New insight into the formation of ammonia in nitrogenase

Albert Þór Þórhallsson

Faculty of Physical Sciences
University of Iceland
2016

NEW INSIGHT INTO THE FORMATION OF AMMONIA IN NITROGENASE

Albert Þór Þórhallsson

15 ECTS thesis submitted in partial fulfillment of a
Baccalaureus Scientiarum degree in Chemistry

Advisor
Ragnar Björnsson

Co-advisor
Egill Skúlason

Faculty of Physical Sciences
School of Engineering and Natural Sciences
University of Iceland
Reykjavik, June 2016

New insight into the formation of ammonia in nitrogenase
Nitrogenase ammonia formation
15 ECTS thesis submitted in partial fulfillment of a B.Sc. degree in Chemistry

Copyright © 2016 Albert Þór Þórhallsson
All rights reserved

Faculty of Physical Sciences
School of Engineering and Natural Sciences
University of Iceland
VRII, Hjarðarhagi 2–6
107, Reykjavik
Iceland

Telephone: 525 4000

Bibliographic information:
Albert Þór Þórhallsson, 2016, New insight into the formation of ammonia in nitrogenase, B.Sc. thesis, Faculty of Physical Sciences, University of Iceland.

Reykjavik, Iceland, June 2016

Abstract

The majority of the atmosphere is inert N_2, which has to be converted to NH_3 to be usable for organism. This conversion is achieved by diazotrophs that contain the enzyme nitrogenase. The system has been studied for a long time yet the mechanism is still unknown. A QM/MM model of the MoFe protein of nitrogenase was used to explore mechanistic aspects of substrate reduction at the FeMo cofactor (FeMoco). The project consisted of five parts:

- The spin coupling of the resting state of the protein
- Carbon monoxide inhibition of the protein
- The structure of the singly reduced FeMoco
- Structural aspects of the P-cluster
- Hydrazine reduction by the enzyme

Broken-symmetry solutions of the resting state of FeMoco were explored with different density functionals (BP86, TPSSh and B3LYP). Lowest energy broken-symmetry solutions were in agreement with previous work. Substrate adducts were found to influence spin states energies, and the CO inhibited FeMoco features a spin configuration that differs from the resting state. We also explored structural aspects of the singly reduced state of FeMoco with promising preliminary results. Investigations of the P-cluster led to a QM/MM optimized structure in acceptable agreement with the crystal structure. Attempts to calculate the redox potentials revealed sensitivity with respect to the broken-symmetry solutions that needs to be resolved by future calculations. Lastly, binding and reduction of hydrazine was explored at both Mo and Fe atoms and compared to synthetic MoFeS cubane clusters.

Útdráttur

Meirihluti andrúmslofts er óhvarfgjarnt N_2, sem nýtist ekki lífverum nema því sé umbreytt í NH_3. Þessi umbreyting er framkvæmd af diazotroph bakteríum sem innihalda ensímið nítrógenasa. Þetta kerfi hefur verið mikið rannsakað án þess að virkni þess sé að fullu útskýrð. QM/MM líkan af MoFe próteini var notað til að kanna virkni þess í afoxun hvarfefna með FeMo hjálparþættinum (FeMoco). Þetta verkefni samanstóð af fimm hlutum:

- Spunakúplanir grunnástand FeMoco
- Hindraður kolmónoxíð FeMoco
- Bygging á einnar rafeindar afoxuðum FeMoco
- Byggingareiginleikar P-klasans
- Afoxun hýdrasín með ensíminu

Mismunandi lausnir með rofna spunasamhverfu fyrir grunnástand FeMoco voru kannaðar með nokkrum þéttnifellum (BP86, TPSSh og B3LYP). Orkulægsta lausnin var í góðu samræmi við fyrri rannsóknir. Bundin hvarfefni reyndust hafa áhrif á orkur spunaástands og CO hindraður FeMoco var með spunamynstur sem var frábrugðið grunnástandinu. Einnig könnuðum við byggingareiginleika á einnar rafeindar afoxuðum FeMoco. Þær niðurstöður lofa góðu.Rannsókn á P-klasanum gaf QM/MM lágmarkaða byggingu í ágætu samræmi við þekkta kristalbyggingu; tilraunir til að reikna afoxunarmætti klasans, leiddu í ljós næmni fyrir spunalausnum sem krefst frekari reikninga. Að endingu var afoxun hýdrasíns könnuð bæði á Mo og Fe atómum FeMoco og borin saman við MoFeS kúbanklasa.

Contents

List of Figures

List of Tables

Abbreviations

ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
B3LYP	Becke '88 exchange and Lee-Yang-Parr correlation with 20% HF exchange
BO	Born-Oppenheimer
BP86	Becke '88 exchange and Perdew '86 correlation
BS	Broken symmetry
COSMO	Continuum solvation model
DFT	Density functional theory
ENDOR	Electron nuclear double resonance
EPR	Electron paramagnetic resonance
ET	Electron transfer
EXAFS	Extended X-Ray Absorption Fine Structure
Fd	Ferredoxin
GGA	Generalized Gradient Approximation
HF	Hartree—Fock
HK	Hohenberg-Kohn
HS	Highspin
KS	Kohn-Sham
LDA	Local Density Approximation
LT	Lowe—Thorneley
MD	Molecular dynamics
MM	Molecular Mechanics
MO	Molecular orbital
NEB	Nudge elastic band
PES	Potential Energy Surface
QM	Quantum mechanics
QM/MM	Quantum mechanics/Molecular Mechanics
RHF	Restricted Hartree—Fock
RKS	Restricted Kohn-Sham
SCF	Self-consistent field
SD	Slater-Dirac
SpReAD	Spatially resolved anomalous dispersion refinement
TPSSh	The exchange functional of Tao, Perdew, Staroverov, and Scuseria with 10% HF exchange
UHF	Unrestricted Hartree—Fock
UKS	Unrestricted Kohn-Sham
WFT	Wavefunction Theory
VP	Variational principle
XAS	X-ray absorption spectra

Acknowledgments

I would like to give special thanks to my advisor Ragnar Björnsson for copious amount of support and help with this research. I would also like to thank Egill Skúlason for putting aside time from his ever busy schedule to proofread this work.

I would also like to thank all the people that I have prompted to comment on this project, and have duly done so. Their feedback has surely increased the quality of this work. Special thanks go to Barði Benediktsson which work I am building on, and has also been of much help in the complex literature we have had to explore.

I would specifically like to thank the University of Iceland for the use of their computer clusters, SÓL, housed and serviced by VR-III, and GARÐAR, which is housed and serviced by Reiknistofnun Háskóla Íslands.

Finally, thanks and gratitude go to my family, which have supported me though the time of my studies in so many ways; without there support I could not have done much.

1. Introduction

1.1. Nitrogenase

Nitrogen fixation, the reduction of dinitrogen (N_2), is an essential component of nucleic acids and proteins, therefore all organisms require this nutrient for life. Even though almost 4 out of 5 atmospheric molecules are dinitrogen, the inertness of the elemental form requires a conversion to a usable form like ammonia (NH_3)[1]. The main component of fertilizer is ammonia and it is estimated that around 2% of the current energy production of the world is dedicated to making reduced nitrogen via the Haber-Bosch process ($N_2 + 3\,H_2 \longrightarrow 2\,NH_3$) and while that process, is extremely efficient, is an industrial endeavor, needing efficient transportation and not well suited to a small local production that is easily accessible by all farmers[2]. The carbon footprint of the Haber-Bosch process is also an issue, as most of the H_2 comes from natural gas, e.g. CH_4, and so produces large quantities of CO_2.

Nature "fixes" dinitrogen with the enzyme nitrogenase, found in bacteria, including *A. vinelandii* which has the most studied nitrogenase. Nitrogenase reduces dinitrogen at ambient pressure and temperature using only acidic protons and electrons from a electron-transfer protein e.g. ferredoxin (Fd) coupled with the energy currency of life, ATP, with the reaction:

$$N_2 + 16\,\mathrm{MgATP} + 10\,H^+ + 8\,e^- \longrightarrow 2\,{NH_4}^+ + 16\,\mathrm{MgADP} + 16\,P_i + H_2 \qquad (1.1)$$

The enzyme is composed of two separate metalloproteins, MoFe protein and Fe protein, and as such is one of the most interesting and complex metalloenzymes so far isolated. In fact the complex nature of nitrogenase and the daunting number of unanswered questions about the mechanism of nitrogenase activity has earned it the nickname "the Everest of enzymes"[3].

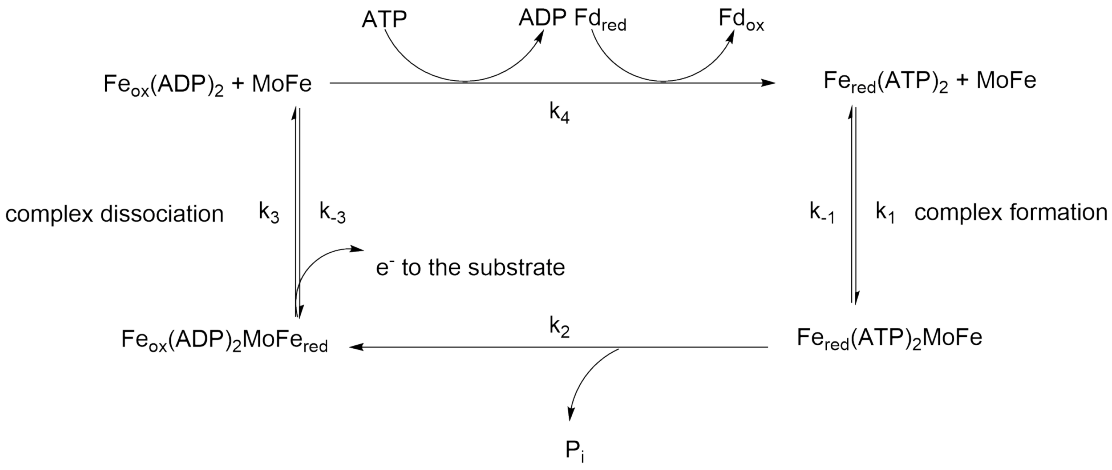


Figure 1.1: Scheme of the Fe protein cycle of nitrogenase adapted from Kästner[6]

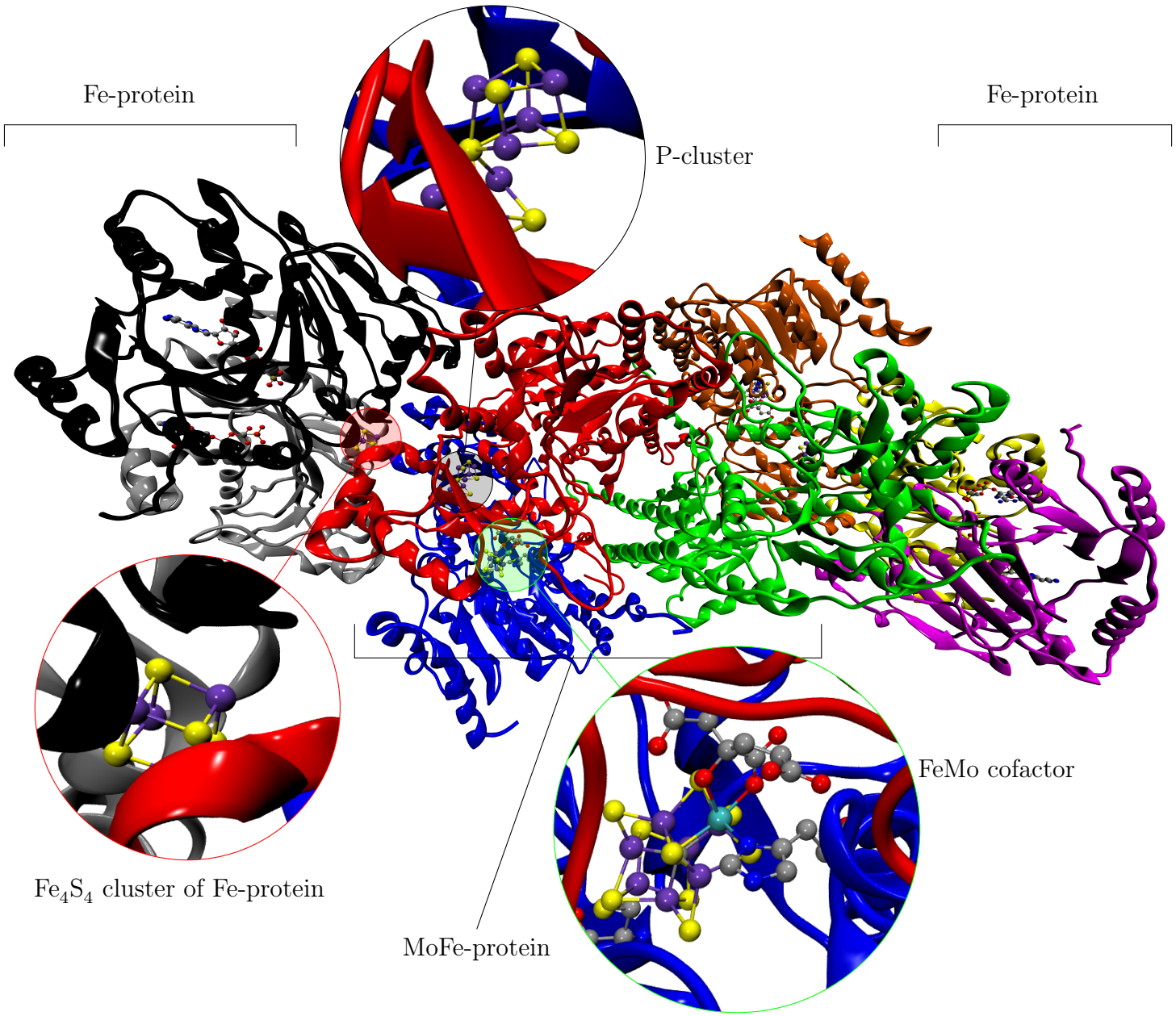
1.1.1. Fe-protein and MoFe protein

Nitrogenase consist of two metalloproteins, the Iron (Fe) protein and the molybdenum iron (MoFe) protein, named after the metal composition of each protein. Alternative systems, homologous to the MoFe system, may be induced if molybdenum is not available (e.g. vanadium nitrogenase). The MoFe protein is a $\alpha_2\beta_2$ tetramer, with each $\alpha\beta$-pair functioning as an catalytic unit, and each unit has the molecular weight of ≈ 120 kDa[4]. The $\alpha\beta$-unit of MoFe protein is the site of substrate reduction, and contains two unique metal clusters, the heterometallic iron-molybdenum cofactor (FeMoco) cluster, and the P-cluster, a homometallic iron sulfur cofactor, which is thought to mediate electron transfer from the Fe-protein to FeMoco. The Fe-protein is a homodimeric protein, folded with a single α/β type domain and contains an $[Fe_4S_4]$ cluster and is known to deliver electrons one-at-a-time to the MoFe protein[5]. The kinetic scheme in figure 1.1 has been proposed[1,6] for the substrate reduction in which these proteins are involved. The complex dissociation step is the rate limiting step and is also the last step of each cycle[5].

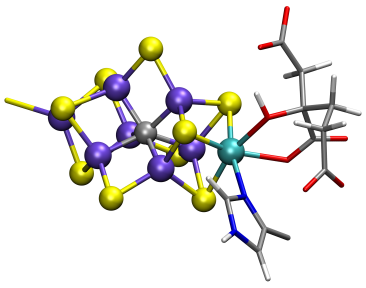
1.1.2. FeMo cofactor

FeMoco is the N_2-fixation catalyst of nitrogenase and can be isolated from the enzyme and retain some of its catalytic activity[7] but not for N_2 reduction. In the absence of dinitrogen, nitrogenase is an excellent hydrogen evolving catalyst[8].

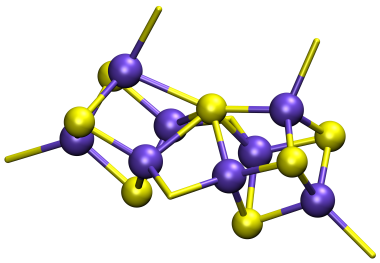
Substrate reduction depends on the electron flux through the nitrogenase enzyme because nitrogenase substrates are reduced by different numbers of electrons. Electron flux through the MoFe protein depends on the ratio of Fe-protein to MoFe protein. Ammonia formation from dinitrogen is favoured at high electron flux but hydrogen at low electron flux[1]. The hydrogen evolution interestingly can never be suppressed by high electron flux or high N_2 pressure and seems to be mandatory[9] for the reaction as indicated in equation 1.1. FeMoco consists of 7 irons, 1 molybdenum, 9 sulfides and a single carbon atom[10,11]. These atoms are shown as violet, cyan, yellow and grey respectively in figure 1.2(a) and 1.2(b). FeMoco is embedded in each α subunit of the MoFe protein as can be seen in the figure 1.2(a) and is ligated to the peptide matrix through one cysteine ligand bound to the unique tetrahedral iron at one end of the cluster and through one histidine ligand (His442, nitrogen shown as blue, carbon grey and hydrogen white) bound to the molybdenum atom at the other end of the cluster. The six irons at the middle part of the cluster are arranged as a prismatic structure around a hexagonally coordinated carbon atom[10,11] with each iron coordinated by three sulfide atoms. To satisfy an octahedral coordination, molybdenum is ligated by the bidentate R-homocitrate (see figure 1.2(b), oxygen is red and carbon and hydrogen are coloured as before).



(a) Nitrogenase



(b) FeMo cofactor



(c) P-cluster

Figure 1.2: (a) Nitrogenase and its eight subunits in the MoFe/Fe protein complex state. The blue, brown and red, green area are the α and β-units of MoFe protein respectively. The black, silver, purple and yellow are α-units of Fe-protein homodimer. Figure are created using VMD[12,13] with the PDB file of nitrogenase from *A. vinelandii* with the code 4WZB[14]. (b) FeMoco with R-homocitrate and part of Cys275 and His442. (c) P-cluster showing only parts of the cystiene ligands

1.1.3. P-cluster

As indicated before the P-cluster supplies FeMoco the electrons needed for reducing reactants. It consist of eight irons and resembles two [Fe_4S_4] cubane merged together and contains 7 sulfur atoms. All iron atoms of the P-cluster are in the ferrous state[15,16] and the central sulfur atom is hexagonally coordinated by six iron atoms (see figure 1.2(c)). Each iron atom is coordinated by two or three sulfide ligands and one terminal or bridging cysteinyl ligand. The P-cluster is oxidised during substrate reduction (P^{+1} state, mixed $S = \frac{1}{2}, \frac{5}{2}$ states[1,17] according to EPR) resulting in reduction of FeMoco (E_1 state) and is in turn[5] reduced by the Fe-protein ($Fe_{red} \rightarrow Fe_{ox}$) to its resting state ($P^N$ state, EPR silent diamagnetic state[1,17]).

1.1.4. Mechanism

The mechanism of nitrogenase reduction is a multi step process involving cycles of Fe-protein electron transfer (ET) and proton transfers to FeMoco and is currently best understood in terms of the Lowe—Thorneley (LT) kinetic model for nitrogenase function[1,18,19].

