wird mit einer Reihe von Intermediaten auf der Basis von Stickstoffmonoxid modelliert (Fe(II)-NO^+, Fe(II)-$NO^\bullet$, Fe(II)-NO^- and Fe(II)-HNO). Die Aktivierungsbarrieren für die verschiedenen Proton- und Elektrontransferschritte werden im Rahmen der Marcus-Theorie bestimmt. Der komplexe Reaktivierungsprozess kann prinzipiell auf zwei möglichen Wegen stattfinden: entweder durch zwei aufeinanderfolgende protonengekoppelte Elektronentransfer-Schritte (engl. Proton Coupled Electron Transfer, PCET), oder durch drei aufeinanderfolgende Elementarreaktionen, die Reduktion, PCET und Protonierung beinhalten. Kinetische Simulationen zeigen, dass der Reaktivierungsmechanismus über zwei PCET-Schritte eine Möglichkeit ist, das tiefe energetische Minimum, in dem das Fe(II)-$NO^\bullet$ Intermediat liegt, zu durchschreiten.

Im Kapitel 3.3 wird der zweite Teil des von C*c*NiR katalysierten Nitrit-Reduktionszyklus im Detail untersucht. Ausgehend vom HNO-Intermediat müssen insgesamt drei Elektronen und vier Protonen zur Verfügung gestellt werden um das Endprodukt Ammoniak zu erreichen. Der erste Schritt in diesem Halbzyklus ist die Reduktion des HNO-Intermediats durch PCET-Reaktion. Dabei werden die isomeren Intermediate $HNOH^{\bullet}$ und $H_2NO^{\bullet}$ gebildet. Beide Intermediate sind aktiv und werden ohne weiteres in Hydroxylamin überführt, wobei der intramolekulare Protonentransfer von Arg_{114} oder His_{277} am günstigsten ist. Die protonierte Seitenkette stellt das Proton zur Verfügung, wodurch die heterolytische N-O-Bindungsspaltung initiiert wird. So wird ein H_2N^+-Intermediat gebildet, welches ohne weiteres ein Elektron aufnehmen kann, sodass H_2N^+* resultiert, welches wiederum mit Tyr_{218} reagiert. Im letzten Schritt des Nitrit-Reduktionsprozesses führt die intramolekulare Reaktion mit Tyr_{218} direkt zu dem Endprodukt Ammoniak. Die Dissoziation des Produkts verläuft unter Änderung des Spinzustands, was auch durch Resonanz-Raman Spektroskopie belegt wurde (Martins, G., et. al. (2010), *J Phys Chem B 114*, 5563).

Kapitel 3.4 befasst sich mit dem *cd1* Nitrit-Reduktase (NIR) Enzym. NIR ist das Schlüsselenzym im Denitrifizierungsprozess, in dem Nitrit zu Stickstoffmonoxid (NO) reduziert wird. Die drei Aminosäuren Tyr_{10}, His_{327} und His_{369} befinden sich auf einer Seite des Häms im aktiven Zentrum. In diesem Kapitel liegt das Hauptaugenmerk auf der Identifizierung und Charakterisierung möglicher Wasserstoffbrückenbindungen, die diese Aminosäuren ausbilden können, da dies die Stabilität des Komplexes beeinflusst. Das NO im Nitrosyl Häm *d1*-Komplex bildet Wasserstoffbrückenbindungen mit Tyr_{10} und His_{369}, während das zweite konservierte Histidin, His_{327}, eine geringere Neigung zur Bildung von Wasserstoffbrückenbindungen zeigt. Zudem wird gezeigt, dass das Wasserstoffbrückennetzwerk im aktiven Zentrum dynamisch ist und die Änderung des Protonierungszustands eines der Liganden die Stärke und Position der anderen Wasserstoffbrückenbindungen beeinflusst. In der Y10F Mutante hat His_{369} einen geringeren Abstand zu NO, während die Mutation der beiden entfernten Histidin-Seitenketten Tyr_{10} verschiebt, sodass eine Wasserstoffbrückenbindung zwischen Tyr_{10} und Edukt nicht mehr möglich ist. Der Einfluss des Wasserstoffbrückennetzwerks wird mit Hinblick auf die Stabilität des Komplexes und den Katalysezyklus diskutiert.

Die Elektronenstruktur des [4Fe-3S]-Clusters in Hydrogenase I (Hase I) wird in Kapitel 3.5 diskutiert. Der Cluster zeigt zwei Redox-Übergänge innerhalb eines sehr kleinen Potentialbereichs und bildet so einen super-oxidierten Zustand über +200 mV vs. SHE. Eine Kristallstruktur hat gezeigt, dass dieser Zustand durch die Koordination eines Rückgrat-Stickstoffatoms an ein Eisenatom stabilisiert wird. Somit durchläuft der [4Fe-3S]-Cluster redox-abhängige Strukturänderungen, die außer dem Transfer von Elektronen weitere Ziele erfüllen. Feldabhängige ^{57}Fe-Mössbauer- und EPR-Daten für Hase I werden präsentiert, die in Verbindung mit kalibrierten DFT-Berechnungen die Verteilung der Fe-Valenzen und die Erstellung von Spinkopplungs-Schemata in den Eisen-Schwefel-Clusters mit verschiedenen Oxidationszuständen ermöglichen. Die Ergebnisse verdeutlichen, dass die Elektronenstruktur des [4Fe-3S]-Clusters in allen drei Oxidationszuständen den konventionellen [4Fe-4S]-Kubanen stark ähnelt, obwohl spezifische Unterschiede für einzelne Eisenzentren vorliegen.

Die Implementierung der zweiten Ableitungen der SCF-Energie wird im Kapitel 3.6 diskutiert. Die zweiten Ableitungen der elektronischen Energie stellen die Grundlage für die Berechnung von Kraftkonstanten, harmonischen Schwingungsfrequenzen und Intensitäten in Infrarot- und Raman-

Spektren dar. Zur Beschleunigung der Hessian-Evaluierung, insbesondere der Zwei-Elektronen-Integrale und deren Ableitungen, werden die Näherungen “Resolution of the Identity” (RI) und “Chain of Spheres” (COS) angewandt. Als Teil der vorliegenden Arbeit werden diese Näherungen an unterschiedlichen Stellen der Hessian-Evaluierung eingeführt, z.B. bei der Berechnung der Referenzenergie, in der abschließenden Evaluierung der Integral-Ableitungen und in zahlreichen Schritten der CPSCF Prozedur (vom engl. Coupled Perturbed SCF). Die Leistung der Näherungen und möglicherweise durch sie erzeugte Abweichungen werden ausführlich diskutiert. Die Anwendbarkeit des Hessian-Programms wird durch zusätzliche Funktionalitäten erweitert, wie beispielsweise Pseudopotentiale, van-der-Waals-korrigierte zweite Ableitungen sowie QM/MM Hessians.

LIST OF PUBLICATIONS

Publications related to the thesis:

Chapter 3.1:

Bykov, D.; Neese, F. *Substrate binding and activation in the active site of cytochrome c nitrite reductase: a density functional study.* J. Biol. Inorg. Chem. (2011) 16:417-430.

(Contribution by DB: All calculations of the electronic structures, data analysis and paper writing)

Chapter 3.2:

Bykov, D., Neese, F. *Reductive activation of the heme-iron-nitrosyl intermediate in the reaction mechanism of Cytochrome c Nitrite Reductase. A theoretical study.* J. Biol. Inorg. Chem. (2012), 17(5):741-60.

(Contribution by DB: All calculations of the electronic structures, data analysis and paper writing)

Chapter 3.3:

Bykov, D., Plog, M., Neese, F. *The nitroxyl, hydroxylamine and ammonia intermediates in the reaction cycle of cytochrome c nitrite reductase. A theoretical study.* (2013), accepted to JBIC.

(Contribution by DB: All calculations of the electronic structures, data analysis and paper writing)

Chapter 3.4:

Radoul M., **Bykov, D.,** Rinaldo, S., Cutruzzola, F., Neese, F., Goldfarb, D. *Dynamic hydrogen-bonding network in the distal pocket of the nitrosyl complex of pseudomonas aeruginosa cd1 nitrite reductase.* J. Am. Chem. Soc. (2011) 133(9):3043-3055.

(Contribution by DB: Support of the experimental studies by the electronic structure calculations, analysis of the calculated data, paper writing)

Chapter 3.5:

Pandelia, M-E., **Bykov, D.,** Izsak, R., Infossi, P., Giudici-Orticoni, M-T., Bill, E., Neese, F., Lubitz, W. *Reply to Mouesca et al.: Electronic structure of the proximal [4Fe-3S] cluster of O_2-tolerant [NiFe] hydrogenases.* PNAS (2013), 110(28):2539.

Pandelia, M-E., **Bykov, D.,** Izsak, R., Infossi, P., Giudici-Orticoni, M-T., Bill, E., Neese, F., Lubitz, W. *The electronic structure of the unique [4Fe-3S] cluster in O_2-tolerant hydrogenases characterized by ^{57}Fe Mössbauer and EPR spectroscopy.* PNAS (2013), 110(2):483-488.

(Contribution by DB: Support of the experimental studies by the electronic structure calculations, analysis of the calculated data, paper writing)

Chapter 3.6:

Bykov, D., Petrenko, T., Izsak, R., Kossmann, S., Becker U., Valeev, E., Neese, F. *Analytic second derivatives of Hartree-Fock and hybrid DFT energies. A detailed analysis of different approximations.* (2013), to be submitted to JCP.

(Contribution by DB: Leading programmer of the project, running test calculations, data analysis and paper writing)

Publications by the Author not related to the thesis:

Bykov, D. A., Zaderko, A. N., Datsyuk, A. M., Diyuk, V. E., Yatsimirskii, V. K., Lobanov, V. V. *Quantum-chemical modelling of the reactivity of charcoal surface double bonds.* Theoretical and Experimental Chemistry (2008), 44(1), 32-36.

Chekhovskii, A., Tomila, T., Ragulya, A., Timofeeva, I., Ivanchuk, A., **Bykov, D.**, Labunets, T. *Kinetics of CxNy formation on electrode surface through electrochemical method.* Science of Sintering (2007), 39(3), 287-294.

ACKNOWLEDGEMENTS

It is my great pleasure to thank those who contributed to my PhD thesis over the years in Bonn and in Mülheim. There were various scientists from whom I had the opportunity and pleasure to learn from. They all guided me to some extent through my PhD time:

First and most important, I want to express my deep gratitude to my scientific supervisor Prof. Frank Neese, who gave me the opportunity to work in the fields of bioinorganic and theoretical chemistry. Prof. Neese taught me not only the basics of theoretical chemistry but also showed his attitude and passion to work, became great example of scientific excellence for me. He gave me freedom and time to work independently and evolve my skills in the subject. Under his guidance I was working on a variety of projects having great start and support from his site during initial steps of my scientific career.

I also would like to thank Dr. Taras Petrenko and Dr. Robert Izsak for many scientific discussions. It was always a pleasure to get explanations with examples from Taras. His deep understanding of the subject helped me to see the connections between physical picture and mathematical formulations. Robert was always ready to discuss details on the integral evaluations and mathematics related questions. We spent a great deal of time reading and discussing the "great pink book".

I had a nice and productive research visit to Virginia Tech University and had the opportunity to work with Prof. Edward Valeev. My sincere thank goes to Ed, who introduced me to the field of two-electron integrals evaluation and gave a lot of advanced programming tips.

I also wish to thank my collaborators Prof. Wolfgang Lubitz and Prof. Daniella Goldfarb and the members of their groups Dr. Eckhard Bill, Dr. Maria Pandelia and Dr. Marina Radoul, with whom I have done highly successful and interesting projects on the interface between experiment and theory.

Many other people in the physical and theoretical institute in Bonn and in MPI in Mülheim contributed to a pleasant atmosphere over the last few years. First of all, my warm thanks go to Claudia Kronz, for her help in setting lots of questions for me upon arrival. I would like to thank my roommates in Bonn and in Max-Planck institute: Christoph, Shengfa, Michael, Dimitrios, Maylis, Mahesh. People with whom I had many discussions Dimitriy, DimitriosL, Tobi, Ute, Marius, Ragnar, Oliver, FrankW. Separate thanks to Igor and Vera for the proofreading. Last, I would like to express my gratitude to the whole Neese group!

My warm thanks go to my friends who supported me during PhD time. Michailo and Larisa, Irina and Michail, Lena and Valeriy, Roland, Itana and John, Eugeniy and Olga thanks for being great friends!

Thanks to all from the "Energy Converters" band! That was a great pleasure to rehears, perform and have pizza evenings together.

My very special thanks go to my family, who has been supporting me through the whole time of my PhD studies.
Finally, I want to thank the University of Bonn and Max-Planck society for financial support.

TABLE OF CONTENTS

Abstract V
Zusammenfassung IX
List of Publications XIII
Acknowledgements XV
List of figures XX
List of tables XXII
List of abbreviations XXIII
1. **Introduction** 1
*1.1 **Experimental and Computational Methods in Studying Enzymatic Reactions*** *1*
1.1.1 Overview of the Experimental and Computational Methods 1
1.1.2 DFT Cluster Approach 2
*1.2 **Enzymes*** *3*
1.2.1 Cytochrome C Nitrite Reductase 3
1.2.2 cd1 Nitrite Reductase 6
1.2.3 Hase 1 Hydrogenase 8
*1.3 **Analytical Derivatives Method*** *9*
1.3.1 Introduction to the Analytical Derivatives Method 9
1.3.2 Approximations in the Hessian Evaluation 10
2. **Methods** 12
*2.1 **Calculation of the EPR Parameters with DFT*** *12*
2.1.1 Theory of the EPR in Quantum Chemistry 12
2.1.2 Calculation of the EPR Spectroscopic Parameters for cd1 NiR 14
*2.2 **Electron and Proton Transfer in Enzymes*** *14*
2.2.1 Marcus Theory and Proton-Coupled Electron Transfer Theory 14
2.2.2 Calculation of the ET and PT Rates with DFT 16
2.2.3 Calculation of the Standard Redox Potentials and pK_a's 18
*2.3 **Mössbauer Parameters Calculations with DFT*** *19*
2.3.1 Theory of Mössbauer Spectroscopy in Quantum Chemistry 19
2.3.2 Calculation of the Mössbauer Parameters for Hase 1 Hydrogenase 20
*2.4 **Analytical Derivatives Method*** *21*
2.4.1 Analytic gradient and hessian of the ground state energy 21
2.4.2 Exchange-correlation derivatives with respect to nuclear coordinates 26
2.4.3 The Resolution of Identity for Analytical Second Derivatives 30
2.4.4 The Chain-of-spheres Approximation in Calculation of Hartree-Fock Exchange 32
3. **Results** 37
*3.1 **Substrate binding and activation in CcNiR*** *37*
3.1.1 Abstract 37
3.1.2 Computational Details 37
Building up computational model 37
Computational set up 39

3.1.3 General strategy of the investigation. 39
The scope of the investigation 39
Setting the energetic reference point in pK_a calculations 40
3.1.4 Substrate binding 41
Spin State of the Porphyrine Iron 41
Rest state of the CcNiR enzyme 43
Nitrite Binding in the active site of the enzyme 44
Substrate Back-Bonding in ferrous-nitrite complex 45
3.1.5 The N–O bond cleavage. 46
Possible pathways for the first N–O bond cleavage based on protonated His_{277} 46
Possible pathways for the first N–O bond cleavage based on deprotonated His_{277} 50
Summary of N–O bond cleavage 51
3.1.6 Conclusions for the nitrite binding and activation 52
3.2 Reductive activation of the heme-iron-nitrosyl *53*
3.2.1 Abstract 53
3.2.2 Computational details. 53
Computational model 53
Computational details 54
3.2.3 Results and discussion 55
Geometries 55
Electronic structure, bonding 58
Calculation of the reference points, pK_a and redox potentials. 62
Reaction energetics 64
Modeling of the proton supply 67
Kinetic barrier calculation 69
3.2.4 Conclusions 73
3.3 The Nitroxyl, Hydroxylamine and Ammonia Intermediates *74*
3.3.1 Abstract 74
3.3.2 Computational model 75
Computational set up 75
3.3.3 Results 76
General strategy 76
The HNO intermediate. 76
Reactivity of the HNOH and H_2NO radical stage. 80
Reactivity of the hydroxylamine intermediate and second N-O bond leavage. 83
Ammonia stage 85
Dissociation of the final product 88
3.3.4 Discussion 89
3.4 Dynamic hydrogen bonding network in the distal pocket of the nitrosyl complex of Pseudomonas aeruginosa cd_1 nitrite reductase *90*
3.4.1 Abstract 90
3.4.2 Experimental Section 91
Mutagenesis and protein purification 91
Sample Preparation 91
Spectroscopic Measurements. 92
Computational details. 93
3.4.3 Experimental Results 94
Echo-Detected EPR 94
^1H Davies and ^2H Mims ENDOR 95
3.4.4 Theoretical results 98

Geometric Structure 98
Electronic Structure and Bonding 100
3.4.5 Calculated EPR parameters 102
g-tensor 102
^{14}N hyperfine couplings of NO and proximal His_{182} 103
Hyperfine couplings of the Tyr_{10} and His_{369} protons 104
3.4.6 Discussion 107
3.4.7 Conclusions 109
3.5 The electronic structure of the unique [4Fe-3S] cluster in O_2-tolerant hydrogenases characterized by ^{57}Fe Mössbauer and EPR spectroscopy *110*
3.5.1 Abstract 110
3.5.2 Results 110
Electric Mössbauer Parameters and Valence States. 111
Magnetic Mössbauer Spectra and Spin Distribution. 112
Electronic Structure Calculations 116
3.5.3 Discussion 118
3.6 The Hessian implementation *121*
3.6.1 Implementation 121
The program flow 121
Van der Waals (VDW) correction 121
QM/MM hessian 122
Effective core potential (ECP) Hessian 123
The COSX implementation 123
3.6.2 Numerical results 124
Computational details 124
Integration grids 127
Benchmark result: accuracy 128
Benchmark result: efficiency and timing 131
Benchmark result: representative calculations and parallelization 134
3.6.3 Conclusions 136

4. **Conclusions** 138
Appendices *143*
A. Contour plots of important MOs calculated under B3LYP/TZV(2d,2p) level of theory for the $(HNO^*)H_TN_A$ complex 143
B. Calibration of the PBE(a) and B3LYP(b) functionals for the prediction of ^{57}Fe isomer shifts. 144
C1. Statistical analysis of the errors in the calculated vibrational frequencies (ω) corresponding to the ExtremeSCF(ZM) settings, as determined with respect to the results obtained in the approximation-free ExtremeSCF calculations 145
C2. Statistical analysis of the differences between the vibrational frequencies (ω) calculated using the SCFConvN and ExtremeSCF(ZN) settings 145
D. The derivatives of the Becke grid weights 146
Bibliography *150*

LIST OF FIGURES

Figure 1.1 (a) The overall structure of CcNiR (b) The active site of CcNiR (c) The proposed mechanism **4**
Figure 1.2 The active site of Cytochrome c nitrite reductase (CcNiR) **5**
Figure 1.3 Structures of the [4Fe-3S] cluster of H. marinus hydrogenase, ... **9**
Figure 3.1 Computational model of the C*c*NiR active site **38**
Figure 3.2 Overview of the present investigation on CcNiR **41**
Figure 3.3 (a) Backbonding interaction of nitrite molecule with ferrous iron (b) The spatial orientation of the NO_2^- substrate molecule **45**
Figure 3.4 The charge redistribution in NO_2^- molecule upon binding to the active site **46**
Figure 3.5 The conformational potential energy surface of the NO_2^- (a) and HONO (b) adducts **47**
Figure 3.6 Scheme of the first N–O cleavage **49**
Figure 3.7 Potential energy surface of first N–O bond cleavage reaction **51**
Figure 3.8 Computational model for the C*c*NiR active site. The heme was modeled without its ring substituent **54**
Figure 3.9 The optimized structures of the $\{Fe(NO)\}^6$-$\{Fe(NO)\}^8$ and $\{Fe(HNO)\}^7$-$\{Fe(HNO)\}^8$ complexes **56**
Figure 3.10 Contour plots of important MOs. (a) The $\{Fe(NO)\}^6$ MOs are presented, (b) the $\{Fe(NO)\}^7$ MOs and (c) the $\{Fe(NO)\}^8$ MOs **59**
Figure 3.11 The model of the bis-histidine-ligated heme 3 ..... **62**
Figure 3.12 Overall thermodynamic scheme of the recharging process **64**
Figure 3.13 Possible reactions to overcome the kinetic trap (CEH2 complex) **66**
Figure 3.14 Optimized structures of the Ca-site models representing part of the inlet channel **68**
Figure 3.15 The entire recharging process scheme of the C*c*NiR active site **71**
Figure 3.16 Simulated kinetics for both proposed pathways on the basis of the quantum chemical results **72**
Figure 3.17 Computational model for the C*c*NiR active site (taken from pdb structure 2E80) **75**
Figure 3.18 Contour plots of important MOs for the $(HNO)H_TH_A$ complex **78**
Figure 3.19 Possible protonation and reduction steps starting from $(HNO)H_T$ intermediate **79**
Figure 3.20 Contour plots of important MOs for $(HNOH^*)H_TH_A$ (a) and $(H_2NO^*)H_TH_A$ (b) **82**
Figure 3.21 Possible protonation and reduction steps starting from $(HNOH^*)H_TH_A$ (a) and $(H_2NO^*)H_TH_A$ (b) intermediates **82**
Figure 3.22 Contour plots of important MOs for $(H_2NOH)H_TH_A$ **84**
Figure 3.23 Possible protonation and reduction steps starting from $(H_2NOH)H_T$ intermediate **85**
Figure 3.24 Contour plots of important MOs for $(H_2N^+)H_T$ and $(H_2N^*)H_T$ **86**
Figure 3.25 Possible protonation and reduction steps starting from $(H_2N^+)H_T$ intermediate. **87**
Figure 3.26 Dissociation curves for $(H_3N^+{}^*)H_A$ (a) and $(H_3N^+{}^*)H_TH_A$ (b) models **88**
Figure 3.27 (a) Comparison of W-band ED EPR (8 K) spectra of frozen solutions of WT-NO, Y10F-NO and dHis-NO. Simulations (magenta) of the spectra of WT-NO (b) and Y10F-NO (c) **95**
Figure 3.28 Comparison of orientation selective ^1H Davies ENDOR spectra of WT-NO in H_2O (black) with ^2H Mims ENDOR spectra of WT-NO in D_2O (magenta) at the indicated g values **96**
Figure 3.29 Comparison of orientation selective ^1H Davies ENDOR spectra of a frozen solution of WT-NO (magenta) with that of Y10F-NO (black) at the indicated *g* values **97**
Figure 3.30 Comparison of the orientation selective ^2H Mims ENDOR spectra of WT-NO (magenta) with Y10F-NO (black) and dHis-NO (blue) at the indicated g values **98**
Figure 3.31 Calculated geometric parameters of the three selected frozen models (Å and deg). **99**
Figure 3.32 The SOMO orbital (a) and the spin density (b) of the *model A* **101**
Figure 3.33 Orientation of the g-tensor as obtained from DFT calculations on the frozen structures originating from 1nir **102**
Figure 3.34 Experimental (black) and simulated (magenta/blue) ^1H Davies ENDOR spectra (a) and ^2H Mims ENDOR spectra (b) of a frozen solution of the WT-NO **105**
Figure 3.35 Proposed conformational changes in the distal pocket of the d_1-NO complex caused by mutation. **109**
Figure 3.36 Zero-field Mössbauer spectra of Hase I recorded at 160 K **112**
Figure 3.37 Magnetic Mössbauer spectra of Hase I (A) reduced with H_2 and (B) super-oxidized with air **113**

Figure 3.38 Computational model for the proximal iron-sulfur cluster of the membrane-bound respiratory [NiFe]-hydrogenase. **116**
Figure 3.39 Week bonding interactions between iron centers in the model Ox2_24 **117**
Figure 3.40 Three representative examples for the *AnHess* program: Penicillin, Plastocyanin model and Vancomycine **135**
Figure 3.41 The parallelization efficiency of the COSX algorithm for the analytical hessian implementation. **136**

LIST OF TABLES

Table 3.1 Relative Energies (kcal/mol) and selected geometric parameters (Å) for Water- and Hydroxyl- complexes of the active side of C*c*NiR **42**
Table 3.2 Energies (kcal/mol) and Selected Geometric Parameters (Å) for the substrate complexes **44**
Table 3.3 Atomic charges (Mulliken analysis), bond lengths (Å) and bond orders of the NO_2^- substrate molecule in gas phase and in a bound to ferrous center form. **46**
Table 3.4 Selected geometric parameters for the nitrosyl complexes investigated **57**
Table 3.5 Mulliken population analysis of the B3LYP densities for the nitrosyl complexes investigated **61**
Table 3.6 The calculated selected geometric parameters (Å) and relative energies of high- and low-spin states (ΔE in kcal/mol) for the bis-histidine ligated heme model. **63**
Table 3.7 Calculated adiabatic proton affinities (APA in kcal/mol) and ΔH (kcal/mol) of the water dissociation reaction in the presence of the Ca-site. **69**
Table 3.8 The estimated values for reorganization energies (kcal/mol) using a combination of electrostatic continuum modeling and force constant analysis **70**
Table 3.9 Selected geometric parameters (distances in Å, angles in degrees and vibrational frequencies in cm^{-1}) for the HNO derivatives **77**
Table 3.10 Selected geometric parameters (distances in Å, angles in degrees and vibrational frequencies in cm^{-1}) for the HNOH and H_2NO radical derivatives **80**
Table 3.11 Selected geometric parameters (distances in Å, angles in degrees and vibrational frequencies in cm^{-1}) for the hydroxylamine and hydroxylamine radical species **83**
Table 3.12 Selected geometric parameters (distances in Å) for the Ammonia intermediates **85**
Table 3.13 The geometric parameters (Å and deg) of *models A*, *B* and *C* **100**
Table 3.14 Binding energies (kcal/mol) and spin population on Fe, N(NO) and O(NO) **101**
Table 3.15 The DFT calculated g-values of the three models compared with the experimental values. **102**
Table 3.16 DFT (PBE) calculation of the hyperfine (A) and nuclear quadrupole (e^2Qq/h) coupling constants (MHz) of ^{14}N of the NO and proximal His_{182} ligand **103**
Table 3.17 DFT (PBE) calculation of hyperfine (A) and nuclear quadrupole (e^2Qq/h) coupling constants (MHz) of protons of Tyr_{10} and His_{369} which are H-bonded to the NO group **104**
Table 3.18 FeS Cofactors and Redox States of Hase I. **111**
Table 3.19 Mössbauer parameters of the reduced and super-oxidized Hase I obtained from magnetic-field and temperature dependent measurements **114**
Table 3.20 Calculated Mössbauer parameters for the iron sites of the proximal [4Fe-3S] cluster in the reduced and super-oxidized state **118**
Table 3.21 Distribution of formal oxidation states in the FeS centers of *A. aeolicus* Hase I **119**
Table 3.22 Definition of different ORCA program settings specifying the SCF and CP-SCF convergence criteria together with associated integral and basis function thresholds, used for monitoring the accuracy and efficiency of analytic second derivative calculations **125**
Table 3.23 Definition of the standard DFT grids included in the ORCA package **127**
Table 3.24 Statistical analysis of the errors in the calculated vibrational frequencies (ω) corresponding to different ORCA program settings, as determined with respect to the results obtained in the approximation-free calculations with the ExtremeSCF settings **129**
Table 3.25 Statistical analysis of the errors in the calculated vibrational frequencies (ω) corresponding to different XC grids, as determined with respect to the results obtained with Grid 7 ..... **130**
Table 3.26 Statistical analysis of the differences between the vibrational frequencies (ω) calculated analytically (*AnFreq*) and by the finite-difference methods, *NumFreq1* and *NumFreq2*, using XC grids 4 and 7 ..... **131**
Table 3.27 Statistical analysis of the errors in the calculated vibrational frequencies (ω) corresponding to different two-electron flags *HESS*(i_1, i_2, i_3, i_4) , as determined with respect to the results obtained in the approximation-free calculations with the ExtremeSCF settings **132**
Table 3.28 The timings for the *AnHess* program **134**
Table 3.29 The *AnHess* program timings for three representative examples **135**

LIST OF ABBREVIATIONS

Å	**Å**ngstroem
Arg	Arginine
B3LYP	**B**ecke; **3** parameter; **L**ee-**Y**ang-**P**arr
BO	**B**orn-**O**ppenheimer
BP86	**B**ecke; **P**erdew; 19**86**
BS	**B**roken **S**ymmetry
CAS	**C**omplete **A**ctive **S**pace
CAS-CI	**C**omplete **A**ctive **S**pace **C**onfiguration **I**nteraction
CASSCF	**C**omplete **A**ctive **S**pace **S**elf-**S**onsistent **F**ield
CC	**C**oupled **C**luster
C*c*NiR	**C**ytochrome *c* **N**itrite **R**eductase
CD	**C**ircular **D**ichroism
cd1NiR	cd1 **N**itrite **R**eductase
COSMO	**Co**nductor-like **S**creening **Mo**del
CI	**C**onfiguration **I**nteraction
CIS	**C**onfiguration **I**nteraction with **S**ingle excitations
Def2-TZVP(-f)	**Def**inition **2** of **T**riple-**Z**eta **V**alence plus **P**olarization (**-f** functions)
DFT	**D**ensity **F**unctional **T**heory
DOMO	**D**oubly **O**ccupied **M**olecular **O**rbital
ECP	**E**ffective **C**ore **P**otential
EPR	**E**lectron **P**aramagnetic **R**esonance
eV	**e**lectron-**V**olt
ET	**E**lectron **T**ransfer
GGA	**G**eneralized **G**radient **A**pproximation
GTO	**G**aussian-**T**ype **O**rbital
HF	**H**artree-**F**ock
HS	**H**igh **S**pin
K	**K**elvin
LDA	**L**ocal **D**ensity **A**pproximation
LSD	**L**ocal **S**pin-**D**ensity Approximation
PC	**P**oint **C**harges
PCET	**P**roton **C**oupled **E**lectron **T**ransfer
PSII	**P**hoto**s**ystem **II**
PT	**P**roton **T**ransfer
RHF	**R**estricted **H**artree-**F**ock
RI	**R**esolution of the **I**dentity
RIJCOSX	**R**esolution of the **I**dentity for Coulomb (**J**-) integrals and **C**hain **o**f **S**pheres for Exchange (**X**-) integrals
QM	**Q**uantum **M**echanics
QM/MM	**Q**uantum **M**echanics/**M**olecular **M**echanics
rmsd	**r**oot **m**ean **s**quare **d**eviation
SCF	**S**elf-**C**onsistent **F**ield

SOMO	**S**ingly **O**ccupied **M**olecular **O**rbital
TD-DFT	**T**ime-**D**ependent **D**ensity **F**unctional **T**heory
TS	**T**ransition **S**tate
TZVP	**T**riple-**Z**eta **V**alence plus **P**olarization
UHF	**U**nrestricted **H**artree-**F**ock
UV/Vis	**U**ltraviolet and **Vis**ible part of the electromagnetic spectrum
VMO	**V**irtual **M**olecular **O**rbital
XC	E**x**change-**C**orrelation
VDW	**V**an **d**er **W**aals
ZORA	**Z**ero'th **O**rder **R**egular **A**pproximation
ZPE	**Z**ero **P**oint **E**nergy

1. INTRODUCTION

Enzymes are biological macromolecules that catalyze chemical reactions in a living cell.[1] The variety of the catalyzed processes and efficiency in reaction acceleration is a striking feature of these macromolecular substances. For instance, in case of urease, an enzyme that speeds up the urea hydrolysis, the acceleration amounts to ~10^{14} in comparison to the reaction in aqueous solution.[2]

Enzymes that contain metal ions in their structure are called metalloenzymes. This class of enzymes, in vast majority of cases, is involved into catalysis of redox reactions.[3] Probably the most prominent and complex metalloenzyme is Photosystem II. The enzyme catalyzes the process of photosynthesis, which is of central importance for life on our platen.[4]

Understanding of the enzymatic activity on molecular level takes one of the central roles in contemporary biochemical research. A wide range of modern spectroscopy and computational methods pursues this goal. A brief overview of available experimental and computational technics is given below.

1.1 Experimental and Computational Methods in Studying Enzymatic Reactions

1.1.1 Overview of the Experimental and Computational Methods

Rapid progress in studying enzymatic catalytic activity has been achieved owing to the development of spectroscopy methods. The major step forward was the determination of the three-dimensional structure of myoglobin utilizing X-Ray spectroscopy.[5] This work for the first time made it possible to correlate structural information with enzymatic functionality. Nowadays the protein data bank[6] contains almost 90 000 structures (as of Tuesday Apr 02, 2013 at 5 PM PDT there were 89393 Structures). The protein structure analysis studies are also supported by nuclear magnetic resonance (NMR) spectroscopy.[7] Though the NMR method has been developing several decades it is less used than the X-Ray analysis.

Special attention in studying metalloproteins must be paid to spectroscopy methods used to probe specific regions in the entire protein structure (so called cofactors, prosthetic groups, active sites). These special sites are of particular interest because the enzymatic catalytic activity often almost entirely depends on them. For example, prosthetic groups are the binding sites for the substrates, take part in electron and proton supply, activate and transform the substrate molecules. Gaining understanding of the active site functionality is a key ingredient in studying enzymatic catalytic activity. Spectroscopy methods like electron paramagnetic resonance (EPR),[8] resonance Raman (rR),[9,10] magnetic circular dichroism (MCD)[11] and Mössbauer spectroscopy[12,13] are often the only source of information on structural and electronic features of short-living active site intermediates.

The other way to find valuable information on catalytic process of a chosen enzyme is to apply site-directed mutagenesis or chemical modification.[14] A conclusion on the catalytic activity of one particular side chain can be drawn from these experiments. From the kinetic experiments relative stability of the reaction intermediates is usually assessed.[15,16]

A wide spectrum of theoretical methods has been applied to study enzymatic reactions. The simplest approach is classical molecular mechanics (MM).[17,18] Within this method interactions between atoms are treated in the framework of classical Newton mechanics. This makes possible to study conformational changes, protein-protein interactions, docking processes and others. The computational effort of the MM method is very low which allows studying of entire proteins (several hundreds of thousands or even millions of atoms). On the other hand, MM cannot provide information on the electronic structure of a given reaction intermediate and thus shed light on the processes like bond breaking, charge transfer or excitations.

A possibility to model intermediates in a reaction mechanism is given by applying quantum mechanical (QM) method.[18] Using QM method a lot of detailed information on the energetics, structural features, electronic states can be received. Through the modeling of spectroscopy parameters direct comparison with the experiment is possible.[12] Moreover, QM method makes predictions on reactivity and thus entire reaction flow in the catalytical cycle of interest.[19] Unfortunately, the computational effort for the QM methods is high. Depending on the method of choice the formal scaling with the number of basis function (proportional to the size of a molecule) is between $O(N^3)$ to $O(N^7)$.[20] This puts restriction on the possible size of a molecule in the QM methods. Systems of the size of 1-10 atoms can be studied with correlated methods and up to ~300 atoms can be treated with density functional theory (DFT). To overcome this problem the cluster model approach is used. The principles of this approach are discussed in the following section.

1.1.2 DFT Cluster Approach

The cluster approach is one of the most popular methods in computational chemistry nowadays in dealing with protein modeling.[21] The essence of the method lies in the idea that the protein functionality can be modeled only using an essential part of the protein – the active site. Following this logic the active site is cut out of the protein interior. The truncated model is then used in the calculation in most cases with DFT.[22] Crucial question in the cluster approach is what residues should be included in the model. The simplest approach is a uniform inclusion of the surrounding ligands within some radius from the active site. More clever and preferential choice is to rely on mutational studies of the protein of interest. There the importance of each active site residue is deduced explicitly from the experiment.

To mimic missing protein influence continuum solvation models are usually applied. One of the most popular continuum solvation models is COSMO.[23] In the vast majority of cases the COSMO model gives reliable results in the computational studies of the protein reactivity.[21]

1.2 Enzymes

1.2.1 Cytochrome C Nitrite Reductase

The multiheme cytochrome *c* nitrite reductase (C*c*NiR) is a key enzyme involved in the second and terminal step of the dissimilatory nitrate reduction pathway of the nitrogen cycle.[24,25] The dissimilatory process is used by several bacteria to grow anaerobically with nitrate as a terminal electron acceptor.[26,27] An essential part of the dissimilatory process is the reduction of nitrite to ammonia, which can be carried out efficiently without release of any intermediate. In this case it is called the respiratory nitrite ammonification process.[28] The C*c*NiR enzyme catalyzes ammonification.

The X-ray crystal structures of C*c*NiR from *E.coli*,[29] *W. succinogenes*,[30] *S. deleyianum*,[31] *D. desulfuricans*,[32,33] and *D.vulgaris,*[34,35,36] have been solved and revealed the enzyme to be homodimeric. Each monomer contains five covalently attached *c*-type hemes and two six-coordinated calcium atoms. Four heme-irons are bis-histidine ligated, while the fifth, presumably the active site of the protein, has a unique lysine as a proximal ligand (Figure 1.1, (a), (b)). The crystallographic investigations revealed that the active site residues are strongly conserved in all periplasmic cytochrome *c* nitrite reductases (NrfAs) and form an environment of positive electrostatic potential around the active site.[37] Also, positions of the four bis-histidine heme groups in each monomer are strongly conserved. The distribution of the hemes in the protein facilitates fast and effective transfer of electrons to the active site.[38]

Recently, a novel octaheme cytochrome *c* nitrite reductase (TvNiR) has been isolated from the haloalkaliphilic bacterium *T. nitratireducens*.[39] The TvNiR enzyme exists as a stable hexamer containing 48 hemes. The arrangement of the five hemes in the catalytic C-terminal domain is identical to that found in NrfAs. Also the catalytic site of TvNiR resembles that of NrfAs. However, TvNiR has special structural features, such as a covalent bond between the 'catalytic' tyrosine and the adjacent cysteine. It also features an unusual topography of the product channels that open into the void interior space of the protein hexamer.

A reaction mechanism for nitrite reduction by C*c*NiR was proposed by Einsle *et al*.[40] The authors formulated their hypothesis based on the crystallographic analysis of substrate complexes in combination with density functional theory (DFT) calculations. As a model of the active site of the protein an iron-porphyrine with proximal ammonia and various reaction intermediates as distal ligands were considered in the investigation. The proposed reaction mechanism starts with the binding of substrate, which, after heterolytic cleavage of the first N–O bond, proceeds by two one-electron reduction and protonation steps to a Fe(II)–HNO complex. The latter species was proposed to be readily reduced by two electrons and protonated to yield Fe(II)–H_2NOH, a crystallographically characterized intermediate.[41] Further reduction leads to elimination of a second water molecule and release of the product ammonia (Figure 1.1, (c)).

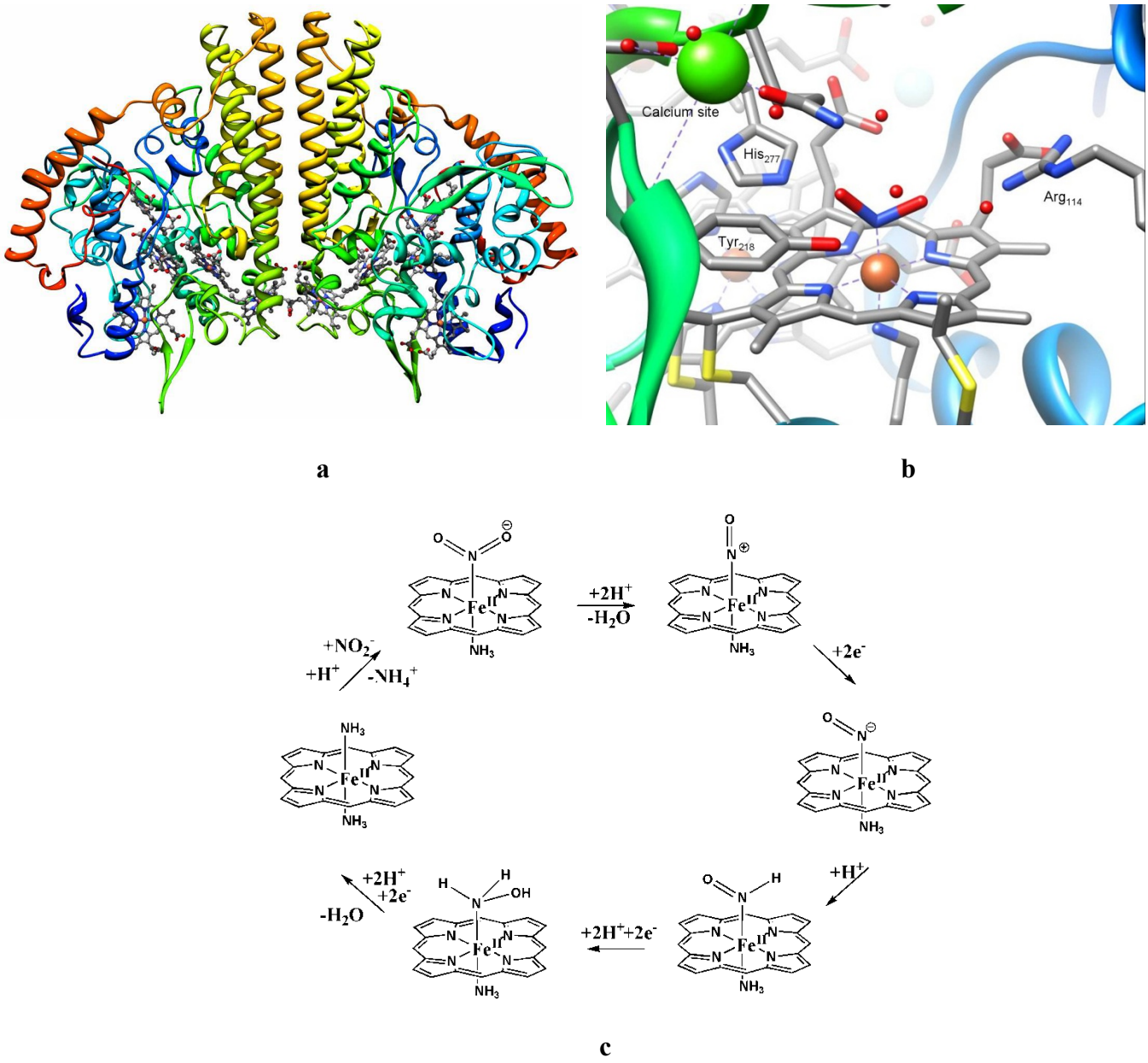


Figure 1.1 (a) The overall structure of CcNiR. Shown is the functional dimer; each monomer contains five c-type hemes. (b) The active site of CcNiR contains His_{277}, Arg_{114} and Tyr_{218} in direct proximity to the NO_2^- substrate molecule. (c) The proposed mechanism of the nitrite reduction by CcNiR adapted from ref.[40]

The ferrous-nitrosyl complex ($\{Fe(NO)\}^7$ in the Enemark and Feltham notation[42]) was characterized spectroscopically and kinetically as well as through DFT calculations.[43,44,45] Hydroxylamine, another putative intermediate, was also characterized.[41] Despite this progress, details of the reaction mechanism still need to be clarified. In particular the work of Einsle *et al.*[40] did not take into account the adjacent residues around the active site and consequently could not provide insight into the role of the side chains in substrate binding and conversion. Furthermore, transition states were not calculated and hence the kinetic plausibility of the proposed mechanism has not yet been assessed.

Obviously, the effectiveness of the reduction critically depends on the electron and proton supply processes. High turnover rates for nitrite reduction by C*c*NiR (415 – 962 s^{-1})[37] suggest very efficient

proton and electron transfer processes that must be guided and supported by the protein. Though there are two calcium ions per monomer only one of them is firmly bound and located close to the active site (here referred to as the first Ca-site). The second calcium site is labile and can be exchanged with one of the cations present in solution. Four heme-irons are bis-histidine ligated, while the fifth, the active site of the protein, has a unique lysine in the proximal coordination site. The number and distribution of the hemes in the proteins are of great importance in facilitating fast and effective transfer of electrons to the active site. In fact, the iron-iron distances of adjacent hemes are all less than 13 Å and one of the hemes is located sufficiently close to the dimer interface to accomplish efficient electron transfer between the subunits.[38] Thus, already the dimeric enzymes, if fully loaded, carry enough electrons to perform the entire reduction cycle. The sequence of reduction events was studied by protein film voltametry (PFV)[46,47] and magnetic circular dichroism spectroscopy (MCD).[48] Although it was speculated that the catalytic heme may act as a site for coordinated exchange of two electrons,[49,50] the later MCD studies of *E.coli* NrfA did not support cooperative reduction.[48] Fast and efficient electron transfer is also ensured by the formation of stable complexes with physiological redox partners (membrame-associated complex with a membrane-anchored tetra-heme cytochrome (NrfH) in case of *W. succinogenes* C*c*NiR[51] or with soluble penta-heme cytochrome NrfB in case of *E.coli*[29]).

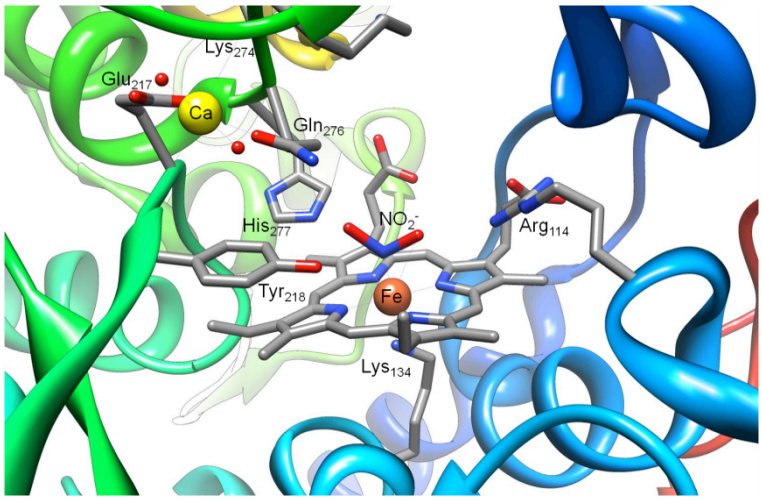


Figure 1.2 The active site of Cytochrome c nitrite reductase (CcNiR). The substrate occupies the distal position of the heme-iron center. His_{277}, Ty_{r218} and Arg_{114} form hydrogen bonds with the substrate. The Ca^{2+} coordination sphere is formed by Gln_{276}, Lys_{274}, Glu_{217}, Tyr_{218} and two water molecules. Gln_{276} and Glu_{217} coordinate calcium through side chains oxygens (Gln_{276}-Oε1, Glu_{217}-Oε1 and Glu_{217}-Oε2) whereas Lys_{274} and Tyr_{218} provide main-chain carbonyl oxygen atoms. (Taken from PDB structure 2E80).

During the catalytic reduction of nitrite, protons are necessarily consumed. These protons are presumably supplied through a positively charged substrate channel.[32] An interesting feature of the substrate channel is the presence of the firmly bound Ca-site in direct proximity to the active site of the enzyme (Figure 1.2). It was established that calcium is essential for activity. Upon addition of Ca^{2+} to the C*c*NiR solution an increase of reductase activity was detected.[52] Moreover, treatment of C*c*NiR with chelating agents led to a significant decrease of activity, which can however be fully restored by addition of Ca^{2+}.[52] The possible role of the Ca-site was also discussed in literature[32,52] where they assign an important physiological role to the Ca-site in terms of structure and

electrostatic interactions. Conceptually, the positive charge has an impact on the electrostatic potential of the inlet channel and the highly conserved coordination sphere helps to stabilize the protein conformation. However, there is another major function that may be assigned to the Ca-site, namely assistance in proton transport events. The water molecules in the substrate channel could donate their protons to the active site while the hydroxide ions, which are then formed, would be stabilized by the Ca^{2+} dication. Keeping in mind that the Ca-site is directly connected to the active site Tyr_{218} and that the distance to the catalytic iron center is only 10.7 Å, the Ca-site may also play an important role in the proton supply. Moreover, a similar proton transfer role has been previously established and clearly discussed for the Zn-containing active site of Carbonic Anhydrase.[53]

The catalytic cycle proceeds through the heme-nitrosyl intermediate formation stage. Heme-nitrosyl derivatives are a subject of continued scientific interest, due to their unique electronic structures and the important role that NO plays in biological processes.[54] In particular NO is involved in biosynthetic processes in mammals (accomplished in the cytochrome P450 family), the cardiovascular regulation and biological transport properties. Moreover, NO is an intermediate in the denitrification cycle used by bacteria for energy production.[44] In fact, nitric oxide reductase activity was studied for C*c*NiR from *E. coli*.[55] The authors established high turnover rates of 840 s^{-1} for NO and predicted the protein to play a role also as an NO detoxification agent.

A number of model systems have been synthesized and characterized.[56] Despite this progress there is little known about the recharging process in heme-Fe-nitrosyl biological catalytic systems. For instance, special attention must be paid to the reduction of the $\{Fe(NO)\}^7$ adduct which was proposed to represent a kinetic trap on a basis of its high stability. In this work, the role of a PCET reaction as a means to overcome the kinetic trap is discussed.

The hydroxylamine adduct, which is considered to be a terminal reaction intermediate, was also characterized spectroscopically.[41] Furthermore, a pathway for the reaction mechanism of the hydroxylamine oxidoreductase has been published.[57] This enzyme catalyzes the back transformation of hydroxylamine into nitrite.

Despite the recent progress in the investigation of the reaction mechanism of C*c*NiR the general picture lacks a lot of details. In particular, proton and electron supply, the second N–O bond cleavage, dissociation of the ammonia final product and other muss be considered in detail. All this makes C*c*NiR a challenging investigation object for the modern computational chemistry.

1.2.2 cd1 Nitrite Reductase

The denitrifying enzyme cd_1 nitrite reductase (cd_1 NiR) catalyzes the reduction of nitrite (NO_2^-) to nitric oxide (NO) as a part of the denitrification process.[58,59] The cd_1 NiR purified from the periplasm of the denitrifying bacteria *Pseudomonas aeruginosa* (*P. aeruginosa*) is a homodimer, each monomer of the enzyme containing one c-heme and one unique d_1-heme.[60,61] The c-heme accepts electrons from external electron donors[62] and transfers them to the d_1-heme, where nitrite binds and is converted to NO.[60,62,63] In the past the d_1 Fe (II)-NO bound state of cd_1 NiR has been considered as a "dead end" product and therefore it was believed that the NO release must take place from the Fe(III)-NO state.[64,45] However, a recent kinetic study[65] on *P. aeruginosa* cd_1 NiR has

shown, unexpectedly, that NO can be rather rapidly released from the fully reduced d_1-heme despite its high affinity ($K_d \approx 10^7$ M) for the heme Fe(II). The NO dissociation rate of *P. aeruginosa* cd_1 NiR is ~ 70 s^{-1}, which is 100-fold faster than that measured for any other heme in the ferrous state.[65] These new findings reveal the uniqueness of the behavior of the cd_1 NiR among other hemeproteins.[66,67,68] On the basis of these studies it was proposed that the unique d_1-heme structure might be a pre-requisite for the faster rate of NO dissociation from the ferrous form, a property which cannot be achieved with a standard b-type heme.[69] The catalytic activity of *P. aeruginosa* cd_1 NiR therefore depends on the unique chemical structure of the d_1-heme. An important role is also played by the two conserved histidines (His_{327} and His_{369}), located in the active pocket site.[65,70,71] Interestingly, a relatively fast NO dissociation rate has been observed also for the H369A mutant (where His_{369} is replaced by alanine).[65]

The three dimensional (3D) crystal structures of *P. aeruginosa* cd_1 NiR of the wild-type (WT) in the oxidized (pdb 1nir),[70] reduced and reduced-NO bound (pdb 1nno) [62] states have been determined. A distinctive feature of *P. aeruginosa* cd_1 NiR enzyme is the so-called "domain swapping" of its N-terminal tail that brings tyrosine (Tyr_{10}) of one monomer close to the d_1-heme site of the other monomer. Tyr_{10} is hydrogen bonded to the hydroxide axial ligand of the oxidized d_1-heme. The reduced and NO-bound WT *P. aeruginosa* cd_1 NiR structures were obtained by soaking the crystals of the oxidized protein with sodium ascorbate and potassium nitrite.[62] Reduction of the WT *P. aeruginosa* cd_1 NiR was found to lead to conformational changes[62] involving tightening of the 56-62 loop in the c-heme domain and rotation of the Tyr_{10} side chain resulting in a 4.2 Å shift of the tyrosine OH group. This movement opens the active site pocket giving access to the substrate. In the Fe(II) NO-bound state the Fe-N-O angle was found to be 135° and the two conserved histidines were found within hydrogen bond distance to the oxygen of NO (3.4 Å to His_{327} and 2.6 Å to His_{369}). Tyr_{10} is shifted away, with a distance of 4.9 and 4.1 Å between its O(OH) and the nitrogen and oxygen of the NO, respectively.

The location of His_{327}, His_{369} and Tyr_{10} in the d_1-heme pocket and the formation of H-bonds between these residues and the substrate (NO^{2-}) and the product (NO) are strategic factors in catalysis. The H327A (where His_{327} was replaced by alanine) and H369A mutants of *P. aeruginosa* cd_1 NiR showed a 100-fold decrease in the nitrite reductive activity compared to the WT. [71] It was also found that the contribution of the two histidines is not equivalent; the H369A mutant showed a more significant decrease in the affinity for nitrite. On the other hand, the replacement of Tyr_{10} by phenylalanine (Y10F mutant) in *P. aeruginosa* cd_1 NiR showed no effect on the nitrite reductase activity.[72] The nitrosyl complex of the H327A (pdb 1hzu) and H369A (pdb 1hzv) mutants[71,73] showed significant differences with respect to the nitrosyl complex of the WT. The N-terminal region of the H327A and H369A mutants is disordered and therefore it is difficult ascertain whether Tyr_{10} is located in the vicinity of the d_1-heme distal pocket; the d_1-heme of the both mutants is remarkably more accessible to the solvent than the WT enzyme.[72] In the H369A-NO complex, the position and orientation of the NO is significantly different from that of the NO bound to the reduced WT structure. Moreover, in this mutant, NO is within H-bonding distance to a water molecule which replaces His_{369} (2.7 Å). These studies demonstrate the structural flexibility of the active site of the cd_1 NiR[17] and raise the question of whether the structural changes observed in the crystal takes place also in solution and whether they are relevant for the activity of the enzyme.

To date, a direct observation and characterization of the proposed hydrogen bonds in the distal active site pocket of the nitrosyl d_1-heme in solution have not been reported. These may play a role in the stabilization/destabilization of the NO bound state. Thus, a detailed study of the distal heme pocket of the nitrosyl d_1-heme complex focusing on the hydrogen bonding to the NO muss be conducted. Specifically, the nitrosyl complexes of WT, Y10F and dHis (a double mutant H327A/H369A,, where the both His_{327} and His_{369} were replaced by alanines) are under consideration. This can be achieved by a combination of high field electron-nuclear double resonance (ENDOR) techniques (W-band, 95 GHz) and density function theory (DFT) calculations. ENDOR spectra yield the hyperfine interactions of the unpaired electron with magnetic nuclei in its close vicinity and their analysis gives structural information. Relating the hyperfine couplings to structural parameters is often non-trivial and requires quantum chemical calculations.[74,75,76] The DFT calculations are often required for both signal assignment and interpretation of the spin Hamiltonian parameters in terms of structure because of the overlapping ENDOR signals of several H-bonds. The effectiveness of this approach was demonstrated in the investigation of the H-bond characteristics of the nitrosyl complex of myoglobin (Mb-NO).[77]

1.2.3 Hase 1 Hydrogenase

Catalysis of hydrogen conversion in [NiFe] hydrogenases relies on electron transfer between the active site and a chain of iron-sulfur clusters.[78] The most abundant components of such chains are low-potential tetranuclear [4Fe-4S] centers of the ferredoxin-type.[79] In the so-called Group I hydrogenases, such [4Fe-4S] clusters are found close to the active [NiFe] site near the enzyme's surface (in the so-called proximal and distal positions, respectively).[78] A third cluster in medial position, typically of the [3Fe-4S] type, completes the electron transfer pathway, in which the cofactors are less than 10 Å apart from each other.

Recently, a new subclass of O_2-tolerant, membrane-bound enzymes was recognized to possess an uncommon proximal iron-sulfur cluster with unique redox properties that enable a new protection mechanism against oxidative inactivation of the [NiFe] site. Typical representatives of such hydrogenases are those from *Aquifex aeolicus*,[80] *Hydrogenovibrio marinus*,[81] *Ralstonia eutropha*[82] and *Escherichia coli*.[83] Three independent crystal structure investigations[81,82,84] revealed that the proximal cluster in these enzymes is not a classical [4Fe-4S] but a novel [4Fe-3S] cluster with an unusual six-cysteine binding motif ($CX\underline{C}CX_{94}CX_4\underline{C}X_{28}C$) that is unique for the sub-class of O_2-tolerant enzymes.[85,86] Four of the six cysteine residues (C) coordinate the cluster in the usual manner as terminal ligands to the [4Fe-3S] core, whereas the sulfur (S^-) atom of one of the supernumerary cysteines ($\underline{C}$) substitutes for an inorganic sulfide (S^{2-}) in the cubane core, and the other serves as a second terminal ligand to one of the Fe atoms[81,82,84] (Figure 1.3). In contrast, the structures of the medial and distal iron-sulfur clusters and their arrangement relative to the [NiFe] site are almost identical to those of standard hydrogenases.[78,82]

Conventional [4Fe-4S] clusters perform either a single-electron 1+/2+ redox transition (in low-potential ferredoxins) or a 2+/3+ transition (in HiPIPs).[79] Both forms share the same isoelectronic 2+ oxidation state with spin $S = 0$, albeit with some minor differences in structures and electron

delocalization.[87,88] In contrast, the proximal [4Fe-3S] cluster in Hase I from *A. aeolicus* can attain *three* redox states within a redox potential span of only 150 mV,[80] two of which are paramagnetic with $S = 1/2$, as we could show by EPR-detected potentiometric titrations. Particularly, a new super-oxidized state, not known for low-potential [4Fe-4S] clusters,[89] manifests at high potentials.[80] According to the crystal structure of *H. marinus* hydrogenase[81] and EPR-detected electrochemical titrations on *A. aeolicus* Hase I,[80] this super-oxidized $[4Fe\text{-}3S]^{5+}$ state is stabilized by structural changes arising from deprotonation of a backbone-nitrogen atom and concomitant nitrogen coordination to one of the Fe sites.[80,81] Also a glutamate group was found within bonding distance to the cluster.[84]

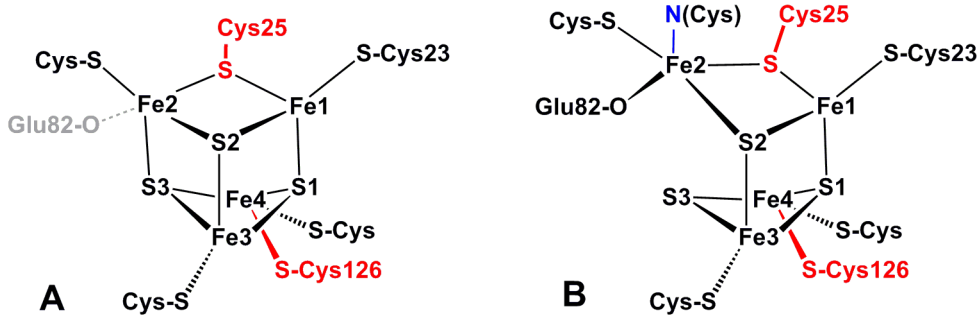


Figure 1.3 Structures of the [4Fe-3S] cluster of H. marinus hydrogenase, (A) in the reduced (3+) state and (B) in the super-oxidized (5+) state (Labeling of iron sites is like in [81]; conversion to the alternative scheme.of (5,[84]is achieved by mapping Fe1→1, Fe2→4, Fe3→2, Fe4→3.) The two 'supernumerary' cysteines Cys-25 and Cys-126 are indicated in red, Glu82 is sketched in grey (A) or black (B). The cysteine closest to the [NiFe] site (Cys-23), and the bond to the backbone nitrogen in the superoxidized state (blue) are also indicated. The distance between Fe2 and S3 varies from 240 pm in the reduced-state (A) to 401 pm in the super-oxidized state (B).

In this contribution, we present a detailed electronic structural investigation of the iron-sulfur cofactors in *A. aeolicus* Hase I based on EPR and applied-field Mössbauer spectroscopy in conjunction with quantum-chemical calculations. In particular, the proximal [4Fe-3S] cluster has been thoroughly examined based on the results of the new crystallographic studies. Local properties of the Fe atoms involved in the redox-induced structural changes have been identified and their spin- and charge distribution has been compared to that of conventional [4Fe-4S] clusters[90,91,92] in view of their crucial role of the [4Fe-3S cluster] toward O_2-tolerance within this class of enzymes.

1.3 Analytical Derivatives Method

1.3.1 Introduction to the Analytical derivatives method

The importance of the derivatives methods in contemporary computational chemistry has been long recognized.[93] Relying on this technique a whole variety of properties can be calculated.[22] Among others, the first self-consistent field (SCF) energy derivatives with respect to nuclear coordinates found their use in the determination of equilibrium geometry and saddle points and subsequent

utilization in studying of potential surfaces and chemical kinetics. The second derivatives of electronic energy are the base for the calculation of force constants, harmonic vibration frequencies, infra red (IR) and Raman intensities.

Starting from pioneering work of Pople *et al.*[94] on the implementation of the analytic second derivatives of the Hartree-Fock SCF energy, the topic stays of great importance. Nowadays, the work mainly goes on in two fields: (1) Implementation of the analytic derivatives for new computational schemes like QM/MM hybride methods;[95,96] (2) development of more efficient implementations.[97,98,99,100] The high need of the efficient implementations straightly follows from the fact that the energy derivatives put much more demand on memory and storage capacities of a computer than the evaluation of the total electronic energy or even the gradients. For instance, the original implementation of Pople *et al.*[94] followed simple strategy of saving integrals and their derivatives on disk (conventional method). This strategy allows treating only small molecules as the storage need quickly exceeds any reasonable computer capacities. In the direct method, first introduced by Almlöf,[101] the two-electron integrals are calculated "on the fly" avoiding integral storage. Apart from the integral storage problems the computational effort in the calculation of the second derivatives is high. As it was discussed many times in literature[102] the most time consuming step in the second SCF energy derivative evaluation is the solution of the coupled-perturbed SCF equations (CP-SCF). The solution of the CP-SCF gives the first derivative of the molecular density, needed for the Hessian construction. The CP-SCF set of equation must be solved for each degree of freedom in a molecule. The dimension of the CP-SCF equations is $i*a$, where i is the number of occupied molecular orbitals (MO) and a is the number of unoccupied MO. As the unoccupied space is usually much larger then the unoccupied the dimension of the CP-SCF equation is much bigger then the SCF equations. Obviously, within the direct integral method the most time goes into the evaluation of the two-electron integrals on each iterative step of the CP-SCF equations solution. The second time-consuming step in the second SCF energy derivatives is the evaluation of the integral second derivatives in the Hessian final terms. Hence, the two-electron integral evaluation needs to be effectively accelerated.

1.3.2 Approximations in the Hessian Evaluation

To speed up the evaluation of the two-electron integrals and their derivatives the resolution of the identity (RI) and the Chain of Spheres (COS) approximations can be applied. The RI method has its roots in work of Baerends *et. al.*[103] It was further elaborated by Dunlop[104,105] and brought to its final form by Almlöf[106] and Ahlrichs.[107,108] The RI-J method (where the resolution of the identity is applied to the Coulomb two-elector part of the Fock operator) was also applied to the calculation of the second energy derivatives.[102] In this implementation by Deglmann *et. al.* the RI-J approximation was applied only in the solution of the CP-SCF equations. Integral first and second derivatives were calculated without the RI technique. The reasoning, given by the authors, was that the integral derivatives "are not the dominant constituents of computation times".

The COS approximation[109] is a seminumerical approach originally introduced for the efficient evaluation of the Hartree-Fock exchange (hence the pseudonym COSX). In this theoretical framework, the integration over coordinates of one electron are evaluated numerically in a two

electron (exchange) integral. It is closely related to the Pseudospectral approach of Friesner;[110] the relationship was discussed in more detail in our earlier work.[111] Since it's original implementation, COS has been successfully used at the second-order Møller-Plesset (MP2) level, along with SCF and MP2 gradients.[112] Various TDDFT methods also benefit enormously from speedups offered by the COS algorithm.[113] Recently, COS was also introduced for the evaluation of various intermediates of higher level correlation methods: the singles Fock term of the localized pair natural orbital coupled cluster (LPNO-CC) method[114] and the external exchange term of spin component scaled MP3 method.[115] As part of the present work, the COS approximation is introduced at various stages of the molecular Hessian evaluation procedure, e.g., the reference energy calculation, various steps of the CPSCF procedure, and the final integral derivative evaluation.

2. METHODS

2.1 Calculation of the EPR Parameters with DFT

2.1.1 Theory of the EPR in Quantum Chemistry

The EPR g-tensor is a property that can be calculated as a second derivative of the energy and it is implemented as such in ORCA for the SCF methods, e.g. HF and DFT. At the SCF level four contributions arise:

$$g_{\mu\nu}^{(SZ)} = \delta_{\mu\nu} g_e \tag{1}$$

$$g_{\mu\nu}^{(RMC)} = -\frac{\alpha^2}{S} \sum_{k,l} P_{kl}^{\alpha-\beta} \left\langle \phi_k \left| \hat{T} \right| \phi_l \right\rangle \tag{2}$$

$$g_{\mu\nu}^{(DSO)} = \frac{1}{2S} \sum_{k,l} P_{kl}^{\alpha-\beta} \left\langle \phi_k \left| \sum_A \xi\left(r_A\right) \left[\mathbf{r}_A \mathbf{r}_O - \mathbf{r}_{A,\mu} \mathbf{r}_{O,\nu} \right] \right| \phi_l \right\rangle \tag{3}$$

$$g_{\mu\nu}^{(PSO)} = \sum_{k,l} \frac{\partial P_{kl}^{\alpha-\beta}}{\partial B_\mu} \left\langle \phi_k \left| h_\nu^{SOC} \right| \phi_k \right\rangle \tag{4}$$

Here, g_e is the free-electron g-value (=2.002319...), S is the total spin, α - the fine structure constant, $P^{\alpha-\beta}$ is the spin density matrix, [116] is the basis set, $\hat{T}$ is the kinetic energy operator, $\xi(r_A)$ an approximate radial operator, h^{SOC} the spatial part of an effective one-electron spin-orbit operator and B_μ is a component of the magnetic field. The derivative of the spin-density matrix is calculated from coupled-perturbed SCF theory with respect to a magnetic field perturbation (orbital Zeeman term).

G-tensor calculations at the SCF level are not highly demanding in terms of basis set size. Basis sets that give reliable SCF results (at least valence double zeta plus polarization) usually also give reliable g-tensor results. For many molecules the Hartree-Fock approximation will give reasonable predictions. In a number of cases, however, it breaks down completely. DFT is more robust in this respect and the number of molecules where it fails is much smaller. Among the density functionals, the hybrid functionals seem to be the most accurate.

The hyperfine coupling has three contributions:

(a) The isotropic Fermi contact term that arises from the finite spin density on the nucleus under investigation. It is calculated for nucleus N from:

$$a_{iso}\left(N\right)=\left(\tfrac{4}{3}\pi\left\langle S_z\right\rangle^{-1}\right)g_e g_N\beta_e\beta_N\rho\left(\vec{R}_N\right) \tag{5}$$

Here, $\left\langle S_z\right\rangle$ is the expectation value of the z-component of the total spin, g_e and g_N are the electron and nuclear g-factors and β_e and β_N are the electron and nuclear magnetons respectively. $\rho\left(\vec{R}_N\right)$ is the spin density at the nucleus. The proportionality factor $P_N=g_e g_N\beta_e\beta_N$ is commonly used and has the dimensions MHz bohr^3 in ORCA.

(b) The spin dipole part that arises from the magnetic dipole interaction of the magnetic nucleus with the magnetic moment of the electron. It is also calculated as an expectation value over the spin density as:

$$A^{dip}_{\mu\nu}\left(N\right)=P_N\sum_{kl}\rho_{kl}\left\langle\phi_k\left|r_N^{-5}\left(3\vec{r}_{N\mu}\vec{r}_{N\nu}-\delta_{\mu\nu}r_N^2\right)\right|\phi_l\right\rangle \tag{6}$$

where ρ is the spin density matrix and $\vec{r}_N$ is a vector of magnitude r_N that points from the nucleus in question to the electron ($\{\varphi\}$, as usual, is the set of basis functions).

(c) The second order contribution that arises from spin-orbit coupling. Presently ORCA can calculate all these contributions. The first two are calculated as simple expectation values of the appropriate operators over the self-consistent spin density, but the second order contribution requires the solution of the coupled-perturbed SCF equations and is consequently computationally more demanding. The contribution can be written:

$$A^{orb}_{\mu\nu}\left(N\right)=-\frac{1}{2S}P_N\sum_{kl}\frac{\partial\rho_{kl}}{\partial I_\mu}\left\langle\phi_k\left|h_\nu^{SOC}\right|\phi_l\right\rangle \tag{7}$$

The derivative of the spin density is computed from solving the coupled-perturbed SCF equations with respect to the nucleus-orbit coupling as perturbation. The nucleus-orbit coupling is represented by the operator:

$$h_\nu^{NOC}\left(A\right)=\sum_i r_{iA}^{-3}l_{i,\nu}^{(A)} \tag{8}$$

where the sum is over electrons and A is the nucleus in question.

The field gradient tensor is closely related to the dipole contribution to the hyperfine coupling. The main differences are that the electron instead of the spin density enters its calculation and that it contains a nuclear contribution due to the surrounding nuclei. It is calculated from

$$\begin{aligned}V_{\mu\nu}\left(N\right)=&-\sum_{kl}P_{kl}\left\langle\phi_k\left|r_N^{-5}\left(3\vec{r}_{N\mu}\vec{r}_{N\nu}-\delta_{\mu\nu}r_N^2\right)\right|\phi_l\right\rangle\\&+\sum_{A\neq N}Z_A\vec{R}_{AN}^{-5}\left(3\vec{R}_{AN\mu}\vec{R}_{AN\nu}-\delta_{\mu\nu}R_{AN}^2\right)\end{aligned} \tag{9}$$

with Z_A as the nuclear charge of nucleus A and $\vec{R}_{AN}$ as a vector of magnitude R_{AN} that points from nucleus A to nucleus N. $\mathbf{P}$ is the first order density matrix.

