SCUOLA DI DOTTORATO
UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

Department of Earth and Environmental Sciences
PhD program in Chemical, Geological and Environmental Sciences
Cycle XXXII - Curriculum in Chemical Sciences

A COMPUTATIONAL OUTLOOK ON THE CATALYSIS EXERTED BY THE UNIQUE ACTIVE SITE OF MoCu CO DEHYDROGENASES

Rovaletti Anna

Registration number 786723

Tutor: Prof. Claudio Greco

Coordinator: Prof. Maria Luce Frezzotti

ACADEMIC YEAR 2019/2020

Dedicated to the people with whom I had the honour
of sharing the effort and joy of research

Abstract

Production and consumption processes in soil ecosystems contribute to the global biochemical cycles of many trace gases (CH_4, CO, H_2, N_2O and NO) that are relevant for atmospheric chemistry and climate. Such small gas molecules play different role into the metabolism of microorganisms placed in soil that rely on specific metalloenzymes for their transformation. Among these, molybdenum-based metalloenzymes were evidenced to be crucial in such context. In particular, a specific molybdoenzyme was reported to be involved in atmospheric CO oxidation. MoCu CO dehydrogenases (MoCu CODH) is an enzyme found in aerobic carboxido-bacteria, such as *Oligotropha carboxidovorans* which represent one of the essential components in the biogeochemical carbon monoxide (CO) consumption. In fact, they contribute to maintenance of sub-toxic concentration of CO in the lower atmosphere by processing approximately $2{\times}10^8$ tons of it annually. This bacterial metalloprotein catalyses the oxidation of CO to CO_2, while it can also split H_2 in two protons and two electrons. Such reactions are performed thanks to a unique active site composed of two metals, a copper ion and a molybdenum one, linked together through a sulphur atom. Despite extended theoretical and experimental studies had been carried out concerning this enzyme, several aspects related to its reactivity have not been unravelled.

In the present thesis, we focused on the *in silico* description of MoCu CODH in order to deepen the understanding of the reaction mechanisms catalysed by the enzyme. To do so, in the framework of density functional theory (DFT), we applied models of different sizes to obtain an accurate description of the system.

In the context of CO oxidation catalysis, we evidenced that if a previously proposed thiocarbonate-like intermediate is formed along the catalytic path, it does not represent a rate-limiting species on the enzymatic energy landscape, differently from results of previous theoretical studies. Moreover, we were able to suggest an alternative catalytic mechanism for the oxidation of CO that involves the direct role of a water molecule, activated by the sourrounding active site.

As for the MoCu CODH hydrogenase activity, two plausible mechanisms for the splitting of H_2 were presented. For the first time we suggested that the MoCu CODH active site may be viewed as a Frustrated Lewis Pair (FLP), and we proposed a FLP-like mechanism for oxidation of the dihydrogen. Alternatively, a protonation event –e.g. Cu-bound cysteine residue protonation – prior to binding of H_2 to the active site proved to be necessary to present a plausible reactive channel.

Contents

List of publications

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

I **The challenging in silico description of carbon monoxide oxidation as catalyzed by molybdenum-copper CO dehydrogenase**
A. Rovaletti, M. Bruschi, G. Moro, U. Cosentino and C. Greco
Frontiers in chemistry 2019, 6, 630

II **A thiocarbonate sink on the enzymatic energy landscape of aerobic CO oxidation? Answers from DFT and QM/MM models of MoCu CO-dehydrogenases**
A. Rovaletti, M. Bruschi, G. Moro, U. Cosentino, U. Ryde, and C. Greco
Journal of catalysis 2019, 372, 201–205

III **Theoretical Insights into the Aerobic Hydrogenase Activity of Molybdenum–Copper CO Dehydrogenase**
A. Rovaletti, M. Bruschi, G. Moro, U. Cosentino, C. Greco and U. Ryde
Inorganics 2019, 7.11, 135

IV **QM/MM study of the binding of H_2 to MoCu CO dehydrogenase: development and application of improved H_2 parameters for a correct treatment of gas-protein interaction**
A. Rovaletti, C. Greco and U. Ryde
Submitted to the Journal of Molecular Modelling

V **The direct role of water in CO oxidation catalysed by Mo-Cu CO-dehydrogenase. Answers from theory.**
A. Rovaletti, M. Bruschi, U. Cosentino, G. Moro, U. Ryde and C. Grecoo
In preparation

List of Publications not included in the Thesis

1. A. Rovaletti and C. Greco;
   "Organophosphorous ligands in hydrogenase-inspired iron-based catalysts: A DFT study on the energetics of metal protonation as a function of P-atom substitution"
   Journal of Physical Organic Chemistry 2018, 31.1, e3748.

