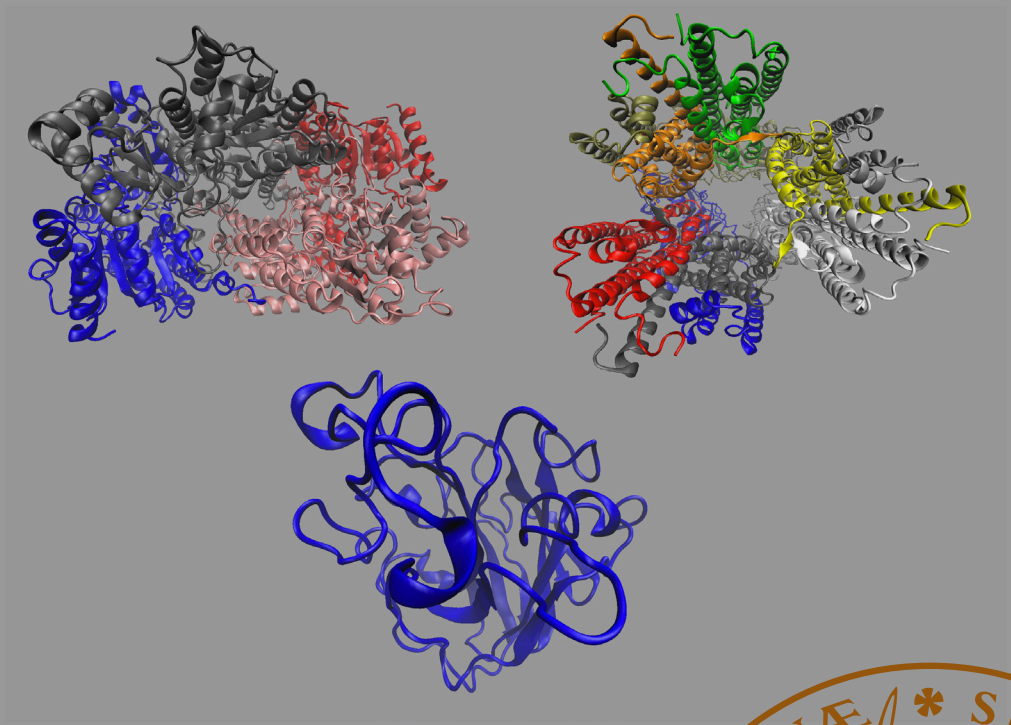
Computational Studies of Metalloenzymes

Torbjörnsson, Magne

2023

Computational Studies of Metalloenzymes

MAGNE TORBJÖRNSSON | DIVISION OF THEORETICAL CHEMISTRY | LUND UNIVERSITY



Computational Studies of Metalloenzymes

Computational Studies of Metalloenzymes

by Magne Torbjörnsson

LUND
UNIVERSITY

Doctoral Dissertation
Thesis advisors: Prof. Ulf Ryde and Ass. Prof. Erik D. Hedegård
Faculty opponent: Prof. Ran Friedman, Linnaeus University

Doctoral Dissertation
To be presented with the permission of the Faculty of Science of Lund University, for public defense at the Centre for Chemistry and Chemical Engineering, Lecture Hall A, Naturvetarvägen 14, Lund, on Friday 28th of April 2023 at 13.00

Computational Studies of Metalloenzymes

by Magne Torbjörnsson

LUND
UNIVERSITY

Cover illustration front: Magne Torbjörnsson

Division of Theoretical Chemistry, Department of Chemistry

isbn: 978-91-7422-938-7 (print)
isbn: 978-91-7422-939-4 (pdf)

Printed in Sweden by Media-Tryck, Lund University, Lund 2023

Media-Tryck is a Nordic Swan Ecolabel certified provider of printed material. Read more about our environmental work at www.mediatryck.lu.se

MADE IN SWEDEN

To my family

Contents

List of publications

This thesis is based on the following publications, referred to by their Roman numerals:

I **Comparison of the accuracy of DFT methods for reactions with relevance to nitrogenase**
M. Torbjörnsson, U. Ryde
Electron. Struct., *3*, 034005, **2021**

II **Calibration of DFT functionals for a minimal nitrogenase $[\mathbf{Fe}(\mathbf{SH})_4\mathbf{H}]^-$ model employing state-of-the-art ab initio methods**
V. P. Vysotskiy, **M. Torbjörnsson**, H. Jiang, E. D. Larsson, L. Cao, U. Ryde, H., Zhai, S. Lee, G. Chan
Manuscript

III **Histidine oxidation in lytic polysaccharide monooxygenase**
M. Torbjörnsson, M. Hagemann, U. Ryde, E. D. Hedegård
J. Biol. Inorg. Chem. DOI:10.1007/500775-023-01993-4, **2023**

IV **Computational study of three metal sites in particulate methane monooxygenase**
M. Torbjörnsson, L. Cao, E. D. Hedegård, U. Ryde
Manuscript

List of publications not included in the thesis

V **Exciton coupling induces vibronic hyperchronism in light-harvesting complexes**
J. Schulze, **M. Torbjörnsson**, O. Kühn, T. Pullerits
New J. Phys, *16*, 045010, **2014**

VI **Fast monolayer adsorption and slow energy transfer in CdSe quantum dot sentized ZnO nanowires**
K. Zheng, K. Zidek, M. Abdellah, **M. Torbjörnsson**, P. Chabera, S. Shao, F. Zhang, T. Pullerits
J. Phys. Chem. A, *117*, 5919-5925, **2013**

VII **Digital cavities and their potential applications**
J. K. Karki, **M. Torbjörnsson**, R. J. Widom, A. H. Marcus, T. Pullerits
J. Instr, *8*, T05005, **2013**

VIII **Protein configuration landscape fluctuations revealed by exciton transition polarizations in single light harvesting complexes**
S. Tubasum, **M. Torbjörnsson**, D. Yadov, R. Camacho, G. Söderlind, I. G. Scheblykin, T. Pullerits
J. Phys. Chem. B., *120*, 724-732, **2016**

Acknowledgements

First of all, I would like to thank my supervisors Ulf and Erik for supporting me through this project and for always being a nice company and very helpful. I would also thank all the senior scientists at the division, Jan, Bo, Gunnar, Mikael, Per-Åke, Esko, Petter, Marie, Martin, Magnus, Valera and Per-Olof for all your efforts running the division over the years and welcoming me to be a part of it. I also thank Maria for help with administrative matters.

Secondly, I would like to thank my coauthors Victor, Hao, Ernst, Lili, Huanchen, Seunghoon, Garnet and Marlisa. Also the members in our group and in Ulf´s weekly journal clubs, current members, Marcos, Meiting, Hao, Vilhelm, Kristoffer, Victor, Iria, Joel and Eijla as well as earlier members, Majda, Octav and Justin, deserves to be thanked for nice company and interesting discussions.

I also thank Ulf, Emma, Karin, Olle, Maria and Joakim at physical chemistry, Johan and Viveka at CAS and Sven at the faculty for great help and support. I would also like to thank Tõnu and coworkers for my time at chemical physics.

I would like to thank all the rest of the PhD students and postdocs at the divisions of theoretical chemistry and physical chemistry for all the fun things we have done during my years here.

Finally, I thank family and friends but you are fortunately too many to be mentioned individually here. However, I am glad to have all of you in my life!

Populärvetenskaplig sammanfattning på svenska

Enzymer spelar en viktig roll för många reaktioner i naturen. De är proteiner som som ökar hastigheten för kemiska reaktioner utan att själva reagera. Detta kallas för att de katalyserar reaktionerna. I reaktionerna verkar enzymerna på andra molekyler som kallas substrat. De ämnen som bildas kallas produkter. Nästan alla processer i en levande cell katalyseras av enzymer för att de ska gå tillräckligt snabbt. Bara i den mänskliga kroppen finns över 700 olika enzymer. Enzymer är också viktiga för industriella tillämpningar, till exempel i produktion av biobränslen och konstgödsel. Det är därför viktigt att förstå hur de fungerar.

För att studera enzymer används både experiment och datorsimuleringar. I den här avhandlingen fokuserar vi på datorsimuleringar. Datorerna har utvecklats mycket kraftigt de senaste årtiondena vilket har lett till att datorsimuleringar av kemiska reaktioner har fått en stor betydelse. De teoretiska grunderna är dock äldre än så.

För kroppar i den storleksnivå vi normalt sett möter i vår vardag är är vi vana vid vilka fysiska regler som gäller. Detta kallas klassisk mekanik. För mycket mindre partiklar gäller delvis andra och ibland överaksakande regler. Detta kallas för kvantmekanik.

Det är kvantmekaniken som ligger till grund för vad som idag kallas kvantkemi. Grunderna lades fram av forskare för ungefär hundra år sedan. Enligt kvantmekaniken snurrar elektronerna runt atomkärnan och enbart vissa energinivåer är tillåtna. Man säger att energin är kvantiserad. För att förstå vilken energi en elektron har löser man den så kallade Schrödingerekvationen. Den är vad matematikerna kallar för en partiell differential ekvation, det vill säga en ekvation som innehåller derivator och där lösningen är en funktion som beror av flera variabler. Dessa är tiden och alla partiklarnas koordinater i rummet. Lösningarna kallas för vågfunktioner och ger information om sannolikheten att en partikel har en viss position.

Tyvärr är Schrödingerekvationen bara möjlig att lösa analytiskt i enkla fall, vilket betyder att man måste göra förenklingar för att kunna lösa Schrödinger ekvationen för mer realistiska fall med en hjälp av en dator. Detta kallas för numeriska lösningar.

I den här avhandlingen används i huvudsak två numeriska metoder, molekylmekanik och täthetsfunktionalteori (DFT). I DFT beräknas ett systems energi som en funktional av elektrontätheten.

Vi har använt dessa metoder för att studera tre olika enzymer. Nitrogenas är det enda enzym som kan klyva den starka trippelbindningen i molekylärt kväve för att bilda ammoniak och göra atmosfäriskt kväve tillgängligt för växter. Den innehåller ett komplicerat $MoFe_7S_9C$-kluster, kallat FeMo. Tidigare undersökningar har visat att resultaten starkt beror på av vilken DFT-metod som används, så att man inte vet vil-

ken metod man kan lita på. Därför har vi använt två olika metoder för att kalibrera DFT-beräkningarna. I den första har vi använt experimentell data för strukturer och reaktioner relaterade till nitrogenas. I den andra använder vi i stället mer avancerade kvantmekaniska metoder på en minimal modell av FeMo-klustret.

Lytiskt polysackaridmonooxygenas är ett enzym som bryter ned cellulosa genom en oxidativ reaktion. Det reaktiva centrat innehåller en kopparjon bunden till två histidingrupper. En av de senare har ofta en ovanlig metylering, men orsaken till denna modifiering har varit okänd. Våra beräkningar visar att denna grupp med största sannolikhet skyddar enzymet mot självoxidation.

Partikulärt metanmonooxygenas (pMMO) är ett av två enzymer som kan hydroxylera metan och bilda metanol, som kan användas som ett bättre bränsle än naturgas. Enzymet är membranbundet och därför svårt att studera experimentellt. Därför är sammansättningen och positionen för det reaktiva centrat fortfarande inte känd. Vi har studerat reaktiviteten hos tre möjliga reaktiva centra i pMMO, alla bestående av en kopparjon. Vi visar att alla tre kan utföra liknande kemiska reaktioner, men bara ett av dem, $\mathrm{Cu_C}$, har en gynnsam energetisk profil

Popular science summary in english

Enzymes are important for many reactions in nature. They are proteins that increase the speed of chemical reactions without reacting themselves. Thus, they catalyze the reactions. In the reactions, the enzymes act on other molecules called substrates. The created molecules are called products. Almost all processes in a living cell are catalyzed by enzymes ensuring that they are fast enough. In the human body alone, there are more than 700 different enzymes. Enzymes are important also for industrial applications e. g. the production of bio-fuels and artificial fertilizers. Hence, it is important to understand how enzymes work.

Both experiment and computer simulations are used to study enzymes. The computer power has increased much during the last decades and therefore computer simulations have become increasingly important. The fundamental theories behind the simulations are older.

For objects we meet in daily life we know which physical rules applies. They are called classical mechanics. For much smaller particles, partly others, sometimes surprising, rules are relevant. This is called quantum mechanics.

Quantum mechanics is the foundation of what we today call quantum chemistry. The theory was developed by researches around 100 years ago. According to quantum mechanics, electrons are rotating around atomic nuclei and only certain energy levels are allowed, the energy is quantized. To find the energy of an electron we have to solve the so-called Schrödinger equation. It is what the mathematicians call a partial differential equation, i. e. an equation containing derivatives and the solution is a function of several variables. The variables are the time and the coordinates of all particles. The solutions are called wave functions and indicate the probability that a particle is at a certain position.

Unfortunately, the Schrödinger equation can be solved analytically only for a few simple cases. Therefore simplifications are needed to be able to solve the Schrödinger equation with a computer. This is called numerical solutions.

In this thesis, we use mainly two numerical methods, molecular mechanics and density functional theory (DFT). DFT states that energy of the ground state can be found as a functional of the electron density

We have used these methods to study three different enzymes. Nitrogenase is the only enzyme that can cleave the strong triple bond in molecular nitrogen to form two molecules of ammonia and make atmospheric nitrogen available to plant life. It contains in its active site a complicated $MoFe_7S_9C$ cluster, called the FeMo cluster. Previous DFT investigations have shown that the results depend strongly on which

DFT method is used, so that it is not known which method can be trusted. Therefore, we have used two different approaches to calibrate the DFT calculations. In the first, we use experimental data of structures and reactions related to nitrogenase. In the second, we instead use more advanced quantum mechanical methods on a minimal model of the FeMo cluster.

Lytic polysaccharide monooxygenase is an enzyme that degrades cellulose by an oxidative reaction. The active site contains a copper ion bound two histidine ligands. One of them often has an unusual methylation, but the reason for this modification has been unknown. Our calculations show that this group most likely protects the enzyme against self-oxidation.

Particulate methane monooxygenase (pMMO) is one of two enzymes that can hydroxylate methane, forming methanol, which can be used as a more convenient fuel than natural gas. The enzyme is membrane-bound and therefore hard to study experimentally. Therefore, the nature and position of the active site is still not known. We have studied the reactivity of three putative active sites in pMMO, all mononuclear copper sites. We show that all three sites can perform similar chemical reactions, but only one of them, the $\mathrm{Cu_C}$ site, has favourable energetics.

Computational Studies of Metalloenzymes

Chapter 1

Introduction

In this thesis, we study the reactions of three enzymes, nitrogenase, lytic polysaccharide monooxygenase (LPMO) and particulate methane monooxygenase (pMMO) by computer simulations. They are all metalloenzymes, meaning that they are proteins with one or several bound metal ions. Metalloenzymes are common in nature and constitute about one third of all enzymes known today. In this work, we study metalloenzymes with metal ions located in the active site which perform redox reactions. Enzymes with catalytic functions are called biocatalysts.