2.1.2 Calculation of the EPR Spectroscopic Parameters for cd1 NiR

For calculation of the EPR parameters a flexible basis set was used for iron (CP(PPP)[117]); nitrogen and oxygen atoms were treated with EPR-II basis set,[118] carbons and hydrogens with SV(P).[119] To get reliable results from DFT calculations the size of the integration grid was increased for the iron atom.[120] Fermi contact terms and spin-dipole contributions to the hyperfine coupling (HFC) tensor were calculated as expectation values using B3LYP in the preliminary set of calculations of the nine frozen structures of the oxidized form (1nir). For the ***A***, ***B***, ***C*** frozen and relaxed structures of 1nir and relaxed structures of the reduced NO-bound form (1nno) PBE has been used[121] The PBE functional was chosen because it has shown the best agreement with the experiment in the recent extensive investigation of MbNO.[77] Second order contributions to the hyperfine couplings that arise from spin-orbit coupling (SOC) as well as ***g***-tensors were obtained using coupled perturbed Kohn-Sham theory.[122,123] The spin orbit coupling (SOC) operator was treated by spin-orbit mean-field (SOMF) approximation to the Breit-Pauli operator.[124]

Nuclear quadruple coupling constants, e^2qQ/h, were calculated from the electric field gradients V_{ii} according to the equation $e^2qQ/h = const*V_{ii}*Q$, where Q is the nuclear quadrupole moment ($Q(^{14}\mathrm{N})$ = 0.019 barn).[125] The factor $const$ = 234.96 serves to convert e^2qQ/h from atomic to MHz units. The deviation of the nuclear quadrupole tensor from axial symmetry is given by the asymmetry parameter $h = (V_{xx} - V_{yy})/V_{zz}$ in a coordinate system where $|V_{zz}|>|V_{yy}|>|V_{xx}|$.

2.2 Electron and Proton Transfer in Enzymes

2.2.1 Marcus Theory and Proton-Coupled Electron Transfer Theory

In biology, the transfer of protons (PT) and electrons (ET) is usually carried out in a concerted way through proton coupled electron transfer (PCET).[126] The coupling between PT and ET strongly influences both the kinetic and thermodynamic behavior of the process. Generally, a reaction pathway in which electrons and protons are transferred synchronously (the PCET process) is energetically much more favorable than a stepwise mechanism.[127] The underlying reason is that high-energy intermediates that would arise from individual PT and ET steps are avoided. By avoiding these highly endothermic reaction steps, the likelihood of maintaining a reaction cycle with overall low kinetic barriers is increased. Many examples of PCET exist for different biological[128,129] and non-biological systems.[130,131] They have been studied both experimentally[132,133] and theoretically.[134,135,136] The PCET formalism can be readily applied in the case of C*c*NiR since the electrons are supplied from the assisting hemes at the same time as protons are supplied through the inlet channel.

The concept of PCET is rooted in Marcus theory.[137,138] Important further developments are due to Hammes-Schiffer and co-workers[139] and Cukier and co-workers.[140] By closely following the original Marcus derivation, the rate constant, k_{ET}, is given as

$$k_{ET} = k_{ET}(0)\exp\left(\frac{-(\lambda + \Delta G^0)^2}{4\lambda RT}\right) \tag{10}$$

where ΔG^0 is the driving force for the reaction, R and T are the universal gas constant and temperature, λ is the energy needed to adjust the geometric structure of the reagents and reorient the surrounding solvent molecules from the equilibrium configuration to the configuration required for the transfer of an electron. The exponential prefactor $k_{ET}(0)$ is an averaged transition probability for electron transfer, which depends mostly on the distance and degree of electron coupling between reactants.

The rate constant for the PCET, k_{PCET}, was derived by Cukier starting from Marcus picture of charge transfer reactions:

$$k_{CT} = \frac{V_{CT}^2}{\hbar}\sqrt{\frac{\pi}{\lambda_s k_B T}}\exp\left(\frac{-(\lambda_s + \Delta G^0)^2}{4\lambda_s k_B T}\right) \tag{11}$$

where the characteristic parameters that come from the coupling of the reactant and product charge distribution to the surrounding medium are the solvent reorganization energy, λ_s, and reaction free energy, ΔG^0. In the case where the reactant and product states are weakly coupled (nonadiabatic reactions), the exponential prefactor V_{CT} is the electronic (protonic) matrix element V_{ET} (V_{PT}) which evaluates the coupling between two electronic(protonic) states. For stronger coupling (adiabatic reactions), the prefactor is replaced by $\omega_s / 2\pi$, where ω_s characterizes the rate of solvent orientational polarization fluctuations.[140] For the PCET the coupling element V_{PCET} was derived as:

$$V_{PCET} \equiv V_{el}\left\langle \chi_{in'} \mid \chi_{fn} \right\rangle \tag{12}$$

indicating the result for PCET as the product of the ET electronic coupling V_{el} and the Franc-Condon (FC) factors $\left\langle \chi_{in} \mid \chi_{fn} \right\rangle$ connecting the proton in its initial χ_{in} and final states χ_{fn}. Using this simplified coupling expression the rate constant can be formulated as:

$$k_{PCET} = \frac{V_{el}^2}{\hbar^2}\sqrt{\pi\hbar^2 / \lambda_S k_B T}\sum_{n_i}\rho_{n_i}^i\sum_{n_f}\left|\left\langle \chi_{n_f}^f \middle| \chi_{n_i}^i \right\rangle\right|^2 \exp\left(\frac{-(\lambda_S + \Delta E^{el} + \varepsilon_{nf}^f - \varepsilon_{ni}^i)^2}{4\lambda_S k_B T}\right) \tag{13}$$

here the initial states are thermally weighted over the equilibrium proton distribution in its initial state, $\rho_{n_i}^{i}$ and the driving force of the reaction is represented by the sum of ΔE^{el}, the electronic structure contribution, and $\varepsilon_{nf}^{f} - \varepsilon_{ni}^{i}$, the level difference of the proton is its initial and final states.

A similar expression was derived by Hammes-Schiffer in the limit of electronically adiabatic PT and electronically nonadiabatic ET:[141]

$$k_{PCET} = \frac{2\pi}{\hbar} \sum_{\mu} P_{I\mu} \sum_{\nu} V_{\mu\nu}^{2} \sqrt{\frac{1}{4\pi\lambda_{\mu\nu} k_B T}} \exp\left(\frac{-(\lambda_{\mu\nu} + \Delta G_{\mu\nu}^{0})^{2}}{4\lambda_{\mu\nu} k_B T} \right) \qquad (14)$$

Following this formalism, each vibrational mode or channel in the reactant well couples to the product well with a different electronic coupling, $V_{\mu\nu}$, and the overall rate for each vibrational mode is weighted by the Boltzmann factor for the thermal population of that channel.

2.2.2 Calculation of the ET and PT Rates with DFT

According to Marcus theory (eq. 10) the free energy of activation (ΔG^{*}) is related to the driving force of the electron transfer step (ΔG^{0}) and to the reorganization energy (λ) through the equation:

$$\Delta G^{*} = \frac{(\lambda + \Delta G^{0})^{2}}{4\lambda} \qquad (15)$$

The reorganization energy can be expressed as a sum of inner λ_{is} and outer λ_{os} sphere reorganization energies:

$$\lambda = \lambda_{is} + \lambda_{os} \qquad (16)$$

where λ_{os} can be readily calculated applying continuum theory as in equation:

$$\lambda_{os} = (\Delta q)^{2} \left(\frac{1}{2r_1} + \frac{1}{2r_2} - \frac{1}{R} \right) \left(\frac{1}{D_{op}} - \frac{1}{D_s} \right) \qquad (17)$$

Here Δq is the transferred charge, r_1 and r_2 are the effective radii of the reactants, R is the effective distance for electron transfer and D_{op} and D_s are the optical and static dielectric constants of the intervening medium. The D_{op} value is approximately 4 for a protein interior,[142] the radius of the active site is 4 Å and the radius of the electron donating heme is 4.5 Å (according to structure from ref.[40]). The distance between the active site and the donating heme was estimated to be 9.6 Å (the distance between the two iron centers).

The inner sphere reorganization energy λ_{is} is a consequence of the geometric changes in the active site and the heme-donor to the configuration appropriate for electron transfer. We used the procedure described in detail by Rosso *et al.*[143] to estimate the inner sphere reorganization energy in the ET process. It was suggested that the λ_{is} can be calculated from total energy differences of four gas-phase clusters. Namely, the λ_{is} consist of two components:

$$\lambda_{is} = \lambda_{is1} + \lambda_{is2} \tag{18}$$

where λ_{is1} and λ_{is2} are:

$$\lambda_{is1} = E(Fe(III)_{Fe(II)}) - E(Fe(III)_{OPT}) \tag{19}$$

$$\lambda_{is2} = E(Fe(II)_{Fe(III)}) - E(Fe(II)_{OPT}) \tag{20}$$

here the E denotes the total electronic energy and the subscript "OPT" indicates the minimum energy structure; $Fe(III)_{Fe(II)}$ represents a single point energy of iron(III) complex in the optimized iron(II) geometry, and vice versa for $Fe(II)_{Fe(III)}$.

The inner sphere reorganization λ_{is} term for PCET and PT was calculated according to the theory developed by Hammes-Schiffer et al.[144] The inner sphere reorganization energy, assumed to originate from the Ca-site ligand bond reorganization and Fe-N-O core reorganization, is approximated as

$$\lambda_{is} = \sum_j \frac{f_j^r f_j^p}{f_j^r + f_j^p} (\Delta q_j)^2 \tag{21}$$

where the summation runs over Ca-ligand modes (these modes were chosen as metal-ligand bonds of the Ca coordination sphere), Fe-N(NO) and N-O(NO) modes, f_j^r and f_j^p are the equilibrium force constants of the *j*th mode in the reactant and product, respectively, and Δq_j is the difference between reactant and product equilibrium bond lengths. The force constants were calculated at BP86/TZVP level of theory according to procedure developed by Petrenko *et al.*[145]

The corresponding reaction rates for each reaction step were calculated using transition state theory according to the equation:

$$k = \frac{k_b T}{h} \exp\left(-\frac{E_a}{RT}\right) \tag{22}$$

where k_b is the Boltzmann constant, h - Plank constant, R - universal gas constant, T - temperature and E_a - activation energy. Using calculated rate constants, concentrations of every intermediate were deduced as a function of time. The set of differential equations were solved numerically using the Tenua program,[146] that is based on KINSIM by Barshop, Wrenn and Frieden.[147]

2.2.3 Calculation of the Standard Redox Potentials and pK$_a$'s

In addition to the calculated free energy differences (protonation energies, electron affinities and bond formation energies), it is convenient to analyze calculated pK$_a$ values and redox potentials. The values of pK$_a$ and redox potentials were estimated from the calculated free energy differences ΔG according to the following standard equations:

$$\Delta G_{PT} = -RT \ln K_a \tag{23}$$

$$pK_a = -\log K_a \tag{24}$$

$$\Delta G_{ET} = -FE^0 \tag{25}$$

where R is the universal gas constant, T - temperature, K_a - equilibrium constant and F - Faraday's constant. Inserting eq. 24 into 23 one obtains a useful relation between ΔG and pK$_a$:

$$\Delta G_{PT} = 2{,}303RT\ pK_a = -1{,}37(kcal/mol) * pK_a(at\ 298\ K) \tag{26}$$

The eq. 23 and 25 connect absolute dissociation constant and redox potential with free energy changes in one proton/electron reaction. The most common way to find free energy changes is to apply Born-Haber thermodynamic cycle. The method is widely used in pK$_a$ and redox potential calculations *d*-metal complexes.[148] Using the Born-Haber cycle ΔG_{PT} can be expressed:

$$\Delta G_{PT} = \Delta G_g + \Delta G_{sol}(A^-) + \Delta G_{sol}(H^+) - \Delta G_{sol}(AH) \tag{27}$$

where ΔG_g is free energy change of the AH deprotonation in gas phase; $\Delta G_{sol}(A^-)$, $\Delta G_{sol}(AH)$, $\Delta G_{sol}(H^+)$- solvation free energies for AH, $A^-$and a proton H^+. The $\Delta G_{sol}(A^-)$ and $\Delta G_{sol}(AH)$ were calculated using COSMO, whereas the solvation free energy of a proton (275 kcal/mol) was taken from the reference.[149]

The Born-Haber cycle applied for the redox potential calculation gives following equation for the free energy of the one electron reduction:

$$\Delta G_{ET} = \Delta G_g^{ET} + \Delta G_{sol}(OX) - \Delta G_{sol}(RED) \tag{28}$$

where ΔG_g^{ET} is free energy of the one electron reduction in gas phase; $\Delta G_{sol}(OX)$ and $\Delta G_{sol}(RED)$ are the solvation free energy of oxidized and reduced forms respectively. For the sake of comparison to experimental values the absolute potential of the standard hydrogen electrode (SHE) should be added to the calculated values (the absolute electrode potential of SHE is 4.28 V according to Kelly *et al*.[150]).

Although the Born-Haber thermodynamic cycle is utilized frequently in pK_a and redox potential calculations, one must be aware that the *ab initio* prediction of pK_a values and redox potentials is exceedingly difficult. The redox potential of heme-iron complexes is influenced by local charges, ligand type, their basicity and spatial alignment as well as more subtle effects like directed hydrogen bonds or nonplanar distortions. Unfortunately, the calculations also need to be of very high accuracy as an error of 1 kcal in the calculated free energy translates to an error of approximately 40 mV in the predicted redox potential. As the DFT methods that we use here have at least an error of 2-3 kcal,[22] one cannot expect errors of less than 100 mV in these calculations. Similar comments apply to the calculation of pK_a values. A number of strong local interaction such as hydrogen bonds, charge and dipole interactions, or interaction with metal center in the protein environment can shift pK_a values significantly.[151]

2.3 Mössbauer Parameters Calculations with DFT

2.3.1 Theory of Mössbauer Spectroscopy in Quantum Chemistry

Calibration Approach in Calculation of Isomer Shifts

The isomer shift is linearly dependent to the electron density at the Mössbauer absorbing nucleus. Thus, one can write:

$$\delta_{MB} = a + b[\rho(0) - c] \tag{29}$$

where a and b are fit parameters to be determined by linear regression and c is a number merely introduced for convenience. This equation implies that the only quantity that must be known in order to predict the isomer shift is the electron density at the nucleus. Within first-order property framework using DFT the density at the nucleus is readily calculated as:

$$\rho(0) = \sum_{\mu\nu} P_{\mu\nu} \left\langle \phi_\mu \middle| \delta(r-0) \middle| \phi_\nu \right\rangle = \sum_{\mu\nu} P_{\mu\nu} \phi_\mu(0) \phi_\nu(0) \tag{30}$$

The fitting parameters are best obtained form a linear regression analysis of calculated electron densities versus observed isomer shifts. In most cases the predictions on the basis of the simple regression analysis are successful. Importantly, the experience gained from these applications indicates that the quality of the calibration does not depend on the charge-state of the iron centers, not on their spin state, not on their coordination number of the nature of ligands or whether the iron is involved in spin-coupling or not. Thus, these calculations, despite their simplicity, are successful and robust. Usually, regression analysis in the prediction of isomer shifts gives values of uncertainty that is smaller than 0.1 mm/s^{-1}.

The field gradient tensor calculation

The field gradient tensor is closely related to the dipole contribution to the hyperfine coupling described earlier (see 2.1.1). The main differences are that the electron instead of the spin density

enters its calculation and that it contains a nuclear contribution due to the surrounding nuclei. It is calculated from

$$V_{\mu\nu}\left(N\right) = -\sum_{kl} P_{kl}\left\langle \phi_k \left| r_N^{-5}\left(3\vec{r}_{N\mu}\vec{r}_{N\nu} - \delta_{\mu\nu} r_N^2\right)\right| \phi_l \right\rangle + \sum_{A\neq N} Z_A \vec{R}_{AN}^{-5}\left(3\vec{R}_{AN\mu}\vec{R}_{AN\nu} - \delta_{\mu\nu} R_{AN}^2\right) \tag{31}$$

with Z_A as the nuclear charge of nucleus A and $\vec{R}_{AN}$ as a vector of magnitude R_{AN} that points from nucleus A to nucleus N. $\mathbf{P}$ is the first order density matrix.

Once available, and supplemented by the nuclear contribution, the EFG tensor can be diagonalized. The numerically largest element $V_{\max}$ (in atomic units) defines the value of q which is in turn used to calculate the quadrupole splitting parameter as $e^2qQ = 235.28V_{\max}Q$, where Q is the quadrupole moment of the nucleus in barn. Transformed to its eigensystem the quadrupole splitting enters the nuclear Hamiltonian in the following form:

$$\hat{H}_Q = \hat{I}Q\hat{I} = \frac{e^2qQ}{4I(2I-1)}\hat{I}\begin{pmatrix} -(1-\eta) & 0 & 0 \\ 0 & -(1-\eta) & 0 \\ 0 & 0 & 2 \end{pmatrix}\hat{I} \tag{32}$$

The asymmetry parameter η is defined as:

$$\eta = \frac{\left|V_{yy} - V_{xx}\right|}{V_{zz}} \tag{33}$$

The quadrupole splitting is then given by:

$$\Delta E_Q = \frac{1}{2}eqQV_{zz}\sqrt{1+\frac{1}{3}\eta^2} \tag{34}$$

V_{xx}, V_{yy} and V_{zz} are the principal components of the EFG tensor in a coordinate system with $\left|V_{zz}\right| \geq \left|V_{yy}\right| \geq \left|V_{xx}\right|$, e is the positive elementary charge, and $Q(^{57}Fe)$ is the nuclear quadrupole moment of the I=3/2 excited state of ^{57}Fe (measured in barn).

2.3.2 Calculation of the Mössbauer Parameters for Hase 1 Hydrogenase

Various Mössbauer parameters were obtained from PBE/TZVP[18] and B3LYP[152,153]/TZVP single point calculations on the optimized structures. For the calculation of isomer shifts, the test set from Neese *et al.*[120] was used to calibrate our choice of DFT functional and basis set. The calibration procedure consists of calculating the electron density at the nuclei of interest and comparing to the experimentally known isomer shift values. The linear correspondence is then fitted to a straight line

using the least squares method (see Appendix B 1). To obtain reliable Mössbauer parameters form DFT calculations an additional flexibility of the basis set in the nuclear region and higher integration accuracy is required. For this purpose the ORCA "core properties" CP(PPP) basis set was used with the radial integration accuracy parameter increased to 9.0 for iron centers.

Nuclear quadruple coupling constants, e^2qQ/h, were calculated from the electric field gradients V_{ii} according to the equation $e^2qQ/h = const*V_{ii}*Q$, where Q is the nuclear quadrupole moment (Q(Fe) = 0.16 barn). The factor const = 234.96 serves to convert e^2qQ/h from atomic to MHz units (whereas 11.6248 MHz correspond to a Doppler velocity of 1mm/s for the 14.412 keV γ-radiation of ^{57}Fe). The deviation of the nuclear quadrupole tensor from axial symmetry is given by the asymmetry parameter $\eta = (V_{xx} - V_{yy})/V_{zz}$ in a coordinate system where $|V_{zz}|>|V_{yy}|>|V_{xx}|$.

Fermi contact terms and spin-dipole contributions to the hyperfine coupling (HFC) tensor were calculated as expectation values using PBE and B3LYP. Second order contributions to the hyperfine couplings that arise from spin-orbit coupling (SOC) were obtained using coupled perturbed Kohn-Sham theory.[22] The spin orbit coupling (SOC) operator was treated by the spin-orbit mean-field (SOMF) approximation to the Breit-Pauli operator.[23]

2.4 Analytical derivatives method

2.4.1 Analytic gradient and hessian of the ground state energy

In general, for hybrid DFT methods the ground state energy within a given atomic orbital (AO) basis set $\{\varphi_\mu(\mathbf{r})\}$ can be written as

$$E = V_{NN} + \sum_{\mu\nu\sigma} P_{\mu\nu\sigma} h_{\mu\nu} + \frac{1}{2}\sum_{\mu\nu\sigma\eta}\left(P_{\mu\nu\sigma}P_{\kappa\tau\eta} - c_x\delta_{\sigma\eta}P_{\mu\tau\sigma}P_{\nu\kappa\eta}\right)(\mu\nu\,|\,\kappa\tau) + E_{XC}\left[\rho_\sigma(\mathbf{r})\right] \qquad (35)$$

where V_{NN} is the nuclear repulsion energy, and $\hat{h}$ is the one-electron core Hamiltonian operator. Indices μ,ν,κ,τ denote basis functions, $(\mu\nu\,|\,\kappa\tau)$ are four-center Coulomb integrals over basis functions in Mulliken notation, σ and η are the spin labels, E_{XC} is the exchange-correlation (XC) energy, and $\mathbf{P}_\sigma$ is the density matrix that depends on the expansion coefficients of the molecular orbitals (MOs), as follows:

$$P_{\mu\nu\sigma} = \sum_i C_{\mu i\sigma} C_{\nu i\sigma} \qquad (36)$$

As usual, indices i,j denote occupied, a,b virtual, and p,q,r,s general ground state Kohn-Sham (KS) MOs. The electron density is given by

$$\rho_\sigma(\mathbf{r}) = \sum_{\mu\nu} P_{\mu\nu\sigma}\varphi_\mu(\mathbf{r})\varphi_\nu(\mathbf{r}) \qquad (37)$$

The KS equations, as well as the energy derivative properties, can easily be obtained using the Lagrangian formulation of the ground state energy. The Lagrangian is defined as

$$L\left[\{\mathbf{C}_{i\sigma}\},\varepsilon_{ij\sigma}\right]=E\left[\{\mathbf{C}_{i\sigma}\}\right]-\sum_{ij\sigma}\varepsilon_{ij\sigma}\left[\sum_{\mu\nu}C_{\mu i\sigma}C_{\nu j\sigma}S_{\mu\nu}-\delta_{ij}\right] \tag{38}$$

where **S** is the overlap matrix. The Lagrange multipliers $\varepsilon_{ij\sigma}$ are introduced in order to keep orthonormality constraints upon the MOs. The stationary condition with respect to MO coefficients leads to the KS equations, which for the canonical case ($\varepsilon_{ij\sigma}=F_{ij\sigma}=\varepsilon_{i\sigma}\delta_{ij}$) read

$$\sum_{\nu}F_{\mu\nu\sigma}(\mathbf{C})C_{\nu i\sigma}=\varepsilon_{i\sigma}\sum_{\nu}S_{\mu\nu}C_{\nu i\sigma} \tag{39}$$

The matrix elements of the Fock operator are given by

$$F_{\mu\nu\sigma}=h_{\mu\nu\sigma}+\sum_{\kappa\tau\eta}P_{\kappa\tau\eta}\left[(\kappa\tau\,|\,\mu\nu)-c_x\delta_{\sigma\eta}(\kappa\nu\,|\,\mu\tau)\right]+V^{xc}_{\mu\nu\sigma} \tag{40}$$

where $V^{xc}_{\mu\nu\sigma}$ is a matrix element of the XC potential:

$$V^{xc}_{\mu\nu\sigma}=\int\varphi_{\mu}(\mathbf{r})\frac{\delta E^{xc}}{\delta\rho_{\sigma}(\mathbf{r})}\varphi_{\nu}(\mathbf{r})d\mathbf{r} \tag{41}$$

The ground state energy gradient can straightforwardly be derived taking into account that for any external perturbation λ

$$E\left[\{\mathbf{C}_{i\sigma}(\lambda)\}\right]=L\left[\{\mathbf{C}_{i\sigma}(\lambda)\},\varepsilon_{ij\sigma}(\lambda)\right] \tag{42}$$

whereby

$$E^{\lambda}=\sum_{\mu i\sigma}\frac{\partial L}{\partial C_{\mu i\sigma}}C^{\lambda}_{\mu i\sigma}+L^{(\lambda)} \tag{43}$$

The superscript λ denotes the full derivative with respect to λ, and superscript (λ) indicates the derivative with fixed MO coefficients. The first term in the right-hand side of eq. (43) is zero by virtue of the stationary condition for the Lagrangian. Thus, the ground state energy gradient can be evaluated according to the formula:

$$\begin{aligned}E^{\lambda}=&V^{\lambda}_{NN}+\sum_{\mu\nu\sigma}P_{\mu\nu\sigma}h^{\lambda}_{\mu\nu}-\sum_{\mu\nu\sigma}W_{\mu\nu\sigma}S^{\lambda}_{\mu\nu}+\frac{1}{2}\sum_{\mu\nu\sigma\eta}\left(P_{\mu\nu\sigma}P_{\kappa\tau\eta}-c_x\delta_{\sigma\eta}P_{\mu\tau\sigma}P_{\nu\kappa\eta}\right)(\mu\nu\,|\,\kappa\tau)^{\lambda}+\\&+\sum_{\sigma}\int V^{xc}_{\sigma}\left(\rho_{\eta}(\mathbf{r})\right)\rho^{(\lambda)}_{\sigma}(\mathbf{r})d\mathbf{r}\end{aligned} \tag{44}$$

where $\mathbf{W}_{\sigma}$ is the energy-weighted density matrix defined as

$$W_{\mu\nu\sigma}=\sum_{ij}\varepsilon_{ij\sigma}C_{\mu i\sigma}C_{\nu j\sigma} \tag{45}$$

For the following derivation it is convenient to start with a different form of the energy gradient expression:

$$E^{\lambda} = E^{(\lambda)} - \sum_{ij\sigma} F_{ij\sigma} S^{(\lambda)}_{ij\sigma} \tag{46}$$

where

$$S^{(\lambda)}_{pq\sigma} = \sum_{\mu\nu} C_{\mu p\sigma} C_{\nu q\sigma} S^{\lambda}_{\mu\nu} \tag{47}$$

Taking second order derivatives with respect to the parameters λ and ϑ leads to

$$E^{\lambda\theta} = E^{(\lambda)(\vartheta)} + E^{(\lambda)\{\vartheta\}} - \sum_{ij\sigma} F^{(\vartheta)}_{ij\sigma} S^{(\lambda)}_{ij\sigma} - \sum_{ij\sigma} F^{\{\vartheta\}}_{ij\sigma} S^{(\lambda)}_{ij\sigma} - \sum_{ij\sigma} F_{ij\sigma} S^{(\lambda)(\vartheta)}_{ij\sigma} - \sum_{ij\sigma} F_{ij\sigma} S^{(\lambda)\{\vartheta\}}_{ij\sigma} \tag{48}$$

where $D^{\{\vartheta\}}$ $\left(D = E, F_{pq}, S_{pq}\right)$ denotes the MO derivative terms:

$$D^{\{\vartheta\}} = \sum_{\mu i\sigma} \frac{\partial D}{\partial C_{\mu i\sigma}} C^{\vartheta}_{\mu i\sigma} \tag{49}$$

For the perturbed MO coefficients we use the usual parameterization: [94]

$$C^{\vartheta}_{\mu p\sigma} = \sum_{q} C_{\mu q\sigma} U^{\vartheta}_{qp\sigma} \tag{50}$$

The symmetric part of the $\mathbf{U}^{\vartheta}_{\sigma}$ matrix is obtained from the invariance of the MO orthogonality condition upon nuclear displacements,[94] whereby

$$U^{\vartheta}_{qp\sigma} + U^{\vartheta}_{pq\sigma} = -S^{(\vartheta)}_{pq\sigma} \tag{51}$$

The antisymmetric part of the occupied/occupied block of the $\mathbf{U}^{\vartheta}_{\sigma}$ matrix represents a pure rotation between the occupied orbitals, for which the energy is invariant. Thus, we choose[94]

$$U^{\vartheta}_{ij\sigma} = -\frac{1}{2} S^{(\vartheta)}_{ij\sigma} \tag{52}$$

The virtual/occupied block of $\mathbf{U}^{\vartheta}_{\sigma}$ can be obtained from the requirement that the canonical condition $F_{ai\sigma} = 0$ is invariant with respect to external perturbations:

$$F^{\vartheta}_{ai\sigma} = F^{(\vartheta)}_{ai\sigma} + \left(\varepsilon_a - \varepsilon_i\right) U^{\vartheta}_{ai\sigma} - \varepsilon_i S^{(\vartheta)}_{ai\sigma} + G_{ai\sigma}\left[\mathbf{P}^{\vartheta}\right] = 0 \tag{53}$$

where $G_{pq\sigma}[\mathbf{D}]\left(\mathbf{D} = \mathbf{Q}^{\vartheta}, \mathbf{P}^{\vartheta}, \mathbf{U}^{\vartheta}\right)$ is the linear transformation:

$$G_{pq\sigma}[\mathbf{D}] = \sum_{\mu\nu} C_{\mu p\sigma} C_{\nu q\sigma} G_{\mu\nu\sigma}[\mathbf{D}] \tag{54}$$

$$G_{\mu\nu\sigma}[\mathbf{D}] = \sum_{\kappa\tau\eta} \frac{\partial F_{\mu\nu\sigma}}{\partial P_{\kappa\tau\eta}} D_{\kappa\tau\eta} = \sum_{\kappa\tau\eta} \frac{\partial^2 E}{\partial P_{\kappa\tau\eta} \partial P_{\mu\nu\sigma}} D_{\kappa\tau\eta} =$$
$$= \sum_{\kappa\tau\eta} D_{\kappa\tau\eta} \left((\mu\nu \mid \kappa\tau) - c_x \delta_{\sigma\eta} (\kappa\nu \mid \mu\tau) \right) + \sum_{\eta} \int f^{xc}_{\sigma\eta}(\mathbf{r}_1, \mathbf{r}_2) D_\eta(\mathbf{r}_2) \varphi_\mu(\mathbf{r}_1) \varphi_\nu(\mathbf{r}_1) d\mathbf{r} \quad (55)$$

$f^{xc}_{\sigma\eta}(\mathbf{r}_1, \mathbf{r}_2)$ in eq. (55) is the XC kernel:

$$f^{xc}_{\sigma\eta}(\mathbf{r}_1, \mathbf{r}_2) = \frac{\delta^2 E}{\delta\rho_\sigma(\mathbf{r}_1) \delta\rho_\eta(\mathbf{r}_2)} \quad (56)$$

the pseudodensity $D_\sigma(\mathbf{r}_2)$ is defined as usual:

$$D_\sigma(r) = \sum_{\mu\nu} D_{\mu\nu\sigma} \varphi_\mu(\mathbf{r}) \varphi_\nu(\mathbf{r}) \quad (57)$$

The explicit Fock matrix derivatives, $F^{(\vartheta)}_{\mu\nu\sigma}$, are given by

$$F^{(\vartheta)}_{\mu\nu\sigma} = h^\vartheta_{\mu\nu} + \sum_{\kappa\tau\eta} P_{\kappa\tau\eta} \left[(\kappa\tau \mid \mu\nu)^\vartheta - c_x \delta_{\sigma\eta} (\kappa\nu \mid \mu\tau)^\vartheta \right] + V^{xc(\vartheta)}_{\mu\nu\sigma} \quad (58)$$

Taking into account eqs. (36) and (50)-(51), the perturbed density matrices $\mathbf{P}^\vartheta_\sigma$ can be expressed as

$$P^\vartheta_{\mu\nu\sigma} = -Q^\vartheta_{\mu\nu\sigma} + \sum_{ai} U^\vartheta_{ai\sigma} \left(C_{\mu i\sigma} C_{\nu a\sigma} + C_{\nu i\sigma} C_{\mu a\sigma} \right) \quad (59)$$

where we defined the pseudo-density matrices $\mathbf{Q}^\vartheta_\sigma$:

$$Q^\vartheta_{\mu\nu\sigma} = \sum_{\sigma} P_{\mu\kappa\sigma} S^{(\vartheta)}_{\kappa\tau\sigma} P_{\tau\nu\sigma} \quad (60)$$

Substituting eq. (59) into eq. (53) leads to the conventional form of the coupled perturbed SCF (CPSCF) linear equations for $\sum_\sigma N^{occ}_\sigma N^{virt}_\sigma$ unknown coefficients $U^\vartheta_{ai\sigma}$:[94]

$$(\varepsilon_i - \varepsilon_a) U^\vartheta_{ai\sigma} - 2G_{ai\sigma}\left[\mathbf{U}^\vartheta\right] = F^\vartheta_{RHS,ai\sigma} \quad (61)$$

The right-hand sides of the CPSCF equations, $\mathbf{F}^\vartheta_{RHS}$, are given by

$$F^\vartheta_{RHS,ai\sigma} = F^{(\vartheta)}_{ai\sigma} - G_{ai\sigma}\left[\mathbf{Q}^\vartheta\right] - \varepsilon_i S^{(\vartheta)}_{ai\sigma} \quad (62)$$

Given the expression for the perturbed densities (eq. (59)), the individual terms in eq. (48) can then be modified as follows:

$$E^{(\lambda)\{\vartheta\}} = \sum_{\mu\nu\sigma} \frac{\partial E^{(\lambda)}}{\partial P_{\mu\nu\sigma}} P^\vartheta_{\mu\nu\sigma} = -\sum_{\mu\nu\sigma} F^{(\lambda)}_{\mu\nu\sigma} Q^\vartheta_{\mu\nu\sigma} + 2\sum_{ai\sigma} F^{(\lambda)}_{ai\sigma} U^\vartheta_{ai\sigma} \quad (63)$$

$$\sum_{ij\sigma} F_{ij\sigma}^{\{\vartheta\}} S_{ij\sigma}^{(\lambda)} = -\sum_{ij\sigma} \varepsilon_{i\sigma} S_{ij\sigma}^{(\vartheta)} S_{ij\sigma}^{(\lambda)} + \sum_{\mu\nu\sigma\eta} \frac{\partial^2 E}{\partial P_{\kappa\tau\eta} \partial P_{\mu\nu\sigma}} P_{\kappa\tau\eta}^{\vartheta} Q_{\mu\nu\sigma}^{\lambda} = \\ = -\sum_{ij\sigma} \varepsilon_{i\sigma} S_{ij\sigma}^{(\vartheta)} S_{ij\sigma}^{(\lambda)} - \sum_{\mu\nu\sigma} G_{\mu\nu\sigma}\left[\mathbf{Q}^{\lambda}\right] Q_{\mu\nu\sigma}^{\vartheta} + 2\sum_{\mu\nu\sigma} G_{ai\sigma}\left[\mathbf{Q}^{\lambda}\right] U_{ai\sigma}^{\vartheta} \quad (64)$$

$$\sum_{ij\sigma} F_{ij\sigma} S_{ij\sigma}^{(\lambda)\{\vartheta\}} = -\sum_{ij\sigma} \varepsilon_{i\sigma} S_{ij\sigma}^{(\vartheta)} S_{ij\sigma}^{(\lambda)} + 2\sum_{ia\sigma} \varepsilon_{i\sigma} S_{ai\sigma}^{(\lambda)} U_{ai\sigma}^{(\vartheta)} \quad (65)$$

Finally, the second derivatives of the total SCF energy are calculated according to the formula:

$$E^{\vartheta\lambda} = V_{NN}^{\vartheta\lambda} + E^{(\vartheta)(\lambda)} + H\left(\{\varepsilon_i\}, \mathbf{S}^{(\vartheta)}, \mathbf{S}^{(\lambda)}\right) + H\left(\mathbf{F}^{(\vartheta)}, \mathbf{Q}^{\lambda}\right) + H\left(\mathbf{Q}^{\vartheta}, \mathbf{Q}^{\lambda}\right) + H\left(\mathbf{F}_{RHS}^{\lambda}, \mathbf{U}^{\vartheta}\right) \quad (66)$$

where the individual contributions to the Hessian matrix are defined as:

$$H\left(\{\varepsilon_{i\sigma}\}, \mathbf{S}^{(\vartheta)}, \mathbf{S}^{(\lambda)}\right) = 2\sum_{ij\sigma} \varepsilon_{i\sigma} S_{ij\sigma}^{(\vartheta)} S_{ij\sigma}^{(\lambda)} \quad (67)$$

$$H\left(\mathbf{F}^{(\vartheta)}, \mathbf{Q}^{\lambda}\right) = -\sum_{\mu\nu\sigma} F_{\mu\nu\sigma}^{(\vartheta)} Q_{\mu\nu\sigma}^{\lambda} - \sum_{\mu\nu\sigma} F_{\mu\nu\sigma}^{(\lambda)} Q_{\mu\nu\sigma}^{\vartheta} \quad (68)$$

$$H\left(\mathbf{Q}^{\vartheta}, \mathbf{Q}^{\lambda}\right) = \sum_{\mu\nu\sigma} G_{\mu\nu\sigma}\left[\mathbf{Q}^{\lambda}\right] Q_{\mu\nu\sigma}^{\vartheta} \quad (69)$$

$$H\left(\mathbf{F}_{RHS}^{\lambda}, \mathbf{U}^{\vartheta}\right) = 2\sum_{ai\sigma} U_{ai\sigma}^{\vartheta} F_{RHS,ai\sigma}^{\lambda} \quad (70)$$

$$E^{(\vartheta)(\lambda)} = \sum_{\mu\nu\sigma} P_{\mu\nu\sigma} h_{\mu\nu}^{\lambda\theta} - \sum_{\mu\nu\sigma} W_{\mu\nu\sigma} S_{\mu\nu}^{\lambda\theta} + \frac{1}{2}\sum_{\mu\nu\sigma\eta}\left(P_{\mu\nu\sigma} P_{\kappa\tau\eta} - c_x \delta_{\sigma\eta} P_{\mu\tau\sigma} P_{\nu\kappa\eta}\right)(\mu\nu \mid \kappa\tau)^{\lambda\theta} \\ + \sum_{\sigma\eta} \int f_{\sigma\eta}^{xc}(\mathbf{r}_1, \mathbf{r}_2) \rho_{\sigma}^{(\lambda)}(\mathbf{r}_1) \rho_{\eta}^{(\vartheta)}(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 + \sum_{\sigma} \int V_{\sigma}^{xc}(\mathbf{r}_1) \rho_{\sigma}^{(\lambda)(\vartheta)}(\mathbf{r}) d\mathbf{r} \quad (71)$$

Thus, analytical calculation of the second derivatives of the SCF total energy proceeds via the following computational steps:

1. Evaluation of the $\mathbf{S}_{\sigma}^{(\vartheta)}$ (eq. (47)) and $\mathbf{Q}_{\sigma}^{\vartheta}$ matrices (eq. (60)).
2. Formation of the $H\left(\{\varepsilon_{i\sigma}\}, \mathbf{S}^{(\vartheta)}, \mathbf{S}^{(\lambda)}\right)$ terms (eq. (67))
3. Calculation of the $\mathbf{F}_{\sigma}^{(\vartheta)}$ matrices (eq. (58)).
4. Addition of the $H\left(\mathbf{F}^{(\vartheta)}, \mathbf{Q}^{\lambda}\right)$ terms to the Hessian (eq. (68))
5. Calculation of the $\mathbf{G}_{\sigma}\left[\mathbf{Q}^{\lambda}\right]$ matrices (eq.(55)).
6. Addition of the $H\left(\mathbf{Q}^{\vartheta}, \mathbf{Q}^{\lambda}\right)$ terms to the Hessian (eq. (69)).

7. Formation of the right-hand sides of the CPSCF equations (eq. (62)).
8. Solution of the $3N$ CPSCF equations for the orbital rotation coefficients $\mathbf{U}_{\sigma}^{\vartheta}$.
9. Addition of the $H\left(\mathbf{F}_{RHS}^{\lambda}, \mathbf{U}^{\vartheta}\right)$ terms to the Hessian (eq. (70)).
10. Calculation of the integral second derivative contributions (eq. (71))

2.4.2 Exchange-correlation derivatives with respect to nuclear coordinates

The XC contributions to the Hessian (eqs. (55), (58), and (71)) are derived for a general form of the XC functional $E_{XC}\left[\rho_{\sigma}(\mathbf{r})\right]$. In the local gradient-corrected approximation, which is commonly used nowadays, the XC functional is given by the integral of a function f that depends on the local densities and their gradient invariants $\gamma_{\alpha\alpha}, \gamma_{\alpha\beta}, \gamma_{\beta\beta}$:

$$E_{XC}\left[\rho_{\alpha}(\mathbf{r}), \rho_{\beta}(\mathbf{r})\right] = \int f(\rho_{\alpha}, \rho_{\beta}, \gamma_{\alpha\alpha}, \gamma_{\alpha\beta}, \gamma_{\beta\beta}) d\mathbf{r} \tag{72}$$

$$\gamma_{\alpha\alpha} = \left|\nabla\rho_{\alpha}\right|^2, \quad \gamma_{\alpha\beta} = \nabla\rho_{\alpha}\cdot\nabla\rho_{\beta}, \quad \gamma_{\beta\beta} = \left|\nabla\rho_{\beta}\right|^2 \tag{73}$$

For this functional form, the XC contributions to the energy derivatives were considered in some detail in several studies[100,154] We note that despite the consistency of the previous implementations, the corresponding working formulae are somewhat intricate, which obscures simple aspects of the evaluation of the XC derivatives. Below we present a unified and simple strategy for calculating various XC terms that appear in DFT single point and derivative calculations.

Due to the complex functional form of the XC energy, it is evaluated by quadrature methods.[154] The integral in eq. (72) is approximated with a sum over a finite number of grid points $\left\{\mathbf{r}_g\right\}$:

$$\begin{aligned} &E_{XC} \approx \sum_g w_g f_g \\ &f_g = f(\mathbf{r}_g) \quad w_g = w(\mathbf{r}_g) \end{aligned} \tag{74}$$

where w_g are the quadrature weights.

Taking the first derivative with respect to a parameter x leads to

$$\begin{aligned} &E_{XC}^{x} = \sum_{\sigma g} w_g \left(V_{0\sigma g}\rho_{\sigma g}^{x} + \mathbf{V}_{\sigma g}\nabla\rho_{\sigma g}^{x}\right) + \sum_g w_g^{x} f_g \\ &V_{0\sigma g} = V_{0\sigma}(\mathbf{r}_g) \quad \mathbf{V}_{\sigma g} = \mathbf{V}_{\sigma}(\mathbf{r}_g) \quad \rho_{\sigma g} = \rho_{\sigma}(\mathbf{r}_g) \end{aligned} \tag{75}$$

where $V_{0\sigma}(\mathbf{r})$ and $\mathbf{V}_{\sigma}(\mathbf{r})$ $(\sigma = \alpha, \beta)$ are the components of the XC potential defined as

$$V_{0\sigma}(\mathbf{r}) = \frac{\partial f}{\partial \rho_\sigma} \tag{76}$$

$$\mathbf{V}_\sigma(\mathbf{r}) = 2\frac{\partial f}{\partial \gamma_{\sigma\sigma}}\nabla\rho_\sigma + \frac{\partial f}{\partial \gamma_{\alpha\beta}}\left(\delta_{\sigma\alpha}\nabla\rho_\beta + \delta_{\sigma\beta}\nabla\rho_\alpha\right) \tag{77}$$

Taking into account that the quadrature weights depend only on the molecular geometry, for $x = P_{\mu\nu\sigma}$, eq. (75) leads to the well-known expression for the matrix elements of the exchange-correlation potential:[155]

$$V^{xc}_{\mu\nu\sigma} = \frac{\partial E^{xc}}{\partial P_{\mu\nu\sigma}} = \sum_g \left(w_g V_{0\sigma g}\varphi_{\mu g}\varphi_{\nu g} + w_g \mathbf{V}_{\sigma g}\nabla\left(\varphi_{\mu g}\varphi_{\nu g}\right)\right) \tag{78}$$

Likewise, the second derivatives of the XC energy with respect to parameters x and y are straightforwardly evaluated as

$$\begin{aligned} E^{xy}_{XC} = &\sum_{\sigma g} w_g\left(V_{0\sigma g}\rho^{xy}_{\sigma g} + \mathbf{V}_{\sigma g}\nabla\rho^{xy}_{\sigma g}\right) + \sum_{\sigma g} w_g\left(V^x_{0\sigma g}\rho^{y}_{\sigma g} + \mathbf{V}^x_{\sigma g}\nabla\rho^{y}_{\sigma g}\right) + \\ &\sum_{\sigma g} w^y_g\left(V_{0\sigma g}\rho^{x}_{\sigma g} + \mathbf{V}_{\sigma g}\nabla\rho^{x}_{\sigma g}\right) + \sum_{\sigma g} w^x_g\left(V_{0\sigma g}\rho^{y}_{\sigma g} + \mathbf{V}_{\sigma g}\nabla\rho^{y}_{\sigma g}\right) + \sum_g w^{xy}_g f_g \end{aligned} \tag{79}$$

Thus, taking $x = P_{\mu\nu\sigma}$ and $y = P_{\kappa\tau\eta}$, the quadrature formula for the XC contribution to the $G_\sigma[\mathbf{D}]$ can be written in the following way:

$$G^{xc}_{\mu\nu\sigma}[\mathbf{D}] = \sum_{\kappa\tau\eta}\frac{\partial^2 E_{XC}}{\partial P_{\kappa\tau\eta}\partial P_{\mu\nu\sigma}}D_{\kappa\tau\eta} = \sum_{\sigma g}\left(w_g V^1_{0\sigma g}\varphi_{\mu g}\varphi_{\nu g} + w_g \mathbf{V}^1_{\sigma g}\nabla\left(\varphi_{\mu g}\varphi_{\nu g}\right)\right) \tag{80}$$

where $V^1_{0\sigma g} = V^1_{0\sigma}(\mathbf{r}_g, D(\mathbf{r}_g))$ and $\mathbf{V}^1_{\sigma g} = \mathbf{V}^1_\sigma(\mathbf{r}_g, D(\mathbf{r}_g))$ $(\sigma = \alpha, \beta)$ are the components of the effective first order XC potential defined as

$$V^1_{0\sigma}(\mathbf{r}, D(\mathbf{r})) = f^1\left(\rho_\sigma; \{\rho, \gamma, D\}\right) \tag{81}$$

$$\begin{aligned} \mathbf{V}^1_\sigma(\mathbf{r}, D(\mathbf{r})) = &2\frac{\partial f}{\partial \gamma_{\sigma\sigma}}\nabla D_\sigma + \frac{\partial f}{\partial \gamma_{\alpha\beta}}\left(\delta_{\sigma\alpha}\nabla D_\beta + \delta_{\sigma\beta}\nabla D_\alpha\right) + \\ &+ 2f^1\left(\gamma_{\sigma\sigma}; \{\rho, \gamma, D\}\right)\nabla\rho_\sigma + f^1\left(\gamma_{\alpha\beta}; \{\rho, \gamma, D\}\right)\left(\delta_{\sigma\alpha}\nabla\rho_\beta + \delta_{\sigma\beta}\nabla\rho_\alpha\right) \end{aligned} \tag{82}$$

The function $f^1\left(x_p(\mathbf{r}); \{\rho(\mathbf{r}), \gamma(\mathbf{r}), D(\mathbf{r})\}\right)$ $(x_p = \rho_\alpha, \rho_\beta, \gamma_{\alpha\alpha}, \gamma_{\alpha\beta}, \gamma_{\beta\beta})$ is defined as:

$$\begin{aligned} f^1\left(x_p; \{\rho, \gamma, D\}\right) = &\frac{\partial^2 f}{\partial x_p \partial \rho_\alpha}D_\alpha + \frac{\partial^2 f}{\partial x_p \partial \rho_\beta}D_\beta + 2\frac{\partial^2 f}{\partial x_p \partial \gamma_{\alpha\alpha}}\left(\nabla\rho_\alpha \cdot \nabla D_\alpha\right) + 2\frac{\partial^2 f}{\partial x_p \partial \gamma_{\beta\beta}}\left(\nabla\rho_\beta \cdot \nabla D_\beta\right) \\ &+ \frac{\partial^2 f}{\partial x_p \partial \gamma_{\alpha\beta}}\left(\nabla\rho_\alpha \cdot \nabla D_\beta + \nabla D_\alpha \cdot \nabla\rho_\beta\right) \end{aligned} \tag{83}$$

The formula for $G^{xc}_{\mu\nu\sigma}[\mathbf{D}]$(eq. (80)) has the same form as the expression in eq. (78). Therefore, $\mathbf{G}^{xc}_{\sigma}[\mathbf{D}]$ is formed in an analogous fashion to the XC contributions to the Fock matrices in conventional SCF calculations. The only difference consists in the evaluation of the effective XC potential (eqs. (76)-(77) for $V_{0\sigma}$, $\mathbf{V}_{\sigma}$, and eqs. (81)-(83) for $V^{1}_{0\sigma}, \mathbf{V}^{1}_{\sigma}$).

Calculation of XC derivative terms with respect to nuclear coordinates has several important aspects that have been scrutinized in detail in the literature.[100,154,156,157] If during numerical integration the atoms are considered to be embedded in a fixed Cartesian grid, then it is necessary to take into account only the explicit dependence of E_{XC} upon nuclear coordinates via atom-centered Gaussian basis functions. In such a case, the derivatives of the basis function centered on atom *A*, with respect to nuclear displacements, are given by:

$$\frac{\partial^n \varphi_\mu(\mathbf{r}_g - \mathbf{R}_A)}{\partial \mathbf{R}_B^n} = (-1)^n \delta_{AB} \frac{\partial^n \varphi_\mu(\mathbf{r}_g - \mathbf{R}_A)}{\partial \mathbf{r}_g^n} \tag{84}$$

In practice, the numerical grid used to evaluate the XC energy is dependent on the molecular geometry. As the standard technique, the quadrature methods employ the atomic partitioning scheme proposed by Satoko,[158] and subsequently developed by Becke.[159] In this approach, the multicenter integration of eq. (74) is reduced to a sum of single-center integrations, so that a given grid point $\mathbf{r}_g$ belongs to a spherical grid centered on an atom *A*:

$$\mathbf{r}_g = \mathbf{r}_{g(A,i)} = \mathbf{R}_A + \mathbf{r}_i \tag{85}$$

where $\{\mathbf{r}_i\}$ define a suitable one-center spherical integration grid centered at the origin. According to the Becke's scheme, the quadrature weights are given by

$$w_g = w_i w_A(\{\mathbf{R}_B\}) \tag{86}$$

where w_i are the constant weights of the one-center quadrature formula, and the atomic weights w_A depend on the entire molecular geometry.[159] Thus, in order to be consistent with the quadrature approximation of the XC energy, *both* the weight and grid point derivatives must be taken into account in differentiating E_{XC} with respect to nuclear coordinates.[154,157] In this case, the basis function derivatives, which are needed for the evaluation of the XC contributions to the Hessian, should be calculated as follows:

$$\frac{\partial \varphi_\mu(\mathbf{r}_{g(C,i)} - \mathbf{R}_A)}{\partial \mathbf{R}_B} = \frac{\partial \varphi_\mu(\mathbf{r}_{g(C,i)} - \mathbf{R}_A)}{\partial \mathbf{r}_g}(\delta_{BC} - \delta_{AB}) \tag{87}$$

$$\frac{\partial^2 \varphi_\mu(\mathbf{r}_{g(C,i)} - \mathbf{R}_A)}{\partial \mathbf{R}_B \partial \mathbf{R}_D} = \frac{\partial^2 \varphi_\mu(\mathbf{r}_{g(C,i)} - \mathbf{R}_A)}{\partial \mathbf{r}_g^2}(\delta_{BC} - \delta_{AB})(\delta_{DC} - \delta_{AD}) \tag{88}$$

According to Baker et al.,[157] if the weight derivatives are taken into account, then the corresponding grid point derivatives must be included to obtain correct gradients, and vice versa. Otherwise, the

resulting gradients are wholly incorrect if only one of these terms is left out. As established empirically, this conclusion can be extended to any order derivative.

As indicated in several studies, the importance of the grid weight derivatives diminishes with increasing grid size.[154,157] Neglect of these terms typically leads to very small errors in the calculated energy gradients, if properly designed modest grids are employed. As concluded by Baker,[157] essentially identical results, in the sense of chemical accuracy, are obtained with these grids, whether or not quadrature weight derivatives are included in the DFT energy gradient. For smaller grids the neglect of the weight derivatives can result in failure of geometry optimizations, which is related to non-vanishing gradient at an energy minimum.[154] According to our own experience, the inclusion of these terms results in more favorable convergence rates in geometry optimizations incorporating many floppy degrees of freedom, even with the larger standard grids used for this purpose. To account for the weight derivatives is also necessary to maintain rigorously the translational invariance in the calculated gradient.[154] However, if used with smaller grids, the inclusion of the weight derivatives is not able to compensate a poor quality grid.[157] Another important consideration is that the accuracy of the quadrature changes as the molecule is distorted.[156] As a consequence, the computed potential energy surface wobbles around the true one. In this case, it turns out that the energy gradient obtained by assuming the fixed grid is more accurate and smoother than the formally correct gradient from the "moving" grid scheme. Handy et al.[156] supported their arguments by the results of numerical frequency calculations on a few small molecules. This study was based on finite difference between analytic gradients with and without weight derivatives. It was concluded that differentiation of weights gives poorer agreement with the frequencies obtained with large grids having small quadrature errors.

There are basically three methods of weight derivative treatment in the calculation of an analytic Hessian, which are without weight derivatives (denoted 0), and with weight derivatives up to the first (1) and second (2) order. The last two schemes are equivalent to the finite difference between analytic gradients without and with weight derivatives, respectively. Scheme (1) leads to a symmetry breaking of the resulting Hessian matrix. Such an error is typically small for general purpose grids, and decreases with larger grids. This strategy requires symmetrization of the Hessian for frequency calculations. In ref.[154] it was concluded that the error introduced upon going from scheme (1) to (0) is substantially larger than that involved in going from scheme (2) to (1). In our analytic Hessian implementation, the differentiation of weights is performed according to scheme (1). There are two principal advantages of this approach compared to scheme (2). First, as noted in ref.[98], the computational effort for first weight derivatives needed in $\mathbf{F}_{\sigma}^{(\vartheta)}$ and $E_{XC}^{(\vartheta)(\lambda)}$ is negligible, whereas the cost of second weight derivatives is much higher, and comparable to the effort required in the corresponding Coulomb part. Second, we have checked Handy's conclusion that scheme (1) is more accurate than (2), when the corresponding numerical vibrational frequencies via the finite difference of analytic gradients, obtained with extremely tight SCF settings without prescreening, are compared with the results for larger grids (vide infra).

As we could see from numerous test calculations, namely the explicit XC derivatives $E_{XC}^{(\vartheta)(\lambda)}$, this, requires differentiation of weights to yield accurate analytic Hessians, whereas for the rest of the terms, $\mathbf{F}_{\sigma}^{(\vartheta)}$, it is less important, if general purpose grids are used. For the analytic Hessian to be fully equivalent to the numeric one obtained with the finite difference between analytic gradients

without weight derivatives, the differentiation with respect to the displacement parameters λ and ϑ in eq. (66) should be performed without and with differentiation of weights, respectively. However, we make a small deviation from this approach, and include first weight derivatives for both $\mathbf{F}_\sigma^{(\vartheta)}$ and $\mathbf{F}_\sigma^{(\lambda)}$, which contribute to $H\left(\mathbf{F}^{(\vartheta)},\mathbf{Q}^{\lambda}\right)$ (eq. (68)) and $H\left(\mathbf{F}_{RHS}^{\lambda},\mathbf{U}^{\vartheta}\right)$ (eq. (70)). In this way, the latter terms in the Hessian matrix become exactly symmetric, as they are supposed to be, which can thus be considered as an improvement of their accuracy.

The quadrature formula for the XC contribution to the $\mathbf{F}_\sigma^{(\lambda)}$ can be obtained from eq. (79) by taking $x=\lambda$ and $y=P_{\mu\nu\sigma}$:

$$\begin{aligned} F_{\mu\nu\sigma}^{xc,(\lambda)} = &\sum_{\sigma g} w_g \left(V_{0\sigma g}+\mathbf{V}_{\sigma g}\nabla\right)\left(\varphi_{\mu g}\varphi_{\nu g}\right)^{\lambda} + \sum_{\sigma g} w_g \left(V_{0\sigma g}^{\lambda}+\mathbf{V}_{\sigma g}^{\lambda}\nabla\right)\left(\varphi_{\mu g}\varphi_{\nu g}\right) \\ &+ \sum_{\sigma g} w_g^{\lambda} \left(V_{0\sigma g}+\mathbf{V}_{\sigma g}\nabla\right)\left(\varphi_{\mu g}\varphi_{\nu g}\right) \end{aligned} \tag{89}$$

where

$$V_{0\sigma g}^{\lambda} = V_{0\sigma}^{1}(\mathbf{r}_g,\rho^{(\lambda)}(\mathbf{r}_g)) \qquad \mathbf{V}_{\sigma g}^{\lambda} = \mathbf{V}_{\sigma}^{1}(\mathbf{r}_g,\rho^{(\lambda)}(\mathbf{r}_g)) \tag{90}$$

$\rho_\sigma^{(\lambda)}(\mathbf{r}_g)$ is the explicit derivative of the electron density on the grid:

$$\rho_\sigma^{(\lambda)}(\mathbf{r}_g) = \sum_{\mu\nu} P_{\mu\nu\sigma}\left(\varphi_\mu(\mathbf{r}_g)\varphi_\nu(\mathbf{r}_g)\right)^{\lambda} \tag{91}$$

The basis function derivatives in the latter formula should be calculated according to eq. (87).

The explicit second derivatives of the XC energy are obtained from eq. (79) by taking $x=\lambda$, $y=\vartheta$:

$$\begin{aligned} E_{XC}^{(\lambda)(\vartheta)} = &\sum_{\sigma g} w_g \left(V_{0\sigma g}\rho_{\sigma g}^{(\lambda)(\vartheta)}+\mathbf{V}_{\sigma g}\nabla\rho_{\sigma g}^{(\lambda)(\vartheta)}\right) + \sum_{\sigma g} w_g \left(V_{0\sigma g}^{\vartheta}\rho_{\sigma g}^{(\lambda)}+\mathbf{V}_{\sigma g}^{\vartheta}\nabla\rho_{\sigma g}^{(\lambda)}\right) + \\ &+ \sum_{\sigma g} w_g^{\vartheta} \left(V_{0\sigma g}\rho_{\sigma g}^{(\lambda)}+\mathbf{V}_{\sigma g}\nabla\rho_{\sigma g}^{(\lambda)}\right) \end{aligned} \tag{92}$$

According to scheme (1), the weight and grid point derivatives are included only for the parameter ϑ. In particular, the basis function derivatives which contribute to $\rho_\sigma^{(\lambda)}(\mathbf{r}_g)$ and $\rho_\sigma^{(\vartheta)}(\mathbf{r}_g)$ are calculated by eqs. (84) and (87), respectively.

2.4.3 The Resolution of Identity for Analytical Second Derivatives

The rate determining steps in the evaluation of analytic Hessian are:[98,102] 1) the formation of the Fock-like matrices $\mathbf{G}_\sigma\left[\mathbf{Q}^{\vartheta}\right]$ and $\mathbf{F}_\sigma^{(\vartheta)}$ to right-hand sides of the CPSCF equations; 2) formation of the trial vectors $\mathbf{G}_\sigma\left[\mathbf{U}^{\vartheta}\right]$ in the iterative solution of the CPSCF equations; 3) calculation of explicit

energy second derivatives $E^{(\vartheta)(\lambda)}$. For larger molecules, most of the computational time is due to the evaluation of the two-electron integral contributions, if no approximations are employed. Also, a significant computational effort goes into the calculation of the XC terms, which requires a careful optimization of the integration procedures. In this section, we consider efficient approximations for the two-electron terms, which appear in analytic Hessian calculations (eqs. (54)-(71)).

Similar to earlier studies,[98,102] the Fock-like matrices $\mathbf{G}_\sigma[\mathbf{D}]$ and $\mathbf{F}_\sigma^{(\vartheta)}$ are formed in the AO basis using a direct integral algorithm, followed by back transform to the MO basis, as outlined in eqs. (54)-(55).

The Coulomb, $\mathbf{J}_\sigma[\mathbf{D}]$, and exchange-type contributions, $\mathbf{K}_\sigma[\mathbf{D}]$, to the Fock response matrices $\mathbf{G}_\sigma[\mathbf{D}]$ have the same form as in KS Fock matrices:

$$J_{\mu\nu\sigma}[\mathbf{D}] = \sum_{\kappa\tau\eta} D_{\kappa\tau\eta}(\mu\nu\,|\,\kappa\tau) \tag{93}$$

$$K_{\mu\nu\sigma}[\mathbf{D}] = -c_x \sum_{\kappa\tau} D_{\kappa\tau\sigma}(\mu\tau\,|\,\kappa\nu) \tag{94}$$

The Coulomb and exchange parts of the $\mathbf{F}_\sigma^{(\vartheta)}$ matrices, $\mathbf{J}_\sigma^{(\vartheta)}$ and $\mathbf{K}_\sigma^{(\vartheta)}$, respectively, are given by

$$J_{\mu\nu\sigma}^{(\vartheta)} = \sum_{\kappa\tau\eta} P_{\kappa\tau\eta}(\mu\nu\,|\,\kappa\tau)^{\vartheta} \tag{95}$$

$$K_{\mu\nu\sigma}^{(\vartheta)} = -c_x \sum_{\kappa\tau} P_{\kappa\tau\sigma}(\mu\tau\,|\,\kappa\nu)^{\vartheta} \tag{96}$$

For non-hybrid density functionals, the efficiency of analytic second derivative calculations can be substantially improved by using the RI-J approximation for the evaluation of Coulomb terms in eqs. (93), (95), and (71). In ref. [98,102], the RI-J algorithm was applied for the formation of $\mathbf{J}_\sigma[\mathbf{D}]$, which in the case of the Hessian calculations with pure density functionals results in a total speedups of up ~2-2.5 for larger molecules. In the present paper, we also consider the RI-J approximation for the explicit second derivatives of the Coulomb energy and the derivative matrices $\mathbf{J}_\sigma^{(\vartheta)}$.

In the RI-J method, the density is approximated by an expansion in a set of atom-centered auxiliary basis functions $\{\eta_p(\mathbf{r})\}$:

$$D(\mathbf{r}) = \sum_{\mu\nu\eta} D_{\mu\nu\eta}\varphi_\mu(\mathbf{r})\varphi_\nu(\mathbf{r}) = \sum_p d_p \eta_p(\mathbf{r}) \tag{97}$$

The minimization of the residual Coulomb self-repuslion leads to a linear equation system for $\mathbf{d}$:[107]

$$\mathbf{V}\mathbf{d} = \mathbf{g} \tag{98}$$

where

$$V_{pq} = \int \eta_p(\mathbf{r}_1)\eta_q(\mathbf{r}_2) r_{12}^{-1} d\mathbf{r}_1 d\mathbf{r}_2 \tag{99}$$

$$g_p = \sum_{\mu\nu\eta} D_{\mu\nu\eta} (\mu\nu \mid p) \tag{100}$$

Thus, $\mathbf{J}_\sigma[\mathbf{D}]$ in the RI-J approximation becomes:[98]

$$J_{\mu\nu\sigma}[\mathbf{D}] \approx \sum_{\kappa\tau\eta} d_p (p \mid \mu\nu) \tag{101}$$

Taking into account the relation between $\mathbf{d}$ and $\mathbf{g}$ (eq. (98)), explicit differentiation of the expansion coefficients with respect to the displacement parameter ϑ leads to

$$d_p^{(\vartheta)} = \sum_q \left(g_q^{(\vartheta)} - \sum_r d_r V_{rq}^{\vartheta} \right) V_{qp}^{-1} \tag{102}$$

Having calculated $d_p^{(\vartheta)}$, the derivative matrix $\mathbf{J}_\sigma^{(\vartheta)}$ is formed in a simple way:

$$\mathbf{J}_\sigma^{(\vartheta)} = \sum_{\kappa\tau\eta} d_p^{(\vartheta)} (p \mid \mu\nu) + \sum_{\kappa\tau\eta} d_p (p \mid \mu\nu)^{(\vartheta)} \tag{103}$$

Using the density expansion from eq. (97), the Coulomb energy, E_C, reads

$$E_C = \sum_{pq} d_p d_q V_{pq} \tag{104}$$

in which case the expansion coefficients $\mathbf{d}$ are determined from eqs. (98)-(100) with $\mathbf{D}_\sigma = \mathbf{P}_\sigma$. After simple manipulations with eqs. (98), (102), and (104), we arrive at the following formulae for explicit first and second derivatives of the Coulomb energy:

$$E_C^{(\vartheta)} = \sum_p d_p g_p^{(\vartheta)} - \frac{1}{2} \sum_{pq} d_p d_q V_{pq}^{(\vartheta)} \tag{105}$$

$$E_C^{(\lambda)(\vartheta)} = \sum_{pq} d_p^{(\lambda)} d_q^{(\vartheta)} V_{pq} + \sum_p d_p g_p^{(\lambda)(\vartheta)} - \frac{1}{2} \sum_{pq} d_p d_q V_{pq}^{\lambda\vartheta} \tag{106}$$

2.4.4 The Chain-of-spheres Approximation in Calculation of Hartree-Fock Exchange

For an efficient approximation of the Fock-like matrices and explicit energy derivatives in the case of HF and hybrid DFT calculations we combine the RI approximation for the evaluation of the Coulomb terms and seminumerical COSX algorithm for the formation of the exchange-type contribution.

The Hartree-Fock exchange in the AO basis can be explicitly written as

$$K_{\mu\nu\sigma}[\mathbf{D}] = -c_x \sum_{\kappa\tau} D_{\kappa\tau\sigma} \int \mu(\mathbf{r}_1) \kappa(\mathbf{r}_1) \int \frac{\nu(\mathbf{r}_2) \tau(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2. \tag{107}$$

In COS, one of these integrations is performed numerically as

$$K_{\mu\nu\sigma}[\mathbf{D}] \approx -c_x \sum_{\kappa\tau} D_{\kappa\tau\sigma} \sum_g \mu(\mathbf{r}_g)\kappa(\mathbf{r}_g) w_g \int \frac{\nu(\mathbf{r})\tau(\mathbf{r})}{|\mathbf{r}-\mathbf{r}_g|} d\mathbf{r} \tag{108}$$

where κ(rg) denotes the value of κ evaluated at grid point g, and wg is the quadrature weight associated with g. Introducing the intermediates

$$A_g^{\nu\tau} = \int \frac{\nu(\mathbf{r})\tau(\mathbf{r})}{|\mathbf{r}-\mathbf{r}_g|} d\mathbf{r}, \tag{109}$$

and

$$X_{\kappa g} = w_g^{1/2} \kappa(\mathbf{r}_g), \tag{110}$$

the following equation is obtained

$$K_{\mu\nu\sigma}[\mathbf{D}] \approx -c_x \sum_g X_{\mu g} \sum_\tau A_g^{\nu\tau} \sum_\kappa D_{\kappa\tau\sigma} X_{\kappa g} \tag{111}$$

where the terms are ordered to reflect the order of contraction in the implementation.

For gradients, COS is used to evaluate the exchange contribution $\mathbf{K}_\sigma^{(\vartheta)}$ as part of the following contraction with the density

$$E_x^{(\vartheta)} = \sum_{\mu\nu\kappa\tau\sigma} P_{\mu\nu\sigma} K_{\kappa\tau\sigma}^{(\vartheta)} = \sum_{\mu\nu\kappa\tau\sigma} P_{\mu\nu\sigma} P_{\kappa\tau\sigma} (\mu\kappa|\nu\tau)^\vartheta = 4 \sum_{\mu\nu\kappa\lambda\sigma} P_{\mu\nu\sigma} P_{\kappa\tau\sigma} (\mu^\vartheta\kappa|\nu\tau), \tag{112}$$

Note that the permutation symmetry of integrals and relabeling of summation indices are used here in the exact case in order to avoid the evaluation of the derivatives of the quantity $A_g^{\nu\tau}$ in the corresponding COS expression (where the permutation symmetry is broken). This approximation is consistent with the earlier Pseudospectral approach. Thus, the COSX gradient term reads

$$E_x^{(\vartheta)} = -4c_x \sum_g P_{\mu\nu\sigma} X_{\mu^\vartheta g} \sum_\tau A_g^{\nu\tau} \sum_\kappa P_{\kappa\tau\sigma} X_{\kappa g} \tag{113}$$

where

$$X_{\mu^\vartheta g} = w_g^{1/2} \mu^\vartheta(\mathbf{r}_g). \tag{114}$$

Note that this expression corresponds to evaluating of the derivative function on the grid, and therefore no grid weight derivatives occur (see discussion below).

In a Hessian calculation, the construction of the CPSCF right hand sides requires the evaluation of $K_\sigma^{(\vartheta)}$ directly (i.e. not contracted with a second density), while the Hessian contains $K_\sigma^{(\vartheta)}$ is contracted with the perturbed density. When the COS approximation is used, it must be kept in mind that as a consequence of the lack of two identical densities in these cases, the permutation and relabeling approach can only be partially performed in the exact case to yield simplified COS

expressions. One could still avoid derivatives of $A_g^{\nu\tau}$ and use those of μ and κ, which would correspond to evaluating derivative function values in $K_\sigma^{(\vartheta)}$ as

$$K_{\mu\nu\sigma}^{(\vartheta)} \approx -c_x \sum_g X_{\mu^\vartheta g} \sum_\tau A_g^{\nu\tau} \sum_\kappa P_{\kappa\tau} X_{\kappa g} - c_x \sum_g X_{\mu g} \sum_\tau A_g^{\nu\tau} \sum_\kappa P_{\kappa\tau} X_{\kappa^\vartheta g} \tag{115}$$

However, one cannot avoid the $A_g^{\nu\tau}$ first derivative for the integral second derivatives, and thus for a Hessian calculation this term must be implemented and evaluated anyway. Furthermore, each contraction in the second term scales with the fourth power of the system size; the contraction with the density now has to be carried out as many times as many nuclear coordinates there are. Since the $A_g^{\nu\tau}$ derivatives are available, it seems worthwhile to use them here, even though this approach is somewhat more expensive, it is also expected to be more accurate, since the derivative functions appear in the analytic integration part (these have terms with higher angular momenta and numerically would require a larger grid for the same given accuracy). Thus, the expression eventually used is

$$K_{\mu\nu\sigma}^{(\vartheta)} \approx -c_x \sum_g X_{\mu g} \sum_\tau A_g^{(\nu\tau)^\vartheta} \sum_\kappa P_{\kappa\tau} X_{\kappa g} \tag{116}$$

where

$$A_g^{(\nu\tau)^\vartheta} = \int \frac{1}{|\mathbf{r}-\mathbf{r}_g|} (\nu(\mathbf{r})\tau(\mathbf{r}))^\vartheta d\mathbf{r}. \tag{117}$$

Note that the derivative of the Coulomb operator is not involved here (see discussion below).