2. A. Rovaletti, G. Moro, U. Cosentino, and C. Greco;
   "How general is the effect of organic ligands bulkiness on the basicity on metalorganic catalysts? Steps away from the metal and through the transition series by DFT modeling"
   In preparation.

Acknowledgements

I would like to thank my supervisor, Prof. Claudio Greco, who guided me so positively and who always made me feel confident in my abilities. You are the best strategist I have ever met, without your careful planning I would never have been able to achieve these results.

I would also like to thank Prof. Ulf Ryde, who hosted me in his research group in Lund, Sweden. I would like to express my sincere gratitude for providing guidance and feedback throughout most of my PhD projects.

Finally, I would like to thank my family and friends for supporting me during the darkest times of my research as well as during the compilation of my publications and of this dissertation.

In conclusion, I would like to acknowledge myself for never having lost hope!

A computational outlook on the catalysis exerted by the unique active site of MoCu CO dehydrogenases

1 Introduction

It is well established that soil ecosystems can consume the trace concentrations of several gases (H_2, CO, CH_4, N_2O, NO, ...) found in the lower atmosphere (troposphere)[1]. As both very significant and least understood processes in the context of global biochemical cycles, these phenomena have inspired interest among atmospheric chemists for several decades. Actually, in general the composition of the atmosphere is controlled by natural and anthropogenic emission of gases and aerosols, chemical and microphysical transformations, wet scavenging, surface uptake or emission by the land ecosystems and by the ocean and its ecosystems. Changes in the concentration of atmospheric trace gases, mainly due to human activity, are contributing to climate warming. Significant increase of the amount of carbon dioxide (CO_2), methane (CH_4), carbon monoxide (CO), oxidised nitrogen compounds (NO and N_2O) and decrease of dimethyl sulphide (DMS) has been observed since the middle of the last century. CO_2, CH_4, NO and NO_2 contribute to a rise in the global temperature due to their ability of absorbing infrared radiation and their persistence in the atmosphere[2,3]. Meanwhile, CO can influence the concentration of some of the above mentioned greenhouse gases by controlling the oxidising capacity of the atmosphere –e.g. by reacting with OH radicals.[4]. Finally, emissions of DMS have a prominent role since it represents a precursor gas for cloud condensation nuclei, thereby affecting the radiation reflected back to space by clouds (the cloud albedo)[5,6]. As above anticipated, it has been reported that most of the production and consumption of these molecular gases take place in soil by means of microorganisms[1,7,8,9,10]. These small molecules play different roles in the metabolism of the microorganisms, e.g. they act as metabolites, co-metabolites or products, thanks to specific enzymes that catalyse their transformation in the presence of metal ion cofactors. Understanding the principles at the basis of the action of these metalloenzymes is crucial as they perform reactions that are critical from both a biological and an environmental point of view, and also for possible advancements of the chemical industry towards a biomimetic and eco-compatible future. Interestingly, enzymes containing molybdenum as an essential cofactor play a major role in such context, as better specified in the next sections

of the present thesis. They will provide a general overview on relevant Mo-based metalloenzymes, and a specific focus on the molybdoenzymes involved in atmospheric CO and H_2 oxidation, the latter being the main topics of the PhD research work here described.

