

Computational Chemistry studies on metallo-enzymes for
the splitting of hydrogen and the interconversion between
carbon monoxide and carbon dioxide

vorgelegt von
MSc. Chemistry
Daria Tombolelli
geb. in Civitavecchia, Rom, Italien

von der Fakultät II - Mathematik und Naturwissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades
Dr.rer.nat.

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr. Nediljko Budisa

Gutachterin: Prof. Dr. Maria Andrea Mroginski

Gutachter: Prof. Dr. Claudio Greco

Tag der wissenschaftlichen Aussprache: 13 December 2017

Berlin 2018

Parts of the following journal articles are included into the dissertation:

Ciaccafava, A. *, Tombolelli, D.*, Domnik, L., Fessler, J., Jeoung, J.-H., Dobbek, H., Mroginski, M.-A., Zebger, I., Hildebrandt, P. (2016). When the inhibitor tells more than the substrate: the cyanide-bound state of a carbon monoxide dehydrogenase. *Chemical Science*, **7** (5), 3162 - 3171.

<http://pubs.rsc.org/en/content/articlepdf/2016/sc/c5sc04554a>

Ciaccafava, A., Tombolelli, D., Domnik, L., Jeoung, J.-H., Dobbek, H., Mroginski, M.-A., Zebger, I., Hildebrandt, P. (2017). Carbon Monoxide Dehydrogenase Reduces Cyanate to Cyanide. *Angewandte Chemie (International Ed. in English)*, **56** (26), 7398 - 7401.

<https://doi.org/10.1002/anie.201703225>

Abstract

Carbon monoxide dehydrogenases (CODH) are metallo-enzymes which perform the interconversion between CO and CO₂. In particular, in the present work, nickel iron containing CODH from *Carboxydotherrmus hydrogenoformans*, is studied. Its catalytic mechanism is still unclear, because of the difficulties at the experimental level to trap substrate-bound intermediates and pure redox states. For this reason, we concentrated on the mechanism of inhibition as performed by two inhibitors, CN and NCO, which are structural and electronic mimickers of the substrates, CO and CO₂. To address this issue, we employed quantum mechanical/molecular mechanics (QM/MM) theoretical method. We were able to clarify both the structures of CN and NCO inhibited *Ch*CODH, and the mechanism of inhibition as performed by those two. The use of QM/MM methods, where complex metallic clusters are studied explicitly in the protein environment, turns out to be fundamental to clarify the specific characteristics of such systems. This work was possible thanks to a joint collaboration with crystallographers and spectroscopists. Moreover, we have been interested in another well known metallo-enzyme, also containing [NiFe] active site and iron-sulfur clusters. In particular, we studied the mechanism of protection against oxygen which takes place in the [NiFe] membrane bound hydrogenase (MBH) from *Ralstonia eutropha*. We first studied the structural changes of the unique iron-sulfur cluster, proximal to the active site, upon oxidation. To achieve this goal, we performed a series of QM/MM calculations on different models of the proximal cluster (PC). In a second step, we clarified the role and function of two proton transfer pathways, connecting the PC cluster and the active site. These pathways are part of a proton coupled electron transfer (PCET) mechanism, which consists in an exchange of proton and electrons between PC and active site clusters, and contributes to the protection of the catalytic activity of the active site against O₂. We studied this by means of molecular dynamic (MD) simulations. Furthermore, a preliminary study of the mutant D117S

of *ReMBH*, by means of QM/MM calculations, is presented in this thesis. Asp117 is part of a proton transfer pathway connecting the active site and the surface of the protein, and its mutation for a serine residue hinders the proton exchange, resulting in a unprecedented structure of the [NiFe] active site. The application of different theoretical methodologies in the study of *ReMBH*, gives us the opportunity to better understand the interconnection between electronic and structural changes taking place at the metallic centers, and the translocation of protons, within the protein.

Zusammenfassung

Kohlenmonoxid Dehydrogenasen (CODH) sind Metalloproteine, welche CO und CO₂ ineinander umwandeln. In der vorliegenden Arbeit wird die nickel- und eisenhaltige CODH von *Carboxydotherrmus hydrogenoformans* (Ch) im Speziellen untersucht. Der katalytische Mechanismus ist weiterhin unklar, da es experimentell schwierig ist, Substrat gebundene Intermediate und reine Redoxzustände zu erfassen. Aus diesem Grund haben wir uns auf den Mechanismus der Inaktivierung durch zwei Inhibitoren, CN und NCO, konzentriert, welche strukturell und elektronisch die Substrate, CO und CO₂, imitieren. Um diese Thema zu bearbeiten, haben wir die theoretische Methode der quantenmechanischen / molekularmechanischen (QM/MM) Berechnungen verwendet. Wir waren in der Lage, beide Strukturen der CN und CO inhibierten ChCODH Zustände und den Mechanismus der jeweiligen Inhibierung aufzuklären. Die Verwendung der QM/MM Methode, bei welcher komplexe Metallcluster explizit in ihrer Proteinumgebung untersucht wurden, erwies sich als wichtig, um die besonderen Charakteristiken eines solchen Systems zu erfassen. Diese Arbeit wurde durch die enge Zusammenarbeit mit Kristallografen und Spektroskopikern möglich. Darüber hinaus waren wir an einem weiteren bekannten Metalloenzym interessiert, welches ebenfalls ein Nickel-Eisen haltiges aktives Zentrum und Eisen-Schwefel Cluster beinhaltet. Dabei haben wir den Mechanismus der Sauerstofftoleranz genauer untersucht, welcher in der Membran gebundenen [NiFe] Hydrogenase (MBH) von *Ralstonia eutropha* (Re) statt findet. Wir haben zunächst die Änderungen in der Struktur des strukturell einzigartigen Eisen-Schwefel Clusters, der sich proximal zum aktiven Zentrum befindet, während der Oxidation untersucht. Dazu haben wir mehrere QM/MM Berechnungen an unterschiedlichen Modellen des proximalen Clusters (PC) vorgenommen. In einem zweiten Schritt, haben wir die Rolle und Funktion von zwei Protonentransferwegen aufgeklärt, welche den proximalen Cluster und das aktive Zentrum verbinden. Diese Transferwege sind Teil eine Protonen gekoppelten Elek-

tronentransfermechanismus (PCET), welcher aus dem Austausch von Protonen und Elektronen zwischen dem proximalen Cluster und dem aktiven Zentrum besteht und zur Sauerstofftoleranz des aktiven Zentrums beiträgt. Dies wurde mit Hilfe von molekular dynamischen (MD) Simulationen untersucht. Weiterhin wird in dieser Arbeit eine vorläufige Studie der D117S Mutante von ReMBH unter Anwendung von QM/MM Berechnungen vorgestellt. Asp117 ist Teil eines Protonentransferweges, welcher das aktive Zentrum mit der Proteinoberfläche verbindet. Seine Mutation durch eine Serinaminosäure behindert den Protonenaustausch, was zu einer ungewöhnlichen Struktur des aktiven Zentrums führt. Die Anwendung verschiedener theoretischer Methoden bei der Untersuchung der *Re*MBH gibt uns die Möglichkeit, den Zusammenhang zwischen elektronischen und strukturellen Änderungen an den Metallzentren und der Translokation von Protonen innerhalb des Proteins besser zu verstehen.

Contents

1	Introduction	1
1.1	Carbon Monoxide Dehydrogenase from <i>Carboxydotherrnus hydrogenoformans</i> . . .	2
1.2	[NiFe]-Hydrogenase from <i>Ralstonia eutropha</i>	5
2	Theoretical Background	9
2.1	Overview	9
2.2	Quantum Mechanics	9
2.2.1	Hartree Fock Method	10
2.2.2	Density Functional Method	12
2.2.3	Broken Symmetry method	15
2.3	Molecular Mechanics	19
2.3.1	Classical Force Fields	20
2.3.2	Molecular Dynamics Simulation	20
2.4	QM/MM method: a combined approach	22
2.4.1	Energy of the system	23
3	Carbon Monoxide Dehydrogenase from <i>Carboxydotherrnus hydrogenoformans</i>	26
3.1	The CN ⁻ inhibited state of <i>Ch</i> CODH	26
3.1.1	Introduction and Methods	26
3.1.2	Results	32
3.1.3	Discussion	37
3.2	The NCO inhibited state of <i>Ch</i> CODH	40

3.2.1	Introduction and Methods	40
3.2.2	Results	44
3.2.3	Discussion and Outlook	45
4	Membrane Bound Hydrogenase from <i>Ralstonia eutropha</i>	47
4.1	QM/MM method and the primary event of proton transfer	47
4.1.1	Results	47
4.1.2	Discussion	54
4.2	Proton Transfer Pathways study by means of MD simulations	57
4.2.1	Introduction and Methods	57
4.2.2	Results	62
4.2.3	Conclusion and Outlook	68
4.3	MBH mutant at the Active Site side	70
4.3.1	Introduction and Methods	70
4.3.2	Results	73
4.3.3	Discussion and Outlook	75
5	Conclusions	79

Chapter 1

Introduction

There are two major energy-related problems the world is facing in the next 50 years, 1) the increased competition for fossil fuels reserves because of their depletion and 2) the increasing level of atmospheric CO_2 which could produce large and uncontrollable impacts on the climate. A solution to these problems is to provide a future energy supply that is secure and CO_2 -neutral, switching to non fossil energy sources, which implies the development of new methods and technologies to transform the energy produced by these new sources in a form that can be stored, transported and used upon demand [1]. Carbon dioxide, through carbon fixation process of inorganic carbon to organic compounds by living organisms, is the ultimate source of the fossil fuels. The most prominent process that drives carbon fixation into these fuels is photosynthesis, the biological conversion of sunlight, water, and carbon dioxide into reduced organic materials. Chemosynthesis is another form of carbon fixation that can take place in the absence of sunlight. Organisms that grow by fixing carbon are called autotrophs. Autotrophs include photoautotrophs, which synthesise organic compounds using the energy of sunlight, and lithoautotrophs, which synthesise organic compounds using the energy of inorganic oxidation. Different pathways for CO_2 fixation exist, and they use different mechanisms and enzymes to process CO_2 making C–H and C–C bonds. Research on homogeneous and heterogeneous catalysts for CO_2 and CO reduction has also contributed to our understanding of C–C and C–H bond formation reactions involved in the production of synthetic fuels. Moreover, one of the major obstacles to an efficient conversion of CO_2 into fuels is the lack of catalysts. In the light of this becomes clear

the importance of a synergistic contribution of the catalysis and biological communities to the problem of converting carbon dioxide directly into fuels. Biological systems must extract energy from their environments in order to complete their metabolic processes. Evolution of living organisms made it possible for them to exist in a wide variety of environments so that they can use different sources to interconvert energy. In fact, in addition to the metabolic pathways used for the CO₂ fixation there exist also important pathways that produce and use H₂, reduce N₂ to ammonia, oxidise water and reduce O₂. All these metabolic pathways involve the storage and use of energy in the form of chemical bonds. Hence, the ability of the scientists to carry out these same transformations in a controlled manner becomes essential. This fact means that catalysis science will benefit from an in-depth look at these important metabolic processes while biochemistry science can get hold of knowledge about how synthetic catalysts promote these same chemical transformations. A notable example of this interaction between different scientific areas can be found in the study of synthetic and natural catalysis of H₂ production and oxidation. The research on this field has involved a large number of scientists from different fields like enzymologists and crystallographers, spectroscopists, computational and organometallic chemists, from biological, physical and chemical communities. The same productive interaction is evolving in the challenging research about CO₂ reduction to fuels. In the present work, different aspects of two of the above mentioned research subjects, CO₂ fixation and H₂ production catalytic processes, carry out by two enzymes, Carbon monoxide dehydrogenase (CODH) and Hydrogenase (H₂-ASE), have been investigated from the computational chemistry point of view. The work has been done in close collaboration with scientists of other fields, within UniCat project.

1.1 Carbon Monoxide Dehydrogenase from *Carboxydotherrmus hydrogenoformans*

Carbon monoxide dehydrogenases (CODH) are enzymes responsible for the interconversion between CO and CO₂ following the water-gas shift reaction: $CO + H_2O \leftrightarrow CO_2 + 2H^+ + 2e^-$.

CODHs play crucial roles in various metabolic pathways, including the Wood-Ljungdahl pathway that allows anaerobic microorganisms, such as *Moorella thermoacetica* (*Mt*) and *Carboxydotherrmus hydrogenoformans* (*Ch*) to grow on H₂ and CO₂ as electron and carbon source, or to oxidise CO using it as a source of electrons [1]. Furthermore, the reverse process, that

is, the enzymatic reduction of CO_2 , is very interesting since it may serve as a template for the bioinspired development of catalysts for CO_2 degradation. These enzymes exhibit turnover frequencies for CO oxidation as high as 40000 s^{-1} (at 70° C using methyl viologen cation as oxidant) and 45 s^{-1} for CO_2 reduction. These enzymes also operate at potentials near the thermodynamic potential of the CO_2/CO couple, $\sim 0.52\text{ V}$ vs the SHE at pH 7 [2]. Here the O_2 -sensitive [NiFe] CODH from *Carboxydotherrmus hydrogenoformans* (*Ch*) is studied. *Ch*CODH is an homodimeric enzyme which harbors a total number of five metal clusters, two [4Fe4S] B-clusters, one single [4Fe4S] D-cluster that bridges the two subunits and two unique [Ni(4Fe4S)O(H_x)] C-clusters, which is the active site where the enzymatic reaction takes place. Moreover, the C-cluster of one subunit and the B-cluster of the other are closer than those from the same subunit. Thus, a functional dimer is required for rapid electron transfer. The C-cluster resembles a classical cubane shape [4Fe4S] from which one Fe, the so called Fe_1 , is pulled out and let to dangle. The vacant Fe position has been filled by a Ni atom, which is a coordinatively unsaturated NiII species and binds three S ligands in an apparent planar T-shaped environment. The first coordination sphere of the Fe_1 center, instead, consists of a histidine ligand (His261), a cysteine (Cys295), a μ -3-sulfido ligand, and a fourth light atom, possibly water or hydroxide, depending on the oxidation state of the cluster. This fourth ligand is also close to Ni.

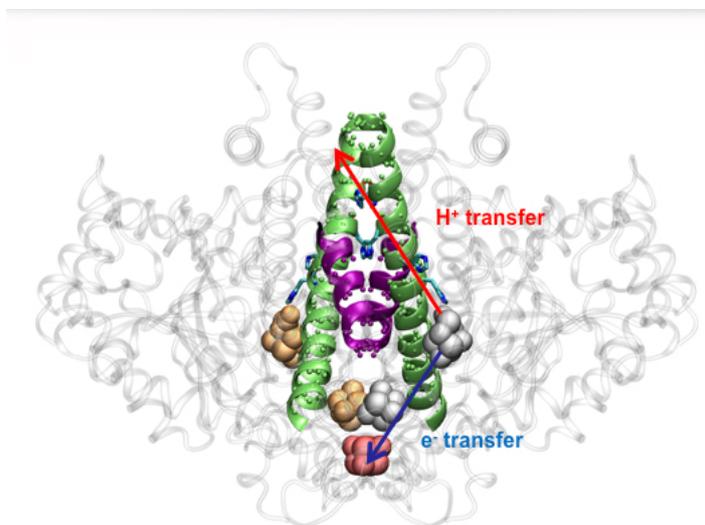


Figure 1.1: *Ch*CODH protein in the biological assembly dimerical structure.

C-cluster adopts at least two different oxidation states during catalysis. The Cred1 redox

state is able to bind CO while Cred2, obtained by a further two-electron reduction, targets CO₂ [3],[4]. Despite a vast number of biochemical, spectroscopic, and crystallographic investigations [5], the mechanism of the catalytic process of CODH is still unclear, mostly because of the difficulty to trap substrate-bound intermediate or pure redox states. A possible approach to overcome this limitation is to use inhibitors mimicking the substrates, which has been the choice in the present study. On the basis of the crystal structure (PDB code: 4UDX), which has been recently published at true-atomic resolution [6], CO₂ binds to Ni via the C atom to form a Ni–C bond and with one of the carboxylate oxygen atoms (O₁) bound to Fe₁, as well as forming a hydrogen bond to a lysine residue (K563). The second oxygen atom, O₂, appears to be hydrogen bonded with a protonated histidine residue (H93). A second structure (PDB code: 4UDY) has been resolved [6], also at very high resolution of 1.09 Å where NCO, an inhibitor of the activity of the protein, is bound to the Ni site. In both these structures, CO₂ and NCO had undergone two-electron reduction step, leading to the formation of Ni-bound carboxylate and a carbamoyl H₂NCO⁺ respectively, products that were stabilised via π -backbonding. This provided insight into the binding and activation of CO₂ and its inhibitory counterpart NCO. Moreover, in analogy to the CO₂ reduction to CO, NCO reduction to CN by CODH was recently postulated, supported by the analogous NCO reduction to CN observed in nitrogenases. To further elucidate the degradation of NCO by *Ch*CODH, IR spectroscopy has been employed to probe the substrate bound to the C-cluster of the enzyme in solution and in the crystalline state, while quantum-mechanical/molecular mechanics (QM/MM) calculations have been used to predict the characteristic vibrational modes of the carbamoyl moiety [4]. The picture for CO binding, instead, when we started this study, was still under debate. Structural analysis pointed to a bent geometry of CO bound to the terminal Ni in the CODH-CO complex of *Methanosarcina barkeri* [7]. This finding could be explained stating that a bent Ni–CO geometry would lower the energetic activation barrier for CO oxidation. Moreover, in a complementary study, cyanide was found in the same bent conformation [8] (PDB: 3I09). However, indications were pointing in different direction. First, an independent re-evaluation of the electron density maps of these two structures did not provide clear evidence in support of the orientation of CO or CN ligands [9]. In another crystal structure from *Ch*CODH (PDB: 3I39), a different binding mode of CN was present [10] (Fig. 1). A powerful tool to rationalise this conflicting results might be IR spectroscopy technique, since it probes the C \equiv N stretching mode in a spectral region free of

any interfering bands of the protein. However, the results previously reported in IR spectroscopy were not able to uniquely assign the multiple signals detected. Joint work with crystallographers (AG Dobbek, HU Berlin) and spectroscopists (AG Hildebrandt, TU Berlin) allowed us to shed light on this long debate. IR spectroscopic measurements were carried out with the wild-type, and mutant *Ch*CODH variant and this analysis served as a reference for the QM/MM calculations which were performed on the basis of new high-resolution crystal structures to elucidate details of the active site structure of the *Ch*CODH-CN complex [3].

In the next sections the computational chemistry work on both *Ch*CODH-CN and $-NCO$ inhibited states will be presented.

1.2 [NiFe]-Hydrogenase from *Ralstonia eutropha*

The cleavage of H_2 in biological systems is performed by a diverse class of metalloenzymes named hydrogenases. The dihydrogen molecule is converted into protons and electrons following the reaction:



This reaction takes place at a specialised metal center that increases the acidity of H_2 and leads to a heterolytic splitting of the molecule, strongly accelerated by the presence of a nearby base [11]. The reverse-direction reaction, the heterogenesis of H_2 , consists in the coupling of H^+ and H^- , thus consuming protons and protecting the cell from an excess of reductive power. This bidirectionality makes this enzyme an excellent candidate for energy converting devices. One of the biggest challenges of chemists nowadays is the development of technologies for capturing and storing renewable energy as a fuel that can be used on demand. Hydrogen is a direct answer for such energy economy. Electricity can be easily obtained from any renewable and energy stored as H_2 produced from water by electrolysis. When this energy is required back, oxidation of H_2 in fuel cells serves as source of electricity. The mostly used electrocatalyst in this case is platinum. However, this precious metal has an insufficient availability and is easily poisoned by H_2S and CO , typical contaminants of H_2 from steam reforming. In this sense, hydrogenases solve these limitations, since they use earth abundant metals in their active site (Fe and Ni), with activity comparable to Pt and their inhibition by H_2S and CO is mostly reversible. Hydrogenases are divided into three classes according to the metal composition of their active sites, which can be

NiFe, [FeFe] or [Fe] hydrogenases. The second class is the most efficient in H_2 production with $k_{cat} \approx 104s^{-1}$. However hydrogenases of this group are irreversibly inactivated by even trace amounts of O_2 . Anyhow, within the NiFe class there exist a subgroup which is O_2 -tolerant. In particular, the present work focuses on the Membrane-Bound NiFe hydrogenase (MBH) from *Ralstonia eutropha* (*Re*). *Re*MBH, represented in figure 1.2, is made up of two subunits, large

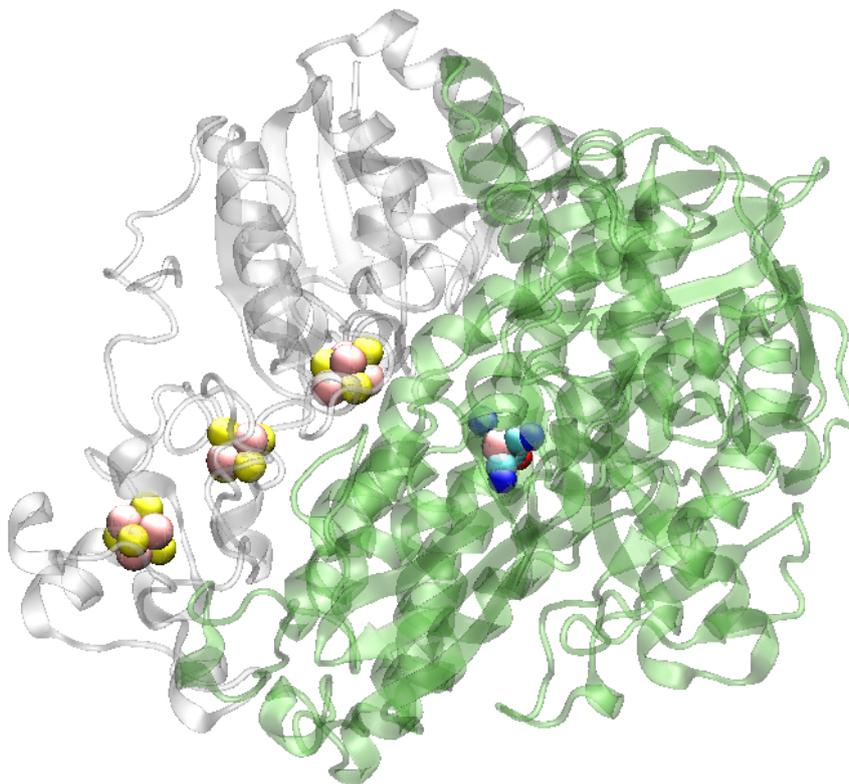


Figure 1.2: NiFe-hydrogenase protein structure in the H_2 -reduced state (pdb entry: 3RGW).

and small. The active site is located in the large one and consists of a bimetallic NiFe center coordinated by four cysteine residues (Cys 75, Cys 78, Cys 597, Cys 600) and CO/CN^- ligands to Fe. A bridging position is alternatively occupied or unoccupied by different ligands (I.e. H^- , OH^-) during the catalytic cycle. Protons are released during the catalytic activity and their transport to the surface is initiated by an arginine residue which is forming hydrogen bond to one of the CN^- ligands to Fe center of the Active Site [12]. In standard hydrogenases, upon oxidation, two EPR-active redox states called Ni-A and Ni-B are detected. Ni-B is characterised by a NiIII–FeII center with a bridging hydroxide ion, and is a "ready" state, because it reactivates

rapidly under reducing conditions. Ni-A state, instead, a NiIII–FeII center is an "unready" state and requires prolonged reactivation. This latter state has not been observed with EPR and FTIR studies on *Re*MBH and AaHase-1 [13]. In the small subunit instead three iron-sulfur clusters, which serve as electron relay, are present. They are named after their distance to the active site as distal, medial and proximal cluster. The distal is a standard cubane FeS-cluster with a three-cysteine and one-histidine coordination $[4\text{Fe}4\text{S}](\text{His})(\text{Cys})_3$ first ligation sphere, while the medial is a $[3\text{Fe}4\text{S}](\text{Cys})_3$ cluster. Finally, the proximal cluster is a cubane cluster where one sulfur is substituted by one cystein (Cys19) thiolate, bridging μ_2 -bridges Fe_1 and Fe_4 site, and a second supernumerary cysteine (Cys120) completes the tetrahedral coordination of the Fe_3 site. It is known that in MBH the tolerance toward O_2 is not due to a modification of the active site, where the coordination of the metal pair remains the same to the standard O_2 -inactivated NiFe enzymes. Also the access of O_2 through the hydrophobic gas channel is not limited. Recently [14], the publication of the X-ray structure of superoxidized (PDB: 4IUC, 4IUB) and partially reduced (PDB: 4IUD) forms of *Re*MBH shed new light into the dispute about its tolerance upon oxygen. The surprising finding of an hydroxyl group attached to the Fe_1 of the iron-sulfur proximal cluster which, together with the bond formed between the Fe_4 and nitrogen of the backbone of Cys20, contributes to the stabilization of two electron superoxidized state, raised new puzzling questions about the overall mechanism taking place into the protein, not specifically at the active site itself. Moreover, while in the superoxidized state structures the oxygen species has between 70% and 90% occupancy, the oxo-ligand at Fe_1 in partially reduced structure was present in a fraction of $\sim 30\%$ of the protein and due to potential X-ray induced damage, the fraction of clusters containing the O ligand can only be taken as an estimation. The formation of the mentioned bond and the attachment of the OH group are still under debate. Ni-A state has been described recently [15] with a structure where a hydroxide occupies the bridging position between Ni and Fe and a coordinating cysteine is oxidated to bridging sulfenate. Previously, crystallographic studies had assigned to it a hydroperoxide ligand. A recent theoretical work from Bregli [16] shows that the only ligand which can give a comprehensive explanation to both structures [17],[18] is a peroxide O_2^{2-} . However the sulfenate species described in ref. [15] is the most stable among all, so that the species coordinating peroxide and hydroperoxide ligand are probably intermediates along the mechanism of the oxidation in the presence of oxygen, leading to the Ni-A sulfenate state. Ni-A state is an inactive state that can be only very slowly reactivated. Electrons and protons

necessary for the reactivation, are thus provided by the proximal cluster. This is an example of proton-coupled electron transfer (PCET), which is a very important mechanism, often found in proteins. Different computational studies have been published where the structural changes this unique cluster in different O_2 -tolerant hydrogenases encounters upon one or two electron oxidation are analysed. However, the new finding of an hydroxyl group attached to the Fe_1 of *ReMBH* was not taken into account, and the structure of the oxidized state of the proximal cluster is still under debate. Moreover, proton transfer pathways, translocating protons between the iron-sulfur proximal cluster and the nickel-iron active site, have been proposed. In particular, assuming that the Fe_1 -bound oxo-ligand, which has been demonstrated being an hydroxyl group [14] in the superoxidized crystal structure, is derived from the dissociation of a water molecule, a proton would be liberated and transferred via a His residue, which resides close to Fe_1 , to the NiFe active site. This proton transfer pathway is called His-path. A second proton transfer pathway, described before for *HmMBH* by Shomura et al. [19] has also been proposed for *ReMBH* in the recent Nature work by Frielingsdors et al. [14]. According to it, a proton is delivered from N(Cys20) to the NiFe site. This second proton transfer pathway is called Glu-path since the first proton transfer event is involving residue Glu76. Besides, in case the Fe_1 -bound oxygen ligand is derived from O_2 , these postulated proton pathways might also be involved in the delivery of protons from the active site to the [4Fe3S] cluster. One of the purposes of this PhD work has been to present an insight into the mechanism of the transformation encountered by the iron-sulfur proximal cluster upon one electron oxidation and combine it with the study of the proposed proton transfer pathways and their role in the protection of the catalytic function of the active site in aerobic conditions. Therefore, a structural study of the iron sulfur proximal cluster during the first proton transfer event, with Quantum Mechanical/Molecular Mechanic combined approach, and a dynamic analysis of the two aforementioned pathways, employing molecular dynamic simulation methods, have been carried out. A fundamental assumption of the approach here presented is that a proton can be translocated between two points in space, namely a donor and an acceptor, when an hydrogen bond is formed between them. The proton can thus be translocated between protonable side chains and with the help of the formation of transient adducts of water molecules and polarizable atoms.