Enzymes are involved in many reactions in nature and take part in many processes of industrial interest, e.g.in agriculture, metabolism and chemical synthesis. It should also be pointed out that the human body has more than 700 enzymes involved in various reactions. Because of the many applications, enzymes have been thoroughly studied, traditionally by experiments, but today also with computer simulations.

Traditionally, the prime source of structural information of enzymes is X-ray crystallography. In this, the pattern of scattered X-rays by crystalline samples is studied. Unfortunately, hydrogen atoms are not seen and it is impossible to measure the phase. Chemical processes are typically too fast to be studied.

A way to get around some of these problems is to use computer simulations, and this is the main goal of this thesis. With increasing computer power, computational methods have become an important alternative and supplement to experimental techniques. We use mainly two techniques, density functional theory (DFT) and molecular mechanics (MM). It is also possible to combine them in a method called QM/MM. This is often done for performance reasons, because MM methods are less computationally expensive than QM methods. Different DFT methods sometimes give diverging results. Naturally, this is problematic when modelling chemical reactions.

Therefore, we have performed two benchmarking studies in this thesis discussing the accuracy of different DFT methods.

The first studied enzyme is nitrogenase. It is the only enzyme that can cleave the triple bond in N_2 to form ammonia and make nitrogen available for plant metabolism [1]. This is a very important enzyme and has therefore been thoroughly studied both with experiments and computer simulations. Still, there are no consensus regarding the reaction mechanism. Crystallographic studies have shown that nitrogenase is a heterotetramer that contains a $MoFe_7S_9C$ cluster in the active site and this cluster is bound to the protein by a histidine and a cysteine residue [2, 3].

The second studied enzyme is lytic polysaccharide monooxynegase (LPMO). This enzyme is a copper dependent enzyme that is used for degradation of polysaccharides, such as cellulose and chitin [4–9]. The active cite Cu-ion in LPMO has six ligands, two water molecules, two histidines (one binding in a histidine brace) and one tyrosine. One of the major problems with this enzyme is that it can self-oxidise and deactivate. One of the histidines is methylated and we have investigated if this methyl group can protect the enzyme from self-oxidation.

The third studied enzyme is pMMO, which is one of two enzymes in nature that can break the C–H bond in methane [10–12]. pMMOs are the dominant enzyme in methanotrophic bacteria [10, 13]. For this enzyme, there is not any consensus regarding the nature of the active site. Various studies have suggested different putative active sites , e. g. the Cu_B site (which may contain one or two Cu ions), the Cu_C site, or a nearby Cu_{C2} site, or even a trinuclear cluster, which is still not fully observed in any crystal or cryogenic electron-microscopy structure. We have examined three of these sites to see if they can perform the oxidation of methane.

Chapter 2

Methods

2.1 Quantum Mechanics

Quantum mechanics is a fundamental theory of the behaviour of matter. It was developed in the early part of the 20th century. According to quantum mechanics, all measurable properties of a system can be deduced from the wave function Ψ, which is the solution of the Schrödinger equation. In its time-independent form, it can be written

$$\mathbf{H}\Psi = E\Psi. \tag{2.1}$$

Here, $\mathbf{H}$ is the Hamilton operator and E is the energy of the system. This can be seen as an eigenvalue problem where Ψ is the eigenfunction and E the eigenvalue to the Hamilton operator. Ψ does not have any direct interpretation but $|\Psi|^2$ gives the probability density that the particle is at a certain position. The Hamiltonian operator involves several terms. For a molecular system,

$$\mathbf{H} = \mathbf{V}_{en} + \mathbf{V}_{ee} + \mathbf{V}_{nn} + \mathbf{T}_e + \mathbf{T}_n \tag{2.2}$$

where, $\mathbf{V}_{en}$ is the attractive electrostatic potential between the electrons and the nuclei, $\mathbf{V}_{ee}$ and $\mathbf{V}_{nn}$ are the repulsive electrostatic interactions between all electrons and nuclei, respectively. $\mathbf{T}_e$ is the kinetic energy of the electrons and $\mathbf{T}_n$ is the kinetic energy of the nuclei. The various potential energies are given by

$$\mathbf{V}_{en} = -\frac{e^2}{4\pi\varepsilon_0}\sum_{i=1}^{N}\sum_{A=1}^{n}\frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|}, \tag{2.3}$$

$$\mathbf{V}_{ee} = \frac{e^2}{4\pi\varepsilon_0}\sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \tag{2.4}$$

and

$$\mathbf{V}_{nn} = \frac{e^2}{4\pi\varepsilon_0} \sum_{A=1}^{n} \sum_{B>A}^{n} \frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|}. \tag{2.5}$$

Here, i and j are the various electrons, A and B the various nuclei and Z_A and Z_B are the atomic number of the nuclei. e is the electron charge and ε_0 is the permittivity of vacuum. The total number of electrons is denoted N and the total number of nuclei is denoted n. $\mathbf{r}_i$ and $\mathbf{R}_A$ are the coordinates of the electrons and nuclei. The kinetic energy contribution are given by

$$\mathbf{T}_e = -\frac{\hbar}{2m_e} \sum_{i=1}^{N} \nabla_i^2 \tag{2.6}$$

and

$$\mathbf{T}_n = -\frac{\hbar}{2M_A} \sum_{A=1}^{n} \nabla_A^2. \tag{2.7}$$

In these equations, $\hbar$ is Planck´s constant divided by 2π, m_e is the mass of an electron and M_A the corresponding mass for nuclei A. ∇_i^2 and ∇_A^2 are the Laplacian operator for the electron and the nuclei, respectively.

Solving the Schrödinger equation is complicated but can be simplified by the so called Born–Oppenheimer approximation. This approximation is based on the fact that the mass of the nucleus is three to five magnitudes larger than the mass of the electron, so that the electrons move much faster than the nuclei. Therefore the kinetic energy of the nuclei can be omitted, simplifying Eq. 2.2 to

$$\mathbf{H} = \mathbf{V}_{en} + \mathbf{V}_{ee} + \mathbf{V}_{nn} + \mathbf{T}_e. \tag{2.8}$$

According to the Born–Oppenheimer approximation, the nuclei are stationary and only the electrons move.

2.2 Slater Determinants and Hartree–Fock Theory

To be able to efficiently solve the Schrödinger equation for more than one electron, approximations need to be used. One is that the total wave function can be written as a product of one-electron wave functions, ϕ_i. To fulfil the Pauli principle, it is normally written as a Slater determinant

$$\Phi_{SD} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(x_1) & \phi_2(x_1) & \dots & \phi_N(x_1) \\ \phi_1(x_2) & \phi_2(x_2) & \dots & \phi_N(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(x_N) & \phi_2(x_N) & \dots & \phi_N(x_N) \end{vmatrix}. \tag{2.9}$$

The one-electron wave function has a spatial and a spin part. The spatial part is here denoted $\psi(x)$. Electrons with a positive spin (spin up) are denoted α and those with negative spin (spin down) are called β. Thus, the one-electron wave function is given by

$$\phi = \psi(x) \cdot \begin{cases} \alpha(\omega) \\ \beta(\omega) \end{cases} \quad . \tag{2.10}$$

The next approximation is to write the one-electron wave function as a linear combination of atomic orbitals (basis functions) χ:

$$\phi_i = \sum_{\alpha=1}^{M} c_{\alpha,j}\chi_\alpha, \tag{2.11}$$

where M is the number of basis functions. The coefficients $c_{\alpha,j}$ in Eq. 2.11 can be determined by the variational principle, which state that the Rayleigh ratio $\mathcal{E}$ for a trial function, is given by,

$$\mathcal{E} = \frac{\langle \phi_{trial}|\mathbf{H}|\phi_{trial}\rangle}{\langle \phi_{trial}|\phi_{trial}\rangle} \tag{2.12}$$

is larger that lowest eigenvalue to the Hamiltonian, E_0.

One of the simplest methods to solve the Schrödinger equation is the so called Hartree–Fock method. This method assumes that each electron moves in an average field of all other electrons, which is called the mean field approximation, and that the total wave function is a Slater determinant. The Fock operator is defined as

$$\mathbf{F}_i = \mathbf{h}_i + \sum_{j=1}^{N} (\mathbf{J}_j - \mathbf{K}_j) \tag{2.13}$$

with

$$\mathbf{h}_i = -\frac{\hbar}{2m_e}\nabla_i^2 - \frac{e^2}{4\pi\varepsilon_0}\sum_{A=1}^{n}\frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|}. \tag{2.14}$$

In Eq. 2.13, $\mathbf{J}$ is the Coulomb operator describing the electron–electron repulsion between two electrons and $\mathbf{K}$ is the exchange operator giving the energy of exchanging two electrons. This can be treated as an eigenvalue problem according to

$$\mathbf{F}_i\boldsymbol{\phi}_i = \varepsilon_i\boldsymbol{\phi}_i, \tag{2.15}$$

where ε_i is the energy for orbital ϕ_i and is given by

$$\varepsilon_i = \langle \phi_i|\mathbf{F}_i|\phi_i\rangle = h_i + \sum_{j=1}^{N}\left(J_{i,j} - K_{i,j}\right). \tag{2.16}$$

The total energy is given by

$$E = \sum_{i=1}^{N} \varepsilon_i - \frac{1}{2} \sum_{i,j=1}^{N} (J_{i,j} - K_{i,j}) + V_{nn}, \tag{2.17}$$

where $V_{nn} = \langle \Phi | \mathbf{V}_{nn} | \Phi \rangle$, with $\mathbf{V}_{nn}$ from Eq. 2.5.

In practise, the calculation is an iterative process where the so called Roothan equations play an important role. They are given by

$$\mathbf{Fc} = \boldsymbol{\varepsilon} \mathbf{Sc}, \tag{2.18}$$

where $\mathbf{F}$ is the Fock matrix, $\mathbf{S}$ is the overlap matrix, $\mathbf{c}$ is a matrix with the values $c_{\alpha,j}$ defined in Eq. 2.11 and $\boldsymbol{\varepsilon}$ is a diagonal matrix containing the orbital energies. In a calculation, this equation is solved iteratively until the energy has converged.

2.3 Basis sets

As mentioned above in Eq.2.11, the one-electron wavefunction are expressed as a linear expansion of known mathematical functions, the basis set. The latter should be similar to hydrogen-atom wavefunctions, but also mathematically convenient (easy to calculate the needed integrals). In practise, two different types of basis set are used. These are Slater-type orbitals (STO) and Gaussian-type orbitals (GTO). The GTOs are given by

$$\chi_{\zeta,n,l,m}(r,\theta,\varphi) = N Y_{l,m}(\theta,\varphi) r^{2n-2-l} e^{-\zeta r^2} \tag{2.19}$$

and the STOs are given by

$$\chi_{\zeta,n,l,m}(r,\theta,\varphi) = N Y_{l,m}(\theta,\varphi) r^{n-1} e^{-\zeta r}. \tag{2.20}$$

Here, n, l,and m are quantum numbers, r is the distance between the electron and the nuclei, $Y_{l,m}(\theta,\varphi)$ are the spherical harmonics and N is a normalisation constant. The parameter ζ specifies the broadening of the function. A large value of ζ gives a sharper function and a small value gives a broader function.

The STOs have the same functional dependence as the hydrogen wavefunction. The GTOs do not have the correct shape of the wavefunction compared to the analytical solution of the wave function, and therefore more GTOs are needed to obtain a good accuracy. However, GTOs are computational more convenient and therefore most quantum chemistry software uses GTOs.

The accuracy of the basis set depends on the number of terms in Eq. 2.11, (an infinite sum gives a perfect result).

The smallest possible basis set contains only one basis function for each electron pair. This is called single zeta (SZ). It is normally necessary to add another basis function to get a better accuracy, using two basis functions for each electron pair (DZ). Even better results are obtained with three basis functions for each electron pair (TZ) or four (QZ). It is often enough to use one a single basis function for the core orbitals, and two for the valence orbitals. In this case, the letter V is added to the name. One example is VDZ, which also is called split valence (SV).

In this thesis, three basis set from the Karlsruhe family are used [14]. These are def2-SV(P), def2-TZVPD and def2-QZVPD. In paper II, instead Dunning's correlation consistent basis sets are used, cc-pVXZ, where X is D, T or Q.

2.4 Density Functional Theory

In 1964 Hohenberg and Kohn introduced two theorems that are the basis for the so-called density functional theory (DFT). The first theorem states that the ground-state energy can be obtained from the electronic density, ρ, and the second states that there exists a functional, $F[\rho]$, that gives the ground-state energy E. Thus, the ground-state energy can be obtained by a functional that depends only on the electron density [15, 16].

This is a bit surprising as the electron density is a function of only the three Cartesian coordinates, whereas the wavefunction is a function of $3N$ coordinates, where N is the number of electrons.

In practise, the use of DFT in computer calculations is based on Kohn–Sham theory [17]. This is an iterative self-consistent-field approach. Here, the energy is split up in two parts. One that can be solved exactly, and is calculated from the Schrödinger equation, and one correction term. The total energy is given by

$$E[\rho] = T_S[\rho] + E_{ne}[\rho] + J[\rho] + E_{XC}[\rho]. \tag{2.21}$$

Here, $T_S[\rho]$ is the kinetic energy for non-interacting electrons, $E_{ne}[\rho]$ is the attraction between nuclei and electrons, $J[\rho]$ is the Coulomb repulsion between the electrons and $E_{XC}[\rho]$ is the kinetic energy-correction and the exchange–correlation energy. $E_{XC}[\rho]$ is not known exactly but can be approximated in various ways.

2.4.1 Exchange–Correlation Functionals

In the local-density approximation (LDA), E_{XC} is constructed from the Dirac expression for the energy and is given by

$$E_{XC}^{LDA}[\rho] = A \int \rho(\mathbf{r})^{4/3} \mathrm{d}\mathbf{r}, \tag{2.22}$$

with

$$A = \frac{9}{8}\alpha \left(\frac{3}{\pi}\right)^{1/3} j_0. \tag{2.23}$$

Here α is an adjustable parameter, normally set to 2/3, and

$$j_0 = \frac{e^2}{4\pi\varepsilon_0}. \tag{2.24}$$

Another approach is the generalised gradient approximation (GGA). Here, the gradient, ∇, of the electronic density is also included,

$$E_{XC}^{GGA}[\rho] = \int f(\rho(\mathbf{r}), \nabla\rho(\mathbf{r})) \mathrm{d}\mathbf{r}, \tag{2.25}$$

where f is a suitable function.