1.1 Gas processing Mo-enzymes and their environmental role

Molybdenum is the only second-row transition metal required by most living organisms, and it is nearly universally distributed in biology[5]. Although rare on the Earth's crust, Mo is abundant in water where it is present as molybdate, MoO_4^{2-}. Such bioavailability and the remarkable chemical versatility exhibited by the metal, make Mo a crucial component for enzymatic catalysis. At present, over 50 molybdenum-containing enzymes, accomplishing widely diverse biological processes in plants, animals and microorganisms, have been purified and biochemically characterised[11]. According to the organometallic cofactor into which the molybdenum ion is incorporated, the corresponding enzyme can fall into two classes: nitrogenase, if the metal is incorporated into a unique [$MoFe_7S_9$] cluster, or the mononuclear molybdenum enzyme class, if Mo is bound to a pyranopterin-dithiolene ligand[12]. The role of the cofactor is to serve as a scaffold for a correct metal positioning, to modulate the redox potential and to facilitate the electron transfer to or from the metal active site[13]. The mononuclear molybdenum enzyme class can be further subdivided into three large and non-overlapping families, namely the xanthine oxidase (XO) family, the sulfite oxidase (SO) family, and the dimethylsulfoxide (DMSO) reductase family. Such families are distinguished by similarities in the structure of the Mo-containing active site. The xanthine oxidase family members have a $LMo^{VI}OS(OH)$ active site with a distorted square pyramidal coordination geometry[14]. The apical position is occupied by an oxygen atom which does not play a direct role in catalysis. The pyranopterin-dithiolene ligand, labelled as L, is found as molybdopterin cytosine dinucleotide (MCD) cofactor and it coordinates the Mo ion via a dithiolene linkage in one half of the equatorial plane. The remaining equatorial positions are occupied by a catalytically essential sulphur atom and by a catalytically labile –OH group. Exceptions to the $LMo^{VI}OS(OH)$ pattern are rare and they are represented by substitution of the Mo=S group by a Mo=Se one in nicotinate dehydrogenase, or by a Mo–S–Cu–S-Cys moiety in the case of carbon monoxide dehydrogenase (MoCu CODH)[15]. Enzymes belonging to this family catalyse hydroxylation and oxo-transfer reactions with water as the ultimate source of the oxygen atom incorporated into the product. The reducing equivalents generated in the course of reaction proceed from the molybdenum centre, through two [2Fe-2S] clusters and usually a flavin adenine nu-

cleotide (FAD) cofactor. The electron acceptors are generally represented by NAD^+, $NADP^+$, molecular oxygen or quinones.

Members of the sulfite oxidase family are characterised by an active site similar to the one of the XO family[11,12]. In analogy to the latter, it assumes a distorted square pyramidal geometry with an apical Mo=O group and a bidentate dithiolene ligand from the molybdopterin cofactor in the equatorial plane. The Mo–OH group of XO is substituted by a Mo=O one and a cysteine ligand replaces the equatorial sulphur atom and links the active site to the protein matrix. Reactions catalysed by the enzymes of this family are the addition or removal of oxygen to the substrate via a lone pair of electrons.

The dimethylsulfoxide reductase family is the most diverse molybdenum-containing enzyme family in terms of structure and catalysis[11,12]. All members present two equivalents of the bidentate pyranopterin cofactor and an overall $L_2Mo^{VI}Y(X)$ trigonal prismatic coordination geometry. The nature of the Mo=Y ligand of the oxidised form of the enzyme can be represented by O, S or even Se. Meanwhile, the X ligand can be found as a serine, a cysteine, a selenocysteine, an aspartate residue or an hydroxide. From a catalytic viewpoint, members of the DMSO reductase family mainly catalyse oxygen atom transfer reactions, but oxidation-reduction and hydroxylation/hydration reactions can also occur.

As above illustrated, molybdoenzymes are nearly ubiquitous on Earth and they are directly involved in the carbon, nitrogen, and sulphur cycles. In particular, three specific Mo-enzymes, mainly found in soil bacteria, contribute to biochemically essential reactions for regulation of atmospheric gases. Examples of this include fixation of atmospheric N_2, formation of dimethylsulfide and transformation of carbon monoxide.