Chapter 2

Theoretical Background

2.1 Overview

The application of computational methods in chemistry is having a large boost in the recent years, thanks to the large improvement of computers ability in performing faster and larger scale calculations. A quite large pool of methods are available nowadays, which can be applied with different purposes from the prediction of electronic structure of small molecules to dynamics of big systems like proteins, or even viruses and cells. In this work three different methods have been used and will be described in their theoretical background in Chapter 2, while their application in the calculation of different properties of the two enzymes presented in the introduction will be presented in Chapters 3 and 4.

2.2 Quantum Mechanics

Quantum mechanical methods are based on the solution of the Schrödinger Equation:

$$\hat{H}\Psi_i = E_i\Psi_i \quad (2.1)$$

where Ψ_i is the wave function of the system in state i of energy E_i . \hat{H} is the Hamiltonian operator which contains five contributions:

$$\hat{H} = -\sum_i \frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_k \frac{\hbar^2}{2m_k} \nabla_k^2 + \sum_{i<j} \frac{e^2}{r_{ij}} + \sum_{k<l} \frac{e^2 Z_k Z_l}{r_{kl}} - \sum_i \sum_k \frac{e^2 Z_k}{r_{ik}}. \quad (2.2)$$

First term, \hat{T}_e , is the kinetic energy of the electrons and second term, \hat{T}_N , is the kinetic energy of the nuclei. The other three terms, \hat{V}_{ee} , \hat{V}_{NN} and \hat{V}_{Ne} are the potential energy operators of attractive and repulsive electrostatic interactions between the electrons, the nuclei and between each other. Within the Born Oppenheimer approximation, which presume that is possible to separate the nuclear motion from the electronic motion since nuclei of atoms move much slower compared to electrons, is possible to consider the positions of the nuclei fixed. In this way kinetic energy of the nuclei, \hat{T}_N , is zero and the potential energy of the nucleus-nucleus repulsion, \hat{V}_{NN} is constant. It is then possible to simplify the expression of the Hamiltonian to:

$$\hat{H} = \hat{T}_e + \hat{V}_{Ne} + \hat{V}_{ee} = - \sum_i \frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_i \frac{e^2 Z_i}{r_i} + \frac{1}{2} \sum_{i \neq j} \frac{e^2}{4\pi\epsilon_0 r_{ij}}. \quad (2.3)$$

In eq. 2.3, ∇^2 is the Nabla Operator, Z is the effective nuclear charge, e is the elementary charge, ϵ_0 is the vacuum permittivity, r_i is the nucleus-electron distance and r_{ij} the distance between two electrons.

2.2.1 Hartree Fock Method

Hartree Fock method is based on the separation of many-electron wave function into a product of monoelectronic wavefunctions:

$$\Psi = \psi_{a,\alpha}(1)\psi_{a,\beta}(2)\dots\psi_{z,\beta}(N) \quad (2.4)$$

Equation 2.4 is the wavefunction for an N-electron closed-shell molecule, where electron 1 occupies molecular orbital ψ_a with spin α , electron 2 occupies molecular orbital ψ_a with spin β , and so on. Since the wavefunction must satisfy the Pauli principle changing sign under the permutation of any pair of electrons, the wavefunction has to be written as a sum of all possible permutations, with appropriate sign:

$$\Psi = \psi_{a,\alpha}(1)\psi_{a,\beta}(2)\dots\psi_{z,\beta}(N) - \psi_{a,\alpha}(2)\psi_{a,\beta}(1)\dots\psi_{z,\beta}(N) + \dots \quad (2.5)$$

The entire sum, considering $N!$ terms can be written as a determinant, the so called Slater Determinant:

$$\Psi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_{a,\alpha}(1) & \psi_{a,\beta}(1) & \cdots & \cdots & \psi_{z,\beta}(1) \\ \psi_{a,\alpha}(2) & \psi_{a,\beta}(2) & \cdots & \cdots & \psi_{z,\beta}(2) \\ \vdots & \vdots & & & \vdots \\ \vdots & \vdots & & & \vdots \\ \psi_{a,\alpha}(N) & \psi_{a,\beta}(N) & \cdots & \cdots & \psi_{z,\beta}(N) \end{vmatrix} \quad (2.6)$$

This one-electron wave functions are composed of a spatial orbital and a spin function and they must satisfy the Hartree-Fock equations:

$$f_1 \psi_{a,\sigma}(1) = \epsilon \psi_{a,\sigma}(1) \quad (2.7)$$

where σ is either α or β and f_1 is the Fock operator, which can be written as:

$$f_1 = h_1 + \sum_j \{2J_j(1) - K_j(1)\} \quad (2.8)$$

The first term is the core hamiltonian:

$$h_1 = -\frac{\hbar^2}{2m_e} \nabla_1^2 - \sum_n \frac{Z_n e^2}{4\pi\epsilon_0 r_{ni}} \quad (2.9)$$

the second term is the Coulomb operator J :

$$J_j(1) \psi_a(1) = \int \psi_j^*(2) \psi_j(2) \left(\frac{e^2}{4\pi\epsilon_0 r_{12}} \right) \psi_a(1) d\tau_2 \quad (2.10)$$

an the third is the exchange operator K , where:

$$K_j(1) \psi_a(1) = \int \psi_j^*(2) \psi_a(2) \left(\frac{e^2}{4\pi\epsilon_0 r_{12}} \right) \psi_j(1) d\tau_2 \quad (2.11)$$

Although the Hartree-Fock equations (see 2.7) looks simple, the Fock operator f depends on the wavefunctions of all the electrons. To proceed is then necessary to guess the initial form of the wavefunctions, and use them in the definition of Coulomb and exchange operators. This process is then continued using the latest wavefunctions and the entire procedure is called self-consistent field (SCF). In order to solve the Hartree-Fock equations the wavefunctions have to be expressed as linear combination of M atomic orbitals χ_i :

$$\psi_a = \sum_{i=1}^M c_{i\alpha} \chi_i \quad (2.12)$$

Replacement of this latter one into eq. 2.12 gives:

$$f_i \sum_{i=1}^M c_{i\alpha} \chi_i(1) = \epsilon_\alpha \sum_{i=1}^M c_{i\alpha} \chi_i(1) \quad (2.13)$$

then multiplication from the left by $\chi_j^*(1)$ and integration over the coordinates of electron 1 gives:

$$\sum_{i=1}^M F_{ji} c_{i\alpha} = \epsilon_\alpha \sum_{i=1}^M S_{ji} c_{i\alpha} \quad (2.14)$$

which in a matrix form is:

$$\mathbf{FC} = \mathbf{SC}\epsilon \quad (2.15)$$

where \mathbf{F} is the matrix formed from the Fock operator:

$$F_{ij} = \int \chi_i^*(1) f_1 \chi_j(1) d\tau \quad (2.16)$$

and \mathbf{S} is the matrix of overlap integrals:

$$S_{ij} = \int \chi_i^*(1) \chi_j(1) d\tau \quad (2.17)$$

2.2.2 Density Functional Method

Density functional theory (DFT) [20] is a QM modeling method used in order to investigate the electronic structure of many-body systems. With this theory, the properties of a many-electron system can be determined by using functionals, i.e. functions of another function, which in this case is the spatially dependent electron density. Hence the name density functional theory comes from the use of functionals of the electron density. DFT has been very popular for calculations in solid-state physics since the 1970s. However, DFT was not considered accurate enough for calculations in quantum chemistry until the 1990s, when the approximations used in the theory were greatly refined to better model the exchange and correlation interactions. In many cases the results of DFT calculations for solid-state systems agree quite satisfactorily with experimental data. This technique is less demanding in terms of computational effort, computer time and in some cases, i.e. with d-metal complexes, which are the most represented in this work, DFT has better agreement with experimental values than, for example, is obtained from Hartree-Fock procedures. This theory has a completely different approach with respect to the classical methods used as the HF method. In fact these methods are based on the total wavefunction of the system, which depends on the number of the electrons and nuclei forming the system. Thus,

calculations of the total wavefunction of bigger systems becomes a challenge. DFT methods, instead, developed in the first years of the 1960s by Pierre Hohenberg and Walter Kohn [21], are based on the electron density, ρ , which substitutes the wavefunction ψ , and depends only on three variables despite of the number of the electrons. Hence, the simplification of the problem results evident. Hohenberg e Kohn (HK) demonstrated that the density of the non degenerate ground-state (GS) of a system of electrons gives the possibility to determine univocally all the properties of the fundamental state itself. Thus any physic observable of the GS, as its total energy, can be expressed as a function of the electron density, $E[\rho]$. The electron density is itself a function of position, $\rho(\mathbf{r})$. Hence, knowing $E[\rho]$ in its functional form and applying the minimisation principle to the functional energy, the problem can be easily resolved obtaining the density of the GS. Then, from the GS density it could be possible in principle calculate all the other observable of the GS knowing the relative form of the functional. The exact ground-state energy of a molecule with n -electron molecules is:

$$E[\rho] = E_K + E_{P;e,N} + E_{P;e,e} + E_{XC}[\rho] \quad (2.18)$$

where E_K is the total electron kinetic energy, $E_{P;e,N}$ the electron-nucleus potential energy, $E_{P;e,e}$ the electron-electron potential energy, and $E_{XC}[\rho]$ the exchange-correlation energy, which counts for the effects due to spin. The theory is based on the Thomas-Fermi model developed in 1927. It permits to calculate the energy of a multielectronic system representing its kinetic energy as a functional of the electronic density, combining it with the classical expression that takes into account the nuclei-electron and electron-electron interactions. The accuracy of the Thomas-Fermi equation was limited by the fact that it does not consider the exchange energy, i.e. the energy caused by the interaction of the superimposed wavefunctions, predicted by HF theory. In 1928 a significative improvement to the theory was made by Paul Dirac adding the functional for this latter. The Thomas-Fermi-Dirac, however, still lack in accuracy because of its representation of the kinetic energy, followed by the error in the exchange energy, which still was not taking into account the interelectronic repulsion. Although density functional theory has its conceptual roots in the Thomas-Fermi model, DFT was put on a firm theoretical footing by the two Hohenberg-Kohn theorems. The two fundamental theorems of Hohenberg and Kohn refers to the GS without considering the magnetic field. 1. All the properties of the non-degenerate GS of a multi-electron system are described by functionals of the monoelectronic density; 2. The GS

energy is the minimum value of $E[\rho]$ with respect to all the electronic densities for which it is defined. Within the framework of Kohn-Sham DFT (KS-DFT) [22], the intractable many-body problem of interacting electrons in a static external potential is reduced to a tractable problem of non-interacting electrons moving in an effective potential. The effective potential includes the external potential and the effects of the Coulomb interactions between the electrons, e.g., the exchange and correlation interactions. Modeling the latter two interactions becomes the difficulty within KS-DFT. The simplest approximation is the local-density approximation (LDA), which is based upon exact exchange energy for a uniform electron gas, which can be obtained from the Thomas-Fermi model, and from fits to the correlation energy for a uniform electron gas. Non-interacting systems are relatively easy to solve as the wavefunction can be represented as a Slater determinant of orbitals. Further, the kinetic energy functional of such a system is known exactly. The exchange-correlation part of the total-energy functional remains unknown and must be approximated. As usual in many-body electronic structure calculations, the nuclei of the treated molecules or clusters are seen as fixed (the Born-Oppenheimer approximation), generating a static external potential V in which the electrons are moving. There are many sophisticated methods for solving the many-body Schrödinger equation based on the expansion of the wavefunction in Slater determinants. While the simplest one is the Hartree-Fock method, more sophisticated approaches are usually categorised as post-Hartree-Fock methods. However, the problem with these methods is the huge computational effort, which makes it virtually impossible to apply them efficiently to larger, more complex systems. These orbitals:

$$\rho(\mathbf{r}) = \sum_{i=1}^N |\psi_i(\mathbf{r})|^2 \quad (2.19)$$

are used to construct the electron density from. They are calculated from the Kohn-Sham equations, which can be found applying the variation principle to the electron energy, and are like the Hartree-Fock equations except for a term V_{XC} which is the exchange-correlation potential:

$$\left\{ -\frac{\hbar^2}{2m_e} \nabla_1^2 - \sum_{j=1}^N \frac{Z_j e^2}{4\pi\epsilon_0 r_{j1}} + \int \frac{\rho(\mathbf{r}_2) e^2}{4\pi\epsilon_0 r_{12}} d\mathbf{r}_2 + V_{XC}(\mathbf{r}_1) \right\} \psi_i(\mathbf{r}_1) = \epsilon_i \psi_i(\mathbf{r}_1) \quad (2.20)$$

The exchange-correlation potential is the 'functional derivative' of the exchange-correlation energy:

$$V_{XC}[\rho] = \frac{\partial E_{XC}[\rho]}{\partial \rho} \quad (2.21)$$

The Kohn-Sham equations are then solved iteratively and self-consistently. First, a guess for the electron density is used. It is commonly used a superposition of atomic electronic densities. Then the exchange-correlation potential is calculated by assuming an approximate form of the dependence of the exchange-correlation energy on the electron density and evaluating the functional derivative in equation 2.21. For this step, the simplest approximation is the local-density approximation (LDA) and is possible to write

$$E_{XC}[\rho] = \int \rho(\vec{r})\varepsilon_{XC}[\rho(\mathbf{r})]d\mathbf{r} \quad (2.22)$$

where ε_{XC} is the exchange-correlation energy per electron in a homogeneous gas of constant density. Next, the Kohn-Sham equations are solved to obtain an initial set of orbitals, which will be then used to obtain a better approximation to the electron density (from equation 2.19) and the process is repeated until the density and the exchange-correlation energy are constant to within some tolerance. The main problem with DFT is that the exact functionals for exchange and correlation are not known except for the free electron gas. However, approximations exist which permit the calculation of certain physical quantities quite accurately. In physics the most widely used approximation is the local-density approximation (LDA), where the functional depends only on the density at the coordinate where the functional is evaluated. This can be generalised to include electron spin (LSDA - local spin-density approximation). Another way is to use the generalised gradient approximations (GGA) taking into account the gradient of the density at the same coordinate. Difficulties in expressing the exchange part of the energy can be relieved by including a component of the exact exchange energy calculated from Hartree-Fock theory. Functionals of this type are known as hybrid functionals. The BP86 functional, applied to the biological systems in this work, is a GGA functional and it is a combination of the Becke88 [23] correlation functional and the Perdew86 [24] functional for the exchange energy. It is a pure functional well suited for metals, as they are in the systems studied in this work.

2.2.3 Broken Symmetry method

The two proteins which are subject of this PhD thesis, Carbon monoxide dehydrogenase and Hydrogenase, are metalloenzymes. Their catalytic [NiFe] centers and iron-sulfur clusters, have particular electronic structure, which needs to be described with a specific method, at a quantum mechanical level. In fact, these transition metal clusters are open-shell systems which can

occupy different BS states. The electronic BS-state occupied has great influence on their structure, energies and spectroscopical features, which makes the use of this method necessary. Early quantum mechanical studies on polynuclear models of iron-sulfur clusters employed a spin restricted molecular orbital framework but such models had deficiencies in describing ground and low-lying excited states [25]. As for H_2 at large internuclear distances, such a formulation fails when the overlap of atomic orbitals on two or more centers becomes very small and a molecular orbital description, which assigns electrons of both spins to each center, becomes inappropriate. In iron-sulfur clusters, the iron d-d interactions are weak, and the closed-shell solution is unstable. In fact perturbations act allowing electron spins to localise at the iron centers. Hence, there is a lack of correlation between electrons of opposite spin, provoking important effects, like the failure in predicting the existence of paramagnetic states. In order to circumvent these problems and average between molecular orbital and valence bond limits, a method which uses the idea of spin and space unrestricted wavefunctions, thus providing an improved description of exchange correlation effects, has to be employed. For this reason the Broken Symmetry formulation within DFT formulation (BS-DFT), has been used in our work and its theoretical formulation is described here in the present section. The BS method was first introduced by the pioneering work of Louis Noodleman (Scripps Research Institute, CA, USA) at the beginning of the '80s [26],[27] and became, in the years, a very common method in this field, especially since it has been formulated within DFT theory. BS is the appropriate method to compute Energetics and some Spectroscopic properties such as exchange coupling constants J in molecular magnetism, g-tensors and hyperfine coupling tensors A in EPR or isomer shifts δ and quadrupolar tensors ΔEQ in Mössbauer spectroscopy for poly-radicals and bioinorganic clusters. This section will concentrate on the calculation of energy values upon description of electronic configurations. The description of antiferromagnetically coupled clusters requires the use of a multideterminantal wavefunction, while DFT approach is a monodeterminantal method. The basic idea behind the BS method is to model the coupling where α or β spin excess is localised on the different spin centers of the system, in such a way that the wavefunction $\Psi_{unrestricted}$ can be written as a linear combination of pure spin states. To describe what a broken symmetry is, an example of a simple case will be used. In a dimer of copper Cu(II) with d^9 configuration, each atom bears one $\frac{1}{2}$ spin electron and can be described as:

$$|s, ms\rangle \tag{2.23}$$

where s is the spin of the electron and ms its magnetic quantum number. A spatial extension $a(\vec{r})$ and $b(\vec{r})$ is assigned to each unpaired electron ($u.e^-$), in such a way that two local monomer states are defined as:

$$|s_a, ms_a \rangle; |s_b, ms_b \rangle \quad (2.24)$$

Starting from these two local monomer states is possible to build the determinant Ψ_{ab} with total $Ms = ms_a + ms_b = 0$. Here the spin arrangement is such that monomer a has spin α , $s_a = \frac{1}{2}$ and $ms_a = +\frac{1}{2}$, while monomer b has spin β , $s_b = \frac{1}{2}$ and $ms_b = -\frac{1}{2}$. Broken Symmetry state is thus define when the α spin electron of the orbital extension $a(\vec{r})$ is spatially localised on one side and β spin electron of the orbital extension $b(\vec{r})$ is spatially localised on the other side. $\Psi_{ab} = |\frac{1}{2}, +\frac{1}{2} \rangle \otimes |\frac{1}{2}, -\frac{1}{2} \rangle$ represents what is called a spatially "broken symmetry" state. It is important to notice that it does not represents a pure spin state, in fact it is not an eigenvalue of the operator \mathbf{S}^2 . A very important point to address is that such BS determinants can not be built for all molecules. In fact, they are constructed for chemically valence-bond (VB) systems, i.e., for systems in which, to a first approximation, the dimer (trimer, etc.) can really be considered as the sum of its two (or more) weakly interacting monomers. For example, in the simple case of the H_2 molecule, this occurs when the two hydrogen atoms are sufficiently far away from one another (i.e., during dissociation), while in the case of the Cu(II) dimer, this usually occurs when the two metal sites are indirectly linked via intermediate bridges. This latter is the same as happens in metalclusters of enzymes treated in this work, where iron centers are bridged through sulfur atoms. BS states, above described, are related to pure spin states (i.e. singlet and triplet states). Within the wave function theory (WFT) pure spin states can be constructed mixing various determinants, obtaining an eigenfunction of the operator \mathbf{S}^2 . For example, in the Cupper dimer is possible, starting from the two monomers with $s = \frac{1}{2}$, to built two simple spin-coupled magnetic states which are defined with $\Psi_{S, Ms} = |S, Ms \rangle$.

In this way a singlet Σ state will be:

$$\begin{aligned} \Psi_{0,0} &\rightarrow \mathbf{S}^2 |0,0 \rangle = S(S+1) = 0 \\ &\rightarrow S_z |0,0 \rangle = Ms = 0 \end{aligned} \quad (2.25)$$

and a triplet T will be:

$$Ms = -1, 0, +1$$

$$\begin{aligned}
\Psi_{1,M_s} &\rightarrow \mathbf{S}^2 |1, M_s\rangle = S(S+1) = 2 \\
&\rightarrow S_z |1, M_s\rangle = M_s
\end{aligned}
\tag{2.26}$$

Hence we can define both this states in terms of the localised BS states.

For the singlet state:

$$I : \Psi_{0,0} = |0, 0\rangle = (2^{-\frac{1}{2}})(\Psi_{ab} - \Psi_{ba}) \tag{2.27}$$

For the triplet state:

$$II : \Psi_{1,+1} = |1, +1\rangle = \left| \frac{1}{2}, +\frac{1}{2} \right\rangle \otimes \left| \frac{1}{2}, +\frac{1}{2} \right\rangle \tag{2.28}$$

$$III : \Psi_{1,0} = |1, 0\rangle = (2^{-\frac{1}{2}})(\Psi_{ab} - \Psi_{ba}) \tag{2.29}$$

$$IV : \Psi_{1,-1} = |1, -1\rangle = \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \tag{2.30}$$

In this way is clear that Σ and $T_{1,0}$ can be expanded in terms of the localised BS states Ψ_{ab} and Ψ_{ba} , while $T_{1,+1}$ and $T_{1,-1}$ can not and are described instead with one determinant when both spins are parallel. This means that they can be computed with DFT methods and that energies are $E(II) = E(IV)$, which is none but the energy of the triplet state $E(\Psi_{1,M_s})$ called E_T . Moreover the three magnetic components of the triplet state (II to IV) are energetically degenerate in the absence of magnetic field. It is thus possible to indirectly compute the energy of the two-determinant state $\Psi_{1,0} = |1, 0\rangle$: $E(III) = E_T$. The next step is to calculate the energy of the singlet state $E_{\Sigma_{0,0}}$. This can be done combining $\Psi_{0,0}$ and $\Psi_{1,0}$ so that:

$$\Psi_{ab} = (2^{-\frac{1}{2}})(\Psi_{0,0} + \Psi_{1,0}) \tag{2.31}$$

and since $E(\Psi_{ab}) = E_{BS}$ we will have:

$$E_{BS} = \left(\frac{1}{2}\right)(E_{\Sigma} + E_T) \tag{2.32}$$

and

$$E_{\Sigma} = 2E_{BS} - E_T \tag{2.33}$$

Although most pure spin states, being multi-determinantal, cannot be directly computed within DFT, the energies of some of them (here the singlet state Σ) can be indirectly computed through appropriate combinations of pure spin states (here the triplet state) to yield broken symmetry states accessible by DFT. This last statement can also be generalised to dimers whose local

monomer spins are larger than $\frac{1}{2}$. In such systems the triplet state of the previous case will correspond to an high spin (HS) state where all the electrons have the same spin. This state is in fact a mono-determinantal state since it exists only one way to represent a configuration with all unpaired electrons. Again in such a system the energy of the singlet state E_Σ has to be calculated and this can be done indirectly defining an appropriate BS state with $Ms = 0$, of the general form $\Psi_{ab} = |s_a, ms_a\rangle \otimes |s_b, ms_b\rangle$. The key relation in such a case is the following general relation linking global pure spin states $\Psi_{S,Ms} = |S, Ms\rangle$ and localised broken symmetry states $|s_a, ms_a\rangle \otimes |s_b, ms_b\rangle$:

$$|S, Ms\rangle = \sum_{ms_a, ms_b} C_{S \rightarrow BS}(ms_a, ms_b) \cdot |s_a, ms_a\rangle \otimes |s_b, ms_b\rangle \quad (2.34)$$

where $ms_a + ms_b = Ms$ and $\sum_{ms_a, ms_b} C_{S \rightarrow BS}(ms_a, ms_b)^2 = 1$ and $C_{S \rightarrow BS}$ are Clebsch-Jordan coefficients [28], defined for each combination of spin states. For tabulated values of Clebsch-Jordan coefficients see the book Atoms and molecules, page 28, Table 1.4 [28] where the most common are resumed. Knowing that a pure spin state has energy:

$$E_S = (-J/2)[S(S+1) - (s_a(s_a+1) - (s_b(s_b+1)))] \quad (2.35)$$

where J is the exchange coupling constant, and reversing the eq. 2.34, which becomes:

$$|s_a, ms_a\rangle \otimes |s_b, ms_b\rangle = \sum_S C_{BS \rightarrow S}(S, Ms) \cdot |S, Ms\rangle \quad (2.36)$$

we can obtain expressions of the energy of BS and therefore Σ states.

2.3 Molecular Mechanics

The Molecular Mechanical (MM) methods are based on classical physics, where the atoms are considered as spheres connected each other by springs. Here electrons are ignored and energies are calculated only depending on the positions of the nuclei of the system, which is possible within the Born-Oppenheimer approximation. From this simple idea several force fields have been developed, with different degrees of complexity. This method is significantly faster than QM methods and can be applied to larger systems such as enzymes, proteins and polymers. One of the problems in the use of the MM theory is that each atom, bond, angle and torsional angle as well as the different forces between them has to be parametrized, and usually the more general

the parameters are, less accurate is the result for MM and vice versa. Among others, frequently force fields are for example CHARMM [29] (applied in this work), AMBER [30] and GROMOS [31].