Thorneley-Lowe kinetic model

The LT scheme consists of eight Fe-protein ET cycles as depicted in figure 1.3 (denoted here as E_n where n is number of transferred electrons) and eight proton transfers (denoted E_nH_m where m is number of bound protons) to each $\alpha\beta$-unit of the MoFe protein. This results in two mole NH_3 and one mole H_2 for each mole of N_2. Many studies have revealed that N_2 binds to FeMoco at a more reduced level than E_0. N_2 binds either at the E_3H_3 or E_4H_4 state and is only further reduced at the E_4H_4 state as is shown in the LT scheme in figure 1.3 and requires[21] a reversible exchange of N_2 for H_2. The system relaxes by $E_4H_4 \rightarrow E_2H_2$+$H_2$, $E_3H_3 \rightarrow E_1H$+H_2 and $E_2H_2 \rightarrow E_0 + H_2$. It is often assumed that a proton-transfer follows a ET but this is not necessarily the case as this has not been firmly established. Thus the redox state could have a different number of H than indicated.

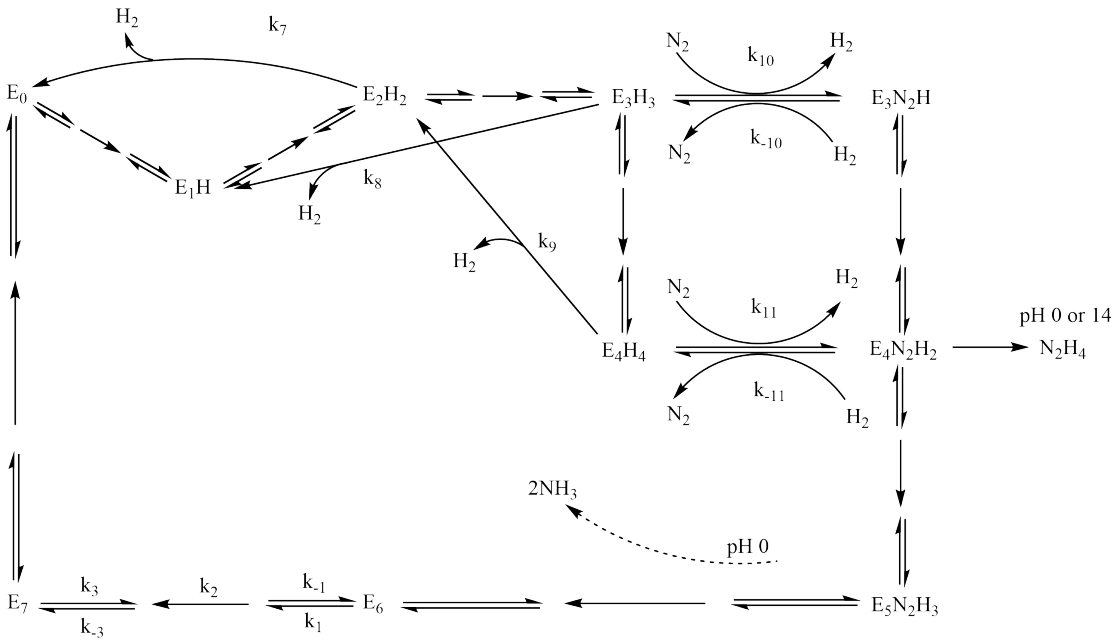


Figure 1.3: The Thorneley-Lowe kinetic mechanism adapted from Orme-Johnson[20]

1.1.5. Substrates in addition to the reduction of dinitrogen

Nitrogenase can reduce a wide variety of substrate e.g. alkynes, azide, cyanides, nitrile analogues, carbon disulfide, carbonyl sulfide, diazines, hydrazine, nitrous oxide, nitrate, thiocyanate, cyanate and dinitrogen (for a more complete list see Seefeldt *et al.*[22]). An altered nitrogenase can reduce even more substrates[23] as can the isolated cofactor with an lanthanide-based reductant[24]. Various inhibitors work for various reactions. All reactions of wild-type nitrogenase except proton reduction can be inhibited by carbon monoxide[9]. A few inhibition cases merit special mention. For example acetylene reduction can completely inhibit proton reduction to hydrogen gas[9]. This is complemented by the fact that cyanate is an inhibitor of proton reduction and if combined with CO, stops all reductions[9]. In contrast to that, dinitrogen can never completely inhibit the proton reduction[9] as has been noted before.

Reaction of deuterium gas and acetylene with dinitrogen as co catalyst has the side-product of C_2H_3D ($\approx$ 2% total ethylene formed) with a minor quantity of $C_2H_2D_2$[25,26]. This is interesting as acetylene reduction with D_2 does not form any deuterated product, which mean something in the mechanism of using N_2 as a co-catalyst, affect the hydrides that reduce C_2H_2. This is coupled with the fact that

this reaction has the percularity that an increase in the partial pressure of dinitrogen (p_{N_2}) results in an *increased quantity of deuterated ethylene products* [25,26]. If electron flux is reduced, the deuterated ethylene products quantity decreases [26]. This indicates that turnover of N_2 forms at some point a FeMoco state that can bind H_2, in the form of hydrides as this deuteration is separate and distinct from incorporation of D^+ (see table 1.1). It is also of special note that reducing C_2D_2, follows cis-addition [27] which suggests a symmetric hydride of equal hydricity.

As an inhibitor and in some cases a substrate [23,24] carbon monoxide is of special interest. CO-bound nitrogenase is the first X-ray crystal structure [28] with a substrate bound to the cofactor. As such it could elucidate much insight into the mechanism of nitrogenase and was thus a part of this work.

Various inhibitory and reaction information can be seen in table 1.1.

Table 1.1: Inhibitors of reactions of nitrogenase. Red mark, ⊗, means it inhibits the reaction while green mark, ⍟, means it enhances the reaction. The reactions are shown with the substrate important to that reaction's product indicated after the double arrow as there may be more product that are not necessarily affected by the inhibitor/enhancer. A down arrow signifies a lowering of amount and an up arrow a addition of an amount

Reaction	Additional substrate and effect
$H^+ \longrightarrow H_2$	C_2H_2⊗[9],CN^-⊗[9,29], SCN^-⊗, CS_2⊗[29], N_2H_2⊗, N_2⊗[9,*] N_3^-⊗, CH_3NC⊗[9]
$H^+ + e^- + N_2 + D_2 \longrightarrow HD$	$p_{N_2}\uparrow \frac{n_{HD}}{n_{H_2}}\uparrow$ [30], $p_{CO}\uparrow 1\% \frac{n_{HD}}{n_{H_2}}\downarrow$ [30],N_2O⊗[30]
$N_2 \longrightarrow NH_4^+$	C_2H_2⊗[9], N_2H_4⊗[30], CO⊗[9], NO⊗[9], N_2O⊗[9], C_2H_2⊗, CN^-⊗[9], SCN^-⊗, $SeCN^-$⊗[31], H_2⊗[30]
$N_2H_4 \longrightarrow NH_4^+$	CO⊗[9], N_2⊗
$HCN \longrightarrow CH_4$	N_3^-⊗[9], CH_3NC⊗, CO⊗[9], CN^-⊗[9], C_2H_2⍟[9], H_2⍟[9], N_2O⍟[9]
$C_2D_2 \longrightarrow 96\%cis\text{–}C_2H_2D_2, 4\%trans\text{–}C_2H_2D_2$ [27]	CO⊗[9], CN^-⊗[9], SCN^-⊗[29], COS⊗[29], CS_2⊗[29], N_2⊗[9,25], N_3^-⊗[9], N_2O⊗[9], $e^-\downarrow \frac{n_{cis\text{–}C_2H_2D_2}}{n_{trans\text{–}C_2H_2D_2}}\downarrow$ [27]
$C_2H_2 + D_2 + N_2 + H^+ + e^- \longrightarrow minor\text{–}C_2H_2D_2, major\text{–}C_2H_3D$ †[26]	CO⊗[9], CN^-⊗[9], SCN^-⊗[29], COS⊗[29], CS_2⊗[29], N_2⊗[9,25], N_3^-⊗[9], N_2O⊗[9], $p_{N_2}\uparrow n_{C_2H_3D}$ & $n_{C_2H_2D_2}\uparrow$ [25,26]
$C_2H_2 + D^+ \longrightarrow cis\text{–}C_2H_2D_2, trans\text{–}C_2H_2D_2$ [22,27]	CO⊗[9], CN^-⊗[9], SCN^-⊗[29], COS⊗[29], CS_2⊗[29], N_2⊗[9,25], N_3^-⊗[9], N_2O⊗[9]
$N_2H_2 \longrightarrow NH_4^+$	CO⊗[9], H_2⊗[25,32], N_2H_4⊗[32]
$SCN^- \longrightarrow CH_4$	N_2⊗[29], CO⊗[9,29]

*At infinite p_{N_2} evolution of H_2 is 13–23% of its maximum value [9]

†Side products of ethylene (≈ 98% of acetylene reductant). [26]

1.1.6. Aim of this research

The main aim of this research was to gain structural and mechanistic insights into the problem of nitrogen reduction by quantum chemical calculations.

The resting state of FeMoco was analysed by exploring the many different broken-symmetry solutions of FeMoco. This work is discussed in chapter 3.

Another topic of interest is the unusual CO bound FeMoco structure as revealed in the recent crystal structure. As it is known that the EPR signal changes when nitrogenase is CO inhibited, we sought to gain insight into this observation. This work can be seen in chapter 4.

The structure of the singly reduced protein is also of special interest to us as EXAFS data suggests that FeMoco, "breathes", as it is reduced. We wanted to see if this could be reproduced by QM/MM models of possible singly reduced states. These result can be seen in chapter 5.

The P-cluster has much significance as it is the natural redox agent of nitrogenase. We wanted to see if we could calculate the redox potential of the one electron oxidation and compare it to experimental data. The progress we made can be seen in chapter 6.

Hydrazine could well be the last stage of dinitrogen reduction in nitrogenase. As we had done prior work on hydrazine reduction on synthetic [$MoFe_3S_4$] cubanes, we wanted to compute the analogous mechanism in nitrogenase on the similar molybdenum iron cofactor. This work is discussed in chapter 7.

2. Theory and methods

2.1. Quantum mechanics

A complete description of the underlying cause of chemical phenomena are rooted in quantum mechanics. All particles can be described with a *wavefunction*, Ψ, which in essence is the de Broglie wave associated with the particle. Information about quantities of interest can be computed from the wavefunction with quantum operators. The non-relativistic Schrödinger equation describes how the wavefunction changes with respect to time, t:

$$i\hbar\frac{\partial\Psi(r,t)}{\partial t} = \hat{H}\Psi(r,t) \tag{2.1}$$

where the Hamiltonian operator is:

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial r^2} + V(r,t) \tag{2.2}$$

and m is the mass of particle, $\bar{h}$ is the reduced ($\frac{h}{2\pi}$) Planck's constant and $V(r,t)$ is the potential energy of the system that is dependent on both the spatial coordinate, r, of the particle and time, t.

The *wavefunction* is an eigenfunction, and the energy, E, is an eigenvalue of the $\hat{H}$ operator. When a time-independent potential, $V(r)$, is defined, the time-independent form of the Schrödinger equation can be used instead:

$$\hat{H}\psi(r) = E\psi(r) \tag{2.3}$$

If the wavefunction of a system is known the square of the time-independent wavefunction, ψ, is interpreted as a probability density, ρ, signifying the probability that a particle is in a given position in a volume of space, dr:

$$\rho = |\psi(r)|^2\, dr \tag{2.4}$$

While the time-independent Schrödinger equation can be solved analytically for the hydrogen atom and a few other simple one-particle problems, no such solutions exists for many-body systems. This difficulty can be addressed with numerical methods and in principle the equation can be solved to yield the ground and exited state wavefunctions for any quantum system with a time-independent potential but numerical methods are computationally very demanding and thus only useful for small systems as can be seen from the form $\hat{H}$ takes for a system of nuclei and electrons:

$$\hat{H} = -\sum_i \frac{\hbar^2}{2m_e}\nabla_i^2 - \sum_k \frac{\hbar^2}{2m_k}\nabla_k^2 - \sum_i \sum_k \frac{e^2 Z_k}{4\pi\epsilon_0 r_{ik}} + \sum_{i<j} \frac{e^2}{4\pi\epsilon_0 r_{ij}} + \sum_{k<l} \frac{e^2 Z_k Z_l}{4\pi\epsilon_0 r_{kl}} \tag{2.5}$$

where i and j are indices of electrons and k and l are the indices of nuclei and the Laplace operator (del squared) is:

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{2.6}$$

The m stands again for mass, r is the inter-particle distance, e the elementary charge, Z is the nucleus charge and $4\pi\epsilon_0$ is the permittivity of free space (also known as vacuum permittivity). The first two terms of eq. 2.6 are the kinetic energy of the electron and nuclei respectively, the third term describes the potential energy between the electrons and nuclei, and the fourth and fifth terms describe the potential energy of electron-electron and nuclei-nuclei interactions respectively.

By recognising that electron move much faster than nuclei (about 1800 times faster) one can make a crucial approximation, the Born-Oppenheimer (BO) approximation, which ignores the second term in eq. 2.6 and only includes the fifth term as a constant classical energy term with respect to nuclear positions:

$$(\hat{H}_{elec} + V_{NN})\psi_{elec} = E_{elec}\psi_{elec} \tag{2.7}$$

$$\hat{H}_{elec} = \sum_{i=1}^{N} -\frac{\hbar^2}{2m_e}\nabla_i^2 + \sum_{i=1}^{N} V_{ext}(r_i) + \frac{e^2}{4\pi\epsilon_0}\sum_{i<1}^{N}\frac{1}{r_{ij}} \tag{2.8}$$

$$V_{NN} = \sum_{k<l}^{N} \frac{e^{2Z_k Z_l}}{4\pi\epsilon_0 r_{kl}} \tag{2.9}$$

As the nuclear are constant parameters to be tuned, the Schrödinger equation is solved for certain geometry with fixed nuclear position and then changed to get the Born-Oppenheimer potential energy surface (PES) which has proven to be one

of the most useful tools of chemistry. The energy of the system can be often found by the variational principle (VP), which states[33]:

$$\langle E_{trial} \rangle = \frac{\langle \Phi_{trial} | \hat{H} | \Phi_{trial} \rangle}{\langle \Phi_{trial} | \Phi_{trial} \rangle} \leq E_0 \tag{2.10}$$

This principle opens up the possibility to try all possible trial wavefunctions as the energy will always be higher or equal to the groundstate wavefunction energy, E_0. Thus a systematic route to lower a trial wavefunction's energy opens up. A convenient starting point is approximating the N-electron wavefunction as a linear combination of one electron wavefunctions and the Slater determinant is one way to make such an approximation as it obeys the Pauli exclusion principle[34,35]:

$$\psi_{SD} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(x_1) & \phi_2(x_1) & \cdots & \phi_N(x_1) \\ \phi_1(x_2) & \phi_2(x_2) & \cdots & \phi_N(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(x_N) & \phi_2(x_N) & \cdots & \phi_N(x_N) \end{vmatrix} \tag{2.11}$$

where $\phi(x)$ are spin orbitals. When the VP is also used, the Hartree-Fock (HF) energy equation can be derived[33]:

$$E_{HF} = \langle \psi_{SD} | H_{elec} + V_{NN} | \psi_{SD} \rangle \tag{2.12}$$

Minimisation of the energy can be achieved by varying the orbitals and this is achieved by solving the HF equations[36]:

$$\hat{F}\phi_i = \epsilon_i \phi_i \tag{2.13}$$

for each electron where ϵ_i are the eigenvalues to the Fock operator $\hat{F}$ which is defined as[33]:

$$f = -\frac{1}{2}\nabla_i^2 - \sum_k \frac{Z_k}{r_{ik}} + V_{HF} \tag{2.14}$$

where V_{HF} is the HF potential which is an average repulsive potential of all electrons except for the one which is being solved. The Fock operator depends on the orbitals and operates on them and has thus have to be solved through the self-consistent method. First initial orbitals are guessed, and iterating the HF equations until the orbitals stop changing. Then the HF energy equation gives the total energy. HF can be systematically improved by including exited Slater determinants as a basis for a more flexible wavefunction. This is referred to as post-HF and is computationally costly. As HF theory is not accurate enough for most chemistry and post-HF is very costly, an alternative approach that has become very popular will be described next: Density functional theory (DFT)

2.2. Density functional theory

Modern DFT builds on the Hohenberg-Kohn (HK) theorems and the Kohn-Sham (KS) approach which provide a connection between the electronic density and the total energy of the molecule[33].

The first HK states that the exact groundstate energy is a functional of the density and given the Born-Oppenheimer approximation the equation is as follows[36]:

$$E[\rho(r)] = T_e[\rho(r)] + V_{eK}[\rho(r)] + V_{ee}[\rho(r)] + V_{KK} \tag{2.15}$$

where T_e is the kinetic energy of the electron density, V_{eK} is the potential energy between the electron density and nuclei, V_{ee} is the potential energy between the electron density and the electron density and V_{KK} the potential between nuclei. There should be a universal functional valid for all system based on the HK theorem but unfortunately, the exact form is unknown as of this time. The KS approach make it possible to calculate the kinetic term via orbitals and the V_{eK} is computed in a manner exactly to Hartree-Fock (HF) theory and V_{KK} is calculated as classical Coulomb potential energy but the potential of electron-electron interactions is problematic. To simplify the problem the term is broken up into the classical Coulomb potential energy, $J[\rho(r)]$, and the exchange-correlation energy, $E_{XC}[\rho(r)]$. The exchange-correlation energy term has to be approximated as it is unknown and the simplest way forward is to use the exchange and correlation functionals that have been derived for a uniform electron gas. This is called the local density approximation (LDA):

$$E_{XC}^{LDA}[\rho(r)] = E_X^{SD}[\rho(r)] + E_C^{LDA}[\rho(r)] \tag{2.16}$$

where E_X^{SD} is the Slater-Dirac (SD) exchange functional and E_C^{LDA} the LDA correlation functional. As this approach is not sufficiently accurate for molecules (with more complicated densities), another approximation is commonly used, the Generalized Gradient Approximation (GGA)[33]:

$$E_{XC}^{GGA}[\rho(r)] = \int E_{XC}^{LDA}[\rho(r)] F_{XC}(\rho(r), s) dr \tag{2.17}$$

where F_{XC} function is called an enhancement factor and can be divided up into F_X and F_C. There have been many attempts of constructing the enhancement factors and thus there are many functionals available. The following functionals were used in this work: BP86, B3LYP and TPSSh. The Becke '88 exchange (B88)[37] and Perdew '86 (P86)[38] correlation functionals uses the Becke '88 exchange

enhancement factor (F_X^{B88}):

$$F_X^{B88}(s) = 1 + \frac{as^2}{1 + bs\ln(s + \sqrt{1 + s^2})} \tag{2.18}$$

where a and b are fitted to noble gas atom HF exchange energies and Perdew '86 correlation takes a more complicated form (not shown). The B3LYP hybrid functional uses a blend of the B88 exchange[37,39] with a mixture of Lee-Yang-Parr (LYP)[40] and LDA correlation functionals and a scaled down HF exchange term:

$$E^{B3LYP} = E_{XC}^{LDA} + 0.2(E_X^{HF} - E_X^{LDA}) + 0.72(E_X^{B88} - E_X^{LDA}) + 0.81(E_C^{LYP} - E_C^{LDA}) \tag{2.19}$$

This functional is thus more expensive than regular GGA-DFT due to the HF exchange term. The TPSS functional by Tao, Perdew, Staroverov, and Scuseria is a meta-GGA functional; this functional depends in addition to $\rho(r)$ and $\nabla\rho(r)$ (normal GGA), on the kinetic energy density, $\nabla^2\rho(r)$ and is thus slightly more computationally expensive than other GGA functionals, like LYP and BP86.

