SCUOLA DI DOTTORATO
UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

Dipartimento di Scienze dell'Ambiente e della Terra
Dottorato di Ricerca in Scienze - Ciclo XXIX
Curriculum in Scienze Chimiche

QUANTUM-MECHANICAL STUDY OF THE STEREOELECTRONIC AND CATALYTIC PROPERTIES OF METALLO-ENZYMES INVOLVED IN REACTIONS OF ENVIRONMENTAL AND TECHNOLOGICAL RELEVANCE

Breglia Raffaella

Matricola 720068

Tutor: Prof. Maurizio Bruschi

Coordinatore: Prof.ssa Maria Luce Frezzotti

ANNO ACCADEMICO 2015/2016

SUMMARY

The use of enzymes in the industrial catalysis is one of the most promising strategies to develop sustainable technologies and processes with minimal impact on the environment. Apart from being catalysts in vivo for numerous biochemical reactions, enzymes can also act as catalysts in vitro for various reactions relevant for industrial applications. Since enzymes are highly specific, biodegradable and usually act under milder conditions than traditional catalysts, enzyme-based processes allow to minimize waste generation and save large amounts of raw materials and energy. However, enzymatic technologies are generally not able to compete economically with the well-established chemical processes, that have been optimized for years, and enzymes are fragile and active only under very narrow reaction conditions. One solution to such problems could be the development of bioinspired synthetic catalysts based on reverse engineering studies of metallo-enzymes. To rationalize the stereoelectronic features required for the design of novel synthetic bioinspired catalysts, a theoretical study of the enzyme active site aimed at understanding the catalytic power of the enzyme and characterizing the catalytic mechanism may be extremely helpful.

In this context, the subject of this dissertation is the computational investigation of the stereoelectronic and catalytic properties of metallo-enzymes involved in reactions of technological and environmental relevance. In particular, the research focused on Hydrogenase and Carbon Monoxide Dehydrogenase (CODH) enzymes, that catalyse the reversible interconversion of protons and reducing equivalents into molecular hydrogen ($2H^+ + 2e^- \rightarrow H_2$) and the interconversion of CO and CO_2 ($CO + H_2O \rightarrow CO_2 + 2H^+ + 2e^-$), respectively. Interest in these enzymes has increased dramatically in recent years due to the increasing uncertainty about the availability of fossil fuels and concerns about CO_2 emissions and their effects on climate. The possibility of producing H_2 through hydrocarbon-independent processes and development of new methods for CO_2 reduction, make these enzymes very promising target for reverse engineering studies.

Quantum mechanics calculations have been carried out in the framework of the Density Functional Theory (DFT) on models of the enzyme active sites. Models of different sizes, ranging from the minimal metal clusters to very large systems, including the second coordination sphere, have been developed to investigate the role of the protein environment in tuning the electronic properties and the geometry of the active site. Several potential intermediate species along reaction pathways have been further investigated by calculating their spectroscopic properties. The comparison of such calculated properties with available experimental data allows to get a better understanding of the catalytic mechanism and to explore the stereoelectronic properties of the enzyme.

The thesis is divided into six chapters. Chapter 1 presents the context and the relevance of the PhD research project. It also provides an introductory overview on the investigated biological systems and outlines the main research objectives. Chapter 2 introduces the theoretical fundamentals of the employed methodology, whereas Chapter 3, 4 and 5 are devoted to the description and the discussion of the results obtained by our computational study on Hydrogenases and Carbon Monoxide Dehydrogenases.

In particular, Chapter 3 is focused on the investigation of the stereoelectronic and catalytic properties of the Ni,Fe-containing CODH enzyme, whose catalytic mechanism is still poorly understood. With the aim of unrevealing which stereoelectronic properties of the active site (known as the C-cluster) are crucial for the efficient binding of the substrates CO_2 and CO, binding of these molecules have been investigated to different redox states of the C-cluster using different sized models of the active site. In Chapter 4 is presented the study carried out on [NiFe]-hydrogenases aimed at elucidating the mechanism of the aerobic and anaerobic oxidative inactivation and reactivation of the active site. While under aerobic conditions oxidation is promoted by O_2, under anaerobic conditions two water molecules have been proposed to provide the oxygen atoms inserted in the active site in the oxidized forms of the enzyme. In Chapter 5, the reactivity of the bimetallic [Mo-(μ_2-S)-Cu] active site of Mo,Cu-containing CODHs towards H_2 is investigated by DFT calculations. With the aim of deepening insights into the nature of a H_2-bound paramagnetic enzyme form experimentally observed during the reaction of the enzyme with H_2, EPR parameters have been calculated for this species and compared with the values obtained from experiments.

The thesis closes in Chapter 7, in which the research contributions of this dissertation are reviewed and directions for future research are outlined.

CONTENTS

ABBREVIATIONS

Enzymes and Molecules

[NiFe]-hydrogenase	Ni,Fe-containing hydrogenase
[FeFe]-hydrogenase	Fe,Fe-containing hydrogenase
H-cluster	Active site of [FeFe]-hydrogenase
$[2Fe]_H$	Bimetallic subunit in the H-cluster
Fe_p	Proximal iron in $[2Fe]_H$
Fe_d	Distal iron in $[2Fe]_H$
adt	Azadithiolate
pdt	Propandithiolate
edt	Etandithiolate
CODH	Carbon Monoxide Dehydrogenase
Ni-CODH	Ni,Fe-containing CODH
C-cluster	Active site of Ni-CODH
Mo-CODH	Mo,Cu-containing CODHs
Acetyl-CoA	Acetyl Coenzyme A
ACS	Acetyl-CoA Synthase
ACDS	Acetyl-CoA Decarbonylase/Synthase
ATP	Adenosine Tri-Phosphate
ADP	Adenosine Di-Phosphate
FAD	Flavin Adenine Dinucleotide

Spectroscopic techniques

IR	Infrared
FTIR	Fourier Transformed Infrared spectroscopy
UV	Ultraviolet
EPR	Electron Paramagnetic Resonance
ENDOR	Electron Nuclear Double Resonance
XAS	X-ray Absorption Spectroscopy

Quantum Chemistry and Methods

QM	Quantum Mechanics
DFT	Density Functional Theory
KS	Kohn-Sham
HF	Hartree-Fock
UHF	Unestricted HF
RHF	Restricted HF
ROHF	Restricted Open-shell HF
MP	Møller-Plesset perturbation theory

MP2	Second order MP
CI	Configuration interaction
QCI	Quadratic CI
CISD	CI with single and double excitations
CC	Coupled Cluster
CCSD(T)	CC with singles, doubles and estimation of triple excitations
BO	Born-Oppenheimer approximation
SD	Slater Determinant
MO	Molecular Orbital
AO	Atomic Orbital
LCAO	Linear Combination of Atomic Orbitals
NBO	Natural Bond Orbital
SIE	Self-Interaction Error
STO	Slater Type Orbital
GTO	Gaussian Type Orbital
LDA	Local Density Approximation
LSDA	Local Spin Density Approximation
GGA	Generalized Gradient Approximation
ACF	Adiabatic Connection Formula
ACM	Adiabatic Connection Model
SCF	Self Consistent Field
B3LYP	Becke’s three parameter nonlocal exchange and Lee-Yang-Parr 1988 nonlocal correlation functional
BP86	Becke’s nonlocal exchange and Perdew 1986 nonlocal correlation functional
SVP	double-ζ basis set with polarization
TZVP	triple-ζ basis set with polarization
RI	Resolution of Identity
COSMO	Conductor-like screening model
BS	Broken-Symmetry

General and miscellaneous

TS	Transition State
PES	Potential Energy Surface
PDB	Protein Data Bank
RMSD	Root Mean Square Deviation
δ	Isomer shift
ΔE_q	Quadrupole splitting
E'_0	Midpoint oxidation-reduction potential

PUBLICATIONS

Raffaella Breglia, Manuel Antonio Ruiz-Rodriguez, Alessandro Vitriolo, Rubén Francisco Gonzàlez-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*, **2016**. DOI: 10.1007/s00775-016-1416-1

Raffaella Breglia, Maurizio Bruschi, Ugo Cosentino, Luca De Gioia, Claudio Greco, and Giorgio Moro, **"A theoretical study on the reactivity of the Mo/Cu containing carbon monoxide dehydrogenase with dihydrogen,"** *Protein Engineering, Design & Selection*, pp. 1–6, **2016**. DOI: 10.1093/protein/gzw071

Maurizio Bruschi, Raffaella Breglia, Federica Arrigoni, Piercarlo Fantucci, and Luca De Gioia, "**Computational approaches to the prediction of the redox potentials of iron and copper bioinorganic systems,"** *International Journal of Quantum Chemistry*, pp. 1–11, **2016**. DOI: 10.1002/qua.25228

Antonio Magri, Giovanni Tabbi, Raffaella Breglia, Luca De Gioia, Piercarlo Fantucci, Maurizio Bruschi, Raffaele Bonomo, Diego La Mendola, **"Copper ion interaction with the RNase catalytic site of the Angiogenin protein: an experimental and theoretical investigation,"** Submitted to *Inorganic Chemistry*.

Raffaella Breglia, Claudio Greco, Luca De Gioia, and Maurizio Bruschi, **"DFT investigation of aerobic and anaerobic oxidative inactivation of the [NiFe]-hydrogenase active site"** To be submitted.

Raffaella Breglia, Claudio Greco, Luca De Gioia, and Maurizio Bruschi **"Theoretical investigation of the reactivation mechanism of the Ni-A and Ni-B oxidized states of [NiFe]-hydrogenases"** To be submitted.

Raffaella Breglia, Matteo Sensi, Claudio Greco, Luca De Gioia, and Maurizio Bruschi **"Theoretical insights into stereolectronic and catalytic properties of the Ni,Fe-containing carbon monoxide dehydrogenases"** To be submitted.

Raffaella Breglia, Claudio Greco, Piercarlo Fantucci, Luca De Gioia, and Maurizio Bruschi **"DFT investigation of ^{57}Fe Mössbauer parameters for [FeFe]-hydrogenase models"** To be submitted.

CHAPTER 1

INTRODUCTION

An electrocatalyst catalyzes a redox “half reaction” in which a chemical transformation is coupled to electron transfer at an electrode. In surface electrocatalysts, such as platinum, the active sites are integral to the electrode, whereas molecular electrocatalysts can be attached to the electrode surface or diffuse freely in solution, but depend upon interfacial electron transfer. Enzymes, a special category of molecular electrocatalysts, are distinguished by their extraordinary activities, yet limited by their size and fragility. Driven by industrial and technological needs for significant improvements in rates and efficiency, enzymes can provide crucial insights into the principles underpinning the design and performance of synthetic molecular electrocatalysts.

The increased understanding of how enzymes work, how they can be manipulated, stabilized and engineered, is likely to allow fine chemical synthesis for previously inaccessible synthetic pathways. Or, equally important, to make existing ones ‘greener’. Indeed, using enzymes in catalytic processes is perfectly in line with the philosophy of Green Chemistry [1]. Not only enzymes are specific, stereoselective and efficient, they work best in the least toxic solvent of all: water. Finally, enzymes are biodegradable and can effectively be cultivated in bacteria. It is easy to envision future chemical plants where organic solvents have been replaced by water, and either bacteria or immobilized enzymes doing the job previously done by rare-earth metals. Complete understanding and control of enzyme structure-activity relationships and dynamics may allow us to design biocatalysts for any given chemical transformation. Of course, science is still far from such a goal. However, there is an ongoing, rapid development of experimental methods and computational tools, that have lead in recent years to remarkable progress towards the understanding of catalytic and stereoelectronic properties of enzymes.

1.1 Towards a “green” economy: Learning from Nature

Enzyme engineering is considered one of the most promising approaches to develop chemical and biotechnological solutions to two of the biggest challenges of our society; the sustainable production of chemical fuels from non-petrochemical sources and the reduction of greenhouse gas emissions.

Fossil fuels like coal, oil and natural gases, which were accumulated for millions of years on earth, are not renewable on human time scale; the rate of their formation is much slower than the rate of their exploitation. The search for renewable fuels to satisfy the increasing worldwide energy demand, induced by the increase of the population and the economic development, has become a major issue in recent years [2]. The emphasis on alternative and sustainable fuel resources has gained importance due to the environmental concerns about the increasing CO_2 emissions in the atmosphere, due to the intensive use of conventional fossil fuels, and their effects on climate [3]. In this context, the understanding of the nature and the action mechanism of biological systems involved in reactions of technological and environmental relevance can be extremely helpful for the development of alternative energy technologies. In particular, enzymes that catalyse reactions involving H_2 and CO_2 have attracted increasing interest in recent years.

Molecular hydrogen has been proposed as an ideally clean fuel storing large amounts of energy, which can be released without producing greenhouse gases [4]. However, dihydrogen is currently generated from fossil fuels in highly energy-consuming processes which release greenhouse gases in the atmosphere. Carbon-neutral routes to produce H_2 are mainly based on the electrolysis of water, based however on the use of the expensive and relatively scarce noble metal platinum [5]. The possibility of producing molecular hydrogen through hydrocarbon-independent processes of electrochemical proton reduction has been stimulating chemists towards the synthesis of inexpensive catalytic materials that could replace the currently used platinum-containing electrocatalysts. The search for cheaper, but still efficient catalysts could take advantage from the study of the hydrogenase enzymes, that are able to catalyse the reversible interconversion of H_2 into protons and electrons at high rates [6–9]. Hydrogenases have attracted considerable interest, not only as biological targets for the design of biomimetic or bioinspired catalysts [10–18], but also for their direct utilization in biotechnological hydrogen production processes [19–24]. Understanding the catalytic process of the H_2 evolution and oxidation at the active site of these enzymes is one of the key topics for the development of a future “hydrogen economy”.

Efficient electro- and/or photo-chemical reduction of CO_2 to provide useful carbon resources is also considered a promising approach for overcoming the limited supply of fossil fuels and the resultant increase of greenhouse gas [25–27]. Selective CO_2 reduction at low activation energy, however, is a scientifically critical challenge, because of versatile product formation and single/multi-electron reactions [28]. The design of efficient and selective base metal catalysts for CO_2 electrochemical reduction is a key requirement for the development of large-scale and eco-friendly processes for conversion of the atmospheric CO_2 into sustainable fuels [29]. A better understanding of the biological conversion of CO_2 may provide useful insight into the chemical principles underlying catalysis of CO_2 activation and reduction. In particular, carbon monoxide dehydrogenase enzymes (CODHs) have attracted great attention in recent years for their capability to catalyze the interconversion of CO_2 and CO [30].

1.2 Gas-processing metalloenzymes

The most popular autotrophic theory of the origin of life postulates that primordial metabolisms developed on mineral iron–sulphur surfaces under reducing conditions [31]. During this period of the Earth's evolution, between 4.6 and 3.5 billion years ago, the atmosphere was probably rich in gases such as H_2, CO and CO_2, and its hot oceans contained relatively high concentrations of transition-metal ions such as Fe^{2+} and Ni^{2+} [32].

Along with H_2 as an energy vector, CO_2 was most probably used as a carbon source through an ancestral version of the Wood–Ljungdahl pathway generating acetyl–coenzyme A ($CH_3C(O)–SCoA$) to be used in anabolic processes [33]. These reactions took place in an anoxic environment, as it was not until about 2.2 billion years ago that oxygen produced by photosynthetic cyanobacteria entered the atmosphere in significant amounts [32]. As a consequence, extant anaerobic H_2 and/or CO oxidizers and CO_2 reducers are currently confined in anoxic settings, for example the bottoms of lakes and oceans and the digestive systems of animals. A notable feature of these gas-based redox reactions is that they are catalysed by enzymes containing complex metallic active sites often reminiscent of FeS-containing mineral structures. This analogy constitutes the basis for the concept of primordial surface metabolism on transition-metal sulphides of the 'iron–sulphur–world' theory for the origin of life [31]. Although the catalysis of H_2 and CO_2 involving reactions is based on the metallic active site, the protein matrix determines the specificity of the reaction providing the ligands coordinated to the metals, modulating the redox potential and, in several cases, allowing gas diffusion to/from the active site through hydrophobic tunnels.

1.3 Hydrogenases

Hydrogenases enzymes, expressed by several eukaryotic and prokaryotic microorganisms, catalyse the reversible interconversion of protons and reducing equivalents into molecular hydrogen, with minimal electrochemical overpotential [34];

$$2H^+ + 2e^- \rightleftharpoons H_2 \tag{1.1}$$

Due to the many different metabolic contexts in which hydrogenases work, very diverse functions are carried out by these enzymes. The principals ones are to provide energy for organisms by oxidation of molecular hydrogen and to balance the redox potential of the cell. Depending on the redox potential and location in the cell, hydrogenases can provide reducing equivalents and protons by splitting of molecular hydrogen or remove electrons by its production [7]. They may also be involved in establishing transmembrane proton gradients [35].

Hydrogenases, according to the metal composition of the active site, are classified into [FeFe], [NiFe] and [Fe]-hydrogenases (see Figure 1.1) [7, 36, 37]. The presence of iron atoms coordinated by biologically unusual CO or/and CN^- ligands, firstly detected by FTIR spectroscopy [38][39], is a characteristic feature common to all classes.

The active site of [FeFe]-hydrogenases (Figure 1.1a), known as H-cluster, contains a [6Fe-6S] organometallic unit composed of a typical [4Fe–4S] cubane covalently linked through a cysteinyl sulphur group to a [2Fe-2S] sub-site (usually termed $[2Fe]_H$), where catalysis takes place [40, 41]. The two iron atoms of the $[2Fe]_H$ cluster, referred as to proximal (Fe_p) and distal (Fe_d) with respect to the $[Fe_4S_4]$ moiety, are bridged by a dithiolate ligand, which has been identified as a azadithiolate (adt) [42–44]. One terminal CO and CN^- ligands are bound to each Fe atom of the $[2Fe]_H$ unit. A third CO ligand is bridged or semibridged between the two metal atoms, depending on the redox state of the enzyme. Conversely, in the active site of NiFe]-hydrogenases (Figure 1.1b) one iron and one nickel atoms are bridged by two cysteine residues. Two further cysteine residues are terminally bound to the Ni atom while two CN and one CO ligands are coordinated to the Fe atom. While [FeFe] and [NiFe]-hydrogenases contain a sulfur–bridged bimetallic core, [Fe]-hydrogenases harbours only one iron atom with an unusual coordination sphere (Figure 1.1c) [45–47]. In addition, [Fe]-hydrogenases differ from [FeFe] and [NiFe]-hydrogenases because they require a second substrate, the methenyl tetrahydromethanopterin, to activate H_2 [48].

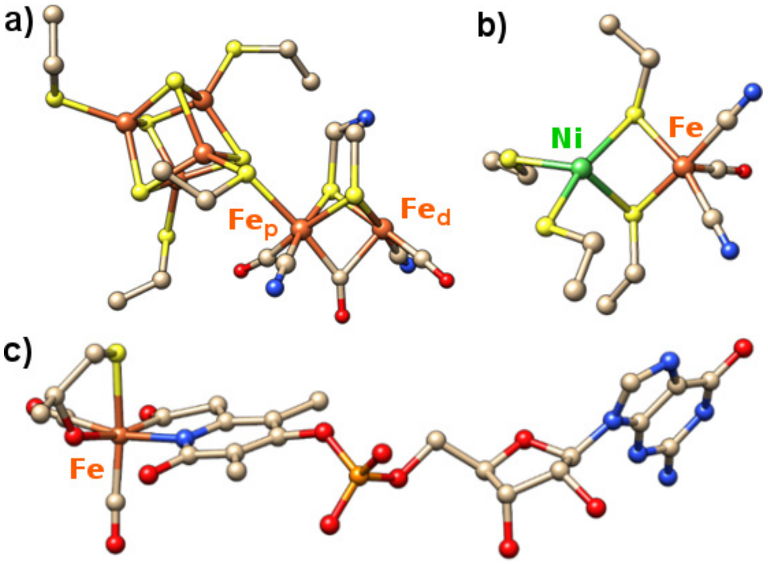


Figure 1.1 Schematic representation of the active site of a) [FeFe]-hydrogenases (pdb code: 1FEH), b) [NiFe]-hydrogenases (pdb code: 3MYR) and c) [Fe]-hydrogenases (pdb code: 3H65). Atoms are coloured according to the following scheme: nickel, green; iron, orange; nitrogen, blue; oxygen, red; sulfur, yellow; carbon, tan.

[FeFe]-hydrogenases are more active in H_2 production than [NiFe]-hydrogenases, that are usually involved in H_2 oxidation. However, the former are irreversibly inactivated upon O_2

exposure; oxidation of the proximal [Fe_4S_4] cluster prevents the electron transfer through the enzyme to the active site [49]. Conversely, reaction of [NiFe]-hydrogenases with O_2 can lead one or more oxidized inactive states, that can be reactivated upon reduction [50–52]. The latter property has stimulated numerous experimental and theoretical studies addressed to the development of O_2-tolerant biomimetic synthetic catalysts and the direct utilization of [NiFe]-hydrogenases in biotechnological hydrogen production processes. A more detailed discussion of these enzymes, on which my research focused, is presented in the following section.

1.3.1 [NiFe]-hydrogenases

[NiFe]-hydrogenases represent the most widespread hydrogenase class in Nature [53–55]; this fact, as well as the very interesting reactivity promoted by such enzymes has led researchers to spend increasing efforts in the study of [NiFe]-hydrogenases.

Genetically, [NiFe]-hydrogenases can be classified into five groups [35]. First group (group **I**) is the [NiFeSe]-hydrogenase from *Desulfovibrio baculatum*. Group **II** comprises instead the hydrogenases from *D. gigas*, *D. vulgaris* Miyasaki F and *D. fructosovorans*. Group **III** hydrogenases (i.e. *Rhodobacter capsulatus*) are slightly more divergent and display less homology. The structural difference between hydrogenases of group **II** and **III** is manifested in the different role: group **II** enzymes donate electrons to a soluble periplasmic cytochrome while electrons from group **III** hydrogenases are delivered directly to a membrane bound electron transport chain. All enzymes in groups **I**, **II** and **III** consist of a large (56-68 kDa) and a small subunit (28-35 kDa). The former contains the NiFe-catalytic site, whereas the latter contains auxiliary FeS clusters having an electron transfer function. Conversely, enzymes in group **IV** (i.e. *Ralstonia eutropha* and *Methanobacterium autotrophicum*) and **V** (E. coli hydrogenase-3) have more subunits. The [NiFe]-hydrogenases with higher homology could be evolved from a common ancestor. On the other hand, the not significant homology between the polypeptides encoding the active site of [NiFe] and [FeFe]-hydrogenases suggests that these two enzymes families evolved independently.

[NiFe]-hydrogenases can also be classified into three groups according to their O_2-induced inactivation properties; (**1**) O_2-tolerant enzymes that maintain catalytic activity under atmospheric levels of O_2 [56, 57]; (**2**) O_2-resistant [NiFeSe]-hydrogenases that lose their activity under air but reactivate rapidly upon exposure to H_2 [58] and (**3**) O_2-sensitive enzymes that are much more difficult to reactivate after air exposure.

Crystallographic studies have shed light on the structural reasons behind the dissimilar behaviour respect to the O_2-exposure. The active site of the simplest [NiFe]-hydrogenase is located in the large subunit and is composed, as described above, by one Ni and one Fe atom bridged by two cysteinyl thiolates. A third ligand bridges the two metal atoms in several enzyme forms: an oxygen-containing ligand in oxidized inactive states and a hydride in reduced active states [37, 55, 59–63]. Therefore, the coordination environment of Fe and Ni in these enzyme forms can be described as octahedral and distorted trigonal bipyramidal, respectively. Differing from the H-cluster of [FeFe]-hydrogenases, the active site of [NiFe]-hydrogenases consists only of a bimetallic core; a FeS cluster bounded to it is not required for the catalytic activity. In the small subunit are generally contained a proximal [Fe_4S_4] cluster, a medial [Fe_3S_4] and a distal [Fe_4S_4] cluster, which transfer electrons between the NiFe site and the external electron carriers (see Figure 1.2). Exceptions are the O_2-tolerant enzymes that have a modified proximal [Fe_4S_3] cluster [64–67] and the [NiFeSe]-hydrogenases that contain a medial [Fe_4S_4] cluster and a selenocysteine moiety in place of one of the terminal cysteine of the conventional [NiFe] enzyme [68].

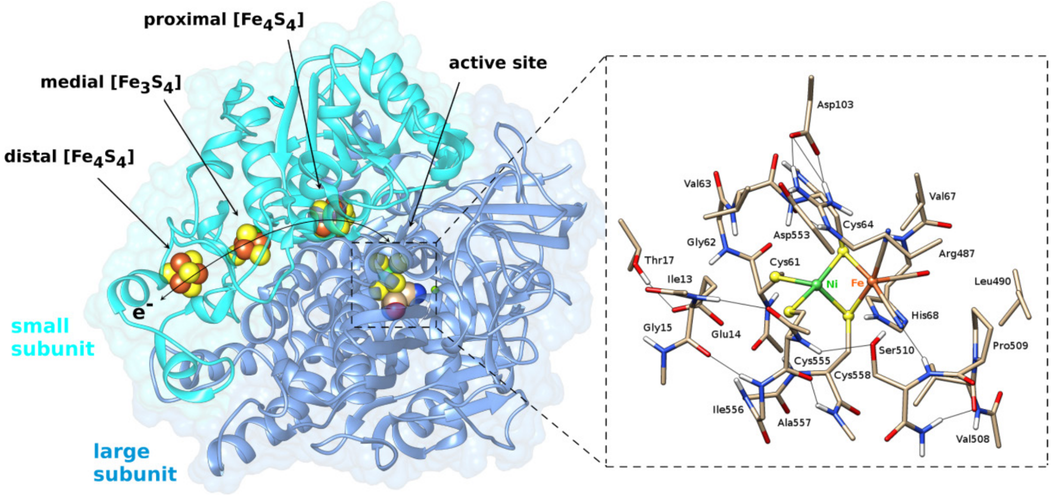
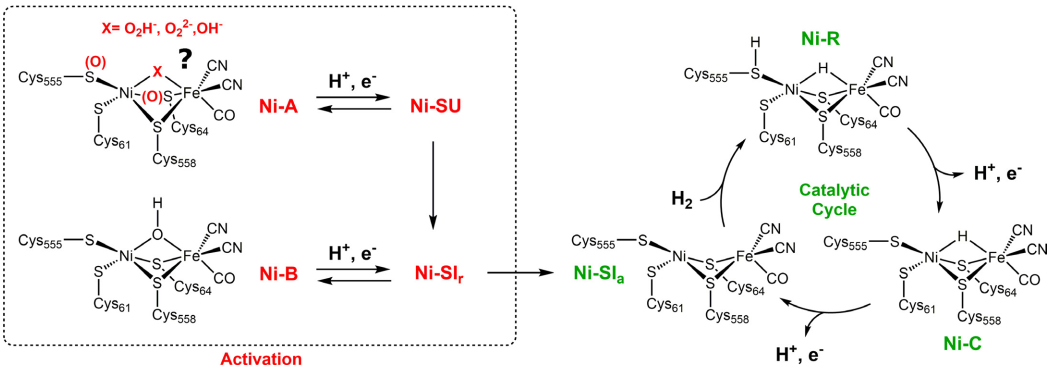


Figure 1.2 X-ray crystal structure of the [NiFe] hydrogenase from Allochromatium vinosum (PDB: 3MYR). The small subunit (shown in cyan) harbours two [Fe4S4] clusters and one [Fe3S4] cluster which transfer electrons to and from the active site. The small subunit (shown in blue) contains the [NiFe] cluster. The [NiFe] cluster and its protein environmen are shown enlarged on the right. Atoms are coloured according to the following scheme: nickel, green; iron, orange; nitrogen, blue; oxygen, red; sulfur, yellow; carbon, tan; hydrogen, white.

Several forms of the active site of O_2-sensitive hydrogenases, differing for the nature of the ligand in the third bridging site between the metal ions and/or redox and protonation state, have been identified by infrared (IR) and electron paramagnetic resonance (EPR) spectroscopy (see Scheme 1.1) [53, 62, 69, 70]. The EPR-silent Ni-SI_a (silent active) state has been suggested to correspond to the redox species that binds and activates H_2. In this form, both the Ni and Fe atoms attain the +2 oxidation state and are characterized by one vacant coordination position, where H_2 might bind. However, H_2 binding is recently proposed to be regiospecific taking place exclusively at the Ni center [71]. Heterolytic cleavage of H_2 followed by proton transfer to the sulfur atom of one of the terminally coordinated cysteine residues lead to the most reduced active form of the enzyme, which is usually referred to as Ni-R. One-electron oxidation and transfer of the sufur-bounded proton to the acceptor of the H^+ transfer network yield the catalytically active paramagnetic Ni-C state, which corresponds to a Ni(III)Fe(II) species and a hydride ligand bridging the two metal ions. Release of the second proton and the further one-electron oxidation converts Ni-C into the active Ni-SI_a state.

Oxidized and inactive forms have been also identified for the active site of [NiFe]-hydrogenases (see Scheme 1.1). Two different paramagnetic states, generally referred to as Ni-A and Ni-B, have been recognized by their spectroscopic signatures and rates of reactivation under reducing conditions [72]. In both these forms, the Fe and Ni ions attain the +2 and +3 redox states [73, 74], respectively; however, as far as reactivation kinetics is concerned, Ni-A is much slower than Ni-B [75]. The Ni-B state, upon addition of H_2 under anaerobic conditions, takes few minutes to restore the enzyme activity and therefore it is also called the “ready” state. The one-electron reduction of the Ni-B form yields a EPR-silent species referred to as Ni-Si$_r$ (silent ready), which is promptly converted to the Ni-SI$_a$ form, which is an intermediate of the catalytic cycle. The distinction between these two silent states (ready and active) has been attributed to protonation and subsequent dissociation of the oxygen containing ligand from the active site [76]. Conversely, the Ni-A is also known as the ‘unready’ state, because it is reduced within minutes to the EPR-silent Ni-SU (silent unready) state, but several hours under H_2 are required to return to the active Ni-SI$_a$ state. The different activation kinetic of Ni-B and Ni-A [72] suggests that the two resting states feature key structural differences, the nature of which is still unclear. While there is a general consensus on the Ni-B state as containing a hydroxide ligand bridging Ni(III) and Fe(II) atoms, the nature of the Ni-A state is still object of debate among researchers. Several X-ray crystallographic structures have been assigned to Ni-A, which however differ for the nature of the ligands associated to the active site [61, 63, 77–79]. In addition, X-ray analyses were characterized in most cases by disorder as different forms with partial occupancies were resolved in the same crystal.

Scheme 1.1 Activation mechanism and catalytic cycle of [NiFe]-Hydrogenases.



Depending on the conditions of exposure to O_2, different ratios of the Ni-A and the Ni-B states are obtained [51, 72, 77, 80–82]. Interestingly, the redox potential of the aerobic inactivation seemed to be the major parameter that determines the proportions of the two distinct inactive states. The NiA-to-NiB ratio increases when reaction takes place under more oxidizing conditions [51]; the Ni-A is preferentially formed upon exposure to O_2 at high redox potential, which should favour partial reduction of O_2, whereas the Ni-B state is preferentially formed when the O_2 reacts with the enzyme at low redox potentials, that should favour complete reduction of O_2 to water. Therefore, it has been suggested that [NiFe]-hydrogenases are inactivated because their active site is coordinated to partially or completely reduced O_2, resulting in the formation of the Ni-A and Ni-B, respectively. According to this hypothesis, an hydroxide, formed by the complete reduction of O_2, bridges the Ni and the Fe atoms in the Ni-B state, as confirmed by experimental and theoretical studies [61, 77, 78, 83–87], whereas in the Ni-A state a peroxide ligand has been suggested to coordinate the metallic cluster as a result of incomplete O_2 reduction. However, although some crystallographic and theoretical studies have identified the Ni-A state as containing a peroxide ligand [61, 77, 88], experimental evidences do not support this assignment. In fact, although peroxide species are generally sensitive to UV light, no change of EPR signal was observed after exposure of Ni-A to this radiation [89]. In addition, also the interpretation of the crystallographic structures of the Ni-A state, recently characterized by X-ray diffraction does not support the presence of a peroxide species [63]; a bridging hydroxide and a coordinating cysteine oxidized to bridging sulfenate have been confidently assigned to this state.

A characteristic feature of the [NiFe]-hydrogenases is that the same ready and unready enzyme states could be generated even under strictly anaerobic conditions at high overpotential [75, 90, 91], rising additional questions on the nature of the inactivation mechanism of [NiFe]-hydrogenases. Experiments showed that the proportions of Ni-A and Ni-B inactive states were not strongly dependent on the O_2 addition despite its presence accelerated the inactivation [91]. Also in the absence of O_2, more oxidizing conditions increase the proportion of Ni-A with respect to that of Ni-B.

1.4 Carbon monoxide dehydrogenases

Carbon monoxide dehydrogenases (CODHs) are found in both aerobic and anaerobic carboxidotrophic bacteria, that grow on CO, a toxic gas to animals, as a sole source of carbon and energy. In these organisms, CO is used as a metabolic building block; it is converted into CO_2, that is then fixed into cellular carbon. Thanks to their metabolism, approximately 10^8 tons of CO is removed annually from the lower atmosphere to maintain the concentration of CO at sub-toxic levels.

Two chemically distinct types of CODH are distinguished by their metal composition and distribution: Ni,Fe-containing CODH, known as Ni-CODHs, employed by diverse groups of anaerobic bacteria and archaea, and Mo,Cu-containing CODH, henceforth referred to as Mo-CODHs, functioning in aerobic carboxidotrophic bacteria (see Figure 1.3). In the former, a highly asymmetric [Ni-Fe-S] cluster (Figure 1.3a) catalyses CO oxidation with turnovers of up to 40000 s^{-1} and CO_2 reduction with rates of 12 s^{-1} [92, 93], according to the following equation:

$$CO + H_2O \rightleftharpoons CO_2 + 2H^+ + 2e^- \tag{1.2}$$

In Mo-CODHs, the catalytic site is instead a bimetallic [Mo-(μ_2-S)-Cu] system (see Figure 1.3b). Differing from the O_2 sensitive Ni-CODHs that catalyse the reversible interconversion between CO and CO_2, the Mo-CODHs [94, 95] catalyze only the CO oxidation, at a moderate rate (100 s^{-1}), and are tolerant to O_2.

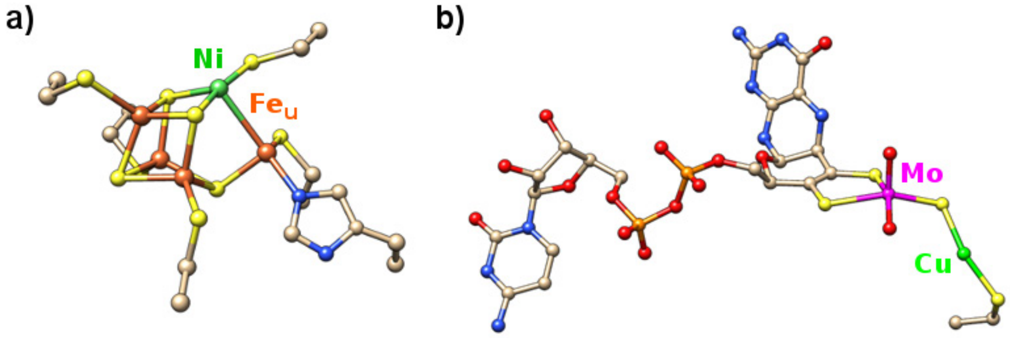


Figure 1.3 Schematic representation of the active site of: a) Ni-CODH (pdb code: 3B52) and b) Mo-CODH (pdb code: 1N5W). Atoms are coloured according to the following scheme: nickel, green; iron and phosphor, orange; molybdenum, magenta; copper, light green; nitrogen, blue; oxygen, redus; sulfur, yellow; carbon, tan.

1.4.1 Ni-CODHs

The Ni-containing carbon monoxide dehydrogenases play a fundamental role in the global carbon cycle in anaerobic microorganisms.

They can be found as unifunctional or/and bifunctional enzymes in a variety of microorganisms, including acetogenic, methanogenic, photosynthetic and sulfate-reducing bacteria [96]. Unifunctional Ni-CODHs are found in anaerobic microbes such as the photosynthetic bacterium *Rhododspirillum rubrum* [97] and the thermophilic hydrogenogenic eubacterium *Carboxydothermus hydrogenoformans* [98], in which the CODH works in concert with a membrane-bound CO-induced hydrogenase to generate a proton motive force by coupling CO oxidation to H_2 evolution [99–101]; this allows the organisms to grow with CO as the only energy source. In acetogenic bacteria (e.g. *Moorella thermoacetica*), the Ni-CODH is instead strongly associated with the acetyl CoA synthase (ACS) in the bifunctional CODH/ACS (acetyl-CoA synthase) enzyme, which couples the CODH-catalyzed reduction of CO_2 to CO with the synthesis of acetyl-CoA [102, 103]. Here, CO produced from CO_2 is a gaseous intermediate that travels approximately 70 Å through an extraordinary hydrophobic tunnel within the enzyme complex to the ACS A-cluster, where it becomes the carbonyl of acetyl-CoA. Acetyl-CoA is then either converted into cellular biomass, or its high energy thioester bond can be cleaved to drive phosphorilation of ADP to ATP in supplying energy for the cell, producing acetate as a waste product. Additionally, CODH and ACS components are present in the acetyl-CoA decarbonylase/synthase (ACDS) complex, a multymeric machine that is a major route to methane production in methanogenic archaea (e.g. *Methanosarcina thermophila* and *M. barkeri*). While in acetogenic bacteria the synthesis of acetyl-CoA allows autotrophic growth on CO_2 or CO by the Wood–Ljungdahl carbon fixation pathway, in methanogenic archaea the disassembly of acetyl-CoA facilitates the utilisation of acetate in metabolism and methane generation. It is estimated that 10^{11} tons of acetate and 10^9 tons of methane are produced globally from CO_2 every year by anaerobic acetogens and methanogens [104–106].

Representative structures of unifunctional and bufunctional carbon monoxide dehydrogenases are shown in Figure 1.4. The unifunctional CODH from *C. hydrog*enoformans (Figure 1.4a) is a homodimeric protein of approximately 130 kDa with five metal clusters [107]. Each subunit contains a $[Fe_4S_4]$ cubane (B-cluster) and an asymmetrical [Ni-Fe-S] cluster (C-cluster), which is the active site for the reversible interconversion of CO to CO_2. An additional $[Fe_4S_4]$ cubane (D-cluster) is located at the interface between the two monomers. The bifunctional CODH/ACS from *M. thermoacetica* (Figure 1.4b) is instead a 310 kD $\alpha_2\beta_2$ heterotetramer [103]. The β_2 domains, responsible for CO_2/COchemistry, form the center of the complex and closely resemble the structures of *C. hydrogenoformans* CODH. On each end of the complex is an α subunit containing the A-cluster responsible for acetyl-CoA synthesis. The catalytic A-cluster contains a standard $[Fe_4S_4]$ cluster bridged to a proximal metal site by a cysteine thiolates that is, in turn, linked to a distal (with respect to the $[Fe_4S_4]$ cubane) metal site through an additional pair of cysteine thiolates. Whereas in all ACS structures the distal site has been found to bind Ni in a square-planar coordination, either Ni, Zn and Cu have been shown to occupy the proximal site of the cluster [102, 103]. In particular, electron density maps of the *M. thermoacetica* CODH/ACS indicate a Cu ion coordinated to the $[Fe_4S_4]$ cubane.

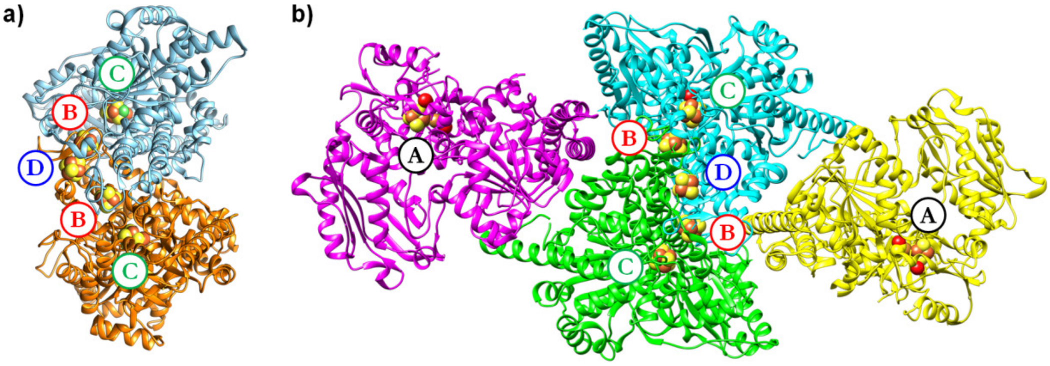


Figure 1.4 X-ray CODH crystal structures. a) Homodimeric Ni-CODH from *Carboxydothermus hydrogenoformans* (PDB code 3B52), in which each subunit contains one $[Fe_4S_4]$ cubane and one $[NiFe_4S_4]$ cluster, termed cluster B and C, respectively. A further $[Fe_4S_4]$ cluster, referred as cluster D, is located at the monomer/monomer interface. b) Bifunctional CODH/ACS from *Moorella thermoacetica* (PDB code 1MJG), in which the β_2 domains (cyan and green) resemble the *C. hydrogenoformans* CODH. Additionally, each α subunit (yellow and magenta) contains a Ni,Cu-$[Fe_4S_4]$ cluster, known as A-cluster.

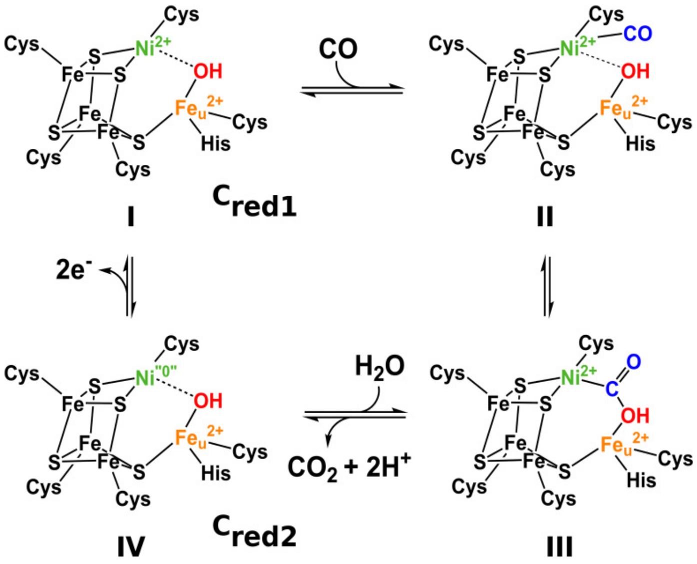
Among all the metal clusters contained in unifunctional and bifunctional CODHs, the catalytic C-cluster at which occurs the reversible CO_2 reduction is of particular interest. It is covalently bound to the protein by five cysteine residues and one histidine residue and is composed by an unusual structure formed by Ni, Fe and S atoms (see Figure 1.3a). Three Fe atoms and one Ni atom form a $[NiFe_3S_4]$ cluster, with a structure very similar to that of a 'canonical' $[Fe_4S_4]$, to which is bound an additional Fe atom extraneous to the cuboidal-like core (unique Fe or Fe_u). The high resolution (1.6 Å) crystal structure of the reduced CODH protein from *C. hydrogenoformans* shows a further bridging sulfide between the Ni and the Fe_u metal ions [98]. Conversely, in the crystallographic structure of the non-reduced, CO preincubated CODH C-cluster from *R. rubrum* (2.8 Å resolution) no bridging sulfide at the Ni-Fe_u is observed [97]. The X-ray geometry reported for the C-cluster in the bifunctional CODH/ACS from *M. thermoacetica* with (at 1.9 Å resolution) [102] and without (at 2.2 resolution)[103] CO ligand confirm the positions of the five metal ions observed in the C-cluster from monofunctional enzymes and the absence of a bridging sulfide ligand between the Ni and Fe_u. Currently, it is generally accepted that the catalytic form of the active site does not contain a fifth sulfide ligand [96, 108].

Three redox states of the C-cluster have been characterized by spectroscopic data: a fully oxidized inactive state (C_{ox}), an active state obtained from the monoelectronic reduction of C_{ox} (C_{red1}) and one obtained from the bielectronic reduction of C_{red1} (C_{red2}). A further undetected diamagnetic state (C_{int}) is postulated to arise from the one-electron reduction of C_{red1} or from the one-electron oxidation of C_{red2}. C_{ox} has a spin state S = 0 and exhibits a Mössbauer spectra typical of $[Fe_4S_4]^{2+}$ with no evidence of Fe_u [109]. At potentials below -200 mV, C_{ox} is reduced by one electron to the S = 1/2 C_{red1} state (g values at 2.01, 1.81, and 1.65, g_{av} = 1.82) [110]. C_{red1} Mössbauer parameters suggest high spin Fe(II), Fe(II), Fe(III) formal oxidation states for the $[Fe_3S_4]$ subsite and the high spin Fe(II) state for Fe_u [109], whereas L-Edge X-ray absorption spectroscopy (XAS) studies indicate a low-spin diamagnetic Ni(II) ion [111]. In this picture, the antiferromagnetic coupling between a mixed-valence Fe(II)Fe(III) pair (S = 9/2) and a ferrous Fe(II)Fe(II) pair (S = 4) results in a pure spin S = 1/2 for the C-cluster in the C_{red1} state. The lack of definite ^{61}Ni hyperfine coupling in the C_{red1} EPR signal is consistent with a Ni site electronically isolated from the cluster that does not participate in the spin coupling mechanism. The two-electron reduction of C_{red1} leads to the paramagnetic C_{red2} state (g values at 1.97, 1.87, and 1.75; g_{av} = 1.86), whose electronic structure is still uncertain [96, 107, 108, 112, 113].

C_{red1} and C_{red2} states differ by two electrons but where those electrons localize on the C-cluster is unknown. The similar EPR spectrum of these two states indicate that the electronic structure of the $[Fe_3S_4]$ core fragment is unchanged, whereas Ni K- and L-edge XAS studies suggest a low-spin diamagnetic Ni(II) for both C_{red1} and C_{red2} [111, 114]. However, the accommodation of two electrons at the Fe_u atom appears unlikely. On the basis of these considerations, two alternative descriptions of Ni in the C_{red2} state have been proposed: Ni(0) or the isoelectronic protonated site formulated as a nickel hydride species Ni(II)-H [107, 113].

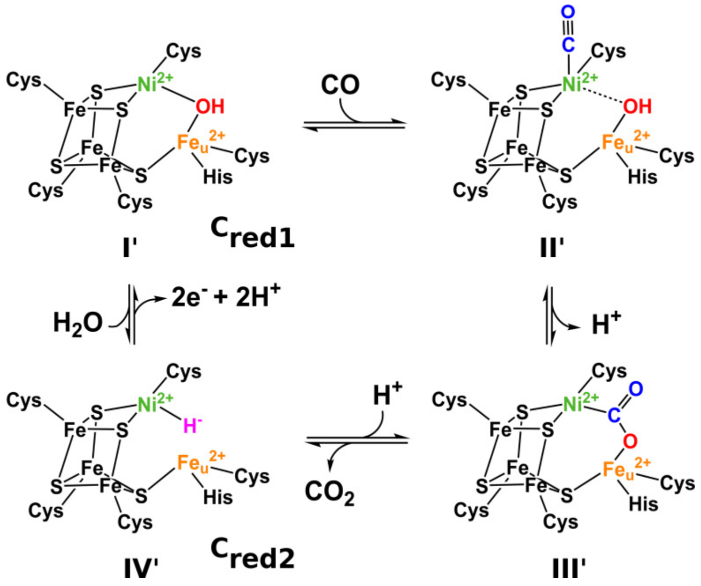
Despite contradicting scenarios for the mechanism of CODHs since the 2000s [115–117], it has been accepted that the catalysis involves the reductive conversion of the inactive C_{ox} state to C_{red1} and C_{red2}. Since C_{red1} and C_{red2} differ by two electrons and have an operational midpoint potential of -530 mV, that coincides with the values found for the CO_2/CO pair ($E^{\circ\prime}$ = -558 mV), they are respectively proposed as the redox states competent for CO oxidation and CO_2 reduction [96]. A catalytic mechanism involving the formation of a formal Ni(0) at the C-cluster (see Scheme 1.2) was firstly proposed by Jeoung and Dobbek on the basis of high resolution crystal structures of CODH with relevant ligands bound to the C-cluster [107]. Two structures from samples poised at -320 mV and -600mV, in which Ni is coordinated by three sulfur ligands with distorted T-shaped coordination geometry and Fe_u is bound to a H_2O/OH^- ligand, were interpreted as corresponding to the C_{red1} and C_{red2} states, respectively (models **I** and **IV** in Scheme 1.2). Another sample poised at -600 mV and treated with $NaHCO_3$ led to a remarkable structure with a bridging CO_2 between the Ni and Fe_u ions (model **III** in Scheme 1.2); the C atom is bound at the Ni center, with one O atom of the carboxylate group hydrogen-bonded to a conserved His residue and the other O atom coordinated to Fe_u and hydrogen-bonded to a conserved Lys group. All these structures have been assigned to catalytic relevant intermediates. When the catalytic mechanism is considered from the direction of CO oxidation, a water molecule binds to the Fe_u atom, as suggested by ENDOR spectroscopic [118] and X-ray crystallographic [97, 98] studies. This binding promotes the loss of a proton, which is transferred to the solvent via a series of His residues leading to the formation of the C_{red1} state. The substrate CO should then bind to an open coordination site of the Ni ion in the C_{red1} state and the reaction with the hydroxide ligand coordinated to Fe_u should lead to the formation of a CO_2H-bound intermediate. Loss of the proton from the carboxylate group should lead to the CO_2 formation; two electrons are then delivered to the C-cluster, reducing it from C_{red1} to C_{red2} and generating a Ni(0) state. Finally, two electrons are transferred, one at a time, from the C-cluster to the B-cluster, then to the D-cluster, and ultimately to external electron acceptors. This returns the C-cluster to the C_{red1} state.

Scheme 1.2 Ni-CODH Catalytic mechanism proposed by Jeoung and Dobbek [107], involving the formation of Ni(0) in the C_{red2} state.



A revised version of the catalytic cycle of Ni-containing CODH, proposed by Fontecilla-Camps *et al.* on the basis of combined structural and theoretical data [113], is shown in Scheme 1.3. According to this mechanism, CO reaches the active site via the hydrophobic tunnel apical to the Ni square planar coordination environment consisting of three S atoms and the Ni-Fe_u bridging hydroxide (intermediate **I'**). Binding of CO (intermediate **II'**) and reaction with OH^- would form the carboxylate species (intermediate **III'**). Reductive elimination of CO_2 leads to the C_{red2} state, featuring a hydride bound to the Ni(II) ion that adopts a square pyramidal coordination (intermediate **IV'**).

Scheme 1.3 Ni-CODH Catalytic mechanism proposed by Fontecilla-Camps *et al* [113], in which the Ni subsite in C_{red2} is bonded to a H atom, making what is formally a Ni(II)-H species.



According to both proposed mechanisms and experimental evidences, CO and CO_2 bind C_{red1} and C_{red2}, respectively. However, important aspects related to the substrate binding to the C-cluster are still unclear. In Scheme 1.2 and Scheme 1.3, CO binding is proposed to occur according to an associative mechanism that consists in the CO binding to the Ni-atom when a hydroxide ligand is coordinated to Fe_u. Conversely, a dissociative mechanism is assumed for the CO_2 binding; CO_2 binds the C-cluster after the dissociation of the H_2O ligand from Fe_u. Associative mechanism for CO binding is supported by the X-ray structure of *Methanosarcina barkeri* CODH [119], that depicts CO bound to Ni of the C-cluster in a position adjacent the water molecule/hydroxide, which is still bound to Fe_u. On the other hand, dissociative CO_2 binding is supported by the crystal structures in which CO_2 bridges the Ni and Fe_u atoms of the C-cluster [107, 120]. However, the crystal structure of CODH/ACS from *M. thermoacetica* [102] showing a putative CO ligand at the apical coordination Ni site and no ligands bound to Fe_u suggest a dissociative mechanism for CO binding, whereas the CODH structure from *C. hydrogenoformans* with the inhibitor n-butyl isocyanate [121] is indicative of an associative CO_2 binding. Indeed, it is possible to suppose that CO_2 terminally binds the Ni atom when an hydroxide ligand is coordinated to Fe_u, with a similar binding mode to that observed for n-butyl isocyanate.

1.4.2 Mo-CODHs

The Mo-CODHs [94, 95], belonging to the xanthine oxidase family of molybdoenzymes, are the enzymes responsible for CO oxidation in aerobic CO metabolism of carboxydotrophic bacteria. In these aerobic microbes, a fraction of the CO_2 produced by the CODH-catalyzed CO oxidation is fixed in the course of the pentose-phosphate pathway [122–124], whereas the derived reducing equivalents are transferred to O_2 through an electron transport chain involving quinones [124].

The CODH of *Oligotropha carboxidovorans* is the most thoroughly characterized Mo-CODH enzyme [94, 95, 123, 125]. As shown in Figure 1.5, it is a homodimer consisting of three subunits: the molybdoprotein domain (89 kDa, represented in cyan), a FAD containing domain (30 kDa, in green) and a FeS domain containing two [Fe_2S_2] clusters (18 kDa, in orange).

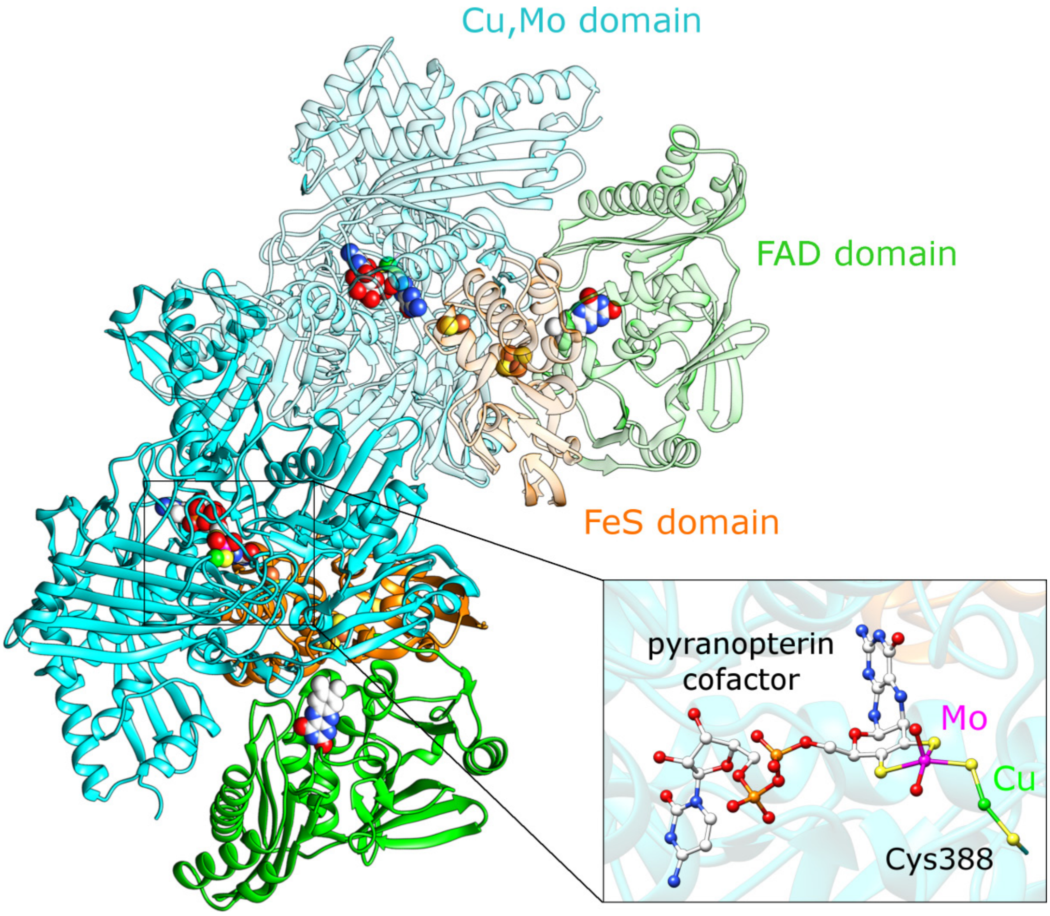
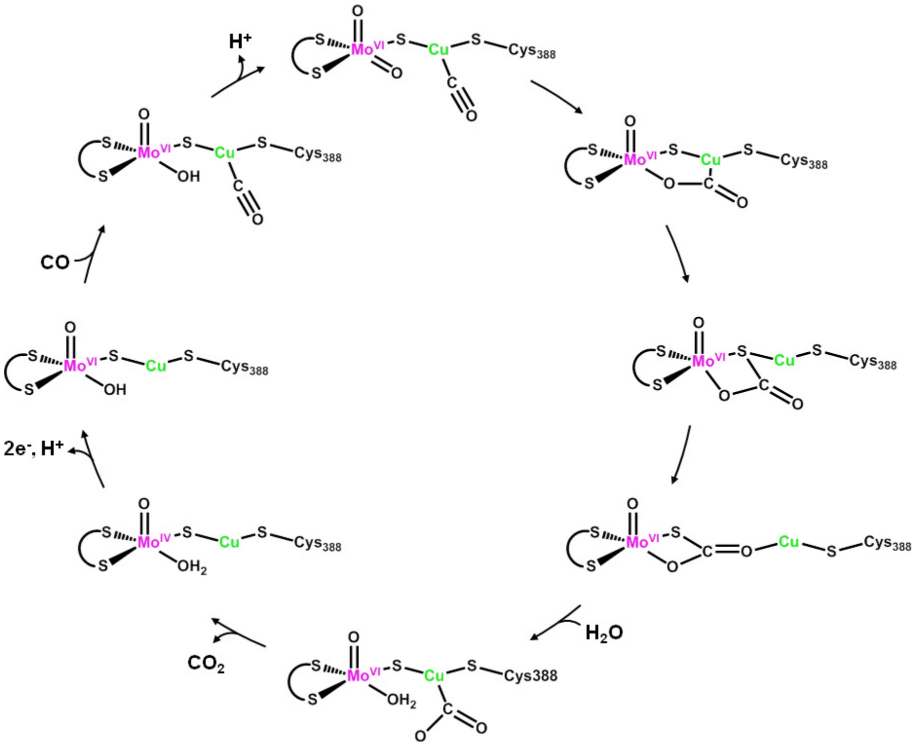


Figure 1.5 Schematic representation of the Mo-CODH strucure from *O. carboxidovorans* (pdb code: 1N5W) and zooming on the structure of the binuclear Mo/Cu active site.

The molybdoprotein domain contains the bimetallic active site, in which a molybdenum and a copper atoms are bridged by a sulfide ion (enlargement of Figure 1.5). The Cu ion is directly connected to the protein matrix by means of the sulfur atom of the Cys388 residue, while Mo is also bound to a dithiolene ligand having pterin and cytosin dinucleotide substituents. The coordination environment of the Mo ion is completed by one (or two, depending on the enzyme state) oxo ligands; a hydroxo ligand may substitute one of the oxo groups along the catalytic process of CO oxidation. Reducing equivalents, derived from the CO oxidation, are transferred from the Mo-Cu active site to the [Fe_2S_2] cluster, to the distal [Fe_2S_2] cluster and finally to the FAD cofactor.

A very high-resolution (1.1 Å) structure of CO dehydrogenase in complex with the inhibitor n-butylisonitrile has been reported [94]. In this structure the bond between the bridging sulfur atom and Cu is cleaved and the inhibitor is inserted between Mo and Cu. A reaction mechanism has been proposed in which in a similar way CO is inserted between bridging sulfur and copper of the binuclear center to yield a bridging thiocarbamate and the reduced molybdenum (see Scheme 1.4). The thiocarbamate is then proposed to be hydrolyzed by solvent, with the regeneration of the sulfur bridge to complete the catalytic cycle. This mechanism for CO oxidation is also supported by theoretical investigations [126, 127].

Scheme 1.4 Mechanism of CO oxidation in Mo-containing carbon monoxide dehydrogenase as described by Siegbahn and Shestakov [126].



As shown in Scheme 1.4, the proposed catalytic cycle implies the initial binding of CO to the Cu center, when the active site attains the Mo(VI)Cu(I) state. Notably, CO coordination to Cu(I) at the outset of catalysis is supported by ENDOR experiments [128]. New C-O bond formation using one of the Mo-bound oxo ligands precedes the insertion of a water molecule between the Mo and the newly formed CO_2 group; this then leads to the detachment of carbon dioxide and the closure of the catalytic cycle. Electrons are transferred via the sulfido ligand to yield the reduced Mo(IV) state. Cu(I) does not change the oxidation state along the mechanism, as evidenced by EPR measurements. The chain of redox active cofactors (i.e. two $[Fe_2S_2]$ clusters and a FAD molecule) facilitates the reoxidation of the reduced Mo ion through electron transfer from the active site to external redox partners like cytochrome b561.

CODH enzymes are very promising targets for reverse engineering studies aimed at developing efficient catalysts for CO_2 reduction, that plays an important role in converting CO_2 to carbon-based fuel in a clean and sustainable way. In addition, the biological CO_2/CO conversion catalysed by these enzymes may also inspire alternative routes for similar reactions in chemical industries. Indeed, it was experimentally observed that both Ni-CODHs and Mo-CODHs are also able to catalyse the oxidation of molecular hydrogen at significant rates [129–131]. The anaerobic Ni-CODHs catalyse the oxidation as well as the reverse H_2 evolution, with rates of ~ 3 s^{-1} for H_2 production and a rate of 0.1 s^{-1} for H_2 oxidation.

The hydrogenase activity of Mo-CODHs has recently been the object of a detailed experimental study, that confirmed the reduction of the enzyme by H_2 over a broad range of pH values [132]. Important mechanistic clues for H_2 binding and H_2 splitting at the metallic active site are provided by a new EPR signal observed in the course of the reaction with H_2, which is attributable to the partially reduced Mo(V)Cu(I) enzyme in which the Cu ion is coordinated by H_2. Displacement of a Cu-bound H_2O ligand by H_2 has been proposed as first step of the H_2 oxidation. Subsequent deprotonation through a copper-hydride intermediate leads to formal reduction of the binuclear center to a Mo(IV)Cu(I) state. Despite these advances, the mechanism of H_2 oxidation by Mo-CODHs has not yet been clearly elucidated and further investigations of the reactivity of the active site towards H_2 are required.