2.3.1 Classical Force Fields

A general formulation for a force field is:

$$\begin{aligned}
V(r^N) = & \sum_{bonds} \frac{k_{l,i}}{2} (l_i - l_{i,0})^2 \\
& + \sum_{angles} \frac{k_{\theta,i}}{2} (\theta_i - \theta_{i,0})^2 \\
& + \sum_{torsions} \frac{V_n}{2} (1 + \cos(n\chi - \gamma)) \\
& + \sum_{improper} \frac{k_{\omega,i}}{2} (\omega_i - \omega_{i,0})^2 \\
& + \sum_{i=1}^N \sum_{j=i+1}^N \left(4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right).
\end{aligned} \tag{2.37}$$

The bonds and angles are represented by a simple harmonic expression, torsion around a dihedral bond by the Pitzer potential, energies arising from improper or out of plane motions, the van der Waals (VdW) interaction represented by a Lennard-Jones potential and electrostatic interactions by a Coulombic interaction of atom-centered point charges. Electrostatic and VdW interactions are calculated only between atoms in different molecules or for atoms in the same molecule separated by at least three bonds. Those non-bonded interactions separated by exactly three bonds (1-4 interactions) are reduced by the application of a scale factor.

2.3.2 Molecular Dynamics Simulation

The Molecular Dynamics (MD) simulation method is based on the solution of Newton's second law:

$$\frac{d^2 x_i}{dt^2} = \frac{F_{x_i}}{m_i} \tag{2.38}$$

where F is the force exerted on particle i of mass m_i in x direction. From the knowledge of the force on each atom, it is possible to determine the acceleration of each atom in the system. Integration of the equations of motion yields a trajectory that describes the variation of positions, velocities and accelerations with time. From this trajectory, the average values of properties can

be determined. The method is deterministic, so that the structure at any future time can be determined from the initial one. A non analytical solution can be achieved by applying a finite difference method which is done by dividing the calculation into very small timesteps ($10^{-15} - 10^{-14}$ s). Different algorithms can be used. The most common are the Verlet, Verlet Velocity and the Leap Frog Method.

2.3.2.1 Verlet Method

In the Verlet Method [32] for the actual timestep t , positions $x(t)$ and forces $f(t)$ or accelerations $a(t)$ are available. To predict the atom positions for the next timestep $t + \delta t$, positions from the previous timestep $t - \delta t$ are required:

$$x(t + \delta t) = x(t) + \delta t v(t) + \frac{1}{2} \delta t^2 a(t) + \dots \quad (2.39)$$

$$x(t - \delta t) = x(t) - \delta t v(t) + \frac{1}{2} \delta t^2 a(t) - \dots \quad (2.40)$$

The combination of equations 2.39 and 2.40 gives:

$$x(t + \delta t) = 2x(t) - x(t - \delta t) + \delta t^2 a(t). \quad (2.41)$$

From the new positions $x(t + \delta t)$ new forces and accelerations can now be calculated. Drawbacks are the missing velocity and the necessity to include positions from the previous step into the calculation, meaning that approximations need to be applied to guess the positions $x(t - \delta t)$ for the first step.

2.3.2.2 Velocity Verlet Method

Velocity Verlet is an implementation of the Verlet algorithm, where positions, forces and velocities at time $(t + \delta t)$ are calculated in three steps. First, the velocity at half timestep:

$$v(t + \frac{1}{2} \delta t) = v(t) + \frac{1}{2} \delta t a(t) \quad (2.42)$$

using current velocity and accelerations at time t . In the next step this half timestep velocity is used to calculate the new positions:

$$\begin{aligned} x(t + \delta t) &= x(t) + \delta t v(t) + \frac{1}{2} \delta t^2 a(t) \\ &= x(t) + v(t + \frac{1}{2} \delta t) \delta t \end{aligned} \quad (2.43)$$

from which the accelerations $a(t + \delta t)$ and forces can be computed within a force field method. Then velocities for the next timestep are computed as:

$$\begin{aligned} v(t + \delta t) &= v(t) + \frac{1}{2}\delta t a(t) + \frac{1}{2}\delta t a(t + \delta t) \\ &= v(t + \frac{1}{2}\delta t) + \frac{1}{2}\delta t a(t + \delta t) \end{aligned} \tag{2.44}$$

2.3.2.3 Leap Frog algorithm

Also the Leap Frog algorithm [33] uses the half timestep $t + \frac{1}{2}\delta t$ but it computes the positions $x(t + \delta t)$ for the next timestep using the last half timestep of the velocities $v(t - \frac{1}{2}\delta t)$. Hence, velocities are not calculated at the actual timestep but are half shifted, so the method appears to have leap-frog character. Velocities and positions are determined according to:

$$\begin{aligned} x(t + \delta t) &= x(t) + \delta t v(t + \frac{1}{2}\delta t) \\ v(t + \frac{1}{2}\delta t) &= v(t - \frac{1}{2}\delta t) + \delta t a(t) \end{aligned} \tag{2.45}$$

2.4 QM/MM method: a combined approach

Molecular Mechanics (MM) methods simulate atoms as rigid charged spheres, use simple potentials to describe bonds, angle bending and torsions and account for electrostatic intermolecular interactions using a simple Coulombic potential and the steric interactions by means of a Lennard-Jones potential. Those are quite simple approximations and hence fast enough to be applied for simulations of large and complex organic and biological systems, like those studied in the present work. However, the MM force fields, are unable to describe the changes in the electronic structure of a system undergoing a chemical reaction, simple forming or cleavage of a bond. They are not applicable for unusual metal clusters which are present in the enzymes studied here. Such changes in electronic structure and such complex systems, require quantum mechanics (QM) methods for a proper treatment. Unfortunately, due to the very demanding computational cost, the application of QM is still limited to relatively small systems consisting of up to tens or several hundreds of atoms, or even smaller systems, depending on the level of theory employed. Thus, despite the fast increasing of computational capability available nowadays, molecular modelling and simulation of large, complex systems at the atomic level remain a challenge to computational chemists. Algorithms that combine QM and MM methods provide a solution to this problem, combining the accuracy of a quantum mechanical description with

the low computational cost of molecular mechanics. The incorporation of QM into MM can be accomplished in various ways, and the one used largely in the present work is the so-called QM/MM methodology. The basic idea behind this technique is the partition of a system in, at least, two regions: one, containing for example atoms involved in catalytic activity and in general the metal cluster with first and/or second solvation sphere, which is treated quantum mechanically (QM region) and one treated with molecular mechanics (MM region). This enables to speed-up the calculation and at the same time simulate the influence of the environment on the reactive core [34]. In the present work the systems are divided into three parts. One is the region treated quantum mechanically, usually centred on the electronically relevant of the system, which is embedded in the MM region. This second is usually defined up to a distance of 15 to 20 Å while the remaining outer part of the system is only contributing as point charges.

2.4.1 Energy of the system

Despite the easy idea behind the QM/MM scheme, the Energy formulation brings already the first problems. Within QM/MM framework exist two schemes which can be used to compute energies. One is the subtractive scheme:

$$E_{tot} = E_{MM}^{all} - E_{MM}^{QMreg} + E_{QM}^{QMreg} \quad (2.46)$$

in which the energy of the all system E^{all} is calculated at the MM level then the energy for the QM region E^{QMreg} calculated on the QM level is added while the same calculated at the MM level is subtracted, in order to avoid the double counting. A major disadvantage of the approach is that MM parameters for the QM region need to be assigned and in many cases, specially in the case of this work, where active sites include metals for which those parameters are not accurate enough. Moreover the interaction between QM and MM region is treated at MM level, which can lead to larger errors in the energy calculation. Another one is the additional scheme:

$$E_{tot} = E_{MM}^{MMreg} + E_{QM}^{QMreg} + E_{QM/MM} \quad (2.47)$$

where the last term $E_{QM/MM}$ is added to account explicitly for the coupling between the two regions, in terms of electrostatic, van der Waals and bonded interactions, as shown in the next equation:

$$E_{QM/MM} = E_{QM/MM}^{el} + E_{QM/MM}^{vdW} + E_{QM/MM}^b \quad (2.48)$$

Electrostatic interactions First term in equation 2.48 represents the electrostatic interaction between the QM and the MM region and can be evaluated in several ways, like the mechanical embedding, electrostatic embedding or polarized embedding methods. In the mechanical embedding method the QM calculation is performed for the QM region as in the gas-phase, while the electrostatic interaction between QM and MM regions is evaluated by the MM code, using a classical point charge model for the QM charge distribution. This is computationally efficient since QM energies, gradients and the Hessian are calculated at the same cost as in gas phase, plus it has the advantage that produces two values for the energies of QM and MM regions which are separable. However, there exist major drawbacks. First, charges in the MM region do not interact with the QM density, which is then not polarised by the electrostatic environment. Second, if the charge distribution of the QM region changes also charges should be updated, causing discontinuities in the potential energy surface (PES), since the charges of this region are simply handled with the MM method charge model. The electrostatic embedding, instead, consists in the inclusion of the QM region into a cloud of point charges representing the MM region. This is achieved including the MM point charges as one electron terms in the QM Hamiltonian, which then can be written as:

$$\hat{H}_{QM/MM}^{el} = - \sum_i^N \sum_{J \in MMreg}^L \frac{q_J}{|\vec{r}_i - \vec{R}_J|} + \sum_{a \in QMreg}^M \sum_{J \in MMreg}^L \frac{q_J Q_a}{|R_a - R_J|} \quad (2.49)$$

where q_J are the MM point charges located at \vec{R}_J , Q_a are the nuclear charges of the QM atoms at \vec{R}_a and \vec{r}_i are the electron positions. The indices i, J and a run over the N electrons, L point charges and M QM nuclei respectively. Now the QM/MM electrostatic interaction is treated at QM level which gives possibility to get a more advanced and accurate description of it. Moreover, the electronic structure of the inner region can adapt to changes in the charge distribution of the environment, even though the MM charge distribution near the boundary region still can not be polarized by the QM region, which is the major drawback of the method. In order to overcome this disadvantage, the polarized embedding scheme has been implemented. Several approaches allow to obtain a polarized MM charge distribution. For example, one way is to calculate the charges of the MM atoms at the QM/MM boundary region quantum mechanically in the presence of the QM part. Another way is a fully self-consistent definition where the polarizable MM model is integrated into the QM Hamiltonian.

Van der Waals interactions The atoms of the QM part are described by parameters from related atom types of the MM part, and the energy is then calculated at MM level. One of the drawbacks of this method is that errors can arise from van der Waals parameters not suitable for the individual QM atoms. However, as those are short range interactions, it is possible to reduce the errors associated by departing the QM/MM boundary region from the site of interest.

Bonded interactions When the boundary region cuts through a covalent bond, interactions from the bonded parts, count for the coupling. Different approaches are available to handle these covalent bonds cuts. One is the localized-orbital scheme, where hybrid orbitals, kept fixed during the calculations, are placed at the border between the two subsystems, simulating a saturated bond. A second possible method is the boundary atom scheme, where the first atom of the MM part is included in both the QM calculation, where it mimics a saturated bond, and in the MM calculation as a normal MM atom. The third method, which is the one used in this work, is the link atom scheme. In this method an additional atom, called the link atom, which is normally a hydrogen atom, is added to saturate the broken covalent bond. In this way the QM part is complete without any free bonds, while for the MM region the link atom is omitted so that the covalent bond between the boundary MM atom and the boundary QM atom is treated molecular mechanically. However, the presence of an additional atom causes overpolarization due to the close proximity to the MM boundary atom.

Chapter 3

Carbon Monoxide Dehydrogenase from *Carboxydotherrmus* *hydrogenoformans*

In this chapter, the results obtained for Carbon Monoxide Dehydrogenase are presented. A very tight interconnection between spectroscopy, crystallography and computation, has been fundamental to bring an exhaustive interpretation of the structure and function of a specific CODH protein from *Carboxydotherrmus hydrogenoformans*, which has been introduced in Chapter 1. IR spectroscopic experiments here discussed were performed by Dr. Ciaccafava at TU Berlin, while crystal structures of WT and mutants for *Ch*CODH were provided by Dr. Fessler working in the group of Prof. Dr. Dobbek at HU Berlin.

3.1 The CN^- inhibited state of *Ch*CODH

3.1.1 Introduction and Methods

A detailed picture for CO binding to the active site cluster of *Ch*CODH is not available, even though a bent geometry has been suggested of this ligand bound to the terminal Ni in the CODH-CO complex of *Methanosarcina barkeri* [7]. This observation is consistent with the fact

that bent geometry would lower the activation barrier for CO oxidation. Moreover, a crystal structure was resolved where cyanide, which is a structural analogue of CO and an inhibitor of the catalytic activity of CODH, has been found in a similar bent conformation [8]. More recently re-evaluation of those electronic density maps disproved this type of conformation for both CO and CN ligands [9]. Moreover, a new structure has been published for *Ch*CODH where cyanide ligand had a different binding mode [10]. The two possible binding modes are presented in figure 3.1.

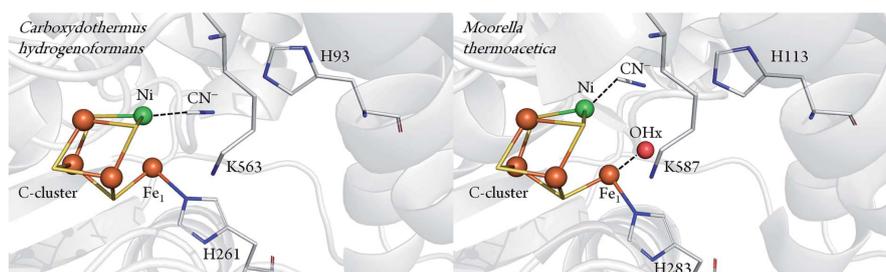


Figure 3.1: The two binding modes of cyanide to Ni site as in the resolved crystal structures of *Ch*CODH (pdb: 3I39) and *Mt*CODH (pdb: 3I04). The picture is reproduced by permission of The Royal Society of Chemistry from Ciaccafava, A., Tombolelli, D., Domnik, L., Fessler, J. *Chem. Sci.*, **2016**, 7, 3162 - 3171. <http://pubs.rsc.org/en/content/articlepdf/2016/sc/c5sc04554a>

3.1.1.1 Vibrational Spectroscopy study on *Ch*CODH-CN

To shed light on these controvert results, IR spectroscopy is a very useful technique since it is capable of probing the $C \equiv N$ stretching mode in a region where no other interferences by the protein environment are present. However the results reported previously for *Clostridium thermoaceticum* CODH showed multiple signals in this region. Therefore, a clarification was needed, and spectroscopic IR measurements have been carried out with wild-type and different engineered *Ch*CODH variants to assess the parameters that control the CN stretching frequency [3]. This analysis served as a reference for the quantum-mechanical/molecular mechanics (QM/MM) calculations which are here presented starting from different high-resolution crystal structures to elucidate details of the active site structure and to clarify some important characteristics of its protein environment. The IR spectroscopy has been used to identify the CN inhibited state.

After exposure of the protein to CN a sharp band appears in the C \equiv N stretching region at 2110 cm⁻¹. Free CN⁻ anion is expected to present a band at 2080 cm⁻¹ and this upshift is expected since Ni, to which cyanide ligand is bound, is a weak π -donor metal. The nature of the band was further proven with labelled ¹³CN⁻ revealing isotopic shift of 45 cm⁻¹. Moreover, the complex was also analysed in the crystalline state, following the same protocol used for the crystallographic analysis. The IR spectra obtained also revealed a band at 2110 cm⁻¹ that shifts down by 45 cm⁻¹ to 2065 cm⁻¹ when crystals were incubated with ¹³CN⁻. Thus, the frequency (and isotopic shift) of the Ni-bound cyanide of the crystalline complex is the same as for the sample in solution. In order to compare this new result with previous controvert publications, where other frequencies have been assigned to the CN stretching mode, some control experiments, which resembles the protocol used in those previous experiments, have been performed. In particular, a second protocol has been used which does not contemplate the removal of the buffer at the end of the incubation time. In previous experiments, performed on *Ch*CODH, a band at 2037 cm⁻¹ (1995 cm⁻¹ with ¹³CN⁻) and a second absorption appearing only at higher CN concentration at 2079 cm⁻¹ (2037 cm⁻¹ with ¹³CN⁻) were observed. Applying the second protocol, in fact, different bands at 2124, 2093 and 2037 cm⁻¹, shifting to 2078, 2061 and 1995 cm⁻¹ upon exposure to ¹³CN⁻, have been detected. The band at 2093 cm⁻¹ and its isotopically shifted counterpart at 2061 cm⁻¹ has been assigned straightforwardly to HCN and H¹³CN. Other two bands, instead, were not expected since there is only one binding site available for CN⁻ at the active site in *Ch*CODH. Since CN⁻ has high affinity towards metal ions like Fe and Ni a prolonged exposure of the protein to cyanide might lead to the formation of various metallic complexes. Control experiments have been therefore performed. When nickel chloride NiCl₂ is incubated with cyanide an absorption band at 2124 cm⁻¹ appeared, thus assigned to [Ni(CN)₄]²⁻ complex. Besides, the cyano complex formed with iron, [Fe(CN)₆]⁴⁻, absorbed at 2037 cm⁻¹. Hence, the control experiments allowed assignment of the three bands detected with this second protocol. Bands at 2080 cm⁻¹ (2036 cm⁻¹ for ¹³CN⁻) and 2037 cm⁻¹ (1995 cm⁻¹ for ¹³CN⁻) at pH 8, are attributed to aqueous CN⁻ and [Fe(CN)₆]⁴⁻, respectively, while the third band at 2124 cm⁻¹ (2078 cm⁻¹ for ¹³CN⁻) is assigned to the C \equiv N stretching of [Ni(CN)₄]²⁻. Ni-bound cyanide is in close vicinity to His93 and Lys563 which may interact with the cyanide-bound metal center. Addition of cyanide to a Lys563A variant displayed a wild-type like IR spectrum and the corresponding CN stretching band shows a downshift from 2110 to 2102 cm⁻¹.

3.1.1.2 The setup of QM/MM calculations

Based on these experimental observations QM/MM calculations became necessary in order to clarify which is the binding mode of CN detected by IR experiments (2110 cm^{-1} , downshifted to 2065 cm^{-1} when crystals were incubated with $^{13}\text{CN}^-$). Moreover, different calculations have been performed in order to validate the assignment of the bands appearing upon prolonged exposition to cyanide and assigned to Ni and Fe cyanide metal complexes, as shown before. A new crystal structure for *Ch*CODH-CN inhibited protein, at a very high resolution of 1.23 \AA , was available (PDB code: 5FLE). However, this structure did not differ in a substantial way from the previously published crystal structure of cyanide inhibited *Ch*CODH (PDB code: 3I39, $d_{min} = 1.35\text{ \AA}$), having RMSD for C_α of only 0.14 \AA . Starting from this crystal structure (PDB code: 5FLE) and the inhibitor free structure at -320 mV (PDB code: 3B53), eight different structural models, listed in table 3.1 were built. First, the protein was solvated in a box of $130 \times 130 \times 130\text{ \AA}^3$ made of explicit TIP3P waters. Hydrogens were inserted in their positions for residues at their standard state at pH 7, except for those residues in the vicinity of the C-cluster or involved in suggested proton transfer pathways [35]. Moreover, different protonation states for His93 and Lys563 were tested with the aim to investigate their interaction with the non-proteic ligands at the active site, with which they interact through hydrogen bond interaction. However, unless otherwise indicated, His93 was set double protonated on both N_δ and N_ϵ , and Lys563 side chain is positively charged ϵ -amino group (NH_3^+). Then, 20000 steps of energy minimisation were performed using NAMD [36] with the CHARMM22 [37] force field, in order to optimise hydrogen positions and orientation of water molecules. Heavy atoms, instead, were kept fixed to their initial positions. In this way a good starting point for the subsequent geometry optimisation can be obtained and, at the same time, the structure remains as close as possible to the crystal structure conformation. Specific force field parameters for the five metal clusters are lacking so they were kept fixed during the energy minimisation steps. Only charges for the metal clusters had to be computed with DFT methods, while van der Waals parameters were taken from values reported in the literature for similar systems [38]. Finally, structures resulting from the last minimisations step were refined through QM/MM calculations, employing ChemShell program [39]. This latter environment is an interface between a variety of QM and MM codes, but in this work Turbomole was chosen as quantum chemical program while the MM calculation was

performed using DLPOLY with CHARMM22 force field [37]. The system was divided in three layers, each of which was treated at a different level of theory and accuracy. The QM part consists of (1) the [Ni₃Fe₄S] C-cluster and the iron Fe1; (2) Cys295, Cys333, Cys446, Cys476 and Cys526, which coordinate the Fe atoms of the cluster; (3) His261, which is interacting with Fe1; (4) all the other non-proteic ligands, including cyanide on Ni in different positions and orientations and hydroxyl group on Fe1 in some of the models; (5) His93, Lys563 interacting directly with the non-proteic ligands, Gln332 interacting via hydrogen bonds with Lys563, and Ile567 whose role is important because it is responsible for the steric hindrance exerted on the cyanide ligand. DFT was employed in the optimisation of the QM part, with a pure density functional BP86 and combined with the resolution-of-the-identity (r̃) technique [23]. Two different basis set were used for metals (Ni and Fe) and non-metals. In particular, def2-TZVP and 6-31G* were employed, respectively. The reason of this choice was that the combination of high quality basis set and the BP86 functional [23] gives accurate molecular structures and harmonic vibrational frequencies, without the need of any additional correction [40]. This is possible thanks to an error compensation that finally allows direct comparison of calculated harmonic frequencies and experimentally detected ones. The second layer consists of water molecules and protein residues included in a sphere of 15 Å around the Ni atom of the C-cluster. All the atoms of this layer were treated explicitly at the MM level, and they are allowed to move during the optimization steps. Finally, a third layer was defined where all the remaining atoms, not included in the previous two layers, were enclosed. These latter were held fixed during the QM/MM geometry optimisation. Covalent bonds at the QM/MM border were cut and saturated by hydrogen link atoms, while the coupling between QM and MM was computed within electrostatic embedding and charge shift scheme [39]. Also, QM calculations on two models for the metal complexes of Ni and Fe which are formed from the degradation of the C-cluster upon prolonged incubation with cyanide. The calculations have been performed both in vacuum and applying PCM (Polarized Continuum Model) [41] to model the environment, using Gaussian09 [42] DFT calculations, with BP86 functional [23] and two different basis sets for metals (def2-TZVP) and non-metals (6-31G*), with the same set-up used for the QM part in the QM/MM calculations. The calculated vibrational frequencies, with implicit water environment, are in very good agreement with the experimental ones, thus providing a further confirmation of the assignment of the bands to the two metal complexes. CODH protein functions as a dimer. In fact, the C-cluster of one subunit

and the B-cluster of the other are closer than those from the same subunit in such a way that C- and B- cluster of different subunits exchange electrons more rapidly. This is reason why the biological assembly of *Ch*CODH has been employed here in this calculations. Moreover, a cavity below the C-cluster, which constitutes the electrostatic environment of the area where the enzymatic reaction takes place, is very close to the interface between the two monomeric parts of the protein. If a monomeric protein were employed this cavity would be filled with water molecules thus changing in a significant way the electrostatic of the catalytic center. In particular Lys563 forms a very strong hydrogen bond with Gln332, which is at the interface between the monomers and is stabilized by an hydrogen bond interaction with His195 of the second monomeric part. If this is not present in the model, Gln332 is solvent exposed creating instability in the nickel iron sulfur core environment. After optimisation, the structures were

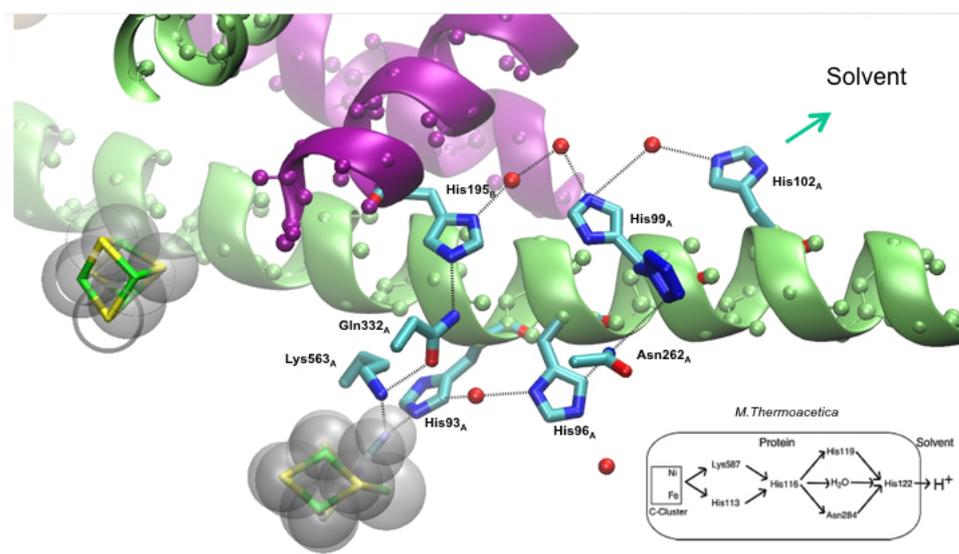


Figure 3.2: The interface between the monomeric units of CODH are presented in the figure. Also a possible proton transfer, proposed for *Mt*CODH [35] and hence respecting the numbering of this second structure, is quoted as a reference.

analysed and compared with the crystal structure. This comparison was based on relevant geometrical parameters (bond lengths, bond angles) and a tilt angle, which was defined as the angle between a unit vector originated in the center of a plane formed by Ni - S - Fe₄ - S and a vector directed along the axis of Ni-CN bond. The last step was the calculation of vibrational

frequencies of the quantum mechanical part in the optimised conformation surrounded by the point charges of the atoms of the second and third layers. In this case Gaussian09 program was employed with the same DFT functional and basis set as for the QM/MM optimisation [42].