TPSSh is a hybrid functional of TPSS which uses 10% HF exchange and is defined in a similar way to the B3LYP functional without the scaling of the correlation term[41].

$$E^{TPSSh} = 0.9E_X^{TPSS} + 0.1E_X^{HF} + E_C^{TPSS} \tag{2.20}$$

The Kohn-Sham equations are the DFT equivalent of the HF equations and solving them minimizes the KS energy expression[33]:

$$\hat{F}^{KS}\phi_i = \epsilon_i\phi_i \tag{2.21}$$

where the KS operator is:

$$\hat{F}^{KS}[\rho(r)] = h_i[\rho(r)] + \sum_j J_{ij}[\rho(r)] + v_{XC}[\rho(r)] \tag{2.22}$$

where $v_{XC}[\rho(r)] = \frac{\delta E_{XC}[\rho(r)]}{\delta\rho(r)}$

2.2.1. Unrestricted DFT and broken-symmetry

As many molecules have all electrons paired, up-spin paired with down-spin, in the same molecular orbital (MO), the spin term of the spin orbitals in HF orbitals cancels out and results in simpler, cheaper HF equations, called Restricted Hartree-Fock equations (RHF) or if using DFT, Restricted Kohn-Sham (RKS). If on the

other hand there is an odd number of electrons for any atom in a molecule or open-shells are present this simplification must be abandoned. That is called Unrestricted Hartree-Fock (UHF)/Unrestricted Kohn-Sham (UKS) and is more costly but correctly describes spin-polarization effects in molecules. UHF/UKS has the disadvantage however that UHF/UKS wavefunctions are not eigenfunctions of the $\hat{S}^2$ operator. This is known as spin contamination and the energy can be artificially lowered beyond the ground state because of it[33].

When open shells are present in the molecule on multiple atoms, the electronic structure becomes more complicated due to possible magnetic/spin coupling between the open shell transition metal ions. The unpaired electrons on different sites couple weakly to produce a ladder of spin states[36]. This ladder may phenomenologically be described by a so-called Heisenberg Hamiltonian (also called a spin Hamiltonian)[36]:

$$\hat{H} = -2J * \hat{S}_A * \hat{S}_B \tag{2.23}$$

where $\hat{S}_A$ and $\hat{S}_B$ are the spin-operators for sites A and B, respectively and J is the exchange coupling constant. A positive J refers to ferromagnetic coupling where the coupled electrons have the same spin. If J is negative, the spins of the coupled electrons act as if paired and the coupling in the system is said to be antiferromagnetic. Ferromagnetic coupling will result in a high spin state while antiferromagnetic coupling will give a low spin state. The system can also be in a spin state in between. A molecule with multiple open shell metal ions result in a complicated spin coupling problem. A molecule with many unpaired electrons will thus be able to couple either way. To calculate spin coupled systems with DFT, where many self-consistent field (SCF) solutions exist for many possible spin states, the system is first calculated as the high spin (HS) state with all unpaired electrons the same spin. Next the best way is to generate a so called Broken Symmetry (BS) solution in which some electrons localizes with up spin (also called α-electrons) on one site and other electrons with down spin (also called β-electrons) by flipping the spin density on site of interest. This is in many way a trick to emulate the complexity of the system. The BS solution is then found by solve the SCF equations to find the solution with M_S value corresponding to the spin of interest. BS-DFT in practice does not always find the BS solution it is guided to and thus the choice which atoms are flipped is more guidelines for the DFT computations than a setting. BS solutions are also not eigenfunctions of the $\hat{S}^2$ operator and thus spin contamination is noticeable and the relevance of the BS solution to reality is not always clear and has to be evaluated, case by case.

2.3. QM/MM model

As a full QM description of a large system can be prohibitively costly, an approximation of the full system can be used where the area of interest (FeMoco in this case) is treated with a QM method while the nearest environment are approximated with molecular mechanics (MM). This is called QM/MM. This can be described with the following additive equation:

$$E_{QM/MM} = E_{QM}(I) + E_{MM}(II) + E_{\text{QM-MM}}(I, II) \tag{2.24}$$

where $E_{QM/MM}$ is the total energy of the system, $E_{QM}(I)$ is the QM energy of the QM region (I), $E_{MM}(II)$ is the MM energy of the MM region (II) and $E_{\text{QM-MM}}(I, II)$ is the QM-MM interaction term between the two regions. The MM energy consist of both bonded and nonbonded terms:

$$\begin{aligned} E_{MM} = & \sum_{\text{bonds}} k_d(d - d_0)^2 + \sum_{\text{angles}} k_\theta(\theta - \theta_0)^2 + \sum_{\text{dihedrals}} k_\phi[1 + \cos(n\phi + \delta)] \\ & + \sum_{\text{nonbonds}} \epsilon_{AB}[\left(\frac{\sigma_{AB}}{r_{AB}}\right)^{12} - \left(\frac{\sigma_{AB}}{r_{AB}}\right)^{6}] + \frac{1}{4\pi\epsilon_0}\frac{q_a q_B}{r_{AB}} \end{aligned} \tag{2.25}$$

where the constants $k, d_0, \theta_0, \delta, \epsilon, \sigma, q_A, q_B$ are fitted so that the force field energies of the system for which the force field is intended. Several parameterised force-fields are available to describe proteins and was CHARMM36[42,43] force field used in this work. The QM and MM region are typically connected by bonds and thus a scheme is needed to correctly represent the boundary. One of those scheme is the link atom approach which was used in this work. The link atom method terminates the dangling bond in the QM region by the addition of link atoms (usually hydrogen atom) along the QM-MM bonds, thereby mimicking the effect of real atom. This deals with the QM-MM boundary problem as the dangling bond of the QM region is saturated with H atoms as part of the QM region in E_{QM} calculation and a so called charge-shifting correction used to get rid of a strong artificial electrostatic interaction due to the first MM atom being to close to the link atom. The QM/MM calculations were performed with ChemShell[44], a computational chemistry program that uses a modular approach. Chemshell links together the different external programs and was ORCA[45] used for the QM calculation and DL_POLY[46] for the MM calculations. Both the MM and QM/MM model used in this work were modelled by Benediktsson[47].

All DFT computations were carried out using the ORCA program, version 3.0.3.[45], using the ZORA relativistic approximation[48,49] and all-electron relativistically re-contracted def2-SVP basis sets[50,51] on all atoms except Mo, Fe and S which used

def2-TZVP basis sets[50,51] (using def2-XVP/J auxiliary basis sets[52] for use with the RI-J approximation). When B3LYP and TPSSh hybrid functionals were used the RIJCOSX approximation[53–56] was also used. All geometry optimizations included a dispersion correction[57,58].

2.3.1. ORCA gas phase computations

Geometries of gas phase clusters were optimized with the BP86 functional using the ZORA relativistic approximation and all-electron relativistically recontracted def2-SVP basis sets[50,51] on all atoms except Mo, Fe and S which used def2-TZVP basis sets[50,51] (using def2-XVP/J auxiliary basis sets[52] for use with the RI-J approximation). A dielectric field ($\epsilon = 4$) was introduced using the COSMO approximation[59] in order to account for environmental effects. Geometry optimizations included a dispersion correction[57,58].

3. E_0 structure

As noted before the Fe-protein adds electrons to the MoFe protein. The state of added electrons to the MoFe protein are the LT states with E_0 as the ground state before any added electrons. As the Fe-protein binds, the P-cluster reduces FeMoco ($P^N \rightarrow P^{1+}$) and then goes to back the neutral state with the electron from Fe-protein ($Fe_{red} + P^{1+} \rightarrow Fe_{ox} + P^N$). This happens faster than complex dissociation[5] ($k_{ET} = 140\ s^{-1}$ vs. $k_3 = 6\ s^{-1}$, see table A.1 in appendix) so to a good approximation a change in LT state is mostly describing the redox state of FeMoco. The complete nature of the resting form of nitrogenase (E_0) is not fully known. The E_0 state is EPR active with a spin of $S = \frac{3}{2}$[19]. The oxidation state of FeMoco has been thought to be $[MoFe_7S_9C]^{-3}$, $[MoFe_7S_9C]^{-1}$ or $[MoFe_7S_9C]^{+1}$[60]. Given that sulfur and carbon are thought to be in their usual closed-shell forms S^{2-} and C^{4-}, the respective oxidation state discussed in the literature are $[6\,Fe^{2+}{:}1\,Fe^{3+}{:}Mo^{4+}]$, $[4\,Fe^{2+}{:}3\,Fe^{3+}{:}Mo^{4+}]$, $[2\,Fe^{2+}{:}5\,Fe^{3+}{:}Mo^{4+}]$ where the molybdenum atom has been assigned as Mo^{4+} based on ^{95}Mo Electron nuclear double resonance (ENDOR) studies[61,62].

However, new studies[63] have indicated that the molybdenum atom is actually Mo^{3+} which in respect to the resting state, necessitates reconsidering the previously assigned iron oxidation state models that resulted from the assuming of Mo^{4+}. This would translate the three oxidation states of FeMoco into:

$[5\,Fe^{2+}{:}2\,Fe^{3+}{:}Mo^{3+}]$, $[3\,Fe^{2+}{:}4\,Fe^{3+}{:}Mo^{3+}]$, $[1\,Fe^{2+}{:}6\,Fe^{3+}{:}Mo^{3+}]$. Recent spatially resolved anomalous dispersion refinement (SpReAD) data by Einsle *et al.*[64] have narrowed down the possibilities to $[3\,Fe^{2+}{:}4\,Fe^{3+}{:}Mo^{3+}]$ or the $[MoFe_7S_9C]^{-1}$ with the oxidation states arranged as depicted in figure 3.1.

Noodleman *et al.*[65] performed DFT calculation and found ten broken-symmetry (BS) solutions on the bases on an approximate C_3 axis of rotational symmetry of the FeMoco cluster. In that work Noodleman *et al.* assumed FeMoco to without the central carbon atom (as it had not been discovered at that time) and assigned a Mo^{4+} oxidation state on molybdenum. Later work[66,67] used an interstitial nitride and found a configuration referred to as BS7 to be most stable. Szilagyi *et al.*[68], DeBeer *et al.*[11] and Björnsson *et al.*[63] have confirmed that configuration to be stable with a carbide.

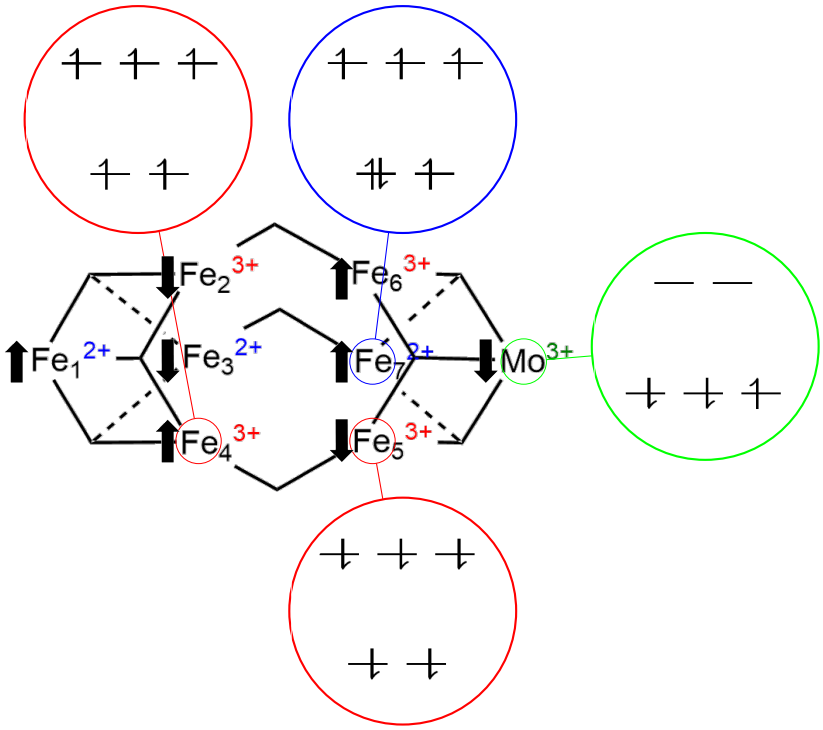


Figure 3.1: Simplified spin diagram[‡] of FeMoco metallocluster with metal oxidation states from Einsle et al.[64] and the lowest BS DFT state according to our result. Mo^{3+} d^3 configuration is shown as suggested by Björnsson et al.[63]. Tetrahedral d-electron configuration are shown for Fe^{2+} and Fe^{3+}. Configurations are oversimplified as electrons may be delocalized as is known with spin coupled iron-sulfur systems[69]

[‡]Diagram made with ChemDraw®.

3.1. Results

We decided to explore the BS solution problem with the now more accepted Mo^{3+} and the recently verified $[MoFe_7S_9C]^{-1}$ total charge. We performed gas phase single point energy calculations on the X-ray geometry from Spatzal *et al.*[10] with BP86, B3LYP and TPSSh functionals and final $M_S = \frac{3}{2}$ and $M_S = \frac{1}{2}$ with all BS solutions that have 3 irons flipped and then repeated with molybdenum flipped. The $M_S = \frac{1}{2}$ solutions were explored to check how close in energy such solutions are as $S = \frac{1}{2}$ FeMoco states are known to appear in substrate reduction studies. The lowest energy results are depicted in figure 3.2 and the complete results are in table A.3 in the appendix. The spins that were flipped are indicated by the number of iron shown in figure 3.1. It is of special note how close in energy the lowest $M_S = \frac{3}{2}$ and the lowest $M_S = \frac{1}{2}$ BS solutions are in these calculations. For BP86 this spin energy gap is around 5 kcal/mol and widened considerable when we used TPSSh and B3LYP. These results confirms work by Noodleman *et al.*[67] that the symmetric BS solutions, $Fe_2\downarrow Fe_4\downarrow Fe_7\downarrow$, $Fe_3\downarrow Fe_4\downarrow Fe_6\downarrow$ and $Fe_2\downarrow Fe_3\downarrow Fe_5\downarrow$ are the lowest in energy with the last being lowest in all functionals. It should be noted that spin flip with or without flipping molybdenum in figure 3.2 are the same BS solution.

In QM/MM calculation performed, in this work the $Fe_3\downarrow Fe_4\downarrow Fe_6\downarrow$ solution was mostly used as previous geometric analysis[§] suggested better agreement with experiment. Recent work[§] and SpReAD experiments[64] suggest $Fe_2\downarrow Fe_3\downarrow Fe_5\downarrow$ may be the state favoured in the protein. All three functionals used, agree on the stability of the three BS solutions mentioned.

[§]Unpublished work by B. Benediktsson

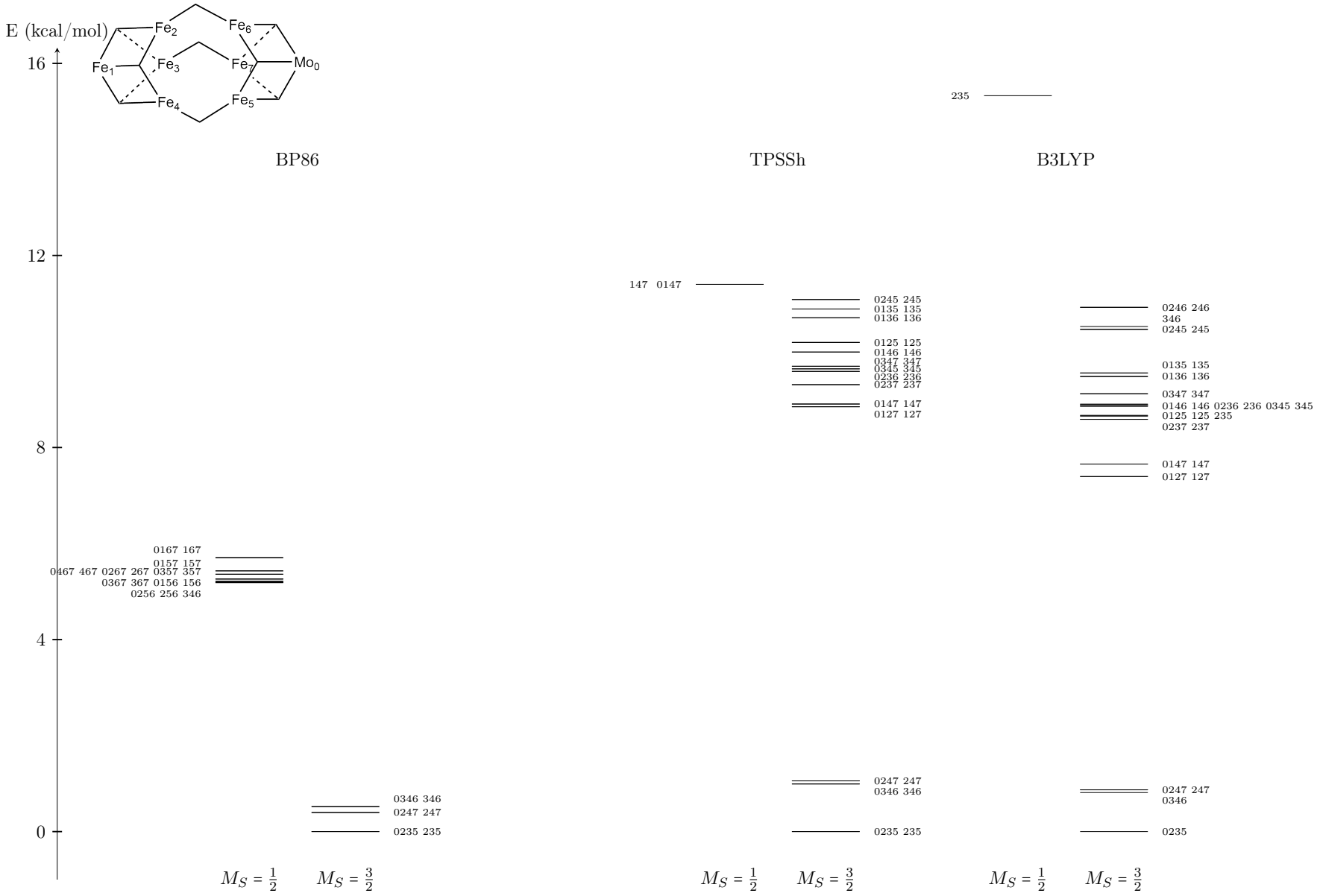


Figure 3.2: Comparison of multiple BS solutions found by spin flips on Mo and Fe with different functionals. Both from HS solution. The HS solutions were: BP86 282.45 *kcal/mol, TPSSh* 208.45 *kcal/mol, B3LYP* 182.70 *kcal/mol*

4. CO inhibited MoFe protein

Carbon monoxide is an excellent non-competitive inhibitor (suggesting a different binding site than N_2) of all reactions of nitrogenase, except proton reduction to dihydrogen which is does not inhibit at all[9]. N_2 and CO are isoelectronic, both good π-acids (π acceptors) and only bind to a reduced state of FeMoco. CO is a reversible inhibitor and is a substrate to VFe protein[70] (with vanadium instead of molybdenum in the catalytic cofactor) and some cases a substrate for the MoFe protein[23,24]. The EPR signal changes from $S = \frac{3}{2}$ to $S = \frac{1}{2}$ in adducts of CO to nitrogenase and there have been found two distinct signals depending on the pressure of CO involved, called lo-CO if there is low CO pressure and hi-CO if there is high CO pressure[3,71,72]. There has even been a third EPR signal with $S = \frac{3}{2}$ detected with high electron flux and $p_{CO} \geq 50$ kPa called hi(5)-co[71,72]. Interestingly many other intermediates also show $S = \frac{1}{2}$ spin signal and a mutated nitrogenase has even been found to have $S = \frac{1}{2}$ while reducing proton to dihydrogen[73]. The same spin state has also been discovered in the photoinduced $E_4H_4 \rightarrow E_2^*H_2$ reaction[74]. Different CO adduct geometries have been proposed[75,76] but have not been on firm ground until 2014 when an X-ray crystal structure of the CO-adduct was solved[28]. The crystal structure shows curiously that a bridging sulfide has been replaced with a bridging CO adduct and is believed to be of a lo-CO structure. Other CO adduct structure are not known.