In the meta generalised gradient approximation (mGGA), a term describing either the Laplacian of the electron density ($\nabla^2\rho(\mathbf{r})$) or the kinetic energy density, τ, is included, giving:

$$E_{XC}^{mGGA}[\rho] = \int f(\rho(\mathbf{r}), \nabla\rho(\mathbf{r}), \nabla^2\rho(\mathbf{r}), \tau(\mathbf{r})) \mathrm{d}\mathbf{r}, \tag{2.26}$$

with $\tau(\mathbf{r})$ as

$$\tau(\mathbf{r}) = \frac{\hbar^2}{2m_e} \sum_{i=1}^{N} \nabla\phi_i^*(\mathbf{r}) \cdot \phi_i(\mathbf{r}). \tag{2.27}$$

It is also possible to include a certain amount of Hartree–Fock exchange ,γ, giving a hybrid functional according to

$$E_{XC}^{hybrid}[\rho] = E_{XC}^{GGA}[\rho] + \gamma E^{HF}[\rho], \tag{2.28}$$

One of the most serious problem with many DFT functionals is the self-interaction error, i.e. that each electron interacts electrostatically with the electron density from itself. To reduce this problem, sometimes so-called range-separated functionals are

used. In these, the electron–electron Coulomb operator for the exchange energy is separated into two parts according to

$$\frac{1}{r_{i,j}} = \frac{1 - \mathrm{erf}(\omega r_{i,j})}{r_{i,j}} + \frac{\mathrm{erf}(\omega r_{i,j})}{r_{i,j}}, \tag{2.29}$$

where erf is the error function, ω is a parameter that controls the partitioning between the two parts and $r_{i,j} = |\mathbf{r}_i - \mathbf{r}_j|$. If HF exchange is used for the long-range component and the density exchange functional as the short-range part, the functionals is called range-separated functionals. If the opposite is true, they are called screened-exchange functionals.

In this thesis 23 functionals are used. These are six GGA-functionals with no admixture of Hartree–Fock exchange, PBE [18], BP86 [19, 20], BLYP [19, 21], ωB97 (range-separated) [22], ωB97x (range-separated) [22] and B97D [23], four meta GGA functionals with no Hartree–Fock exchange, SCAN [24], TPSS [25], M06-L [26] and MN15L [27] and the hybrid functionals TPSSh (10%) [28], B3LYP* [29] (15%), B3LYP (20%) [19, 21, 30], SCAN0 (25%) [31], HSE06 (25%, screened-exchange) [32], PBE0 (25%) [33], M06 (27%) [34], PW6B95 (28%) [35], HSE12s (42.5%, screened-exchange) [36], MN15 (44%) [27], BHLYP (50%) [37], M06-2X (54 %) [34], and M06-HF (100%) [38]. The values within the parentheses is the amount of Hartree–Fock exchange, γ, in Eq. 2.28.

2.5 Coupled-cluster calculations

Coupled-cluster theory is a method to improve Hartree–Fock calculations by performing a series expansion in excited conformations. The first step is to define the cluster operator C. It relates the electronic wave function Ψ to the HF wave function Ψ_0 according to

$$\Psi = e^C \Psi_0, \tag{2.30}$$

where e^C is given by

$$e^C = 1 + C + \frac{1}{2!}C^2 + ... \tag{2.31}$$

and

$$C = C_1 + C_2 + ... + C_N. \tag{2.32}$$

N is the total number of electrons and the C_i operators generate all determinants with i excitations from the HF-wavefunction. If these operators act on the wave function we get the following results.

$$C_1 \Psi_0 = \sum_{a,p} t_a^p \Psi_a^p \tag{2.33}$$

and

$$C_2\Psi_0 = \sum_{a,b,p,q} t_{ab}^{pq}\Psi_{ab}^{pq}. \tag{2.34}$$

Here, t_a^p is single-excitation amplitudes and t_{ab}^{pq} is double-excitation amplitudes. It is possible to continue to higher orders. The methods we used in my calculations are CCSD (coupled cluster singles and doubles), where $C = C_1 + C_2$, and CCSD(T), (coupled cluster singles, doubles and perturbatively treated triples) where $C = C_1 + C_2 + C_3$. CCSD(T) has become the gold-standard method in QM, owing to its accurate results.

To calculate the energy, it is necessary to start with the Schrödinger equation. Here, we use the simplest CCD equation with $C = C_2$.

$$\mathbf{H}e^C\Psi_0 = Ee^C\Psi_0 \tag{2.35}$$

which can be rewritten as

$$\mathbf{H}\left(1 + C_2 + \frac{1}{2}C_2^2 + ...\right)\Psi_0 = E\left(1 + C_2 + \frac{1}{2}C_2^2 + ...\right)\Psi_0. \tag{2.36}$$

After multiplication of Eq. 2.36 with Ψ_0^* and integrating over all space we get

$$E_{HF} + \langle\Psi_0|\mathbf{H}C_2|\Psi_0\rangle = E. \tag{2.37}$$

Unfortunately this equation contains two unknowns, the energy E and the amplitudes in C_2. A way to get around this is to multiply the Schrödinger equation with Ψ_{ij}^{kl*} to get

$$\langle\Psi_{ij}^{kl}|\mathbf{H}|\Psi_0\rangle + \langle\Psi_{ij}^{kl}|\mathbf{H}C_2|\Psi_0\rangle + \frac{1}{2}\langle\Psi_{ij}^{kl}|\mathbf{H}C_2^2|\Psi_0\rangle = E\langle\Psi_{ij}^{kl}|C_2|\Psi_0\rangle. \tag{2.38}$$

After inserting Eq. 2.37 in Eq. 2.38 we get

$$\begin{aligned}\langle\Psi_{ij}^{kl}|\mathbf{H}|\Psi_0\rangle + \langle\Psi_{ij}^{kl}|\mathbf{H}C_2|\Psi_0\rangle + \frac{1}{2}\langle\Psi_{ij}^{kl}|\mathbf{H}C_2^2|\Psi_0\rangle = \\ = (E_{HF} + \langle\Psi_0|HC_2|\Psi_0\rangle)\,\langle\Psi_{ij}^{kl}|C_2|\Psi_0\rangle\end{aligned} \tag{2.39}$$

Now we have m equations and after computing relevant Hamiltonian matrix element, it is possible to get both the energy and the wave function.

2.6 Molecular Mechanics

In standard QM methods, the Schrödinger equation is solved. Unfortunately this can be time consuming and typically becomes prohibitive for systems with more than

$\sim$ 1000 atoms. In many applications it is necessary to study more atoms. A way to get around this problem is to use so-called molecular mechanics (MM) methods. In these, the Schrödinger equation is no longer considered and the electrons are ignored. Instead, atoms are modelled as balls and bonds as strings. The total energy is calculated from an empirical energy function, which gives the energy as a function of the coordinates. For a biological macromolecule it typically consists of five energy terms:

$$E_{tot} = E_{bonds} + E_{angles} + E_{dihedrals} + E_{VdW} + E_{el}. \tag{2.40}$$

Here, the first three terms correspond to the internal energy as a function of the bonds, angles and dihedrals. The last two correspond to the intermolecular interactions from Van der Waals and electrostatic contributions. The bond energy is

$$E_{bonds} = \sum k_b \left(r - r_0\right)^2, \tag{2.41}$$

where k_b is the spring force constant, r is the actual bond length and r_0 the ideal bond length. This is a harmonic potential. The energy of the bond angles are given by

$$E_{angles} = \sum k_a \left(\theta - \theta_0\right)^2, \tag{2.42}$$

where, k_a is the angle force constant, θ the actual angle and θ_0 is the ideal angle. Also this expression is a harmonic potential. The dihedral term uses a periodic function to model torsion angle rotation and is given by

$$E_{dihedrals} = \sum \frac{V_n}{2} \left(1 + \cos\left(\phi n - \delta\right)\right), \tag{2.43}$$

where n is the periodicity of the angle ϕ, δ is the phase shift (typically 0^o or 180^o) and V_n is the force constant. The Van der Waals interaction is described by a Lennard-Jones potential

$$E_{VdW} = \sum_i \sum_{j \neq i} 4\varepsilon_{ij} \left(\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right) \tag{2.44}$$

where ε_{ij} is the depth of the potential, r_{ij} is the distance between atoms and σ_{ij} is the distance where the energy is zero. This expression describes energies between atoms that are not covalently bonded. At large distances it goes to zero and for small distances it is strongly repulsive. The last term, giving the electrostatic contribution, is given by

$$E_{el} = \sum_i \sum_{j \neq i} \frac{q_i q_j}{4\pi\varepsilon\varepsilon_0 r_{ij}}, \tag{2.45}$$

where q_i and q_j are the partial charges of each atom, ε_0 is the permittivity of vacuum, ε is the permittivity of the medium (the dielectric constant) and r_{ij} is the distance between the atoms.

It can be seen from the equations above that many parameters need to be known to do a MM calculation. The existing standard force fields have parameters obtained from experimental data or QM calculations.

In this thesis, we have used the Amber ff14SB force field.

2.7 Molecular Dynamics

In molecular mechanics (MD), the main idea is that the motion of a molecule follows the laws of classical mechanics. Newton's second law states that

$$\mathbf{F}_i(t) = m_i \mathbf{a}_i(t), \tag{2.46}$$

with

$$\mathbf{a}_i(t) = \frac{\mathrm{d}^2 \mathbf{r}_i(t)}{\mathrm{d}t^2} \tag{2.47}$$

and $\mathbf{F}_i$ is the force on atom i, m_i its mass, $\mathbf{a}_i(t)$ its acceleration and $\mathbf{r}_i(t)$ the position in space. The force $\mathbf{F}_i$ is calculated from the potential energy U according to

$$\mathbf{F}(t) = -\frac{\mathrm{d}U(t)}{\mathrm{d}\mathbf{r}(t)}. \tag{2.48}$$

The potential can be calculated from a QM or a MM force-field calculation.

In practise, MD calculations are done using an iterative algorithm starting from an initial guess and thereafter calculating velocities and positions by integrating Eq. 2.46 for a time step Δt. For small time steps, this can be done numerically by using a Taylor expansion according to

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \frac{\mathrm{d}\mathbf{r}(t)}{\mathrm{d}t}\Delta t + \frac{1}{2}\frac{\mathrm{d}^2\mathbf{r}(t)}{\mathrm{d}t^2}\Delta t^2 + ..., \tag{2.49}$$

with

$$\mathbf{v}(t) = \frac{\mathrm{d}\mathbf{r}(t)}{\mathrm{d}t}. \tag{2.50}$$

For the velocity, $\mathbf{v}(t)$, the corresponding Taylor expansion is

$$\mathbf{v}(t + \Delta t) = \mathbf{v}(t) + \frac{\mathrm{d}\mathbf{v}(t)}{\mathrm{d}t}\Delta t + \frac{1}{2}\frac{\mathrm{d}^2\mathbf{v}(t)}{\mathrm{d}t^2}\Delta t^2 + ..., \tag{2.51}$$

with

$$\mathbf{a}(t) = \frac{\mathrm{d}\mathbf{v}(t)}{\mathrm{d}t}. \tag{2.52}$$

If Δt is small, Eq. 2.49 can be approximated by

$$\mathbf{r}(t + \Delta t) \approx \mathbf{r}(t) + \mathbf{v}(t)\Delta t + \frac{1}{2}\mathbf{a}(t)\Delta t^2, \tag{2.53}$$

and Eq. 2.51 to

$$\mathbf{v}(t + \Delta t) \approx \mathbf{v}(t) + \mathbf{a}(t)\Delta t. \tag{2.54}$$

The algorithm starts from an initial guess of $r(0)$ and $v(0)$ and thereafter F and a is calculated and the atoms are moved to $r(t + \Delta t)$ and the velocities to $v(t + \Delta t)$. This procedure is then repeated for a number of time steps.

A normal time step for a MD simulation is around 0.5 fs. However, this can be increased by using the SHAKE algorithm [39], where bond lengths involving hydrogen atoms are fixed. This allows the time step to be increased to 2 fs.

In most MD simulations, we want to study a system in water solution. This is done by adding a box of water molecules. There are many different water models, TIP3P [40], SPC/E [41] or TIP4P [40]. It is also necessary to assign protonation states to the residues in the protein by studying their hydrogen-bonding pattern, solvent accessibility, surrounding residues, etc.

A small box or droplet of a solvated protein is often a poor model of the real protein in water solution. A better model can be obtained by using periodic boundary conditions. Then, the box is imagined to be surrounded by identical copies of the box, arranged symmetrically so if an atom leaves the box on one side, it enters the box on the opposite side. Thereby, an infinite system is modelled at essentially no extra computational cost. Moreover, then Ewald summation can be used to describe long-ranged electrostatic interactions [42].

2.8 QM/MM

QM methods are more accurate than MM methods but they are computationally much more expensive. A compromise can be obtained with the so-called QM/MM method [43–45]. Here, the most interesting part of the studied system is described with QM methods and the rest with MM methods. For a biomacromolecule, the most interesting part is the active site and this is normally described by the QM method and the rest of the proteins with the more efficient MM method. A typical QM/MM energy function is given by

$$E_{QM/MM} = E_{MM12} + E_{QM1} - E_{MM1}. \tag{2.55}$$

Here, the total system is divided into two subsystems, 1 and 2. Subsystem 1 is described by a QM calculation and both subsystems are described by a MM calculation. To avoid double counting, the MM energy for subsystem 1 needs to be removed.

In most calculations, there are covalent bonds between subsystems 1 and 2. Since QM calculations need to have full valences of all atoms, it is necessary to truncate the QM system. In most QM/MM approaches this is accomplished by replacing a carbon atom at the boundary (the carbon link atom, CL) by a hydrogen atom (the hydrogen link atom, HL) [46, 47].

A way to improve the result of a QM/MM calculation is to include the electrostatic interactions between the QM and the MM regions in the QM calculations. This is called electrostatic embedding. Then, the QM/MM energy is given by

$$E_{QM/MM} = E^{CL}_{MM12,q_1=0} + E^{HL}_{QM1+ptch2} - E^{HL}_{MM1,q_1=0}. \tag{2.56}$$

Here $E^{CL}_{MM12,q_1=0}$ is the MM energy for the entire system with CL atoms and with the charges of subsystem 1 zeroed, $E^{HL}_{QM1+ptch2}$ is the QM energy of subsystem 1 with truncated HL atoms and including a point-charge model of system 2, and $E^{HL}_{MM1,q_1=0}$ is the MM energy for the QM region with HL atoms and no electrostatic interactions.

In this thesis, an interface, called ComQum, is used to perform the QM/MM simulations [47, 48]. This program combines the QM software Turbomole and the MM software AMBER. It uses the energy function in Eq. 2.56.

Chapter 3

Studied systems

In this thesis, we have performed studies related to three enzymes. They are shortly introduced here.

3.1 Nitrogenase

Nitrogenase is the only enzyme in nature that can cleave the triple bond in N_2 to form ammonia and make nitrogen available for plant metabolism [1]. This is important in many applications, for example in agriculture and the production of artificial fertilizers. Nitrogenase has been thoroughly studied with both theoretical and experimental methods. Crystallographic studies have shown that nitrogenase is a heterotetramer (Figure 3.1) that contains a $MoFe_7S_9C$ cluster in the active site called the FeMo cluster [2, 3] (Figure 3.2). This cluster is bound to the protein by a histidine and a cysteine residue [1].