N_2 is the most abundant gas in the atmosphere (78%) but the triple bond connecting the two N atoms of the molecule makes it an incredibly inert gas. In order to enter into most organisms metabolism and to maintain the effective nitrogen content in the soil, N_2 must be converted into its N "fixed" forms such as ammonia (NH_3) or nitrate (NO_3^-). In nature, the largest amount of nitrogen fixation is accomplished by nitrogenase enzymes that convert atmospheric nitrogen into ammonia[16]. Such reduction reaction is performed by two iron–sulfur clusters: the $MoFe_7S_9$ homocitrate iron-molybdenum cluster and the Fe_8S_7 also called P cluster, which are presumed to be the active site for substrate binding and reduction and the electron transfer mediator, respectively. N_2 fixation can alternatively occur through much more infrequent geochemical processes such as lightning. Industrially, reduction of N_2 to NH_3 is accomplished by the Haber-Bosch process, which, differently from the biological reaction, requires high temperature and pressure, as well as iron catalysts.

Dimethylsulfide (DMS) is a volatile compound mainly formed by enzymatic reaction,

starting from the water-soluble dimethylsulfoxide (DMSO) molecule. Organisms such as *Rhodobacter capsulatus* express a periplasmic DMSO reductase, characterised by a L_2MoO^{VI}(O-Ser) active site in the oxidized form, which reduce DMSO to the anti-greenhouse gas, DMS[11,17,18].

Carbon monoxide is a fatal gas for many living organisms and it is present at the trace level in the atmosphere where it constitutes an indirect greenhouse gas. Maintenance of sub-toxic levels of carbon monoxide in the troposphere is directly correlated to the activity of CO-oxidising bacteria in soil, accounting for the annual clearance of about 2×10^8 tons of CO[4,8,19,20,21]. Only one type of CO dehydrogenase (CODH) enzyme has been found to date that is able to catalyse the conversion of CO to CO_2 in aerobic conditions: the MoCu CODH[22].

The subject of this dissertation is to examine the mechanisms involved in the oxidation reactions carried out by MoCu CODH via different computational techniques.

1.2 MoCu CO dehydrogenase

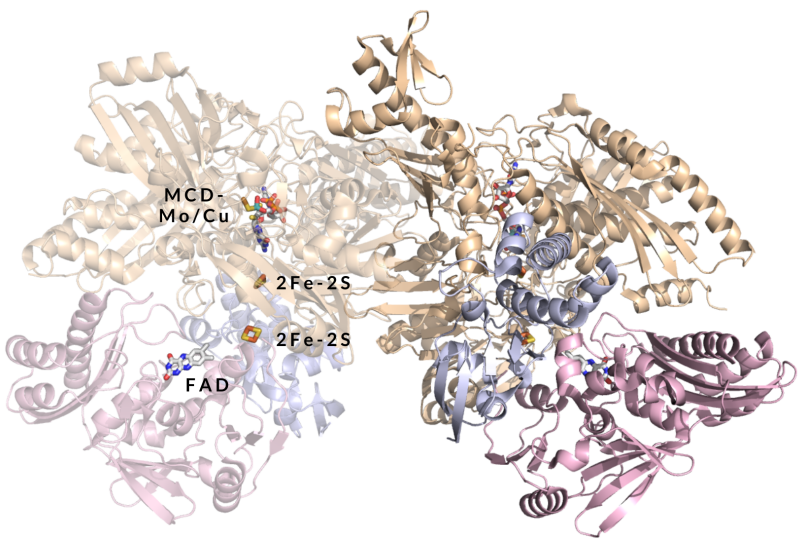


Figure 1.1: Representation of the $(\alpha\beta\gamma)_2$ heterohexamer (PDB: 1N5W). Each 89 kDa CoxL subunit (wheat) contains a Mo/Cu centre. Each 18 kDa CoxS subunit (blue) contains two [2Fe-2S] clusters. Each 30 kDa CoxM subunit (pink) contains a FAD cofactor.

The Mo/Cu-dependent CO dehydrogenase (MoCu CODH) enzyme is expressed by the aerobic soil bacteria *Oligotropha carboxidovorans* and it is a dimer of two heterotrimers with an overall $(\alpha\beta\gamma)_2$ structure (Figure 1.1). Three subunits are found in each protomer: a large subunit (CoxL; 89 kDa) containing the active site, a medium

subunit (CoxM, 30 kDa) containing the FAD cofactor and a small subunit (CoxS; 18 kDa) containing two [2Fe–2S] iron-sulphur clusters. The reducing equivalents deriving from the oxidation of the substrate pass on to the FAD via the two iron-sulphur clusters, and finally on to the quinone pool.