In order to understand how the spin state changes when CO binds to FeMoco, we explored different BS solutions for the FeMoco adduct and a possible reaction mechanism for the CO bridging state and H_2S loss. Understanding CO binding to FeMoco may aid in understanding N_2 binding.

4.1. Results

Similar gas phase single point energy computations as before were performed, now on the geometry from Spatzal *et al.*[28] that features a μ–CO-bound structure, again with the BP86, B3LYP and TPSSh functionals and spin states $M_S = \frac{3}{2}$ and $M_S = \frac{1}{2}$. Calculation were performed with all BS solutions that have 3 irons flipped and repeated with molybdenum flipped. The results that are lowest in energy are depicted in figure 4.1 and the complete results are in table A.3 in the appendix.

The first interesting point here was that BP86 functional did not predict $S = \frac{1}{2}$ to be the ground state of the CO-bound structure that is most likely the spin state of this adduct. Also of interest was how many of the spin-flip calculations converged on same BS solution. The Mulliken charges and spin densities showed clearly that all of the lowest TPSSh functional computations were converging on the same solution, namely $Fe_2\downarrow Fe_4\downarrow Fe_6\downarrow Fe_7\downarrow$, and it seems that CO as a π-acid promotes ferromagnetic coupling between iron Fe_2 and Fe_6 leading to a different final spin state of the cluster. B3LYP led to a slightly different picture but again the $Fe_2\downarrow Fe_4\downarrow Fe_6\downarrow Fe_7\downarrow$ spin-flip pattern was found as the lowest BS solution with $M_S = \frac{1}{2}$.

A separate project consisted of exploring the mechanism of CO binding to the FeMoco cluster. At some time in the mechanism the total spin must change from the $S = \frac{3}{2}$ to $S = \frac{1}{2}$ and so we tried every step of the mechanism with both spins. All these calculations were gas-phase geometry optimisations using a continuum solvation model[59] (COSMO) with a dielectric constant, $\epsilon = 4$, and the BP86 functional. All calculation used the $Fe_3\downarrow Fe_4\downarrow Fe_6\downarrow$ BS solution. Finding a suitable proton donor for the calculation posed a problem. We decided that a glutamic acid in gas phase with COSMO, $\epsilon = 4$, would suffice to start with and can it be reasoned that some of the nearby glutamic acid might be protonating the cluster though the exact means of protonation is unknown[6,77–79].

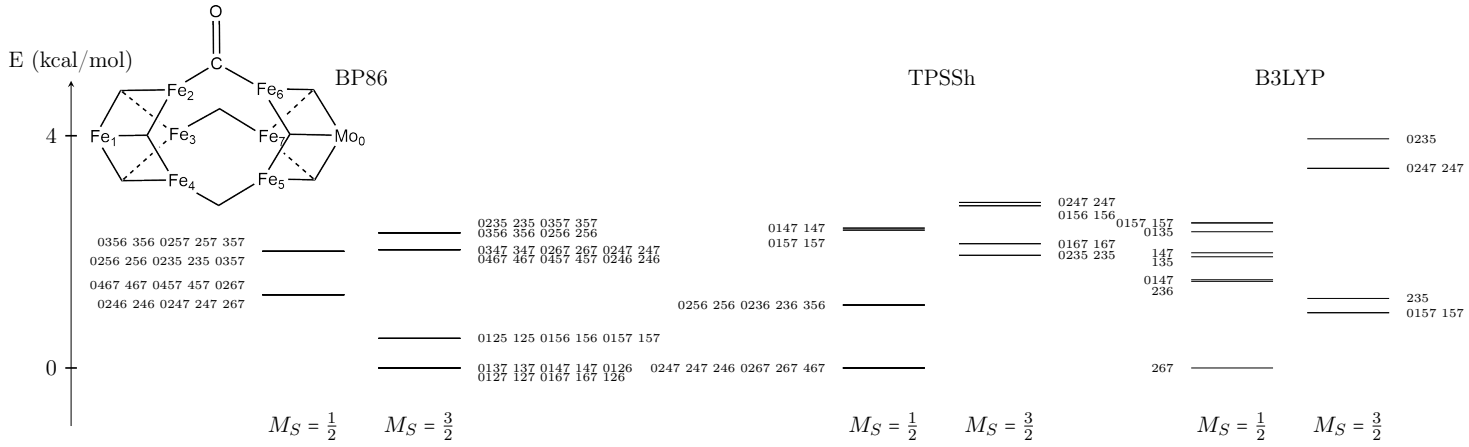


Figure 4.1: Comparison of multiple BS solutions found by spin flips on Mo and Fe on the CO adduct geometry with different functionals; *The HS solutions were: BP86* 328.22 *kcal/mol, TPSSh* 237.67 *kcal/mol, B3LYP* 206.94 *kcal/mol*

Interestingly $M_S = \frac{1}{2}$ becomes lower in energy right away when the sulfide bridge connected to iron Fe_6 is protonated or CO binds to iron Fe_6 as an adduct (see figure 4.2). Protonation of the sulfide bridge is probable as the first step followed by the CO adduct on Fe_6 and is the second protonation of the sulfide the rate limiting step with an energy barrier of approximately 30 kcal/mol, suggesting that the second protonation probably can not occur on a non-reduced cofactor. Our calculation indicate that CO can favourably bind to the ground state without any added electrons to FeMoco which is not consistent with experiment. However we note that our calculations do not take into account zero-point energy and entropy that will affect the binding energy and only employ a crude small cluster model. future QM/MM calculations will describe the environment more accurately. It would be interesting to see if the spin change holds true for other irons in the cluster, for example Fe_2 which is probable coupled to Fe_6 and could also bind substrates. We will continue these calculations with a suitable reduction agent and explore more BS solutions in this mechanism to explore every avenue of inquire.

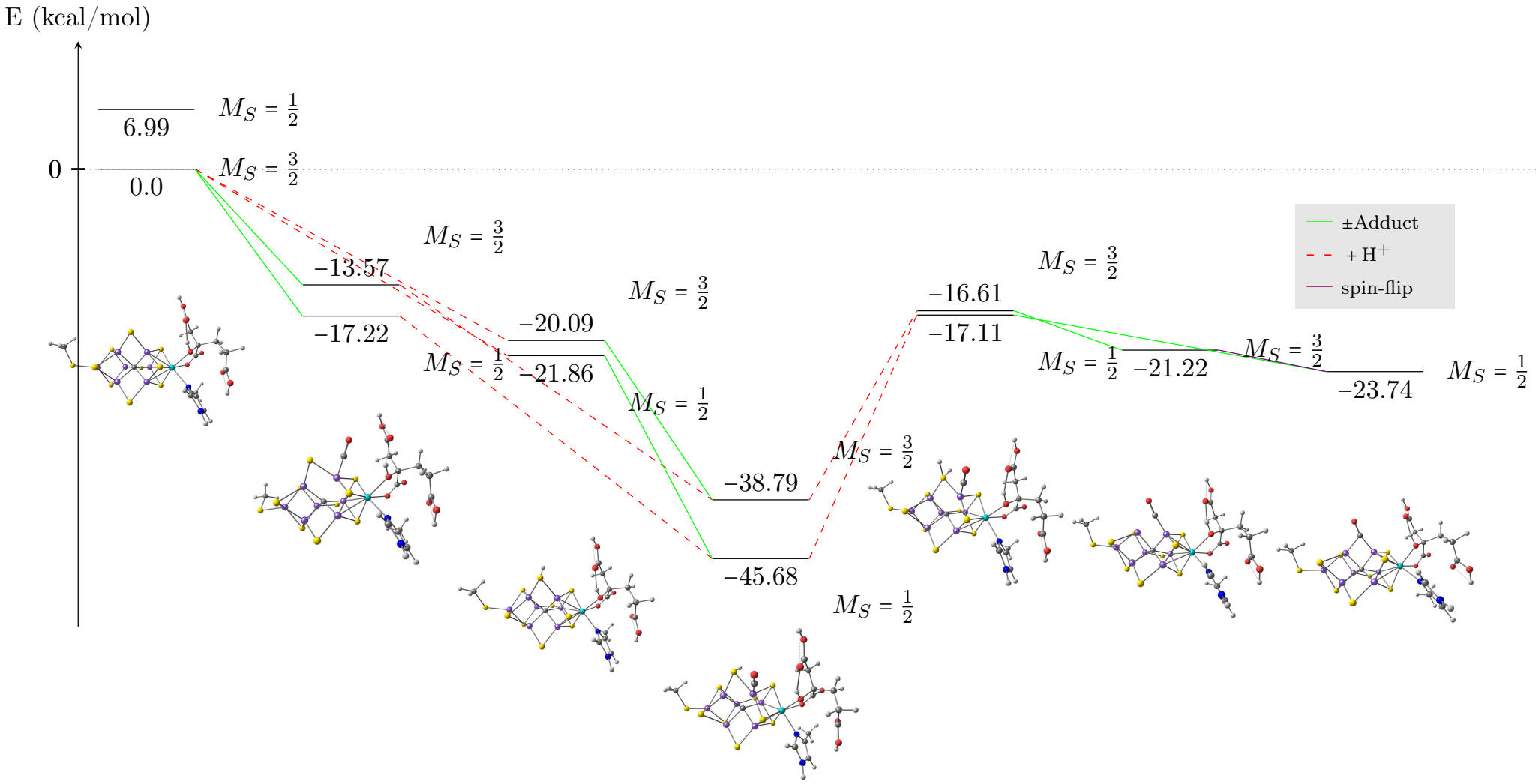


Figure 4.2: Free energy diagram for CO binding to FeMoco cluster with BP86 functional. Images of the cluster were made with ChemCraft[80]. ± adduct refers to adding or removing a molecule, in this case adding CO and removing H_2S. The red dashed lines is an addiction of a proton. The violet lines involve a change in M_S. All energies are relative to the ground state of FeMoco. All calculations used $Fe_3{\downarrow}\,Fe_4{\downarrow}\,Fe_6{\downarrow}$ BS solution

5. E_1 structure

Little is known about the E_1 state other than that it is presumed to build up during turnover due to a inability to relax to the ground state[21,81]. All reactions of nitrogenase have E_1 as a part of the reaction and information gathered on the nature of this state will have immense value. E_1 is EPR silent and paramagnetic[82] ($S \geq 1$) but it is unknown what the total spin is. Studies on E_1 have been carried out by Cramer *et al.*[83,84] and Münck *et al.*[85]. While E_1 is often assumed to be accompanied by proton addition (the label E_1H_1 is often used), direct evidence for this is lacking.

Cramer *et al.*[83,84] found by EXAFS that the average Mo–O/N and Mo–Fe bond length shorten by the one electron reduction (Mo–O/N by 0.070 Å and Mo–Fe by 0.050 Å). This is opposite to what one would expect as more electrons would be expected to destabilize chemical bonds and hence the bonds would elongate. Indeed this is more akin to oxidation of the cluster as that would contract the metal-to-ligand/metal bonds. This is coupled with the fact that Münck *et al.* found that the change in Mössbauer isomer shift between the one electron reduced FeMoco to ground state is smaller then the isomer change between the one electron oxidized FeMoco to the ground state (δ_{av} = 0.02 mm/s vs. δ_{av} = 0.06 mm/s, respectively). This was interpreted as an indicator that the one electron reduction of FeMoco could be centered on Mo. However, little evidence suggests Mo changes oxidation state upon oxidation or reduction of FeMoco[60]. Another way of interpreting this small isomer shift change would be possible hydride formation instead of metal reduction.

The EXAFS and Mössbauer data of course already suggests that the E_1 state is something more complicated than simple metal based redox. In this study we built up several models of E_1, including models based on metal hydride formation and sulfur protonation accompanying one electron reduction. We also tried the recently proposed carbon protonation[86,87].

5.1. Results

The bond length change that Cramer *et al.*[83] found is noticeable missing from our first computations, calculations made with both BP86 and TPSSh, which assumed that *no proton transfer in going from* E_0 *to* E_1. As these calculations show bond lengths elongate rather than shorten (BP86 Mo–O/N by 0.038 Å and Mo–Fe by 0.066 Å; TPSSh Mo–O/N by 0.045 Å and Mo–Fe by 0.072 Å), something was amiss.We decided to investigate proton migration on FeMoco in the E_1 state. Previous calculation by Bjornsson[?] suggest that the Mo-bound alkoxy group of homocitrate is protonated in the E_0 state. Deprotonation or proton migration could explain the Mo–O/N bond contraction. In our calculations, both M_S = 2 and M_S = 1 spin states were explored but all possibilities have not been explored. All energies are QM energies of ORCA as the changes in the protein made large contributions in the MM energy which changed between optimization, indicating that the QM region may not be large enough. The QM/MM model was made by Benediktsson[47].

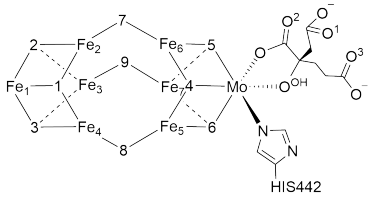


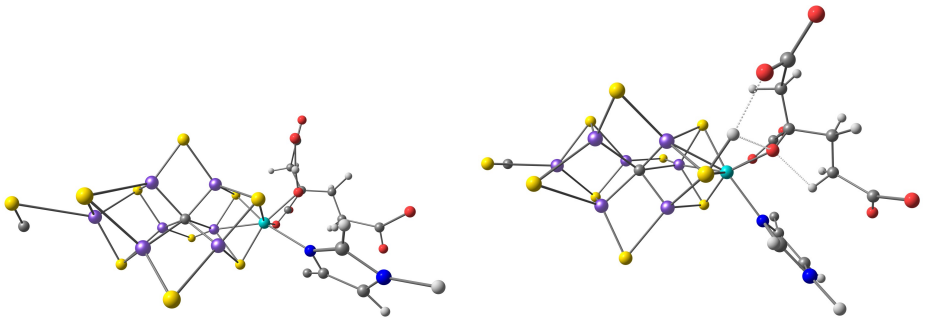
Figure 5.1: Simplified diagram of FeMoco metallocluster with sulfur and carboxylic acid protonation places marked in. Hexavalent carbon is not shown for clarity

We used $Fe_3\downarrow Fe_4\downarrow Fe_6\downarrow$ BS solution and TPSSh for this work.

Table 5.1: Geometric data and relative energies of different E_1 models compared to E_0; All Bond lengths in Å. Different BS solution of the ground state are indicated by what irons were flipped otherwise $Fe_3\downarrow Fe_4\downarrow Fe_6\downarrow$ was used. All E_1 calculations used the final spin $M_S = 2$ unless indicated otherwise. All E_0 calculations had $M_S = \frac{3}{2}$. Proton position are according to figure 5.1 with the H after the position it was added unless it was bridging. Then hapticity is followed by atoms it bridged. E_0 with different charges are also shown. Dotted means no data. The difference in the ave. metal-metal bond or metal-ligand bond was calculated as bonds of E_1 state minus E_0

	ΔE (kcal/mol)	Mo–Fe_6	Mo–Fe_7	Mo–Fe_5	Average Fe–Mo	ΔFe–Mo	Mo–HIS	Mo–O^{OH}	Mo–COO^-	Average MoO/N	ΔMoO/N
Experimental crystal E_0[10]	··	2.670	2.680	2.730	2.693	··	2.330	2.160	2.190	2.227	··
Experimental EXAFS E_0[83]	··	··	··	··	2.710	··	··	··	··	2.200	··
Experimental EXAFS E_1[83]	··	··	··	··	2.65–2.66	−0.050	··	··	··	2.130	−0.070
BP86 E_0 $O^{OH}H$	··	2.667	2.633	2.668	2.754	··	2.297	2.138	2.191	2.209	··
BP86 E_1 $O^{OH}H$	··	2.653	2.634	2.677	2.820	0.066	2.309	2.208	2.224	2.247	0.038
TPSSh E_0 $O^{OH}H$	··	2.721	2.646	2.676	2.738	··	2.293	2.111	2.170	2.191	··
TPSSh E_1 $O^{OH}H$	2.09	2.700	2.639	2.675	2.810	0.072	2.307	2.198	2.205	2.236	0.045
TPSSh E_1 O^1H	7.93	2.716	2.662	2.705	2.834	0.096	2.311	2.077	2.224	2.204	0.013
TPSSh E_1 O^2H	55.57	2.716	2.677	2.715	2.784	0.046	2.311	1.977	2.210	2.166	−0.025
TPSSh E_1 O^3H	57.32	2.707	2.673	2.716	2.783	0.045	2.307	1.987	2.184	2.159	−0.032
TPSSh E_1 Fe_6H	23.07	2.723	2.678	2.720	2.789	0.051	2.305	1.978	2.181	2.155	−0.036
TPSSh E_1 S_1H	10.22	2.706	2.684	2.719	2.769	0.031	2.318	1.987	2.183	2.163	−0.029
TPSSh E_1 μ^2–HFe_2–S_2	24.59	2.699	2.671	2.742	2.779	0.041	2.309	1.976	2.190	2.158	−0.033
TPSSh E_1 S_3H	8.42	2.724	2.687	2.703	2.775	0.037	2.324	1.985	2.182	2.164	−0.028
TPSSh E_1 S_4H	24.08	2.688	2.675	2.737	2.718	−0.020	2.319	2.006	2.178	2.168	−0.023
TPSSh E_1 μ^2–HFe_6–S_5	24.23	2.756	2.713	2.709	2.773	0.035	2.268	1.973	2.195	2.145	−0.046
TPSSh E_1 S_6H	20.08	2.692	2.726	2.769	2.765	0.027	2.319	1.955	2.189	2.154	−0.037
TPSSh E_1 S_7H	5.57	2.650	2.674	2.721	2.789	0.051	2.322	1.993	2.190	2.168	−0.023
TPSSh E_1 S_8H	21.28	2.733	2.697	2.655	2.771	0.033	2.319	1.972	2.175	2.155	−0.036
TPSSh E_1 S_9H	18.57	2.731	2.641	2.741	2.783	0.046	2.291	1.973	2.191	2.152	−0.039
TPSSh $E_1, M_S = 1$ $O^{OH}H$	0	2.923	2.682	2.727	2.800	0.062	2.270	2.189	2.191	2.217	0.025
TPSSh $E_1, M_S = 1$ O^1H	10.00	2.738	2.659	2.704	2.826	0.088	2.305	2.080	2.217	2.201	0.010
TPSSh $E_1, M_S = 1$ S_3H	11.65	2.717	2.676	2.704	2.773	0.035	2.314	1.982	2.181	2.159	−0.032
TPSSh $E_1, M_S = 1$ S_7H	6.61	2.676	2.676	2.730	2.772	0.034	2.311	1.997	2.175	2.161	−0.030
TPSSh $E_1, M_S = 1$ Fe_6H	34.06	2.700	2.686	2.725	2.787	0.050	2.310	1.973	2.183	2.155	−0.036
TPSSh E_1 μ^6–$CHFe_{2456}$ face	17.35	2.732	2.689	2.713	2.755	0.017	2.315	1.976	2.179	2.157	−0.035
TPSSh E_1 μ^2–HFe_6–C & $O^{OH}H$ dehydrated	··	2.809	2.658	2.697	2.755	0.017	2.300	2.105	2.189	2.198	0.007
TPSSh E_1 S_7H & μ^2–HFe_6–Fe_2 dehydrated	··	2.660	2.706	2.798	2.772	0.034	2.314	1.937	2.176	2.142	−0.049
TPSSh E_1 μ^2–HFe_6–Fe_2 no S_7	··	2.654	2.730	2.754	3.869	1.132	3.343	2.122	4.140	3.202	1.011
TPSSh E_1 no proton	··	2.700	2.639	2.675	2.810	0.072	3.355	2.198	2.205	2.586	0.395
TPSSh E_0 $Fe_2\downarrow Fe_3\downarrow Fe_5\downarrow$ $O^{OH}H$	··	2.654	2.655	2.757	2.733	−0.005	2.301	2.102	2.168	2.191	−0.001
TPSSh E_0 $Fe_2\downarrow Fe_4\downarrow Fe_7\downarrow$ $O^{OH}H$	··	2.647	2.724	2.687	2.734	−0.004	2.301	2.098	2.176	2.192	0.001
TPSSh $[MoFe_7S_9C]^{-3}$	··	2.715	2.631	2.685	2.840	0.102	2.312	2.217	2.232	2.253	0.062
TPSSh $[MoFe_7S_9C]^{+1}$	··	2.664	2.721	2.737	2.688	−0.050	2.277	2.030	2.113	2.140	−0.051

It is of interest that the lowest local minimum found had $S = 1$, see figure 5.2(a), a sign that more computations with that spin should be explored. Of the geometry which shows a difference in bonds in accordance with experiment, only one shows a shortening of the average Mo–Fe bonds as can be seen in table 5.1 and in figure 5.2(b). A proton on S_4 that can be seen in figure 5.2(b) had a prominent effect on the average Mo–O/N bonds as did a bridging hydride between Fe_6–Fe_2 when accompanied with a protonated S_7 which is shown in figure 5.3(a). It has to be noted that to keep the system with the same total charge, the proton on His195 was removed. This complicates the calculation of relative energies of the system and will require additional computations.