Finally, COS is used to evaluate the second integral derivative contributions to the Hessian,

$$\begin{aligned} E_x^{(\lambda)(\vartheta)} = & -c_x \sum_{\mu\nu\kappa\tau\sigma} P_{\mu\nu\sigma} P_{\kappa\tau\sigma} (\mu\kappa|\nu\tau)^{\vartheta\lambda} = \\ & -4c_x \sum_{\mu\nu\kappa\tau\sigma} P_{\mu\nu\sigma} P_{\kappa\tau\sigma} ((\mu^{\vartheta\lambda}\kappa|\nu\tau) + (\mu^\vartheta\kappa^\lambda|\nu\tau) + (\mu^\vartheta\kappa|\nu^\lambda\tau) + (\mu^\vartheta\kappa|\nu\tau^\lambda)) \end{aligned} \tag{118}$$

where again symmetries are exploited to obtain favorable COS expressions

$$\begin{aligned} & E_x^{(\lambda)(\vartheta)} = \\ & -c_x \sum_g \sum_{\mu\nu} (X_{\mu g} P_{\mu\nu} \sum_\tau A_g^{\nu^{\vartheta\lambda}\tau} \sum_\kappa P_{\kappa\tau} X_{\kappa g} + \\ & X_{\mu^\vartheta g} P_{\mu\nu} \sum_\tau A_g^{\nu\tau} \sum_\kappa P_{\kappa\tau} X_{\kappa^\lambda g} + X_{\mu^\vartheta g} P_{\mu\nu} \sum_\tau A_g^{(\nu\tau)^\lambda} \sum_\kappa P_{\kappa\tau} X_{\kappa g}) \end{aligned} \tag{119}$$

For the first term containing the second derivative functions, these are kept within the analytic integration part as

$$A_g^{\nu^{\vartheta\lambda}\tau} = \int \frac{\tau(\mathbf{r})}{|\mathbf{r}-\mathbf{r}_g|} \nu(\mathbf{r})^{\vartheta\lambda} d\mathbf{r}, \tag{120}$$

since these are expected to be the most sensitive to numerical evaluation.

Earlier in the section, the lack of grid weight derivatives was referred to. This is a consequence of the fact that our COS Hessian is a numerical approximation to the exact one rather than the derivative of the numerical energy formula. The significance of the grid weight derivatives was discussed earlier in the literature in a DFT context. At this point it is worth reconsidering the problem from the COS point of view, to clarify why the grid weight correction was eventually abandoned.

The essence of numerical integration is the evaluation of the integral of a function as

$$\int f\,dx = \sum_g w_g f_g$$

as discussed earlier. In the followings we will assume that the grid is "well behaved", i.e. that the above sum converges to the analytic value of the integral, as the number of grid points go to infinity. A similar assumption is made about the derivatives of the function.

If one considers the nuclear derivative of a function depending on the difference, $\Delta X = X_A - X_e$, of a nuclear and an electronic coordinate, one arrives at

$$\frac{\partial}{\partial X_N}\int f(X_A - X_e)d\tau = \int \frac{\partial f}{\partial \Delta X}\frac{\partial \Delta X}{\partial X_N}d\tau = \int f'\delta_{AN}d\tau$$

Evaluating the derivative functions on a grid is what one does in COS, yielding

$$\sum_g f'_g \delta_{AN} w_g$$

Alternatively, one can derivate the numerical integral formula to get

$$\frac{\partial}{\partial X_N}\sum_g f(X_A - X_g)w_g = \sum_g f'_g(\delta_{AN} - \delta_{P(g)N})w_g + \sum_g f_g w'_g$$

where $P(g)$ is the parent atom of gridpoint g. Since the two approaches should both converge to the analytic result using a well behaved grid, one finds that their difference should disappear at the infinite grid limit,

$$\lim_{N_g\to\infty}\left(\sum_g f_g w'_g - \sum_g \delta_{P(g)N} f'_g w_g\right) = 0$$

In words, this means that at the infinite grid limit, the grid weight derivative contributions (first term) cancel out the artificial grid point derivative contributions arising from the discretization of electronic coordinates (second term). The practical conclusion is that one should either use a larger grid using the first approach, neglecting all terms that disappear in the above limit, or invest some additional effort into weight derivatives, and use all these additional terms with a smaller grid. This also explains the result observed by Baker and discussed above that grid point derivative terms and

weight derivative terms should be used together, since only their difference disappear at the analytic limit.

Considering molecular integrals, these can in general be evaluated numerically as

$$\sum_g \mu_{Ag} \nu_{Bg} A_g w_g \qquad (121)$$

where Ag is a different function for different integrals. For the two electron integrals, the above results need to be modified somewhat, since the Coulomb operator (1/r12) does not depend on any nuclear coordinates in the exact case, while numerically a dependence is introduced via grid point coordinates (1/r1g),

$$A_g^{\kappa\tau} = \int \frac{\kappa_{Cg} \tau_{Dg}}{r_{1g}} d\mathbf{r}_1. \qquad (122)$$

This means that other than derivative contributions from κCg and τDg, an additional term depending on the derivative of 1/r1g also has to be added to the grid point derivative part, so that the above limit would properly approach zero. This is a purely artificial term that appears due to the numerical procedure and which is also cancelled by grid weight derivative contributions at the infinite grid limit. The present implementation of the COSX gradient (and Hessian) for the HF exchange corresponds to the simpler approach, where all weight and grid point derivatives are neglected, since in order to add weight derivatives, one also has to implement the Coulomb operator derivatives. At the cost of a larger grid, the present procedure is simpler, while in the weight derivative scheme the additional operator term may also increase the cost. For larger grids the increase in the computational effort needed can be significant, where in fact this contribution is expected to be cancelled. However, grid weight derivatives may be useful in the efficient evaluation of DFT functionals, where no such formal problems arise. A derivation of the Becke grid weights is therefore provided in the Appendix D.

3. RESULTS

3.1 Substrate binding and activation in C*c*NiR

3.1.1 Abstract

Cytochrome *c* nitrite reductase is a homodimeric enzyme, containing five covalently attached *c*-type hemes per subunit. Four of the heme-irons are bis-histidine ligated, while the fifth, the active site of the protein, has an unusual lysine coordination and calcium site nearby. A fascinating feature of this enzyme is that the full six-electron reduction of the nitrite is achieved without release of any detectable reaction intermediate. Moreover, the enzyme is known to work over a wide pH range. Both findings suggest a unique flexibility of the active site in the complicated six-electron, seven-proton reduction process. In the present work, we employed density functional theory to study the energetics and kinetics of the initial stages of nitrite reduction.

The possible role of second-sphere active site amino acids as proton donors is investigated by taking different possible protonation states and geometrical conformations into account. It was found that the most probable proton donor is His_{277}, whose spatial orientation and fine-tuned acidity lead to energetically feasible, low-barrier protonation reactions. However, substrate protonation may also be accomplished by Arg_{114}. The calculated barriers for this pathway are only slightly higher than the experimentally determined value of 15.2 kcal/mol for the rate-limiting step. Hence, having proton donating side chains of different acidity within the active site may increase the operational pH range of the enzyme. Interestingly, Tyr_{218}, which was proposed to play an important role in the overall mechanism, appears to not take part in the reaction during the initial stage.

3.1.2 Computational Details

Building up computational model

Starting structures for the enzyme active site models were taken from the protein data bank[6] (PDB code 2E80[40]). Relying on the X-ray, voltammetry[46] and mutational studies,[160,161] the model for the theoretical investigation was constructed in the following way: His_{277}, Arg_{114} and Tyr_{218} were taken as the first sphere distal ligands. These side chains were cut at α-carbon atoms. The heme is included in the model without its substituents. Lys_{134}, the proximal ligand of iron, was approximated by methylamine. Also, one crystallographic water molecule was included in the active site model (Figure 3.1).

Constraints were imposed on the initial structures in order to mimic the steric influence of the protein backbone on the active site. To avoid unrealistic rearrangements of the protein side chains, we have chosen a constraint fragment based optimization strategy. The molecular fragments that

constitute the cluster model are connected with each other. The positions of all atoms within the fragments are fully optimized by the program and only the shortest distance between connected fragments, angles and dihedrals, which contain this distance, will be constrained to keep the fragments together and retain their spatial orientation. The necessary algorithms for these calculations are implemented in the ORCA program.

In this investigation Tyr_{218}, His_{277} and Arg_{114} form three fragments connected to active site that itself was defined to consist of four fragments: The heme group, the substrate molecule, the axial Lys_{134} and an additional water molecule. All fragments are fully optimized but their mutual relative orientation in space is kept fixed to prevent unreasonable movements of the side chains that cannot happen in the true protein.

To verify this constraint type, complexes with NO_2^-, HNO_2 and $\{Fe(NO)\}^6$ were also calculated with a weaker set of constraints where only the α-carbons of Tyr_{218}, His_{277} and Arg_{114} side chains were fixed. It was found that the energetic differences between two constraint types for the first protonation reaction $NO_2^- + H^+ \rightarrow HNO_2$ are within 2 kcal/mol. By contrast, the further protonation $Fe\text{-}HNO_2 + H^+ \rightarrow \{Fe(NO)\}^6 + H_2O$ has an energetic differences between two constraint types that can reach 6 kcal/mol. We believe that the energy difference arises largely from the repulsive interaction of side chains with the positively charged $Fe(II)\text{-}NO^+$ adduct. As a result of this interaction, the active site side chains, if not constraint, change their positions drastically and move out of the active site. This is obviously impossible in the real protein. Consequently, the fragment constraints reflect the real situation more realistically.

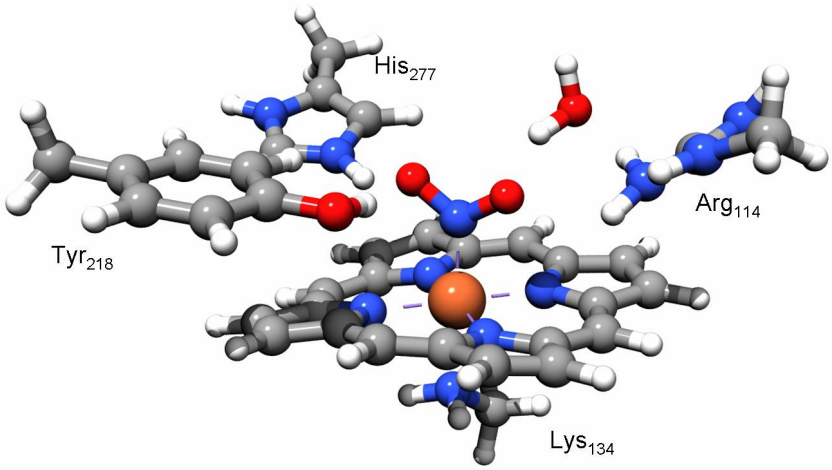


Figure 3.1 Computational model of the C*c*NiR active site. The heme was taken without its ring substituents; side chains His_{277}, Arg_{114}, Tyr_{218} and Lys_{134} were cut at α-carbon atoms; one crystallographic water was fit into the model.

Computational set up

The calculations were performed with the ORCA electronic structure package[117] using DFT.[22] The models of the enzyme active site were optimized with the BP86 functional[162,163] together with the resolution-of-the-identity (RI) approximation (in the Split-RI-J variant)[164] procedure for the Coulomb term.[103,105] Geometry optimizations were carried out using a triple-ζ basis set TZVP[165] together with a matching auxiliary basis set (TZV/J[107,108] in ORCA notation) applied to all atoms. Subsequent numerical frequency calculations were performed for all models for the sake of verifying founded structures as local minima or transition states and obtaining thermochemical corrections. The frequencies were computed through two-sided numerical differentiation of analytic gradients using increments of 0.002 bohr.

Single point energy calculations for each model were performed with the B3LYP functional[152,153] that is thought to provide higher accuracy than BP86 for reaction and activation energies.[21] Moreover, an additional set of calculations for resting state of the enzyme and nitrite-bound state was performed using the B3LYP* functional[166] to investigate the sensitivity of the calculated spin states energies with respect to the fraction of Hartree-Fock exchange involved. The 'chain-of-spheres' (RIJCOSX) approximation[109] was invoked in these calculations. The basis set was of triple-ζ quality with two sets of first polarization functions on all atoms TZV(2d, 2p)[167] in ORCA notation).

In order to model the long-range dielectric effects of the protein, the COSMO[23] solvation model was employed with a dielectric constant of 4.9. All other COSMO parameters were left at their default values. Moreover, a semi-empirical method to account for van der Waals interactions was employed.[168,169]

For the visualization purposes the unrestricted Kohn-Sham orbitals were transformed to quasi-restricted orbitals[170] (QROs), which were localized according to the Pipek-Mezey intrinsic localization criterion.[171] Orbitals, densities and structures were visualized with the Chimera program.[172]

3.1.3 General strategy of the investigation.

The scope of the investigation

The general strategy of the investigation is shown in Figure 3.2. The complexes RS1-RS4 represent the resting state of the enzyme. Different protonation states of His_{277} together with distal water and hydroxyl ligands and all possible spin states were investigated. Substrate binding was considered for both feasible coordination modes (Fe-$\underline{O}NO^-$ and Fe-$\underline{N}OO^-$; complexes A1-A4). Again, all possible spin- and protonation states of His_{277} were considered. Oxygen coordination was found to be energetically less favorable by 4.3 kcal/mol. Thus, the proton transfer reaction was only modeled for complexes A1 and A4. As proton donors, all three side chains of the active site were considered. The results of the first proton transfer reaction are complexes B1-B6. From all six variants the most

thermodynamically favorable are reactions A1→B2 and A4→B6. Therefore B2 and B6 were taken as starting points for the second proton transfer reaction.

Starting from complex B2, three scenarios are possible. The first is direct transfer of the proton from Arg_{114} (B2→D2), where the second is a recharging of His_{277} (complex C2) and subsequent proton transfer from it (C2→D3). The final conceivable way is the formation of the HONOH adduct C1. A similar scheme was considered for the B6 complex. Direct transfer from Tyr_{218} (B6 to D4) was compared with the recharging schemes of Arg_{114} (B6→C3→D2 and C3→D1). In all cases the N–O bond is broken and the final product $\{Fe(NO)\}^6$ complex is formed (complexes D1-D4).

Setting the energetic reference point in pK_a calculations

Considering the recharging of His_{277} (B2→C2) and Arg_{114} (B6→C3), the importance of the energetic reference point should be mentioned. In this case the energy of the incoming proton is implied as a 'reference point'. As the source of protons is an open question for this system one way to estimate the proton's energy is to take the proton reference energy from cluster-ion solvation data.[149] Another possibility is to calculate the protonation enthalpy of a side chain of interest in solution (while relying on a continuum solvation model) and compare it to the protonation enthalpy of the same side chain in the active site under investigation. The difference in computed enthalpies then provides an estimate of the shift in the pK_a value. Both approaches were used in present investigation.

The quantitatively more satisfying procedure is the first approach. Using the calculated free energy of dissociation of histidine in the complex of interest, ΔG_D^{His}, one can calculate the corresponding equilibrium constant and the pK_a value by eqs. (123) and (124). ΔG_D^{His} is calculated as a difference between the free energy of formation of the complex with a protonated histidine side chain (ΔG_C^{HisH}) and the complex with deprotonated histidine (ΔG_C^{His}) minus the proton energy from cluster-ion solvation data ($\Delta G_{sol}^{H^+}$), see eq (3). The pK_a of His_{277} was calculated to be 8.8, which suggests that the protonated form of His_{277} should be considered on the initial stage of the reaction.

$$\Delta G_D^{His} = -RT \ln K_a \tag{123}$$

$$pK_a = -\log K_a \tag{124}$$

$$\Delta G_D^{His} = \Delta G_C^{HisH^+} - \Delta G_C^{His} - \Delta G_{sol}^{H^+} \tag{125}$$

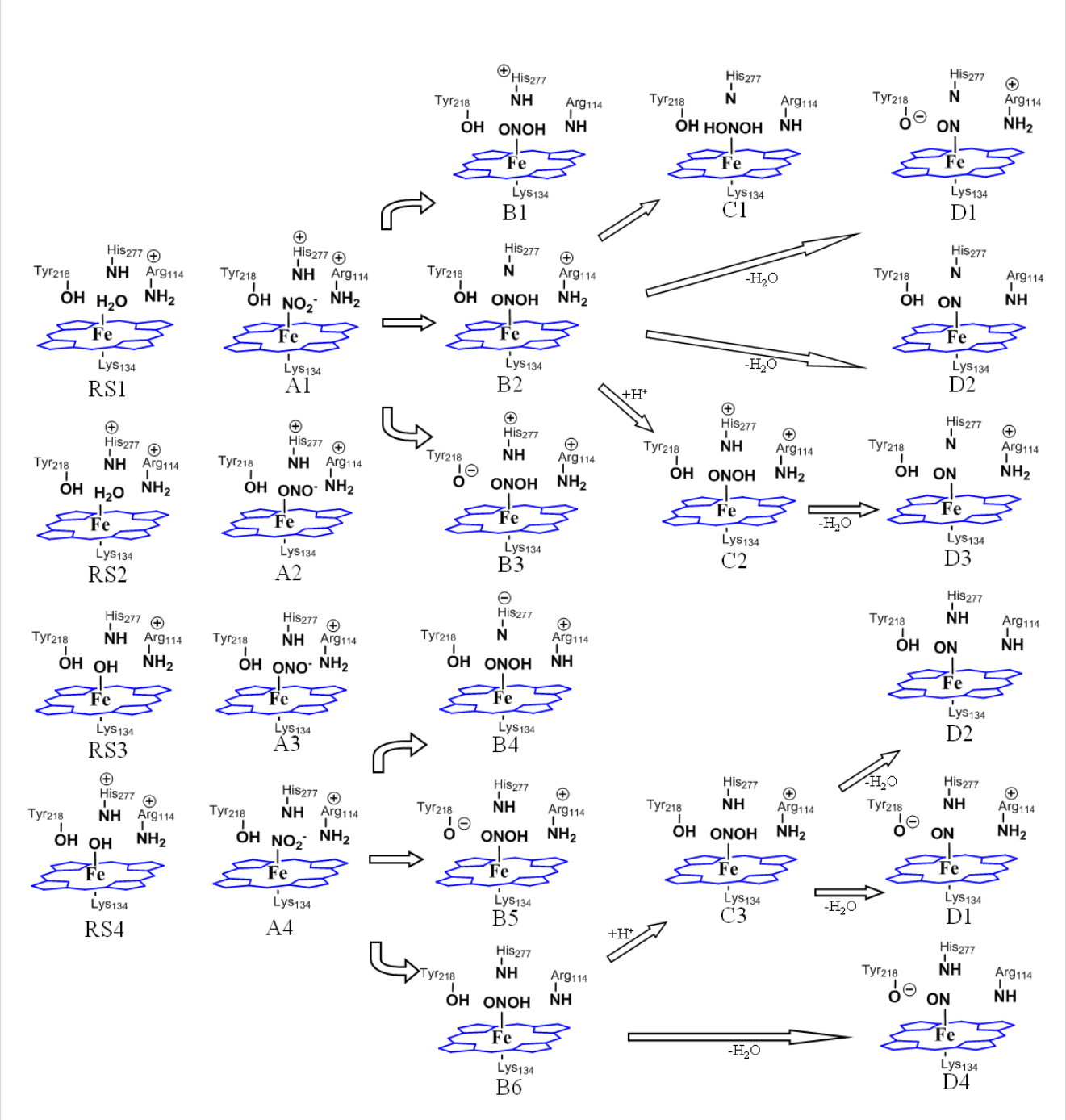


Figure 3.2 Overview of the present investigation. Complexes RS1-RS4 represent possible resting states of the enzyme. A1-A4 represent substrate complexes. There are six possible proton transfer reactions leading to the HONO adduct (complexes B1-B6), which may than accept a second proton to form the $\{Fe(NO)\}^6$ intermediate (D1-D4) either directly or through a recharging stage (C2, C3). The scheme also evaluates the hypothetical formation of a HONOH adduct (reaction B2→C1).

3.1.4 Substrate binding

Spin State of the Porphyrine Iron

In order to explore the nitrite binding process, different possible structures, oxidation states and spin states for the resting state of the enzyme and the nitrite adduct were investigated. Protonated and deprotonated forms of His_{277} were considered. The distal ligand was modeled as H_2O, OH^- and NO_2^- respectively. Both, N- and O- coordination modes of nitrite were examined.

The role of the axial ligands in hemoproteins is a longstanding issue. We refer to a recent publication by Rydberg *et al.* in this respect.[173] Using DFT, the authors provided an overview over spin-state energies, geometries, electronic structure and bonding, and the reaction energies of the intermediates involved in the reaction cycles of cytochrome P450, peroxidase, and catalase. At variance with experimental results,[174] it was concluded that ferric water complexes are unstable, and that the ferric hydroxyl doublet state is much more stable than higher spin states. It is known that pure functionals always overestimates the stability of lower spin states, when hybrid functionals perform better relative to high-accuracy wavefunction based calculations and experiment.[175] However, no single functional has been found that is uniformly accurate for all metals oxidation- and spin-states. The single point calculations in this work were conducted using the B3LYP hybrid functional that is among the better ones for predicting relative spin state energetics. Moreover, an additional set of calculations using the B3LYP* functional was performed which differs from B3LYP by a reduced fraction of Hartree-Fock exchange (15% vs 20%). While no single fraction of HF exchange is universally optimal for the calculation of spin state energetics, the B3LYP* functional has previously given reasonable results for a variety of transition metal complexes.[176]

According to our results, low spin ferrous (ferric)-hydroxocomplexes are favored by 8 kcal/mol in comparison to the respective triplet and quartet states and by more than 9 kcal/mol relative to the high-spin quintet and sextet forms respectively (Table 3.1). The B3LYP* functional predicts even larger level splittings of 9.2 (10.5) kcal/mol for triplet (quartet) and 12.8 (14.7) kcal/mol for the quintet (sextet) states. In the case of the protonated His_{277} side chain, the hydroxide ligand dissociates from low-spin ferrous complexes upon optimization. In these calculations, the OH^-group receives a proton from His_{277} and rebinds to the metal center in form of an axial water ligand.

Table 3.1 Relative Energies (kcal/mol, B3LYP/TZV(2d,2p)) and selected geometric parameters (Å) for Water- and Hydroxyl- complexes of the active side of C*c*NiR. "Unst" stands for Unstable.

His_{277} side chain is NOT protonated					His_{277} side chain is protonated				
Complex	Spin State	Relative Energy	Fe – O distance	Fe – N(porph.) distance	Complex	Spin State	Relative Energy,	Fe – O distance	Fe – N(porph.) distance
	S=0	0	1.961	1.984, 2.003, 2.016, 1.992		S=0	Unst*	Unst*	
Fe^{2+} - OH	S=1	8.1(9.2)**	1.862	2.004, 2.019, 2.030, 2.006	Fe^{2+} - OH	S=1	0	1.889	2.004, 2.029, 2.027, 2.005
	S=2	11.0(12.8)	1.991	2.030, 2.037, 2.049, 2.018		S=2	Unst*	Unst*	
	S=1/2	0	1.855	1.996, 2.011, 2.027, 1.999		S=1/2	0	1.879	1.997, 2.017, 2.019, 1.998
Fe^{3+} - OH	S=3/2	8.0(10.5)	1.974	2.016, 2.026, 2.031, 2.005	Fe^{3+} - OH	S=3/2	Unst*	Unst*	

	S=5/2	9.4(14.7)	1.935	2.069, 2.088, 2.073, 2.088		S=5/2	9.8(14.8)	2.094	2.069, 2.078, 2.070, 2.061
	S=0	0	2.134	1.996, 2.001, 1.999, 1.999		S=0	0	2.169	1.984, 2.016, 2.005, 1.993
Fe^{2+} - H_2O	S=1	Unst	Unst		Fe^{2+} - H_2O	S=1	Unst	Unst	
	S=2	Unst	Unst			S=2	Unst	Unst	
	S=1/2	0	2.106	1.998, 2.003, 1.998, 2.009		S=1/2	0	2.133	1.986, 2.024, 2.004, 2.005
Fe^{3+} - H_2O	S=3/2	Unst	Unst		Fe^{3+} - H_2O	S=3/2	Unst	Unst	
	S=5/2	Unst	Unst			S=5/2	Unst	Unst	

* the dissociating hydroxyl group receives a proton from the nearby histidine to form water.

** energies in parentheses were calculated with B3LYP* functional.

Rest state of the CcNiR enzyme

As expected, there is no additional π-bonding taking place in case of either ferric or ferrous complexes with an axial water ligand. Hence, water is only weakly bound with binding energies of -10.8 and -5.6 kcal/mol for the ferrous and ferric complexes respectively. To obtain more accurate binding energies, one should take into account the binding energy of water in bulk solution. Two separate effects must be considered here: the enthalpy effect related to the formation of multiple hydrogen bonds in bulk solution and the entropy effect associated with the immobilization of the water molecule at active site. It has been established that the enthalpy effect can reach a value of 7 kcal/mol.[177] For the particular case of C*c*NiR, the enthalpy difference between bulk water and water bound to the active site should not be as large because of the substrate channel is filled with water. The water molecules in the substrate channel are obviously also involved in multiple hydrogen bonding networks that include the water molecule bound to the metal centre. The size of the entropy penalty was discussed in the literature as well.[178] It was concluded that the value should not exceed 2 kcal/mol at 300 K. Consequently, the impact on the binding energies should be even smaller. In summary, it worth emphasizing that although the binding process is a complex issue, the calculated values do reflect the relative stability of the Fe-H_2O bonds well, which was the primary goal of our H_2O binding energy calculations.

The low spin Fe(III) – H_2O species exhibits an iron-oxygen bond of 2.106 Å, which is in good agreement with earlier crystallographically determined value of 2.1 Å.[40] The low spin Fe(III) – H_2O complex is assumed to be the resting state of the enzyme.[40] Upon reduction of the resting state to the Fe(II) – H_2O form, the iron – oxygen distance increases to 2.134 Å, thus indicating a weakening of this bond which is also reflected in high driving force of -39 kcal/mol for the substitution of the bound water by nitrite. Therefore, the axial water in the ferrous water active site can be readily displaced by nitrite. For the high and intermediate spin ferric and ferrous active sites, water dissociates from the active site during the optimization (Table 3.1).

Nitrite Binding in the active site of the enzyme

A recent resonance Raman investigation by Hildebrand and co-workers provided evidence that nitrite binding to the enzyme from *D. vulgaris* induces a transition from the high-spin to the low-spin configuration in the catalytic heme center.[179] The authors concluded that the low-spin/high-spin ratio rises from about 2 to 3.35. This has been interpreted as arising from the binding of nitrite to the free axial position of some of the five-coordinated high-spin hemes.

In summary, on the basis of experimentally observed Fe–O distance and the recent resonance Raman investigation, the resting state of the enzyme is a ferric – H_2O complex, from which the axial water ligand can be readily replaced by the substrate nitrite upon reduction of the protein at the entrance point of the catalytic turnover cycle.

In a recent investigation of Yi *et al*. the binding of nitrite through oxygen was studied for horse heart myoglobin.[180] This protein features an iron-porphyrin active site and a histidine axial ligand. It was shown that the distal histidine residue in the binding pocket forces the oxygen binding mode of nitrite through formation of a specific hydrogen bond. Thus, the possibility of oxygen coordination of nitrite was also examined for C*c*NiR. The results in Table 3.2 show that oxygen coordination is energetically unfavorable, however, not by a large margin. Thus, while being energetically disfavored, the small energy differences indicate that this binding mode may occur under certain circumstances as implied in the case of nitrite-myoglobin. The Fe-O(NO_2) bond distances are significantly longer compared to Fe-N(NO_2) and the nitrite ligand is only weakly bound in the intermediate and high spin complexes or dissociates upon optimization.

In summary, it is concluded that binding through oxygen or the change to a higher spin states is unfavorable for C*c*NiR. This result is consistent with the previous study by Einsle *et al*.[40] using a smaller chemical model.

Table 3.2 Energies (kcal/mol, B3LYP/TZV(2d,2p)) and Selected Geometric Parameters (Å) for the substrate complexes. His_{277} was assumed to be in a protonated form.

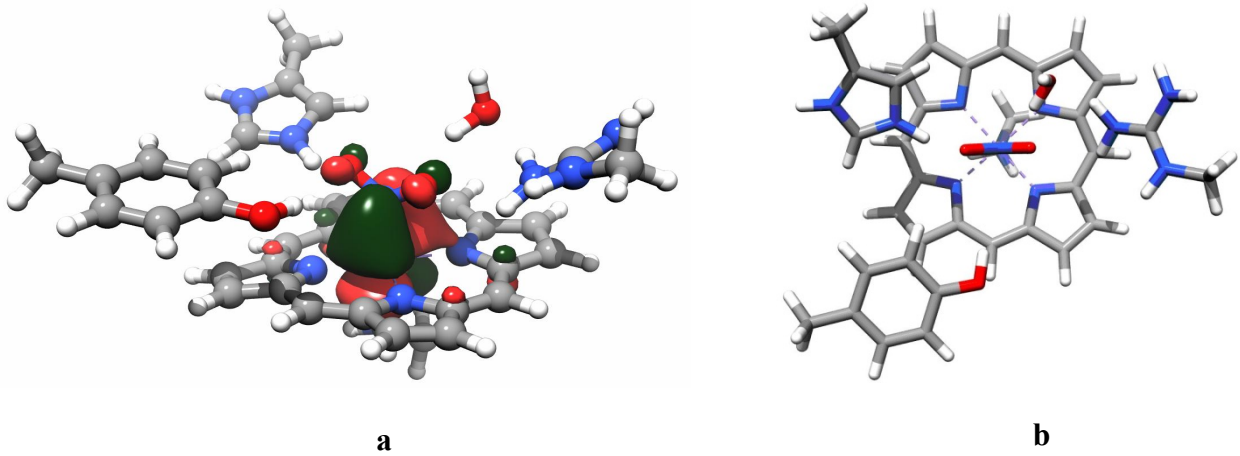
Complex	Spin State	Relative Energy	Fe – O(NO_2) or Fe – N(NO_2) distance	Fe – N(porph.) distance	E(O-binding)-E(N-binding)
Fe^{2+} - ONO^-	S=0	0	2.052	1.995, 2.002, 2.013, 1.996	4.3
	S=1	unstable	-	-	-
	S=2	unstable	-	-	-
Fe^{3+} - ONO^-	S=1/2	0	2.016	2.009, 1.996, 1.997, 2.014	4.1
	S=3/2	0.6(2.5)*	2.597	1.994, 2.002, 2.009, 1.998	1.8
	S=5/2	5.5(10.4)	2.340	2.077, 2.053, 2.039, 2.070	0.4
Fe^{2+} - NO_2^-	S=0	0	1.877	2.009, 2.003, 2.012, 1.996	-
	S=1	unstable	-	-	-
	S=2	4.1(10.9)	2.342	2.072, 2.070,	-

				2.054, 2.073	
	S=1/2	0	1.942	1.999, 2.018, 2.012, 1.983	-
Fe^{3+} - NO_2^-	S=3/2	2.8(6.0)	2.398	1.997, 2.006, 2.017, 1.996	-
	S=5/2	9.2(15.0)	2.353	2.056, 2.056, 2.058, 2.055	-

the energies in braces are calculated with B3LYP functional.

Substrate Back-Bonding in ferrous-nitrite complex

The N-bound ferrous-nitrite complex is the energetically most favorable state. Furthermore, it exhibits the shortest Fe-N(NO_2) bond distance (1.877 Å). As discussed at length previously,[40] the short distance suggests a pronounced back-bonding effect, which reveals itself in the interaction of the low lying d_{xz} orbital of iron and the LUMO orbital of the NO_2^- molecule (Figure 3.3, (a)). The LUMO has π-antibonding character with respect to the N–O bond. Hence, receiving electron density from the heme-iron to its π-antibonding orbital makes the N–O interaction weaker, with a long N–O bond and a decreased bond order. As pointed out by Einsle *et al.*,[40] this interaction is of key importance for the initial activation of the N–O bond.



a **b**

Figure 3.3 (a) Backbonding interaction of nitrite molecule with ferrous iron. The d_{xz}-orbital of iron interacts with the LUMO orbital of the NO_2^- molecule, which has antibonding character with respect to the N–O bond. (b) The spatial orientation of the NO_2^- substrate molecule. NO_2^- nearly bisects the N(porph) – Fe – N(porph) angle. This gives the strongest interaction with the iron d_{xz} back-bonding orbital, which is perpendicular to the plane of the NO_2^- molecule.

The ferrous-nitrite back-bonding has been observed and clearly discussed for a variety of model complexes.[181,182,56] It was emphasized that not only shortening of the Fe–N(NO_2^-) is diagnostic of back-bonding but also the spatial orientation of the nitrite relative to the porphyrin plane.[56] This orientation was studied by means of Mössbauer and EPR spectroscopy.[183] Both methods suggest an orientation where the plane of the π-interacting metal orbital (either d_{xz} or d_{yz}) is perpendicular to the nitrite plane such that the π-acceptor capabilities of the N-bound nitrite are maximized (Figure 3.3, (b)). The same orientation is also found in our calculations. It should be noted that nature located His_{277} in the best possible place to maximize the back-bonding interaction and at the same

time to keep the distance to His_{277} short thus facilitating the donation of protons (Figure 3.4). Table 3.3 demonstrates how the active site residues assist in redistributing the charge in the nitrite substrate molecule and thereby contribute to its activation. Compared to the gas phase, the oxygen atoms gain additional negative charge and hence become susceptible to electrophilic attack (Figure 3.4). Moreover, both N–O bonds become longer and their bond orders decrease significantly, thus indicating bond weakening (Table 3.1).

Table 3.3 Atomic charges (Mulliken analysis), bond lengths (Å) and bond orders of the NO_2^- substrate molecule in gas phase and in a bound to ferrous center form.

	Gas phase NO_2^-			ferrous-nitrite complex in C*c*NiR		
	Oxygen 1*	Nitrogen	Oxygen 2	Oxygen 1	Nitrogen	Oxygen 2
Atomic charges	-0,48	-0,03	-0,48	-0,49	+0.51	-0,51
	Ox. 2 – Nit. 1		Nit. 1 – Ox. 3	Ox. 2 – Nit. 1		Nit. 1 – Ox. 3
Bond lengths	1,278		1,278	1,291		1,269
Bond orders	1,48		1,48	1,23		1,27

*Oxygen 1 forms a hydrogen bond with His_{277}; oxygen 2 hydrogen bonds to Arg_{114}.

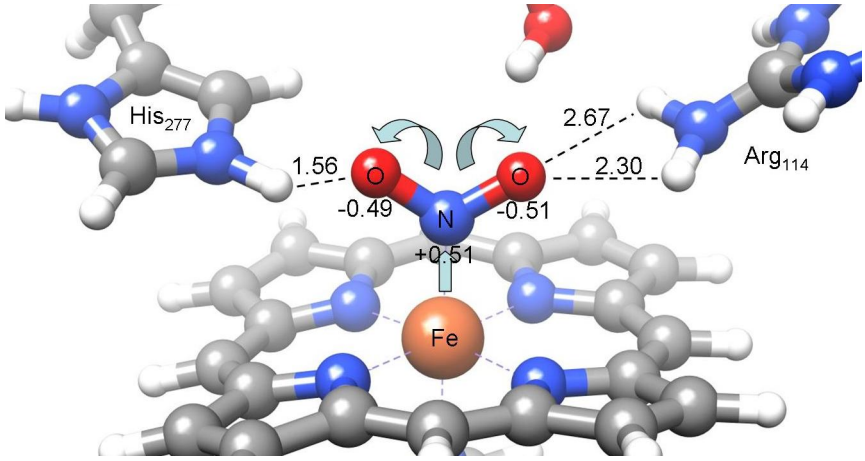


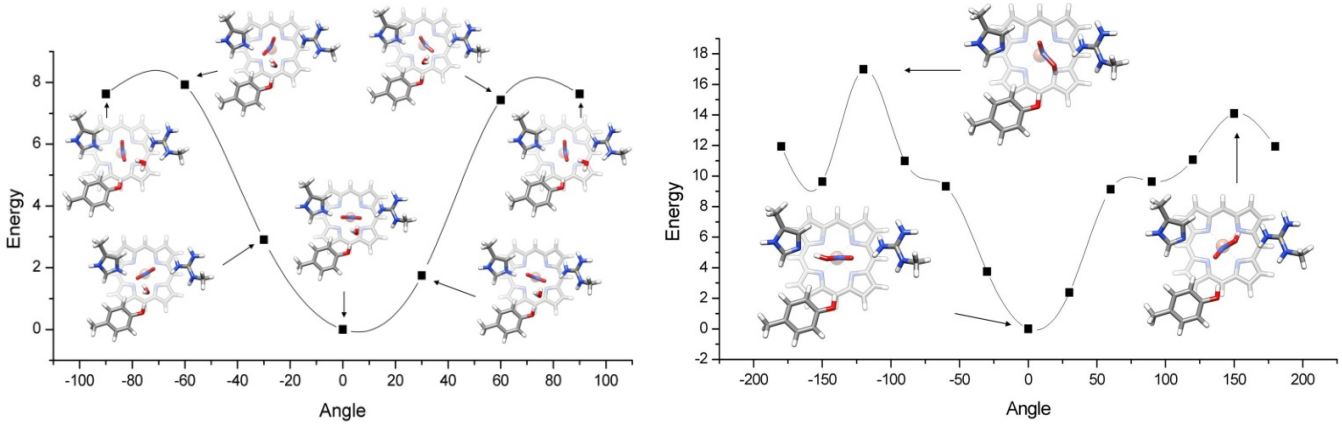
Figure 3.4 The charge redistribution in NO_2^- molecule upon binding to the active site. The oxygen atoms of the substrate bear negative charge and hence become susceptible to the electrophilic attack by His_{277} and Arg_{114}.

3.1.5 The N–O bond cleavage.

Possible pathways for the first N–O bond cleavage based on protonated His_{277}

Initially, the first N–O bond cleavage in the presence of protonated His_{277} was considered. Starting from complex A1, the hydrogen transfer reaction leads to three possible products (species B1– B3). The second proton transfer takes place either in the "recharged" complex C2 or directly in B2 where Arg_{114} serves as a proton donor. As a consequence of the second proton transfer, the N–O bond is cleaved, water is released and the $\{Fe(NO)\}^6$ intermediate is formed (D2 and D3). (Figure 3.2).

Before modeling the proton transfer reactions, the conformational flexibility of the NO_2^- adduct was studied. Interestingly, for the TvNiR[39] enzyme a conformation where the substrate molecule forms direct hydrogen bonds to Tyr_{218} and His_{277} rather than to Arg_{114} and His_{277} was proposed. In contrast to the proposed TvNiR conformation, a deep energy minimum along the NO_2^- rotation coordinate, which corresponds to the conformation found in NrfAs, was observed in the calculations (Figure 3.5, (a)). The origin of this deep minimum is the substantial back-bonding between the nitrite and the Fe-porphyrin system in conjunction with the hydrogen bonds to neighboring groups. Hence, the lowest energy conformer is the one in which the NO_2^- molecule nearly bisects the N(porph) – Fe – N(porph) angle. The conformation also gives the strongest interaction with the iron d_{xz} back-bonding orbital, which is in this case exactly perpendicular to the plane of the NO_2^- molecule. Moreover, the two positively charged side chains His_{277} and Arg_{114} establish strong hydrogen bonds with NO_2^-. The lowest energy conformer was taken as initial structure for modeling of the proton transfer reaction 'A1' in Figure 3.2.



a **b**

Figure 3.5 The conformational potential energy surface of the NO_2^- (a) and HONO (b) adducts (relaxed surface scan in 30° steps using B3LYP/TZVP(2d,2p)). The rotation around Fe – N(NO_2^-) and Fe – N(HONO) was studied.

Taking into account the assumption that His_{277} is in the doubly protonated form, the first step of the reaction can be accomplished in three possible ways, since each of the side chains, His_{277}, Arg_{114} or Tyr_{218}, could act as proton donor. In the calculations, the transfer A1→B3, where Tyr_{218} is the proton donor, leads to an unstable product while stable HONO adducts are formed when His_{277} or Arg_{114} acts as a proton donor. However, the thermodynamic driving forces for the two feasible protonation reactions differ significantly. Reaction A1→B1 is predicted to be uphill by prohibitive 17,7 kcal/mol, while A1→B2 is calculated uphill by only +4,9 kcal/mol and is hence more

favorable. The calculated transition state energy for the proton transfer reaction from His_{277} was calculated to be +5,5 kcal/mol. This indicates that the reaction is certainly kinetically feasible but that the equilibrium is on the educt side. However, since, the protonated product is rapidly consumed the reaction still moves in forward direction.

The second proton could either be transferred from Arg_{114} (B2→D2 and B2→C1), Tyr_{218} (B2→D1) or from recharged His_{277} (B2→C2→D3). All scenarios were investigated computationally. The calculations reveal that a hypothetical HONOH adduct C1 would be highly unstable. It spontaneously dissociates upon optimization to the HONO adduct and a proton, which moves onto one of the donor side chains.

The possibility of His_{277} side chain to be recharged with a new proton was considered. Comparing the proton affinities of the complex with a histidine in solution, the reprotonation energy was estimated to be +9,6 kcal/mol. This suggests that His_{277} remains in the deprotonated state unless a significantly endothermic reprotonation step takes place.

As the reprotonation is not a favorable process, Arg_{114} and Tyr_{218} were investigated as possible proton donors. However, in order to obtain a favorable spatial orientation for the interaction of the nitrite molecule with these side chains, HONO must first rotate around the N–Fe bond. Thus, the energetic profile around the rotation coordinate was computationally investigated through a relaxed surface scan (steps of 30° were performed and the remaining degrees of freedom were optimized in each step). The resulting potential energy surface is shown in Figure 3.5, (b).

According to our calculations, the most stable conformer is the one where the hydroxyl group is pointing out to His_{277} and the second oxygen of HONO forms a hydrogen bond with Arg_{114}. Approximately 8 kcal/mol are needed to force the hydroxyl group to interact with Tyr_{218} or Arg_{114}. Consequently, this increases the transition state energy of the second proton abstract reaction. However, the conformer where the -OH group of the nitrite substrate interacts with Tyr_{218} can be significantly stabilized if only the hydrogen at δ-nitrogen of His_{277} moves to the ε-nitrogen. In this case only 4.2 kcal/mol are needed to create such a conformer and therefore, the possibility of the proton transfer from Tyr_{218} was considered computationally. The calculated driving force of the reaction B2→D1 is +24,5 kcal/mol. A transition state was located and has energy of +37.5 kcal/mol, which makes the transfer from Tyr_{218} thermodynamically and kinetically forbidden under physiological conditions. Interestingly, for the proton transfer from Tyr_{218} the empirical van der Waals correction has the large impact (3.9 kcal/mol) thus resulting in a reduced reaction barrier of 33.6 kcal/mol. Nevertheless, this is barrier is still too high to represent a realistic pathway.

There is one more possible way of transferring the second proton onto the nitrite substrate, namely the B2→D2 proton transfer from Arg_{114}. The energetics of this proton transfer reaction were calculated to be +7,7 kcal/mol with a corresponding transition state energy of +24,4 kcal/mol.

Thus, the computational results indicate that the second proton transfer is a rather energy demanding process. Moreover, the calculated activation barriers are too high, since the experimental kinetic studies show that the barrier of the rate-limiting step should not exceed 15,2 kcal/mol.[38,38,52] Most likely, the endothermic recharging of His_{277} B2→C2 could still be more favorable. Keeping this in mind, the complex with recharged His_{277} was optimized. It was found that after His_{277} protonation, the proton transfer reaction C2→D3 can proceed without any

additional activation barrier and immediately yields the {Fe(NO)}^{6} intermediate and a water upon optimization. Thus, enough energy must be given to the B2 complex in order to create the unstable protonated complex C2, which will pass its proton barrierless to the HONO adduct to form a water and {Fe(NO)}^{6}. One way to estimate this energy is to create a conformer of the B2 complex, which will be recharged with a proton to form C2. The proton reference energy is taken from cluster-ion solvation data.[149] The calculations predict that an energy of +4,2 kcal/mol is necessary to rotate the HONO molecule towards Tyr_{218} in order to form a stable conformer and facilitate recharging of His_{277}. The protonation energy of His_{277} using the reference energy adds another +7,9 kcal/mol to the energetic demand of this process. Taken together, the calculations predict an energy demand of +12,1 kcal/mol to form complex C2, which renders this scenario the most plausible one of the ones investigated (Figure 3.6).

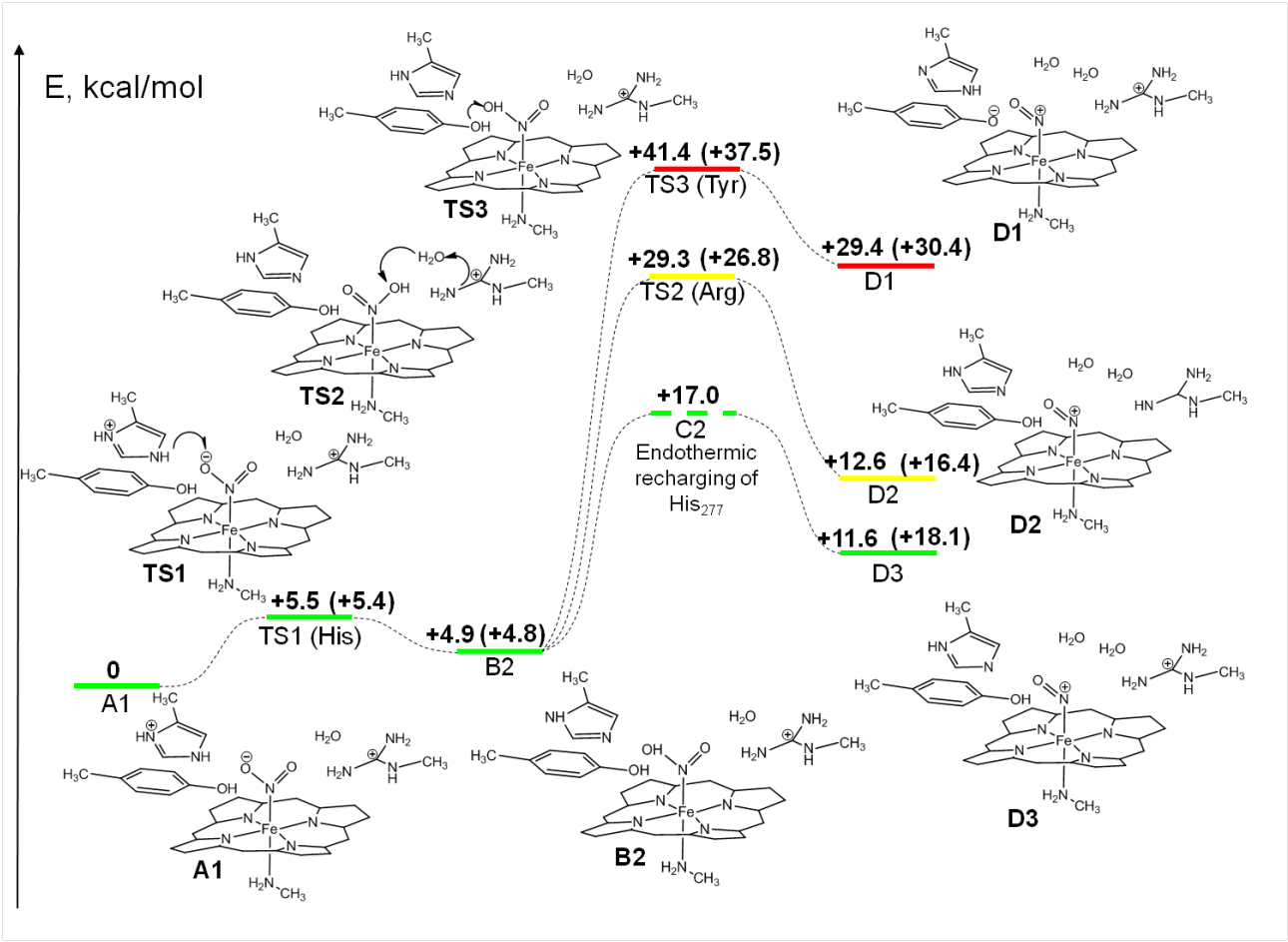


Figure 3.6 Scheme of the first N–O cleavage (Histidine is protonated, B3LYP/TZV(2d,2p), values in parentheses contain a semi-empirical van der Waals correction). The first protonation is accomplished by His_{277}, leading to the HONO adduct. The later has three possible ways of receiving the second proton. The transfers from Tyr_{218} and Arg_{114} have prohibitively high barriers. Thus, the most feasible pathway is the endothermic recharging of His_{277}, which leads to a barrierless reaction with the HONO adduct thus forming a water and {Fe(NO)}^{6}.

Possible pathways for the first N–O bond cleavage based on deprotonated His_{277}

Keeping in mind the possibility that His_{277} might also be in the deprotonated neutral state, another possible pathway for the first proton transfer was considered. In this case, the proton donor was considered to be Arg_{114}.

Besides the transfer from Tyr_{218} (A4→B5) and from neutral His_{277} (A4→B4), who were both found to be thermodynamically highly unfavorable, there are two possibilities for a proton to be abstracted by the substrate molecule in reaction A4→B6. The first is the direct proton transfer from Arg_{114} and the second is via an intermediate water molecule that acts as proton relay. An energetic estimate of both pathways was performed and both transition states were found. According to the calculations, the direct transfer reaction is marginally more favorable (by ~0.3 kcal/mol), which is negligible within the uncertainties of the method. Hence, both routes may be feasible. The calculated activation energy of reaction A4→B6 is +16,7 kcal/mol. This value is only slightly higher than that obtained by kinetic investigations[38,52] and therefore the possibility that more than one residue may act as a proton donor in the active site of C*c*NiR appears to be realistic.

In this alternative pathway, the HONO adduct would rearrange to the lowest energy conformer. From calculations, it was determined that the conformer which has the OH group of the HONO adduct pointing towards Tyr_{218} is lowest in energy. However, the energy difference is not as large as predicted for protonated His_{277}. In addition, there appears to be no significant rotational barrier for the HONO molecule. Hence, rotation will be essentially free and the second proton could come either from Tyr_{218} or Arg_{114}. In contrast to the case with protonated His_{277}, the HONO adduct has no significant charge-charge interactions since Arg_{114} and His_{277} are not charged (compare complexes B3 and B6). It should also be mentioned that the HONO adduct can receive the second proton also from Arg_{114}.

The first possibility to be considered is the transfer of the second proton from Tyr_{218} in the reaction B6→D4. The calculated reaction enthalpy equals +19.2 kcal/mol with a transition state energy of +31.1 kcal/mol. Hence, this pathway is energetically not feasible. In the case of reprotonated Arg_{114} the energy of reaction C3→D1 is almost equal (+20.2 kcal/mol) and the barrier is even higher (+32.1 kcal/mol). Although the van der Waals correction provides a significant decrease in the reaction energy (-7.9 kcal/mol), both scenarios in which Tyr_{218} acts as a proton donor are not energetically feasible. Moreover, the recharging energy of Arg_{114}, using the proton reference energy from cluster-ion solvation data,[149] was estimated to be -13 kcal/mol, which makes reaction C3→D2 the most plausible pathway for the second proton transfer. Indeed, the calculated barrier of +17.2 kcal/mol (14.1 kcal/mol with van der Waals correction) is in a good agreement with what is inferred from kinetic experimental studies.[38,52] It should be noted that the experimentally determined rate refers to the overall reduction process. Our interpretation obviously assumes that the N–O bond cleavage is the rate limiting step in both cases - proton transfer from His_{277} at low pH or alternatively from Arg_{114} at pH > 8. The energy profile for the first N–O cleavage based on deprotonated His_{277} is summarized in Figure 3.7.

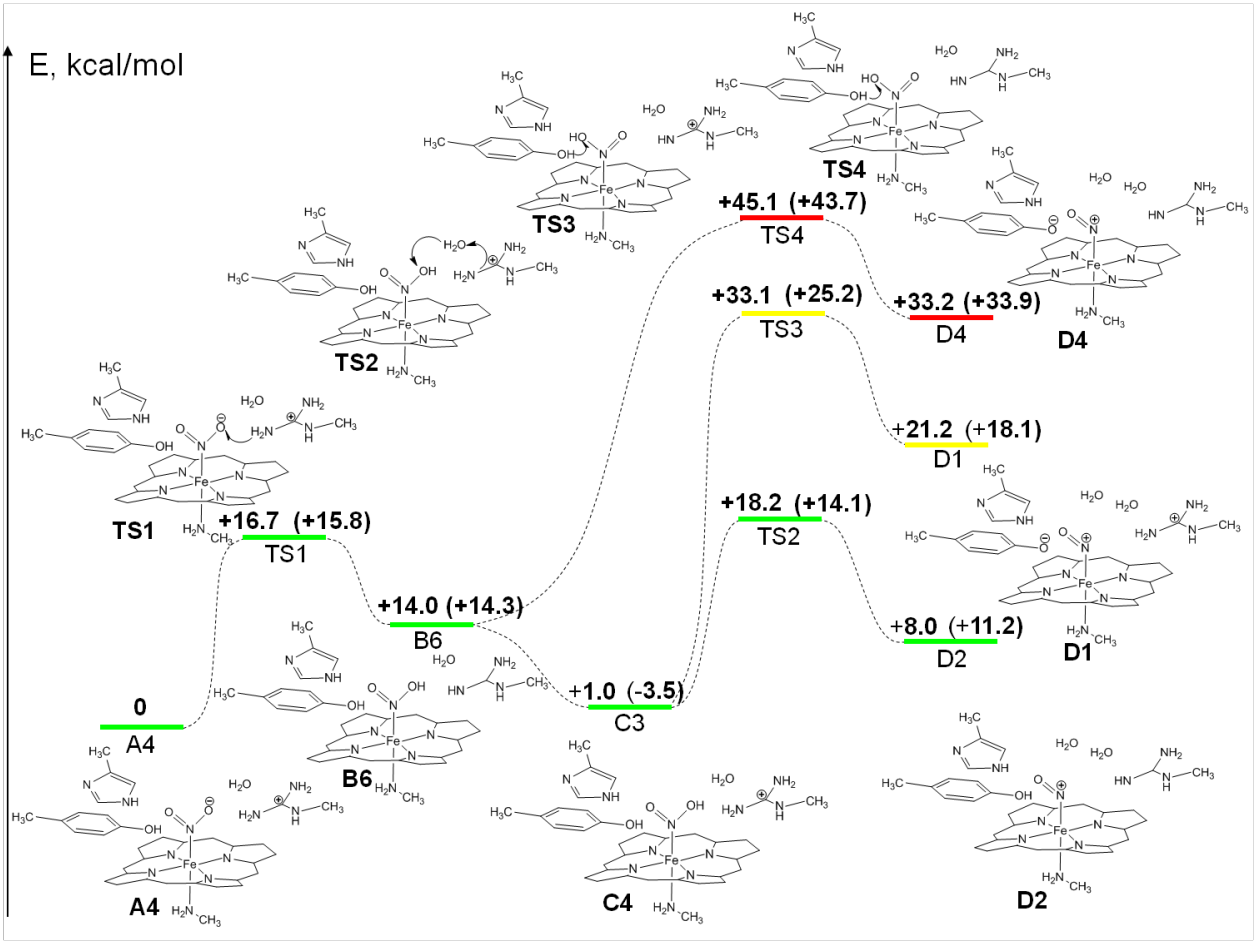


Figure 3.7 Potential energy surface of first N–O bond cleavage reaction. In the case of neutral His_{277}, proton donation can be adopted by Arg_{114}. Proton donation by Tyr_{218} is thermodynamically and kinetically impossible. The values in parentheses contain an empirical van der Waals correction.

Summary of N–O bond cleavage

Based on our calculations, that are consistent with previous results,[40] the cleavage of the first N–O bond is heterolytic. This conclusion can be drawn from the analysis of atomic partial charges of the HONO adduct and the cleavage transition state electronic structure. The hydroxyl oxygen of the HONO fragment carries a negative charge of -0.72 (Mulliken population analysis), whereas the partial charge of the second oxygen is only -0.2; the nitrogen is positively charged (+0.40). Hence, the N–OH bond in the HONO adduct is polarized and weakened by the interaction with the active site iron and thereby can be readily cleaved in the presence of proton donors. In case of the more acidic His_{277}, the endothermic formation of the protonated form directly leads to N–O bond cleavage and no additional activation energy is needed to initiate the process. However, proton donation from Arg_{114} for the first and second proton transfer has almost the same activation energy (+16.7 and +17.2 kcal/mol respectively). The energy of 17,2 kcal/mol is needed almost entirely for the H – N(Arg_{114}) bond breaking process; the N–OH bond breaking features an almost vanishing barrier. Analogous conclusions are drawn from the analysis of the transition state complex geometry and the molecular orbitals. The N–OH bond length is 2.21 Å, which shows that the bond

is actually cleaved in the transition state. The proton donated from Arg_{114} side chain is interacting with the oxygen lone pair.

3.1.6 Conclusions for the nitrite binding and activation

The influence of the second coordination sphere in the active site of C*c*NiR has been assessed. Several charge-, spin- and protonation states were investigated in order to elucidate the most likely series of events leading to the formation of the $\{Fe(NO)\}^6$ adduct that accompanies the cleavage of the first N–O bond. Consistent with earlier results, it is concluded that the reaction proceeds on the low-spin surface and must originate from the ferrous nitrite adduct.

Two plausible scenarios were calculated that are both compatible with the experimentally determined rate (suggesting an effective barrier of 15.9 kcal/mol).[38,52] The two pathways differ in the assumed protonation state of His_{277}. The reaction proceeds relatively smoothly if one assumes that this residue is doubly protonated and can be recharged during the reaction cycle. Alternatively, or perhaps at higher pH, Arg_{114} may serve as proton donor in a pathway that is energetically similarly feasible as the one proceeding via His_{277}.

There are two major effects which facilitate substrate activation in the Fe(II) – nitrite complex: a) back-bonding and b) interaction of nitrite with second sphere residues. As a consequence of the back-bonding effect, the NO_2^- based LUMO antibonding orbital becomes populated and hence the N–O bond is weakened. Moreover, the interaction of the nitrite oxygens with charged side chains causes an asymmetric charge distribution which breaks the equivalence of the N–O bonds (both with a formal bond order of 1.5) and hence activates the more single-bond like N–O bond for reductive cleavage. This was already hypothesized in the earlier study[40] and could now be substantiated using larger computational models of the active site.

In addition, it was found here that the conformation with which the NO_2^- molecule bonds to the active site is optimal for the proton attack from His_{277} and simultaneously for strong back-bonding. Thus, the most acidic side chain in the active site of the C*c*NiR, His_{277}, is located optimally for carrying out protonation steps. Both features are reflected in the only moderately endothermic reaction energy of +4,9 kcal/mol and the calculated barrier of 5,5 kcal/mol for proton transfer. The formed HONO adduct represents the lowest conformer with respect to HO-N=O rotation. Since the rotational potential energy surface is rather rigid, the only feasible way for the transfer of the second proton is the recharging of His_{277} in this pathway. This can be achieved in a reasonably endothermic step and is rendered irreversible by a barrierless reaction leading to N–O bond cleavage and formation of the $\{Fe(NO)\}^6$ species. Ultimately, the source of protons in the active site is an unresolved question. The Ca-site could play a significant role not only as a structural motif but also in shuttling protons to the active site. These issues deserve further investigation.

Alternatively, or at higher pH, another pathway is considered to be feasible. Above pH ~8 His_{277} is probably in its deprotonated form. Hence, the proton donor functionality must be adopted by Arg_{114}. Our calculations suggest that this is energetically and kinetically feasible.

Although the acidity of Tyr_{218} is certainly much lower than that of histidine, it was considered as a proton donor in each step of the reaction. All calculated reaction and activation energies are very much higher than the ones found for His_{277} and Arg_{114}. Hence, we exclude a significant role of this residue in the initial N–O bond cleavage chemistry. Thus, the role of Tyr_{218} in the reaction cycle of C*c*NiR remains to be resolved. A radical step involving hydrogen atom transfer appears to be a plausible alternative. However, neither experimentally not theoretically has evidence for such a possibility been obtained so far. The present results, however, implicate that Tyr_{218} is ineffective as proton donor.

3.2 Reductive activation of the heme-iron-nitrosyl

3.2.1 Abstract

Cytochrome *c* nitrite reductase catalyzes the six-electron, seven-proton reduction of nitrite to ammonia without release of any detectable reaction intermediate. This implies a unique flexibility of the active site combined with a finely tuned proton and electron delivery system. In the present work, we employed density functional theory (DFT) to study the recharging of the active site with protons and electrons through the series of reaction intermediates based on nitrogen monoxide (Fe(II)-NO^+, Fe(II)-$NO^\bullet$, Fe(II)-NO^- and Fe(II)-HNO). The activation barriers for the various proton and electron transfer steps were estimated in the framework of Marcus theory. Using the obtained barriers, the kinetics of the reduction process was simulated. It was found that the complex recharging process can be accomplished in two possible ways: either through two consecutive proton coupled electron transfers (PCET) or in the form of three consecutive elementary steps involving reduction, PCET and protonation. Kinetic simulations revealed the recharging through two PCETs to be a means of overcoming the predicted deep energetic minimum that is calculated to occur at the stage of the Fe(II)-$NO^\bullet$ intermediate. The radical transfer role for the active site Tyr_{218}, as proposed in the literature, cannot be confirmed on the basis of our calculations. The role of the highly conserved calcium located in the direct proximity to the active site in proton delivery has also been studied. It was found to play an important role in the substrate conversion through the facilitation of the proton transfer steps.

3.2.2 Computational details.

Computational model

Initial structures for the enzyme active site models were taken from the protein data bank[6] (PDB code 2E80[40]). The model used in the theoretical investigation consists of His_{277}, Arg_{114} and Tyr_{218} as the first sphere distal ligands. These side chains were truncated at the α-carbon atoms. The heme is included in the model without its substituents. Lys_{134}, the proximal ligand of iron, was approximated by methylamine. The model also includes two water molecules (Figure 3.8).

Constraints were imposed on the initial structures in order to mimic the steric influence of the protein backbone on the active site. Thereby, the fragment constraints type was chosen. The side chains Tyr_{218}, His_{277} and Arg_{114} form three fragments connected to the active site that itself was defined to consist of four fragments: the heme group, the substrate molecule, the axial Lys_{134} and two additional water molecules. All fragments are fully optimized but the mutual relative orientation of distal side chains in space is kept fixed to prevent unreasonable movement of the side chains that would not be possible in the full protein environment. The performance and errors introduced by this type of constraints have been studied in our previous investigation.

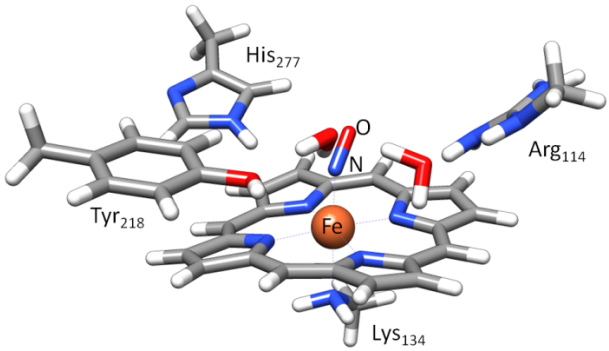


Figure 3.8 Computational model for the C*c*NiR active site. The heme was modeled without its ring substituents; side chains His_{277}, Arg_{114}, Tyr_{218} and Lys_{134} (taken from pdb structure 2E80) were truncated at α-carbon atoms;

Computational details

The calculations were performed with the ORCA electronic structure package version 2.7[117] using density functional theory (DFT).[22] The models of the enzyme active site were optimized with the BP86 functional[162,163] together with the resolution-of-the-identity (RI) approximation (in the Split-RI-J variant). [164] A triple-ζ basis set (TZVP)[165] together with a matching auxiliary basis set (TZV/J[107,108] in ORCA notation) was applied for all atoms. Subsequent numerical frequency calculations were performed for all models for the sake of verifying structures found as local minima and obtaining thermochemical corrections. Frequencies were computed through two-sided numerical differentiation of analytic gradients using increments of 0.002 Bohr.

Single point energy calculations were performed on the optimized structures with the B3LYP functional[152,153] that is thought to provide higher accuracy than BP86 for reaction and activation energies.[21] The 'chain-of-spheres' (RIJCOSX) approximation[109] was invoked in these calculations. The basis set was of triple-ζ quality with two sets of first polarization functions on all atoms (TZV(2d, 2p) [167] in ORCA notation). In order to crudely model the long-range dielectric effects of the protein, the COSMO[23] solvation model was employed with a dielectric constant of 4.9. All other COSMO parameters were left at their default values. Additionally, Grimme's DFT-D2 semi-empirical correction was employed to account for dispersion forces as implemented in the ORCA package.[168,169]

The model structures and orbitals were visualized with the Chimera program.[172] The spin-restricted Kohn-Sham occupied orbitals were localized according to Pipek-Mezey (PM) algorithm. The unoccupied orbitals are not localized. The spin-unrestricted Kohn-Sham orbitals were transformed into quasi-restricted orbitals[170] (QRO) and then localized with PM algorithm.[171] Simulation of the kinetics was conducted using Tenua program,[146] based on KINSIM by Barshop, Wrenn and Frieden.[147]

3.2.3 Results and discussion

Geometries

The set of optimized complexes is shown in Figure 3.9. Starting from complex C, the first step is either protonation or reduction leading to the formation of complexes CH and CE, where H stands for an additional proton and E stands for an electron. Proceeding further along these lines in a systematic fashion, possible oxidation states $\{Fe(NO)\}^6$-$\{Fe(NO)\}^7$ with different protonation states were taken into consideration. The complexes in the last row in Figure 3.9 have protonated Fe-HNO cores and represent a probable protonation state of the substrate. The electronic structure of each of these species was then analyzed in order to deduce how the protonation state of the surrounding ligands influences the electronic structure of the Fe-NO/Fe-HNO core.

The first set of optimized structures (complexes C – CH3) have analogous $\{Fe(NO)\}^6$ electronic cores and only differ in the overall active site protonation state. In the case of complex C, all side chains are in their deprotonated states, whereas in complex CH Tyr_{218} is in its protonated neutral form. Obviously, the thermodynamics of protonation and deprotonation is governed by the relative acidities of the side chains. To estimate relative acidities the PROPKA3.1 program was used.[184,185,186]

Our calculations showed that the most basic residue in the active site for all resolved structures is Tyr_{218} with estimated pKa 14-15. For Arg_{114} the predicted pKa is 12-13 and the pKa of His_{277} was predicted to be significantly shifted to the very acidic region of 1-2. Accordingly to these predictions, the sequence of the protonation is following: the first proton goes to Tyr_{218}, the second to Arg_{114} and the third one to His_{277}. Additionally, as the predicted difference in the pKa of Tyr_{218} and Arg_{114} isn't large, the formation of the CH_Arg complex, where the first additional proton is placed on the Arg_{114} residue, was also considered (Table 3.4).

The reduction of the $\{Fe(NO)\}^6$ complex leads to the formation of the paramagnetic $\{Fe(NO)\}^7$ species, which is also considered in four different protonation states: CE, CEH, CEH2 and CEH3 (second row in Figure 3.9). Further reduction leads to the $\{Fe(NO)\}^8$ complex and the corresponding protonated states are denoted as CEE, CEEH, CEEH2 and CEEH3 (the third row in Figure 3.9). Additionally, the formation of the HNO containing adducts is considered with the CEH_HNO, CEH2_HNO, CEEH_HNO and CEEH2_HNO complexes.

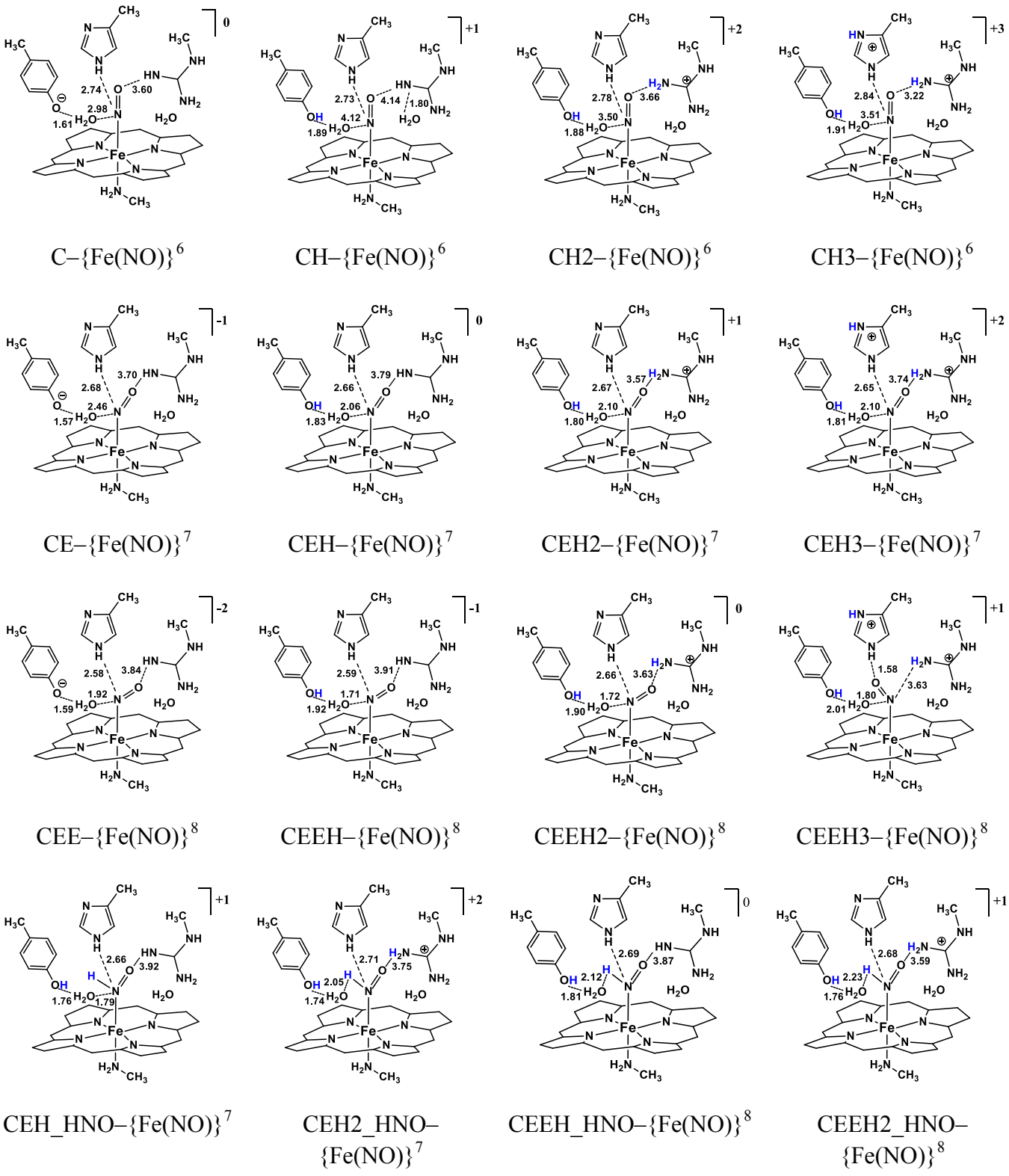


Figure 3.9 The optimized structures of the $\{Fe(NO)\}^6$-$\{Fe(NO)\}^8$ and $\{Fe(HNO)\}^7$-$\{Fe(HNO)\}^8$ complexes with different protonation states of the active site side chains. The extra protons are highlighted with blue color.