Nitrogenase catalyzes the reaction [1]

$$N_2 + 8e^- + 8H^+ + 16ATP \rightarrow 2NH_3 + H_2 + 16ADP + 16P_i. \qquad (3.1)$$

The reaction is normally described by the so-called Lowe–Thornely cycle [49]. This has nine intermediates, E_0 to E_8, differing in the number of added electrons and protons.

Even if nitrogenase has been thoroughly studied both experimentally [1, 50–57] and by computer simulations [1, 51, 58–72], there is still no consensus regarding the reaction mechanisms or the structure of the involved intermediates. The E_4 intermediate is believed to bind to N_2. It has been suggested by Hoffman et al. that the E_4 structure

contains two hydride ions bridging two pairs of Fe ions [55, 64]. These are Fe2–Fe6 and Fe3–Fe7. Atom names are shown in Figure 3.2. This has been supported by DFT calculations and ENDOR experiments [55, 64]. Einsele et al. has suggested that one of the the sulfide ligands, S2B, dissociates and that the two hydride ions both bridge the Fe2–Fe6 ions [73]. Björnsson has suggested a similar structure but with the difference that S2B dissociates only from the two Fe ions [74]. Finally, Siegbahn suggested that E_4 instead contains a triply protonated carbide ion moving out of the cluster [75–77].

This disagreement is partly caused by the problem that different DFT methods give different results. In papers I and II, we compare DFT functionals to decide which gives the most accurate results [78].

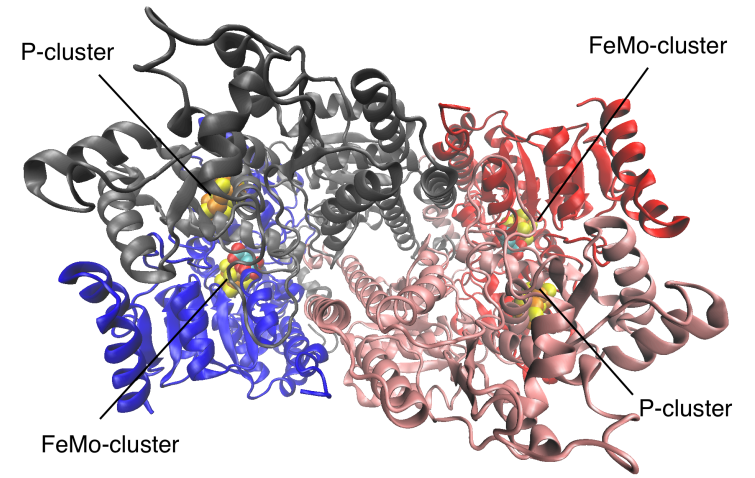


Figure 3.1: *The nitrogenase protein with the FeMo and P-clusteres marked*

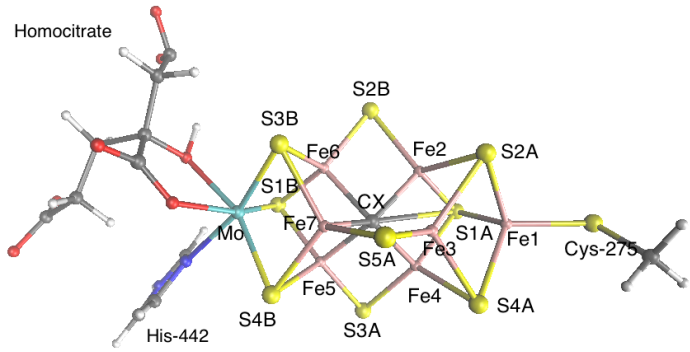


Figure 3.2: *The FeMo-cluster with atom names.*

3.2 LPMO

Lytic polysaccharide monooxygenases (LPMO), (Figure 3.3), are copper dependent enzymes that can be used for degradation of polysaccharides such as cellulose and chitin [4–9]. LPMOs are divided into five families, AA11, AA13, AA14, AA15 and AA16 [4, 5, 79–85] but LPMOs belonging to different families are not necessarily very different in function. LPMOs catalyse an oxidative reaction,

$$\mathrm{RH} + \mathrm{O_2} + 2\mathrm{H^+} + 2e^- \rightarrow \mathrm{ROH} + \mathrm{H_2O}, \tag{3.2}$$

illustrating the oxidation of a substrate RH. They can also use $\mathrm{H_2O_2}$ as the cosubstrate

$$\mathrm{RH} + \mathrm{H_2O_2} \rightarrow \mathrm{ROH} + \mathrm{H_2O}, \tag{3.3}$$

which avoids the need of an external reductant. Recent findings show that $\mathrm{H_2O_2}$ leads to significantly faster reactions than $\mathrm{O_2}$ [86–89].

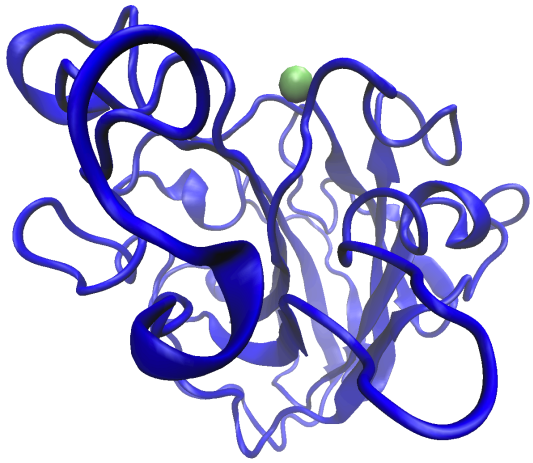


Figure 3.3: *The LPMO enzyme with the Cu-atom in the active site showed.*

The active site of the LPMOs consists of a copper ion ligated to three nitrogen donor atoms in a histidine brace involving one histidine ligand binding by the imidazole side chain and another histidine ligand binding by both the imidazole group and the amino-terminal group Figure 3.4. The copper ion also binds one or two water molecules, and in some enzymes also to a tyrosine.

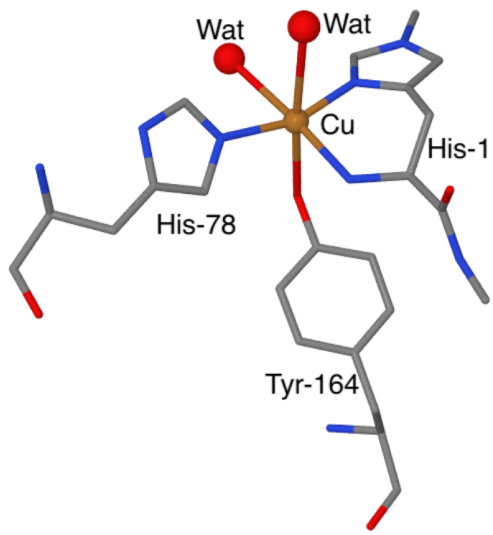


Figure 3.4: *The active site of LPMO.*

Industrial applications of LPMO have already begun to appear but are far from their full potential [90]. One major problem is that LPMOs self-oxidise and deactivate. Therefore the mechanism of this deactivation is important to understand. So far, computational studies of LPMOs have focused either on the substrate binding process [91–93] or on the reaction with the bound substrate [94–99]. In the latter case it is believed that the reaction with the substrate needs a strong oxidant, probably a Cu(II)–oxyl species [97–99].

Still, there are no investigations of the self-oxidation reaction, but it has been noted to occur on the histidine residues binding to copper [88]. However, the mechanism of the oxidation is unknown. The substrate is known to protect against oxidative damage [88].

In may LPMOs , it is observed that one of the His ligands is methylated on NE2. The function of this methylation is not clear, but two studies have suggested that it may have a protective role, avoiding self-oxidation [79, 100].

3.3 pMMO

There exist two enzymes in nature that can break the C–H bond in methane [10], soluble methane monooxygenases [11, 12] and particulate (i.e. membrane-bound) methane monooxygenases [10, 13]. They are referred to as sMMO and pMMO, respectively. sMMOs are known to have a diiron centre in the active site [101, 102]. The mechanism of this enzyme is quite well understood.

pMMOs are the dominant enzyme in methanotrophic bacteria [10, 13]. However, they are harder to study and the mechanism is not fully understood. In fact, there is not even any consensus regarding the nature of the active site. Crystallographic studies have shown that the enzyme is a trimer of three chains pmoA, pmoB and pmoC (Figure 3.5).

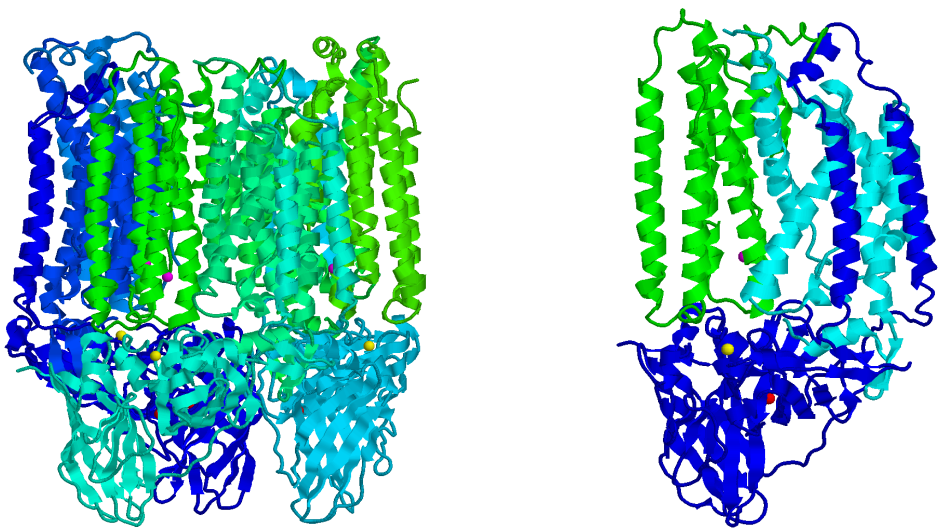
The first X-ray studies identified a monomeric and a dimeric Cu site in the pmoB subunit and a Zn site in the pmoC subunit [12]. These three metalsites are denoted A, B and C in the following. The identification of site B as a dinuclear site was mainly based in extended X-ray absorption fine-structure (EXAFS) studies [103–105]. Later crystallographic studies sometimes found only a mononuclear site and a detailed quantum refinement study gave no support for any dinuclear Cu_B site in any structure [106]. Instead, a single Cu ion fitted the crystallographic raw data best.

It has been much discussed which is the active site of the pMMO. The crystallographers originally argued that it is the Cu_B site. This was tempting especially as it is

coordinated by three histidine residues, one of which coordinates with both the terminal amino group and the imidazole side chain, i.e. with a histidine brace, similar to the one found in the active site of the LPMOs. However, later they changed their mind and suggested instead that $\mathrm{Cu_C}$ is the active site [107, 108]. In contrast, early Mössbauer spectroscopy studies instead suggested that site C is the active site, but they argued that it contains a diiron site, similar to that in the sMMOs [109]. On the other hand, Chan and coworkers have long argued that the active site is a trimeric copper site that is not observed in the crystal structures [110–114].

Recently, two cryogenic electron microscopy studies of pMMO have been presented. The study of Rosenzweig and coworkers observed the same Cu ions as in the crystal structures and suggested that the active site is $\mathrm{Cu_C}$ or possibly another nearby mononuclear Cu site (called $\mathrm{Cu_{C2}}$ in the following) [115]. On the other hand, the structure by Chan and coworkers showed several new Cu sites. They suggested that the active site is a trinuclear cluster, but they observed only two Cu ions there in their structure [116].

Several computational studies of the pMMOs have been performed. One suggested that the active species is a bis(μ-oxo)Cu(I)Cu(II) site (based on $\mathrm{Cu_B}$) or a Cu(III)–oxo species (based on $\mathrm{Cu_C}$) [117–119]. Another study compared the reactivity of the $\mathrm{Cu_B}$ and $\mathrm{Cu_C}$ sites with a duroquinol substrate [120]. They argued that only the $\mathrm{Cu_C}$ site is reactive.



(a) (b)

Figure 3.5: *pMMO with three metal sites, $\mathrm{Cu_A}$ in red, $\mathrm{Cu_B}$ in yellow and $\mathrm{Cu_C}$ in magenta (from the 3RGB structure. (a) shows the entire trimer of trimers, whereas (b) shows only one subunit with pmoA in cyan, pmoB in blue and pmoC in green.*

Chapter 4

Summary of the Papers

4.1 Paper I: Comparison of the Accuracy of DFT Methods for Reactions with Relevance to Nitrogenase

As described before in this thesis, nitrogenase is an enzyme that can cleave the triple bond in N_2 to form ammonia. This process has applications in agriculture and chemical synthesis. Therefore, nitrogenase has been the subject of many studies, both experimental and theoretical.

A major problem with theoretical studies of nitrogenase is that different DFT methods give very different results. For example, different functionals give different predictions of which structure is most favourable for the key E_4 intermediate in the reaction mechanism and relative energies of different protonation states may differ by ~600 kJ/mol between standard GGA and hybrid functionals. To solve this problem it is important to calibrate different DFT methods using experimental data. This has been attempted by Prof. Dance. Unfortunately, he used an uncommon software, DMol3, with numerical basis sets and without any dispersion correction. We have therefore extended this study using more standard and state-of-the art methods and also more modern DFT functionals.

We employed 16 DFT methods: PBE, BP86, BLYP, B97D (GGA functionals), TPSS, M06-L, MN15-L (meta GGA functionals), TPSSh, B3LYP, PBE0, M06, PW6B95, MN15, BHLYP, M06-2X and M06-HF (hybrid functionals). In the calculations, three basis sets were used, def2-SV(P), def2-TZVP and def2-QZVP. The resolution-of-identity approximation (RI) was used to speed up the calculations. Dispersion corrections were included for most methods, using the DFT-D4 approach.

Since most experiments are measured in solution it is necessary to include solvent effects. This was done by using the conductor-like screening model (COSMO) implemented in Turbomole [121, 122].

We calculated bond lengths, angles, vibrational energies, enthalpies, and entropies for the binding of N_2, H_2, and CO to various transition-metal complexes with Fe, Ni, Cr, Mo and W. They were chosen because they are related to the reactions catalysed by of nitrogenase. We compared our calculations with experimental data and Dance's results to evaluate the various DFT methods.

The ranking was done by four methods: the first was to sum the mean absolute deviation (MAD) for the various DFT functionals. The second was to weight the errors with the largest absolute deviation from experiments. The third was to weight the MADs with a chosen acceptable error for each property. The fourth was to see how often each functional gives a results within twice the experimental uncertainty.