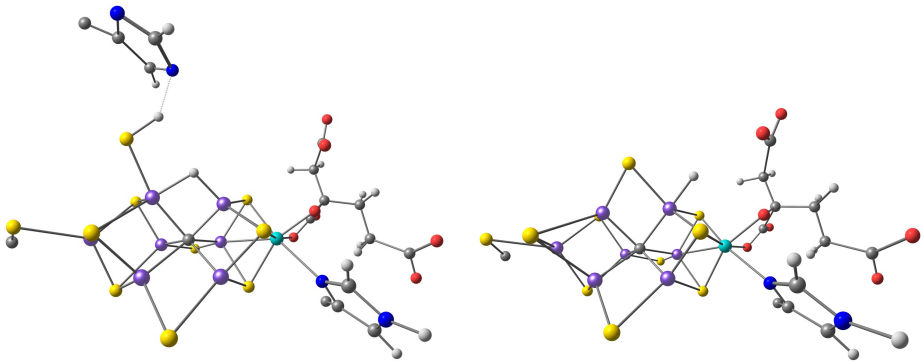


(a) E_1 state with $M_S = 1$ and a proton on the usual alcohol oxygen (O^{OH})

(b) E_1 with S_4 protonated

Figure 5.2: Interesting E_1 models

This was also the case for the bridging iron-carbon hydride with a proton on the usual alcohol oxygen (O^{OH}), shown in figure 5.4. This geometry is of special interest as carbon-iron bridging hydrogen have been found before and that complex consisted of low-spin irons[88,89]. Without knowing how favourable the protonation of the hexavalent carbon is one can only speculate that this bridging hydride can perhaps transfer to Fe_6, the geometry of which can be seen in figure 5.3(b), and from there partake in reactions or just go straight to a substrate to reduce it, but as can be seen in table 5.1 protonating Fe_6 is uphill, energy wise.



(a) E_1 state with a proton on S_7 and a bridging hydride between iron Fe_2 and Fe_6

(b) E_1 with a terminal hydride on Fe_6

Figure 5.3: Interesting E_1 models continuation

One might also speculate that the longest carboxylic arm of the homocitrate, lowers the barrier for a proton to travel to S_4, figure 5.2(b), as homocitrate plays a significant role for dinitrogen reduction[71] and CO can inhibit proton reduction if it homocitrate changed to citrate[1]. A hydride on Fe_6, as in figure 5.3(b), at the same time can attack the proton to release dihydrogen if the alkoxy group is reprotonated. So this is a hypothesis to be tested and could possible explain the hydrogen evolution.

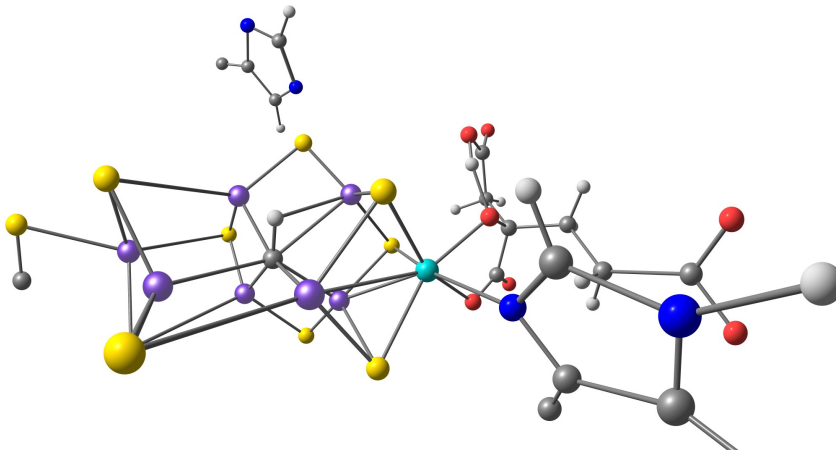
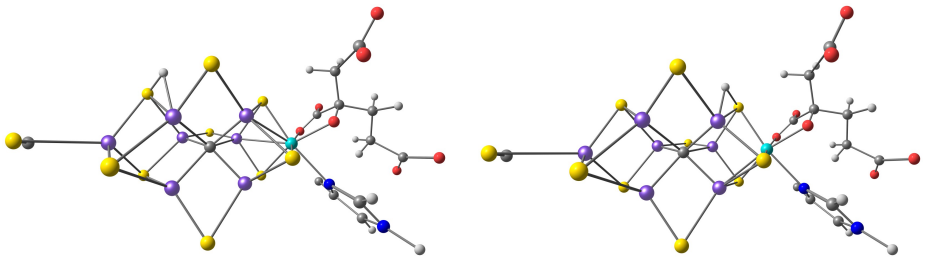


Figure 5.4: E_1 state with a bridging iron-carbon hydride, μ^2–HFe_6–C, and a proton on the alcohol group of homocitrate

5. E_1 structure

Another interesting avenue of inquire is a composition of bridging sulfur-iron hydrogen as both μ^2–HFe_2–S_2, figure 5.5(a), and μ^2–HFe_6–S_5, figure 5.5(b), hydrides showed some contraction of the average Mo–O/N bonds and are symmetrical on the cluster. This could relate to the reversible photoinduced reductive elimination of Hoffman *et al.*[74] where symmetrical hydrides played a part in their explanation of the photoinduced effect of the E_4 state.



(a) E_1 state with a bridging iron-sulfur hydride, μ^2–HFe_2–S_2

(b) E_1 state with a bridging iron-sulfur hydride, μ^2–HFe_6–S_5

Figure 5.5: Interesting E_1 models continuation

As the lowest energy geometry did not show changes it is likely that more calculations are required to locate the true E_1 state consistent with the EXAFS data. Computation of the Mössbauer isomer shifts for all E_1 models and comparison to the experimental data from Münck *et al.*[85] is also something to explore. We would also like to check if the spin-flip pattern has perhaps changed which was the case in previous sections. Such a investigation might lead to a lower BS solution and perhaps the true E_1 state.

6. P-cluster

The P-cluster is a homometallic iron sulfur cofactor which is thought to mediate electrons from the Fe-protein to FeMoco. In the process it goes from its resting state, P^N, to a singly oxidised state, P^{1+}. Under turnover conditions the reaction is as follows:

$$\underset{S=0}{\mathrm{P^N}} \underset{Fe_{red}}{\overset{FeMoco}{\rightleftharpoons}} \underset{S=\frac{5}{2},S=\frac{1}{2}}{\mathrm{P^{1+}}} \tag{6.1}$$

All iron atoms of the resting state P-cluster are in the ferrous state[15,16]. At pH 8 the redox potential is −309 mV and is pH dependent[90]. A singly oxidised P-cluster has a mixed EPR signal with $S = \frac{5}{2}$ and $S = \frac{1}{2}$ with the former in majority and based on g-factors of the signal there are two $S = \frac{5}{2}$ states[91,92]. The P-cluster is slower to give FeMoco the electron then to get another from Fe-protein[5,91] and is thus a very short time oxidised. This cycle is known to repeat one electron at a time[92]. This is puzzling as isolated studies find that at normal nitrogenase pH (7.0–7.4)[79] it is favourable for the P-cluster to give two electrons[90]. The binding of Fe-protein might thus have a local pH effect to prevent this as it is apparent that under turnover the P-cluster only undergoes one electron oxidation.

As the P-cluster is the reducing agent for FeMoco we attempted to calculate the redox potentials from DFT calculations of the resting state and singly oxidised P^{1+} states of the cluster. This included increasing the size of the QM region to improve the accuracy of the computations. We also looked for any changes in the protein environment.

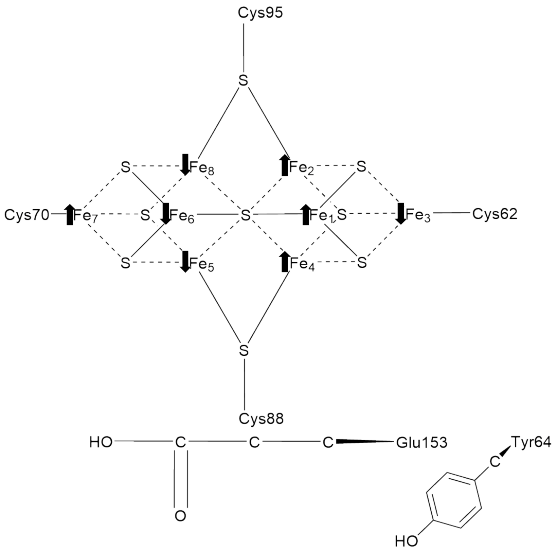
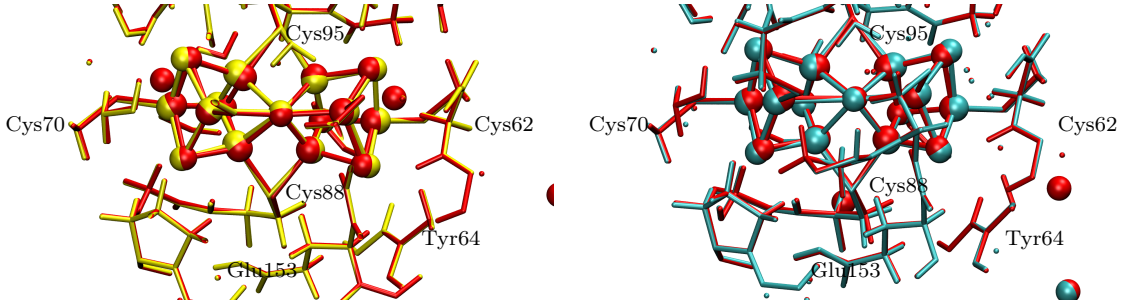


Figure 6.1: Simplified spin diagram of P-cluster with some important amino acid residues. For clarity, residue Cys154 ligating Fe_1 and Cys153 ligating Fe_6 are omitted from diagram

6.1. Results

A P-cluster QM/MM model was made similar to the FeMoco also by B. Benediktsson[47]. The QM region included the nearest residues such as serine, threonine and glutamic acid. The spin-flip $Fe_3\downarrow Fe_5\downarrow Fe_6\downarrow Fe_8\downarrow$ was used for all P-cluster calculations and the resting state, P^N, was calculated as $M_S = 0$ and P^{1+} as $M_S = \frac{1}{2}$ and $M_S = \frac{5}{2}$.

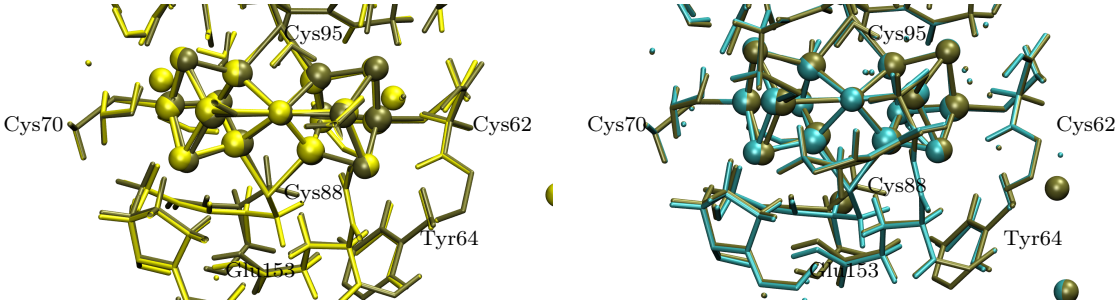


(a) P^{1+} $M_S = \frac{1}{2}$ BP86 (red) vs. TPSSh (yellow)

(b) P^{1+} $M_S = \frac{5}{2}$ BP86 (red) vs. TPSSh (cyan)

Figure 6.2: Overlay figures of the P-cluster: Oxidised states with different functionals

The geometry of the one electron oxidised P-cluster is similar when comparing the functionals as can be seen in figures 6.2(a) and 6.2(b). A comparison of the spin states versus the resting state with the TPSSh functional showed little as these geometry are almost the same as can be seen in figure 6.3(a) and 6.3(b). It should be noted that we started with a smaller QM region and the BP86 functional which did not predict almost degenerate $M_S = \frac{1}{2}$ and $M_S = \frac{5}{2}$ states. This energy gap got smaller when the QM region was enlarged. TPSSh on the other hand with the larger QM region predicted that the $M_S = \frac{1}{2}$ was the lowest of the P^{1+} states in clear contradiction to experiment. This might be interpreted thus that TPSSh does not delocalize the electrons enough, and in this case BP86 is giving a better description of the system. It is also possible that we are not finding the correct $M_S = \frac{5}{2}$ or $M_S = \frac{1}{2}$ BS solutions as EPR reveal two different $S = \frac{5}{2}$ signals. It is of note that the P^N state with TPSSh had a peculiar Mulliken spin population which no other calculation showed. Fe_5 and Fe_1 had almost no spin population and clearly coupled together (same spin density, opposite sign). Most likely we have yet to find the right spin-flip configuration for this system and this issue will be explored more thoroughly in the future.



(a) P^{1+} $M_S = \frac{1}{2}$ (yellow) vs. P^N (brown) TPSSh

(b) P^{1+} $M_S = \frac{5}{2}$ (cyan) vs. P^N (brown) TPSSh

Figure 6.3: Overlay figures of the P-cluster: Oxidised vs. resting state

As can be seen in figure 6.3(a) and 6.3(a), there is an interesting change in the position of the Cys88 ligand and the slight alteration of the carboxylic arm of Glu153 residue is also noteworthy. Interestingly the Tyr64 residue deviated noticeable but the tyrosine residue was not included in the QM region. In all these calculations the MM energy changed between the optimizations and this is indicative that we do not have a large enough QM region and we will enlarge the region systematically as we continue this research. The comparison with crystal structures by Spatzal *et al.*[10] and Tezcan *et al.*[14] in figure 6.4 was done by manually superimposing the structures to best fit the P^N TPSSh optimized QM/MM geometry. The crystal structures by Spatzal *et al.*[10] is the crystal structure our model was built from and is of a MoFe protein from *A. vinelandii*, while Tezcan *et al.*[14] is of $Fe_{red}(ATP)_2MoFe$ complex also from *A. vinelandii*.

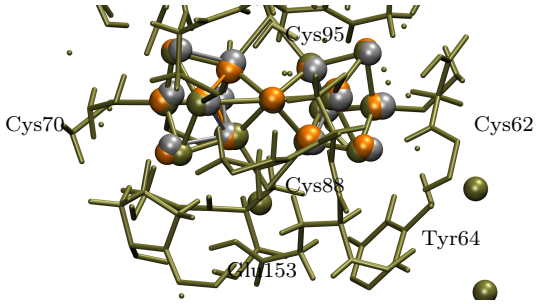


Figure 6.4: Overlay figure of the P-cluster: P^N TPSSh (brown) vs. crystal structures (Spatzal[10], orange and Tezcan[14], grey)

As can be seen in figure 6.4 there are some noticeable difference when the crystal structures are overlaid on the P^N TPSSh geometry and might it indicate that the

spin-flip configuration we used was inadequate since it deviates from both of the crystal structures and are further studies under way. Overall the geometry of the P^N state is in agreement with the crystal structures as the resolution of these structures is limited.

7. Hydrazine reduction

Nitrogenase reduces hydrazine[93] with two protons and two electrons as shown by the following equation:

$$N_2H_4 + 2\,e^- + 2\,H^+ \xrightarrow{N_2ase} 2\,NH_3 \tag{7.1}$$

Hydrazine is the only single bond species that nitrogenase reduces. It is probable that hydrazine is also an intermediate in dinitrogen reduction as hydrazine is released with acid or base quenching during N_2 turnover[1] and mutated MoFe protein results in hydrazine as a byproduct as does VFe protein[94]. It has also been reported that hydrazine inhibits both dinitrogen[30], diazene[32] (another likely intermediate of dinitrogen reduction) and hydrazine[30] reduction. It inhibits these reactions by binding to the active site[93] and lowering the electron flux[93] through FeMoco. One other aspect of hydrazine inhibition is because it can be protonated to yield hydrazinium (pK_a = 8.1)[95], which is not a substrate[96], and hydrazine is thus a better substrate at higher pH. Hydrazine binds to the E_1 LT state[25] and the only intermediate detected is a terminal amide (M–NH_3), based on EPR signal during turnover[97]. It is also of note that EPR shows a $S = \frac{1}{2}$ signal with a mutated MoFe protein[93] during turnover of hydrazine.

Prior work we did on a system similar to FeMoco, synthesized by Coucouvanis *et al.*[98], could be insightful. Both these systems have a ground state with $S = \frac{3}{2}$ EPR signal and consist of tetrahedral irons, μ^3-sulfides and a octahedral molybdenum. The [$MoFe_3S_4$] cubanes are also catalytic and can reduce some of the same substrates as nitrogenase[98–101]. Our result was that hydrazine would only bind to Mo but not to Fe[102].

The binding site and mechanism of hydrazine in nitrogenase is not known and this was explored by QM/MM calculations.

7.1. Results

In these calculations we used the BP86 functional with $M_S = \frac{3}{2}$ for the ground state and $M_S = 2$ for the singly reduced FeMoco. The $Fe_3\downarrow Fe_4\downarrow Fe_6\downarrow$ BS solution was used. We investigated possible hydrazine binding to molybdenum and found that weak binding was possible with both a singly reduced FeMoco and the ground state of FeMoco. As can be seen in figure 7.1 the Mo–N bond length between was 2.29 Å where one of the homocitrate arms has dislocated to allow access to Mo. We approximated the binding energy by calculating hydrazine with the same level of theory in the gas phase and using relative energies of product minus the reactants. The binding energy, when calculated, is indicated by ΔE. This turned out to be slightly downhill for the ground state; ΔE = −9.13 kcal/mol, and also for singly reduced FeMoco; ΔE = −6.78 kcal/mol. Interestingly the Mo–N bond length is not affected by reduction as can be seen in figure 7.2.