1.5 Outline of the thesis

The interest in hydrogenases and carbon monoxide dehydrogenases has increased significantly over the past years, since the disclosure of stereoelectronic and catalytic properties of these enzymes is important not only in the context of the efforts aimed at elucidating structure-function relationships but also for the rational design of novel bioinspired catalysts for H_2 production and CO_2 reduction. However, despite the large amount of experimental and theoretical data available for these enzymes, several aspects related to their structure and reactivity are not yet fully understood. Further investigations are required to shed more light on the stereoelectronic and catalytic properties of such enzymes. In this context, theoretical methods, particularly those based on density functional theory (DFT), have been proved to be useful tools to explore molecular properties of the enzyme active sites providing complementary information to experimental techniques. Knowledge from crystallographic and spectroscopic studies often can't disclose all structural and redox features of the enzymes. In addition computational studies allow to perform calculations on model systems avoiding the costs and time employed in the process of synthesis.

In this context, in order to provide significant insights into molecular properties of hydrogenases and carbon monoxide dehydrogenases, quantum mechanics calculations have been carried out in the framework of the Density Functional Theory (DFT) on models of the active site of these enzymes. Models of different sizes, ranging from the minimal metal clusters to very large systems, including the second coordination sphere, have been developed. Several intermediate species along the enzymes reaction pathways have been further investigated calculating spectroscopic properties. The comparison of calculated and experimental spectroscopic parameters allowed to get a better understanding of the catalytic mechanisms and to explore the stereoelectronic properties of the active sites in the different forms of the enzymes.

Different issues for hydrogenases and carbon monoxide dehydrogenases have been addressed, depending on the current state of knowledge and the still open questions concerning these enzymes. Important achievements have been made over the past decade in understanding hydrogenases. However, several issues concerning their practical use as catalysts for H_2 production remain unsolved. On the other hand, compared to hydrogenases, the knowledge achieved on carbon monoxide dehydrogenases is still very limited. Further kinetic, spectroscopic and computational studies are required to improve our understanding of the CODHs reactivity. In light of these considerations, quantum mechanical calculations on hydrogenases have been carried out to provide insights into the development of bioinspired catalysts and their direct utilization in biotechnological hydrogen production processes. Conversely, the study of CODHs was aimed at elucidating stereoelectronic properties and reactivity of their active site.

In Chapter 3 will be presented a study on Ni-CODHs, whose catalytic mechanism is still poorly understood. With the aim of unrevealing which characteristics of the active site are crucial for the efficient binding of the substrates, CO_2 and CO binding to different forms of the C-cluster has been investigated using both a minimal model of the C-cluster and a very large model of the active site. A comparative analysis of results obtained using these two different size models highlighted the fundamental role of the protein in assisting the substrates binding to the C-cluster.

Conversely, stereoelectronic factors for efficient binding and oxidation of H_2 at the active site of [NiFe]-hydrogenases have been already disclosed. Oxidation and consequent inactivation of the active site are instead not yet well understood. A better understanding of such processes could be exploited in the design of novel O_2-tolerant biomimetic synthetic catalysts and in the direct utilization of [NiFe]-hydrogenases in biotechnological hydrogen production processes. In this context, to outline the key features of the peculiar reactivity towards O_2, interconversion mechanisms between active and inactive forms of [NiFe]-hydrogenases have been investigated using a large-size DFT model of the active site and its protein environment (Chapter 4). The nature of the oxidized Ni-A state has been first investigated by systematic explorations of conformational and configurational space relying on accurate energy calculations, and by comparisons of theoretical geometries with the X-ray structures currently available. Since it was demonstrated that the inactive Ni-B and the Ni-A states can be generated even in the absence of O_2, the oxidative inactivation mechanism was investigated by simulating both aerobic and anaerobic conditions. While under aerobic conditions oxidation is promoted by O_2, under anaerobic conditions two water molecules have been proposed to provide the oxygen atoms that are inserted in the active site in the oxidized forms of the enzyme.

Finally, DFT has been used to explore the reactivity of the bimetallic [Mo-(μ_2-S)-Cu] active site of Mo-CODHs towards H_2 (Chapter 5). With the aim of deepening insights into the nature of a H_2-bound paramagnetic form of the enzyme experimentally observed during the reaction of the enzyme with H_2, EPR parameters have been predicted for this species and compared with the values obtained from experiments.

CHAPTER 2

THEORETICAL BACKGROUND AND METHODS

Metalloproteins are involved in a plethora of catalytic processes, and therefore have attracted the attention of inorganic and metallo-organic chemists, which have been inspired by the structure of active sites of metalloproteins in the design and synthesis of bio-mimetic coordination compounds [133]. In this context, the role of computational chemistry, and more specifically of quantum mechanical (QM) methods, cannot be underestimated. These methods, when giving reliable predictions, may nicely complement the experimental investigations [134], especially to elucidate the catalytic mechanisms of metallo-enzymes, and of related bio-mimetic coordination compounds. In fact, most QM studies have been focused on the chemical steps of the catalytic cycles, disclosing the structure and relative energy of key intermediate species and corresponding transition states connecting them. Nowadays, calculations of chemical properties of complex systems as enzymes through a quantum mechanical approach are possible thanks to the availability of efficient computational formalisms and to the outstanding technological development of computing facilities.

2.1 The Schrödinger equation

Since the revelations of phenomena such as energy quantization and particle-wave duality of light in the early 1900's, has been know that an accurate description of chemical processes requires the treatment of the atoms by their own constituents, nucleus and electrons, that cannot be described using the Newton's equations of classical mechanics. Valuable mathematical tools for describing the motion of electrons in a molecule are provided by quantum mechanics.

The foundation of quantum mechanics is a product of many great contributions but the most famous and also most relevant for chemists is the one made by Erwin Schrödinger in 1926 [135]. Schrödinger theory, which extended the de Broglie concept of matter waves providing a formal method of treating the dynamics of physical particles in terms of associated waves, is in fact the basis of quantum chemistry.

Schrödinger proposed that the state of a physical system is completely described by its time-dependent wave function Ψ, i.e. Ψ contains all information about the system's state. While no physical property can be drawn directly from Ψ, the complex conjugate $\Psi*\Psi$ (or $|\Psi|^2$ for short) represents a probability density. For example, the probability $\rho(\boldsymbol{r})$ of finding an electron described by Ψ in a volume element dr, is found by:

$$\rho(\boldsymbol{r}) = |\Psi(\boldsymbol{r})|^2 dr \tag{2.1}$$

Since the probability of finding a particle somewhere is 1, the wave function, to represent a physically observable system, must satisfy the normalization criterion:

$$\int |\Psi|^2 \, d\tau = \langle\Psi|\Psi\rangle = 1 \tag{2.2}$$

where dτ represents integration over the entire space.

The wave function Ψ of a physic system, associated to its energy, is found by solving the relativistic time-dependent Schrödinger equation:

$$\hat{H}(\boldsymbol{r},\boldsymbol{R},t)\Psi(\boldsymbol{r},\boldsymbol{R},t) = i\hbar\frac{\partial}{\partial t}\Psi(\boldsymbol{r},\boldsymbol{R},t) \tag{2.3}$$

where $\hat{H}$ is the Hamiltonian or energy operator, t is the time, **r** and **R** are respectively the position vectors of electrons and nucleus containing in the system, i is the imaginary unit and ħ is the Planck constant divided by 2π. However, it is often impracticable to use a complete relativistic formulation of this equation. Therefore, Schrödinger himself postulated a non-relativistic approximation which is nowadays often used, especially in quantum chemistry. A special cases are the solutions of the time-independent Schrödinger equation, where the Hamiltonian itself has no time dependency, and the solutions therefore describe standing waves which are called stationary states. The time-independent Schrödinger equation is not only easier to treat, but the knowledge of its solutions also provides crucial insight to handle the corresponding time-dependent equation. The time-independent equation is obtained by the approach of separation of variables, i.e. the spatial part of the wave function is separated from the temporal part via:

$$\Psi(\boldsymbol{r},\boldsymbol{R},t) = \psi(\boldsymbol{r},\boldsymbol{R})\tau(t) = \psi(\boldsymbol{r},\boldsymbol{R}) \cdot e^{-i\omega t} \tag{2.4}$$

Furthermore, the left-hand side of the resultant equation reduces to the energy eigenvalue of the Hamiltonian multiplied by the wave function, leading to the general eigenvalue equation:

$$\hat{H}(\boldsymbol{r},\boldsymbol{R})\Psi(\boldsymbol{r},\boldsymbol{R}) = E\Psi(\boldsymbol{r},\boldsymbol{R}) \tag{2.5}$$

where $\hat{H}(\boldsymbol{r})$ is the time-independent Hamiltonian. $\Psi(\boldsymbol{r})$ is the eigenfunction of $\hat{H}$ while E is its eigenvalue.

The general many-particle time-independent Hamiltonian describing a system of N electrons and M nuclei takes the form:

$$\hat{H}(\boldsymbol{r},\boldsymbol{R}) = -\sum_{i=1}^{N}\frac{\hbar^2}{2m_e}\nabla_i^2 - \sum_{A=1}^{M}\frac{\hbar^2}{2M_A}\nabla_A^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{e^2 Z_A}{\boldsymbol{r}_{iA}} + \frac{1}{2}\sum_{i=1}^{N}\sum_{j>i}^{N}\frac{Z_A}{\boldsymbol{r}_{ij}} + \frac{1}{2}\sum_{A=1}^{M}\sum_{B>A}^{M}\frac{e^2 Z_A Z_B}{\boldsymbol{R}_{AB}} \tag{2.6}$$

where i and j run over the N electrons and A and B over the M nuclei. M_A and m_e are respectively the mass of the nucleus and the electron and Z_A is the atomic number. $\boldsymbol{R}$ and $\boldsymbol{r}$ denote interparticular distances ($\boldsymbol{r}_{iA} = |\boldsymbol{r}_i - \boldsymbol{R}_A|$, $\boldsymbol{r}_{ij} = |\boldsymbol{r}_i - \boldsymbol{r}_j|$, $\boldsymbol{R}_{AB} = |\boldsymbol{R}_A - \boldsymbol{R}_B|$) whereas ∇_i^2 and ∇_A^2 represent respectively the Laplace operator of the electron and nucleus, which denote the sum of differential operators with respect to the Cartesian coordinate:

$$\nabla_i^2 = \frac{\partial}{\partial x_i^2} + \frac{\partial}{\partial y_i^2} + \frac{\partial}{\partial z_i^2} \tag{2.7}$$

The five terms composing the Hamiltonian consist of the operators corresponding to, in order, the kinetic energy of the electrons ($\hat{T}_e$) and nuclei ($\hat{T}_N$), the attractive electron–nuclei interaction ($\hat{V}_{eN}$), the electron–electron repulsion ($\hat{V}_{ee}$) and the repulsive nuclei–nuclei interaction ($\hat{V}_{NN}$). Hence, the Hamiltonian operator can be written according to the following compact form:

$$\hat{H}(\boldsymbol{r},\boldsymbol{R}) = \hat{T}_e(\boldsymbol{r}) + \hat{T}_N(\boldsymbol{R}) + \hat{V}_{eN}(\boldsymbol{r},\boldsymbol{R}) + \hat{V}_{ee}(\boldsymbol{r}) + \hat{V}_{NN}(\boldsymbol{R}) \tag{2.8}$$

where the various kinetic $\hat{T}$ and potential $\hat{V}$ operators depend on the electrons ($\boldsymbol{r}$) and nucleus ($\boldsymbol{R}$) positions.

All the properties of a system, and particularly the energy, can, in principle, be derived from the solution of the non-relativistic time-independent Schrödinger equation. However, only for one-electron systems, such as an isolated hydrogen atom, this equation can be solved exactly. The interactions of a many-body system are too complicated to be solved analytically; even a simple molecule such as water contain 13 particles, and according to Equations (2.6) and (2.8) is described by an Hamiltonian of 126 terms. Therefore, to handle chemical problems a series of approximations need to be made.

2.2 The Born-Oppenheimer approximation

The so-called Born-Oppenheimer approximation (BO) is central to Quantum Chemistry and is applied routinely because of its simplicity and high retained accuracy. It states that since electrons move orders of magnitude faster than the nuclei, they can be viewed as moving in a field of fixed nuclear potential. This approximation is justified by the significant difference of mass of electrons and nuclei and it practically allows the separation of the motion of the electrons from the motion of the nuclei. Applying this approximation, the total wave function of the system can be written as:

$$\Psi(\boldsymbol{r},\boldsymbol{R}) = \phi_N(\boldsymbol{R})\psi_{el}(\boldsymbol{r};\boldsymbol{R}) \tag{2.9}$$

where $\phi_N(\boldsymbol{R})$ describes the nuclei and $\psi_{el}(\boldsymbol{r};\boldsymbol{R})$, which depends parametrically on the positions of the nuclei, the electrons. Consequently, the Schrödinger equation can be split up into an electronic and a nuclear part:

$$\hat{H}_N(\boldsymbol{R})\phi_N(\boldsymbol{R}) = E_N(\boldsymbol{R})\phi_N(\boldsymbol{R}) \tag{2.10}$$

$$\hat{H}(\boldsymbol{r};\boldsymbol{R})\psi_{el}(\boldsymbol{r};\boldsymbol{R}) = E(\boldsymbol{r};\boldsymbol{R})\psi_{el}(\boldsymbol{r};\boldsymbol{R}) \tag{2.11}$$

where $\hat{H}_N(\boldsymbol{R})$ and $\hat{H}(\boldsymbol{r};\boldsymbol{R})$ represent respectively the nuclear and electronic Hamiltonians:

$$\hat{H}_N(\boldsymbol{R}) = \hat{T}_N(\boldsymbol{R}) \tag{2.12}$$

$$\hat{H}(\boldsymbol{r};\boldsymbol{R}) = \hat{T}_e(\boldsymbol{r}) + \hat{V}_{eN}(\boldsymbol{r};\boldsymbol{R}) + \hat{V}_{ee}(\boldsymbol{r}) + \hat{V}_{NN}(\boldsymbol{R}) \tag{2.13}$$

Especially for problems of molecular physics and quantum chemistry, the electronic Schrödinger equation is of major interest. But despite all simplifications, a simple look at the last equation indicates that there are still a few more crucial points left to deal with until a useful solution can be obtained. The nuclei-nuclei repulsion $\hat{V}_{NN}(\boldsymbol{R})$ is reduced to a potential constant for a fixed geometry and therefore can be considered separately. The kinetic energy term $\hat{T}_e(\boldsymbol{r})$ and the electron-electron repulsion $\hat{V}_{ee}(\boldsymbol{r})$ don't depend on the nuclear coordinates $\boldsymbol{R}$. The nuclei-electron interaction $\hat{V}_{eN}(\boldsymbol{r};\boldsymbol{R})$ is therefore the only part of the electronic Hamiltonian which depends on nuclear positions. Hence, the purely electronic Hamiltonian and the electronic part of the Schrödinger equation read:

$$\hat{H}_{el}(\boldsymbol{r};\boldsymbol{R}) = \hat{T}_e(\boldsymbol{r}) + \hat{V}_{eN}(\boldsymbol{r};\boldsymbol{R}) + \hat{V}_{ee}(\boldsymbol{r}) \tag{2.14}$$

$$\hat{H}_{el}(\boldsymbol{r};\boldsymbol{R})\psi_{el}(\boldsymbol{r};\boldsymbol{R}) = E_{el}(\boldsymbol{r};\boldsymbol{R})\psi_{el}(\boldsymbol{r};\boldsymbol{R}) \tag{2.15}$$

The Schrödinger equation can be then solved only for the electronic part, keeping the positions of the nuclei as external parameters in solving it. The electronic energy $E_{el}(\boldsymbol{r};\boldsymbol{R})$ is therefore obtained for a frozen conformation of the nuclei. By calculating the total energy for different nuclear arrangements as the sum of $E_{el}(\boldsymbol{r};\boldsymbol{R})$ and $\hat{V}_{NN}(\boldsymbol{R})$, a *potential energy surface* (PES) can be obtained, which defines the equilibrium conformations of a molecule. The nuclei can be considered moving on this energy surface on which every point is the solution to the electronic Schrödinger equation at a given nuclear geometry $\boldsymbol{R}$.

It should be noted that our general perception of chemistry as being the interaction of electrons about a set of localized nuclei becomes rather blurred as one goes beyond the Born-

Oppenheimer approximation, since entities such as 'bond lengths' or 'transition states' become meaningless.

Unfortunately, the BO approximation is not enough to solve the electronic Schrödinger equation for a many-body system. Further approximations for wave functions are needed.

2.3 The variational principle

In order to find a suitable strategy to approximate the analytically not accessible solutions of many-body problems, a very useful tool is the variational principle, that allows to approach the ground state wave function ψ_0, which corresponds to the lowest energy E_0 of a given system. Within the Born-Oppenheimer approximation, this means finding the electronic distribution that minimizes E_{el}. Luckily, it can be shown that the energy of an model wave function ψ cannot become lower than the energy of the exact (unknown) wave function. The following equation, corresponding to the variational principle, provides a way to systematically improve an approximate solution to the Schrödinger equation.

$$E_{el} = \frac{\int \psi^* \hat{H} \psi dx}{\int \psi^* \psi \, dx} = \frac{\langle \psi | \hat{H} | \psi \rangle}{\langle \psi | \psi \rangle} \geq E_0 \tag{2.16}$$

2.4 An antisymmetric many-body representation

The most simple ansatz of a many-body wave function $\psi(\boldsymbol{r}_1, \boldsymbol{r}_{2,}, \ldots, \boldsymbol{r}_N)$, whose form is unknown, is the so called Hartree product; a product of a number of one-partcle wave functions $\varphi_i(\boldsymbol{r}_i)$.

$$\psi^{HP}(\boldsymbol{r}_1, \boldsymbol{r}_{2,}, \ldots, \boldsymbol{r}_N) = \varphi_1(\boldsymbol{r}_1)\varphi_2(\boldsymbol{r}_2) \ldots \varphi_N(\boldsymbol{r}_N) = |\varphi_1 \varphi_2 \ldots \varphi_N\rangle \tag{2.17}$$

However, the Hartree product fails to comply with the Pauli exclusion (or antisymmetry) principle. It states that for fermions (particles with half-integer spin), the wavefunction must be antisymmetric with respect to a permutation of two particles ($\hat{P}$), i.e. an exchange of two particles must cause a sign change of the wave function. That is:

$$\hat{P}_{ik}\Psi(x_1, \ldots, x_{i,}, \ldots, x_k, \ldots) = \Psi(x_1, \ldots, x_k, \ldots, x_i, \ldots) = -\Psi(x_1, \ldots, x_{i,}, \ldots, x_k, \ldots) \tag{2.18}$$

It turns out that a forth quantum number is needed to describe a particle and its wav function, namely the spin number. The now familiar concept of electron spin was first invoked in 1925 to explain the fine structure of electronic spectra [136]. This intrinsic property appears as a forth quantum number in relativistic quantum mechanics, as shown by Dirac in 1928 [137]. In non-relativistic representations, spin is postulated, and is introduced by using *spin orbitals*, denoted $\chi(\mathbf{x})$. A spin orbital is a one-electron wave function given by the product of a spatial orbital $\varphi(\boldsymbol{r})$, that depends on the coordinate of the electron, and the spin function $\alpha(\omega)$ or $\beta(\omega)$, that depends on its spin:

$$\chi(\boldsymbol{x}) = \begin{cases} \varphi(\boldsymbol{r})\alpha(\omega) \\ \quad or \\ \varphi(\boldsymbol{r})\beta(\omega) \end{cases} \tag{2.19}$$

Using a normalization factor of $(N!)^{-1/2}$ an N-electron wave function ψ that fulfils the antisymmetry requirement can be constructed as a so-called Slater Determinant [138], in which each column contains a spin orbital whereas the rows are labelled by the electron coordinates. The Slater determinant can be seen as a linear combination of permuted Hartree products. Permutation of two rows, which corresponds to interchange of two electrons, changes the sign of the determinantal wave function ψ, thus satisfying the requirement of antisymmetry. If two electrons occupy the same spin orbital, two rows in the determinant become equal, i.e. the determinant will vanish (Pauli Exclusion Principle).

$$\psi^{SD}(x_1, x_{2,}, \dots, x_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\boldsymbol{x}_1) & \chi_2(\boldsymbol{x}_1) & \cdots & \chi_N(\boldsymbol{x}_1) \\ \chi_1(\boldsymbol{x}_2) & \chi_2(\boldsymbol{x}_2) & \cdots & \chi_N(\boldsymbol{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\boldsymbol{x}_N) & \chi_2(\boldsymbol{x}_N) & \cdots & \chi_N(\boldsymbol{x}_N) \end{vmatrix} \equiv \tag{2.20}$$

$$\equiv |\varphi(x_1, x_{2,}, \dots, x_N)\rangle \equiv |\chi_1\chi_2 \cdots \chi_N\rangle$$

2.5 The basis set approximation

Since an analytical expression for the one-electron spin orbitals composing the Slater determinant $\psi^{SD}(\mathbf{x}_i)$, as well as for the many-electrons eigenfunction, is not accessible, a further wave function approximation is needed. The spatial functions $\varphi(\boldsymbol{r})$ of the spin orbitals $\chi(\mathbf{x})$ can be constructed as a linear combination of basis functions, whose analytical forms are known a priori.

$$\varphi(\boldsymbol{r}) = \sum_{\mu}^{M} c_{\mu i}\phi_{\mu}(\boldsymbol{r}) \tag{2.21}$$

Expanding an unknown function, such as a molecular orbital, in a set of known functions is not an approximation, if the basis is complete. However, a complete basis means that an infinite number of basis functions must be used, which is impossible in actual calculations. The introduction of a finite basis set consists of one of the inherent approximations in essentially all ab initio methods.

The molecular orbitals (MO) in a molecule are usually given as a linear combination of the atomic orbitals (AO) of the atoms contained in the molecule (LCAO, Linear Combination of Atomic Orbitals). An unknown MO can be thought of as a function in the infinite coordinate system spanned by the complete basis set. When a finite basis is used, only the components of the MO along those coordinate axes corresponding to the selected basis can be represented. The smaller the basis, the poorer the representation. The type of basis functions used also influences the accuracy. The better a single basis function is able to reproduce the unknown function, the fewer basis functions necessary for achieving a given level of accuracy. Knowing that the computational effort of ab initio methods scales formally at least as M^4, it is of course of prime importance to make the basis set as small as possible without compromising the accuracy.

2.5.1 Slater and Gaussian Type Orbitals

There are two types of basis functions (also called Atomic Orbitals, AO, although in general they are not solutions to an atomic Schrödinger equation) commonly used in electronic structure calculations: Slater Type Orbitals (STO) and Gaussian Type Orbitals (GTO).

A procedure that has come into wide use is to fit a Slater-type orbital (STO) to a linear combination of n = 1,2,3,… primitive Gaussian functions. This is the STO-nG procedure. In particular, STO-3G basis sets are often used in polyatomic calculations, in preference to evaluating integrals with Slater functions. Slater type orbitals have the functional form:

$$\chi_{\zeta,n.l,m}(r,\theta,\varphi) = NY_{l,m}(\theta,\varphi)r^{n-1}e^{-\zeta r} \tag{2.22}$$

where *n*, *l*, *m* are the quantum numbers, N is a normalization constant and $Y_{l,m}$ are the spherical harmonic functions effectively determining the shape of the GTO (s, p, d, etc.). ζ is instead a parameter determining the spread of the orbital. The exponential dependence on the distance between the nucleus and the electron mirrors the exact decay behavior of the orbitals for the hydrogen atom. However, since STOs do not have any radial nodes, nodes in the radial part are introduced by making linear combinations of STOs. The exponential dependence ensures a fairly rapid convergence with increasing number of functions, however, the calculation of three- and four-centre two electron integrals cannot be performed analytically. STOs are primarily used for atomic and diatomic systems where high accuracy is required, and in semiempirical methods where all three- and four- center integrals are neglected.

Gaussian type orbitals can be written in terms of polar or cartesian coordinates:

$$\chi_{\zeta,n.l,m}(r,\theta,\varphi) = NY_{l,m}(\theta,\varphi)r^{2n-2-l}e^{-\zeta r^2} \tag{2.23}$$

$$\chi_{\zeta,l_x.l_y,l_z}(x,y,z) = Nx^{l_x}y^{l_y}z^{l_z}e^{-\zeta r^2} \tag{2.24}$$

where the sum of l_x, l_y and l_z determines the type of orbital (for example $l_x + l_y + l_z = 1$ is a p-orbital).

The r^2 dependence in the exponent makes the GTOs inferior to the STOs in two aspects. At the nucleus the GTO has zero slope, in contrast to the STO which has a "cusp" (discontinuous derivative), and GTOs have problems representing the proper behaviour near the nucleus. The other problem is that the GTO falls off too rapidly far from the nucleus compared with an STO, and the "tail" of the wave function is consequently represented poorly. Both STOs and GTOs can be chosen to form a complete basis, but the above considerations indicate that a linear combination of GTOs is required to reproduce the accuracy of STOs. A rough guideline says that three times as many GTOs as STOs are required for reaching a given level of accuracy. The increase in number of basis functions, however, is more than compensated for by the ease by which the required integrals can be calculated. In terms of computational efficiency, GTOs are therefore preferred, and used almost universally as basis functions in electronic structure calculations.

2.5.2 Classification of basis sets

Having decided on the type of basis function (STO/GTO) and their location (nuclei), the most important factor is the number of functions to be used. The smallest number of functions possible is a minimum basis set. It can typically be systematically improved by changing the number of basis functions, the degree of polarizability and by allowing electronic 'smearing' by explicit inclusion of diffuse GTOs with very small values of ζ.

The first improvement in the basis sets is a doubling of all basis functions, producing a Double Zeta (DZ) type basis. The term zeta stems from the fact that the exponent of STO basis functions is often denoted by the greek letter ζ. A DZ basis thus employs two s-functions for hydrogen (1s and 1s'), four s-functions (1s, 1s', 2s and 2s') and two p-functions (2p and 2p') for first row elements, and six s-functions and four p-functions for second row elements. Doubling the number of basis functions allows for a much better description of the fact that the electron distribution in molecules can differ significantly from the one in the atoms and the chemical bond may introduce directionalities which cannot be described by a minimal basis.

The next step up in basis set size is a Triple Zeta (TZ) basis. Such a basis contains three times as many functions as the minimum basis, i.e. six s-functions and three p-functions for the first row elements. Some of the core orbitals may again be saved by only splitting the valence, producing a triple zeta split valence basis set. The names Quadruple Zeta (QZ) and Quintuple Zeta (5Z, not QZ) for the next levels of basis sets are also used, but large sets are often given explicitly in terms of the number of basis functions of each type.

In most cases higher angular momentum functions are also important, these are denoted polarization functions. Consider for example a CH bond which is primarily described by the hydrogen s-orbital(s) and the carbon s- and p_z-orbitals. It is clear that the electron distribution along the bond will be different than that perpendicular to the bond. If only s-functions are present on the hydrogen, this cannot be described. However, if a set of p-orbitals is added to the hydrogen, the p component can be used for improving the description of the H-C bond. The p-orbital introduces a polarization of the s-orbital(s). Similarly, d-orbitals can be used for polarizing p-orbitals, f-orbitals for polarizing d-orbitals etc. Adding a single set of polarization functions (p-functions on hydrogens and d-functions on heavy atoms) to the DZ basis forms a Double Zeta plus Polarization (DZP) type basis. Similarly to the sp-basis sets, multiple sets of polarization functions with different exponents may be added. If two sets of polarization functions are added to a TZ sp-basis, a Triple Zeta plus Double Polarization (TZ2P) type basis is obtained.

One disadvantage of all energy optimized basis sets is the fact that they primarily depend on the wave function in the region of the inner shell electrons. However, chemistry is mainly dependent on the valence electrons. Furthermore, many properties (for example polarizability) depend mainly on the wave function "tail" (far from the nucleus), which energetically is unimportant. Therefore, energy optimized basis sets are usually augmented explicitly with diffuse functions (basis functions with small exponents).

The fact that many basis functions go into describing the energetically important, but chemically unimportant, core electrons is the foundation for contracted basis sets.

Combining the full set of basis functions, known as the primitive GTOs (PGTOs), into a smaller set of functions by forming fixed linear combinations is known as basis set contraction, and the resulting functions are called contracted GTOs (CGTOs).

$$\chi(CGTO) = \sum_{i}^{k} a_i \, \chi_i(PGTO) \tag{2.25}$$

The previously introduced acronyms DZP, TZ2P etc., refer to the number of contracted basis functions. Contraction is especially useful for orbitals describing the inner (core) electrons, since they require a relatively large number of functions for representing the wave function cusp near the nucleus, and furthermore are largely independent of the environment. Contracting a basis set will always increase the energy, since it is a restriction of the number of variational parameters, and makes the basis set less flexible, but will also reduce the computational cost significantly. The decision is thus how much loss in accuracy is acceptable compared to the gain in computational efficiency.

2.6 The Hartree-Fock method

The Hartree-Fock (HF) method was developed by Hartree and Fock soon after the birth of quantum mechanics [139]. Briefly, its aim is to iteratively find a set of molecular orbitals (MOs) that minimize the energy of the wave function ψ, approximated as a single Slater determinant. The minimization can safely be carried out due to the variational principle, until self-consistency is reached. The HF method finally invokes the ‘mean field’ approximation, which means that each electron does not experience other individual electrons, but a mean field from their charges. In other terms, HF method reduces an N-electron problem to N one-electron problems. The one-electron hamiltonian can be rewritten as the Fock operator $\hat{f}(i)$:

$$\hat{f}(i) = \overbrace{-\sum_{i=1}^{N} \frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_{i=1}^{N} \sum_{A=1}^{M} \frac{e^2 Z_A}{r_{iA}}}^{\hat{h}(i)} + \overbrace{\sum_{i=1}^{N} \sum_{j>i}^{N} \hat{J}_j(i) - \hat{K}_j(i)}^{v_{HF}(i)} \tag{2.26}$$

The core Hamiltonian $\hat{h}(i)$ collects the terms dependent only on electron *i*, i.e. kinetic and potential energy due to nucleus-electron interaction, while the HF potential $v_{HF}(i)$ involves the electron-electron interactions in an average potential. $\hat{J}_j(i)$ and $\hat{K}_j(i)$ are the Coulomb and exchange operators respectively. Here, the major approximation of Hartree-Fock can be seen;

the two electron repulsion operator from the original Hamiltonian is exchanged by a one-electron operator $v_{HF}(i)$ which describes the repulsion in average.

$\hat{J}_{\mathrm{j}}(i)$ and $\hat{K}_{\mathrm{j}}(i)$ are defined by the following equations:

$$\hat{J}_j(i) = \int \left|\chi_j(j)\right|^2 r_{ij}^{-1} d\boldsymbol{r}_j \tag{2.27}$$

$$\hat{K}_j(i)\chi_i(i) = \int \left[\chi_j^*(i) r_{ij}^{-1} \chi_i(j)\right] \chi_j(i)\, d\boldsymbol{r}_j \tag{2.28}$$

While $\hat{J}_{\mathrm{j}}(i)$ represents the ordinary electrostatic repulsion between electrons i and j, the exchange integral $\hat{K}_{\mathrm{j}}(i)\chi_i(i)$ has no classical representation and arises due to the antisymmetry of ψ. Using the Fock operator, one can construct a set of eigenvalue problems called the Hatree-Fock equations:

$$\hat{f}(i)\chi_i(i) = \varepsilon_i \chi_i(i) \tag{2.29}$$

where ε_i is the energy of the *i*-th orbital. As seen from Equations (2.27)-(2.29), $v_{HF}(i)$ is functionally dependent on the spin orbitals $\chi_i(i)$, which are not known a priori since they are solutions to Equation (2.29). Hence the iterative approach, referred to as a self-consistent-field or SCF procedure.

Defining each spin orbital as a basis set expansion (see above), the Hartree-Fock equations can be written as:

$$\hat{f}(i) \sum_{\mu}^{M} c_{\mu i} \phi_{\mu} = \varepsilon_i \sum_{\mu}^{M} c_{\mu i} \phi_{\mu} \tag{2.30}$$

Multiplying from the left by a specific basic set and integrating yields M equations in the atomic orbital basis, known as the Roothan-Hall equations, that can be collected in a matrix notation.

$$\boldsymbol{FC} = \boldsymbol{\varepsilon SC} \tag{2.31}$$

where F is the Fock matrix with the elements $F_{\mu\nu} = \langle \chi_\mu | \hat{f} | \chi_\nu \rangle$, S is an overlap matrix with the elements $S_{\mu\nu} = \langle \chi_\mu | \chi_\nu \rangle$ and C contains all basis set coefficients. The goal of the iterative calculation is now to find C such that the norm of the diagonal eigenvalue matrix $\boldsymbol{\varepsilon}$ is minimized. The HF ground state energy of a N-electron system can be written as:

$$E_{HF} = \sum_{i}^{N} \varepsilon_i - \frac{1}{2} \sum_{ij}^{N} \left(J_{ij} - K_{ij}\right) + V_{NN} \tag{2.32}$$

where $\varepsilon_i = \langle \chi_i | \hat{f}(i) | \chi_i \rangle$ is a diagonal element of F, and J_{ij} and K_{ij} represent the two-electron integrals $\langle \chi_i | \hat{J}(i) | \chi_j \rangle$ and $\langle \chi_j | \hat{K}(i) | \chi_j \rangle$, respectively. Note that the summation in the last equation runs over *i*'s and *j*'s, and hence contains unphysical coloumbic self-*interaction*. However the J_{ii} terms are exactly cancelled since $K_{ii} = J_{ii}$. As will be discussed in section 2.13, this is an important feature of HF theory

2.7 Limitations and failings of the HF approach: post-HF methods

Atoms as well as molecules can have an even or odd number of electrons. If the number of electrons is even and all of them are located in double occupied spatial orbitals φ_i, th compound is in a singlet state. Such systems are called closed-shell systems. Compounds with an odd number of electrons as well as compounds with single occupied orbitals, i.e. species with triplet or higher ground state, are called open-shell systems. These two types of systems correspond to two different approaches of the Hartree-Fock method. In the restricted HF-method (RHF), all electrons are considered to be paired in orbitals whereas in the unrestricted HF-method (UHF) this limitation is lifted totally. It is also possible to describe open-shell systems with a RHF approach where only the single occupied orbitals are excluded which is then called a restricted open-shell HF (ROHF) which is an approach closer to reality but also more complex and therefore less popular than UHF [140].

There are also closed-shell systems which require the unrestricted approach in order to get proper results. For instance, the description of the dissociation of H_2 (i.e. the behaviour at large internuclear distance), where one electron must be located at one hydrogen atom, can logically not be obtained by the use of a system which places both electrons in the same spatial orbital. Therefore the choice of method is always a very important point in HF calculations [139].

The size of the investigated system can also be a limiting factor for calculations. Kohn states a number of $M = p^5$ with $3 \leq p \leq 10$ parameters for a result with sufficient accuracy in the investigation of the H_2 system [141]. For a system with $N = 100$ (active) electrons the number of parameters rises to

$$M = p^{3N} = 3^{300} \;\; to \;\; 10^{300} \approx \; 10^{150} \;\; to \;\; 10^{300} \tag{2.33}$$

The latter equation states, that the minimization of the energy would have to be performed in a space of at least 10^{150} dimension which exceeds the computational possibilities nowadays by far. HF-methods are therefore restricted to systems with a small number of involved electrons ($N \approx 10$). Referring to the exponential factor in Equation (2.33) this limitation is sometimes called *exponential wall* [141].

The major limit of the Hartree-Fock method is however due to the description of the many electron wave function by a single Slater determinant. It treats the motions of individual electrons as independent of one another, since the instantaneous interactions between individual pairs of electrons are replaced by interactions between one electron and the average field created by all the other electrons (independent-particle model). The loss of flexibility causes electrons to get in each other's way to a greater extent than would actually be the case, leading to an overestimation of the overall electron repulsion energy. Hence, the energy obtained by HF calculations is always higher than the exact ground state energy, in accordance with the variational principle. The most accurate energy obtainable by HF-methods (E_{HF}) is called the Hartree-Fock-limiting energy [140]. The difference between E_{HF} and E_0, that would result from solution of the exact Schrödinger equation, is called correlation energy and can be denoted as

$$E_{corr}^{HF} = E_0 - E_{HF} \tag{2.34}$$

Despite the magnitude of E_{corr}^{HF} is typically only 1% of E_0, it is comparable to (and influential on) the energy of a typical chemical reaction. For example, in a N_2 molecule

$$E_{corr}^{HF} = 14.9\ eV < 0.001 \cdot E_0 \tag{2.35}$$

can have a huge influence. For instance, the experimental dissociation energy of the N_2 molecule is

$$E_{diss} = 9.9\ eV < E_{corr}^{HF} \tag{2.36}$$

which corresponds to a large contribution of the correlation energy to relative energies such as reaction energies which are of particular interest in quantum chemistry.

The electron correlation between electrons with different spins, not treated explicitly in the HF method, has to do with the so-called Coulomb hole. This hole is connected to the interaction of two unlike spins, while the so-called Fermi hole is connected to two like spins. As electrons are Fermi particles, two electrons with the same spin can not be present at the same time at the same position (Fermi hole). Conversely, two electrons with differing spin (one alpha, the other beta) can be present at the same time at the same place, but are less likely to do so due to electrostatic repulsion (Coulomb hole). As the wavefunction is built up of Slater determinants, the Fermi hole is correctly taken care of. But the Coulomb hole is in the Hartree-Fock theory only taken into account in an average way.

Many efforts have been made to include the correlation energy in the calculations in the so-called post-Hartree-Fock methodologies. Some of them seek to improve the results by mixing in virtual orbitals in the wave function. Two examples are the configuration interaction (CI) and coupled cluster (CC) methods. In CI the wave function is defined as a linear combination of excited Slater determinants while in CC it is defined as an excitation operator, expanded in a power series acting on the HF wave function. Conversely, the Møller-Plesset (MP) methods correct the HF energy adding correlation effects by means of the Rayleigh–Schrödinger perturbation theory.

Accurate energies can be obtained from these methods which, however, are computationally very expensive: the more accurate the method, the higher the computational cost. With standard ab initio methods, reaching a high accuracy (i.e. the so called chemical accuracy of about 1-2 kcal mol^{-1}) is very demanding and thus only relatively small molecular systems can be treated. A radically different approach to lower computational cost of molecular calculations is employed in methods based on Density Functional Theory (DFT), that also take into account the correlation energy.

2.8 Density Functional Theory: a new base variable

In the early 1990s the application of density functional theory (DFT) for the understanding of chemical problems provoked a remarkable revolution in quantum chemistry. The development of DFT methods opened up a new era where relatively large systems, containing transition metals as well, could be investigated quantum mechanically [142]. The success of DFT originates from the implicit consideration of the electron correlation and their moderate computational costs with respect to post-HF methods. Modern (hybrid) DFT has accuracy at least comparable to MP2 (second-order Møller-Plesset) while requiring a computational effort in line with a HF calculation [140]. Although formally exact, all known DFT protocols need to be fed with some empirical expressions, meaning they are not strictly ab initio methods although sometimes referred to as such.

In HF theory the N-electrons wave function is used as the central quantity. Although its knowledge (or the knowledge of its good approximation) allows to access to all information that can be known about a particular electronic state, it depends on 4N variables (three spatial and one spin for each of the N electrons) and cannot be probed directly through experiment. In DFT methods, the many-body interacting system is instead described by its electron sity $\rho(\boldsymbol{r})$, that only depends on 3 spatial variables, independently of the number of electrons. Therefore, while the complexity of a wave function increases with the number of the electrons, the electron density has the same number of variables, independently of the system size. Furthermore, unlike the wave function, the electron density of a molecule is a physical observable that can be obtained experimentally by X-ray crystallography. It can be also calculated by the normalized N-electrons wave function describing the system, according to the following equation, in which spin coordinates are neglected:

$$\rho(\boldsymbol{r}_1) = N \int d\boldsymbol{r}_2 \dots \int |\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_N)|^2 d\boldsymbol{r}_2 \dots d\boldsymbol{r}_N \tag{2.37}$$

$\rho(\boldsymbol{r})$ is a measure of the probability of finding a particular electron with arbitrary spin in the infinitesimal volume element $d\boldsymbol{r}_1$, while the remaining N-1 electrons have arbitrary positions in the rest of the space. $\rho(\boldsymbol{r})$ must fulfil two conditions; it must vanish when $\boldsymbol{r}$ tends to infinite, whereas has maxima at the positions of the nuclei, and integrate the total number of electrons in the system:

$$\begin{gathered} \rho(\boldsymbol{r} \to \infty) = 0 \\ \int \rho(\boldsymbol{r}_1) d\boldsymbol{r}_1 = N \end{gathered} \tag{2.38}$$

2.9 Forerunners to current DFT methods

The idea of using the electron density rather than the wave function as the variational quantity is almost as old as quantum chemistry; the first attempt to express the total energy of a system as a functional of the electron density comes from calculations made independently by

Enrico Fermi and P.A.M Dirac in the 1920s on an ideal electron gas, work now well-known as the Fermi-Dirac statistics. In independent work by Thomas [143] and Fermi [144] in 1927, atoms were modelled as systems with a positive potential (the nucleus) located in a uniform (homogenous) electron gas. This obviously unrealistic idealization, the Tomas-Fermi model, or with embellishments by Dirac, the Thomas-Fermi-Dirac model, gave surprisingly good results for atoms, but failed completely for molecules: it predicted all molecules to be unstable toward dissociation into their atoms.

The so-called $X\alpha$ method ($X=$ exchange, α is a parameter in the $X\alpha$ equation, see section 2.12.1), introduced in 1951 by Slater [140], gives much better results. It can be regarded as a more accurate version of the Thomas-Fermi model, and is probably the first chemically useful DFT method. The $X\alpha$ method, which was developed mainly for atoms and solids, has also been used for molecules. Theoretical foundation for modern DFT was however laid by Hohenberg and Kohn in 1964 [145]. DFT, as a tool of electronic structure calculations has received recognition by the award of the Nobel Prize in chemistry in 1998.

2.10 The Hohenberg-Kohn theorems

DFT was proven to be an exact theory for many-body systems by Hohenberg and Kohn, which demonstrated that the electronic energy of a given system is determined completely by the electron density $\rho(\boldsymbol{r})$. In other words, the energy of a system is a unique functional of its electron density, i.e. there exists a one-to-one correspondence between the electron density and the energy. Furthermore, Hohenberg and Kohn showed that, given an external potential, the ground state energy can be obtained by minimizing the energy functional, with respect to the electron density. When the density is the true ground state electron density, it minimizes the energy functional. The Hohenberg-Kohn theory is based upon two theorems [145].

A note on semantics: a function is a prescription for producing a number from a set of variables (coordinates). A functional is similarly a prescription for producing a number from a function, which in turn depends on variables. A wave function and the electron density are thus functions, while an energy depending on a wave function or an electron density is a functional. A function depending on a set of variables is denoted with parenthesis, $f(x)$, while a functional depending on a function is denoted with brackets, $F[f]$.

Theorem 1: For any system of multiple interacting particles in an external potential $V_{ext}(\boldsymbol{r})$, the potential $V_{ext}(\boldsymbol{r})$ is determined uniquely, except for a constant C, by the ground state particle density $\rho_0(\boldsymbol{r})$.

Proof of Theorem 1: It is assumed that two external potentials $V_{ext}(\boldsymbol{r})$ and $V'_{ext}(\boldsymbol{r})$, which differ by more than a constant, lead to the same ground-state electron density $\rho_0(\boldsymbol{r})$. Two different external potentials imply that the two Hamiltonian operators, $\hat{H}$ and $\hat{H}'$, and the corresponding ground-state wave-functions, ψ and ψ', are different. Since ψ' is not the ground-state wave-function of $\hat{H}$, it follows that:

$$\begin{aligned} &\langle\psi'|\hat{H}|\psi'\rangle > E_0 \\ &\langle\psi'|\hat{H}'|\psi'\rangle + \langle\psi'|\hat{H} - \hat{H}'|\psi'\rangle > E_0 \\ &E'_0 + \langle\psi'|\hat{V}_{ext} - \hat{V}'_{ext}|\psi'\rangle > E_0 \\ &E'_0 + \int [V_{ext}(\mathbf{r}) - V'_{ext}(\mathbf{r})]\rho(\mathbf{r})dr > E_0 \end{aligned} \tag{2.39}$$

Similarly, taking ψ as an approximate function of $\hat{H}'$ yields:

$$E_0 + \int [V_{ext}(\mathbf{r}) - V'_{ext}(\mathbf{r})]\rho(\mathbf{r})dr > E'_0 \tag{2.40}$$

By addition of Equations 2.39 and 2.40 the inequality $E'_0 + E_0 > E_0 + E'_0$ is obtained, which represent an inconsistency and therefore provides by *reduction ad absurdum* that there is a one-to-one correspondence between the electron density $\rho(\mathbf{r})$ and the external potential $\hat{V}_{ext}(\mathbf{r})$, and thereby also with the Hamiltonian operator and the energy. In the language of the DFT theory, the Hamiltonian and the energy are unique functional of the electron sity; $\hat{H} = \hat{H}[\rho]$ and $E_0 = E_0[\rho]$. From the first Hohenberg-Kohn theorem it is obvious that also the ground-state wave function is a unique functional of the ground-state electron density, $\psi_0 = \psi_0[\rho]$. This means that, if the ground-state electron density is known, all other properties of the many–body system can be uniquely determined.

Since the original publication of Hohenberg and Kohn, $\hat{H}[\rho]$ is resembled by the electronic Hamiltonian operator $\hat{H}_{el}$ introduced in Equation (2.14), with the one difference that the nuclear contribution potential $\hat{V}_{eN}$ is replaced by a general external potential $\hat{V}_{ext}$. Therefore, the Hamiltonian of the system is denoted as:

$$\hat{H}[\rho] = \hat{T}_e + \hat{V}_{ext} + \hat{V}_{ee} \tag{2.41}$$

Within the Born-Oppenheimer approximation, the operators $\hat{T}_e$ and $\hat{V}_{ee}$ depend only on the number of electrons and therefore are the same for any N-electron system. Since that $\rho(\mathbf{r})$ determines the number of electrons in the system, it follows that the kinetic energy T_e and the electron-electron repulsion E_{ee} are functionals of the electron density $\rho(\mathbf{r})$. $\hat{V}_{ext}$ is instead system dependent. Therefore, the total energy can be written as:

$$E[\rho] = \langle\psi|\hat{T}_e + \hat{V}_{ext} + \hat{V}_{ee}|\psi\rangle = F_{HK}[\rho] + \int V_{ext}(\mathbf{r})\rho(\mathbf{r})dr \tag{2.42}$$

where $F_{HK}[\rho]$ is defined as $F_{HK}[\rho] = T_e[\rho] + E_{ee}[\rho]$. Since $T_e[\rho]$ and $E_{ee}[\rho]$ do not depend on the external potential which represents the particular system of interest, $F_{HK}[\rho]$ is a universal functional, which is identical for every system. However, the explicit forms of $T_e[\rho]$ and $E_{ee}[\rho]$, and therefore that of $F_{HK}[\rho]$, lie completely in the dark.

Theorem 2: An Universal functional for energy $E[\rho]$ in terms of the density $\rho(\boldsymbol{r})$ can be defined, valid for any external potential $V_{ext}(\boldsymbol{r})$. For any particular $V_{ext}(\boldsymbol{r})$, the exact ground state energy of the system is the global minimum value of this functional, and the density $\rho(\boldsymbol{r})$ that minimizes the functional is the exact ground state density $\rho_0(\boldsymbol{r})$.

Proof of Theorem 2: The second theorem has been proved by variation calculus. According to variational principle, for any wave function ψ', the energy functional $E[\psi']$:

$$E[\psi'] \equiv \langle\psi'|\hat{H}|\psi'\rangle = \langle\psi'|\hat{T}_e + \hat{V}_{ext} + \hat{V}_{ee}|\psi'\rangle \tag{2.43}$$

has its global minimum value only when ψ' is the ground-state wave function ψ_0, with the constraint that the total number of the particles is conserved. According to the first Hohenberg-Kohn theorem, ψ' must correspond to a ground state with electron density $\rho'(\boldsymbol{r})$ and external potential $V'_{ext}(\boldsymbol{r})$, then $E[\psi']$ is a functional of $\rho'(\boldsymbol{r})$. It is follow that:

$$\begin{aligned} &E[\psi'] = E[\rho'(\boldsymbol{r})] = \int V'_{ext}\rho'(\boldsymbol{r})d\boldsymbol{r} + F_{HK}[\rho'(\boldsymbol{r})] > E[\psi_0] \\ &E[\psi_0] = E[\rho_0(\boldsymbol{r})] = \int V'_{ext}\rho_0(\boldsymbol{r})d\boldsymbol{r} + F_{HK}[\rho_0(\boldsymbol{r})] \\ &E[\rho'(\boldsymbol{r})] > E[\rho_0(\boldsymbol{r})] \end{aligned} \tag{2.44}$$

Thus the energy functional $E[\rho(\boldsymbol{r})]$ evaluated for the correct ground-state density $\rho_0(\boldsymbol{r})$ is indeed lower than the value of this functional for any other density $\rho(\boldsymbol{r})$. Therefore by minimizing the total energy functional of the system with respect to variations in the density $\rho(\boldsymbol{r})$, the exact ground-state density and energy.

Recapitulating, it has been shown that density functional theory provides a clear and mathematical exact framework for the use of the electron density as base variable. Nevertheless, nothing of what has been derived is of practical use. Or in other words, the Hohenberg-Kohn theorems, as important as they are, do not provide any help for the calculation of molecular properties and also don't provide any information about approximations for $F_{HK}[\rho]$ functionals. In the direct comparison to the variational approach of the Hartree-Fock method, the variational principle introduced in the second theorem of Hohenberg and Kohn is even more tricky. Whereas in wave-function based approaches like Hartree-Fock the obtained energy value provides information about the quality of the trial wave function (the lower energy, the better the wave function), this is not the case in the variational principle based on the electron density. More than that, it can even happen that some functionals provide energies lower than the actual ground state energy in particular calculations.

2.11 The Kohn-Sham equations

The Hohenberg-Kohn theorems imply the existence of a universal functional $F_{HK}[\rho]$, but says nothing about its form nor how it can be constructed, since the explicit forms of $T_e[\rho]$ and $E_{ee}[\rho]$ are unknown. It is the procedure proposed by Kohn and Sham in 1965 [146] that puts the Hohenberg-Kohn theorems into practical use. Although the strength of DFT is that only the total density needs to be considered, it is the orbitals reintroduction in their procedure that helped find the universal functional $F_{HK}[\rho]$.

The main idea in the Kohn-Sham (KS) approach is splitting the unknown kinetic energy functional $T_e[\rho]$ into two terms, one of which can be calculated exactly, and a small correction term. For this purpose, similarly to HF theory, a non-interacting reference system is introduced as a route to find the ground-state density $\rho_0(\boldsymbol{r})$ and energy E_0. It is possible assume the Hamiltonian operator given by Equation (2.41) as:

$$\hat{H}_\lambda = \hat{T}_e + \hat{V}_{ext}(\lambda) + \lambda\,\hat{V}_{ee} \tag{2.45}$$

where $0 \le \lambda \le 1$ and the external potential $\hat{V}_{ext}(\lambda)$ is adjusted so that the same density $\rho(\boldsymbol{r})$ is obtained for both $\lambda = 1$ (the real system) and $\lambda = 0$ (a hypothetical system with non-interacting electrons). For the $\lambda = 0$ case the exact solution to the Schrödinger equation is given as a Slater determinant for which the exact kinetic energy T_S (the subscript S denotes that it is calculated from a Slater determinant) is given as:

$$T_S = \sum_{i=1}^{N} \left\langle \phi_i^{KS} \middle| -\frac{1}{2}\nabla^2 \middle| \phi_i^{KS} \right\rangle \tag{2.46}$$

where ϕ_i^{KS}, called Kohn-Sham (KS) orbitals, are the spatial functions of the spin orbitals composing the Slater determinant. If instead $\lambda = 1$, electrons are interacting and using a single determinant neglects the electron correlation. Therefore, for a real system Equation (2.46) is only an approximation and does not provide the total kinetic energy. However, just as HF theory provide ~99% of the correct answer, the difference between the exact kinetic energy and that calculated by assuming non-interacting orbitals is small. The remaining kinetic energy can be absorbed into exchange-correlation term $E_{xc}[\rho]$, and the expression of $F_{HK}[\rho]$ can be rewritten as:

$$F_{HK}[\rho] = T_e[\rho] + E_{ee}[\rho] = T_e[\rho] + J[\rho] + E_{ncl}[\rho] = T_S + J[\rho] + E_{xc}[\rho] \tag{2.47}$$

where $E_{ee}[\rho]$ is divided into the classical and non-classical contributions to the electron-electron interaction, $J[\rho]$ and $E_{ncl}[\rho]$, respectively. The exchange-correlation energy $E_{xc}[\rho]$ contains everything that is unknown:

$$E_{xc}[\rho] = (T[\rho] - T_S) + (E_{ee}[\rho] - J[\rho]) \tag{2.48}$$

The first parenthesis may be considered the kinetic correlation energy, which is not covered by the non-interacting system, while the second contains both exchange and potential correlation energy, i.e. the non-classical electrostatic interaction energy $E_{ncl}[\rho]$.

Hence, the total energy expression of the real interacting system can be written as:

$$
\begin{aligned}
E[\rho] &= T_S + J[\rho] + E_{xc}[\rho] + E_{Ne}[\rho] \\
&= \sum_{i=1}^{N} \left\langle \phi_i^{KS} \middle| -\frac{1}{2}\nabla^2 \middle| \phi_i^{KS} \right\rangle + \frac{1}{2}\iint \frac{\rho(\boldsymbol{r}_i)\rho(\boldsymbol{r}_j)}{\boldsymbol{r}_{ij}} d\boldsymbol{r}_i d\boldsymbol{r}_j + E_{xc}[\rho] + \sum_{A=1}^{M} \int \frac{Z_A \rho(\boldsymbol{r})}{\boldsymbol{r}_{iA}} dr
\end{aligned}
\tag{2.49}
$$

where all terms are functionals of the electron density $\rho(\boldsymbol{r})$. The only functional for which no explicit form can be given is $E_{xc}[\rho]$. Also T_S is unknown, but it can be obtained exactly from Equation (2.46). Therefore, it is not required to express it as functional of the electron density. The electron-electron and electron-nuclei Coulomb interactions, $J[\rho]$ and $E_{Ne}[\rho]$, are instead defined by their classical expressions.

On the other hand, also the electron density $\rho(\boldsymbol{r})$ of the real system is unknown. However, it can be defined under the assumption that it is the same as that of the non-interacting reference system, for which the wave function is a Slater determinant. It is follows that:

$$
\rho(\boldsymbol{r}) = \rho_S(\boldsymbol{r}) = \sum_{i=1}^{N} \left|\phi_i^{KS}(\boldsymbol{r}_i)\right|^2 \tag{2.50}
$$

Substituting the above expression into Equation (2.49) leads to Kohn-Sham functional energy $E^{KS}[\rho]$, in which all contribution terms are given in terms of the one-electron orbitals ϕ_i^{KS}. The wave functions ϕ_i^{KS} which minimize $E^{KS}[\rho]$ satisfy the following set of one-electron eigenvalue equations, known as Kohn-Sham equations:

$$
\hat{h}_i^{KS}\phi_i^{KS} = \varepsilon_i^{KS}\phi_i^{KS} \quad i = 1,2,\dots,N \tag{2.51}
$$

where ε_i^{KS} are the KS energy levels and $\hat{h}_i^{KS}$ is the Kohn-Sham operator, given by:

$$
\hat{h}_i^{KS} = -\frac{1}{2}\nabla_i{}^2 + \hat{V}_{class} + \hat{V}_{xc} + \hat{V}_{Ne} = -\frac{1}{2}\nabla_i{}^2 + \int \frac{\rho(\boldsymbol{r}_i)}{\boldsymbol{r}_{ij}} d\boldsymbol{r}_i + \frac{\delta E_{xc}[\rho(\boldsymbol{r})]}{\delta\rho(\boldsymbol{r})} + \sum_{A}^{M} \frac{Z_A}{\boldsymbol{r}_{iA}} \tag{2.52}
$$

where $\hat{V}_{class}$ and $\hat{V}_{Ne}$ are the potentials corresponding to the classical electron-electron and electron-nuclei Coulomb interactions, while $\hat{V}_{xc}$ is the exchange-correlation energy potential defined as the functional derivative of $E_{xc}[\rho(\boldsymbol{r})]$ with respect to $\rho(\boldsymbol{r})$.

Similarly to the HF method, the unknown KS orbitals ϕ_i^{KS} can be expanded in a set of basis functions according to the LCAO approximation.

$$
\phi_i^{KS} = \sum_{s=1}^{m} c_{si}\vartheta_s \qquad i = 1,2,\dots,m \tag{2.53}
$$

Substituting the basis set expansion into the KS equations, multiplying by a specific basic set and integrating yields a set of equations in the atomic orbital basis, which can be collected into a single matrix equation analogous to the Equation (2.31).

$$
\boldsymbol{H}^{KS}\boldsymbol{C} = \boldsymbol{\varepsilon}\boldsymbol{S}\boldsymbol{C} \tag{2.54}
$$

The key to solving the KS equations then becomes, as in the standard HF method, the calculations of elements and diagonalization of the resultant matrix $\boldsymbol{H}^{KS}$.

Since $\hat{h}_i^{KS}$ depends on $\rho(\boldsymbol{r})$ through the $\hat{V}_{xc}$ potential, the Kohn-Sham equations have to be solved self-consistently. The iterative procedure in determining the density that minimizes $E[\rho]$, begins with an initial guess of the density. This guess is usually a non-interacting atoms guess, obtained by the summing the electron densities of the individual atoms of the molecule, at the molecular geometry. From this guess, it is possible to calculate the $\boldsymbol{H}^{KS}$ elements $H_{rs}^{KS} = \langle \vartheta_r | \hat{h}^{KS} | \vartheta_s \rangle$ and obtain initial guesses of the coefficients in the basis expansion of Equation (2.53). These coefficients are used in Equation (2.51) to calculate a better density function, which is used to calculated improved matrix elements H_{rs}^{KS}, that in turn give improved coefficients. The self-consistent procedure is iterated a number of times until the convergence of the density is reached. The final density and final KS orbitals are used to calculate the energy from Equation (2.49).

2.12 The exchange-correlation energy functional

Although it is clear that there are many similarities between HF theory and DFT, there is an important difference. If the exact $E_{xc}[\rho]$ functional was known, DFT provide the exact total energy. DFT methods have the potential of including the correlation energy, at a computational effort similar to that for determining the uncorrelated HF energy. Although this is certainly the case for approximations to $E_{xc}[\rho]$, this is not necessarily true for the exact $E_{xc}[\rho]$. It may well be that the exact $E_{xc}[\rho]$ functional is so complicated that the computational effort for solving the KS equations will be similar to that required for solving the Schrödinger equation (exactly). Devising good approximations to $E_{xc}[\rho]$functional is therefore the main challenge for DFT.

It is possible to prove that the exchange-correlation energy potential $\hat{V}_{xc}$ is unique functional, valid for all systems, but an explicit functional form of this functional has been elusive. The difference between DFT methods is the choice of the functional form of the exchange-correlation energy. There is little guidance from theory how such functionals should be chosen, and consequently many different potentials have been proposed. Functional forms are often designed to have a certain limiting behaviour (e.g. including the uniform gas limit), and fitting parameters to known accurate data. Which functional is the better will have to be settled by comparing the performance with experiments or high-level wave mechanics calculations.

It is customary to separate $E_{xc}[\rho]$ into two parts, a pure exchange $E_x[\rho]$ and a correlation part $E_c[\rho]$, although it is not clear that this a valid assumption. Each of these energies is often written in terms of the energy per particle (energy density), ε_x and ε_c:

$$E_{xc}[\rho] = E_x[\rho] + E_c[\rho] = \int \rho(\boldsymbol{r})\, \varepsilon_x[\rho(\boldsymbol{r})] d\boldsymbol{r} + \int \rho(\boldsymbol{r})\, \varepsilon_c[\rho(\boldsymbol{r})] d\boldsymbol{r} \qquad (2.55)$$

The corresponding potential, given as the derivative of the energy with respect to the density, results:

$$\hat{\boldsymbol{V}}_{xc}(\boldsymbol{r}) = \frac{\delta E_{xc}[\rho(\boldsymbol{r})]}{\delta \rho(\boldsymbol{r})} = \varepsilon_{xc}[\rho(\boldsymbol{r})] + \rho(\boldsymbol{r}) \frac{\delta \varepsilon_{xc}(\boldsymbol{r})}{\delta(\boldsymbol{r})} \tag{2.56}$$

The exchange energy is "by definition" given as a sum of contributions from the α and β spin density, as exchange energy only involves electrons of the same spin. The correlation between electrons of parallel spin is different from that between electrons of opposite spin.

$$E_x[\rho] = E_x^{\alpha}[\rho_\alpha] + E_x^{\beta}[\rho_\beta] \tag{2.57}$$

$$E_c[\rho] = E_x^{\alpha\alpha}[\rho_\alpha] + E_x^{\beta\beta}[\rho_\beta] + E_x^{\alpha\beta}[\rho_\alpha\, , \rho_\beta] \tag{2.58}$$

The total density is the sum of the α and β contributions, $\rho = \rho_\alpha + \rho_\beta$, and for a closed shell singlet these are identical ($\rho_\alpha = \rho_\beta$). Functionals for the exchange and correlation energies may be formulated in terms of separate spin-densities; however they are often given instead as functions of the spin polarization ζ (normalized difference between ρ_α and ρ_β), and the radius of the effective volume containing one electron, r_S, also known as the Wigner-Seitz radius (radius of a sphere with constant ρ and a total charge of one electron):

$$\zeta = \frac{\rho_\alpha - \rho_\beta}{\rho_\alpha + \rho_\beta} \tag{2.59}$$

$$r_S = \left(\frac{4\pi\rho}{3}\right)^{-1/3} \tag{2.60}$$

In the formulas below it is implicitly assumed that the exchange energy is a sum over both α and β densities.