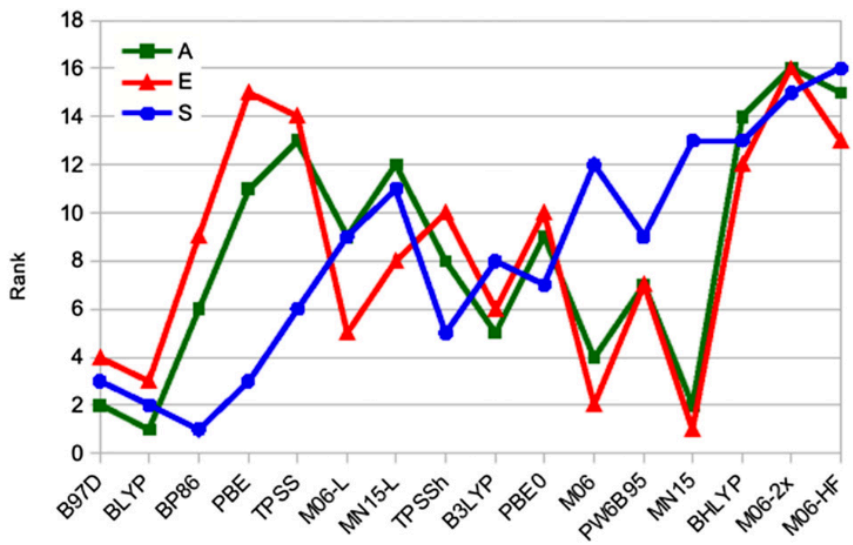


Figure 4.1: *The sum of the ranking for all properties (A), energetic properties (E) and structural properties (S). The lower value the better performance.*

In general, the four scoring methods gave quite similar results. As can be seen in Figure 4.1, we saw a significant difference between energetic and structural results. For structural properties, GGA and meta-GGA functionals give the best results, except M06-L and MN15-L. TPSSh also gives quite good results but B97D, BLYP, BP86 are the best. For energies, MN15 gives the best results but M06, BLYP, B97D and M06-L show also good scoring. Including both types of properties give results quite similar to those for the energetic properties. Hence, BLYP seems to behave well both for structures and energies. B97D also gives rather good results followed by MN15, M06

and B3LYP. Our results are quite different from those by Dance (who pointed out PBE as the best functional), partly because we use more data and tested more DFT methods, and partly because we include dispersion correction in our calculations.

4.2 Paper II: Calibration of DFT functionals for a minimal nitrogenase $[Fe(SH)_4H]^-$ model employing state-of-the-art ab initio methods

As said in the previous section, different QM methods give diverging results for calculation on nitrogenase. This is of course a serious problem when doing computer calculations on this enzyme. We have therefore constructed a minimal model of the FeMo cluster in nitrogenase shown in Figure 4.2. We study the energy difference between protonation states of the two models, one with hydrogen atom placed on the Fe ion (i. e. formally a hydride ion) and the other with a proton on the S atom.

We use DFT, coupled cluster methods, multiconfigurational CASSCF/CASPT2[123, 124] and DMRG-CASSCF to calculate this energy difference, ΔE. We use Dunning´s correlation consistent basis sets of two sizes.

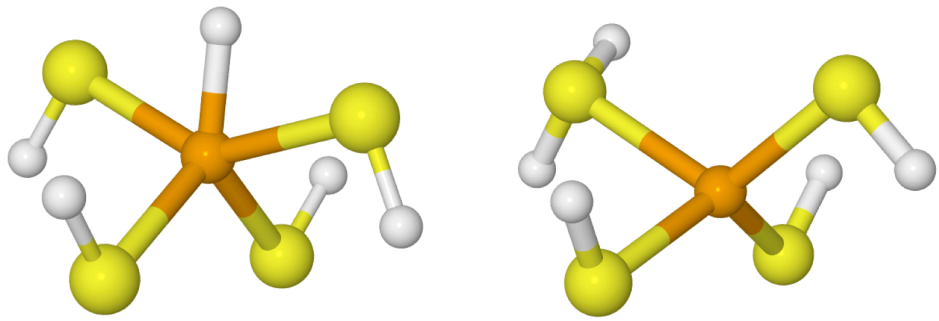
For the DFT calculations we used TURBOMOLE 7.6 with 17 functionals: PBE, BLYP, B97D, BP86 (pure GGA), TPSS, SCAN (meta GGA), ωB97 (range separated pure GGA), B3LYP, B3LYP*, PBE0, M06 (hybrid GGA), TPSSh, SCAN0, MN15 (hybrid meta-GGA), ωB97X (range separated GGA), HSE06 and HSE12S (hybrid GGA screened-exchange). As for the full FeMo cluster, we see large differences between the various DFT methods, (25–162 kJ/mol) mainly depending on the amount of Hartree–Fock exchange, see Figure 4.3.

The model systems are so small that we can perform high-level coupled-cluster methods to calibrate the DFT methods. With the small basis set, we can afford to run full CCSDT calculations. To seek a cheaper alternative to use with larger basis sets, we test four different methods with approximate treatment of the triples. Among these, BCCSD(T) shows the best performance. Our best estimate of ΔE is obtained by extrapolation of the HF–CCSD–CCSDT results. We also use the stochastic heath-bath configuration interaction (SHCI) method [125–127] with extrapolation to obtained an independent second estimate of ΔE. Quit satisfactorily, the results of the two methods agree within 1 kJ/mol. CCSD(T) with Kohn–Sham (DFT) orbitals turns out to be a rather accurate approach, but the results depend somewhat on what DFT method used to obtain the orbitals.

The D_1 diagnostics indicate that the FeH model is quite multiconfigurational. Therefore, we perform multiconfigurational CASSCF/CASPT2 and DMRG calculations. Unfortunately such calculations depend strongly on the active space, as well as on the parameters of the method. However, by comparing the results of DMRG-CASCI and CCSD(T) calculations with the same orbitals and active space, we show that the CCSD(T) ΔE estimate is affected by only 5 kJ/mol by the multiconfigurational

nature of the complexes.

Consequently, we can conclude that our extrapolated CC results are quite accurate and our best estimate of ΔE is 101 kJ/mol. The DFT method that comes closests to this result is M06 and B3LYP (with errors of 3 and 5 kJ/mol), followed by B3LYP*.



(a) (b)

Figure 4.2: Our model system with the extra hydrogen on Fe (a) or on S (b).

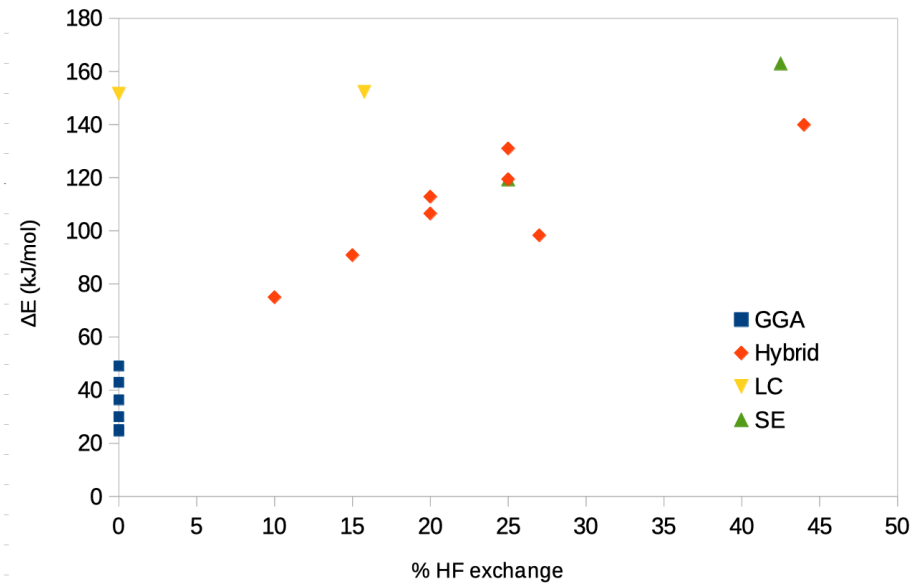


Figure 4.3: *Energy differences ΔE between the two protonated model structures computed with the various DFT approaches and the cc-pVXZ basis set as a function of the amount of HF exchange in the exchange functional. The methods are divided into (meta-)GGA, hybrid (meta-)GGA, long-range corrected (LC) and screened-exchange (SE) methods.*

4.3 Paper III: Histidine Oxidation in Lytic Polysaccharide Monooxygenase

The mechanism of LPMOs has been described earlier in this thesis. One of the major questions regarding this enzyme is why one of its histidine ligands is methylated, as is shown in Figure 4.4. In paper III, we study the effect of this group and whether it protects the enzyme from self oxidation by QM/MM studies.

The QM calculations were performed with Turbomole 7.5.1. We use the functionals TPSS and B3LYP and the basis sets def2-SV(P) and def2-TZVPD. The resolution-of-identity approximation (RI) was used to speed up the calculations. DFT-D3 dispersion correction were also included. We optimised geometries at the TPSS/def2-SV(P) level and then performed single-point calculations at the B3LYP/def2-TZVPD level.

We study a reaction in which the active site, either $[\mathrm{CuO}]^{+}$ or $[\mathrm{CuOH}]^{2+}$, abstracts a hydrogen atom from either His-1 or His-78 to form the intermediate $[\mathrm{CuOH}]^{+}$ or $[\mathrm{CuOH_2}]^{2+}$. Then, the OH^{-} group binds back to the histidines to get a hydroxylated product.

It can be seen in Figure 4.5 that the intermediate is less stable than the reactant, whereas the product is much more stable than the reactant. Reactions with $[\mathrm{CuOH_2}]^{2+}$ give the largest energy barriers.

The most important result is how the methyl group affects the reaction energies. It can be seen in Figure 4.5 that methylation increases the maximum activation barrier with 10 kJ/mol. This might seem small, but it would increase the life-time of the enzyme from 0.3 to 18 h. Hence, our results suggests that the methyl group on His-1 protects His-1 from self oxidation. For unmethylated His-78, the reaction barrier is approximately the same as for methylated His-1. Therefore, there is no reason to methylate His-78 since the net activation energy is the same anyway.

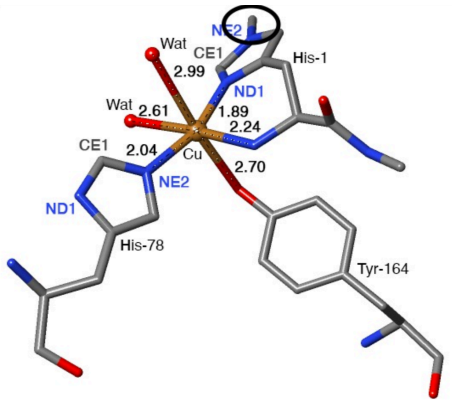


Figure 4.4: *LPMO with two histidines, one tyrosine and two water molecules binding to the Cu active site. The methylation of His-1 is emphasised with the black circle.*

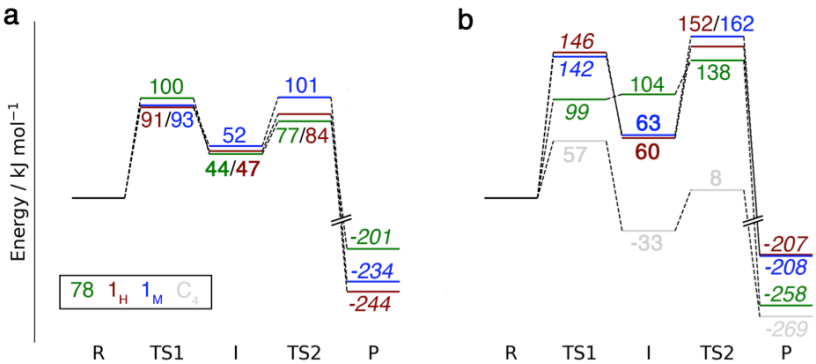


Figure 4.5: *Energy diagrams for the hydrogen-abstraction and rebound reactions of LPMO. Reaction with His-78, His-1 (with or without the methylation) and of the substrate are shown in green, red, blue and gray, respectively.*

4.4 Paper IV: Computational study of three metal sites in particulate methane monooxygenase

Methane is the hardest organic molecule to hydroxylate because it is hard to break the C–H bond. In nature, there is two enzymes that can catalyse that reaction, sMMO and pMMO. The latter is the more active enzyme, but it is membrane-bound. Therefore, it is hard to study experimentally. In fact, there is still no consensus regarding the nature and position of the active site in the enzyme. In Paper IV, we use QM/MM calculations to study the reactivity of three putative active sites, viz. the Cu_B, Cu_C and Cu_{C2} sites.

The study is based on a 2.8 Å crystal structure from Methylococcus capsulatus (PDB code 3RGB; Cu_B and Cu_C) and a 2.14 Å cryogenic electron microscopy structure from the same organism (PDB code 7S4H; Cu_{C2}).

All MD calculations were performed with the Amber software using the ff14SB force field and TIP3P water molecules. Charges of the metal sites in the MD simulations were obtained from TPSS/def2-SV(P) calculations. QM/MM calculations was performed with the ComQum software. The DFT calculations were done using Turbomole 7.1, 7.2 and 7.6.1. We used TPSS-D3/def2-SV(P) for geometry optimisations and B3LYP/def2-TZVPD for single-point energy calculations.

The calculations show that three sites can perform the same reactions (shown in Figure 4.6 and can form a reactive $[CuO]^+$ state. $[CuO]^+$ can abstract a proton from methane, which gives a Cu-bound OH^- group. This group can then recombine with CH_3 giving methanol. The reactive $[CuO]^+$ states forms from a O–O bond cleavage in the $[CuH_2O_2]^+$ structure, giving $[CuO]^+ + H_2O$. However, for Cu_B and Cu_C, the $Cu^I + HOOH$ structure, which is another isomer of $[CuH_2O_2]^+$ is 53 and 19 kJ/mol more stable than $[CuO]^+ + H_2O$, respectively, whereas for Cu_{C2} the opposite is true by 25 kJ/mol. For Cu_B, HOOH also dissociates from Cu and may diffuse away.

The activation energy for the hydrogen-abstraction reaction is 59, 57 and 93 kJ/mol for the Cu_B, Cu_C and Cu_{C2} sites, respectively (cf. Figure 4.7. The activation energy for the rebound reactions are 49, 9 and 63 kJ/mol, respectively, and involves a crossing from triplet or open-shell singlet to the closed-shell singlet for Cu_B and Cu_C. If the energies are measured from the best $[CuH_2O_2]^+$ structure, the net activation barriers are 112, 76 and 93 kJ/mol for Cu_B, Cu_C and Cu_{C2}, respectively. Thus, our results suggests that Cu_C is proper the active site of the enzyme.