The calculated and available experimental structural parameters are in good agreement. For instance, the Fe-N(NO and trans-Lys_{134}) distances are within 0.04 Å of the crystallographically determined values. Interestingly, the largest deviation is observed for the C and CH_Arg

complexes. In general, introduction of a positive charge into the system leads to a shortening of the metal-ligand bonds. However, the CH_Arg complex differs insignificantly from complex C. As the Arg_{114} – NO distance is rather large (Figure 3.9), the protonated Arg_{114} has less impact on the metal-ligand distances or on the N-O bond. Significant elongation of the latter bond is observed, however, for complex CH_Arg. In addition, the linear arrangement of the Fe–N–O unit, which is usually observed in $\{Fe(NO)\}^6$ model hemes and hemoproteins, is well reproduced in models CH_Tyr, CH2 and CH3 while the models C and CH_Arg have slightly different bent conformations. The bending mode is, of course, well known to be particularly soft.[187]

Upon reduction to the $\{Fe(NO)\}^7$ form, the Fe-NO bond is calculated to lengthen by about 0.09 Å is calculated. This Fe-NO distance is similar for all protonation states. At the same time the N-O bond is elongated in some cases by up to 1.202 Å while Fe-N(Lys_{134}) may be shortened significantly by 0.056 Å while moving from the CE to CEH3 form. The Fe-N-O angle is bent with an angle of approximately 138° calculated for all protonated states of the $\{Fe(NO)\}^7$ complex, which is consistent with what is usually observed in such systems.[187] The protonated forms CEH_HNO and CEH2_HNO are characterized by significant stretching of the Fe-NO bond (by 0.08 Å) and the N-O bond (by 0.03 Å), although the Fe-N(Lys_{134}) distance is shortened. The Fe-N-O angle is calculated to be 129°.

A further reduction leads to the $\{Fe(NO)\}^8$ electronic configuration. This reduction is accompanied by a further lengthening of the Fe-NO bond by 0.06 Å and the N-O bond by approximately 0.04 Å compared to the corresponding $\{Fe(NO)\}^7$ form (complexes CE, CEH, CEH2 and CEH3). The bending of the NO unit to an angle of 123° for $\{Fe(NO)\}^8$ (Table 3.4) is consistent with earlier studies on related systems.[187]

Table 3.4 Selected geometric parameters (distances in Å and angles in degrees) for the nitrosyl complexes investigated in this work. $\{Fe(NO)\}^6$ complexes with different protonation state are denoted as C, CH, CH2, CH3; $\{Fe(NO)\}^7$ complexes as CE, CEH, CEH2, CEH3 and $\{Fe(NO)\}^8$ complexes as CEE, CEEH, CEEH2, CEEH3. The protonated $\{FeHNO\}^7$ form is denoted as CEH_HNO and CEH2_HNO and $\{FeHNO\}^8$ form as CEEH_HNO and CEEH2_HNO. The two conformers CH_Tyr and CH_Arg differ in the position of the proton (on Tyr_{218} or on Arg_{114} respectively).

Complex	*Fe-NO*	*N-O*	*FeNO angle*	*Fe-N(lys)*	*Fe-N(porph)*	*N(p.)-Fe-NO dihedral*
C	1.697	1.163	153	2.124	2.022, 2.025, 2.010, 2.007	24
CH_Tyr	1.651	1.147	178	2.056	2.021, 2.018 2.016, 2.013	52
CH_Arg	1.698	1.171	152	2.076	2.020, 2.019, 2.017, 2.010	36
CH2	1.648	1.149	176	2.054	2.026, 2.023, 2.011, 2.010	17
CH3	1.648	1.151	173	2.053	2.026, 2.025, 2.010, 2.010	-12
CE	1.742	1.192	138	2.229	2.029, 2.029, 2.005, 2.003	42
CEH	1.745	1.197	138	2.186	2.030, 2.027, 2.002, 2.002	43
CEH2	1.748	1.199	138	2.177	2.032, 2.022, 2.005, 2.002	42
CEH3	1.746	1.202	138	2.173	2.028, 2.024, 2.007, 2.002	48
CEE	1.802	1.232	122	2.327	2.025, 2.022, 1.990, 1.993	43
CEEH	1.799	1.235	123	2.301	2.029, 2.026, 1.991, 1.987	45
CEEH2	1.804	1.238	122	2.284	2.029, 2.019, 1.991, 1.988	49
CEEH3	1.779	1.261	123	2.264	2.026, 2.025, 2.004, 1.989	123

CEH_HNO	1.821	1.232	129	2.116	2.016, 2.012, 2.005, 1.982	52
CEH2_HNO	1.830	1.235	129	2.109	2.017, 2.013, 1.995, 1.995	54
CEEH_HNO	1.788	1.249	131	2.126	2.003, 2.005, 2.014, 2.019	65
CEEH2_HNO	1.790	1.249	130	2.120	1.999, 2.026, 2.016, 1.998	48

The final set of values represent the dihedral angle between the Fe-N-O plane and the plane formed by N(NO)-Fe-N(porph). The rotation of the NO fragment above the porphyrin plane in the $\{Fe(NO)\}^6$ complex is essentially independent of the protonation states of the surrounding side chains as the Fe-N-O unit adopts an almost linear conformation. Contrary to the $\{Fe(NO)\}^6$ case, the conformational behavior of the NO ligand in $\{Fe(NO)\}^7$ is of particular interest since the orientation of the NO guides and supports further protonation. According to our results, the most stable conformation is obtained when the N-O axis of the NO ligand points in the direction of the meso-carbon of the porphyrine ring. This conformation was also reported as an energy minimum by Patchkovskii *et al*.[188] although the potential surface calculated by these authors is very flat. Also, similar potential energy surface was obtained in a recent investigation of Sundararajan *et al.*[189] To study this issue a relaxed surface scan calculation has been carried out. Indeed, the highest calculated barrier was found to be 3 kcal/mol (the graph on energy-angle dependence for the N(porph)-Fe-N(NO) unit in CEH2 complex is presented in the supporting material).

The $\{Fe(NO)\}^8$ adducts show a similar trend. As expected, the dihedral angle remains close to 45° except for the CEEH3 complex, where it flips to the value of 123°. The conformation of the CEEH3 complex is essentially very similar: the N-O axis also points towards the meso-carbon, but in the direction of the protonated His_{277}. Obviously, the flip in the NO orientation is driven by an additional strong hydrogen-bond formed with H-N(His_{277}) (Table 3.4). These findings can be rationalized through closer examination of the electronic structure and bonding in the $\{Fe(NO)\}^6$-$\{Fe(NO)\}^8$ series.

Electronic structure, bonding

The electronic structure changes along the $\{Fe(NO)\}^6$-$\{Fe(NO)\}^8$ series were extensively discussed previously[190] and are consistent with a wide range of experimental results.[191] Hence, this section will concentrate on the electronic structure changes brought about by the second sphere residues.

The electronic structure of the low-spin $\{Fe(NO)\}^6$ complexes is best described as a Fe(II) center bound to a closed shell NO^+ ligand.[54] This is consistent with the presence of a set of three iron-based doubly occupied t_{2g}-derived orbitals. The set includes an essentially nonbonding d_{xy} based orbital and two π-bonding orbitals formed from the Fe-$d_{xz,yz}$ and NO-π^* fragment orbitals (Figure 3.10, (a)). The first two low-lying unoccupied orbitals have mainly NO-π^* character and are strongly antibonding with respect to the Fe-$d_{xz,yz}$ orbitals. This picture nicely rationalizes the observed changes of the geometry and charge within the Fe-NO unit. As was discussed above, the C and CH_Arg complexes are characterized by elongated Fe-N(NO) and N-O distances. Moreover, a distortion of the Fe-N-O angle is observed. This behavior can be traced back to the interaction of the oxygen(Tyr_{218})-based lone pair with the NO-π^* based unoccupied MOs. As the NO-π^* based MO has antibonding character with respect to both the Fe-N(NO) and the N-O bonds, electron

donation into these orbitals results in their elongation. In addition, since one of the quasi-degenerate NO-π^* based orbitals preferentially obtains electron density, an increased bending of the Fe-N-O unit is expected. The polarization of one of the NO-π^* based MOs is rather significant since $\{Fe(NO)\}^6$ complexes are known to be strong electrophiles.[56]

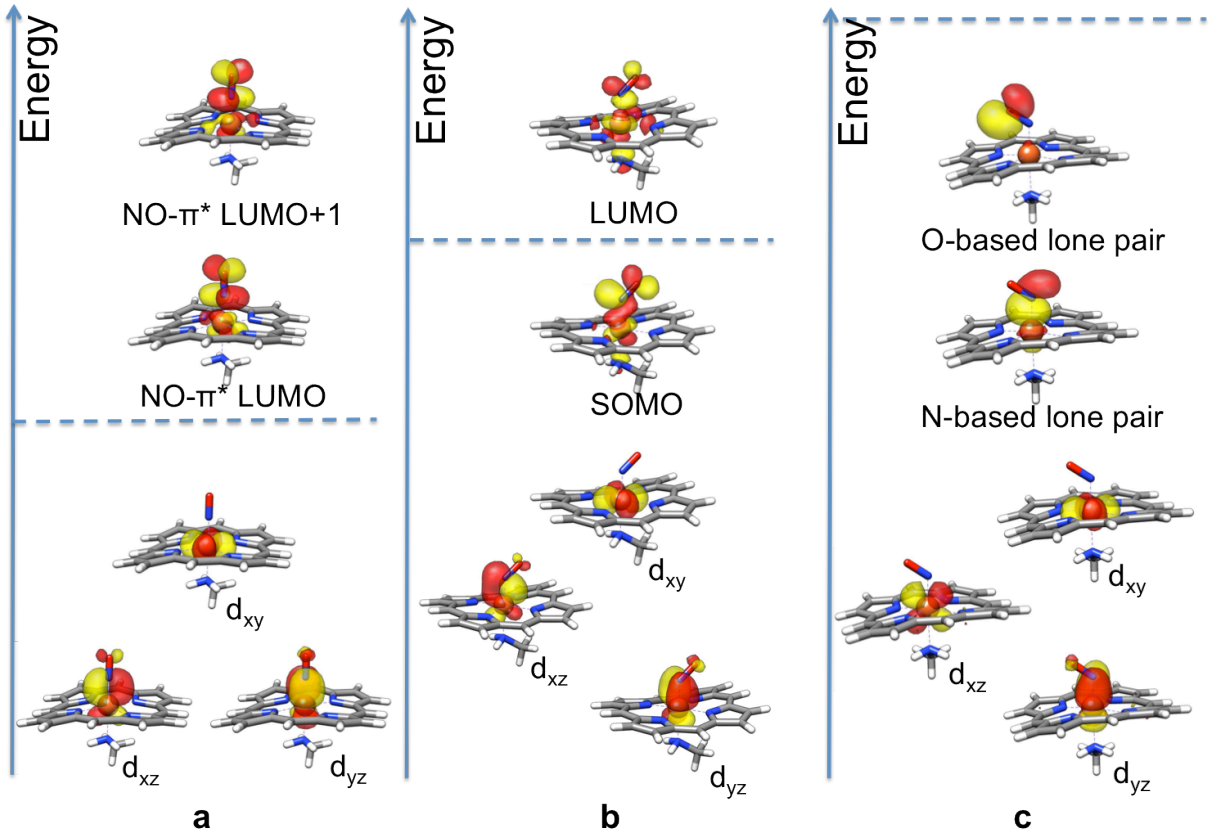


Figure 3.10 Contour plots of important MOs calculated under B3LYP/TZV(2d,2p) level of theory. Active site side chains are not depicted for simplicity. (a) The $\{Fe(NO)\}^6$ MOs are presented. The occupied orbitals were localized according to Pipek-Mezey (PM) algorithm. The unoccupied orbitals are not localized. (b) The $\{Fe(NO)\}^7$ MOs are depicted. The spin-unrestricted Kohn-Sham orbitals were transformed into quasi-restricted orbitals (QRO) and then localized with PM algorithm. (c) The $\{Fe(NO)\}^8$ MOs are considered.

Upon reduction to $\{Fe(NO)\}^7$, the unpaired electron is placed into one of the strongly antibonding Fe-NO orbitals with mainly NO-π^* character. The Jahn-Teller effect then leads to a pronounced bending of the Fe-NO unit. In the $\{Fe(NO)\}^7$ complex the two NO-π^* orbitals are split into an MO with σ-bonding-character with respect to the Fe – N(NO) bond (π^*_σ SOMO orbital) and a π-antibonding MO (π^*_π LUMO orbital).(Figure 3.10, (b)) The unpaired electron occupies the π^*_σ orbital. Occupation of the SOMO which has σ-bonding-character with respect to the Fe – N(NO) bond greatly contributes to the observed stability of the Fe-NO bond in $\{Fe(NO)\}^7$ complex.[191] It should be noted that the composition of this orbital is highly sensitive to the theoretical method used and has been the subject of much debate in the literature.[192] In this specific case, pure functionals (such as BP86) are known to give generally better results than hybrid functionals.[189]

Since the SOMO is of antibonding character with respect to the N-O bond of the NO ligand, it causes the weakening and elongation of this bond (Table 3.4). The Fe-N(NO) bond is also elongated because of the less interaction of the metal d_{xz} and NO-π* orbitals in the tilted Fe-N-O core of $\{Fe(NO)\}^7$ in comparison to the almost linear Fe-N-O core in $\{Fe(NO)\}^6$. The pronounced ligand-centered character of the reduction can also be deduced from the analysis of the Mulliken atomic charges (Table 3.5). The charge of the central iron atom changes by only 0.01 e^- upon reduction, while the charge on the NO fragment becomes more negative by 0.25 e^-. For the CE-CEH3 series the gradual increase of the Fe-N(NO) and decrease of Fe-N(Lys_{134}) distances is a consequence of the increasingly positive charge of the proximal second sphere residues. This positive charge attracts electron density and hence causes electron flow from the iron-porphyrin fragment to the NO-ligand.

Further reduction of the $\{Fe(NO)\}^7$ complex leads to the formation of the $\{Fe(NO)\}^8$ complex. Again, one of the t_{2g}-derived orbitals is essentially nonbonding while the other two have π-bonding character. Though the interaction of the metal d_{xz} and NO-π* orbitals becomes even smaller than in $\{Fe(NO)\}^7$ complex (the MO contains 93% of metal character). The additional electron enters the π^*_σ orbital. Consequently, the Fe-NO bond is further elongated although not to the same extent as for the first reduction step. Double occupation of π^*_σ orbital leads to the splitting of one of the N-O bonds and formation of two separate lone pairs on N and O (Figure 3.10, (c)). The lone pair on N is involved in σ-bonding with the formally unoccupied d_z^2 orbital of iron. The formation of the lone pair at the NO-nitrogen is of fundamental importance for further reaction, since it serves as a site of electrophilic attack. Also, as the electron density from π^*_σ orbital is delocalized over the entire Fe-N-O core in the $\{Fe(NO)\}^8$ complex the N-O bond becomes stronger. The strengthening of the N-O bond was also deduced in the investigation of Lehnert *et al.*[193] based on the DFT calculations and observed strong shift of the N-O stretching frequency down to 1500cm^{-1} (in comparison to 1290 cm^{-1} for free NO^-). Based on the calculation and experimental data available the electron density distribution in the $\{Fe(NO)\}^8$ complex can be described as a mixture of Fe(II)-NO^- and Fe(I)-NO descriptions. An interesting effect was observed for the CEEH3 complex. As mentioned above, the NO ligand rotates to form a hydrogen bond with the protonated His_{277} in this complex (the distance is 1.58 Å). This interaction leads to charge redistribution in the Fe-NO unit. The positive charge on iron increases and the partial charge of the N(NO) becomes significantly more negative, thus indicating that electron density is transferred to the NO ligand. At the same time, the charge on the O(NO) becomes more positive which is a consequence of hydrogen bond formation. Xu *et al.*[194] have argued that H-bonding to O pulls electron density into the NO π*-orbital, thus strengthening the Fe-N bond while weakening the N-O bond. However, hydrogen bonding to N withdraws bonding electrons from both the Fe-N and the N-O bonds into a (partial sp^2) N based nonbonding orbital. Hence, both bonds are weakened by this interaction. This trend is confirmed by our investigation for complexes CEEH2 – CEEH3. Upon the protonation of His_{277} in CEEH2 the NO ligand rotates and O(NO) forms a strong hydrogen bond with His_{277}. As a consequence the Fe-N(NO) bond is shortened whereas the N-O bond is elongated, indicating the strengthening of the Fe-N bond and the weakening the N-O bond (Table 3.4 and Figure 3.9).

The electronic structure of the CEH_HNO and CEH2_HNO ($\{FeHNO\}^7$) complexes resembles the electronic structure of the $\{Fe(NO)\}^8$ series rather than $\{Fe(NO)\}^7$. Separate lone pairs on the NO nitrogen and oxygens are formed and the lone pair on the N(NO) is engaged in a sigma-bond with the additional hydrogen. This finding implies that the NO radical in $\{Fe(NO)\}^7$ is protonated and

the HNO which is formed oxidizes Fe(II) resulting in an Fe(III)-HNO complex. Indeed, the SOMO in the CEH_HNO and CEH2_HNO complexes is almost purely (93.3%) iron based. The electronic structure can best be described as consisting of a neutral HNO species bound to a low-spin Fe(III) center. As a consequence, the amount of back-donation drops sharply and the Fe-N(NO) distance increases while at the same time the splitting of one of the N-O bonds leads to the elongation of this bond. Further reduction leads to CEEH_HNO and CEEH2_HNO complexes. An additional electron enters an iron-based SOMO orbital. The resulting electronic configuration can be described as Fe(II)-HNO.

Table 3.5 Mulliken population analysis of the B3LYP densities. $\{Fe(NO)\}^6$ complexes with different protonation states are denoted as C, CH, CH2, CH3; $\{Fe(NO)\}^7$ complexes are denoted as CE, CEH, CEH2, CEH3 and $\{Fe(NO)\}^8$ complexes as CEE, CEEH, CEEH2, CEEH3. The $\{FeHNO\}^7$ complexes are denoted as CEH_HNO and CEH2_HNO, the $\{FeHNO\}^8$ complexes as CEEH_HNO and CEEH2_HNO. The two conformers CH_Tyr and CH_Arg differ in the position of the proton (on Tyr_{218} or on Arg_{114} correspondently).

Complex	*Fe*	*N(NO)*	*O(NO)*	*N(lys)*	*SD Fe/NO*
C	-0.20	0.32	-0.18	-0.35	
CH_Tyr	-0.16	0.35	-0.19	-0.33	
CH_Arg	-0.20	0.24	-0.21	-0.33	
CH2	-0.15	0.34	-0.20	-0.33	
CH3	-0.16	0.32	-0.20	-0.33	
CE	-0.16	0.17	-0.30	-0.42	0.32/0.67
CEH	-0.17	0.16	-0.30	-0.41	0.38/0.62
CEH2	-0.17	0.14	-0.30	-0.41	0.36/0.63
CEH3	-0.17	0.13	-0.30	-0.41	0.37/0.62
CEE	-0.16	0.03	-0.47	-0.47	
CEEH	-0.18	0.05	-0.47	-0.46	
CEEH2	-0.18	0.03	-0.47	-0.46	
CEEH3	-0.09	-0.06	-0.44	-0.46	
CEH_HNO	-0.29	0.22	-0.29	-0.36	1.24/-0.13
CEH2_HNO	-0.28	0.18	-0.28	-0.35	1.27/-0.21
CEEH_HNO	-0.24	0.14	-0.42	-0.39	
CEEH2_HNO	-0.24	0.13	-0.41	-0.39	

As described above, a gradual elongation of the Fe-N(NO) bond is calculated along the electron transfer series $\{Fe(NO)\}^6$-$\{Fe(NO)\}^8$. This was discussed in detail earlier,[190] and implies that the back-bonding in the series decreases. Nevertheless, the N(p)-Fe-N-O dihedral angle stays constant for $\{Fe(NO)\}^7$ and $\{Fe(NO)\}^8$ complexes. This can be explained by the fact that the rotational barrier of the NO ligand mostly depends on the strength and spatial position of the available hydrogen bonds rather than on the formation of directed π-bonds. In fact, in the $\{Fe(NO)\}^6$ complex the NO ligand is positively charged (+0.13 e^-) and only forms a weak hydrogen bond with His_{277} (Figure 3.9). In the reduced complexes NO is negatively charged (-0.15 and -0.42 e^- for $\{Fe(NO)\}^7$ and $\{Fe(NO)\}^8$, respectively). Multiple hydrogen bonds are formed with the surrounding amino acid side chains in both oxidation states.

Taken together, these results show that it is mandatory to model the surrounding protein ligands in order to obtain reliable results. The electronic and conformational behavior of the NO ligand is greatly affected by surrounding charged side chains. The positively charged side chains polarize the FeNO core and thereby modulate its redox potential. Obviously, the polarization effect depends on the distance. By taking the active site side chains which form hydrogen bonds to the substrate into account in the modeling one includes at least the most important sources of redox potential modulation by short-range charge interactions in the calculations. The modified redox potential of the core in turn guides further proton supply and establishes multiple hydrogen bonds that play a key role in the subsequent steps of the reaction cycle.

Calculation of the reference points, pK_a and redox potentials.

Prior to discussing the energetics of the different possible electron and proton transfer reactions, the selection of a reference point for the active site recharging with electrons and protons must be made. The electron supply is accomplished by the transport of electrons from the external electron donor (presumably from a membrane-anchored tetra-heme cytochrome (NfrH) for the *W. succinogenes* C*c*NiR) through a cascade of redox-active hemes to the active site. Hence, the energy of the incoming electrons can be estimated from the adiabatic electron affinity of the neighboring heme 3. For this purpose a model of heme 3 has been constructed (Figure 3.11).

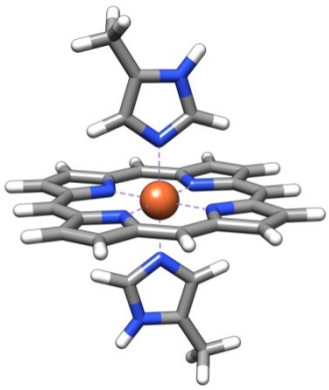


Figure 3.11 The model of the bis-histidine-ligated heme 3. The histidine side chains are approximated by 4(5)-methylimidazoles. The heme appears in the model without its ring substituents.

The role of the bis-histidine-ligated redox active hemes as electron donors has been studied both experimentally and theoretically.[195,196,197] In the studies by Smith *et al*.[196,197] the authors predicted the values of the adiabatic electron affinity to be 121 and 119 kcal/mol for the high-spin and low-spin states, respectively. According to our calculations the low-spin state is more favorable by 7.9 kcal/mol and 7.0 kcal/mol for Fe(II) and Fe(III) correspondently, in good agreement with previous calculations (Table 3.6). Moreover, the calculated geometric parameters are also consistent with those of Smith *et al*. However, the adiabatic electron affinities predicted by our calculations are noticeably smaller and amount to 94 and 93 kcal/mol for low-spin and high-spin hemes,

respectively. A possible source of discrepancy in the calculated absolute values could come from the COSMO and van der Waals corrections applied in this investigation. In our calculations, when the van der Waals correction was not applied the heme affinities were 102 and 100 kcal/mol for low-spin and high-spin hemes, respectively. Without the COSMO correction, the adiabatic electron affinities increase to up to 120 kcal/mol, consistent with Smith *et al*. Despite these discrepancies, the reference point for the electron reduction of the active site was set at 94 kcal/mol for our calculations. The reason for this choice is the effective error cancellation that takes place due to the fact that the reduction of the active site is calculated as the difference between the electron affinity of the active site and the (adiabatic) ionization potential of the electron donor, both of which are calculated at the same level of theory. Additionally, it is possible to test the calculated reference values proposed by Siegbahn *et al.*[198] using the following approach: The standard redox potential (E^0) of the redox pair NO_2^-/NH_4^+ is +0.34 V. Assuming NrfB as redox partner (E^0=-0.42 V), one readily calculates that the reduction of NO_2^- to ammonia is exergonic by 119 kcal/mol. Using adiabatic electron affinity of 94 kcal/mol and the proton absolute hydration enthalpy from cluster-ion solvation (-275 kcal/mol)[149] as reference values, one calculates the reduction energetics of the reaction $NO_2^- + 6e^- + 8H^+ = NH_4^+ + 2H_2O$ to be 122 kcal/mol. This value is in reasonable agreement with the value calculated from electrochemically measured redox potentials.

Table 3.6 The calculated selected geometric parameters (Å) and relative energies of high- and low-spin states (ΔE in kcal/mol) for the bis-histidine ligated heme model.

Complex	ΔE	Fe-N(His)	Fe-N(Porph)
Fe(II) S=0	0	2.003/2.003	2.007
Fe(II) S=2	7.9	2.298/2.298	2.084
Fe(III) S=1/2	0	1.988/2.003	2.005
Fe(III) S=5/2	7.0	2.230/2.235	2.071

Proton supply is accomplished through the transfer of protons from the positively charged inlet channel identified by Einsle *et al.*[40] Assuming the proton donor to be a water molecule (or a corresponding charged hydronium ion) from the substrate channel, one can take the proton absolute hydration enthalpy from cluster-ion solvation data as a reference point.[149] This choice is justified by the fact that the inlet channel is filled with water which is involved in multiple hydrogen bonding interactions. Hence, the dissociation in the inlet channel and in bulk solution should be comparable.

Having defined the energetic reference point, we are now in a position to study the recharging of the active site i.e. consecutive proton and electron addition to the active site of the protein. The calculated reaction scheme is summarized in Figure 3.12. In addition to the calculated free energy differences (protonation energies, electron affinities and bond formation energies), calculated pK_a values and redox potentials are also given. The values of pK_a and redox potentials were estimated from the calculated free energy differences ΔG according to the standard equations presented in chapter 2.2.3.

Despite difficulties in computing accurate numbers for these quantities from first principles, we are confident that the calculations presented here correctly represent the trends. In particular, as we compare species of very similar chemical structure error compensation is expected to be effective.

Some trust in the presented results can be gained by calculating the redox potential for the bis-histidine ligated heme-iron and the pK_a's of the active site side chains in solution. Bis-histidine ligated heme-iron complexes are well studied and the measured redox potential for biomimetic bis-histidine mesoporphyrin IX complex is -0.28 V.[199] Our calculations predict a redox potential of -0.34 V, in a good agreement with the experimental value.

The Born-Haber cycle based approach that we employed for pK_a calculations also gives reasonable results in test calculations. For example, the pK_a values for tyrosine, arginine and histidine in solution were estimated to be 8.2, 12.1, and 3.2 respectively, which is in qualitative agreement with standard values of 10.3, 12.3 and 6.5 for these amino acid side chains.[151] Taken together, while we do not claim to have achieved high accuracy in the calculation of redox potentials and pK_a values we expect the trends in these numbers to be correctly captured by the calculations as we proceed through the catalytic cycle and will go on to discuss our results below.

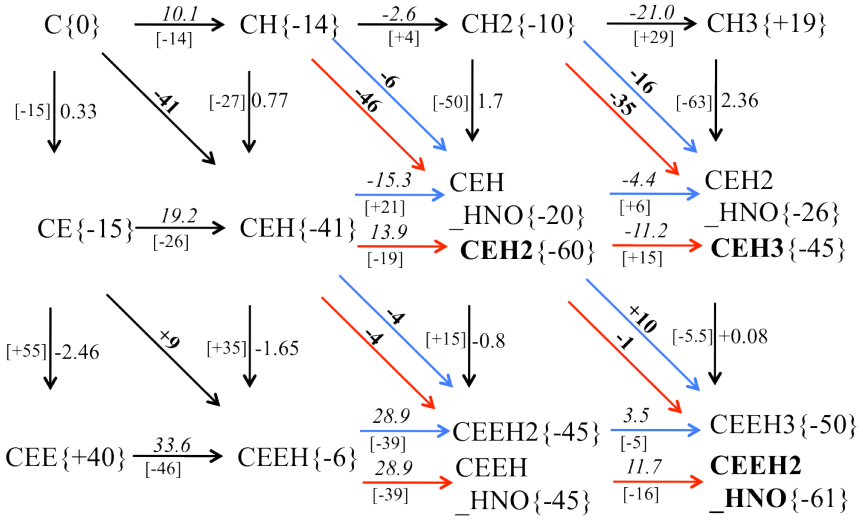


Figure 3.12 Overall thermodynamic scheme of the recharging process. There are free energy changes (ΔG in kcal/mol) in the square brackets and relative energies with respect to C complex in curved brackets (all energies are B3LYP-D2/TZV(2d,2p). Calculated pK_a values (in *italics*) for the protonation reactions are located above horizontal arrows. The redox potentials (in V) for reduction reactions are placed on the right of the vertical reaction arrows. Formation free energy changes (in kcal/mol) for PCET reactions are printed boldface. Red arrows denote reactions where a proton move to a side chain while blue arrows highlight reactions when a proton is transferred to the NO ligand. In case two complexes have the same number of protons, the complex that has a lower relative energy is highlighted in bold and is taken as starting point for the consecutive reactions.

Reaction energetics

The data presented in Figure 3.12 span a wide range of pK_a values. However, only some of the values should be taken into consideration under biologically relevant conditions. For example, for Tyr_{218} there are three protonation reactions: C→CH, CE→CEH and CEE→CEEH. Only the first

two are relevant since the CEE adduct is essentially impossible to reach. Hence, the pK_a for Tyr_{218} is conservatively estimated to lie between 10 and 19. For Arg_{114} the only energetically possible protonation is CEH→CEH2 and the estimated pK_a for Arg_{114} is 13.9. His_{277} can be protonated in the reaction CEEH→CEEH3 having an estimated pK_a of 3.5.

In our previous investigation of the initial reaction steps in C*c*NiR, it was shown that the N–O bond cleavage could be carried out in two possible ways.[200] The most probable proton donor is His_{277}, although the proton donation function can be adopted by Arg_{114} for higher pH. According to these results, the starting point for this investigation is either the complex with deprotonated Arg_{114} and His_{277} (CH complex), or the one with protonated Arg_{114} and deprotonated His_{277} (CH2 complex). The fully deprotonated complex C, which was taken into consideration for consistency, was revealed to be unstable as it readily abstracts a proton (the ΔG of the reaction is -13.8 kcal/mol) and forms the CH complex. The calculated pK_a value of 10.1 is in good agreement with the experimental value for Tyrosine. Moreover, PROPKA3.1 predicts even larger pKa values of 14-15 for Tyr_{218}, and implies that Tyr_{218} is likely to be protonated. A probable role of the protonated Tyr_{218} may be in facilitating the next reduction step (compare the calculated redox potentials of C and CH) by maintaining a higher redox potential of the active site. Alternatively, the first proton could go to Arg_{114} forming CH_Arg. Interestingly, the protonation of Arg_{114} is 7 kcal/mol less favorable than protonation of Tyr_{218} although the acidity of Tyrosine is much higher than that of Arginine.

Analyzing the results in Figure 3.12, it is clear that all complexes of the first row can easily oxidize any of the neighboring *c*-type hemes, whose measured redox potentials range from +22 to -275 mV.[48] However, only the protonation of C to CH is calculated to be relevant under physiological conditions. Indeed, the formation of doubly and triply protonated complexes CH2 and CH3 is endothermic and can only be accomplished under pH conditions that are unbearable for living organisms. Thus, remaining possibilities must be considered: electron uptake CH→CEH and PCET reactions CH→CEH2, CH→CEH_HNO.

The CH complex is a relatively strong oxidant (+0.77 V). Hence, the reduction CH→CEH is likely. After reduction, the CEH complex can be readily further protonated to form CEH2. At the same time, a large energy gain is estimated for the PCET reaction CH→CEH2 (-46.2 kcal/mol). Hence, a competition between two conceivable pathways might be anticipated. Alternative PCET CH→CEH_HNO where both particles, a proton and electron, are transferred to the Fe-NO core is energetically less favorable (-6 kcal/mol).

The PCET reaction CH→CEH2, leads to the formation of a very stable intermediate. According to our results and consistent with previous studies, the CEH2 complex ($\{Fe(NO)\}^7$ core with protonated Tyr_{218} and Arg_{114}) can be a kinetic trap in the reduction process.[40] The only thermodynamically possible reaction starting from CEH2 is PCET to CEEH2_HNO. In case of this almost thermonuetral PCET the final product of $\{FeHNO\}^8$ electronic configuration is formed. Thus the complete recharging of the active site through two consecutive PCET reactions seems to be a feasible pathway.

Alternatively, a possible way to overcome the kinetic trap would be an intramolecular radical transfer reaction from Tyr_{218} to the Fe-NO core in CEH2 (Figure 3.13, (a)). Intramolecular

hydrogen transfer from Arg_{114} in CEH2 was also considered (Figure 3.13, (b)). As a product of the intromolecular proton transfer from Arg_{114} a CEH_HNO complex is formed.

It was found that, in contrast to radical reactions involving Tyrosine side chains,[30] our calculations predict that such a reaction is thermodynamically (reaction energy +22.3 kcal/mol) and kinetically (transition state barrier +85 kcal/mol) impossible. Also, intramolecular proton transfer from Arg_{114} was also calculated to be impossible (reaction energy of +39,2 kcal/mol).

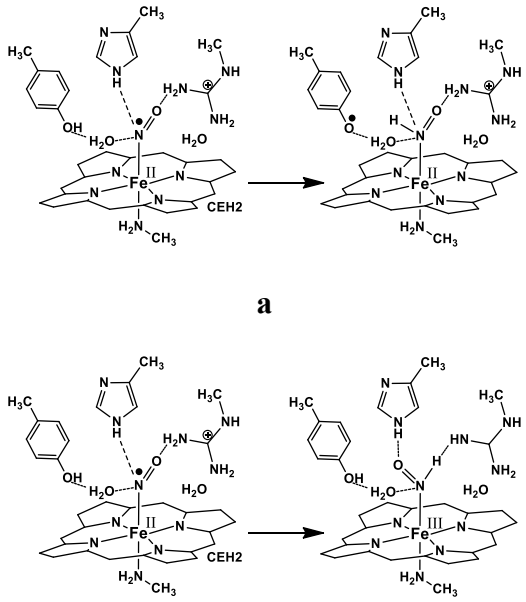


Figure 3.13 Possible reactions to overcome the kinetic trap (CEH2 complex). Intramolecular radical transfer from Tyr_{218} (a) and intramolecular proton transfer from Arg_{114} (b) in complex CEH2.

A closer look at the thermodynamic scheme reveals one more probable pathway for recharging of the active site: the starting complex CH could be reduced (CH→CEH) before a PCET reaction takes place (CH→CEH2). Reduced CEH complex turns over to CEEH2 or CEEH_HNO through PCET and a further protonation would lead to the final product complex CEEH2_HNO. Also CEEH3 complex can be formed by protonation from CEEH2. In this pathway CH→CEH→CEEH2/CEEH_HNO→CEEH2_HNO/CEEH3 all reactions are exothermic and thus are competitive to the recharging scheme CH→CEH2→CEEH2_HNO, where the latter step is almost thermoneutral.

Finally, in case the CEEH3 complex is formed, another proton is required to form the HNO adduct. The $\{Fe(NO)\}^8$ adduct is a very strong nucleophile. Our calculations show large negative reactions energies for proton transfer from either His_{277} or Arg_{114}. The calculated free energy differences are -19,3 and -13,7 kcal/mol, respectively. In fact, $\{Fe(NO)\}^8$ could even abstract a proton from Tyr_{218} in an almost thermoneutral reaction.

All transition state energies corresponding to the aforementioned reactions were located. Transferring a proton from Tyr_{218} has a very high barrier of +65 kcal/mol. The two alternative

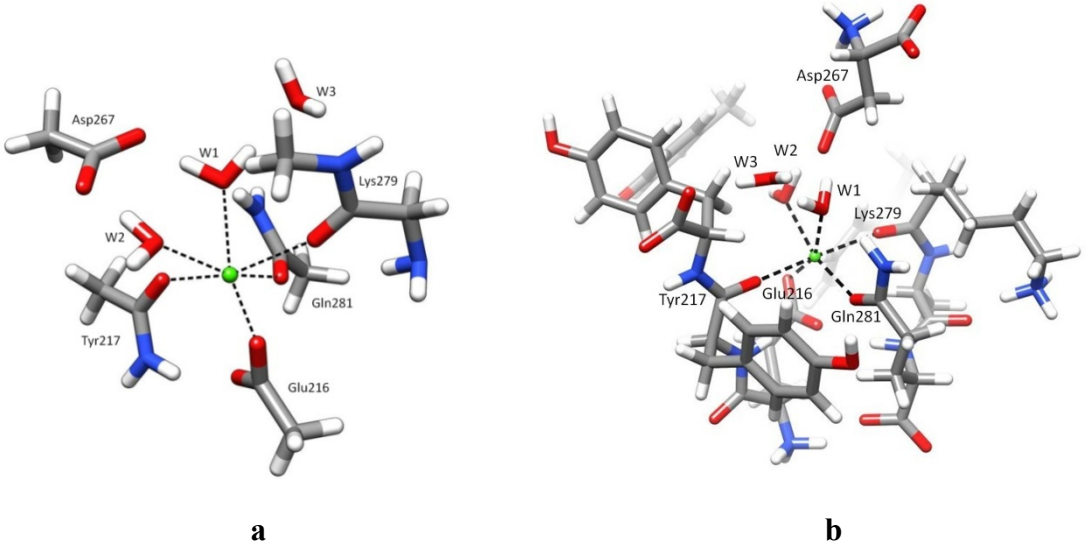
reactions have almost identical barriers: protonation from His_{277} has a barrier of 5,2 kcal/mol while the protonation involving Arg_{114} is slightly lower (4,4 kcal/mol). Both processes are considered as possible pathways for HNO formation.

In summary, the thermodynamic analysis revealed two possible pathways. In contrast to the previously assumed reaction scheme,[40] the two electron reduction of the $\{Fe(NO)\}^6$ intermediate is predicted to be unfavorable. Instead, there are two possible reaction pathways; 1) involving two PCET reactions (CH→CEH2→CEEH2_HNO) or 2) involving reduction (CH→CEH) followed by PCET (CEH→CEEH2/CEEH_HNO) and protonation (CEEH2→CEEH2_HNO/CEEH3). The $\{Fe(NO)\}^8$ CEEH3 complex is predicted to readily react with His_{277} or Arg_{114} to form the $\{FeHNO\}^8$ product. The alternative second pathway involves only exothermic steps with small driving forces, which is beneficial in biochemical reactions. Possibly, the competition between the two pathways is governed by kinetic factors.

Modeling of the proton supply

The modeling is based on the plausible assumption that the proton is supplied from the positively charged inlet channel. The Ca-site is a part of the inlet channel and is situated in direct proximity of the active site. As was discussed previously,[32] the Ca-site may play an important role as a structural element. In addition, the Ca^{2+} dication obviously carries two positive charges as well as two water molecules in its coordination sphere making an impact into the maintaining of the positive charge and structural stability of the inlet channel. This also implies that the Ca-site may modulate the proton supply process by acidifying the coordinated water molecules. To further study this issue two models of the inlet channel including the Ca-site were constructed (Figure 3.14). The first model is comprised of only the first coordination sphere of the Ca^{2+} ion. The Lys_{274} and Tyr_{218} amino acids coordinate the Ca^{2+} ion through backbone oxygen atoms, whereas the amino acids Glu_{217} and Gln_{276} coordinate through their carboxyl acid groups. Asp_{262} was included to neutralize the positive charge of the central atom. Three water molecules, two of which have direct bonds with Ca^{2+}, were added to complete the coordination sphere and form multiple hydrogen bonds with the surrounding amino acids. It is important to note that the charge of the system is kept at zero in the initial complex in order to study the adiabatic proton affinity (Table 3.7). The larger model includes side chains within 6 Å of the central Ca cation (Figure 3.14). Besides the first coordination sphere ligands, three tyrosine amino acids Tyr_{255}, Tyr_{219} and Tyr_{218} together with their side chains were included. As was discussed by Einsle,[30] the tyrosine side chains may play an important role in dealing with the stabilization of radical intermediates of the reduction cycle. In addition the Trp_{391}, Leu_{273} and Ala_{275} residues were also included in order to guarantee that the hydrogen bonding network with the Ca ligands is complete. The goal of these calculations was to obtain a reliable reference energy for the adiabatic proton affinities for the proton transfer from the inlet channel into the active site and to estimate the reorganization energy in the reactions involving proton transfer. A comparison of the adiabatic proton affinities of the Ca-site to those of the active site, allows one to judge if a proton can be supplied to the substrate from the inlet channel or not. Two possibilities were examined: 1) the transfer of a proton localized at the Ca-site (an extra proton was added to coodinated water forming H_3O^+) and 2) the dissociation reaction of one water molecule in the coordination sphere of the Ca ion. In the first case the ΔH of the reaction Ca-site + H^+ = Ca-siteH^+

gives the adiabatic proton affinity of the inlet channel, which can be compared to the experimentally available proton solvation energy. According to our calculation an extra proton is localized on the Asp_{262} side chain in both models and form two hydrogen bonds with two coordinated water molecules. In the second case one of the coordinated waters was ionized and the ΔH of the water dissociation reaction $H_2O = H^+ + OH^-$ was studied.



a **b**

Figure 3.14 Optimized structures of the Ca-site models representing part of the inlet channel. (a) The model comprises first sphere ligands: Asp_{262}, Tyr_{218}, Glu_{217}, Gln_{276}, Lys_{274} (the numeration of the side chains is taken from pdb structure 2E80) together with 3 water molecules. Two of water molecules are in the distorted octahydral coordination sphere of Ca atom. The total number of atoms is 56. (b) The model comprises residues within 6 Å from the central atom. 170 atoms are included in the model. In both models W1 stands for the water molecule which is protonated/ionized in the investigation.

According to our calculations the small model can reproduce the solvation energy of a proton reasonably well (Table 3.7). The calculated adiabatic proton affinity of -280 kcal/mol is in reasonable agreement with the value deduced from cluster-ion solvation data (-275 kcal/mol). Moreover, the proton affinity for the large model is -269 kcal/mol, which is slightly less than the predicted value of -275 kcal/mol. Conceptually, the proton localized in the inlet channel on Ca-site can be transferred to the active site with lower energy expenditure than a proton from bulk solution. A differential effect of ~6 kcal/mol might be assigned to the role of the Ca-site in the proton supply to the C*c*NiR active site. Thus, the effect of the Ca-site is to render all stages where a proton transfer is involved more favorable by this amount.

Additionally, although almost no changes in the ΔH of water dissociation are observed in moving from the small to the large model of the Ca-site, it is worthwhile to point out that both values are significantly smaller than the corresponding ΔH of a pure water dissociation calculated at the same level of theory. Consequently, the Ca-site might not only act as an effective proton mediator but also as an acid/based catalytic subunit, which can efficiently produce protons from coordinated water molecules and supply the protons directly to the active site. The distances between the active site and Ca-site assume only 2 to 3 water molecules as a bridge to the active site residues, thus a proton can go in single step.

Table 3.7 Calculated adiabatic proton affinities (APA in kcal/mol) and ΔH (kcal/mol) of the water dissociation reaction in the presence of the Ca-site.

Model	APA	ΔH
Small model in neutral form $+H^+$	-280	
Small model in neutral form $-H^+$		+317
Large model in neutral form $+H^+$	-269	
Large model in neutral form $-H^+$		+320

Kinetic barrier calculation

All kinetic barrier calculation were carried out according to formulas described in section 2.2.2.

Using an optical dielectric constant of 2, the outer sphere reorganization energy for ET was estimated to be 11 kcal/mol. The calculated values for the inner sphere reorganization energy lie between 16 and 37 kcal/mol, which is in good agreement with what was measured for the self-exchange reorganization energy of cytochrome *c*.[201] The calculated values suggest adiabatic barriers ranging from an effectively barrierless reaction CH2→CEH2 to barriers as high as 58 kcal/mol for the hypothetically possible reduction of CE to CEE (Table 3.8).

Using the same electrostatic continuum model, the outer sphere reorganization energies for PCET and PT can be estimated (27 kcal/mol and 12 kcal/mol were obtained for PCET and PT respectively). The calculated data is summarized in Table 3.8.

The BP86 functional is known to occasionally describe metal-ligand bonds incorrectly which could result in errors in the calculated force constants. Comparing our results with the experimentally adjusted force field of Lehnert *et al.*[193] deviations of up to 1 mdyn/Å are observed. Though the force constants from[193] are more accurate for the species for which such information is available, it is impossible to obtain experimentally adjusted values for all oxidation/protonation states considered in this work. Nevertheless, changes of the force constants of up to 1 mdyn/Å cause only minor changes on the order of 0.05 kcal/mol in the estimated free energy barriers. The underlying reason is that the free energy barrier in the Marcus theory depends linearly on the reorganization energy and quadratically on the reaction free energy (eq. 10).

Table 3.8 The estimated values for reorganization energies (kcal/mol) using a combination of electrostatic continuum modeling and force constant analysis. The corresponding activation energies are calculated according to eq. 10.

	Electron transfer				
	C→ CE	CE→ CEE	CH→ CEH	CEH→ CEEH	CH2→ CEH2
λ_{is}	16.3	21.1	37.5	20.5	36.7
λ	27.3	32.1	48.5	31.5	47.7
ΔG^*	1.5	58.9	2.3	35.3	0

	CEH2→ CEEH2	CH3→ CEH3	CEH3→ CEEH3	CEH_HNO→ CEEH_HNO	CEH2_HNO→ CEEH2_HNO
λ_{is}	20.5	37.1	26.5	3.8	4.2
λ	31.5	48.1	37.5	14.8	15.2
ΔG^*	16.9	1.2	6.8	1.8	6.5

	Proton transfer								
	C → CH	CH → CH2	CH2 → CH3	CE → CEH	CEH → CEH2	CEEH2 → CEEH3	CEH2 → CEH3	CEE → CEEH	CEEH → CEEH2
λ_{is}	9.7	5.1	4.5	7	5.7	4.9	4.5	7.2	5.9
λ	21.7	17.1	16.5	19	17.7	16.9	16.5	19.2	17.9
ΔG^*	0.7	6.3	31.1	0.7	0.2	2.2	15.4	9.2	6.6

	CEH → CEH_HNO	CEH_HNO→ CEH2_HNO	CEH2 → CEH2_HNO	CEEH → CEEH_HNO	CEEH_HNO→ CEEH2_HNO	CEEH2 → CEEH2_HNO
λ_{is}	1.6	1	1.7	0.2	1.8	4.5
λ	13.6	13	13.7	12.2	13.8	16.5
ΔG^*	22	1	41.5	14.7	0.1	2.1

	PCET					
	C→ CEH	CH→ CEH2	CH2→ CEH3	CE→ CEEH	CEH→ CEEH2	CEH2→ CEEH3
λ_{is}	5.8	5.1	4.3	3.8	2.6	2.5
λ	32.8	32.1	31.3	30.8	29.6	28.8
ΔG^*	0.5	1.6	0.1	12.9	5.4	13.2

	CH→ CEH_HNO	CH2→ CEH2_HNO	CEH→ CEEH_HNO	CEH_HNO→ CEEH2_HNO	CEH2→ CEEH2_HNO
λ_{is}	10.3	11	2	1.4	1.9
λ	37.3	38	29	28.4	28.9
ΔG^*	6.5	3.2	5.4	1.4	6.5

The calculated barriers provide a qualitative picture of the entire process. Importantly, both pathways that were proposed on the basis of the thermodynamic analysis appear to be kinetically feasible. In fact, the recharging via two PCET reactions proceeds through activation barriers of 1.6 and 6.5 kcal/mol. The alternative scheme, judging from the obtained barriers, includes an electron

transfer barrier of 2.3 kcal/mol (CH→CEH), followed by PCET (CEH→CEEH_HNO) and proton transfer (CEEH_HNO→CEEH2_HNO) with barriers 5.4 and 0.1 kcal/mol, respectively.

Using the calculated activation barriers the combined potential energy surface for the entire recharging process can be constructed (Figure 3.15). In case two intermediates have the same amount of protons but differ in the proton distribution in the active site (CEH2 and CEH_HNO for example), the intermediate that has lower energy was taken for the potential energy surface construction.

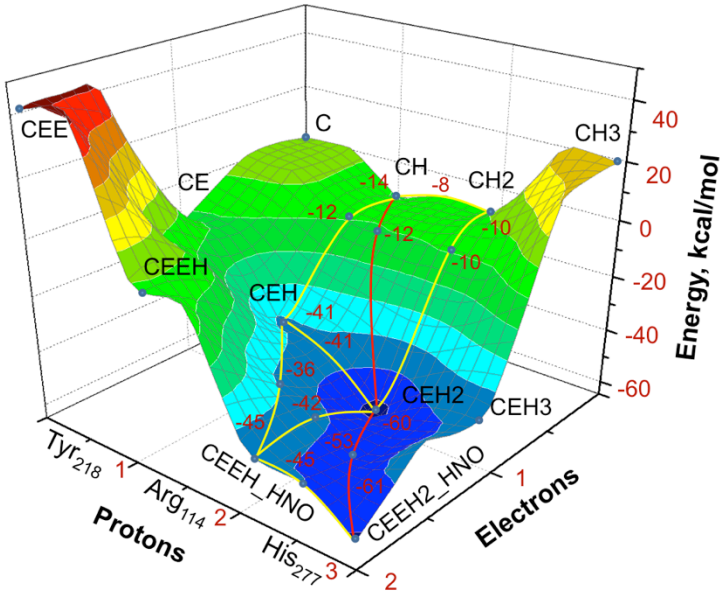


Figure 3.15 The entire recharging process scheme of the C*c*NiR active site. Feasible steps are highlighted with curved lines (red for PCET, yellow for ET and black for PT).

Several conclusions emerge from these results. First of all, there are two highly exothermic reactions that involve the starting complex CH. The first one is reduction to CEH (CH→CEH), while the second is the PCET reaction CH→CEH2. The PCET is thermodynamically more favorable. However, the ET can still take place, as the estimated activation barriers are very close in energy (Table 3.8). Continuing from the CEH complex, again two possibilities have to be considered. The PT reaction CEH→CEH2 can compete with the PCET reaction CEH→CEEH_HNO. The PCET barrier is 5.4 kcal/mol (Table 3.8), whereas the protonation is almost barrierless. As a consequence the CEH2 intermediate is preferentially formed. Formation of CEEH_HNO and its subsequent very fast turn over to CEEH2_HNO is considered to be less favorable. Thereby, the predicted kinetic trap CEH2 is the main intermediate after the first reduction. Nevertheless, the enzyme is known to convert NO with high turnover rates,[55] showing that the kinetic trap does not prevent the reaction from occurring. The most probable way out of the kinetic trap is the PCET CEH2→CEEH2_HNO, which has an activation barrier of 6.5 kcal/mol and is almost thermoneutral. Alternative reactions are endothermic and would require activation barriers of 16.9 or 15.4 kcal/mol for the reduction CEH2→CEEH2 or protonation CEH2→CEH3 respectively. Thus the PCET reaction CEH2→CEEH2_HNO is a rate-limiting step of the entire recharging process.

Using the calculated rate constants, the kinetics of the process was simulated by solving the coupled differential equations describing the various interconversions (Figure 3.16). The most probable reaction pathway CH→CEH2→CEEH2_HNO and the alternative parallel pathway: CH→CEH→CEEH_HNO→CEEH2_HNO, including the parallel protonation reaction CEH→CEH2, were treated in the simulation. The corresponding reaction rates for each reaction step were calculated using transition state theory according to the equation:

$$k = \frac{k_b T}{h} \exp\left(-\frac{E_a}{RT}\right) \qquad (126)$$

where k_b is the Boltzmann constant, h - Plank constant, R - universal gas constant, T - temperature and E_a - activation energy. Using calculated rate constants, concentrations of every intermediate were deduced as a function of time. The set of differential equations were solved numerically using the Tenua program,[146] that is based on KINSIM by Barshop, Wrenn and Frieden.[147] Owing to the small barriers, the overall effective reaction rate is very fast and is predicted to proceed on a microsecond timescale. The simulations demonstrate that the first reduction step is very fast. The substrate CH is readily consumed and entirely turned over to the CEH2 and CEH intermediates (Figure 3.16). The latter then reacts to form the CEH2 intermediate, whereas the CEEH_HNO intermediate does not build up to appreciable steady state concentrations. As a result of the first reduction the CEH2 complex is the only intermediate that is formed (Figure 3.16). The further reaction proceeds only through the reversible PCET CEH2→CEEH2_HNO step, which reaches its quasi-equilibrium after 200 μS. According to our simulations, the CEH2 intermediate forms in small equilibrium concentrations (~ 5% in the reaction mixture). In fact, trace quantities of CEH2 were also detected in the EPR experiments of Einsle *et al*.[40]

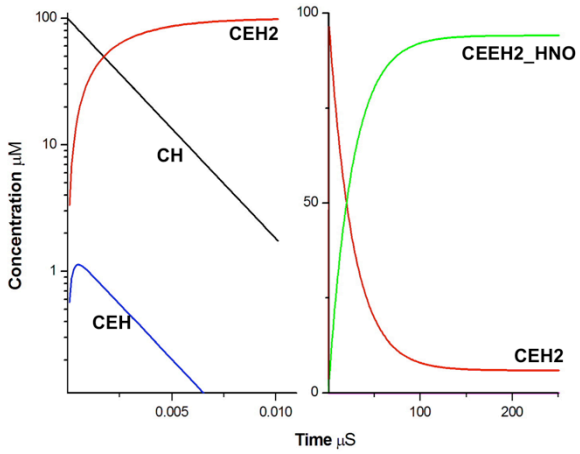


Figure 3.16 Simulated kinetics for both proposed pathways on the basis of the quantum chemical results. The enzyme concentration was assumed to be 100 μM. The first pathway consists of two consecutive PCET reactions CH→CEH2→CEEH2_HNO. The second pathway includes a cascade of slightly exothermic reactions CH→CEH→CEEH_HNO→CEEH2_HNO.

Finally, as it is seen from the kinetic simulation, that there is no inhibition of the reaction by the formation of the overly stable $\{Fe(NO)\}^7$ intermediate CEH2, despite the presence of a deep energy

minimum in the energy landscape (Figure 3.15). The recharging of the active site proceeds through two consecutive PCET reactions CH→CEH2→CEEH2_HNO.

3.2.4 Conclusions

The present work has provided detailed insights into the recharging process of the C*c*NiR active site following the first N-O bond cleavage. Taking into account the possible redox states of the active site core ($\{Fe(NO)\}^6$-$\{Fe(NO)\}^8$) together with possible protonation states of the adjacent active site ligands and the substrate makes it possible to trace a consecutive series of events leading to the fully recharged complex $\{FeHNO\}^8$. The donation of two electrons and three protons is needed to reach the $\{FeHNO\}^8$ intermediate with protonated Tyr_{218} and Arg_{114} side chains.

It was shown that including the relevant active site side chains is important for obtaining reliable results. For instance, the surrounding protein ligands are responsible for hydrogen bond formation and charge distribution in the active site, which in turn influences the reduction of the substrate. An analysis of the conformational behavior and the changes in the electronic structure of the $\{Fe(NO)\}^6$-$\{Fe(NO)\}^8$ series revealed mostly ligand-centered reduction processes. As a consequence of two electron reduction, one of the NO π-bonds is broken and two separate lone pairs on N(NO) and O(NO) are formed. The N(NO) centered lone pair is the center for the subsequent nucleophilic attack forming HNO.

The reduction of the active site is accomplished by transport of electrons from the external electron donor (a membrane-anchored tetra-heme cytochrome (NfrH) redox partner in case of *W. succinogenes* and soluble penta-heme cytochrome NrfB in case of *E.coli*) through a cascade of redox-active hemes. Theoretically, the energy of the incoming electron can be estimated from the adiabatic electron affinity of the neighboring heme. The calculations showed that the heme can easily be oxidized by the active site either in the ET reaction CH→CEH or in the PCET process CH→CEH2. The second electron, however, can only be supplied upon addition of an extra proton.

The proton supply issue was addressed. The fact that the Ca-site is placed in the inlet channel in direct proximity of the active site and, as was discussed previously, plays an important role in the substrate reduction (presumably as a structural element), implies that it can also effectively modulate the proton supply process. Indeed, a proton localized on the Ca-site is approximately 6 kcal/mol higher in energy than a proton in pure water solution. Thus, the stages involving a proton transfer become thermodynamically more favorable. Moreover, the calculated ΔH of water dissociation in the Ca-site is significantly smaller than the corresponding ΔH of pure water dissociation calculated at the same level of theory (a well known effect of water coordination to a metal center). Consequently, the Ca-site could not only act as an effective proton mediator but also as an acid/based catalytic subunit, which supplies the protons required for the reduction process.

Detailed considerations regarding the thermodynamics of the recharging process have been made. For the purpose of analysis, the calculated Gibbs free energy differences were transformed to pK_a's and redox potentials. Our calculations confirm the existence of a deep potential energy minima for the CEH2 complex ($\{Fe(NO)\}^7$ electronic configuration), which was also proposed experimentally to represent a thermodynamic sink.[40] Consequently, the system must either overcome or avoid this

energy minimum as it is known that the reduction proceeds further to the fully reduced product, NH_3, without the release or accumulation of any detectable intermediates. According to our calculations two scenarios are possible. The first one consists of two consecutive PCET reactions (CH→CEH2→CEEH2_HNO) and the alternative one is a reduction (CH→CEH) followed by PCET (CEH→CEEH_HNO) and protonation (CEEH_HNO→CEEH2_HNO)(Figure 3.13). The first pathway is characterized by a highly exothermic first PCET step and almost thermoneutral second PCET step. The calculated free energy difference for the first PCET is -46.2 kcal/mol and the following second step needs ~0.7 kcal/mol to be accomplished. The second pathway does not contain endothermic reactions and involves only moderately exothermic reactions.

Additional insight into the competition between the two pathways can be obtained by analyzing the kinetics of the process. To this end, Marcus theory[138] combined with methods developed by Rosso[143] and Hammes-Schiffer[139] were used to estimate the activation barriers. These calculations make a detailed consideration of the entire recharging process possible. Importantly, both pathways proposed from the thermodynamic analysis scheme also appear to be kinetically feasible. Recharging due to two PCET reactions proceeds through two moderate activation barriers of 1.6 and 6.5 kcal/mol. The alternative scheme includes an electron transfer barrier of 2.3 kcal/mol (CH→CEH), followed by PCET (CEH→CEEH2) with a barrier of 5.4 kcal/mol and proton transfer (CEEH2→CEEH3) with a barrier 0.1 kcal/mol. In either pathway CEEH2_HNO (electronic configuration $\{FeHNO\}^8$) is formed. Simulation of the reaction kinetics showed that the first pathway is more probable. This is based on the fact that the first reduction proceeds rapidly and leads to the formation of the CEH2 complex. The latter is then undergoes conversion to the CEEH2_HNO in the rate-limiting equilibrium reaction with an activation barrier of 6.5 kcal/mol (Figure 3.16).

It should be noted that the complex recharging process in the active site of C*c*NiR can only be accomplished through a well organized network of proton and electron donors. Neighboring low-redox potential hemes are assumed to be effective electron donors. The change in the reduction state is then immediately accompanied by proton transfer from the inlet channel, part of which is the Ca-site. It is this coupling between electron and proton transfer, (a characteristic of biological processes) that takes place in the largely hydrophobic interior of proteins, which makes efficient recharging of the active site possible.

3.3 The Nitroxyl, Hydroxylamine and Ammonia Intermediates

3.3.1 Abstract

In this chapter we considered in details the second half cycle of the nitrite reduction catalyzed by Cytochrome *c* nitrite reductase. In total 3 electrons and 4 protons must be provided to reach the final product ammonia starting form HNO intermediate. According to our results, the first event in this half cycle is the reduction of the HNO intermediate accomplished by two PCET reactions. The isomeric intermediates HNOH* and H_2NO* are formed (* denotes an electron). Both intermediates are active and are readily transformed into hydroxylamine most likely through intramolecular

proton transfer either from Arg_{114} or His_{277}. The protonated side chain then provides its proton to initiate a heterolytic cleavage reaction of the N-O bond. As a result the H_2N^+ intermediate is formed. The letter readily picks up an electron forming H_2N^+* which in turn reacts with Tyr_{218}. Interestingly, an evidence for Tyr_{218} activity was provided by the mutational studies of Lukat (Lukat, P. et. al., 2008, *Biochemistry 47*, 2080) but have never been observed on initial stages of the overall reduction process. According to our results, intramolecular reaction with Tyr_{218} in the final step of the nitrite reduction process leads directly to the H_3N^+* ammonia final product. The product dissociation was found to proceed through the change of spin state, which was also observed in resonance Raman investigation of Martins (Martins, G., et. al. (2010), *J Phys Chem B 114*, 5563)

3.3.2 Computational model

Initial structures for the enzyme active site models were taken from the protein data bank[6] (PDB code 2E80[40]). The model used in the theoretical investigation consists of His_{277}, Arg_{114} and Tyr_{218} as the first sphere distal ligands. These side chains were truncated at the α-carbon atoms. The heme is included in the model without its substituents. Lys_{134}, the proximal ligand of iron, was approximated by methylamine. The model also includes two water molecules (Figure 3.17).

Constraints were imposed on the initial structures in order to mimic the steric influence of the protein backbone on the active site. Thereby, the fragment constraints type was chosen. The side chains Tyr_{218}, His_{277} and Arg_{114} form three fragments connected to the active site that itself was defined to consist of four fragments: the heme group, the substrate molecule, the axial Lys_{134} and two additional water molecules. All fragments are fully optimized but the mutual relative orientation of distal side chains in space is kept fixed to prevent unreasonable movement of the side chains that would not be possible in the full protein environment. The performance and errors introduced by this type of constraints have been studied in our previous investigation.[200]

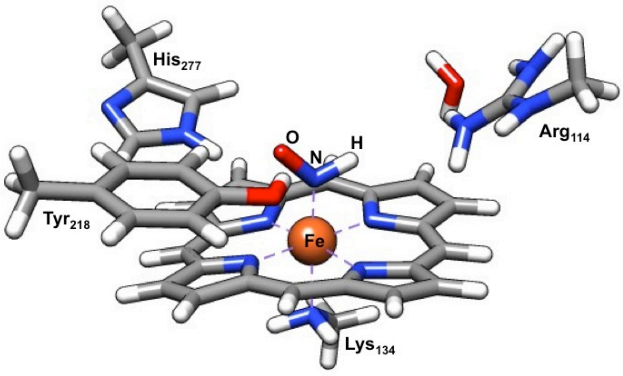


Figure 3.17 Computational model for the C*c*NiR active site. The heme was modeled without its ring substituents; side chains His_{277}, Arg_{114}, Tyr_{218} and Lys_{134} (taken from pdb structure 2E80) were truncated at α-carbon atoms.

Computational set up

The calculations were performed with the ORCA electronic structure package[117] using density functional theory (DFT).[22] The models of the enzyme active site were optimized with the BP86 functional[162,163] together with the resolution-of-the-identity (RI) approximation (in the Split-RI-J variant[164]).[103,105] A triple-ζ basis set TZVP[165] together with a matching auxiliary basis set (TZV/J[107,108] in ORCA notation) was applied for all atoms. Subsequent numerical frequency calculations were performed for all models for the sake of verifying structures found as local minima or transition states and obtaining thermochemical corrections. Frequencies were computed through two-sided numerical differentiation of analytic gradients using increments of 0.002 Bohr. Single point energy calculations were performed on the optimized structures with the B3LYP functional[152,153] that is thought to provide higher accuracy than BP86 for reaction and activation energies.[21] The 'chain-of-spheres' (RIJCOSX) approximation[109] was invoked in these calculations. The basis set was of triple-ζ quality with two sets of first polarization functions on all atoms (TZV(2d, 2p)[167] in ORCA notation). In order to crudely model the long-range dielectric effects of the protein, the COSMO[23] solvation model was employed with a dielectric constant of 4.9. All other COSMO parameters were left at their default values. Additionally, Grimme's semi-empirical correction was employed to account for dispersion forces as implemented in the ORCA package.[168,169]

For visualization purposes the unrestricted Kohn-Sham orbitals were transformed to quasi-restricted orbitals[170] (QROs), which were localized according to the Pipek-Mezey intrinsic localization criterion.[171] The model structures were visualized with the Chimera program.[172]

3.3.3 Results

General strategy

In the complicated multistage reduction process taking place in C*c*NiR several possibilities for the reaction proceeding must be taken into account at each step: (1) Protonation of the surrounding side chains and protonation of the metal-substrate core; (2) Redox properties of the heme-iron core; (3) Intramolecular proton transfer; (4) Intramolecular proton coupled electron transfer (PCET). The analysis of the results starts with the discussion of the geometrical parameters and electronic structure of the reaction intermediates on a given stage. Then the reactivity of the intermediates on this stage of the reduction is considered. The notation used throughout this paper shows in brackets the intermediate bound to the metal core on a certain stage (with an extra charge or electron if needed) and the protonation state of the active site side chains. For example, $(H_2NO^*)H_TH_A$ denotes H_2NO intermediate with an extra electron bound to Iron and Tyr_{218} with Arg_{114} are in protonated state.

The HNO intermediate.

Calculated geometrical parameters of the Fe-HNO adducts are in a good agreement with the available experimental reports,[202,203] as well as with synthetic[204,190] and theoretical models.[205] Two characteristic distances, Fe-N(NO) and N-O, and an angle Fe-N-O are well reproduced. For example, the metal – NO distances reported for Myoglobin and computational models[205] vary in the range between 1.795 – 1.820 Å. This distance is shorter in our models (Table 3.9), which should probably be attributed by the presence of the positively charged side chains and their polarizing effect on the Fe-NO core. As a consequence of the polarization the Fe-NO distance becomes shorter in the series $(HNO)H_T$ - $(HNO)H_TH_A$ - $(HNO)H_TH_AH_H$. The N–O bond distance and the calculated vibrational stretching frequencies are also consistent with previous investigations.[205,193] Here, the correlation between protonation states of the side chains and bond length/frequency is not immediately obvious but becomes intelligible through an analysis of the calculated electronic structures.

Table 3.9 Selected geometric parameters (distances in Å, angles in degrees and vibrational frequencies in cm^{-1}) for the HNO derivatives. An additional proton is put either to the HNO core or to one of the surrounding side chains: Tyr_{218} (denoted H_T), Arg_{114} (H_A) or His_{277}(H_H). Possible reduction is denoted by the star. In the column with Fe-N(porph) distances an average of the four bond lengths is shown.

Complex	*Fe-NO*	*N-O*	*∠Fe-N-O*	*Fe-N(lys)*	*Fe-N (porph)*	*ν(NO)*
$(HNO)H_T$	1.78	1.26	130	2.13	2.01	1374
$(HNO)H_TH_A$	1.78	1.25	133	2.12	2.01	1398
$(HNO)H_TH_AH_H$	1.77	1.26	133	2.11	2.01	1352
$(HNOH^+)H_A$	1.83	1.39	131	2.09	2.01	984
$(HNOH^+)H_TH_A$	1.78	1.36	126	2.07	2.01	1038
$(HNOH^+)H_TH_H$	1.78	1.38	126	2.08	2.01	1009
$(H_2NO^+)H_T$	1.96	1.32	73	2.04	2.02	1177
$(HNO^*)H_T$	1.85	1.31	132	2.12	2.01	1200
$(HNO^*)H_TH_A$	1.83	1.28	132	2.12	2.01	1275

The electronic structure of the Fe-HNO complex is exhibited on Figure 3.18. In the localized MOs scheme one can identify three orbitals of mainly iron character, one of which (labeled d_{yz} in Figure 3.18) is involved in a strong *π-backbonding* interaction with the HNO ligand. The other two metal-based MOs (d_{xy} and d_{xz}) have nonbonding character. A strong *σ-bonding* interaction is formed between the metal (involving the metal d_z^2 based orbital) and the NO ligand. The ligand based MOs reveal a double bond between N and O. An important feature of the electronic structure is the presence of an oxygen-based lone pair, which, together with the N-O *π-bonding* orbital, should determine hydrogen bonding interactions with the active site side chains. The lowest unoccupied MO features an empty mainly HNO-π* orbital. Hence, the description of this species can be considered as low-spin Fe(II) bound to a neutral HNO ligand. These results are in agreement with other calculations on coordinated HNO species.[190]

In the series $(HNO)H_T$ - $(HNO)H_TH_A$ - $(HNO)H_TH_AH_H$ the N-O bond length becomes shorter for the $(HNO)H_TH_A$ but for $(HNO)H_TH_AH_H$ it the bond is longest. Obviously, the interaction between the $\{FeHNO\}^8$ unit and Arg_{114} in $(HNO)H_TH_A$ is of different nature than with His_{277} in $(HNO)H_TH_AH_H$. It should be noted, that the HNO ligand is in the same conformation in all three models: the oxygen points towards His_{277}. Thus, the interaction in $(HNO)H_TH_AH_H$ with protonated

His_{277} is stronger and leads to the formation of a strong hydrogen bond with the O(HNO) lone pair. Furthermore, this hydrogen bonding interaction polarizes the d_{yz} MO, which has antibonding character with respect to N-O bond. As a consequence the d_{yz} MO is more strongly populated which weakens and lengthens the N-O bond. Protonated Arg_{114} in $(HNO)H_TH_A$ has less impact on the NO-ligand because of the larger distance between the side chain and the ligand. Also since an extra proton of Arg_{114} forms a hydrogen bond with N(HNO) an interaction of the opposite character is expected: less electronic density populates the d_{yz} MO.