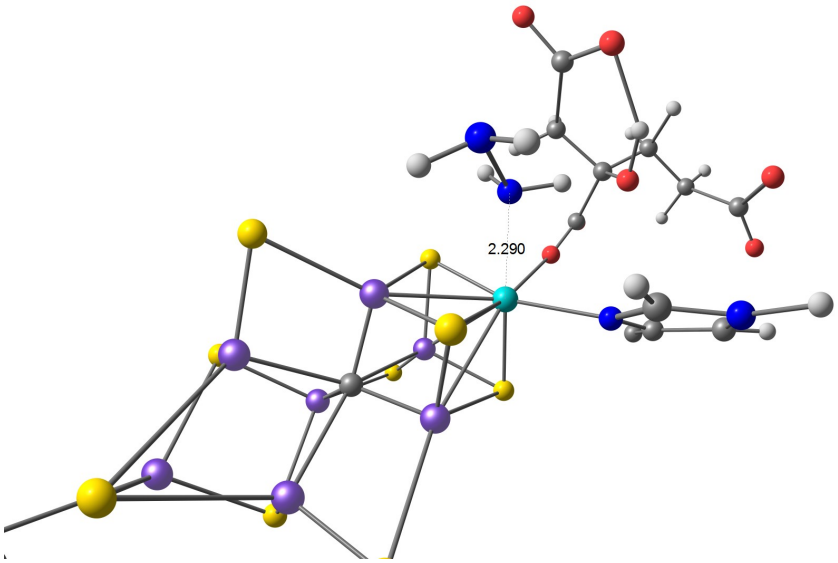


Figure 7.1: Ground state FeMoco N_2H_4–Mo adduct; ΔE = −9.13 kcal/mol. Bond lengths in Å

Surprisingly we found hydrazine able to bind, weakly, to Fe_6 in the ground state as can be seen in figure 7.3. Calculation showed this to be slightly uphill; ΔE = 2.27 kcal/mol. This is in stark contrast to the seemingly weak binding (based on the Fe–N bond length), as can be seen from the bond length in figure 7.4, to iron Fe_6 that was calculated to be downhill; ΔE = −11.49 kcal/mol. This larger binding energy is probably due to stronger hydrogen bonding and dispersion effects. In

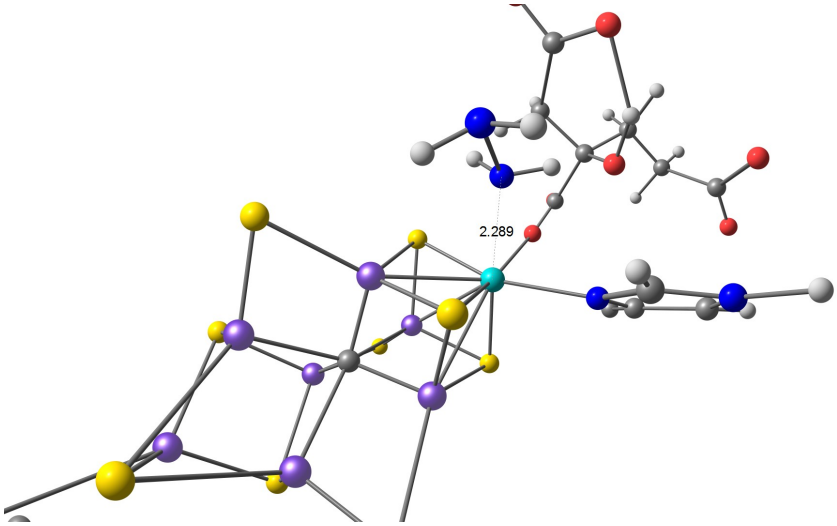


Figure 7.2: Singly reduced FeMoco N_2H_4–Mo adduct; ΔE = −6.78 kcal/mol. Bond lengths in Å

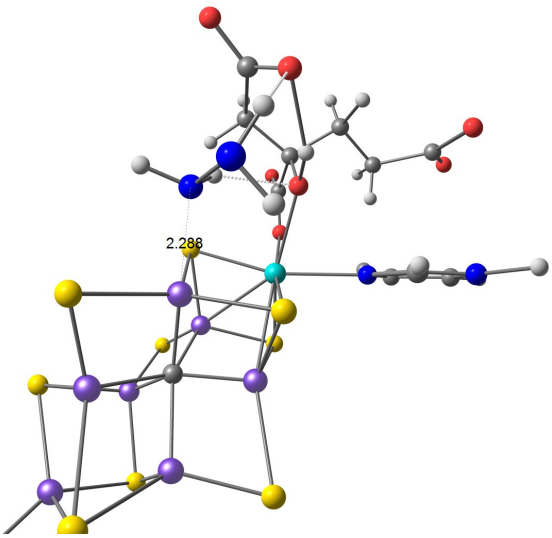


Figure 7.3: Ground state FeMoco N_2H_4–Fe adduct; ΔE = 2.27 kcal/mol. Bond lengths in Å

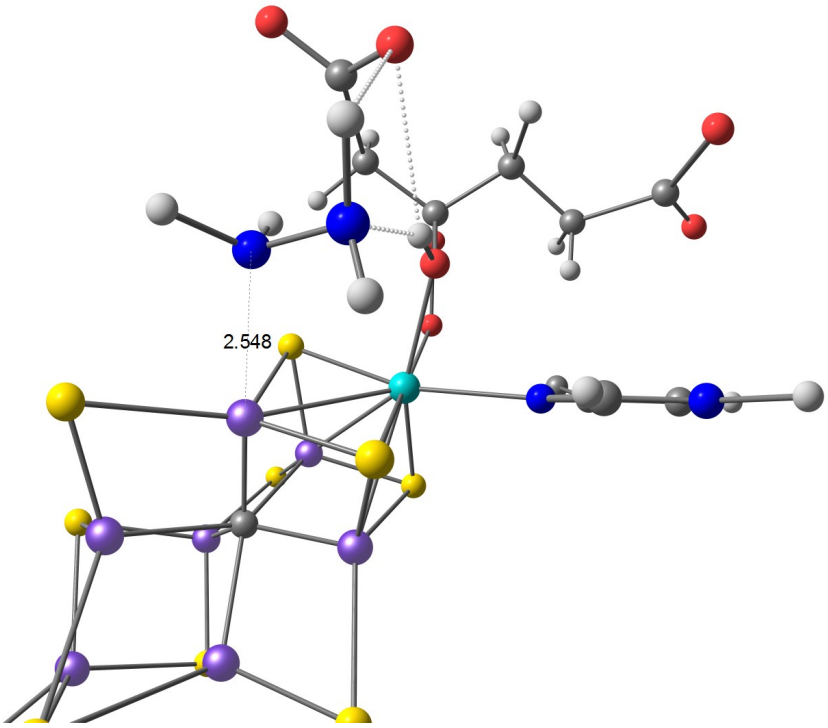
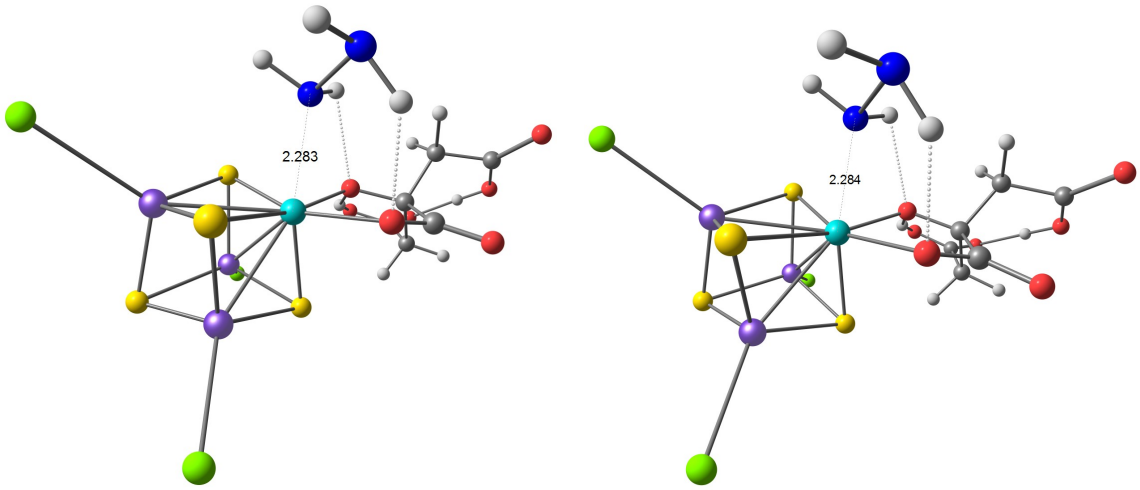


Figure 7.4: Ground state FeMoco with weakly bound N_2H_4; ΔE = −11.49 kcal/mol. Bond lengths in Å



(a) Ground state $[MoFe_3S_4Cl_3(Hcit)]^{3-}$ with N_2H_4–Mo adduct; ΔE = −16.37 kcal/mol

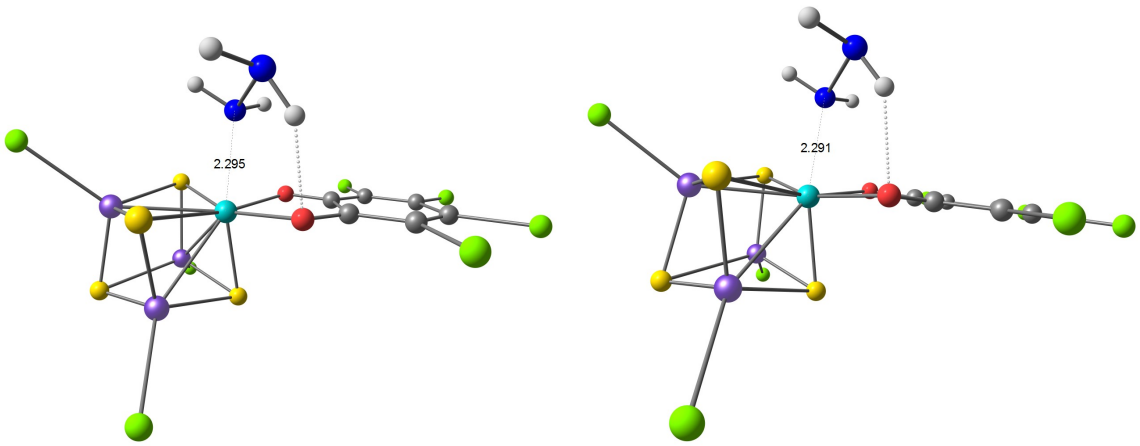
(b) Singly reduced $[MoFe_3S_4Cl_3(Hcit)]^{3-}$ with N_2H_4–Mo

Figure 7.5: $[MoFe_3S_4]$ cubanes synthesized by Coucouvanis et al.[98]. Computations by us[102]. Bond lengths in Å

all probability all hydrazine would be thus bound in the ground state of FeMoco, which is not consistent with experiment. Lack of thermal and entropic effects could explain this.

Our work with the $[MoFe_3S_4]$ cubanes showed quite different results[102] and comparison between systems is informative. Despite being very similar there are differences. For example although the bond lengths were the same (see figures 7.5(a), 7.5(b), 7.6(a) and 7.6(b)) the relative energy of binding was intriguingly different. Hydrazine binding energy to the $[MoFe_3S_4Cl_3]$ cubane depends on the ligand and was much more downhill when citrate (called Hcit here) was the ligand than tetrachlorocatecholate (3,4,5,6-tetrachlorobenzene-1,2-bis(olate) called Cl_4–cat here), ΔE = −16.37 kcal/mol vs. ΔE = −12.13 kcal/mol respectively.

Another contrasting difference is that with $[MoFe_3S_4]$ cubanes, hydrazine would not bind to any of the irons with or without added electrons. In FeMoco, however, we found the strongest hydrazine binding to be to the singly reduced FeMoco on Fe_6, as can be seen in figure 7.7, calculated to be downhill; ΔE = −8.53 kcal/mol and with a short bond length of 2.12 Å. As this was a promising candidate as the first step of the mechanism we calculated a better approximation of a non bound hydrazine in the protein as can be seen in figure 7.9. With this approxima-



(a) Ground state $[MoFe_3S_4Cl_3(Cl_4\text{–cat})(N_2H_4)]^{2-}$; ΔE = −12.13 kcal/mol

(b) Singly reduced $[MoFe_3S_4Cl_3(Cl_4\text{–cat})(N_2H_4)]^{2-}$

Figure 7.6: $[MoFe_3S_4]$ cubanes synthesized by Coucouvanis et al.[98]. Computations by us[102]. Bond lengths in Å

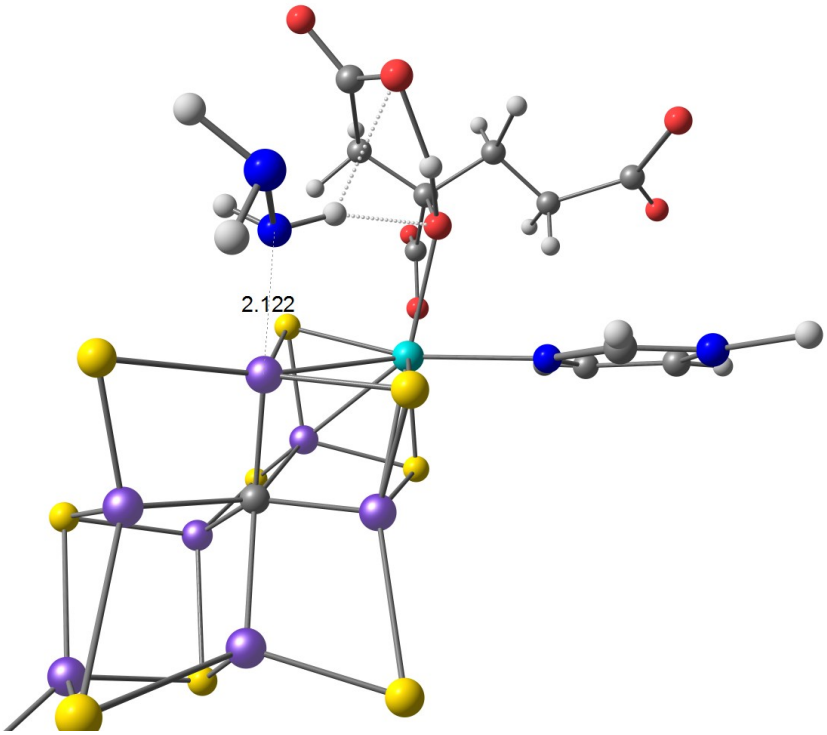


Figure 7.7: Singly reduced FeMoco N_2H_4–Fe adduct; ΔE = −4.59 kcal/mol. Bond lengths in Å

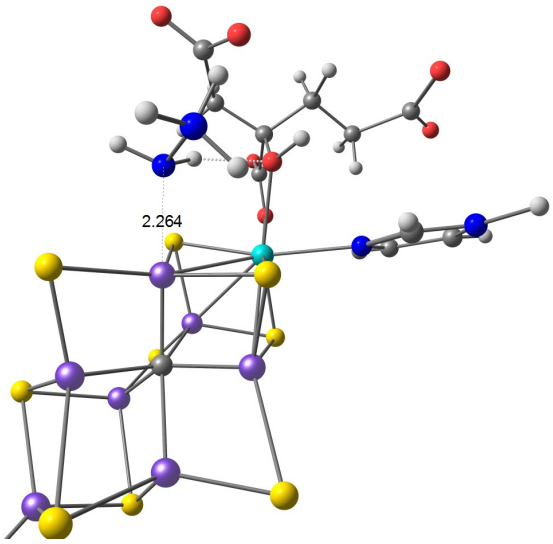


Figure 7.8: Singly reduced FeMoco $N_2H_5^+$–Fe adduct; ΔE_{proton} = −22.59 kcal/mol. Bond lengths in Å

tion the binding turned out to be downhill as before; ΔE = −4.59 kcal/mol which is much more reliable than previous calculation. Protonation of this Fe bound hydrazine was according to the glutamic acid approximation described earlier, is downhill; ΔE_{proton} = −22.59 kcal/mol. We tried to see if at this stage the N–N bond could be broken and found the minimum in figure 7.10 which gave a bond breaking energy of; ΔE_{bond} = −20.30 kcal/mol. Future calculation will attempt to locate the saddle point and estimation of the barrier with nudge elastic band (NEB) calculations[103–107]. It seems from the geometry of the M–NH_2 that protonation may be favourable as the amine adduct appears to have a lone pair as was the case in our $[MoFe_3S_4Cl_3(Cl_4\text{–cat})(NH_2)]^{1-}$ calculation (see figure 7.11). Interestingly this was not as apparent with citrate as a ligand (see figure 7.12) so further studies must be done to see if this is the case.