2.12.1 Local density methods

The simplest approximation to $E_{xc}[\rho(\boldsymbol{r})]$ is within the framework of the *Local Density Approximation* (LDA); this assumes that the local density can be treated as a uniform electron gas, or equivalently that the electron density varies only very slowly with position. The exchange energy for a uniform gas is given by the Dirac formula:

$$E_x^{LDA}[\rho] = -C_x \int \left[\rho^{4/3}(\boldsymbol{r})\right] d\boldsymbol{r} \quad \text{where} \quad C_x = \frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3} \tag{2.61}$$

$$\varepsilon_x^{LDA}[\rho] = -C_x \rho^{1/3}(\boldsymbol{r}) \tag{2.62}$$

In the general more case, where the α and β densities are not equal, LDA (where the sum of the α and β densities is raised by to the 4/3 power) has been virtually abandoned and replaced by the *Local Spin Density Approximation* (LSDA), which is given as the sum of the individual densities raised to the 4/3 power:

$$E_x^{LSDA}[\rho] = -2^{1/3}C_x \int \left[\rho_\alpha^{4/3}(\boldsymbol{r}) + \rho_\beta^{4/3}(\boldsymbol{r})\right] d\boldsymbol{r} \tag{2.63}$$

$$\varepsilon_x^{LSDA}[\rho] = -2^{1/3}C_x \left[\rho_\alpha^{1/3}(\boldsymbol{r}) + \rho_\beta^{1/3}(\boldsymbol{r})\right] \tag{2.64}$$

LSDA may also be written in terms of the total density and spin polarization:

$$\varepsilon_x^{LSDA}[\rho] = -\frac{1}{2}C_x\rho^{1/3}(\boldsymbol{r})\left[(1+\zeta)^{4/3} + (1-\zeta)^{4/3}\right] \tag{2.65}$$

For closed shell systems LSDA is equal to LDA, and since it is the most common case, LDA is often used interchangeably with LSDA, although it is not true in the general case. The $X\alpha$ method of Slater (section 2.9), developed before the KS approach, can be considered as an LDA method where the correlation energy is neglected and the exchange term is given as:

$$\varepsilon_{X\alpha}[\rho] = -\frac{3}{2}\alpha C_x\rho^{1/3}(\boldsymbol{r}) \tag{2.66}$$

With $\alpha = 2/3$ this is identical to the Dirac expression. The original $X\alpha$ method used $\alpha = 1$, but a value of $3/4$ has been shown to give better agreement for atomic and molecular systems.

The correlation energy of a uniform electron gas has been determined by Monte Carlo methods for a number of different densities. In order to use these results in DFT calculations, it is desirable to have a suitable analytic interpolation formula. This has been constructed by Vosko, Wilk and Nuisar (VWN) [147] and is in general considered to be a very accurate fit. It interpolates between the unpolarized ($\zeta = 0$) and spin polarized ($\zeta = 1$) limits by the following functional:

$$\varepsilon_c^{VWN}(r_S,\zeta) = \varepsilon_c(r_S,0) + \varepsilon_a(r_S)\left[\frac{f(\zeta)}{f''(\zeta)}\right][1-\zeta^4] + [\varepsilon_c(r_S,1) - \varepsilon_c(r_S,0)]f(\zeta)\zeta^4 \tag{2.67}$$

$$f(\zeta) = \frac{(1+\zeta)^{4/3} + (1-\zeta)^{4/3} - 2}{2(2^{1/3}-1)} \tag{2.68}$$

The $\varepsilon_c(r_S,\zeta)$ and $\varepsilon_a(r_S)$ functionals are parameterized as:

$$\begin{aligned}\varepsilon_{c/a}(r_S,\zeta) = A\Bigg\{ & ln\frac{x^2}{X(x)} + \frac{2b}{Q}tan^{-1}\left(\frac{Q}{2x+b}\right) - \\ & -\frac{\vartheta x_0}{X(x_0)}\left[ln\frac{(x-x_0)^2}{X(x)} + \frac{2(b+2x_0)}{Q}tan^{-1}\left(\frac{Q}{2x+b}\right)\right]\Bigg\}\end{aligned} \tag{2.69}$$

where $x = \sqrt{r_S}$, $X(x) = x^2 + bx + c$ and $Q = \sqrt{4c - b^2}$. The parameters A, x_0, b and c are fitting constants, different for $\varepsilon_c(r_S,0)$, $\varepsilon_c(r_S,1)$ and $\varepsilon_a(r_S)$. A modified form for $\varepsilon_{c/a}(r_S)$ has given by Perdew and Wang [148] and is used in connection with the PW91 functional described in section 2.12.2.

$$\varepsilon_{c/a}^{PW91}(x) = -2a\rho(1+\alpha x^2)\,ln\left[1 + \frac{1}{2a(\beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 x^2)}\right] \tag{2.70}$$

Here, a, $\alpha, \beta_1, \beta_2, \beta_3$ and β_4 are suitable constants.

The LSDA approximation in general underestimates the exchange energy by ~10%, thereby creating errors which are larger than the whole correlation energy. Electron correlation is furthermore overestimated, often by a factor close to 2, and bond strengths are as a consequence overestimated. Despite the simplicity of the fundamental assumption, LSDA methods are often fund to provide results with accuracy similar to that obtained by HF methods.

2.12.2 Gradient corrected methods

Improvements over the LSDA approach have to consider a non-uniform electron gas. A step in this direction is to make the exchange and correlation energies dependent not only on the electron density, but also on derivatives of the density. Such methods are known as *Gradient Corrected* or *Generalized Gradient Approximation* (GGA) methods. They are also sometimes referred as *non-local* methods, although this is somewhat misleading since the functionals depend only on the density (and derivatives) at a given point, not on a space volume as for example the Hartree-Fock exchange energy.

Perdew and Wang (PW86) [149] proposed modifying the LSDA exchange expression to that shown in Equation (2.71), where x is a dimensionless gradient variable, and a, b and c being suitable constants (summation over equivalent expressions for the α and β densities is implicitly assumed).

$$\varepsilon_x^{PW86} = \varepsilon_x^{LDA}(1 + ax^2 + bx^4 + cx^6)^{1/15} \ , \ x = -\frac{|\nabla\rho|}{\rho^{4/3}} \tag{2.71}$$

Becke [150] proposed a widely used correction (B or B88) to the LSDA exchange energy, which has the correct $-r^{-1}$ asymptotic behaviour for the energy density (but not for the exchange potential).

$$\varepsilon_x^{B88} = \varepsilon_x^{LDA} + \Delta\varepsilon_x^{B88} \ , \ \Delta\varepsilon_x^{B88} = -\beta\rho^{1/3}\frac{x^2}{1 + 6\beta x \, sinh^{-1} x} \tag{2.72}$$

The β parameter is determined by fitting to known atomic data and x is defined in Equation (2.71).

Perdew and Wang have proposed an exchange functional similar to B88 to be used in connection with the PW91 correlation functional given below (Equation (2.76)).

$$\varepsilon_x^{PW91} = \varepsilon_x^{LDA}\left(\frac{1 + a_1 \, sinh^{-1}(xa_2) + (a_3 + a_4 e^{-bx^2})x^2}{1 + xa_1 \, sinh^{-1}(xa_2) + a_5 x^2)}\right) \tag{2.73}$$

where a_{1-5} and b again are suitable constants and x is defined in Equation (2.71).

There have been various gradient corrected functional forms proposed for the correlation energy. One popular functional (not a correction) is due to Lee, Yang and Parr [151] (LYP) and has the form:

$$
\begin{aligned}
\varepsilon_c^{LYP} = -a\frac{\gamma}{\left(1+d\rho^{-\frac{1}{3}}\right)} - ab\frac{\gamma e^{-c\rho^{-1/3}}}{9\left(1+d\rho^{-\frac{1}{3}}\right)\rho^{8/3}}\Big[18\big(2^{2/3}\big)C_F\left(\rho_\alpha^{8/3}+\rho_\beta^{8/3}\right) - \\
-18\rho t_w + \rho_\alpha(2t_w^\alpha + \nabla^2\rho_\alpha) + \rho_\beta(2t_w^\beta + \nabla^2\rho_\beta)\Big] \\
\gamma = 2\left[\frac{\rho_\alpha^2+\rho_\beta^2}{\rho^2}\right] \quad , \quad t_w^\sigma = \frac{1}{8}\left(\frac{|\nabla\rho_\sigma|^2}{\rho_\sigma} - \nabla^2\rho_\sigma\right)
\end{aligned}
\tag{2.74}
$$

where the a, b, c and d parameters are determined by fitting to data for the helium atom. The t_w functional is known as the local Weizsacker kinetic energy density. Note that the γ-factor becomes zero when all the spin are aligned $(\rho = \rho_\alpha, \rho_\beta = 0)$, i.e. the LYP functional does not predict any parallel spin correlation in such case (e.g. the LYP correlation energy in triplet He is zero).

Perdew proposed a gradient correction to the LSDA result. It appeared in 1986 [149] and is known by the acronym B86.

$$
\begin{aligned}
\varepsilon_c^{P86} = \varepsilon_c^{LDA} + \Delta\varepsilon_c^{P86} \quad , \quad \Delta\varepsilon_c^{P86} = \frac{e^{\Phi}C(\rho)|\nabla\rho|^2}{f(\zeta)\rho^{7/3}} \quad , \quad \Phi = a\frac{C(\infty)|\nabla\rho|}{C(\rho)\rho^{\frac{7}{6}}} \\
f(\zeta) = 2^{1/3}\sqrt{\left(\frac{1+\zeta}{2}\right)^{5/3} + \left(\frac{1-\zeta}{2}\right)^{5/3}} \quad , \quad C(\rho) = c_1\frac{c_2 + c_3 r_s + c_4 r_s^2}{1 + c_5 r_s + c_6 r_s^2 + c_7 r_s^3}
\end{aligned}
\tag{2.75}
$$

where the a and c_{1-7} are numerical constants. This functional was later modified to the following form (also a correction of to the LSDA energy) by Perdew and Wang in 1991 (PW91 or P91) [152]:

$$
\begin{aligned}
\varepsilon_c^{PW91}[\rho] = \rho(H_0(t, r_s, \zeta) + H_1(t, r_s, \zeta)) \\
H_0(t, r_s, \zeta) = b^{-1}f(\zeta)^3\, ln\left[1 + a\frac{t^2 + At^4}{1 + At^2 + A^2t^4}\right] \\
H_1(t, r_s, \zeta) = \frac{16}{\pi}(3\pi^2)^{1/3}[C(\rho) - c]f(\zeta)^3 t^2 e^{-dx^2/f(\zeta)^2} \\
f(\zeta) = \frac{1}{2}\left((1+\zeta)^{\frac{2}{3}} + (1-\zeta)^{\frac{2}{3}}\right) \\
t = \left(\frac{192}{\pi^2}\right)^{1/6}\frac{|\nabla\rho|}{2f(\zeta)\rho^{7/6}} \quad , \quad A = a\left[e^{-b\varepsilon_c(r_s,\zeta)/f(\zeta)^3} - 1\right]^{-1}
\end{aligned}
\tag{2.76}
$$

where $\varepsilon_c(r_s, \zeta)$ is the PW91 parameterization of the LSDA correlation energy functional (Equation (2.73)), x and $C(\rho)$ are as defined in Equations (2.71) and (2.75), and a, b, c and d are suitable constants.

It should be noted that several of the proposed functionals violate fundamental restrictions, such as predicting correlation energy for one-electron systems (for example P86 and PW91) or failing to have exchange energy cancel the Coulomb self-repulsion. One of the more recent functionals which does not have these problems is due to Becke (B95) [153] which has the form:

$$
\begin{gathered}
\varepsilon_c^{B95} = \varepsilon_c^{\alpha\beta} + \varepsilon_c^{\alpha\alpha} + \varepsilon_c^{\beta\beta} \\
\varepsilon_c^{\alpha\beta} = \left[1 + a(x_\alpha^2 + x_\beta^2)\right]^{-1} \varepsilon_c^{PW91,\alpha\beta} \ , \ \varepsilon_c^{\sigma\sigma} = [1 + b x_\sigma^2]^{-2} \frac{D_\sigma}{D_\sigma^{LDA}} \varepsilon_c^{PW91,\sigma\sigma} \\
D_\sigma = \sum_i^N |\nabla \phi_i|^2 - \frac{(\nabla \rho_\sigma)^2}{4\rho_\sigma} \ , \ D_\sigma^{LDA} = 2^{5/3} C_F \rho_\sigma^{5/3}
\end{gathered}
\tag{2.77}
$$

Here σ runs over α and β spins, x_σ is defined in Equation (2.71), a and b are fitting parameters and ε_c^{PW91} is the Perdew-Wang parameterization of the LSDA correlation functional (Equation(2.71)).

2.12.3 Hybrid methods

Since the E_x is a much larger contribution to E_{xc} than E_c, it would seem loical to use the formally exact Hartree-Fock exchange energy, defined in Equation (2.28). However due to discrepancies in how HF and KS theory define exchange, a full HF treatment of E_x in DFT yields awful results [140]. Instead, much effort has been devoted to develop *hybrid functionals*, employing linear combination of HF and KS exchange.

From the Hamiltonian in Equation (2.45) and the definition of the exchange-correlation energy in Equation (2.48), an exact connection can be made between the exchange-correlation energy and the corresponding potential connecting the non-interacting reference and the actual system. The resulting equation is called *Adiabatic Connection Formula* (ACF) [154] and involves an integration over the parameter λ which "turns on" the electron-electron interaction.

$$
E_{xc} = \int_0^1 \langle \Psi_\lambda | \hat{V}_{xc}(\lambda) | \Psi_\lambda \rangle \, d\lambda \tag{2.78}
$$

In the crudest approximation, taking $\hat{V}_{xc}(\lambda)$ to be linear in λ, the integral is given as the average of the values at the two-end points.

$$
E_{xc} \cong \frac{1}{2} \langle \Psi_0 | \hat{V}_{xc}(0) | \Psi_0 \rangle + \frac{1}{2} \langle \Psi_1 | \hat{V}_{xc}(1) | \Psi_1 \rangle \tag{2.79}
$$

In the $\lambda = 0$ limit, the electrons are non-interacting and there is consequently no correlation energy, only exchange energy. Furthermore, since the exact wave function in this case is a single Slater determinant composed of KS orbitals, the exchange energy is exactly that given by Hatree-Fock theory (Equation (2.32)). If the KS orbitals are identical to the HF orbitals, the "exact" exchange is precisely the exchange energy calculated by HF methods. The last term in Equation (2.79) is still unknown. Approximating it by the LSDA result defines the *Half-and-Half* (H+H) method [155].

$$
E_{xc}^{H+H} = \frac{1}{2} E_x^{HF} + \frac{1}{2} (E_x^{LSDA} + E_c^{LSDA}) \tag{2.80}
$$

Since the GGA methods give a substantial improvements over LDA, a generalized version of the Half-and-Half method may be defined by writing the exchange energy as a suitable combination of LSDA, exact exchange and a gradient correction term. The correlation energy may similarly be taken as the LSDA formula plus a gradient correction term.

$$E_{xc}^{B3} = (1-a)E_x^{LSDA} + aE_x^{HF} + b\Delta E_x^{B88} + (1-c)E_c^{LSDA} + cE_c^{GGA}) \tag{2.81}$$

Adiabatic Connection Model (ACM) and *Becke 3 parameter functional* (B3) are examples of hybrid models which include exact exchange. The a, b and c parameters are determined by fitting to experimental data and depend on the term chosen for E_c^{GGA}. Owing to the substantially better performance of such parameterized functionals the Half-and-Half model is rarely used anymore.

The most well-known B3 gradient-corrected hybrid method was first proposed by Becke in 1993 [156] and was a three parameter mix of B88 and HF exchange together with a PW91 correlation. Its most employed form uses VWN LDA and LYP GGA correlation instead of PW91 and is denoted B3LYP [157]:

$$E_{xc}^{B3LP} = (1-a)E_x^{LSDA} + aE_x^{HF} + b\Delta E_x^{B88} + (1-c)E_c^{VWN} + cE_c^{LYP}) \tag{2.82}$$

with the empirical coefficients a, b and c fitted 0.20, 0.72 and 0.81, respectively. The B3LYP hybrid functional thus contains 20% Hartree-Fock exchange.

2.13 Advantages and limits of DFT

The number of applications of DFT is growing rapidly by the year and some of the latest and most significant studies include the understanding and design of catalytic processes in enzymes and zeolites, electron transport, solar energy harvesting and conversion, drug design in medicine, as well as many other problems in science and technology.

The story behind the success of DFT is the search for the exchange-correlation functional $E_{xc}[\rho]$ that uses the electron density to describe the intricate many-body effects within a single particle formalism. If the exchange-correlation functional that is used was exact, then DFT would correctly describe the quantum nature of matter. However, in practice the explicit form of $E_{xc}[\rho]$ is never known. It is the approximate nature of the exchange-correlation functional that is the reason both for the success and the failure of DFT applications [140, 141].

The greatest advantage of DFT is certainly its low computational costs. At a cost lower than Hartree-Fock calculations, DFT includes at least partly electron correlation and is therefore in principle more accurate. Hartree-Fock and DFT formally scale with the size of the treated system N of the order of the fourth and third power, N^4 and N^3, respectively. When electron correlation is considered in Post-Hartree-Fock approaches the computation time increases to N^5 for second-order Møller-Plesset perturbation theory, and N^7 for quadratic configuration interaction (QCI) with single and double excitations and perturbatively added triples (CISD(T)) and coupled-cluster (CCSD(T)). This formal scaling behaviour reduces quickly for larger systems. The last two methods are the most accurate one but their demands make them hardly feasible for systems larger a few atoms.

Further important advantages offered by DFT are the following:

- The DFT method keeps a one-electron structure similar to the HF one thus offering a much easier interpretation of the results while incorporating the correlation effects.
- The DFT description of open-shell systems is much more balanced than the HF description, which is known to enhance the importance of high-spin contributions.
- When open-shell systems are treated at the *unrestricted* level (different MOs for different spin), the DFT solution is generally contaminated by high multiplicity contributions to a much lesser extent than the HF one. Thus, DFT wave functions are usually quite close to a pure *spin representation*.
- DFT methods are found to give excellent results for molecular systems characterized by near-degenerate states, which are generally not adequately described by ab initio mono-determinantal perturbational theories (for example, MP2).

However, DFT methods are not free from limitations. The perhaps most obvious shortcoming is that DFT is a ground-state method, and as such incapable of dealing with excited states. For optical transitions, this problem can be overcome by performing time-dependent density functional theory calculations. There, the molecule is perturbed by an external electromagnetic field and the response of the electron density is monitored. The wavelength at which the electron density oscillates in resonance with the field corresponds to an optical excitation. However, it is still impossible to find the electron density of excited states (in particular of states with the same symmetry as the ground state), and therefore some properties remain elusive.

Apart from this fundamental issue, there are also some other failures of DFT (or, more exactly, the currently employed approximations in DFT), that all follow as a consequence of the approximate treatment of exchange and correlation energies. Perhaps the most spectacular artefact is the *self-interaction error* (SIE), which effectively means that an electron experiences a repulsion from itself [140]. This clearly unphysical situation arises because the exchange part of the $E_{xc}[\rho]$ term does not exactly cancel the $J_{ii}[\rho]$ repulsion (c.f Equations (2.49) and (2.52)) in general. In HF theory the exchange is exact, as mentioned, and the Coulomb and exchange parts of the self-inteaction cancel exactly (Equation (2.32)). The classic illustration of this error is the H_2^+ ion, for which DFT predicts a severely underestimated dissociation energy, and furthermore that each separated nucleus obtains half an electron each [140]. While the SIE is only well-defined for one-electron systems, such H_2^+, the phenomenon is general [158], and is perhaps better described as a *delocalization error* [159] because the electron self-repulsion promotes a superficial delocalization. Typical consequences in larger systems are underestimations of reaction barriers [160, 161] and incorrect response to the addition or removal of charges [159]. Attempts have been made to construct 'self-interaction corrected' DFT [160, 162], but have had limited impact due to practical difficulties in implementation [140].

The KS theory tries to capture the effects of the correlations without building them explicitly into the wave function. There are certain systems and properties where explicit treatment of the correlations is needed. Van der Waals interaction between molecule is a well-known phenomenon depending on correlation. It is an in attraction between particles that falls off as $1/r^6$ at moderately large distances, which is due to fluctuations in dipole fields associated with electronic transitions. Thus it is completely absent from a theory that represent the electron wave function by a single Slater determinant. Whether DFT can predict the small energy differences associated with van der Waals interactions or if additional corrections or nonlocal functionals of the density are needed has been the subject of much debate and current research. This interaction, although one of the weakest, is key to the accurate understanding of the biological processes involved in many drug-protein and protein-protein interactions.

Finally, but perhaps the most important weakness of DFT is the lack of a systematic procedure to progressively improve the $E_{xc}[\rho]$ functionals. Consequently, the accuracy of DFT results cannot be increased systematically, as it occurs with the fully ab initio approaches making use of strictly variational methods and wave function expansions. In addition, the limitations inherent to the adopted exchange-correlation functional cannot be a priori removed, since the exact form of the functional remains unknown. As a consequence, the reliability of a given functional can be only established in a heuristic way, by comparison of the DFT results with the experiments or, when possible, with extremely accurate results obtained from highly correlated ab initio methods. This approach has led in many instances to the introduction in the formalism of the functional of different arbitrary parameters (as seen in section 2.12), in order to achieve a better reproduction of a large set of experimental data.

Furthermore, also determining the most appropriate functional for a particular application is a important DFT challenge. It is not always clear what the best choice of functional is for the system of interest. Many DFT methods perform poorly for hydrogen-bonded systems and other weakly-bound intermolecular complexes, so care must be taken in selecting the DFT functional.

2.14 Using DFT methods to explore enzymatic reactivity

In spite of the above limitations, DFT has become a general tool to investigate models of metal-containing proteins. The recently developed GGA functionals and hybrid methods, give accuracies similar or higher than ab initio MP2 methods, but with the advantage of a computational cost that is comparable to Hartree–Fock calculations. In addition, recent techniques developed to calculate the Coulomb energy, such as the Resolution of Identity (RI) approximation [163], have further improved the computational efficiency. The reliability of quantum chemical methods to study simple models of metal enzymes is nowadays well grounded [164–166]. In fact, theoretical results can be very useful for the interpretation and the rationalization of experimental data. In addition, quantum chemical investigations can discriminate among alternative reaction mechanisms and identify intermediate species too short-lived to be characterized experimentally. Moreover, even when a complete picture of the chemistry is available, theoretical investigations can help to explain why a particular reaction path is favoured with respect to others. Finally, in silico coordination chemistry can allow, in a relatively cheap way, to explore synthetic pathways and chemical properties of coordination compounds not yet experimentally characterized, stimulating and possibly driving synthetic efforts.

2.15 Geometry optimization

In order to characterize chemical steps of reaction pathways, including catalytic cycles, the disclosure of the structure and the relative energy of key intermediate species and corresponding transition states connecting them is crucial. Geometry optimization, therefore, is a key component of most computational chemistry studies that are concerned with the structure and/or reactivity of molecules and thereby the finding of the nuclear coordinates leading to the stationary points on the PES.

The potential energy surface is usually described as a hilly landscape, with valleys, peaks, and mountain passes. The valleys of a potential energy surface represent reactants, intermediates, and products of a reaction. The position of the minimum in a valley represents the equilibrium structure. The energy difference between the product valley and reactant valley minima represents the energy of the reaction. The lowest energy pathway between the reactant valley and the product valley is the reaction path. The highest point on this lowest energy reaction path is the transition state (TS) for the reaction, and the difference between the energy of the TS and that of the reactant is the energy barrier for the reaction. A TS is a maximum in one direction (the direction connecting reactant and product along the reaction path) and a minimum in all other directions (directions perpendicular to the reaction path). A TS is also termed as a first-order saddle point.

The first and second derivatives of the energy with respect to the geometrical parameters can be used to construct a local quadratic approximation to the potential energy surface:

$$E(\boldsymbol{x}) = E(\boldsymbol{x}_0) + \boldsymbol{g}_{\boldsymbol{0}}^T \Delta \boldsymbol{x} + \frac{1}{2} \Delta \boldsymbol{x}^T \boldsymbol{H}_{\boldsymbol{0}} \Delta \boldsymbol{x} \tag{2.83}$$

where $\boldsymbol{g}_0$ is the gradient (dE/dx) at $\boldsymbol{x}_0$, $\boldsymbol{H}_0$ is the Hessian $(d^2E/d\boldsymbol{x}^2)$ and $\Delta \boldsymbol{x} = \boldsymbol{x} - \boldsymbol{x}_{\boldsymbol{0}}$. The gradient and the Hessian can be used to confirm the character of minima and TSs. The negative of the gradient is the vector of forces on the atoms in the molecule. Because the forces are zero for minima, TSs, and higher order saddle points, these structures are also termed stationary points. The Hessian or matrix of second derivatives of the energy is also known as the force constant matrix. The eigenvectors of the mass-weighted Hessian in Cartesian coordinates correspond to the normal modes of vibration. For a structure to be characterized as a minimum, the gradient must be zero and all of the eigenvalues of the Hessian corresponding to molecular vibrations must be positive; equivalently, the vibrational frequencies must be real (the vibrational frequencies are proportional to the square root of the eigenvalues of the mass-weighted Hessian). For a TS, the potential energy surface is a maximum in one direction (along the reaction path) and a minimum in all other perpendicular directions. Therefore, a TS is characterized by a zero gradient and a Hessian that has one (and only one) negative eigenvalue; correspondingly, a TS has one and only one imaginary vibrational frequency. The vibrational mode corresponding to the imaginary frequency is also known as the transition vector. At the TS, the transition vector is tangent to the reaction path in mass-weighted coordinates.

Most nonlinear optimization algorithms are based on a local quadratic approximation of the potential energy surface (Equation (2.83)). Differentiation with respect to the coordinates yields an approximation for the gradient, given by:

$$\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{g}_0 + \boldsymbol{H}_{\boldsymbol{0}}\Delta\boldsymbol{x} \tag{2.84}$$

At a stationary point, $\boldsymbol{g}(\boldsymbol{x}) = 0$; thus, in the local quadratic approximation to the potential energy surface, the displacement to the minimum is given by:

$$\Delta\boldsymbol{x} = -\boldsymbol{H}_0^{-1}\boldsymbol{g}_0 \tag{2.85}$$

This is known as the Newton or Newton–Raphson step. Newton and quasi-Newton methods are the most efficient and widely used procedures for optimizing equilibrium geometries and can also be used effectively to find TSs. For each step in the Newton method, the Hessian in Equation (2.84) is calculated at the current point. For quasi-Newton methods, Equation (2.84) is used with an approximate Hessian that is updated at each step of the optimization. Because actual potential energy surfaces are rarely quadratic, several Newton or quasi-Newton steps are required to reach a stationary point. At each step, Newton methods require the explicit calculation of the Hessian, which can be rather costly. Conversely, Quasi-Newton methods start with an inexpensive approximation to the Hessian. The difference between the calculated change in the gradient and the change predicted with the approximate Hessian is used to improve the Hessian at each step in the optimization.

$$\boldsymbol{H}^{new} = \boldsymbol{H}^{old} + \Delta\boldsymbol{H} \tag{2.86}$$

For a quadratic surface, the updated Hessian must fulfill the Newton condition,

$$\Delta\boldsymbol{g} = \boldsymbol{H}^{new}\Delta\boldsymbol{x} \tag{2.87}$$

where $\Delta\boldsymbol{g} = \boldsymbol{g}(\boldsymbol{x}^{new}) - \boldsymbol{g}(\boldsymbol{x}^{old})$ and $\Delta\boldsymbol{x} = \boldsymbol{x}^{new} - \boldsymbol{x}^{old}$. However, there are an infinite number of ways to update the Hessian and fulfill the Newton condition. Therefore, numerous updates have been proposed.

The problem with the Newton-Raphson methods is that for large systems, computing, diagonalizing and/or storing the Hessian becomes too demanding. It is therefore common practice to use different approximations such as conjugate-gradient methods. The concept behind these methods is to choose a new search direction that will lower the energy while remaining at or near the minimum in the previous search direction. If the Hessian has coupling between the coordinates, the optimal search directions are not orthogonal but are conjugate, in the sense that $\boldsymbol{H}\Delta\boldsymbol{x}^{old} = 0$.

In quantum chemistry, geometry optimization within the Born-Oppenheimer approximation is equally straightforward albeit more tedious. One minimizes the electronic energy for a given set of nuclear coordinates, followed by an update of the latter by one of the aforementioned techniques based on the forces on the nuclei for that particular electronic configuration. The next iteration starts with a new electronic structure minimization and so on until both the energy and nuclear forces are below some threshold value. It should be noted that the only way to assure that an optimized minimum is a global minimum is to find all minima. This is rarely done in quantum chemistry calculations, and one typically relies on prior knowledge about the system to make a qualified starting guess. For larger systems, the meaning of global minima becomes blurred as the conformational degrees of freedom will be distributed over many states, and the macroscopic state is best represented by an ensemble average. The reason why quantum chemistry methods are often successful at predicting properties such as reaction energies and activation barriers is that contributions from thermodynamic distributions largely cancel, so that the few considered states can be seen as representative of the whole ensemble. Again, such assumptions must be made with care.

TS optimization is much more difficult. Since the TS is a first-order saddle point, the diagonalized Hessian should have exactly one negative eigenvalue. Although Newton-Raphson may find such points by chance, there is no fail-safe strategy to do this. The numerous available TS optimization methods can roughly be divided into local and interpolation methods. The former are the most common but require a good guess of the actual TS geometry to work, as well as an explicit initial calculation of the Hessian. The TS guess can be found by searching the PES by various constrained optimization techniques, but still takes a good deal of chemical intuition and experience. Calculating the Hessian is very demanding and limits the size of systems feasible for TS optimization.

2.16 Solvent effects

Basic calculations in quantum chemistry are done in gas-phase manner, assuming that the interaction between the model complex and surrounding medium is negligible. However, most natural and laboratory chemistry does not occur in vacuum. There exists an interaction between solute and solvent. The solute properties as structure, stability, spectra, and reactivity depend on the solvent, particularly a polar one.

The long-range or polarization effects induced by the solvent (e.g., the protein environment in enzymatic catalysis) can be reproduced by modeling the solvent as a macroscopic continuum with a given dielectric permittivity constant ε and the solute as contained in a cavity of this continuous medium. In the dielectric cavity methods the solvent acts as a perturbation on the gas-phase behaviour of the system since the solute is subjected to the electrostatic potential created by the continuum, which in turn is polarized by the influence of the solute itself. The cavity model is a good approximation, provided that essential hydrogen bonding interactions (i.e., the short-range effects due to individual solvent molecules) are included in the model. A dielectric constant of 4 is predicted to be the more appropriate for treating polarization effects of the protein environment on the stereoelectronic properties of the active sites by a variety of theoretical methods based on normal mode analysis and on molecular dynamics simulations [167] as, for example, it accounts for electronic polarization and small backbone fluctuations [168].

One of the extensively used "continuum solvation" models is the conductor-like screening model, abbreviated as COSMO [169, 170]. In the COSMO approach, the details of the cavity construction differ in different COSMO implementations. In most cases it is constructed as an assembly of atom-centered spheres with radii approximately 20% larger than the Van der Waals radius. For the actual calculation the cavity surface is approximated by segments, e.g., hexagons, pentagons, or triangles. COSMO derives the polarization charges of the continuum, caused by the polarity of the solute, from a scaled-conductor approximation. If the solvent was an ideal conductor, the electric potential on the cavity surface must disappear. If the distribution of the electric charge in the molecule is known, e.g. from quantum chemistry, it is possible to calculate the charge q^* on the surface segments. For solvents with finite dielectric constant, this charge q is lower by approximately a factor $f(\varepsilon)$:

$$q = f(\varepsilon)q^* \quad , \quad f(\varepsilon) = \frac{\varepsilon - 1}{\varepsilon + x} \tag{2.88}$$

where $x = 1/2$ has been found most useful based on theoretical arguments. From the thus determined solvent charges q and the known charge distribution of the molecule, the energy of the interaction between the solvent and the solute molecule can be calculated.

2.17 Chemical models

An important factor that can really determine the overall quality of a given quantum chemical calculation is the accuracy and the adequacy of the molecular model used. The choice of reliable and sufficiently large model may be thus very important for accurate description of enzymatic active sites. However the computational cost of correlated to quantum mechanical methods mostly prevents their applications to very large systems.

Thus, the real enzyme systems should be reduced to a minimum-sized active site model, where those atoms undergoing chemical changes during catalysis are included. Since activity of a enzyme is mainly attributed to the active site, the real size of the system is often reduced to minimal models including only the first coordination sphere of the reactive center. Nevertheless, larger models including all aminoacids in the second coordination sphere of the active sites are required to investigate properties that might be influenced by the protein environment. Therefore, computational models are usually the result of a compromise between a realistic representation of the system and the necessity to treat relatively small models in order to reduce the computation time.

2.18 Applied Software Packages

Various software were employed to obtain the results discussed in this dissertation. The Turbomole program [171] was used for electronic structure calculations with DFT methods, geometry optimizations of both ground states and transition states, as well as for computation of frequencies. EPR parameters calculations were carried out using the program package ORCA [172, 173].

For visualization of X-ray protein structures and development of models of their active sites was used the PyMOL Molecular Graphics System [174]. Conversely, initial geometries for the computations were created, the resulting structures were inspected and further edited with Molden [175]. Visualization of molecules, electron densities and orbitals, as well as figure generations, were carried out using Molekel [176] and Chimera [177] programs.

CHAPTER

3

STEREOELECTRONIC AND CATALYTIC PROPERTIES OF THE NI-CODH ACTIVE SITE

Several aspects concerning structure and reactivity of the Ni-CODH enzyme are still poorly understood. The well-characterized catalytic intermediates C_{red1} and C_{red2} have been proposed as the redox forms of the active site competent, respectively, for CO oxidation and CO_2 reduction [96]. However, structure and electronic nature of these states is still under discussion [96, 107, 108, 112, 113]. Metal sites at which occurs the bielectronic reduction of the C-cluster from C_{red1} to C_{red2} are still unknown. Whereas the Ni(II) redox state has been accepted for the Ni atom in C_{red1}, contrasting spectroscopic data (see section 1.4.1) suggest two alternative descriptions of the Ni atom for the C_{red2} state; Ni(0) and the isoelectronic Ni(II)-hydride species [107, 113]. Consequently, different mechanisms can be proposed for the CODH enzyme. According to the Ni(II)/Ni(0) assignment for the Ni atom in the C_{red1}/C_{red2} state, Jeoung and Dobbek's high resolution structures suggest a catalytic mechanism involving the formation of a nonbridging hydroxide ligand bounded to the Fe_u atom in both active states [107]. The C_{red1} and C_{red2} states have been respectively assigned to samples poised at -320 mV and -600mV ([OH-320] and [OH-600]). A third structure from sample poised at -600 mV and treated with $NaHCO_3$, was instead resolved as relevant intermediate of the catalytic cycle in which CO_2 bridges the Ni and Fe_u ions ([COO-600]). However, Fontecilla-Camps and co-workers [113] noticed that the [OH-320] and [OH-600] structures are almost identical and suggested that the latter one is a mixture of [OH-320] and [COO-600]. Furthermore, employing hybrid quantum mechanical/molecular mechanical techniques, they proposed a bridging hydroxide and an hydride bounded to a Ni(II) ion as the putative redox active ligands in the C_{red1} and C_{red2} states, respectively. The absence of the hydroxide ligand in the C_{red2} state is supported by spectroscopic experiments; the loss in C_{red2} of the strong ENDOR signal observed for the C_{red1} state was attributed to the release of OH^- [118]. According to this observation, the C_{red2} state was also proposed to consist of a Ni in the 0 oxidation state with no exogenous ligand [108, 178].

The uncertainty in the oxidation states, the nature of the active ligands and their coordination mode in C_{red1} and C_{red2} is at the origin of the controversy around the mechanism of the Ni-CODH. A large number of potential intermediates can be proposed on the basis of the current spectroscopic and theoretical data. Moreover, according to the available X-ray structures [102, 107, 119–121], both an associative and dissociative mechanism can be proposed for the substrates binding. In order to contribute to provide significant insights in the unveiling the stereoelectronic and catalytic properties of the Ni-CODH enzyme, quantum mechanical calculations have been performed on a very large and a minimal model of the active site. CO_2 and CO binding to the C-cluster, also in the presence of other substrates, have been investigated in different redox states.

3.1 Computational details

3.1.1 Models of the Ni-CODH active site

The starting structure for the DFT calculations was based on the X-ray geometry of the *Carboxydothermus hydrogenoformans* Ni-CODH (pdb code: 3B52) [107], in which a CO_2 molecule bridges the Ni and the Fe_u atoms of the active site. In the following, the residues are numbered according to this structure.

Two models of different size (schematically represented in Figure 3.1) have been considered to investigate the effect of the protein environment on the stereoelectronic properties of the C-cluster. The validity of such an approach (cluster models) for the DFT investigation of transition metal containing enzymes has been extensively discussed [179–181]. The smallest model (SM), which contains up to 64 atoms, includes the $[Fe_4NiS_4]$ core of the C-cluster and the side chains of the residues forming its first coordination sphere. The five cysteine residues coordinated to the nickel and iron atoms (Cys295, Cys333, Cys446, Cys476, Cys526) and the histidine residue coordinated to the Fe_u atom (His261) are terminated at the Cα atoms and saturated with hydrogens. During the geometry optimization, terminal atoms are constrained to their crystallographic positions, in order to avoid unrealistic distortions of the C-cluster. The largest model (LM) contains up to 234 atoms and has a size of 24 Å. In this model have been included selected atoms of 16 residues belonging to the second coordination sphere (Ala91, Gly92, His93, Ser94, His96, Asp219, Cys223, Asp 231, Glu299, Arg303, Gln332, Ser525, Val527, Lys563, Ile567 and Trp570) and three water molecules, apart from all the atoms contained in the small model.

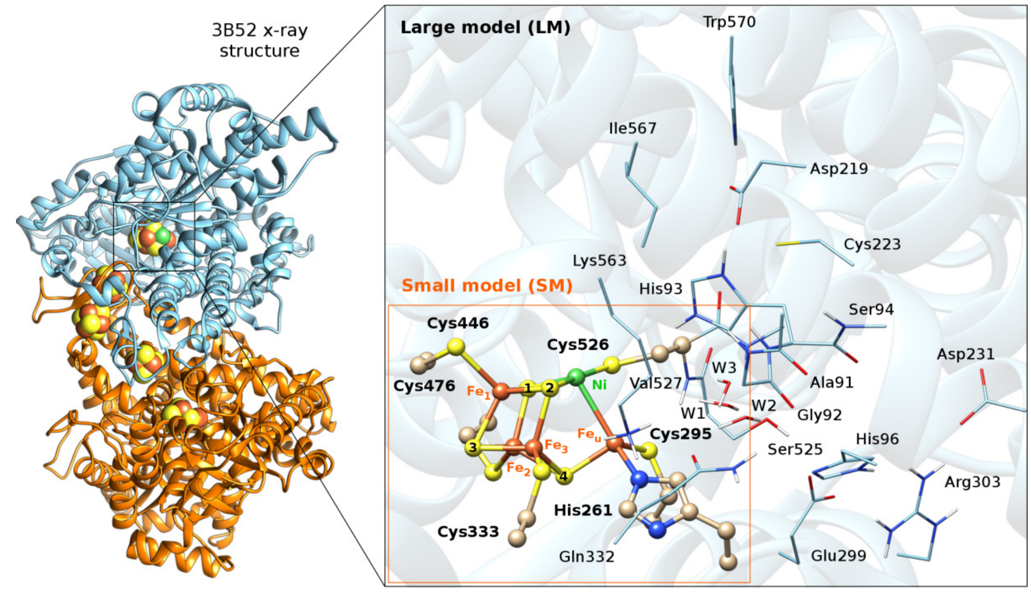


Figure 3.1 Two different size models of C-cluster obtained from the X-ray crystal structure of the *C. hydrogenoformans* Ni-CODH (PDB: 3B52)[107]. The metallic cluster and selected atoms of its protein environment, included in the large DFT model (LM), are shown on the right half of the figure (black box, 'stick' representation); atoms belonging also to the small model (SM) are represented as 'ball and stick' and relative residue names are indicated with bold labels (orange box). For the sake of clarity, aliphatic hydrogen atoms and the CO_2 ligand bridging the Ni and Fe_u ions are not shown.

In particular, the entire residue His93 and the side chain of Lys563 have been included in the model because they should be directly involved in the catalytic cycle, interacting with the ligands bounded to the C-cluster [107, 120] and participating in acid-base reactions [97, 98]. Notably, His93 is positioned at the top of a cationic tunnel composed of histidine residues located on sequential turns of a helix starting near the C-cluster and ending at the protein surface, that is proposed to facilitate transfer of protons during the reaction [97, 182]. The importance of His93 and Lys563 in catalysis is confirmed by the loss of the enzymatic activity after their mutation [182].

Three protonation states are possible for His93 depending on whether δN, εN, or both atoms are protonated. Since proton on δN strongly interacts with the carboxylate group of Asp219, it was always included in the model. Conversely, proton on εN interacts with nonprotein ligands at the active site. Therefore, it is possible to assume that protonation state of εN plays a crucial role in the binding and release of the substrates. On the basis of these considerations, His93 was modeled as either doubly protonated or singly protonated at the δN atom.

The carbonyl and the Cα atoms of Ala91, the entire residue Gly92 and the N and the Cα atoms of Ser94 have been included in the model because they form a small α-helix in which is contained His93, whereas the side chain of Asp219 terminated at the Cα atom has been selected because its carboxylate group is H-bonded to it. Conversely, the side chain of Trp570 terminated at the Cβ atom, interacting with Asp219, and the side chain of Cys223 terminated at the Cα atom have been included in the model in order to avoid unrealistic conformation changes of the side chain of His93. The side chain of Ile567 terminated at the Cα atom has been added to the model, since it is close to the vacant coordination site on Ni, whereas the side chain, the carbonyl and the Cα atoms of Ser525 and the N and the Cα atoms of Val527 have been selected because they form the peptide chain containing the residue Cys526, belonging to the first coordination sphere. Conversely, the side chain of Glu299 terminated at the Cα atom has been included in the model because it is H-bonded to Ser525. Finally, the side chains of His96, Gln332 and Cys223 terminated at the Cα atoms, the side chain of Arg303 truncated at the Cγ atom, the side chain of Asp231 terminated at the Cβ atom and three water molecules, selected among those reported in the crystallographic structure, have been included in the model in order to mimic the entire H-bonds network in the C-cluster environment. The truncated residues have been saturated with hydrogen atoms. During the geometry optimisations, 31 atoms have been constrained to the crystallographic positions, in order to avoid unrealistic distortions at the boundary of the model.

The results obtained using the minimal DFT model of the C-cluster (SM) and the very large model of the active site (LM) will be compared in order to investigate the importance of the protein environment in tuning the coordination geometry of the C-cluster and its reactivity towards the substrates CO_2 and CO.

3.1.2 DFT calculations

Quantum mechanics (QM) calculations have been carried out in the DFT framework with the Turbomole program suite [171], using the BP86 exchange-correlation functional [149, 150] in conjunction with the resolution-of-the-identity (RI) technique [163]. The BP86 functional is commonly used to study metal-containing molecular systems as metallo-enzymes, due to the increasing available computational data which indicate that BP86 is one of the most accurate pure functionals to study transition metal compounds [183]. BP86, coupled with an appropriate basis set, predicts reaction energies with a reasonable accuracy and reproduces experimental geometries within a few hundredths of an Å [140].

In the small model, an all-electron valence triple-ζ basis set with polarization functions (def-TZVP) [184] has been used for all atoms. In the large model the def-TZVP basis set is used for the $[Fe_4NiS_4]$ core of the C-cluster, the sulphur and the Cβ atoms of Cys526 and all atoms of Cys295, Cys333, Cys446, Cys476 and His261. For all other atoms, the double-ζ basis set SVP [185] has been used. The effects of the protein environment have been modeled by placing the molecular cluster in a polarizable continuum medium with $\varepsilon = 4$, according to the conductor-like screening model (COSMO) [169, 170].

In the framework of the single-determinant DFT approach, the antiferromagnetic coupling of the Fe atoms in the C-cluster has been treated within the spin-unrestricted broken-symmetry formalism (BS) introduce by Noodleman [186]. The resulting BS state is not a pure spin state, but rather a mixed state in which the majority spin and minority spin are arranged either spin-up and spin- down to give a spin coupling pattern with the correct net total spin and either an overall antiferromagnetic or ferromagnetic alignment. To construct a desired BS state, a calculation on the high-spin state is first completed, which is a pure spin state described by a single determinant with all unpaired electrons aligned in the same direction (spin-up) to adopt the highest possible total spin state. The density of the high spin state is then manipulated by exchanging designated blocks of α and β electron densities. In this way, the starting density for the desired spin-flipped state, from which BS states are obtained by SCF convergence, is created.

In the C_{red1} state a mixed valence Fe(II,HS)Fe(III,HS) pair of spin $S = 9/2$ and a ferrous Fe(II,HS)Fe(II,HS) pair of spin $S = 4$ are coupled antiferromagnetically to give an overall low-spin ground state which exhibits an $S = 1/2$ EPR signal [110, 178]. Conversely, the Ni atom is a low spin diamagnetic Ni(II) [111, 114]. The $S = 1/2$ BS state for C_{red1} was therefore obtained by manipulating the density of the highest possible total spin state $S = 17/2$. Analogously, the desired BS states for the one- and two-electron reduced states, C_{int} ($S = 0$) and C_{red2} ($S = 1/2$), were obtained from the $S = 18/2$ and $S = 19/2$ high spin states, respectively.

The resulting electronic structures of the six possible non equivalent spin coupling schemes for the C-cluster which satisfy $S = 1/2$ for C_{red1} and C_{red2} and $S = 0$ for C_{int}. (Figure 3.2) were checked by computing Mulliken spin densities and NBO atomic charges.

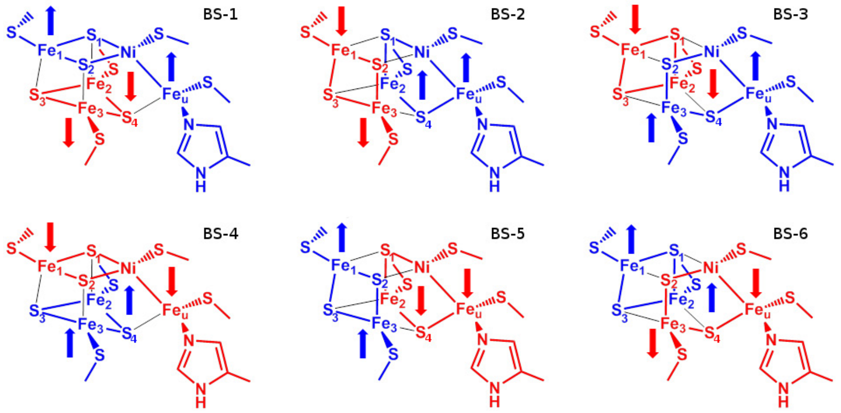


Figure 3.2 Schematic representation of all possible non equivalent spin coupling schemes for the C-cluster, in which two pairs of Fe atoms are coupled antiferromagnetically. Layers containing Fe atoms with α and with β electron density are respectively indicated in blue and red. All results presented in the following refers to the BS-1 spin coupling.

Since the resolution of the x-ray structures of the enzyme is not sufficiently high to establish the broken-symmetry state of the C-cluster, the geometry of several forms of the enzyme have been optimized according to all possible spin coupling schemes. The geometries of the different broken-symmetry states are very similar, whereas their relative stabilities are strongly influenced by the electronic structure of the C-cluster. In particular, the spin alignment pattern corresponding to the BS-1 state in Figure 3.2 is generally found of lower energy. For example, BS-1 for the free form of the C-cluster in the C_{red1} state optimized using the small model is more stable than all other BS states by more than 8 kcal/mol (see Table 3.1). Therefore, all results reported in this work refer to structures optimized aligning the Fe site spin vectors according to the BS-1 spin coupling scheme.

Table 3.1 Relative energies (in kcal/mol) for all possible spin coupling schemes of the C_{red1} state of the C-cluster, optimized using the small model (SM), NBO atomic charges and spin densities (in parenthesis) of selected atoms of the C-cluster.

	ΔE	**Ni**	**Fe_u**	**Fe_1**	**Fe_2**	**Fe_3**
BS-1	0	0.54(0.27)	0.96(3.40)	0.73(3.12)	0.71(-3.00)	0.71(-3.02)
BS-2	+12.5	0.54(0.52)	0.87(3.26)	0.72(-3.06)	0.76(3.12)	0.67(-3.02)
BS-3	+8.0	0.53(0.42)	0.96(3.38)	0.70(-3.01)	0.74(3.09)	0.75(3.16)
BS-4	+9.3	0.50(0.39)	0.96(-3.36)	0.53(-2.44)	0.74(3.06)	0.74(3.11)
BS-5	+14.8	0.53(0.31)	0.91(-3.19)	0.78(3.24)	0.68(-2.91)	0.75(3.14)
BS-6	+12.7	0.56(0.40)	0.88(-3.12)	0.78(3.20)	0.77(3.14)	0.72(-3.03)

Nomenclature: in the following, computational models will be labelled according the general scheme **RS-X-M**, where **RS** is the formal redox state of the C-cluster (C_{red2}, C_{int} or C_{red1}), **X** is the specific chemical nature of the complex and **M** is the size of the model (LM or SM). The protonation state of His93 is indicated by the superscript after the **LM** label; **LM^a** and **LM^b** denote the doubly- and the singly- protonated state, respectively.

3.2 CO_2-bound intermediates of the C-cluster

Numerous X-ray structures of Ni-CODHs, with different ligands bounded to the C-cluster, have been reported. Among these, the species featuring a CO_2 ligand bridged between Ni and Fe_u [107, 120] is perhaps the most interesting one because it corresponds to the key catalytic intermediate in which CO_2 is activated by the C-cluster.

With the aim of providing significant insights on the structure and the electronic nature of this intermediate, the coordination of CO_2 to the C-cluster has been investigated in different redox states (C_{red2}, C_{int} and C_{red1}) by DFT calculations. Two different size models of the C-cluster have been used in order to disclose the crucial role of the protein environment in efficient binding of CO_2 to the C-cluster.

3.2.1 Geometry of the CO_2-bound C-cluster

Two high resolution crystal structures of the C-cluster featuring a CO_2 ligand bridging Ni and Fe_u in a μ_2,η^2 coordination geometry have been reported by Jeoung and Dobbek (see Figure 3.3). The 3B52 structure has been solved at 1.5 Å from a sample poised at -600 mV in presence of $NaHCO_3$ as the CO_2 source [107]. The 4UDX structure, which is also determined at -600 mV in the presence of HCO_3^-/CO, has been solved at 1.03 Å [120]. Since the true atomic resolution of the 4UDX structure, reported in 2015, was not yet available when we started our study on Ni-CODH, the 3B52 structure was used as starting geometry of the C-cluster for DFT calculations.

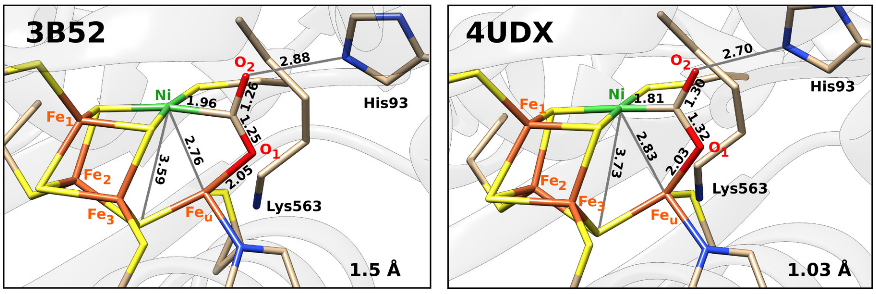


Figure 3.3 Schematic representation of the active site geometry in the 3B52 and 4UDX crystal structures, solved respectively at 1.5 Å and 1.03 Å. Selected distances are given in Å.

Geometry optimization of the large model LM^a of the CO_2-bound C-cluster in the C_{red2} state does not lead to a significant structural rearrangement of the C-cluster relative to the 3B52 structure (see Figure 3.4 and Table 3.2) [107]. In both structures, the carboxyl carbon atom of CO_2 is coordinated to the Ni ion (η^1_C coordination), completing its distorted square planar geometry, whereas one carboxyl oxygen (hereafter, referred as O1) is bound to the Fe_u metal (η^1_O coordination), resulting in a μ_2-$\eta^2_{C,O}$ binding mode of CO_2 bridging the Ni-Fe_u site. The computed Ni-C and Fe_u-O1 bond lengths of 1.87 Å and 2.16 Å are in good agreement with the experimental values of 1.96 Å and 2.05 Å, and the same is for the Ni-C-O1 and the Fe_u-O1-C angles of 119.4° and 105.0° compared to 115.6° and 103.4°. The predicted C-O2 (hereafter, the O2 label refers to the nonbridging oxygen atom) distance of 1.26 Å perfectly reproduce the experimental value, whereas the C-O1 bond distance of 1.32 Å is slightly longer than that observed in the X-ray 3B52 structure (1.25 Å). The CO_2 coordinated to the C-cluster in the optimized structure is also slightly more bent than that reported in the X-ray structure. The computed O1-C-O2 angle is 122.0°, while it is 132.6° in the 3B52 structure. Consequently, the computed Ni-C-O2 angle (122.4°) results 14° larger than that in the X-ray structure (108.0°). However, it is worth noting that the bent CO_2 geometry predicted by our calculations is very similar to that found in the true atomic resolution 4UDX structure [120]. The latter differs substantially from that determined at 1.5 Å resolution in the position of the carbon atom of bound CO_2. The Ni-C bond (1.81 Å) is substantially shorter than determined earlier, whereas the two C-O bonds are considerably elongated (C-O1 and C-O2 distances are 1.32 Å and 1.30 Å, respectively). Notably, the O1-C-O2 angle (117.2°) is 15° smaller than that estimated in the 3B52 structure (132.6°) and only 5° larger than predicted by our model. In addition, the calculated distances between the oxygen atoms of the CO_2 ligand and the nitrogen atoms of His93 and Lys563 residues, which form H-bonds with such oxygen atoms, better reproduce those found in the 4UDX structure. Indeed, the computed O1-N(Lys563) and O2-εN(His93) distances are 2.83 Å and 2.57 Å, whereas they are equal to 2.72 Å and 2.70 Å in the 4UDX structure.

No significant structural differences have been found in the CO_2 geometry of the CO_2-bound C-cluster in the C_{int} and C_{red1} states, optimized using the large model, with respect to that predicted in the C_{red2} state (see Figure 3.4 and Table 3.2). On the other hand, substantial geometry changes have been determined for the [$NiFe_4S_4$] cluster and side chain of the Cys526 residue. In particular, oxidation of C_{red2} to C_{int} and then to C_{red1} leads to the progressive contraction of the C-cluster with the sulphur atom of Cys526 approaching the Fe_u ion, as indicated by the Ni-S4 and Fe_u-S(Cys526) distances that are 3.66 Å and 3.42 Å in C_{red2}, 3.53 Å and 3.38 Å in C_{int} and 3.27 Å and 2.65 Å in C_{red1}.

In order to detect the role of the protein environment in modelling the geometry of the active site, the structures of the CO_2-bound C-cluster in the C_{red2}, C_{int} and C_{red1} states have been optimized also using the small-size model. The CO_2 geometry is only slightly affected by the absence in this model of the residues belonging to the second coordination sphere (see Figure 3.4 and Table 3.2). In particular, the C-O1 and C-O2 bonds are shorter by 0.03 Å than those predicted for the corresponding structures optimized using the large model, whereas the O1-C-O2 angles are larger not more than 3°. A major rearrangement has been predicted for the [$NiFe_4S_4$] cluster. Also in the case of the small model, oxidation of the C-cluster leads to a progressive contraction of the [$NiFe_4S_4$] cluster and the movement of Cys526 towards Fe_u, but the Ni-S4 and Fe_u-S(Cys526) distances are significantly shorter than those calculated using the large model as well as than those of the X-ray structures.

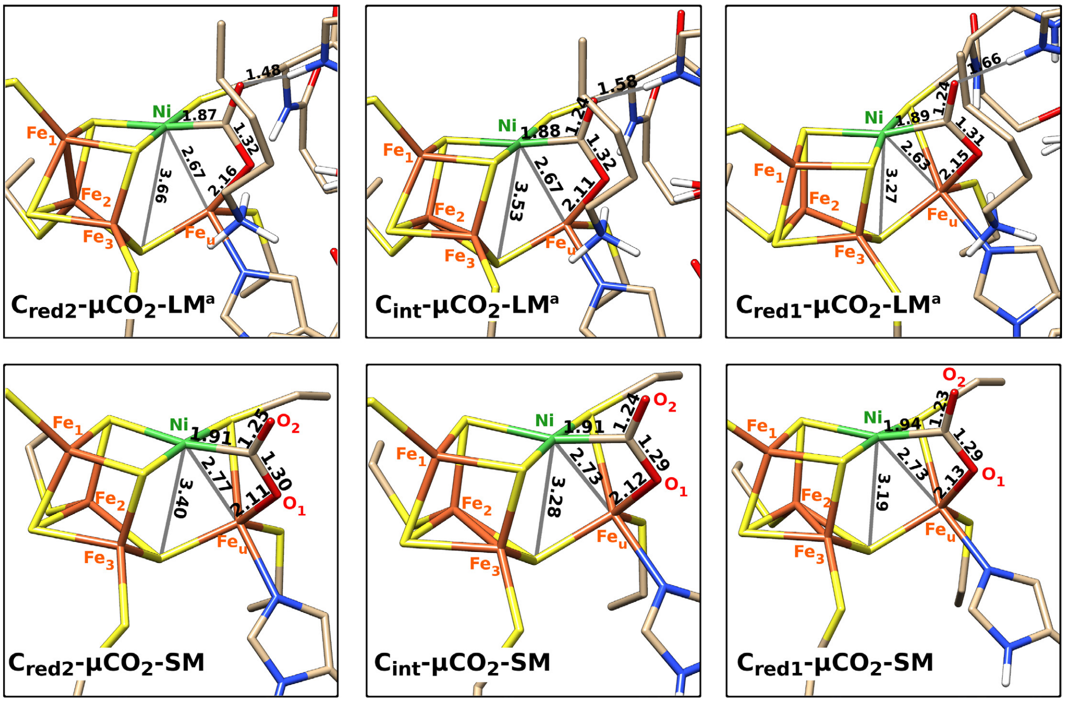


Figure 3.4 Schematic representation of the DFT geometries of the μCO_2-bound C-cluster in C_{red2}, C_{int} and C_{red1} states, optimized using the large (LM^a) and small (SM) model. Selected distances are given in Å. For the sake of clarity, aliphatic hydrogen atoms are not shown.

Table 3.2 Selected distances (Å) and angles (°) for the large (LM[a]) and small (SM) models of the CO_2-bound C-cluster in the C_{red2}, C_{int} and C_{red1} states and for the x-ray structures 3B52, 4UDX and 2iYV. The O2 label in 2YIV structure denotes the N atom of the n-butyl isocyanate ligand.