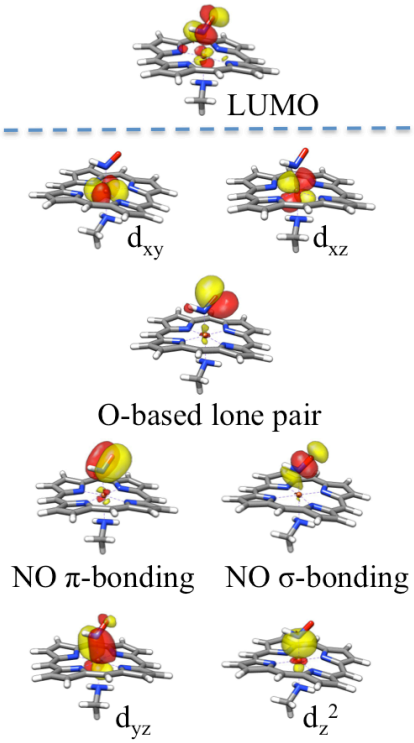


Figure 3.18 Contour plots of important MOs calculated under B3LYP/TZV(2d,2p) level of theory for the $(HNO)H_TH_A$ complex. Active site side chains are not depicted for simplicity. The occupied orbitals were localized according to Pipek-Mezey (PM) algorithm. The unoccupied orbitals are not localized.

The Fe-HNO intermediate is transformed to Hydroxylamine (H_2NOH) in a 2-electron 2-proton reaction step. The elementary steps of this process are considered below. To this end the issue of the reference energetics for the required incoming protons and electrons must be considered. We follow the same route as in our earlier work on C*c*NiR.[206] Briefly, the electron supply in C*c*NiR is carried out by the transport of electrons from the external electron donor (presumably from a membrane-anchored tetra-heme cytochrome (NfrH) for the *W. succinogenes* C*c*NiR) through a cascade of redox-active hemes to the active site. Hence, the energy of the incoming electrons can be estimated from the adiabatic electron affinity of the neighboring heme which was previously calculated to be 94 kcal/mol.[206]

Proton supply is presumably accomplished through the transfer of protons from the positively charged inlet channel identified by Einsle *et al.*[40] Assuming the proton donor to be a water molecule

(or a corresponding charged hydronium ion) from the substrate channel, one can take the proton absolute hydration enthalpy from cluster-ion solvation data as a reference point.[149] This choice is justified by the fact that the inlet channel is filled with water which is involved in multiple hydrogen bonding interactions. Hence, the dissociation in the inlet channel and in bulk solution should be comparable.

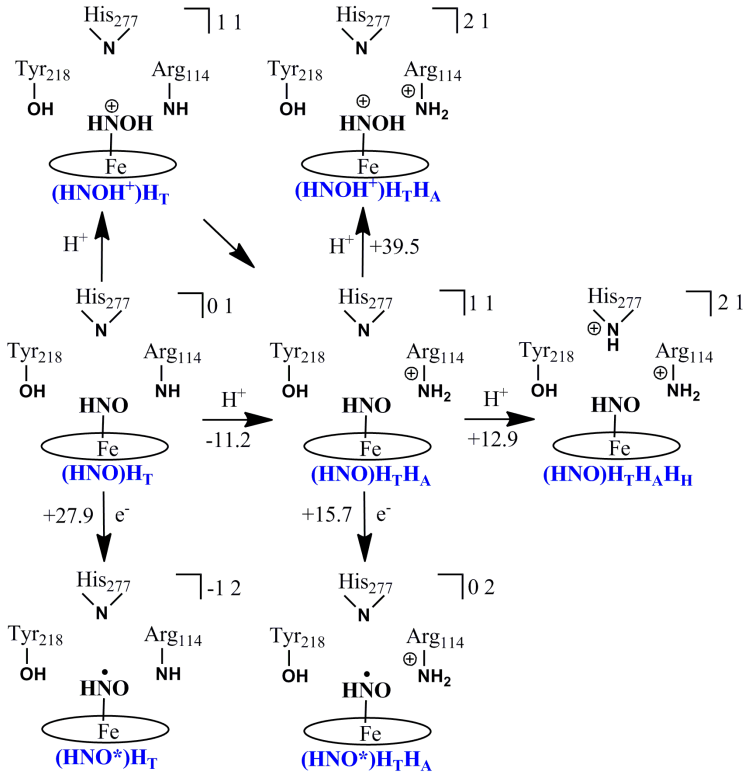


Figure 3.19 Possible protonation and reduction steps starting from $(HNO)H_T$ intermediate. An extra proton can go to one of the side chains or directly to the Fe-HNO core.

Starting from the $(HNO)H_T$ intermediate several protonation and reduction elementary steps are possible (Figure 3.19). For instance, $(HNO)H_T$ can be reduced to $(HNO^*)H_T$ or protonated to form either $(HNOH^+)H_T$ or $(HNO)H_TH_A$. In the case of reduction $(HNO)H_T + e^- = (HNO^*)H_T$ an extra electron is put on the antibonding LUMO orbital causing a drop in the Fe-N(NO) and N-O distances and in the characteristic N-O stretching vibrational frequency (Table 3.9). An analysis of the electronic structure (see App. A) showed that the SOMO in $(HNO^*)H_T$ has a shape that is comparable to the LUMO of $(HNO)H_T$. The reduction $(HNO)H_T + e^- = (HNO^*)H_T$ is energetically very unfavorable (Figure 3.19). Moreover, the protonation of the substrate: $(HNO)H_T + H^+ = (HNOH^+)H_T$ is also an energetically prohibited reaction. An extra proton on $(HNOH^+)H_T$ is transferred to the Arg_{114} upon optimization. Hence, the only possibility, which is energetically feasible, is $(HNO)H_T + H^+ = (HNO)H_TH_A$ (Figure 3.19).

After Arg_{114} protonation in the $(HNO)H_TH_A$ complex no extra protons or electrons can be transferred to the active site at a low energetic costs. As a possible mean to proceed with the reduction process through exothermic elementary stages intramolecular reactions in $(HNO)H_TH_A$ were considered:

1. Fe-HNO + HO-Tyr(**$(HNO)H_TH_A$**) = Fe-$HNOH^+$ + $^-$O-Tyr(**$(HNOH^+)H_A$**) +47.1 kcal/mol

2. Fe-HNO + H_2N^+-Arg(**$(HNO)H_TH_A$**) = Fe-$HNOH^+$ + HN-Arg(**$(HNOH^+)H_T$**) (H^+ jumps back to Arg_{114})
3. Fe-HNO + H_2N^+-Arg(**$(HNO)H_TH_A$**) = Fe-H_2NO^+ + HN-Arg(**$(H_2NO^+)H_T$**) +44.7 kcal/mol

The notation for the intramolecular reactions shows the intermediate bound to the metal core and the active site side chain, which acts as a proton donor on this stage. In the intramolecular reactions 1-3 proton donation role of Arg_{114} and Tyr_{218} was considered. Resulting intermediates on this stage were $HNOH^+$ and H_2NO^+. Both intermediates have significantly elongated N-O bond and its vibrational frequency dropped (Table 3.9), which suggests the N-O bond activation through the protonation. However, it was found that none of the three reactions should be considered as realistic for the reduction step under the conditions of a living cell because of either unfavorable reaction energetics or insufficient product stability. Thus, analyzing all possible results, the endothermic protonation of His_{277} could still be an option. Importantly, the cleavage of the first bond N-O bond in NO_2^- substrate on the initial stage of the process is done similarly through the endothermic reprotonation of the His_{277} side chain.[200] The resulting fully protonated $(HNOH^+)H_TH_AH_H$ complex can react in intramolecular reactions 4-5. Both reactions are highly endothermic, though the proton transfer from His_{277} in reaction 5 is energetically much more favorable. Because of the endothermicity of reactions 4 and 5 additional possibilities have to be considered. Therefore, the PCET reactions 6 and 7 were studied. Here His_{277} acts as the proton donor and an extra electron comes from neighboring bis-histidine-ligated heme. In this case both PCET reactions 6 and 7 are exothermic and can be considered as possible reaction pathway. The isomeric intermediates $(HNOH^*)H_TH_A$ and $(H_2NO^*)H_TH_A$ are formed and considered as starting points for the following reaction stage.

4. Fe-HNO + H_2N^+-Arg(**$(HNO)H_TH_AH_H$**) = Fe-$HNOH^+$ + HN-Arg(**$(HNOH^+)H_TH_H$**) +50.6 kcal/mol
5. Fe-HNO + HN^+-His(**$(HNO)H_TH_AH_H$**) = Fe-$HNOH^+$ + N-His(**$(HNOH^+)H_TH_A$**) +26.6 kcal/mol
6. ET: Fe-HNO + HN^+-His(**$(HNO)H_TH_AH_H$**) + e^- = Fe-H_2NO^* + N-His(**$(H_2NO^*)H_TH_A$**) -12.3 kcal/mol
7. ET: Fe-HNO + HN^+-His(**$(HNO)H_TH_AH_H$**) + e^- = Fe-H_2NO^* + N-His(**$(HNOH^*)H_TH_A$**) -16.7 kcal/mol

Reactivity of the HNOH and H_2NO radical stage.

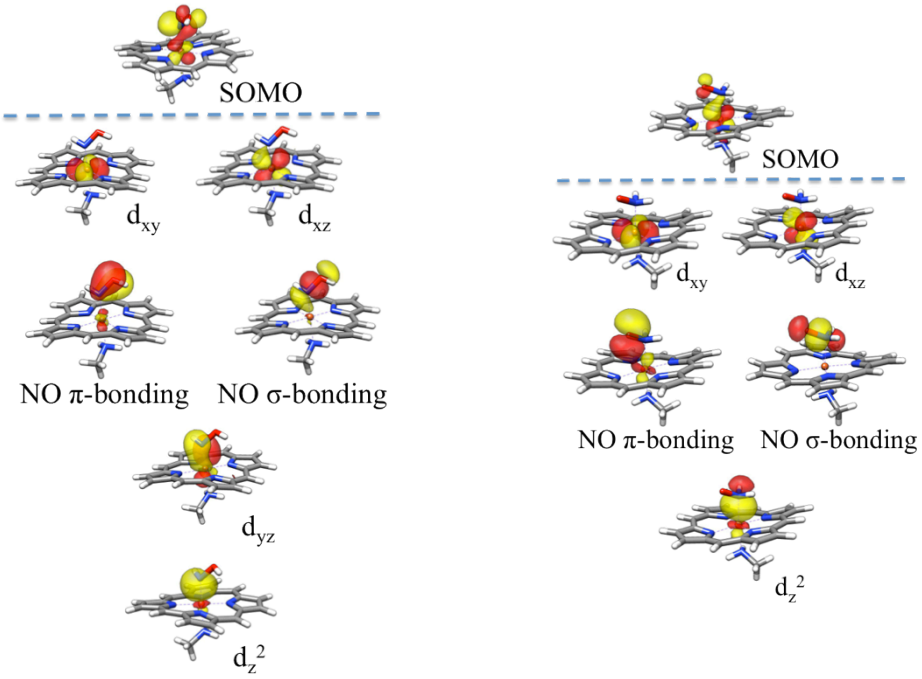
The HNOH and H_2NO species are elusive intermediates in the second half cycle of the 6-electron 7-protons reduction process. In fact, HNOH and H_2NO iron-porphyrin complexes were characterized in the electrochemical reduction experiments of Liu *et. al.*[207] In addition, HNOH was proposed to be an intermediate in siroheme Nitrite Reductase[204] and in Cytochrom P450 NO Reductase.[208] Comparison of the geometric parameters reveals further elongation of the Fe-N(NO) and N-O bonds in $(HNOH^*)H_TH_A$ and $(HNOH^*)H_TH_AH_H$ in comparison to $(HNOH^+)H_TH_A$ and $(HNOH^+)H_TH_AH_H$ oxidized complexes from the previous stage (Table 3.10).

Table 3.10 Selected geometric parameters (distances in Å, angles in degrees and vibrational frequencies in cm^{-1}) for the HNOH and H_2NO radical derivatives. An additional proton is transferred either to the metal-substrate core or to one of the surrounding side chains: Tyr_{218}

(denoted H_T), Arg_{114} (H_A) or His_{277}(H_H). The star denotes possible reduction. In the column with Fe-N(porph) distances an average of the four bond lengths is shown.

Complex	*Fe-NO*	*N-O*	*∠Fe-N-O*	*Fe-N(lys)*	*Fe-N(porph)*	*ν(NO)*
$(HNOH^{*})H_TH_A$	1.87	1.39	129	2.11	2.01	999
$(HNOH^{*})H_TH_AH_H$	1.87	1.44	123	2.1	2.01	884
$(H_2NOH^{+*})H_T$	2.01	1.43	121	2.06	2.0	935
$(H_2NOH^{+*})H_TH_A$	1.98	1.45	119	2.07	2.01	882
$(HNOH^{-})H_TH_A$	1.97	1.47	116	2.18	2.0	837
$(H_2NOH)H_TH_A$	1.99	1.46	117	2.04	2.0	821
$(H_2NO^{*})H_TH_A$	2.03	1.36	111	2.1	2.01	1010
$(H_2NO^{*})H_TH_AH_H$	2.04	1.35	116	2.11	2.01	971
$(H_2NO^{-})H_TH_A$	2.04	1.41	124	2.11	2.0	891

Consideration of the electronic structure of $(HNOH^{*})H_TH_A$ (Figure 3.20, (a)) reveals a similar bonding picture as was found for the NO ligand. As anticipated, there are *σ-* and *π-bonding* interactions between the N and O atoms. The pronounced elongation relative to the unprotonated species is due to the SOMO orbital, which has antibonding character with respect to the N-O bond but has bonding character with respect to the Fe-NO interaction. Despite this bonding interaction, the metal – ligand distance is longer than in $(HNOH^{+})H_TH_AH_H$. The reason is the weakening of the back-donation interaction. Furthermore, the back-donation is totally absent in $(H_2NO^{*})H_TH_A$ complex (Figure 3.20, (b)) as both p-orbitals of nitrogen are involved into covalent bonds with two hydrogen atoms. As a consequence, the Fe-NO bond in $(H_2NO^{*})H_TH_A$ is elongated to 2.03 Å, (Table 3.10), which is more in line with a standard coordination bond.



a **b**

Figure 3.20 Contour plots of important Pipek-Mezey localized MOs (B3LYP/TZV(2d,2p)) for $(HNOH^*)H_TH_A$ (a) and $(H_2NO^*)H_TH_A$ (b). Active site side chains are not depicted for simplicity but are included in the calculations.

A Mulliken population analysis reveals a ligand-centered reduction for $(HNOH^*)H_TH_A$. The spin density is mostly located on the ligand (64%). Also the charge on the $HNOH^*$ drops by -0.31*e* upon reduction while the charge on Fe becomes even more positive by +0.05*e*. For the H_2NO^* core the situation is different: the spin density is almost exclusively located on Fe. Hence, $(HNOH^*)H_TH_A$ is best viewed as a ligand radical bound to low-spin iron II, whereas $(H_2NO^*)H_TH_A$ has as leading resonance structure a positive H_2NO^+ core bound to a low-spin Fe^{III} center. This represents a dramatic difference in the electronic structures of the two species and must have noticeable influence on the further reduction reactions.

The proton and electron uptake in the $(HNOH^*)H_TH_A$ and $(H_2NO^*)H_TH_A$ complexes is depicted in Figure 3.21 It was found that both intermediates have high proton affinities. There is number of possibilities to reach the H_2NOH intermediate. Among them: (a) protonation of $(HNOH^*)H_TH_A$ in an almost thermoneutral step $(HNOH^*)H_TH_A + H^+ = (H_2NOH^{*+})H_TH_A$ and a consecutive exothermic reduction $(H_2NOH^{*+})H_TH_A + e^- = (H_2NOH)H_TH_A$ (Figure 3.21, (a)); (b) exothermic protonation $(H_2NO^*)H_TH_A + H^+ = (H_2NOH^{*+})H_TH_A$ and the same reduction to hydroxylamine $(H_2NOH^{*+})H_TH_A + e^- = (H_2NOH)H_TH_A$ (Figure 3.21, (b)); (c) Reduction of $(H_2NO^*)H_TH_A$ leads directly to $(H_2NOH)H_T$ as the proton from Arg_{114} is transferred to the H_2NO^* without additional energy barrier.

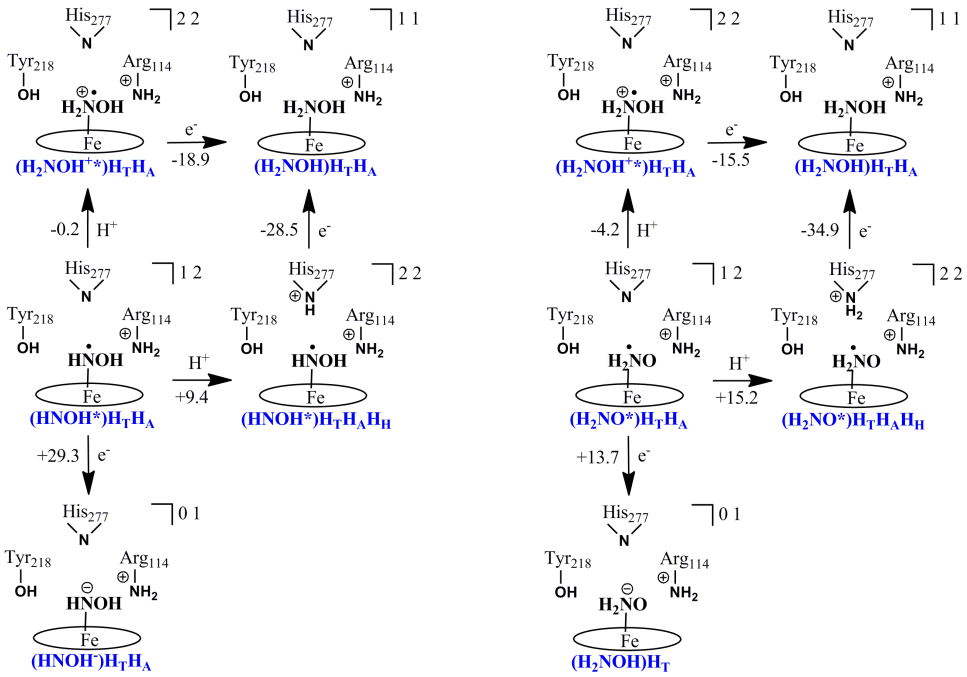


Figure 3.21 Possible protonation and reduction steps starting from $(HNOH^*)H_TH_A$ (a) and $(H_2NO^*)H_TH_A$ (b) intermediates. An extra proton can go to one of the side chains or directly to the Fe-HNOH core.

In both reaction schemes (a) and (b) the $(H_2NOH^{*+})H_TH_A$ intermediate is most likely formed as a first stage. The latter then can be readily transformed further to the hydroxylamine in case an extra proton is available. Besides protonation and reduction of the $(HNOH^{*})H_TH_A$ and $(H_2NO^{*})H_TH_A$ complexes, three possible intramolecular reactions were considered:

8. Fe-HNOH* + H_2N^+-Arg(**$(HNOH^{*})H_TH_A$**) = Fe-H_2NOH^{+}* + HN-Arg(**$(H_2NOH^{+*})H_T$**) -4.8 kcal/mol (TS 33.9 kcal/mol)
9. Fe-HNOH* + HN^+-His(**$(HNOH^{*})H_TH_AH_H$**) = Fe-H_2NOH* + N-His(**$(H_2NOH^{+*})H_TH_A$**) - 9.6 kcal/mol (TS 0.5 kcal/mol)
10. Fe-H_2NO* + H_2N^+-Arg(**$(H_2NO^{*})H_TH_A$**) = Fe-H_2NOH^{+}* + HN-Arg(**$(H_2NOH^{+*})H_T$**) -5.7 kcal/mol (TS 1.9 kcal/mol)

For reactions 8-10 true reaction barriers were located. Reaction 8 has an activation barrier of 33.9 kcal/mol and hence can be excluded from further consideration since experimental data suggest an activation barrier not higher than 15.2 kcal/mol in the rate limiting step.[37,52] Although reaction 9 is exothermic and almost barrierless, the educt $(HNOH^{*})H_TH_AH_H$ of reaction 9 is difficult to reach since His_{277} reprotonation is endothermic by 9.4 kcal/mol. The most likely process is proton transfer from Arg_{114} in reaction 10. The reaction is the most exothermic among all possibilities 8-10 and has a very low activation barrier. The product of the proton transfer from Arg_{114} is then reduced to $(H_2NOH)H_T$ in an exothermic reduction step. Hence, consistent with experiment, the most likely scenario is the formation of a hydroxylamine intermediate in form of $(H_2NOH)H_T$ or $(H_2NOH)H_TH_A$.

Reactivity of the hydroxylamine intermediate and second N-O bond leavage.

The hydroxylamine intermediate found computationally is also implied as catalytically competent intermediate by experiments on C*c*NiR[52,209] and other nitrite reductases.[210,211] In the initial work of Einsle *et al.*[40] the X-Ray structure of the bound hydroxylamine intermediate was solved. Good agreement between experimental and calculated structures was observed. The N-O bond distance of 1.46 - 1.48 Å is very close to free hydroxylamine which implies pure single bond character and a relatively weak interaction with the metal center (Table 3.11). These conclusions are also supported by the analysis of electronic structure (Figure 3.22), which reveals no additional *π-back-bonding* interactions between iron and hydroxylamine and only one conventional *σ*-donor coordination type interaction.

Table 3.11 Selected geometric parameters (distances in Å, angles in degrees and vibrational frequencies in cm^{-1}) for the hydroxylamine and hydroxylamine radical species. An additional proton is put to one of the surrounding side chains: Tyr_{218} (denoted H_T), Arg_{114} (H_A) or His_{277}(H_H). The star denotes possible reduction. In the column with Fe-N(porph) distances the average of the four bond lengths is shown.

Complex	*Fe-NO*	*N-O*	*∠Fe-N-O*	*Fe-N(lys)*	*Fe-N(porph)*	*ν(NO)*
$(H_2NOH)H_T$	2.02	1.46	125	2.05	2.0	812
$(H_2NOH)H_TH_A$	1.99	1.46	117	2.04	2.0	821
$(H_2NOH)H_TH_AH_H$	1.98	1.48	115	2.04	2.01	773

$(H_2NOH^*)H_T$	1.99	1.46	126	2.04	2.02	786
$(H_2NOH^*)H_TH_A$	1.98	1.47	121	2.04	2.02	762
$(H_2NOH^*)H_TH_AH_H$	1.97	1.49	120	2.04	2.02	717

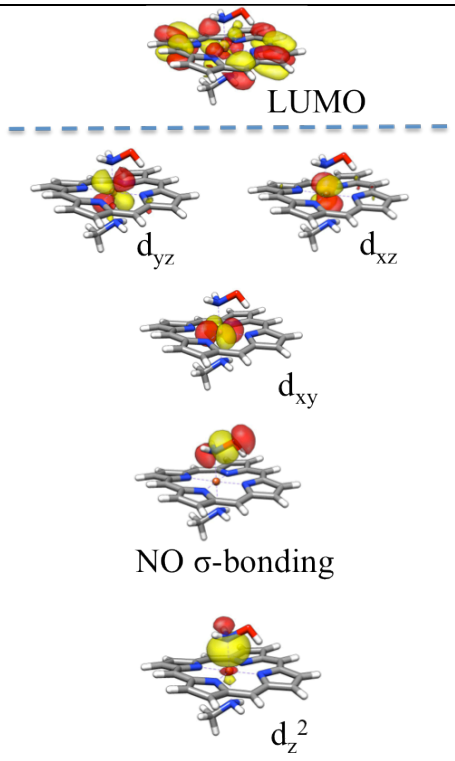


Figure 3.22 Contour plots of important MOs (B3LYP/TZV(2d,2p)) for $(H_2NOH)H_TH_A$. Active site side chains are not depicted for simplicity. The occupied orbitals were localized according to the Pipek-Mezey (PM) algorithm. Unoccupied orbitals are canonical.

The reactivity of the hydroxylamine intermediate is depicted in Figure 3.23. In the same way as in previous sections all possible electron or proton transfer reactions were considered. The protonation of Arg_{114} in reaction $(H_2NOH)H_T + H^+ = (H_2NOH)H_TH_A$ is highly exothermic and is likely to take place in the overall pathway. Further protonation leading to $(H_2NOH)H_TH_AH_H$, fully protonated complex) is slightly endothermic. Hence, both complexes must be considered as possible reactants in intramolecular reactions. By contrast, the reduction was found to be energetically very unfavorable regardless of the protonation state of the active site. For all three complexes of the first row in Figure 3.23 highly endothermic reaction energetics were calculated.

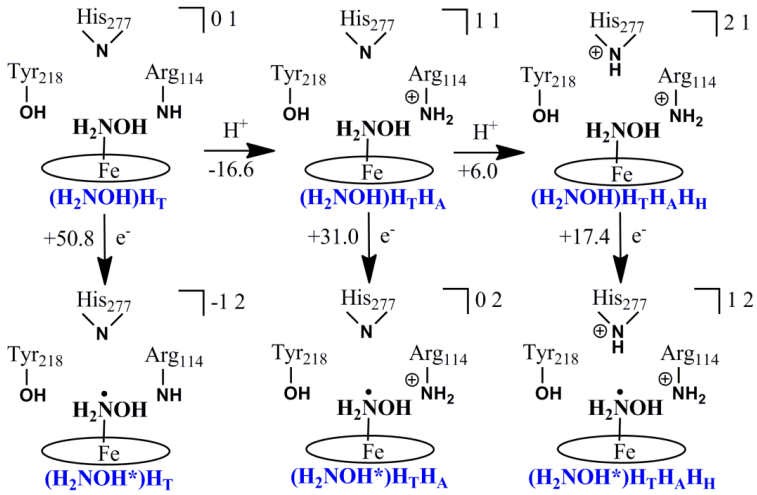


Figure 3.23 Possible protonation and reduction steps starting from $(H_2NOH)H_T$ intermediate.

As intramolecular reactions with hydroxylamine ligand, the following possibilities were taken into consideration:

11. Fe-H_2NOH + HO-Tyr(**$(H_2NOH)H_T$**) = Fe-H_2N^+ + $^-$O-Tyr(**(H_2N^+)**) +27.2 kcal/mol
12. Fe-H_2NOH + HO-Tyr(**$(H_2NOH)H_TH_A$**) = Fe-H_2N^+ + $^-$O-Tyr(**$(H_2N^+)H_A$**) +24.3 kcal/mol
13. Fe-H_2NOH + H_2N^+-Arg(**$(H_2NOH)H_TH_A$**) = Fe-H_2N^+ + HN-Arg(**$(H_2N^+)H_T$**) +11.4 kcal/mol (TS 33.2 kcal/mol)
14. Fe-H_2NOH + HN^+-His(**$(H_2NOH)H_TH_AH_H$**) = Fe-H_2N^+ + N-His(**$(H_2N^+)H_TH_A$**) +9.0 kcal/mol (TS 12.2 kcal/mol)

In the intramolecular reaction 11, Tyr_{218} was probed as a possible proton donor but was found to be energetically not competitive. The same holds for reaction 12. Reaction 13 could be feasible from a thermodynamic point but must be excluded from consideration on the basis of the high barrier. Hence, the only possible reaction, leading to cleavage of the second N-O bond, is reaction 14. Interestingly, the His_{277} amino acid was already implicated to play a key role in the first N-O bond cleavage.[200]

Ammonia stage

Ammonia complexes of heme-containing enzymes have been reported and structurally characterized.[212,213] Porphyrin complexes ligated by ammonia were synthesized and characterized.[214] The calculated geometric parameters for Fe^{III} ammonia complexes (H_3N^+*), $(H_3N^+*)H_T$, $(H_3N^+*)H_A$, $(H_3N^+*)H_TH_A$ (Table 3.12) are in excellent agreement with the ultra-high-resolution structure of the ammonia complex of Nitrophorin 4.[212] Noticeably, the metal-substrate distance is much shorter for the oxidized (H_2N^+) and deprotonated $(H_2N*)H_T$ species. This implies completely different bonding situation in those intermediates.

Table 3.12 Selected geometric parameters (distances in Å) for the Ammonia intermediates. An additional proton is put to one of the surrounding side chains: Tyr_{218} (denoted H_T), Arg_{114} (H_A) or His_{277}(H_H). The star denotes possible reduction by one e^-. In the column with Fe-N(porph) distances an average of the four bond lengths is shown.

Complex	*Fe-N(sub)*	*Fe-N(lys)*	*Fe-N(porph)*
$(H_2N^+)H_T$	1.77	2.09	2.01
$(H_2N^+)H_TH_A$	1.76	2.12	2.01
$(H_2N*)H_T$	1.81	2.09	2.01
$(H_2N*)H_TH_A$	1.86	2.13	2.01
$(H_2N^-)H_T$	1.86	2.11	2.01
$(H_2N^-)H_TH_A$	1.94	2.11	2.01
(H_3N^+*)	1.96	2.15	1.99
$(H_3N^+*)H_A$	2	2.13	1.99
$(H_3N^+*)H_T$	2.05	2.1	2.0
$(H_3N^+*)H_TH_A$	2.03	2.05	2.0

$(H_3N)H_A$	2.05	2.04	2.0

The electronic structure of the $(H_2N^+)H_T$ and $(H_2N^*)H_T$ intermediates is shown in Figure 3.24. The short calculated metal-substrate distance caused by additional *π-back-bonding* interactions, as indicated by the shape of the d_{yz} based MO in Figure 3.24, (a). The *π-back-bonding* interaction plays a key role in the initial substrate activation stage of the reduction process.[200] Also, strong *π-back-donation* is responsible for the high stability of the $\{Fe(NO)\}^7$ complex in the later stages of the process. However, high stability of the ammonia intermediate could be counterproductive by acting as a thermodynamic sink. In the case of $(H_2N^+)H_T$ its high electron affinity prevents the reaction from being stalled. The reduction yields -36.2 kcal/mol and the product $(H_2N^*)H_T$ has the other bonding pattern (Figure 3.24, (b)). The Fe-N(H_2N^*) distance is stretched indicating bond weakening (Table 3.12). Mulliken spin-density in the $(H_2N^*)H_T$ is mostly assigned to the metal, but N atom of the ligand bears also significant spin and the shape of the SOMO suggests rather delocalized distribution of the unpaired electron (Figure 3.24, (b)). Important is that the SOMO has antibonding character and is mostly responsible for the Fe-N bond weakening.

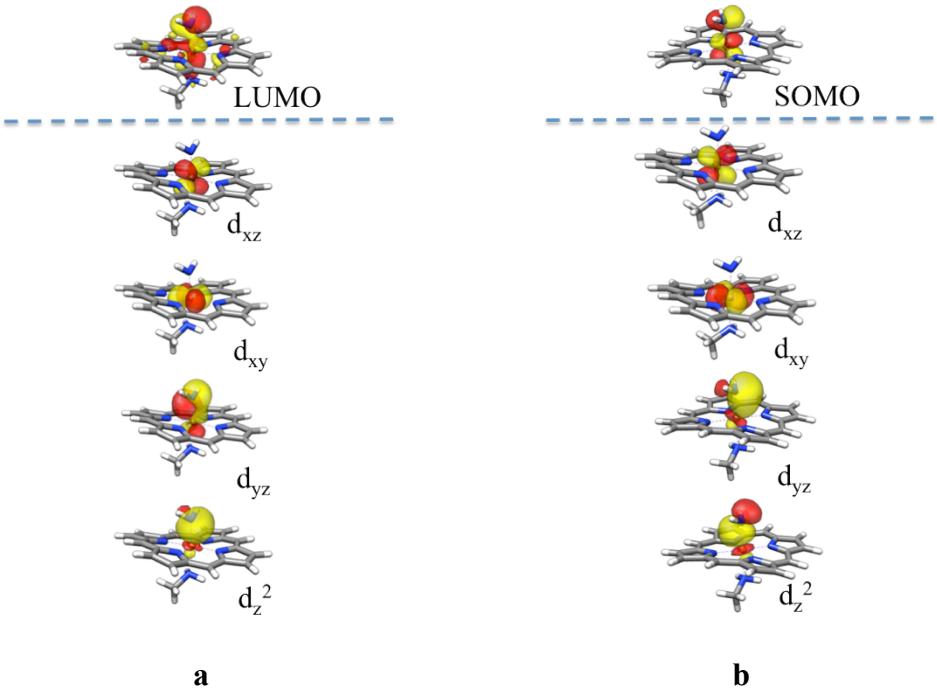


Figure 3.24 Contour plots of important Pipek-Mezey localized MOs (B3LYP/TZV(2d,2p) for $(H_2N^+)H_T$ and $(H_2N^*)H_T$. Active site side chains are not depicted for simplicity.

Figure 3.25 illustrates the reactivity of the $(H_2N^+)H_T$ species. $(H_2N^+)H_T$ is highly electrophilic and abstracts an electron in a highly exothermic reaction: $(H_2N^+)H_T + e^- = (H_2N^*)H_T$. /ΔG = -36.2 kcal/mol. $(H_2N^*)H_T$ also readily accepts a proton from Arg_{114} (Figure 3.25). However, proceeding from $(H_2N^*)H_TH_A$ there is no reaction with negative ΔG, although the protonation $(H_2N^*)H_TH_A + H^+ = (H_2N^*)H_TH_AH_H$ is uphill by only 2.8 kcal/mol.

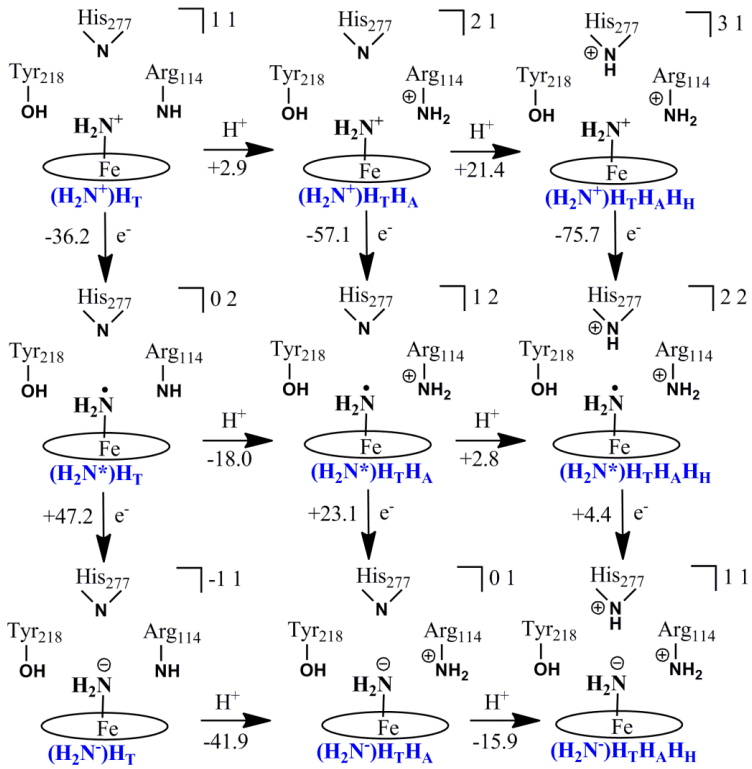


Figure 3.25 Possible protonation and reduction steps starting from $(H_2N^+)H_T$ intermediate.

As a means to reach the final product, ammonia, a number of intramolecular reactions were considered (reactions 15 – 18). The two reactions of $(H_2N^*)H_TH_A$ and $(H_2N^*)H_TH_A$ with Tyr_{218} are particularly interesting (reaction 15 and 16). Here, Tyr_{218}, which was shown to be inactive in either electron or proton transfer in the previous stages of the nitrite reduction,[200,206] is acting as a competent proton donor. A barrier of only 8 kcal/mol was calculated for reaction 16. The other plausible reaction pathway is through intramolecular protonation, where Tyr_{218} donates its proton already at the stage of $(H_2N^*)H_T$, before the Arg_{114} protonation (reaction 15). In this case, the barrier is even slightly lower and amounts to 5.9 kcal/mol. Reaction 18 could be an option if the endothermic protonation of His_{277} could take place. Reaction 17 is not taken into further consideration because of the high activation barrier.

15. Fe-NH_2* + HO-Tyr(**$(H_2N^*)H_T$**) = Fe-NH_3* + $^-$O-Tyr(**(H_3N^{+*})**) + 4.1 kcal/mol (TS 5.9 kcal/mol)
16. Fe-NH_2* + HO-Tyr(**$(H_2N^*)H_TH_A$**) = Fe-NH_3* + $^-$O-Tyr(**$(H_3N^{+*})H_A$**) + 4.0 kcal/mol (TS 8.0 kcal/mol)
17. Fe-NH_2* + H_2N^+-Arg(**$(H_2N^*)H_TH_A$**) = Fe-NH_3* + HN-Arg(**$(H_3N^{+*})H_T$**) -3.3 kcal/mol (TS 32.1 kcal/mol)
18. Fe-NH_2* + HN^+-His(**$(H_2N^*)H_TH_AH_H$**) = Fe- NH_3* + N-His(**$(H_3N^{+*})H_TH_A$**) -16.6 kcal/mol (TS 3.3 kcal/mol)

Importantly, it was postulated previously[40] and then shown by mutation studies[160] that Tyr_{218} is indispensable for enzyme activity. However, its suspected role as a hydrogen atom donor during the early stages of the reaction could never be convincingly confirmed. Thus, reactions 15 and 16 are

the first theoretical evidence that hint at the possible role of the Tyr_{218} in the overall reaction pathway.

Dissociation of the final product

A recent investigation by Martins *et al.*[179] provided the first experimental evidence that nitrite binding induces a transition from the high-spin to the low-spin configuration of the central iron. Moreover, detailed DFT investigations of the initial stage of the reduction process[200] also suggested that water is bound to low-spin Fe^{III} in the resting state of the enzyme. To study weather the dissociation takes place on the low- or high- spin surface, a relaxed surface scan of the Fe – $N(NH_3)$ distance was performed. The results are depicted in Figure 3.26. First, the dissociation directly following the intramolecular reaction with Tyr_{218} was probed (Figure 3.26, (a)). It was found that the high- and low- spin energy surfaces are energetically well separated and the dissociation takes place from the low-spin Fe^{III} metal configuration. The dissociation energy was estimated to be approximately 13 kcal/mol. A second relaxed surface scan was performed for the $(H_3N^+*)H_TH_A$ complex (Figure 3.26, (b)). The protonation state of the Tyr_{218} active site residue is the only difference between the two sets of calculations. In this scan the low-spin surface is very similar to that of $(H_3N^+*)H_A$, but the high-spin surface is now much lower in energy. Thus a low- to high-spin transition can happen and the dissociation may takes place in an almost thermoneutral fashion. The latter can bind water to form the resting state of the enzyme or it could be directly replaced by the substrate nitrite upon reduction of the protein to initiate the next catalytic cycle.

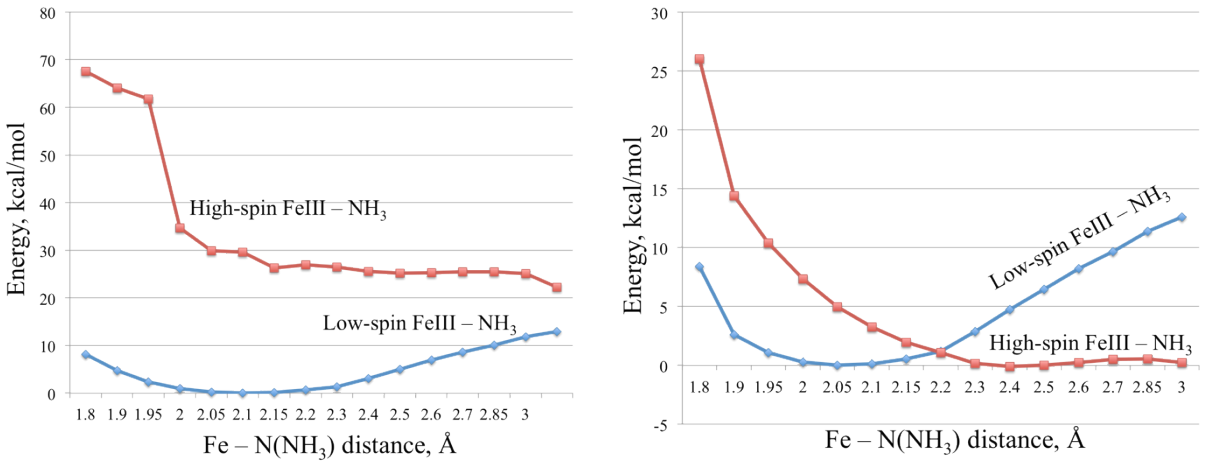


Figure 3.26 Dissociation curves for $(H_3N^+*)H_A$ (a) and $(H_3N^+*)H_TH_A$ (b) models. The relaxed surface scan was carried out for Fe – $N(NH_3)$ distance. The difference between (a) and (b) is caused by the protonation of Tyr_{218}. Only when Tyr_{218} is protonated can the final product switch to the high-spin energy surface and dissociate readily.

3.3.4 Discussion

The 6-electron 7-proton reduction of nitrite to ammonia is a highly complex process. Many factors play crucial roles in maintaining high turnover rates of the nitrite reduction. Among these factors are:

1) Substrate binding and activation through specific bonding and back-bonding interactions;

2) Electron and proton supply processes;

3) Fine-tuned hydrogen bonding networks in the active site as well as in the inlet and outlet channels;

4) Presence of Tyr_{218} which potentially is able to stabilize radical intermediates and take part in Lewis acid/base chemistry;

5) Spin-state changes during the reaction, e.g. upon product dissociation.

All these factors turn C*c*NiR into a challenging object for modern computational chemistry.

In this article we considered in details the second half cycle of the nitrite reduction. In total 3 electrons and 4 protons must be provided to reach the final product ammonia starting form HNO intermediate. At each stage the probable protonation states of the adjacent ligands and metal-substrate core were considered. According to our results, the first event in this half cycle is the reduction of the HNO intermediate. However, our calculations showed that this reduction is energetically not possible unless an additional proton enters the active site. The situation when the reduction is accomplished by protonation (usually referred as proton coupled electron transfer) is very characteristic for biological processes. In the case of HNO reduction two PCET reactions are exothermic and can be considered as possible reaction pathways. The isomeric intermediates $(HNOH^*)H_TH_A$ and $(H_2NO^*)H_TH_A$ are formed during the PCET reactions and are considered as starting points on the subsequent reduction steps.

Consideration of the electronic structure of the $(HNOH^*)H_TH_A$ and $(H_2NO^*)H_TH_A$ intermediates revealed crucial differences. The calculations predict $(HNOH^{*})H_TH_A$ to be a radical bound to low-spin Fe^{II}, whereas the $(H_2NO^{*})H_TH_A$ is best described as H_2NO^+ bound to low-spin Fe^{III}. Despite the differences in the electronic structure both intermediates were found to be highly active on further stages of the process. The most probable reaction (10), where Arg_{114} accomplishes the proton donation role, has a very low activation barrier as well as the highest energy gain of all investigated alternatives. Moreover, the proton of Arg_{114} is sterically in the best position for the substrate protonation.

The hydroxylamine intermediate is particularly interesting. It was found that the second N-O bond is cleaved in a very similar fashion as the first N-O bond.[200] In order to achieve this step, the active site must be fully protonated, thus implying the slightly endothermic protonation of His_{277}. The latter then provides its proton to initiate a heterolytic cleavage reaction of the N-O bond. As a result the $(H_2N^+)H_TH_A$ intermediate and a water molecule are formed. The activation barrier of this process is 12.2 kcal/mol, which is still lower then the experimentally determined rate-limiting barrier of 15.2 kcal/mol.

$(H_2N^+)H_TH_A$ features a strong metal–substrate *π-back-bonding* interaction, which could lead to a strong bond and hence an undesired thermodynamic sink in the overall reaction pathway. The problem was discussed in details for nitroxyl intermediate.[206] In case of $(H_2N^+)H_TH_A$ This is less of an issue because the intermediate is highly electrophilic and is readily transformed to $(H_2N^+*)H_TH_A$.

Further reduction of $(H_2N^+*)H_TH_A$ revealed the proton donor activity of Tyr_{218}. The issue of the most probable role of Tyr_{218} was discussed several times in literature.[40,160] Evidence for its importance was provided by the mutational studies of Lukat *et. al.*[160] Our calculations support the essential role of Tyr_{218} in the final step of the nitrite reduction process. Intramolecular reaction with Tyr_{218} leads to the $(H_3N^+*)H_A$ ammonia final product. The reaction is characterized by a moderately endothermic reaction free energy and a low activation barrier. Consideration of electronic structure of the $(H_3N^+*)H_A$ intermediate explains the pronounced Tyr_{218} effect: stabilization of the radical intermediate due to electron delocalization in the Tyr_{218} aromatic ring stabilize the $(H_3N^+*)H_A$ intermediate.

The issue of product dissociation was addressed by conducting relaxed surface scans of the Fe-$N(NH_3)$ distance in two model complexes: $(H_3N^+*)H_A$ and $(H_3N^+*)H_TH_A$. It was found that the dissociation of the final product ammonia from $(H_3N^+*)H_A$ is an energy consuming process. A recent resonance Raman investigation of Martins *et. al.*[179] provided evidence that the reduction process starts with a transition from the high-spin to the low-spin configuration. This spin state change was found to take place in the dissociation of product from $(H_3N^+*)H_TH_A$. Here, the elongation of the Fe-$N(NH_3)$ distance leads to the switching from the low- to high- energy surface. As a consequence of the spin state change, the dissociation is an overall much less energetically demanding process. According to our results, ammonia is released from active site in the Fe^{III} state. The open coordination site is then occupied by water molecule in the resting state or another nitrite molecule enters the active site to initiate next catalytic cycle.

3.4 Dynamic hydrogen bonding network in the distal pocket of the nitrosyl complex of Pseudomonas aeruginosa cd_1 nitrite reductase

3.4.1 Abstract

cd_1 nitrite reductase (NIR) is a key enzyme in the denitrification process that reduces nitrite to nitric oxide (NO). It contains a specialized d_1-heme cofactor, found only in this class of enzymes, where the substrate, nitrite, binds and is converted to NO. For a long time it was believed that NO must be released from the ferric d_1-heme to avoid enzyme inhibition by the formation of ferrous-nitroso complex which was considered as a dead-end product. However, recently an enhanced rate of NO dissociation from the ferrous form, not observed in standard b-type hemes, has been reported and attributed to the unique d_1-heme structure (Rinaldo, S.; Arcovito, A.; Brunori, M.; Cutruzzolà, F. *J Biol Chem.* **2007**, *282*, 14761–14767). Here we report on a detailed study of the spatial and electronic structure of the ferrous d_1-heme NO complex from *Pseudomonas aeruginosa* cd_1 NIR and two mutants Y10F and H369A/H327A in solution, searching for the unique properties that are responsible for the relatively fast release. There are three residues at the "distal" side of the heme

(Tyr_{10}, His_{327} and His_{369}) and in this work we focus on the identification and characterization of possible H-bonds they can form with the NO, thereby affecting the stability of the complex. For this purpose we have used high field pulse electron-nuclear double resonance (ENDOR) combined with density functional theory (DFT) calculations. The DFT calculations were essential for assigning and interpreting the ENDOR spectra in terms of geometric structure. We have shown that the NO in the nitrosyl d_1-heme complex of cd_1 NIR forms H-bonds with Tyr_{10} and His_{369} whereas the second conserved histidine, His_{327}, appears to be less involved in NO H-bonding. This is in contrast to the crystal structure that shows that Tyr_{10} is removed from the NO. We have also observed a larger solvent accessibility to the distal pocket in the mutants compared to the wild-type. Moreover, it was shown that the H-bonding network within the active site is dynamic and that a change in the protonation state of one of the residues does affect the strength and position of the H-bonds formed by the others. In the Y10F mutant His_{369} is closer to the NO, whereas mutation of both distal histidines displaces Tyr_{10} removing its H-bond. The implications of the H-bonding network found in terms of the complex stability and catalysis are discussed.

3.4.2 Experimental Section

The experimental section gives shortly the information on protein purification, sample preparation and EPR measurments. *This material is kept for the consistency and is not real part of the thesis!*

Mutagenesis and protein purification

Pseudomonas aeruginosa cd_1 NIR wild-type was purified as described elsewhere.[215] Mutagenesis, expression in *Pseudomonas putida*, and purification of the Y10F and dHis (double mutant, where His_{327} and His_{369} were replaced by alanines) were performed as described.[72,216,217] Since *Pseudomonas putida* expression system could not produce the protein that contains the d_1-heme, but only the *c*-heme, this semiapo-NIR was reconstituted *in vitro* with the d_1-heme extracted from the *P. aeruginosa* cd_1 NIR wild type as detailed elsewhere.[77]

Sample Preparation

Nitrosyl d_1-heme complexes of the WT and Y10F were obtained under anaerobic conditions in 50 mM bis-Tris buffer, pH 7.0. The enzyme was reduced with an excess (~ 200-fold) of sodium ascorbate and after 1 h of incubation, ~ 50-fold of sodium nitrite was added to the enzyme. The final protein concentrations varied between 0.25–0.5 mM. In case of the dHis-NO complex preparation a nitric oxide solution was added instead of the nitrite solution,[218,219] keeping the same relative amounts of ascorbate. Reduction and NO binding were monitored via UV-Vis.[218] For a good glass formation upon freezing, glycerol was added (20–30 % of the final volume). After mixing the components, quartz capillaries were rapidly filled and the solution was frozen by immersing the samples into liquid nitrogen.

For the deuterium exchange experiments D_2O (D, 99.9 %) and glycerol-$(O\text{-}D)_3$ (D, 98 %), both from Cambridge Isotope Laboratories, Inc., were used. All solutions were prepared under anaerobic conditions using the mixture of 50 mM bis-Tris (pH 7.0) in D_2O and 20–30% of glycerol. The stock solutions of sodium ascorbate and sodium nitrite were prepared, using the mixture of the buffer (D_2O) with deuterated glycerol, and degassed. The initial protein solution was diluted 10-fold by the

mixture of the deuterated buffer and glycerol, and then concentrated. Sodium ascorbate (~ 60 mM) was added to the concentrated D_2O exchanged protein and incubated for 1h. Then sodium nitrite/nitric oxide (for dHis-NO) (~ 15 mM) was added to give the final concentration of cd_1 NIR ~ 0.25–0.4 mM. The obtained nitrosyl complex was rapidly frozen in liquid nitrogen. The W-band EPR samples were prepared using the glove box, where quartz capillaries were anaerobic filled with NO-complex and immediately frozen in liquid nitrogen. The W-band EPR samples were prepared using the Glove box, where quartz capillaries were anaerobic filled with NO-complexes and immediately frozen in liquid nitrogen.

Spectroscopic Measurements.

Pulse EPR and ENDOR measurements were carried out on a home-build W-band spectrometer operating at 94.9 GHz described elsewhere.[220] The main advantages of using high field are the enhancement of both the absolute sensitivity and the resolution. The increased sensitivity permits measurements of relatively small samples (2–3 μL). While high resolution is essential to obtain an accurate determination of the principal g values, it also allows performing orientation selective ENDOR experiments. Furthermore, the appreciably larger Zeeman interaction allows the separation of different nuclei which have close gyromagnetic ratios (g) and facilitate the detection of low g nuclei such as ^2H. In addition, the overlap of signals from strongly coupled ^{14}N and weakly coupled ^1H, which is notorious at X-band is not encountered at W-band.

Echo detected (ED) EPR spectra were recorded using the two-pulse echo sequence, p/2-t-p-t-echo, the microwave (mw) pulse lengths were $t_{p/2}$ = 12.5 ns and t_p = 25 ns; and interpulse time t of 350 ns. The magnetic field was calibrated using the ^1H Larmor frequency obtained from the ENDOR spectra. ENDOR spectroscopy is alternative way to determine the NMR frequencies. It involves the application of both mw and radio frequency (RF) pulses. The standard experiments for measuring the ENDOR frequencies are the Mims[221] and Davies[222] sequences. The ENDOR spectrum is recorded by scanning the RF frequency, while monitoring the echo intensity of a stimulated (Mims) or a two-pulse echo (Davies). The ^1H ENDOR spectra were measured using the Davies ENDOR pulse sequence, p-T-p/2-t-p-t-echo, with RF p pulse applied during the time interval T. Here $t_{p/2}$ = 100 ns, t_p = 200 ns, t = 500 ns and t_{RF} = 25 μs. The ^2H ENDOR spectra were measured using the Mims ENDOR sequence p/2-t-p/2-T-p/2-t-echo, with an RF pulse applied during the time interval T. The experimental conditions for Mims ENDOR were $t_{p/2}$ = 12.5 ns, t = 350 ns (typical t value that place the blind spots well outside the spectral range) and t_{RF} = 45 μs. All ENDOR spectra were recorded using the random acquisition mode,[223] with one shot for each point and the total number of scans was varied 300–12000 depending on the S/N. The repetition time was 1 ms to which about 12 ms should be added that accounts for experimental parameters update and data transfer. All EPR measurements were carried out at 8 K.

Spectral simulations were performed to derive the spectroscopic parameters (**g** tensor, hyperfine tensor, nuclear quadrupole tensor, *etc*). The echo detected (ED) EPR spectrum was simulated using EasySpin.[224] The ^1H Davies and ^2H Mims ENDOR were simulated using the Simbud[116] software taking into account the effect of the blind spots.

Computational details.

Starting geometries for the cd_1 NIR active site model were created using the crystal structure of both the oxidized form (pdb 1nir[15]) and the reduced NO-bound form (pdb 1nno[7]). Though the resolution of 1nno is rather low (> 2.65 Å) we considered this structure because of the changes observed in the organization of the amino acid residues in the distal side of the d_1-heme pocket, which may play an important role in the substrate attraction and release. The models comprises the d_1-heme, the NO group, the proximal histidine and the side chain residues, His_{327}, His_{369} and Tyr_{10}, each of which can form direct hydrogen bonds to the NO molecule. All side chain residues were truncated at the a-carbon. For each of these starting structures nine cluster models, varying in the protonation states of the two nitrogens of the two distal histidines His_{327} and His_{369}, starting geometries were generated and optimized. The aim was to determine the effect of the protonation states of these histidines on the structure and the EPR parameters.

Constraints were imposed on the initial structures in order to mimic the steric influence of the protein backbone. Because the protein backbone is not included in the cluster models the side chains can drastically change their initial positions moving into free space, which is filled with protein chains in a real enzyme. To avoid such unrealistic rearrangements, side chains of interest were chosen and assigned to fragments. The fragments can then be connected with each other. The position of all atoms within the fragments are fully optimized by the program while the shortest distance between connected fragments, bond angles and dihedrals angles, which contain this distance, were frozen (as in the crystal structure) to keep the fragments together and retain their orientation. Such calculations are referred to as "frozen". This preliminary set of calculations using the 1nir as the initial structure showed that His_{327} does not form H-bonds with the NO and practically does not affect the EPR parameters. Therefore we further concentrated on three representative structures ***A***, ***B*** and ***C*** with different protonation states of His_{369}. We also carried out calculations on these three initial models with more relaxed constrains, where only the side chains α-carbons were frozen, to ensure that a large enough range of structures are tested. These are referred to as relaxed structures. For the structures starting from the 1nno structure we considered all nine structures and in the structure optimization we fixed only the position of the α-carbons, namely relaxed structures.

The calculations were performed with the ORCA electronic structure package version 2.7[117] using various levels of DFT.[22] The models of the enzyme active site were optimized applying the BP86 functional[163,162] together with the resolution-of-the-identity (RI) approximation procedure within the Split-RI-J variant[164] for the Coulomb term.[105,103] The geometry optimizations were carried out using the triple-ζ quality TZVP basis set[165] applied for all atoms. An auxiliary basis set TZV/J[107,108] was employed for the RI approximation.

To consider the long range electrostatic effects, self-consistent reaction field (SCRF) computations for each model were conducted using the COSMO model.[225] The dielectric constant of chloroform (4.9) was used in the calculation which is roughly representative of typical values used for protein environments.[21] All other COSMO parameters were default values. Moreover, to assess explicitly the hyperfine couplings for solvent's protons each model was supplemented with three water

molecules, which positions were fully optimized. The hyperfine couplings of solvent's protons were then calculated under the same level of theory.

For the purpose of analysis the unrestricted Kohn-Sham orbitals were transformed into the quasi-restricted orbitals[170] (QROs) which were then localized according to Pipek-Mezey localization procedure.[171] Orbitals, densities and structures were visualized with the Chimera program.[172]

3.4.3 Experimental Results

All experimental measurements have been done by our collaborators in Israel. *This section is kept for the consistency of the presented material and is not a real part of the thesis!*

Echo-Detected EPR

The nitrosyl-heme complexes of the cd_1 NIR construct studied in this work were prepared under conditions where only the d_1-heme-NO complex is formed.[219] The echo-detected (ED) EPR spectrum of a frozen solution of d_1-heme-NO complex of WT cd_1 NIR of *P. aeruginosa* (WT-NO) at pH 7.0 and its simulation are shown in Figure 3.27, (a). The simulation of the spectrum revealed the present of two paramagnetic species; a major species (~ 90 %) with a rhombic g-factor [g_{min}, g_{mid}, g_{max}] = [1.960 2.004 2.062] and a minor contribution (~ 10 %) of rhombic conformation with [g_1, g_2, g_3] = [2.076, 2.025, 1.953]. The symmetry of the minor species is not unambiguous. Accordingly we chose the symmetry which gives a best fit to the experimental data while noticing a possibility of the contribution of more than one minor species, i.e. an axial and a rhombic one. A close look shows that the agreement between the simulation at the experimental spectrum at the high field edges of the spectrum is not perfect, indicating the presence of a very small amount of a third species that we have ignored. The g-values are in good agreement with earlier reports in the literature where g_{mid} is referred to as g_{zz}, g_{max} as g_{xx} and g_{min} as g_{yy}.[226,227,228,229,230] Different preparations, however, may differ in the relative amounts of the species, which nonetheless is always low.

In our earlier report[231] the nitrosyl complex was prepared in a phosphate buffer at pH 8.0 and its spectrum has shown a lower amount of the minor species for both WT-NO and the nitrosyl complex of the Y10F mutant compared to the preparation in this work. The spectra of the nitrosyl-heme complex of the Y10F (Y10F-NO) and dHis mutants (dHis-NO) were similar to that of the WT-NO, as compared in Figure 3.27, (b), and may contain different amounts of the minor species (particularly dHis-NO). The simulation of the Y10F-NO spectrum, shown in the Figure 3.27, (c), exhibits the contribution of a major rhombic species (~ 85%) with [g_{min}, g_{mid}, g_{max}] = [1.970, 2.003, 2.062] and similarly to the spectrum of the WT-NO, a minor rhombic species with [g_1, g_2, g_3] = [2.090, 2.029, 1.955].

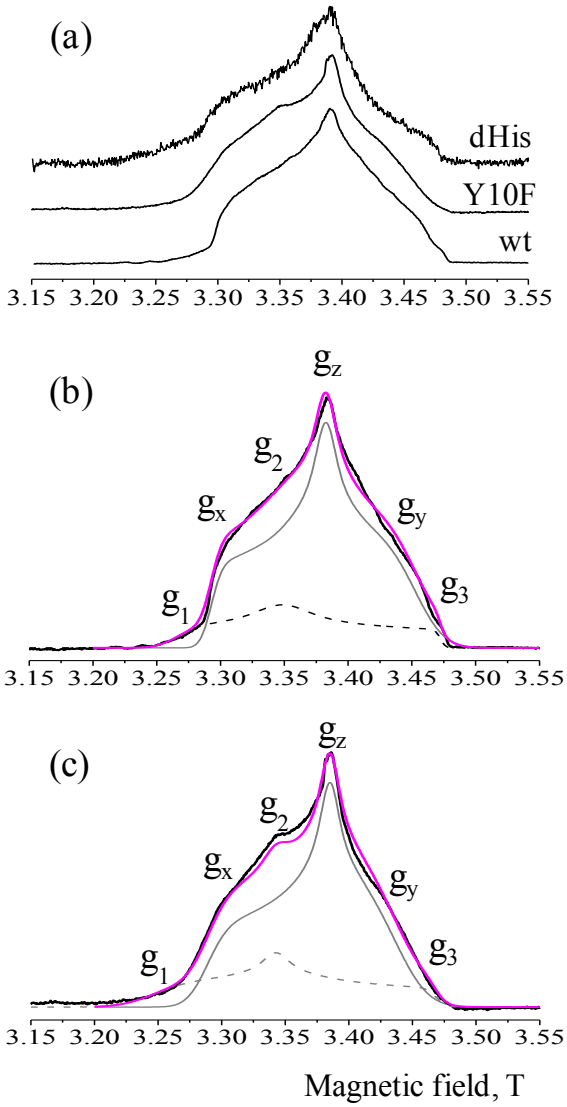


Figure 3.27 (a) Comparison of W-band ED EPR (8 K) spectra of frozen solutions of WT-NO, Y10F-NO and dHis-NO. Simulations (magenta) of the spectra of WT-NO (b) and Y10F-NO (c) as compared to the experimental spectra (black). The spectra of the individual components of major species (solid grey) and minor species (dashed grey) are shown as well. The simulation parameters are given in the text.

^{1}H Davies and ^{2}H Mims ENDOR

In order to identify the protons forming hydrogen bonds with the NO group and to determine their hyperfine and quadrupole (for ^{2}H nuclei) couplings, we have performed a set orientation selective ^{1}H Davies and ^{2}H Mims measurement at seven magnetic field positions within the EPR spectrum. The chosen field positions correspond to the principal g-values of the rhombic species and to some points in between. Because the rhombic species comprises 90% of the sample, in the following we have neglected any possible contributions of the minor species to the ENDOR spectra. The signals of H-bonded protons are identified by comparing the ^{1}H ENDOR spectra of samples prepared in H_2O with the ^{2}H signals of a sample prepared in D_2O as shown in Figure 3.28. The scale of the ^{2}H

spectra was multiplied by $g^1H/g^2H = 6.5$ for easy comparison with the ^1H spectra. This comparison clearly shows signals of exchangeable protons with a maximum hyperfine splitting of ~ 8 MHz appearing at g = 2.002 (marked with arrows). The difference in the outer edges of the spectra between the ^1H and ^2H spectra is attributed to the ^2H nuclear quadrupole interaction that effectively broadens the lines and occasionally quadrupolar splitting can be resolved. The largest detected splitting, ~8 MHz, is similar to that observed recently for the H-bond between the distal histidine and the nitrogen of the NO in Mb-NO.[77] This suggests a similar H-N(O) distance, provided that the spin density distribution is similar.

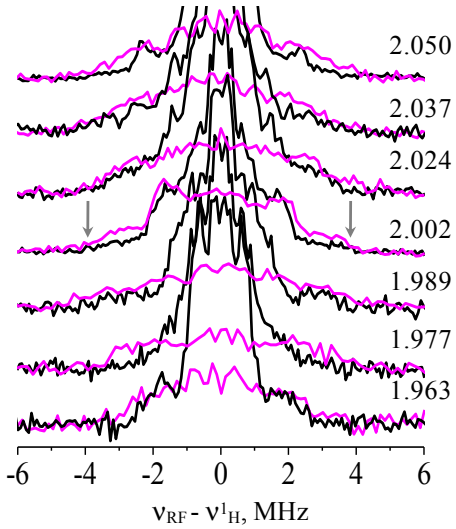


Figure 3.28 Comparison of orientation selective ^1H Davies ENDOR spectra of WT-NO in H_2O (black) with ^2H Mims ENDOR spectra of WT-NO in D_2O (magenta) at the indicated g values. The frequency scale of the ^2H spectra was multiplied by $g^1H/g^2H = 6.5$. Arrows mark the large splitting of ~ 8 MHz.

A similar set of ENDOR measurements was carried out on Y10F –NO. In Figure 3.29 we compare ^1H Davies ENDOR spectra of WT-NO with those of Y10F-NO. We observe differences (marked by arrows) that are most significant at field positions corresponding to g = 2.002 and g = 2.050.

The disappearance of the feature at ± 2.5 MHz (hyperfine coupling, A = 5 MHz, marked with arrows at g = 2.050) upon replacing Tyr_{10} with Phe provides a tentative assignment of these signals to the OH proton of Tyr_{10} that forms H-bond with the NO group. This comparison also reveals the appearance of new signals at ± 1.5 MHz in the Y10F-NO spectrum at g = 2.050 (marked with asterisk). These are tentatively assigned to the protons of one of the histidines, which were shifted due to the mutation.

This assignment, however, does not agree with the crystal structure of WT-NO[63] where Tyr_{10} is situated at a distance to far too form an H-bond with the NO. An alternative assignment is that the 5 MHz doublet at the g-2.050 spectrum of WT-NO is due to a histidine proton and the change in coupling represents the response of the histidine to the removal of the Tyr_{10}. These two alternatives were resolved using DFT calculations (see below).

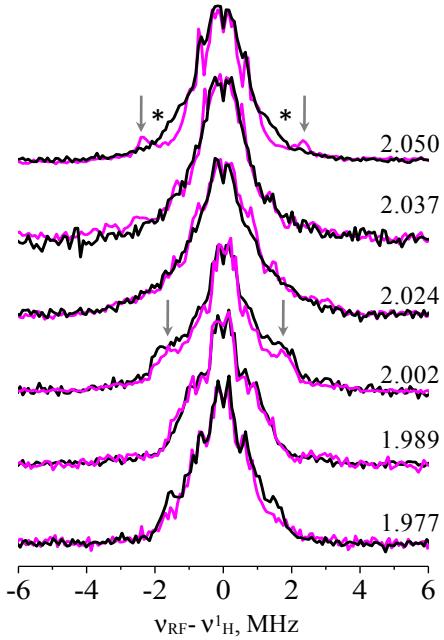


Figure 3.29 Comparison of orientation selective ^{1}H Davies ENDOR spectra of a frozen solution of WT-NO (magenta) with that of Y10F-NO (black) at the indicated *g* values. Signals marked with arrows are tentatively assigned as H-bonded proton of Tyr_{10}, while those marked with asterisk are assigned as H-bonded proton of a histidine. Here the spectra were normalized to the strongest signal.

To further highlight the differences between the WT-NO and its Y10F-NO and dHis-NO mutants, we compared their ^{2}H Mims ENDOR spectra as well (see Figure 3.30). When ENDOR spectra are compared one has to pay a special attention to how the spectra were normalized, especially when the differences are associated with intensity changes, as opposed to line shifts. The simplest way is to normalize to the most intense signal, which often is the signal of distant protons at the nuclear Larmor frequency, particularly in Mims ENDOR. This approach is problematic, if there are changes in this region. Another way, the seemingly most proper one, is to present the spectra using the ENDOR effect (ε) given by:

$$\varepsilon = [I(RFoff) - I(RFon)] / I(RFoff)$$

where ((*RFon*) and I(*RFoff*) are the echo intensity with *RF on* and *RF off* resonance, respectively. Here the problem is that ε depends on experimental conditions such as probe tuning and sample position and therefore require special care in the adjustment of the experimental conditions. Therefore, only large changes that are consistent with other observations can be considered as reliable, unless some internal standard is used. The spectra in Figure 3.30 were normalized according to the ENDOR effect. The EPR signal intensity of dHis-NO was significantly lower than in the other two (see Figure 3.27, (b)) and therefore we could not acquire a complete series of orientation selective ENDOR spectra for this sample. The comparison shown in Figure 3.30 suggests that the signals marked with arrows (g = 2.002) and corresponding to A(^{2}H) = 0.5 MHz (~ 3.3 MHz for ^{1}H) belong to the H-bonded OH proton of Tyr_{10} according to the first assignment alternative or one of the histidines according to the second alternative. In Y10F-NO the splitting of this doublet is somewhat smaller and weaker and it is assigned to an exchangeable proton of one of

the histidines because it is absent in the spectrum of dHis-NO. The spectra of the mutants also show a significantly larger intensity at the Larmor frequency particularly at g = 2.024. This suggests that in the mutants there are more distant exchangeable protons, probably solvent molecules. The total width of the dHis spectrum is only slightly smaller than the other two suggesting that there may be a water molecule in H-bond distance to the NO.

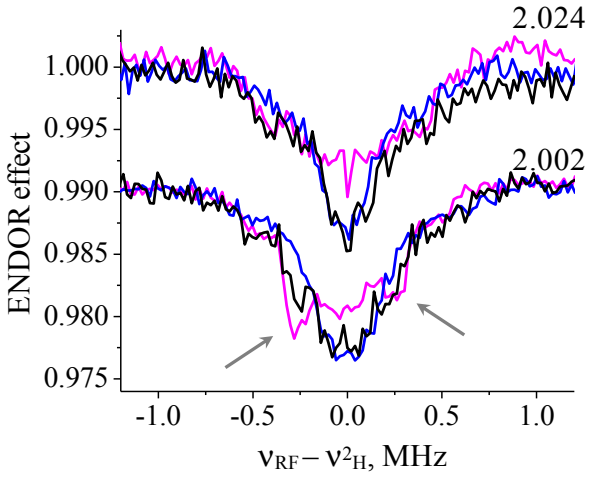


Figure 3.30 Comparison of the orientation selective ^{2}H Mims ENDOR spectra of WT-NO (magenta) with Y10F-NO (black) and dHis-NO (blue) at the indicated g values. Signals marked with arrows are assigned to H-bonded proton of Tyr_{10} (first alternative). Here the spectra are normalized according to the ENDOR effect and the g = 2.002 spectra are shifted down by 0.01.

Through the ENDOR spectra of WT-NO and the mutants we identified two H-bonds to the NO group with substantial hyperfine coupling but unlike Mb-NO,[77] where the orientation selection ^{2}H spectra were well resolved and showed a clear lineshape evolution in the orientation selection spectra (that could be well reproduced by simulation), here the ^{2}H spectra do not show well resolved distinct spectral features throughout the series due to the presence of the two H-bonds (or even more). Consequently, we did not attempt to simulate these spectra, which would require a large number of parameters (6 for each proton) but rather turned to DFT predictions to aid the spectral analysis. In our earlier study on Mb-NO[77] such calculation were found most helpful in spectral interpretation and assignment.