As previously mentioned, MoCu CODH belongs to the XO family of which it constitutes an exception because of its noncanonical binuclear active site (Figure 1.2). The latter is characterised by a MCD–MoO_2(μS)Cu bimetallic centre which switches between $Mo^{VI}Cu^{I}$ and $Mo^{IV}Cu^{I}$ redox states during catalysis. A cysteine residue (Cys388) binds the copper ion and connects the active site to the protein matrix. The protonation state of the equatorial oxygen ligand of Mo has been object of debate. In fact, experimental X-ray diffraction (XRD) results were interpreted as indicative of a Mo(=O)OH core both for the oxidised and the reduced form of the enzyme[22]. Differently, extended X-ray absorption fine structure (EXAFS) spectroscopy suggested a completely deprotonated equatorial Mo-oxo ligand for the oxidised state, Mo^{VI}(=O)O, and a completely reduced one for the reduced state, Mo^{IV}(=O)OH_2[23]. Recent data in support of the EXAFS-based hypothesis come from electron paramagnetic resonance (EPR) analysis of Mo^{V} analogue and from theoretical calculations[24,25].

In addition to its CO oxidising activity, it was reported that the MoCu CODH also exhibits hydrogenase activity, oxidising H_2 to protons and electrons[26,27]. Binding and oxidation of the substrate is believed to occur at the Cu^{I} centre[27,28]. The process is favoured by the presence of a highly delocalised redox-active orbital over the Mo–μS–Cu unit that allows a facile electron transfer to the Mo centre[29].

Crucial residues belonging to the active-site second coordination sphere are conserved in MoCu CODH and homologous of the XO family[11]. The highly conserved glutamate residue, Glu763, located in proximity to the equatorial oxygen ligand of Mo, is considered to act as a base to facilitate deprotonation events[27]. By comparison, the aromatic ring of a phenylalanine residue, Phe390, locate in a flexible loop in front of the Cu ion, is believed to play a role in the correct positioning of the substrates within the active site[25] and to allow the access of rather bulky molecules – e.g., *n*-butylisocyanide, L-cysteine, coenzyme A or glutathione – representing reversvible and irreversible enzyme inhibitors[22,30].

1.2.1 CO and H_2 oxidation catalytic mechanisms

Mechanistic mode characterising the biotransformation of CO to CO_2 in MoCu CODHs still represents a controversial theme. Important contributions in this debate has been reported both by experimental and theoretical groups over the last two

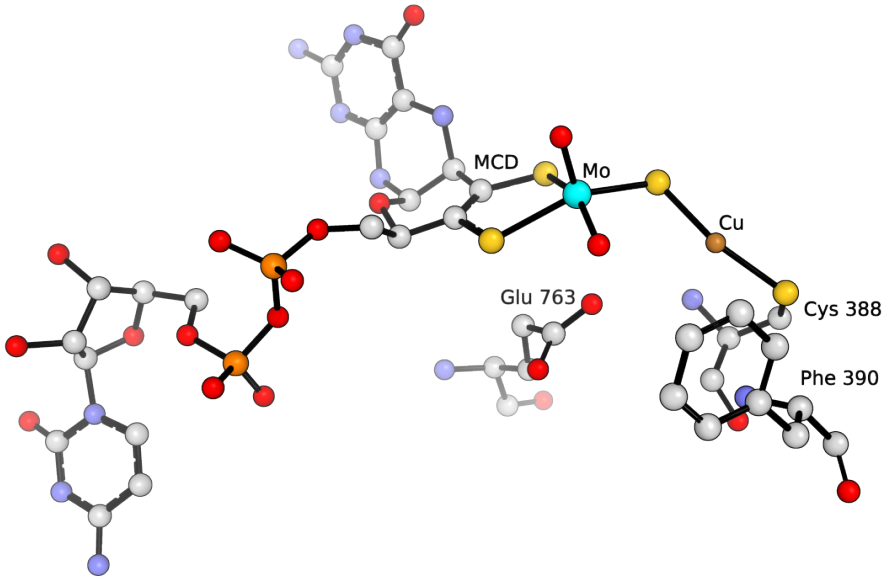


Figure 1.2: Active site structure of MoCu CO dehydrogenase from *Oligotropha carboxidovorans* (PDB: 1N5W). The Mo and Cu atoms are bridged by a μS, and the Cu is bound to the protein via the Cys388 thioether linkage. Crucial conserved residues, Glu763 and Phe390, are also shown in figure.