	X-ray structures			**μCO_2**						**tCO_2**					**tCO_2-OH**					
				LM[a]			**SM**			**LM[a]**		**SM**			**LM[a]**			**SM**		
	3B52	4UDX	2YIV	C_{red2}	C_{int}	C_{red1}	C_{red2}	C_{int}	C_{red1}	C_{red2}	C_{int}	C_{red2}	C_{int}	C_{red1}	C_{red2}	C_{int}	C_{red1}	C_{red2}	C_{int}	C_{red1}
Ni-Fe_u	2.76	2.83	3.10	2.67	2.67	2.63	2.77	2.73	2.73	2.65	2.66	2.98	3.01	3.33	3.04	3.01	2.80	3.31	3.27	2.98
Ni-S_{Cys526}	2.10	2.11	2.09	2.24	2.22	2.23	2.25	2.24	2.24	2.23	2.22	2.23	2.24	2.22	2.26	2.23	2.23	2.28	2.25	2.25
Ni-S4	3.59	3.73	3.99	3.66	3.53	3.27	3.40	3.28	3.19	3.69	3.55	3.45	3.49	3.33	3.70	3.61	3.25	3.54	3.40	2.94
S2-Ni-S_{Cys526}	168.3	166.2	161.0	172.4	169.1	169.0	175.0	175.0	173.0	166.6	167.2	161.8	165.4	165.7	157.4	150.3	156.6	153.5	154.9	165.2
S1-Ni-C	171.3	174.9	147.8	174.3	173.2	167.2	170.2	167.4	165.0	158.1	159.4	149.1	147.6	142.3	141.0	140.3	128.1	134.8	137.8	118.6
Fe_u-S4	2.19	2.33	2.32	2.26	2.25	2.22	2.27	2.26	2.23	2.24	2.23	2.41	2.36	2.33	2.32	2.31	2.25	2.34	2.33	2.28
Fe_u-S_{Cys295}	2.34	2.29	2.30	2.34	2.31	2.31	2.45	2.41	2.36	2.34	2.31	2.23	2.25	2.20	2.38	2.36	2.37	2.54	2.47	2.45
Fe_u-N_{His261}	2.03	2..15	2.05	2.16	2.16	2.21	2.26	2.25	2.22	2.19	2.19	2.13	2.13	2.11	2.25	2.27	2.27	2.35	2.33	2.28
Fe_u-S_{Cys526}	3.33	3.34	3.44	3.42	3.38	2.65	2.69	2.58	2.52	2.65	2.55	2.34	2.33	2.37	2.71	2.66	2.58	2.52	2.52	2.49
Ni-Fe_u-S4	92.3	92.1	93.7	95.1	91.1	84.2	84.3	81.5	79.4	97.6	92.5	81.6	81.8	85.1	86.3	84.3	79.4	75.6	72.6	68.2
Ni-C	1.96	1.81	1.95	1.87	1.88	1.89	1.91	1.91	1.94	1.89	1.90	1.95	1.99	2.04	1.92	1.92	1.95	1.94	1.94	1.99
Fe_u-O1	2.05	2.03	4.06	2.16	2.11	2.15	2.11	2.12	2.13	3.19	3.05	4.15	4.12	3.90	4.00	3.99	3.93	4.03	4.00	3.96
C-O1	1.25	1.32	1.34	1.32	1.32	1.31	1.30	1.29	1.29	1.32	1.31	1.25	1.24	1.23	1.30	1.28	1.26	1.27	1.26	1.25
C-O2	1.26	1.30	1.24	1.26	1.24	1.24	1.25	1.24	1.23	1.24	1.24	1.26	1.25	1.23	1.24	1.24	1.23	1.26	1.25	1.24
O1-C-O2	132.6	117.2	121.1	122.0	123.8	126.3	124.1	126.7	128.9	122.4	125.6	130.6	134.0	138.9	125.8	130.4	135.0	127.2	130.6	134.5
Ni-Fe_u-O1	69.5	64.6	43.0	67.4	68.0	68.4	64.9	65.6	65.9	56.6	58.3	42.6	42.9	46.6	43.8	44.3	44.9	43.5	43.9	43.8
Ni-C-O1	119.4	119.8	113.0	115.6	115.0	114.8	111.2	111.5	111.3	120.6	120.6	121.0	119.0	118.5	116.3	120.5	118.0	120.4	120.9	113.9
Ni-C-O2	108.0	124.5	125.9	122.4	121.2	118.9	124.6	121.7	119.8	117.0	113.8	108.2	107.0	102.6	117.9	109.1	107.0	112.4	108.5	111.6
Fe_u-O	-	-	1.96	-	-	-	-	-	-	-	-	-	-	-	2.03	2.01	1.97	1.92	1.92	1.90
O1-O	-	-	2.51	-	-	-	-	-	-	-	-	-	-	-	2.66	2.69	2.72	2.89	2.88	2.87
O1-H	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.66	1.70	1.73	1.91	1.89	1.89
Fe_u-O-H	-	-	-	-	-	-	-	-	-	-	-	-	-	-	110.9	111.1	110.9	109.6	110.9	112.2
S4- Fe_u-O	-	-	148.4	-	-	-	-	-	-	-	-	-	-	-	143.7	141.0	134.2	143.5	139.7	134.8
O1-N_{Lys563}	2.64	2.72	3.94	2.83	2.86	2.86	-	-	-	2.78	2.80	-	-	-	3.97	4.08	4.22	-	-	-
O1-H_{Lys563}	-	-	-	1.78	1.81	1.82	-	-	-	1.72	1.74	-	-	-	3.15	3.26	3.39	-	-	-
O2-εN_{His93}	2.88	2.70	4.39	2.57	2.64	2.70	-	-	-	3.03	3.07	-	-	-	4.14	4.41	4.63	-	-	-
O2-εH_{His93}	-	-	-	1.48	1.58	1.66	-	-	-	2.37	2.40	-	-	-	3.18	3.46	3.70	-	-	-
O1-εN_{His93}	3.93	3.92	2.85	3.79	3.86	3.82	-	-	-	2.62	2.69	-	-	-	2.59	2.68	2.77	-	-	-
O1-εH_{His93}	-	-	-	-	-	-	-	-	-	1.52	1.62	-	-	-	1.49	1.62	1.75	-	-	-
O-N_{Lys563}	-	-	2.67	-	-	-	-	-	-	-	-	-	-	-	2.55	2.59	2.63	-	-	-
O-H_{Lys563}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.40	1.47	1.53	-	-	-

The validity of our DFT models for the prediction of the geometry of the Ni-CODH active site has been checked by superimposing structures of the $[NiFe_4S_4]$-CO_2 cluster extracted by theoretical and experimental geometries and calculating the corresponding root-mean-square deviations (RMSD) (see Table 3.3). Interestingly, the smaller RMSD value has been computed for the CO_2-bound C-cluster in the C_{red2} state optimized using the large model compared to the true atomic resolution 4UDX crystal structure (0.107 Å), despite the starting geometry for DFT calculations has been based on the 3B52 structure. When the analysis is performed on the small model the lower RMSD value has also been calculated for the C_{red2} state (0.201 Å), but it is much higher than that calculated using the large model highlighting the key role of the protein environment in tuning the geometry of the active site. In addition, the analysis of the RMSD values calculated for the $[NiFe_4S_4]$ core and CO_2 separately confirm that the protein environment mainly affects the geometry of the metallic cluster rather than that of the bound CO_2, as discussed above. The oxidized states C_{int} and C_{red1}, optimized using both large and small model, were characterized by larger deviations. These results support the assignment of the X-ray 3B52 and 4UDX structures poised at -600 mV to the C_{red2} state.

Table 3.3 RMSD values calculated for the structure of the metallic cluster and the bound CO_2 ($[NiFe_4S_4]$-CO_2), the $[NiFe_4S_4]$ core ($[NiFe_4S_4]$), and the CO_2 ligand (CO_2) extracted by the theoretical geometries, compared to the 3B52 and 4UDX crystal structures.

		LM^a			**SM**		
		$\mathbf{C_{red2}}$	$\mathbf{C_{int}}$	$\mathbf{C_{red1}}$	$\mathbf{C_{red2}}$	$\mathbf{C_{int}}$	$\mathbf{C_{red1}}$
3B52	$[NiFe_4S_4]$-CO_2	0.133	0.131	0.167	0.207	0.220	0.230
	$[NiFe_4S_4]$	0.125	0.138	0.181	0.225	0.247	0.260
	CO_2	0.062	0.054	0.041	0.051	0.037	0.027
4UDX	$[NiFe_4S_4]$-CO_2	0.107	0.118	0.174	0.201	0.223	0.239
	$[NiFe_4S_4]$	0.102	0.120	0.175	0.205	0.233	0.253
	CO_2	0.047	0.058	0.071	0.057	0.072	0.084

Although the μ_2-$\eta^2_{C,O}$ binding of CO_2 is the only one that has been observed in X-ray structures of the C-cluster [107, 120], other coordination modes have been investigated. Interestingly, all species in which CO_2 is terminally bounded to Ni or Fe_u by one oxygen atom are unstable, resulting in CO_2 release from the active site or isomerisation to the more stable structure in which CO_2 bridges the Ni-Fe_u site ($\boldsymbol{\mu}\mathbf{CO_2}$ species in Figure 3.4) during geometry optimization. Conversely, species in which the CO_2 ligand is terminally bounded to the Ni ion through the carbon atom ($\mathbf{tCO_2}$ species in Figure 3.5) is predicted as stable isomer in all redox states when corresponding structures have been optimized using the small model (SM). In the case of the large model (LM^a), it turns out that the $\mathbf{tCO_2}$ species is a stable isomer only in C_{red2} and C_{int} states. $\mathbf{C_{red2}}$**-**$\mathbf{tCO_2}$**-LM**^{a} and $\mathbf{C_{int}}$**-**$\mathbf{tCO_2}$**-LM**^{a} are less stable than $\mathbf{C_{red2}}$**-**$\boldsymbol{\mu}\mathbf{CO_2}$**-LM**^{a} and $\mathbf{C_{int}}$**-**$\boldsymbol{\mu}\mathbf{CO_2}$**-LM**^{a}, respectively, by 8.6 and 10.9 kcal/mol (see Figure 3.6), in agreement with the experimental observation of a bridging CO_2 ligand [107, 120]. However, the $\mathbf{tCO_2}$ species could correspond to a potential intermediate along the catalytic pathway. Since the Ni ion is positioned at the end of the substrate channel [98], the **tCO$_2$** species, in which the carbon atom of CO$_2$ completes the distorted tetrahedral coordination geometry of Ni, may play a key role in the binding/dissociation of the CO$_2$ molecule to the C-cluster.

As in the **µCO$_2$** species, in the **tCO$_2$** forms CO$_2$ has a bent geometry with geometrical parameters very similar to those found for **µCO$_2$** adducts; the CO$_2$ angle varies between 122.4° and 125.6° and the C-O1 bond (connecting the carbon atom to the oxygen atom closer to Fe$_u$) is longer than the C-O2 one. Conversely, in small models a less bent CO$_2$ geometry is observed and C-O1 and C-O2 bonds are approximately characterized by the same bond length (Table 3.2). In **tCO$_2$-LM^a** only one oxygen atom (O1) interact favourably with the protein environment, whereas in **µCO$_2$-LM^a** both oxygen atoms of CO$_2$ establish strong H-bonds with the surrounding amino acids residues (Lys563 and His93). The more favourable H-bond network in the **µCO$_2$-LM^a** isomers is probably the main reason for their larger stability with respect to the **tCO$_2$-LM^a** species.

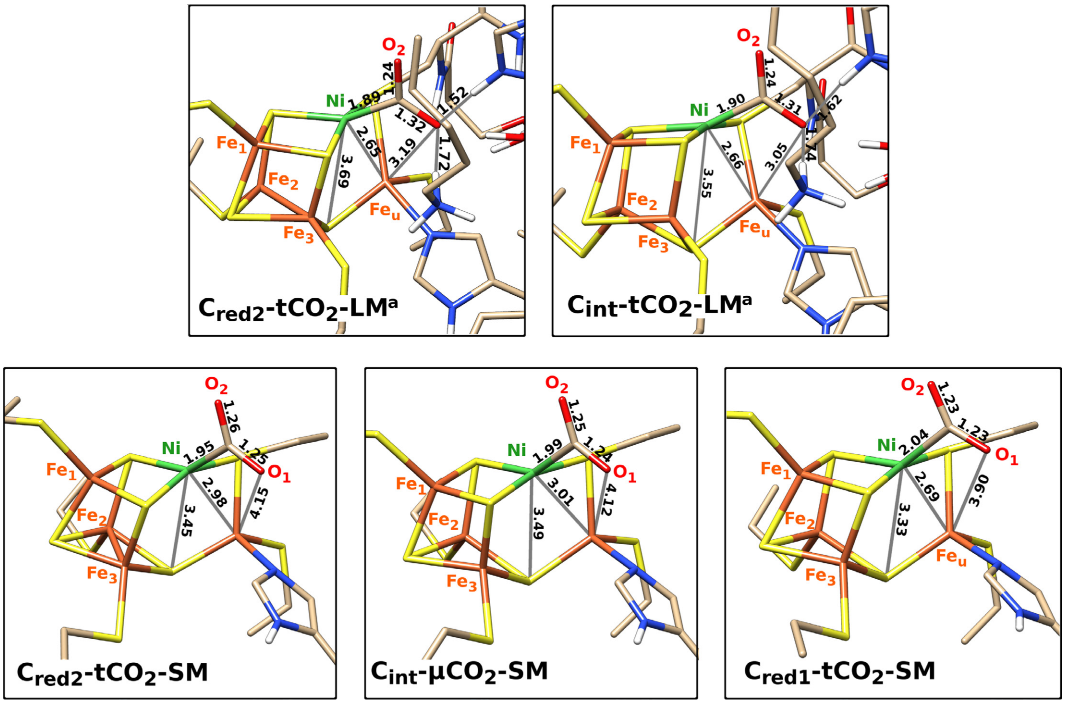


Figure 3.5 Schematic representation of the DFT geometries of LM^a and SM models of the tCO$_2$-bound C-cluster. Noting that **C$_{red1}$-tCO$_2$-LM^a** is unstable and during geometry optimization converges to **C$_{red1}$-µCO$_2$-LM^a**. Selected distances are given in Å.

The crucial role of the protein environment on the relative stability of the two isomers is also highlighted by the reversed order of stability calculated with the small model in which interactions with Lys563 and His93 are not considered; **C$_{red2}$-tCO$_2$-SM**, **C$_{int}$-tCO$_2$-SM** and **C$_{red1}$-tCO$_2$-SM** are predicted to be more stable than **C$_{red2}$-µCO$_2$-SM**, **C$_{int}$-µCO$_2$-SM** and **C$_{red1}$-µCO$_2$-SM** respectively by 1.5, 3.6 and 4.8 kcal/mol (see Figure 3.6).

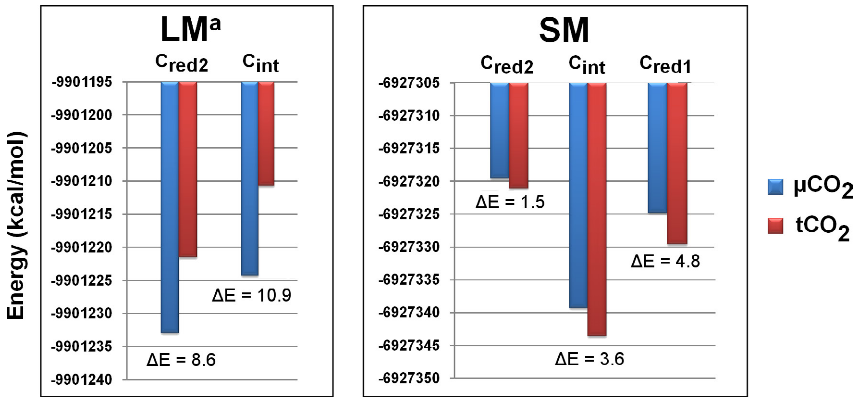


Figure 3.6 Histogram format of the calculated energies (in kcal/mol) of the small (SM) and the large (LM^a) models of the CO_2 adducts of the C-cluster in the C_{red2}, C_{int} and C_{red1} redox states.

Finally, a last CO_2-bound form of the active site has been characterized. X-ray structure of *C. hydrogenoformans* CODH [121], featuring the inhibitor n-butyl isocyanate terminally coordinated to Ni and an hydroxide ligand bound to Fe_u, prompted us to investigate the CO_2 binding to the OH-bound C-cluster. Interestingly, the resulting **tCO_2-OH** adducts (Figure 3.7) be stable complexes in all redox states, using both large and small model. A very similar binding mode to that observed for the n-butyl isocyanate has been predicted for CO_2; in all isomers CO_2 binds to Ni with a distorted tetrahedral coordination geometry, which contrasts with the square-planar geometry observed in the μCO_2-bound C-cluster, and one oxygen atom of CO_2 is at a suitable distance and in a suitable orientation to form hydrogen bonds with His93 and the OH^- ligand on Fe_u, which in turns is H-bonded to Lys563 (see Table 3.2).

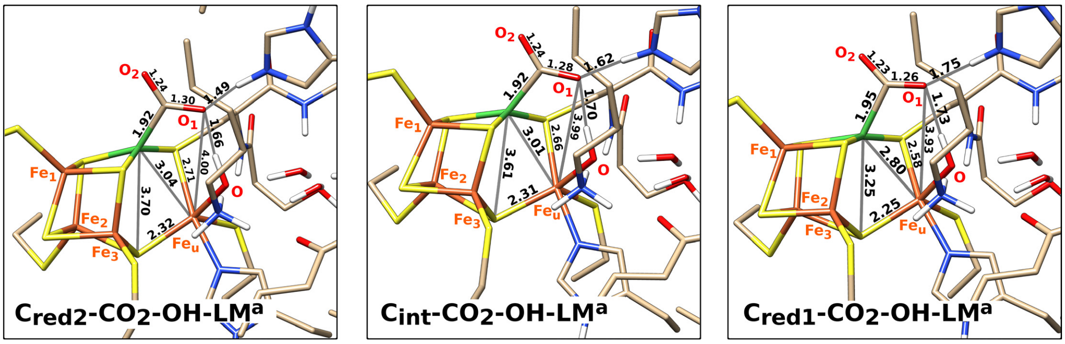


Figure 3.7 Schematic representation of the DFT geometries of the tCO_2-OH isomers of the C-cluster in C_{red2}, C_{int} and C_{red1} states, optimized using the large (LM^a) model. Selected distances are given in Å. For the sake of clarity, aliphatic hydrogen atoms are not shown.

3.2.2 Binding of CO_2 to the C-cluster

Crystallographic characterization of CO_2-bound C-cluster has challenged the previous interpretations of spectroscopic, redox and catalytic properties of the C-cluster. X-ray structures of CO_2-bound C-cluster have been obtained in conditions for which the enzyme should be in the C_{red2} state [107, 120]; however, CO_2 shifts the EPR signal of C_{red1} but not that of C_{red2}, as would be expected if CO_2 binds to this state of the enzyme [187]. On the other hand, as excepted for the binding of CO_2 to C_{red2}, the C_{red1}/C_{red2} reduction potential increases in presence of CO_2 [188].

To explain these contrasting data, it was proposed that CO_2 binds the C_{red2} and C_{red1} states of the C-cluster in two different ways [189]. In the C_{red2} state, CO_2 may bind to the electronically isolated Ni site, without affecting the electronic structure of the C-cluster. Conversely, since CO_2 shifts the C_{red1} EPR signal, which arise predominately from the [Fe_4S_4] moiety [109], CO_2 can bind in this state to both the Ni and an Fe of the cluster. However, in X-ray structures of the CO_2-bound C-cluster in the C_{red2} state the CO_2 ligand bridges the Ni and Fe_u atoms [107, 120], and our calculations show that Lys563 and His93 provide more stability to the **μCO₂** adduct than to the **tCO₂** isomer. Therefore, in order to shed more light on the reactivity of the C-cluster towards CO_2, binding energies of CO_2 to C_{red2}, C_{int} and C_{red1} redox states of the active site have been calculated.

The binding of CO_2 in the μ_2-$\eta^2_{C,O}$ mode to the large model of the C-cluster (LM^a) is always calculated to be energetically favoured. However, the CO_2 binding energy decreases with the oxidation of the C-cluster from C_{red2} to C_{red1}. In particular, the computed binding energies are equal to -35.0, -22.8 and -8.9 kcal/mol for C_{red2}, C_{int} and C_{red1} states, respectively. Comparison of these results with those obtained using the small-size model (see Table 3.4) highlighted the fundamental role of the protein environment in tuning the CO_2 binding to the C-cluster. Indeed, the CO_2 binding energies computed using such small model are significantly smaller than those calculated using the large model. In particular, only the binding of CO_2 to the C_{red2} state is still exoenergetic (-12.9 kcal/mol), whereas the binding of CO_2 to C_{int} and C_{red1} turns out to be endoenergetic by 1.4 and 6.4 kcal/mol, respectively. On the other hand, the trend in the binding energies as a function of the oxidized state is preserved when compared to that observed with the large model.

Table 3.4 Energies differences (in kcal/mol) for binding reaction of CO_2 to the C-cluster, calculated using LM^a, LM^b and SM models of the active site. Superscripts a and b respectively denote double and single protonation of His93. Asterisked energy values refer to the reaction represented in Figure 3.8.

	μCO₂			**tCO₂**		**tCO₂-OH**	
	LM^a	**LM^b**	**SM**	**LM^a**	**SM**	**LM^a**	**SM**
C_{red2}	-35.0	-14.2	-12.9	-26.3	-17.7	+2.8*	-26.2
C_{int}	-22.8	-3.4	+1.4	-12.7	-2.2	+9.5*	-14.1
C_{red1}	-8.9	+10.8	+6.4	-	+4.9	+5.3	-5.0

These results show that the C_{red2} redox state of the enzyme features the highest affinity for CO_2 and that the interaction of this molecule to the C-cluster is assisted by the network of H-bonds formed nearby the active site. In particular, His93 and Lys563 should be crucial for the CO_2 coordination. The elongation of H-bonds between these residues and oxygen atoms of

CO_2 with the oxidation of the active site from C_{red2} to C_{red1} (see Table 3.2) reflects the smaller stabilization of the CO_2-bound adduct by the protein environment.

In order to further investigate the stabilizing interaction due the formation of a H-bond between the O2 atom of CO_2 and the protonated His93, large models of the C-cluster were also optimized modeling His93 as neutral residue (LM^b). As expected, deprotonation at the εN atom strongly affects the binding of CO_2 to the C-cluster (see Table 3.4). Due to the unfavourable interaction between the oxygen atoms of CO_2 and the unprotonated εN atom of His93, which yields the latter atom more than 3.2 Å away from O2, binding energy of CO_2 decreases significantly. It is worth noting that when His93 is singly protonated on the δN atom the CO_2 binding to the C_{red1} state is endoenergetic by 10.8 kcal/mol, whereas when His93 is doubly protonated it is energetically favoured in all redox states. These results confirm the fundamental role of His93 in tuning the stability of the CO_2-bound form of the enzyme. Therefore, we can suggest that this residue, depending on its protonation state, may favor the binding or the release of the CO_2 molecule from the C-cluster. CO_2 can bind to the active site when His93 is doubly protonated, assisted by a strong interaction with the proton on the εN atom. CO_2 may be instead released when the histidine residue is deprotonated on εN. This conclusion is based on the plausibility of different protonation states for His93 during catalysis; this residue was in fact proposed to be involved in proton transfer from/to the C-cluster [97, 98].

Binding energies of CO_2 to the terminal position of the Ni atom are significantly lower than those calculated for **μCO_2** isomers (Table 3.4). However, formation of **C_{red2}-tCO_2-LM^a** and C_{int}-**tCO_2-LM^a** is energetically favored by 26.3 and 12.7 kcal/mol supporting the hypothesis that the **tCO_2** species corresponds to a catalytic intermediate in which a first covalent interaction is established between CO_2 and the metallic cluster. As indicated by the comparison of binding energies computed using the LM^a and SM models, also the terminal binding of CO_2 to Ni is assisted by favourable interactions with the protein environment.

Since binding of CO_2 can also occur through an associative mechanism to the C-cluster in which OH^- is coordinated to Fe_u, reaction energies for formation of **tCO_2-OH** adducts have been calculated (see Table 3.4). However, the binding energy of CO_2 to the large model of the OH-bound C-cluster can be calculated only for the C_{red1} state (+5.3 kcal/mol) because the OH-bound form of the enzyme in C_{red2} and C_{int} states is unstable, resulting in the deprotonation of His93 by the OH^- ligand that is released from Fe_u as H_2O during geometry optimization (see section 3.4). Reaction energies reported in Table 3.4 for C_{red2} and C_{int} states refer to the formation of **C_{red2}-tCO_2-OH-LM^a and C_{int}-tCO_2-OH-LM^a** according to the reaction represented in Figure 3.8. It is worth noting that the formation of the **tCO_2-OH-LM** adducts, in which CO_2 and OH^- bind simultaneously to the C-cluster, is predicted to be an endoenergetic process in all redox states.

In light of these results, an associative mechanism for CO_2 binding to the C-cluster can be ruled out. Therefore, CO_2 binds to the unbound form of the active site according to a dissociative mechanism. In this case, the binding of CO_2 to both C_{red2} and C_{red1} states of the C-cluster is energetically favored, although CO_2 binds preferentially to C_{red2}. The reason why the exposure of the enzyme to CO_2 shifts the EPR signal of C_{red1} but not that of C_{red2} remains to be explained [187].

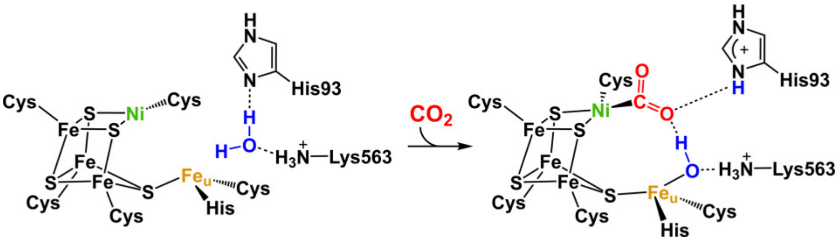


Figure 3.8 Schematic representation of formation reaction of **C_{red2}-tCO_2-OH-LM^a** and **Cint-tCO_2-OH-LM^a**.

3.2.3 Electronic structure of the CO_2-bound C-cluster

As stated in the previous section CO_2 binds to the C-cluster with a bent geometry to which corresponds an elongation of both C-O bonds (see Table 3.2, Figure 3.4 and Figure 3.5). This arrangement is a consequence of the CO_2 reductive activation, induced by the $\eta^1{}_C$ coordination to the Ni atom. The reduction of the C-O bond orders is explained by the Dewar-Chatt-Duncanson scheme [190, 191], as due to the σ-donation from the HOMO of CO_2 to unoccupied orbitals of the C-cluster and to the backdonation from high-energy occupied orbitals of the metallic cluster to the LUMO of CO_2 with a π* anti-bonding character.

A net electron transfer from the C-cluster to the CO_2 ligand is confirmed by the partial NBO charge of the bound CO_2 (see Table 3.5). In the **μCO_2-LM** species, the charge of CO_2 in the C_{red2}, C_{int} and C_{red1} states is equal to -1.04, -0.98 and -0.91, respectively. For the CO_2 ligand in **C_{red2}-tCO_2-LM** and **C_{int}-tCO_2-LM** species, slightly smaller charges, equal to -1.00 and -0.92, respectively, have been calculated. The role of His93 in the electron transfer from the C-cluster to CO_2 is highlighted by the decrease of the negative charge of CO_2 in the species in which this residue is deprotonated at the εN atom. Indeed, in this case, the charge of CO_2 is equal to -0.95, -0.88 and -0.79 for the C_{red2}, C_{int} and C_{red1} states, respectively. As shown in Table 3.5, the charge of the oxygen atom closer to Fe_u (O1), which is involved in stronger interactions with Lewis acid centers of the protein environment is always more negative than that calculated for O2.

The effect of the protein environment in promoting the electron transfer from the cluster to CO_2 is further highlighted by considering the charge of CO_2 in **μCO_2-SM** and **tCO_2-SM**, which, indeed, is less negative than that calculated for the corresponding species optimized using the large model.

The reductive activation of CO_2 corresponds to the oxidation of the C-cluster. A detailed analysis of the electronic structure of CO_2-bound and unbound species allowed us to identify the metal sites which are mainly involved in the oxidation (see Table 3.5). In this respect, due to the well known electron spin and charge delocalization in FeS clusters, analysis of the electronic structure of the C-cluster can be still more insightful considering the net charges of the two subunits $\{Fe_1NiFe_uS_1S_2\}$ (hereafter labelled layer L1) and $\{Fe_3Fe_4S_3S_4\}$ (hereafter labelled layer L2). L1 and L2 correspond, respectively, to the blue and red layers of the BS-1 state reported in Figure 3.2. Note that each layer in which is partioned the C-cluster also includes the covalently bound atoms of the residues coordinating the metal sites contained in it.

Comparing the electronic structure of the **C_{red2}-μCO_2-LM** model, corresponding to X-ray structures of CO_2-bound C-cluster [107, 120], with that of **C_{red2}-LM** unbound complex, it is interesting to note that only the atomic charge on Fe_u is significantly affected by the CO_2 binding, as in **C_{red2}-μCO_2-LM** it is 0.20 higher than that in **C_{red2}-LM**. Conversely, atomic charges on Ni and other Fe atoms do not differ by more than 0.06. Accordingly, the net charges on L1 and L2 suggest that the oxidation of the C-cluster occurs predominantly on the $\{Fe_1NiFe_uS_1S_2\}$ layer. In **C_{red2}-μCO_2-LM**, its charge is in fact 0.75 more positive than that in C_{red2}-LM. An increase of only 0.20 has been instead computed for the charge of the $\{Fe_3Fe_4S_3S_4\}$ moiety. These data suggest that the CO_2 binding to the C_{red2} state of the C-cluster promotes the transfer of one electron from the Fe_u-containing layer to the CO_2 ligand. Hence, the formal **C_{red2}-μCO_2-LM** model can be described as the adduct between the C_{int} redox state of the C-cluster and the anionic CO_2^-. Accordingly, electronic structure of the C-cluster in **C_{red2}-μCO_2-LM** is more similar to that of the unbound form of the active site in the C_{int} state than that in C_{red2}.

Analogously, also in **C_{int}-μCO_2-LM**, **C_{red1}-μCO_2-LM**, **C_{red2}-tCO_2-LM** and **C_{int}-tCO_2-LM**, the CO_2 ligand is reduced mainly at the expense of the $\{Fe_1NiFe_uS_1S_2\}$ layer of the C-cluster (see Table 3.5). However, it is worth noting that in **C_{red1}-μCO_2-LM**, differing from other complexes in which oxidation occurs predominately on Fe_u, the electron transferred to the CO_2 molecule arises mainly from Fe_1. The fact that oxidation promoted by CO_2 binding to C_{red2} and C_{red1} occurs at different metal sites may be the reason why exposure of the enzyme to CO_2 has different effect on EPR signal of C_{red2} and C_{red1}.

Despite in all models the CO_2 computed charge is always close to the unity, the mono-electronic reduction of CO_2 to the $CO_2^{\bullet -}$ radical is ruled out. The bi-eletronic reduction to the formal diamagnetic CO_2^{2-} species is considered more likely, as indicated by the CO_2 Mulliken spin population, which is always equal to zero (see Table 3.5). The CO_2 bi-electronic reduction is also suggested by the comparison of the geometry of the bound CO_2 with those computed for the free CO_2 and its mono- and di-reduced species. The C-O bond lengths and the O1-C-O2 angle of the CO_2 ligand coordinated to large models of the C-cluster, as well as in the true atomic resolution crystallographic structure [120], best match with those of CO_2^{2-} (1.32 Å and 117.9°, respectively). For the $CO_2^{\bullet -}$ radical and the linear CO_2 molecule have been computed shorter C-O distances (1.25 Å and 1.17 Å, respectively) and larger O1-C-O2 angles (135.2° and 180.0°). On the basis of geometric parameters for the CO_2 molecule in the CO_2-bound C-cluster, it is possible conclude that the binding reaction promotes a flow of electrons from the cluster to CO_2, thereby rendering the ligand more similar to the two-electron-reduced carboxylate group than to CO_2. The CO_2 charge of about -1, lower than the expected value of -2, can be attributed to the large charge delocalization provided by the covalent linkage between metal sites, sulfide ions and sulphur atoms of coordinating cysteine residues.

Table 3.5 Computed NBO charges and, in parenthesis, Mulliken spin population for selected atoms of the C-cluster, atoms of the CO_2 ligand and layers in which is partioned the active site. L1 and L2 correspond respectively to the blue and red layers of the BS-1 state of Figure 3.2. Reported values are calculated using the large models (LM[a]) of CO_2-bound and unbound forms of the C-cluster.

	C_{red2}-μCO_2	C_{int}-μCO_2	C_{red1}-μCO_2	C_{red2}-tCO_2	C_{int}-tCO_2	C_{red2}	C_{int}	C_{red1}
Ni	0.50(0.12)	0.51(0.02)	0.54(0.17)	0.49(0.15)	0.50(0.06)	0.47(0.33)	0.53(-0.03)	0.53(0.20)
Fe_u	1.06(3.37)	1.06(3.37)	0.69(3.55)	1.00(3.36)	1.02(3.40)	0.86(3.16)	0.84(3.13)	0.95(3.35)
Fe_1	0.76(3.19)	0.70(3.06)	1.14(3.04)	0.74(3.14)	0.66(2.92)	0.81(3.28)	0.75(3.15)	0.72(3.10)
Fe_2	0.81(-3.17)	0.81(-3.22)	0.73(-3.02)	0.79(-3.15)	0.80(-3.19)	0.76(-3.09)	0.72(-3.05)	0.67(-2.91)
Fe_3	0.67(-2.82)	0.77(-3.17)	0.74(-3.08)	0.67(-2.83)	0.78(-3.17)	0.73(-3.01)	0.74(-3.15)	0.70(-2.99)
C	0.53(0.01)	0.54(0.00)	0.55(0.00)	0.53(0.00)	0.56(0.00)	-	-	-
O1	-0.86(0.02)	-0.84(0.02)	-0.82(0.04)	-0.84(0.00)	-0.82(0.00)	-	-	-
O2	-0.71(0.00)	-0.68(0.00)	-0.64(0.00)	-0.69(0.00)	-0.65(-0.01)	-	-	-
L1	-1.11(7.19)	-0.73(6.90)	-0.17(7.42)	-1.17(7.17)	-0.79(6.83)	-1.86(7.17)	-1.40(6.56)	-0.81(7.13)
L2	-1.49(-6.26)	-1.05(-6.91)	-0.74(-6.51)	-1.50(-6.21)	-1.06(-6.68)	-1.69(-6.34)	-1.39(-6.50)	-1.04(-6.14)
CO_2	-1.04(0.03)	-0.98(0.02)	-0.91(0.04)	-1.00(0.00)	-0.92(-0.02)	-	-	-

3.3 CO-bound forms of the C-cluster

While only two crystallographic structures of CO_2-bound C-cluster have been reported, many structures of CODH that identified, with partial or nearly complete occupancy, a CO molecule bound to the active site have been solved. In particular, the structure of CODH/ACS from *M. thermoacetica* [102] shows a CO ligand at the apical coordination site of Ni and the unique Fe free from coordination from exogenous ligands. Conversely, the structure of *M. barkery* [119] shows a CO bound to the Ni atom in a position adjacent to a OH^-/H_2O ligand coordinated to Fe_u. Analogously to, Ni-CODH from *R. rubrum* [97] and *C. hydrogenoformans* [192] shows an additional ligand to Ni, that has been assigned to CO, and no exogenous ligands on Fe_u, similarly to the structure of *M. thermoacetica.* Interestingly, the 3B52 structure of the CO_2-bound C-cluster [113], discussed in section 3.2.1, features a conformation with 35% occupation with a Ni-bound CO located at the entrance/exit tunnel for gaseous substrates/products [98, 102] and a H_2O molecule interacting with a conserved water network. Since Ni-bound CO and H_2O are potential substrates for CO_2 production, this conformation may correspond to a catalytic intermediate that precedes the μCO_2 adduct.

3.3.1 Geometry of CO-bound C-cluster with and without OH^- bound to Fe_u

The Ni atom in the *M. thermoacetica* 1OAO structure (1.9 Å resolution, Figure 3.9) [102] is four coordinated with a distorted tetrahedral geometry, including three endogenous ligands (two bridging sulfide ions and a Cys residue) and one exogenous ligand modelled as CO. The unique Fe coordinated only by endogenous ligands (a bridging sulfide ion, and His and Cys residues). In this crystal, the Ni ion is either disordered or only partially occupied because an electron density map containing no negative peak at the Ni site is obtained only when Ni occupancy is set to 0.5. Even the $[Fe_3S_4]$ subsite and the unique Fe are best modeled with 0.75 occupancies. As 1OAO, the 3CF4 structure of CO-treated *M. barkery* CODH crystals (2.0 Å resolution, Figure 3.9) [119] shows electron density consistent with a diatomic ligand terminally bound to the Ni atom, that on the basis of the shape of the electronic density and the interactions with the C-cluster and the non-coordinating residues near the active site, was assigned to a CO molecule. In this structure, however, CO is bound to Ni in an unexpected bent conformation, with a Ni-C-O bond angle of 102.8°; the linear binding of CO seems to be sterically impeded by the steric hindrance of a conserved isoleucine residue close to the active site. The 3CF4 structure also shows a monoatomic ligand, compatible with a water/hydroxyl group, terminally bound to Fe_u.

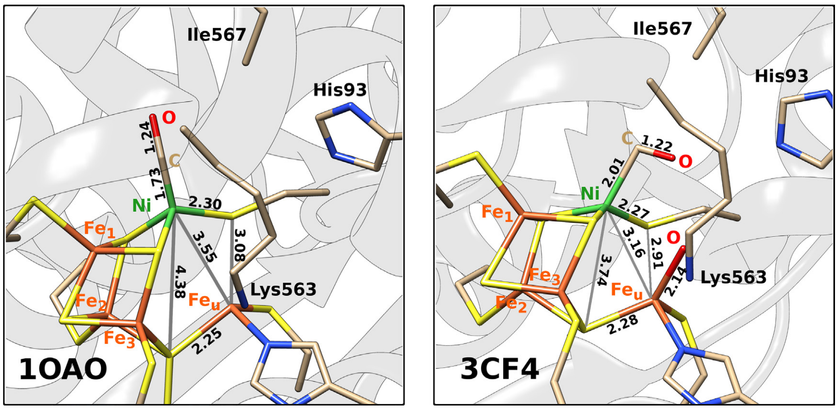


Figure 3.9 Schematic representation of the active site geometry in 3CF4 and 1OAO crystal structures. Selected distances are given in Å. A conserved isoleucine residue (Ile641 in 3CF4 and Ile591 in 1OAO), that is believed to sterically impede linear CO binding, is also shown in 'stick'.

On the basis of such structures, binding of CO has been investigated both as unique exogenous ligand coordinated to the C-cluser, and in the presence of OH bound to the C-cluster. Geometry optimizations of large models of CO-bound forms of the enzyme have been carried out considering both singly and doubly protonated states of His93, although the long distance between the O atom of CO and the εN of His93 in crystal structures (6.00 in 1OAO and 3.34 in 3CF4) suggests that there are not direct interaction between CO and this residue. In fact, optimized geometries of CO-bound C-cluster with no exogenous ligands on Fe_u and in which His93 is doubly protonated are very similar to those found when this residue is not protonated at the εN atom. No significant structural changes are also observed by comparing small and large models of CO-bound C-cluster (see Figure 3.10 and Table 3.6).

As found in the experimental structure 1OAO all computed geometries feature a CO ligand coordinated to the apical position of Ni completing a distorted tetrahedral coordination geometry. However, in optimized structures the Ni-C-O angle is equal to about 170°, whereas it is 145.2° in the crystal. Moreover, the Ni atom, is displaced about 1 Å above the plane formed by the three coordinating S atoms, while in the X-ray structure such displacement is only 0.74 Å. This discrepancy may originate from the movement of the sulphur atom of Cys526 toward the Fe_u ion. Indeed, the Ni-S_{Cys526} is always significantly underestimated in the computed geometries compared the X-ray structures. As for the μCO_2-adducts, oxidation of the active site leads to the progressive contraction of the C-cluster, as indicated by the shortening of the Ni-S4 and Ni-Fe_u distances.

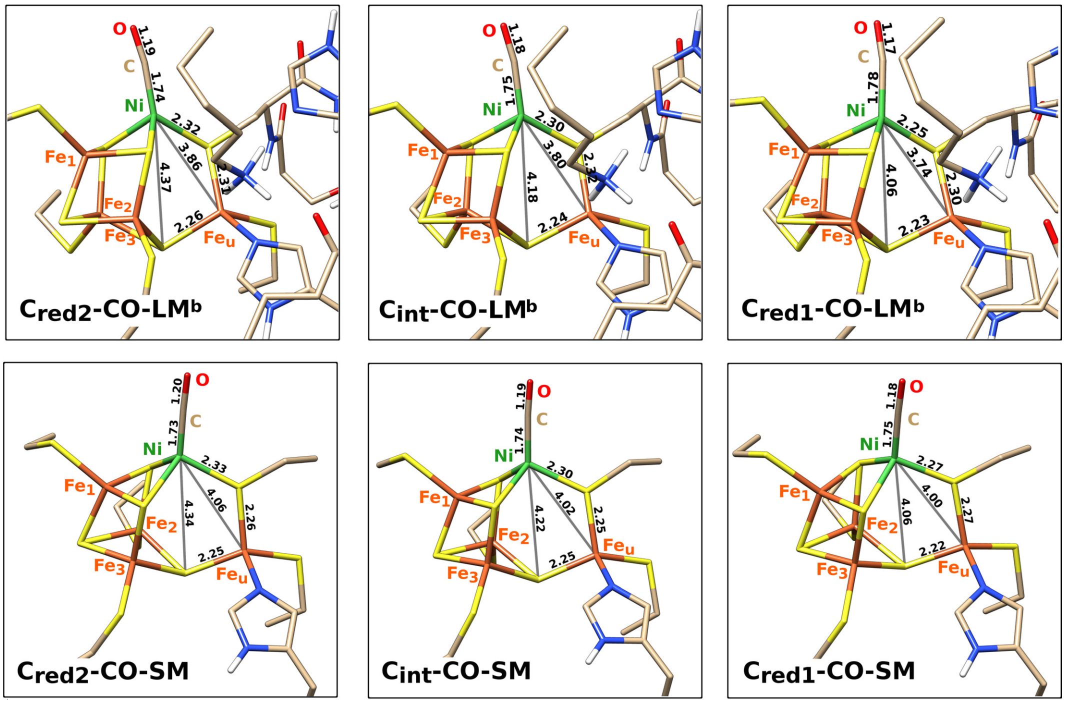


Figure 3.10 DFT geometries of large and small models of the CO-bound C-cluster. Singly protonated state of His93 is considered in the large model (LM^b). Selected distances are given in Å.

Two possible isomers have been investigated for the CO-bound C-cluster with an hydroxide ligand coordinated to Fe_u. In the first isomer CO binds apically to Ni (**apCO-OH**), while in the second isomer CO is terminally coordinated to Ni in a bent conformation (**tCO-OH**). While the first isomer was predicted to be a stable isomer in all redox states using the small model of the active site, only the C_{red1} state was characterized as a genuine minimum on the PES using the large model, independently of the protonation state of His93; **C_{int}-apCO-OH-LM** and **C_{red2}-apCO-OH-LM**, during geometry optimization, converged to species in which the hydroxide ligand was protonated, forming H_2O, through the transfer of a proton from Lys563 (see Figure 3.11), or His93 in the cases it was originally protonated at εN. Conversely, **tCO-OH** adducts were identified as stable isomers only using the large model in which His93 is doubly protonated (see Figure 3.12). Therefore, formation of an H-bond between the oxygen atom of CO and the proton on the δN atom of His93 may be crucial for the stabilization of the bent geometry of CO in this isomer.

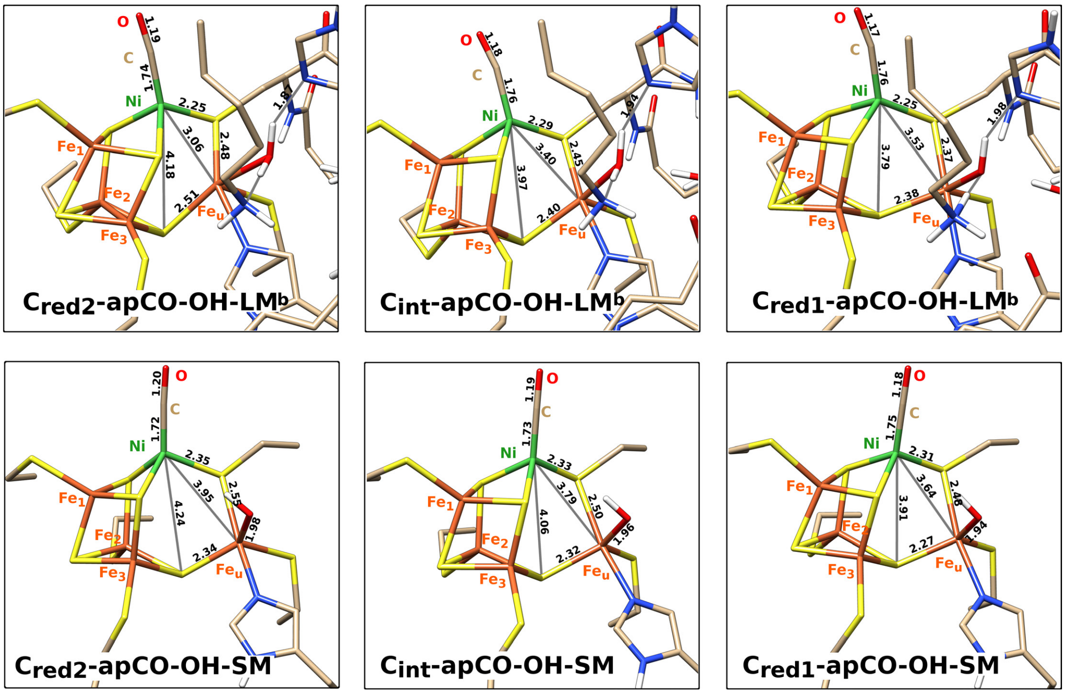


Figure 3.11 Optimized geometries of large and small models of apCO-OH adducts of the C-cluster. Singly protonated His93 is considered in the large model (LM^b). Selected distances are given in Å.

In **tCO-OH** species S1-Ni-S_{Cys256} and S2-Ni-C angles are significantly larger than in the corresponding **apCO-OH** isomers (see Table 3.6), indicating that the distorted tetrahedral geometry of Ni is approaching the square planar geometry. The Ni-C-O angle is equal to 146.7, 147.5 and 154°, respectively in C_{red2}, C_{int} and C_{red1}, in good agreement with the value reported in the experimental 1OAO structure (145.2°), but considerably lower than that observed in the 3CF4 crystal (102.8°) in which a hydroxide is bound to Fe_u as in **tCO-OH** species.

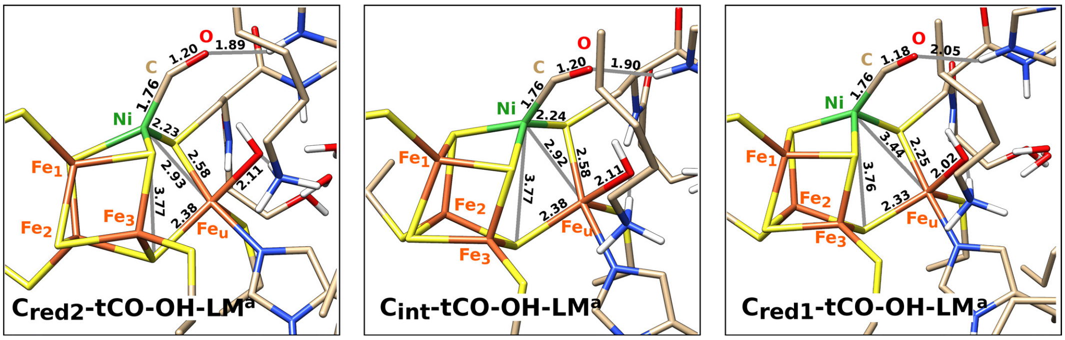


Figure 3.12 DFT geometries of tCO-OH adducts of the C-cluster, obtained using the large model of the active site in which His93 is doubly protonated.. Selected distances are given in Å.

The different arrangement of CO found in the computed structures with respected to that of the 3CF4 structure should be due to the different interactions established by CO; it forms an H-bond with His93 in **tCO-OH** species, whereas in the 3CF4 structure it is H-bound to the hydroxide ligand on Fe$_u$, as indicated by O$_{CO}$-εN$_{His93}$ and O$_{CO}$-O$_{OH}$ distances (see Table **3.6**). The peculiar CO bent geometry found in 3CF4 is not likely to be stable. Indeed, it is not predicted to be a stable isomer by our calculations; in all redox states it converges to **tCO-OH** or **apCO-OH** during geometry optimizations. In addition, the bent geometry of CO calculated in the **tCO-OH** species is less stable than that in which the CO ligand moves to the apical position as, indeed, **C$_{red1}$-tCO-OH** is 3.3 kcal/mol higher in energy than **C$_{red1}$-apCO-OH**.

Table 3.6 Selected distances (Å) and angles (°) for large models of the CO-bound C-cluster, with and without OH$^-$ coordinated to the Fe$_u$ atom, in the C$_{red2}$, C$_{int}$ and C$_{red1}$ states and for 1OAO and 3CF4 X-ray structures. LM^a and LM^b denote respectively the doubly and singly protonated states of His93.

	apCO-LM^b			**X-ray**	**apCO-OH-LM^b**			**tCO-OH-LM^a**			**X-ray**
	C$_{red2}$	**C$_{int}$**	**C$_{red1}$**	**1OAO**	**C$_{red2}$**	**C$_{int}$**	**C$_{red1}$**	**C$_{red2}$**	**C$_{int}$**	**C$_{red1}$**	**3CF4**
Ni-Fe$_u$	3.86	3.80	3.74	3.55	3.06	3.40	3.53	2.93	2.92	3.44	3.16
Ni-S$_{Cys526}$	2.32	2.30	2.25	2.30	2.25	2.29	2.25	2.23	2.24	2.27	2.27
Ni-S4	4.37	4.18	4.06	4.38	4.18	3.97	3.79	3.77	3.77	3.76	3.74
S2-Ni-$_{Cys526}$	102.2	108.8	108.6	129.7	107.4	118.4	122.1	137.5	138.4	118.2	143.5
S1-Ni-C	111.8	108.9	111.0	108.6	112.3	108.7	108.6	140.8	140.7	145.9	125.8
Fe$_u$-S4	2.26	2.24	2.23	2.25	2.51	2.40	2.38	2.38	2.38	2.33	2.28
Fe$_u$-S$_{Cys295}$	2.43	2.40	2.30	2.24	2.60	2.39	2.36	2.41	2.27	2.35	2.35
Fe$_u$-N$_{His261}$	2.14	2.15	2.14	2.07	2.39	2.29	2.18	2.27	2.41	2.02	2.04
Fe$_u$-S$_{Cys526}$	2.31	2.32	2.30	3.08	2.48	2.45	2.37	2.58	2.58	2.33	2.91
Ni-Fe$_u$-S4	86.4	83.3	81.7	95.2	96.8	84.6	77.2	90.0	89.9	78.8	85.4
Ni-C	1.74	1.75	1.78	1.73	1.74	1.76	1.76	1.76	1.76	1.76	2.01
C-O$_{CO}$	1.19	1.18	1.17	1.24	1.19	1.18	1.17	1.20	1.20	1.18	1.22
Ni-C-O$_{CO}$	171.8	170.4	169.0	145.2	171.7	164.4	165.2	146.7	147.5	154.5	102.8
O-δC$_{Ile567}$	3.47	3.45	3.47	3.50	3.51	3.46	3.38	2.98	3.04	4.18	3.28
C-δC$_{Ile567}$	3.66	3.69	3.84	3.98	3.81	3.62	3.55	3.54	3.57	3.34	3.29
Fe$_u$-O$_{OH}$	-	-	-	-	2.40	2.19	2.00	2.11	2.11	2.02	2.14
S4-Fe$_u$-O$_{OH}$	-	-	-	-	156.2	148.0	141.8	149.6	149.7	138.8	127.9
O$_{CO}$-O$_{OH}$	-	-	-	-	6.19	6.17	5.98	3.27	3.28	3.51	2.75
O$_{OH}$-N$_{Lys563}$	-	-	-	-	2.71	2.63	2.60	2.63	2.63	2.65	2.68
O$_{OH}$-H$_{Lys563}$	-	-	-	-	1.04	1.07	1.47	1.51	1.50	1.56	-
O$_{CO}$-εN$_{His93}$	6.40	6.28	6.38	6.00	5.81	5.92	5.73	2.82	2.82	2.93	3.34
O$_{CO}$-εH$_{His93}$	-	-	-	-	-	-	-	1.89	1.90	2.05	-
H$_{OH}$- εN$_{His93}$	-	-	-	-	1.87	1.94	1.98	3.20	3.21	3.30	-

According to our results, infrared spectroscopy had suggested that there is no single, stable site for CO binding to the C-cluster [193]. We can argue that CO initially binds to the apical coordination site of Ni, located at the end of the substrate channel [98]. Subsequent interaction between CO and His93 may trigger isomerisation reaction of **apCO-OH** to **tCO-OH**, which should correspond to a potential intermediate along the catalytic pathway, and in which CO approaches the hydroxide ligand coordinated to Fe_u. A further CO shift is expected to follow. Gong et al. proposed that before the nucleophilic attack of the hydroxide ion on CO, the CO carbon moves toward the OH^- ligand, causing Ni to switch from the tetrahedral-like geometry observed in the 3CF4 structure to a more square planar geometry, while the CO oxygen remains in its original position [119]. Such a "carbon shift" causes CO to be bent in the opposite direction, allowing the carbon to be closer to the attacking hydroxide, without movement of the CO oxygen. This proposal stemmed from the observation that when the structure of the CO-bound active site is compared with that of the CO_2-bound one, the position of the CO oxygen remains essentially constant after conversion to CO_2, while the CO_2 carbon was shifted, modifying the Ni coordination geometry.

3.3.2 CO binding to the C-cluster

Since CO shifts the EPR signal of C_{red1} but not that of C_{red2}, it was proposed that CO binds and is oxidized by the enzyme in the C_{red1} state and that it does not bind to C_{red2} [187, 194]. However, it was also proposed that CO can bind noncatalytically to the electronically isolated Ni atom of C_{red2} without affecting the electronic nature of the C-cluster [189].

According to our calculation, CO binding to the unbound form of the C-cluster is energetically favoured in all redox states (see Table 3.7). In addition, using both the small and the large models of the active site, binding of CO to the more electron-rich C-cluster in the C_{red2} state is calculated to be more favoured than the binding of CO to C_{red1}, as excepted considering the π-acceptor nature of this ligand. These results are consistent with the binding of CO to both C_{red1} and C_{red2} states as proposed by Anderson and Lindahl [189]. CO binding energies calculated with the large model in which His93 is doubly protonated (LM^a) are significantly smaller than those obtained with singly protonated His93 (LM^b). In particular, CO binding to LM^a is less favoured than CO binding to LM^b by 9.7, 14.8, and 9.3 kcal/mol in C_{red2}, C_{int} and C_{red1}, respectively. Accordingly, the large distance between the CO oxygen and the εN-nitrogen of His93 observed in the CO-bound C-cluster X-ray structures suggests the singly protonated state for this residue [102, 119]. Therefore, CO binding is hindered by protonation of the εN atom of His93. CO binding energies similar to those obtained with LM^b have been calculated using the small model, suggesting that protein environment doesn't have a stabilizing effect for the CO adducts of the C-cluster. Indeed, no significant interactions have been identified in CO-LM species (Figure 3.10).

Our results indicate that both CO and CO_2 ligands can bind to the unbound form of the enzyme in the C_{red2} and C_{red1} states. However, when His93 is doubly protonated, the C_{red2} redox state shows an higher affinity for CO_2; binding energies of CO and CO_2 to **C_{red2}-LM^a** are indeed -25.5 and -35.0 kcal/mol, respectively. Conversely, if His93 is singly protonated at the δN atom, C_{red2} binds preferentially CO (-35.2 kcal/mol) than CO_2 (-14.2 kcal/mol). On the other hand, the binding of CO to C_{red1} is energetically more favoured than that of CO_2, independently of the protonation state of His93; CO and CO_2 binding energies to **C_{red1}-**

LM^a/C_{red1}-LM^b are respectively -14.3/-23.6 and -8.9/+10.8 kcal/mol (see Figure 3.13). Interestingly, the apical binding of CO to the Ni atom of the C-cluster in the C_{red1} state is energetically favored also when a hydroxide is terminally coordinated to Fe_u (see Table 3.7). However when His93 is doubly protonated, the binding of CO to the OH-bound C-cluster (-3.3 kcal/mol) is energetically less favoured than the binding of CO to the unbound form of the enzyme (-14.3 kcal/mol). On the other hand, when His93 is singly protonated, the binding energy of CO in the presence or absence of the hydroxide ligand is very similar (-22.6 and -23.6 kcal/mol, respectively).

Table 3.7 Energies differences (in kcal/mol) for binding reaction of CO to the C-cluster, calculated using the large (LM) and small (SM) model of the active site. The **LM^a** and **LM^b** labels denote the doubly and singly protonated states for His93, respectively. Binding energies in parenthesis also involve reaction energy associated with deprotonation of H_2O by the singly protonated His93 residue and migration of the resulting hydroxide to the Fe_u site.

	apCO			**apCO-OH**			**tCO-OH**
	LM^a	**LM^b**	**SM**	**LM^a**	**LM^b**	**SM**	**LM^a**
C_{red2}	-25.5	-35.2	-32.8	-	-	-29.8	(-3.8)
C_{int}	-21.1	-35.9	-20.9	-	-	-26.1	(+4.1)
C_{red1}	-14.3	-23.6	-16.8	-3.3	-22.6	-22.4	-0.1

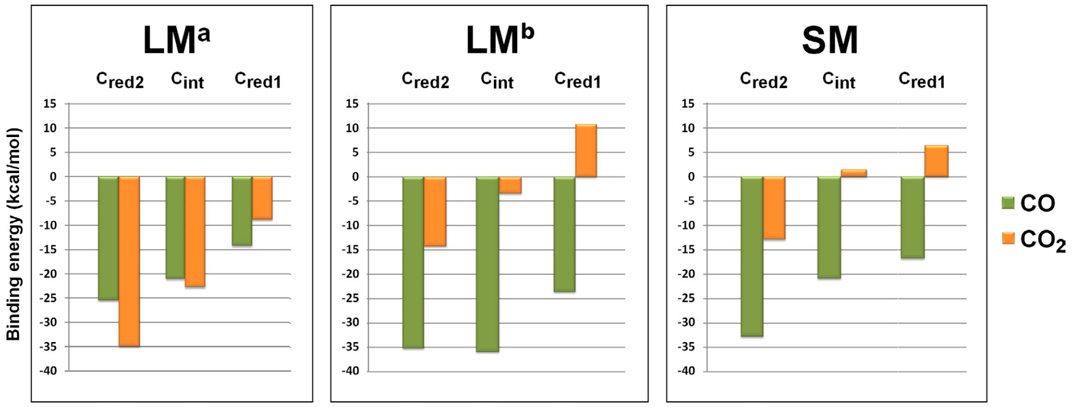


Figure 3.13 Histogram format of of CO and CO_2 binding energies (in kcal/mol) to the C_{red2}, C_{int} and C_{red1} redox states of the C-cluster, calculated using the LM^a and LM^b models of the active site.

3.3.3 Electronic structure of the CO-bound C-cluster

The same analysis of the electronic structure carried out for the CO_2-adducts, has been carried out for the CO-bound forms of the C-clsuter. (see Table 3.8). Interestingly, the CO ligand is always negatively charged when bound to the C-cluster. The charge of CO varies between -0.27 and -0.63 indicating a significant electron transfer from the C-cluster to the CO ligand. The comparison of computed charges for the active site with and without CO ligand (see Table 3.5 and Table 3.8) shows a partial oxidation of the Fe_u atom when CO is bound to the C-cluster. As expected, the value of the CO charge depends on the redox state of the C-cluster;

by reducing the active site, the negative charge of CO increases. The CO charge is also affected by the CO coordination mode; the CO terminally bound to the Ni atom with a bent geometry is always more negative with respect to the apical CO. Interestingly, in the **C_{red2}-tCO-OH** species, the charge on CO (-0.63) is even more negative than that computed for the hydroxide ion (-0.59). Clearly, the activation of the CO ligand is correlated to its geometry. Longer C-O distances and lower Ni-C-O angles have been determined for the CO ligands with a more negative charge. In particular, CO in **C_{red2}-tCO-OH**, with the highest negative charge (-0.63), features the longest C-O bond (1.20 Å) and lowest Ni-C-O angle (146.7°). The electron transfer from the C-cluster to the CO ligand should be critical for the CO in order to adopt the bent geometry necessary for the carbon atom to approach the oxygen atom of the hydroxide ligand. It is worth noting that the nucleophilic attack of the Fe_u-bound OH^- on CO is also facilitated by the more electrophilic character of the carbon atom in CO ligands with a bent geometry. For the latter is in fact always calculated a more positive charge than that observed for the carbon atom in the apical CO ligands.

Table 3.8 Computed NBO charges and, in parenthesis, Mulliken spin population for selected atoms of large models of CO-bound forms of the C-cluster, in the C_{red2}, C_{int} and C_{red1} states. L1 and L2 refers respectively to atoms composing the blue and red layers of the BS-1 state of Figure 3.2. LM^a denotes he doubly protonated state of His93.

	apCO-LM^a			**apCO-OH-LM^a**	**tCO-OH-LM^a**		
	C_{red2}	**C_{int}**	**C_{red1}**	**C_{red1}**	**C_{red2}**	**C_{int}**	**C_{red1}**
Ni	0.53(0.29)	0.54(0.17)	0.56(0.26)	0.57(0.25)	0.56(0.18)	0.57(-0.05)	0.57(0.15)
Fe_u	1.03(3.40)	1.04(3.40)	1.08(3.51)	1.19(3.58)	1.09(3.41)	1.12(3.38)	1.18(3.57)
Fe_1	0.81(3.23)	0.76(2.98)	0.71(3.06)	0.69(3.04)	0.75(3.15)	0.73(3.12)	0.69(3.06)
Fe_2	0.64(-3.23)	0.73(-3.14)	0.73(-3.00)	0.76(-3.07)	0.72(-2.93)	0.77(-3.11)	0.72(-2.98)
Fe_3	0.83(-2.85)	0.82(-3.23)	0.70(-2.99)	0.78(-3.09)	0.83(-3.19)	0.84(-3.23)	0.79(-3.13)
C_{CO}	0.14(-0.02)	0.18(-0.02)	0.22(0.00)	0.21(0.00)	0.06(-0.04)	0.13(-0.09)	0.19(-0.02)
O_{CO}	-0.57(-0.02)	-0.53(-0.02)	-0.49(-0.02)	-0.51(0.06)	-0.69(-0.01)	-0.63(-0.02)	-0.58(0.00)
O_{OH}	-	-	-	-1.18(-0.01)	-1.09(0.02)	-1.14(0.04)	-1.14(0.07)
H_{OH}	-	-	-	0.47(0.00)	0.50(0.00)	0.48(0.00)	0.47(0.00)
L1	-1.58(7.32)	-1.22(6.85)	-0.72(7.33)	-0.66(7.44)	-1.40(7.19)	-1.06(6.79)	-0.57(7.21)
L2	-1.64(-6.34)	-1.22(-6.82)	-0.88(-6.35)	-1.09(-6.52)	-1.70(-6.40)	-1.41(-6.77)	-1.11(-6.44)
CO	-0.43(-0.04)	-0.35(-0.03)	-0.27(-0.02)	-0.30(0.06)	-0.63(-0.05)	-0.51(-0.11)	-0.39(-0.02)
OH	-	-	-	-0.71(-0.01)	-0.59(0.02)	-0.66(0.04)	-0.68(0.07)

3.4 C_{red1} and C_{red2} states: hydroxide, hydride or no ligands bound to the C-cluster?

With the aim of contributing to shed more light on the structure and the electronic properties of C_{red1} and C_{red2}, a detailed analysis of geometry and electronic nature of the C-cluster has been carried out on structures that in previous works were tentatively assigned to these species. In particular, unbound and OH-bound forms of the active site have been investigated for C_{red1}, C_{int} and C_{red2}. For the latter state, also an hydride has been considered as putative active ligand of the C-cluster as suggested by the mechanism proposed by Fontecilla *et al.* [113].

On the basis of experimental data and previous proposal, two species with an hydroxide ligand coordinated to the C-cluster have been investigated. In the first species (hereafter, labelled **tOH**) OH$^-$ is terminally coordinated to Fe_u while in the second (indicated as **bOH**) OH$^-$ occupies the bridging position between Ni and Fe_u. Formation of the first isomer is supported by the crystal structures of *C. hydrogenoformans* CODH [107] and *M. thermoacetica* CODH/ACS [195] in which a OH$^-$/H_2O ligand is terminally bound to the unique Fe site completing a distorted tetrahedral geometry of Fe_u (structures 3B53 and 3I01 in Figure 3.14, respectively). Bridging OH$^-$ coordination is instead suggested for the C_{red1} state by hybrid quantum mechanical/molecular mechanical techniques [113].

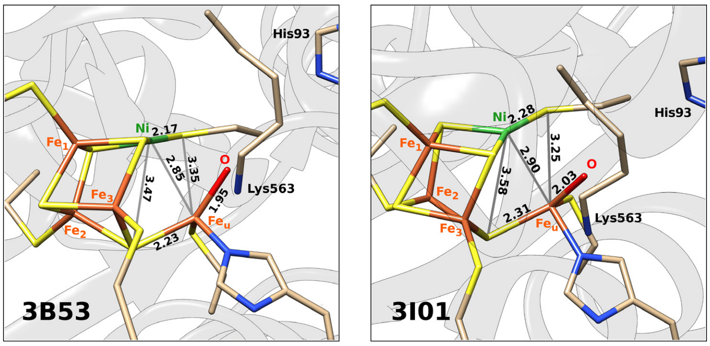


Figure 3.14 Schematic representation of the active site geometry in 3B53 and 3I01 crystal structures. Selected distances are given in Å.