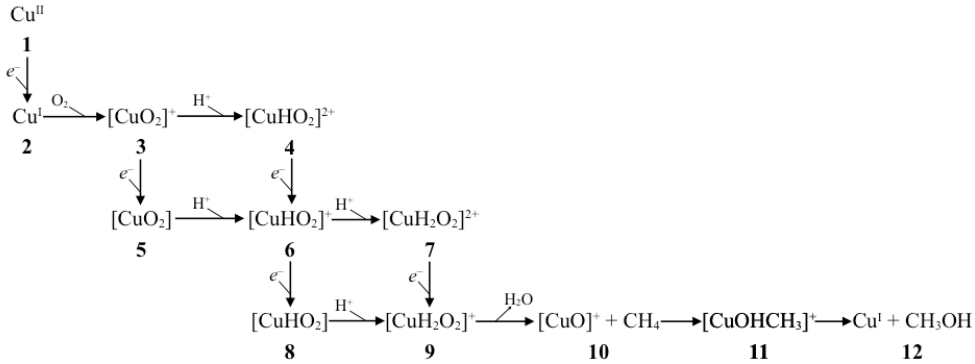


Figure 4.6: *Reactions in the activation of* O_2 *by pMMO*

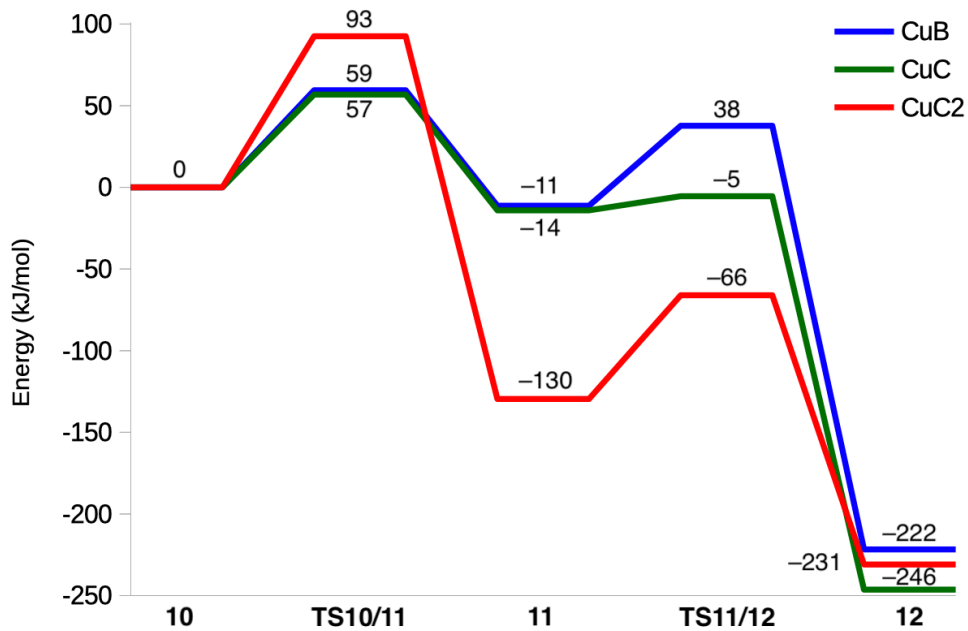


Figure 4.7: *Reaction energies of the* Cu_B, Cu_C *and* Cu_{C2} *sites in pMMO, relative to the best state of* 10.

Chapter 5

Conclusions and Outlook

In this thesis, I have used computational chemistry to study enzyme reactions. In such studies, it is mandatory to know that the employed methods are accurate and reliable. Therefore, it is necessary to calibrate the methods. This can be done in two different ways, as is illustrated by the studies by Papers I and II. In the first, we use experimental data to calibrate DFT methods to be used to study reactions with relevance to nitrogenase. Experimental data constitute ideal reference points, because they in principle represent the truth. On the other hand, almost no experimental data are available for true nitrogenase-like reactions and models. Therefore, it is hard to find data that is fully relevant for nitrogenase. Moreover, the experimental conditions are often quite different from those used in the calculations, making the comparison harder. Finally, also experimental measurements have a limited accuracy and in several cases in Paper I, we have good reasons to suspect errors in the experimental data.

In Paper II, we instead try to use more accurate QM methods to calibrate DFT methods to use for nitrogenase calculations. Such an approach has the advantage that the various methods measure exactly the same thing. On the other hand, the active-site FeMo cluster of nitrogenase is much too large to treat with the most accurate QM methods. Therefore, we have designed a small model that can be treated with advanced QM methods, but still shows some of the characteristics from nitrogenase (an iron-sulfur cluster with large difference between DFT methods for the energy difference between different protonation states).

Unfortunately, the two studies do not give consistent results. Paper II points out M06 and B3LYP as the best DFT functionals. In Paper I, M06 is the second best method for energies, but among the worst for structures. Likewise, B3LYP ranks only as 5–8 among the tested functionals (cf. Figure 4.1). Moreover, it is known that B3LYP gives

quite poor structures for the FeMo cluster of nitrogenase [128, 129]. Conversely, BLYP, B97D and MN15 (the three methods with the best ranks in Paper I) give errors of 71, 58 and 39 kJ/mol for the model in Paper II (ranking 9–16 among the 17 tested methods). The reason for this is of course that Papers I and II deal with different systems and properties. The model in Paper II is of course highly relevant for nitrogenase, but it is only a single, very simplified model. However, the good thing is that we also test out cheaper QM methods that can be used for larger and better models of the FeMo cluster. In particular, BCCD(T) and CCSD(T) with DFT orbitals seem to be promising methods to use in future studies of iron–sulfur clusters with 2–8 metal ions.

In Papers III and IV, we instead study reaction mechanism of two metalloenzymes of highest scientific and bioengineering interest, LPMO and pMMO. Both enzymes involve mononuclear Cu sites, a reactive $[CuO]^+$ state and similar reaction mechanisms. The reaction mechanism of LPMO is rather well-known and we study the self-oxidation mechanism and for the first time give a credible explanation to why one of the histidine ligands of the Cu ion is methylated in many LPMO enzymes. Future studies may investigate further self-oxidation products of LPMO.

For pMMO, the situation is very different. For this enzyme, there is not even any consensus regarding the location and nature of the active site. Therefore, we compare the reactivity of three putative active sites. Moreover, we compare the reactivity of the Cu_B site in pMMO with the active site in LPMO, which both have a histidine-brace motif. We show that the Cu_C site has the most favourable energetics. This is in agreement with recent bioinformatic and experimental studies that indicate that Cu_B probably is not the active site for methane oxidation. However, it is possible that it is instead used for the production of H_2O_2, which may be used in the true active site. Our results support such a suggestion, showing that formation and dissociation of H_2O_2 is favourable for the Cu_B site. Future studies, should also investigate the reactivity of the trinuclear Cu_D site, once a reliable structure can be found.

References

[1] B. M. Hoffman, D. Lukoyanov, Z.-Y. Yang, D. R. Dean, and L C Seefeldt. Mechanism of nitrogen fixation by nitrogenase: The next stage. *Chem. Rev.*, 114(8):4041–4062, 2014. doi: https://doi.org/10.1021/cr400641x.

[2] T. Spatzal, M. Aksoyoglu, L. Zhang, S. L. A. Andrade, E. Schleicher, D. C. Rees, and O. Einsel. Evidence for interstitial carbon in nitrogenase femo cofactor. *Science*, 334(6058):940, 2011. doi: 10.1126/science.1214025.

[3] O. Einsle and D. C. Rees. Structural enzymology of nitrogenase enzymes. *Chem. Rev.*, 120(12):4969–5004, 2020. doi: https://doi.org/10.1021/acs.chemrev.0c00067.

[4] G. Vaaje-Kolstad, B. Westereng, S. J. Horn, Z. Liu, M Sorlie, and V. G. H. Eijsink. An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science*, 330(6001):219–222, 2010. doi: 10.1126/science.1192231.

[5] P. V. Harris, D. Welner, K. C. McFarland, E. Re, J.-C. Navarro Poulsen, K. Brown, R. Salbo, H. Ding, E. Vlasenko, S. Merino, F. Xu, J. Cherry, S. Larsen, and L. Lo Leggio. Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: Structure and function of a large, enigmatic family. *Biochemistry*, 49(15):3305–3316, 2010. doi: https://doi.org/10.1021/bi100009p.

[6] K. K. Meier, S. M. Jones, T. Kaper, H. Hansson, M. J. Koetsier, S. Karkehabadi, E. I. Solomon, M. Sandgren, and B. Kelemen. Oxygen activation by cu lpmos in recalcitrant carbohydrate polysaccharide conversion to monomer sugars. *Chem. Rev*, 118(5):2593–2635, 2018. doi: https://doi.org/10.1021/acs.chemrev.7b00421.

[7] Z. Forsberg, M. Sorlie, D. Petrovic, G. Courtade, F. L. Aachman, G. Vaaje-Kolstad, B. Bissaro, A. Rohr, and V. G H. Eijsink. Polysaccharide degradation

by lytic polysaccharide monooxygenases. *Current Opinion in Structural Biology*, 59:54–64, 2019. doi: https://doi.org/10.1016/j.sbi.2019.02.015.

[8] T. Tandrup, K. E. H Frandsen, K. S. Johansen, J.-G. Berrin, and L. Lo Leggio. Recent insights into lytic polysaccharide monooxygenases (lpmos). *Biochem Soc Trans*, 46(6):1431–1447, 2018. doi: https://doi.org/10.1042/BST20170549.

[9] P. H. Walton and G. J. Davies. On the catalytic mechanisms of lytic polysaccharide monooxygenases. *Current Opinion in Chemical Biology*, 31:195–207, 2016. doi: https://doi.org/10.1016/j.cbpa.2016.04.001.

[10] M. A. Culpepper and A. C. Rosenzweig. Architecture and active site of particulate methane monooxygenase. *Crit. Rev. Biochem. Mol. Biol.*, 47(6):483–492, 2012. doi: 10.3109/10409238.2012.697865.

[11] M. Merkx, D. A. Kopp, M. H. Sazinsky, M. H. Blazyk, J. Muller, and S. J. Lippard. Dioxygen activation and methane hydroxylation by soluble methane monooxygenase: A tale of two irons and three proteins. *Angew. Chemie Int. Ed.*, 40(15):2782–2807, 2012. doi: 10.1002/1521-3773(20010803)40:15<2782::AID-ANIE2782>3.0.CO;2-P.

[12] A. Messerschmidt. *Handbook of Metalloproteins*. Wiley, 2001.

[13] E. I. Solomon, D. E. Heppner, E. M. Johnston, J. W. Ginsbach, J. Ciera, M. Qayyum, M. T. Kieber-Emmons, C. H. Kjaergaard, R. G. Hadt, and L. Tian. Copper active sites in biology. *Chem. Rev.*, 114(7):3659–3853, 2014. doi: https://doi.org/10.1021/cr400327t.

[14] F. Weigend and R. Ahlrichs. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for h to rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys*, 7:3297–3305, 2005. doi: DOIhttps://doi.org/10.1039/B508541A.

[15] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136(3B): B864–B871, 1964.

[16] T. Ziegler. Density functional theory as a practical tool in studies of organometallic energetics and kinetics. beating the heavy metal blues with dft. *Can. J. Chem.*, 73(6):743–761, 1995. doi: https://doi.org/10.1139/v95-095.

[17] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140(4A):A1133–A1138, 1965. doi: http://dx.doi.org/10.1103/PhysRev.140.A1133.

[18] J. P. Perdew, K. Burke, and M. Ermzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 78:3865, 1996. doi: https://doi.org/10.1103/PhysRevLett.77.3865.

[19] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev.*, 38(6):3098–3100, 1988. doi: https://doi.org/10.1103/PhysRevA.38.3098.

[20] J. P. Perdew. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev.*, 33(12):8822–8824, 1988. doi: https://doi.org/10.1103/PhysRevB.33.8822.

[21] C. Lee and R. G. Parr. Development of the colle-salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B*, 37(2):785–789, 1988. doi: https://doi.org/10.1103/PhysRevB.37.785.

[22] J. D. Chai and M. Head-Gordon. Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys.*, 128(8):084106, 2008. doi: 10.1063/1.2834918.

[23] S. Grimme. Semiempirical gga-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.*, 27(15):1787–1799, 2006. doi: https://doi.org/10.1002/jcc.20495.

[24] J. Sun, A. Ruzsinsky, and J .P. Perdew. Strongly constrained and appropriately normed semilocal density functional. *Physical Review Letters*, 115(3):03642, 2015. doi: https://doi.org/10.1103/PhysRevLett.115.036402.

[25] J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria. Climbing the density functional ladder: Nonempirical meta–generalized gradient approximation designed for molecules and solids. *Phys. Rev. Lett.*, 91(14):146401, 2003. doi: https://doi.org/10.1103/PhysRevLett.91.146401.

[26] Y. Zhao and D. G. Truhlar. A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. *J. Chem. Phys*, 125(19):194101, 2006. doi: https://doi.org/10.1063/1.2370993.

[27] H. S. Yu, X. He, S. L. Li, and D. G. Truhlar. Mn15: A kohn–sham global-hybrid exchange–correlation density functional with broad accuracy for multi-reference and single-reference systems and noncovalent interactions. *Chem. Sci*, 7:5032–505, 2016. doi: https://doi.org/10.1039/C6SC00705H.

[28] V. N Staroverov, G. E. Scuseria, J. Tao, and J. P. Perdew. Comparative assessment of a new nonempirical density functional: Molecules and hydrogen-bonded complexes. *J. Chem. Phys.*, 119:12129–37, 2003. doi: https://doi.org/10.1063/1.1626543.

[29] O. Salommon, S. Reiher, and B. A. Hess. Assertion and validation of the performance of the b3lyp∗ functional for the first transition metal row and the g2 test set. *The Journal of Chemical Physics*, 117(10):4729–4737, 2002. doi: https://doi.org/10.1063/1.1493179.

[30] A. D. Becke. Density-functional thermochemistry. iii. the role of exact exchange. *J. Chem. Phys.*, 98:5648–5652, 1993. doi: https://doi.org/10.1063/1.464913.

[31] K. Hui and J.-D. Chai. Scan-based hybrid and double-hybrid density functionals from models without fitted parameters. *The Journal of Chemical Physics*, 144(4):044114, 2016. doi: https://doi.org/10.1063/1.4940734.

[32] A. V. Krukau, O. A. Vydrov, and A.-F. Izmaylov. Influence of the exchange screening parameter on the performance of screened hybrid functionals. *The Journal of Chemical Physics*, 125(22):224106, 2006. doi: https://doi.org/10.1063/1.2404663.

[33] J. P. Perdew, M. Ernzerhof, and K. Burka. Rationale for mixing exact exchange with density functional approximations. *J. Chem. Phys.*, 105:9982, 1996. doi: https://doi.org/10.1063/1.472933.

[34] Y. Zhao and D. G. Truhlar. The m06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: Two new functionals and systematic testing of four m06-class functionals and 12 other functionals. *Theor. Chem. Acc.*, 120:215–241, 2008. doi: 10.1007/s00214-007-0310-x.

[35] Y. Zhao and D. G. Truhlar. Design of density functionals that are broadly accurate for thermochemistry, thermochemical kinetics, and nonbonded interactions. *J. Phys. Chem. A*, 109(25):5656–5667, 2005. doi: https://doi.org/10.1021/jp050536c.

[36] J. E. Moussa, P. A. Schultz, and J. R. Chelikowsky. Analysis of the heyd-scuseria-ernzerhof density functional parameter space. *The Journal of Chemical Physics*, 136(20):204117, 2012. doi: https://doi.org/10.1063/1.4722993.

[37] A. D. Becke. A new mixing of hartree-fock and local density-functional theories. *J. Chem. Phys*, 98:1372, 1993. doi: https://doi.org/10.1063/1.464304.