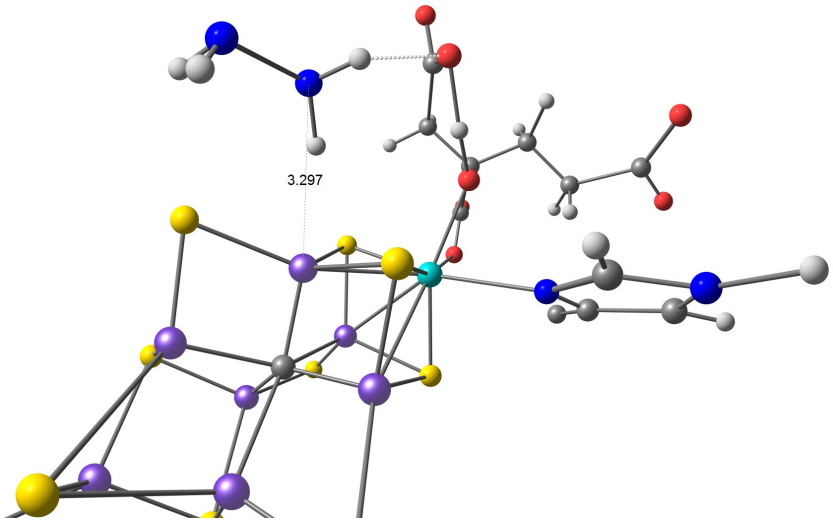


Figure 7.9: Singly reduced FeMoco + unbound N_2H_4. Bond lengths in Å

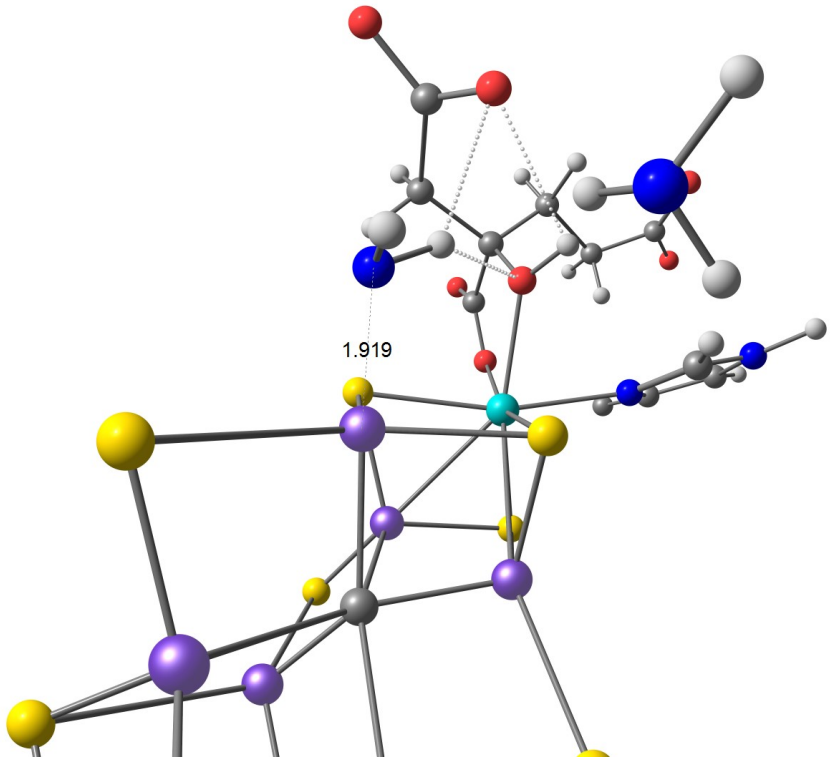


Figure 7.10: Structure of singly reduced FeMoco NH_2–Fe + NH_3 after N–N bond breaking. ΔE_{bond} = −20.30 kcal/mol. Bond lengths in Å

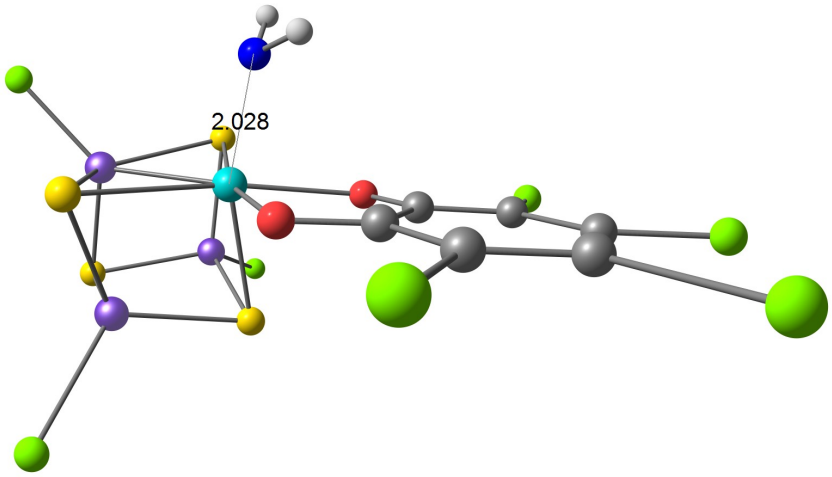


Figure 7.11: Structure of singly reduced $[MoFe_3S_4Cl_3(Cl_4\text{-}cat)(NH_2)]^{2-}$ (After N–N bond break without NH_3). $[MoFe_3S_4]$ cubanes synthesized by Coucouvanis et al.[98]. Computations by us[102]. Bond lengths in Å

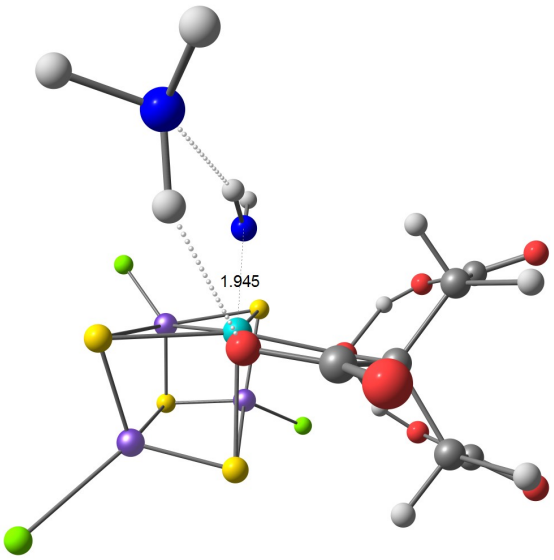


Figure 7.12: Structure of singly reduced $[MoFe_3S_4Cl_3(Hcit)(NH_2)(NH_3)]^{2-}$ (After N–N bond break with NH_3). $[MoFe_3S_4]$ cubanes synthesized by Coucouvanis et al.[98]. Computations by us[102]. Bond lengths in Å

8. Conclusions

In conclusion we have successfully used a QM/MM model of the MoFe protein of nitrogenase to explore the electronic structure of FeMoco in multiple redox states and the binding of various substrates. In investigating the electronic structure we found that the BP86 functional appears not to compute the spin states as well as TPSSh and further endeavour should probably use the TPSSh functional to research nitrogenase.

We uncovered a change in the lowest energy BS solution if an adduct binds to FeMoco, at sulfur S_7 and iron Fe_6, resulting in a magnetic coupling between irons Fe_6 and Fe_2, changing the total spin of FeMoco from $S = \frac{3}{2}$ to $S = \frac{1}{2}$, as revealed by experiments. Such binding may explain $S = \frac{1}{2}$ states found in other substrate reduction experiments.

The structure of the singly reduced E_1 state was explored. We believe future work will reveal the nature of this singly reduced state but unfortunately this work could not be completed. Nevertheless preliminary results show that most likely a bridging hydride can possibly account for a shortening of metal-metal and metal-ligand bonds making FeMoco appear as if oxidised.

Our study of the structure and chemistry of the P-cluster show that a thorough investigation into the BS solutions is needed to get insight into the workings of the P-cluster. Another interesting avenue in that regard is the question of geometric change, both in the structure of the cluster and the protein environment, which will be further studied and might have significance in understanding this system.

The comparison of synthetic $[MoFe_3S_4]$ cubanes and FeMoco showed that the protein environment plus the hexavalent carbon change the system to a large extent. Our preliminary result thus show that hydrazine can bind to both Fe_6 and Mo in the ground state (possible explaining inhibition effects) but is only reduced by binding to the singly reduced FeMoco's Fe_6. Future work will include, finishing the cycle of hydrazine reduction and explore new BS solutions of FeMoco in accordance with the preceding work presented here.

This enzyme remains enigmatic and attempts to understand how it accomplishes dinitrogen reduction has prompted many questions. In our view, careful examination of the experimental data with complementary computational methodology can reunite enzyme and model chemistry in such a way as to clear the way to understanding the "Everest of enzymes" and thus finding the holy grail of chemistry[108].

Bibliography

[1] B. K. Burgess and D. J. Lowe, *Chem. Rev.*, 1996, **96**, 2983–3012.

[2] J. W. Erisman, M. A. Sutton, J. Galloway, Z. Klimont and W. Winiwarter, *Nat. Geosci*, 2008, **1**, 636–639.

[3] B. M. Hoffman, D. R. Dean and L. C. Seefeldt, *Acc. Chem. Res.*, 2009, **42**, 609–619.

[4] P. W. Ludden, in *Nitrogenase Complex*, John Wiley & Sons, Ltd, 2001.

[5] S. Duval, K. Danyal, S. Shaw, A. K. Lytle, D. R. Dean, B. M. Hoffman, E. Antony and L. C. Seefeldt, *Proc. Natl. Acad. Sci.*, 2013, **110**, 16414–16419.

[6] J. Kästner, *PhD thesis*, *Biological Nitrogen Fixation — Simulation of the Reaction Mechanism of Nitrogenase from First Principles*, Clausthal University of Technology, 2004.

[7] B. Smith, M. Durrant, S. Fairhurst, C. Gormal, K. Grönberg, R. Henderson, S. Ibrahim, T. L. Gall and C. Pickett, *Coord. Chem. Rev.*, 1999, **185–186**, 669 – 687.

[8] D. N. Pham and B. K. Burgess, *Biochem.*, 1993, **32**, 13725–13731.

[9] J. M. Rivera-Ortiz and R. H. Burris, *J. Bacteriol.*, 1975, **123**, 537–545.

[10] T. Spatzal, M. Aksoyoglu, L. Zhang, S. L. A. Andrade, E. Schleicher, S. Weber, D. C. Rees and O. Einsle, *Science*, 2011, **334**, 940–940.

[11] K. M. Lancaster, M. Roemelt, P. Ettenhuber, Y. Hu, M. W. Ribbe, F. Neese, U. Bergmann and S. DeBeer, *Science*, 2011, **334**, 974–977.

[12] W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph.*, 1996, **14**, 33–38.

[13] J. Stone, *MSc thesis*, Computer Science Department, University of Missouri-Rolla, 1998.

[14] F. A. Tezcan, J. T. Kaiser, D. Mustafi, M. Y. Walton, J. B. Howard and D. C. Rees, *Science*, 2005, **309**, 1377–1380.

[15] M. K. Johnson, A. J. Thomson, A. Robinson and B. E. Smith, *Biochimica et Biophys. Acta (BBA) - Protein Struct.*, 1981, **671**, 61 – 70.

[16] P. A. Lindahl, V. Papaefthymiou, W. H. Orme-Johnson and E. Münck, *J. Biol. Chem.*, 1988, **263**, 19412–19418.

[17] R. R. Eady, *Chem. Rev.*, 1996, **96**, 3013–3030.

[18] G. J. Leigh, *Eur. J. Biochem.*, 1995, **229**, 14–20.

[19] R. N. F. Thorneley and D. J. Lowe, in *Molybdenum Enzymes; Spiro*, ed. T. G. Spiro, Wiley, pp. 221–284.

[20] W. H. Orme-Johnson, in *The Molybdenum—Iron Protein of Nitrogenase*, ch. 18, pp. 257–270.

[21] D. Lukoyanov, Z.-Y. Yang, N. Khadka, D. R. Dean, L. C. Seefeldt and B. M. Hoffman, *J. Am. Chem. Soc.*, 2015, **137**, 3610–3615.

[22] L. C. Seefeldt, Z.-Y. Yang, S. Duval and D. R. Dean, *Biochimica et Biophys. Acta (BBA) - Bioenerg.*, 2013, **1827**, 1102 – 1111.

[23] Z.-Y. Yang, V. R. Moure, D. R. Dean and L. C. Seefeldt, *Proc. Natl. Acad. Sci.*, 2012, **109**, 19644–19648.

[24] C. C. Lee, Y. Hu and M. W. Ribbe, *Angew. Chem. Int. Ed.*, 2015, **54**, 1219–1222.

[25] B. M. Hoffman, D. Lukoyanov, Z.-Y. Yang, D. R. Dean and L. C. Seefeldt, *Chem. Rev.*, 2014, **114**, 4041–4062.

[26] Z.-Y. Yang, N. Khadka, D. Lukoyanov, B. M. Hoffman, D. R. Dean and L. C. Seefeldt, *Proc. Natl. Acad. Sci.*, 2013, **110**, 16327–16332.

[27] P. M. C. Benton, J. Christiansen, D. R. Dean, and L. C. Seefeldt, *J. Am. Chem. Soc.*, 2001, **123**, 1822–1827.

[28] T. Spatzal, K. A. Perez, O. Einsle, J. B. Howard and D. C. Rees, *Science*, 2014, **345**, 1620–1623.

[29] M. E. Rasche and L. C. Seefeldt, *Biochem.*, 1997, **36**, 8574–8585.

[30] B. K. Burgess, S. Wherland, W. E. Newton and E. I. Stiefel, *Biochem.*, 1981, **20**, 5140–5146.

[31] T. Spatzal, K. A. Perez, J. B. Howard and D. C. Rees, *eLife*, 2015, **4**, e11620.

[32] B. M. Barney, J. McClead, D. Lukoyanov, M. Laryukhin, T.-C. Yang, D. R. Dean, B. M. Hoffman and L. C. Seefeldt, *Biochem.*, 2007, **46**, 6784–6794.

[33] *The Computational Chemistry Course at University of Iceland*, 2015-12-12, `https://sites.google.com/site/compchemui/`.

[34] W. Koch and M. C. Holthausen, in *A Chemist's Guide to Density Functional Theory*, Wiley-VCH Verlag GmbH, 2001, pp. 3–18.

[35] I. N. Levine, in *Quantum Chemistry*, Prentice-Hall of India, 1999, pp. 290–295.

[36] F. Neese, *Coord. Chem. Rev.*, 2009, **253**, 526 – 563.

[37] A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098–3100.

[38] J. P. Perdew, *Phys. Rev. B*, 1986, **33**, 8822–8824.

[39] A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.

[40] C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785–789.

[41] J. Tao, J. P. Perdew, V. N. Staroverov and G. E. Scuseria, *Phys. Rev. Lett.*, 2003, **91**, 146401.

[42] J. A. D. MacKerell, D. Bashford, M. Bellott, J. R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B*, 1998, **102**, 3586–3616.

[43] A. D. Mackerell, M. Feig and C. L. Brooks, *J. Comput. Chem.*, 2004, **25**, 1400–1415.

[44] S. Metz, J. Kästner, A. A. Sokol, T. W. Keal and P. Sherwood, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2014, **4**, 101–110.

[45] F. Neese, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2012, **2**, 73–78.

[46] I. T. Todorov, W. Smith, K. Trachenko and M. T. Dove, *J. Mater. Chem.*, 2006, **16**, 1911–1918.

[47] B. Benediktsson, *B.Sc. thesis*, *Understanding Nitrogenase: A Computational Model of the MoFe Protein from Azotobacter vinelandii*, University of Iceland, 2015.

[48] E. van Lenthe, E. J. Baerends and J. G. Snijders, *J. Chem. Phys.*, 1993, **99**, 4597–4610.

[49] C. van Wüllen, *J. Chem. Phys.*, 1998, **109**, 392–399.

[50] F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3297–3305.

[51] D. A. Pantazis, X.-Y. Chen, C. R. Landis and F. Neese, *J. Chem. Theory Comput.*, 2008, **4**, 908–919.

[52] K. Eichkorn, F. Weigend, O. Treutler and R. Ahlrichs, *Theor. Chem. Accounts*, 1997, **97**, 119–124.

[53] F. Neese, F. Wennmohs, A. Hansen and U. Becker, *Chem. Phys.*, 2009, **356**, 98 – 109.

[54] S. Kossmann and F. Neese, *Chem. Phys. Lett.*, 2009, **481**, 240 – 243.

[55] R. Izsák and F. Neese, *J. Chem. Phys.*, 2011, **135**,.

[56] T. Petrenko, S. Kossmann and F. Neese, *J. Chem. Phys.*, 2011, **134**,.

[57] S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem.*, 2011, **32**, 1456–1465.

[58] S. Grimme, *J. Chem. Phys*, 2010, **132**, 154104.

[59] A. Klamt and G. Schüürmann, *J. Chem. Soc., Perkin Trans. 2*, 1993, 799–805.

[60] R. Bjornsson, F. Neese, R. R. Schrock, O. Einsle and S. DeBeer, *JBIC J. Biol. Inorg. Chem.*, 2015, **20**, 447–460.

[61] S. P. Cramer, K. O. Hodgson, W. O. Gillum and L. E. Mortenson, *J. Am. Chem. Soc.*, 1978, **100**, 3398–3407.

[62] S. P. Cramer, W. O. Gillum, K. O. Hodgson, L. E. Mortenson, E. I. Stiefel, J. R. Chisnel, W. J. Brill and V. K. Shah, *J. Am. Chem. Soc.*, 1978, **100**, 3814–3819.

[63] R. Bjornsson, F. A. Lima, T. Spatzal, T. Weyhermuller, P. Glatzel, E. Bill, O. Einsle, F. Neese and S. DeBeer, *Chem. Sci.*, 2014, **5**, 3096–3103.

[64] T. Spatzal, J. Schlesier, E.-M. Burger, D. Sippel, L. Zhang, S. L. Andrade, D. C. Rees and O. Einsle, *Nat. Commun.*, 2016, **7**, 10902.

[65] T. Lovel, J. Li, T. Liu, D. A. Case and L. Noodleman, *J. Am. Chem. Soc.*, 2001, **123**, 12392–12410.

[66] T. Lovell, R. A. Torres, W.-G. Han, T. Liu, D. A. Case and L. Noodleman, *Inorg. Chem.*, 2002, **41**, 5744–5753.

[67] D. Lukoyanov, V. Pelmenschikov, N. Maeser, M. Laryukhin, T. C. Yang, L. Noodleman, D. R. Dean, D. A. Case, L. C. Seefeldt and B. M. Hoffman, *Inorg. Chem.*, 2007, **46**, 11437–11449.

[68] T. V. Harris and R. K. Szilagyi, *Inorg. Chem.*, 2011, **50**, 4811–4824.

[69] L. Noodleman, C. Peng, D. Case and J.-M. Mouesca, *Coord. Chem. Rev.*, 1995, **144**, 199 – 244.

[70] C. C. Lee, Y. Hu and M. W. Ribbe, *Science*, 2010, **329**, 642–642.

[71] Z. Maskos, K. Fisher, M. Sørlie, W. E. Newton and B. J. Hales, *JBIC J. Biol. Inorg. Chem.*, 2005, **10**, 394–406.

[72] H.-I. Lee, L. M. Cameron, B. J. Hales and B. M. Hoffman, *J. Am. Chem. Soc.*, 1997, **119**, 10121–10126.

[73] R. Y. Igarashi, M. Laryukhin, P. C. D. Santos, H.-I. Lee, D. R. Dean, L. C. Seefeldt and B. M. Hoffman, *J. Am. Chem. Soc.*, 2005, **127**, 6231–6241.

[74] D. Lukoyanov, N. Khadka, Z.-Y. Yang, D. R. Dean, L. C. Seefeldt and B. M. Hoffman, *J. Am. Chem. Soc.*, 2016, **138**, 1320–1327.

[75] L. Yan, C. H. Dapper, S. J. George, H. Wang, D. Mitra, W. Dong, W. E. Newton and S. P. Cramer, *Eur. J. Inorg. Chem.*, 2011, **2011**, 2064–2074.

[76] L. Yan, V. Pelmenschikov, C. H. Dapper, A. D. Scott, W. E. Newton and S. P. Cramer, *Chem.-Eur. J.*, 2012, **18**, 16349–16357.

[77] D. Smith, K. Danya, S. Raugei and L. C. Seefeldt, *Biochem.*, 2014, **53**, 2278–2285.

[78] I. Dance, *Sci. Reports*, 2013, **3**, 3237.

[79] K. Vichitphan, *PhD thesis*, *Azotobacter vinelandii Nitrogenase: Effect of Amino-Acid Substitutions at the αGln-191 Residue of the MoFe Protein on Substrate Reduction and CO Inhibition*, Virginia Polytechnic Institute and State University, 2001.

[80] *Chemcraft - Graphical program for visualization of quantum chemistry computations*, 2015-10-22, `http://www.chemcraftprog.com/`.

[81] D. Lukoyanov, Z.-Y. Yang, S. Duval, K. Danyal, D. R. Dean, L. C. Seefeldt and B. M. Hoffman, *Inorg. Chem.*, 2014, **53**, 3688–3693.

[82] B. Huynh, M. Henzl, J. Christner, R. Zimmermann, W. Orme-Johnson and E. Münck, *Biochimica et Biophys. Acta (BBA) - Protein Struct.*, 1980, **623**, 124 – 138.

[83] J. Christiansen, R. C. Tittsworth, B. J. Hales and S. P. Cramer, *J. Am. Chem. Soc.*, 1995, **117**, 10017–10024.

[84] J. Chen, J. Christiansen, R. C. Tittsworth, B. J. Hales, S. J. George, D. Coucouvanis and S. P. Cramer, *J. Am. Chem. Soc.*, 1993, **115**, 5509–5515.

[85] S. J. Yoo, H. C. Angove, V. Papaefthymiou, B. K. Burgess, and E. Münck, *J. Am. Chem. Soc.*, 2000, **122**, 4926–4936.

[86] L. Rao, X. Xu and C. Adamo, *ACS Catal.*, 2016, **6**, 1567–1577.

[87] M. L. McKee, *J. Phys. Chem. A*, 2016, **120**, 754–764.

[88] M. A. Beno, J. M. Williams, M. Tachikawa and E. L. Muetterties, *J. Am. Chem. Soc.*, 1981, **103**, 1485–1492.