Both isomers of the OH-bound C-cluster have been identified as genuine minima on the PES by using the small model of the active site. However, in all redox states, **bOH** species are slightly less stable than the **tOH** isomers (see Figure 3.15); in particular **C_{red2}-bOH-SM**, **C_{int}-bOH-SM** and **C_{red1}-bOH-SM** are less stable than **C_{red2}-tOH-SM**, **C_{int}-tOH-SM** and **C_{red1}-tOH-SM** by 3.5, 3.0 and 2.9 kcal/mol. On the other hand, using the large model LM^a, only **C_{red1}-tOH-LM^a** has been identified as stable isomer. All other species are unstable since during geometry optimizations the proton of His93 is transferred to the OH$^-$ ligand, which further dissociate from the active site as water molecule. As shown in Figure 3.15, the release of water is favoured by the formation of a strong H-bond network with Lys563, His93 and a conserved water. The OH$^-$ ligand in **C_{red2}-tOH-LM^a** and **C_{int}-tOH-LM^a** dissociates from the C-cluster as water molecule also in the models in which His93 is singly protonated at δN (LM^b). In this case is the Lys563 residue to act as a proton donor for the protonation of the OH$^-$ ligand.

On the basis of the results presented above, only the C_{red1} state is compatible with the OH-bound form of the C-cluster, since in the C_{int} and C_{red2} redox states the hydroxide ligand dissociates from the cluster as a water molecule. In the C_{red1} state the hydroxide ligand is terminally coordinated to the Fe$_u$ atom, completing it distorted tetrahedral geometry, in agreement with the 3B53 X-ray structure [107], which has been assigned to the C$_{red1}$ state of the C-cluster. Bridging coordination of OH$^-$ between Ni and Fe$_u$, as proposed for C$_{red1}$ by Fontecilla-Camps and co-workers [113], does not correspond to a stable isomer in our calculations.

The crystal structure showing a OH/H$_2$O ligand bound to Fe$_u$ (PDB code 3B51), solved at a potential compatible with the C$_{red2}$ state (-600 mV) is apparently in contrast with our results [107]. However, a subsequent refinement of the crystallographic data showed that such structure more confidently corresponds to a mixture of the OH-bound C-cluster in the C$_{red1}$ state (structure 3B53) and CO$_2$-bound C$_{red2}$ state (structure 3B52) [113]. The absence of the hydroxide ligand in the C$_{red2}$ state is also supported by spectroscopic experiments; the loss in C$_{red2}$ of the strong ENDOR signal observed for the C$_{red1}$ state was in fact attributed to the absence of the OH$^-$ ligand coordinated to the active site [118]. Release of the hydroxide ligand in the C$_{red2}$ state was also suggested by QM/MM calculations [113].

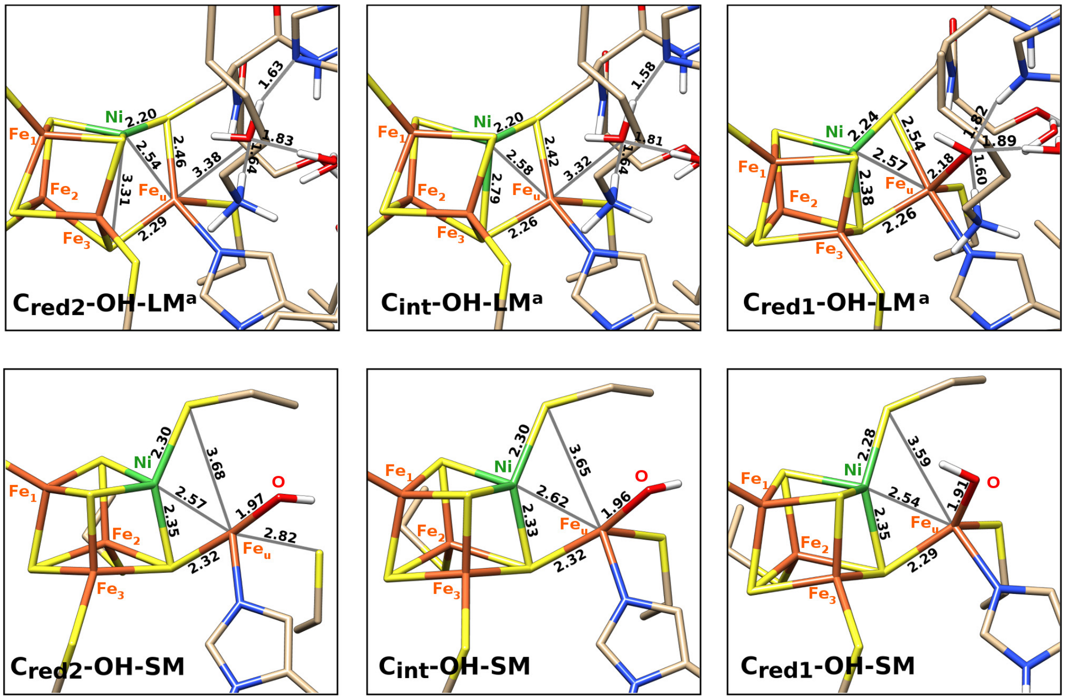


Figure 3.15 Schematic representation of the DFT geometries of large and small models of the tOH-bound C-cluster. Selected distances are given in Å. Aliphatic hydrogen atoms are not shown.

The possibility that no exogenous ligands are contained in C$_{red1}$, C$_{int}$ and C$_{red2}$ was subsequently investigated by optimizing unbound forms of the C-cluster with the small and the large model of the active site (see Figure 3.16). Interestingly, the geometry of the 1MGJ X-ray structure from *M. thermoacetica* containing only endougenous ligands (see Figure 3.17) [103] is better reproduced by the **C$_{red2}$-LM^a** model (see Table 3.9). For instance, the Ni-S4 and Ni-Fe$_u$ distances, respectively equal to 2.99 and 2.57 Å, are in perfect agreement with the experimental ones (2.99 and 2.61 Å).However, in all redox states, the sulphur atom of Cys526 is coordinated to the Fe_u ion. It should be noted that this structural rearrangement is also observed in optimized geometries of CO_2- and CO-bound forms of the C-cluster. This discrepancy can be attributed to the large flexibility of the active site, rather than to the inadequacy of the model used. Since the C-cluster is coordinated to two different protein domains, and the orientation of these domains is not fixed with respect to each other, a slight movement of one domain with respect to the other could compress or expand the C-cluster, altering side-chain metal interactions [97, 114]. Such flexibility is confirmed by the 1JQK X-ray structure in which the Fe_u-S distance is about 2.4 Å and by the higher B-factors values for the C-cluster when compared to those of other metallic clusters [97]. Although potentially important in catalysis, flexibility could be one source of protein sample heterogeneity that complicates X-ray and spectroscopic data analysis [109].

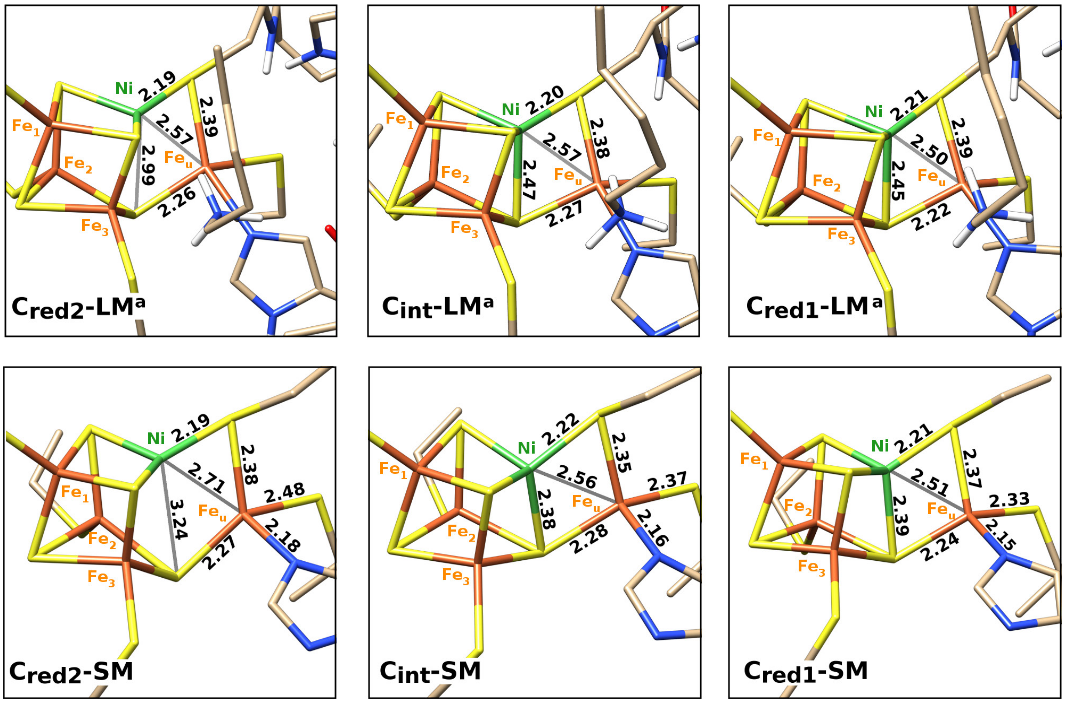


Figure 3.16 Schematic representation of the DFT geometries of large (LM^a) and small models of unbound forms of the C-cluster. Selected distances are given in Å.

Finally, the model of the C_{red2} state in which an hydride is bound to the C-cluster has been investigated. The hypothesis that H^- may be coordinated to C_{red2} has been proposed by Fontecilla-Camps and co-workers. Since the unusual coordination of Ni and Fe_u in the 1MJG structure (Figure 3.17), they initially suggest the presence of a bridging hydride between these atoms [196], not detectable at 2.2 Å resolution. Afterwards, on the basis of QM/MM calculations, the same authors proposed a terminal hydride ligand on Ni [113]. However, only the bridging isomer is predicted to be a stable conformation by our calculations. During geometry optimization of the terminal isomer, the hydride moved to a distances of 1.61 and 1.74 Å from Ni and Fe$_u$, respectively (**Cred$_2$-H-LM^b** in Figure 3.17). The hydride binding to the C-cluster leads to a shorter Ni-Fe$_u$ distance (2.49 Å) with respect to that in the X-ray 1MJG structure (2.61 Å) and that calculated in **Cred$_2$-LM** (2.57 Å). Moreover, the S4 atoms moves away from Ni causing an elongation of the Ni-S4 distance (3.19 Å) with respect to the unbound form of the C-cluster (see Table 3.9). Interestingly, while the Fe$_u$-S(Cys526) distance is generally underestimated in the species discussed above, it is significantly longer in **Cred$_2$-H-LM^b** (2.58 Å).

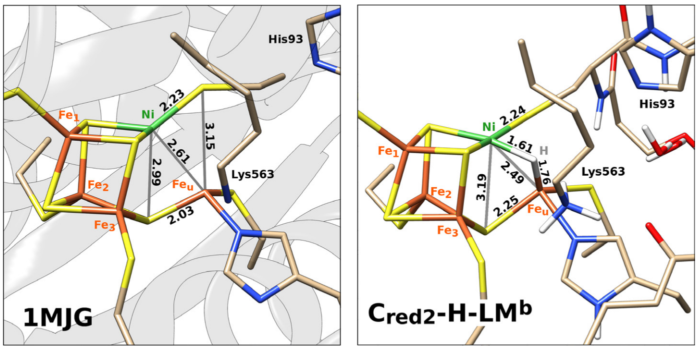


Figure 3.17 Schematic representation of the active site geometry in 1MJG X-ray structure and DFT large model of the C-cluster bound to an hydride ligand in the C$_{red2}$ state (**C$_{red2}$-H-LM^b**). Selected distances are given in Å.

Table 3.9 Selected distances (Å) and angles (°) for large models (LM^a and LM^b) of OH-bound, unbound and H-bound C-cluster and for X-ray structures 3B53, 3I01 and 1MJG.

	C$_{red1}$-tOH-LM^a	3B53	3I01	C$_{red2}$-LM^a	C$_{int}$-LM^a	C$_{red1}$ -LM^a	C$_{red2}$-H-LM^b	1MJG
Ni-Fe$_u$	2.57	2.85	2.90	2.57	2.57	2.50	2.49	2.61
Ni-S$_{Cys526}$	2.24	2.17	2.28	2.19	2.20	2.21	2.24	2.23
Ni-S4	2.38	3.47	3.58	2.99	2.47	2.45	3.19	2.99
Fe$_u$-S4	2.26	2.23	2.31	2.26	2.27	2.22	2.25	2.03
Fe$_u$-S$_{Cys295}$	2.36	2.90	2.37	2.53	2.44	2.37	2.36	2.33
Fe$_u$- N$_{His261}$	2.19	1.82	2.03	2.24	2.18	2.19	2.19	1.95
Fe$_u$-S$_{Cys526}$	2.54	3.35	3.25	2.39	2.38	2.39	2.58	3.15
Fe$_u$-O	2.18	1.95	2.03	-	-	-	-	-
Fe$_u$-O-H	87.2	-	-	-	-	-	-	-
S4-Fe$_u$-O	129.7	139.2	132.8	-	-	-	-	-
O-N$_{Lys563}$	2.70	2.51	2.77	-	-	-	-	-
O-H$_{Lys563}$	1.60	-	-	-	-	-	-	-
Ni-H	-	-	-	-	-	-	1.61	-
Fe$_u$-H	-	-	-	-	-	-	1.76	-

In order to identify which structure among those proposed for C$_{red1}$ (unbound and hydroxide-bound C-cluster) and C$_{red2}$ (unbound and hydride-bound C-cluster) shows the best agreement with spectroscopic results and to contribute to disclose electronic properties of the C-cluster, charge on metallic atoms and layers composing the C-cluster have been computed for different redox states and forms of the enzyme (see Table 3.10).

Table 3.10 Computed NBO charges and, in parenthesis, Mulliken spin population for selected atoms of OH-bound, unbound and hydride-bound forms of large models of the C-cluster. L1 and L2 correspond respectively to the blue and red layers of the BS-1 state of Figure 3.2.

	C$_{red1}$-tOH-LM^a	**C$_{red1}$-LM^a**	**C$_{int}$-LM^a**	**C$_{red2}$-LM^a**	**C$_{red2}$-H-LM^b**
Ni	0.53(0.12)	0.53(0.20)	0.53(-0.03)	0.47(0.23)	0.52(0.10)
Fe$_u$	1.12(3.52)	0.95(3.35)	0.84(3.13)	0.86(3.16)	0.95(3.27)
Fe$_1$	0.73(3.13)	0.72(3.10)	0.75(3.15)	0.81(3.28)	0.76(3.20)
Fe$_2$	0.70(-2.99)	0.67(-2.91)	0.72(-3.05)	0.76(-3.09)	0.74(-2.99)
Fe$_3$	0.73(-3.07)	0.70(-2.99)	0.74(-3.15)	0.73(-3.01)	0.72(-2.94)
O	-1.19(0.05)	-	-	-	-
H(OH)	0.48(0.00)	-	-	-	-
L1	-0.83(7.26)	-0.81(7.13)	-1.40(6.56)	-1.86(7.17)	-1.16(7.10)
L2	-1.14(-6.35)	-1.04(-6.14)	-1.39(-6.50)	-1.69(-6.34)	-1.41(-6.11)
OH	-0.72(0.05)	-	-	-	-
H(hydride)	-	-	-	-	-0.31(-0.01)

Interestingly, the binding of an hydroxide ligand to the C$_{red1}$ state does not affect the electronic structure of the C-cluster, as indicated by the very similar charges computed for **C$_{red1}$-tOH** and **C$_{red1}$**. According to experimental results, atomic charges on iron atoms in both species suggest the assignment of C$_{red1}$ to a species containing three Fe(II) and one Fe(III) ions. The positive charge of Fe$_u$ is in fact significantly higher than that of Fe$_1$, Fe$_2$ and Fe$_3$; it is equal to 1.12 in **C$_{red1}$-tOH** and 0.95 in C$_{red1}$, whereas charges of about 0.7 are calculated for Fe$_1$, Fe$_2$ and Fe$_3$. However, the Fe(III) assignment to Fe$_u$ is in contrast to the interpretation given by Mössbauer experiments that assign the Fe(II) oxidation state for this atom [109].

The comparison of the atomic charges of **C$_{red1}$-tOH** and **C$_{red1}$** with those of **C$_{int}$**, suggests that the one-electron reduction of C$_{red1}$ corresponds to the Fe$_u$ redox-change from Fe$_u$(III) to Fe$_u$(II), as the Fe$_u$ charge in **C$_{int}$** is respectively 0.28 and 0.17 lower than that in **C$_{red1}$-tOH** and **C$_{red1}$**, whereas the atomic charges of Ni, Fe$_1$, Fe$_2$ and Fe$_3$ do not differ by more than 0.05. Accordingly, the net charges on L1 and L2 layers indicate that the reduction of the C-cluster occurs predominantly on the {Fe$_1$NiFe$_u$S$_1$S$_2$} layer. Indeed, in C$_{int}$, the charge of the L1 layer is 0.57 and 0.59 more negative than those in **C$_{red1}$** and **C$_{red1}$-tOH**, respectively. A reduction of only 0.25/0.35 has been instead computed for the charge of the {Fe$_2$Fe$_3$S$_3$S$_4$} (L2) layer.

Assuming that C$_{red2}$ corresponds to the unbound C-cluster, the identification of the metal sites at which occurs the second reduction of the C-cluster is more complicate. No significant changes have been observed between the charges of the metal atoms in the C$_{int}$ and C$_{red2}$

states (see Table 3.10). Since the charges of both L1 and L2 decreases after reduction, we can assume that the added electron is largely delocalized on both layers. On the other hand, assuming the $NiFe_u$-bridging H^- containing C-cluster as possible structure for C_{red2}, the analysis of the electronic structure favours the Ni(II)-H^- assignment for **C_{red2}-H**, rather than Ni(0)-H^+ or Ni(I)-H.

In both **C_{red2}** and **C_{red2}-H** species the formal redox state of the Ni atom is predicted by the analysis of the electronic structure to be +2. This is in accord with Ni K- and L-edge XAS studies that indicate a low-spin diamagnetic Ni(II) for both C_{red1} and C_{red2} [109, 111, 114]. The two extra electrons are located on Fe atoms of the C-cluster in **C_{red2}** and on the hydride ligand in **C_{red2}-H**. Therefore, the assignment of the C_{red2} state to the C-cluster with no exogenous ligand is against experimental evidence from UV-visible [194] and Mössbauer studies [109], indicating that the electronic structure of the $[Fe_uFe_3S_4]$ cluster does not change upon reduction and from EPR data [110, 178], suggesting that it remains in the +1 oxidation state with a total spin S = 1/2 in C_{red1} and C_{red2}. On the other hand, the electronic structure of **C_{red2}-H** satisfies the experimental data discussed above but does not explain the absence of a strong ENDOR signal corresponding to the hydride ligand in C_{red2} [118].

On the basis of these results, both **C_{red2}** and **C_{red2}-H** are considered plausible candidates for the C_{red2} state. In particular, it is possible to assume that if His93 is singly protonated at the δN atom, the C_{red2} corresponds to the **C_{red2}** species in which no ligands are coordinated to the C-cluster. On the other hand, doubly protonation state of His93 favours the formation of **C_{red2}-H**. The proton transfer from His93 to the NiFe site, according to the reaction in Figure 3.18, is indeed energetically favoured by 16.3 kcal/mol. Interestingly, the formation of the hydride-bound C-cluster in the C_{red1} state is instead disfavored by 6.6 kcal/mol.

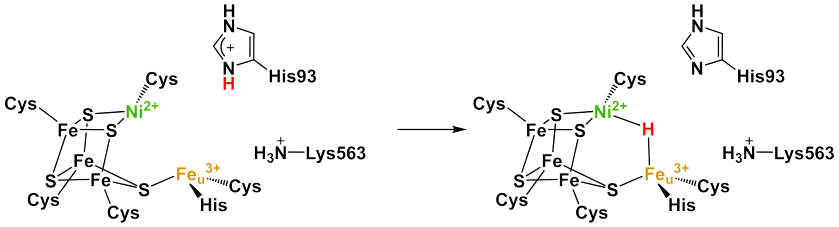


Figure 3.18 Schematic representation of formation reaction of the **C_{red2}-H-LM^b** adduct from the **C_{red2}-LM^a** species.

3.5 Conclusions

The disclosure of the stereoelectronic and catalytic properties of the active site of Ni,Fe-containing carbon monoxide dehydrogenases is important not only in the context of the efforts aimed at elucidating structure-function relationships in CODHs but also for the development of bioinspired catalysts for CO_2/CO interconversion that may be used for the removal of such gases from the environment. To contribute to such an effort, different forms of the C-cluster in C_{red2}, C_{int} and C_{red1} redox states have been investigated.

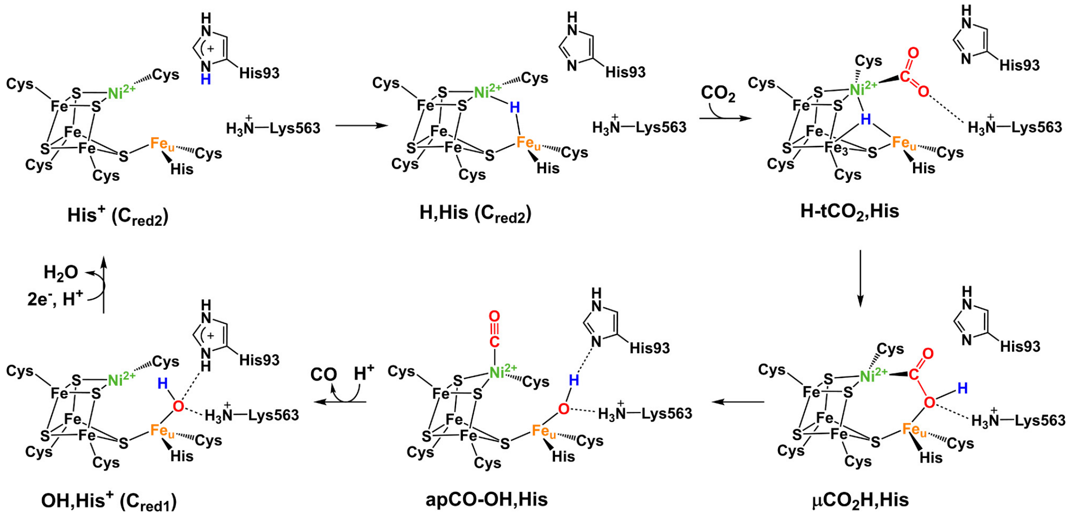
The comparison of results obtained using a minimal DFT model (metal ions and first coordination sphere) and a very large model of the active site (metal cluster, first and second coordination sphere − 270 atoms) highlights the crucial role of the protein environment in tuning the coordination geometry of the C-cluster and the stability of CO_2 and CO adducts. In particular, the protonation state of the His93 is predicted to be fundamental for the binding of CO_2 and CO substrates; binding of CO_2 is favoured by the charged protonated form of His93, whereas CO binding is favoured by the neutral form of this residue, in which only the δN atom is protonated. Another factor affecting the energetics of the chemical binding of the substrates is the redox state of the C-cluster. In particular, binding of both CO_2 and CO is favoured by the reduction of the active site from C_{red1} to C_{red2}. However, when His93 is doubly protonated, CO binding to C_{red2} is less favoured than CO_2 binding by about 10 kcal/mol. On the other hand, independently of the protonation state of His93, C_{red1} binds preferentially to CO. Interestingly, for both CO and CO_2 we have identified a stable conformation in which the exogenous ligand is terminally bound to the Ni atom, located at the end of the substrate channel [98]. Such conformation may play a key role in the binding/dissociation of the CO_2 and CO molecules to/from the C-cluster.

Our theoretical study also provides significant insights in the structure and electronic nature of C_{red1} and C_{red2}. The C-cluster with an hydroxide ligand coordinated to Fe_u is considered the more plausible candidate for C_{red1}, whereas the nature of the C_{red2} state is assumed to depend on the protonation state of His93. If His93 is singly protonated at the δN atom, C_{red2} may correspond to the C-cluster with no exogenous ligands. If His93 is instead doubly protonated, the energetically favoured proton transfer from this residue to the C-cluster leads to the C_{red2} state in which a hydride bridges the Ni-Fe_u site. Computed charges and spin densities on metallic atoms of the C-cluster suggest low-spin diamagnetic Ni(II) for both redox states, according to Ni K- and L-edge XAS studies [109, 111, 114]. One-electron reduction of C_{red1} corresponds to the reduction of the Fe_u ion from Fe(III) to Fe(II), whereas the second electron required for the reduction to C_{red2} is largely delocalized on the metal cluster or on the hydride ligand. The different redox state of the Fe_u atom in C_{red1} and that in C_{int} and C_{red2} explains the instability of the OH-bound forms of the C-cluster in the latter states; the hydroxide binds to Fe_u(III) in C_{red1} and dissociates from Fe_u(II) in C_{int} and C_{red2}, as water molecule. In **tCO_2-OH** and **tCO-OH** adducts OH^- was found to bind the Fe_u atom also in the C_{int} and C_{red2} redox states. However, we can note that in these species a net electron transfer from the C-cluster to the CO_2 and CO ligands has been shown to oxidize the Fe_u from Fe(II) to Fe(III), as indicated by the computed charges on metallic atoms and substrates.

Combination of the DFT results described above leads to a more detailed version of the catalytic mechanism of the Ni-CODH enzyme. Proposed mechanism for the CO_2 reduction is presented in Scheme 3.1, whereas that for the CO oxidation is shown in Scheme 3.2. Since the reaction mechanism is reversible, the two catalytic mechanisms contain the same intermediates. However, for sake of clarity, the mechanism of CO_2 reduction and CO oxidation are depicted in two different schemes because the direction of the catalytic reaction is differently influenced by the doubly and singly protonated state of His93.

In the CO_2 reduction direction, it is possible to assume that His93 is doubly protonated since the reduction of the CO_2 molecule requires that two protons are supplied to the active site and the neutral form of His93 can act as donor of a proton. The transfer of the hydrogen atom from this residue to the C-cluster in the C_{red2} state, leading to the formation of a hydride ligand bridging the NiFe site (intermediate **H,His** in Scheme 3.1), is indeed energetically favoured by 16.3 kcal/mol. According to this hypothesis, once **H,His** is formed, CO_2 reaches the active site via the hydrophobic tunnel apical to the Ni atom and terminally binds to the Ni(II) ion, with the concomitant change of the coordination mode of the hydride ligand from $NiFe_u(\mu^2\text{-}H)$ to $NiFe_uFe_3(\mu^3\text{-}H)$ (intermediate **H-tCO₂,His**). Interestingly, the binding of CO_2 to hydride-bound C_{red2} promotes an electron transfer from the cluster to CO_2, as already observed for the CO_2 binding to the unbound forms of the enzyme. However, the negative charge of the CO_2 ligand in **H-tCO₂,His** (-0.68) is significantly smaller than those calculated for the **tCO₂** species of the C-cluster in the absence of the hydride ligand (about -1), resulting in a lower activation of the CO_2 molecule (O-C-O angle = 134.3° and C-O distances of 1.24 and 1.25 Å).

Scheme 3.1 Proposed mechanism for CO_2 reduction at the C-cluster.

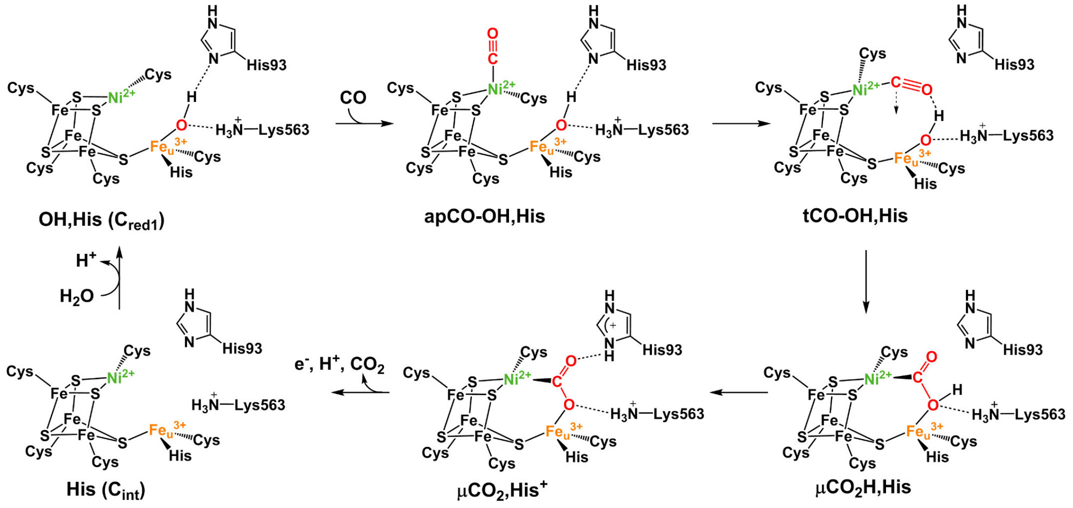


The CO_2 binding and the hydride ligand displacement is a slightly endoenergetic process (3.1 kcal/mol). On the other hand, the isomerisation reaction leading to the formation of the $\mu_2\text{-}\eta^2_{C,O}$ CO_2H adduct, strongly stabilized by the H-bond network with His93, Lys563 and a water molecule, is exoenergetic by 3.3 kcal/mol (intermediate **μCO₂H,His**). Although the CO_2 binding is favoured by the charged protonated form of His93, it is assumed that CO_2 binds to the C-cluster when His93 is singly protonated because if this residue is doubly protonated the formation of the CO_2H adduct is endoenergetic by 9.9 kcal/mol.

Finally, the cleavage of the carboxylate C-O bond, assisted by the transfer of the μ^3-H atom to the CO_2 ligand (as indicated by the increase of the C-O distance from 1.25 Å in **H-tCO₂,His** to 1.42 Å in **μCO₂H,His**), yields the **apCO-OH,His** species (-28.8 kcal/mol) in which both CO and OH^- ligand are coordinated to the active site. Once **apCO-OH,His** is formed, protonation of His93 favours the release of CO from the active and the formation of the C_{red1} state (+3.3 kcal/mol). One electron reduction of C_{red1} gives rise to C_{int} from which the hydroxide ligand dissociates as H_2O. An additional one-electron reduction and protonation regenerates the C_{red2} species.

In the CO oxidation direction, CO binds apically to the Ni atom in the C_{red1} state in which an hydroxide ligand is coordinated to Fe_u (**apCO-OH,His** intermediate in Scheme 3.2). Although our calculations show that the doubly protonation state of His93 favours the movement of CO ligand towards the hydroxide ligand necessary for the CO_2 formation, His93 is believed to be singly protonated on the δN atom. Since the oxidation of CO requires that two protons are released from the active site, the presence of a nearby residue that can act as proton acceptor is more plausible. In addition, the singly protonated state for His93 is supported by the long distance between the CO oxygen atom and the εN of His93 in crystal structures of CO-bound C-cluster [102, 119]. Accordingly, our calculations indicate that CO binding is energetically more favoured if His93 is singly protonated (-22.6 kcal/mol). It can be supposed that the rearrangement of the CO geometry from linear to bent (**tCO-OH,His** intermediate) is induced by the formation of an elaborate and extended network of hydrogen bonds in which His93 participates as proton acceptor rather than hydrogen bond donor. The hydroxide ligand, Lys563 and conserved water molecules are also involved in this network which stabilizes the bent geometry of CO, initially induced by the electron transfer from the C-cluster to the CO ligand.

Scheme 3.2 Proposed mechanism for CO oxidation at the C-cluster.



Once **tCO-OH,His** is formed, a ‘carbon shift’ may occur to place the carbon in proximity to the hydroxide, and promoting the nucleophilic attack yielding the **μCO_2H,His** species. Subsequent deprotonation of the carboxylate ligand lead to the **μCO_2,His^+** intermediate in which Hi93 is doubly protonated (-17.3 kcal/mol). According to our calculations, the CO_2 release from **μCO$_2$,His$^+$** is endoenergetic by 35.0 kcal/mol. However, after deprotonation of His93 and one-electron oxidation of the C-cluster this process is endoenergetic by only 3.4 kcal/mol. Further one-electron oxidation and binding of H_2O regenerates the C_{red1} state.

As shown in see Scheme 3.1 and Scheme 3.2, on the basis of the computed reaction energies and the assignment of C_{red1} and C_{red2}, an associative mechanism is proposed for the binding and the release of CO, whereas a dissociative mechanism is considered more plausible for those of CO_2. Moreover, the role of the His93 residue in tuning the direction of the catalytic reaction is elucidated; CO_2 reduction is favoured by the doubly protonated state of His93, which promotes the CO_2 binding and the CO release from the active site. Conversely, the CO oxidation is facilitated by the singly protonated state on the δN atom of His93 that promote CO biding and CO_2 release. However, several aspects of the catalytic mechanism are still not clear. In particular, further investigations are required to characterize the H-bond network established between the CO ligand and the protein environment that may trigger the rearrangement of the CO geometry and its movement towards the hydroxide. Moreover, the nature of the ‘carbon shift’ necessary for an adequate positioning of the carbon atom in proximity to the hydroxide to allow the nucleophilic attack of OH^- on CO remains to be elucidated.

CHAPTER 4

REVERSIBLE INACTIVATION OF [NIFE]-HYDROGENASES

While the catalytic pathway of [NiFe]-hydrogenases is in general understood, oxidation and reactivation mechanisms of such enzymes are still not clear. Exposure of [NiFe]-hydrogenases to O_2 leads to the formation of the Ni-A and Ni-B inactive oxidized states, which feature distinctive IR and EPR signatures [6]. In both these forms, the Fe and Ni ions attain the +2 and +3 redox states [73, 74], respectively; however, as far as reactivation kinetics is concerned, Ni-A is much slower than Ni-B [75]. The latter observation implies that the two resting states feature key structural differences, the nature of which is still question of debate. Interestingly, recent results show that Ni-A and Ni-B can form also in the presence of oxidants other than dioxygen [91], rising additional questions on their nature.

The characterization of the oxidized forms of [NiFe]-hydrogenases through the calculation of spectroscopic parameters is a challenging task. A very recent combined experimental and theoretical study by Lubitz and co-workers [197] evidenced how difficult it is to mark a distinction between the Ni-A and the Ni-B states based on a comparison between computed and experimental EPR and FT-IR parameters. As for the former, the drawback of g values calculated by means of density functional theory (DFT) is that they are usually found to fluctuate in a non-systematic way and are significantly underestimated when compared with experimental values; comparison between computed and experimental vibrations of CO and/or CN ligands bound to the Fe ion in Ni-A and Ni-B also proved to be only partially informative, a fact which again depends not only on the somehow limited accuracy of theoretical spectroscopic parameters, but also on the high degree of similarity shared by the experimental spectra of the two enzyme forms. Nonetheless, insights provided by Lubitz and coworkers lent support to the hypothesis that both the Ni-A and Ni-B states feature an OH^- ligand placed in bridging position between Ni and Fe ions, thus occupying the site where substrate binding is believed to occur when the enzyme is in the catalytically active state.

Another key issue in [NiFe]-hydrogenases inhibition is the redox modification of metal-bound cysteine ligands [61, 63, 77–79]. Crystallographic studies on the *D. fructosovorans* enzyme in oxidizing conditions provided results which are best explained by the presence of partly oxidized cysteine ligands [109]; in other words, this X-ray structure features a mixture of states, a conclusion that is supported by detailed quantum refinement studies [109]. More favourable conditions for the specific investigation of the Ni-A state were recently achieved by Fontecilla-Camps and co-workers [63]. According to the crystallographic study by these authors, a bridging hydroxyl ligand is actually present in the enzyme active site and is accompanied by oxidation of metal-bound sulphur. In particular, it turned out that in Ni-A, only one of the bridging cysteine side chains carries clear signs of oxidation (Cys64 of *Allochromatium vinosum*).

Although being an extremely powerful technique for structural characterization, protein crystallography does not come without shortcomings. In particular, X-ray irradiation of metalloproteins can give place to photoreduction processes of redox-active cofactors [198]. Notably, theoretical calculations–usually based on DFT–represent in principle an ideal approach for corroborating crystallographic results. In fact, DFT models with a pre-defined overall charge, corresponding to specific formal redox states of the metal-containing cofactor under investigation, can be geometry optimized thus providing reliable structures to be compared with crystallographic data. However, quite surprisingly, in the case of the Ni-A state of [NiFe]-hydrogenases, the various DFT studies described in literature [78, 199–202] have increased the uncertainty on the structure of this state, due to relevant discrepancies between the computational results and those given by X-ray experiments. Clearly, such discrepancies may be due to the inadequacy of the chemical models used so far [203]. In our laboratories, a large-size model of [NiFe]-hydrogenases and its surroundings has been recently developed and validated. Based on results from a previous study [71], this model is expected to provide reliable energy differences between isomers, such that speciation and mechanistic routes leading to the occurrence of a certain enzyme form can be discussed with reasonable confidence. This model also enabled to discuss fine details of the coordination geometry around the metal atoms, including the subtle effects exerted by the protein environment on the first coordination shell of the Ni ion.

The above premises stimulated us to perform a DFT study on the oxidation of the enzyme active site, based on accurate energy calculations on [NiFe]-hydrogenases models, and on comparisons of theoretical geometries with the X-ray structures currently available [61, 63, 77–79]. Results obtained by this study will be discussed in three separate sections, after a brief description of the DFT model and the methods employed. In the first section (section 4.2) will be discussed the theoretical characterization of the Ni-A state, while the second section

(section 4.3) will focus on the aerobic and anaerobic oxidative inactivation pathway of the active site. The reactivation mechanism of the oxidized states will be discussed in the third section (section 4.4).

4.1 Computational details

4.1.1 Computational model of the [NiFe]-hydrogenases active site

The starting structure for the DFT calculations was based on the X-ray geometry of the *Allochromatium vinosum* [NiFe]-hydrogenase (pdb code: 3MYR) [79], in which the oxo ligand bridging the Ni and Fe ions was removed. In the following, the residues are numbered according to the 3MYR structure (all residues included in the models belong to the large chain of the enzyme, with the exception of Thr17 which belongs to the small chain and, for the sake of clarity, is labeled as Thr17(S)). Three models of different size have been considered.

The largest model (LM) contains 288 atoms and has a size of about 18 Å (see Figure 4.1). It includes the bimetallic NiFe cluster, with two CN^- and one CO ligands coordinated to the Fe atom, two cysteine residues terminally coordinated to the Ni atom, and two further cysteine residues bridging the two metallic ions. In addition, the model also includes the residues of the first and second coordination spheres that in the following are numbered according to the 3MYR structure. In the enzyme, the two terminal cysteine residues, Cys61 and Cys555, belong to two different Ni-binding motifs that also include the two bridging cysteine residues, Cys64 and Cys558, respectively. Each motif contains two cysteines, located at the N- and the C-termini and separated by two other residues which are also contained in the model. For the motif Cys555-Ile556-Ala557-Cys558, only the backbone atoms of Ile556 and Ala557 are included in the model, since their side chains are oriented far from the active site, whereas for the peptide chain Cys61-Gly62-Val63-Cys64, the side chain of Val63 is included in the model, since it is close to the vacant coordination site on Ni. These two peptide chains are terminated by selected backbone atoms of residues Ile60, Thr65, Pro554, and Ala559.

In addition, the model also includes the side chains of residues Thr17(S), Asp103, Leu490, Arg487, and Asp553, terminated at the Cα atoms; a peptide chain containing the residues His68 and Val67; a peptide chain containing the carbonyl and the Cα atoms of Val507, and the entire residues Val508, Pro509, and Ser510. Finally the model includes a peptide chain containing the carbonyl and Cα atoms of Ile13, the entire residue Glu14, and the N and Cα atoms of Gly15. The truncated residues were saturated with hydrogens. Arg487, Asp553, and Asp103 have been included in the model, since they form a network of H-bonds which also involves one CN^-ligand, whereas Ser510 is H-bonded to the other CN^- ligand through its hydroxyl and amide hydrogens. Val67, Leu490, Val508, and Pro509 have been selected because they form a hydrophobic pocket surrounding the $Fe(CN)_2(CO)$ group. His68 is close to the bimetallic cluster, and depending on its protonation state, it can be involved in different H-bond interactions with the active site and/or with nearby residues. The three possible protonation states at δN, εN, or both atoms were investigated in detail by Neese *et al.* [204] using a model based on the X-ray structure of the *Desulfovibrio vulgaris Miyazaki F* [NiFe]-hydrogenase; it was concluded that the best comparison between experimental and theoretical data is obtained when His68 is protonated at the εN atom by forming a H-bond with the sulphur atom of Cys555. On the basis of these results, His68 is always modeled as protonated at the εN atom. Conversely, Glu14 was modeled both as glutamate and glutamic acid, since it was proposed to be involved in proton transfer from/to the bimetallic cluster [85, 205, 206]. In the X-ray structure of the enzyme from *Allochromatium vinosum*, it forms H-bonds with Ala557 and Thr17(S), which were also retained in the model. Selected backbone atoms of the adjacent residues Ile13 and Gly15 were also added to the model, since they form H-bonds with residues Pro554 and Ile556 belonging to the Cys555/Cys558 binding motif.

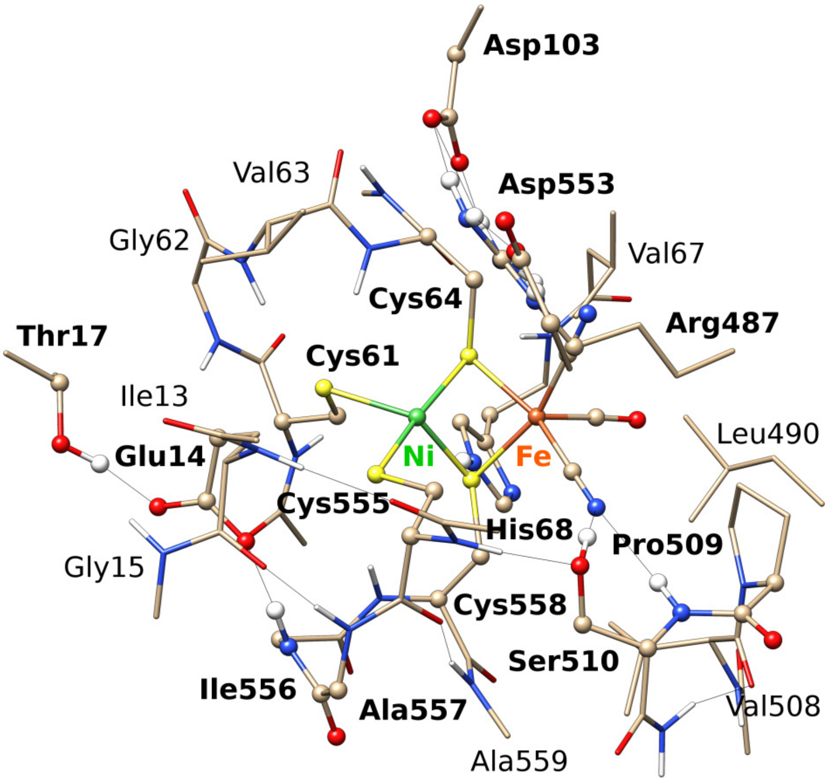


Figure 4.1 Schematic representation of the [NiFe] cluster and selected atoms of the protein environment included in the large DFT model (LM) of the active site, derived from the X-ray crystal structure of the *Allochromatium vinosum* [NiFe]-hydrogenase (pdb code 3MYR). Atoms belonging also to the medium DFT model (MM) are represented as 'ball and stick' and relative residue name are indicated with bold labels.

The medium size model (MM), containing up to 135 atoms, contains the first coordination sphere of the NiFe cluster. In the second coordination sphere, the model includes Glu14, Thr17, His68, Asp103, Arg487, Pro509, Ser510, Asp553, Ile556 and Ala557 (Figure 4.1). Initial coordinates were taken from the corresponding optimized structure of the large-size model. A detailed list of the atoms composing the large (LM) and medium (MM) models has been reported in Table 4.1. During geometry optimizations, selected atoms have been constrained to the crystallographic positions, in order to avoid unrealistic distortions at the boundary of the model. These atoms have been also indicated in Table 4.1.

Table 4.1 Detailed information of the atoms included in the large (LM) and small (SM) models. Hydrogen atoms are not explicitly reported in the list. C, O and N are backbone carbonylic carbon and oxygen, and amidic nitrogen, respectively. Atoms constrained during geometry optimizations are marked with an asterisk. Conversely, atoms in parenthesis indicate atoms at which has been terminated the side chain of the residue.

Residue	**Atoms in LM**	**Atoms in SM**	**Basis set**
$NiFe(CN)_2(CO)$	all	all	def-TZVP
Ile13	C, O and Cα*	none	def-SVP
Glu14	all	side chain (Cγ*)	def-TZVP (carboxylic group), def-SVP (other atoms)
Gly15	N and Cα*	none	def-SVP
Thr17(S)	side chain, Cα*	side chain (Cβ*)	def-SVP
Cys61	all	side chain, Cα*	def-TZVP (side chain), def-SVP (other atoms)
Gly62	zll, H_N*[a]	none	def-SVP
Val63	all	none	def-SVP
Cys64	all	all	def-TZVP (side chain and Cα) def-SVP (other atoms)
Thr65	N, Cα*	none	def-SVP
Val67	side chain, C, O, Cα*	none	def-SVP
His68	side chain, N, Cα*	side chain (Cβ*)	def-SVP
Asp103	side chain, Cα*	side chain (Cβ*)	def-SVP
Arg487	side chain, Cα*	side chain (Cδ*)	def-SVP
Leu490	side chain, Cα, all C atoms*[b]	none	def-SVP
Val507	C, O, Cα*	none	def-SVP
Val508	all	none	def-SVP
Pro509	all	C, O, Cα*	def-SVP
Ser510	all	side chain, N, Cα*	def-SVP
Thr511	N*	none	def-SVP
Asp553	side chain, Cα*	side chain (Cβ*)	def-SVP
Pro554	C, O, Cα*	none	def-SVP
Cys555	all	side chain, Cα*	def-TZVP (side chain and Cα), def-SVP (other atoms)
Ile556	N, C, O*[c], Cα	C, O, Cα*	def-SVP
Ala557	N, C, O*[d], Cα	N, Cα*	def-SVP
Cys558	all,O*[e]	side chain, Cα*	def-TZVP (side chain and Cα), def-SVP other atoms)
Ala559	N, Cα*	none	def-SVP

[a]amidic hydrogen of Gly1062 forms an H-bond with the carbonilic oxygen of Cys16. This H atom has been added at 1.02 Å from N in the C-N-Cα plane, and constrained at that position; [b]all carbon atoms are constrained to the X-ray positions to avoid an unrealistic large rotation of the sidechain during geometry optimization; [c]carbonyl oxygen forms an H-bond with protonated δN of His1561in the α-helix at the C-terminius; [d]carbonyl oxygen forms an H-bond with the amidic hydrogen of His1561 in the α-helix at the C-terminius; [e]carbonyl oxygen forms an H-bond with the amidic hydrogen of Ala1560 in the α-helix at the C-terminius.

A minimal model (SM) of the active site including only the ligands of the first coordination sphere of the two metal atoms, with formula $[(CH_3CH_2S)_4NiFe(CN)_2(CO)]$, was also considered. In this model, the four cysteine residues are terminated at the Cα atoms and they are therefore modeled as ethylthiolates. Energy values were calculated by constrained geometry optimizations, in which all atoms with the exception of hydrogens were constrained at the position computed for the large model. Comparison between large models and constrained small models allows to highlight the effect of the protein environment on the stereoelectronic properties of the active site of [NiFe]-hydrogenases, whereas comparison between constrained and unconstrained small models allows to investigate the geometric factors that affect the stability of the different species.

4.1.2 DFT calculations

Geometry optimizations of all minimum energy structures have been performed on the larger model (LM) of the active site, whereas for the optimizations of the transition states the medium model (MM) has been used. Only some structures have been optimized using the small model (SM).

All calculations were carried out in the framework of the density functional theory by using the BP86 [149, 150] functional in conjunction with the resolution-of-the-identity (RI) [163] technique, with the TURBOMOLE 6.4 program suite [171]. An all-electron valence triple-ζ basis set with polarization functions TZVP [184] has been used for Ni, Fe, the atoms of the first coordination sphere, and the residue Glu14, while the SVP double-ζ basis set [185] has been used for all other atoms. For details, see Table 4.1. To model the effects of the protein environment on the stereoelectronic properties of the active site, geometry optimizations carried out using the large model were performed using the conductor-like screening model (COSMO) [169, 170] with a polarizable continuum medium at $\varepsilon = 4$ [167, 168, 207]. Conversely, calculations carried out using the small model, which is meant not to include the effect of the environment on the NiFe cofactor, have been performed in a vacuum.

Dispersion interactions were not included in the calculations because the results obtained by using the BP86-COSMO scheme are consistent with those obtained by previous investigations on the same large model of the [NiFe]-hydrogenases active site carried out by including the dispersion interactions correction given by the DFT-D3 Grimme scheme [208–210] as implemented in TURBOMOLE [171].

The optimization of transition state structures has been carried out according to a procedure based on a pseudo-Newton-Raphson method. Initially, a geometry optimization of a transition state guess geometry, in which selected internal coordinates are constrained, was performed. Vibrational analysis of such optimized geometry is than carried out. If one negative eigenmode corresponding to the reaction coordinate is found, the geometry is used as the starting point for the successive optimization steps applying the trust-region image minimization method [211], which is designed to maximize the energy along one of the eigenvectors (i.e., the chemically relevant one) and minimize it in all other directions. The presence of constrained atoms resulted in the occurrence of several imaginary frequencies, that correspond to torsional vibrational modes involving the constrained atoms. However, the eigenvector corresponding to the reaction coordinate was easily identified as the one with an eigenvalue much lower than the other negative eigenvalues.

4.1.3 Validation of the DFT model

In a previous study on the same DFT model of the active site of [NiFe]-hydrogenases, the large model (LM) was shown to correctly describe the peculiar see-saw conformation of the Ni atom, which is crucial for the binding of H_2 in the Ni-S$_{Ia}$ state [71]. Here, the adequacy of the model and the quality of the adopted level of theory have been also supported by the agreement of the predicted geometry with the X-ray structure of the Ni-B state (Figure 4.2).

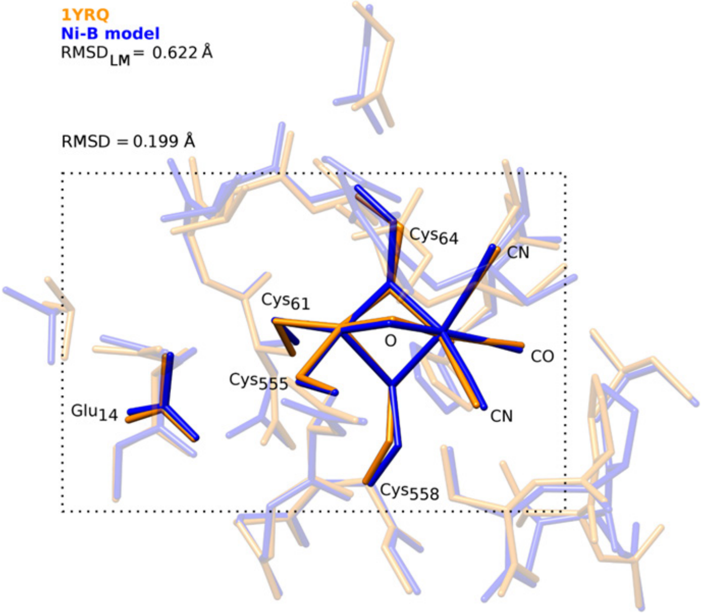


Figure 4.2 RMSD values and superimposition of the optimized geometry of the Ni-B species (blue) with the corresponding 1YRQ X-ray structures (orange) [61]. The RMSD$_{LM}$ value is calculated including all of the (non-hydrogen) atoms in the large model. Conversely, the RMSD value of 0.199 Å refers to the structure of the bimetallic cluster only including the first coordination sphere (cysteine residues as $CH_3CH_2S^-$) and the side chain of the Glu14 residue (as CH3COO$^-$), extracted by the theoretical geometry optimized using the large model, compared to the 1YRQ geometry.

The root-mean-square deviation (RMSD) of the bimetallic cluster (including the side chain of Glu14) compared with the crystallographic structure of Ni-B (pdb code 1YRQ) [61] is 0.20 Å. The arrangement of the ligands around the metal ions is very well reproduced by the model. In particular, the computed SCys61-Ni-SCys558 and SCys555-Ni-SCy64 angles in the geometry optimized model are 109.0° and 173.8°, respectively (107.1° and 169.3° in the crystal structure). The computed Ni–Fe distance, equal to 2.88 Å, also matches the experimental value (2.82 Å). Other relevant geometrical parameters are in reasonable agreement with the X-ray values, also considering the heterogeneity of these parameters observed in the different X-ray structures assigned to Ni-B (see Table 4.2). The only exception is given by the Ni-SCys558 distance which is underestimated by more than 0.1 Å when compared to all of the X-ray structures considered. The RMSD calculated including all of the (non- hydrogen) atoms in our model is equal to 0.62 Å, which can be deemed acceptable in consideration of the very small number of atoms constrained to the X-ray positions in the geometry optimizations.

Table 4.2 Selected distances (Å) of the Ni-Fe active site for the optimized Ni-B model and several crystallographic structures that can be assigned to Ni-B. Sulphur atoms are numbered progressively according to the order of the Cysteine residue in the protein sequence.

	Ni-B	**1YRQ**[b] [61]	**1WUJ** [77]	**1WUK** [77]	**3MYR**[a] [79]
Ni-Fe	2.88	2.81	2.69	2.67	3.01
Ni-S1	2.28	2.15	2.23	2.22	2.24
Ni-S2	2.34	2.31	2.38	2.28	2.58
Ni-S3	2.22	2.17	2.07	2.14	2.25
Ni-S4	2.40	2.51	2.55	2.54	2.62
Fe-S2	2.40	2.26	2.26	2.28	2.32
Fe-S4	2.35	2.28	2.34	2.33	2.31
Ni-O	1.93	1.88	1.67	1.63	1.89
Fe-O	2.04	1.91	2.14	1.98	1.97
Resolution		2.10	1.40	1.10	2.10

[a] average value of four molecules; [b] average value of five molecules.

4.2 Theoretical characterization of the Ni-A state

Several X-ray crystallographic structures have been obtained in oxidizing conditions and assigned to the Ni-A state, which, however, differ for the nature of the ligands associated with the active site. These structures can be roughly classified into three groups depending on the modification of the active site: (1) structures with a diatomic oxygenic ligand coordinated to the active site; (2) structures with a bridging monoatomic oxygenic ligand coordinated to the active site; and (3) structures with a bridging monoatomic oxygenic ligand accompanied by oxidation of a cysteine residue coordinated to the Ni atom.

In this context, the nature of Ni-A has been investigated by DFT calculations, considering systematically a large number of plausible forms in which oxygenic ligands featuring different protonation states bind the active site, with the final aim of comparing our results with those from earlier theoretical and crystallographic studies. The different X-ray structures assigned to the Ni-A state are discussed separately depending on the nature of the oxygenic ligands.

Nomenclature: in the following, computational models of the investigated species will be labelled according the general scheme **KHj**, where **K** is a generic letter referring to the binding mode of the oxygenic ligands and **j** is the number of protons bound to the oxygenic ligands. Additional tags may also be present, as follows: isomers differing for protonation regiochemistry will be referred to by including progressive numbers after the **KHj** tag; oxidation at the level of a cysteine residue occupying the **X** position in the amino-acid sequence is indicated with the **CysXO$_b$** or **CysXO$_t$** tags in the case of bridging and terminal sulfenates, respectively (*vide infra*); finally, when the Glu14 residue is protonated (i.e., it attains neutral state), the label **GluH** is added. The computational model of the Ni-B state, featuring a hydroxyl ligand bridging the two metal ions, will be labelled **Ni-B**.

4.2.1 Characterization of Ni-A as containing a diatomic oxygenic ligand

Two different crystallographic studies of Fontecilla-Camps *et al.* [61] and Ogata *et al.* [77] identified in Ni-A a diatomic oxygenic ligand which, on the basis of the O-O distance, was proposed to be a (hydro)peroxide (see Figure 4.3). In both crystallographic structures, this oxygenic ligand bridges the two metal ions with a side-on, bidentate binding mode to nickel, an unusual arrangement in bioinorganic systems.

Sample preparation for the X-ray analysis in both structural determinations has been carried out in the presence of O_2; the diatomic oxygenic ligand in this putative Ni-A state can then be derived from the binding of O_2 to Ni(II) followed by an intramolecular electron transfer to give Ni(III) and superoxide O_2^-, and a further monoelectronic reduction of the ligand to give the (hydro)peroxide. Since hydrogen atoms could not be identified in the crystallographic experiments here considered, the diatomic oxygenic ligand has been modeled as peroxide (O_2^{2-}) hydroperoxide (O_2H^-), and hydrogen peroxide (H_2O_2) (see Scheme 4.1).

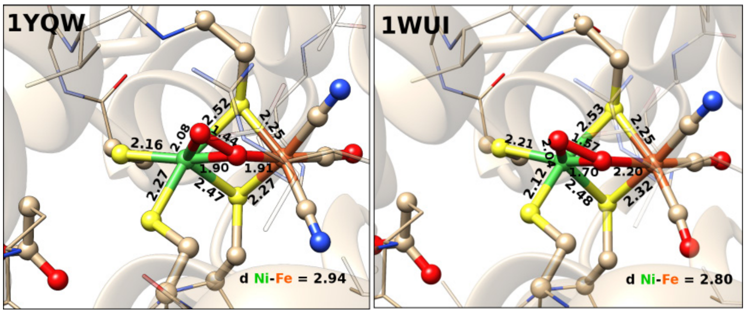


Figure 4.3 Schematic representation of the X-ray structures of the Ni-A state reported in Refs. [61] and [77]. Selected interatomic distances are given in Å. For the sake of clarity, aliphatic hydrogen atoms are not shown.

Three different coordination modes of the peroxide ligand (O_2^{2-}) were characterized as minima on the Potential Energy Surface (PES), all featuring a very similar stability. In the most stable isomer, the peroxide ligand is bound side-on to Ni (species **A** in Figure 4.4, Scheme 4.1) with an arrangement similar to that of the experimental X-ray structures. The coordination of the Ni atom, by taking the position of the O_2^{2-} ligand at the midpoint of the O-

O bond, is trigonal bipyramidal, with the two oxygen atoms (hereafter, referred to as O1 and O2 according to the numbering in Scheme 4.1 and Scheme 4.2) asymmetrically coordinated to Ni.

In particular, the computed Ni-O1 distance of 1.88 Å was in good agreement with the experimental value of Ref. [61] (1.90 Å), but significantly longer with respect to the value reported in Ref. [77] (1.70 Å). The Ni-O2 distance of 2.18 Å was also closer to the X-ray value of Ref. [61] (2.11 Å) than that of Ref. [77] (2.04 Å). The Ni-Fe distance equal to 2.90 Å well reproduces the experimental values (2.93 Å), whereas the Fe-O1 distance of 2.04 Å is longer than that reported in Ref. [61] (1.90 Å[61] and 2.80 Å [77]) and shorter than that reported in Ref. [77] (2.20 Å). The calculated O-O bond length equal to 1.42 Å is compatible with a peroxide group; moreover, this interatomic distance is in good agreement with the experimental value (1.45 Å) determined in the crystallographic study described in Ref. [61], but it is significantly shorter when compared with the O-O distance in the structure reported by Ogata *et al.* (1.57 Å) [77]. The bond lengths and the orientation of the O_2^{2-} ligand are also in close agreement with those obtained by the QM/MM calculations reported in Ref. [78]. The unpaired electron is mainly localized on Ni (0.56), the Sγ atom of Cys558 (0.13), and the O_2^{2-} ligand (0.24). The stability of **A** is favoured by the formation of H-bonds between the two oxygen atoms of the O_2^{2-} ligand and the guanidinium group of Arg487.

In the second isomer, the O_2^{2-} ligand bridges the two metal atoms with a μ-η^1:η^1-O_2 binding mode (species **C** in Figure 4.4, Scheme 4.1). This species is less stable than **A** by only 0.5 kcal/mol, however, the orientation of the peroxide ligand in **C** is clearly different from that reported in the crystallographic structures of Refs. [61] and [77].

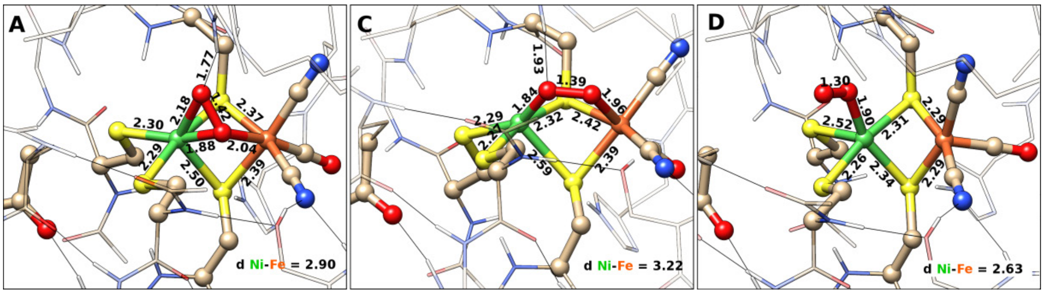


Figure 4.4 Schematic representation of the optimized geometries of the [NiFe] hydrogenase active site coordinated by a peroxide ligand (O_2^{2-}). Selected interatomic distances are given in Å. For the sake of clarity, aliphatic hydrogen atoms are not shown.

In addition, the calculated Ni-Fe distance, equal to 3.22 Å, is significantly longer than the corresponding experimental values (2.93 Å [61] and 2.80 Å [77]). The unpaired electron, as in **A**, is mainly localized on Ni (0.56), the Sγ atom of Cys558 (0.14), and the O1 atom of the peroxide ligand (0.18). End-on coordination of O_2^{2-} to the Ni atom, resulting in an η^1-O_2 adduct, characterizes the third isomer (species **D** in Figure 4.4, Scheme 4.1) which is 4.1 kcal/mol higher in energy than **A**. The unpaired electron in **D** is almost totally localized on the diatomic oxygenic ligand that should, therefore, be represented as a superoxo O_2^- species, leaving the Ni atom in the diamagnetic +2 redox state. This observation is also supported by the short O1-O2 distance equal to 1.30 Å. Another possible coordination mode of the peroxide ligand, in which one oxygen atom bridges the two metals (species **B** in Scheme 4.1) has been tested, but during geometry optimization, it converges to species **A**.