[38] Y. Zhao and D. G. Truhlar. Density functional for spectroscopy: no long-range self-interaction error, good performance for rydberg and charge-transfer states, and better performance on average than b3lyp for ground states. *J. Phys. Chem. A*, 110(49):13126–13130, 2006. doi: https://doi.org/10.1021/jp066479k.

[39] J. P. Ryckaert, G. Cigotti, and H. J. C. Berendesen. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys*, 23(3):327–341, 1977. doi: https://doi.org/10.1016/0021-9991(77)90098-5.

[40] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys*, 79(2):926–935, 1983. doi: https://doi.org/10.1063/1.445869.

[41] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma. The missing term in effective pair potentials. *J. Phys. Chem*, 91(24):6269–6271, 1987. doi: https://doi.org/10.1021/j100308a038.

[42] T. Darden, D. York, and L. Pedersen. Particle mesh ewald: An nlog(n) method for ewald sums in large systems. *J. Chem. Phys.*, 98(12):10089, 1987. doi: https://doi.org/10.1063/1.464397.

[43] A. Warshel and M. Levitt. *J. Mol. Biol.*, 103(2):227–249, 1976. doi: https://doi.org/10.1016/0022-2836(76)90311-9.

[44] H. M. Senn and W. Thiel. Qm/mm methods for biomolecular systems. *Angew. Chemie Int. Ed.*, 48(7):1198–1229, 2009. doi: https://doi.org/10.1002/anie.200802019.

[45] U. Ryde. Chapter six - qm/mm calculations on proteins. *Methods Enzymol.*, 577:119–158, 2016. doi: https://doi.org/10.1016/bs.mie.2016.05.014.

[46] N. Reuter, A. Dejaegare, B. Maigert, and M. Karplus. Frontier bonds in qm/mm methods: A comparison of different approaches. *J. Phys. Chem. A*, 104(8):1720–1735, 2000. doi: https://doi.org/10.1021/jp9924124.

[47] U. Ryde. The coordination of the catalytic zinc ion in alcohol dehydrogenase studied by combined quantum-chemical and molecular mechanics calculations. *J. Comput. Aided. Mol. Des.*, 10(2):153–164, 1996. doi: 10.1007/BF00402823.

[48] U. Ryde and M. H. M. Olsson. Structure, strain, and reorganization energy of blue copper models in the protein. *Int. J. Quantum Chem.*, 81(5):335–347, 2001. doi: https://doi.org/10.1002/1097-461X(2001)81:5<335::AID-QUA1003>3.0.CO;2-Q.

[49] R. N. Thorneley and D. J. Lowe. Kinetics and mechanism of the nitrogenase enzyme system. *Molybdenum enzymes*, 7:89–116, 1985.

[50] K. M. Lancaster, M. Roemelt, P. Ettenhuber, Y. Hu, M. W. Ribbe, F. Neese, U. Bergmann, and S. DeBeer. X-ray emission spectroscopy evidences a central carbon in the nitrogenase iron-molybdenum cofactor. *Science*, 334(6058):974–7, 2001. doi: 10.1126/science.1206445.

[51] R. Bjornsson, F. A. Lima, T. Spatzal, T. Weyhermuller, P. Glatzel, E. Bill, O. Einsle, F. Neese, and D. DeBeer. Identification of a spin coupled mo(iii) in the nitrogenase iron-molybdenum cofactor. *Chem. Sci.*, 5:3096, 2014. doi: https://doi.org/10.1039/C4SC00337C.

[52] B. Burgess and D. J. Lowe. Mechanism of molybdenum nitrogenase. *Chem. Rev.*, 96(7):2983–3012, 1996. doi: https://doi.org/10.1021/cr950055x.

[53] B. Schmid, H.-J. Chiu, V. Ramakrishnan, Jb B. Howard, and D. C. Rees. *Nitrogenase Handbook of Metalloproteins pp 1025-1036*. New York: Wiley, 2006.

[54] R. Y. Igarashi, M. Laryukhin, P. C. Dos Santos, H.-I. Lee, D. R. Dean, L. C. Seefeldt, and B. M. Hoffman. Trapping H bound to the nitrogenase femo-cofactor active site during H_2 evolution: characterization by endor spectroscopy. *J. Am. Chem. Soc.*, 127(17):6231–6241, 2005. doi: https://doi.org/10.1021/ja043596p.

[55] V. Hoeke, L. Tociu, D. A. Case, L. C. Seefeldt, S. Raugei, and B. M. Hoffman. High-resolution endor spectroscopy combined with quantum chemical calculations reveals the structure of nitrogenase janus in- termediate $E_4(4H)$. *J. Am. Chem. Soc.*, 141(30):11984–11996, 2019. doi: https://doi.org/10.1021/jacs.9b04474.

[56] L. E. Roth and F. A. Tezcan. X-ray crystallography. *Methods. Mol. Biol.*, 766: 147–164, 2011. doi: 10.1007/978-1-61779-194-9_10.

[57] C. Van Stappen, L. Decambs, G. E. Cutsail, R. Bjornsson, J. T. Henthorn, J. A. Birrell, and S. DeBeer. The spectroscopy of nitrogenases. *Chem. Rev.*, 120(12): 5005–5081, 2020. doi: https://doi.org/10.1021/acs.chemrev.9b00650.

[58] P. P. Hallmen and J. Kastner. N_2 binding to the femo-cofactor of nitrogenase. *Z. Anorg. Allg. Chem.*, 641(1):118–122, 2015. doi: https://doi.org/10.1002/zaac.201400114.

[59] I. Dance. Activation of N_2 the enzymatic way. *Z. Anorg. Allg. Chem.*, 641(1): 91–99, 2015. doi: https://doi.org/10.1002/zaac.201400120.

[60] J. B. Varley, Y. Wang, K. Chan, F. Studt, and J. K. Norskov. Mechanistic insights into nitrogen fixation by nitrogenase enzymes. *Phys. Chem. Chem. Phys.*, 17(1): 29541–29547, 2015. doi: https://doi.org/10.1039/C5CP04034E.

[61] P. E. M. Siegbahn. Model calculations suggest that the central carbon in the femo-cofactor of nitrogenase becomes protonated in the process of nitrogen fixation. *J. Am. Chem. Soc.*, 138(33):10485–10495, 2016. doi: https://doi.org/10.1021/jacs.6b03846.

[62] L. M. McKee. A new nitrogenase mechanism using a CFe_8S_9 model: does H_2 elimination activate the complex to N_2 addition to the central carbon atom? *J. Phys. Chem. A*, 120(5):754–764, 2016. doi: https://doi.org/10.1021/acs.jpca.5b10384.

[63] L. Rao, X. Xu, and C. Adamo. Theoretical investigation on the role of the central carbon atom and close protein environment on the nitrogen reduction in mo nitrogenase. *ACS Catal.*, 6(3):1567–1577, 2016. doi: https://doi.org/10.1021/acscatal.5b02577.

[64] S: Raugei, L. C. Seefeldt, and B. M. Hoffman. Critical computational analysis illuminates the reductive-elimination mechanism that activates nitrogenase for N_2 reduction. *Proc. Natl Acad. Sci.*, 115(45):10521–10530, 2018. doi: https://doi.org/10.1073/pnas.1810211115.

[65] I. Dance. Computational investigations of the chemical mechanism of the enzyme nitrogenase. *ChemBioChem*, 21(12):1671–1709, 2020. doi: https://doi.org/10.1002/cbic.201900636.

[66] F. Tuczek. *Nitrogen fixation in nitrogenase and related small-molecule models: results of DFT calculations RSC Metallobiology Series 7 ed R Hille, C Schulzke and M L Kirk pp 223-274*. Cambridge: Royal Society of Chemistry, 2017.

[67] I. Dance. Theoretical investigations of the mechanism of biological nitrogen fixation at the femo cluster site. *J. Biol. Inorg. Chem*, 1:581–586, 1996. doi: https://doi.org/10.1007/s007750050096.

[68] K. K. Stavrev and M. C. Zerner. Studies on the hydrogenation steps of the nitrogen molecule at the azotobacter vinelandii nitrogenase site. *J. Quantum Chem.*, 70(6):1159–1168, 1998. doi: https://doi.org/10.1002/(SICI)1097-461X(1998)70:6<1159::AID-QUA5>3.0.CO;2-Z.

[69] P. E. M. Siegbahn, J. Westerberg, M. Svensson, and R. H. Crabtree. Nitrogen fixation by nitrogenases: a quantum chemical study. *J. Phys. Chem. B*, 102(9): 1615–1623, 1998. doi: https://doi.org/10.1021/jp972207t.

[70] T. Lovell, J. Li, T. Liu, D. A. Case, and L. Noodleman. Femo cofactor of nitrogenase: a density functional study of states MN, M^{OX}, M^{R} and M^{I}. *J. Am. Chem. Soc.*, 123(49):12392–12410, 2001. doi: https://doi.org/10.1021/ja011860y.

[71] H. Xie, R. Wu, Z. Zhou, and Z. Cao. Exploring the interstitial atom in the femo cofactor of nitrogenase: insights from qm and qm/mm calculations. *J. Phys. Chem. B*, 112(36):11435–11439, 2008. doi: https://doi.org/10.1021/jp803616z.

[72] J. Kastner and P. E. Blochl. Ammonia production at the femo cofactor of nitrogenase: results from density functional theory. *J. Am. Chem. Soc.*, 129(10): 2998–3006, 2007. doi: https://doi.org/10.1021/ja068618h.

[73] M. Rohde, Sippel. D., C. Trncik, S. L. A. Andrade, and O. Einsle. The critical e(4) state of nitrogenase catalysis. *Biochemistry*, 57(38):5497–5504, 2018. doi: https://doi.org/10.1021/acs.biochem.8b00509.

[74] A. T. Thorhallsson, B. Benediktsson, and R. Bjornsson. A model for dinitrogen binding in the e(4) state of nitrogenase. *Chem. Sci.*, 10:11110–11124, 2019. doi: https://doi.org/10.1039/C9SC03610E.

[75] P. E. M. Siegbahn. Model calculations suggest that the central carbon in the femo-cofactor of nitrogenase becomes protonated in the process of nitrogen fixation. *J Am Chem Soc*, 138:1533747–15759, 2016. doi: 10.1021/jacs.6b03846.

[76] P. E. M. Siegbahn. Is there computational support for an unpro- tonated carbon in the e(4) state of nitrogenase? *J. Comput. Chem.*, 39(12):743–747, 2018. doi: https://doi.org/10.1002/jcc.25145.

[77] P. E. M. Siegbahn. The mechanism for nitrogenase including all steps. *Phys. Chem. Chem. Phys.*, 21:15747–15759, 2019. doi: https://doi.org/10.1039/C9CP02073J.

[78] L. Cao, O. Caldararu, and U. Ryde. Protonation and reduction of the femo cluster in nitrogenase studied by quantum mechanics/molecular mechanics (qm/mm) calculations. *J. Chem. Theory Comput.*, 14:6653–6678, 2018. doi: https://doi.org/10.1021/acs.jctc.8b00778.

[79] R. J. Quinlan, M. D. Sweeny, L. L. Leggio, H. Otten, J. C. N. Poulsen, K. S. Johansen, K. B. R. M. Krogh, C. I. Jorgensen, M. Tovborg, A. Anthonsen, T. Tryfona, C. P. Walter, P. Durpee, F. Xu, G. J. Davise, and P. H. Walton. Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proceedings of the National Academy of Sciences of the United States of America*, 108(37):15079–15084, 2011. doi: https://doi.org/10.1073/pnas.1105776108.

[80] G. R. Hemsworth, B. Henrissat, G. J. Davies, and P. H. Walton. Discovery and characterization of a new family of lytic polysaccharide monooxygenases. *Nature Chemical Biology*, 10(2):122–126, 2014. doi: 10.1038/nchembio.1417.

[81] W. T. Vu, V. abd Beeson, E. A. Span, E. R. Fraquar, and M. A. Marletta. A family of starch-active polysaccharide monooxygenases. *Proceedings of the National Academy of Sciences*, 111(38):13822–13827, 2014. doi: https://doi.org/10.1073/pnas.1408090111.

[82] L. Leggio, T. J. Simmons, J. C. N. Poulsen, K. E. H. Frandsen, G. R. Hemsworth, M. A. Stringer, P. von Freisesleben, M. Tovborg, K. S. Johansen, L. De Maria, P. V. Harris, C. L. Soong, P. Dupree, T. Tryfona, N. Lenfant, B. Henrissat, G. J. Davies, and P. H. Walton. Structure and boosting activity of a starch-degrading lytic polysaccharide monooxygenase. *Nature Communications*, 6(1):5961, 2015. doi: 10.1038/ncomms6961.

[83] M. Couturier, S. Ladeveze, G. Sulzenbacher, L. Ciano, M. Faunel, C. Moreau, A. Villares, B. Cathala, F. Chaspoul, K. E. Frandsen, A: Labourel, I. Herpoel-Gimbert, S. Grisel, M. Haon, N. Lenfant, Roginaux H., D. Ropartz, G. J. Davies, M.-N. Rosso, P. H. Walton, B. Henrissat, and J.-G. Berrin. Lytic xylan oxidases from wood-decay fungi unlock biomass degradation. *Nature Chemical Biology*, 14(3):306–310, 2018. doi: 10.1038/nchembio.2558.

[84] F. Sabbadin, G. R. Hemsworth, L. Ciano, B. Henrissat, P. Dupree, T. Tryfona, R. D. S. Marques, S. T. Sweeney, K. Besser, L. Elias, G. Pesante, Y. Li, A. A. Dowle, R. Bates, L. D. Gomez, R. Simister, G. J. Davies, P. H. Walton, N. C. Bruce, and S. J. McQueen-Mason. An ancient family of lytic polysaccharide monooxygenases with roles in arthropod development and biomass digestion. *Nature Chemical Biology*, 9:756, 2018. doi: 10.1038/s41467-018-03142-x.

[85] C. Filiatrault-Chastel, D. Navarro, M. Haon, S. Grisel, I. Herpoel-Gimbert, D. Chevret, M. Faunel, B. Henrissat, S. Heiss-Blanquet, A. Margeot, and J.-G. Berrin. Aa16, a new lytic polysaccharide monooxygenase family identified in fungal secretomes. *Biotechnology for Biofuels*, 12:55, 2019. doi: 10.1186/s13068-019-1394-y.

[86] B. Bissaro, A. K. Rohr, G. Muller, P. Chylenski, M. Skaugen, S. Forsberg, Hornm S. J., G. Vaaje-Kolstad, and V. G. H. Eijsink. Oxidative cleavage of polysaccharides by monocopper enzymes depends on H_2O_2. *Nature Chemical Biology*, 13(10):1123–1128, 2017. doi: 10.1038/nchembio.2470.

[87] J. A. Hangasky and M. A. Marletta. A random-sequential kinetic mechanism for polysaccharide monooxygenases. *Biochemistry*, 57(22):1391–3199, 2018. doi: 10.1021/acs.biochem.8b00129.