Name	pdb	Origin	Presence of OH	Addition of cyanide
<i>Ch</i> CODH-CN _{5FLE} exp	5FLE	Exp.	No	-
<i>Ch</i> CODH-CN _{5FLE}	5FLE	Calc.	No	From cryst
<i>Ch</i> CODH-CN _{5FLE} /tilt	5FLE	Calc.	No	From cryst
<i>Ch</i> CODH-CN _{5FLE} /tilt/OH	5FLE	Calc.	Yes	From cryst
<i>Ch</i> CODH-CN _{5FLE} /K563dep	5FLE	Calc.	No	From cryst
<i>Ch</i> CODH-CN _{5FLE} /K563A	5FLE	Calc.	No	From cryst
<i>Ch</i> CODH-CN _{3B53}	3B53	Exp.	No	-
<i>Ch</i> CODH-CN _{3B53} /tilt	3B53	Calc.	No	in silico
<i>Ch</i> CODH-CN _{3B53} /tilt/OH	3B53	Calc.	Yes	in silico

Table 3.1: Models used in the QM/MM calculations

3.1.2 Results

First, the optimised structure of the CN inhibited state is compared to the starting crystal structure (pdb: 5FLE, $d = 1.2 \text{ \AA}$). The main structural parameters of C-cluster and residues around the organic ligand are taken into account. Only Ni-C(CN) bond length is predicted shorter by 0.11 \AA compared to the experimental value. A previous QM/MM study [9] also reported a comparably short Ni-C bond, also observed in the structure with CO₂ and NCO bound. Moreover, in a previous published *Ch*CODH-CN structure (pdb: 3I39, $d = 1.36 \text{ \AA}$), a shorter bond length was determined, in line with the results of an earlier EXAFS study. Ni atom has a square-planar geometry, coordinated to CN, arranged with a tilt angle of 6.40° of CN ligand. Cyanide, in this configuration, forms hydrogen bond with side chain of His93 ($d = 2.72 \text{ \AA}$) and Lys563 ($d = 2.98 \text{ \AA}$).

In a second calculation, the model *Ch*CODH-CN_{3B53}, generated starting from the *Ch*CODH structure (pdb: 3B53) by adding a CN⁻ ligand, was optimised. Also, prior to optimisation, a water molecule in the vicinity of the CN binding site was inserted to keep in line with the

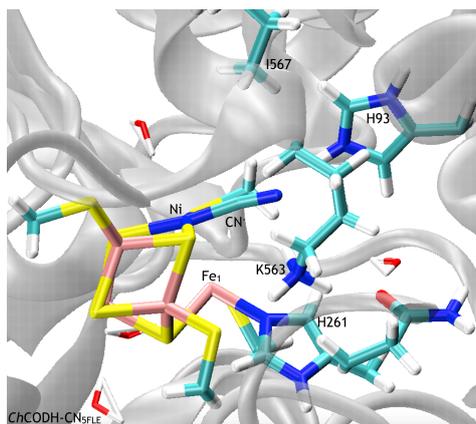


Figure 3.3: The optimised structure of the model $ChCODH-CN_{5FLE}$. For clarity, only side chains of the residues are represented.

experimental structure of $ChCODH-CN$. There is good agreement with respect to the geometry of the C-cluster and CN environment. The tilt angle is small (3.03 \AA) for the CN ligand and, as in the above mentioned structure, it is in hydrogen bond distance to His93 (2.74 \AA) and Lys563 (2.88 \AA). CN stretching frequencies have been calculated to be 2074 cm^{-1} and 2058 cm^{-1} , for $ChCODH-CN_{5FLE}$ and $ChCODH-CN_{3B53}$, respectively. These values are both lower than the experimental value of 2110 cm^{-1} .

Model	Ni-Fe1	Ni-C(CN)	N(CN)-N(K563)	N(CN)-N(H93)	CNtilt	Ni-C-N	ν_{CN}/cm^{-1}	(exp-calc)
$ChCODH-CN_{5FLE}^{exp}$	2.67	1.92	3.00	2.66	5.98	168	2110	-
$ChCODH-CN_{5FLE}$	2.65	1.81	2.98	2.72	6.40	169	2074	-36
$ChCODH-CN_{5FLE}/\text{tilt}$	2.66	1.82	3.79	2.69	18.42	171	2107	-3
$ChCODH-CN_{5FLE}/\text{tilt}/\text{OH}$	2.94	1.85	4.59	2.82	29.07	165	2054	-56
$ChCODH-CN_{5FLE}/\text{K563dep}$	2.73	1.83	3.97	2.72	24.44	173	2109	-1
$ChCODH-CN_{5FLE}/\text{K563A}$	2.71	1.83	-	2.70	23.52	173	2112	+10
$ChCODH-CN_{3B53}$	2.63	1.80	2.88	2.74	3.03	175	2058	-52
$ChCODH-CN_{3B53}/\text{tilt}$	2.68	1.81	3.94	2.66	17.71	175	2129	+19
$ChCODH-CN_{3B53}/\text{tilt}/\text{OH}$	2.86	1.83	4.65	2.76	37.35	165	2054	-56

Table 3.2: Structural parameters and vibrational frequencies of the CN stretching mode from experiments and calculations on different models

3.1.2.1 The tilted CN position

As previously postulated the initial binding conformation of cyanide to Ni may complete a distorted tetrahedral coordination geometry of the ligands around the metal atom. For this reason two models (*Ch*CODH-CN_{5FLE}/tilt and *Ch*CODH-CN_{3B53}/tilt) were built with an initial tilt angle of 60.71°. Upon optimisation they reached different minima, with higher tilt angles of $\sim 18^\circ$. The N-N distance between the CN ligand and the side chain of Lys563 is increased to 3.8-3.9 Å, while the distance to His93 remains unchanged. For these two models, the CN stretching frequencies calculated are closer to the experimental value, being 2107 cm⁻¹ and 2129 cm⁻¹, respectively.

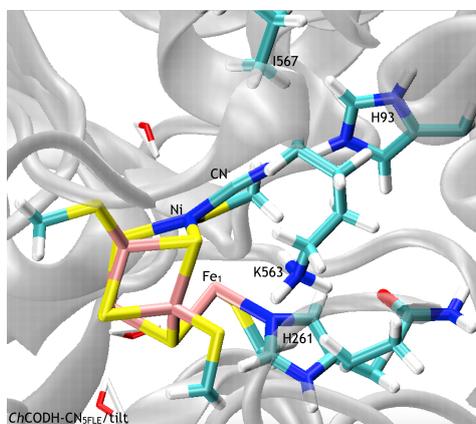


Figure 3.4: The optimised structure of the model *Ch*CODH-CN_{5FLE}/tilt. For clarity, only side chains of the residues are represented.

3.1.2.2 An hydroxyl is added to Fe1

In the two previous models (*Ch*CODH-CN_{5FLE}/tilt and *Ch*CODH-CN_{3B53}/tilt) an open cavity, allowing binding of hydroxyl group to Fe1, is left free. To probe the influence of OH ligand on the frequency of the vibrational stretching of cyanide, and to model a structure equivalent to the previous published CN-inhibited state of *Mt*CODH (pdb: 3I04), an hydroxyl group was added. Upon optimization the tilt angle increases from $\sim 18^\circ$ to $\sim 35^\circ$, consistent with the experimentally determined tilt angle in *Mt*CODH ($\sim 30^\circ$). When the hydroxyl group is present, an extra hydrogen bond to the CN molecule is formed, and this contributes to a decrease of the stretching frequency to 2054 cm⁻¹.

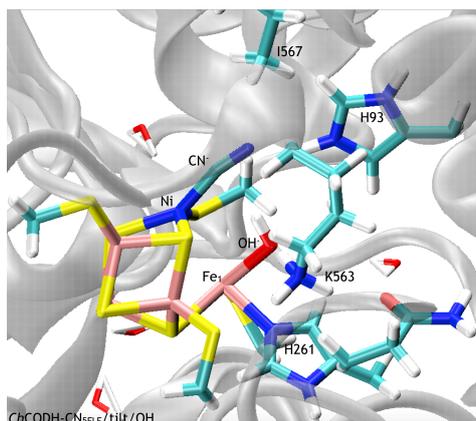


Figure 3.5: The optimised structure of the model *Ch*CODH-CN_{5FLE}/tilt/OH. For clarity, only side chains of the residues are represented.

3.1.2.3 The role of Lys563

In all geometries analysed so far, the hydrogen bond interaction between cyanide ligand and His93 is stable and always present. However, the tilt angle geometry seems to depend mainly on the interaction with Lys563. In order to clarify the role of this residue, a model for the K563A mutant was generated *in silico*, starting from the wild-type protein in the CN inhibited state (pdb: 5FLE). Upon optimisation, the CN assumes a configuration with a tilt angle of $\sim 24^\circ$. The calculated CN stretching frequency is 2112 cm^{-1} . The agreement with the corresponding experimental frequency of 2102 cm^{-1} , measured on the K563A mutant of *Ch*CODH, is within the error range of this type of calculation.

Finally, to study the interaction of Lys563 with cyanide and the subsequent changes of the corresponding CN vibrational frequency, a model with deprotonated Lys563, was investigated. The starting structure is again the wild-type of the *Ch*CODH-CN where a proton of the side chain of Lys563 is not present. This proton was the one in the proximal position to cyanide and thus participating in the hydrogen bond interaction, with the nitrogen of CN as acceptor and nitrogen of the side chain of the lysine as donor. This interaction is thus not established, and the calculated frequency for this latter optimised structure is 2109 cm^{-1} , which is in very good agreement with the experimentally measured of 2110 cm^{-1} .

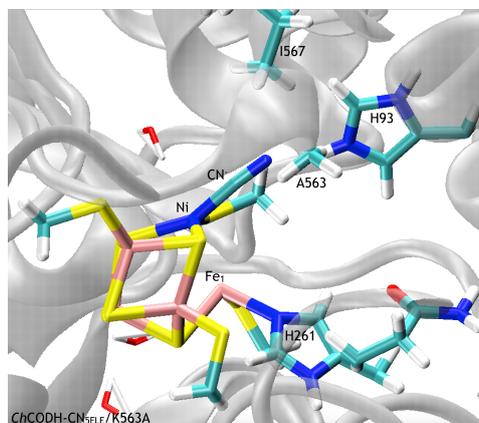


Figure 3.6: The optimised structure of the model *Ch*CODH-CN_{5FLE}/K563A. For clarity, only side chains of the residues are represented.

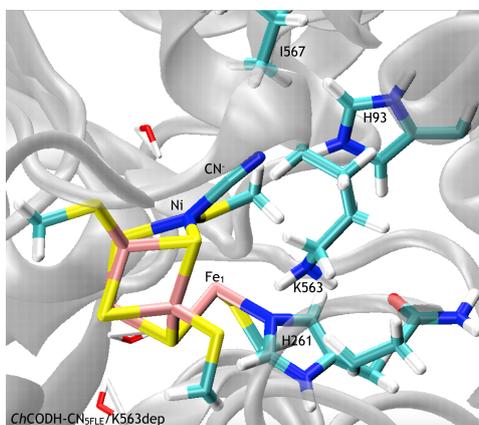


Figure 3.7: The optimised structure of the model *Ch*CODH-CN_{5FLE}/K563dep. For clarity, only side chains of the residues are represented.

3.1.3 Discussion

In early IR studies and, also, based on two different crystal structures of the inhibited state of CODH from *Ch* and *Mt* organisms, the idea of two different stretching modes for cyanide has been explained in terms of different modes of binding [8],[10]. In an attempt to reconcile crystallographic and spectroscopic data [43],[44] a two-step inhibition mechanism was proposed [45]. The idea was that in the first step cyanide reversibly binds to Ni, forming a distorted tetrahedral geometry around the metal, in a tilted mode as in the *Mt*CODH-CN X-Ray structure, similar to CO binding. In a second step, instead, the hydroxyl ligand on Fe1 is released, and a slow reorganization of the cofactor follows a more tight binding of the cyanide to the C-cluster. This step would then lead to the square-planar coordination geometry of Ni, as observed in *Ch*CODH-CN. The reassignment of the vibrational frequencies from the IR spectroscopy experiments, made by our collaborators reopened the discussion regarding the mode of binding of the cyanide ligand and its interaction with the environment. In particular, as already stated above, the band at 2110 cm^{-1} has been assigned to the $\text{C}\equiv\text{N}$ stretching, disregarding of the pH or temperature of the experiment and the state of the protein (solid crystal state or in aqueous solution). Moreover, bands at 2037 cm^{-1} (1995 cm^{-1} for $^{13}\text{CN}^-$) and 2124 cm^{-1} (2078 cm^{-1} for $^{13}\text{CN}^-$) were assigned to the $\text{C}\equiv\text{N}$ stretching of $[\text{Fe}(\text{CN})_6]^{4-}$ and $[\text{Ni}(\text{CN})_4]^{2-}$ at pH 8, respectively. The different models analysed in the present work intended to reproduce computationally the vibrational frequency assigned to cyanide and to determine the structure and the details of the environmental influence [3].

QM/MM calculations indicate that the cofactor-CN adduct is characterised by the lack of an hydroxyl ligand at the Fe1 as the hydrogen bonding interactions would cause a downshift of this vibrational mode by $\sim 50\text{ cm}^{-1}$ compared to the experimental value of 2110 cm^{-1} . Such a band has not been observed at any of the experimental conditions used in IR spectroscopy, so the adduct $\text{Fe}_1(\text{OH}^-)\text{Ni}(\text{CN}^-)$ is not formed. Of course, this result does not rule out the existence of such an intermediate state, but if it exists, it is not accumulated even at sub-stoichiometric *Ch*CODH/CN ratios and decays quite rapidly on the sub-minute time scale to its more stable $\text{Fe}_1\text{Ni}(\text{CN})$ form. Stretching frequency of metal-bound cyanide ligands is highly sensitive to the geometry of the cofactor, which affects its electronic configuration, as well as to the electrostatics in the environment, in particular, hydrogen bonding interactions. This has been already shown

also in previous studies on active site of [NiFe]-hydrogenases, where two cyanide ligands are bound to the Fe atom and provide, together with a third CO ligand a fingerprint for the interpretation of structures, even beyond what X-Ray cristallography can distinguish. Here, much stress has to be put on the positions occupied by protons, which are labile and not detectable in the crystal structures at the available resolution.

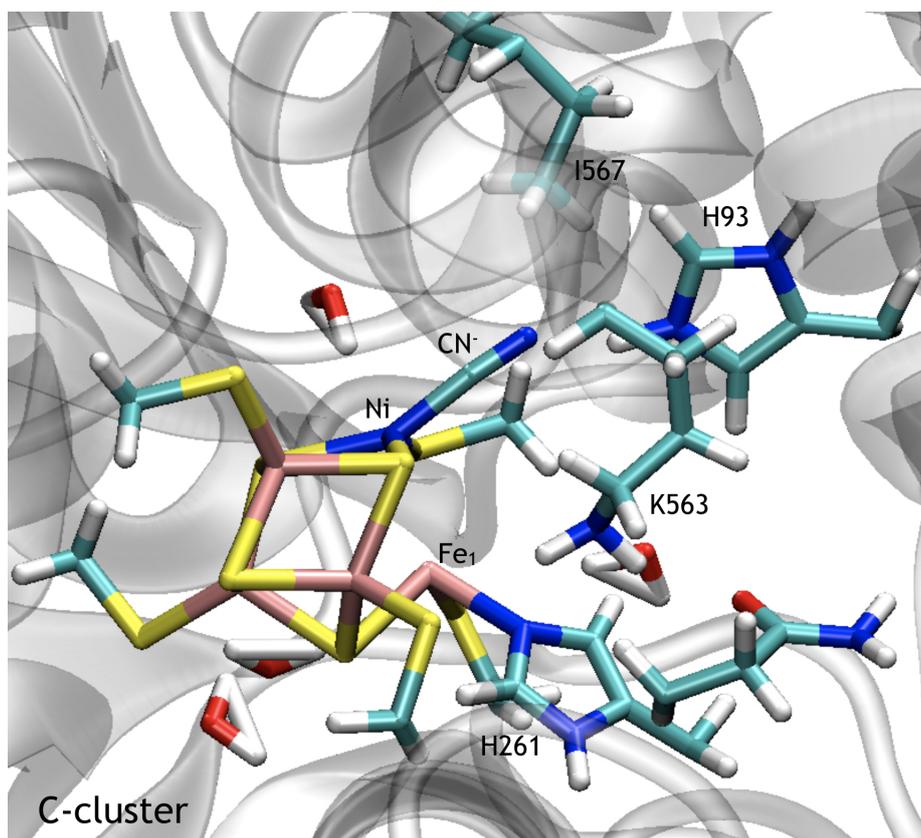


Figure 3.8: The structure of the active site of *Ch*CODH in the CN inhibited form.

Hence, to assess the importance of the parameters mentioned above, influencing the cyanide, and to relate the observed frequency with the crystallographic data, QM/MM calculations are here employed. Starting with the coordinates of the high-resolution structural model of CODH-CN inhibited state, the QM/MM-optimised cofactor does not show substantial displacement from the crystal structure (see *Ch*CODH-CN_{5FLE}^{exp} and *Ch*CODH-CN_{5FLE} in table 3.2). However, the corresponding CN stretching frequency is underestimated by 36 cm⁻¹, with a much larger deviation than the intrinsic inaccuracy of the theoretical approach [3]. A good frequency match,

instead, was obtained for the optimised structure of the Lys563A mutant in which the positively charged side chain in the vicinity of the cyanide ligand is deprotonated. Furthermore, the optimised geometry of the C-cluster displays a large tilt angle, of $\sim 24^\circ$. The same result is also obtained upon optimisation of the high-resolution WT structure but using a deprotonated Lys563. In this latter case, the increase of the tilt angle comes together with an increased nitrogen-nitrogen distance of the cyanide ligand to the Lys563 from 3 to 4 Å. A substantial weakening of the electrostatic interactions with the cyanide ligand is then produced. Also in this case, the calculated frequency of 2109 cm^{-1} agrees very well with the experimental value. The same large nitrogen-nitrogen distance but a somewhat smaller tilt angle is obtained for the wild-type protein upon raising the initial value for the tilt angle before optimisation. Again, an excellent agreement with the experimental CN stretching frequency is obtained with $\Delta\nu_{exp} - \nu_{calc} = +3$. In all the optimised structures, the N_ϵ atom of His93 remains at a distance of 2.7 Å from N of cyanide, which, together with the nearly linear $(\text{Ni})\text{CN}\ddot{\text{N}}\text{H}\ddot{\text{N}}\text{N}_\epsilon(\text{His93})$ geometry, points to strong hydrogen bonding interactions. The comparison of calculated and experimental CN stretching frequencies indicates that the interactions between the cyanide ligand and Lys563 must be weak, corresponding to a relatively large nitrogen-nitrogen distance (~ 4 Å). Such weak interactions point to a deprotonated side chain of Lys563. Hydroxyl ligand at Fe1 is identified as the potential proton acceptor, released upon CN binding to the Ni.

Time-independent frequency and intensity of the CN stretching probed with various IR experiments, indicates that Lys563 remains deprotonated over the time scale of hours. This is interpreted by the fact that the active site of the cyanide-bound C-cluster in the Cred1 state, is not accessible for protons from the solution. These protons should be translocated following a proton transfer pathway, identified in *MtCODH* [35], which is conserved also in *ChCODH*. It might be that structural changes along this proton channel, caused by the binding of cyanide may impair re-protonation of Lys563. As an alternative interpretation, a redox-state dependent proton transfer has been hypothesised. This view is confirmed by a very high turnover for CO oxidation, at the Cred1 state. This requires in fact a rapid unidirectional proton translocation from C-cluster to the solution. The reverse proton translocation, instead, occurs during enzymatic CO_2 reduction which involves the Cred2 state. Accordingly, the oxidation state of the C-cluster controls the directionality of the proton transfer such that re-protonation of Lys563 is inhibited in the cyanide complex of Cred1. A deprotonated Lys563, however, is not consis-

tent with the present high-resolution structure of the *Ch*CODH-CN complex which displays a distance of only 3 Å, concomitant to a smaller tilt angle, and thus points to appreciable Lys563-CN interactions. Whereas X-ray crystallography reflects the protein including a protonated side chain, the IR spectra that are identical in the crystalline state at low temperature and in solution at ambient temperature indicate a deprotonated Lys563 residue. To reconcile these findings, we suggest that, during accumulation of the diffraction data, radiation damage at the C-cluster may cause generation of a hydroxyl radical from a water molecule in the vicinity of Lys563, which is associated with a distinctly lower pKa of ca. 11 [46]. As a consequence, protonation of the Lys side chain could occur.

3.2 The NCO inhibited state of *Ch*CODH

3.2.1 Introduction and Methods

In the previous section the inhibition of *Ch*CODH by CN was studied, while in this section the binding of NCO, which is an analogue for CO₂, is presented. In particular, recent high resolution structure with $d < 1.1$ Å suggested that the binding of CO₂ and of the homologue NCO to crystals of *Ch*CODH would lead to the formation of a Ni-bound carboxylate and a carbamoyl H₂NCO⁺, respectively [47]. It has been hypothesised that, in analogy to what happens in the catalytic mechanism when CO₂ is reduced to CO, NCO would be thus reduced to CN by CODH. The same reduction of NCO to CN has been detected in nitrogenase thus strengthening what has been postulated. Moreover, a slow oxidative turnover of n-butyliisocyanide into n-butyliisocyanate for this specific enzyme has been observed, which has been interpreted as the corresponding reverse reaction. [47],[48],[49] In a recently published paper with our collaborators [4], IR spectroscopy has been employed to probe the substrate bound to the C-cluster of the enzyme, both in solution and in the crystalline state. The system has also been optimised, starting from the available high resolution crystal structure (pdb: 4UDY). On the optimised structure at QM/MM level, the vibrational spectra has been calculated to help out spectroscopic studies to make interpretation on the results obtained experimentally. All the IR experiments here presented have been performed by Dr. Alexandre Ciaccafava.

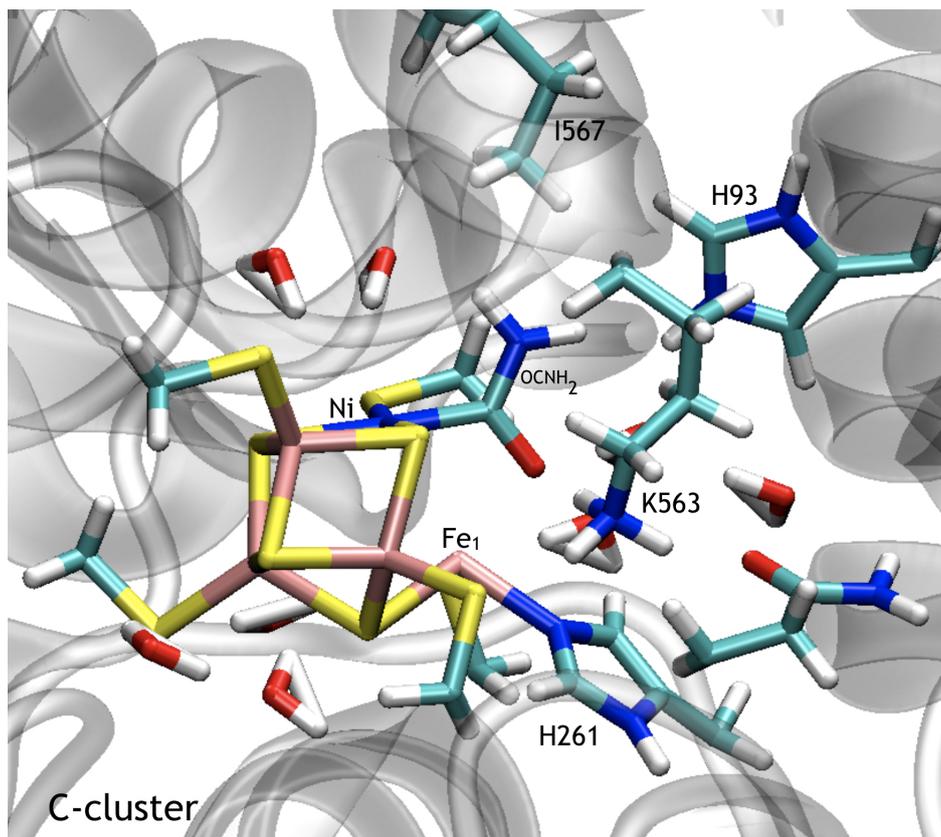


Figure 3.9: The structure of the active site of *Ch*CODH in the NCO inhibited form.

3.2.1.1 IR spectroscopy results

First, various experimental IR spectra have been measured for the NCO inhibited enzyme in solution. The dithionite reduced *Ch*CODH protein was exposed to a solution of NCO and the IR spectra obtained. Two bands were observed in the region between 2270 and 2000 cm^{-1} . One at 2168 cm^{-1} , attributed to a stretching mode of NCO in solution and the second one at 2110 cm^{-1} , in the same position of the one previously observed CN adduct and attributed to the CN stretching of Ni-bound cyanide [3]. Based on this observation, has been hypothesised that the C-cluster in the Cred2 state has an active role in the reduction of NCO to CN. More control experiments have been thus performed in order to eventually confirm this hypothesis. First, a Ni-deficient *Ch*CODH has been exposed to NCO and the registered spectra did not present the band at 2110 cm^{-1} . In a second experiment a DTT-reduced protein has been used, thus starting from Cred1 electronic state. Also in this second case, no band corresponding to the CN stretching band has been detected after incubation with NCO. Further experiments were later carried out in order to exclude the presence of any contamination of the NCO or as the product of NCO reduction to CN, which could be possible because of the presence of metal ions in the solution coming from the enzyme. Thus IR spectra of aqueous NCO with DT was detected and it did not present any band at 2080 or 2093 cm^{-1} . This two bands, as explained in the above section, have been assigned to the CN and HCN in solution. Further spectra in the presence of NiCl_2 and FeCl_2 have been studied, and they again did not show the two aforementioned footprint bands, thus allowing to rule out free CN as a contaminant. The IR spectroscopic studies have been further extended to NCO-bound enzyme in the crystalline state. In this second case, no signal due to the CN enzyme could be detected at 2110 cm^{-1} . The absence of this band in the crystalline state spectrum is not due to intrinsically lower spectral quality. These spectroscopic data allow to conclude that NCO is converted into CN at the C-cluster and that this transformation occurs in the enzyme in solution while is otherwise impaired in the crystalline state.