It is worth noting that proton transfer from the protonated Glu14 residue–which should be the first acceptor in the proton transfer chain–to the species **A** to give **AH-1** (Figure 4.5, Scheme 4.1) is an endoenergetic process by about 2.5 kcal/mol. The proton transfer turns out to be favoured by about 7 kcal/mol when the formed hydroperoxide adduct is in the most stable configuration (**BH**; *vide infra*).

Several DFT studies have already investigated the binding of a diatomic oxygenic ligand modeled as a hydroperoxide (O_2H^-). Investigations performed on models containing only the bimetallic cluster were not able to reproduce the crystallographic structure [202, 212], whereas a theoretical structure with the side-on coordination mode as observed experimentally (**AH** in Scheme 4.1) as identified by Siegbahn [88] using a larger model, derived from the X-ray structure of *D. gigas* [NiFe]-hydrogenase, and by Söderhjelm and Ryde using a QM/MM model of the enzyme [78] from *D. fructosovorans*. However, this species was found by Siegbahn to be less stable than other two species; one in which O1 symmetrically bridges the two metal atoms, and the other featuring the two oxygen atoms terminally coordinating Ni and Fe in the μ-η^1:η^1-O_2 mode (species **BH** and **CH-2**, respectively, in Scheme 4.1) [88].

It turned out that also using our large model, side-on binding mode of the O_2H^- ligand to Ni is less favoured with respect to that in which O1 is symmetrically coordinated to both metal atoms (see Scheme 4.1). Protonation of each of the two oxygen atoms was investigated; the species **AH-2** was unstable as it converged to **BH** during geometry optimization, while the species **AH-1** in which the proton was placed on O1 is a minimum on the PES, even though less stable than **BH** by 10.7 kcal/mol . In addition, even if **AH-1** features the crystallographically observed side-on binding mode, the distances between O1 and the metal ions Ni-O1 = 2.51 Å and Fe-O1 = 2.32 Å (see Figure 4.5) differ significantly from the experimental ones (1.90 Å [61] and 1.70 Å [77]; 1.90 Å [61] and 2.20 Å [77]).

Other isomers, such as **CH-1** and **CH-2**, with a μ-η^1:η^1 coordination mode of the HO_2^- ligand, and **DH-2** with the end-on coordination of the same ligand were also identified in this work as minima on the PES (see Figure 4.5, Scheme 4.1). These species are higher in energy than **BH** by 11.0, 6.8, and 3.5 kcal/mol, respectively, but they are all more stable than **AH-1**.

In the most stable hydroperoxide species **BH**, the Ni centre has a distorted square pyramidal geometry with the Sγ atom of Cys558 in apical position. The larger stability of this species can be explained considering the features of the hydrogen-bonding network at the interface between the bimetallic cluster and the protein environment. The hydrogen atom of the hydroperoxide ligand forms an H-bond with the guanidinium group of Arg487 and a second H-bond with a cyanide ion coordinated to the iron atom. The formation of these stabilizing interactions is likely to correspond to the driving force which directly leads to adduct **BH** when the optimization is started from the side- on coordination mode of the hydroperoxide. The unpaired electron in **BH** is prevalently localized on the Ni(III) ion (0.59) and on the apical Sγ atom of Cys558 (0.24). Charge and spin population of the O_2H^- ligand, equal to -0.44 and 0.06, respectively, clearly support its assignment to a hydroperoxide.

A comparison with relevant X-ray structures evidences that the Ni-Fe distance in adduct **BH** is close to the experimental X-ray values (2.92 vs. 2.93 Å [61] and 2.80 Å [77]). The same holds true for the Fe-O bond length (2.03 Å compared with 1.90 Å [61] and 2.20 Å [77]), whereas the key HO_2^--Ni distances are not well reproduced by the calculations due to the different hydroperoxide-binding modes. In particular, the Ni-O_2 distance is 2.69 Å, about 0.7 Å shorter than in the experimental structures.

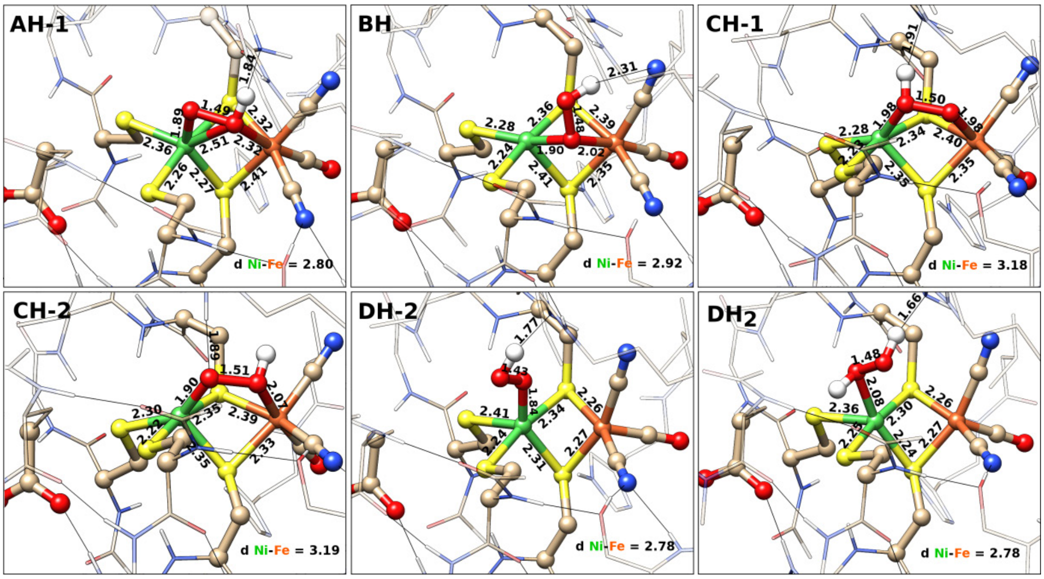
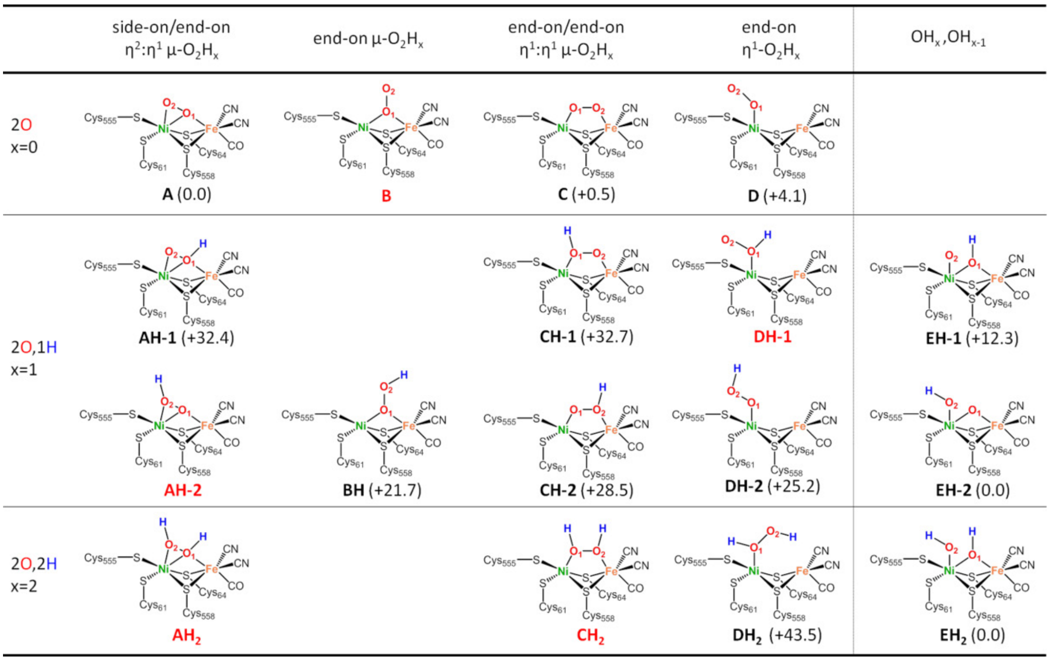


Figure 4.5 Schematic representation of the optimized geometries of the [NiFe] hydrogenase active site coordinated by a hydroperoxide (O_2H^-) and a hydrogen peroxide (H_2O_2) ligand. Selected interatomic distances are given in Å. For the sake of clarity, aliphatic hydrogen atoms are not shown.

Coordination of the neutral hydrogen peroxide (H_2O_2) ligand was investigated in this study as well. However, all species in which the H_2O_2 ligand bridges the two metal ions were unstable and, during geometry optimization, converged to more stable structures in which the H_2O_2 terminally binds the Ni atom or the O-O bond is spontaneously cleaved. The latter species, in which one hydroxide bridges the Ni-Fe site and the other one terminally binds the Ni atom, is more stable than the former by as much as 43.5 kcal/mol, and it will be discussed in detail in the following section. The species with H_2O_2 terminally coordinated to Ni (**DH$_2$** in Figure 4.5, Scheme 4.1) has a distorted square pyramidal arrangement of ligands, in which, however, the apical position is occupied by the Sγ atom of Cys61, and with the Ni-O1 and O1-O2 distances equal to 2.08 and 1.48 Å, respectively. This form is stabilized by the formation of a strong H-bond between the hydrogen atom bounded to O_2 and the guanidinium group of Arg487. The unpaired electron is mainly localized on Ni (0.53) and the apical Sγ atom of Cys61, therefore indicating that the spin is partly delocalized on the apical ligand, independently of the specific nature of the latter. Clearly, this form, as shown in Figure 4.5, differs significantly from the X-ray structures reported in Refs. [61] and [77]. It should also be noted that proton transfer from the protonated Glu14 to the hydroperoxide species in **DH-2** to give the species **DH$_2$** is an endoenergetic process by more than 12 kcal/mol, indicating that protonation of the hydroperoxide ligand to give the coordinated hydrogen peroxide is a strongly disfavoured process.

On the basis of the results discussed above, the coordination of a diatomic oxygenic ligand is compatible with the X-ray structures only in the case of the O_2^{2-} ligand. However, experimental evidences do not support the conclusion that it corresponds to Ni-A. In fact, although peroxide species are generally sensitive to UV light, no change of EPR signal was observed after exposure of Ni-A to this radiation [89]. Moreover, it is rather unlikely that anaerobic oxidation to Ni-A, in which oxygen atoms are provided by water molecules [80], can actually lead to the formation of an O-O bond in the potential range used in electrochemical experiments or with the chemical oxidants used to obtain Ni-A.

Scheme 4.1 Schematic representation of the investigated species containing a diatomic oxygenic ligand or two monoatomic oxygenic ligands. Geometries tagged with a red label refer to unstable species. The relative energies (values in parenthesis) are in kcal/mol.



4.2.2 Characterization of Ni-A as containing monoatomic oxygenic ligands

Another plausible hypothesis for Ni-A is that the bridging ligand features a single oxygen atom. This hypothesis is supported by the crystal structure of the [NiFe]-hydrogenase assigned to the Ni-A state from the photosynthetic bacterium *Allochromatium vinosum* [79], in which the extra bridging ligand between the Ni and the Fe atoms in the active site has been identified as a hydroxide ligand, as found for the Ni-B state, but probably holding a different conformation thus leading to the observed peculiar spectroscopic properties.

On the basis of these experimental evidences and recent crystallographic studies, several structures characterized by a bridging single oxygen-containing ligand were investigated as candidates for the Ni-A state. In particular, structures in which a hydroxide ligand bridges the two metal ions and an additional mono-oxygenic species binds the active site were considered. The likely presence of a second O atom in the Ni-A state is also supported by the very dissimilar reactivation properties of the Ni-B and the Ni-A states that may not be due only to the different conformations of the bridging hydroxide. In this context, the species **EH-1** and **EH-2**, that can be ideally considered as deriving from the cleavage of the O-O bond of the hydroperoxide ligand, have been first investigated (see Figure 4.6, Scheme 4.1).

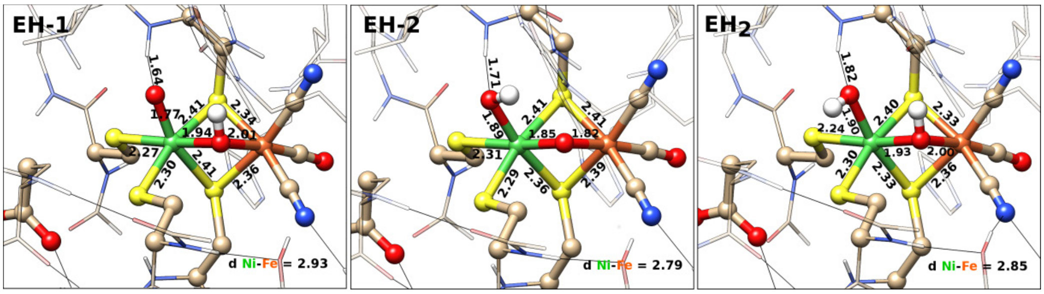


Figure 4.6 Schematic representation of the geometries in which the [NiFe] hydrogenase active site is coordinated by monoatomic oxygenic ligands. Selected interatomic distances are given in Å. For the sake of clarity, aliphatic hydrogen atoms are not shown

Notably, these species are more stable than **BH** by 9.4 and 21.7 kcal/mol, respectively, revealing that cleavage of the O-O bond of the hydroperoxide ligand is a strongly favoured process. The stability of **EH-1** and **EH-2** with respect to the hydroperoxide-bounded species **BH** should also be due to the formation of a new strong interaction between the oxygen atom coordinated to Ni and the back- bone amide hydrogen atom of Cys64 (in **EH-1** and **EH-2** the O2···HN distance is 1.64 and 1.71 Å, respectively, while it is longer than 4 Å in **BH**). In these species, in which the coordination of both metal ions is octahedral, the Ni–O1, Fe–O1, and Ni-Fe distances equal, respectively, to 1.94, 2.01 and 2.93 Å for **EH-1** and 1.85, 1.82, and 2.79 Å for **EH-2** reproduce the experimental structures reported in Ref. [61] (1.90, 1.90, and 2.93 Å) and [77] (1.70, 2.20, and 2.80 Å). However, the O-O distance (2.59 and 2.65 Å in **EH-1** and **EH-2**, respectively) is much longer compared with that in the crystallographic structures (1.45 Å [61] and 1.57 Å [77]), suggesting that **EH-1** and **EH-2** do not correspond to any of the crystallographic structures resolved so far. The species **EH-1** and **EH-2** are also interesting for their peculiar electronic structures. In fact, in **EH-2**, the unpaired electron turned out to be localized on the oxygen atom bridging the metal ions and on the Fe atom (spin populations: 0.47 and 0.40, respectively); a similar picture is observed in the case of **EH-1**, where the O2 atom terminally coordinated to Ni features a relatively large atomic spin value (0.59). In other words, atomic charge and spin population analyses show that in the latter two species, the Ni ion is diamagnetic and attains the Ni(IV) state, whereas the oxo ligand is partially reduced. Partial oxidation also occurs at the level of the Fe ion in **EH-2**.

Protonation of both oxygen atoms leads to species **EH$_2$** (Figure 4.6, Scheme 4.1) which is also more stable than the hydrogen peroxide bounded species (**DH$_2$**) by as much as 43.5 kcal/mol. In **EH$_2$**, the Ni atom features an octahedral coordination with the terminal OH group forming an H-bond with the amide hydrogen atom of Val63. The Ni-O1 and Fe-O1 distances as well as the O1···O2 distances are very similar to those calculated for the **EH-2** species. The unpaired electron is also largely localized on the Fe atom and the bridging OH group.

4.2.3 Characterization of Ni-A as containing a monoatomic oxygenic ligand and oxidized cysteine residues

The more recent X-ray structure of *Desulfovibrio fructosovorans* [NiFe]-hydrogenase assigned to the Ni-A state [63] was characterized by a hydroxide ligand bridging the two metal ions, and the bridging Cys64 residue oxidized to bridging sulfenate, with an oxygen atom inserted between sulphur and Ni atoms. In several previous crystallographic investigations, partial occupancy states were also resolved by attributing residual density peaks to oxi- dative modifications of the bridging Cys64 or the terminal Cys555 residues [61, 77, 213, 214]. Notably, the species corresponding to the putative Ni-A state in Ref. [63] (species **FH-Cys64O$_b$** in Figure 4.7, Scheme 4.2) turned out to be more stable than any other investigated isomer with two oxygens and one proton associated with the active site. In particular, **FH-Cys64O$_b$** is more stable than **BH** and **EH-2** by as much as 42.1 and 20.4 kcal/mol, respectively, in the former case in agreement with the previous theoretical calculations [78].

Furthermore, the optimized structure of **FH-Cys64O$_b$** reproduces very well the geometry of the crystallographic structure [78]. The RMSD of the bimetallic cluster (including the side chain of Glu14) compared with the crystallographic structure of putative Ni-A (pdb code 4UPE) [63] (is 0.24 Å. The Ni atom features a distorted square pyramidal coordination, in which the Sγ atom of Cys558 occupies the apical position. This apical ligand is slightly bent along the SCys61-Ni-O(H) axis with the SCys61-Ni-SCys558 and O(H)-Ni-SCys558 angles equal to 110.5 and 82.5°, respectively. The calculated Ni-Fe distance is 3.11 Å (expt. = 3.10 Å), whereas the Ni–OH and the Fe-OH distances are equal to 1.94 Å (expt. = 1.92 Å) and 2.03 Å (expt. = 2.05 Å). The Ni-OSCys64 distance equal to 1.95 Å is also in agreement with the experimental one (1.89 Å). **FH-Cys64O$_b$** is stabilized by the formation of a strong H-bond between the bridging sulfenate oxygen atom and the amidic hydrogen of Cys64 (O···HN = 1.80 Å) which replaces the weaker interaction of the Sγ atom of Cys61 with the same amidic group observed in other oxidized and active forms of the enzyme.

Analysis of the electronic structure of **FH-Cys64O$_b$** shows that the unpaired electron is prevalently localized on the Ni ion (0.60) and the Sγ atom of the apical Cys558 (0.26), in accord with the Ni(III)Fe(II) assignment. Atomic NBO charges show that the two oxygen atoms contained in **FH-Cys64O$_b$** are fully reduced with the concomitant oxidation of Sγ of Cys64, the charge of which is, indeed, equal to +0.7, a value to be compared with the -0.2 charge units computed for the S atoms of the other coordinating Cys residues. This finding is in line with protein film voltammetry experiments, suggesting that Ni-A differs from Ni-B in having two extra oxidizing equivalents [51], stored as a trapped O atom, yielding a sulfoxide [215], sulfenic acid [61], or as a coordinated peroxo-ligand.

Oxidation of the other cysteines coordinating the Ni atom (Cys61, Cys555, and Cys558) to bridging sulfenate (**FH-Cys61O$_b$**, **FH-Cys555O$_b$**, and **FH-Cys558O$_b$** in Figure 4.7 and Scheme 4.2) was also investigated to unveil if such oxidations can compete with that of Cys64.

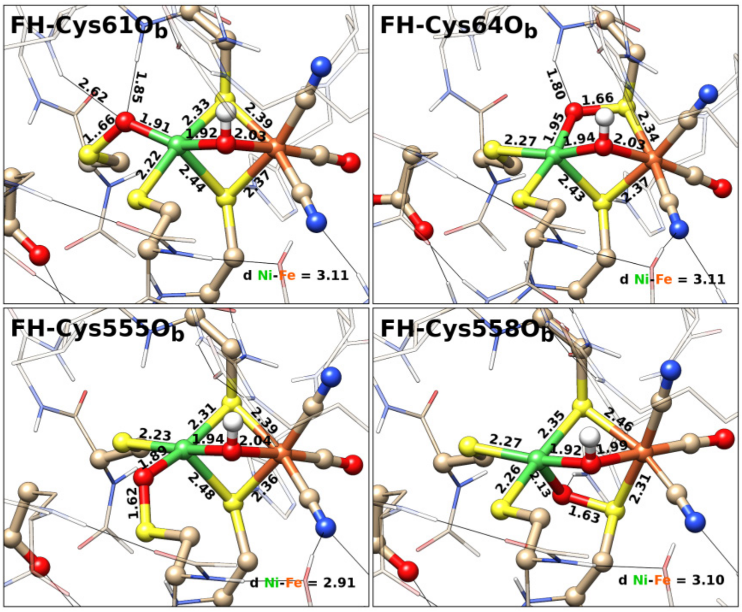
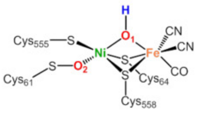
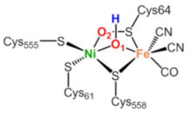
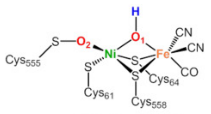
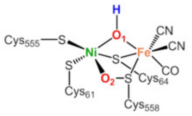
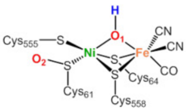
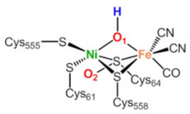
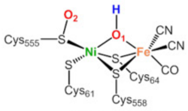
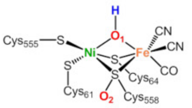
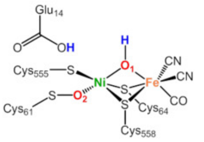
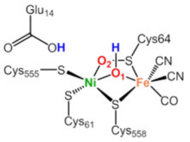
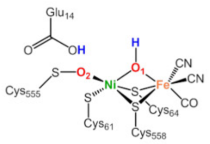
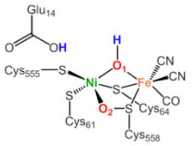
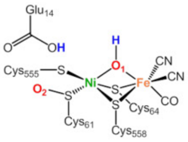
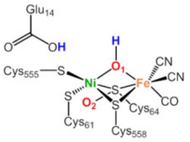
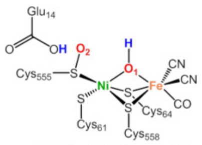
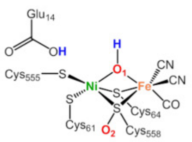


Figure 4.7 Schematic representation of the DFT calculated geometries of **FH-Cys61O$_b$**, **FH-Cys64O$_b$**, **FH-Cys555O$_b$**, and **FH-Cys558O$_b$** species. Selected interatomic distances are given in Å. For the sake of clarity, aliphatic hydrogen atoms are not shown

Notably, all these species are less stable than **FH-Cys64O$_b$**, by more than 9 kcal/mol, corroborating the experimental finding on the selective oxidation of Cys64. All these species feature a pseudo-square pyramidal coordination of the Ni atom with the same distortion of the apical ligand observed in **FH-Cys64O$_b$**, with the exception of **FH-Cys558O$_b$** for which the sulfenate oxygen at the apical position is roughly perpendicular to the base of the pyramid Figure 4.7. The electronic structure of the other three bridging sulfenate isomers is also similar to that of **FH-Cys64O$_b$**, with the unpaired electron mainly localized on the Ni ion (0.60-0.65), and on the atom at the apical position of the square pyramid, which is SCys558 for **FH-Cys61O$_b$** and **FH-Cys555O$_b$**, and the bridging sulfenate oxygen for **FH-Cys558O$_b$**.

Insertion of the bridging sulfenate oxygen atom in **FH-Cys61O$_b$** and **FH-Cys558O$_b$** leads to the formation of strong H-bonds with the amidic hydrogen of Cys64 (d O···HN = 1.85 Å) and with the protonated nitrogen of His68 (d O···HN = 1.96 Å), respectively, that, in the two cases, replace the interaction of SCys64 and SCys558 with the same H-bond donors. Conversely, in **FH-Cys555O$_b$**, the sulfenate oxygen does not form novel H-bonds with neighbouring residues (see Figure 4.7). For the most stable **FH-Cys64O$_b$** species, other protonation states were investigated to verify whether they are more compatible with the recent crystallographic structure of Ref. [63]. In this respect, when the bridging ligand is deprotonated–independently of the protonation state of the bridging sulfenate group–the distances between the oxo ligand and the Ni and the Fe ions are significantly shorter than the experimental values. On the other hand, species containing protonated sulfenate groups may not be considered good candidates for the Ni-A state due to their very long Ni-O(S) bond distance. Furthermore, the species containing an oxo bridging ligand and a protonated Cys64 sulfenate group is less stable than in the differential stabilization of the bridging sulfenate isomers can be fur- ther isomer **FH-Cys64O$_b$** by as much as 36.3 kcal/mol.

Scheme 4.2 Schematic representation of the investigated species containing a monoatomic oxygenic ligand and oxidized cysteine residues. The relative energies (values in parenthesis) are in kcal/mol.

	OH, S(Cys$_{61}$)O,GluH$_x$	OH, S(Cys$_{64}$)O,GluH$_x$	OH, S(Cys$_{555}$)O,GluH$_x$	OH, S(Cys$_{558}$)O,GluH$_x$
2O,1H x=0	 **FH-Cys61O$_b$** (+9.6; **+10.3**)	 **FH-Cys64O$_b$** (0.0 ; **0.0**)	 **FH-Cys555O$_b$** (+11.9 ; **+10.7**)	 **FH-Cys558O$_b$** (+12.0 ; **+14.8**)
	 **FH-Cys61O$_t$** (+1.5 ; **+7.4**)	 **FH-Cys64O$_t$** (+3.6 ; **+10.1**)	 **FH-Cys555O$_t$** (+5.7 ; **+9.0**)	 **FH-Cys558O$_t$** (+4.2 ; **+11.3**)
2O,2H x=1	 **FH-Cys61O$_b$-GluH** (+14.1; **+11.3**)	 **FH-Cys64O$_b$-GluH** (+2.5; **0.0**)	 **FH-Cys555O$_b$-GluH** (+15.2; **+14.7**)	 **FH-Cys558O$_b$-GluH** (+15.6; **+14.3**)
	 **FH-Cys61O$_t$–GluH** (+3.7; **+5.3**)	 **FH-Cys64O$_t$-GluH** (+8.6; **+11.2**)	 **FH-Cys555O$_t$-GluH** (0.0; **+1.5**)	 **FH-Cys558O$_t$-GluH** (+8.4; **+12.1**)

The role of the protein environment inspected by considering small models of the active site, which only includes the first coordination sphere, and by comparing their relative stabilities to those calculated with the large model. In fact, when single point energy calculations were carried out on such small models in the geometries of large models, **FH-Cys64O$_b$** still results to be the most stable isomer (see Table 4.3). **FH-Cys558O$_b$**, having the other bridging cysteine oxidized to sulfenate, is about 4.5 kcal/mol less stable than **FH-Cys64O$_b$** whereas the other two bridging sulfenate isomers **FH-Cys61O$_b$** and **FH-Cys555O$_b$** in which oxidation is carried out on the terminal cysteines are higher in energy with respect to **FH-Cys64O$_b$** by about 8 kcal/mol. These results indicate that the stability of **FH-Cys64O$_b$** is not only due to the formation of a more favourable H-bond network, but also to intrinsic stereoelectronic features of the bimetallic cluster in this configuration. In particular, the constraints imposed by the protein on the geometry of the active site should favour **FH-Cys64O$_b$** with respect to **FH-Cys558O$_b$**, as Cys64 is in the basal position, whereas Cys558 occupies the apical position of the square pyramid configuration. The insertion of an oxygen atom in **FH-Cys558O$_b$** may be disfavoured by the rigidity imparted by the Cys555/Cys558-binding motif that is included in a small α-helix at the C-terminus of the large subunit. Structural effects on the relative stability of **FH-Cys64O$_b$** and **FH-Cys558O$_b$**, can also be inspected by comparing their geometries to that of **Ni-B**, which contains only the bridging hydroxide as exogenous ligand. Figure 4.8 shows the structures of **FH-Cys64O$_b$** (orange) and **FH-Cys558O$_b$** (blue) overlaid on the **Ni-B** geometry (gray), underlining the effect exerted by the bridging cysteine modifications on the geometry of the active site.

Table 4.3 Relative stabilities (in kcal/mol) of the sulfenate isomers calculated for the large model (LM) and the small model (SM). In the latter case the relative stabilities refer to single point energy calculations on the geometry of the large model.

	Cys61O$_b$	Cys64O$_b$	Cys555O$_b$	Cys558O$_b$	Cys61O$_t$	Cys64O$_t$	Cys555O$_t$	Cys558O$_t$
LM	+9.6	0.0	+11.9	+12.0	+1.5	+3.6	+5.7	+4.3
SM	+8.1	0.0	+8.2	+4.8	+9.2	+15.2	+10.3	+12.3

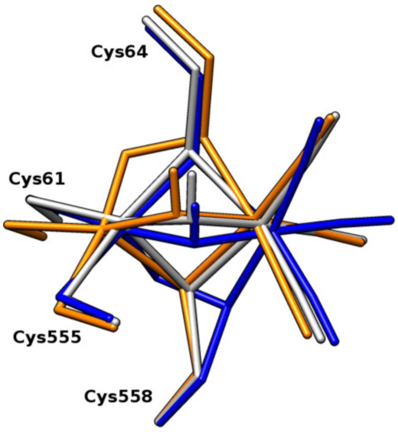


Figure 4.8 Comparison between the **FH-Cys64O$_b$** (orange), **FH-Cys558O$_b$** (blue), and **Ni-B** (gray) structures, optimized using the large model.

It can be noted that in **FH-Cys64O$_b$**, the position of SCys64 is not significantly affected by the oxidation (the deviation from the position of the corresponding atom in **Ni-B** is only 0.43

Å), whereas oxidation of Cys558 leads to the displacement of the corresponding sulphur atom by as much as 0.92 Å, compared with its position in Ni-B. This structural rearrangement in **FH-Cys558O$_b$** promotes a shift of the bridging hydroxide from its original position (deviations of the oxygen and hydrogen atoms with respect to the corresponding hydroxide positions in **Ni-B** are 0.62 and 0.63 Å, respectively), leading to the cleavage of the H-bonds of the hydroxide ligand with Arg487 (d OH···N = 3.20 Å), and with the NH group of Cys64.

Oxidation of the cysteine residues coordinated to the Ni atom can also occur to give terminal sulfenates, in which oxygen is terminally bounded to a S atom. Indeed, X-ray structures of [NiFe]- and [NiFeSe]- hydrogenases identified low occupancy states with sulphur atoms of cysteines coordinated to Ni oxidized by terminal oxygen atoms [216–218]. The species in which Cys61, Cys64, Cys555, and Cys558 have been oxidized to the corresponding terminal sulfenates are labelled **FH-Cys61O$_t$**, **FH-Cys64O$_t$**, **FH-Cys555O$_t$**, and **FH-Cys558O$_t$**, respectively, and they are shown in Figure 4.9 and Scheme 4.2. Notably, all terminal sulfenate species are slightly less stable than **FH-Cys64O$_b$** but more stable than all other bridging sulfenate isomers. In particular, all terminal sulfenates are more stable than the corresponding bridging sulfenate isomers by more than 6 kcal/mol, with the exception of **FH-Cys64O$_t$**, which is less stable than **FH-Cys64O$_b$** by 3.6 kcal/mol. On the other hand, the most stable terminal sulfenate **FH-Cys61O$_t$** is only 1.5 kcal/mol higher in energy than **FH-Cys64O$_b$**.

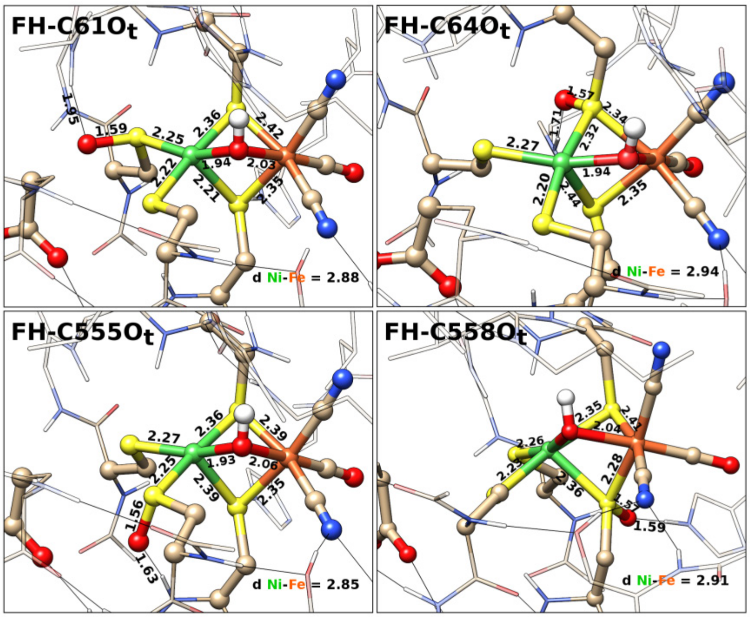


Figure 4.9 Schematic representation of the DFT optimized geometries of **FH-Cys61O$_t$**, **FH-Cys64O$_t$**, **FH-Cys555O$_t$**, and **FH-Cys558O$_t$** species. Selected interatomic distances are given in Å. For the sake of clarity, aliphatic hydrogen atoms are not shown.

In all of the terminal sulfenate isomers, the Ni atom preserves the distorted square pyramidal coordination discussed above, with Sγ atom of Cys558 occupying the apical position (see

Figure 4.9). The S-Ni and S-Fe distances are not significantly affected by the oxidation of the sulphur atom to sulfenate. The electronic structure of the terminal sulfenate isomers is similar to that of the corresponding bridging sulfenate species, but with the unpaired electron slightly delocalized also on the oxidized sulphur.

Terminal sulfenate species are stabilized by the formation of novel strong H-bonds between the sulfenate oxygen atom and donors of the nearby residues. In particular, in **FH-Cys61O$_t$**, a novel H-bond is formed between the sulfenate oxygen and the hydrogen amide of Val63 (d O···HN = 1.95 Å), in **FH-Cys64O$_t$** and **FH-Cys558O$_t$** between the sulfenate oxygen and the protonated nitrogen of the His68 imidazole ring (d O···HN = 1.71 and 1.60 Å, respectively), and in **FH-Cys555O$_t$** between the sulfenate oxygen and the amidic hydrogen of Cys558 (d O···HN = 1.63 Å). It is interesting to note that when single point calculations were carried out on small models in the geometries of the large models, the terminal sulfenate isomers are less stable than **FH-Cys64O$_b$** by more than 9 kcal/mol (see Table 4.3). These values are larger than those calculated using the large model, indicating that terminal sulfenates in the geometry of the active site constrained by the protein environment are less stable than the corresponding bridging sulfenates. The stabilities of the terminal sulfenate isomers are significantly increased due to the formation of a more favourable H-bond network with the residues in the nearby of the active site.

Protonation of the Glu14 residue may promote the formation of a strong H-bond with the bridging or terminal sulfenate oxygen atom of Cys555, therefore affecting the relative stability of the corresponding sulfenate isomers. Indeed, according to our calculations, **FH-Cys555O$_t$-GluH** (Figure 4.10, Scheme 4.2), in which Glu14 is protonated and the terminal Cys555 is oxidized to terminal sulfenate, results to be more stable than any other sulfenate in such conditions. **FH-Cys64O$_b$-GluH** (Figure 4.10, Scheme 4.2)–the most stable isomer when Glu14 is not protonated–is higher in energy with respect to **FH-Cys555O$_t$-GluH** by 2.5 kcal/mol.

It is worth noting that residual density peaks in the X-ray structures determined by Ogata *et al.* [77] were assigned to an oxygen atom bounded to Cys555 at a position very similar to that observed in **FH-Cys555O$_t$-GluH**. Indeed, the RMSD of the bimetallic cluster calculated by comparing **FH-Cys555O$_t$-GluH** with the corresponding X-ray structure (pdb code 1WUK) [77] is as small as 0.19 Å. This value increases to 0.28 Å when Glu14 is included in the calculation, due to slightly different orientations of the side chain of this residue between the theoretical and the X-ray structures. It is interesting to note that oxidation of the Sγ atom of Cys555 to terminal sulfenate does not alter the geometry of the cluster as all geometrical parameters remain almost identical to those calculated for Ni-B. Therefore, the formation of this species can be explained by assuming that crystals were grown with a partial protonation of Glu14.

The protonation of the Glu14 does not appear to affect the stability of the species in which Cys555 is oxidized to bridging sulfenate (**FH-Cys555O$_b$-GluH**) as it still remains 12.9 kcal/mol less stable than **FH-Cys64O$_b$-GluH**, a value very similar to the one calculated when Glu14 is not protonated. It is also important to note that the proton transfer from Glu14 to the sulfenate oxygen atoms is always a strongly endoenergetic process, and in the case of **FH$_2$-**

Cys555O$_b$, it does not occur as the pro- ton spontaneously migrates from Sγ to Glu14 to give the **FH-Cys555O$_b$-GluH** species (see Scheme 4.3).

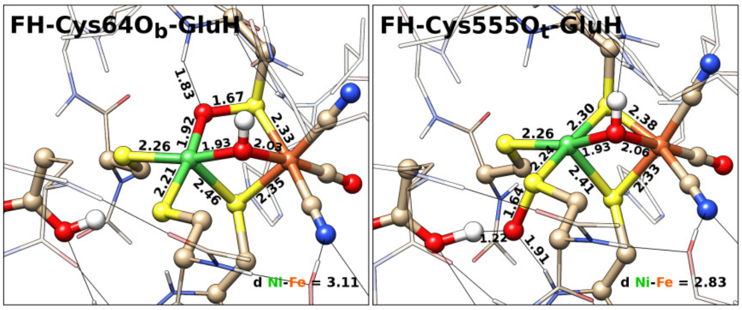
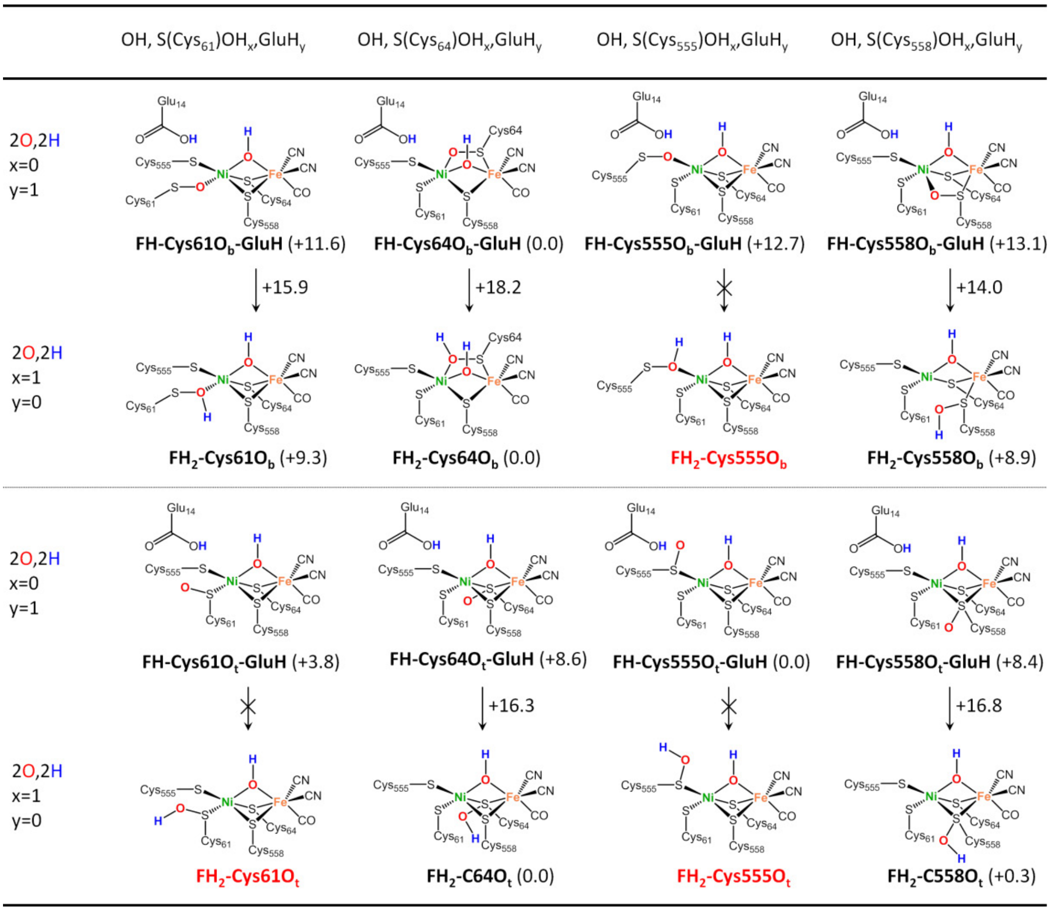


Figure 4.10 Schematic representation of the DFT calculated geometries of **FH-Cys64O$_b$**-GluH and **FH-Cys555O$_t$-GluH** species. Selected interatomic distances are given in Å. For the sake of clarity, aliphatic hydrogen atoms are not shown terminal.

Scheme 4.3 Schematic representation of the proton transfer reaction from the protonated Glu14 residue to the sulfenate oxygen atom. Red labels refer to unstable species, resulting in proton release from the sulfenate oxygen atom towards Glu14 upon geometry optimization. Energy differences (values in parenthesis) and reaction energies are in kcal/mol.



According to the results discussed above, oxidation of the [NiFe]-hydrogenase enzymes should lead to the selective oxidation of Cys64 to bridging sulfenate in agreement with the X-ray structure recently proposed by Fontecilla-Camps *et al.* [63]. It is worth noting that the insertion of the oxygen atom into the Ni-S bond in **FH-Cys64O$_b$** does not alter the spin density distribution with respect to **Ni-B**, as in both cases the spin is almost totally localized on Ni and the apical Sγ atom of Cys558. Indeed, atomic spin densities as well as the spin isodensity surfaces and the SOMOs of **FH-Cys64O$_b$** and **Ni-B** almost perfectly match (see Table 4.4 and Figure 4.11). This analysis may explain why Ni-A and Ni-B feature very similar ENDOR spectra [14], despite the fact that the Ni atom has a different coordination sphere in the two structures. In addition, the calculated Cα-Cβ-Sγ-Ni dihedral angles between **FH-Cys64O$_b$** and **Ni-B**, equal to 94.7 and 102.9°, respectively, span a difference which is in perfect agreement with the value predicted by the analysis of the experimental hyperfine coupling constants of the β-CH2 protons of Cys558 for the Ni-A and Ni-B species [197].

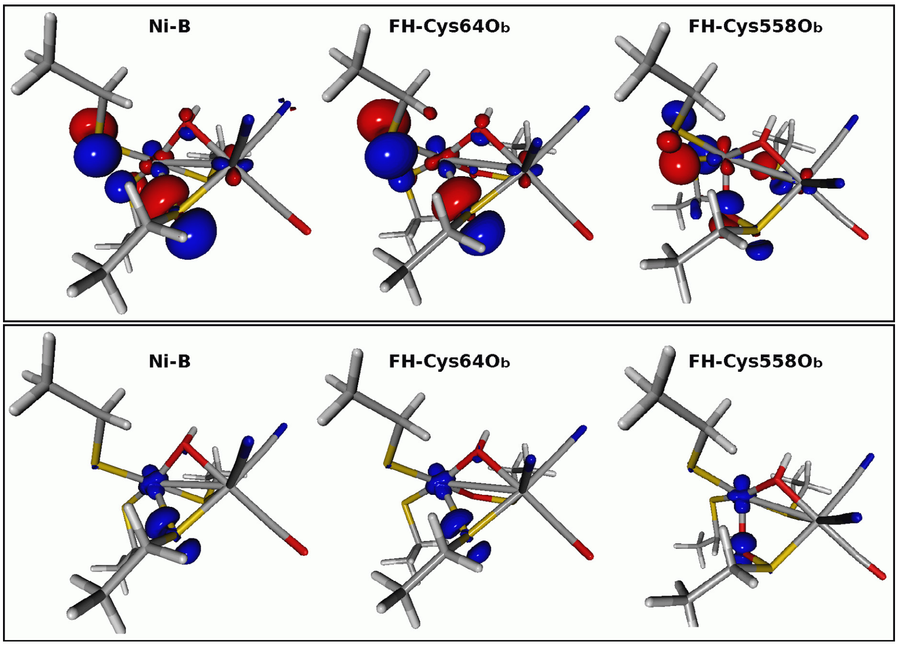


Figure 4.11 SOMO (top) and spin isodensity surface (bottom) of the a) Ni-B, b) FH-Cys64Ob and c) FH-Cys558Ob species calculated using the small model, in which all atoms with the exception of hydrogens have been constrained at the position computed for the large model.

Protonation of Glu14 should promote the formation of a terminal sulfenate on Cys555, as, indeed, observed in minor occupancy forms of several X-ray structures (*vide supra*). Oxidation of the other cysteines to sulfenates should be ruled out as the corresponding species are less stable in both conditions. The orientation of the side chain of Glu14 in the different X-ray structures might give some insight into its protonation state. In this regard, our calculations show that when Glu14 is deprotonated, it forms an H-bond with the amide hydrogen atom of Ala557, whereas when Glu14 is protonated, it approaches the Sγ atom of Cys555 by changing the orientation of the carboxylate group. Comparison of the geometry in the models of **FH-Cys64O$_b$** and **Ni-B** with the corresponding X-ray structures shows that the agreement is much better when Glu14 is deprotonated, strongly suggesting such protonation state for the Glu14 in the X-ray structures. However, when **FH-Cys555O$_t$** is considered, the picture is less clear, as explained in the following. In the X-ray structures featuring terminal sulfenate on Cys555, the Glu14 residue has an orientation in between the one observed in the protonated model and the alternative one in the unprotonated model, with the latter being in slightly better agreement with experiments. In this respect, it should be noted that the quantum refinement of a crystallographic structure of a [NiFe] hydrogenase from *Desulfovibrio fructosovorans* is compatible with the presence of a minor occupancy form featuring Cys555 oxidized to sulfenate and a concomitant rearrangement of the side chain of Glu14 to an alternative position, indicating that Glu14 is disordered in the crystal structure in a way that depends by the oxidation of Cys555 [78].

Table 4.4 Energy of the SOMO orbital (eV) and atomic Mulliken spin populations of the species **FH-Cys64O$_b$**, **FH-Cys558O$_b$** and the **Ni-B** state, calculated for the large (LM) and the small model (SM). In the former case the geometries are those of the large model where only hydrogen atoms have been optimized.

	SM			LM		
	FH-Cys64O$_b$	**FH-Cys558O$_b$**	**Ni-B**	**FH-Cys64O$_b$**	**FH-Cys558O$_b$**	**Ni-B**
SOMO	-2.225	-2.293	-2.200	-2.058	-2.198	-2.014
Spin Ni	0.59	0.63	0.56	0.60	0.65	0.58
spin Fe	0.01	0.02	0.02	0.00	0.02	0.02
spin O(H)	0.02	-0.01	0.01	0.02	0.00	0.02
spin S(Cys61)	0.02	0.02	0.03	0.02	0.00	0.02
spin S(Cys64)	0.00	0.03	-0.02	0.00	0.02	-0.02
spin S(Cy555)	0.06	0.07	0.07	0.07	0.08	0.10
spin S(Cy558)	0.29	0.04	0.30	0.26	0.04	0.27
spin O$_b$(Cys64)	0.00	-	-	0.00	0.02	-
spin O$_b$(Cys558)	-	0.19	-	-	0.16	-

4.2.4 Conclusions

The nature of the oxidized inactive Ni-A state in [NiFe]-hydrogenases has remained elusive until recently when a crystallographic structure containing a bridging hydroxide and a coordinating cysteine oxidized to bridging sulfenate has been confidently assigned to this state [63]. The previous crystallographic analysis has identified the Ni-A state as containing a (hydro)peroxide ligand coordinated to the Ni atom. The presented results show that only a peroxide (O_2^{2-}) ligand is compatible with the two crystallographic structures reported in Refs. [61] and [77], corroborating the conclusions of previous theoretical investigations. Hydroperoxide ligand (HO_2^-) should be excluded as the most stable coordination mode differ significantly from the crystallographic structures and, in addition, the hydroperoxide species are at least 25 kcal/mol less stable than the isomer in which the O-O bond is cleaved and one oxy- gen atom moves to a terminal position on Ni. On the other hand, the species coordinating peroxide or hydroperoxide ligand, or a couple of oxygen atoms originating from O-O bond cleavage are probably intermediates along the mechanism of aerobic oxidation leading to the Ni-A state.

The main focus of our theoretical characterization of the Ni-A state was to investigate the relative stabilities of the isomers with coordinating cysteines oxidized to bridging and terminal sulfenates. Indeed, all these species are more stable than any hydroperoxide forms by at least 30 kcal/mol, as well as than any species in which the bound O_2 is dissociated by at least 8 kcal/mol. The sulfenate species corresponding to the X-ray structure recently assigned to Ni-A [63] is more stable than all other sulfenate isomers, supporting the experimental assignment to Ni-A for such a structure. Its stability is due to both the peculiar stereoelectronic feature of the oxygen atom inserted into that position and the formation of a more favourable H-bond network. On the other hand, the species in which the coordinating cysteines are oxidized to terminal sulfenates are only slightly less stable then putative Ni-A, and their relative stabilities are largely affected by the network of H-bonds they form. In particular, the protonation of Glu14 promotes the formation of a strong H-bond with the Cys555 terminal sulfenate, making this species the most stable one among the considered isomers. This can explain the X-ray determination of partial occupancy states in which Cys555 is oxidized [77] with the oxygen atom at a position very close to the one obtained from theoretical calculations. The protonation of Glu14 can clearly depend on the conditions in which the crystals were grown. In this regard, it has to be noted that analysis of the orientation of the side chain of Glu14 residue in the X-ray structures does not allow to confidently assign the protonation state of this residue. The picture described above is compatible with the possible coexistence of two different forms of the active site in the crystallographically characterized Ni-A samples: one featuring a bridging sulfenate on Cys64, and one featuring a terminal sulfenate on Cys555.

The better understanding of the origin of stability of bridging and terminal sulfenate isomers presented in this study will hopefully drive mutagenesis experiment aimed at altering the relative stabilities of the sulfenate isomers to favour the formation of a terminal sulfenate species which may feature faster reactivation kinetics.

4.3 Oxidative inactivation of [NiFe]-hydrogenases

Under oxidizing (aerobic and anaerobic) conditions, active forms of standard [NiFe]-hydrogenases are reversibly converted into the inactive Ni-A and Ni-B states [50–52]. While under aerobic conditions oxidation is prompted by O_2, under anaerobic conditions oxidation can be achieved by either modulation of the redox potential or addition of strong oxidants different from O_2 [75, 90, 91]. Depending on the reaction conditions, different ratios of Ni-A and Ni-B are obtained. In general, more reducing conditions (lower potential of the oxidizing agent or of the electrode potential) favour the formation of Ni-B over Ni-A [51, 72, 77, 80–

82, 91]. In this context, in order to fully elucidate inactivation pathways of [NiFe]-hydrogenases, both aerobic and anaerobic oxidation have been investigated.

Our theoretical study on aerobic inactivation was guided by the following experimental evidences: if the enzyme is initially reduced to the paramagnetic Ni-C state, its treatment with a saturated solution of O_2 at 0 °C leads to the formation of a pure Ni-B oxidized state [61]; conversely, when a less reduced enzyme is exposed to O_2, the Ni-A state is formed in large amounts [51]. On the basis of these experimental data, the two chemically distinct Ni-C and Ni-SI$_a$ forms have been considered as 'representative' initial states in the aerobic oxidation. In particular, Ni-C and Ni-SI$_a$ are here considered as the more reduced and more oxidized states, respectively, that, upon O_2 exposure, preferentially form the Ni-B or the Ni-A states.

Since the enzyme inactivation by exposure to O_2 is fast, oxidation has been suggested to be an inner-sphere process [51]. Direct O_2 binding to the active site in the Ni-C and the Ni-SI$_a$ forms have been therefore investigated as the first step of the aerobic oxidation of the [NiFe]-hydrogenases. The hypothesis of a direct attack of O_2 to the NiFe binuclear cluster is supported by the simple consideration of the kinetic diameters of H_2 (2.71 Å), O_2 (3.55 Å), and Xe (4.78 Å). Since the latter has been used to study the H_2 channel using X-ray diffraction [219], the O_2 kinetic diameter does not exclude the possibility that O_2 can use the route taken by H_2 to reach the active site [220].

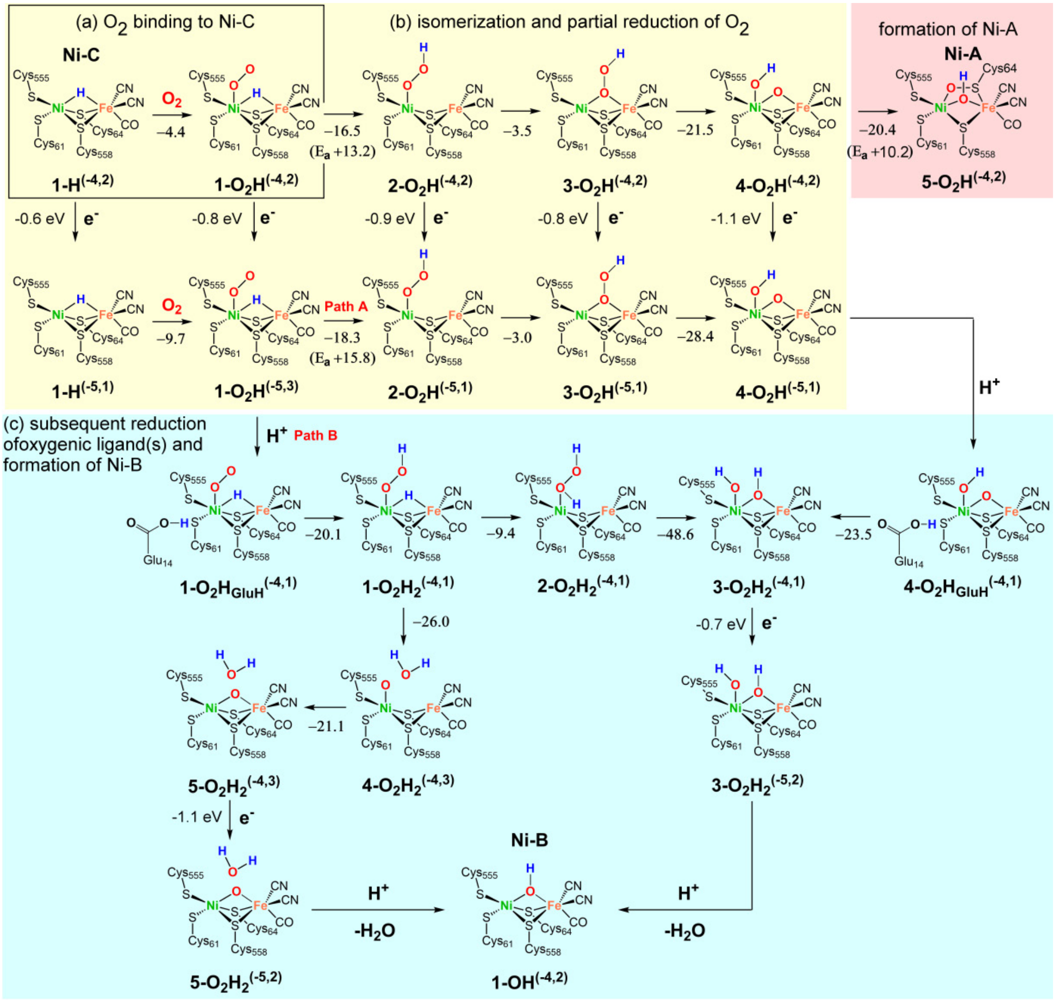
While under aerobic conditions oxidation is promoted by O_2 which also provides the oxygen atoms of the oxygenic species included in Ni-B and Ni-A, under anaerobic conditions these oxygen atoms have been proposed to be provided by two water molecules. This hypothesis is supported by the ^{17}O ENDOR study carried out by Carepo *et al.* [80] that demonstrates that the bridging oxygenic ligand of the Ni-A state originates from a water molecule. Here, also the oxygen atom of the sulfenate group is proposed to take its origin from a water molecule. Thus, direct binding of two H_2O molecules to the active site of [NiFe]-hydrogenases has been considered to study the anaerobic oxidation of the enzyme.

Nomenclature: in the following, computational models of the investigated species will be labelled according to the general scheme $\mathbf{n\text{-}O_xH_y^{(k,j)}}$, where **n** is a generic number referring to the binding mode of the various ligands, containing **x** oxygens and **y** hydrogens, coordinated to the active site in the considered species. Conversely, the **k** and **j** superscripts indicate the charge and the spin state of the model complex, respectively. Additional tag may be also present; when the Glu14 residue is protonated (i.e. it attains neutral state) the label **GluH** is added after the $\mathbf{n\text{-}O_xH_y}$ label.

4.3.1 Aerobic oxidation of Ni-C

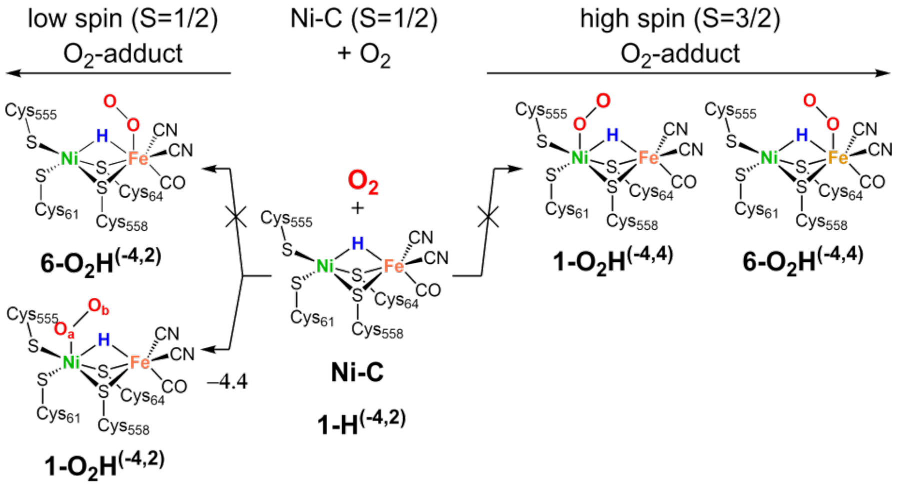
Aerobic oxidation of the active site to the Ni-B form requires fast delivery of sufficient electrons to the active site to completely reduce the attacking O_2 to the bridging hydroxide coordinated to the binuclear cluster, and to a water molecule which is proposed to leave the active site. If O_2 reacts with the active site at the level of Ni-C, corresponding to a $Ni^{III}Fe^{II}$ state with a μ-bridging hydride, two of the four electrons required for the reduction of O_2 may be supplied by the hydride. Two further electrons should be provided by the auxiliary FeS clusters which are assumed to attain the reduced state. The most plausible pathways for the oxidation of Ni-C to Ni-B with concomitant O_2 reduction are shown in Scheme 4.4. For the sake of clarity, the proposed mechanism will be discussed subdividing it into three steps: (a) O_2 binding to Ni-C (highlighted in the boxed region), (b) isomerisation and partial reduction of O_2 (yellow background) and (c) subsequent reduction of the oxygenic ligand(s) and formation of Ni-B (cyan background).

Scheme 4.4 Most plausible mechanisms for aerobic oxidation of Ni-C. (a) O_2 binding to Ni-C (boxed region) and (b) isomerisation and partial reduction of O_2 are marked by a yellow background, while (c) subsequent reduction of the oxygenic ligand(s) and formation of Ni-B is depicted with a cyan background. Conversely, red background denotes the (d) formation of Ni-A. Energy differences and activation energies (E_a, calculated using the MM model) are in kcal/mol while E_{red}-E_{ox} values are in eV.



(a) **O_2 binding to Ni-C.** Binding of O_2 to the EPR active Ni-C (S=1/2) species, **1-H$^{(-4,2)}$**, has been investigated by considering both low (S=1/2) and high (S=3/2) spin states of the O$_2$-adduct (see Scheme 4.4 and Scheme 4.5). O$_2$ can bind to the Ni^{III} atom to give the low-spin O$_2$-adduct **1-O$_2$H$^{(-4,2)}$** with a binding energy equal to −4.4 kcal/mol. Coordination of O$_2$ occurs end-on to Ni, which is then characterized by a pseudo octahedral arrangement of ligands. The Ni-O$_a$-O$_b$ angle is 121.6°, whereas the Ni-O$_a$ and O$_a$-O$_b$ bond lengths are equal to 1.98 and 1.28 Å, respectively (a and b labels indicate the oxygen atoms of the O$_2$ ligand as in Scheme 4.4 and Scheme 4.5). The latter distance is about 0.06 Å longer than in the O$_2$ molecule. Interestingly, spin and charge populations of the O$_2$ ligand in **1-O$_2$H$^{(-4,2)}$**, equal to 1.16 and -0.34, respectively, clearly suggest that O$_2$ is reduced to the superoxo species. The corresponding high-spin adduct **1-O$_2$H$^{(-4,4)}$** is unstable, resulting in O$_2$ release during geometry optimization. It is also worthy to note that O$_2$ does not bind to the Fe^{II} atom in both spin states (geometry optimizations of **6-O$_2$H$^{(-4,2)}$** and **6-O$_2$H$^{(-4,4)}$** lead to dissociation of O$_2$).

Scheme 4.5 Possible reaction pathways for binding of O$_2$ to Ni-C. Energy differences are in kcal/mol.