[88] S. Kuusk, B. Bissaro, P. Kuusk, Z. Forsberg, V. G. H. Eijsink, M. Sorlie, and P. Valjamae. Kinetics of H_2O_2-driven degradation of chitin by a bacterial lytic polysaccharide monooxygenase. *J. Biol. Chem.*, 193(2):523–531, 2018. doi: 10.1074/jbc.M117.817593.

[89] R. Kont, B. Bissaro, V. G. H. Eijsink, and P. Valjamae. Kinetic insights into the peroxygenase activity of cellulose-active lytic polysaccharide monooxygenases (lpmos). *Nature, Communications*, 11(1):5786, 2020. doi: 10.1038/s41467-020-19561-8.

[90] E. A. Span and M. A. Marletta. The framework of polysaccharide monooxygenase structure and chemistry. *Current Opinion in Strucural Biology*, 35:93–99, 2015. doi: https://doi.org/10.1016/j.sbi.2015.10.002.

[91] G. Courtade, L. Ciano, A. Paradisi, P. J. Lindley, Z. Forsberg, M. Sorlie, R. Wimmer, G. J. Davies, V. G. H Eijsink, P. H. Walton, and F. L. Aachman. Mechanistic basis of substrate-O_2 coupling within a chitin-active lytic polysaccharide monooxygenase: An integrated nmr/epr study. *Proceedings of the National Academy of Sciences,*, 117(32):19178–19189, 2020. doi: https://doi.org/10.1073/pnas.2004277117.

[92] Y. A. Theibich, S. P. A. Sauer, L. L. Leggio, and E. D. Hedegard. Estimating the accuracy of calculated electron paramagnetic resonance hyperfine couplings for a lytic polysaccharide monooxygenase. *Computational and Structural Biotechnology Journal*, 19:555–567, 2021. doi: 10.1016/j.csbj.2020.12.014.

[93] B. Bissaro, I. Isaksen, G. Vaaje-Kolstad, V. G. H. Eijsink, and A. K. Rohr. How a lytic polysaccharide monooxygenase binds crystalline chitin. *Biochemistry*, 57(12):1893–1906, 2018. doi: https://doi.org/10.1021/acs.biochem.8b00138.

[94] S. Kim, J. Stahlberg, M. Sandgren, and R. S. Paton. Quantum mechanical calculations suggest that lytic polysaccharide monooxygenases use a copper-oxyl, oxygen-rebound mechanism. *Proceedings of the National Academy of Sciences*, 111(1):149–154, 2013. doi: https://doi.org/10.1073/pnas.1316609111.

[95] E. D. Hedegard and U. Ryde. Targeting the reactive intermediate in polysaccharide monooxygenases. *J. Biol. Inorg. Chem.*, 22:1029–1037, 2017. doi: https://doi.org/10.1007/s00775-017-1480-1.

[96] L. Bertini, R. Breglia, M. Lamburghi, P. Fantucci, L. De Gioia, M. Borsari, M. Sola, C. A. Bortolotti, and M. Bruschi. Catalytic mechanism of fungal lytic polysaccharide monooxygenases investigated by first-principles calculations. *Inorg. Chem.*, 57(1):86–97, 2018. doi: https://doi.org/10.1021/acs.inorgchem.7b02005.

[97] B. Wang, E. M. Johnston, P. Li, S. Shaik, G. J. Davies, P. H. Walton, and C. Rovira. Qm/mm studies into the H_2O_2-dependent activity of lytic polysaccharide monooxygenases: Evidence for the formation of a caged hydroxyl radical intermediate. *ACS Catal.*, 8(2):1346–1351, 2018. doi: https://doi.org/10.1021/acscatal.7b03888.

[98] E. D. Hedegard and U. Ryde. Molecular mechanism of lytic polysaccharide monooxygenases. *Chem. Sci.*, 9:3866–3880, 2018. doi: https://doi.org/10.1039/C8SC00426A.

[99] B. Bissaro, B. Streit, I. Isaksen, V. G. H. Eijsink, G. T. Beckham, J. L. DuBois, and A. K. Rohr. Molecular mechanism of the chitinolytic peroxygenase reaction. *Proceedings of the National Academy of Sciences*, 117(3):1504–1513, 2020. doi: https://doi.org/10.1073/pnas.1904889117.

[100] D. M. Petrovic, B. Bissaro, P. Chylenski, M. Skaugen, M. Sorlie, M. S. Jensen, F. L. Aachman, G. Courtade, A. Varnai, and V. G. H. Eijsink. Methylation of the n-terminal histidine protects a lytic polysaccharide monooxygenase from auto-oxidative inactivation. *Protein Sci.*, 27(9):1636–1650, 2018. doi: 10.1002/pro.3451.

[101] N. Elango, R. Radhakrishnan, W. A. Froland, B. J. Wallar, C. A. Earhart, J. D. Lipscomb, and D. H. Ohlendorf. Crystal structure of the hydroxylase component of methane monooxygenase from methylosinus trichosporium ob3b. *Angew. Chemie Int. Ed.*, 6(3):556–568, 1997. doi: 10.1002/pro.5560060305.

[102] A. Rosenzweig, C. Frederick, S. Lippard, and P. Nordlund. Crystal structure of a bacterial non-haem iron hydroxylase that catalyses the biological oxidation of methane. *Nature*, 366:537–543, 1993. doi: https://doi.org/10.1038/366537a0.

[103] M. M. Martinho, D. W. Choi, A. A. DiSpirito, W. E. Antholine, J. D. Semaru, and E. Munck. Mossbauer studies of the membrane-associated methane monooxygenase from methylococcus capsulatus bath: evidence for a diiron center. *J. Am. Chem. Soc.*, 129(51):15783–15785, 2007. doi: https://doi.org/10.1021/ja077682b.

[104] R. L. Lieberman, Shrestha D. B., P. D. Doan, B. M. Hoffman, T. L. Stemmler, and A. C. Rosenzweig. Purified particulate methane monooxygenase from methylococcus capsulatus (bath) is a dimer with both mononuclear copper and a copper-containing cluster. *Proc. Natl. Acad. Sci. U. S. A.*, 100(7):3820–3825, 2003. doi: 10.1073/pnas.0536703100.

[105] A. C. Rosenzweig. The metal centres of particulate methane mono-oxygenase. *Biochem Soc. Trans.*, 36(6):1134–1137, 2008. doi: 10.1042/BST0361134.

[106] R. L. Lieberman, K. C. Kondapalli, Shrestha. D. B., A. S. Hakemian, S. M. Smith, J. Tesler, J. Kuzelka, R. Gupta, A. S. Borovik, S. J. Lippard, B. M. Hoffman, and A. C. Rosenzweig. Characterization of the particulate methane monooxygenase metal centers in multiple redox states by x-ray absorption spectroscopy. *Inorg. Chem.*, 45(20):8372–8381, 2006. doi: https://doi.org/10.1021/ic060739v.

[107] M. O. Ross, F. MacMillan, J. Wang, A. Nisthal, T. J. Lawton, B. D. Olafson, S. L. Mayo, A. C. Rosenzweig, and B. M. Hoffman. Particulate methane monooxygenase contains only mononuclear copper centers. *Science*, 364:566–570, 2019. doi: 10.1126/science.aav2572.

[108] S. Y. Ro, L. F. Schachner, C. W. Koo, R. Purohit, J. P. Remis, G. E. Kenney, B. W. Liauw, P. M. Thomas, S. M. Patrie, N. L. Kelleher, and et al. Native top-down mass spectrometry provides insights into the copper centers of membrane-bound methane monooxygenase. *Nat Commun*, 10(1):2675, 2019. doi: 10.1038/s41467-019-10590-6.

[109] R. L. Lieberman and A. C. Rosenzeweig. Crystal structure of a membrane-bound metalloenzyme that catalyses the biological oxidation of methane. *Nature*, 434(7030):177–182, 2005. doi: 10.1038/nature03311.

[110] V. C. C. Wang, S. Maji, P. P. Y. Chen, H. K. Lee, S. S. F. Yu, and S. I. Chan. Alkane oxidation: Methane monooxygenases, related enzymes, and their biomimetics. *Chemical Reviews*, 117:8574–8621, 2017. doi: 10.1021/acs.chemrev.6b00624.

[111] W. H. Chang, H. H. Lin, I. K. Tsai, S. H. Huang, S. C. Chung, I. P. Tu, S. F. F Yu, and S. I. Chan. Copper centers in the cryo-em structure of particulate methane monooxygenase reveal the catalytic machinery of methane oxidation. *J. Am. Chem. Soc.*, 143:9922–9932, 2021. doi: 10.1021/jacs.1c04082.

[112] S. I. Chan, V. V. Wang, J. C. Lai, S. S. Yu, P. P. Chen, C. L. Chen, and M. K. Chan. Redox potentiometry studies of particulate methane monooxygenase: support for a trinuclear copper cluster active site. *Angew. Chem. Int. Ed. Engl.*, 46(12):1992–1994, 2007. doi: 10.1002/anie.200604647.

[113] S. I. Chan, V. C. C. Wang, P. P. Y. Chen, and S. F. F. Yu. Methane oxidation by the copper methane monooxygenase: Before and after the cryogenic electron microscopy structure of particulate methane monooxygenase from methylococcus capsulatus. *Journal of the Chinese Chemical Society*, 69(8):1147–1158, 2022. doi: 10.1002/jccs.202200166.

[114] S. I. Chan, Chang W.-H., S.-H. Huang, H.-H. Lin, and S. S. F. Yu. Catalytic machinery of methane oxidation in particulate methane monooxygenase (pmmo). *Journal of Inorganic Biocemistry*, 2021(225):1147–1158, 2022. doi: 10.1016/j.jinorgbio.2021.111602.

[115] W. K. Koo, F. J. Tucci, Y. He, and A. C. Rosenzweig. Recovery of particulate methane monooxygenase structure and activity in a lipid bilayer. *Science*, 375: 1287–1291, 2022. doi: 10.1126/science.abm3282.

[116] W.-H. Chang, H.-H. Lin, I. K. Tsai, S.-H. Huang, S.-C. Chung, I.-P. Tu, S. S.-F. Yu, and S. I. Chan. Copper centers in the cryo-em structure of particulate methane monooxygenase reveal the catalytic machinery of methane oxidation. *J. Am. Chem. Soc.*, 143(26):9922–9932, 2021. doi: https://doi.org/10.1021/jacs.1c04082.

[117] L. Cao, O. Caldararu, A. C. Rosenzweig, and U. Ryde. Quantum refinement does not support dinuclear copper sites in crystal structures of particulate methane monooxygenase. *Angew. Chemie-Int. Ed.*, 57(1):162–166, 2018. doi: https://doi.org/10.1002/anie.201708977.

[118] K. Yoshizawa and Y. Shiota. Conversion of methane to methanol at the mononuclear and dinuclear copper sites of particulate methane monooxygenase (pmmo): a dft and qm/mm study. *J. Am. Chem. Soc.*, 128(30):9873–9881, 2006. doi: https://doi.org/10.1021/ja061604r.

[119] Y. Shiota and K. Yoshizawa. Comparison of the reactivity of bis(-oxo)$Cu^{II}Cu^{III}$ and $Cu^{III}Cu^{III}$ species to methane. *Inorg. Chem.*, 48(3):838–845, 2009. doi: https://doi.org/10.1021/ic8003933.

[120] W. Peng, X. Qu, S. Shaik, and B. Wang. Deciphering the oxygen activation mechanism at the cuc site of particulate methane monooxygenase. *Nature Catalysis*, 4:266–273, 2021. doi: 10.1038/s41929-021-00591-4.

[121] A. Klamt and G. Schuurmann. Cosmo: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc, Perkin Trans*, 2:799–805, 1993. doi: https://doi.org/10.1039/P29930000799.

[122] A. Schafer, A. Klamt, D. Sattel, J. C. W. Lohrenz, and F. Eckert. Cosmo implementation in turbomole: extension of an efficient quantum chemical code towards liquid systems. *Phys. Chem. Chem. Phys*, 2:2187–2193, 2000. doi: https://doi.org/10.1039/B000184H.

[123] K. Andersson, P. A. Malmqvist, and B Roos. Second-order perturbation theory with a complete active space self-consistent field reference function. *The Journal of Chemical Physics*, 96(2):1218–1226, 1992. doi: 10.1063/1.462209.

[124] K. Andersson, P. A. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski. Second-order perturbation theory with a casscf reference function. *The Journal of Physical Chemistry*, 94(14):5483–5488, 2002. doi: 10.1021/j100377a012.

[125] A. A. Holmes, N. M. Tubman, and C. J. and Umriga. Heat-bath configuration interaction: An efficient selected configuration interaction algorithm inspired by heat-bath sampling. *J. Chem. Theory Comput.*, 12(8):3674–3680, 2016. doi: 10.1021/acs.jctc.6b00407.

[126] S. Sharma, A. A. Holmes, G. Jeanmairet, A. Alavi, and C. J. Umrigar. Semistochastic heat-bath configuration interaction method: Selected configuration interaction with semistochastic perturbation theory. *Journal of Chemical Theory and Computation*, 13(4):1595–1604, 2017. doi: 10.1021/acs.jctc.6b01028.

[127] J. Li, M. Otten, A. A. Holmes, S. Sharma, and C. J. Umrigar. Fast semistochastic heat-bath configuration interaction. *The Journal of Chemical Physics*, 149(21): 214110, 2018. doi: 10.1063/1.5055390.

[128] L. Cao and U. Ryde. Extremely large differences in dft energies for nitrogenase models. *Phys. Chem. Chem. Phys*, 21(5):2480–2488, 2019. doi: https://doi.org/10.1039/C8CP06930A.

[129] B. Benediktsson and R. Bjornsson. Analysis of the geometric and electronic structure of spin-coupled iron-sulfur dimers with broken-symmetry dft: Implications for femoco. *J. Chem. Theory Comput.*, 18(3):1437–1457, 2022. doi: https://doi.org/10.1021/acs.jctc.1c00753.

Scientific publications

Author contributions

Paper I: Comparison of the accuracy of DFT methods for reactions with relevance to nitrogenase

I did all the calculations and participated in the writing of the manuscript.

Paper II: Calibration of DFT functionals for a minimal nitrogenase $[Fe(SH)_4H]^-$ model employing state-of-the-art ab initio methods

I did some preliminary coupled cluster calculations and did all the DFT calculations. I participated in the writing of the manuscript.

Paper III: Histidine oxidation in lytic polysaccharide monooxygenase

I did all the calculations. I participated in the writing of the manuscript.

Paper IV: Computational study of three metal sites in particulate methane monooxygenase

I did the setup of the cryo-EM structure as well as all QM/MM calculations on Cu sites C and C2. I participated in the writing of the manuscript.

Printed by Media-Tryck, Lund 2023 NORDIC SWAN ECOLABEL 3041 0903

ISBN: 978-91-7422-938-7

Theoretical Chemistry
Department of Chemistry
Faculty of Science
Lund University