3.2.1.2 QM/MM calculation on the *Ch*CODH-NCO inhibited state

To explore the NCO vibrational spectra, calculations have been performed on the *Ch*CODH-NCO inhibited state. The starting point is the crystal structure of this inhibited state with pdb code 4UDY. First, it was solvated in a box of explicit TIP3P waters, hydrogens inserted in their

position for the protonable side chains at the standard state at pH 7. Two conformations of the His93 residue are resolved in the crystal structure. The binding of NCO to C-cluster changes the conformation of His93, shifted by 1.36 Å compared to the position found in free and CO₂-bound CODH. The distance between the nitrogen of NCO and N_ε of His93 to 3.63 Å is thus increased. The second conformation of this residue has a lower occupancy of 35%, where the His93 is in the same position as in the CO₂-bound and free CODH. It has been stated that the driving force for the conformational change of His93 is the need to create the required space for the two hydrogen atoms of the carbamoyl group. Hence, the most occupied conformations of His93 has been employed. The residue has been double protonated, since on the N_ε side it has interaction with NCO ligand, while on the other side forms a bond with Asp219, which is necessary to stabilise the position of this residue in the crystal structure. From this starting set-up, 20000 steps of energy minimisation were performed using NAMD [36] with the CHARMM22 force field [37], in order to optimise hydrogen positions and orientation of water molecules. Heavy atoms, instead, were kept fixed to their initial positions. Charges for the metal clusters were computed with DFT methods, while van der Waals parameters were taken from values reported in the literature for similar systems. Finally, the resulting structures are refined through QM/MM calculations, with ChemShell program, employing Turbomole for the quantum chemical part, and DLPOLY for the MM calculation. The system is divided into three layers, treated at a different level of theory. The QM part consists of (1) the [Ni₃Fe₄S]C-cluster, with NCO ligand and the iron Fe1; (2) Cys295, Cys333, Cys446, Cys476 and Cys526, which coordinate the Fe atoms of the cluster; (3) His261, which is interacting with Fe1; (4) His93, Lys563 interacting directly with the non-proteic ligands, Gln332 interacting via hydrogen bonds with Lys563, and Ile567. DFT was employed in the optimisation of the QM part, with a pure density functional BP86 [23] and two different basis sets for metals (Ni and Fe) and non-metals, namely def2-TZVP and 6-31G*, respectively. The second layer is made up of water molecules and protein residues within a sphere of 15 Å around Ni atom. Atoms in this layer are treated explicitly at the MM level and kept free to move during the optimisation steps. Finally, in the third layer, all the remaining atoms were included. The obtained structure is analysed, comparing it to the crystal structure. The last step is the calculation of the vibrational frequencies for the QM part surrounded by point charges, employing Gaussian09 program [42].

3.2.2 Results

3.2.2.1 Analysis of the optimised structure

The optimised structure has been analysed by comparing it to the crystal structure. Recently the results of DFT calculations with B3LYP functional and LACV3P** basis set on this system, have been published [6]. This quantum chemical calculation will be used as a reference for the analysis of the QM/MM optimised structure, studied in the present work. A short Ni-C bond length of 1.85 Å is observed, confirming a substantial π backbonding effect, which is equivalent to a metal-to-ligand charge transfer. This effect increases the flux of electrons from the metal to the ligand. There is a very good agreement of the bond lengths and angles between the published crystal structure, DFT and QM/MM optimised structures. Some important parameters are listed in table 3.3. Defining a plane passing through Ni-C-N(CO) a slight out-of-plane (OOP) distortion of the oxygen of the ligand is present in the crystal structure (2.8°). The OOP angle is less pronounced in our calculations, compared to what has been observed previously [6]. The dihedral angle, in fact, assumes a value of $\sim 1^\circ$. NCO usually appears as a linear ligand in both monometallic and bimetallic complexes. Here, the bent geometry of NCO, observed in the crystal structure and confirmed in the present work, is atypical and further underlines the unique reactivity of C-cluster. QM/MM calculation confirms that NCO is best described in the carbamoyl form. At the QM level [6], the biggest deviation was observed for the C-N(H₂) bond length which decreased from 1.44 Å to 1.39 Å after DFT optimisation. In the QM/MM optimised structure instead, this bond is shorter with a value of 1.34 Å.

	Ni-C	N-C	C-O	O-Fe1	N-C-O	C-O-Fe1	Ni-C-O	OOP O(NCO)	OOP Fe1
Crystal Structure (4YUD)	1.79	1.44	1.33	2.04	114	106	123	2.81	5.03
DFT optimised	1.81	1.39	1.31	2.04	115	107	122	2.30	4.90
QM/MM optimised	1.85	1.34	1.31	2.17	119	105	119	0.53	3.91

Table 3.3: Structural parameters from experiments and calculations on *NCO*-CODH. Bond lengths and distances are expressed in Angstrom while angles in degrees.

3.2.2.2 Vibrational Frequency calculation on the QM/MM optimised structure

Frequency calculation has been performed, employing Gaussian09 [42], on the optimised structure of the NCO bound protein. The frequency is calculated on the QM part surrounded by the point charges of the atoms which were part of the other two layers. Characteristic vibrational modes of the carbamoyl moiety have been calculated to be 1598, 1432 and 1293 cm^{-1} . The vibrational mode at 1598 cm^{-1} involves the cyanide stretching along the CN bond and NH_2 unit bending, the mode at 1432 cm^{-1} involves the NCO asymmetrical stretching, while the mode at 1293 cm^{-1} involves the NCO symmetric stretching and HNC bending movements. Calculation for the N^{13}CO has also been performed and the vibrational frequencies calculated were 1595, 1395 and 1263 cm^{-1} . These band positions fall into a spectral region of saturating absorption such that their detection was impossible in the IR experiments performed by Dr. Ciaccafava [4].

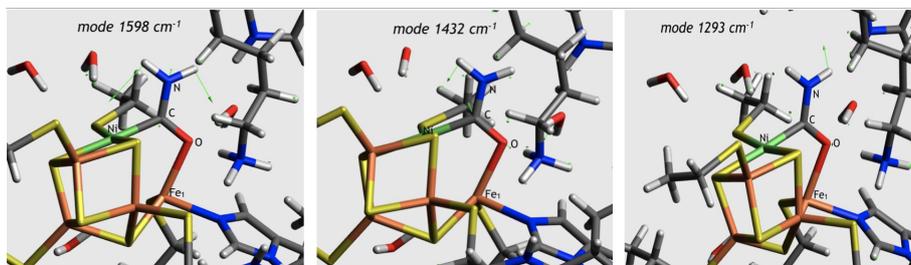


Figure 3.10: The normalised displacement vectors of the vibrational modes at the NCO ligand at the active site of *Ch*CODH-NCO.

3.2.3 Discussion and Outlook

The QM/MM calculation and the comparison with the crystal structure confirm the assignment of the NCO ligand to a double protonated NH_2CO moiety, bound to Ni site of C-cluster in *Ch*CODH. Also the position of the different residues located in the first shell environment of the C-cluster on the Ni side, is confirmed to be in agreement with such a ligand. Lys563 is keeping the interaction with NH_2CO and, most important, also His93 maintains its conformation, confirming the hypothesis that the A occupation, which was 65% occupied in the crystal structure, corresponds to the presence of NH_2CO ligand. Also the double protonation state of His93 is reflected by a substantial stability of its position with respect to the crystal structure upon optimisation.

The calculated frequency values for NCO-CODH confirmed that none of the bands observed in IR experiments could be assigned to the intrinsic vibrations of the NH_2CO moiety, bound to the C-cluster. Hence, we confirm that the peak observed at 2110 cm^{-1} in IR experiments can be assigned to the C-cluster bound cyanide vibrational mode. Therefore, where the 2110 cm^{-1} band has been detected, the conversion of NCO to CN took place.

In order for the conversion of the Ni-bound carboxylate to CO to take place, the uptake of two protons from the solution and a subsequent release of a water molecule from the active site are necessary. As assessed before [3], the directionality of the proton channel depends on the redox-state of the C-Cluster. In Cred1 state, only proton transfer to the solution is possible, while this direction is reversed in Cred2 because protons are necessary for the degradation of CO_2 to CO. The same mechanism should rule the NCO to CN conversion. However, the IR experiments showed a discrepancy between the solvated and crystalline state. We hypothesised that crystal lattice effects may prevent the formation of cyanide.

Chapter 4

Membrane Bound Hydrogenase

from *Ralstonia eutropha*

4.1 QM/MM method and the primary event of proton transfer

4.1.1 Results

The structural models of the Membrane Bound Hydrogenase (MBH) used for the QM/MM calculations were built starting from the crystallographic structure of reduced state (PDB: 3RGW) [50]. Unless specified otherwise, titratable side chains were protonated according to their standard state at pH 7. The entire system was then solvated in a box of TIP3P water molecules. The position of hydrogen atoms and the orientation of water molecules were optimised, in a first step, by means of 20000 steps of classical molecular mechanics energy minimisation using NAMD code [36] in combination with the CHARMM22 force field [37]. During this procedure, heavy atoms of the protein were kept fixed to their initial positions. Mulliken charges for the metal clusters were computed with DFT, where the BP86 functional [23], 6-31G* basis set for non-metals and def2-TZVP for metals are applied. Their Van der Waals parameters, instead, were taken from values reported for similar systems in the literature [38]. In a second step, the geometry of the [4Fe-3S] proximal cluster and its environment, made up of the protein and a sphere of explicit

TIP3P water molecules, with a radius of 40 Å around the center of the protein, were refined via hybrid QM/MM calculations. Accordingly, the system was divided into three layers, each of which treated at a different level of theory and accuracy. The QM part consists of (1) the [4Fe-3S] PC-cluster; (2) Cys19-Cys20 moiety, Cys17, Cys115, Cys120 and Cys149, which coordinate the Fe atoms of the cluster; (3) His229, which faces Fe₁; (4) Glu76 which is the first acceptor of the proton. All atoms in this region were described by broken symmetry density functional theory (BS-DFT) using BP86 functional [23] combined with the resolution-of-the-identity (r̃RĨ) technique. The def2-TZVP basis set was used for iron, while the 6-31G* basis set was employed for all other atoms of the QM part. The second layer was made up of protein residues and water molecules included in a sphere of 17 Å around the Fe₁ atom of the PC-cluster. All atoms in this section were treated in an explicit manner at a molecular mechanical level and allowed to move during the optimisation cycles. The third layer which includes the remaining atoms was held fixed during the optimisation procedure. As in the previous step, the CHARMM22 force field [37] was applied for describing the classical bonded and non-bonded interactions. Furthermore, no specific cut-off for the non-bonded interactions was used in the QM/MM calculations. Covalent bonds at the QM/MM border were cut and saturated by hydrogen link atoms. The coupling between QM and mobile-MM regions was computed using electrostatic embedding with a charge-shift scheme [39],[51]. In order to investigate the molecular mechanism which triggers the opening of the [4Fe-3S] cluster upon reduction and the subsequent proton release, three structural models of PC were built. As depicted in figure 4.1, these structural models differ in the composition and oxidation state of the proximal [4Fe3S] cluster. Model 1, represents the [4Fe3S] core in the oxidized state (4+) with the protonated nitrogen of Cys20 and anionic Glu76. Model 2 is similar to model 1 but with a charge neutral Glu76 and covalent bond between the Cys20 and Fe₄. Finally, model 3 resembles model 2 but with a one electron reduced [4Fe3S]³⁺ core. Moreover, a fourth model (model 0) was analysed, which is in the reduced state, with standard protonation, representing the crystal structure of the reduced protein (PDB: 3RGW). For these models, the broken symmetry state was set to BS12 according to Pelmenchikov, Kaupp, JACS, 2013, 135, 11809 [52] using the notation introduced by Volbeda et al. [53], in which the indexes denote Fe atoms carrying excess local spin antiparallel to the total spin. Reduced state of the proximal cluster has 3 Fe(II) [d6, S=2] and 1 Fe(III) [d5, S=5/2], while the oxidized state has 2 Fe(II) and 2 Fe(III). Within the BS formulation is possible to define two couples of Fe_a

and Fe_β , resulting $4 \times 4 \times 4 \times 4 / 2 = 6$ possible selections of the pairs among the four Fe ions. Since the identity of the Fe(III)/Fe(II) oxidation states within the α/β pairs is often lost during the calculation, mixed-valence pairs are formed with formal oxidation state of 2.5+, on average. This means that $\text{BSab} = \text{BSba}$, and only 6 configurations are available, which reduce to 3 for the diamagnetic oxidized state. In BS12, Fe_1 and Fe_2 have β spin density excess, resulting in a configuration such as $[-4/2, -4/2, +5/2, +4/2]$ for the reduced state and $[-4/2, -5/2, +4/2, +5/2]$ for the oxidized one.

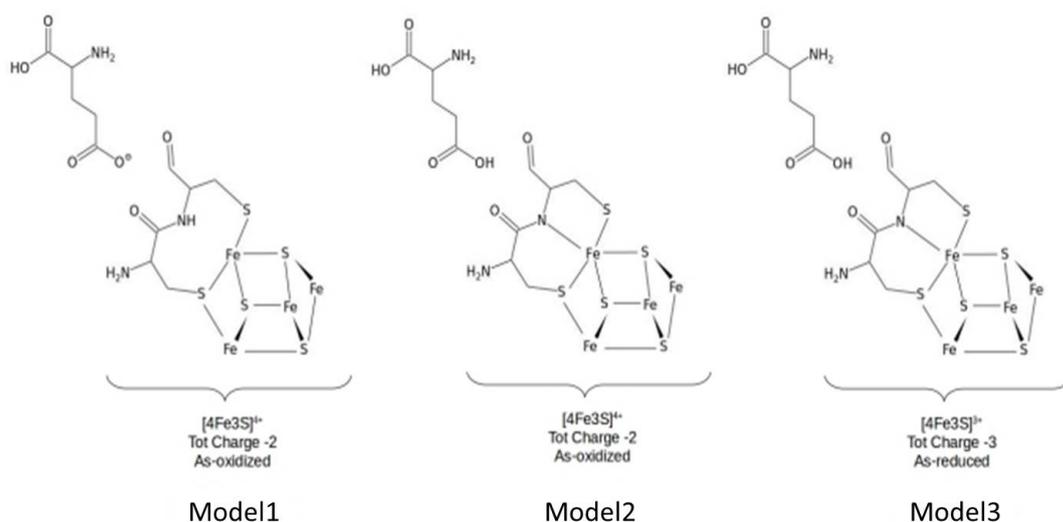


Figure 4.1: Schematic representation of the three models built and used as a starting point for three QM/MM structural optimisation.

All QM/MM calculations were performed with the ChemShell interface [54] linking the quantum chemical package Turbomole [55] with the molecular mechanics engine DLPOLY. Geometry optimisation used the limited memory quasi Newton L-BFGS algorithm working with delocalized internal coordinates and implemented in the HDLC optimiser [56].

First, model 0 is optimised to let the structure reach the equilibrium geometry within the aforementioned set-up. In table 4.1 RMSD (Root Mean Square Deviation) of the optimised model 0 with respect to the crystal structure for different parts of the QM region, are presented, while relevant data of the two structures are shown in table 4.2. Large displacements are observed around Fe_4 , Cys20 and Glu76, with a maximum of 0.50 Å at distance C20NH-E76O. Here proton

is adapting to the environment, establishing the hydrogen bond interaction, with an A-D distance (Acceptor-Donor) of 2.79 Å and A-HD angle of 144°. Anyway, no big changes are observed, and the model 0 optimised geometry is used as starting point for comparison with the other models.

RMSD	H ₂ -reduced MBH (model 0)
	[4Fe ₃ S] ₃ ⁺ E76 (COO ⁻) C20 (NH)
[4Fe3S]	0.087
QM-region	0.256
H229	0.256
E76	0.341
C20	0.213
C19	0.103
C17	0.251
C115	0.248
C120	0.157
C149	0.141

Table 4.1: RMSD values of QM/MM optimised model 0 with respect to crystal structure (PDB: 3RGW)

Later, model 1, 2 and 3 were optimised and they are compared with optimised model 0 structure. Some relevant structural parameters are listed in table 4.3. The optimised QM/MM structures of all the four models are depicted in figure 4.2.

Model 1 optimisation results in a closed form cluster. In fact, distances between Fe₄ - S₃ and Fe₄ - C20N change from 2.41 to 2.29 Å and from 3.54 to 3.63 Å, respectively, in comparison to model 0. Also in optimised model 1 displacements are mainly located around Fe₄ but no proton translocation is observed upon optimisation. The strength of the hydrogen bond interaction which is established between Cys20 and Glu76 is not affected by the extraction of an electron from the cluster core, maintaining similar structural parameters as observed in model 0, with an A-D distance of 2.84 Å and A-HD angle of 148°. Upon one electron oxidation the change

	X-Ray (Å)	Model 0
Fe1-S1	2.288	2.326
Fe1-S2	2.266	2.280
Fe1-C17S	2.304	2.213
Fe1-C19S	2.310	2.277
Fe2-S1	2.280	2.277
Fe2-S2	2.281	2.315
Fe2-S3	2.269	2.217
Fe2-C115S	2.267	2.267
Fe3-S2	2.293	2.224
Fe3-S3	2.289	2.378
Fe2-C120S	2.337	2.406
Fe2-C149S	2.303	2.311
Fe4-S1	2.279	2.275
Fe4-S3	2.315	2.413
Fe4-C20SG	2.302	2.288
Fe4-C19SG	2.310	2.355
Fe4-C20N	3.286	3.538
Fe1-Fe2	2.595	2.503
Fe1-Fe3	3.516	3.415
Fe1-Fe4	2.668	2.682
Fe2-Fe3	2.714	2.714
Fe2-Fe4	2.845	2.705
Fe3-Fe4	3.961	4.025
C20N-E76O1	4.891	4.606
C20N-E76O2	3.080	2.79
C20N-H-E76O2	–	144

Table 4.2: Structural parameters of QM/MM optimised model 0 with respect to crystal structure (PDB: 3RGW)

	Model 0	Model 1	Model 2	Model 3
Fe1-S1	2.326	2.314	2.300	2.300
Fe1-S2	2.280	2.222	2.226	2.293
Fe1-C17S	2.213	2.214	2.236	2.242
Fe1-C19S	2.277	2.352	2.350	2.298
Fe2-S1	2.277	2.239	2.331	2.383
Fe2-S2	2.315	2.310	2.314	2.330
Fe2-S3	2.217	2.294	2.273	2.211
Fe2-C115S	2.267	2.240	2.286	2.340
Fe3-S2	2.224	2.317	2.340	2.239
Fe3-S3	2.378	2.355	2.275	2.258
Fe2-C120S	2.406	2.358	2.359	2.404
Fe2-C149S	2.311	2.327	2.341	2.331
Fe4-S1	2.275	2.308	2.371	2.337
Fe4-S3	2.413	2.289	4.040	4.112
Fe4-C20SG	2.288	2.293	2.299	2.320
Fe4-C19SG	2.355	2.390	2.379	2.359
Fe4-C20N	3.538	3.630	2.027	2.082
Fe1-Fe2	2.503	2.563	2.620	2.512
Fe1-Fe3	3.415	3.494	3.576	3.532
Fe1-Fe4	2.682	2.736	3.000	2.908
Fe2-Fe3	2.714	2.629	2.653	2.728
Fe2-Fe4	2.705	2.711	3.763	3.610
Fe3-Fe4	4.025	3.868	5.205	5.213
C20N-E76O1	4.606	4.449	4.866	4.883
C20N-E76O2	2.79	2.83	3.37	3.48
C20N-H-E76O2	144	148	85	82

Table 4.3: Structural parameters of QM/MM optimised model 1, 2 and 3 with respect to model0.

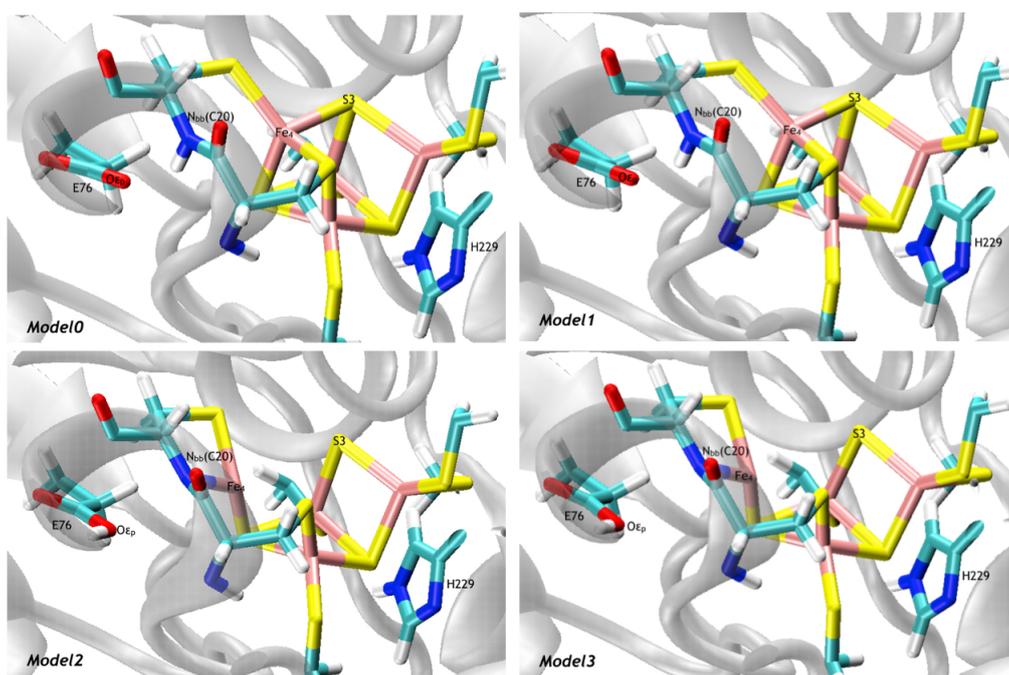


Figure 4.2: QM/MM optimised structures of the four models 0, 1, 2 and 3.

in charge is only of $+0.03$ and a positive partial charge is located on the iron atom. The total $+1$ formal charge is, among iron atoms, mainly distributed on Fe_3 . The starting geometry of model 2 has the proton moved to the glutamate carboxylic oxygen which then assumes its acidic form. This optimised model assumes the so called open form, which recalls the geometry found fully occupied in the superoxidized form crystal structure and partially occupied in the oxidized one, where Fe_4 is moving toward N of Cys20 backbone and a bond between these two is formed. Distances between $Fe_4 - S_3$ and $Fe_4 - C20N$ are changed by $+1.7 \text{ \AA}$ and -1.3 \AA , respectively. Here, the hydrogen bond interaction of protonated Glu76 and deprotonated nitrogen of Cys20 is not maintained. In fact, even though the A-D distance (3.37 \AA) could still be considered within the range of the hydrogen bond interaction the angle formed by the acceptor, proton and donor is too distorted (85°). Moreover, the translocation of H^+ away from the cluster core to Glu76 lets the Fe_4 atom bear a partial positive charge of 0.24 , the highest among all iron atoms of the core. Finally, structure of model 3 which also presents the translocated proton is found in the open form upon optimisation with distance $Fe_4 - S_3$ increased of 1.8 \AA and $Fe_4 - C20N$ decreased of 1.2 \AA . Here, the hydrogen bond interaction between $N_{bb}(Cys20)$ and $OH(Glu76)$ is lost, with

a A-D distance of 3.48 Å but and A-HD angle of 82°. In this case, where a proton is transferred, but only a vertical transition is modelled, the charge distribution is favouring a partial positive charge on Fe₄ which also acquires a $\Delta\delta$ charge of +0.32. All the changes in partial charges of model 1, model 2 and model 3, with respect to model 0, are listed in table 4.4. Finally, in figure

	Model 0	Model 1	Model 2	Model 3
	ρ_{charge}	$\Delta\delta_{charge(M1-M0)}$	$\Delta\delta_{charge(M2-M0)}$	$\Delta\delta_{charge(M3-M0)}$
Fe ₁	-0.23	+0.02	+0.04	+0.01
Fe2	-0.12	-0.06	+0.08	-0.01
Fe3	+0.16	+0.05	+0.06	+0.09
Fe4	-0.08	+0.03	+0.32	+0.32
S1	-0.26	+0.13	-0.03	-0.18
S2	-0.22	+0.08	+0.07	0.00
S3	-0.24	0.00	-0.25	-0.03

Table 4.4: Partial Mulliken charges changes in QM/MM optimised iron-sulfur core of model 1, 2 and 3 with respect to model 0.

4.3 the electron density (ρ_{elec}) distributions on the iron sulfur core and the residues included in the QM part are shown. In model 0 and model 1 the absence of ρ_{elec} between Fe₄ and NCys20 indicates that this bond is not formed, while the Fe₄ - S₃ bond remains stable. In model 2 and model 3, instead, upon optimisation the electron density rises in the area between iron and cysteine backbone nitrogen, where now a bond is formed, while it decreases between Fe₄ and S₃, indicating a bond breaking.