(b) Isomerisation and partial reduction of O$_2$. As shown in Scheme 4.4, the O$_2$ ligand of **1-O$_2$H$^{(-4,2)}$** can abstract the hydrogen atom of the bridging hydride, with the concomitant formation of a terminal (OOH) group coordinated to Ni (species **2-O$_2$H$^{(-4,2)}$**, corresponding to **DH-2** in section 4.2), followed by the migration of the hydro-oxygenic ligand to a bridged position between the two metal atoms, yielding the **3-O$_2$H$^{(-4,2)}$** intermediate (**BH** in section 4.2). These two processes are exoenergetic by 16.5 and 3.5 kcal/mol, respectively. The activation energy equal to +13.2 kcal/mol for the migration of the hydrogen atom to O$_2$ supports the plausibility of such pathway. In the corresponding transition state, the O$_a$-O$_b$ distance (1.32 Å) is about 0.04 Å longer than in the **1-O$_2$H$^{(-4,2)}$** reagent, while the hydrogen atom moves to an asymmetric bridging position between Fe and O$_b$, in which it is closer to O$_b$ than to Fe (H−Ob = 1.64 Å; H−Fe = 1.75 Å). The formation of an H-bond between the hydrogen atom of the (OOH) group and the guanidinium group of Arg487 in **2-O$_2$H$^{(-4,2)}$** contributes to the stabilization of this intermediate. Notably, the O$_a$-O$_b$ bond length in **2-O$_2$H$^{(-4,2)}$** increases up to 1.43 Å, a value compatible with an hydroperoxide species. Analysis of charge and spin populations (qO_2H=-0.34, spin O_2H=0.42) corroborates this assignment with the O_2H ligand described as hydroperoxide and the Ni atom as Ni(III). The Ni atom in **2-$O_2H^{(-4,2)}$** features a pseudo square pyramidal coordination in which the S atom of Cys61 occupies the apical position. In **3-$O_2H^{(-4,2)}$**, the O_2H ligand moves to an asymmetric bridging position between Ni and Fe, in which the O_a atom is closer to Ni than Fe (O_a-Ni= 1.90 Å; O_a-Fe= 2.03 Å). The hydrogen atom of the OOH ligand forms an H-bond with the CN ion coordinated to the Fe atom. The O_a-O_b and Ni-Fe bond distances are equal to 1.48 and 2.92 Å, respectively.

In the proposed mechanism, **3-$O_2H^{(-4,2)}$** further evolves by dissociating the O-OH bond in a more stable intermediate characterized by a terminal hydroxide coordinated to Ni (O_a-Ni = 1.90 Å) and an oxygen atom bridging the two metal ions (**4-$O_2H^{(-4,2)}$**, corresponding to **EH-2** in section 4.2). In **4-$O_2H^{(-4,2)}$** the Ni atom has a pseudo octahedral coordination, with the bridging oxygen atom being slightly closer to Fe than Ni (O_b-Fe= 1.82 Å; O_b-Ni= 1.85 Å) and the O_a-O_b distance equal to 2.65 Å. The **4-$O_2H^{(-4,2)}$** species is more stable than **3-$O_2H^{(-4,2)}$** by as much as 21.7 kcal/mol. The formation of H-bonds with the protein environment should contribute to the larger stability of **4-$O_2H^{(-4,2)}$**; in particular the hydrogen and oxygen atoms of the hydroxide ligand forms an H-bond with the guanidinium group of Arg487 and the amide group of the Val63, respectively.

First monoelectronic reduction of the oxygenic ligand(s) may occur at all of the intermediates discussed above. In order to identify the step for which such process is more favourable, energy differences between reduced and oxidized species (E_{red}-E_{ox}) have been calculated. Since they are expected to follow the same trend as the experimental redox potentials, the E_{red}-E_{ox} values provide an estimate of the propensity of a oxidized species to be reduced; the more negative the E_{red}-E_{ox} value (the more positive the potential), the greater its tendency to be reduced. Notably, energy differences show that O_2 binding promotes the reduction of the active site since the E_{red}-E_{ox} value of **1-$O_2H^{(-4,2)}$** is about 0.2 V lower than that of **1-$H^{(-4,2)}$**. A further significant decrease of E_{red}-E_{ox}, of about 0.3 V, is observed after the O_2 cleavage in **4-$O_2H^{(-4,2)}$**. These results suggest that, depending on the reducing conditions, monoeletronic reduction of the active site can occur after the O_2 binding or the O_2 cleavage leading, respectively, to the formation of **1-$O_2H^{(-5,3)}$** and **4-$O_2H^{(-5,1)}$** (see Scheme 4.4). The **1-$O_2H^{(-5,3)}$** species is characterized by the O_2 ligand still terminally coordinated to Ni and a bridging H atom. For this intermediate the triplet spin state is more stable than the singlet spin state by 4.0 kcal/mol. Conversely in the **4-$O_2H^{(-5,1)}$** species, featuring a terminal hydroxide and a bridging oxygen atom, the singlet spin state is more stable than the triplet one by 4.8 kcal/mol.

(c) Subsequent reduction of the oxygenic ligand(s) and formation of Ni-B. Reaction from **1-$O_2H^{(-5,3)}$** or **4-$O_2H^{(-5,1)}$** can proceed following two different mechanisms to give Ni-B. In both cases, since electron and proton transfers are commonly coupled in enzymatic redox mechanisms, first monoelectronic reduction was proposed to be followed by the protonation of the oxygenic ligands (Scheme 4.4). For the protonation of the O_2 ligand in **1-$O_2H^{(-5,3)}$** two different pathways were considered: (A) intramolecular hydrogen transfer from the bridging hydride to the O_2 ligand (Path A in Scheme 4.4); (B) proton transfer from Glu14 residue, modelled as glutamic acid to the O_2 ligand (Path B in Scheme 4.4), as this residue was already proposed to be involved in proton transfer from/to the bimetallic cluster of the active site [85, 205, 206]. According to path A, the bridging hydrogen atom of **1-$O_2H^{(-5,3)}$** moves to the O_b atom of the O_2 ligand, yielding the **2-$O_2H^{(-5,1)}$** intermediate (-18.3 kcal/mol). This step is characterized by an energy barrier of +15.8 kcal/mol, a value very similar to that computed for the same process in the oxidized state (**1-$O_2H^{(-4,2)}$** → **1-$O_2H^{(-4,2)}$**). Therefore, it is not possible to discriminate among which redox state more favour the hydrogen transfer through the comparison of the energy barriers. **2-$O_2H^{(-5,1)}$** can then evolve to the **4-$O_2H^{(-5,1)}$** species discussed above in a few exoenergetic steps featuring the same (reduced) intermediates considered for the oxidized species. Protonation of **4-$O_2H^{(-5,1)}$** preferentially yields the **3-$O_2H_2^{(-4,1)}$** species, containing one terminal and one bridged hydroxide.

On the other hand, in path B, the protonation of the O_b atom in **1-$O_2H^{(-5,3)}$** yields the **1-$O_2H_2^{(-4,1)}$** species, characterized by a terminal (OOH) group coordinated to Ni and a bridged hydrogen atom, that can evolve by a very simple mechanism to Ni-B (proton transfer from Glu14 to the O_a atom coordinated to Ni is energetically disfavoured and leads to the dissociation of the oxygenic OOH ligand). In the subsequent step, the bridging hydrogen atom can be abstracted by O_b (transfer to O_a is disfavoured as the intermediate product **2-$O_2H_2^{(-4,1)}$** is less stable than **4-$O_2H_2^{(-4,3)}$** by 16.6 kcal/mol) leading to the formation of the O-OH_2 group, from which during geometry optimization a water molecule dissociates and moves away from the active site, leaving the O_a atom terminally coordinated to Ni (**4-$O_2H_2^{(-4,3)}$**). Then the O_a atom moves to the bridging position (**5-$O_2H_2^{(-4,3)}$**). Both reaction steps are exoenergetic by 26.0 and 21.1 kcal/mol, respectively.

As shown in Scheme 4.4, formation of the Ni-B state (**1-$OH^{(-4,2)}$**) from the potential intermediates **3-$O_2H_2^{(-4,1)}$** and **5-$O_2H_2^{(-4,3)}$** requires further reduction and protonation.

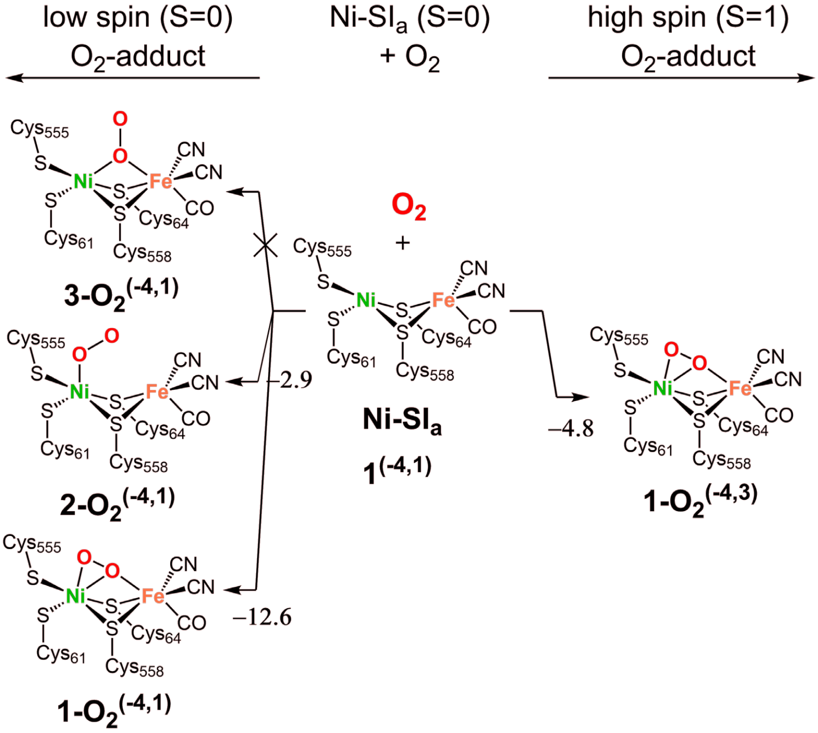
In this scenario, upon O_2 exposure, Ni-B is preferentially formed from Ni-C if the reaction conditions are compatible with the transfer of two electrons from external reducing agents to the active site. If oxidizing conditions do not allow the active site reduction, the two further electrons required for the full O_2 reduction can be supplied by the sulphur atom of Cys64, which oxidized to bridging sulfenate [63], leading to the formation of the Ni-A state. According our calculations, Ni-C can be converted into Ni-A (**5-$O_2H^{(-4,2)}$**, corresponding to **FH-Cys64O_b** in section 4.2) trough only isomerisation reactions; once **4-$O_2H^{(-4,2)}$** is formed, it can evolve to Ni-A (-20.4 kcal/mol), by insertion of the O_a atom in the Ni-S(Cys64) bond (Scheme 4.4, red background). Activation energy of 10.2 kcal/mol, computed for this isomerisation reaction, support the plausibility of the mechanism proposed for the Ni-A formation starting from the Ni-C state. In the corresponding transition state the Ni-O_a and O_a-S(Cys64) bond lenghts are 2.39 and 1.73 Å, respectively. Conversely the O_a-H and O_b-H distances are respectively equal to 1.01 and 1.95 Å.

4.3.2 Aerobic oxidation of Ni-SI_a

Geometry optimization of the $Ni^{II}Fe^{II}$ Ni-SI_a model, either in the low (S=0) or high (S=1) spin states, leave the Ni atom in a peculiar seesaw coordination geometry, that was already demonstrated to be a crucial factor to establish a favourable energetic profile and a convenient regiochemistry for H_2 binding to the NiFe metal cluster [71]. The energy difference was computed to be about 6 kcal/mol in favour of the low spin state [221].

An extensive search of the coordination modes of O_2 to the NiFe metal cluster in the low spin state of Ni-SI$_a$ (**1**$^{(-4,1)}$) has revealed two stable O_2-adducts, one in which O_2 bridges the two metal ions (**1-O**$_2^{(-4,1)}$) with a side-on coordination to Ni, and one in which O_2 is terminally coordinated to the Ni atom (**2-O**$_2^{(-4,1)}$) (see Scheme 4.6). The O_2 binding energies are equal to -12.6 and -2.9 kcal/mol for the two modes, respectively. Therefore, binding of O_2 in **1-O**$_2^{(-4,1)}$ is favoured by as much as 10 kcal/mol with respect to the terminal binding in **2-O**$_2^{(-4,1)}$. It is interesting to note that, for the bridging isomer, the O_2–adduct formation is a process energetically favoured also in the high spin state. However, the triplet spin state (**1-O**$_2^{(-4,3)}$) is less stable than the singlet one by about 8 kcal/mol (Scheme 4.6).

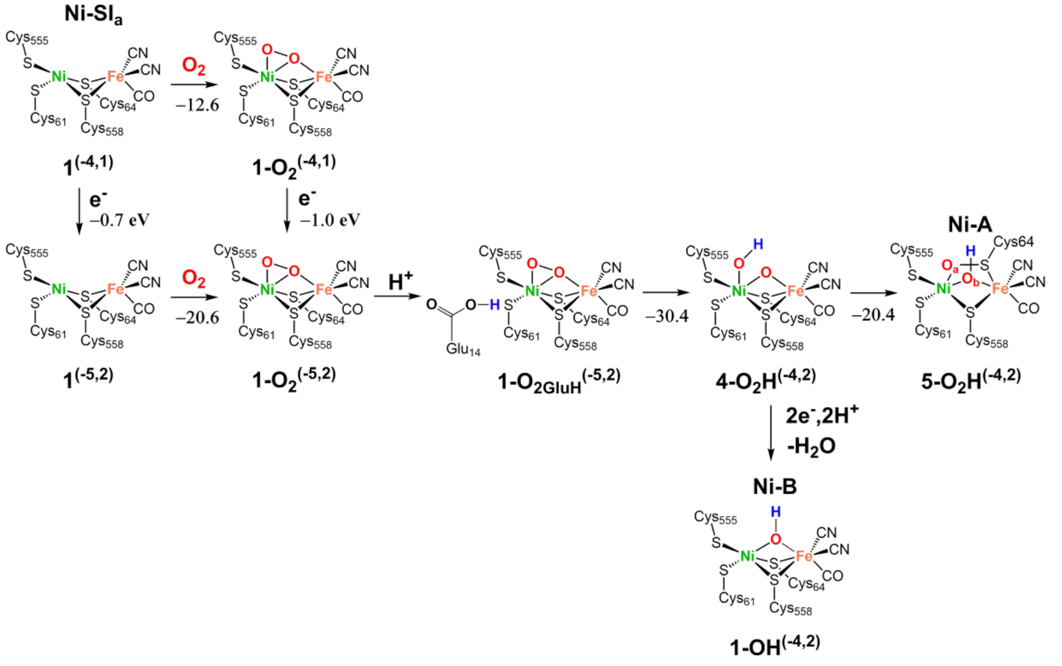
Scheme 4.6 Schematic representation of the possible reaction pathways for the O_2 binding to the Ni-Si$_a$ state of the active site of [NiFe]-hydrogenases. Energy differences are in kcal/mol.



As observed in the oxidation of the Ni-C state, binding of O_2 to the bimetallic cluster favours an electronic transfer to the active site as the E_{red}-E_{ox} value of the O_2-adduct is about 0.3 V lower than that of Ni-SI$_a$ (see Scheme 4.7). Subsequent proton transfer from the Glu14 residue to the oxygen atom terminally coordinated to Ni, promotes the bond cleavage of O_2 and leads to the formation of the **4-O**$_2$**H**$^{(-4,2)}$ species, already described in the oxidation mechanism starting from Ni-C. This process is favoured by the formation of strong H-bonds with Val67 and Arg487. Interestingly, the O_2 cleavage immediately after the O_2 binding is instead energetically disfavoured by 17.7 kcal/mol.

The **4-$O_2H^{(-4,2)}$** intermediate can evolve to Ni-A or Ni-B (Scheme 4.7), according to the same mechanisms proposed for the aerobic oxidation of the Ni-C state (see Scheme 4.4). However, it should be noted that if O_2 attacks the active site at the level of the Ni-SI$_a$ state its full four electron reduction requires more external electrons than those required if O_2 attacks the Ni-C state. Since Ni-SI$_a$ is one electron more oxidized than Ni-C, O_2 reduction at the active site of the former requires one more electron from the FeS relay.

Scheme 4.7 Schematic representation of the most plausible aerobic pathway for Ni-A and Ni-B formation starting from the Ni-SI$_a$ state. Energy differences are in kcal/mol while E_{red}-E_{ox} values are in eV.



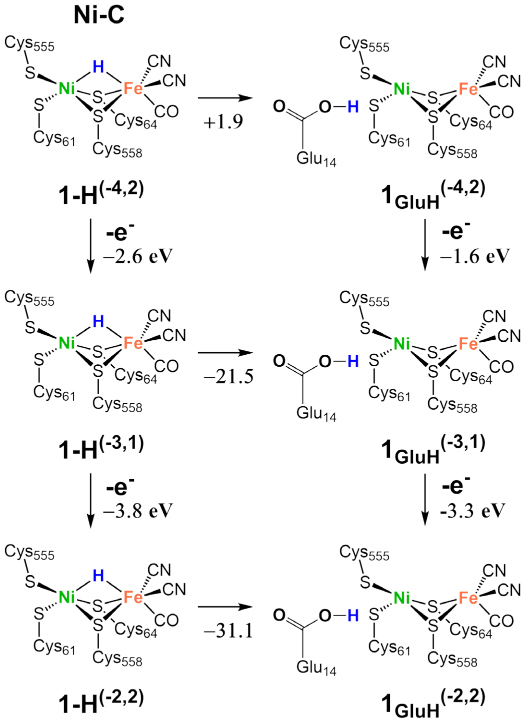
Starting from Ni-SI$_a$, the formation of Ni-A requires one external electron, since further three electrons are supplied by the active site; one electron is provided by the oxidation of Ni^{II} to Ni^{III}, while two electrons are provided by the oxidation of the S atom of Cys64 which in the Ni-A state is oxidized to the formal S^0 redox state. Conversely, three external electrons are required to ensure that Ni-B is formed, avoiding oxidation of cysteine residues and formation of Ni-A. However, fast delivery of three electrons to the active site is critical. In fact, when the enzyme in the Ni-SI$_a$ form is exposed to O_2 it converts preferentially to Ni-A [51].

4.3.3 Anaerobic oxidation

As already discussed, under anaerobic conditions the oxygen atoms included in Ni-B and Ni-A have been proposed to be provided by two water molecules [80]. According to this assumption, H_2O binding to the bimetallic cluster of [NiFe]-hydrogenases has been investigated as a first step of the anaerobic oxidation.

As for the O_2 binding, the H_2O binding to the Ni-C and the Ni-SI$_a$ forms should be considered. However, under the strong oxidizing conditions required for the anaerobic oxidation of the active site, the Ni-C state is oxidized to species for which the transfer of the bridging hydride to the carboxylate group of Glu14 is strongly favoured. In fact, whereas in the case of Ni-C (**1-H$^{(-4,2)}$**) this process is almost isoenergetic (+1.9 kcal/mol), in the case of the mono- (**1-H$^{(-3,1)}$**) and di-oxidized (**1-H$^{(-2,2)}$**) species of the Ni-C state it is exoenergetic, by as much as -21.5 and -31.1 kcal/mol, respectively (see Scheme 4.8).

Scheme 4.8 Proton transfer from the bridged position between the two metal atoms to the Glu14 residue. Energy differences are in kcal/mol while E$_{red}$-E$_{ox}$ values are in eV.



In addition, the computed E$_{red}$-E$_{ox}$ values show that proton migration to Glu14 promotes the oxidation of the active site, as the E$_{red}$-E$_{ox}$ value for **1$_{GluH}$$^{(-4,2)}$** and **1$_{GluH}$$^{(-3,1)}$** are 1.0 V and 0.5 V less negative than those of **1-H$^{(-4,2)}$** and **1-H$^{(-3,1)}$** (see Scheme 4.8). These results are compatible with a picture such that, under oxidizing conditions, the Ni-C state releases the bridging hydride as H$^+$ and it is oxidized, depending on the imposed potential, either to the Ni^{II}Fe^{II} (Ni-SI$_a$) form or to the Ni^{III}Fe^{II} redox state. Therefore, only the Ni-SI$_a$ state and its oxidized species have been considered as starting forms for the anaerobic oxidation of the [NiFe]-hydogenases.

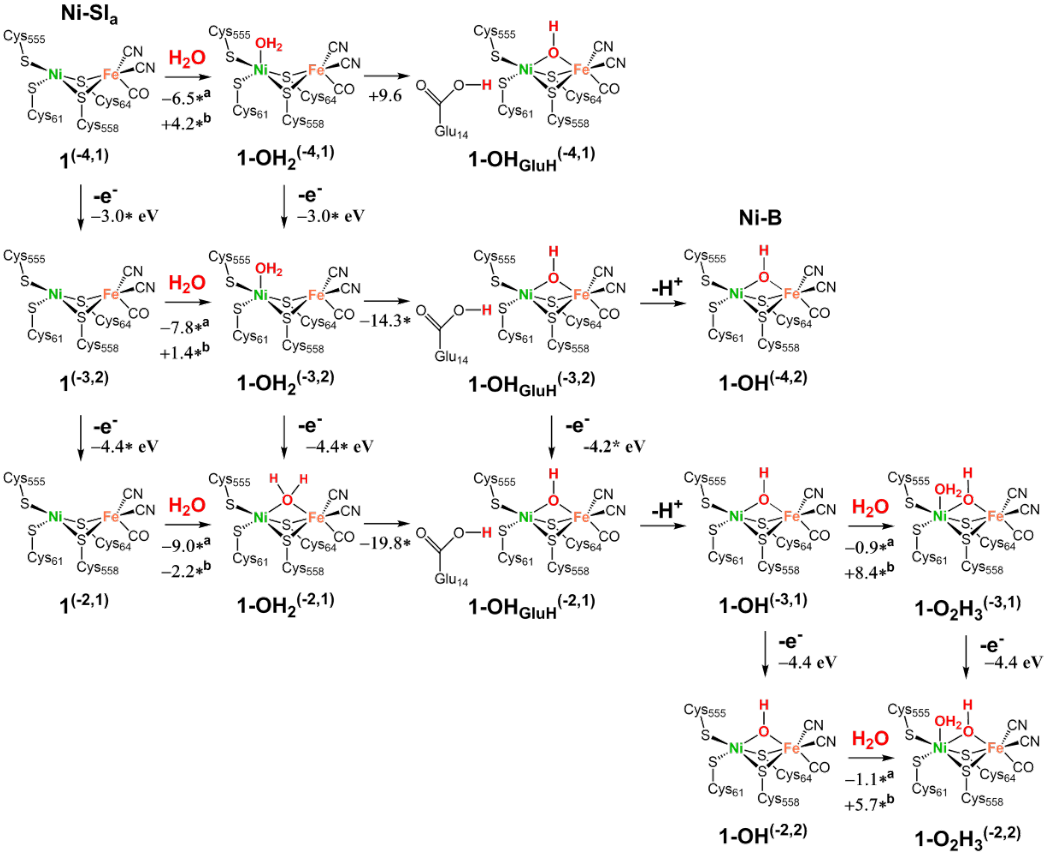
Before describing the mechanism of anaerobic oxidation proposed on the basis of our calculations, it should be specified that the geometry optimization of several potential intermediates has been carried out on models of the active site in which Asp553 is modelled as protonated. The application of this expedient was necessary to avoid unrealistic oxidation of such residue featuring a negatively charged carboxylic group close to the boundary of the model. Energy differences calculated using this expedient are marked with an asterisk in the main text and in the corresponding figures.

Oxidation to Ni-B. The most plausible pathway for formation of Ni-B (Ni^{III}-OH-Fe^{II}) starting from the Ni-SI$_a$ state ($Ni^{II}Fe^{II}$) involves the coordination of H_2O to the active site, one-electron oxidation and deprotonation of such water molecule. Such steps are schematically shown in Scheme 4.9. Since H_2O binding can occur either to the Ni-SI$_a$ form or its one-electron oxidized state (**1**$^{(-4,1)}$ and **1**$^{(-3,2)}$, respectively), both possibilities have been been investigated. Two coordination modes of H_2O have been identified for the two forms of the enzyme; one in which H_2O is terminally coordinated to Ni, and one in which H_2O bridges the two metal ions. However, in both redox states, the former isomer is considerably more stable than the latter. Interestingly, H_2O binding is energetically favoured, with a binding energy of -6.5* and -7.8* kcal/mol for **1-OH$_2$**$^{(-4,1)}$ and **1-OH$_2$**$^{(-3,2)}$, respectively, when considering as a reactant a H_2O molecule in the bulk solvent. On the other hand, H_2O binding to the Ni atom is less favourable when as a reactant is considered a H_2O molecule hydrogen bonded to Cys61, Cys64 and His68, as in this case the binding energies are +4.2* and +1.4* kcal/mol for **1-OH$_2$**$^{(-4,1)}$ and **1-OH$_2$**$^{(-3,2)}$, respectively. These values suggest that an equilibrium condition is established between the H_2O molecule H-bonded to Cys61, Cys64 and His68 and the H_2O coordinated to Ni, which is shifted to the product after one-electron oxidation of the active site. Such equilibrium between unbound and H_2O-coordinated forms in the enzyme may explain the very slow formation of the Ni-B state from the anaerobic oxidation of Ni-SI$_a$ [54]. [The presence of a crystallization water molecule very close to Cys61, Cys64 and His68 in the X-ray 3MYR structure [79], used to build the DFT models, supports these considerations.

The above results suggest that H_2O binds to the Ni atom of the active site after mono-electron oxidation of the Ni-SI$_a$ state. However, H_2O binding to Ni-SI$_a$ cannot be ruled out by the analysis of the redox properties of the two states. Indeed, the E_{red}-E_{ox} values calculated for **1**$^{(-4,1)}$ and **1-OH$_2$**$^{(-4,1)}$ are very similar and cannot be used to discriminate between water binding or oxidation as the first step of the anaerobic inactivation. Conversely, H_2O deprotonation and migration of the resulting terminal hydroxide to the bridging position to form the Ni-B state takes place necessarily after oxidation of the metallic site. Indeed, the conversion of **1-OH$_2$**$^{(-3,2)}$ to **1-OH$_{GluH}$**$^{(-3,2)}$ is energetically favoured by 14.3* kcal/mol, whereas the release of a proton from **1-OH$_2$**$^{(-4,1)}$ to give **1-OH$_{GluH}$**$^{(-4,1)}$ is endoenergetic by 9.6 kcal/mol.

Oxidation to Ni-A. Binding of an additional water molecule, two-electrons oxidation and two deprotonation steps should further be required to generate the Ni-A state from Ni-B. In this respect, it is worth noting that the second water molecule cannot bind the active site in the Ni-B state; the H_2O-adduct of the Ni-B form is unstable and, during geometry optimization, H_2O dissociates from the active site. Therefore, a second mono-electron oxidation of the cluster may be required for the binding of the second water molecule. This oxidation can occur before or after the binding of the first H_2O as indicated by the comparable E_{red}-E_{ox} values of **1**$^{(-3,2)}$, **1-OH**$_2^{(-3,2)}$ and **1-OH**$_{GluH}^{(-3,2)}$ (see Scheme 4.9).

Scheme 4.9 Mechanism of the Ni-Si$_a$/Ni-B conversion under anaerobic conditions, involving the following steps: (1) one-electron oxidation; (2) hydration and (3) deprotonation. Second oxidation of the active site and binding of second H_2O are also shown. Asterisked energy differences are calculated using the LM model in which Asp553 is modelled as protonated. Superscripts a and b denote binding energies to the metallic site of a water molecule in the bulk solvent or hydrogen bonded to Cys61, Cys64 and His68, respectively. Energy differences are in kcal/mol, while E_{red}-E_{ox} values are in eV.

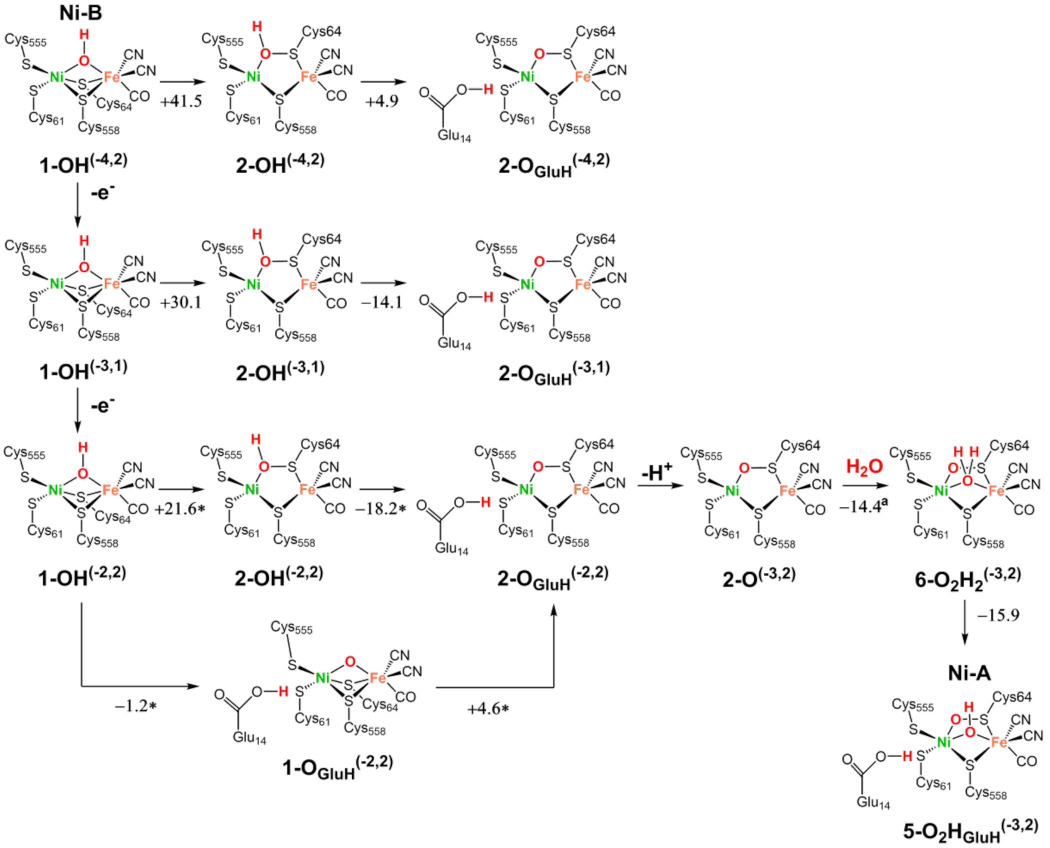


Assuming that the anaerobic formation of Ni-A takes place by a mechanism not involving Ni-B as intermediate, as suggested by Fontecilla-Camps and co-workers [63], the second oxidation of the cluster should occur before the binding of the first H_2O molecule leading to the **1**$^{(-2,1)}$ species shown in Scheme 4.9. The increase of the NiA-to-NiB ratio under more oxidizing conditions [91] supports the plausibility of the two-electron oxidation of the Ni-SI$_a$ state. According to this hypothesis, H_2O binding to **1**$^{(-2,1)}$ is energetically favoured both considering as reactant a H_2O molecule in the bulk solvent and a H_2O molecule hydrogen bonded to Cys61, Cys64 and His68, as the corresponding binding energies are -9.0* and -2.2* kcal/mol, respectively. Interestingly, in this form a particularly strong H-bond between Arg487 and the H_2O ligand favours the formation of the bridging H_2O-adduct (**1-OH**$_2^{(-2,1)}$) with respect to the terminal isomer. The deprotonation of this water molecule and the transfer of the proton to the Glu14 residue to give the **1-OH$_{GluH}$$^{(-2,1)}$** species is exoenergetic by about 19.8* kcal/mol.

The second water molecule can bind the Ni atom of **1-OH$^{(-3,1)}$** leading to **1-O$_2$H$_3$$^{(-3,1)}$**. However, the binding of this water molecule from the bulk solvent is almost isoenergetic (-0.9* kcal/mol) and it is even disfavoured by 8.4* kcal/mol considering the H_2O molecule H-bonded to Cys61, Cys64 and His68 as reactant. H_2O binding is slightly more favoured by a further oxidation of the cluster, but not enough to make this process exoenergetic (see Scheme 4.9). In this case, the binding energies to **1-OH$^{(-2,2)}$** of a water molecule from the bulk solvent and the H_2O nearby the active site are -1.1* and +5.7* kcal/mol, respectively.

These results suggest that the binding of a second water molecule to a form of the active site in which an hydroxide occupies the bridging position between the Ni and Fe atoms is slightly disfavoured. On the basis of these results, we considered plausible pathways in which the first H_2O molecule provides the oxygen atom of the sulfenate group of Ni-A, leaving a vacant position between Ni and Fe for binding of second water molecule (see Scheme 4.10).

Scheme 4.10 Schematic representation of plausible steps that lead to the formation of Ni-A under anaerobic conditions. Energy differences are in kcal/mol while E_{red}-E_{ox} values are in eV.



The oxidation of the active site to the bridging sulfenate may occur through the insertion into the Ni-S(Cys64) bond of the bridging hydroxide (see Scheme 4.10). However, the migration of this hydroxide is strongly endoenergetic for all redox states investigated. The oxidation level of the active site make the process less unfavourable as shown by the decrease of the reaction energies from +41.5 kcal/mol for the formation of **2-OH**$^{(-4,2)}$ to the value of +21.6* for the formation of **2-OH**$^{(-2,2)}$. On the other hand, the subsequent proton transfer from the hydroxide to Glu14 is energetically favoured for **2-OH**$^{(-3,1)}$ and **2-OH**$^{(-2,2)}$ by about -14 and -18 kcal/mol, respectively, leading to the species **2-O**$_{\mathbf{GluH}}^{(-3,1)}$ and **2-O**$_{\mathbf{GluH}}^{(-2,2)}$, which however are still less stable than the corresponding bridging hydroxide **1-OH**$^{(-3,1)}$ and **1-OH**$^{(-2,2)}$ reactants. In particular, for the more oxidized species investigated the insertion of the bridging hydroxide into the Ni-S(Cys64) bond and the subsequent proton transfer to the Glu14 residue is an endoenergetic process by only about 3.4 kcal/mol. The same reaction is significantly more favourable by considering first the proton transfer from the bridging hydroxide followed by the oxygen insertion into the Ni-S(Cys64) bond (see Scheme 4.10). In this case the first step to give **1-O**$_{\mathbf{GluH}}^{(-2,2)}$ is exonergetic by -1.2 kcal/mol, whereas the second step to give **2-O**$_{\mathbf{GluH}}^{(-22)}$ is endoenergetic by +4.6 kcal/mol. Once **2-O**$_{\mathbf{GluH}}^{(-2,2)}$ is formed, it can evolve to Ni-A by the energetically favoured binding of a second water molecule and the release of a proton from the active site as shown in Scheme 4.10.

It is worth noting that our calculations do not support the mechanism for the Ni-A anaerobic formation proposed by Fontecilla-Camps and co-workers [63], according to which the insertion of the bridging oxygenic ligand into the Ni-S(Cys64) bond should occur at the Ni^{III}-OH level. This isomerisation reaction is predicted to be endoenergetic by as much as 41.5 kcal/mol in our calculations. An additional oxidation of the cluster is required to make this process feasible.

4.3.4 Discussion

Upon O_2 exposure the enzyme is reversibly converted into oxidized inactive states. According to our proposed mechanism, O_2 reacts directly at the Ni-Fe site. The energetically favoured O_2 binding to the Ni atom leads to the formation of intermediates that can evolve in different ways depending on the number of reducing equivalents available for the process. If the enzyme in the initial redox state contains four reducing equivalents, it can convert to Ni-B; otherwise, Ni-A is formed. Starting from the Ni-C state, two of the four electrons required for the full O_2 reduction can be directly supplied by the bridging hydride, whereas further two reducing equivalents can be delivered by the electron relay of the enzyme, that in such reducing condition is proposed to contain two or three reduced Fe-S clusters [222]. Therefore, the mechanism proposed in this thesis explain the predominant formation of Ni-B under the reducing conditions in which Ni-C is the prevalent form of the enzyme. The formation of small amounts of Ni-A from Ni-C can also be satisfactorily explained by our calculations; if two reducing equivalents are not available to the active site for the full O_2 reduction by external reducing agents these two electrons can be supplied by the sulphur atom of Cys64, and Ni-C, exclusively through isomerisation reactions, converts into Ni-A.

According to the experimental data, Ni-A forms in larger amount by supplying O_2 to the enzyme in more oxidized conditions such as those in which Ni-SI$_a$ is the prevalent form [222]. Because Ni-B requires three reducing equivalent to be formed from the Ni-SI$_a$ state, whereas the enzyme can only provide two reducing equivalents (one localized on Ni^{2+} and the other in the reduced medial cluster) in such state, Ni-B should not be formed. Conversely, two reducing equivalents are sufficient for the conversion of Ni-SI$_a$ into Ni-A.

Therefore, according to our proposed mechanism for the aerobic oxidation of the [NiFe]-hydrogenases and the previous experimental results [51, 61], electron-availability of the enzyme controls the Ni-A to Ni-B ratio (see Scheme 4.4); ‘electron-poor’ environments favour the Ni- A state, whereas Ni-B is favoured under ‘electron-rich’ conditions. The dependence of the Ni-B formation on the number of reducing equivalents available for the process is supported by the fact that all O_2-tolerant [NiFe]-hydrogenases, which are characterized by a modified proximal [Fe_4S_3] cluster [223–225], generate only the Ni-B state in the reaction with O_2; this modified [Fe_4S_3] cluster can deliver two electrons at close potentials (see Figure 4.12) and allow to reverse the normal electrons flow through the electron transfer chain during O_2 attack enabling its complete reduction and avoiding the Ni-A formation [64, 65].

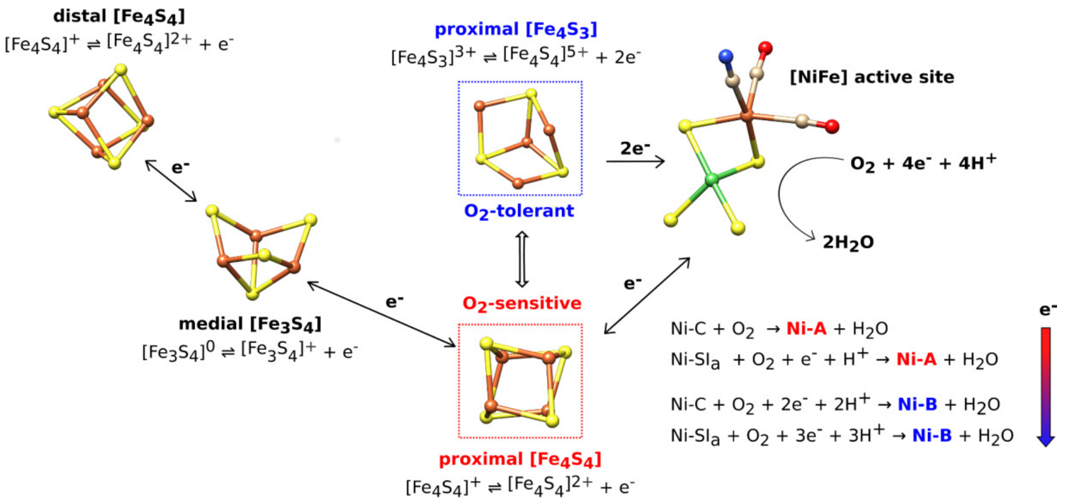


Figure 4.12 The minimal functional unit of a O_2-sensitive NiFe-hydrogenase (pdb code: 1WUL) consisting of a [NiFe] active site and three FeS clusters that transfer electrons to and from the catalytic site. During O_2 attack, FeS clusters can be oxidised to provide one or more electrons required for full O_2 reduction. According our proposed mechanism, the Ni-A to Ni-B ratio depend on the number of reducing equivalents available for the process. In O_2-tolerant NiFe-hydrogenase, the proximal [Fe_4S_4] cluster is replaced by a [Fe_4S_3] cluster (pdb code: 3RGW), that is able to release two electrons allowing selective Ni-B formation.

That O_2 reaches the active site and then binds at Ni-Fe site is supported by the experimental observation that slowing down the rate of diffusion along the gas channel connecting the enzyme surface with the active site decreases the rate of O_2 inhibition [226]; moreover, an EPR study carried out using labelled O_2 suggested that an oxygenic species originating from

O_2 is bounded close to nickel in both the ready and unready enzyme forms [227]. However, O_2 binding to the active site is not the only mechanism proposed for the enzyme aerobic oxidation. According to EPR and FTIR experiments [91], a second possibility is that the attacking O_2 is not incorporated as an active site ligand but acts as an electron acceptor in the vicinity of the active site. Evidence for such mechanism has been obtained by the reoxidation of H_2-reduced *D. gigas* enzyme in $H_2{}^{17}O$ with $^{16}O_2$, where ^{17}O binding to nickel in the Ni-A state was observed by means of ENDOR spectroscopy [80]. According to our calculations, O_2 binding to the active site is energetically favored. However, this result doesn't rule out the possibility that the O_2 molecule acts as an external oxidizing agent and the oxygen species bound to the NiFe site in the inactive states originates from water molecules of the bulk solvent. In this latter case the reaction pathway can follow the mechanism proposed for the anaerobic oxidation, in which the O_2 molecule acts only as an external electron acceptor. According to this hypothesis, water molecules provide the oxygen atoms contained in the Ni-B and the Ni-A states.

The Ni-A-to-Ni-B ratio in the anaerobic oxidation is also strongly dependent on the oxidizing conditions, such as the potential of the electrode or the used oxidizing agent [91]. On the basis of our results, Ni-A is formed in oxidizing conditions such that a fast two-electron oxidation of the Ni-SI_a state can take place; otherwise, equilibrium between the one-electron oxidized state of Ni-SI_a and its H_2O adduct lead to the Ni-B state, whose anaerobic formation is known to be very slow [54].

Two alternative pathways for the anaerobic formation of Ni-A are suggested by our results. In the first pathway, hydration and deprotonation of the two-electron oxidized state of Ni-SI_a lead to the formation of an intermediate in which the bridging position between the Ni and Fe atoms of the active site is occupied by an hydroxide ligand. The second water molecule binds to the Ni atom of the one-electron oxidized state of this species (see Scheme 4.9). Deprotonation of H_2O and insertion of the resulting oxygenic species in the Ni-S(Cys64) bond lead to Ni-A. Conversely, according to the second pathway, the first H_2O molecule provides the oxygen atom of the sulfenate group of Ni-A; the H_2O binding is followed by the release of two protons and the insertion of the oxygen atom into the Ni-S(Cys64) bond, leaving a vacant position between Ni and Fe for binding of the second water molecule (see Scheme 4.10). Since both pathways involve a slightly endoenergetic step with a similar reaction energy (+5.7* and +4.6* kcal/mol for first and second pathway, respectively), it is not possible to discriminate between the two proposed mechanism for anaerobic oxidation of the Ni-SI_a state.

Another topic related to the inactive states of the [NiFe]-hydrogenases is still unclear. Experimental studies suggested that the Ni-B and Ni-A states do not directly interconvert. This evidence is in accord with the proposed pathway for the aerobic oxidation; two different mechanisms have been proposed for the Ni-B and the Ni-A formation. Conversely, under anaerobic conditions, although the proposed mechanism for formation of Ni-A does not imply the formation of Ni-B as intermediate, it cannot be ruled out that under sufficiently strong oxidizing conditions, once formed, Ni-B may be oxidized to Ni-A. This consideration is supported by an ESR study of reoxidation of hydrogen-activated *D. gigas* hydrogenase with dichlorophenolindophenol (DCIP), which shows that after prolonged exposure to oxidizing conditions, the Ni-B signal decreased and that assigned to Ni-A increased, in parallel with the decrease in enzyme activity [72]. Since this phenomenon was not observed on reoxidation by O_2, this experiment support our proposed mechanism that in the absence of O_2, the Ni-B state can be oxidized to Ni-A under strong oxidizing conditions.

4.4 Reactivation mechanism of Ni-A and Ni-B

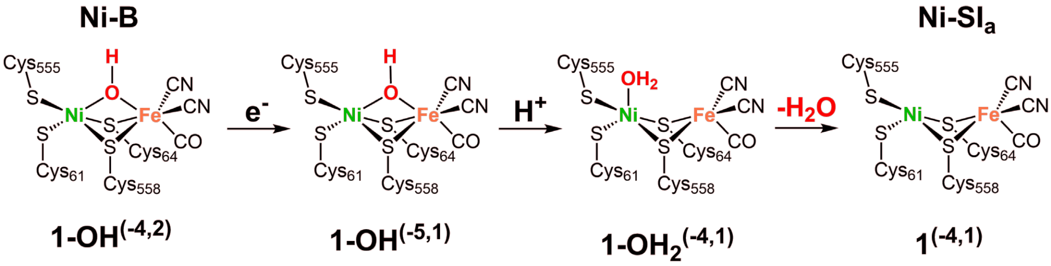
Oxidized [NiFe]-hydrogenases require reductive activation that involves a structural reorganization of the active site from a form stable in air to the form that is catalytically active for H_2-uptake. The difference between the Ni-A and Ni-B states was associated with a difference in the rate of this activation process; the Ni-B state is activated within seconds while the Ni-A state, at room temperatures, takes hours [72, 75, 227]. However, the mechanism of the reactivation of the oxidized [NiFe]-hydrogenases is still not clear. Prompted by these consideration we decided to investigate the mechanism of the activation mechanisms for the Ni-B and the Ni-A states in order to rationalise their different reactivation kinetics.

For both inactive states, the first step of the activation corresponds to the one-electron reduction of Ni^{III} to Ni^{II} as observed by EPR experiments [72, 228, 229]. The mono-reduced states of Ni-B and Ni-A, namely Ni-SI_r and Ni-SU respectively, are both EPR-silent and catalytically inactive but they can be distinguished on the basis of the IR bands of their CO and CN ligands [76]. X-ray diffraction studies showed that the main structural difference at the level of the active site between the oxidized and the reduced active enzymes is that, in the latter, the oxygen species that bridges the metals is absent, and the Ni-Fe distance is ca. 0.25 Å shorter [68]. These results support the hypothesis that the mono-electronic reduction of the oxidized states promote the removal of the bridging oxygen species, allowing H_2 binding to the active site and catalytic turnover [68].

The Ni-B to Ni-Si_r conversion is a reaction involving a mono-electron reduction [228, 230]. Infrared studies suggest that there are two forms of the Ni-Si_r state that differ for their protonation state [231]. Reduction of Ni-B at 2 °C shifts the υ(CO) band from 1943 cm^{-1} to 1931 cm^{-1} (at pH 6) or 1910 cm^{-1} (at pH 9). At low pH, EXAFS studies indicated that the removal of the bridging oxygen species and the shortening of the Ni-Fe distance from 2.85 to 2.60 Å had already occurred [232]. Albracht *et al.* [82, 222] suggested that the bridging OH^- is still present in the Ni-SI_r state prepared at 2 °C and pH 9, but it becomes protonated at low pH leading to the formation of a water molecule. It was assumed that, at 2 °C, this H_2O molecule remains captured in the active-site pocket and sterically hinders the reaction with H_2, while at room temperature it is released allowing the enzyme activation.

These experimental observations suggested that the reactivation mechanism between the inactive Ni-B and the active Ni-SIa state occurs following the mechanism shown in Scheme 4.11. Mono-electronic reduction and protonation of the Ni-B state and subsequent dissociation of the water molecule from the active site yields the active Ni-SI_a state. The two forms of the Ni-Si_r state should correspond to the species **1-OH$^{(-5,1)}$**, in which a hydroxide ion is still coordinated to the Ni^{II} atom, and **1-OH$_2$$^{(-4,1)}$** in which, after protonation of **1-OH$^{(-5,1)}$**, H_2O is terminally coordinated to the Ni^{II} atom. Our calculations show that the reduction and the protonation of the Ni-B state to form the **1-OH$_2$$^{(-4,1)}$** species lead to the removal of the bridging ligand (Ni-O distance = 2.1 Å) and the shortening of the Ni-Fe distance from 2.9 to 2.6 Å, in agreement with EXAFS data [232]. The isomer in which H_2O bridges the two metal ions is less stable by 16.7 kcal/mol with respect to the terminal one. As discussed in the previous section H_2O dissociation and migration to a nearby position H-bonded to hydrogen bonded to Cys61, Cys64 and His68 is an almost isoenergetic process (-2.1 kcal/mol). This result is compatible with the experimentally observed easy interconversion between Ni-Si$_r$ and Ni-SI$_a$ at 25 °C [222].

Scheme 4.11 Schematic representation of the activation pathway of the Ni-B state, leading to the Ni-SI$_a$ state. Energy differences are in kcal/mol.



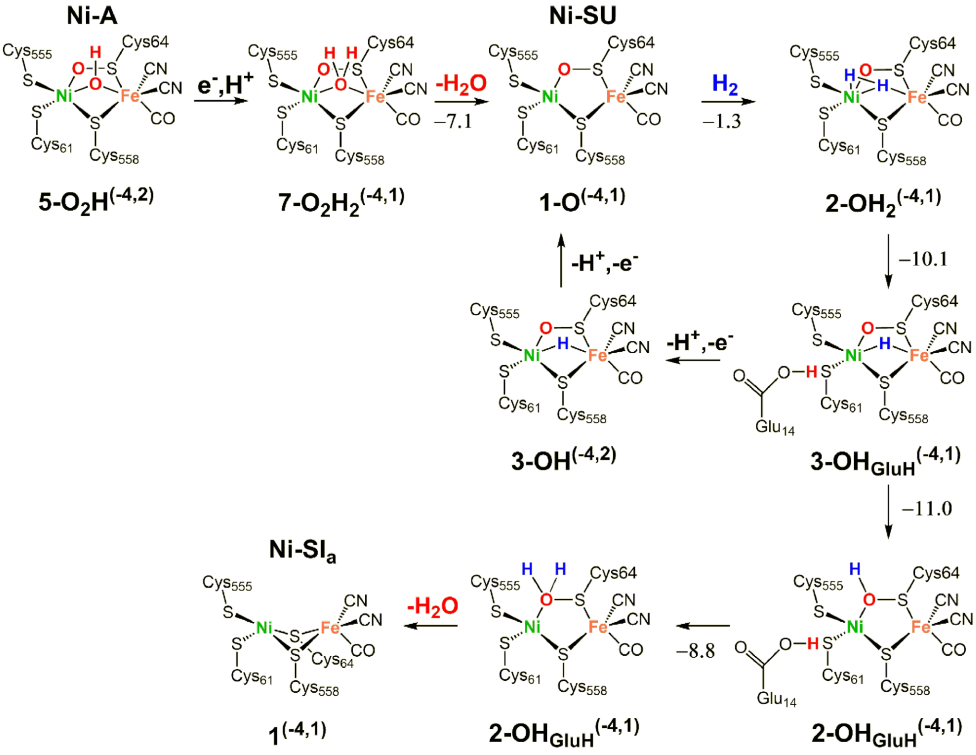
Since reductive activation processes can take place upon H_2 exposure, the reaction of H_2 with the active site of the Ni-B state was also investigated. Albracht *et al.* [82] suggested that the bridging hydroxyl group acts as the base for the heterolytic cleavage of the H_2 molecule on Ni. However, our DFT calculations showed that the adduct in which H_2 binds to the Ni^{III} atom of Ni-B is unstable, resulting in H_2 release, while H_2 binding to the Fe atom is energetically disfavoured by as much as 15.4 kcal/mol. On the other hand, the process represented by the reaction:

$$Ni - B + H_2 \rightleftharpoons Ni - C + H_2O \tag{4.1}$$

which corresponds to the H_2 cleavage and the subsequent dissociation of a water molecule to form the Ni-C state is exoenergetic by about 5 kcal/mol. This mechanism, however, can be definitively ruled out as it does not imply the formation of the diamagnetic Ni-Si$_r$ state, further supporting the activation mechanism proposed in Scheme 4.11.

In the case of the Ni-A state, the first step of the activation process should also correspond to a one-electron reduction [228, 230]. The pH dependence of the midpoint potential of the Ni-A/Ni-SU couple indicates that the mono-electron reduction is accompanied by a proton transfer. The subsequent removal of the bridging oxygen species was confirmed by the experiments based on ^{17}O labelling of Ni-A and its reductive activation [80]. The activation mechanism proposed for Ni-A is schematically represented in Scheme 4.12. The mono-electron reduction and protonation of the oxidized enzyme leads to the **7-O$_2$H$_2$$^{(-4,1)}$** species in which, analogously to **1-OH$_2^{(-4,1)}$** (mono-reduced and protonated state of Ni-B), a water molecule is still coordinated to the Ni^{II}Fe^{II} site. In this species the Ni-O and the Fe-O distances are 2.8 and 2.1 Å, respectively. The adduct in which the protonation occurred on the oxygen atom coordinated to the sulphur atom of Cys64 is less stable than **7-O$_2$H$_2^{(-4,1)}$** by 5.3 kcal/mol. Since dissociation and migration of the water molecule to form an H-bond network is an exoenergetic process (-7.1 kcal/mol), it is possible to suggest that the Ni-SU state corresponds to the **1-O$^{(-4,1)}$** species in which the H_2O has already left the active site.

Scheme 4.12 Schematic representation of the activation pathway of the Ni-A state to the Ni-SI$_a$ state. Energy differences are in kcal/mol.



As in the Ni-SI$_a$, the metal ions in this state are both characterized by one vacant coordination position, where H_2 might bind. However Ni-SU is a catalytically inactive state. Therefore, with the aim of explaining the different properties of Ni-SU and Ni-SI$_a$, dihydrogen binding and cleavage at the active site of the putative Ni-SU state were investigated. According to the catalytic mechanism previously described for [NiFe]-hydrogenases [71], the oxidative addition of H_2 on the Ni-SI$_a$ state is followed by the H_2 cleavage and proton transfer from Ni to the carboxylate group of Glu14. Subsequently, two electrons and two protons have to be transferred from the active site to reform the Ni-SI$_a$ state. Starting from Ni-SU the binding and activation processes may follow a similar pathway. Binding of H_2 to the Ni atom of the **1-O$^{(-4,1)}$** species is slightly exoenergetic (-1.3 kcal/mol), albeit not as much as the H_2 binding to Ni-SI$_a$ (-4.4 kcal/mol) [71]. The subsequent H_2 cleavage and proton transfer to Glu14 to give the **3-OH$_{GluH}^{(-4,1)}$** complex are exoenergetic processes (-10.1 kcal/mol). Conversely, the formation of the paramagnetic intermediate **$3\text{-OH}^{(-4,2)}$**, corresponding to the Ni-C state in the catalytic mechanism, is predicted to be disfavoured by the presence of the bidging sulfenate. This is suggested by the comparison of the HOMO energy of **$3\text{-OH}_{\text{GluH}}^{(-4,1)}$** (-0.96 eV) and that of the corresponding catalytic species Ni-R (-0.69 eV). This significant energy difference indicates that oxidation of **$3\text{-OH}_{\text{GluH}}^{(-4,1)}$** requires more positive potentials than the oxidation of the Ni-R species, suggesting that the presence of an O atom coordinated to Ni decreases makes much more difficult the electron transfer from the the active site preveting the formation of the **$3\text{-OH}^{(-4,2)}$** paramagnetic species. According to these data, EPR experiments have identified only three EPR-active states: Ni-A, Ni-B and Ni-C.

Differently from the Ni-B activation, a one-electron reduction of the active site, accompanied by single proton transfer, is not therefore sufficient to reactivate the Ni-A state; the reduction of the oxygenic species trapped between the Ni and the S atom of Cys64 is necessary to form an active form of the enzyme. Thus, the activation of the Ni-SU state requires further reducing equivalents that should be simultaneously available in order to avoid the formation of paramagnetic states, whose occurrence was actually not experimentally observed. These electrons can be provided by an H_2 molecule that binds the active site in the Ni-SU state (*vide supra*). According to this hypothesis, the cleavage of H_2 and the transfer of the two H atoms to the oxygen atom of the sulfenate group yields the **$3\text{-OH}_2^{(-4,1)}$** species, in which the S-O distance increases up to 2.8 Å and a water molecule is coordinated to the Ni atom. The detachment of this H_2O leads to the shortening of the Ni-S distance from 3.8 to 2.3 Å and the consequent formation of the active Ni-SI$_a$ state (see Scheme 4.12).

4.4.1 Conclusions

According to our proposed mechanism for [NiFe]-hydrogenases reactivation, the conversion of Ni-B to Ni-SI$_a$ is achieved by mono-electron reduction and protonation of the active site, which promote the removal of the bridging hydroxide. This process is predicted to be no sufficient to reactivate the Ni-A state; H_2 binding on the resulting Ni-SU state is required to form the active Ni-SI$_a$ state. The two electrons provided by the H_2 molecule reduce the sulfenate group and subsequent proton transfers allow the release of a second water molecule from the active site. The hypothesis that an H_2 molecule provides two of the three electrons required for Ni-A activation is supported by the fact that no paramagnetic species are formed during the activation process. Moreover, dynamic electrochemical kinetic studies showed that in the absence of H_2 only the reversible equilibrium between Ni-A and Ni-SU can be established [52], suggesting that H_2 is required to make this reductive transformation irreversible.

On the basis of our results, we can suggest that the slower activation of Ni-A compared to that of Ni-B is due to the conversion of Ni-SU to Ni-SI$_a$. This conclusion is supported by experimental data. FTIR-spectroelectrochemical and electrochemical studies of Ni-A reactivation [52, 54] showed that Ni-A can be converted to Ni-SU in a fast and reversible electrochemical step and that the rate-determining step, independent of the redox potential, is the conversion of Ni-SU to Ni-SI$_a$.

CHAPTER 5

REACTIVITY OF MO-CODHS TOWARDS DIHYDROGEN

Interaction of H_2 with the Mo-CODH enzyme, and the mechanism of the dihydrogen splitting, has been recently the object of a detailed experimental study by Wilcoxen and Hille [132]. These authors identified an H_2-bound paramagnetic form of the active site and characterized it in terms of copper and hydrogen hyperfine coupling constants. On the basis of these results, the incipient H_2 molecule was proposed to directly bind to the Cu ion, and not the Mo ion.

In this context, a theoretical investigation of the reactivity of the enzyme active site toward H_2 by using DFT calculations has been carried out, also with the aim of deepening insights into the nature of the H_2-bound paramagnetic enzyme form experimentally observed. Moreover, in order to rationalize the loss of H_2-oxidizing activity caused by the substitution of copper with silver in the protein active site [132], geometry optimizations of H_2-bound forms have been also performed for Ag-substituted models.

5.1 Computational details

Geometries were fully optimized at the BP86/def2-TZVP level [149, 150, 233], using the resolution of identity (RI) technique [163]; the chosen basis set was used in conjunction with the corresponding effective core potential in the case of Mo [233]. The energy difference values reported in the following were computed on the basis of total energies obtained at the BP86-RI/ def2-TZVP level. The TURBOMOLE suite of programs was used for all unconstrained energy minimizations and potential energy surface (PES) scans, which were performed in vacuo [171]; all unconstrained minimizations were followed by harmonic vibrational frequency calculations to confirm the minimum character of each of the stationary points localized.

It is noticeable that, in the enzyme, the covalent contact between the Cu/Mo-containing cofactor and the apoprotein is established through a single cysteinyl sulfur ligand, analogously to the case of the substrate-binding site in [FeFe] hydrogenases [40]; the setup here adopted for optimizations–based on an in vacuo, unconstrained DFT representation of the metal ions and the inner sphere ligands–has been successfully applied for the theoretical investigation of the PES associated with dihydrogen binding and splitting on models of the latter metalloenzyme [234].

Transition state structures were searched for according to a procedure based on a pseudo-Newton-Raphson method. At the initial stage, geometry optimization of a guessed transition state structure is carried out constraining the distance(s) corresponding to the reaction coordinate. Vibrational analysis, at the BP86/def2-TZVP level, of the constrained minimum energy structures is then carried out, and if one negative eigenmode corresponding to the reaction coordinate is found, the curvature determined at such point is used as a starting point in the transition state search. The location of the transition-state structure is carried out using an eigenvector-following search: the eigenvectors in the Hessian are sorted in ascending order, with the first one being that associated with the negative eigenvalue. After the first step, however, the search is performed by choosing the critical eigenvector with a maximum overlap criterion, which is based on the dot product with the eigenvector followed in the previous step. Finally, the analytical Hessian matrix is calculated to carry out the vibrational analysis of the stationary point.

The ORCA program [173] was used to perform DFT calculations of g factors and hyperfine coupling constants. Such computations were preceded by single point calculations at the BP86/DGDZVP level [235, 236], using the geometries optimized at the BP86/def2-TZVP level as described above.

5.2 Model composition and preliminary structural investigations

Calculations were carried out on a model of the Mo/Cu-dependent CO dehydrogenase active site, which includes the first shell ligands of the Cu and Mo ions: in particular, Cys388 side chain is represented in the form of a CH_3S^- fragment, whereas the pterin is replaced by an 1,2-ethenedithiolate ligand (see Figure 5.1, and in particular the schematic structure on the left side of this figure).

The atomic composition of the adducts considered in the following can differ with respect to such parent structure in terms of: (i) presence of substrates –CO or H_2–bound to Cu; (ii) protonation of one of the Mo-bound oxo ligands or of the S atom of the CH_3S^- fragment; (iii) detachment of the CH_3SH fragment from the Cu center.

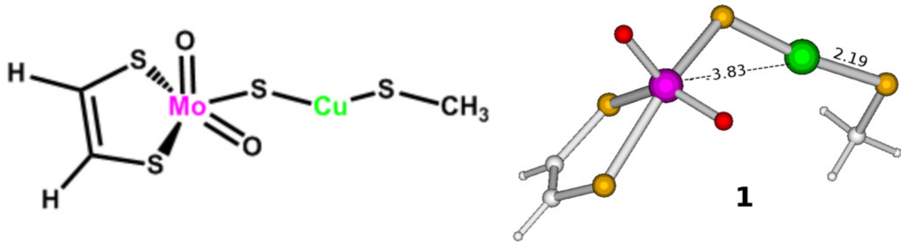


Figure 5.1 The parent complex **1** in a schematic representation (left) and its DFT-optimized form. All distances in Å and (in parentheses) bond angles in degree.

As a preliminary step, in the present study, we investigated the structural effects of CO binding to the enzyme active site: species **1** and **2** (see Figure 5.1 and Figure 5.2, respectively) are the geometry-optimized unbound and bound states of the active site, respectively, both attaining the Mo^{VI}/Cu^{I} redox state. CO binding leads to a Mo-μS-Cu bond angle reduction by ca. 10°; moreover, a shortening of the Cu-Mo distance from 3.83 to 3.68 Å is observed, an outcome in line with previous theoretical and experimental results which indicate that the complex within the protein is in general rather flexible, and that the Cu–Mo distance in particular is not energetically very relevant (notably, elongations of the Cu–Mo distance up to 6.1 Å have been discussed in the previous literature) [126]. CO binding turned out to be favored by 15.2 kcal/mol, and the Cu-C distance was computed to be 1.81 Å long; finally, the Cu-S(Cys) bond length in adduct **2** is 2.26 Å long (a value to be compared with the 2.19 Å distance for the same bond in **1**, see Figure 5.1).