decades[22,25,28,31,32,33,34,35]. The first mechanism was hypothesised in 2002 by Dobbek *et al.*, based on 1.1 Å resolution inhibitor-bound crystal structure (Figure 1.3**A**)[22]. According to this proposal, after CO coordination to the enzyme resting state, the substrate would integrate between the bridging sulphur and copper but still bound to the Mo equatorial oxygen forming a thiocarbonate derivative. The structure of the latter is analogous to the thiocarbamate intermediate formed during inhibition of the enzyme by *n*-butylisocyanide. Upon hydrolysis by water solvent, the oxygenic molybdenum ligand would be regenerated and the CO_2 product would be released. Finally, completion of the reaction cycle occurs upon reoxidation of Mo^{VI} to Mo^{IV} via transfer of two electrons.

Starting from this proposal, all subsequent proposed mechanisms deal with the possible occurrence and the role of the thiocarbonate catalytic intermediate. In 2005, inspired by Dobbek proposal, two independent theoretical works were published (Figure 1.3**B** and 1.3**C**)[31,32]. Both groups employed similar small cluster models for the description of the enzyme active site and evidenced that the thiocarbonate adduct would occupy a deep well in the energy profile of the investigated reaction mechanisms. Consequently, Hofmann *et al.* proposed that the thiocarbonate intermediate corresponds to a dead-end, the occurrence of which seems not to be functional for catalysis. Alternatively, Siegbahn and co-workers integrated the thermodynamically stable thiocarbonate intermediate in their proposed CO oxidation catalytic cycle, sug-

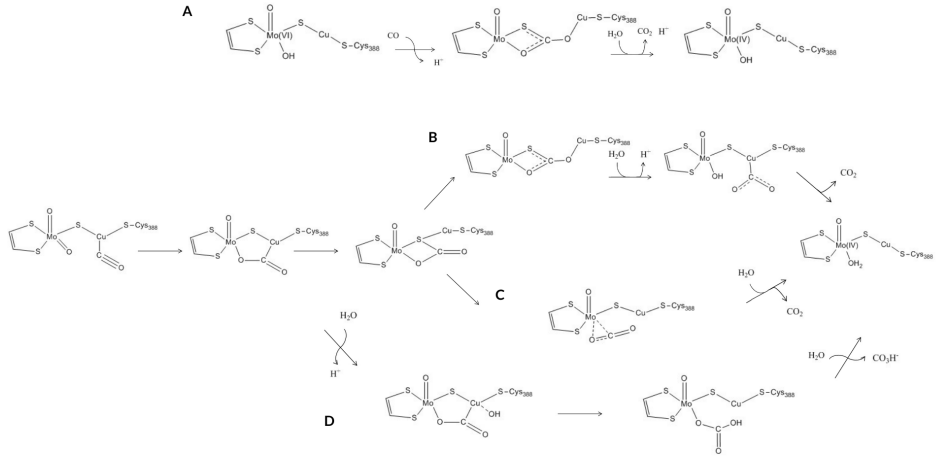


Figure 1.3: Proposed mechanisms of carbon monoxide oxidation by MoCu CO dehydrogenase. **A** shows the reaction mechanism based on the structure of *n*-butylisocyanide-inhibited enzyme proposed by Dobbek's experimental group. **B** and **C** show reaction mechanisms involving the formation of a thiocarbonate species, predicted by computational work of Siegbahn and Shestakov, and by Hofmann *et al*, respectively. **D** shows a reaction mechanism proposed by Stein and Kirk that sees bicarbonate, not CO_2, as the product of the reaction.

gesting that the SCO-bond of the thiocarbonate can be broken upon insertion of a water molecule in the first coordination shell of molybdenum. However, such step was reported not to be facile and the barrier for the release of CO_2 was estimated to be rather high. More recent theoretical studies based on much larger models evidenced that the thiocarbonate species still behave as a thermodynamic sink along the catalytic energy profile, although not as stable as previously suggested[34]. Anyway, its formation still remains problematic for the CO-oxidation catalysis since the authors were not able to locate a transition state for CO_2 releasing directly from the thiocarbonate intermediate.