3.4.4 Theoretical results

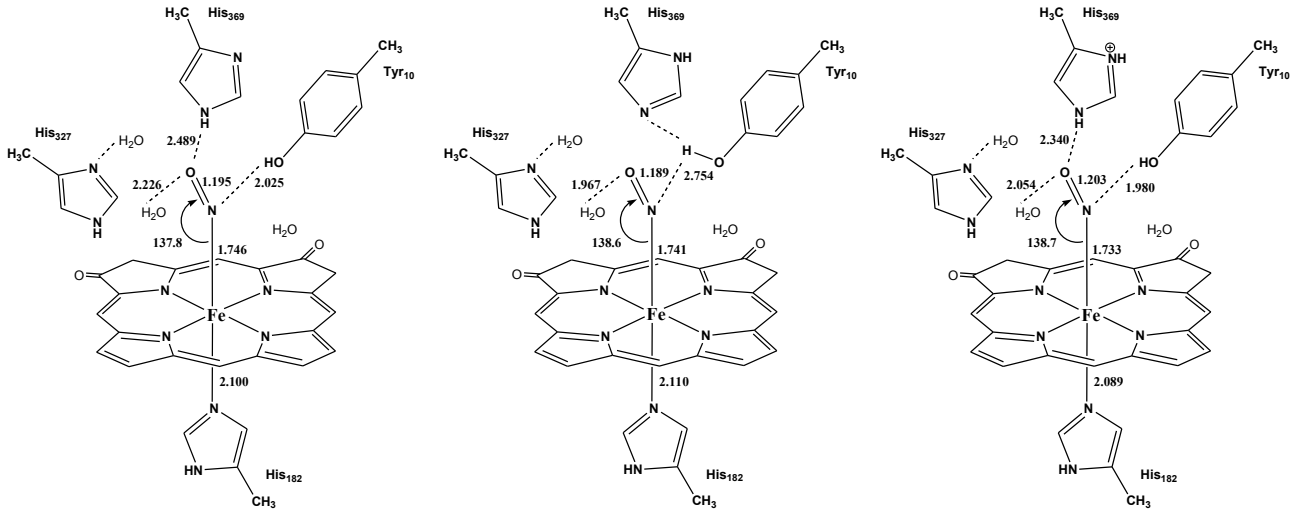
Geometric Structure

In the DFT calculations, the initial structural models were based on two crystal structures, one of the oxidized cd_1 NIR from *P. aeruginosa* (pdb 1nir,[70] resolution 2.15 Å) replacing the OH ligand with NO and the second one is the reduced NO-bound form (pdb 1nno,[63] resolution 2.65 Å). The major difference between the two structures that is relevant to this study is the location of Tyr_{10}, which is located away from the NO in 1nno whereas in 1nir it is at a H-bond distance to the OH

ligand of the d_1-heme. For each structure nine models with different protonation states of the nitrogens of His_{327} and His_{369} were constructed. The models were optimized and EPR parameters were calculated.

The preliminary calculations based on the 1nir structures showed that the protonation state of His_{327} does not significantly influence the ^{14}N hyperfine couplings of the NO and the coordinated nitrogen of the proximal His_{182}. Also, the ^{1}H hyperfine couplings of the OH of Tyr_{10} and of His_{369} are not affected by the protonation state of His_{327}. Furthermore, the ^{1}H and ^{14}N nuclei of His_{327} exhibit small hyperfine couplings. Therefore we proceeded with the calculations focusing only on three models, here after referred to as *models **A***, ***B***, and ***C*** (they are shown in Figure 3.31) where three water molecules were added. The final set of calculations was done with the PBE functional because in our previous extensive study on Mb-NO we found that this is most appropriate choice for calculating the EPR parameters of Fe(II) nitrosyl systems.[77]

In general, the predicted geometrical parameters (bond lengths, bond angles and dihedral angles) are in good agreement with the X-ray geometry of the *P. aeruginosa* WT-NO complex (1nno) [63] (see Figure 3.31). A slight underestimation of the Fe – N(NO) bond (1.746 vs 1.803 Å) and an overestimation of the Fe – N(His_{182}) (2.100 vs 2.039 Å) distances were found. The Fe-N-O angle is also in good agreement with experiment (see Table 1). A major difference between the calculated and experimental structures concerns the distance of Tyr_{10} from the NO. In the 1nno crystal structure O_{Tyr10} is 4.9 Å away from the N(NO) whereas in our 1nir based *models*, particularly ***A*** and ***C***, O_{Tyr10} is closer to the NO. To make sure that this discrepancy is not an outcome of the fragment constraints applied an additional set of calculations was carried out, where constraints were relaxed and only the position of the a-carbon of the amino acids were frozen in the geometry optimization. Also in this set the Tyr_{10} remains close to the NO ligand, which is also reflected in the calculated ^{1}H-O(Tyr_{10}) hyperfine coupling that will be discussed later. In the following, unless otherwise stated models ***A, B*** and ***C*** refer to the frozen structures.



*model **A*** *model **B*** *model **C***

Figure 3.31 Calculated geometric parameters of the three selected frozen models (Å and deg).

In *model **B*** His_{369} is only protonated at N^{δ} and this leads to conformational changes in the distal pocket (Figure 3.31). The hydroxyl group of Tyr_{10} forms an H- bond with the N^{ε} of His_{369}, consequently the distance between the Tyr_{10} OH and the NO increases significantly. When the N^{ε} of His_{369} is protonated the OH of Tyr_{10} is no longer H-bonded to His_{369} and is available to form an H-bond with the nitrogen of the NO.

Table 3.13 The geometric parameters (Å and deg) of *models A*, *B* and *C*. The values in the *exp* line correspond to crystal structure of WT-NO (1nno).[63]

	Fe-N_{NO}	**N-O_{NO}**	**Fe-N_{heme}**	**Fe-N_{His182}**
*model **A***	1.746	1.195	2.044-2.100	2.100
*model **B***	1.741	1.189	2.010-2.076	2.110
*model **C***	1.733	1.203	2.002-2.086	2.089
exp	1.803	1.153	2.052-2.085	2.039
	Fe-N-O	**O_{NO}-H_{His369}**	**N_{NO}-H_{Tyr10}**	**O_{NO}-H_2O**
*model **A***	137.8	2.489	2.025	2.269
*model **B***	138.6	-	2.754	1.967
*model **C***	138.7	2.340	1.980	2.054
exp	139.9	2.597 (to N_{His369})	5.0 (to O_{Tyr10})	3.310

Electronic Structure and Bonding

The electronic structure of the low-spin $\{Fe(NO)\}^7$ complexes has been a subject of intensive theoretical and experimental studies.[191,190] The three iron-based t_{2g} orbitals are doubly occupied. One of them has nonbonding character and consists of the almost pure iron d_{xy} atomic orbital. The two remaining t_{2g} orbitals are dominated by the iron d_{xz} and d_{yz} and NO-π^* contributions.

The unpaired electron (see Figure 3.32) is situated in a strongly Fe-NO antibonding orbital with iron-3d and NO-π* character. The composition of this orbital dominates the spin density contribution (and hence the EPR parameters) and, as discussed at length previously, is particularly sensitive to the theoretical method used. In this specific case, pure functional, such as PBE, are preferred over hybrid functionals.[192] According to the PBE results, the SOMO consists of mostly d_z^2/d_{xz} iron contributions (65%) with a major admixture of the NO-π^* based orbital (33%) (Figure 3.32), similar to what was found for the Mb-NO complex.[77]

Lehnert and coworkers[232,191] have recently proposed that in six coordinated low-spin $\{Fe(NO)\}^7$ complexes the unpaired electron is mostly localized on the NO ligand (80%). The large spin population on the NO ligand was explained as a consequence of the sixth ligand coordination. Our d_1-heme-NO structures are similar to the model systems discussed by Lehnert and coworkers[232] and therefore the differences in spin populations may be caused by the applied density functional together with steric, electronic and hydrogen bonding effects. Taken together this gives a complicated picture, that is reflected in bonding energies and finally also in observed g-values and hyperfine couplings.

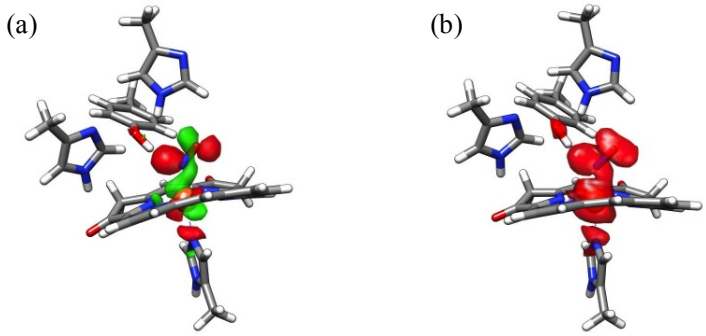


Figure 3.32 The SOMO orbital (a) and the spin density (b) of the *model A*. (The water molecules are not shown). Both the SOMO orbital and the spin density were calculated under the same level of theory as the EPR parameters.

The binding energies of the NO group were studied theoretically by different groups[192,233 ,234] and the results show that pure GGA functionals tend to overestimate the binding energy while hybrid functionals underestimate it. The value of 15 kcal/mol calculated here with the PBE functional for *model **C*** is in good agreement with previous theoretical studies,[234] but still significantly smaller than the experimental value of 20 kcal/mol.[235] The calculated binding energy correlates with the number of the H-bonds in the system (see Table 3.14). *Model **B*** has only one H-bond with a water molecule and hence has the lowest binding energy, whereas *model **C*** has three hydrogen bonds with distances shorter than those of *model **A*** and consequently shows an increase in the binding energy. The physical origin of this effect is that H-bonds stabilize the NO-π* based orbital, shifting it closer to the iron 3d-orbitals, and thus, enhancing the back-bonding effect. Consistent with this notion, the shortest Fe-N(NO) distance is observed for the *model **C***, confirming the strengthening of back-bonding for this complex. On the other hand, the spin population on the NO rises considerably for *model **B*** where the Tyr_{10} does not form a hydrogen bond with the NO. This implies that hydrogen bonds to the O(NO) and N(NO) have a different influence on the electron and spin distribution in the complex as reported by Xu *et al*.[194] They have shown that H-bonding to O pulls electrons into the FeNO π* orbital, thus strengthening the Fe–N bond while weakening the N–O bond. By contrast, H-bonding to N withdraws bonding electrons from both the Fe–N and the N–O bonds into a (partial sp^2) N based nonbonding orbital. Hence, both bonds are weakened by this interaction. This effect is manifested in the lengthening of the Fe–N and the N–O bonds in *model **A*** relative to ***B***.

Table 3.14 Binding energies (kcal/mol) and spin population on Fe, N(NO) and O(NO). Calculations were performed using the B3LYP functional for bonding energies, and the PBE functional for spin populations.

model	binding energies	Mulliken analysis			Loewdin analysis		
		Fe	N(NO)	O(NO)	Fe	N(NO)	O(NO)
*model **A***	13	0.654	0.192	0.137	0.617	0.204	0.144
*model **B***	10	0.553	0.301	0.128	0.541	0.280	0.146
*model **C***	15	0.658	0.197	0.124	0.624	0.206	0.132

3.4.5 Calculated EPR parameters

g-tensor

The calculated g-values, listed in Table 3.15, are in reasonable agreement with the experimental values, except for g_{max} (g_{xx}) which is too low for all three computational models. The underestimation of the g_{max} is commonly found in DFT g-tensor calculations on iron nitrosyl and is attributed to the limitations of the density functionals in predicting magnetic response properties. In particular, the excitation energies, calculated using DFT, enter implicitly in g values calculation and are usually not highly accurate and hence often cause errors in the computed g-tensors.[122,236] M*odel **B*** has lower g_{min} and g_{mid} values which may reflect the decrease of spin density on the Fe as a consequence of the removal of the H-bond of Tyr_{10}. The principal direction of the g-tensor is related to the NO geometry, with the orientation of g_{yy}(g_{min}) approximately along the NO bond (13, 9 and 19° off for *models **A***, ***B*** and ***C*** respectively), g_{zz}(g_{mid}) is perpendicular to it, (in the plane of the Fe-NO unit) and g_{xx} (g_{max}) is perpendicular to these two and 6, 0 and 3° off the heme plane for *models **A***, ***B*** and ***C***, respectively, pointing towards the meso-carbon (see Figure 3.33). This orientation of the g-tensor is similar to that obtained in our earlier calculations on Mb-NO and has been confirmed experimentally.[77]

Table 3.15 The DFT calculated g-values of the three models compared with the experimental values.

	exp	*model **A***	*model **B***	*model **C***
g_{max}(g_{xx})	2.062	2.030	2.030	2.030
g_{min}(g_{yy})	1.960	1.983	1.969	1.984
g_{mid}(g_{zz})	2.004	2.005	2.000	2.007

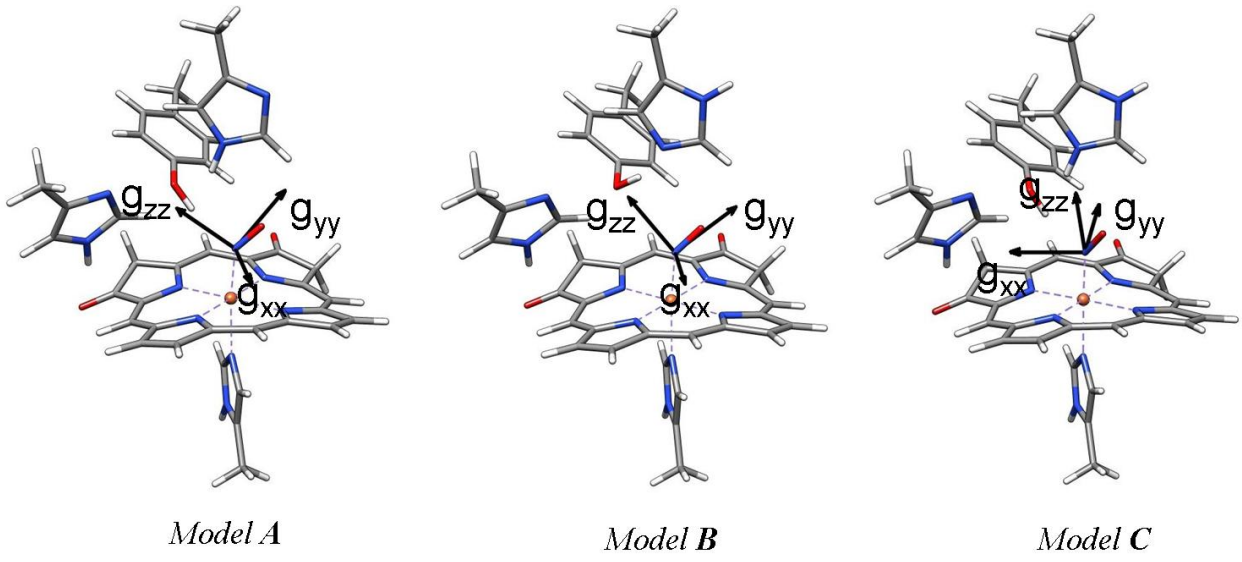


Figure 3.33 Orientation of the g-tensor as obtained from DFT calculations on the frozen structures originating from 1nir. The water molecules are not shown.

^{14}N hyperfine couplings of NO and proximal His_{182}

The calculated ^{14}N hyperfine and nuclear quadrupole parameters of the NO and of the coordinated ^{14}N of the proximal His_{182} of the three models are listed in Table 3.16. The most significant effect is the increase in the ^{14}N(NO) hyperfine interaction when the Tyr_{10} H-bond is removed, due to the shift of spin density onto the NO ligand. The values of *models **A*** and ***C*** are underestimated and are overall similar to those calculated for Mb-NO.[77] In *model **B*** they are overestimated. The principal axis system of the ^{14}N(NO) hyperfine tensor is collinear with that of **g**. The calculated $^{14}N(His_{182})$ hyperfine coupling increases slightly upon removal of the H-$N^{\varepsilon}(His_{369})$ proton which induces the removal of H-bond of Tyr_{10} as well. Consequently, the spin population of $^{14}N(His_{182})$ increases slightly as reflected in the larger calculated hyperfine coupling of $^{14}N(His_{182})$.

The experimental values, listed in Table 3.16, were determined earlier by high field ENDOR and HYSCORE spectroscopy.[231] The spectra of WT-NO and Y10F-NO were the same suggesting that removal of Tyr_{10} does not affect the spin density of $^{14}N(His_{182})$ and most probably also of the bonding properties of the Fe-N_{His} bond. The calculated principal hyperfine components of *models **A*** and ***C*** are underestimated compared to the experimental values, whereas these of ***B*** match the experimental values better, except for A_{yy}. The principal values were derived from earlier simulations. They were carried out under the assumption that the orientation of the ^{14}N hyperfine and quadrupole interactions, which are related to the Fe-N_{His} direction, coincide with the orientation of **g**, where g_{zz} (g_{mid}) was assumed to be along the Fe-N(O) direction. Our recent Mb-NO[77] and current DFT calculations showed that this is not the case and that g_{zz} (g_{mid}) is perpendicular to the NO direction,[231] (Figure 3.33). This wrong relative orientation of **g** and **A** tensors (due to wrong **g** orientation) led to an overestimated A_{yy}. Taking into account the correct orientation (which is similar in all models) we repeated the simulations of the HYSCORE and ENDOR spectra. The new values, (16.1, 16.3, 20.5) MHz match those of *model **B*** in Table 3.16) and are close to the values obtained for Mb-NO, $A^{14}N(His_{93})$ = (17.0, 17.3, 21.5) MHz.[77] The difference only amounts to ~ 6 %, even though the heme structure and the residues in the distal side are quite different. To fit the experimental values the hyperfine couplings values for *model **A*** would have to be increased by ~ 15 – 18 %.

Table 3.16 DFT (PBE) calculation of the hyperfine (A) and nuclear quadrupole (e^2Qq/h) coupling constants (MHz) of ^{14}N of the NO and proximal His_{182} ligand. The corresponding Euler angles (a, b, g) (deg) relate the hyperfine and quadrupole tensor's orientation with g (g_{min}, g_{mid} g_{max}). The values in parentheses were calculated with inclusion of COSMO.

	Exp	*model **A***	*model **B***	*model **C***
^{14}N(NO):				
A_{xx}, MHz	26.9	16.4 (14.6)	36.8 (34.8)	15.9 (15.3)
A_{yy}, MHz	33.5	20.9 (19.3)	39.0 (37.1)	20.9 (20.2)
A_{zz}, MHz	64.5	55.0 (53.5)	82.7 (81.1)	55.5 (55.1)
a_{iso}, MHz	41.6	30.8 (29.2)	52.9 (51.0)	30.8 (30.2)
$(a, b, g)_A$		[90 119 11]	[109 93 2]	[86 59 343]
e^2qQ/h, MHz		-2.499 (-2.554)	-2.437 (-2.486)	-2.720 (-2.715)
h		0.848 (0.823)	0.921 (0.899)	0.733 (0.737)
$(a, b, g)_Q$		[89 119 97]	[110 94 92]	[83 58 252]
$^{14}N(His_{182})$:				
A_{xx}, MHz	16.0	13.1 (13.1)	16.1 (16.0)	11.5 (11.6)

A_{yy}, MHz	19.5	13.4 (13.3)	16.3 (16.3)	11.7 (11.8)
A_{zz}, MHz	19.5	17.2 (17.4)	20.4 (20.6)	15.4 (15.7)
a_{iso}, MHz	18.3	14.6 (14.6)	17.6 (17.7)	12.9 (13.1)
$(a, b, g)_A$		[213 87 4]	[49 90 352]	[31 93 86]
e^2qQ/h, MHz	2.6	-2.268 (-2.112)	-2.347 (-2.209)	-2.110 (-2.024)
h		0.597 (0.687)	0.555 (0.628)	0.701 (0.758)
$(a, b, g)_Q$		[140 95 87]	[54 90 266]	[35 87 355]

Hyperfine couplings of the Tyr_{10} and His_{369} protons

The calculated hyperfine and quadrupolar interactions of the H-bonded protons of Tyr_{10} and His_{369} for *models **A***, ***B***, ***C*** are listed in Table 3.17. The larger spin density on the N(NO) compared to the O(NO) (see Table 3.15) makes ENDOR more sensitive to H-bonds to the N(NO) because of their larger resolved ^{1}H hyperfine splitting. Remarkably, out of the three residues in the distal part of the heme pocket only the H-bonded proton of Tyr_{10} gives rise to large hyperfine couplings in all models. Simulations of the ^{1}H Davies and ^{2}H Mims ENDOR spectra with the DFT parameters of *models **A*** and ***C*** (listed in Table 3.17), resulted in splittings that were significantly larger than those observed experimentally. When we reduced A_{xx}, A_{yy} and A_{zz} of Tyr_{10} OH of *model **A*** by a factor of 1.8 without changing the orientation and the quadrupolar interaction parameters a satisfactory fit was obtained for both the ^{1}H and ^{2}H ENDOR spectra, as shown in Figure 3.34 (a) and (b) respectively. Simulations using the parameters of *Model **C***, also with scaling of 1.8 for the hyperfine of Tyr_{10} gave a reasonable fit, though not as good as for *model* **A**. In *model **B*** one of the water protons has an A_{zz} value close to 8 MHz, which is close to the experimentally observed A_{max}. Simulations using the calculated hyperfine parameters of this proton did not reproduce the lineshape evolution of the orientation selective spectra.

On the basis of the DFT calculations we assign the large couplings observed in WT-NO to the H-bonded proton of Tyr_{10}. We hypothesized that the rather large scaling factor required to fit the DFT results to the experimental spectra might originate from problems with the crystal derived structure constraints on our model. To address this question, a rigid scan of the N(NO)-H(O-Tyr_{10}) distance was conducted while keeping the relative orientations fixed. The distance was increased in steps of 0.2 Å. For a distance of 2.57 Å the obtained hyperfine values, A= (-3.5,-2.9, 8) MHz, are close to the experimental values (A=(-4.2,-3.4, 7.7) MHz). This distance would amount to a change of ~0.5 Å in the N(NO)-H(O-Tyr_{10}) distance. We do think that movements of this size relative to the crystal structure may well take place and may help to explain the apparent discrepancy between theoretical and experimental values.

Table 3.17 DFT (PBE) calculation of hyperfine (A) and nuclear quadrupole (e^2Qq/h) coupling constants (MHz) of protons of Tyr_{10} and His_{369} which are H-bonded to the NO group. The corresponding Euler angles (a, b, g) (deg) relate the hyperfine and quadrupole tensor's orientation with respect to g (g_{min}, g_{mid} g_{max}). The values in parentheses were calculated with COSMO.

	*model **A***	*model **B***	*model **C***
^{1}H/^{2}H(Tyr_{10}):			
A_{xx}, MHz	-7.9 (-7.6)	-3.1 (-3.1)	-8.6 (-8.4)
A_{yy}, MHz	-6.5 (-6.2)	-1.9 (-1.9)	-6.8 (-6.6)
A_{zz}, MHz	13.6 (14.0)	4.3 (4.4)	14.6 (14.9)
a_{iso}, MHz	-0.3 (0.1)	-0.2 (-0.2)	-0.3 (0.0)
$(a, b, g)_A$	[81 124 127]	[77 106 300]	[73 53 75]

e^2qQ/h, MHz	0.249 (0.248)	0.166 (0.166)	0.241 (0.241)
h	0.12 (0.119)	0.175 (0.175)	0.124 (0.124)
$(a, b, g)_Q$	[74 103 195]	[0 53 348]	[40 55 104]
$^1H/^2H(His_{369})$:			
A_{xx}, MHz	-2.4 (-2.4)	-	-2.7 (-2.7)
A_{yy}, MHz	-1.6 (-1.6)	-	-1.6 (-1.6)
A_{zz}, MHz	4.0 (4.0)	-	4.2 (4.2)
a_{iso}, MHz	0.0 (0.0)	-	0.0 (0.0)
$(a, b, g)_A$	[58 95 296]		[42 71 50]
e^2qQ/h, MHz	0.26 (0.258)	-	0.219 (0.221)
h	0.088 (0.087)	-	0.082 (0.085)
$(a, b, g)_Q$	[21 85 107]		[202 76 345]

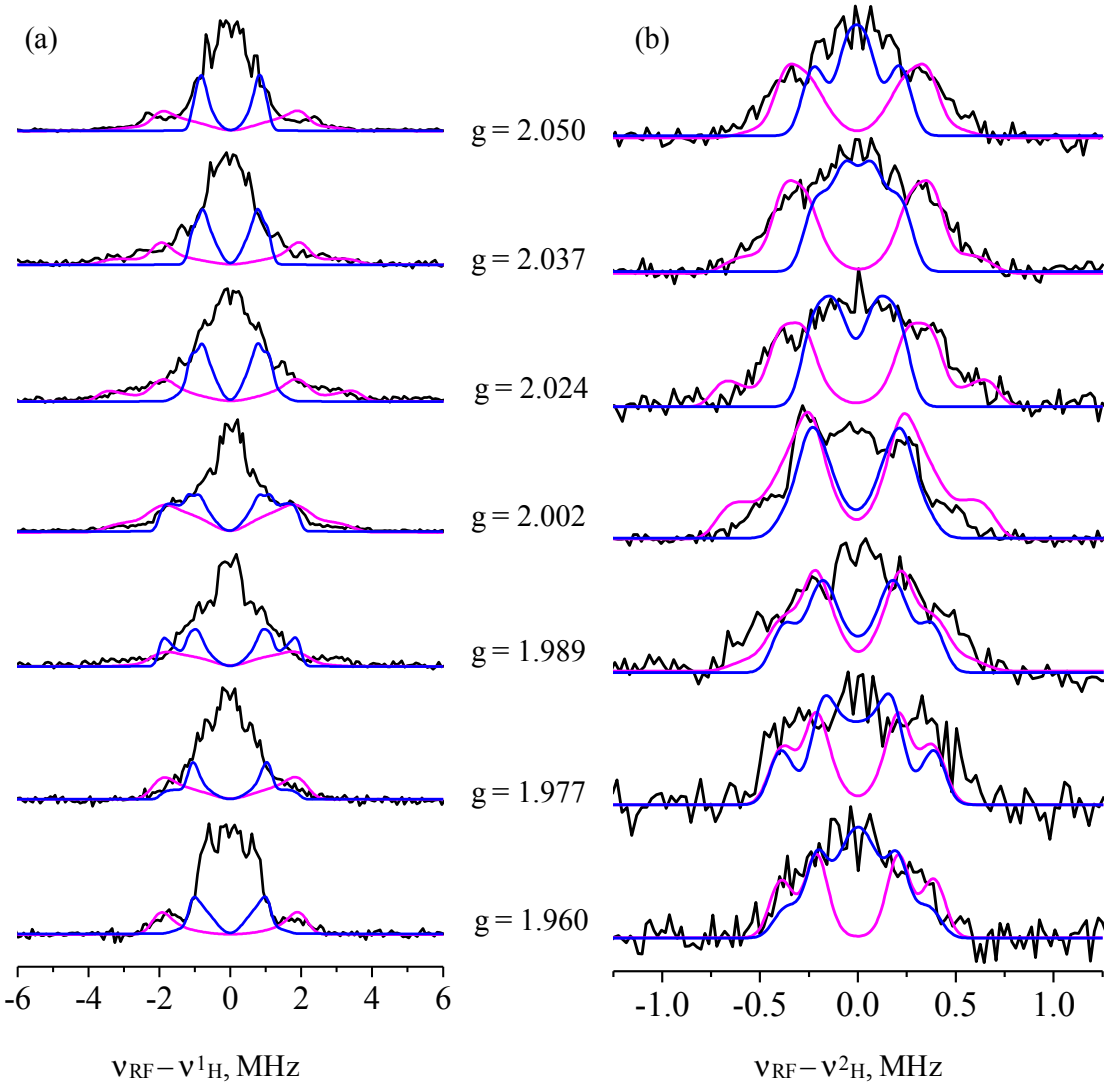


Figure 3.34 Experimental (black) and simulated (magenta/blue) ^{1}H Davies ENDOR spectra (a) and ^{2}H Mims ENDOR spectra (b) of a frozen solution of the WT-NO. The simulations of the H-bonded proton of Tyr_{10} (magenta) and His_{369} (blue) are shown individually. The parameters used for the simulation are given in Table 3.17 for *Model* **A** except the hyperfine couplings of Tyr_{10} that were scaled down by a factor of 1.8.

In all models, except for one water molecule in *model* **B**, the calculated hyperfine couplings of the water molecules were relatively small and therefore they were not included in the simulations. To summarize, of the three *models **A***, ***B*** and ***C***, we do not find a single model that best fits *all* of the relevant experimental parameters: the principal values of **g**, of the hyperfine interaction of ^{14}N(NO) and ^{14}N(His_{182}) and the ^{1}H/^{2}H ENDOR spectra. *Model **B*** provides better agreement with the ^{14}N(His_{182}) data while in *model **A*** these hyperfine values were off by 15–18 %. In contrast, *model **A*** could reproduce the orientation selective ^{1}H/^{2}H ENDOR spectra reasonably well (better than *model **B***) after invoking a single scaling factor for the hyperfine coupling of the exchangeable protons of Tyr_{10}. The necessity of this scaling factor may well have to do with problems with the crystal structure derived constraints. For *model **B*** it would be difficult to account for the spectra of Y10F as His_{369} does not have a proton to contribute to the structure. Therefore we chose *model **A*** as the preferred model.

We have also checked the agreement between the geometrically relaxed *models **A***, ***B*** and ***C*** originating from 1nir structure and optimized without constraints. The behavior, in general, is similar to the frozen models and the ^{1}H-O(Tyr_{10}) in *models **A*** and ***C*** is close to the NO. While the ^{14}N(His_{182}) hyperfine values remained almost unchanged, the agreement of the calculated orientation selective ^{1}H/^{2}H ENDOR spectra with the experimental ones was inferior to the frozen structures because of changes in ligand orientation that would be prevented by the protein backbone in the real system.

Unlike the structures that originated from the oxidized enzyme, where His_{327} did not show any significant H-boning to the NO, in the relaxed structures that originated for the reduced NO-bound structure (pdb 1nno) all three residues, His_{327}, His_{369} and Tyr_{10}, depending on the histidines' protonation states, formed H-bonds with the NO. The EPR parameters obtained from the nine structures gave in general similar agreement to that observed in *models **A**-**C*** for the ^{14}N hyperfine couplings of NO and His_{182}. Here too, protonation of N^{ε} of His_{369} permits forming of an H-bond between the ^{1}H-O(Tyr_{10}) and NO. Out of the nine models we selected those that gave the largest ^{1}H hyperfine coupling, *model **5*** and ***9***, and calculated the orientation selective ^{1}H/^{2}H ENDOR spectra of all three residues, His_{327}, His_{369} and Tyr_{10}. These are compared with the experimental spectra.

The agreement for *model **5*** is not satisfactory. M*odel **9***, is rather similar to *model **A*** in terms of the positions of distal residues with respect to the NO. In this model all nitrogens of His_{369} and His_{327} are protonated and His_{369} and Tyr_{10} form H-bonds with the NO both generating comparable significant hyperfine couplings. The main difference relative to *model **A*** is the longer distance of the ^{1}H-O(Tyr_{10}) to the N(NO) and a closer distance to the O(NO) in *model **9***. Here the magnitude of the hyperfine couplings are close to the experimental values but the lineshape evolution of the orientation selective spectra is not well reproduced and is inferior to *model* **A**. Interestingly, in this *model* the ^{14}N(His_{182}) hyperfine couplings are also underestimated compared to the experiment (by ~ 10 %).

3.4.6 Discussion

In this work we focused on the identification and characterization of the H-bonds formed by amino acid residues in the distal pocket of the reduced d_1-heme with the NO ligand in WT-NO and two mutants, Y10F –NO and dHis-NO, in frozen solutions. The protonation state of the histidines is unknown because the resolution of the X-ray determined structure of WT-NO is too low to allocate the position of the protons. We have used W-band ENDOR to detect the H-bonds and carried out detailed DFT calculations on structures with various possible protonation states of the distal histidines to interpret the experimental results and related them to structure. These revealed a correlation between the histidines' protonation states and the position of the residues and the H-bonds to the NO. The calculations also highlighted expected conformational changes upon variations in the H-bond network that were in line with the experimental observations found in the mutants. The DFT calculations showed that for the structure that agrees best with the experimental results *(model* **A**) the conserved His_{327} residue has no H-bonds to the NO. It was, however, previously shown to play an important role in catalysis.[65,71] A possible explanation of this finding can be that His_{327} maintains a positive charge in the active site, which in turm helps to guide negatively charged substrates like nitrite to the active centre. In WT-NO (pH = 7.0) we identified two H-bonds to the NO group, one is from the OH of Tyr_{10} and the other from the protonated N^{ε} of His_{369}. This is in contrast to the crystal structure of WT-NO which shows that upon reduction and NO binding the Tyr_{10} is removed from the close vicinity of the iron and is too far to form an H-bond with the NO.

The experimental results on WT-NO and its mutants, together with the DFT calculations indicate that the H-bond network is dynamic. Comparison of the ^{1}H Davies ENDOR spectra of WT-NO and Y10F-NO clearly reveals a disappearance of signals and appearance of new ones (see Figure 3.29). Because in Y10F Tyr is removed, we assigned the disappearing signals in WT-NO to the H(OH) of Tyr_{10} and the new signals in the Y10F-NO ENDOR spectra to an H-bond of the exchangeable proton of His_{369}. From the analysis of the ENDOR spectra and the DFT results we suggest a cooperative behavior for Tyr_{10} and His_{369}. The H-bond formed between the protonated N^{ε} of His_{369} to O(NO) allows positioning Tyr_{10} next to the NO. Upon mutation of the Tyr_{10} to Phe, His_{369} becomes closer to the NO yielding a larger ^{1}H hyperfine splitting for the H-bonded proton. Unlike Y10F-NO, a distinct substantial hyperfine coupling corresponding to a proton H-bonded to the NO has not been observed for the dHis mutant. This suggests that in this mutant, the Tyr_{10} is displaced such that it cannot form an H-bond with the N(NO). The total width of the spectrum, however, suggests that there may be a water molecule (or more) coordinated to the NO ligand.

Figure 3.35 summarizes the dynamic H-bonding network and conformational changes caused by mutations based on EPR experimental data and the DFT calculations. It is clear that a change in one of the residues within the d_1-heme active site does affect the strength and position of the H-bonds formed by the other residues. In fact, in Y10F-NO His_{369} moves closer to the NO and its hydrogen bond becomes shorter, whereas mutation of the both distal histidines displaces Tyr_{10} removing its H-bond. Accordingly, one may speculate that any event that changes the location or protonation state of one residue is likely to affect the whole active site structure. Moreover, the DFT calculations of the bonding energies, charge and spin distributions suggest that the hydrogen bond

network of the active site may play some role in the product release. As discussed at length previously, the low-spin $\{Fe(NO)\}^7$ complexes are known to be very stable and thus the NO release should be the very slow if it occurs from the Fe(II) state. Nevertheless, a recent kinetic investigation of Rinaldo *et. al.*[65] has shown that NO dissociates from the reduced d_1-heme considerably faster than from other hemes.

Several electronic effects that are responsible for the stability of the ferrous-nitroso complexes: the amount and strength of the H-bonds in the active site, the back bonding between the ferrous iron and the NO adduct and the spatial conformation of the NO on the porphyrine plane. In the initial step of the reduction process, namely the substrate attraction and activation stage, the active site histidines are assumed to be in a protonated charged form to create positive electrostatic potential responsible for the substrate attraction. After the substrate binding process the reduction proceeds through one of the N-O bonds cleavage and the formation of a water molecule, converting ferrous-nitrite complex to the $\{Fe(NO)\}^6$ adduct (iron is in ferric state). At this stage the NO can be released from the Fe(III) state not forming the stable $\{Fe(NO)\}^7$. However, in the case that there is an electron transfer from the educed c-heme to the d_1-heme before the NO release the stable $\{Fe(NO)\}^7$ complex forms. In this case, if N^{ε}-His_{369} remain deprotonated and Tyr_{10} is not H-bonded (*model* ***B***) NO release may facilitated by the small number of H-bonds. Alternatively, if the N^{ε} His_{369} does reprotonate fast (*models* ***A*** and ***C*** that fit the experimental results), forming an H-bond to the O(NO) that stabilize the Fe(II)-NO bond, this stabilization is counter acted by the H-bond of the Tyr_{10} with the N(NO). The latter populates the (partial sp^2) N based nonbonding orbital thus weakening the Fe–N(NO) bond and this in turn should facilitate NO release.

The last effect is a formation of unfavorable NO orientation on the porphyrine plane which is believed to be largely due to the H-bonded Tyr_{10}. It is known that rotation around Fe-N(NO) bond is hindered by the back-donation interactions between metal and NO ligand. As a consequence the potential energy surface is rather rigid and even small displacement around the metal-ligand bond can cause significant drop in binding energy.

To summarize, the above described electronic effects, although none of them is strong, combined they may have a significant joint impact on the NO release process from the ferrous active site. It is worthwhile to point out that only a highly mobile and flexible active site H-bonding network is suitable for accomplishing the effective reduction process. The positively charged histidines attract the substrate, donate protons, thereby cleaving one N-O bond in NO_2^-, and work in concert with Tyr_{10}, which, after quick reduction of the metal, forms H-bond to N(NO), distorting the NO spatial conformation and lowering the amount of back-donation from the Fe(II). Hence, we suggest that role of Tyr_{10} can be considered as a measure against the formation of a too stable $\{Fe(NO)\}^7$ product by facilitating NO release. To corroborate this suggestion the NO release rates from Y10F-NO and WT from *Pseudomonas stutzeri,* which that does not have an equivalent of Tyr_{10} should be measured.

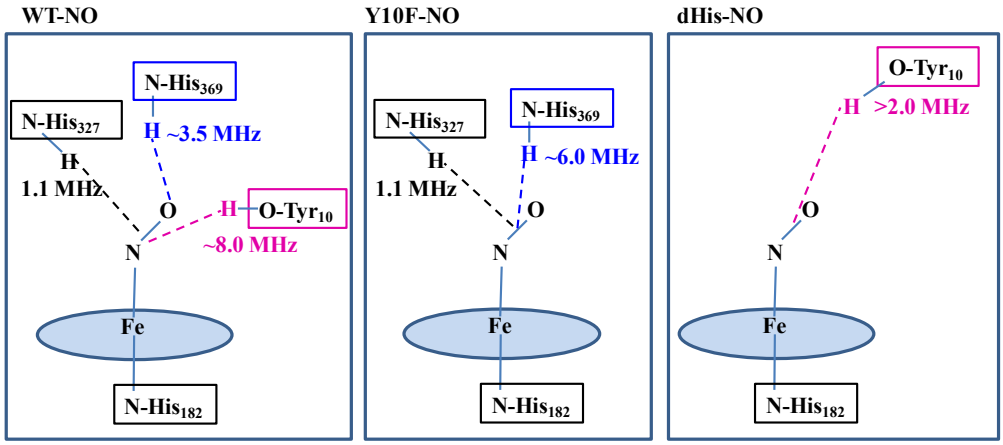


Figure 3.35 Proposed conformational changes in the distal pocket of the d_1-NO complex caused by mutation. The dashed lines correspond to the H-bonds. When it is not known whether the bond is with O(NO) or N(NO) we mark it midway.

Another interesting observation of this work is the rather insensitivity of the hyperfine coupling of the ^{14}N of the axial ligand and the NO to variation in the H-bond network to the NO at the distal pocket. Even more surprising is the similarity to Mb-NO (within ~ 6 %) which has a different heme altogether and only one H-bond to the NO. Finally, both Y10F-NO and dHis-NO were found to exhibit weak interactions with distant water molecules, suggesting some opening of the pocket compared to the WT. This observation implies that the conformational changes observed in the crystal structure of the single mutants (H327A and H369A), involving the c-heme domain relocation and opening of the distal pocket, are likely to take place also in solution. On the other hand, in the WT enzyme, all available data suggest that the active site is more "closed" to solvent.

3.4.7 Conclusions

In this work we have once more demonstrated the value of high field ENDOR combined with DFT calculations for the characterization of H-bonds and highlighted the important role of DFT in assigning and interpreting ENDOR spectra in terms of structural details for complicated systems. We have shown that the NO in the nitrosyl d_1-heme complex of cd_1 NIR forms H-bonds with Tyr_{10} and His_{369} that affect the stabilization of the NO complex. The second conserved His_{327}, appears to be less involved in NO stabilization by H-bonding. This is in contrast to the WT-NO crystal structure where Tyr_{10} is not at an H-bond distance from the NO. We have also observed a larger solvent accessibility to the distal pocket in the mutants compared to the WT. Finally, it is clear from this work that the H-bonding network within the active site is dynamic and that a change in one of the residues does affect the strength and position of the H-bonds formed by the other ones. In the Y10F mutant His_{369} moves closer to the NO and its hydrogen bond is shorter, whereas mutation of both distal histidines displaces Tyr_{10} removing its H-bond.

3.5 The electronic structure of the unique [4Fe-3S] cluster in O_2-tolerant hydrogenases characterized by ^{57}Fe Mössbauer and EPR spectroscopy

3.5.1 Abstract

Iron-sulfur clusters are ubiquitous electron transfer cofactors in hydrogenases. Their types and redox properties are important for H_2-catalysis, but recently their role in a protection mechanism against oxidative inactivation has also been recognized for a new [4Fe-3S] cluster in O_2-tolerant Group 1 [NiFe] hydrogenases. This cluster, which is uniquely coordinated by six cysteines, is situated in the proximity of the catalytic [NiFe] site and exhibits unusual redox versatility. The [4Fe-3S] cluster in Hydrogenase I (Hase I) from *Aquifex aeolicus* performs two redox transitions within a very small potential range, forming a super-oxidized state above +200 mV *vs* SHE. Crystallographic data has revealed that this state is stabilized by the coordination of one of the iron atoms to a backbone nitrogen. Thus, the proximal [4Fe-3S] cluster undergoes redox-dependent changes to serve multiple purposes beyond classical electron transfer. Here we present field-dependent ^{57}Fe-Mössbauer and EPR data for Hase I, which in conjunction with spectroscopically calibrated DFT calculations reveal the distribution of Fe valences and spin-coupling schemes for the iron-sulfur clusters. The data demonstrate that the electronic structure of the [4Fe-3S] core in its three oxidation states closely resembles that of corresponding conventional [4Fe-4S] cubanes, albeit with distinct differences for some individual iron sites. The medial and distal iron-sulfur clusters have similar electronic properties as the corresponding cofactors in standard hydrogenases, although their redox-potentials are higher.

3.5.2 Results

Hase I from *A. aeolicus* grown on ^{57}Fe enriched media[80] was treated with hydrogen, ascorbate, and porphyrexide or air to obtain fully reduced, oxidized and superoxidized samples, as listed in Table 3.18. The redox states of the cofactors were assigned according to our previous titration data,[80] whereby EPR measurements on aliquots of the Mössbauer samples were used to verify the spin counts of the (paramagnetic) cofactors, as well as EPR-silence of the proximal $[4Fe\text{-}3S]^{4+}$ cluster in the ascorbate-treated sample.

The UV-visible spectra of super-oxidized Hase I prepared in air show broad bands around 400 nm, arising from S-*to*-Fe charge transfer transitions.[237] The spectra resemble those of other oxidized O_2-tolerant Group I enzymes,[238] but they cannot be distinguished from those of standard [NiFe] hydrogenases without super-oxidized [4Fe-3S] clusters. Interestingly, however, these features could not be discriminated from those of oxidized HiPIP proteins[239] or alkylated ferredoxin:thioredoxin reductase (NEM-FTR) with cubane clusters in the corresponding 3+ state.[240] Likewise, EPR spectra of the [4Fe-3S] cluster in 3+ and 5+ states, measured previously[80] resemble those of [4Fe-4S] centers in the corresponding 1+ and 3+ states.[240,241,242,243] This is not completely unexpected, since the 4Fe cores are formally isoelectronic and the modified [4Fe-3S] cores differ from classical cubanes only by an inorganic S^{2-} ligand. Although the cluster oxidation numbers are no observables,

they could be substantiated from electric and magnetic Mössbauer hyperfine interactions, probing the local oxidation and spin states of the metal ions.

Electric Mössbauer Parameters and Valence States.

Mössbauer spectra were measured with reduced, oxidized, and super-oxidized samples of Hase I. The zero-field spectra recorded at 160 K could be decomposed in a number of quadrupole doublets without paramagnetic broadening due to fast spin relaxation (Figure 3.37). Five or six distinct subspectra could be identified that arise from [NiFe] and distinct groups of the 11 iron sites within the valence-delocalized FeS cofactors. Primarily we used three subspectra with constrained relative intensities for the iron sites, which have the formal oxidation states "Fe^{2+}", "$Fe^{2.5+}$", and "Fe^{3+}". In addition, a subspectrum with 8% intensity was adopted for the low-spin Fe(II) site in the [NiFe] cluster (1/12 of the total iron content).

Table 3.18 FeS Cofactors [a)] and Redox States of Hase I.

prep./site	*proximal*	*medial*	*distal*	*[NiFe]*
H_2-red[b)]	$[4Fe\text{-}3S]^{3+}$ *S*=1/2	$[3Fe\text{-}4S]^{0}$ *S*=2	$[4Fe\text{-}4S]^{1+}$ *S*=1/2	Ni-R *S*=0
oxid.[c)]	$[4Fe\text{-}3S]^{4+}$ *S*=0	$[3Fe\text{-}4S]^{1+}$ *S*=1/2	$[4Fe\text{-}4S]^{2+}$ *S*=0	Ni-B *S*=1/2
super-oxid.[d,e)]	$[4Fe\text{-}3S]^{5+}$ *S*=1/2	$[3Fe\text{-}4S]^{1+}$ *S*=1/2	$[4Fe\text{-}4S]^{2+}$ *S*=0	Ni-B *S*=1/2

a) Cluster oxidation numbers are obtained from the spectroscopically derived oxidation numbers of Fe with the core sulfides taken in their closed shell form (S^{2-}). b) Fully reduced with H_2 gas; c) Partially reduced with ascorbate d) Preparation in air yields ~20% contamination with the proximal $[4Fe\text{-}3S]^{4+}$ cluster, accounted for in Mössbauer simulations; e) full super-oxidation with porphyrexide.

Moreover, a unique feature of the Hase I spectra in all three different preparations are resolved lines at the wings of the absorptions, which reveal an additional special subspectrum (S) with moderate isomer shift, but remarkably large quadrupole splitting ($\delta/|\Delta E_Q|$ = 0.51/2.61 mms^{-1}, 0.48/2.23 mms^{-1}, and 0.46/2.42 mms^{-1} for reduced, oxidized, and super-oxidized samples). Its relative intensity, ~7.3 to 8.6%, accounts for a fully occupied, genuine iron site. Since this subspectrum was not observed for standard hydrogenases[244,245,246] (although the distal cluster in all known [NiFe] hydrogenases has an uncommon histidine-coordinated iron corner), we assign it to the modified [4Fe-3S] cluster, present only in O_2-tolerant hydrogenases. This is corroborated by the applied-field spectra given below, as such a large quadrupole splitting could neither be assigned to a part of the diamagnetic distal cluster in oxidized samples nor to the integer-spin *S*=2 component of the medial 3Fe cluster.

In addition to the five subspectra introduced so far ("Fe^{2+}", "$Fe^{2.5+}$", "Fe^{3+}", special site "S", and Fe(II) low-spin), we had to add a sixth subspectrum with markedly larger isomer shift ($\delta/\Delta E_Q$ = 0.72/1.56 mms^{-1}) to account for a second distinct site with rather localized ferrous character, occurring only in the spectrum of the H_2-reduced enzyme (Figure 3.36). This component (F) was

also required for consistent simulation of the magnetic Mössbauer spectra and was likewise attributed to a site of the reduced novel [4Fe-3S]$^{3+}$ cluster (see below).

In summary, the zero-field Mössbauer spectra of Hase I obtained under three different redox conditions could be well simulated by using a scheme of subspectra for one 3Fe and two 4Fe clusters, of which the proximal [4Fe-3S] cluster can be super-oxidized and occur in three oxidation states. These states are isoelectronic to the formal [3Fe^{2+}, Fe^{3+}], [2Fe^{2+}, 2Fe^{3+}] and [1Fe^{2+}, 3Fe^{3+}] motifs recognized for classical [4Fe-4S] centers (although physiologically these three never occur in one system). The overall increase of ferric character in the series is reflected in the Mössbauer subspectra by a decrease of the average isomer shifts ($<\delta>_{[4Fe\text{-}3S]}$ = 0.52, 0.43, 0.36 mms^{-1}).

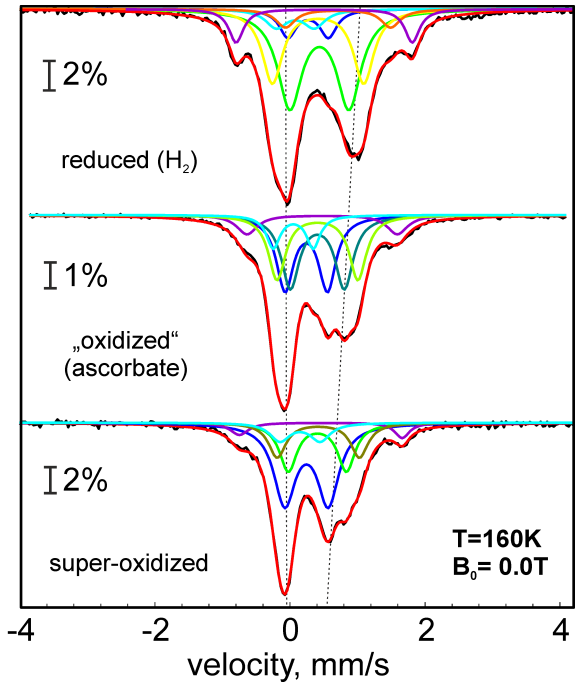


Figure 3.36 Zero-field Mössbauer spectra of Hase I recorded at 160 K. The subspectra are Lorentzian doublets (yellow: "Fe^{2+}"; green: "Fe$^{2.5+}$"; blue: "Fe^{3+}"; purple: special site (S) of [4Fe-3S] clusters; orange: ferrous site (F) of reduced [4Fe-3S] cluster in the H_2-reduced sample; light-blue: low-spin Fe(II) in [NiFe]; red: sum of subspectra). The dotted vertical lines represent the ferric character increasing with oxidation state.

Magnetic Mössbauer Spectra and Spin Distribution.

Applied-field Mössbauer spectra of fully reduced and super-oxidized Hase I recorded at 4.2, 80, and 160 K are shown in Figure 3.37. Since the magnetic spectra exhibit higher site selectivity than the zero-field spectra, not least because the signs of electric field gradients (*efg*) and induced internal fields can be resolved from the field dependence of the spectra,[245,246,247] separate subspectra had to be introduced for all 12 iron sites. To this end, the 5-6 collective subspectra introduced above for the different classes of mixed-valence irons sites seen in the zero-field spectra were broken up into their cluster contributions. For the sake of clarity we present the result in four collective subspectra corresponding to the three FeS clusters and the [NiFe] center.

Spin relaxation was found to be slow for the paramagnetic centers at 4.2 K and fast at 80 and 160 K, which means that the internal fields virtually average to zero at the high temperatures. The condition is valuable for assigning *efg* parameters, *i.e.* the sign of ΔE_Q and the asymmetry parameter η. The zero-field values of δ and ΔE_Q were taken in this program as starting points and valuable constraints for global optimization of the magnetic spectra (allowing only for the usual temperature variations).

In contrast, the large magnetic splitting observed at 4.2 K indicates strong internal fields at some iron nuclei due to strong magnetic hyperfine coupling with static spin expectation values, $\vec{B}_i^{\text{int}} = -\tilde{A}_i / g_N \mu_N \cdot < \vec{S}_i >$ [247] The temperature and field dependence of the hyperfine coupling could be simulated well with $< \vec{S}_i >$ values derived from the usual spin Hamiltonian for clusters with strong intrinsic spin coupling. However, for super-oxidized enzyme also weak isotropic coupling between the proximal cluster and Ni-B as well with the medial cluster had to be introduced with J values up to ~0.02 cm^{-1}. This coupling, which causes reduction of magnetic splittings at weak fields, had been established first by X- and W-band EPR.[80]

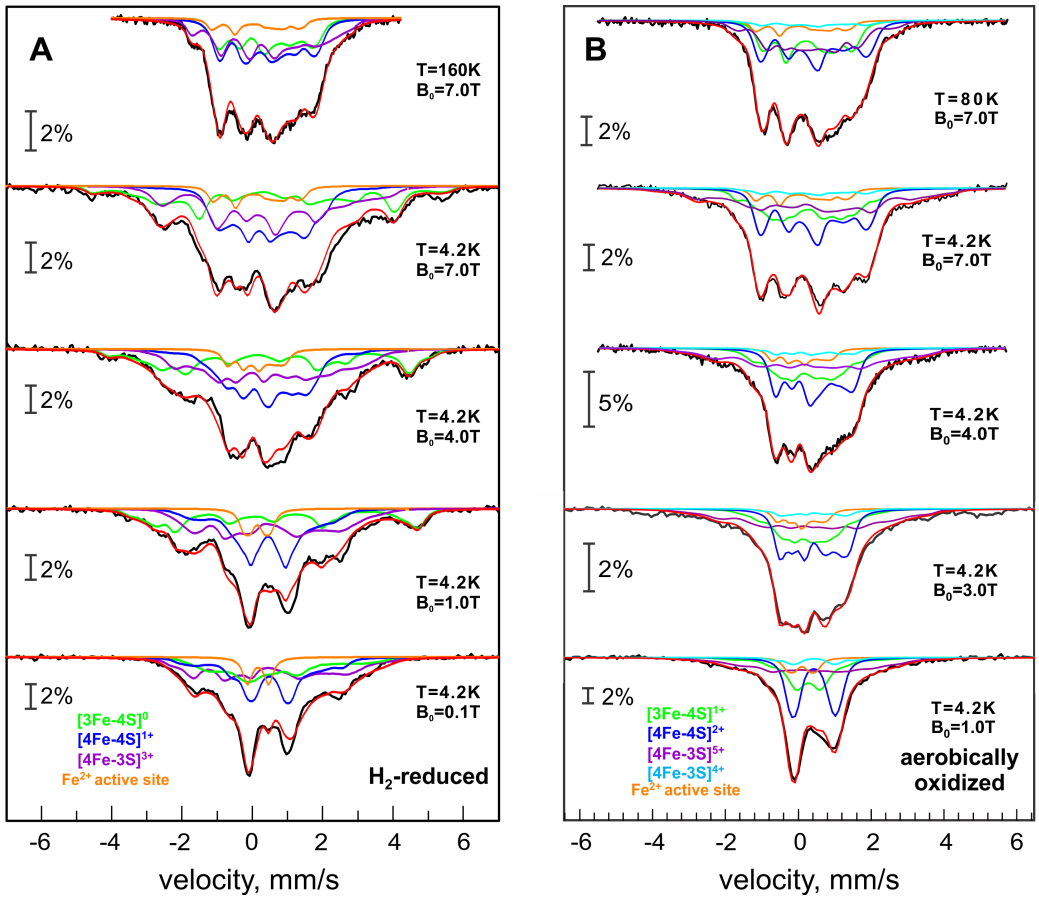


Figure 3.37 Magnetic Mössbauer spectra of Hase I (A) reduced with H_2 and (B) super-oxidized with air. Fields of 0.1 to 7 T were applied perpendicular to the γ-ray beam. The solid lines represent global simulations with the following contributions, (A): proximal $[4Fe\text{-}3S]^{3+}$ (purple), medial $[3Fe\text{-}4S]^{0}$ (light green), distal $[4Fe\text{-}4S]^{1+}$ (blue); (B): proximal $[4Fe\text{-}3S]^{5+}$ (purple), medial $[3Fe\text{-}4S]^{1+}$ (light green), distal $[4Fe\text{-}4S]^{2+}$ (blue). Fe(II) of the [NiFe] site (orange) occurs in both panels.

The signs of the local magnetic hyperfine coupling constants found for the ^{57}Fe-nuclei are consistent with the usual spin distribution known for 3Fe and 4Fe clusters in their different magnetic ground states and spin coupling schemes. Iron sites with majority spin density (↓) have negative *A*-tensors with respect to the cluster spin, whereas minority spin sites (↑) have positive *A*-values.[247] Interestingly, the classical spin distribution pattern also holds for the proximal [4Fe-3S] cluster, which in this respect resembles its isoelectronic unmodified analogs.

In the H_2-reduced enzyme, all iron-sulfur centers are paramagnetic and exhibit appreciable hyperfine splitting at 4.2 K (Figure 3.37A). Only the low-spin Fe(II) site of the [NiFe] center shows weak magnetic splitting from the applied field only, as expected for the diamagnetic Ni-R state (orange trace). The *distal* $[4Fe\text{-}4S]^{1+}$ cluster (blue) may be described by a delocalized mixed-valence pair $Fe^{2.5+}$-$Fe^{2.5+}$ with sub-spin (S^*=9/2) and a ferrous pair with (S^*=4).[91,92] Accordingly, the spin minority ferrous pair has positive hyperfine parameters *A*, whereas the $Fe^{2.5+}$- $Fe^{2.5+}$ pair has negative *A*-values (Table 3.19A). Although the cluster is known to have a modified Fe subsite with a fifth N_ε nitrogen ligand from a histidine,[78,81,82,84] the spectra do not exhibit unusually large δ or ΔE_Q. Similar to previous studies on [NiFe] hydrogenases,[244,245,246] the effect may be small and not resolved in the present spectra.

Table 3.19 Mössbauer parameters of the reduced and super-oxidized Hase I obtained from magnetic-field and temperature dependent measurements (Figure 3.37).

Center	***Mössbauer Parameters***					
	δ, mms^{-1}	ΔE_Q, mms^{-1}	η [a]	$A_{xx}/g_N\mu_N$ T	$A_{yy}/g_N\mu_N$ T	$A_{zz}/g_N\mu_N$ T
			(**A**) H_2-reduced			
	+0.42	+0.84	0.3	-15.0	-30.0	-30.0
[4Fe-3S]$^{3+}$ (proximal)	+0.44	+1.23	0.9	+12.0	+12.0	+26.0
	+0.71	+1.52	0.3	-30.0	-32.0	-32.0
	+0.50	+2.60	0.1	+12.0	+15.0	+15.0
	+0.30	+0.60	1.0	+6.0	+12.0	+12.7
[3Fe-4S]^{0} (medial) [b]	+0.43	+1.44	0.5	-16.0	-15.0	-12.3
	+0.45	+1.32	0.5	-16.0	-15.0	-12.3
	+0.46	+1.06	0.0	+2.8	+2.8	+2.8
[4Fe-4S]$^{1+}$ (distal)	+0.44	+1.00	0.0	+2.8	+2.8	+2.8
	+0.43	+0.95	0.4	-8.0	-30.0	-30.0
	+0.43	+0.96	0.0	-6.0	-28.0	-30.0
[NiFe] (Ni-R state)	+0.13	+0.56	-	-	-	-
			(**B**) Super-oxidized with air			
	+0.39	+0.70	1.0	-9.3	-51.0	-43.9
[4Fe-3S]$^{5+}$ (proximal)	+0.40	+1.00	0.7	-43.0	-13.0	-17.0
	+0.28	+0.60	0.3	+2.2	+25.4	+17.4
	+0.46	+2.45	0.4	+4.4	+27.7	+18.9
[3Fe-4S]$^{1+}$ (medial)	+0.28	+0.60	0.9	-18.0	-27.8	-29.0
	+0.27	+0.60	0.2	+9.0	+9.0	+9.0
	+0.27	+0.60	0.2	+2.5	+2.5	+2.5
	+0.44	+1.25	1.0			
[4Fe-4S]$^{2+}$ (distal)	+0.43	+1.06	0.0			
	+0.40	+0.95	1.0			
	+0.40	+0.95	1.0			
[NiFe] (Ni-B state)	+0.10	+0.60	-	-	-	-

[a] $\eta = (V_{xx}\text{-}V_{yy})/V_{zz}$ is the asymmetry of the electric field gradient; [b] The zero-field splitting parameters D and E/D for S = 2 of the $[3Fe\text{-}4S]^0$ cluster were taken as D = -2.20 cm-1, E/D= 0.26.

The medial $[3Fe\text{-}4S]^0$ cluster (light green, Figure 3.37A) was simulated with S=2, arising from antiferromagnetic coupling of a delocalized mixed-valence $Fe^{2.5+}$-$Fe^{2.5+}$ pair (S*=9/2) and a ferric ion (S=5/2). The effective hyperfine values resemble those of other [3Fe-4S] clusters,[247,248] with negative A for the mixed valence pair (majority spin) and positive A for the ferric site. The zero-field splitting parameters D = -2.20 cm^{-1}, E/D = 0.26 are typical for $[3Fe\text{-}4S]^0$ clusters in hydrogenases.[244]

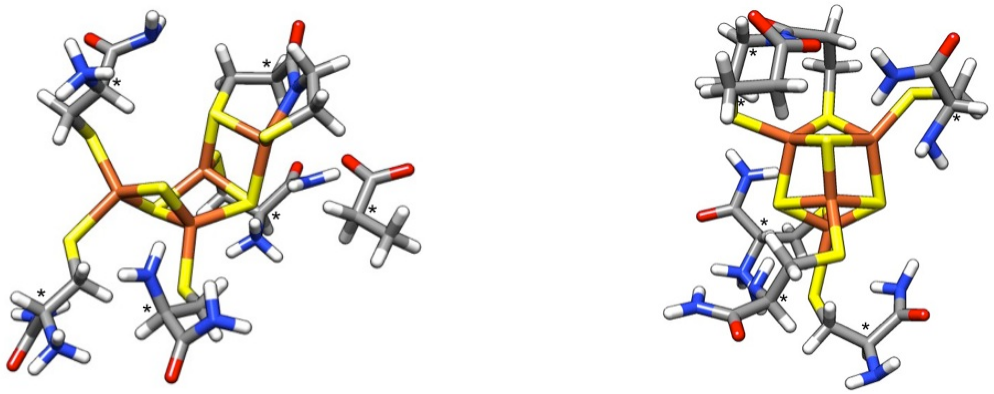
The g-values for the reduced proximal [4Fe-3S] cluster, as taken from previous EPR work,[80] are similar to those of the isoelectronic conventional $[4Fe\text{-}4S]^{1+}$ clusters.[241,249] Correspondingly, the magnetic Mössbauer spectra (purple trace, Figure 3.37A) could be fitted with the same spin coupling scheme found for reduced, low-potential ferredoxins; a delocalized mixed-valence pair $Fe^{2.5+}$-$Fe^{2.5+}$ (S*=9/2) is antiferromagnetically coupled to a ferrous Fe^{2+}-Fe^{2+} pair (S*=4). Based on the signs of the A values, the subspectra with d = 0.50 mms^{-1} / DE_Q = +2.60 mms^{-1} and with d = 0.44 mms^{-1} / DE_Q = +1.23 mms^{-1} are ascribed to the ferrous pair (positive A). The result corroborates the oxidation state +3 for the reduced [4Fe-3S] core (spin coupling scheme and Mössbauer parameters as for 3 Fe^{2+}, 1 Fe^{3+}, completed with 3 S^{2-}; Fig. 1). The A-values for the $[4Fe\text{-}3S]^{3+}$ cluster are generally similar to those of $[4Fe\text{-}4S]^{1+}$ clusters,[248] except that site (F) has relatively large A-values, in accordance with a rather localized ferrous valence (d = 0.72 mms^{-1}) and relatively low covalency (Table 3.19).

In super-oxidized Hase I the medial and proximal FeS clusters are paramagnetic (Table 3.18), whereas the distal $[4Fe\text{-}4S]^{2+}$ cluster is diamagnetic and shows only weak magnetic splitting from the external magnetic field (blue trace in Figure 3.37B). Likewise, the ferrous low-spin site of the [NiFe] center has a vanishing hyperfine field (orange trace, covalently transferred spin density from nickel is negligible. In contrast, the $[3Fe\text{-}4S]^{1+}$ cluster is an all-ferric center with S=1/2 and small g-anisotropy arising from spin frustration.[244,249,250] This is seen from the moderately strong overall splitting in the magnetic Mössbauer spectrum (light green trace). The A-values are characteristic of three-iron sulfur clusters as found previously, with one of the A-tensors being characteristically small (Table 3.19B).[244]

The proximal cluster in its super-oxidized state is oxidized by two electrons more than in the H_2-reduced enzyme. Paramagnetism of the resulting $[4Fe\text{-}3S]^{5+}$ state is supported by the magnetic splitting in the field-dependent Mössbauer spectra. However, predicting a spin coupling scheme is difficult due to the irregular structure of the cluster (Figure 3.35B), which lacks a direct bridge between Fe2 and Fe4. We adopted the general pattern of all mixed-valence four-Fe clusters, that pairs of irons harbor minority and majority spin populations (positive and negative A-values). The best fits performed with this assumption required that the special subspectrum (S) with the largest quadrupole splitting (d /DE_Q = 0.46 / 2.45 mms^{-1}) and a site with small isomer shift (0.28 mms^{-1}) have positive A-values. This appears to be formally consistent with the presence of a mixed-valence *localized* pair (S*=9/2), antiparallel coupled to a ferric pair (2x S=5/2), as recently proposed by a theoretical/crystallographic study.[84] However, the assignment of subspectra to iron sites, obtained from DFT calculations, rules out this simple scheme, as will be discussed below.

Electronic Structure Calculations

Detailed interpretation of the Mössbauer data was obtained from DFT calculations of the electronic structure of the proximal [4Fe-3S] cluster in all three oxidation states. To this end, cluster models have been constructed on the basis of the crystal structures reported for the membrane-bound hydrogenases of *H. marinus* and *R. eutropha*.[81,82] Constrained optimizations have been performed with C-alpha carbons of coordinating amino acids fixed to their crystallographic positions (Figure 3.38). Potential binding of Glu_{82} to Fe2 and its protonation state has been systematically probed. In general, all possible Broken-Symmetry (BS) solutions have been calculated for each structure. The Mössbauer parameters for each of the four iron sites have been used as criteria for the selection of the best BS solution (isomer shift δ, quadrupole splitting ΔE_Q with sign, asymmetry parameter η of the efg, and the sign of the isotropic A-value, A_{iso}). Calculations were carried out with standard functionals (*BP86 and B3LYP*) without adjustments of fractional Hartree-Fock exchange or other *ad hoc* modifications. The agreement between calculated and experimental metric details as well as Mössbauer spectroscopic parameters is reasonable (Table 3.20).



a **b**

Figure 3.38 Computational model for the proximal iron-sulfur cluster of the membrane-bound respiratory [NiFe]-hydrogenase. All amino acids directly interacting with the iron sulfur cluster core Fe_4S_3, plus Glu_{82} are included. The total charge of the superoxidized species (a) was kept neutral, whereas reduced species (b) obtained a charge of -2. The side chain residues were kept frozen at the α-carbons during the geometry optimization (marked with *).

Based on the ΔE_Q values alone, the best BS solution for the super-oxidized cluster is labeled Ox2_14 (spin populations of Fe1 and Fe4 flipped) and corresponds to BS13 in ref.[84] However, this solution does not properly reproduce the experimental δ values or the sign of the A values, of which the latter reveals a basically wrong spin coupling scheme. Instead, our best solution, which is found within 4 kcal/mol of the energy minimum, has the spin populations of Fe2 and Fe4 flipped (Ox2_24), yielding proper values for δ, ΔE_Q, η, and the A values. In both cases, Glu_{82} is required to be unprotonated.

On the basis of the calculations, the special subspectrum (S) of the super-oxidized [4Fe-3S]$^{5+}$cluster with the large quadrupole splitting (ΔE_Q = 2.45 mms^{-1}) is assigned to the nitrogen-coordinated Fe2. This agrees with chemical intuition, where a large *efg* is expected due to the anisotropic covalency that results from asymmetric coordination with a unique, short N-bond.[81] The result is in contrast to the assignment of subspectrum (S) to Fe1,[84] which has quasi-tetrahedral symmetry. Note that our suggested spin distribution (αβαβ spin alignment for Fe1, Fe2, Fe3, Fe4, respectively) does not correspond to a standard coupling scheme observed for iron-sulfur clusters. The standard spin-coupling scheme involving a delocalized mixed-valence pair and two 'ferric' ions of the type [2Fe$^{2.5+}$, 2Fe^{3+}] does not apply since Fe2 and Fe4 with minority spin cannot form a strongly coupled mixed-valence pair with S*=9/2 since they are not connected by bridges that would transmit strong intersite coupling. Consequently, the total spin S=1/2 cannot arise from simple antiparallel coupling of S*=9/2 to two ferric sites with S=5/2 each. Weak bonding interactions found in the DFT model in contrast indicate that Fe1 and Fe3 may be regarded as the valence-delocalized 'Fe^{2+}/Fe^{3+}' pair (Mayer bond order 0.6, Figure 3.39). This makes it more difficult to explain why the remaining 'ferric' sites do not determine the majority spin. Apparently, the asymmetric topology of the [4Fe-3S]$^{5+}$ cluster leads to a more complex spin density distribution pattern compared to standard cubanes, such that in fact the resulting charge and spin distribution is energetically more favorable than in HiPIP clusters, as can be suspected on the basis of the redox potentials.

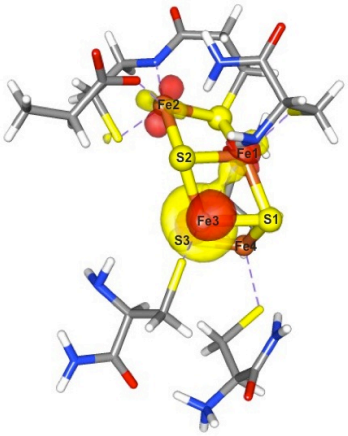


Figure 3.39 Week bonding interactions between iron centers in the model Ox2_24. The delocalized mixed valence pair is formed between Fe1-Fe3. The depicted quasi-restricted orbitals (QROs) are obtained form unrestricted Kohn-Sham orbitals at the B3LYP level of theory. The QROs were localized according to Pipek-Mezey localization procedure.

Selection of the best BS solution for the reduced [4Fe-3S]$^{3+}$ cluster is less clear cut for the super-oxidized state because several reasonable choices exist. Importantly all reasonable choices imply that Glu$_{82}$ is unprotonated. This raises some concerns with respect to its suspected role as a base.[84] Our best result (Red2_24) shows the spin populations on Fe2 and Fe4 to be flipped. Based on this solution, Fe4 and Fe1 are assigned to the special subspectra (S) and (F) found for the reduced proximal cluster, *i.e.* those with largest ΔE_Q and highest δ, respectively (Table 3.20A). Interestingly, the iron site with the largest quadrupole splitting (subspectrum S) has changed from Fe2 in the super-oxidized structure to Fe4 in the reduced structure, in which Fe4 has the unusual two-cysteine coordination in an otherwise rather cuboidal cluster. Unlike for the super-oxidized case, the

optimized structure for Red2_24 forms a tightly closed cubane with a long Fe4 – S-Cys25 separation. The spin distribution obtained from the calculations, in agreement with the experimental A-values, is best explained by the presence of a mixed-valence pair (Fe1 and Fe3) and two 'ferrous' sites (Fe2 and Fe4), which render the special subspectrum (F) with the highest isomer shift part of the mixed-valence pair.

Table 3.20 Calculated Mössbauer parameters for the iron sites of the proximal [4Fe-3S] cluster in the reduced and super-oxidized state Figure 3.35. Tentative assignments to the experimental subspectra from Table 3.19 are shown in brackets.