4.1.2 Discussion

The first point to address here is the structure of the oxidized state of the protein. The crystal structure, in fact, presents the iron sulfur cluster Fe₄ in two configurations, with 50% of occupancy each, while in all the other structures this atom is resolved in a fully occupied position. Moreover, the oxygen bound to Fe₁ in the oxidized state structure (4IUD, resolution 1.45 Å) is present with a reduced occupancy of 30%, as in the higher resolution (1.45 Å) superoxidized structure (4IUC). In the lower resolution (1.61 Å) crystal structure of the superoxidized state (4IUB), instead,

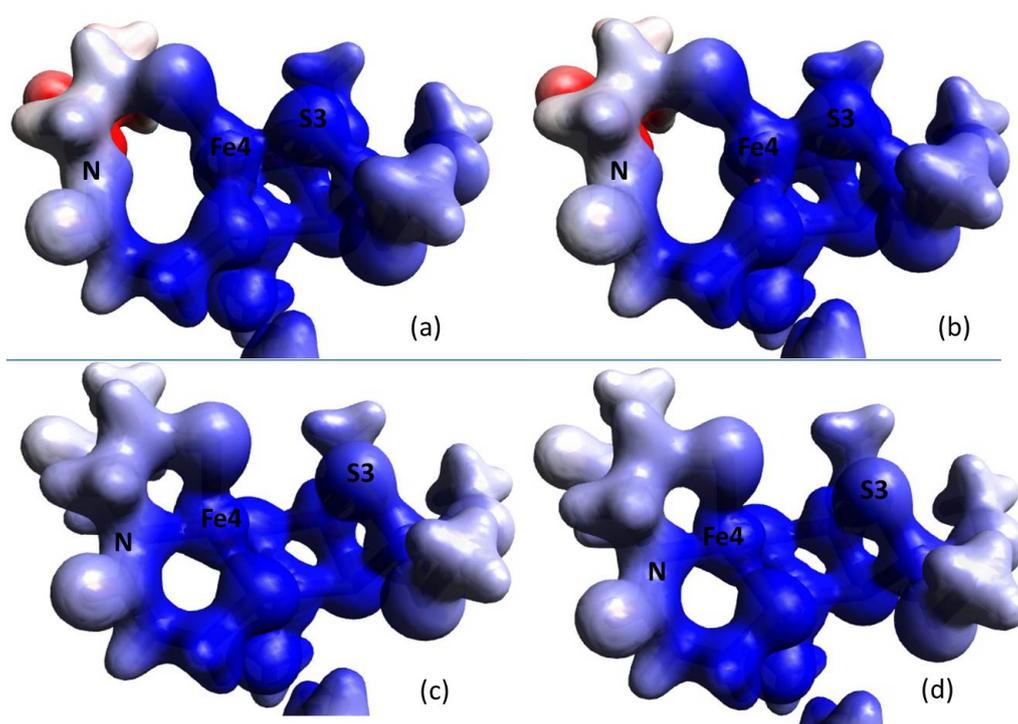


Figure 4.3: Figure 4: Electron density surfaces representations of the QM/MM optimised [4Fe3S] core for (a) H₂-reduced state , (b) model 1, (c) model 2 and (d) model 3.

the oxo-ligand has a higher occupancy of 88%. Since the oxidized crystal structure is resolved modelling a partially reduced state, is not possible to understand whether the features of the cluster are residues of the superoxidized conformation or they are intrinsic to the oxidized state itself. It appears that the crystal structure of the oxidized state is a mixture of two states. One where still the opening movement of the cluster is not completed, and one where a residual oxygen remaining from the superoxidized state partial reduction is still present. In fact, the absence of the opening movement observed in model 1, where only the extraction of one electron took place already, means that one electron oxidation is not providing enough energy to the iron sulfur cluster to rearrange in a way that resembles the superoxidized open structure. What comes out clearly from these calculations instead is that the open-cluster conformation is reachable only when the bond between Fe₄ and nitrogen of the backbone of Cys20 can be formed. This condition is reachable only when a proton is translocated to a glutamate residue in the vicinity of the cluster, and forming a hydrogen bond with N(Cys20). When the proton is moved to its new acceptor site, the glutamate rearranges its position breaking the previous existing Hbond and, de facto, freeing a site on the nitrogen which becomes more electronegative and is open to form a bond with Fe₄. This situation is reproduced in model 2 and model 3, regardless of the electronic state of the cluster. We can thus conclude that the opening movement of the cluster is triggered by the deprotonation of the N(Cys20). The consequent proton transfer to the first proton acceptor, which is identified with Glu76 side chain oxygen, near Fe₄ side of the proximal cluster, is the natural evolution of this deprotonation event. In conclusion, we can say that one electron oxidation is not enough to trigger the opening movement of the cluster, which anyway needs a second step, the proton transfer, in order to reach the configuration of the superoxidized state structure on the Fe₄ side. Proton transfer is not spontaneously taking place with one electron oxidation, meaning that the extrapolation of one electron does not provide the energy necessary to move the proton to Glu76, break the Hbond interaction with NCys20, and consequently break the Fe₄ - S₃ bond interaction and form the N - Fe₄ bond. Looking at resolved crystal structures of different oxidation states, we can speculate that the energy for this rearrangement must come from the abstraction of a second electron to reach the superoxidized state and moreover the formation of an interaction between Fe₁ and an oxygen coming from the protein environment. The optimisation of geometries of the four different models analysed above brings us to the conclusion that one electron oxidation is not enough to let the iron migrate

toward the amide N of Cys20. This movement is not possible since not enough positive charge is located on Fe₄. The extra positive charge cloud which is formed upon extraction of one electron from the cluster moiety, in fact, is distributed on all the atoms and not localised. This finding is a confirmation of what stated previously by Volbeda et al. [53], that the necessary second electron comes directly from the +2/+3 oxidation event of Fe₄. This justifies the need of a second electron oxidation event in order to allow the proximal cluster to protect the NiFe active site against the attack of O₂.

4.2 Proton Transfer Pathways study by means of MD simulations

4.2.1 Introduction and Methods

As stated in the introductory chapter 1.2, two different pathways have been proposed connecting the active site and the proximal cluster, in order to complete a proton-coupled electron transfer mechanism which is believed to be responsible for the protection of the active site against O₂. In the previous section 4.1, the study of the first transfer event which starts the Glutamate pathway was carried out. This section is focused on characterizing the proton transfer pathway from the NiFe active site to the [4Fe3S] PC cluster under the assumption that proton translocation occurs via the Grotthuss mechanism. According to this mechanism, proton migration is explained as a relay process in which one proton binds a certain molecule causing the release of another proton from the molecule who tries to find a new acceptor. Even though Grotthuss shuttling has been at first proposed to occur via water molecules, it can also take place through ionizable molecular groups such as titratable side chains of amino acids in proteins. Thus, this process is usually accompanied by slight structural rearrangements of the residues involved, so that they become ready to host the next proton [57],[58]. Furthermore, proton translocation via Grotthuss mechanism will be efficient if the hydrogen bond network connecting the initial proton donor and final proton acceptor is stable. Thus, the analysis of the proton transfer pathways presented in this work is based on a thorough statistical analysis of hydrogen bond probability and stability monitored in the course of a series of classical molecular dynamical simulations. A similar approach has been applied by Ginovska-pangovska in a 2014 paper [59].

4.2.1.1 The description of the Glu-path

According to the suggested Glu-path [19], a proton is released in the first event when the proximal cluster opens, it translocates to the carboxylate side chain of Glu76, and ends up at the bridge position of the NiFe active site. For this reason, this pathway has been named as the "Glutamate Pathway". The total length of the Glutamate pathway is of 11 Å in both the H₂-reduced and as-isolated forms of the MBH, and it involves five residues connected to each other via six hydrogen bonds (see Figure 4.4). Among these residues, the Glu16 deserves special attention. According to the crystal structures of the MBH in the reduced and superoxidized states, the E16 [50], may adopt two different conformations named as A and B. The A/B ratio is 54/46 and 50/50 in the reduced and superoxidized structures, respectively. Due to its conformational flexibility, the Glu16 has been suggested to act as a gate in proton translocation process. The pathway ends up at the active site where Cys597 receives a proton from Glu27. This latter, located close to Ni bound cysteine is important for proton transfer, based on mutagenesis studies and has been already identified as the active site terminus of various proton transfer pathways [60]. In order to analyze the probability of hydrogen bonds along the Glutamate Pathway four independent MD simulations were performed. While two MD simulations describe the forward proton translocation from the NiFe active site to the [4Fe3S] PC, the remaining two describe the backward translocation of the proton from PC to the active site. The difference between the forward and the backward translocation processes are a) the redox state and structure of the metal cofactors and b) the protonation of the residues involved in the pathway. For the forward translocation we assume that: 1) the active site, with a protonated Cys597, which is able to host a proton [61],[62],[63], is in the NiB state; 2) the [4Fe3S] PC lies in the superoxidized state (+5) as found in the crystal structure of the MBH in the superoxidized state. The backward translocation process, instead, is simulated considering the NiFe active site in the NiC state and the [4Fe3S] PC core in its reduced state (+3) as detected in the crystal structure of the reduced MBH. Furthermore, given the conformational flexibility of the Glu16, the MD simulations for the forward (and the backward) proton translocation processes were performed considering the two possible rotamers of Glu16. The setup used for these four simulations is summarised in table 4.5.

For each of the four MD simulations, the protein has been solvated in a box of explicit TIP3P

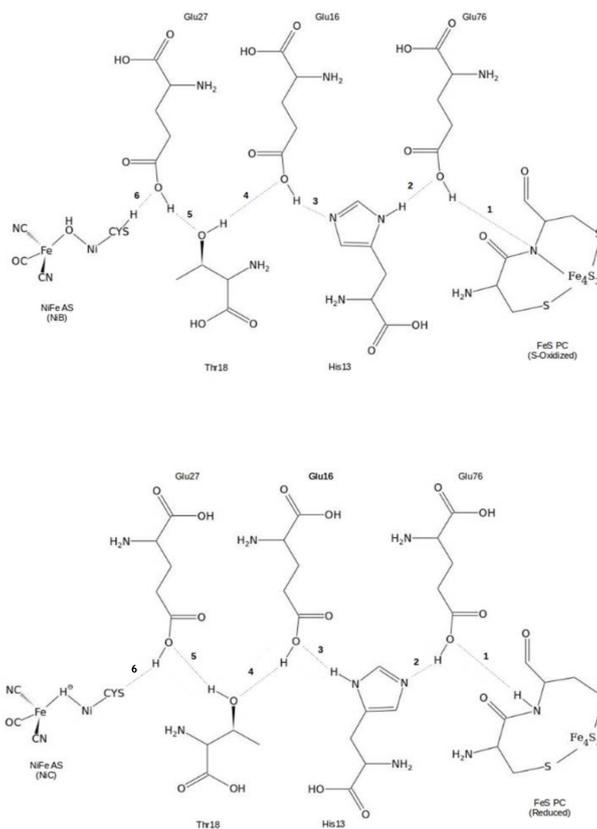


Figure 4.4: Schematic representations of the Glutamate pathway. The upper part of the figure is a two-dimensional schematic representation of the setup used to simulate the Forward translocation (Superoxidized state), while the lower part is the one used for the Backward translocation (Reduced state). In particular, the different protonation states of the residues involved are shown, together with the six H-bonds which will be considered in the analysis (see paragraph 4.2.2.3)

Simulation	[NiFe]	[4Fe3S]PC	X-ray structure	Protonation Pattern	E16-rotamer
Forward translocation					
1	NiB	Sup.Oxid.	4IUC	Scheme 1	A
2	NiB	Sup.Oxid.	4IUC	Scheme 1	B
Backward translocation					
3	NiC	Red.	3RGW	Scheme 2	A
4	NiC	Red.	3RGW	Scheme 2	B

Table 4.5: Summary of the setups used for classical MD simulations

water molecules, with periodic boundary conditions. The systems were first minimised, then gradually heated and, finally equilibrated. The simulations were then started at a $T = 300$ K, and trajectories generated with a length of 100 ns, with a time step of 2 fs. The protein and its environment were free to move during the MD simulations, except for the four metal clusters which have been kept fixed due to the lack of specific parameters. Mulliken charges for those clusters, in the different electronic states, have been calculated employing DFT methods with BP86 functional [23], TZVP basis set for metals and 6-31G* for non-metals. The simulations have been carried out using NAMD program [36] with CHARMM22 force field parameters [37].

4.2.1.2 The His-path can not be described through Grotthus mechanism

A second pathway has been proposed which connects the active site and the proximal cluster.

Two residues are here involved: His82, single protonated on the N_ϵ and forming a H-bond with Cys600, and His229, also single protonated on the ϵ nitrogen which is forming a very short and strong H-bond with OH ligand on the Fe_1 active site [14]. Moreover, a conserved water molecule is supposed to take part into the proton translocation, acting as a shuttle between the two histidines. This movement is hypothetically happening with the help of electrostatic interaction with backbone oxygens of Cys75 and Glu72. However, in order to translocate the proton which resides for both histidine residues on the opposite site with respect to the next acceptor, which is the crystal water molecule, the two residues need to rotate around the $C\beta^-C\gamma$ bond, first breaking their very strong interactions with the respective partners. Hence, it is not possible to treat this pathway through a simple Grotthus mechanism, within our assumptions.

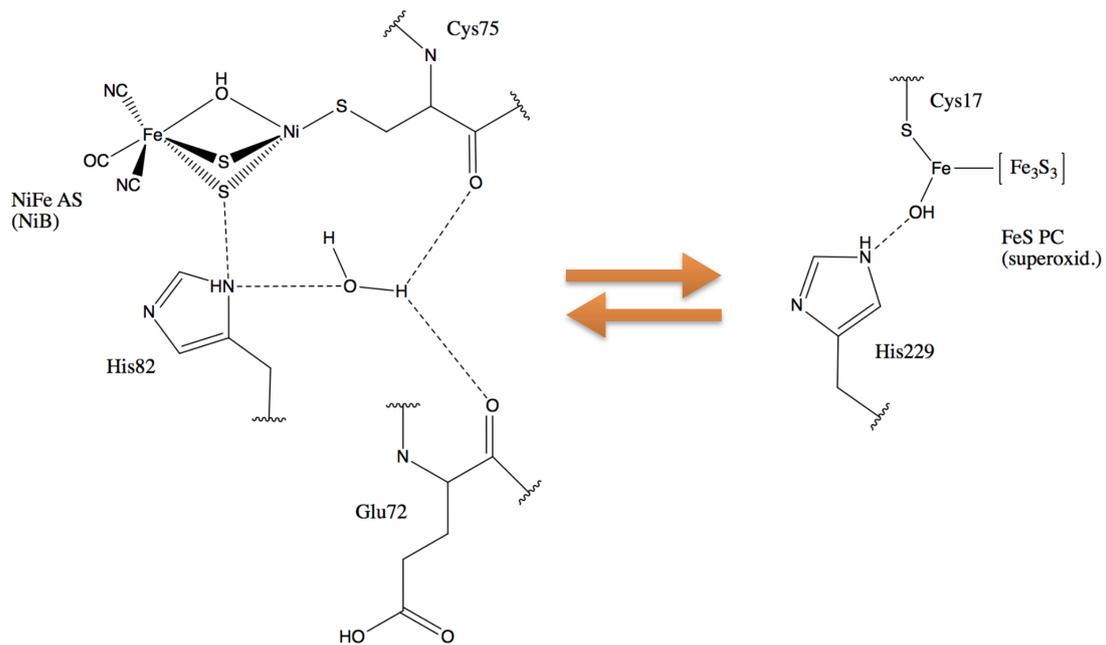


Figure 4.5: The Histidine pathway (His-path) in the *ReMBH* connecting active site and proximal cluster.

For this reason, no specific set-up involving this route has been built. Anyhow, a qualitative analysis was carried out, on this specific area where the residues involved in the His-path are located.

4.2.2 Results

4.2.2.1 The stability of the MD trajectories

In order to analyse the four MD simulations, the RMSDs (Root Mean Square Deviation) (Figure 4.6) of the trajectories have been calculated for the region of interest. This region includes the backbone of the residues which are located between the active site and the proximal iron sulfur cluster and are part of the two proton transfer pathways. Namely, they are Glu16, Glu27, Glu72, Glu76, His13, His82, His229, Cys19, Cys20, Cys75, Cys597 and Thr 18. The reduced structure simulations (Simulation3 and Simulation4, shown in blue and cyan in figure 4.6) are in general less stable than the ones of the oxidized structures (Simulation1 and Simulation2, shown in red and orange in figure 4.6).

RMSD of simulations 1 and 2 have average values R^2 of 0.45 and 0.42 and a maximum of 0.64 and 0.62, respectively. This reflects the structural stability of the region. The deviation in simulation 3 instead has a slightly increasing trend around 70 ns, with an average value of 0.52 and a maximum of 0.85. Finally, simulation 4 presents a more pronounced deviation in the last 10 ns. In this case, in fact, R^2 has an average value of 0.60 and it goes up to 1.33. This high R^2 is attributed to the lost of H-bond contact between Glu27 and Cys597.

4.2.2.2 The role of Glu16

Glu16 residue has at least two different conformations in the crystals. This makes it a good candidate to have gating function in the proton transfer following the Glu-path. In order to analyse its function and stability during MD simulations two dihedrals have been chosen which represent its conformational state during simulation: Ψ which is $C\alpha-C\beta-C\gamma-C\delta$ dihedral angle and Φ which is $C\beta-C\gamma-C\delta-Op$. A statistical analysis of the variation of these two is presented in figure 4.7, for the simulations of the superoxidized state and reduced state. In figure 4.7 the frequency counts for each of the two dihedrals considered are shown, together with their counting combination map (CCM). According to the MD simulation the Glu16, in the reduced

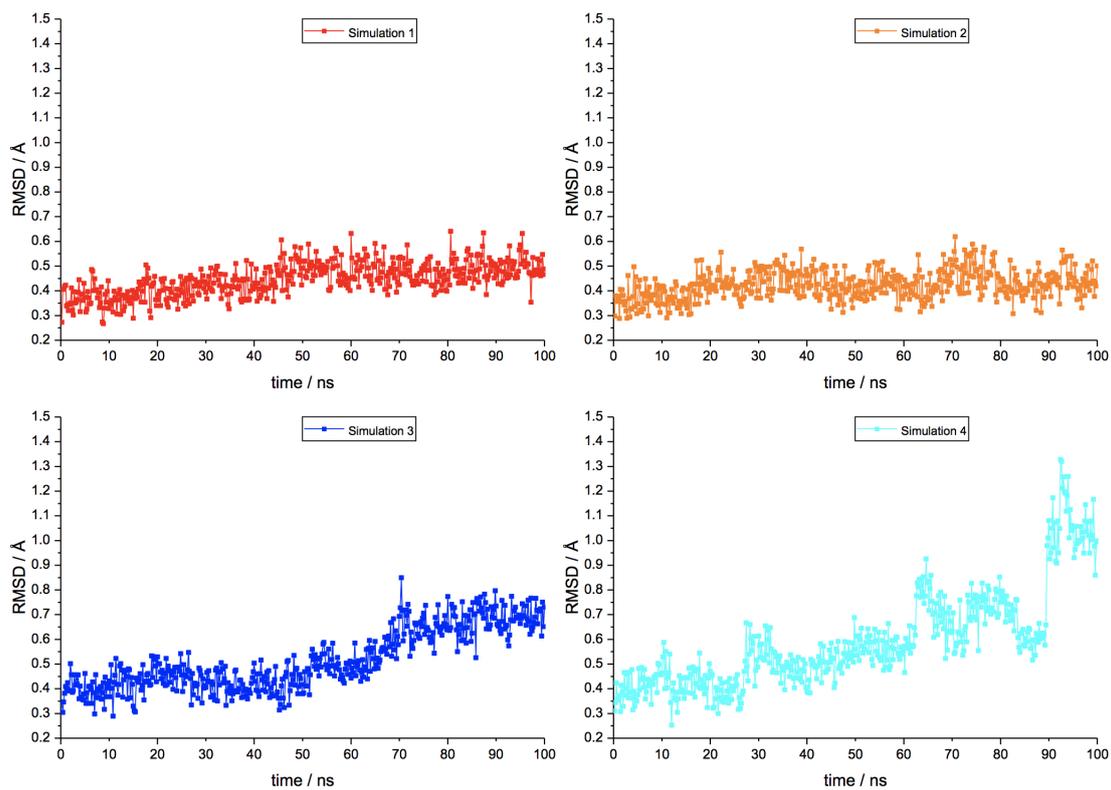


Figure 4.6: RMSD of selected residues of interest (selection A).

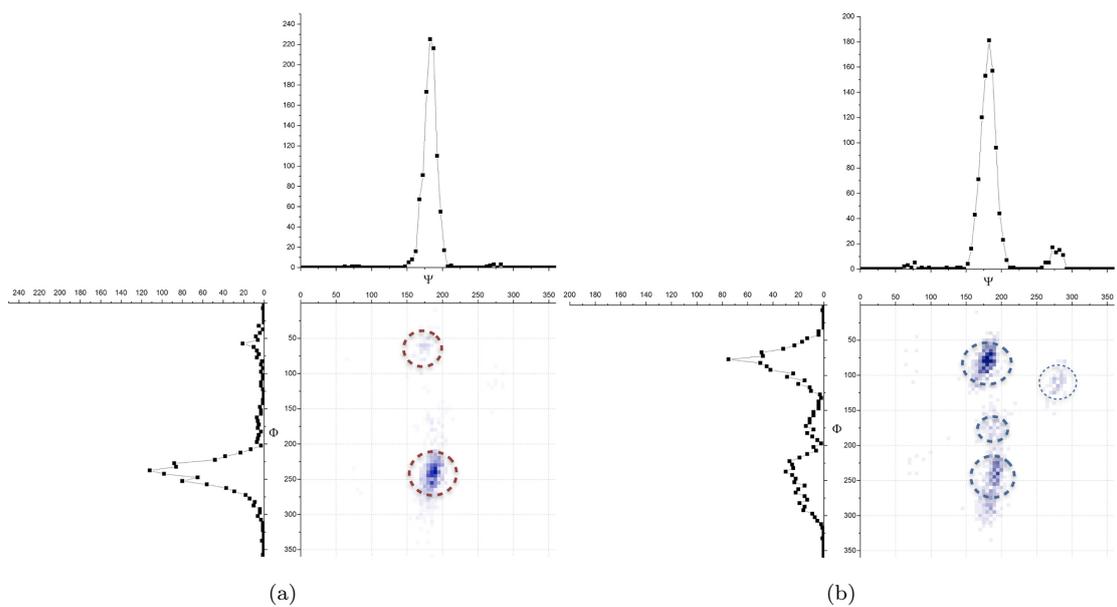


Figure 4.7: Variation of Ψ which is $C\alpha-C\beta-C\gamma-C\delta$ dihedral and Φ which is $C\beta-C\gamma-C\delta-O$ dihedral, of Glu16, during 100 ns MD simulations and the respective Counting Combination Maps (CCMs), for (a) the Superoxidized state simulations and (b) the Reduced state simulations.

state, may adopt four different states, as reflected by the four spots in the combination map in figure 4.7b. Three of them are characterised by the Ψ value of $\sim 175^\circ$ and different values of Φ of $\sim 75^\circ$, $\sim 175^\circ$ and $\sim 250^\circ$. The fourth one has Ψ of $\sim 275^\circ$ and Φ of $\sim 100^\circ$. In the superoxidized state, instead, Glu16 lies only two different states with a Ψ value of $\sim 175^\circ$ and different values of Φ of $\sim 60^\circ$ and $\sim 250^\circ$ (see table 4.6). The residue can assume more and different conformations

Ψ	Φ
Conformation A (PDB:4IUC)	
94°	218°
Conformation B (PDB:4IUC)	
177°	213°
Forward translocation (PDB:4IUC)	
175°	60°
	250°
Conformation A (PDB:3RGW)	
178°	41°
Conformation B (PDB:3RGW)	
168°	176°
Backward translocation (PDB: 3RGW)	
175°	75°
	175°
	250°
275°	100°

Table 4.6: Dihedrals Ψ and Φ of residue Glu16 along the MD simulations.

than what has been resolved in the crystal structure.

In the crystal structure of the Superoxidized state two conformations are resolved, each with 50% of occupancy. Conformation A has Ψ of 94° and Φ of 218° , while conformation B has Ψ of 177° and Φ of 213° . Basically during MD simulations, within the Forward translocation set-up, the most occupied state is that state which is similar to the B-state captured in the crystal

structure. Also in the crystal structure of the reduced state, two conformations A and B are present, with 46% and 54% of occupancy, respectively. Conformation A has Ψ of 178° and Φ of 41° , while conformation B has Ψ of 168° and Φ of 176° . In this case during the MD simulation a more broad distribution of the occupied states is present, with a total number of states equal to four. Among these, one corresponds to state B occupied in the crystal structure of the reduced state protein, which is the mostly occupied one (54%). It comes out, that Glu16 is acting like a valve in the translocation pathway. In fact, as shown above, in the reduced state Glu16 has higher flexibility, occupying four different positions during MD simulations.

4.2.2.3 The Glu-path

Statistical analysis of hydrogen bond formation was done using 500 snapshots extracted from each 100 ns-long MD trajectory. The probability of H-bond formation (PFH) between two residues was computed by counting the number of events during the MD simulation in which a hydrogen bond was formed. A hydrogen bond was considered formed if the distance between hydrogen bond donor atom and the hydrogen bond acceptor atom is less than 3.5 \AA and the donor-hydrogen-acceptor angle is larger than 120° . Only the six hydrogen bonds indicated in figure 4.4 were considered for the analysis. Analysis of the Glu-path has been performed taking into account the probability of formation of the different hydrogen bonds involved in the proposed mechanism, along with the MD simulations. Results are schematically summarised in figure 4.8.

In the upper part of figure 4.8, the red arrows represent the superoxidized state simulations, while the blue arrows indicate the reduced state simulations. Here, the resulting occupancies of the H-bonds which takes part into the translocation of the proton through the Glu-path are presented. The higher the number of times an H-bond is formed, the higher is the probability a proton is transferred between the residues involved. The scheme in figure 4.8, will be analysed and divided in two parts. From the right to the left of the picture, in the direction of the Forward translocation from PC to the Active Site, we define a first part (Part I) from the proximal cluster to the residue Glu16, and a second part (Part II), starting at the Glu16 and ending up at the active site. Part I, presents lower probabilities of formation for H-bonds. Since PFH values for H-bond between Glu76 to His13 are lower than those for Glu76-Glu16, the transfer of proton has more probability to occur via H-bond interaction between the two glutamate residues. However, in the superoxidized state (red arrows), in case of a proton transfer event through His13, the

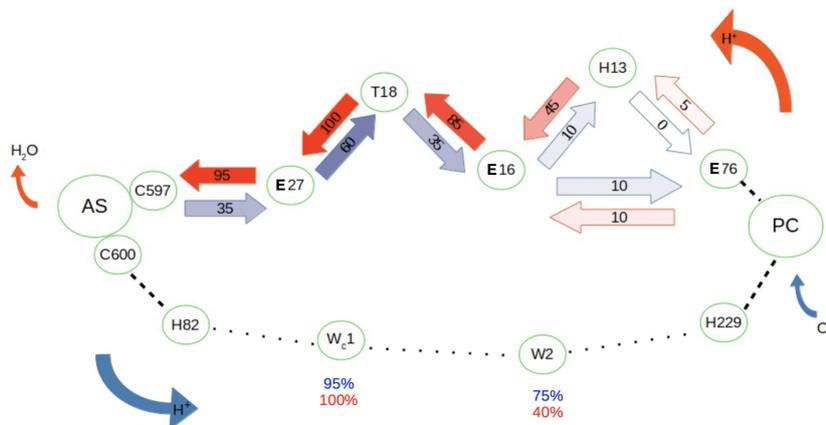


Figure 4.8: Summary of the results on the proposed Proton Transfer Pathways. Upper part: The percentage of the probability of formation of hydrogen bonds (PFH) along Glu-path. Red arrows indicate the *forward translocation* while blue arrows the *backward translocation*. In both cases, the colour intensity is directly proportional to the relative value of PFH. Lower part: The percentage of occupation of the relative cavity, is shown, where red colour corresponds to MD simulations for the *forward translocation*, while blue colour is relative to *backward translocation*.

proton would be transferred to Glu16, while in the reduced state a proton transferred to His13 would be most likely blocked on this latter residue. Part II, instead, presents higher PFH values, which reflects the fact that the hydrogen bond network, formed between Glu16, Thr18, Glu27 and Cys597 is, for both setups, more stable than in the previous state. However, the values of PFH for this second part for the forward translocation have, for all the steps, higher values. The larger difference is noticed for the H-bond between Glu27 and Cys597, which is formed in 95% of the simulation steps for the forward translocation and in only 35% for the backward translocation. This difference remarks the instability of this H-bond, between the reduced and superoxidized setups. This is reflected also in the increased mobility of this part of the protein in the reduced state setup, which causes a spike in the RMSD trend, as shown in section 4.2.2.1. As shown in previous section 4.2.2.2, the Glu16 is acting like a valve. Analysing the PFH values, obtained by the MD simulations, a differentiation between Part I and Part II of the Glu-path is observed and Glu16 clearly shows a gating function. This, in fact, explains the drop of the percentage indicated in the blue arrow between Glu16 and His13 from 45% to 10%, and between

D16 and T18 from 85% to 35%, when comparing the Fort- and Back-translocation setups.