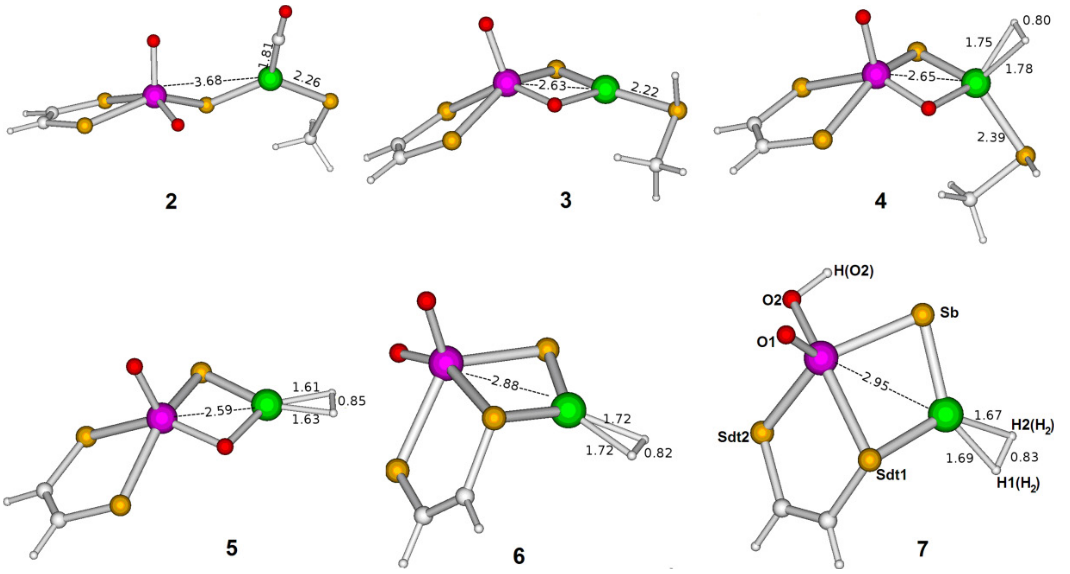


Figure 5.2 Optimized structures of $Mo^{VI}Cu^{I}$ adducts **2**, **3**, **4**, **5** and **6**, and of the $Mo^{V}Cu^{I}$ species **7** whose calculated EPR parameters are reported in the main text and in Table 5.2 and Table 5.3. In all cases, H atoms are represented as small white spheres, C atoms are white middle-sized spheres, whereas O, S, Cu and Mo atoms are represented in red, yellow, green and violet, respectively. All distances in Å and, in parentheses, Mo-μX-Cu (X = S or O) and S-Cu-S' bond angles in degree.

5.3 Investigation of H_2 binding to the Cu center

As a next step, we attempted to optimize Mo^{VI}/Cu^{I} adducts in which the copper ion coordinates H_2: in particular, we initially considered H_2 binding to adduct **1**, i.e. to a thiolato-copper site (S^{2-})-bridged to a MoO_2 core. In such framework, energy minimizations started either from a guess structure featuring H_2 non-classically bound to Cu-which corresponds to the initial adduct previously proposed for H_2 oxidation catalysis [132] (see Figure 5.3) or from the corresponding side-on alternative configuration invariably lead to H_2 detachment from the metal. Such an unexpected result points at the possibility that H_2 binding is preceded by some key structural reorganization of the active site, which favors dihydrogen coordination. In particular, we considered protonation events as plausible reactions to trigger structural rearrangements, in consideration of the fact that the enzyme is believed to mobilize protons in the course of its physiological activity [126].

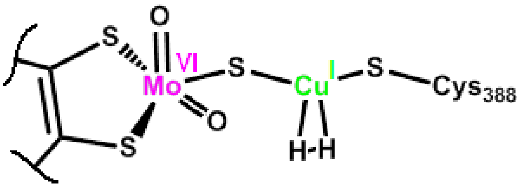


Figure 5.3 Structure of the H_2-bound CO dehydrogenase active site as proposed by Wilcoxen and Hille [132].

Differently from the case of the divalent copper ion, which is known to induce deprotonation of cysteine residues bound to it [237], proton transfer to a Cu^{I}-bound Cys^{-} is expected to be a feasible process. The protonated version of the enzyme active site was geometry optimized, leading to adduct **3** (see Figure 5.2). In **3**, the S(Cys)-Cu distance lengthens slightly as compared with **1** (2.22 vs. 1.84 Å, respectively), whereas the Mo–Cu distance shortens by 1.20 Å (2.63 Å, see Figure 5.2); another notable feature of **3** is that one of the oxo ligands bound to Mo assumes a bridging position, as its distances from the latter ion and from Cu are 1.82 and 2.04 Å, respectively; concomitantly, the Mo-μS-Cu bond angle reduces by 45.4°, as compared with **1**.

Notably, dihydrogen attachment to **3** leads to a bound minimum in this case, i.e. adduct **4** (see Figure 5.2). Such binding reaction features a small kinetic barrier (~5 kcal/mol) and formation of the product turned out to be associated with a variation in total energy of +3.7 kcal/mol. Adduct **4** is likely to represent the key intermediate for the initiation of H_2 oxidation by Mo-CODH, as suggested by the fact that dihydrogen is activated in this complex (H-H distance: 0.80 Å). As compared with **3**, in **4** the distances between the μ-oxo ligand and the metal ions do not change significantly (length differences smaller than 0.02 Å in both cases). The same holds true for the Mo-Cu distance (2.65 Å in **4**), whereas the Cu–S(Cys) bond significantly lengthens upon H_2 binding (2.39 Å in **4**, as compared with the 2.22 Å distance above reported in the case of adduct **3**).

Not differently from complexes **1**-**3**, adduct **4** attains the Mo^{VI}/Cu^{I} redox state; however, as already mentioned, Wilcoxen and Hille identified a paramagnetic H_2-bound adduct in their EPR experiments, which could be attributed to the one-electron reduced form of the encounter complex–an enzyme state that is not expected to take part in the catalytic cycle [132]. Notably, energy minimization of the one-electron reduced counterpart of **4** leads to the detachment of H_2 from the Cu ion, a result incompatible with the experimental outcomes. However, our DFT model is able to rationalize the formation of a Mo^{V}/Cu^{I} adduct capable of coordinating H_2, if the following points are taken into account: (i) detachment of the Cu ion from the protonated cysteinyl ligand as an outcome of the above mentioned weakening of the Cu–S(Cys) bond; (ii) structural rearrangement of the resulting adduct **5** to give place to adduct 6,in which the μ- oxo ligand is substituted by a μ-S ion (see Figure 5.2). The first of the above reactions features a favorable energy profile with a small kinetic barrier of ~1 kcal/mol. As far as the second reaction is concerned, adduct **6** is 7.6 kcal/mol more stable than **5**; notice that the isomerization reaction leading to the latter adduct occurs by overcoming a relatively low barrier of 7.0 kcal/mol. The structure of the transition state connecting the adducts **5** and **6** is shown in Figure 5.4. Remarkably, concomitant one-electron reduction and protonation of **6** gives place to a hydroxo-containing Mo^{V}/Cu^{I} species in which H_2 maintains coordination to the Cu ion (adduct **7**, Figure 5.2).

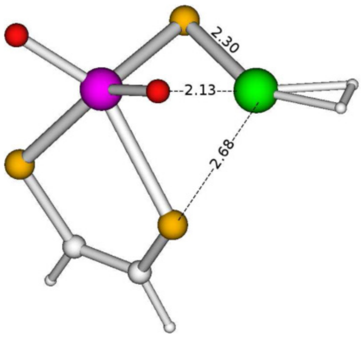


Figure 5.4 Optimized structure of the transition state connecting adducts **5** and **6**; the adopted color code for atoms is the same as in Figure 5.2

It should be pointed out that the eigenvector following approach used for the search of transition state structures (see section 5.1) was fully successful in the search for the TS connecting adducts **5** and **6**, while it did not converge in the case of the $SHCH_3$ detachment reaction from the Cu center, as well as for the reaction of H_2 binding to Cu, probably because of the extreme flatness of the potential energy surface (PES) in the regions of relevance for such reactions. However, reaction barriers could be reasonably estimated for both these regions as mentioned above. In fact, for the H_2 binding reaction ($\mathbf{3} + H_2 \rightarrow \mathbf{4}$), the eigenvector-following procedure ended up on a second-order saddle point (imaginary frequencies at 52.6i and 34.7i cm^{-1}) and it was verified that the first of such imaginary frequencies is associated with the reaction coordinate of interest; this second order saddle point corresponds to an upper bound of 5.3 kcal/mol for the barrier associated to the reaction **3** + H_2 → **4**. As far as the detachment of the SHCH3 moiety from **4** is concerned (reaction **4** → **5** + $SHCH_3$), not even a higher order saddle point was possible to be localized; therefore, we performed a systematic sampling of the PES by carrying out constrained optimizations in which the Cu-S distance of interest was fixed at various, increasing lengths with respect to the bond distance found in **4** (see Figure 5.5); in correspondence of a distance of around 2.8 Å, an energy maximum is found, which is ~1 kcal/mol above the parent adduct.

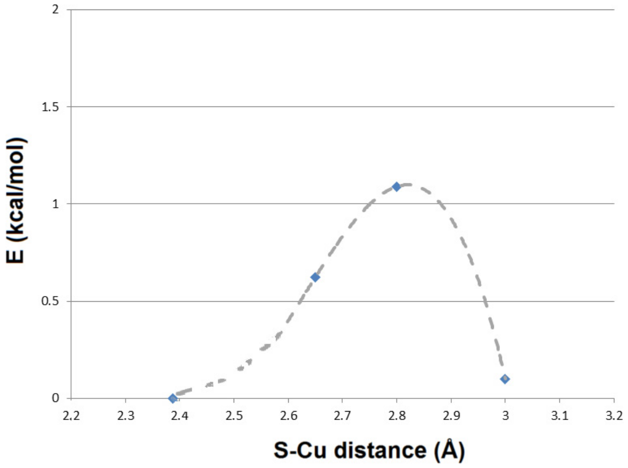


Figure 5.5 Energy profile of $SHCH_3$ detachment from Cu, as determined from geometry optimizations constrained at selected Cu-S($SHCH_3$) distances (see light blue squares at 2.65, 2.80 and 3.00 Å). The grey curve was drawn to help the eye.

5.4 EPR characterization of the paramagnetic H_2-bound adduct

The analysis of the atomic spin densities of **7** (see Table 5.1)shows that the unpaired electron is mainly localized on the Mo atom, corroborating the assignment of the Mo^V/Cu^I redox state. The spin is slightly delocalized on the bridging Sb atom, the Cu atom and the H atom of the hydroxyl ligand. Clearly, the presence of spin density on the copper atom may be in accord with the large ^{63}Cu hyperfine coupling constants observed in the experiments. In addition, the spin localized on the H atom of the hydroxyl ligand suggests that the ^1H hyperfine coupling constants determined from the EPR spectrum can indeed be assigned to this atom. Given its structural features and the spin density distribution, **7** represents a plausible candidate for the paramagnetic dihydrogen-bound form of the enzyme: therefore, we performed calculation of EPR parameters of **7**, and compared the theoretical data with the available experimental results. Notably, g values computed for model **7** as explained in methods are $g_x = 2.003$, $g_y = 1.971$, $g_z = 1.953$, which compare well with the experimentally determined g-tensor ($g_x = 2.0127$, $g_y = 1.9676$, $g_z = 1.9594$). We also performed calculations of the hyperfine coupling constants for Cu^I. As shown in Table 5.2, these computed values are underestimated by up to a factor of 2.9.

Table 5.1 Atomic spin populations of complex **7** calculated using the BP86 functional and two different basis sets. Non-metal atoms are tagged as reported in the optimized structure of adduct 7 in Figure 5.2.

	BP86/def2-TZVP	**BP86/DGDZVP**
Mo	0.96	0.88
Cu	0.02	0.03
Sb	0.06	0.09
Sdt1	0.00	0.00
Sdt2	0.00	0.00
O1	-0.06	-0.06
O2(H)	0.00	0.01
H(O2)	0.03	0.04
H1(H_2)	0.00	0.00
H2(H_2)	0.00	0.01

Table 5.2 Comparison between experimental and computed Cu hyperfine coupling constants (in MHz) and those computed for adduct **7**.

	A_1	**A_2**	**A_3**
Experimental values	169.1	200.2	170.2
Computed§	59	79	82

§Calculation of hyperfine couplings at the B3LYP/DGDZVP level provided similar results (data not shown).

As far as proton hyperfine coupling constants are concerned, experiments show signal at 80, 20 and 130 MHz; notably, the theoretical hyperfine coupling constants of the H atoms in Cu-bound dihydrogen are very small, and a correspondence with experimental values cannot be identified (all computed values smaller than 13 MHz in module, see Table 5.3). However, the hyperfine couplings calculated for the H atom of the hydroxo ligand bound to the Mo center do fit significantly better than the experimental data (theoretical values: 57, 60 and 70MHz, see the last row of Table 5.3), a result which suggests that protonation of the Mo-bound oxide(s) might be at the basis of the experimentally observed hyperfine structure of the EPR spectra.

The low value of computed Cu hyperfine couplings for **7** is a rather surprising outcome, particularly in consideration of the fact that such complex is expected to feature transannullar effects within the Mo(μS)$_2$Cu ring. Since slight changes in the Cu coordination geometry may lead to substantially different Cu hyperfine coupling constants–as discussed in [238]–and a high degree of delocalization was reported in a previous study on a Mo^V(μS)Cu^I biomimetic complex [239], we performed additional geometry optimizations in which the Cu–Sb distance was constrained stepwise at increasing values (2.90, 3.35 and 3.80 Å). Interestingly, the Mulliken spin population at the Cu center initially decreases–it assumes a value of 0.01 spin units when the Cu-Sb is set at 2.90 and 3.35 Å–whereas it significantly increases up to 0.04 spin units when the Cu–Sb distance is further lengthened to 3.80 Å. We unsuccessfully spent extensive efforts to directly localize–by unconstrained geometry optimizations–minimum geometries of the Cu-H_2 bound active site featuring a $Mo^V(\mu S)Cu^I$ framework. Further efforts were therefore spent on alternative energy minimizations and characterization of selected isomers of **7** (or similar adducts featuring a different protonation state, *vide infra*) starting from guess geometries featuring 4-membered-ring core system (see Figure 5.6 and Figure 5.7).

Table 5.3 Comparison between experimental and computed H hyperfine coupling constants (in MHz). Atoms of the model are tagged as reported in the optimized structure of adduct 7 in Table 5.2.

		A_1	A_2	A_3
Experimental values[a]		80	20	130
Computed[b]	H1(H_2)	0	3	4
	H2(H_2)	10	10	1
	H(O2)	57	60	70

[a]Experimental spectra identified two approximately equivalent protons with very similar coupling constants, with a maximum deviation of 10 MHz, which is well below the accuracy of DFT-computed hyperfine couplings. However, such picture could be rationalized within our model either by considering the occurrence of protonation of the second Mo-bound oxo ligands, or based on the co-presence in the sample of different rotamers for the same hydroxo ligand; [b]Calculation of hyperfine couplings at the B3LYP/DGDZVP level provided similar results (data not shown).

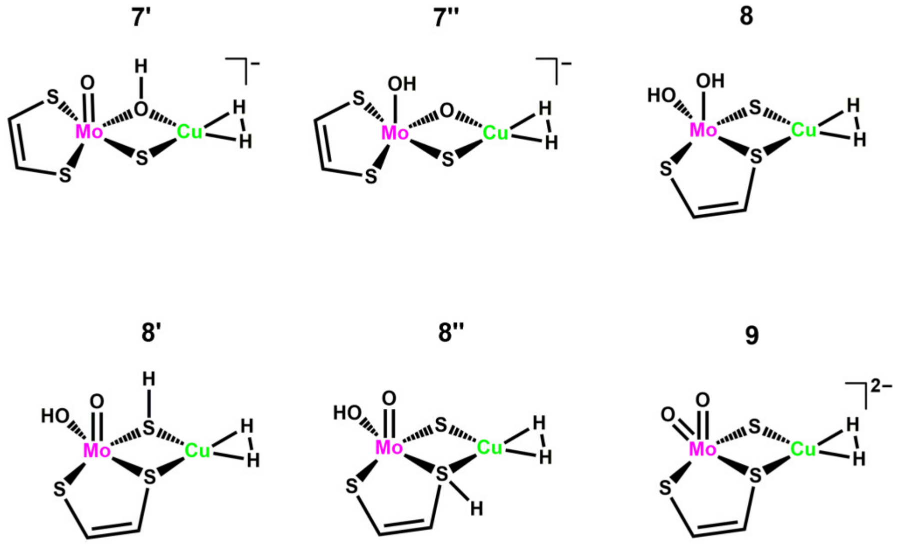


Figure 5.6 Schematic representation of the starting geometries for the optimization of variously protonated H_2 bound adducts.

As a first step, we optimized isomers of **7** in which the H_2-bound Cu center variously coordinates the oxygen and sulfur centers available in the cluster (species **7'** and **7''**, optimized geometries in Figure 5.7). Differently from the case of **7''**, **7'** features a Cu spin population significantly larger than the one of the parent species (see Table 5.4); however, **7** is significantly more stable than **7'** and **7''**, respectively by 10.3 and 18.8 kcal/mol; therefore, neither of such adducts were considered further for EPR parameter calculations. The architecture of 7 was then considered to produce guess geometries resulting from addition of one proton to the available basic sites (see Figure 5.6); the optimized geometries of the corresponding complexes (**8**, **8'** and **8''**) are shown in Figure 5.7. Adduct **8** turned out to correspond to the lowest energy minimum among the three, **8'** and **8''** being higher in energy by 4.0 and 28.6 kcal/mol. All these species feature low spin population values at the Cu ion (see Table 5.4), and therefore they were also not considered for any magnetic properties calculation. Lastly, a hypothetical adduct originating from deprotonation of the Mo- bound OH group in **7** was also considered for geometry optimization (see **9** in Figure 5.6), but it turned out that H_2 detaches from the Cu ion upon geometry optimization.

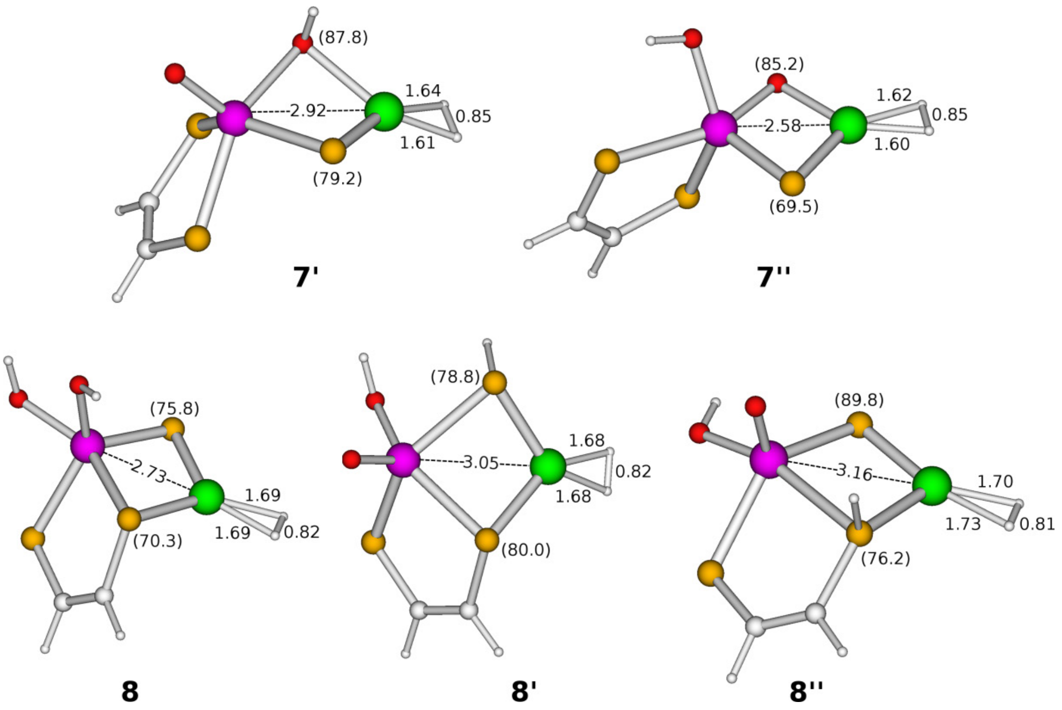


Figure 5.7 Optimized structures of the H_2-bound $Mo^{V}Cu^{I}$ adducts **7'**, **7''**, **8**, **8'** and **8''**; the adopted color code for atoms is the same as in Figure 5.2. All distances in Å and, in parentheses, Mo-µX-Cu (X = S or O) bond angles in degree.

Table 5.4 Mulliken Cu spin populations of adducts **7'**, **7''**, **8**, **8'** and **8''**, calculated at the BP86/def2-TZVP level; the Cu spin population for adduct **7** is also shown for comparison.

	7	7'	7''	8	8'	8''
Spin population	0.02	0.05	0.01	0.01	0	0.02

5.5 Lack of reactivity of the Ag-substituted enzyme toward H_2

Further experimental results that can be used for validation of the picture proposed here for the H_2-enzyme binding process come from the observation that substitution of copper with silver in the protein active site leads to the loss of H_2-oxidizing activity of the active site [132]. In view of this, we also performed geometry optimizations using the geometry of **4** as guess structures, after substitution of Cu with Ag. Notably, no bound complex was possible to be localized in this case, as H_2 detached from the Ag center in the optimization attempts we performed (also in this case, both side-on and end-on attachment of dihydrogen to silver were considered in the guess structures). Not surprisingly, H_2 detachment in our Ag-containing model occurred also when Cys388 is not protonated, analogously to the case of the wild-type enzyme. Such theoretical results are fully consistent with the experimental observations, which indicate that a silver-containing version of the enzyme is not reactive toward H_2 [132].

5.6 Conclusions

In conclusion, our DFT data indicate that H_2 binding to the copper center of Mo-CODH *O. carboxidovorans* follows a route different from the one of CO binding to the same active site: protonation of the cysteinyl sulfur coordinated to the Cu ion is able to trigger a shortening of the Mo-Cu distance, with concomitant reorganization of the copper first coordination sphere. This favors the occurrence of a binding interaction between the latter and dihydrogen. Our model allows also to account for the differences experimentally observed between the wild-type and the Ag-substituted enzyme with regard to H_2-binding abilities: in fact, it turned out that a model in which the Cu^I ion is replaced by Ag^I is unable to coordinate H_2, a result that allows to rationalize the spectroscopic findings on the Ag-containing enzyme. As for the modeling of a peculiar paramagnetic species detected to date only upon H_2 exposure of the wild-type enzyme, PESs sampling of H_2-bound species allowed to identify an adduct featuring calculated g tensors in good agreement with the experimental counterpart. However, marked underestimation of the computed Cu hyperfine coupling constants was observed for such adduct. Future QM/MM studies of the entire enzyme will hopefully allow to unveil the determinants of the EPR properties of this putative side-product occurring in parallel to H_2-oxidation catalysis.

CHAPTER

6

CONCLUDING REMARKS

In view of anthropogenic global warming and the depletion of carbon-based fossil fuels, future decades will face an inevitable demand for alternative energy sources. Hydrogen has been proposed as an ideally clean fuel storing large amounts of energy, which can be released without producing greenhouse gases. However, a hydrogen-based economy is so far impeded by the lack of sustainable approaches for the production and activation of molecular hydrogen.

Hydrogenases are metallo-enzymes catalysing the reversible interconversion of protons and reducing equivalents into molecular hydrogen at high rates and at ambient temperatures. Therefore, they are considered very promising targets for sustainable H_2 evolution and oxidation. On the other hand, carbon monoxide dehydrogenases have attract increasing interest in recent years for their capability to catalyse the conversion of CO_2 and CO. Reverse engineering studies of these enzymes could drive the development of synthetic catalysts for efficient large-scale processes for conversion of atmospheric CO_2 into useful materials, and inspire alternative routes for similar reactions in chemical industries. In this context, quantum mechanical calculations were carried out on models of the active site of these enzymes.

With the aim of contributing to shed more light on the reactivity of Ni-CODHs and, more specifically, disclose which stereoelectronic properties of their active site are crucial for efficient CO_2 and CO binding, a comparative approach was adopted to study a series of systems ranging from minimal models of the active site, to very large models in which all the aminoacids forming the first and second coordination sphere of the C-cluster have been explicitly included. The results of our DFT study contributed to a better understanding of the electronic structure, the nature of the active ligands and their coordination mode in the C_{red1} and C_{red2} states of the C-cluster; C_{red1} is predicted to be a Ni(II)-Fe_u(III) species with a hydroxide ligand coordinated to Fe_u, while C_{red2} is predicted to be a Ni(II)-Fe_u(II) species with no exogenous species coordinated to the C-cluster or a hydride ligand bridging the Ni and Fe atoms, depending on the protonation state of His93. In addition, the DFT investigation of CO_2 and CO binding to the C-cluster in different redox states, also in the presence of other substrates, elucidated several aspects of the Ni-CODH catalytic mechanism and highlighted the crucial role of the protein environment in tuning the coordination geometry and the stability of the intermediates along the reaction pathway. In particular, according to our calculations, CO binding and release take place via an associative mechanism, whereas binding and release of CO_2 follows a dissociative pathway. For both CO_2 and CO was identified a stable conformation in which the substrate is terminally bound to the Ni atom that plays a key role in the binding/dissociation of the CO_2 and CO molecules to/from the C-cluster. Moreover, the direction of the catalytic reaction was found to be strongly influenced by the protonation state of the His93 residue. CO_2 reduction is favored by the doubly protonated state of His93, which promotes the CO_2 binding to the C-cluster and the CO release from the active site. Conversely, the CO oxidation is favored by the singly protonated state on the δN atom of His93, that assists the CO biding and the CO_2 release. Despite the significant insights into the stereoelectronic and catalytic properties of the Ni-CODH active site provided by our study, several steps of the catalytic mechanism are still not clear. In particular, further investigations are required to disclose the key factors for efficient CO_2 cleavage at the active site and the nature of the movement of CO towards the OH^- ligand to form the CO_2 molecule.

Concerning the hydrogenases enzymes, my PhD research was focused on inactive forms of O_2-sensitive [NiFe]-hydrogenases since one of the major challenges for their practical technological uses is the oxidative inactivation of the active site. Aerobic and anaerobic oxidation of [NiFe]-hydrogenases yields two inactive states, called Ni-A and Ni-B, that can be reactivated under reducing conditions. While there is a general consensus on the structure of Ni-B, the nature of Ni-A is still controversial. In this respect, the nature of the Ni-A state was explored by a DFT investigation based on systematic explorations of conformational and configurational space relying on accurate energy calculations, and on comparisons of theoretical geometries with the several X-ray structures (differing for the nature of the bridging ligand in the active site and for the presence of modified cysteine residues) assigned to this state. Among all plausible isomers featuring various protonation patterns and oxygenic ligands, our calculations show that the one corresponding to the crystallographic structure recently reported by Volbeda *et al.* [63] (featuring a bridging hydroxide ligand and the sulphur atom of Cys64 oxidized to bridging sulfenate) is the most stable. However, isomers with cysteine residues oxidized to terminal sulfenate are very close in energy and modifications in the network of H-bond with neighbouring residues may alter the stability order of such species. The better understanding of the origin of stability of bridging and terminal sulfenate isomers presented in this thesis will hopefully drive mutagenesis experiment aimed at favouring the formation of a terminal sulfenate species which may feature faster reactivation kinetics.

Since it was experimentally observed that Ni-B and the Ni-A can be generated in the presence and absence of O_2, the oxidative inactivation mechanism of the active site of [NiFe]-

hydrogenases was investigated by simulating both aerobic and anaerobic conditions. While under aerobic conditions oxidation is promoted by O_2 which also provides the oxygen atoms that are inserted in the active site of inactive forms of the enzyme, under anaerobic conditions two water molecules were proposed to provide these oxygen atoms. Our results supports the experimental evidence that the NiA-to-NiB ratio increases when exposure to O_2 has taken place under more oxidizing conditions[51]. If the enzyme in the initial redox state contains all four electrons required for the full O_2 reduction, under aerobic conditions it converts to Ni-B; otherwise, Ni-A is formed. Aerobic oxidation of the Ni-C form of the enzyme leads preferentially to Ni-B because two electrons are directly supplied by the bridging hydride between the Ni and Fe atoms while further two electrons are delivered by the electron relay of the enzyme. On the other hand, aerobic oxidation of the Ni-SI_a state leads preferentially to Ni-A because it contains only two reducing equivalents (localized on the Ni^{2+} atom and the reduced medial cluster); further two electrons are supplied by the sulphur atom of Cys64 which is oxidized to bridging sulfenate. Although our calculations suggest that O_2 reacts directly at the Ni-Fe site, the attacking O_2 can also act as an electron acceptor without being incorporated in the active site. In this case, the inactivation reaction can follow the same mechanisms proposed for the anaerobic oxidation of the active site.

The DFT investigation of the reactivation mechanism of Ni-B and Ni-A allowed to rationalize the slower activation of Ni-A compared to that of Ni-B. Mono-electron reduction and protonation of the active site is sufficient to reactivate Ni-B, whereas H_2 binding to the Ni-Fe site is also required to reactivate Ni-A. Our theoretical study on [NiFe]-hydrogenases provided significant insights into understanding the stereoelectronic properties of Ni-A and Ni-B and the interconversion pathway between active and inactive forms of the enzyme that could be exploited in the design of novel O_2-tolerant biomimetic synthetic catalysts and in the direct utilization of [NiFe]-hydrogenases in biotechnological hydrogen production processes.

Finally, DFT was used to explore the reactivity of the bimetallic [Mo-(μ_2-S)-Cu] active site of Mo-CODHs. Since it was experimentally observed that these enzymes also catalyse the oxidation of dihydrogen at significant rates, its reactivity towards H_2 was investigated. Our data indicate that H_2 binding to the copper center of the active site follows a route different from that of the CO substrate: protonation of the cysteinyl sulfur coordinated to the Cu ion is able to trigger a shortening of the Mo-Cu distance, with concomitant reorganization of the copper first coordination sphere. This favors the occurrence of a binding interaction between the latter and H_2. Our study allowed also to account for the differences experimentally observed between the wild-type and the Ag-substituted enzyme with regard to H_2-binding abilities. With the aim of deepening insights into the nature of a H_2-bound paramagnetic form of the enzyme experimentally observed during the reaction of the enzyme with H_2, EPR parameters were predicted for this species and compared with the values obtained from experiments. An adduct featuring calculated g tensors in good agreement with the experimental counterpart was identified. However, marked underestimation of the computed Cu hyperfine coupling constants was observed for such adduct. Future QM/MM studies of the entire enzyme will hopefully allow to unveil the determinants of the EPR properties of this putative side-product occurring in parallel to H_2-oxidation catalysis.

BIBLIOGRAPHY

1. Bornscheuer UT, Huisman GW, Kazlauskas RJ, et al (2012)Nature 485:185–194. doi: 10.1038/nature11117
2. Coontz R, Hanson B (2004)Science 305:2004. doi: 10.1126/science.305.5686.957
3. Solomon S, Plattner G-K, Knutti R, Friedlingstein P (2009)Proc Natl Acad Sci U S A 106:1704–9. doi: 10.1073/pnas.0812721106
4. Lubitz W, Tumas W (2007)Chem Rev 107:3900–3903. doi: 10.1021/cr050200z
5. Gómez-Navarro C, Weitz RT, Bittner AM, et al (2007)Nano Lett 7:3499–3503. doi: 10.1021/nl072090c
6. Lubitz W, Reijerse E, van Gastel M (2007)Chem Rev 107:4331–4365. doi: 10.1021/cr050186q
7. Vignais PM, Billoud B (2007)Chem Rev 107:4206–4272. doi: 10.1021/cr050196r
8. Kim JYH, Cha HJ (2013)Korean J Chem Eng 30:1–10. doi: 10.1007/s11814-012-0208-8
9. Jugder BE, Welch J, Aguey-Zinsou KF, Marquis CP (2013)Rsc Adv 3:8142–8159. doi: 10.1039/c3ra22668a
10. Tard C, Pickett CJ (2009)Chem Rev 109:2245. doi: 10.1021/cr800542q
11. Heinekey DM (2009)J Organomet Chem 694:2671–2680. doi: 10.1016/j.jorganchem.2009.03.047
12. Yang JY, Bullock M, DuBois MR, Dubois D (2011)MRS Bull 36:39–47.
13. Fritsch J, Lenz O, Friedrich B (2013)Nat Rev Microbiol 11:106–114. doi: 10.1038/nrmicro2940
14. Simmons TR, Artero V (2013)Angew Chemie - Int Ed 52:6143–6145. doi: 10.1002/anie.201302908
15. Matsumoto T, Kim K, Nakai H, et al (2013)ChemCatChem 5:1368–1373. doi: 10.1002/cctc.201200595
16. Liu X, Ibrahim SK, Tard C, Pickett CJ (2005)Coord Chem Rev 249:1641–1652. doi: 10.1016/j.ccr.2005.04.009
17. Tard C, Liu X, Ibrahim SK, et al (2005)Nature 433:3–6. doi: 10.1038/nature03300.1.
18. Bouwman E, Reedijk J (2005)Coord Chem Rev 249:1555–1581. doi: 10.1016/j.ccr.2004.10.010
19. Mertens R, Liese A (2004)Curr Opin Biotechnol 15:343–348. doi: 10.1016/j.copbio.2004.06.010
20. Melis A, Happe T (2004)Photosynth Res 80:401–409. doi: 10.1023/B:PRES.0000030421.31730.cb
21. Happe T, Hemschemeier A, Winkler M, Kaminski A (2002)Trends Plant Sci 7:246–250. doi: 10.1016/S1360-1385(02)02274-4
22. Schütz K, Happe T, Troshina O, et al (2004)Planta 218:350–359. doi: 10.1007/s00425-003-1113-5
23. Esper B, Badura A, Rögner M (2006)Trends Plant Sci 11:543–549. doi: 10.1016/j.tplants.2006.09.001
24. Elam CC, Padró CEG, Sandrock G, et al (2003)Int J Hydrogen Energy 28:601–607. doi: 10.1016/S0360-3199(02)00147-7
25. Schneider J, Jia H, Muckerman JT, Fujita E (2012)Chem Soc Rev 41:2036. doi: 10.1039/c1cs15278e
26. Qiao J, Liu Y, Hong F, Zhang J (2014)Chem Soc Rev Chem Soc Rev. doi: 10.1039/c3cs60323g
27. Windle CD, Perutz RN (2012)Coord Chem Rev 256:2562–2570. doi: 10.1016/j.ccr.2012.03.010

28. Tanaka K, Ooyama D (2002)Coord Chem Rev 226:211–218. doi: 10.1016/S0010-8545(01)00434-9

29. Thoi VS, Chang CJ (2011)Chem Commun (Camb) 47:6578–6580. doi: 10.1039/c1cc10449g

30. Parkin A, Seravalli J, Vincent KA, et al (2007)J Am Chem Soc 129:10328–10329. doi: 10.1021/ja073643o

31. Wachtershauser G (2006)Philos Trans R Soc B Biol Sci 361:1787–1808. doi: 10.1098/rstb.2006.1904

32. Konhauser KO, Pecoits E, Lalonde S V, et al (2009)Nature 458:750–753. doi: 10.1038/nature07858

33. Ragsdale SW, Pierce E (2008)Biochim Biophys Acta - Proteins Proteomics 1784:1873–1898. doi: 10.1016/j.bbapap.2008.08.012

34. Armstrong F a, Hirst J (2011)Proc Natl Acad Sci U S A 108:14049–54. doi: 10.1073/pnas.1103697108

35. Vignais PM, Billoud B, Meyer J (2001)FEMS Microbiol Rev 25:455–501.

36. Fontecilla-Camps JC, Amara P, Cavazza C, et al (2009)Nature 460:814–22. doi: 10.1038/nature08299

37. Fontecilla-Camps JC, Volbeda A, Cavazza C, Nicolet Y (2007)Chem Rev 107:4273–4303. doi: 10.1021/cr050195z

38. Pierik AJ, Hulstein M, Hagen WR, Albracht SPJ (1998)Eur J Biochem 258:572–578.

39. Happe RP, Roseboom W, Pierik AJ, Albracht SPJ (1997)Nature 385:126.

40. Peters JW, Lanzilotta WN, Lemon BJ, Seefeldt LC (1998)Science (80-) 282:1853–1858. doi: 10.1126/science.282.5395.1853

41. Nicolet Y, Piras C, Legrand P, et al (1999)Structure 7:13–23. doi: 10.1016/S0969-2126(99)80005-7

42. Nicolet Y, De Lacey AL, Vernède X, et al (2001)J Am Chem Soc 123:1596–1601. doi: 10.1021/ja0020963

43. Berggren G, Adamska A, Lambertz C, et al (2013)Nature 499:66–69. doi: 10.1038/nature12239

44. Esselborn J, Lambertz C, Adamska-Venkatesh A, et al (2013)Nat Chem Biol 9:607–9. doi: 10.1038/nchembio.1311

45. Lyon EJ, Shima S, Boecher R, et al (2004)14239–14248.

46. Shima S, Lyon EJ, Sordel-Klippert M, et al (2004)Angew Chemie - Int Ed 43:2547–2551. doi: 10.1002/anie.200353763

47. Shima S, Lyon EJ, Thauer RK, et al (2005)J Am Chem Soc 127:10430–10435. doi: 10.1021/ja051895o

48. Shima S, Ermler U (2011)Eur J Inorg Chem 963–972. doi: 10.1002/ejic.201000955

49. Stripp ST, Goldet G, Brandmayr C, et al (2009)Proc Natl Acad Sci 106:17331–17336. doi: 10.1073/pnas.0905343106

50. De Lacey AL, Fernández VM, Rousset M, Cammack R (2007)Chem Rev 107:4304–4330. doi: 10.1021/cr0501947

51. Lamle SE, Albracht SPJ, Armstrong FA (2004)J Am Chem Soc 126:14899–14909. doi: 10.1021/ja047939v

52. Lamle SE, Albracht SPJ, Armstrong F a. (2005)J Am Chem Soc 127:6595–6604. doi: 10.1021/ja0424934

53. Ogata H, Lubitz W, Higuchi Y (2009)Dalt Trans 37:7577–7587. doi: 10.1039/b903840j

54. de Lacey AL, Hatchikian EC, Volbeda A, et al (1997)J Am Chem Soc 119:7181–7189. doi: 10.1021/ja963802w

55. Volbeda A, Charon MH, Piras C, et al (1995)Nature 373:580–587. doi: 10.1038/373580a0

56. Lauterbach L, Lenz O (2013)J Am Chem Soc 135:17897–17905. doi: 10.1021/ja408420d

57. Goldet G, Wait AF, Cracknell JA, et al (2008)J Am Chem Soc 130:11106–11113. doi: 10.1021/ja8027668

58. Teixeira M, Fauque G, Moura I, et al (1987)Eur J Biochem 167:47–58. doi: 10.1111/j.1432-1033.1987.tb13302.x

59. Happe RP, Roseboom W, Pierik AJ, et al (1997)Nature 385:126.

60. Pierik AJ, Roseboom W, Happe RP, et al (1999)J Biol Chem 274:3331–3337. doi: 10.1074/jbc.274.6.3331

61. Volbeda A, Martin L, Cavazza C, et al (2005)J Biol Inorg Chem 10:239–249. doi: 10.1007/s00775-005-0632-x

62. Pandelia M-E, Ogata H, Lubitz W (2010)ChemPhysChem 11:1127–1140. doi: 10.1002/cphc.200900950

63. Volbeda A, Martin L, Barbier E, et al (2015)JBIC J Biol Inorg Chem 20:11–22. doi: 10.1007/s00775-014-1203-9

64. Fritsch J, Scheerer P, Frielingsdorf S, et al (2011)Nature 479:249–252. doi: 10.1038/nature10505

65. Shomura Y, Yoon K-S, Nishihara H, Higuchi Y (2011)Nature 479:253–256. doi: 10.1038/nature10504

66. Volbeda A, Amara P, Darnault C, et al (2012)Proc Natl Acad Sci 109:5305–5310. doi: 10.1073/pnas.1119806109

67. Volbeda A, Darnault C, Parkin A, et al (2013)Structure 21:184–190. doi: 10.1016/j.str.2012.11.010

68. Garcin E, Vernede X, Hatchikian EC, et al (1999)Structure 7:557–566. doi: 10.1016/S0969-2126(99)80072-0

69. Shafaat HS, Rüdiger O, Ogata H, Lubitz W (2013)Biochim Biophys Acta 1827:986–1002. doi: 10.1016/j.bbabio.2013.01.015

70. Siegbahn PEM, Tye JW, Hall MB (2007)Chem Rev 107:4414–4435. doi: 10.1021/cr050185y

71. Bruschi M, Tiberti M, Guerra A, De Gioia L (2014)J Am Chem Soc 136:1803–1814. doi: 10.1021/ja408511y

72. Fernandez VM, Hatchikian EC, Patil DS, Cammack R (1986)Biochim Biophys Acta 883:145–154.

73. Moura JJG, Moura I, Huynh BH, et al (1982)Biochem Biophys Res Commun 108:1388–1393.

74. Cammack R, Patil DS, Hatchikian EC, Fernandez VM (1987)Biochim Biophys Acta 912:98–109.

75. Fernandez VM, Hatchikian CE, Cammack R (1985)Biochim Biophys Acta 832:69–79. doi: http://dx.doi.org/10.1016/0167-4838(85)90175-X

76. Bleijlevens B, Van Broekhuizen F a., De Lacey AL, et al (2004)J Biol Inorg Chem 9:743–752. doi: 10.1007/s00775-004-0570-z

77. Ogata H, Hirota S, Nakahara A, et al (2005)Structure 13:1635–1642. doi: 10.1016/j.str.2005.07.018

78. Söderhjelm P, Ryde U (2006)J Mol Struct THEOCHEM 770:199–219. doi: 10.1016/j.theochem.2006.06.008

79. Ogata H, Kellers P, Lubitz W (2010)J Mol Biol 402:428–444. doi: 10.1016/j.jmb.2010.07.041

80. Carepo M, Tierney DL, Brondino CD, et al (2002)J Am Chem Soc 124:281–286. doi: 10.1021/ja010204v

81. Jones AK, Lamle SE, Pershad HR, et al (2003)J Am Chem Soc 125:8505–8514. doi: 10.1021/ja035296y

82. Kurkin S, George SJ, Thorneley RNF, Albracht SPJ (2004)Biochemistry 43:6820–6831. doi:

10.1021/bi049854c

83. Pavlov M, Blomberg M, Siegbahn PEM (1999)Int J Quantum Chem 73:197–207.
84. Niu S, Thomson LM, Hall MB (1999)J Am Chem Soc 121:4000–4007. doi: 10.1021/JA983469R
85. Dementin S, Burlat B, De Lacey AL, et al (2004)J Biol Chem 279:10508–10513. doi: 10.1074/jbc.M312716200
86. Pardo A, De Lacey AL, Fernández VM, et al (2006)J Biol Inorg Chem 11:286–306. doi: 10.1007/s00775-005-0076-3
87. Stadler C, De Lacey AL, Montet Y, et al (2002)Inorg Chem 41:4424–4434. doi: 10.1021/ic020016l
88. Siegbahn PEM (2007)Comptes Rendus Chim 10:766–774. doi: 10.1016/j.crci.2007.03.011
89. Van Gastel M (2010)Appl Magn Reson 37:207–218. doi: 10.1007/s00723-009-0044-0
90. Mege RM, Bourdillon C (1985)J Biol Chem 260:14701–14706.
91. Abou Hamdan A, Burlat B, Gutiérrez-Sanz O, et al (2013)Nat Chem Biol 9:15–18. doi: 10.1038/nchembio.1110
92. Ensign S a (1995)Biochemistry 34:5372–8.
93. Svetlitchnyi V, Peschel C, Acker G, Meyer O (2001)J Bacteriol 183:5134–5144. doi: 10.1128/JB.183.17.5134-5144.2001
94. Dobbek H, Gremer L, Kiefersauer R, et al (2002)Proc Natl Acad Sci U S A 99:15971–15976. doi: 10.1073/pnas.212640899
95. Hille R, Dingwall S, Wilcoxen J (2015)J Biol Inorg Chem 20:243–251. doi: 10.1007/s00775-014-1188-4
96. Can M, Armstrong FA, Ragsdale SW (2014)Chem Rev 114:4149–4174. doi: 10.1021/cr400461p
97. Drennan CL, Heo J, Sintchak MD, et al (2001)Proc Natl Acad Sci U S A 98:11973–11978. doi: 10.1073/pnas.211429998
98. Dobbek H, Svetlitchnyi V, Gremer L, et al (2001)Science (80-) 293:1281–1285.
99. Fox JD, Kerby RL, Roberts GP, Ludden PW (1996)J Bacteriol 178:1515–1524.
100. Ensign SA, Ludden PW (1991)J Biol Chem 266:18395–18403.
101. Soboh B, Linder D, Hedderich R (2002)Eur J Biochem 269:5712–5721. doi: 10.1046/j.1432-1033.2002.03282.x
102. Darnault C, Volbeda A, Kim EJ, et al (2003)Nat Struct Biol 10:271–9. doi: 10.1038/nsb912
103. Doukov TI, Iverson TM, Seravalli J, et al (2002)Science 298:567–572. doi: 10.1126/science.1075843
104. Drake HL (1994)In: Drake HL (ed) Acetogenesis. Springer US, Boston, MA, pp 3–60
105. Grahame DA (1991)J Biol Chem 266:22227–22233.
106. Stephenson M (1998)Microbiology 144:2377–2406. doi: 10.1099/00221287-144-9-2377
107. Jeoung JH, Dobbek H (2007)Science (80-) 318:1461–1464.
108. Lindahl PA (2008)Angew Chemie - Int Ed 47:4054–4056. doi: 10.1002/anie.200800223
109. Hu Z, Spangler NJ, Anderson ME, et al (1996)J Am Chem Soc 118:830–845. doi: 10.1021/ja9528386
110. Spangler NJ, Meyers MR, Gierke KL, et al (1998)J Biol Chem 273:4059–4064. doi: 10.1074/jbc.273.7.4059
111. Ralston CY, Wang H, Ragsdale SW, et al (2000)J Am Chem Soc 122:10553–10560. doi: 10.1021/ja0009469
112. Kung Y, Drennan CL (2011)Curr Opin Chem Biol 15:276–283. doi: 10.1016/j.cbpa.2010.11.005

113. Amara P, Mouesca JM, Volbeda A, Fontecilla-Camps JC (2011)Inorg Chem 50:1868–1878. doi: 10.1021/ic102304m

114. Gu WW, Seravalli J, Ragsdale SW, Cramer SP (2004)Biochemistry 43:9029–9035.

115. Feng J, Lindahl PA (2004)Biochemistry 43:1552–1559. doi: 10.1021/bi0357199

116. Heo J, Halbleib CM, Ludden PW (2001)Proc Natl Acad Sci U S A 98:7690–7693. doi: 10.1073/pnas.141230698

117. Heo J, Staples CR, Ludden PW, et al (2001)Biochemistry 40:7604–7611. doi: 10.1021/bi002554k

118. DeRose VJ, Telser J, Anderson ME, et al (1998)J Am Chem Soc 120:8767–8776. doi: 10.1021/ja9731480

119. Gong W, Hao B, Wei Z, et al (2008)Proc Natl Acad Sci U S A 105:9558–9563. doi: 10.1073/pnas.0800415105

120. Fesseler J, Jeoung JH, Dobbek H (2015)Angew Chemie - Int Ed 54:8560–8564. doi: 10.1002/anie.201501778

121. Jeoung JH, Dobbek H (2012)J Biol Inorg Chem 17:167–173. doi: 10.1007/s00775-011-0839-y

122. Cypionka H, Meyer O (1983)J Bacteriol 156:1178–1187.

123. Meyer O, Gremer L, Ferner R, et al (2000)381:865–876.

124. Wilcoxen J, Zhang B, Hille R (2011)Biochemistry 50:1910–1916. doi: 10.1021/bi1017182

125. Dobbek H, Gremer L, Meyer O, Huber R (1999)Proc Natl Acad Sci 96:8884–8889. doi: 10.1073/pnas.96.16.8884

126. Siegbahn PEM, Shestakov AF (2005)J Comput Chem 26:888–898. doi: 10.1002/jcc.20230

127. Hofmann M, Kassube JK, Graf T (2005)J Biol Inorg Chem 10:490–495. doi: 10.1007/s00775-005-0661-5

128. Shanmugam M, Wilcoxen J, Habel-Rodriguez D, et al (2013)J Am Chem Soc 135:17775–17782. doi: 10.1021/ja406136f

129. Santiago B, Meyer O (1996)FEMS Microbiol Lett 136:157–162. doi: 10.1016/0378-1097(95)00498-X

130. Oxidoreductase PF, Menon S, Ragsdale SW (1996)Biochemistry 35:15814–15821.

131. Bhatnagar L, Krzycki J, Zeikus J (1987)FEMS Microbiol Lett 41:337–343.

132. Wilcoxen J, Hille R (2013)J Biol Chem 288:36052–36060. doi: 10.1074/jbc.M113.522441

133. Holm RH, Solomon EI (2014)114:3367–3368.

134. Lupp D, Christensen NJ, Fristrup P (2014)Dalton Trans 43:11093–105. doi: 10.1039/c4dt00342j

135. Schrödinger E (1926)Phys Rev 28:1049–1070. doi: 10.1103/PhysRev.28.1049

136. Uhlenbeck GE (1925)Utrecht

137. Dirac PAM (1928)Cambridge

138. Slater JC (1929)Phys Rev 34:1293–1322. doi: 10.1103/PhysRev.34.1293

139. Szabo A, Ostlund NS (1996)Introd to Adv Electron Struct Theory 480. doi: 10.1119/1.1973756

140. Koch W, Holthausen MC (2001)Neural Networks. doi: 10.1002/3527600043

141. Kohn W (1999)In: Nobel Lect. pp 213–237

142. Raghavachari K (2000)Theor Chem Accounts Theory, Comput Model (Theoretica Chim Acta) 103:361–363. doi: 10.1007/s002149900065

143. Thomas LH (1927)In: Math. Proc. Cambridge Philos. Soc. pp 542–548

144. Fermi E (1928)Zeitschrift für Phys 48:73–79. doi: 10.1007/BF01351576

145. Hohenberg P, Kohn W (1964)Phys Rev 136:864–871. doi: 10.1103/PhysRevB.7.1912

146. Kohn W, Sham LJ (1965)Phys Rev 140:1133–1138. doi: 10.1103/PhysRev.140.A1133

147. Vosko SH, Wilk L, Nusair M (1980)Can J Phys 58:1200–1211. doi: 10.1139/p80-159

148. Perdew JP, Wang Y (1992)Phys Rev B 45:13244–13249. doi: 10.1103/PhysRevB.45.13244

149. Perdew JP (1986)Phys Rev B 33:8822–8824.

150. Becke AD (1988)Phys Rev A 38:3098–3100.

151. Lee C, Yang W, Parr RG (1988)Phys Rev B 37:785–789. doi: 10.1103/PhysRevB.37.785

152. Perdew J, Chevary J, Vosko S, et al (1992)Phys Rev B 46:6671–6687. doi: 10.1103/PhysRevB.48.4978.2

153. Becke AD (1996)J Chem Phys 104:1040–1046. doi: 10.1063/1.470829

154. Harris J (1984)Phys Rev A 29:1648–1659. doi: 10.1103/PhysRevA.29.1648

155. Becke AD (1993)J Chem Phys 98:1372–1377. doi: 10.1063/1.464304

156. Becke AD (1993)J Chem Phys 98:5648–5652. doi: 10.1063/1.464913

157. Stephens PJ, Devlin FJ, Chabalowski CF, Frisch MJ (1994)J Phys Chem 98:11623–11627.

158. Ruzsinszky A, Perdew JP, Csonka GI, et al (2007)J Chem Phys. doi: 10.1063/1.2566637

159. Cohen AJ, P. Mori-Sánchez WY, Mori-Sánchez P, Yang W (2008)Science (80-) 321:792–4. doi: 10.1126/science.1158722

160. Patchkovskii S, Ziegler T (2002)J Chem Phys 116:7806–7813. doi: 10.1063/1.1468640

161. Zhao Y, Lynch BJ, Truhlar DG (2004)J Phys Chem A 108:2715–2719. doi: 10.1021/jp049908s

162. Perdew JP, Zunger A (1981)Phys Rev B 23:5048–5079. doi: 10.1103/PhysRevB.23.5048

163. Eichkorn K, Treutler O, Ohm H, et al (1995)Chem Phys Lett 240:283–290.

164. Lovell T, Himo F, Han W-G, Noodleman L (2003)Coord Chem Rev 238:211–232. doi: 10.1016/S0010-8545(02)00331-4

165. Siegbahn PEM, Blomberg MR a (2000)Chem Rev 100:421–437. doi: 10.1021/cr980390w

166. Niu S, Hall MB (2000)Chem Rev 100:353–406. doi: 10.1021/cr980404y

167. Simonson T, Perahia D (1995)Proc Natl Acad Sci U S A 92:1082–1086. doi: 10.1073/pnas.92.4.1082

168. Gilson M, Honig B (1988)Proteins-Structure Funct Genet 4:7–18.

169. Klamt A, Schuurmann G (1993)J Chem Soc Perkin Trans 2 799–805. doi: 10.1039/P29930000799

170. Klamt A (1995)J Phys Soc 99:2224–2235. doi: 10.1021/j100007a062

171. Ahlrichs R, Bar M, Haser M, et al (1989)Chem Phys Lett 162:165–169.

172. Neese F (2007)Version 2.6, Revision 04, Bonn–Germany.

173. Neese F (2012)Wiley Interdiscip Rev Comput Mol Sci 2:73–78. doi: 10.1002/wcms.81

174. Schrödinger L (2015)

175. Schaftenaar G, Noordik JH (2000)J Comput Aided Mol Des 14:123–134. doi: 10.1023/A:1008193805436

176. Flükiger P, Lüthi HP, Portmann S, Weber J (2000)Swiss Cent. Sci. Comput. Manno, Switz. 2002:

177. Pettersen EF, Goddard TD, Huang CC, et al (2004)J Comput Chem 25:1605–1612. doi: 10.1002/jcc.20084

178. Lindahl PA (2002)Biochemistry 41:2097–2105. doi: 10.1021/bi015932+

179. Siegbahn PEM, Himo F (2009)J Biol Inorg Chem 14:643–651. doi: 10.1007/s00775-009-0511-y

180. Siegbahn PEM, Himo F (2011)Wiley Interdiscip Rev Comput Mol Sci 1:323–336. doi: 10.1002/wcms.13

181. Alberto ME, Marino T, Russo N, et al (2012)Phys Chem Chem Phys 14:14943–14953. doi:

10.1039/c2cp41836c

182. Kim EJ, Feng J, Bramlett MR, Lindahl PA (2004)Biochemistry 43:5728–5734. doi: 10.1021/bi036062u

183. Jonas V, Thiel W (1995)J Chem Phys 102:8474–8484. doi: 10.1063/1.468839

184. Scha̋fer A, Huber C, Ahlrichs R (1994)J Chem Phys 100:5829. doi: 10.1063/1.467146

185. Schäfer A, Huber C, Ahlrichs R (1992)J Chem Phys 97:2571–2577. doi: 10.1063/1.463096

186. Noodleman L (1981)J Chem Phys 74:5737–5743. doi: 10.1063/1.440939

187. Fraser DM, Lindahl PA (1999)Biochemistry 38:15706–15711.

188. Russell WK, Lindahl PA (1998)Biochemistry 37:10016–10026.

189. Anderson ME, Lindahl PA (1996)Biochemistry 35:8371–8380. doi: 10.1021/bi952902w

190. Chatt J, Duncanson LA (1953)J Chem Soc 2939–2947.

191. Dewar JS (1951)Bull Soc Chim Fr 18:C71–C79.

192. Dobbek H, Svetlitchnyi V, Liss J, Meyer O (2004)J Am Chem Soc 126:5382–5387. doi: 10.1021/ja037776v

193. Chen J, Huang S, Seravalli J, et al (2003)Biochemistry 42:14822–14830.

194. Seravalli J, Kumar M, Lu WP, Ragsdale SW (1997)Biochemistry 36:11241–11251. doi: 10.1021/bi970590m

195. Kung Y, Doukov TI, Seravalli J, et al (2009)Biochemistry 48:7432–7440. doi: 10.1021/bi900574h

196. Volbeda A, Fontecilla-Camps JC (2005)Dalt Trans 3443–3450. doi: 10.1039/b508403b

197. Barilone JL, Ogata H, Lubitz W, van Gastel M (2015)Phys Chem Chem Phys 17:16204–16212. doi: 10.1039/C5CP01322D

198. Van Schooneveld MM, DeBeer S (2015)J Electron Spectros Relat Phenomena 198:31–56. doi: 10.1016/j.elspec.2014.12.001

199. Amara P, Volbeda A, Fontecilla-camps JC, Field MJ (1999)J Am Chem Soc 121:4468–4477. doi: 10.1021/ja983971b

200. De Gioia L, Fantucci P, Guigliarelli B, Bertrand P (1999)Int J Quantum Chem 73:187–195. doi: 10.1002/(SICI)1097-461X(1999)73:2<187::AID-QUA12>3.0.CO;2-X

201. Van Gastel M, Stein M, Brecht M, et al (2006)J Biol Inorg Chem 11:41–51. doi: 10.1007/s00775-005-0048-7

202. Pardo A, De Lacey AL, Fernández VM, et al (2007)J Biol Inorg Chem 12:751–760. doi: 10.1007/s00775-007-0227-9

203. Harris T V., Szilagyi RK (2016)J Comput Chem 37:1681–1696. doi: 10.1002/jcc.24384

204. Kampa M, Lubitz W, Van Gastel M, Neese F (2012)J Biol Inorg Chem 17:1269–1281. doi: 10.1007/s00775-012-0941-9

205. Teixeira VH, Soares CM, Baptista AM (2008)Proteins 70:1010–1022. doi: 10.1002/prot.21588

206. Galván IF, Volbeda A, Fontecilla-Camps JC, Field MJ (2008)Proteins 73:195–203. doi: 10.1002/prot.22045

207. Bone S, Pethig R (1985)J Mol Biol 181:323–326.

208. Grimme S, Antony J, Ehrlich S, Krieg H (2010)J Chem Phys 132:154104–154119. doi: 10.1063/1.3382344

209. Grimme S (2010)J Comput Chem 27:1787–1799. doi: 10.1002/jcc

210. Grimme S (2004)J Comput Chem 25:1463–1473. doi: 10.1002/jcc.20078

211. Helgaker T (1991)Chem Phys Lett 182:503–510. doi: 10.1016/0009-2614(91)90115-P

212. Jayapal P, Sundararajan M, Hillier IH, Burton NA (2006)Phys Chem Chem Phys 8:4086–4094. doi: 10.1039/b608069c

213. Matias PM, Soares CM, Saraiva LM, et al (2001)J Biol Inorg Chem 6:63–81. doi: 10.1007/s007750000167

214. Ogata H, Mizoguchi Y, Mizuno N, et al (2002)J Am Chem Soc 124:11628–11635. doi: 10.1021/ja012645k

215. Grapperhaus CA, Darensbourg MY (1998)Acc Chem Res 31:451–459. doi: 10.1021/ar950048v

216. Marques MC, Coelho R, De Lacey AL, et al (2010)J Mol Biol 396:893–907. doi: 10.1016/j.jmb.2009.12.013

217. Marques MC, Coelho R, Pereira IAC, Matias PM (2013)Int J Hydrogen Energy 38:8664–8682. doi: 10.1016/j.ijhydene.2013.04.132

218. Volbeda A, Amara P, Iannello M, et al (2013)Chem Commun (Camb) 49:7061–3. doi: 10.1039/c3cc43619e

219. Montet Y, Amara P, Volbeda A, et al (1997)Nat Struct Biol 4:523.

220. Lide DR (1997)CRC press

221. Bruschi M, De Gioia L, Zampella G, et al (2004)JBIC J Biol Inorg Chem 9:873–884. doi: 10.1007/s00775-004-0588-2

222. George SJ, Kurkin S, Thorneley RNF, Albracht SPJ (2004)Biochemistry 43:6808–6819. doi: 10.1021/bi049853k

223. Lenz O, Ludwig M, Schubert T, et al (2010)ChemPhysChem 11:1107–1119. doi: 10.1002/cphc.200901002

224. Parkin A, Sargent F (2012)Curr Opin Chem Biol 16:26–34. doi: 10.1016/j.cbpa.2012.01.012

225. Evans RM, Parkin A, Roessler MM, et al (2013)J Am Chem Soc 135:2694–2707. doi: 10.1021/ja311055d

226. Liebgott P-P, Leroux F, Burlat B, et al (2010)Nat Chem Biol 6:63–70. doi: 10.1038/nchembio.276

227. van der Zwaan JW, Coremans JMCC, Bouwens ECM, Albracht SPJ (1990)Biochim Biophys Acta - Protein Struct Mol Enzymol 1041:101–110. doi: 10.1016/0167-4838(90)90051-G

228. Cammack R, Patil D, Aguirre R, Hatchikian EC (1982)Febs Lett 142:289–292.

229. Coremans JMCC, van der Zwaan JW, Albracht SPJ (1992)Biochim Biophys Acta 1119:157.

230. Teixeira M, Moura I, Xavier A V., et al (1989)J Biol Chem 264:16435–16450.

231. Bleijlevens B (2002)

232. Davidson G, Choudhury SB, Zhijie G, et al (2000)Biochemistry 39:7468–7479. doi: 10.1021/bi000300t

233. Weigend F, Ahlrichs R (2005)Phys Chem Chem Phys 7:3297–305. doi: 10.1039/b508541a

234. Fan HJ, Hall MB (2001)J Am Chem Soc 123:3828–3829. doi: 10.1021/ja004120i

235. Sosa C, Andzelm J, Elkin BC, et al (1992)J Phys Chem 96:6630–6636. doi: 10.1021/j100195a022

236. Godbout N, Salahub DR, Andzelm J, Wimmer E (1992)Can J Chem 70:560–571. doi: 10.1139/v92-079

237. Dudev T, Lim C (2004)J Am Chem Soc 126:2602–2612. doi: 10.1021/ja038827r

238. Rokhsana D, Large TAG, Dienst MC, et al (2016)JBIC J Biol Inorg Chem 21:491–499. doi: 10.1007/s00775-016-1359-6

239. Gourlay C, Nielsen DJ, White JM, et al (2006)J Am Chem Soc 128:2164–2165. doi: 10.1021/ja056500f