Site	***Calculated Mössbauer Parameters***					
	δ, mms^{-1}	ΔE_Q, mms^{-1}	η	$A_{iso}/g_N\mu_N$ T [a]	spin pop.	charge
			(**A**) Reduced $[4Fe\text{-}3S]^{3+}$			
Fe1 (F)[c]	**0.60** (0.71)	**+1.41** (+1.52)	**0.28** (0.3)	**-45.8** (-31.3)	+3.5	-0.15
Fe2	**0.35** (0.44)	**+1.25**[d] (+1.23)	**0.99** (0.9)	**+31.6** (+20.0)	-3.1	-0.38
Fe3	**0.36** (0.42)	**+0.93** (+0.84)	**0.3** (0.3)	**-36.2** (-25.0)	+3.1	-0.38
Fe4 (S)[b]	**0.46** (0.50)	**+2.13** (+2.60)	**0.51** (0.1)	**+19.2** (+14.0)	-2.7	-0.41
			(**B**) Super-oxidized $[4Fe\text{-}3S]^{5+}$			
Fe1	**0.39** (0.40)	**(-)1.56**[d] (-1.00)	**0.83** (0.7)	**-43.5** (-24.3)	3.3	-0.33
Fe2 (S)[b]	**0.38** (0.46)	**+2.25** (+2.45)	**0.4** (0.5)	**+35.4** (+18.6)	-2.4	-0.09
Fe3	**0.30** (0.39)	**(+)0.81**[d] (+0.70)	**0.74** (1.0)	**-41.8** (-34.7)	3.3	-0.37
Fe4	**0.22** (0.28)	**+0.74** (+0.60)	**0.56** (0.3)	**+42.5** (+24.2)	-3.4	-0.35

[a)]Isotropic experimental values are the average of the anisotropic values given in Table 3.19; [b)] special subspectrum (S); [c)] 'ferrous' subspectrum (F); [d)]The efg tensor has two components of similar size but different sign because of large asymmetry parameter *h* close to 1. For details of tensor orientations with respect to the magnetic axes see SI-Appendix.

3.5.3 Discussion

Tetranuclear iron-sulfur centers are common in biology, exhibiting great diversity in binding motifs, redox properties, and catalytic roles.[79,92,242,251] The proximal [4Fe-3S] cluster found in Hase I from *A. aeolicus* and other organisms[81,82,84] differs from these conventional cubanes by having a distorted core structure as well as a unique six-cysteine coordination mode, common for all Group I O_2-tolerant hydrogenases.[86] Magnetic Mössbauer data presented in this study in conjunction with our previous EPR results, substantiate the notion that the [4Fe-3S] cluster can attain three redox states with formal charges 5+, 4+, and 3+, which are isoelectronic to the 3+, 2+, and 1+ states of conventional [4Fe-4S] clusters (Table 3.21). The oxidation numbers in this formal concept are obtained from spectroscopically probed oxidation states of the iron atoms with the inorganic sulfur ligands taken in their usual closed shell configuration (S^{2-}). Our data suggest that the valence delocalization pattern and the spin-coupling schemes for the [4Fe-3S] cluster are similar to those of conventional [4Fe-4S] clusters, at least for the reduced and oxidized state. The super-oxidized [4Fe-

3S]$^{5+}$ cluster may also be regarded as comprising a mixed-valence $Fe^{2.5+}/Fe^{2.5+}$ and a di-ferric pair, comparable classical oxidized HiPIP-type clusters,[240, 242, 243, 249] but the spin interactions appear to be more complex than in HiPIP clusters. The reduced [4Fe-3S]$^{3+}$ cluster shares the pattern of a mixed-valence and a di-ferrous pair with the reduced cubanes in bacterial ferredoxins.[91, 92] Consistent with these extreme oxidation states, the [4Fe-3S]$^{4+}$ state appears to be well described by two mixed-valence pairs.

Table 3.21 Distribution of formal oxidation states in the FeS centers of *A. aeolicus* Hase I. Arrows indicate the spin orientations of the iron sites arising from exchange and double exchange interaction.

	S	"Fe^{2+}"	"$Fe^{2.5+}$"	"Fe^{3+}"
[4Fe4S]$^{1+}$	1/2	2 (↓↓)	2 (↑↑)	-
[4Fe4S]$^{2+}$	0	-	2+2 (↑↑↓↓)	-
[3Fe4S]^{0}	2	-	2 (↑↑)	1 (↓)
[3Fe4S]$^{1+}$	1/2	-	-	3 (↑↓↑)
[4Fe3S]$^{3+}$	1/2	2 (↓↓)	2 (↑↑)	-
[4Fe3S]$^{4+}$	0	-	2+2 (↑↑↓↓)	-
[4Fe3S]$^{5+}$	1/2	-	2 (↑↑)	2 (↓↓)[a)]

a) Sub-spin of the 'ferric' pair: $S^* = 4$ like in oxidized HiPIP-type clusters

The Mössbauer spectra of the proximal [4Fe-3S] center exhibit throughout a particular subspectrum (S) with large quadrupole splitting, which indicates unusual coordination of a distinct Fe site[252] in all three oxidation states. Although analogous subspectra are known for [4Fe-4S] clusters participating in chemical reactions, site-differentiated Fe atoms with partially trapped valences are usually observed only in one of the two accessible redox states. The perturbations are mostly associated with substrate binding[251] and arise from five-coordination or otherwise perturbed protein ligation.[252] Interestingly, the majority of "perturbed" clusters show signs of a 'special' Fe subsite only in their diamagnetic [4Fe-4S]$^{2+}$ state. Some notable examples are alkylated ferredoxin:thioredoxin reductase (NEM-FTR),[240] the CCG cluster in succinate:quinone oxidoreductase (SQR),[242] the SAM cluster in pyruvate formate lyase (PFL),[251] and the cluster in the LytB from *E. coli*.[253] Though the presence of trapped-valence sites disturb the spin-dependent delocalization,[254] the diamagnetic state (S=0) of these clusters and of the [4Fe-3S]$^{4+}$ in *A. aeolicus* is most likely stabilized through vibronic coupling and/or unequal J coupling constants.

Spectroscopy-oriented quantum chemical calculations provided a detailed description of the electronic structure of the new [4Fe-3S] cluster and correlated the spin and valence distribution derived from EPR and Mössbauer spectra to the recently reported crystallographic data.[81, 82, 84] In particular it could be shown that the unusual Mössbauer subspectra (S) and (F) found for the H_2-reduced state of the [4Fe-3S] cluster can be assigned to Fe4 and Fe1, respectively. While three Fe sites in the reduced structure retain quasi-tetrahedral FeS_4 coordination, Fe4 is unusual in that it is ligated by two terminal cysteine ligands and comes out of the core of the pseudocube (Figure 3.35), which predestines it as the special site (S) with large quadrupole splitting, corroborated by the calculation.

In contrast, the most unusual site in the super-oxidized structure is Fe2. Accordingly, the DFT calculations clearly assign (S) to that iron, although this may have essentially ferric character (if Fe1-Fe3 is regarded as mixed-valence pair). Although ferric compounds are typically characterized by small ΔE_Q due to the vanishing valence contribution of the $3d^5$ configuration to the *efg*, large

ΔE_Q are encountered for such a site if charge asymmetry is caused by a strongly covalent bond, such as with oxo-ligands, for example.[255, 256] Thus the DFT result that Fe2 provides subspectrum (S) in the [4Fe-3S]$^{5+}$ cluster agrees with the crystallographic data, which show that super-oxidation is accompanied by deprotonation of a backbone amide nitrogen of the tandem CC motif and nitrogen coordination to Fe2, forming a tight nitrogen bond (209 pm). The recent crystal structure and theoretical study for the *E.coli* Hyd-1 hydrogenase has shown that Fe2 may be five-coordinate with a nearby deprotonated glutamate. Our DFT results support this interpretation, particularly for the super-oxidized state, since the interpretation of the Mössbauer data appears to be very difficult on the basis of a different model.

We note in passing that the spectroscopic properties of the other clusters, *i.e.* the medial [3Fe-4S] and the distal [4Fe-4S] centers, are in the range of those reported for other hydrogenases[244, 245, 246] which is in agreement with the preserved structural features.[81, 82]

Super-oxidation of the proximal [4Fe-3S]$^{4+}$ cluster to the 5+ state, stabilized by deprotonation and binding of the amide nitrogen of the CC tandem motif and the incorporation of the nitrogen as a ligand to Fe2, is comparable to that of the P-cluster in the *A. vinelandii* nitrogenase upon its two-electron oxidation from the native (P^N) to the oxidized state (P^{OX}).[256] In this step, one of the Fe loses a sulfide bond and binds to the backbone amide nitrogen of an S_γ-coordinated cysteine group. In addition, a second Fe coordinates to the O_γ atom of a serine, thereby avoiding disruption of the cluster. A second case of a backbone nitrogen coordinating a Fe atom is that of nitrile dehydratase.[257] Notably, nitrile dehydratase also yields a large ΔE_Q,[258] however as expected for ferric low-spin ions.

The presence of supernumerary cysteines in the coordination sphere of the [4Fe-3S] cluster appears to be essential for its ability to trigger deprotonation and binding of a peptide amide nitrogen in order to tune its electronic properties for two redox transitions within 150 mV.[80] Exchange of the two 'additional' cysteines Cys-25 and Cys-126, in *R. eutropha* MBH with the usual glycine residues found in standard hydrogenases has shown that the [4Fe-3S] cluster may be involved in a protection mechanism against inactivation of the [NiFe] center by molecular oxygen. The distance and orientation of the [4Fe-3S] cluster with respect to the [NiFe] site in the respective crystal structures[81] are essentially identical in the 3+ and 5+ states, where the ligand of the [4Fe-3S] cluster closest to the [NiFe] site is the conserved Cys-23 (Figure 3.35). It is tempting to assume that electron transfer proceeds *via* covalent bonds from this cysteine to the active site. The exchange coupling constant *J* between the proximal [4Fe-3S]$^{5+}$ and the [NiFe] site is $209 \cdot 10^{-4}$ cm^{-1}, whereas that between the reduced [4Fe-3S]$^{3+}$ and the [NiFe] site is only $36 \cdot 10^{-4}$ cm^{-1}.[80] Electron transfer rates are proportional to the magnitude of the exchange interaction and this significant difference in *J*'s could affect electron transfer rates.[259]

In summary, the [4Fe-3S] cluster uses surplus cysteine ligation to stabilize two redox couples not known to occur in biological [4Fe-4S] cubanes. Based on the present Mössbauer and DFT results, we propose a distribution pattern for Fe valences and local spin populations that basically resembles those of classical cubane clusters, particularly in the lower oxidation states. In detail, however, the electronic properties of the sites participating in the novel chemistry are distinct, particularly in the super-oxidized state. The redox transitions are clearly metal-centered. Due to its proximity to the [NiFe] center, the proximal [4Fe-3S] cluster has in fact been discussed in the past as being to some

degree part of the catalytic core of the enzyme.[260] Therefore considering its very positive redox potential and its appreciable electronic coupling to the [NiFe] site, it is not surprising that the [4Fe-3S] cofactor is able to perform at least two functions through redox-dependent structural changes; to mediate catalytic electron transfer and to serve as an electron-reservoir for the efficient reductive removal of O_2.

3.6 The Hessian implementation

3.6.1 Implementation

The program flow

Calculation of the analytical second derivatives of the SCF total energy proceeds via the following computational steps:

1. Evaluation of the $\mathbf{S}_{\sigma}^{(\vartheta)}$ (eq. (47)) and $\mathbf{Q}_{\sigma}^{\vartheta}$ matrices (eq. (60)).
2. Formation of the $H\left(\{\varepsilon_{i\sigma}\},\mathbf{S}^{(\vartheta)},\mathbf{S}^{(\lambda)}\right)$ terms (eq. (67))
3. Calculation of the $\mathbf{F}_{\sigma}^{(\vartheta)}$ matrices (eq. (58)).
4. Addition of the $H\left(\mathbf{F}^{(\vartheta)},\mathbf{Q}^{\lambda}\right)$ terms to the Hessian (eq. (68))
5. Calculation of the $\mathbf{G}_{\sigma}\left[\mathbf{Q}^{\lambda}\right]$ matrices (eq. (55)).
6. Addition of the $H\left(\mathbf{Q}^{\vartheta},\mathbf{Q}^{\lambda}\right)$ terms to the Hessian (eq. (69)).
7. Formation of the right-hand sides of the CPSCF equations (eq. (62)).
8. Solution of the 3N CPSCF equations for the orbital rotation coefficients $\mathbf{U}_{\sigma}^{\vartheta}$.
9. Addition of the $H\left(\mathbf{F}_{RHS}^{\lambda},\mathbf{U}^{\vartheta}\right)$ terms to the Hessian (eq. (70)).
10. Calculation of the integral second derivative contributions (eq. (71))

In addition to the standard one and two-electron terms, our implementation of the analytic Hessian also includes the number of features for approximate treatments of the dispersion interactions, relativistic effects, core electrons, and point-charge (PC) interactions which are involved in combined quantum mechanic/molecular mechanic (QM/MM) calculations.

Van der Waals (VDW) correction

The most popular variants of Van der Waals (VDW) correction, DFT-D2 and DFT-D3,[168,169,261] which are very effective semiempirical methods, has been implemented. It is widely used for accounting for missing dispersion interactions in a molecule. In this approach, the dispersion

energy, E_{disp}, representing a part of the nonlocal, long- and medium-range, electron correlation, is modeled additively by atom pair interactions:[261]

$$E_{disp} = -\sum_{A>B} \sum_{n=6(8,10)} s_n \frac{C_n^{AB}}{R_{AB}^n} f_{\text{damp}}(R_{AB}) \tag{127}$$

where R_{AB} is the distance between atoms A and B, C_n^{AB} is a dispersion coefficient, $f_{damp}(R_{AB})$ is a damping function. E_{disp} in eq. (127) depends in a very simple way on molecular geometry, and its second derivatives are straightforwardly implemented, similarly to the nuclear repulsion contributions.

QM/MM hessian

The additional interactions considered in QM/MM calculations are represented by the following additive contributions to the total energy:[95,262]

$$E_{QM/MM} = E_{pe} + E_{pn} + E_{pa} + E_{pp} \tag{128}$$

where E_{pe} corresponds to the electrostatic interaction of the electrons with the PCs of the MM region, $\{Z_q\}$:

$$E_{pe} = \sum_q Z_q \sum_{\mu\nu\sigma} P_{\mu\nu\sigma} (\mu \left| \frac{1}{\left| \mathbf{r} - \mathbf{r}_q \right|} \right| \nu) \tag{129}$$

Here, $\mathbf{r}_q$ is the coordinate of the q-th PC. The rest of the terms describe the electrostatic interactions of MM charges with QM nuclei (E_{pn}), the VDW interactions between QM and MM atoms (E_{pa}), and the MM energy of the MM atoms (E_{pp}). As discussed earlier in the literature, even when the number of QM atoms is low, the QM/MM interaction terms, E_{pe}, makes the determination of the full QM/MM Hessian very expensive, since in the additive scheme, every degree of freedom of the MM region leads to an additional CPSCF equation.[262] Moreover, the computational expenses for evaluating explicit first and second derivatives of one-electron integrals in eq. (129) can grow substantially due to a possibly very large number of MM atoms. For solving this problem, it was suggested to omit the CP-SCF equations for MM atoms far from the QM region, which is controlled by an appropriate cutoff parameter.[262] As an efficient alternative, the Mobile Block Hessian formalism has been developed recently that allows to substantially reduce the cost of the QM/MM Hessian calculations.[95] In the present implementation, a simplified approach in which only the QM submatrix of the total Hessian matrix is used. In this case the QM-MM subblock of the Hessian is neglected, whereby the normal coordinates of the QM fragment are considered to be fully decoupled from the surrounding MM region.

In this approach, the additional terms, required in QM/MM analytic Hessian calculation, appear as trivial contributions to $F^{(\vartheta)}$:

$$F^{(\vartheta)}_{QM/MM,\mu\nu\sigma} = F^{(\vartheta)}_{pe} = \sum_q Z_q (\mu \left| \frac{1}{|\mathbf{r} - \mathbf{r}_q|} \right| \nu)^{(\vartheta)} \tag{130}$$

and to the QM Hessian matrix:

$$E^{\vartheta\lambda}_{QM/MM} = \sum_q Z_q \sum_{\mu\nu\sigma} P_{\mu\nu\sigma} (\mu \left| \frac{1}{|\mathbf{r} - \mathbf{r}_q|} \right| \nu)^{(\vartheta)(\lambda)} + E^{\vartheta\lambda}_{pn} + E^{\vartheta\lambda}_{pa} \tag{131}$$

Effective core potential (ECP) Hessian

Effective core potentials (ECP) provide an efficient way for DFT and ab initio studies of molecules containing heavy atoms. They also account for relativistic effects in a simple way. The analytic second derivatives for ECPs were implemented according to refs.[263,264] In the ECP method, the potential due to core electrons of atom C is replaced by the effective core potential, $U_C(r)$:

$$U_C(\mathbf{r}) = U_C^{L+1}(\mathbf{r}) + \sum_{l=0}^{L} \sum_{m=-1}^{l} |lm\rangle \left[U_C^l(\mathbf{r}) - U_C^{L+1}(\mathbf{r}) \right] \langle lm| \tag{132}$$

where L is the largest angular momentum quantum number appearing in the core, and the l-depndent potentials, $U_C^l(r)$ $(l = 0 \div L+1)$, are determined by a fit to a linear combination of Gaussians. The ECPs contributions to the total energy read:

$$E_{ECP} = \sum_C \sum_{\mu\nu\sigma} P_{\mu\nu\sigma} (\mu | U_C | \nu) \tag{133}$$

Thus, ECPs give rise to additional, low cost, terms in $F^{(\vartheta)}$ and $E^{(\vartheta)(\lambda)}$ as follows:

$$F^{(\vartheta)}_{ECP,\mu\nu\sigma} = \sum_C (\mu | U_C | \nu)^{(\vartheta)} \tag{134}$$

$$E^{(\vartheta)(\lambda)}_{ECP} = \sum_C \sum_{\mu\nu\sigma} P_{\mu\nu\sigma} (\mu | U_C | \nu)^{(\vartheta)(\lambda)} \tag{135}$$

The COSX implementation

The evaluation of the COSX energy proceeds through the following intermediates

$$F_{\tau g} = \sum_\kappa P_{\kappa\tau} X_{\kappa g} \tag{136}$$

and

$$G_{\nu g} = \sum_{\tau} A_g^{\nu\tau} F_{\tau g} \tag{137}$$

(Note that F and G in this subsection denote the usual COSX intermediates, and have nothing to do with the Fock matrix related quantities defined previously. This notation is only kept to maintain consistency with the previous COSX literature.

Once the analytic integrals are calculated for a given shell pair they are at once contracted with the corresponding segment $F_{\tau g}$, and thus the direct construction and storage of $A_g^{\nu\tau}$ is avoided. For energy calculations, the exchange matrix is needed, and is contracted with $X_{\mu g}$. For the gradient, $G_{\nu g}$ is then contracted with the second density and finally with $X_{\mu^{\vartheta} g}$.

If there is more than one density (as in CPSCF)

$$F_{\tau g}^{\vartheta} = \sum_{\kappa} P_{\kappa\tau}^{\vartheta} X_{\kappa g} \tag{138}$$

or there are derivative functions present,

$$F_{\tau g}^{\vartheta} = \sum_{\kappa} P_{\kappa\tau} X_{\kappa^{\vartheta} g} \tag{139}$$

this would mean that $G_{\nu g}^{\vartheta}$ should be recalculated for each of these, and, therefore the expensive analytic integration would have to be evaluated multiple times as well. In these cases, $A_g^{\nu\tau}$ is calculated directly (for basis functions that have a significant value on the grid) and then contracted with the remaining quantities later. In the first case, $F_{\tau g}^{\vartheta}$ is still evaluated with the perturbed densities and then it is contacted with $A_g^{\nu\tau}$. If the derivatives appear in the function value part, i.e., one has $X_{\kappa^{\vartheta} g}$, the density is first contracted with $A_g^{\nu\tau}$ and then the result with $X_{\kappa^{\vartheta} g}$ and remaining quantities.

Finally, if there is any number of derivatives occuring in the analytic part, then it is still worth proceeding through the appropriate analogue of $G_{\nu g}$, i.e. rather than storing the whole range of analytical integrals, these are at once contracted into G.

3.6.2 Numerical results

Computational details

All calculations have been performed with a development version of the ORCA electronic structure package.[117] The accuracy and efficiency of different approximations for analytic second derivatives were monitored at the HF and DFT levels. The B3LYP and BP86 functionals as representative

hybrid and pure functionals have been employed. The calculations were performed using the 'def2' variants of split-valence (SV) and triple- valence (TZV) basis sets with the appropriate polarization functions from the Turbomole library.[265] For the fitting basis in the RI-J treatment, the 'def2' fit bases optimized for the SV, TZV and QZV basis sets were used.[266]

The errors in the analytic Hessian due to cumulative effects of different approximations, such as RI-J, COSX, finite convergence in the SCF and CP-SCF calculations, integral prescreening, were characterized by a statistical analysis of the calculated vibrational frequencies relative to the approximation-free results obtained with extremely tight SCF convergence settings (ExtremeSCF), for which any integral prescreening was avoided, whereby all sources of numerical errors were eliminated to a maximum extent. The statistical behavior of errors arising from loosening the SCF and CP-SCF convergence criteria, and increasing the thresholds in the integral prescreening, was examined in a series of the calculations with different settings of increasing quality, labeled as SCFConvN according to the ORCA keyword specifications (-N is the decimal logarithm of the corresponding SCF energy tolerance, N=6–10). Definitions of the ExtremeSCF and SCFConvN settings involving specifications of the SCF and CP-SCF convergence criteria, together with associated integral and basis function prescreening thresholds (Table 3.22).

Table 3.22 Definition of different ORCA program settings specifying the SCF and CP-SCF convergence criteria together with associated integral and basis function thresholds, used for monitoring the accuracy and efficiency of analytic second derivative calculations. The following shorthand notations are used: the integral neglect threshold for the integral prescreening (Thresh), the primitive cutoff factor (TCut), the basis function cutoff value (BFCut), the energy convergence tolerance (TolE), the tolerances for the maximum and root-mean-square changes in the density matrix (TolMaxP and TolRMSP, respectively), DIIS error (TolErr), orbital gradient (TolG), orbital rotation (TolX), and the norm of the residual vector in the solution of the CP-SCF equations.

Convergence settings	*Thresh*	*TCut*	*BFCut*	*TolE*	*TolMaxP*	*TolRMSP*	*TolG*	*TolX*	*TolErr*	*TolZ*
ExtremeSCF	10^{-20}	10^{-20}	0.0	10^{-14}	10^{-14}	10^{-14}	10^{-9}	10^{-9}	10^{-14}	10^{-10}
SCFConv6	10^{-10}	10^{-11}	10^{-10}	10^{-6}	10^{-5}	10^{-7}	3.33×10^{-4}	3.33×10^{-4}	10^{-6}	10^{-6}
SCFConv7	3×10^{-11}	10^{-12}	10^{-11}	10^{-7}	10^{-6}	10^{-8}	10^{-4}	10^{-4}	10^{-7}	10^{-7}
SCFConv8	10^{-11}	10^{-12}	3×10^{-12}	10^{-8}	10^{-7}	10^{-9}	2.5×10^{-5}	2.5×10^{-5}	10^{-8}	10^{-8}
SCFConv9	10^{-12}	10^{-13}	10^{-13}	10^{-9}	10^{-8}	10^{-10}	3.33×10^{-6}	3.33×10^{-6}	10^{-9}	10^{-9}
SCFConv10	10^{-13}	10^{-14}	3×10^{-14}	10^{-10}	10^{-9}	10^{-11}	10^{-6}	10^{-6}	10^{-10}	10^{-10}
SCFConv7(Z8)	3×10^{-11}	10^{-12}	10^{-11}	10^{-7}	10^{-6}	10^{-8}	10^{-4}	10^{-4}	10^{-7}	10^{-8}
SCFConv7(Z9)	3×10^{-11}	10^{-12}	10^{-11}	10^{-7}	10^{-6}	10^{-8}	10^{-4}	10^{-4}	10^{-7}	10^{-9}
ExtremeSCF(Z5)	10^{-20}	10^{-20}	0.0	10^{-14}	10^{-14}	10^{-14}	10^{-9}	10^{-9}	10^{-14}	10^{-5}
ExtremeSCF(Z6)	10^{-20}	10^{-20}	0.0	10^{-14}	10^{-14}	10^{-14}	10^{-9}	10^{-9}	10^{-14}	10^{-6}
ExtremeSCF(Z7)	10^{-20}	10^{-20}	0.0	10^{-14}	10^{-14}	10^{-14}	10^{-9}	10^{-9}	10^{-14}	10^{-7}
ExtremeSCF(Z8)	10^{-20}	10^{-20}	0.0	10^{-14}	10^{-14}	10^{-14}	10^{-9}	10^{-9}	10^{-14}	10^{-8}
ExtremeSCF(Z9)	10^{-20}	10^{-20}	0.0	10^{-14}	10^{-14}	10^{-14}	10^{-9}	10^{-9}	10^{-14}	10^{-9}
ExtremeSCF(Z10)	10^{-20}	10^{-20}	0.0	10^{-14}	10^{-14}	10^{-14}	10^{-9}	10^{-9}	10^{-14}	10^{-10}
TightSCF	2.5×10^{-11}	2.5×10^{-12}	10^{-11}	10^{-8}	10^{-7}	5×10^{-9}	10^{-5}	10^{-5}	5×10^{-7}	10^{-8}
TightSCF_Z7	2.5×10^{-11}	2.5×10^{-12}	10^{-11}	10^{-8}	10^{-7}	5×10^{-9}	10^{-5}	10^{-5}	5×10^{-7}	10^{-7}
TightSCF_Z9	2.5×10^{-11}	2.5×10^{-12}	10^{-11}	10^{-8}	10^{-7}	5×10^{-9}	10^{-5}	10^{-5}	5×10^{-7}	10^{-9}

The particular impact of the CP-SCF convergence criteria on the quality of the calculated vibrational frequencies was studied with the computational settings corresponding to the ExtremeSCF and SCFConvN parameters, among which only the tolerance for the norm of the

residual vector (TolZ) in the solution of the CP-SCF equations was varied. The corresponding settings are collectively referred to as ExtremeSCF(ZM) and SCFConvN(ZM) (Table 3.22), respectively. The parameter M in these notations corresponds to TolZ $=10^{-M}$ (M=6–10). As compared to the previous ORCA implementations, the default program settings which are standardly used for all analytic derivative calculations were additionally optimized based on the statistical error analysis for the frequencies calculated with the SCFConvN and SCFConvN(ZM) settings (vide infra). These program defaults, collectively referred to as TightSCF, are given Table 3.22.

The accuracy calculations were carried out for a test set of 47 molecules that originates from the work of Grimme et al.[267] as used in previous studies.[111,268] The test set mainly consists of small- and medium-sized organic molecules.

Calculations labeled RIJDX used the RI-J approximation for the Coulomb term and exact analytical integration for the exchange contribution. Calculations denoted as RIJCOSX used the COSX approximation for the exchange terms in conjunction with the RI-J approximation. In order to test the effect of the COSX approximation in the individual terms (eqs. (55), (58), (61) and (71)) on the overall accuracy and efficiency, a series of HF calculations collectively labeled as HESS(i_1, i_2, i_3, i_4), where i_k (k=0-2) specify the approximation methods for different two-electron terms has been designed. In these notations, i_1, i_2, i_3, i_4 refer to the two-electron integral contributions in $F^{(\vartheta)}$, $\mathbf{G}\left[\mathbf{Q}^{\vartheta}\right]$, $\mathbf{G}\left[\mathbf{U}^{\vartheta}\right]$, and $E^{(\vartheta)(\lambda)}$, respectively. If i_k=0 the corresponding Coulomb and exchange four-center integrals are calculated exactly; the flags i_k=1 and i_k=2 indicate the RIJDX and RIJCOSX approximations, respectively.

For assessing the accuracy of two different schemes for account of the grid weight derivative contributions in the DFT Hessian, we performed numerical frequency calculations in which the Hessian matrix was evaluated by finite difference between analytic gradients without (NumFreq1) and with (NumFreq2) weight derivatives. These calculations were carried out for XC grids of different quality (see below), using the ExtemeSCF settings, and the central difference scheme with the increment 0.0005 A. The analytic implementation of the Hessian (AnFreq) is equivalent to the finite-difference scheme NumFreq1. The errors in the calculated frequencies, which are due to finite-grid size effects, were determined with respect to the results for the highest quality grid, for both methods of treatment of the grid weight derivatives.

The following definitions were used for the statistical error parameters characterizing the accuracy of the vibrational frequencies within different approximations: ensemble-average of single-molecule root-mean-square, $\overline{|\Delta\omega|}_{\mathrm{rms}}$, mean-absolute, $\overline{|\Delta\omega|}_{\mathrm{mad}}$, and maximum, $\overline{|\Delta\omega|}_{\mathrm{max}}$, deviations of calculated frequencies with respect to reference (the most accurate) values. The ensemble-maxima of the corresponding parameters are denoted as $|\Delta\omega|^{m}_{\mathrm{rms}}$, $|\Delta\omega|^{m}_{\mathrm{mad}}$, and $|\Delta\omega|^{m}_{\mathrm{max}}$, respectively.

Geometry optimizations of the test structures were carried out at the same level of theory as subsequent analytic frequency calculations, using the TightSCF convergence criteria, with the geometry optimization tolerance of 10^{-6} Eh for the energy, 10^{-4} a.u. for the maximum component of the gradient vector, 3×10^{-5} a.u. for the root mean square (RMS) gradient, 10^{-3} Bohr for the maximum step, and 6×10^{-4} Bohr for the RMS step (ORCA keyword TightOpt). Calculations were

run on Intel® Xeon® 2.53GHz machines. All reported times are wall-clock times and have a variability of about 10% due to differing system load.

Integration grids

In accordance with standard ORCA settings for XC numerical quadrature, the radial grids were of Gauss-Chebyshev-type, and the angular integrations were performed with the Lebedev grids. The accuracy of the integration grids is controlled by the number of points of the largest angular grid, N_{ang}, and radial resolution parameter ε that determines the number of radial shells for a given atom, $n_r = 15\varepsilon + 5r - 40$, where r is the row in the periodic table to which the atom belongs.[269] The pruning algorithm of Gill et al.[270] was used to reduce the number of grid points. The errors in the frequency calculations that arise from the finite size of XC integration grids were analyzed by studying the convergence behavior of the frequencies calculated for a series of standard DFT grids of increasing quality as defined in Table 3.23. The grids are labeled as 2-7 according to their specifications in the ORCA program. The highest quality grid 7 ($\varepsilon = 5.67$, $N_{ang} = 770$) was used as a reference in the statistical error analysis. Apart from the analysis of the finite grid-size effects on the accuracy of the calculated analytic frequencies, the other sources of errors in the DFT calculations were examined using ORCA grid 4 ($\varepsilon = 4.67$, $N_{ang} = 302$), for which the grid error was found to be sufficiently small (vide infra).

Table 3.23 Definition of the standard DFT grids included in the ORCA package. N_{ang} is the maximum number of angular grid points for each radial shell, ε is the radial integration accuracy from which the number of radial shells for a given atom is computed.

Grid	N_{ang}	ε
2	110	4.34
3	194	4.34
4	302	4.67
5	434	5.01
6	590	5.34
7	770	5.67

In the implementation of the COSX algorithm for ground state HF and DFT calculations, various integration grids that were derived from the standard DFT grids included in the ORCA package have been extensively tested. Having pursued a goal of reasonable computational cost and an error of about 0.1 kcal/mol in the total ground state exchange energy, it was proposed to use the following strategy: 1) the initial SCF iterations (except the very first one, which is done as 2) are carried out with a small 50 point Lebedev grid and $\varepsilon = \frac{10}{3}$ (grid A); 2) The SCF is converged with a 50 point grid and $\varepsilon = 4$ (grid B); 3) the exchange energy is recomputed on a 194 point angular

grid with $\varepsilon = \frac{13}{3}$ (grid C). For higher accuracy, more accurate standard sets of grids {A,B,C} have been defined. For energy and gradient calculations, the overlap fitting scheme was used.[111]

Following extensive test calculations, the following default grids for the COSX treatment in the analytic second derivative calculations are proposed: 1) the formation of the $\mathbf{F}^{(\vartheta)}$ and $\mathbf{G}\left[\mathbf{Q}^{\vartheta}\right]$ matrices is performed on grid C; 2) the evaluation of the matrices $\mathbf{G}\left[\mathbf{U}^{\vartheta}\right]$ in the CP-SCF equations is performed on grid A. In the test calculations, the COSX approximation was also applied for the evaluation of the explicit second derivatives of the exchange energy (eq. (119)). In this case, the accuracy of the vibrational frequencies calculated with grid C is unsatisfactory (vide infra, J. Chem. Phys. 119, 12763 (2003)). It can be improved by using larger grids, however the efficiency of these calculations is typically noticeably worse compared to the approximation-free integral direct method.

Benchmark result: accuracy

The impact of different program settings on the accuracy of the calculated vibrational frequencies, as determined with respect to the most accurate results obtained with the ExtremeSCF settings, was studied in a series of the SCFConvN and SCFConvN(ZM) calculations. As expected, the accuracy is systematically improved upon going up in the hierarchy of the SCFConvN (N=6-10) settings (Table 3.24).

As a general observation, the statistical errors noticeably depend on the frequency range. Thus, for given computational settings, the error parameters for high-frequency vibrations, $\omega > 200\,\mathrm{cm}^{-1}$, are typically lower by up to 1-2 orders of magnitude than the corresponding errors in the low-frequency vibrations ($\omega < 200\,\mathrm{cm}^{-1}$). Although the ensemble-averages of the calculated errors ($\overline{|\Delta\omega|}_{\mathrm{rms}}$, $\overline{|\Delta\omega|}_{\mathrm{mad}}$, $\overline{|\Delta\omega|}_{\mathrm{max}}$) in the low- and high-frequency ranges appear to be sufficiently small for the SCFConv6 and SCFConv7 settings (Table 3.24), the corresponding ensemble-maxima ($|\Delta\omega|^{m}_{\mathrm{rms}}$, $|\Delta\omega|^{m}_{\mathrm{mad}}$, $|\Delta\omega|^{m}_{\mathrm{max}}$) are unacceptably large. In particular, the maximum errors throughout the test set, $|\Delta\omega|^{m}_{\mathrm{max}}$, are ~9 cm^{-1} (SCFConv6), ~4 cm^{-1} (SCFConv7) for $\omega > 200\,\mathrm{cm}^{-1}$, and are ~29 cm^{-1} (SCFConv6), ~25 cm^{-1} (SCFConv7) for $\omega < 200\,\mathrm{cm}^{-1}$.

The calculations with the SCFConv8 settings already lead to significant improvements in the error parameters, which turn out to be by almost one order of magnitude more accurate than the SCFConv7 results. The corresponding ensemble maxima, $|\Delta\omega|^{m}_{\mathrm{max}} \approx 3.5\,\mathrm{cm}^{-1}$ $\left(\omega < 200\,\mathrm{cm}^{-1}\right)$, $|\Delta\omega|^{m}_{\mathrm{max}} \approx 0.4\,\mathrm{cm}^{-1}$ $\left(\omega > 200\,\mathrm{cm}^{-1}\right)$ are sufficiently low, as compared to the intrinsic errors of HF and DFT methods, and accuracy of numerical second derivative calculations with default (TightSCF) convergence settings (vide infra).

The calculations with the SCFConv9 and SCFConv10 settings indicate that further tightening of the convergence criteria and integral threshold parameters lead to moderate improvements in the

calculated frequencies. More detailed analysis on the basis of the calculation with the SCFConvN and ExtremSCF(ZN) settings has revealed that the tolerance for the norm of the residual vector (TolZ) in the solution of the CP-SCF equations has the most crucial influence on the accuracy of the calculated frequencies. Indeed, the ExtremeSCF(ZM) calculations, in which only TolZ was varied, indicate nearly the same accuracy level as in the corresponding SCFConvM calculations (Table 3.24, App. Table C1), in which, in addition to parameter TolZ, the number of other parameters, such as SCF convergence and integral prescreening thresholds were loosened.

According to the statistical analysis of the differences between the frequencies calculated with the ExtremeSCF(ZM) and SCFConvM settings (App. Table C2), the cumulative effect of tightening of the computational parameters in the hierarchy of settings {SCFConv6,..SCFConv10,ExtremeSCF}, excluding TolZ, is about ~10% of the desired accuracy level (~1 cm^{-1}).

For accessing the influence of the finite-grid size effects on the accuracy of the calculated vibrational frequencies, we performed statistical analysis of the errors in the vibrational frequencies, calculated at the BP86/def2-TZVP level, using XC grids 2-6. Vibrational frequencies obtained with the highest quality grid 7 were used as reference values.

Table 3.24 Statistical analysis of the errors in the calculated vibrational frequencies (ω) corresponding to different ORCA program settings, as determined with respect to the results obtained in the approximation-free calculations with the ExtremeSCF settings. The calculations were performed at the HF/def2-TZVP level of theory.

	ω<200 cm^{-1}				ω>200 cm^{-1}			
	$\lvert\Delta\omega\rvert^m_{max}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{max}$ (cm^{-1})	$\lvert\Delta\omega\rvert^m_{mad}$ (cm^{-1})	$\lvert\Delta\omega\rvert^m_{rms}$ (cm^{-1})	$\lvert\Delta\omega\rvert^m_{max}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{max}$ (cm^{-1})	$\lvert\Delta\omega\rvert^m_{mad}$ (cm^{-1})	$\lvert\Delta\omega\rvert^m_{rms}$ (cm^{-1})
SCFConv6	28.49	4.05	28.49	28.49	8.61	1.43	0.85	1.98
SCFConv7	24.47	2.64	24.47	24.47	3.76	0.5	0.27	0.8
SCFConv7(Z8)	4.05	0.63	2.21	2.88	0.42	0.13	0.05	0.1
SCFConv7(Z9)	3.4	0.28	2.15	2.41	0.45	0.05	0.02	0.07
SCFConv8	3.52	0.58	2.05	2.5	0.41	0.14	0.05	0.09
SCFConv9	3.32	0.27	1.71	2.35	0.51	0.05	0.03	0.08
SCFConv10	2.35	0.21	2.35	2.35	0.51	0.03	0.02	0.08
TightSCF	4.05	0.62	2.21	2.88	0.43	0.13	0.05	0.1
	ω							
	$\lvert\Delta\omega\rvert^m_{max}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{max}$ (cm^{-1})	$\lvert\Delta\omega\rvert^m_{mad}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{mad}$ (cm^{-1})	$\lvert\Delta\omega\rvert^m_{rms}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{rms}$ (cm^{-1})		
SCFConv6	28.49	3.54	1.15	0.43	4.41	0.83		
SCFConv7	24.47	2.07	0.72	0.16	3.92	0.41		
SCFConv7(Z8)	4.05	0.5	0.2	0.04	0.81	0.11		
SCFConv7(Z9)	3.4	0.24	0.15	0.02	0.68	0.05		
SCFConv8	3.52	0.47	0.18	0.04	0.71	0.1		
SCFConv9	3.32	0.22	0.15	0.02	0.67	0.05		
SCFConv10	2.35	0.17	0.09	0.01	0.41	0.03		
TightSCF	4.05	0.48	0.2	0.04	0.81	0.1		

Table 3.25 Statistical analysis of the errors in the calculated vibrational frequencies (ω) corresponding to different XC grids, as determined with respect to the results obtained with Grid 7. The calculations were performed at the BP86/def2-TZVP level of theory.

Grid	ω<200 cm^{-1}				ω>200 cm^{-1}			
	$\lvert\Delta\omega\rvert^m_{\max}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{\max}$ (cm^{-1})	$\lvert\Delta\omega\rvert^m_{\mathrm{mad}}$ (cm^{-1})	$\lvert\Delta\omega\rvert^m_{\mathrm{rms}}$ (cm^{-1})	$\lvert\Delta\omega\rvert^m_{\max}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{\max}$ (cm^{-1})	$\lvert\Delta\omega\rvert^m_{\mathrm{mad}}$ (cm^{-1})	$\lvert\Delta\omega\rvert^m_{\mathrm{rms}}$ (cm^{-1})
6	0.74	0.26	0.52	0.56	0.24	0.09	0.06	0.08
5	1.22	0.57	0.82	0.92	0.42	0.25	0.16	0.2
4	1.06	0.59	1.06	1.06	1	0.73	0.32	0.42
3	13.57	9.7	10.48	10.85	7.61	2.89	1.25	1.98
2	31.84	19.86	20	21.76	29.64	13.85	9.33	13.42
	ω							
	$\lvert\Delta\omega\rvert^m_{\max}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{\max}$ (cm^{-1})	$\lvert\Delta\omega\rvert^m_{\mathrm{mad}}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{\mathrm{mad}}$ (cm^{-1})	$\lvert\Delta\omega\rvert^m_{\mathrm{rms}}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{\mathrm{rms}}$ (cm^{-1})		
6	0.74	0.15	0.11	0.03	0.19	0.05		
5	1.22	0.35	0.22	0.09	0.35	0.13		
4	1.06	0.74	0.34	0.18	0.44	0.26		
3	13.57	4.94	2.1	1.04	3.76	1.51		
2	31.84	16.46	9.33	4.51	13.42	6.23		

As indicated in Table 3.25, the error parameters converge well with the grid size. The errors corresponding to grids 2 and 3 are quite large. The calculations performed with grid 4 lead to substantial improvement in the accuracy. For this grid, the maximum error throughout the entire set of test molecules, $\lvert\Delta\omega\rvert^m_{\max}$, is ~1cm^{-1}. This value turns out to be not very different from the corresponding ensemble-average errors, which also indicates that the error distribution has a rather uniform character. Further improvement of the grid quality leads to more moderate reduction of the error parameters, as indicated in Table 3.25 for grids 5 and 6. Thus, grid 4 appears to be a reasonable choice for the DFT analytic second derivative calculations with the desired accuracy level of ~1cm^{-1} in the vibrational frequencies.

For assessing the accuracy of two different methods for account of the grid weight derivative contributions in the DFT Hessian, we performed numerical frequency calculations in which the Hessian matrix was evaluated by finite difference between analytic gradients without (NumFreq1) and with (NumFreq2) weight derivatives. Methods NumFreq1 and NumFreq2 correspond to the weight derivative schemes (1) and (2), respectively. As indicated in the first 3 rows of Table 3.26, the frequencies calculated by NumFreq1 and NumFreq2 procedures, using the highest quality grid 7 and ExtremeSCF settings, are in close agreement with each other and analytic frequencies calculated with the same computational settings. Small differences between both types of numerical frequencies (~0.5 cm^{-1} for the ensemble-maxima and ~0.1 cm^{-1} for the ensemble-averages) give an evidence that for grid 7 the effects of the second grid weight derivatives are negligible. In contrast with these results, the deviations between the NumFreq1 and NumFreq2 frequencies for grid 4 are about one order of magnitude larger (Table 3.26).

One can see that the numerical, NumFreq1, and analytical frequencies (according to scheme (1)), calculated with grid 4, are very close to those obtained using grid 7, whereas the NumFreq2 frequencies show the deviations which are about one order of magnitude larger, such that the corresponding ensemble-maxima and ensemble-averages are ~8 cm^{-1} and ~2 cm^{-1}, respectively.

Thus, we can conclude that the Hessian calculated according to scheme (1) is in general more accurate than the one in which the second grid weight derivatives were taken into account (scheme (2)), which corroborates the observations by Handy.[156]

Table 3.26 Statistical analysis of the differences between the vibrational frequencies (ω) calculated analytically (*AnFreq*) and by the finite-difference methods, *NumFreq1* and *NumFreq2*, using XC grids 4 and 7. The calculations were performed at the BP86/def2-TZVP level of theory.

	ω<200 cm^{-1}				ω>200 cm^{-1}			
	$\lvert\Delta\omega\rvert^{m}_{\max}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{\max}$ (cm^{-1})	$\lvert\Delta\omega\rvert^{m}_{\mathrm{mad}}$ (cm^{-1})	$\lvert\Delta\omega\rvert^{m}_{\mathrm{rms}}$ (cm^{-1})	$\lvert\Delta\omega\rvert^{m}_{\max}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{\max}$ (cm^{-1})	$\lvert\Delta\omega\rvert^{m}_{\mathrm{mad}}$ (cm^{-1})	$\lvert\Delta\omega\rvert^{m}_{\mathrm{rms}}$ (cm^{-1})
NumFreq1 (Grid 7) / *NumFreq2* (Grid 7)	0.21	0.09	0.14	0.15	0.38	0.18	0.06	0.13
AnFreq (Grid 7) / *NumFreq1* (Grid 7)	0.47	0.21	0.47	0.47	0.61	0.26	0.1	0.17
AnFreq (Grid 7) / *NumFreq2* (Grid 7)	0.51	0.25	0.51	0.51	0.65	0.3	0.13	0.21
NumFreq1 (Grid 4) / *NumFreq2* (Grid 4)	8.67	4.1	8.67	8.67	7.24	3.32	1.82	2.6
NumFreq1 (Grid 4) / *NumFreq2* (Grid 7)	1.03	0.64	1.03	1.03	1.04	0.7	0.34	0.43
AnFreq (Grid 4) / *NumFreq2* (Grid 7)	0.82	0.54	0.58	0.61	1.32	0.84	0.37	0.48
AnFreq (Grid 4) / *NumFreq1* (Grid 4)	0.49	0.22	0.49	0.49	0.59	0.27	0.1	0.17
NumFreq2 (Grid 4) / *NumFreq2* (Grid 7)	7.64	3.84	7.64	7.64	6.6	3.19	1.59	2.28
	ω							
	$\lvert\Delta\omega\rvert^{m}_{\max}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{\max}$ (cm^{-1})	$\lvert\Delta\omega\rvert^{m}_{\mathrm{mad}}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{\mathrm{mad}}$ (cm^{-1})	$\lvert\Delta\omega\rvert^{m}_{\mathrm{rms}}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{\mathrm{rms}}$ (cm^{-1})		
NumFreq1 (Grid 7) / *NumFreq2* (Grid 7)	0.38	0.18	0.06	0.04	0.13	0.06		
AnFreq (Grid 7) / *NumFreq1* (Grid 7)	0.61	0.29	0.09	0.05	0.16	0.09		
AnFreq (Grid 7) / *NumFreq2* (Grid 7)	0.65	0.34	0.13	0.07	0.2	0.11		
NumFreq1 (Grid 4) / *NumFreq2* (Grid 4)	8.67	3.75	2.03	1.08	2.84	1.41		
NumFreq1 (Grid 4) / *NumFreq2* (Grid 7)	1.04	0.73	0.36	0.17	0.46	0.25		
AnFreq (Grid 4) / *NumFreq2* (Grid 7)	1.32	0.84	0.39	0.19	0.5	0.27		
AnFreq (Grid 4) / *NumFreq1* (Grid 4)	0.59	0.3	0.1	0.06	0.17	0.09		
NumFreq2 (Grid 4) / *NumFreq2* (Grid 7)	7.64	3.52	1.81	1.03	2.54	1.35		

Benchmark result: efficiency and timing

For the efficiency and timings we have chosen B3LYP and BP86 as representatives of the hybrid and pure DFT methods. The basis set def2-TZVP in all timing calculations was utilized. The calculations were done for the whole test set, but for the sake of compactness only five molecules with different numbers of atoms and numbers of basis functions were chosen to demonstrate the trends.

Calculations labeled RIJDX used the RI-J approximation for the Coulomb term and exact analytical integration for the exchange contribution. The RIJDX approximation was shown to be very accurate independently from which of the two-electron integral contribution terms it was applied to. The maximum RMS error through the whole test set is only 0.85 cm^{-1}, Table 3.27. Therefore, the RIJDX approximation default refers to the situation where the RI-J is in all two-electron Coulomb terms.

Table 3.27 Statistical analysis of the errors in the calculated vibrational frequencies (ω) corresponding to different two-electron flags *HESS*(i_1, i_2, i_3, i_4) , as determined with respect to the results obtained in the approximation-free calculations with the ExtremeSCF settings. In these notations, i_1, i_2, i_3, i_4 refer to the approximation methods for the calculation the two-electron integral contributions in $\mathbf{F}^{(\vartheta)}$, $\mathbf{G}\left[\mathbf{Q}^{\vartheta}\right]$, $\mathbf{G}\left[\mathbf{U}^{\vartheta}\right]$, and $E^{(\vartheta)(\lambda)}$, respectively. If i_k=0 the corresponding Coulomb and exchange four-center integrals are calculated exactly; the flags i_k=1 and i_k=2 indicate the RIJDX and RIJCOSX approximations, respectively. The calculations were performed at the HF/def2-TZVP level of theory, using the TightSCF settings if not otherwise indicated.

	ω<200 cm^{-1}				ω>200 cm^{-1}			
HESS Flags	$\lvert\Delta\omega\rvert^{m}_{\max}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{\max}$ (cm^{-1})	$\lvert\Delta\omega\rvert^{m}_{\mathrm{mad}}$ (cm^{-1})	$\lvert\Delta\omega\rvert^{m}_{\mathrm{rms}}$ (cm^{-1})	$\lvert\Delta\omega\rvert^{m}_{\max}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{\max}$ (cm^{-1})	$\lvert\Delta\omega\rvert^{m}_{\mathrm{mad}}$ (cm^{-1})	$\lvert\Delta\omega\rvert^{m}_{\mathrm{rms}}$ (cm^{-1})
(0,0,0,0)	0.56	0.16	0.56	0.56	0.29	0.13	0.06	0.09
(1,1,1,1) (RIJDX)	1.28	0.36	1.28	1.28	2.52	1.43	0.68	0.87
(1,1,2,1)	3.3	0.69	1.64	1.82	9.08	2.22	1.02	1.7
(1,2,2,1)	3.3	0.69	1.65	1.81	9.07	2.23	1.01	1.7
(2,1,2,1)	5.92	1.6	5.92	5.92	19.73	3.5	1.72	3.16
(2,2,2,1)	5.89	1.6	5.89	5.89	19.72	3.51	1.73	3.16
(2,2,2,1) (OldGridX)	4.16	0.84	1.97	2.4	19	2.46	1.37	2.92
(2,2,2,2)	392.38	287.74	347.55	348.68	115.49	65.26	44.89	51.2
NumFreq ExtremeSCF	1.42	0.15	1.42	1.42	1.82	0.41	0.45	0.51
NumFreq	5.68	1.95	3.8	3.8	9	3.48	2.13	3.51

	ω					
	$\lvert\Delta\omega\rvert^{m}_{\max}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{\max}$ (cm^{-1})	$\lvert\Delta\omega\rvert^{m}_{\mathrm{mad}}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{\mathrm{mad}}$ (cm^{-1})	$\lvert\Delta\omega\rvert^{m}_{\mathrm{rms}}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{\mathrm{rms}}$ (cm^{-1})
(0,0,0,0)	0.56	0.17	0.06	0.03	0.12	0.05
(1,1,1,1) (RIJDX)	2.52	1.43	0.66	0.45	0.85	0.61
(1,1,2,1)	9.08	2.24	1.02	0.52	1.7	0.73
(1,2,2,1)	9.07	2.25	1.01	0.51	1.7	0.72
(2,1,2,1)	19.73	3.56	1.76	0.69	3.16	1.04
(2,2,2,1)	19.72	3.56	1.77	0.69	3.16	1.04
(2,2,2,1) (OldGridX)	19	2.47	1.4	0.52	2.9	0.75
(2,2,2,2)	392.38	205.37	82	24.6	125.67	49.76
NumFreq ExtremeSCF	1.82	0.46	0.45	0.11	0.5	0.15
NumFreq	9	3.56	2.08	0.69	3.4	1.08

In the introduced notation HESS(i_1, i_2, i_3, i_4), where i_k (i_k=0–2) specify the approximation methods for different two-electron terms $\mathbf{F}^{(\vartheta)}$, $\mathbf{G}\left[\mathbf{Q}^{\vartheta}\right]$, $\mathbf{G}\left[\mathbf{U}^{\vartheta}\right]$, and $E^{(\vartheta)(\lambda)}$ the two-electron flags for the RIJDX are HESS(1, 1, 1, 1).

The efficiency of the RI approximation for the Coulomb derived terms in the evaluation of the second SCF energy derivative can be best illustrated using pure DFT. For the BP86, Table 3.28, the most time consuming step is the solution of the CP-SCF equations, which takes up to 80% of the total hessian evaluation time for the 4-center implementation. This term can be calculated up to 25 times faster using the RIJDX approximation. Overall, the RIJDX implementation is up to 18 times faster than the 4-center implementation for the presented test set of molecules. Obviously, the larger molecules will show even bigger efficiency.

Calculations denoted as RIJCOSX, Table 3.28, used the COSX approximation for the exchange terms in conjunction with the RI-J approximation. From the series of accuracy determining HF calculations the default for the RIJCOSX was chosen to be HESS(1, 2, 2, 1). Namely, the COSX approximation is introduced to the Exchange like underived terms $\mathbf{G}\left[\mathbf{Q}^{\vartheta}\right]$ and $\mathbf{G}\left[\mathbf{U}^{\vartheta}\right]$. Introducing the COSX approximation in the first and second derivative of the Exchange like terms leads to immediate drop in the accuracy. The needed accuracy for the first derivative term $\mathbf{F}^{(\vartheta)}$ can be easily achieved using bigger COSX integration grid (Table 3.27, GridXOld results), but the efficiency is lost in this case. The COSX in $E^{(\vartheta)(\lambda)}$ leads to inaccurate results. The default conservative choice is representing the best balance between accuracy and efficiency.

Important observations were made for the timings of the CP-SCF equations solution. Table 3.28 presents the percentage of the time to solve the equations relatively to the total hessian time. For the approximation free method the statement of Deglmann et. al.,[102] about the fact that the $\mathbf{F}^{(\vartheta)}$, $\mathbf{G}\left[\mathbf{Q}^{\vartheta}\right]$ and $E^{(\vartheta)(\lambda)}$ terms are not the dominant constituents of computation times, is confirmed. The $\mathbf{G}\left[\mathbf{U}^{\vartheta}\right]$ term takes up to 87% of the total time in the 4-center hessian evaluation. Also, in the RIJDX method the CP-SCF solution takes up to 79% of the total time for the molecules from the test set. The situation changes upon applying the COSX approximation. The relative weight of the CP-SCF time drops significantly and constitutes only 33% of the total time. Thus, it is also desirable to introduce the RI in the derivative integral terms $\mathbf{F}^{(\vartheta)}$ and $E^{(\vartheta)(\lambda)}$. Moreover, detailed error analysis proved the error coming from RI to be absolutely insignificant (RMS max error in the test set is less the 1 cm^{-1}).

To compare with the NumFreq implementation the time ratio of NumFreq/AnFreq implementations also was added to the Table 3.28. Of course, the RIJDX and RIJCOSX approximations were applied to both NumHess and AnHess implementations in the calculation of the NumHess/AnHess ratio. For the B3LYP method the analytic implementation for the 4-center and RIJDX is 2-4 times faster than numerical. Introducing RIJCOSX with B3LYP gives even better results of 4 to 7 times. The BP86 RIJDX methods perform also 2-4 times faster in case of analytical implementation, whereas the 4-center approximation is even more favorable in the analytical implementation and amount up to 10-20 times faster for some test set cases.

Table 3.28 The timings for the *AnHess* program. All numbers are in seconds. Corresponding two-electron flags *HESS*(i_1, i_2, i_3, i_4), were set to *HESS*(*1*, *1*, *1*, *1*) for the RIJDX approximation and to *HESS*(*1*, *2*, *2*, *1*) for the RIJCOSX. In these notations, i_1, i_2, i_3, i_4 refer to the approximation methods for the calculation the two-electron integral contributions in $\mathbf{F}^{(\vartheta)}$, $\mathbf{G}\left[\mathbf{Q}^{\vartheta}\right]$, $\mathbf{G}\left[\mathbf{U}^{\vartheta}\right]$, and $E^{(\vartheta)(\lambda)}$, respectively. If i_k=0 the corresponding Coulomb and exchange four-center integrals are calculated exactly; the flags i_k=1 and i_k=2 indicate the RIJDX and RIJCOSX approximations, respectively. The calculations were performed using def2-TZVP basis set and the TightSCF settings. In the calculation of *NumHess/AnHess* ratio the RIJDX and RIJCOSX approximations were applied to both *NumHess* and *AnHess* implementations.

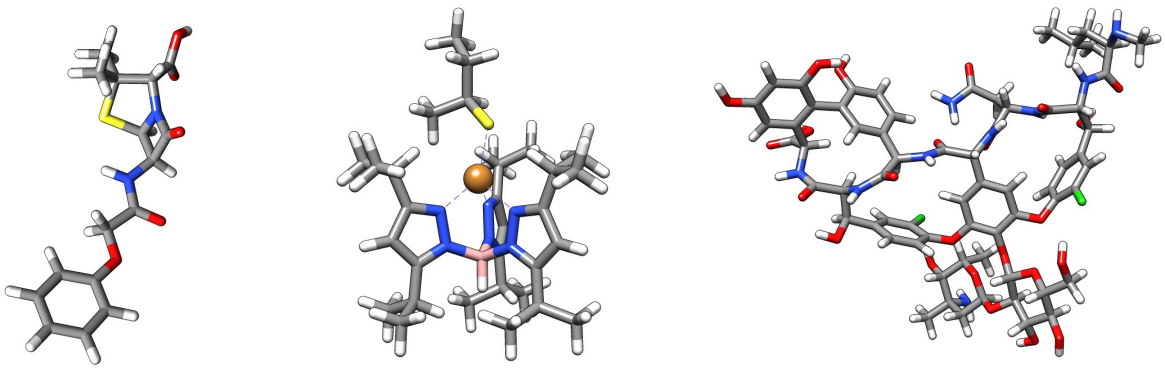
Number of atoms	Number of basis functions	$\mathbf{F}^{(\vartheta)}$		$\mathbf{G}\left[\mathbf{Q}^{\vartheta}\right]$		$\mathbf{G}\left[\mathbf{U}^{\vartheta}\right]$		$E^{(\vartheta)(\lambda)}$		Total Hessian time	Wall Clock time	*AnHess/NumHess* ratio
		abs. time	%	abs. time	%	abs. time	%	abs. time	%			
B3LYP, def2-TZVP, 4 center												
5	105	24	14	12	6	71	40	72	40	178	199	4.0
11	241	234	3	509	6	6271	78	1022	13	8035	8328	2.1
17	227	382	4	1131	11	8013	76	1073	10	10599	10886	2.2
24	444	3507	2	16548	8	162354	83	13131	7	195540	198861	3.6
36	516	5170	1	47843	9	444491	87	14779	3	512282	516298	1.6
B3LYP, def2-TZVP, RIJONX												
5	105	22	14	10	6	52	34	68	45	151	191	4.1
11	241	215	5	287	6	3013	66	1016	22	4531	4861	3.4
17	227	355	6	625	10	4144	69	894	15	6018	6330	3.8
24	444	3114	3	12563	13	72130	72	11992	12	99798	104687	6.1
36	516	4042	2	34518	14	190673	79	11927	5	241160	245597	2.6
B3LYP, def2-TZVP, RIJCOSX												
5	105	23	18	8	7	24	19	71	56	125	145	4.2
11	241	198	14	91	7	321	23	781	56	1391	1547	5.4
17	227	393	19	204	10	472	23	994	48	2063	2227	5.6
24	444	3474	17	1435	7	4079	20	11372	56	20359	21355	5.1
36	516	4713	17	3024	11	9400	33	11372	40	28509	29796	7.1
BP86, def2-TZVP, 4 center												
5	105	21	10	7	3	40	19	143	68	211	255	21.6
11	241	187	6	141	5	1890	62	833	27	3052	3386	5.8
17	227	855	10	681	8	6323	71	1039	12	8898	10547	10.8
24	444	2874	3	9932	10	75635	75	12853	13	101295	108843	9.9
36	516	4056	2	32575	13	201214	80	13654	5	251499	255941	4.0
BP86, def2-TZVP, RIJDX												
5	105	14	30	5	10	16	35	12	25	46	52	3.6
11	241	98	17	54	10	335	60	74	13	562	588	3.3
17	227	250	17	132	9	949	64	154	10	1485	1537	2.3
24	444	1560	21	978	13	4114	56	717	10	7369	7489	2.7
36	516	2737	20	1802	13	7978	57	1364	10	13880	14049	3.2

To demonstrate the AnHess implementation capabilities three representative calculations on biologically relevant molecules were conducted (Table 3.29). The structure of the molecules is depicted in Figure 3.40. All calculations used 16 Intel® Xeon® 2.53GHz processors.

Table 3.29 The *AnHess* program timings for three representative examples. All numbers are in seconds. The B3LYP/RIJCOSX methods with default two-electron flags *HESS*(*1*, *2*, *2*, *1*) was utilized in these calculations. The calculations were performed using def2-TZVP basis set for Penicillin and def2-SVP for Plastocyanin and Vancomycine. The TightSCF settings were used in all calculations.

	$\mathbf{F}^{(\vartheta)}$		$\mathbf{G}\left[\mathbf{Q}^{\vartheta}\right]$		$\mathbf{G}\left[\mathbf{U}^{\vartheta}\right]$		$E^{(\vartheta)(\lambda)}$		Total Hessian time	Wall Clock time
	abs. time	%	abs. time	%	abs. time	%	abs. time	%		
Penicillin	369	9.7	417	11	1881	49.7	1121	29.6	3787	4011
Plastocyanin	1026	5.4	2180	11.4	14589	76.5	1277	6.7	19071	19579
Vancomycine	8569	9.9	13176	15.2	58346	67.2	6765	7.8	86855	87307

Penicillin comprises 42 atoms and has 126 degrees of freedom. Def2-TZVP basis was utilized in these calculations, which means 858 contracted basis functions. The wall clock time for the single point calculation with hessian is less than one hour. The Plastocyanin model compound consists of 95 atoms, including the copper metal center, and 856 contracted basis functions. Def2-svp basis calculation was done in 5 hours and 26 minutes. Vancomycine, already quite challenging molecule with 176 atoms and 1797 basis functions, can be calculated within one day (24 hours, 15 minutes). To handle molecules of this size an efficient batching procedure must be introduced. In our implementation the batching is done separately for each step of the calculation, thus making optimal use of the memory available. Wherever possible data is held in memory to enable speedy processing.



Penicillin Plastocyanin model Vancomycine

Figure 3.40 Three representative examples for the *AnHess* program: Penicillin, Plastocyanin model and Vancomycine. Penicillin has 42 atoms, Plastocyanin model - 96, including the Cu metal center, and Vancomycine - 176.

Figure 3.41 demonstrates the efficiency of the COSX parallelization in the evaluation of different contributions: $\mathbf{F}^{(\vartheta)}$, $\mathbf{G}\left[\mathbf{Q}^{\vartheta}\right]$, $\mathbf{G}\left[\mathbf{U}^{\vartheta}\right]$, and $E^{(\vartheta)(\lambda)}$. The double integral derivative term $E^{(\vartheta)(\lambda)}$

scales very good, the speedup of 14,5 can be achieved for 16 processors. The first derivative terms $\mathbf{F}^{(\vartheta)}$ and $\mathbf{G}\left[\mathbf{Q}^{\vartheta}\right]$ scale also quite well, having speedups of 13,1 and 12,4 for 16 processors, correspondently. The CP-SCF solution term $\mathbf{G}\left[\mathbf{U}^{\vartheta}\right]$ scales somewhat less efficiently, having speedup of only 9,9 for 16 processors. The inferior scaling of the $\mathbf{G}\left[\mathbf{U}^{\vartheta}\right]$ term is due to the fact that for this step a different parallelization scheme had to be applied. In all other cases parallelization was done by distributing the total number of perturbations among the processes, thus allowing independent and fast processing of these calculations. For the CP-SCF solution using RIJCOSX we had to solve each part of the calculation with all processes, which means a lot of communication and synchronization, thereby a loss of time. Nevertheless, the total scaling of 11,9 can be achieved for the parallel hessian calculation on 16 processors.

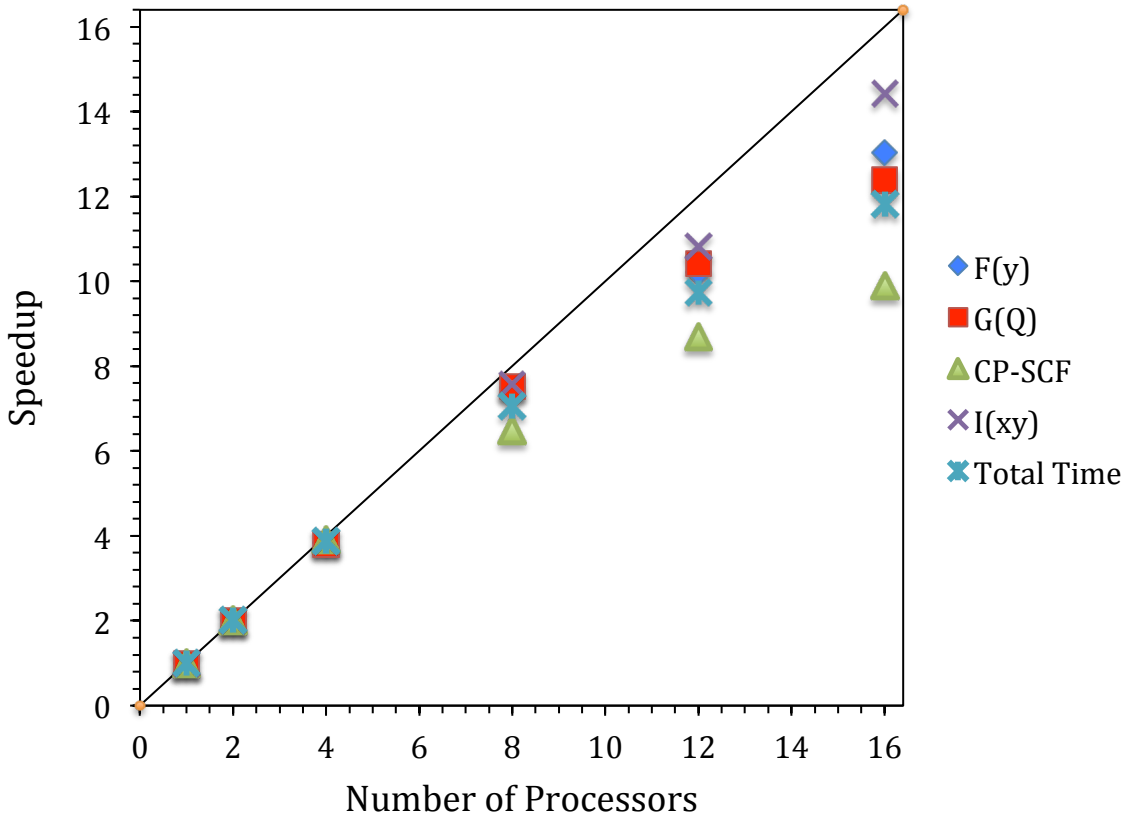


Figure 3.41 The parallelization efficiency of the COSX algorithm for the analytical hessian implementation. Penicillin molecule with 42 atoms was treated under B3LYP/def2-TZVP level of theory. Blue F(y) denotes Fock derivative terms $\mathbf{F}^{(\vartheta)}$, red G(Q) corresponds to $\mathbf{G}\left[\mathbf{Q}^{\vartheta}\right]$, green CP-SCF is $\mathbf{G}\left[\mathbf{U}^{\vartheta}\right]$, and $E^{(\vartheta)(\lambda)}$ represented as violet I(xy). The total hessian time is depicted with turquoise color.

3.6.3 Conclusions

In this work, we have presented efficient approximations for the evaluation of Hartree–Fock and hybrid DFT molecular hessian. The resolution of the identity RI combined with the chain of spheres approximation were applied to different stages of the hessian evaluation. It has been shown that the algorithm is efficient and particularly advantageous for basis sets dominated by high angular momentum for which the observed speedup is up to 2 orders of magnitude relative to conventional 4 center numerical differentiation. The accuracy of the method was studied in details. For the prediction of molecular frequencies the new RIJCOSX method has shown very accurate results (typical errors relative to conventional results <5 cm^{-1}). The efficient parallelization and memory handling combined with the potentially linear scaling with respect to system size makes new hessian implementation highly desirable tool in the ORCA quantum chemistry project. For instance, the method was successfully applied to the calculation of molecular hessian for the vancomycine molecule of 176 atoms and almost 18 hundreds basis functions. The molecular hessian was calculated within one day of total wall clock time.

4. CONCLUSIONS

This thesis is dedicated to three major scientific subjects: (1) modern theoretical methods for studying enzymatic reactivity, (2) support of advanced spectroscopy experiments with theoretical predictions, (3) implementation of second analytical energy derivatives on the self-consistent field (SCF) theory level. The enzyme of interest was cytochrome *c* nitrite reductase. Here special accent was put on kinetics simulation and proton-coupled electron transfer reaction. The theoretical support of EPR and Mössbauer experiments was given in the investigation of Hase 1 hydrogenase and cd1 nitrite reductase. The third part of the thesis includes implementation of SCF second derivatives into ORCA program and its combination with resolution of identity and chain of spheres methods. In the following the conclusions for each of three research directions are presented.

The mechanism of C*c*NiR enzyme has been studied. Two plausible scenarios were calculated that are both compatible with the experimentally determined rate (suggesting an effective barrier of 15.9 kcal/mol). The two pathways differ in the assumed protonation state of His_{277}. The reaction proceeds relatively smoothly if one assumes that this residue is doubly protonated and can be recharged during the reaction cycle. Alternatively, or perhaps at higher pH, Arg_{114} may serve as proton donor in a pathway that is energetically similarly feasible as the one proceeding via His_{277}.

There are two major effects which facilitate substrate activation in the Fe(II) – nitrite complex: a) back-bonding and b) interaction of nitrite with second sphere residues. As a consequence of the back-bonding effect, the NO_2^- based LUMO antibonding orbital becomes populated and hence the N–O bond is weakened. Moreover, the interaction of the nitrite oxygens with charged side chains causes an asymmetric charge distribution which breaks the equivalence of the N–O bonds (both with a formal bond order of 1.5) and hence activates the more single-bond like N–O bond for reductive cleavage. This was already hypothesized in the earlier study[40] and could now be substantiated using larger computational models of the active site.

In addition, it was found here that the conformation with which the NO_2^- molecule bonds to the active site is optimal for the proton attack from His_{277} and simultaneously for strong back-bonding. Thus, the most acidic side chain in the active site of the C*c*NiR, His_{277}, is located optimally for carrying out protonation steps. Both features are reflected in the only moderately endothermic reaction energy of +4,9 kcal/mol and the calculated barrier of 5,5 kcal/mol for proton transfer. The formed HONO adduct represents the lowest conformer with respect to HO-N=O rotation. Since the rotational potential energy surface is rather rigid, the only feasible way for the transfer of the second proton is the recharging of His_{277} in this pathway. This can be achieved in a reasonably endothermic step and is rendered irreversible by a barrierless reaction leading to N–O bond cleavage and formation of the $\{Fe(NO)\}^6$ species. Ultimately, the source of protons in the active site is an unresolved question. The Ca-site could play a significant role not only as a structural motif but also in shuttling protons to the active site. These issues deserve further investigation.