4.2.2.4 The His-path

Moreover, the analysis of the simulation of the residues involved in the His-path, which is another possible translocation pathway connecting active site and proximal cluster (see description in the introductory chapter), has been carried out. It revealed the presence of an additional water which is able to reach an empty cavity where it can assist the translocation of a proton through this pathway. The values indicated in the bottom part of figure 4.8, for the His-path, represent the percentage over the MD simulation of the number of times each water molecule has been observed in the cavity. Blue indicates backward translocation while red indicates forward translocation. W_c1 is the water molecule already present in the crystal structure, while the $W2$ is the position occupied by the water molecules that are able to reach it during the simulations. As expected, values for W_c1 are high, indicating an almost always occupied position and a very high stability of this water molecule, which in fact can be captured during the process of crystallization of the protein. The second position, instead, is transiently occupied by water molecules which are moving toward that cavity escaping from different parts of the environment. This position is detected as unoccupied in the crystal structure since crystallography is able to represent only a static snapshot of the structure itself. Interestingly, we observed that there is a substantial drop in the occupancy of the $W2$ cavity during Forward translocation simulations from 75% to 40%.

4.2.3 Conclusion and Outlook

4.2.3.1 The preferred way of proton translocation and the PCET mechanism

The preferred way of proton translocation in the direction of the active site (red arrows) is taking place through Glu-path (upper part), while the opposite direction of the proton translocation (blue arrows) is more likely to happen through the His-path (bottom part) [see figure 4.8]. When the active site is in the NiC state and oxygen, in the form of a hydroxyl group, is present, there is a major need for protons, to prevent the formation of peroxospecies in the bridging position between Ni and Fe at the active site. The leaving of those species as water, instead, needs to be favoured to preserve the catalytic function of the enzyme in the presence of oxygen. The proximal cluster functions than as provider of protons, through the Glu-path. On the other hand, when the

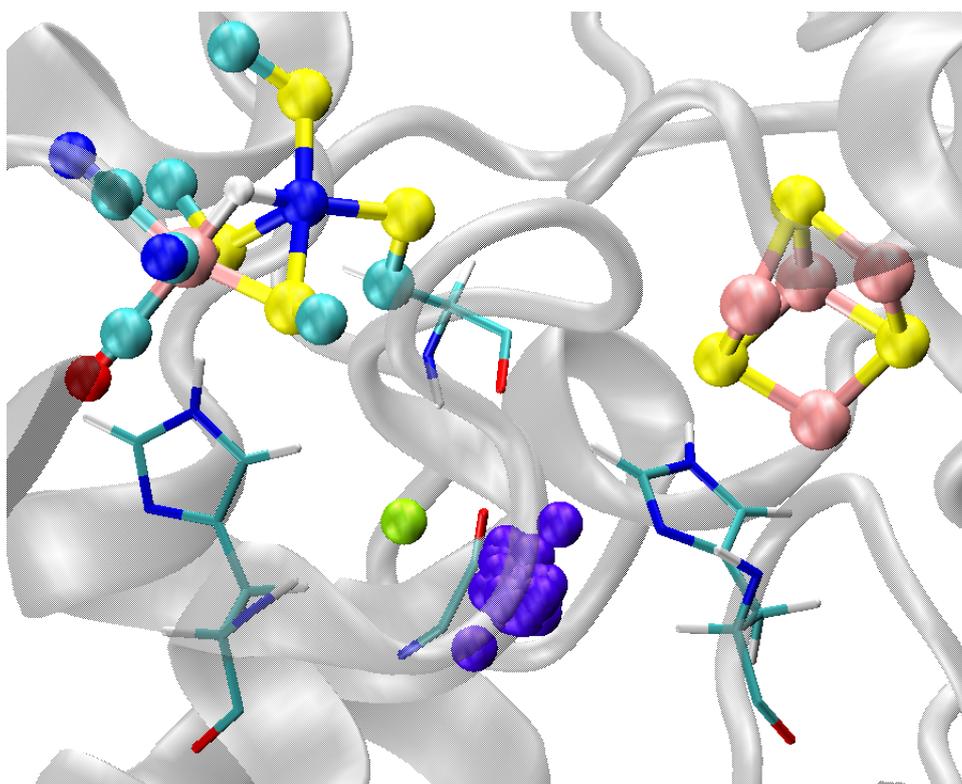


Figure 4.9: The representation of the insertion of water molecules in a free cavity along the MD simulation. The green water is W_c1 in the position occupied in the crystal structure by a water molecule, while violet waters are in the positions occupied by different water molecules along the MD simulations for 50 frames equally distributed along the 100 ns. For clarity, only oxygens of water molecules are represented. The picture is taken from the results of Simulation 3.

protein is in the reduced state an oxygen or a water [14] approaches the PC cluster to attach at Fe_1 , leading to the formation of the superoxidized PC, where an hydroxyl group has been found in the crystal structure. As stated in PhD thesis by Elisabeth Siebert, the OH^- ligand at Fe_1 is preferentially formed from molecular oxygen, which is reduced in a four-electron three-proton process, forming OH^- and H_2O . Hence, in this case, there is a major need for protons in the direction of the proximal cluster and the active site functions as provider of protons, through the His-path. Different mutants have been engineered to explore the role of His229. In particular, when this residue has been substituted by an alanine, three water molecules have been found in the vicinity of the proximal cluster, and at least one of those showed large flexibility in the crystal structure, supporting the hypothesis that H_2O could bind to the PC to form the superoxidized state. Solvent water molecules in the vicinity of the proximal cluster are an alternative source for the hydroxyl ligand, but this binding process is much slower than the formation of the Fe-OH bond from reduced molecular oxygen. In H229A mutant the absence of the His residue breaks the H-bond network, hindering the translocation of a proton via the His-path. This favours the OH^- formation by the dissociation of water into hydroxyl and proton against the reduction of molecular oxygen, confirming that this pathway best performs in the proton translocation in the direction of the proximal cluster. Moreover, as concluded in section 4.1, PC cluster is able to release a proton, only upon two electron oxidation, when it reaches its superoxidized electronic state. In this conditions, the PC is also able to accept the hydroxyl group on Fe_1 . The ability of the proximal cluster to release and uptake protons on needs and to rearrange its geometry upon oxidation and reduction are interconnected. In fact, the stocking of protons is a function of the proximal cluster which is concomitant with the extraction or accommodation of two electrons, resulting in a very efficient PCET mechanism of protection against oxygen attack at the active site.

4.3 MBH mutant at the Active Site side

4.3.1 Introduction and Methods

When MBH performs the catalytic reaction, the active site encounters different electronic states and the bridging position between Ni and Fe is alternatively occupied by various ligands such as hydroxyl (OH^-) and hydride (H^-). During the cycle, protons are uptaken and released and their

transport to the surface is initiated by an arginine residue (Arg530) which is forming hydrogen bond to one of the CN^- ligands to the Fe center. This residue, Arg530 in the wild-type *ReMBH*, is forming a salt bridge with Asp117, as part of a longer proton transfer pathway to the surface. The unpublished crystal structures of a mutant D117S of *ReMBH*, crystalized at different redox potentials, have been resolved in the group of Dr. Patrick Scheerer, at the Institute of Medical Physics and Biophysics in Charite Berlin. Their catalytic function has been experimentally studied by Dr. Stefan Frielingsdorf, in the group of Prof. Dr. Oliver Lenz at TU Berlin, showing interesting characteristics both from the structural and the functional point of view.

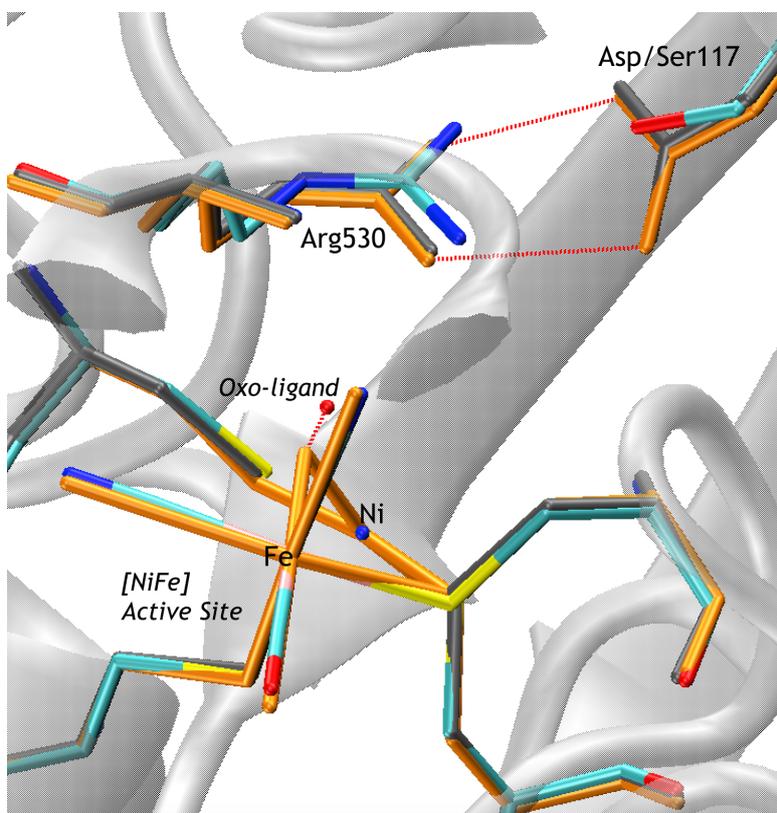


Figure 4.10: NiFe-hydrogenase protein structure comparison between WT (Oxidized (pdb entry: 4IUC), orange and Reduced (pdb entry: 3RGW), grey) and D117S H_2 -reduced mutant (color scheme by atom). The salt bridge between Arg530 and Asp117 in the WT, as well as the translation of the oxo-ligand in the Arg530 direction upon mutation, are shown with red dashed lines.

Here, I will present QM/MM calculations performed on different models made up starting from a partially reduced structure of the D117S-*Re*MBH partially reduced mutant. The goal is to clarify the structure of the active site and its interaction with the environment.

4.3.1.1 The structure of the D117S-MBH in partially reduced state

Despite the reducing conditions due to the exposition to H_2 , the D117S mutant presents an oxygen ligand in the space between Ni and Fe, which, at the same redox conditions, is usually occupied by hydride ligand. The oxygen atom has different occupancies in the samples measured, varying between 50 and 100%, with an average value of 70%. Moreover, the O atom is not bridging between Ni and Fe, but is instead bound to Ni only, with which has a distance of 1.95 Å while it sets at a distance of 2.44 Å to the Fe.

Also an unusual short distance of 2.55 Å between Ni and Fe has been found in this structure. Typically oxidized states (Ni-A and Ni-B) have metal-metal distance around 2.7 and 3.0 Å, whereas the Ni-S species exhibit slightly increased Ni-Fe distances of 2.8 Å or greater. In contrast, for the hydride-containing reduced species (Ni-C and Ni-SR), this distance is strongly reduced to about 2.6 Å.

Interestingly, regarding the first-shell of the environment, Arg530 occupies the same position as in the reduced and oxidized state structure of *Re*MBH (pdb 3RGW, 4IUC respectively). The oxygen atom, interacting with Arg530 through hydrogen bond, is moved up toward the nitrogen atom of arginine side chain. Consequently, the distance between O- $N\omega$ (Arg530) is shortened from 3.09 Å in the oxidized WT state to 2.90 Å in the D117S mutant.

4.3.1.2 The setup of the QM/MM calculation

Different models have been build in order to determine the electronic structure of the active site and the nature of the oxo-ligand found in the crystal structure. In particular, oxidation states 2+ and 3+ have been considered for Ni atom, each with two different possibilities for the oxo ligand, hydroxyl OH^- group or water. For active site structure with Ni^{2+} and water ligand to the Ni atom, two additional models have been considered. One where Cys597 is protonated on the sulfur atom and another one where a water molecule (W_{add}) is added to the structure in an open cavity close in proximity of the oxo-ligand to the Ni. Moreover, an additional model with unusually high 4+ oxidation state for Ni and O_2^- ligand has been used. Starting from

the crystal structure provided for the mutant, standard protonation state at pH 7 for all the residues have been used. Later the protein has been inserted in a TIP3P water box, the position of hydrogen and the orientation of water molecules optimised, with 20000 steps of classical MM energy minimisation, using NAMD [36] with CHARMM22 force field [37], while heavy atoms of the protein were kept fixed. The second step consists in the QM/MM calculation, where the protein in a sphere of waters of 40 Å around the center of the protein is optimised. These calculations are again performed with the ChemShell interface, where the quantum chemical package Turbomole with the molecular mechanics engine DLPOLY are interfaced. The entire system is divided into three layers, which are treated at a different level of theory and accuracy. The QM part consists of (1) the [NiFe] active site; (2) Cys78, Cys600, bridging Ni and Fe, and Cys75, Cys597 coordinating the Ni atom of the cluster; (3) Arg530 and Ser117, which are part of the proton transfer pathway to the surface of interest for this work; (4) Thr553 interacting with cyanide ligand to the Fe atom and contributes to its stabilization and, finally, (5) His82 which interacts with Cys600. This region is described by BS-DFT with BP86 functional [23] and def2-TZVP and basis set for metals and 6-31G* for non-metals. The second region is made of protein residues and water molecules included in a sphere of 17 Å around Ni atom of the active site and is treated at molecular mechanical level, allowed to move during the optimisation. The last layer contains all the atoms not included in the other two and is held fixed.

4.3.2 Results

After optimisation, a comparison between the final structures and the starting crystal structure is made. An overview is presented in the table 4.7. The name of the models refers to the oxidation state of the Ni and the nature of the ligand.

The position of the arginine residue 530 reflects the strength of its interaction with the oxo-ligand. In fact, it can form strong hydrogen bond between O-N ω (Arg530) and the oxo-ligand, and this interaction influences the position of this latter. The last column of table 4.7 describes the change of arginine position after optimisation. There, an equal sign represents the case where Arg530 stays close to the position it assumes in the crystal structure, while Down Shifted (DS) is indicated when this residue is moving toward the active site. Whenever the oxo-ligand is modelled as an hydroxyl group, it tends, upon optimisation, to occupy the bridging position between Ni and Fe. Fe-O bond of the crystal structure (1.95 Å) is very well reproduced in the

	Ni-Fe / Å	Ni-O / Å	Fe-O / Å	Arg530
Crystal Structure D117S - H ₂ reduced	2.55	1.95	2.44	-
Ni(II)-OH	2.98	1.92	2.00	DS
Ni(II)-H ₂ O	2.66	2.20	3.89	=
Ni(II)-H ₂ O-C597H ⁺	2.66	2.14	3.71	=
Ni(II)-H ₂ O-W _{add}	2.62	2.10	3.48	=
Ni(III)-OH	2.85	1.95	2.04	DS
Ni(III)-H ₂ O	2.90	2.11	2.19	=
Ni(IV)-O ₂ ⁻	2.78	1.84	1.93	DS

Table 4.7: Relevant structural parameters of the NiFe active site from experiments and calculations on MBH-D117S mutant.

Ni(II)/Ni(III)-OH models (1.92 and 1.95 Å, respectively). However, the Ni-O distance is reduced of ~ 0.4 Å and the oxo-ligand is occupying the bridging position, almost equidistant to Ni and Fe.

The short Ni-Fe distance is reproduced when Ni is in the oxidation state 2+ and H₂O is present. This electronic state of the active site corresponds to a Ni-S state. Moreover, in the above-mentioned cases, Arg530 does not move down toward the bridging position and stays instead close to the position it assumes in the crystal structure. This happens because the water ligand moves outside the cavity which is located between Ni and Fe atoms, toward an empty cavity which is on the right side of the Ni, opposite to Fe. Therefore, the H-bond between Arg530 and oxo-ligand is broken.

In an attempt to keep the H₂O ligand in the position it occupies in the crystal, two more models were generated. One where the Cys597 sulfur is protonated, as in the Ni-L state [61], and one where an additional water is added *in silico* into a free cavity on the right side of Ni, which is occupied by water ligand upon optimisation of model Ni(II)-H₂O. Results are pointing toward an improvement in the reproduction of the structure of the crystal, i.e. a decreasing of Fe-O distances of Ni(II)-H₂O, Ni(II)-H₂O-C597H⁺ and Ni(II)-H₂O-W_{add} is observed.

The last model with Ni(IV) oxidation state and oxide anion as a ligand has been modeled.

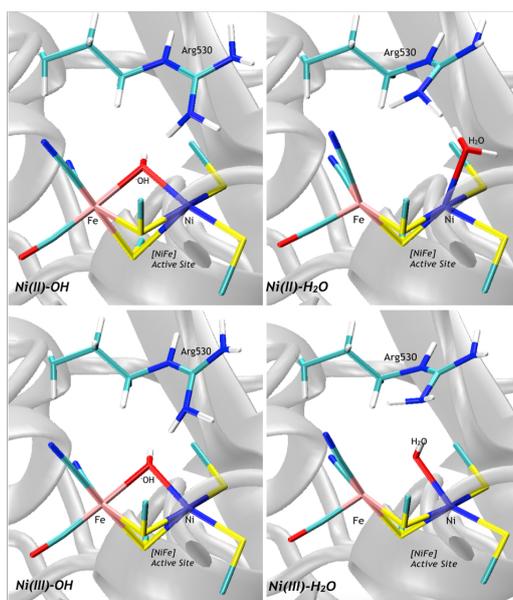


Figure 4.11: QM/MM optimised structures of four models Ni(II)-OH, Ni(II)-H₂O, Ni(III)-OH and Ni(III)-H₂O.

This oxidation state leads to an increment of the Ni-Fe distance of 0.23 Å with respect to the crystal structure. The dianion O²⁻ is tightly bound to the Ni and Fe, to which it is set at very short distance because of the strong ionic interaction.

4.3.3 Discussion and Outlook

From these QM/MM calculations it can be concluded that: 1) there must be no H-bond interaction of the Arg530 with the oxo-ligand, in order to keep the position it occupies in the crystal structure of the mutant. This is achieved when the ligand is modelled as a water and tends to leave the cavity it occupies in the crystal to set into a close by free cavity. In fact, when oxo-ligand is in a position favourable to form H-bond with Arg530 this interaction is pulling down the residue toward [NiFe] site. This happens because the ligand is adjusting its position into the bridging-cavity between Ni and Fe with shorter Ni-O and Fe-O distances; 2) we demonstrated that short distance between Ni and Fe is better reproduced by Ni(II) oxidation state, in agreement with preliminary EPR studies which showed no signal that can be assigned to the active site. A possible interpretation is that there is a discrepancy between the structure captured by

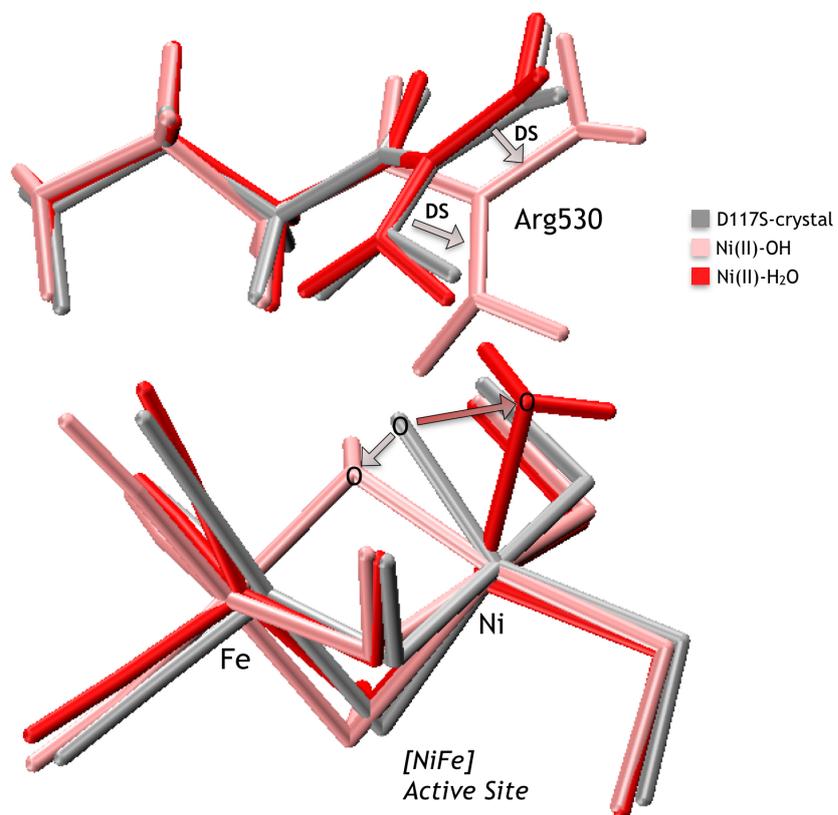


Figure 4.12: optimised structure comparison of two models Ni(II)-OH (pink) and Ni(II)-H₂O (red), against starting crystal structure (silver). The displacements of the water ligand from its crystal structure position and the downshift of Arg530 in Ni(II)-OH are shown.

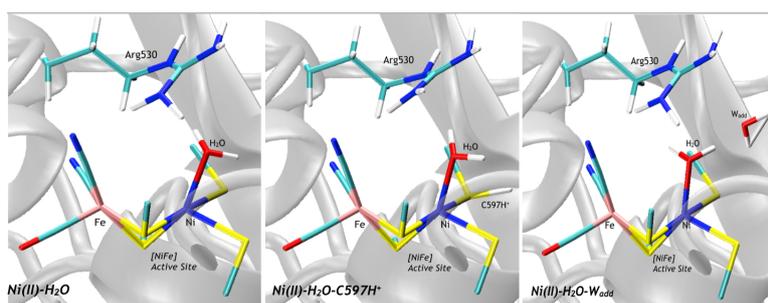


Figure 4.13: QM/MM optimised structures of three models Ni(II)-H₂O, Ni(II)-H₂O-C597H⁺ and Ni(II)-H₂O-W_{add}.

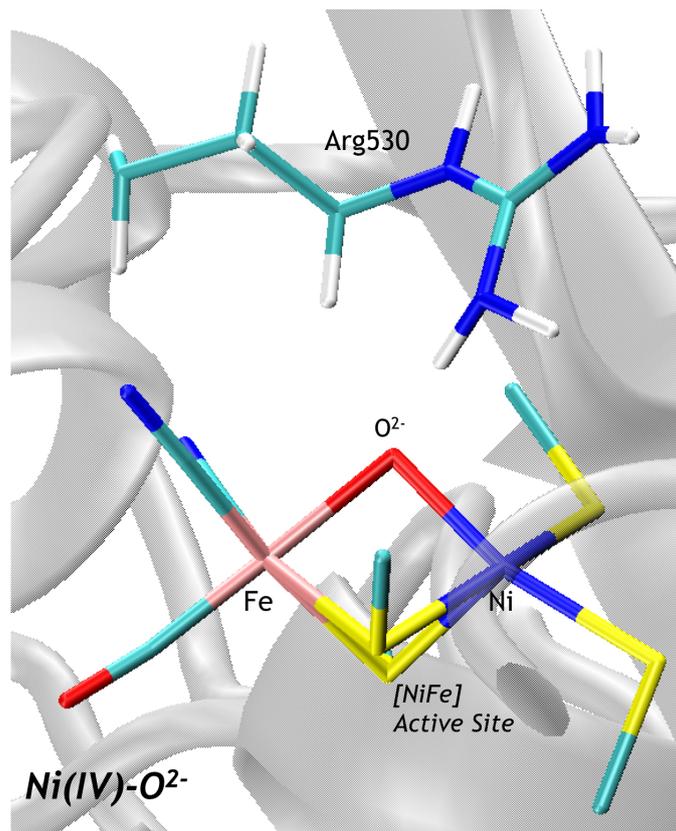


Figure 4.14: QM/MM optimised structure of model Ni(IV)-O²⁻.

means of crystallographic technique and the structure in aqueous solution, where the ligand is free to move and to adjust its position. Alternatively, there could be an element not detectable within the resolution of the crystal structure (i.e. proton), which is blocking the oxo-ligand in its position. However, in this case one could speculate that, whenever an oxo-ligand is present in the proximity of the Arg530, which has lost its strong salt bridge interaction with the mutated residue 117, the arginine would be downshifted in its position because of its H-bond interaction with the oxo-ligand. In the future, different protonation states for Arg530 should be tested, since it is known that in close vicinity of a metal cluster the pK_a of amino acids can vary by many orders of magnitude.