[89] M. Tachikawa and E. L. Muetterties, *J. Am. Chem. Soc.*, 1980, **102**, 4541–4542.

[90] W. N. Lanzilotta, J. Christiansen, D. R. Dean and L. C. Seefeldt, *Biochem.*, 1998, **37**, 11376–11384.

[91] K. Rupnik, Y. Hu, C. C. Lee, J. A. Wiig, M. W. Ribbe and B. J. Hales, *J. Am. Chem. Soc.*, 2012, **134**, 13749–13754.

[92] R. C. Tittsworth and B. J. Hales, *J. Am. Chem. Soc.*, 1993, **115**, 9763–9767.

[93] B. M. Barney, M. Laryukhin, R. Y. Igarashi, H.-I. Lee, P. C. D. Santos, T.-C. Yang, B. M. Hoffman, D. R. Dean and L. C. Seefeldt, *Biochem.*, 2005, **44**, 8030–8037.

[94] M. J. Dilworth and R. R. Eady, *Biochem. J*, 1991, **277**, 465–468.

[95] D. R. Lide, in *CRC handbook of chemistry and physics*, CRC press, 85th edn, 2005, ch. 8, p. 1275.

[96] L. C. Davis, *Arch. Biochem. Biophys.*, 1980, **204**, 270 – 276.

[97] S. Shaw, D. Lukoyanov, K. Danyal, D. R. Dean, B. M. Hoffman and L. C. Seefeldt, *J. Am. Chem. Soc.*, 2014, **136**, 12776–12783.

[98] K. D. Demadis, S. M. Malinak and D. Coucouvanis, *Inorg. Chem.*, 1996, **35**, 4038–4046.

[99] L. J. Laughlin and D. Coucouvanis, *J. Am. Chem. Soc.*, 1995, **117**, 3118–3125.

[100] D. Coucouvanis, K. D. Demadis, S. M. Malinak, P. E. Mosier, M. A. Tyson and L. J. Laughlin, *J. Mol. Catal. A: Chem.*, 1996, **107**, 123–135.

[101] D. Coucouvanis, *JBIC J. Biol. Inorg. Chem.*, 1996, **1**, 594–600.

[102] A. T. Thorhallsson and R. Bjornsson, *Unpublished work*.

[103] D. Sheppard, R. Terrell and G. Henkelman, *The J. Chem. Phys.*, 2008, **128**,.

[104] G. Henkelman, G. Jóhannesson and H. Jónsson, in *Methods for Finding Saddle Points and Minimum Energy Paths*, ed. S. D. Schwartz, Springer Netherlands, Dordrecht, 2002, pp. 269–302.

[105] G. Henkelman, B. P. Uberuaga and H. Jónsson, *The J. Chem. Phys.*, 2000, **113**, 9901–9904.

[106] G. Henkelman and H. Jónsson, *The J. Chem. Phys.*, 2000, **113**, 9978–9985.

[107] H. Jónsson, G. Mills and K. W. Jacobsen, *Nudged Elastic Band Method for Finding Minimum Energy Paths of Transitions*, 385.

[108] A. J. Bard, G. M. Whitesides, R. N. Zare and F. W. McLafferty, *Acc. Chem. Res.*, 1995, **28**, 91–91.

A. Appendix

ChemShell input file example

```
# Name of fragment file
set frag system.c

# Chemshell-script location
set scriptdir /home/ath146/Chemshell-scripts

# Putting fragment file in memory.
fragment $frag old persistent

set numatoms [get_number_of_atoms coords=$frag]
puts "Number of atoms is $numatoms"
# Read in file for lists. This sets up lists $charges, $groups, $types,
# $pdbresidues and $residuegroups
source save-new.chm
# PSF file from PSFgen
set psffile new.psf
# Loading connectivity from PSF file
load_connect_from_psf $frag $psffile
# Sourcing various TCl procs
source $scriptdir/procs.tcl

# Topology
# Here using CHARMM36 files. Modified file
set topmass $scriptdir/top_all36_prot.rtf
set charmmpar $scriptdir/par_all36_prot.prm

set orcapath /scratch/ragnarbj/orca303
source $scriptdir/orca3.0-chemsh-withimage-withbs.tcl
```

```
# sourcing active and frozen lists
source act
puts "Active region is [llength $act] atoms."

# New RB 2 feb 2016
# Now defining frozen list based on act list
set all [seq 1 to $numatoms]
set frozen [listcomp $all $act]

puts "Frozen region is [llength $frozen] atoms."

# QM REGION atoms, charge and multiplicity.
source qmatoms
puts "There are [llength $qmatoms] QM atoms and they are $qmatoms"

# Setting charge and multiplicity.
# If doing Broken-symmetry,
# then high-spin multiplictiy is defined later as well.
set charge -5
set mult 4
####################
# Special BS settings
####################
set brokensym yes
# Multiplicity of High-spin state and Broken-symmetry state.
# Will override $mult. Comment out if not using broken-symmetry.
set hsmult 36
set bsmult $mult

# Selecting which system atom numbers to flip
# (here Fe atoms: Fe2, Fe4 and Fe7 as defined in PDB/PSF)
# Will be converted to ORCA inputfile atom numbers by
# atomnumtoQMregionnum and then converted to comma-sep string.
set atomstoflip {17779 17780 17782}
set spinstofliplist [atomnumtoQMregionnum $qmatoms $atomstoflip]
set spinstoflip [join $spinstofliplist ","]

###################
# ORCA Theory level in simple input line
set orcasimpleinput "! TPSSh RIJCOSX D3BJ def2-SVP def2-SVP/J ZORA
```

```
Grid5 FinalGrid6 tightscf slowconv"
# ORCA block settings
set orcablocks "
%maxcore 2000

%basis
newgto Fe \"ZORA-def2-TZVP\" end
newgto Mo \"ZORA-def2-TZVP\" end
newgto S \"ZORA-def2-TZVP\" end
end

%scf
directresetfreq 1
diismaxeq 20
MaxIter 2500
end

%pal
nprocs 12
end
"
#########################################################################

# Setting up X-H and H-H constraints (TIP3) for optimization.
# Set jobtype to md for MD constraints
set jobtype opt
source $scriptdir/constraints-onlytip3.tcl

# Setting mxlist
set mxlist 38000
puts "mxlist is $mxlist"

# cutoff=1000 groups= $groups
# Optimisation
dl-find \
list_option=full coords=$frag active_atoms= $act constraints= $con maxcycle=1000 \
coordinates=cartesian residues= $pdbresidues result=result.c maxstep=0.1 \
theory=hybrid : [ list \
coupling=shift  debug=no atom_charges= $charges qm_region= $qmatoms conn=$frag \
qm_theory=orca: [ list \
```

```
executable=$orcapath/orca \
brokensym=$brokensym \
hsmult=$hsmult \
bsmult=$bsmult \
spinstoflip=$spinstoflip \
charge=$charge \
mult=$mult \
orcasimpleinput= $orcasimpleinput \
orcablocks= $orcablocks ] \
mm_theory=dl_poly : [ list \
frozen= $frozen \
conn= $frag \
debug=no \
use_pairlist=no \
exact_srf=yes \
mxlist= $mxlist \
cutoff=1000 \
scale14 = { 1.0 1.0 } \
use_charmm_psf=yes \
charmm_psf_file=$psffile \
atom_types= $types \
charmm_parameter_file=$charmmpar \
charmm_mass_file= $topmass ]]


times
```

Table A.1: Rate Constants of the Reactions in Schemes 1 and 2 adapted from Burgess et al.[1]

rate constant	value	comment
k_1	$5 \times 10^7\ M^{-1}s^{-1}$	responsible for lower activity at low protein concentrations
k_{-1}	$15\ s^{-1}$	
k_2	$140\ s^{-1}$	electron transfer from Fe-protein to MoFe protein[5]
k_3	$6\ s^{-1}$	rate-limiting step when substrates and Fe-protein are saturating[5]
k_{-3}	$4.40 \times 10^6\ M^{-1}s^{-1}$	responsible for lower activity at high protein concentrations
k_4	$3 \times 10^6\ M^{-1}s^{-1}$	rate of reduction of $Fe_{ox}(MgADP)_2$ complex
k_6	$1.20 \times 10^9\ M^{-1}s^{-1}$	rate of dissociation of ${S_2O_4}^{2-}$ into $2\,{SO_2}^{\bullet-}$
k_{-6}	$1.75\ s^{-1}$	rate of association of $2\,{SO_2}^{\bullet-}$ into ${S_2O_4}^{2-}$
k_7	$250\ s^{-1}$	gives increased H_2 evolution at low electron flux
k_8	$8\ s^{-1}$	slow to maximize E_3 concentration and hence N_2 binding
k_9	$400\ s^{-1}$	rapid H_2 evolution from most reduced hydridic species
k_{10}	$4 \times 10^5\ M^{-1}s^{-1}$	determine $K_M^{N_2}$ and $K_I^{H_2}$ at low electron flux
k_{-10}	$8 \times 10^4\ M^{-1}s^{-1}$	
k_{11}	$2.20 \times 10^6\ M^{-1}s^{-1}$	determine $K_M^{N_2}$ and $K_I^{H_2}$ at high electron flux
k_{-11}	$3 \times 10^6\ M^{-1}s^{-1}$	

Table A.2: Relative energies in kcal/mol of multiple BS solutions found by spin flips on Mo and Fe (numbered as in figure 3.2) on the crystal structure geometry with different functionals. BS column designates which irons were flipped and a special column show that solution with Mo-flip added. Each spin-flip was calculated with $M_S = \frac{1}{2}$ and $M_S = \frac{3}{2}$ with the different functionals. The lowest energy of a functional is colored red. The HS solutions were: BP86 282.447 kcal/mol, TPSSh 208.453 kcal/mol, B3LYP 182.699 kcal/mol. Method as described before in chapter 2

	BP86				TPSSh				B3LYP			
	$M_S = \frac{1}{2}$		$M_S = \frac{3}{2}$		$M_S = \frac{1}{2}$		$M_S = \frac{3}{2}$		$M_S = \frac{1}{2}$		$M_S = \frac{3}{2}$	
BS		Mo-flip		Mo-flip		Mo-flip		Mo-flip		Mo-flip		Mo-flip
123	20.422	20.423	24.734	24.729	26.472	26.460	33.523	33.522	24.095	24.112	30.666	30.666
124	20.125	20.125	26.036	26.034	26.125	28.227	33.721	33.721	23.833	23.833	40.610	31.056
125	7.643	7.643	14.273	14.291	11.714	11.714	10.189	10.189	15.677	24.886	8.657	8.657
126	15.725	15.725	18.326	18.326	21.631	21.631	17.804	17.804	26.881	26.659	18.247	18.247
127	7.366	7.366	15.273	15.273	12.031	12.031	8.849	8.849	15.936	23.455	7.395	7.395
134	20.077	20.077	25.842	25.843	25.231	25.231	33.507	32.750	23.929	23.513	30.543	30.543
135	8.080	8.080	16.076	16.076	13.197	13.079	10.884	10.885	16.902	24.792	9.552	9.552
136	8.122	8.122	15.392	15.392	13.833	13.833	10.702	10.702	18.169	24.194	9.479	9.479
137	15.715	15.715	18.143	18.144	21.813	21.813	17.971	17.971	26.946	26.866	18.262	18.262
145	15.151	15.151	19.607	19.606	19.307	19.307	17.034	17.034	24.452	36.270	17.285	17.285
146	7.579	7.579	15.517	15.172	12.327	12.327	9.989	9.989	16.534	23.047	8.901	8.901
147	7.073	7.073	15.145	14.953	11.398	11.398	8.906	8.906	15.896	24.062	7.655	7.655
156	5.212	5.212	8.600	8.600	19.598	19.598	13.452	13.452	28.648	49.584	18.268	28.887
157	5.431	5.431	9.183	9.183	17.153	17.153	13.977	13.977	24.671	51.547	18.417	29.606
167	5.707	5.707	8.536	8.537	18.399	18.399	13.528	13.528	25.576	50.743	18.415	30.458
234	9.024	9.023	9.163	9.163	16.369	16.366	15.468	15.465	18.415	18.706	16.709	16.816
235	8.422	8.422	0.000	0.000	16.247	19.551	0.000	0.000	15.328	16.538	8.670	0.000
236	14.783	14.783	9.803	9.803	21.792	21.792	9.585	9.585	25.883	29.033	8.861	8.861
237	14.628	14.629	9.696	9.696	21.730	21.730	9.309	9.309	25.580	29.640	8.587	8.587
245	13.573	13.573	9.736	9.736	21.187	21.187	11.083	11.083	31.726	31.726	10.460	10.460
246	13.797	13.797	10.193	10.193	22.296	22.296	11.442	11.442	34.582	30.950	10.920	10.920
247	6.750	6.750	0.399	0.399	18.940	21.529	1.057	1.057	16.767	18.585	0.870	0.870
256	5.213	5.213	17.324	17.324	34.383	18.684	22.479	22.479	43.706	49.623	27.158	37.591
257	15.991	15.991	10.908	10.909	26.203	27.139	19.277	19.277	37.844	50.944	23.847	26.588
267	5.262	5.262	17.839	17.839	36.033	36.033	24.774	24.774	41.208	57.423	30.433	39.081
345	12.875	12.875	9.473	9.473	19.439	19.439	9.638	9.638	29.840	28.912	8.871	8.871
346	5.186	8.613	0.524	0.524	19.768	19.768	0.992	0.992	17.878	16.865	10.520	0.813
347	13.657	13.657	9.416	9.416	21.144	21.145	9.692	9.692	28.772	28.772	9.119	9.119
356	16.453	16.453	10.102	10.102	27.232	26.837	17.777	19.643	33.740	48.756	26.198	25.459
357	5.361	5.361	17.675	17.675	34.419	34.419	23.696	23.696	36.684	58.561	28.439	38.455
367	5.186	5.186	17.587	17.587	35.792	35.792	24.321	24.321	46.483	48.454	29.027	38.246
456	6.261	6.261	18.020	18.020	29.329	19.954	26.119	24.264	35.177	56.847	28.757	37.682
457	6.750	6.750	18.258	18.262	35.254	35.254	28.832	24.416	44.306	58.717	28.973	38.902
467	5.262	5.262	10.544	10.544	28.219	28.219	18.687	18.687	35.283	51.024	22.348	28.207
567	31.729	31.744	31.630	31.629	48.651	48.666	39.902	39.902	54.242	88.267	52.693	75.011

Table A.3: Relative energies in kcal/mol of multiple BS solutions found by spin flips on Mo and Fe (numbered as in figure 4.1) on the CO adduct geometry with different functionals. BS column designates which irons were flipped and a special column show that solution with Mo-flip added. Each spin-flip was calculated with $M_S = \frac{1}{2}$ and $M_S = \frac{3}{2}$ with the different functionals. The lowest energy of a functional is colored red. The HS solution were: BP86 328.223 kcal/mol, TPSSh 237.669 kcal/mol, B3LYP 206.940 kcal/mol. Method as described before in chapter 2

	BP86				TPSSh				B3LYP			
	$M_S = \frac{1}{2}$		$M_S = \frac{3}{2}$		$M_S = \frac{1}{2}$		$M_S = \frac{3}{2}$		$M_S = \frac{1}{2}$		$M_S = \frac{3}{2}$	
BS		Mo-flip		Mo-flip		Mo-flip		Mo-flip		Mo-flip		Mo-flip
123	7.037	7.037	10.807	10.807	25.051	25.051	29.334	29.340	27.118	27.118	31.111	34.993
124	8.472	8.472	11.946	11.946	28.343	28.343	8.997	8.996	27.750	27.750	35.181	32.118
125	3.661	3.661	0.510	0.510	4.345	4.345	11.282	11.282	12.046	12.046	12.840	12.840
126	3.610	3.610	0.000	0.000	19.320	21.961	13.080	13.080	26.216	35.425	13.629	13.629
127	3.610	3.610	0.000	0.000	5.068	5.068	10.076	10.076	9.171	11.340	12.104	13.295
134	8.472	8.472	11.946	10.807	22.823	22.824	27.575	27.575	23.773	21.067	66.132	29.731
135	4.057	4.057	3.628	3.625	3.289	3.289	13.034	13.034	1.915	2.347	30.175	11.566
136	7.037	7.037	10.909	10.909	7.909	7.909	9.284	9.284	10.980	10.980	14.294	14.294
137	9.776	9.776	0.000	0.000	11.506	14.002	17.864	17.864	10.253	10.253	33.606	18.802
145	12.929	12.929	3.621	3.623	10.881	10.881	17.437	17.437	11.251	9.963	18.116	18.116
146	8.472	8.472	11.946	11.946	7.420	7.420	8.996	8.996	10.706	10.706	13.967	13.967
147	4.781	4.782	0.000	0.000	2.410	2.409	10.609	10.609	1.984	1.520	26.812	9.468
156	3.752	3.766	0.510	0.510	8.566	9.872	2.794	2.794	16.153	20.257	6.610	6.610
157	3.849	3.849	0.510	0.510	2.375	2.375	4.250	4.250	2.498	2.500	0.952	0.952
167	3.610	3.610	0.000	0.000	8.753	11.376	2.140	2.140	14.966	14.966	5.813	5.813
234	3.839	3.839	8.137	8.136	7.177	8.140	6.850	6.850	7.318	7.317	6.364	6.364
235	2.009	2.009	2.328	2.328	4.090	4.090	1.941	1.941	5.261	5.261	1.198	3.947
236	3.839	3.839	9.263	9.263	1.084	1.084	11.473	11.473	1.492	36.553	15.061	15.061
237	11.482	11.482	11.082	11.083	14.239	14.239	9.251	9.251	18.794	19.888	12.746	10.398
245	9.568	9.620	11.304	11.304	13.161	13.161	10.205	10.205	17.613	17.613	11.436	11.308
246	1.262	1.262	2.038	2.038	0.000	6.205	11.868	11.868	9.234	37.300	16.393	16.393
247	1.262	1.262	2.039	2.039	0.000	0.000	2.851	2.852	8.300	8.300	3.439	3.439
256	2.009	2.009	2.328	2.328	1.084	1.084	16.485	16.485	47.036	1.492	20.062	46.722
257	2.009	2.009	15.651	15.652	23.698	23.698	22.725	22.725	27.801	27.801	17.054	17.582
267	1.262	1.262	2.039	2.039	0.000	0.000	22.270	22.270	0.000	56.007	31.478	48.113
345	7.851	7.837	13.783	13.783	9.928	8.520	6.931	6.931	7.357	7.357	5.895	5.895
346	6.993	6.982	7.425	7.426	12.141	12.141	9.941	9.941	17.205	14.063	9.518	7.062
347	6.953	6.953	2.039	2.039	11.394	8.640	6.846	6.846	10.119	10.119	6.014	6.014
356	2.009	2.009	2.328	2.328	1.084	19.432	16.251	15.020	27.624	28.485	20.376	20.376
357	2.009	2.009	2.328	2.328	22.931	22.928	20.803	20.613	23.954	23.954	17.447	17.447
367	12.080	12.080	11.082	11.083	25.767	25.767	20.839	20.839	39.522	49.783	25.448	25.448
456	8.482	8.483	12.093	12.176	26.698	26.698	22.352	22.352	35.019	35.019	28.309	28.309
457	1.262	1.262	2.039	2.039	23.311	23.008	20.600	20.600	24.494	24.492	16.696	16.696
467	1.262	1.262	2.039	2.039	0.000	21.089	5.428	17.031	30.304	12.759	21.215	22.231
567	16.397	16.397	15.651	15.651	33.699	33.699	40.192	40.192	43.209	35.486	29.488	47.411