Alternatively, or at higher pH, another pathway is considered to be feasible. Above pH ~8 His_{277} is probably in its deprotonated form. Hence, the proton donor functionality must be adopted by Arg_{114}. Our calculations suggest that this is energetically and kinetically feasible.

Although the acidity of Tyr_{218} is certainly much lower than that of histidine, it was considered as a proton donor in each step of the reaction. All calculated reaction and activation energies are very much higher than the ones found for His_{277} and Arg_{114}. Hence, we exclude a significant role of this residue in the initial N–O bond cleavage chemistry.

Detailed insights into the recharging process of the C*c*NiR active site following the first N-O bond cleavage have been done. Taking into account the possible redox states of the active site core ($\{Fe(NO)\}^6$-$\{Fe(NO)\}^8$) together with possible protonation states of the adjacent active site ligands and the substrate makes it possible to trace a consecutive series of events leading to the fully recharged complex $\{FeHNO\}^8$. The donation of two electrons and three protons is needed to reach the $\{FeHNO\}^8$ intermediate with protonated Tyr_{218} and Arg_{114} side chains.

An analysis of the conformational behavior and the changes in the electronic structure of the $\{Fe(NO)\}^6$-$\{Fe(NO)\}^8$ series revealed mostly ligand-centered reduction processes. As a consequence of two electron reduction, one of the NO π-bonds is broken and two separate lone pairs on N(NO) and O(NO) are formed. The N(NO) centered lone pair is the center for the subsequent nucleophilic attack forming HNO.

The reduction of the active site is accomplished by transport of electrons from the external electron donor (a membrane-anchored tetra-heme cytochrome (NfrH) redox partner in case of *W. succinogenes* and soluble penta-heme cytochrome NrfB in case of *E.coli*) through a cascade of redox-active hemes. Theoretically, the energy of the incoming electron can be estimated from the adiabatic electron affinity of the neighboring heme. The calculations showed that the heme can easily be oxidized by the active site either in the ET reaction CH→CEH or in the PCET process CH→CEH2. The second electron, however, can only be supplied upon addition of an extra proton.

The proton supply issue was addressed. The fact that the Ca-site is placed in the inlet channel in direct proximity of the active site and, as was discussed previously, plays an important role in the substrate reduction (presumably as a structural element), implies that it can also effectively modulate the proton supply process. Indeed, a proton localized on the Ca-site is approximately 6 kcal/mol higher in energy than a proton in pure water solution. Thus, the stages involving a proton transfer become thermodynamically more favorable. Moreover, the calculated ΔH of water dissociation in the Ca-site is significantly smaller than the corresponding ΔH of pure water dissociation calculated at the same level of theory (a well known effect of water coordination to a metal center). Consequently, the Ca-site could not only act as an effective proton mediator but also as an acid/based catalytic subunit, which supplies the protons required for the reduction process.

Detailed considerations regarding the thermodynamics of the recharging process have been made. For the purpose of analysis, the calculated Gibbs free energy differences were transformed to pK_a's and redox potentials. Our calculations confirm the existence of a deep potential energy minima for the CEH2 complex ($\{Fe(NO)\}^7$ electronic configuration), which was also proposed experimentally to represent a thermodynamic sink.[40] Consequently, the system must either overcome or avoid this energy minimum as it is known that the reduction proceeds further to the fully reduced product, NH_3, without the release or accumulation of any detectable intermediates. According to our calculations two scenarios are possible. The first one consists of two consecutive PCET reactions (CH→CEH2→CEEH2_HNO) and the alternative one is a reduction (CH→CEH) followed by PCET (CEH→CEEH_HNO) and protonation (CEEH_HNO→CEEH2_HNO). The first pathway is

characterized by a highly exothermic first PCET step and almost thermoneutral second PCET step. The calculated free energy difference for the first PCET is -46.2 kcal/mol and the following second step needs ~0.7 kcal/mol to be accomplished. The second pathway does not contain endothermic reactions and involves only moderately exothermic reactions.

Additional insight into the competition between the two pathways can be obtained by analyzing the kinetics of the process. To this end, Marcus theory[138] combined with methods developed by Rosso143 and Hammes-Schiffer139 were used to estimate the activation barriers. These calculations make a detailed consideration of the entire recharging process possible. Importantly, both pathways proposed from the thermodynamic analysis scheme also appear to be kinetically feasible. Recharging due to two PCET reactions proceeds through two moderate activation barriers of 1.6 and 6.5 kcal/mol. The alternative scheme includes an electron transfer barrier of 2.3 kcal/mol (CH→CEH), followed by PCET (CEH→CEEH2) with a barrier of 5.4 kcal/mol and proton transfer (CEEH2→CEEH3) with a barrier 0.1 kcal/mol. In either pathway CEEH2_HNO (electronic configuration $\{FeHNO\}^8$) is formed. Simulation of the reaction kinetics showed that the first pathway is more probable. This is based on the fact that the first reduction proceeds rapidly and leads to the formation of the CEH2 complex. The latter is then undergoes conversion to the CEEH2_HNO in the rate-limiting equilibrium reaction with an activation barrier of 6.5 kcal/mol.

It should be noted that the complex recharging process in the active site of C*c*NiR can only be accomplished through a well organized network of proton and electron donors. Neighboring low-redox potential hemes are assumed to be effective electron donors. The change in the reduction state is then immediately accompanied by proton transfer from the inlet channel, part of which is the Ca-site. It is this coupling between electron and proton transfer, (a characteristic of biological processes) that takes place in the largely hydrophobic interior of proteins, which makes efficient recharging of the active site possible.

The second half cycle of the nitrite reduction catalyzed by Cytochrome *c* nitrite reductase has been considered in details. In total 3 electrons and 4 protons must be provided to reach the final product ammonia starting form HNO intermediate. According to our results, the first event in this half cycle is the reduction of the HNO intermediate accomplished by two PCET reactions. The isomeric intermediates HNOH* and H_2NO* are formed (* denotes an electron). Both intermediates are active and are readily transformed into hydroxylamine most likely through intramolecular proton transfer either from Arg_{114} or His_{277}. The protonated side chain then provides its proton to initiate a heterolytic cleavage reaction of the N-O bond. As a result the H_2N^+ intermediate is formed. The letter readily picks up an electron forming H_2N^+* which in turn reacts with Tyr_{218}. Interestingly, an evidence for Tyr_{218} activity was provided by the mutational studies of Lukat (Lukat, P. et. al., 2008, *Biochemistry 47*, 2080) but have never been observed on initial stages of the overall reduction process. According to our results, intramolecular reaction with Tyr_{218} in the final step of the nitrite reduction process leads directly to the H_3N^+* ammonia final product. The product dissociation was found to proceed through the change of spin state, which was also observed in resonance Raman investigation of Martins (Martins, G., et. al. (2010), *J Phys Chem B 114*, 5563)

As the second scientific direction of the presented thesis EPR and Mösbauer experiments were supported by the theoretical investigations. In particular, was once more demonstrated the value of high field ENDOR combined with DFT calculations for the characterization of H-bonds and highlighted the important role of DFT in assigning and interpreting ENDOR spectra in terms of

structural details for complicated systems. It has been shown that the NO in the nitrosyl d_1-heme complex of cd_1 NIR forms H-bonds with Tyr_{10} and His_{369} that affect the stabilization of the NO complex. The second conserved His_{327}, appears to be less involved in NO stabilization by H-bonding. This is in contrast to the WT-NO crystal structure where Tyr_{10} is not at an H-bond distance from the NO. A larger solvent accessibility to the distal pocket in the mutants compared to the WT has also been observed. Finally, it is clear from this work that the H-bonding network within the active site is dynamic and that a change in one of the residues does affect the strength and position of the H-bonds formed by the other ones. In the Y10F mutant His_{369} moves closer to the NO and its hydrogen bond is shorter, whereas mutation of both distal histidines displaces Tyr_{10} removing its H-bond.

The support of the Mössbauer spectroscopic experiments with theoretical findings has been carried out. Based on the ΔE_Q values alone, the best BS solution for the super-oxidized cluster is labeled Ox2_14 (spin populations of Fe1 and Fe4 flipped) and corresponds to BS13 in ref.[84] However, this solution does not properly reproduce the experimental δ values or the sign of the A values, of which the latter reveals a basically wrong spin coupling scheme. Instead, our best solution, which is found within 4 kcal/mol of the energy minimum, has the spin populations of Fe2 and Fe4 flipped (Ox2_24), yielding proper values for δ, ΔE_Q, η, and the A values. In both cases, Glu_{82} is required to be unprotonated.

On the basis of the calculations, the special subspectrum (S) of the super-oxidized $[4Fe\text{-}3S]^{5+}$cluster with the large quadrupole splitting (ΔE_Q = 2.45 mms^{-1}) is assigned to the nitrogen-coordinated Fe2. This agrees with chemical intuition, where a large *efg* is expected due to the anisotropic covalency that results from asymmetric coordination with a unique, short N-bond.[81] The result is in contrast to the assignment of subspectrum (S) to Fe1,[84] which has quasi-tetrahedral symmetry. Note that our suggested spin distribution ($\alpha\beta\alpha\beta$ spin alignment for Fe1, Fe2, Fe3, Fe4, respectively) does not correspond to a standard coupling scheme observed for iron-sulfur clusters.

Selection of the best BS solution for the reduced $[4Fe\text{-}3S]^{3+}$ cluster is less clear cut for the super-oxidized state because several reasonable choices exist. Importantly all reasonable choices imply that Glu_{82} is unprotonated. This raises some concerns with respect to its suspected role as a base.[84] Our best result (Red2_24) shows the spin populations on Fe2 and Fe4 to be flipped. Based on this solution, Fe4 and Fe1 are assigned to the special subspectra (S) and (F) found for the reduced proximal cluster, *i.e.* those with largest ΔE_Q and highest δ, respectively (Table 3.20A). Unlike for the super-oxidized case, the optimized structure for Red2_24 forms a tightly closed cubane with a long Fe4 – S-Cys25 separation. The spin distribution obtained from the calculations, in agreement with the experimental A-values, is best explained by the presence of a mixed-valence pair (Fe1 and Fe3) and two 'ferrous' sites (Fe2 and Fe4), which render the special subspectrum (F) with the highest isomer shift part of the mixed-valence pair.

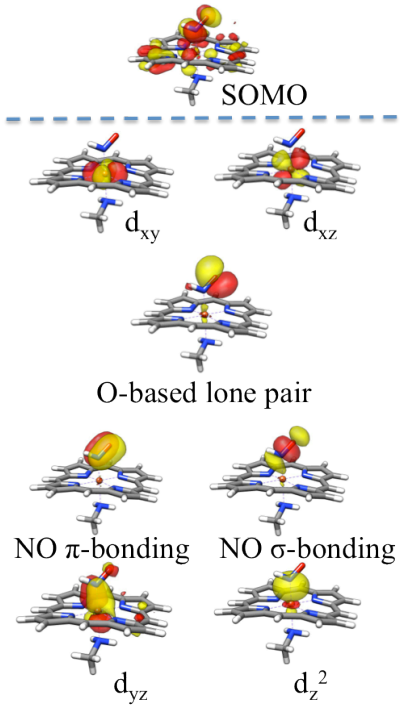
The theoretical third part of the thesis deals with the implementation of the analytical SCF energy derivatives. In this work, efficient approximations for the evaluation of Hartree–Fock and hybrid DFT molecular hessian have been presented. The resolution of the identity (RI) combined with the chain of spheres (COSX) approximation was applied to different stages of the hessian evaluation. It has been shown that the algorithm is efficient and particularly advantageous for basis sets dominated by high angular momentum for which the observed speedup is up to 2 orders of

magnitude relative to conventional 4 center numerical differentiation. The accuracy of the method was studied in details. For the prediction of molecular frequencies the new RIJCOSX method has shown very accurate results (typical errors relative to conventional results <5 cm^{-1}). The efficient parallelization and memory handling combined with the potentially linear scaling with respect to system size makes new hessian implementation highly desirable tool in the ORCA quantum chemistry project. For instance, the method was successfully applied to the calculation of molecular hessian for the vancomycine molecule of 176 atoms and almost 18 hundreds basis functions. The molecular hessian was calculated within one day of total wall clock time.

APPENDICES

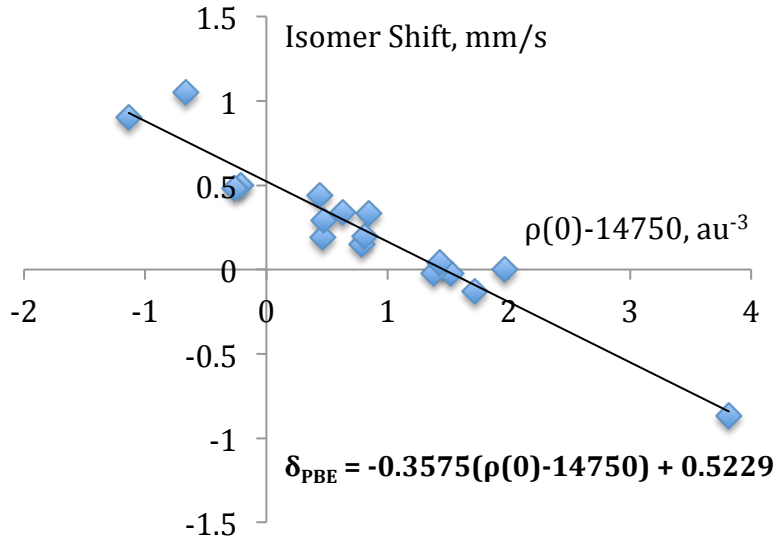
A.

Contour plots of important MOs calculated under B3LYP/TZV(2d,2p) level of theory for the (HNO*)H_TN_A complex. Active site side chains are not depicted for simplicity. The occupied orbitals were localized according to Pipek-Mezey (PM) algorithm. The unoccupied orbitals are not localized.

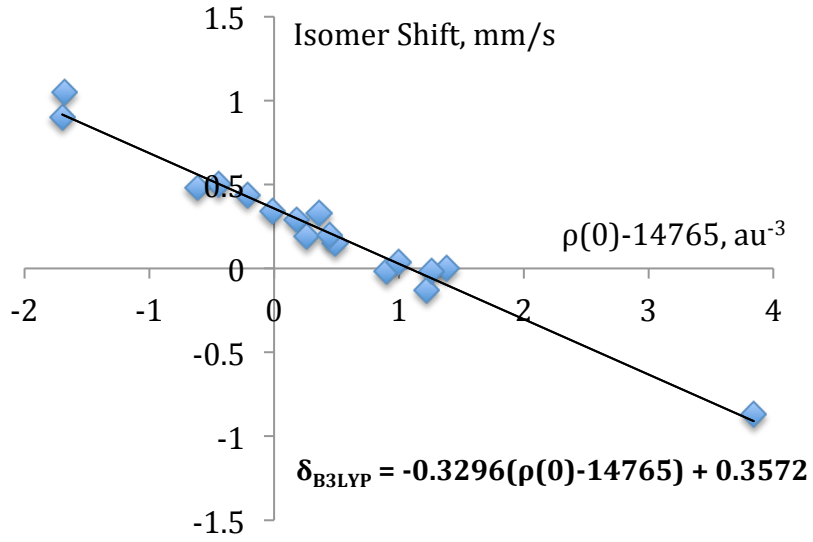


B.

Calibration of the PBE(a) and B3LYP(b) functionals for the prediction of 57Fe isomer shifts. The calibration procedure includes calculation of the electron density at the nuclei of interest and comparison to the experimentally known isomer shift values. The linear correspondence is then fitted to a straight line using the least squares method. Importantly, to achieve better accuracy in isomer shifts calculations, the "core properties" CP(PPP) basis set was used with the radial integration accuracy parameter increased to 9.0 for iron centers.



a



b

C1. Statistical analysis of the errors in the calculated vibrational frequencies (ω) corresponding to the ExtremeSCF(ZM) settings, as determined with respect to the results obtained in the approximation-free ExtremeSCF calculations. The calculations were performed at the HF/def2-TZVP level of theory.

	ω<200 cm^{-1}				ω>200 cm^{-1}			
	$\lvert\Delta\omega\rvert^m_{\max}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{\max}$ (cm^{-1})	$\lvert\Delta\omega\rvert^m_{\mathrm{mad}}$ (cm^{-1})	$\lvert\Delta\omega\rvert^m_{\mathrm{rms}}$ (cm^{-1})	$\lvert\Delta\omega\rvert^m_{\max}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{\max}$ (cm^{-1})	$\lvert\Delta\omega\rvert^m_{\mathrm{mad}}$ (cm^{-1})	$\lvert\Delta\omega\rvert^m_{\mathrm{rms}}$ (cm^{-1})
ExtremSCF(Z6)	28.94	3.99	28.94	28.94	8.62	1.37	0.85	1.99
ExtremSCF(Z7)	24.65	1.89	24.65	24.65	3.76	0.47	0.27	0.8
ExtremSCF(Z8)	3.79	0.61	2.07	2.69	0.41	0.13	0.05	0.09
ExtremSCF(Z9)	3.32	0.27	1.71	2.35	0.51	0.05	0.03	0.08
ExtremSCF(Z10)	2.35	0.14	1.05	1.43	0.51	0.03	0.02	0.08

	ω					
	$\lvert\Delta\omega\rvert^m_{\max}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{\max}$ (cm^{-1})	$\lvert\Delta\omega\rvert^m_{\mathrm{mad}}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{\mathrm{mad}}$ (cm^{-1})	$\lvert\Delta\omega\rvert^m_{\mathrm{rms}}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{\mathrm{rms}}$ (cm^{-1})
ExtremSCF(Z6)	28.94	3.25	1.15	0.41	4.48	0.77
ExtremSCF(Z7)	24.65	1.95	0.73	0.15	3.95	0.39
ExtremSCF(Z8)	3.79	0.45	0.19	0.04	0.76	0.1
ExtremSCF(Z9)	3.32	0.21	0.15	0.02	0.67	0.04
ExtremSCF(Z10)	2.35	0.16	0.09	0.01	0.41	0.03

C2. Statistical analysis of the differences between the vibrational frequencies (ω) calculated using the SCFConvN and ExtremeSCF(ZN) settings. The calculations were performed at the HF/def2-TZVP level of theory.

	ω<200 cm^{-1}				ω>200 cm^{-1}			
	$\lvert\Delta\omega\rvert^m_{\max}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{\max}$ (cm^{-1})	$\lvert\Delta\omega\rvert^m_{\mathrm{mad}}$ (cm^{-1})	$\lvert\Delta\omega\rvert^m_{\mathrm{rms}}$ (cm^{-1})	$\lvert\Delta\omega\rvert^m_{\max}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{\max}$ (cm^{-1})	$\lvert\Delta\omega\rvert^m_{\mathrm{mad}}$ (cm^{-1})	$\lvert\Delta\omega\rvert^m_{\mathrm{rms}}$ (cm^{-1})
SCFConv6/ ExtremeSCF(Z6)	0.45	0.09	0.45	0.45	0.21	0.02	0.02	0.04
SCFConv7/ ExtremeSCF(Z7)	0.27	0.02	0.17	0.17	0.04	0.01	0.01	0.01
SCFConv8/ ExtremeSCF(Z8)	0.27	0.02	0.14	0.19	0.03	0.00	0.01	0.01
SCFConv9/ ExtremeSCF(Z9)	0.05	0.01	0.01	0.03	0.01	0.00	0.00	0.00
SCFConv10/ ExtremeSCF(Z10)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

	ω					
	$\lvert\Delta\omega\rvert^m_{\max}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{\max}$ (cm^{-1})	$\lvert\Delta\omega\rvert^m_{\mathrm{mad}}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{\mathrm{mad}}$ (cm^{-1})	$\lvert\Delta\omega\rvert^m_{\mathrm{rms}}$ (cm^{-1})	$\overline{\lvert\Delta\omega\rvert}_{\mathrm{rms}}$ (cm^{-1})
SCFConv6/ ExtremeSCF(Z6)	0.45	0.07	0.04	0.01	0.08	0.01
SCFConv7/ ExtremeSCF(Z7)	0.27	0.02	0.01	0.00	0.03	0.00
SCFConv8/ ExtremeSCF(Z8)	0.27	0.01	0.01	0.00	0.06	0.00
SCFConv9/ ExtremeSCF(Z9)	0.05	0.00	0.00	0.00	0.01	0.00
SCFConv10/ ExtremeSCF(Z10)	0.00	0.00	0.00	0.00	0.00	0.00

D. In this appendix the derivatives of the Becke grid weights [A. D. Becke: JCP, 88, 2547, 1988] are obtained. Note that a similar derivation was given earlier by [B. G. Johnson, P. M. V. Gill and J. A. Pople: JCP, 98, 5612, 1993]. Here we provide a more detailed derivation reflecting our implementation and obtain the weight derivatives in terms of the underived ones. Contributions from the nuclei in the grid weight function are also separated from those that arise from grid points via their parent atoms.
Within Becke's integration scheme, the molecular grid weights are defined as

$$w_g = \tilde{w}_g w_{A(g)}$$

where $\tilde{w}_g$ is independent of nuclear positions, whereas A(g) indicates the parent atom of gridpoint g. If one does not wish to indicate whether atom A is the parent of g the notation w_{Ag} is used (without brackets). Then, w_{Ag} is defined as

$$w_{Ag} = \frac{P_{Ag}}{\sum_B P_{Bg}}$$

and

$$P_{Ag} = \prod_{B \neq A} s(\mu_{AB}^g) = \prod_{B \neq A} \frac{1}{2}\left(1 - F(\mu_{AB}^g)\right)$$

where

$$F(\mu_{AB}) = f(f(f(g(\mu_{AB}^g))))$$

where *f* and *g* are simple polinomial functions

$$f(x) = \frac{1}{2}(3x - x^3)$$

$$g(x) = x + a_{AB}(1 - x^2)$$

where a_{AB} has the property a_{AB}=-a_{BA}. These functions are defined in terms of coordinates

$$\mu_{AB}^g = \frac{r_{Ag} - r_{Bg}}{R_{AB}}, \qquad A \neq B.$$

Note the following properties

$$\mu_{AB}^g = -\mu_{BA}^g$$

$$F(\mu_{AB}^g) = F_{AB}^g = -F_{BA}^g$$

$$s(\mu_{AB}^g) = s_{AB}^g = 1 - s_{BA}^g$$

The first derivative of $w_{A(g)}$ is then obtained as

$$w_{A(g)}^{X_N} = \frac{P_{A(g)}^{X_N} - w_{A(g)} \sum_B P_{Bg}^{X_N}}{\sum_B P_{Bg}} \qquad (140)$$

where

$$P_{Ag}^{X_N} = \sum_{C \neq A} \frac{\partial s_{AC}^g}{\partial \mu_{AC}^g} \frac{\partial \mu_{AC}^g}{\partial X_N} \prod_{B \neq C,A} s_{AB}^g .$$

This expression can be further simplified by multiplying with s_{AC}^g / s_{AC}^g, using the definition of P_{Ag} and introducing new notations for the derivatives

$$P_{Ag}^{X_N} = P_{Ag} \sum_{C \neq A} \frac{s_{AC}^{g\mu} \mu_{AC}^{gX_N}}{s_{AC}^g}$$

where

$$s_{AC}^{g\mu} = \frac{\partial s_{AC}^g}{\partial \mu_{AC}^g} = -\frac{1}{2} \frac{\partial F_{AC}^g}{\partial \mu_{AC}^g} = -\frac{1}{2} F_{AC}^{g\mu}$$

which can be calculated from the derivative of the polinomial expression defined above; and

$$\mu_{AC}^{gX_N} = \frac{\partial \mu_{AC}^g}{\partial X_N} = (\delta_{AN} - \delta_{CN}) \underbrace{\frac{1}{R_{AC}} \left(\frac{x_N - x_g}{r_{Ng}} - \frac{\mu_{AC}^g (x_A - x_C)}{R_{AC}} \right)}_{M_{AC}^{gX_N}} - \delta_{NP(g)} \underbrace{\frac{1}{R_{AC}} \left(\frac{x_A - x_g}{r_{Ag}} - \frac{x_C - x_g}{r_{Cg}} \right)}_{N_{AC}^{gX_N}}$$

where the nuclear (first term) and grid point (second term) contributions were separated. Since

$F_{AB}^{g\mu} = F_{BA}^{g\mu}$,

it follows that

$s_{AB}^{g\mu} = s_{BA}^{g\mu}$.

Furthermore,

$\mu_{AB}^{gX_N} = -\mu_{BA}^{gX_N}$,

$M_{AB}^{gX_N} = M_{BA}^{gX_N}$,

$N_{AB}^{gX_N} = -N_{BA}^{gX_N}$.

Let us now insert these quantities into (92) and use the definition of the grid weights, starting with the first term

$$\frac{P_{A(g)}^{X_N}}{\sum_B P_{Bg}} = \frac{P_{A(g)} \sum_{B\neq A} \frac{s_{AB}^{g\mu}\mu_{AB}^{gX_N}}{s_{AB}^{g}}}{\sum_B P_{Bg}} = w_{A(g)} \sum_{B\neq A} \frac{s_{AB}^{g\mu}\mu_{AB}^{gX_N}}{s_{AB}^{g}}.$$

Similarly, in the second term

$$\frac{w_{A(g)} \sum_B P_{Bg}^{X_N}}{\sum_B P_{Bg}} = w_{A(g)} \sum_B w_{Bg} \sum_{C\neq B} \frac{s_{BC}^{g\mu}\mu_{BC}^{gX_N}}{s_{BC}^{g}}.$$

Note that one may rewrite the first term as

$$w_{A(g)} \sum_B \delta_{AB} \sum_{C\neq B} \frac{s_{BC}^{g\mu}\mu_{BC}^{gX_N}}{s_{BC}^{g}},$$

and then have the following expression for the weight derivatives

$$w_{A(g)}^{X_N} = w_{A(g)} \sum_B (\delta_{AB} - w_{Bg}) \sum_{C\neq B} \frac{s_{BC}^{g\mu}\mu_{BC}^{gX_N}}{s_{BC}^{g}} =$$

$$w_{A(g)} \sum_B (\delta_{AB} - w_{Bg}) \sum_{C\neq B} \frac{s_{BC}^{g\mu} M_{BC}^{gX_N}}{s_{BC}^{g}} (\delta_{BN} - \delta_{CN}) - w_{A(g)} \sum_B (\delta_{AB} - w_{Bg}) \sum_{C\neq B} \frac{s_{BC}^{g\mu} N_{BC}^{gX_N}}{s_{BC}^{g}} \delta_{AN}.$$

For the first term, evaluating δ_{BN}-δ_{CN}, one gets

$$w_{A(g)} \left((\delta_{AN} - w_{Ng}) \sum_{C\neq N} \frac{s_{NC}^{g\mu} M_{NC}^{gX_N}}{s_{NC}^{g}} - \sum_{B\neq N} (\delta_{AB} - w_{Bg}) \frac{s_{BN}^{g\mu} M_{BN}^{gX_N}}{s_{BN}^{g}} \right).$$

Replacing the running index C in the first term with B, and using the properties of the various quantities with respect to swapping atomic labels, one arrives at

$$w_{A(g)} \sum_{B\neq N} s_{NB}^{g\mu} M_{NB}^{gX_N} \left(\frac{\delta_{AN} - w_{Ng}}{s_{NB}^{g}} - \frac{\delta_{AB} - w_{Bg}}{1 - s_{NB}^{g}} \right).$$

The second term can be written as

$$w_{A(g)} \delta_{AN} \left(\sum_{B<C} (\delta_{AB} - w_{Bg}) \frac{s_{BC}^{g\mu} N_{BC}^{gX_N}}{s_{BC}^{g}} + \sum_{C<B} (\delta_{AB} - w_{Bg}) \frac{s_{BC}^{g\mu} N_{BC}^{gX_N}}{s_{BC}^{g}} \right).$$

Summation indices B and C can than be relabeled in the second term, and using the properties of various intermediates

$$w_{A(g)} \delta_{AN} \sum_{B<C} s_{BC}^{g\mu} N_{BC}^{gX_N} \left(\frac{\delta_{AB} - w_{Bg}}{s_{BC}^{g}} - \frac{\delta_{AC} - w_{Cg}}{1 - s_{BC}^{g}} \right).$$

Using these results, one can finally write the weight derivative expression as

$$w^{X_N}_{A(g)} = w_{A(g)}\left(\sum_{B\neq N} s^{g\mu}_{NB} M^{gX_N}_{NB}\left(\frac{\delta_{AN}-w_{Ng}}{s^g_{NB}} - \frac{\delta_{AB}-w_{Bg}}{1-s^g_{NB}}\right) - \delta_{AN}\sum_{B<C} s^{g\mu}_{BC} N^{gX_N}_{BC}\left(\frac{\delta_{AB}-w_{Bg}}{s^g_{BC}} - \frac{\delta_{AC}-w_{Cg}}{1-s^g_{BC}}\right)\right) =$$

$$w^{N}_{A(g)} - \delta_{AN} w^{g}_{A(g)}.$$

The last equality indicates the separation of nuclear and grid point contributions. In this formulation grid weights are reused in the evaluation of their derivatives, and the implementation only requires a minimal modification of the already existing code.

Bibliography

(1) Lehninger, A. L.; Nelson, D. L.; Cox, M. M. *Wh Freeman* **2005**, *1*.
(2) Jencks, W. P. *Dover Publications* **1987**.
(3) Lippard, S. J.; Berg, J. M. *Univ Science Books* **1994**.
(4) Umena, Y.; Kawakami, K.; Shen, J. R.; Kamiya, N. *Nature* **2011**, *473*, 55.
(5) Kendrew, J. C.; Bodo, G.; Dintzis, H. M.; Parrish, R. G.; Wyckoff, H.; Phillips, D. C. *Nature* **1958**, *181*, 662.
(6) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235.
(7) Wuthrich, K. *Accounts Chem Res* **1989**, *22*, 36.
(8) Hoffman, B. M. *Accounts Chem Res* **1991**, *24*, 164.
(9) Brennan, B. A.; Cummings, J. G.; Chase, D. B.; Turner, I. M.; Nelson, M. J. *Biochemistry* **1996**, *35*, 10068.
(10) Garton, S. D.; Hilton, J.; Oku, H.; Crouse, B. R.; Rajagopalan, K. V.; Johnson, M. K. *Journal of the American Chemical Society* **1997**, *119*, 12906.
(11) Fu, W. G.; Drozdzewski, P. M.; Davies, M. D.; Sligar, S. G.; Johnson, M. K. *Journal of Biological Chemistry* **1992**, *267*, 15502.
(12) Neese, F. *Journal of Inorganic Biochemistry* **2003**, *96*, 38.
(13) Bhatia, S. C.; Ravi, N. *Biomacromolecules* **2003**, *4*, 723.
(14) Seyedsayamdost, M. R.; Argirevic, T.; Minnihan, E. C.; Stubbe, J.; Bennati, M. *Journal of the American Chemical Society* **2009**, *131*, 15729.
(15) Northrop, D. B.; Simpson, F. B. *Bioorgan Med Chem* **1997**, *5*, 641.
(16) Ruzicka, J.; Gubeli, T. *Anal Chem* **1991**, *63*, 1680.
(17) Steinbach, P. J.; Brooks, B. R. *P Natl Acad Sci USA* **1993**, *90*, 9135.
(18) Krogan, N. J.; Cagney, G.; Yu, H. Y.; Zhong, G. Q.; Guo, X. H.; Ignatchenko, A.; Li, J.; Pu, S. Y.; Datta, N.; Tikuisis, A. P.; Punna, T.; Peregrin-Alvarez, J. M.; Shales, M.; Zhang, X.; Davey, M.; Robinson, M. D.; Paccanaro, A.; Bray, J. E.; Sheung, A.; Beattie, B.; Richards, D. P.; Canadien, V.; Lalev, A.; Mena, F.; Wong, P.; Starostine, A.; Canete, M. M.; Vlasblom, J.; Wu, S.; Orsi, C.; Collins, S. R.; Chandran, S.; Haw, R.; Rilstone, J. J.; Gandi, K.; Thompson, N. J.; Musso, G.; St Onge, P.; Ghanny, S.; Lam, M. H. Y.; Butland, G.; Altaf-Ui, A. M.; Kanaya, S.; Shilatifard, A.; O'Shea, E.; Weissman, J. S.; Ingles, C. J.; Hughes, T. R.; Parkinson, J.; Gerstein, M.; Wodak, S. J.; Emili, A.; Greenblatt, J. F. *Nature* **2006**, *440*, 637.
(19) Claeyssens, F.; Harvey, J. N.; Manby, F. R.; Mata, R. A.; Mulholland, A. J.; Ranaghan, K. E.; Schutz, M.; Thiel, S.; Thiel, W.; Werner, H. J. *Angew Chem Int Edit* **2006**, *45*, 6856.
(20) Jensen, F. *Wiley* **2007**.
(21) Siegbahn, P. E. M.; Blomberg, M. R. A. *Chem. Rev.* **2000**, *100*, 421.
(22) Neese, F. *Journal of Biological Inorganic Chemistry* **2006**, *11*, 702.
(23) Klamt, A.; Schuurmann, G. *Perkin Trans* **1993**, 799.
(24) Ye, R. W.; Thomas, S. M. *Curr Opin Microbiol* **2001**, *4*, 307.
(25) Jetten, M. S. *Environ Microbiol* **2008**, *10*, 2903.
(26) Simon, J. *Fems Microbiology Reviews* **2002**, *26*, 285.
(27) Richardson, D. J. *Microbiology-Uk* **2000**, *146*, 551.
(28) Fritz, G.; Einsle, O.; Rudolf, M.; Schiffer, A.; Kroneck, P. M. H. *Journal of Molecular Microbiology and Biotechnology* **2005**, *10*, 223.
(29) Bamford, V. A.; Angove, H. C.; Seward, H. E.; Thomson, A. J.; Cole, J. A.; Butt, J. N.; Hemmings, A. M.; Richardson, D. J. *Biochemistry* **2002**, *41*, 2921.
(30) Einsle, O.; Stach, P.; Messerschmidt, A.; Simon, J.; Kroger, A.; Huber, R.; Kroneck, P. M. H. *Journal of Biological Chemistry* **2000**, *275*, 39608.

(31) Einsle, O.; Messerschmidt, A.; Stach, P.; Bourenkov, G. P.; Bartunik, H. D.; Huber, R.; Kroneck, P. M. H. *Nature* **1999**, *400*, 476.

(32) Cunha, C. A.; Macieira, S.; Dias, J. M.; Almeida, G.; Goncalves, L. L.; Costa, C.; Lampreia, J.; Huber, R.; Moura, J. J. G.; Moura, I.; Romao, M. J. *Journal of Biological Chemistry* **2003**, *278*, 17455.

(33) Almeida, M. G.; Macieira, S.; Goncalves, L. L.; Huber, R.; Cunha, C. A.; Romao, M. J.; Costa, C.; Lampreia, J.; Moura, J. J. G.; Moura, I. *European Journal of Biochemistry* **2003**, *270*, 3904.

(34) Rodrigues, M. L.; Oliveira, T. F.; Pereira, I. A. C.; Archer, M. *Embo Journal* **2006**, *25*, 5951.

(35) Pereira, I. A. C.; LeGall, J.; Xavier, A. V.; Teixeira, M. *Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology* **2000**, *1481*, 119.

(36) Rodrigues, M. L.; Oliveira, T.; Matias, P. M.; Martins, I. C.; Valente, F. M. A.; Pereira, I. A. C.; Archer, M. *Acta Crystallographica Section F-Structural Biology and Crystallization Communications* **2006**, *62*, 565.

(37) Clarke, T. A.; Hemmings, A. M.; Burlat, B.; Butt, J. N.; Cole, J. A.; Richardson, D. J. *Biochemical Society Transactions* **2006**, *34*, 143.

(38) Burlat, B.; Gwyer, J. D.; Poock, S.; Clarke, T.; Cole, J. A.; Hemmings, A. M.; Cheesman, M. R.; Butt, J. N.; Richardson, D. J. *Biochemical Society Transactions* **2005**, *33*, 137.

(39) Polyakov, K. M.; Boyko, K. M.; Tikhonova, T. V.; Slutsky, A.; Antipov, A. N.; Zvyagilskaya, R. A.; Popov, A. N.; Bourenkov, G. P.; Lamzin, V. S.; Popov, V. O. *Journal of Molecular Biology* **2009**, *389*, 846.

(40) Einsle, O.; Messerschmidt, A.; Huber, R.; Kroneck, P. M. H.; Neese, F. *Journal of the American Chemical Society* **2002**, *124*, 11737.

(41) Rudolf, M.; Einsle, O.; Neese, F.; Kroneck, P. M. H. *Biochemical Society Transactions* **2002**, *30*, 649.

(42) Enemark, J. H.; Feltham, R. D. *Coord. Chem. Rev.* **1974**, *13*, 339.

(43) Roncaroli, F.; Videla, M.; Slep, L. D.; Olabe, J. A. *Coord. Chem. Rev.* **2007**, *251*, 1903.

(44) Wasser, I. M.; de Vries, S.; Moenne-Loccoz, P.; Schroder, I.; Karlin, K. D. *Chem. Rev.* **2002**, *102*, 1201.

(45) Averill, B. A. *Chem. Rev.* **1996**, *96*, 2951.

(46) Almeida, M. G.; Silveira, C. M.; Guigliarelli, B.; Bertrand, P.; Moura, J. J. G.; Moura, I.; Leger, C. *Febs Lett* **2007**, *581*, 284.

(47) Gwyer, J. D.; Richardson, D. J.; Butt, J. N. *Journal of the American Chemical Society* **2005**, *127*, 14964.

(48) Marritt, S. J.; Kemp, G. L.; Xiaoe, L.; Durrant, J. R.; Cheesman, M. R.; Butt, J. N. *Journal of the American Chemical Society* **2008**, *130*, 8588.

(49) Gwyer, J. D.; Angove, H. C.; Richardson, D. J.; Butt, J. N. *Bioelectrochemistry* **2004**, *63*, 43.

(50) Mowat, C. G.; Chapman, S. K. *Dalton T* **2005**, 3381.

(51) Simon, J.; Gross, R.; Einsle, O.; Kroneck, P. M. H.; Kroger, A.; Klimmek, O. *Molecular Microbiology* **2000**, *35*, 686.

(52) Stach, P.; Einsle, O.; Schumacher, W.; Kurun, E.; Kroneck, P. M. H. *Journal of Inorganic Biochemistry* **2000**, *79*, 381.

(53) Shinobu, A.; Agmon, N. *J. Phys. Chem. A* **2009**, *113*, 7253.

(54) Goodrich, L. E.; Paulat, F.; Praneeth, V. K. K.; Lehnert, N. *Inorg. Chem.* **2010**, *49*, 6293.

(55) van Wonderen, J. H.; Burlat, B.; Richardson, D. J.; Cheesman, M. R.; Butt, J. N. *Journal of Biological Chemistry* **2008**, *283*, 9587.

(56) Wyllie, G. R. A.; Scheidt, W. R. *Chem. Rev.* **2002**, *102*, 1067.

(57) Fernandez, M. L.; Estrin, D. A.; Bari, S. E. *Journal of Inorganic Biochemistry* **2008**, *102*, 1523.

(58) Yamanaka, T.; Okunuki, K. *Biochimica et Biophysica Acta (BBA) - Specialized Section on Enzymological Subjects* **1963**, *67*, 394.

(59) Horio, T.; Higashi, T.; Sasagawa, M.; Kusai, K.; Nakai, M.; Okunuki, K. *Biochem J* **1960**, *77*, 194.

(60) Zumft, W. G. *Microbiol Mol Biol R* **1997**, *61*, 533.

(61) Allen, J. W. A.; Watmough, N. J.; Ferguson, S. J. *Nat Struct Biol* **2000**, *7*, 885.

(62) Vijgenboom, E.; Busch, J. E.; Canters, G. W. *Microbiol-Sgm* **1997**, *143*, 2853.

(63) Nurizzo, D.; Cutruzzola, F.; Arese, M.; Bourgeois, D.; Brunori, M.; Cambillau, C.; Tegoni, M. *Biochemistry* **1998**, *37*, 13987.

(64) Silvestrini, M. C.; Tordi, M. G.; Musci, G.; Brunori, M. *Journal of Biological Chemistry* **1990**, *265*, 11783.

(65) Rinaldo, S.; Arcovito, A.; Brunori, M.; Cutruzzola, F. *Journal of Biological Chemistry* **2007**, *282*, 14761.

(66) Moore, E. G.; Gibson, Q. H. *Journal of Biological Chemistry* **1976**, *251*, 2788.

(67) Sarti, P.; Giuffre, A.; Forte, E.; Mastronicola, D.; Barone, M. C.; Brunori, M. *Biochem Bioph Res Co* **2000**, *274*, 183.

(68) Borisov, V. B.; Forte, E.; Sarti, P.; Brunori, M.; Konstantinov, A. A.; Giuffre, A. *Biochem Bioph Res Co* **2007**, *355*, 97.

(69) Rinaldo, S.; Brunori, M.; Cutruzzola, F. *Plant Signal Behav* **2008**, *3*, 135.

(70) Nurizzo, D.; Silvestrini, M. C.; Mathieu, M.; Cutruzzola, F.; Bourgeois, D.; Fulop, V.; Hajdu, J.; Brunori, M.; Tegoni, M.; Cambillau, C. *Structure* **1997**, *5*, 1157.

(71) Cutruzzola, F.; Brown, K.; Wilson, E. K.; Bellelli, A.; Arese, M.; Tegoni, M.; Cambillau, C.; Brunori, M. *P Natl Acad Sci USA* **2001**, *98*, 2232.

(72) Cutruzzola, F.; Arese, M.; Grasso, S.; Bellelli, A.; Brunori, M. *Febs Lett* **1997**, *412*, 365.

(73) Brown, K.; Roig-Zamboni, V.; Cutruzzola, F.; Arese, M.; Sun, W.; Brunori, M.; Cambillau, C.; Tegoni, M. *Journal of Molecular Biology* **2001**, *312*, 541.

(74) Kababya, S.; Nelson, J.; Calle, C.; Neese, F.; Goldfarb, D. *Journal of the American Chemical Society* **2006**, *128*, 2017.

(75) Baute, D.; Arieli, D.; Neese, F.; Zimmermann, H.; Weckhuysen, B. M.; Goldfarb, D. *Journal of the American Chemical Society* **2004**, *126*, 11733.

(76) Carmieli, R.; Larsen, T. M.; Reed, G. H.; Zein, S.; Neese, F.; Goldfarb, D. *Journal of the American Chemical Society* **2007**, *129*, 4240.

(77) Radoul, M.; Sundararajan, M.; Potapov, A.; Riplinger, C.; Neese, F.; Goldfarb, D. *Phys Chem Chem Phys* **2010**, *12*, 7276.

(78) Volbeda, A.; Charon, M. H.; Piras, C.; Hatchikian, E. C.; Frey, M.; Fontecillacamps, J. C. *Nature* **1995**, *373*, 580.

(79) Beinert, H. *Journal of Biological Inorganic Chemistry* **2000**, *5*, 2.

(80) Pandelia, M. E.; Nitschke, W.; Infossi, P.; Giudici-Orticoni, M. T.; Bill, E.; Lubitz, W. *P Natl Acad Sci USA* **2011**, *108*, 6097.

(81) Shomura, Y.; Yoon, K. S.; Nishihara, H.; Higuchi, Y. *Nature* **2011**, *479*, 253.

(82) Fritsch, J.; Scheerer, P.; Frielingsdorf, S.; Kroschinsky, S.; Friedrich, B.; Lenz, O.; Spahn, C. M. T. *Nature* **2011**, *479*, 249.

(83) Lukey, M. J.; Roessler, M. M.; Parkin, A.; Evans, R. M.; Davies, R. A.; Lenz, O.; Friedrich, B.; Sargent, F.; Armstrong, F. A. *Journal of the American Chemical Society* **2011**, *133*, 16881.

(84) Volbeda, A.; Amara, P.; Darnault, C.; Mouesca, J. M.; Parkin, A.; Roessler, M. M.; Armstrong, F. A.; Fontecilla-Camps, J. C. *P Natl Acad Sci USA* **2012**, *109*, 5305.

(85) Goris, T.; Wait, A. F.; Saggu, M.; Fritsch, J.; Heidary, N.; Stein, M.; Zebger, I.; Lendzian, F.; Armstrong, F. A.; Friedrich, B.; Lenz, O. *Nat Chem Biol* **2011**, *7*, 648.

(86) Pandelia, M. E.; Lubitz, W.; Nitschke, W. *Bba-Bioenergetics* **2012**, *1817*, 1565.
(87) Dey, A.; Francis, E. J.; Adams, M. W. W.; Babini, E.; Takahashi, Y.; Fukuyama, K.; Hodgson, K. O.; Hedman, B.; Solomon, E. I. *Science* **2007**, *318*, 1464.
(88) Niu, S. Q.; Ichiye, T. *Abstr Pap Am Chem S* **2009**, *238*.
(89) Heering, H. A.; Bulsink, Y. B. M.; Hagen, W. R.; Meyer, T. E. *European Journal of Biochemistry* **1995**, *232*, 811.
(90) Solomon, E. I.; Xie, X. J.; Dey, A. *Chem Soc Rev* **2008**, *37*, 623.
(91) Noodleman, L.; Peng, C. Y.; Case, D. A.; Mouesca, J. M. *Coord. Chem. Rev.* **1995**, *144*, 199.
(92) Beinert, H.; Holm, R. H.; Munck, E. *Science* **1997**, *277*, 653.
(93) Schaefer, H. F.; Yamaguchi, Y. *Theochem-J Mol Struc* **1986**, *28*, 369.
(94) Pople, J. A.; Krishnan, R.; Schlegel, H. B.; Binkley, J. S. *Int J Quantum Chem* **1979**, *16*, 225.
(95) Ghysels, A.; Woodcock, H. L.; Larkin, J. D.; Miller, B. T.; Shao, Y. H.; Kong, J.; Van Neck, D.; Van Speybroeck, V.; Waroquier, M.; Brooks, B. R. *J Chem Theory Comput* **2011**, *7*, 496.
(96) Lipparini, F.; Cappelli, C.; Scalmani, G.; De Mitri, N.; Barone, V. *J Chem Theory Comput* **2012**, *8*, 4270.
(97) Lotrich, V.; Flocke, N.; Ponton, M.; Yau, A. D.; Perera, A.; Deumens, E.; Bartlett, R. J. *J Chem Phys* **2008**, *128*.
(98) Deglmann, P.; Furche, F.; Ahlrichs, R. *Chem Phys Lett* **2002**, *362*, 511.
(99) Wolff, S. K. *Int J Quantum Chem* **2005**, *104*, 645.
(100) Johnson, B. G.; Frisch, M. J. *J Chem Phys* **1994**, *100*, 7429.
(101) Almlof, J.; Faegri, K.; Korsell, K. *J. Comput. Chem.* **1982**, *3*, 385.
(102) Deglmann, P.; May, K.; Furche, F.; Ahlrichs, R. *Chem Phys Lett* **2004**, *384*, 103.
(103) Baerends, E. J.; Ellis, D. E.; Ros, P. *Chem Phys* **1973**, *2*, 41.
(104) Mintmire, J. W.; Dunlap, B. I. *Phys Rev A* **1982**, *25*, 88.
(105) Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. *J Chem Phys* **1979**, *71*, 3396.
(106) Vahtras, O.; Almlof, J.; Feyereisen, M. W. *Chem Phys Lett* **1993**, *213*, 514.
(107) Eichkorn, K.; Treutler, O.; Ohm, H.; Haser, M.; Ahlrichs, R. *Chem Phys Lett* **1995**, *240*, 283.
(108) Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. *Theor. Chem. Acc.* **1997**, *97*, 119.
(109) Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. *Chem Phys* **2009**, *356*, 98.
(110) Friesner, R. A. *Chem Phys Lett* **1985**, *116*, 39.
(111) Izsak, R.; Neese, F. *J Chem Phys* **2011**, *135*.
(112) Kossmann, S.; Neese, F. *J Chem Theory Comput* **2010**, *6*, 2325.
(113) Petrenko, T.; Kossmann, S.; Neese, F. *J Chem Phys* **2011**, *134*.
(114) Izsak, R.; Hansen, A.; Neese, F. *Mol Phys* **2012**, *110*, 2413.
(115) Izsak, R.; Neese, F. *Mol Phys* **2013**, *Accepted for publication*.
(116) http://chem.arizona.edu/rss/epr/software.html.
(117) Neese, F. *Wiley Interdisciplinary Reviews-Computational Molecular Science* **2012**, *2*, 73.
(118) Barone, V. *in Recent Advances In Density Functional Methods* **1995**, *Part I*, 287.
(119) Schafer, A.; Horn, H.; Ahlrichs, R. *J Chem Phys* **1992**, *97*, 2571.
(120) Neese, F. *Inorg Chim Acta* **2002**, *337*, 181.
(121) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys Rev Lett* **1996**, *77*, 3865.
(122) Neese, F. *J Chem Phys* **2001**, *115*, 11080.
(123) Neese, F. *J Chem Phys* **2003**, *118*, 3939.
(124) Neese, F. *The Journal of Chemical Physics* **2005**, *122*, 034107.
(125) Tokman, M.; Sundholm, D.; Pyykko, P.; Olsen, J. *Chem Phys Lett* **1997**, *265*, 60.
(126) Reece, S. Y.; Nocera, D. G. *Annu Rev Biochem* **2009**, *78*, 673.

(127) Mayer, J. M. *Annu. Rev. Phys. Chem.* **2004**, *55*, 363.
(128) Jenson, D. L.; Barry, B. A. *Journal of the American Chemical Society* **2009**, *131*, 10567.
(129) Grove, L. E.; Xie, J.; Yikilmaz, E.; Miller, A. F.; Brunold, T. C. *Inorg. Chem.* **2008**, *47*, 3978.
(130) Mahanta, S.; Singh, R. B.; Kar, S.; Guchhait, N. *Chem Phys* **2008**, *354*, 118.
(131) Yuasa, J.; Fukuzumi, S. *Journal of the American Chemical Society* **2006**, *128*, 14281.
(132) Grosserueschkamp, M.; Friedrich, M. G.; Plum, M.; Knoll, W.; Naumann, R. L. C. *J Phys Chem B* **2009**, *113*, 2492.
(133) Setif, P.; Hirasawa, M.; Cassan, N.; Lagoutte, B.; Tripathy, J. N.; Knaff, D. B. *Biochemistry* **2009**, *48*, 2828.
(134) Hammes-Schiffer, S.; Hatcher, E.; Ishikita, H.; Skone, J. H.; Soudackov, A. V. *Coord. Chem. Rev.* **2008**, *252*, 384.
(135) Blomberg, M. R. A.; Siegbahn, P. E. M. *Bba-Bioenergetics* **2006**, *1757*, 969.
(136) Kaukonen, M.; Soderhjelm, P.; Heimdal, J.; Ryde, U. *J Chem Theory Comput* **2008**, *4*, 985.
(137) Marcus, R. A. *Chem. Rev.* **1988**, *88*, 1475.
(138) Marcus, R. A.; Sutin, N. *Biochim Biophys Acta* **1985**, *811*, 265.
(139) Hammes-Schiffer, S.; Soudackov, A. V. *J Phys Chem B* **2008**, *112*, 14108.
(140) Cukier, R. I. *J Phys Chem B* **2002**, *106*, 1746.
(141) Soudackov, A.; Hammes-Schiffer, S. *J Chem Phys* **2000**, *113*, 2385.
(142) Kaila, V. R. I.; Johansson, M. P.; Sundholm, D.; Wikstrom, M. *P Natl Acad Sci USA* **2010**, *107*, 21470.
(143) Rosso, K. M.; Rustad, J. R. *J. Phys. Chem. A* **2000**, *104*, 6718.
(144) Hatcher, E.; Soudackov, A. V.; Hammes-Schiffer, S. *Journal of the American Chemical Society* **2004**, *126*, 5763.
(145) Petrenko, T.; Sturhahn, W.; Neese, F. *Hyperfine Interact* **2007**, *175*, 165.
(146) D, W. *Available free online at* http://bililite.com/tenua

2011.
(147) Barshop, B. A.; Wrenn, R. F.; Frieden, C. *Anal Biochem* **1983**, *130*, 134.
(148) Wang, T.; Brudvig, G.; Batista, V. S. *J Chem Theory Comput* **2010**, *6*, 755.
(149) Tissandier, M. D.; Cowen, K. A.; Feng, W. Y.; Gundlach, E.; Cohen, M. H.; Earhart, A. D.; Coe, J. V.; Tuttle, T. R. *J. Phys. Chem. A* **1998**, *102*, 7787.
(150) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J Phys Chem B* **2006**, *110*, 16066.
(151) Pace, C. N.; Grimsley, G. R.; Scholtz, J. M. *Journal of Biological Chemistry* **2009**, *284*, 13285.
(152) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B.* **1988**, *37*, 785.
(153) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
(154) Johnson, B. G.; Frisch, M. J. *Chem Phys Lett* **1993**, *216*, 133.
(155) Pople, J. A.; Gill, P. M. W.; Johnson, B. G. *Chem Phys Lett* **1992**, *199*, 557.
(156) Handy, N. C.; Tozer, D. J.; Laming, G. J.; Murray, C. W.; Amos, R. D. *Israel J Chem* **1993**, *33*, 331.
(157) Baker, J.; Andzelm, J.; Scheiner, A.; Delley, B. *J Chem Phys* **1994**, *101*, 8894.
(158) Satoko, C. *Chem Phys Lett* **1981**, *83*, 111.
(159) Becke, A. D. *Phys. Rev. A.* **1988**, *38*, 3098.
(160) Lukat, P.; Rudolf, M.; Stach, P.; Messerschmidt, A.; Kroneck, P. M. H.; Simon, J.; Einsle, O. *Biochemistry* **2008**, *47*, 2080.
(161) Clarke, T. A.; Kemp, G. L.; Van Wonderen, J. H.; Doyle, R.-M. A. S.; Cole, J. A.; Tovell, N.; Cheesman, M. R.; Butt, J. N.; Richardson, D. J.; Hemmings, A. M. *Biochemistry* **2008**, *47*, 3789.
(162) Perdew, J. P. *Phys Rev B* **1986**, *33*, 8822.

(163) Becke, A. D. *J Chem Phys* **1986**, *84*, 4524.
(164) Neese, F. *J. Comput. Chem.* **2003**, *24*, 1740.
(165) Schafer, A.; Huber, C.; Ahlrichs, R. *J Chem Phys* **1994**, *100*, 5829.
(166) Reiher, M.; Salomon, O.; Hess, B. A. *Theor Chem Acc* **2001**, *107*, 48.
(167) Ahlrichs, R.; Co-workers, a. *ftp.chemie.uni-karlsruhe.de/pub/basen (Universität Karlsruhe, Karlsruhe, Germany)* **2001**.
(168) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463.
(169) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787.
(170) Neese, F. *Journal of the American Chemical Society* **2006**, *128*, 10213.
(171) Pipek, J.; Mezey, P. G. *J Chem Phys* **1989**, *90*, 4916.
(172) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *25*, 1605.
(173) Rydberg, P.; Sigfridsson, E.; Ryde, U. *Journal of Biological Inorganic Chemistry* **2004**, *9*, 203.
(174) Goldfarb, D.; Bernardo, M.; Thomann, H.; Kroneck, P. M. H.; Ullrich, V. *Journal of the American Chemical Society* **1996**, *118*, 2686.
(175) Harvey, J. N. *Principles and Applications of Density Functional Theory in Inorganic Chemistry I* **2004**, *112*, 151.
(176) Salomon, O.; Reiher, M.; Hess, B. A. *The Journal of Chemical Physics* **2002**, *117*, 4729.
(177) Xantheas, S. S. **2005**, *116*, 119.
(178) Dunitz, J. D. *Science* **1994**, *264*.
(179) Martins, G.; Rodrigues, L.; Cunha, F. M.; Matos, D.; Hildebrandt, P.; Murgida, D. H.; Pereira, I. A. C.; Todorovic, S. *J Phys Chem B* **2010**, *114*, 5563.
(180) Yi, J.; Heinecke, J.; Tan, H.; Ford, P. C.; Richter-Addo, G. B. *Journal of the American Chemical Society* **2009**, *131*, 18119.
(181) Nasri, H.; Ellison, M. K.; Shang, M. Y.; Schulz, C. E.; Scheidt, W. R. *Inorg. Chem.* **2004**, *43*, 2932.
(182) Nasri, H.; Ellison, M. K.; Krebs, C.; Huynh, B. H.; Scheidt, W. R. *Journal of the American Chemical Society* **2000**, *122*, 10795.
(183) Nasri, H.; Wang, Y. N.; Huynh, B. H.; Walker, F. A.; Scheidt, W. R. *Inorg. Chem.* **1991**, *30*, 1483.
(184) Li, H.; Robertson, A. D.; Jensen, J. H. *Proteins-Structure Function and Bioinformatics* **2005**, *61*, 704.
(185) Bas, D. C.; Rogers, D. M.; Jensen, J. H. *Proteins-Structure Function and Bioinformatics* **2008**, *73*, 765.
(186) Olsson, M. H. M.; Sondergaard, C. R.; Rostkowski, M.; Jensen, J. H. *J Chem Theory Comput* **2011**, *7*, 525.
(187) Ghosh, A. *Accounts Chem Res* **2005**, *38*, 943.
(188) Patchkovskii, S.; Ziegler, T. *Inorg. Chem.* **2000**, *39*, 5354.
(189) Sundararajan, M.; Neese, F. *J Chem Theory Comput* **2012**, *8*, 563.
(190) Serres, R. G.; Grapperhaus, C. A.; Bothe, E.; Bill, E.; Weyhermuller, T.; Neese, F.; Wieghardt, K. *Journal of the American Chemical Society* **2004**, *126*, 5138.
(191) Praneeth, V. K. K.; Nather, C.; Peters, G.; Lehnert, N. *Inorg. Chem.* **2006**, *45*, 2795.
(192) Radon, M.; Pierloot, K. *J. Phys. Chem. A* **2008**, *112*, 11824.
(193) Lehnert, N.; Praneeth, V. K. K.; Paulat, F. *J. Comput. Chem.* **2006**, *27*, 1338.
(194) Xu, C. L.; Spiro, T. G. *Journal of Biological Inorganic Chemistry* **2008**, *13*, 613.
(195) Zoppellaro, G.; Bren, K. L.; Ensign, A. A.; Harbitz, E.; Kaur, R.; Hersleth, H. P.; Ryde, U.; Hederstedt, L.; Andersson, K. K. *Biopolymers* **2009**, *91*, 1064.
(196) Smith, D. M. A.; Dupuis, M.; Vorpagel, E. R.; Straatsma, T. P. *Journal of the American Chemical Society* **2003**, *125*, 2711.
(197) Smith, D. M. A.; Dupuis, M.; Straatsma, T. P. *Mol Phys* **2005**, *103*, 273.

(198) Olsson, M. H.; Siegbahn, P. E.; Blomberg, M. R.; Warshel, A. *Biochim Biophys Acta* **2007**, *1767*, 244.

(199) Cowley, A. B.; Kennedy, M. L.; Silchenko, S.; Lukat-Rodgers, G. S.; Rodgers, K. R.; Benson, D. R. *Inorg. Chem.* **2006**, *45*, 9985.

(200) Bykov, D.; Neese, F. *Journal of Biological Inorganic Chemistry* **2011**, *16*, 417.

(201) Dixon, D. W.; Hong, X.; Woehler, S. E.; Mauk, A. G.; Sishta, B. P. *Journal of the American Chemical Society* **1990**, *112*, 1082.

(202) Ling, Y.; Mills, C.; Weber, R.; Yang, L.; Zhang, Y. *Journal of the American Chemical Society* **2010**, *132*, 1583.

(203) Miranda, K. M. *Coord. Chem. Rev.* **2005**, *249*, 433.

(204) Farmer, P. J.; Sulc, F. *Journal of Inorganic Biochemistry* **2005**, *99*, 166.

(205) Yang, L.; Ling, Y.; Zhang, Y. *Journal of the American Chemical Society* **2011**, *133*, 13814.

(206) Bykov, D.; Neese, F. *Journal of Biological Inorganic Chemistry* **2012**, *17*, 741.

(207) Liu, Y. M.; Ryan, M. D. *J Electroanal Chem* **1994**, *368*, 209.

(208) Riplinger, C.; Neese, F. *ChemPhysChem* **2011**, *12*, 3192.

(209) Angove, H. C.; Cole, J. A.; Richardson, D. J.; Butt, J. N. *Journal of Biological Chemistry* **2002**, *277*, 23374.

(210) Hirasawa, M.; Tripathy, J. N.; Sommer, F.; Somasundaram, R.; Chung, J. S.; Nestander, M.; Kruthiventi, M.; Zabet-Moghaddam, M.; Johnson, M. K.; Merchant, S. S.; Allen, J. P.; Knaff, D. B. *Photosynth Res* **2010**, *103*, 67.

(211) Kuznetsova, S.; Knaff, D. B.; Hirasawa, M.; Setif, P.; Mattioli, T. A. *Biochemistry* **2004**, *43*, 10765.

(212) Roberts, S. A.; Weichsel, A.; Qiu, Y.; Shelnutt, J. A.; Walker, F. A.; Montfort, W. R. *Biochemistry* **2001**, *40*, 11327.

(213) Purwar, N.; McGarry, J. M.; Kostera, J.; Pacheco, A. A.; Schmidt, M. *Biochemistry* **2011**, *50*, 4491.

(214) Kim, Y. O.; Goff, H. M. *Inorg. Chem.* **1990**, *29*, 3907.

(215) Parr, S. R.; Barber, D.; Greenwood, C.; Brunori, M. *Biochem J* **1977**, *167*, 447.

(216) Wilson, E. K.; Bellelli, A.; Liberti, S.; Arese, M.; Grasso, S.; Cutruzzola, F.; Brunori, M.; Brzezinski, P. *Biochemistry* **1999**, *38*, 7556.

(217) Silvestrini, M. C.; Cutruzzola, F.; Dalessandro, R.; Brunori, M.; Fochesato, N.; Zennaro, E. *Biochem J* **1992**, *285*, 661.

(218) Silvestrini, M. C.; Colosimo, A.; Brunori, M.; Walsh, T. A.; Barber, D.; Greenwood, C. *Biochem J* **1979**, *183*, 701.

(219) Nurizzo, D.; Cutruzzola, F.; Arese, M.; Bourgeois, D.; Brunori, M.; Cambillau, C.; Tegoni, M. *Journal of Biological Chemistry* **1999**, *274*, 14997.

(220) Goldfarb, D.; Lipkin, Y.; Potapov, A.; Gorodetsky, Y.; Epel, B.; Raitsimring, A. M.; Radoul, M.; Kaminker, I. *J Magn Reson* **2008**, *194*, 8.

(221) Mims, W. B. *Proc R Soc Lon Ser-A* **1965**, *283*, 452.

(222) Davies, E. R. *Phys Lett A* **1974**, *A 47*, 1.

(223) Epel, B.; Arieli, D.; Baute, D.; Goldfarb, D. *J Magn Reson* **2003**, *164*, 78.

(224) Stoll, S.; Schweiger, A. *J Magn Reson* **2006**, *178*, 42.

(225) Sinnecker, S.; Rajendran, A.; Klamt, A.; Diedenhofen, M.; Neese, F. *J. Phys. Chem. A* **2006**, *110*, 2235.

(226) Muhoberac, B. B.; Wharton, D. C. *Journal of Biological Chemistry* **1980**, *255*, 8437.

(227) Muhoberac, B. B.; Wharton, D. C. *Journal of Biological Chemistry* **1983**, *258*, 3019.

(228) Cheesman, M. R.; Ferguson, S. J.; Moir, J. W. B.; Richardson, D. J.; Zumft, W. G.; Thomson, A. J. *Biochemistry* **1997**, *36*, 16267.

(229) Walsh, T. A.; Johnson, M. K.; Greenwood, C.; Barber, D.; Springall, J. P.; Thomson, A. J. *Biochem J* **1979**, *177*, 29.

(230) Lobrutto, R.; Wei, Y. H.; Mascarenhas, R.; Scholes, C. P.; King, T. E. *Journal of Biological Chemistry* **1983**, *258*, 7437.

(231) Radoul, M.; Centola, F.; Rinaldo, S.; Cutruzzola, F.; Pecht, I.; Goldfarb, D. *Inorg. Chem.* **2009**, *48*, 3913.

(232) Praneeth, V. K. K.; Neese, F.; Lehnert, N. *Inorg. Chem.* **2005**, *44*, 2570.

(233) Blomberg, L. M.; Blomberg, M. R. A.; Siegbahn, P. E. M. *Journal of Inorganic Biochemistry* **2005**, *99*, 949.

(234) Olah, J.; Harvey, J. N. *J. Phys. Chem. A* **2009**, *113*, 7338.

(235) Franke, A.; Stochel, G.; Jung, C.; Van Eldik, R. *Journal of the American Chemical Society* **2004**, *126*, 4181.

(236) Neese, F. *Magn Reson Chem* **2004**, *42*, S187.

(237) Lippard, S. J., Berg, J. M. *University Science Book, Mill Valley, USA* **1994**.

(238) Yoon, K. S.; Fukuda, K.; Fujisawa, K.; Nishihara, H. *Int J Hydrogen Energ* **2011**, *36*, 7081.

(239) Moulis, J. M.; Lutz, M.; Gaillard, J.; Noodleman, L. *Biochemistry* **1988**, *27*, 8712.

(240) Walters, E. M.; Garcia-Serres, R.; Naik, S. G.; Bourquin, F.; Glauser, D. A.; Schurmann, P.; Huynh, B. H.; Johnson, M. K. *Biochemistry* **2009**, *48*, 1016.

(241) Mouesca, J. M.; Lamotte, B. *Coord. Chem. Rev.* **1998**, *178*, 1573.

(242) Hamann, N.; Bill, E.; Shokes, J. E.; Scott, R. A.; Bennati, M.; Hedderich, R. *Journal of Biological Inorganic Chemistry* **2009**, *14*, 457.

(243) Middleton, P.; Dickson, D. P. E.; Johnson, C. E.; Rush, J. D. *European Journal of Biochemistry* **1980**, *104*, 289.

(244) Teixeira, M.; Moura, I.; Xavier, A. V.; Moura, J. J. G.; Legall, J.; Dervartanian, D. V.; Peck, H. D.; Huynh, B. H. *Journal of Biological Chemistry* **1989**, *264*, 16435.

(245) Huynh, B. H.; Patil, D. S.; Moura, I.; Teixeira, M.; Moura, J. J. G.; Dervartanian, D. V.; Czechowski, M. H.; Prickril, B. C.; Peck, H. D.; Legall, J. *Journal of Biological Chemistry* **1987**, *262*, 795.

(246) Surerus, K. K.; Chen, M.; Vanderzwaan, J. W.; Rusnak, F. M.; Kolk, M.; Duin, E. C.; Albracht, S. P. J.; Munck, E. *Biochemistry* **1994**, *33*, 4980.

(247) Schunemann, V.; Winkler, H. *Rep Prog Phys* **2000**, *63*, 263.

(248) Mouesca, J. M.; Noodleman, L.; Case, D. A.; Lamotte, B. *Inorg. Chem.* **1995**, *34*, 4347.

(249) Cammack, R.; Patil, D. S.; Fernandez, V. M. *Biochemical Society Transactions* **1985**, *13*, 572.

(250) Kent, T. A.; Huynh, B. H.; Munck, E. *P Natl Acad Sci-Biol* **1980**, *77*, 6574.

(251) Krebs, C.; Broderick, W. E.; Henshaw, T. F.; Broderick, J. B.; Huynh, B. H. *Journal of the American Chemical Society* **2002**, *124*, 912.

(252) Ciurli, S.; Carrie, M.; Weigel, J. A.; Carney, M. J.; Stack, T. D. P.; Papaefthymiou, G. C.; Holm, R. H. *Journal of the American Chemical Society* **1990**, *112*, 2654.

(253) Seemann, M.; Janthawornpong, K.; Schweizer, J.; Bottger, L. H.; Janoschka, A.; Ahrens-Botzong, A.; Tambou, M. N.; Rotthaus, O.; Trautwein, A. X.; Rohmer, M.; Schunemann, V. *Journal of the American Chemical Society* **2009**, *131*, 13184.

(254) Belinskii, M.; Bertini, I.; Galas, O.; Luchinat, C. *Z Naturforsch A* **1995**, *50*, 75.

(255) Kurtz, D. M. *Chem. Rev.* **1990**, *90*, 585.

(256) Peters, J. W.; Stowell, M. H. B.; Soltis, S. M.; Finnegan, M. G.; Johnson, M. K.; Rees, D. C. *Biochemistry* **1997**, *36*, 1181.

(257) Huang, W. J.; Jia, J.; Cummings, J.; Nelson, M.; Schneider, G.; Lindqvist, Y. *Structure* **1997**, *5*, 691.

(258) Popescu, V. C.; Munck, E.; Fox, B. G.; Sanakis, Y.; Cummings, J. G.; Turner, I. M.; Nelson, M. J. *Biochemistry* **2001**, *40*, 7984.

(259) More, C.; Asso, M.; Roger, G.; Guigliarelli, B.; Caldeira, J.; Moura, J.; Bertrand, P. *Biochemistry* **2005**, *44*, 11628.

(260) Sazanov, L. A.; Hinchliffe, P. *Science* **2006**, *311*, 1430.
(261) Risthaus, T.; Grimme, S. *J Chem Theory Comput* **2013**, *9*, 1580.
(262) Cui, Q.; Karplus, M. *J Chem Phys* **2000**, *112*, 1133.
(263) Cui, Q.; Musaev, D. G.; Svensson, M.; Morokuma, K. *J Phys Chem-Us* **1996**, *100*, 10936.
(264) Breidung, J.; Thiel, W.; Komornicki, A. *Chem Phys Lett* **1988**, *153*, 76.
(265) Weigend, F.; Ahlrichs, R. *Phys Chem Chem Phys* **2005**, *7*, 3297.
(266) Weigend, F. *Phys Chem Chem Phys* **2006**, *8*, 1057.
(267) Grimme, S.; Steinmetz, M.; Korth, M. *J Org Chem* **2007**, *72*, 2118.
(268) Neese, F.; Hansen, A.; Wennmohs, F.; Grimme, S. *Accounts Chem Res* **2009**, *42*, 641.
(269) Krack, M.; Koster, A. M. *J Chem Phys* **1998**, *108*, 3226.
(270) Frisch, M. J.; Johnson, B. G.; Gill, P. M. W.; Fox, D. J.; Nobes, R. H. *Chem Phys Lett* **1993**, *206*, 225.