Chapter 5

Conclusions

In the present work, some aspects of two different metalloenzymes have been analysed. Both systems have important applications in the field of catalysis. First, the oxygen tolerant [NiFe]-hydrogenase from *Ralstonia eutropha* has been studied. In general, [NiFe]-hydrogenases are well-studied systems. Their metallic active site functions with a very well established mechanism in the hydrogen splitting enzymatic reaction. However, in the particular case of oxygen tolerant systems, there exist still many open questions. Some of these have been addressed in this Thesis:

- 1) The iron sulfur proximal cluster structure upon one electron oxidation, and its interaction with the active site, through a complex PCET process, have been studied. In section 4.1, QM/MM optimisations were performed on different models of the proximal cluster to study the dependence of the opening-closing mechanism on their electronic state and their electrostatic environment. In particular, the structure of the so-called oxidized state of the proximal cluster in the oxidized MBH protein has been clarified. In fact, we have been able to demonstrate that the crystal structure of the oxidized state (PDB code: 4IUD) is a mixture of two states: one in which the cluster is in the closed conformation and the other with a residual oxygen remaining from the superoxidized state partial reduction.
- 2) In the attempt to reconstruct the electron and proton movements, happening at the proximal cluster, we could demonstrate that one electron oxidation is not able to provide enough energy to the rearrangement of the iron sulfur cluster, in a conformation which resembles the superoxidized open structure. This open-cluster conformation is possible only upon translocation of a proton from the nitrogen of Cys20 to a glutamate residue

(E76) in the vicinity. This event strengthens the electrostatic interaction between the nitrogen and the Fe4 leading to the formation of the bond Fe4-N(Cys20). The opening movement of the cluster is triggered by the deprotonation of N(Cys20), but the proton transfer is not spontaneously taking place with one electron oxidation. The first acceptor of the proton is identified as Glu76, which slightly changes its conformation after receiving the proton. Consequently, its hydrogen bond interaction with (N)Cys20 is broken and the proton can be further transferred. Thus, we confirm that a second electron is needed in order to allow the proton to be translocated and the proximal cluster to set in its open-conformation. 3) Based on these results, we were interested in studying the fate of the proton after the first event of proton transfer. A PCET mechanism has been thus suggested, connecting the iron sulfur proximal cluster and the nickel iron active site. This process is proposed to be, together with others, responsible for the protection of the active site against oxygen. A proton transfer pathway, which starts with the transfer of H^+ from Cys20 to Glu76, has been studied by means of MD calculations. This Glu-path is the preferred way for proton translocation in the direction of the active site (Forward Translocation). In fact, in order to avoid inhibition of the catalytic activity, the active site needs protons which are provided, together with electrons, by the unique iron sulfur proximal cluster. 4) The occupation of a cavity in the vicinity of the residue His229 by water, observed during the MD simulations, opened up the chance to explain the function of a second proton transfer pathway. In fact, when the protein is reduced, an oxygen or a water can approach the PC cluster and attach at the Fe_1 site, forming the superoxidized geometry of the iron sulfur cluster, where an hydroxyl group is bound to the iron atom. This ligand is preferentially formed from molecular oxygen thus protons and electrons are needed at the PC site for oxygen reduction. We demonstrate in this work that protons are provided by the active site itself. Protons are in fact able to be translocated via the His-path, from the active site to the proximal cluster. According to experimental work, when His229 is mutated by an alanine, water approach to the proximal cluster, to form the hydroxyl ligand to Fe_1 , is favoured. This observation confirms the fact that the proton transfer pathway, which is passing through the histidine 229, plays a fundamental role in providing protons to the proximal cluster. This latter, in this conditions which is also able to trap oxygen molecules. Our interest also spread in the study of an interesting mutant of *ReMBH*. The residue D117, which is in the vicinity of the NiFe active site and is part of a proton transfer chain to the surface, was mutated with a serine, in order to cut the hydrogen bond network. Upon mutation,

interesting features were found at the active site. Analysis of the partially reduced state, where quite short Ni-Fe distance is in contrast, with respect to the WT structures, to the presence of an oxo-species at the Ni site. QM/MM calculations performed on different models of the active site region, with various Ni oxidation state and oxo-ligand types, gave indication for a Ni²⁺ oxidation state, similarly to the Ni_a-SI species. Moreover, we recognised an important role for the hydrogen bond interaction between the oxo-ligand and Arg530, which acts on the position of this latter. In fact, when the oxo-ligand is modelled as a water in the Ni(II)-Fe(II) electronic state model, the water tends to move away from the cavity it occupies in the crystal structure, breaking the H-bond interaction with the Arg530 side chain. However, the identification of the factor which, in the crystalline state brings this oxo-ligand to occupy the cavity between the two transition metals sites, albeit bound only to Ni atom, remains unclear.

The knowledge of the catalytic mechanism of another kind of enzymes, CODH, attracted our interest, because of its increasing importance in technological applications. The interplay between different groups within UniCat project allowed us to reach significant results in order to clarify the interaction of this enzyme with two inhibitors, namely CN and NCO, which are structural analogues of CO and CO₂, respectively. In the study of *Ch*CODH-CN [3], by means of QM/MM calculations, contributed in clarifying the structure of the inhibited state of the active site. We could uniquely assign the C≡N stretching mode of the cyanide bound complex, confirming that a single cyanide is bound to Ni and that this binding occurs upon release of the hydroxyl ligand from the Fe₁. Moreover, we identified the side chain of Lys563 as the proton donor to the hydroxyl ligand, which therefore is able to leave as a water molecule. Nevertheless, the refinement of crystallographic data performed by means of QM/MM calculations, pointed to a protonated Lys563, which is contrast to the reproduction of IR spectroscopic data. We suggest this discrepancy to be due to re-protonation event of Lys563 via a water radical, generated upon X-ray irradiation. We also refined the structure via QM/MM calculation on a recently published structure of the *Ch*CODH-NCO. We were focussed in the assignment of the vibrational modes specifically located at the NCO ligand to Ni atom. Three distinct modes have been assigned to it, and this gave the chance to rule out the presence of NCO bands in the vicinity of the C≡N stretching mode, previously assigned to Ni-bound cyanide. This finding allowed a specific interpretation of the IR spectra and, consequently, demonstrating that CODH is able to reduce cyanite to cyanide [4].

Acknowledgments

Firstly, I would like to express my sincere gratitude to my advisor, Prof. Dr. Maria Andrea Mroginski, for the continuous support of my PhD study and related research, for her patience, motivation, and support. Her guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my PhD study. Besides my advisor, I would like to thank the rest of my thesis committee: Prof. Dr. Nediljko Budisa and Prof. Dr. Claudio Greco, for having accepted to read, comment and their precious time. My sincere thanks also goes to Prof. Dr. Peter Hildebrandt for his support. Without it would not have been possible to conduct this research. Moreover, I would like to thank Dr. Alexandre Ciaccafava, and Dr. Ingo Zebger for their help, continuous support and to having worked hard with me on a substantial part of my PhD project. I thank people in my group for the stimulating discussions, and for all the fun we have had in the last years; Dr. Eleonora Ussano and Dr. Juan Sebastian Toterò, for their help and support during different stages of my scientific career. This work has been supported by the Deutsche Forschungsgemeinschaft (DFG) within the Cluster of Excellence "UJUNICAT" project. The computational power has been provided by the Cluster of the Institute of Mathematic at the Technical University in Berlin and by the North-German Supercomputing Alliance (HLRN).

Last but not least, I would like to thank my families: my parents, my grandmother and my sister's family for supporting me throughout writing this thesis and my life in general; my big-family in Italy, which is always in my heart; my new family here in Berlin, all the people around Le Balene Possono Volare project, Wale Cafe, Yanez Magazine. Above all Francesca, Elena and Mattia, for their help to go over a stressful period with talks, laughs and love in many incredible forms. This thesis is dedicated to my beloved one, with four legs. She never stops to teach me about the meaning of love and being loved, without nothing in exchange.

Bibliography

- [1] Aaron M. Appel, John E. Bercaw, Andrew B. Bocarsly, Holger Dobbek, Daniel L. Dubois, Michel Dupuis, James G. Ferry, Etsuko Fujita, Russ Hille, Paul J A Kenis, Cheryl A. Kerfeld, Robert H. Morris, Charles H F Peden, Archie R. Portis, Stephen W. Ragsdale, Thomas B. Rauchfuss, Joost N H Reek, Lance C. Seefeldt, Rudolf K. Thauer, and Grover L. Waldrop. Frontiers, opportunities, and challenges in biochemical and chemical catalysis of CO₂ fixation. *Chemical Reviews*, 113(8):6621–6658, 2013. ISSN 00092665. doi: 10.1021/cr300463y.
- [2] V. Svetlitchnyi, C. Peschel, G. Acker, and O. Meyer. Two membrane-associated NiFeS-carbon monoxide dehydrogenases from the anaerobic carbon-monoxide-utilizing eubacterium *Carboxydotherrmus hydrogenoformans*. *Journal of Bacteriology*, 183(17):5134–5144, 2001. ISSN 00219193. doi: 10.1128/JB.183.17.5134-5144.2001.
- [3] Alexandre Ciaccafava, Daria Tombolelli, Lilith Domnik, Jochen Fessler, Jae-Hun Jeoung, Holger Dobbek, Maria Andrea Mroginski, Ingo Zebger, and Peter Hildebrandt. When the inhibitor tells more than the substrate: the cyanide-bound state of a carbon monoxide dehydrogenase. *Chem. Sci.*, 2016. ISSN 2041-6520. doi: 10.1039/C5SC04554A.
- [4] Alexandre Ciaccafava, Daria Tombolelli, Lilith Domnik, Jae-Hun Jeoung, Holger Dobbek, Maria-Andrea Mroginski, Ingo Zebger, and Peter Hildebrandt. Carbon Monoxide Dehydrogenase Reduces Cyanate to Cyanide. *Angewandte Chemie (International ed. in English)*, 2017. doi: 10.1002/anie.201703225.
- [5] Mehmet Can, Fraser A. Armstrong, and Stephen W. Ragsdale. Structure, Function, and

- Mechanism of the Nickel Metalloenzymes, CO Dehydrogenase, and Acetyl-CoA Synthase. *Chemical Reviews*, 114(8):4149–4174, 2014. ISSN 0009-2665. doi: 10.1021/cr400461p.
- [6] Jochen Fesseler, Jae Hun Jeoung, and Holger Dobbek. How the [NiFe₄S₄] Cluster of CO Dehydrogenase Activates CO₂ and NCO⁻. *Angewandte Chemie - International Edition*, 54(29):8560–8564, 2015. ISSN 15213773. doi: 10.1002/anie.201501778.
- [7] Weimin Gong, Bing Hao, Zhiyi Wei, Donald J Ferguson, Thomas Tallant, Joseph a Krzycki, and Michael K Chan. Structure of the alpha2epsilon2 Ni-dependent CO dehydrogenase component of the Methanosarcina barkeri acetyl-CoA decarbonylase/synthase complex. *Proceedings of the National Academy of Sciences of the United States of America*, 105:9558–9563, 2008. ISSN 0027-8424. doi: 10.1073/pnas.0800415105.
- [8] Yan Kung, Tzanko I. Doukov, Javier Seravalli, Stephen W. Ragsdale, and Catherine L. Drennan. Crystallographic snapshots of cyanide- and water-bound C-clusters from bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase. *Biochemistry*, 48(31):7432–7440, 2009. ISSN 00062960. doi: 10.1021/bi900574h.
- [9] Patricia Amara, Jean-Marie Mouesca, Anne Volbeda, and Juan C Fontecilla-Camps. Carbon monoxide dehydrogenase reaction mechanism: a likely case of abnormal CO₂ insertion to a Ni-H(-) bond. *Inorganic chemistry*, 50(5):1868–78, mar 2011. ISSN 1520-510X. doi: 10.1021/ic102304m.
- [10] Jae Hun Jeoung and Holger Dobbek. Structural basis of cyanide inhibition of Ni, Fe-containing carbon monoxide dehydrogenase. *Journal of the American Chemical Society*, 131(29):9922–9923, 2009. ISSN 00027863. doi: 10.1021/ja9046476.
- [11] Wolfgang Lubitz, Hideaki Ogata, and Edward Reijerse. Hydrogenases. *Chemical Reviews*, 2014.
- [12] J.C. Fontecilla-Camps, A. Volbeda, C. Cavazza, and Y. Nicolet. Structure/function relationships of [NiFe]- and [FeFe]-hydrogenases. *Chemical Reviews*, 107(10), 2007. ISSN 0009-2665, 1520-6890. doi: 10.1021/cr050195z.
- [13] Jean-Marie Mouesca, Juan C Fontecilla-Camps, and Patricia Amara. The structural plasticity of the proximal [4Fe3S] cluster is responsible for the O₂ tolerance of membrane-bound

- [NiFe] hydrogenases. *Angewandte Chemie (International ed. in English)*, 52(7):2002–6, feb 2013. ISSN 1521-3773. doi: 10.1002/anie.201209063.
- [14] Stefan Frielingsdorf, Johannes Fritsch, Andrea Schmidt, Mathias Hammer, Julia Löwenstein, Elisabeth Siebert, Vladimir Pelmeshnikov, Tina Jaenicke, Jacqueline Kalms, Yvonne Rippers, Friedhelm Lenzian, Ingo Zebger, Christian Teutloff, Martin Kaupp, Robert Bittl, Peter Hildebrandt, Bärbel Friedrich, Oliver Lenz, and Patrick Scheerer. Reversible [4Fe-3S] cluster morphing in an O₂-tolerant [NiFe] hydrogenase. *Nature chemical biology*, 10(MAY): 378–85, 2014. ISSN 1552-4469. doi: 10.1038/nchembio.1500.
- [15] Anne Volbeda, Lydie Martin, Elodie Barbier, Oscar Gutiérrez-Sanz, Antonio L. De Lacey, Pierre Pol Liebgott, Sébastien Dementin, Marc Rousset, and Juan C. Fontecilla-Camps. Crystallographic studies of [NiFe]-hydrogenase mutants: Towards consensus structures for the elusive unready oxidized states. *Journal of Biological Inorganic Chemistry*, 20(1):11–22, 2015. ISSN 14321327. doi: 10.1007/s00775-014-1203-9.
- [16] Raffaella Breglia, Manuel Antonio Ruiz-Rodriguez, Alessandro Vitriolo, Rubén Francisco González-Laredo, Luca De Gioia, Claudio Greco, and Maurizio Bruschi. Theoretical insights into [NiFe]-hydrogenases oxidation resulting in a slowly reactivating inactive state. *Journal of Biological Inorganic Chemistry*, 22(1):137–151, 2017. ISSN 14321327. doi: 10.1007/s00775-016-1416-1.
- [17] Anne Volbeda, Lydie Martin, Christine Cavazza, Michaël Matho, Bart W. Faber, Winfried Roseboom, Simon P.J. Albracht, Elsa Garcin, Marc Rousset, and Juan C. Fontecilla-Camps. Structural differences between the ready and unready oxidized states of [NiFe] hydrogenases. *Journal of Biological Inorganic Chemistry*, 10(3):239–249, 2005. ISSN 09498257. doi: 10.1007/s00775-005-0632-x.
- [18] Hideaki Ogata, Shun Hirota, Asuka Nakahara, Hirofumi Komori, Naoki Shibata, Tatsuhisa Kato, Kenji Kano, and Yoshiki Higuchi. Activation process of [NiFe] hydrogenase elucidated by high-resolution X-ray analyses: Conversion of the ready to the unready state. *Structure*, 13(11):1635–1642, 2005. ISSN 09692126. doi: 10.1016/j.str.2005.07.018.
- [19] Yasuhito Shomura, Ki-seok Yoon, Hirofumi Nishihara, and Yoshiki Higuchi. Structural basis

- for a [4Fe-3S] cluster in the oxygen-tolerant membrane-bound [NiFe]-hydrogenase. *Nature*, 479(7372):253–256, 2011. ISSN 0028-0836. doi: 10.1038/nature10504.
- [20] Wolfram Koch and Max C. Holthausen. *A Chemist’s Guide to Density Functional Theory*, volume 3. 2001. ISBN 3527303723. doi: 10.1002/3527600043. URL <http://doi.wiley.com/10.1002/3527600043>.
- [21] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Physical Review B*, 136(3B):864 – 871, 1964. ISSN 01631829. doi: 10.1103/PhysRevB.7.1912.
- [22] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Physical Review A*, 140(4A):1133 – 1138, 1965. ISSN 0031899X. doi: 10.1103/PhysRev.140.A1133.
- [23] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical Review A*, 38(6):3098–3100, 1988. ISSN 10502947. doi: 10.1103/PhysRevA.38.3098.
- [24] J. P. Perdew. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Physical Review B*, 33(June):8822–8824, 1986.
- [25] L. Noodleman, C. Y. Peng, D. A. Case, and J. M. Mouesca. Orbital interactions, electron delocalization and spin coupling in iron-sulfur clusters. *Coordination Chemistry Reviews*, 144(C):199–244, 1995. ISSN 00108545. doi: 10.1016/0010-8545(95)07011-L.
- [26] Louis Noodleman. Valence bond description of antiferromagnetic coupling in transition metal dimers. *The Journal of Chemical Physics*, 74(10):5737, 1981. doi: 10.1063/1.440939.
- [27] Louis Noodleman and Ernest R. Davidson. Ligand spin polarization and antiferromagnetic coupling in transition metal dimers. *Chemical Physics*, 109(1):131–143, 1986. ISSN 03010104. doi: 10.1016/0301-0104(86)80192-6.
- [28] Mitchel Weissbluth. *Atoms and molecules*. Academic Press, 1978.
- [29] Bernard R. Brooks, Robert E. Bruccoleri, Barry D. Olafson, David J. States, S. Swaminathan, and Martin Karplus. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry*, 4(2):187–217, 1983. ISSN 1096987X. doi: 10.1002/jcc.540040211.

- [30] Wendy D. Cornell, Piotr Cieplak, Christopher I. Bayly, Ian R. Gould, Kenneth M. Merz, David M. Ferguson, David C. Spellmeyer, Thomas Fox, James W. Caldwell, and Peter A. Kollman. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society*, 117(19):5179–5197, 1995. ISSN 00027863. doi: 10.1021/ja00124a002.
- [31] Walter R. P. Scott, Philippe H. Hünenberger, Ilario G. Tironi, Alan E. Mark, Salomon R. Billeter, Jens Fennen, Andrew E. Torda, Thomas Huber, Peter Krüger, and Wilfred F. van Gunsteren. The GROMOS Biomolecular Simulation Program Package. *The Journal of Physical Chemistry A*, 103(19):3596–3607, 1999. ISSN 1089-5639. doi: 10.1021/jp984217f.
- [32] William C. Swope, Hans C. Andersen, Peter H. Berens, and Kent R. Wilson. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *The Journal of Chemical Physics*, 76(1):637–649, 1982. ISSN 0021-9606. doi: 10.1063/1.442716.
- [33] R. W. Hockney. The potential calculation and some applications. *Methods in Computational Physics*, 9:135–211, 1970.
- [34] Hans Martin Senn and Walter Thiel. QM/MM methods for biomolecular systems. *Angewandte Chemie - International Edition*, 48(7):1198–1229, 2009. ISSN 14337851. doi: 10.1002/anie.200802019.
- [35] Eun Jin Kim, Jian Feng, Matthew R. Bramlett, and Paul A. Lindahl. Evidence for a Proton Transfer Network and a Required Persulfide-Bond-Forming Cysteine Residue in Ni-Containing Carbon Monoxide Dehydrogenases. *Biochemistry*, 43(19):5728–5734, 2004. ISSN 00062960. doi: 10.1021/bi036062u.
- [36] James C. Phillips, Rosemary Braun, Wei Wang, James Gumbart, Emad Tajkhorshid, Elizabeth Villa, Christophe Chipot, Robert D. Skeel, Laxmikant Kalé, and Klaus Schulten. Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, 26(16):1781–1802, 2005. ISSN 01928651. doi: 10.1002/jcc.20289.
- [37] A D Mackerell, D Bashford, M Bellott, R L Dunbrack, J D Evanseck, M J Field, S Fischer, J Gao, H Guo, S Ha, 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 Wio, D Yin, and M Karplus. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *Journal of Physical Chemistry B*, 5647 (97):3586–3616, 1998.
- [38] A.K. K Rappé, C.J. J Casewit, K.S. S Colwell, W.A. Goddard III, and W.M. Skiff. UFF, a Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations. *J. Am. Chem. Soc.*, 114(25):10024–10035, 1992. ISSN 0002-7863. doi: 10.1021/ja00051a040.
- [39] Paul Sherwood, Alex H. De Vries, Martyn F. Guest, Georg Schreckenbach, C. Richard A. Catlow, Samuel A. French, Alexey A. Sokol, Stefan T. Bromley, Walter Thiel, Alex J. Turner, Salomon Billeter, Frank Terstegen, Stephan Thiel, John Kendrick, Stephen C. Rogers, John Casci, Mike Watson, Frank King, Elly Karlsen, Merethe Sjøvoll, Adil Fahmi, Ansgar Schäfer, and Christian Lennartz. QUASI: A general purpose implementation of the QM/MM approach and its application to problems in catalysis. *Journal of Molecular Structure: THEOCHEM*, 632:1–28, 2003. ISSN 01661280. doi: 10.1016/s0166-1280(03)00285-9.
- [40] Lian Yu, Claudio Greco, Maurizio Bruschi, Ulf Ryde, Luca De Gioia, and Markus Reiher. Targeting Intermediates of [FeFe] -Hydrogenase by CO and CN Vibrational Signatures. *Inorganic Chemistry*, 50:3888–3900, 2011.
- [41] Jacopo Tomasi, Benedetta Mennucci, and Roberto Cammi. Quantum Mechanical Continuum Solvation Models. *Chemical Reviews*, 105(8):2999–3094, 2005.
- [42] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin,

- K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, \tilde{A} Ū. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox. Gaussian 09 (revision a.1). Gaussian Inc. Wallingford CT 2009.
- [43] Victoria J. DeRose, Joshua Telser, Mark E. Anderson, Paul A. Lindahl, and Brian M. Hoffman. A multinuclear ENDOR study of the C-cluster in CO dehydrogenase from *Clostridium thermoaceticum*: Evidence for H(x)O and histidine coordination to the [Fe₄S₄] center. *Journal of the American Chemical Society*, 120(34):8767–8776, 1998. ISSN 00027863. doi: 10.1021/ja9731480.
- [44] Zhengguo Hu, Nathan J. Spangler, Mark E. Anderson, Jinqiang Xia, Paul W. Ludden, Paul A. Lindahl, and Eckard Münck. Nature of the C-Cluster in Ni-Containing Carbon Monoxide Dehydrogenases. *Journal of the American Chemical Society*, 118(4):830–845, 1996. ISSN 0002-7863. doi: 10.1021/ja9528386.
- [45] Yan Kung and Catherine L. Drennan. A role for nickel-iron cofactors in biological carbon monoxide and carbon dioxide utilization. *Current Opinion in Chemical Biology*, 15(2):276–283, 2011. ISSN 13675931. doi: 10.1016/j.cbpa.2010.11.005.
- [46] G A Poskrebyshev, P Neta, and R E Huie. Temperature Dependence of the Acid Dissociation Constant of the Hydroxyl Radical. *Journal of Physical Chemistry A*, 106:11488–11491, 2002.
- [47] Markus W. Ribbe. Insights into the Mechanism of Carbon Monoxide Dehydrogenase at Atomic Resolution. *Angewandte Chemie - International Edition*, 54(29):8337–8339, 2015. ISSN 15213773. doi: 10.1002/anie.201503979.
- [48] Jae Hun Jeoung and Holger Dobbek. n-Butyl isocyanide oxidation at the [NiFe 4S 4OH x] cluster of CO dehydrogenase. *Journal of Biological Inorganic Chemistry*, 17(2):167–173, 2012. ISSN 09498257. doi: 10.1007/s00775-011-0839-y.
- [49] Madeline E Rasche and Lance C Seefeldt. Reduction of Thiocyanate , Cyanate , and Carbon Disulfide by Nitrogenase : Kinetic Characterization and EPR Spectroscopic Analysis \hat{A} Ū. *Journal of Biological Chemistry*, 296(2):8574–8585, 1997.
- [50] Johannes Fritsch, Patrick Scheerer, Stefan Frielingsdorf, Sebastian Kroschinsky, Bärbel Friedrich, Oliver Lenz, and Christian M T Spahn. The crystal structure of an oxygen-

- tolerant hydrogenase uncovers a novel iron-sulphur centre. *Nature*, 479(7372):249–52, nov 2011. ISSN 1476-4687. doi: 10.1038/nature10505.
- [51] Sebastian Metz, Johannes Kästner, Alexey A. Sokol, Thomas W. Keal, and Paul Sherwood. ChemShell-a modular software package for QM/MM simulations. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 4(2):101–110, 2014. ISSN 17590876. doi: 10.1002/wcms.1163.
- [52] Vladimir Pelmeshnikov and Martin Kaupp. Redox-Dependent Structural Transformations of the [4Fe-3S] Proximal Cluster in O₂ Tolerant Membrane-Bound [NiFe]- Hydrogenase: A DFT Study. *Journal of the American Chemical Society*, 2013.
- [53] Anne Volbeda, Patricia Amara, Claudine Darnault, Jean-marie Mouesca, Alison Parkin, Maxie M Roessler, Fraser A Armstrong, and Juan C Fontecilla-camps. X-ray crystallographic and computational studies of the O₂-tolerant [NiFe]-hydrogenase 1 from *Escherichia coli*. *Proceedings of the National Academy of Sciences*, pages 1–6, 2012. doi: 10.1073/pnas.1119806109.
- [54] It Todorov, W Smith, and Uk Cheshire. The DL POLY 2 user manual, 2011.
- [55] TURBOMOLE V6.2 2010, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from <http://www.turbomole.com>, 2010.
- [56] Salomon R Billeter, J Turner, Walter Thiel, and C H Zu. Linear scaling geometry optimisation and transition state search in hybrid delocalised internal coordinates $\hat{\text{A}}\text{d}$. *Physical Chemistry Chemical Physics*, 2:2177–2186, 2000.
- [57] Noam Agmon. The Grotthuss mechanism. *Chemical Physics Letters*, 244(5-6):456–462, 1995. ISSN 00092614. doi: 10.1016/0009-2614(95)00905-J.
- [58] Samuel Cukierman. Et tu, Grotthuss! and other unfinished stories. *Biochimica et Biophysica Acta - Bioenergetics*, 1757(8):876–885, 2006. ISSN 00052728. doi: 10.1016/j.bbabi.2005.12.001.
- [59] Bojana Ginovska-pangovska, Ming-hsun Ho, John C Linehan, Yuhui Cheng, Michel Dupuis, Simone Raagei, and Wendy J Shaw. Molecular dynamics study of the proposed proton

- transport pathways in [FeFe]-hydrogenase. *Biochimica et biophysica acta*, 1837(1):131–138, 2014. ISSN 0005-2728. doi: 10.1016/j.bbabi.2013.08.004.
- [60] Philip A. Ash, Ricardo Hidalgo, and Kylie A. Vincent. Proton Transfer in the Catalytic Cycle of [NiFe] Hydrogenases: Insight from Vibrational Spectroscopy. *ACS Catalysis*, 7(4): 2471–2485, 2017. ISSN 21555435. doi: 10.1021/acscatal.6b03182.
- [61] Elisabeth Siebert, Marius Horch, Yvonne Rippers, Johannes Fritsch, Stefan Frielingsdorf, Oliver Lenz, Francisco Velazquez Escobar, Friedrich Siebert, Lars Paasche, Uwe Kuhlmann, Friedhelm Lenzian, Maria-andrea Mroginski, Ingo Zebger, and Peter Hildebrandt. Resonance Raman Spectroscopy as a Tool to Monitor the Active Site of Hydrogenases**. *Angewandte Chemie (International ed. in English)*, pages 5162–5165, 2013. doi: 10.1002/anie.201209732.
- [62] Mario Kampa, Maria Eirini Pandelia, Wolfgang Lubitz, Maurice Van Gastel, and Frank Neese. A metal-metal bond in the light-induced state of [nife] hydrogenases with relevance to hydrogen evolution. *Journal of the American Chemical Society*, 135(10):3915–3925, 2013. ISSN 00027863. doi: 10.1021/ja3115899.
- [63] Geng Dong and Ulf Ryde. Protonation states of intermediates in the reaction mechanism of [NiFe] hydrogenase studied by computational methods. *Journal of Biological Inorganic Chemistry*, 21(3):383–394, 2016. ISSN 14321327. doi: 10.1007/s00775-016-1348-9.