Recent experimental ^{13}C and $^{63,65}Cu$ ENDOR studies suggested the possibility that the thiocarbonate adduct and its precursor intermediates might be better thought as readily interconverted species[28].

In order to avoid formation of the controversial thiocarbonate species, Stein and Kirk hypothesised a new different mechanism in which a bicarbonate product can be formed after a nucleophilic attack of a water molecule on μ_2-η^2-CO_2 intermediate (Figure 1.3**D**)[33]. Unfortunately, experimental EPR studies of the partially reduced binuclear centre in complex with bicarbonate showed that bicarbonate coordination to Mo is unlikely to happen during catalysis[36].

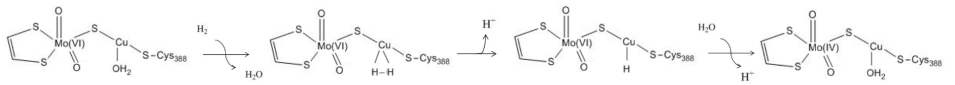


Figure 1.4: Mechanisms of H_2 oxidation exerted by MoCu CO dehydrogenase proposed by Wilcoxen and Hille.

The previously proposed mechanisms share the idea that the equatorial oxo-ligand of Mo is responsible for a direct oxygen transfer to CO performing a nucleophilic attack to the activated carbon monoxide bound to copper. However, a consensus on this key issue has not been reached, also due to the lack of experimental evidences in support of such scenario. In this regard, Hille and coworkers have recently hypothesised that water might conveniently play the role of nucleophile, in place of the Mo-bound oxide[15]. Such a role for water would be favoured by the presence of basic groups that might facilitate its deprotonation upon CO-oxidation.

Concerning the hydrogenase activity of MoCu CODH, the only proposed reaction mechanism for the H_2 splitting was reported by Wilcoxen and Hille based on electron paramagnetic resonance (EPR) spectroscopy studies (Figure 1.4)[27]. According to their proposal, H_2 is believed to coordinate Cu^I in a side-on way. Coordination to metal would activate the dihydrogen molecule and predispose it for splitting by extraction of a proton thanks to a nearby base. Subsequently, a second proton would leave the active site and the two electrons would be transferred to the Mo centre. Only one previous theoretical work was published concerning the hydrogenase activity of MoCu CODH[35]. The latter employed small cluster models for the active site representation and reported the unexpected inhability of the Cu ion in binding the H_2 substrate. In fact, in order to locate a H_2–enzyme adduct, the authors needed to introduce changes in the active site such protonation of the sulphur atom of the cystein ligand bound to Cu.

2 Methods

In this chapter, I give a basic introduction to the theoretical methods used in the thesis.

2.1 Quantum Mechanics

To carry out a reliable theoretical study of metalloproteins and the reactions they catalyse, a good electronic structure description method is required in order to model the metal coordination sphere and the chemical bond breaking and forming. Quantum mechanics (QM) is the only methodological approach that allows for a detailed description of the electron distribution. Within the QM framework, a system is described by the Schrödinger equation, that in its time-independent form is

$$\hat{H}\Psi(R,r) = E\Psi(R,r) \tag{2.1}$$

where the term $\Psi(R,r)$ represents the wave function that depends on all electron (r) and nuclear (R) positions of the system, E is the energetic numerical value of the described state and $\hat{H}$ is the Hamiltonian operator. By knowing the wave function, it is possible to completely describe the system by means of the probability function given as the square of the wave function, $|\Psi|^2$. The Hamiltonian operator is defined

as a sum of kinetic ($\hat{T}$) and potential ($\hat{V}$) operators for all particles

$$\begin{aligned}\hat{H} &= \hat{T}_e + \hat{T}_n + \hat{V}_{en} + \hat{V}_{ee} + \hat{V}_{nn} \\ &= -\sum_{i=1}^{N} \frac{\hbar^2}{2m_e}\nabla_i^2 - \sum_{A=1}^{M} \frac{\hbar^2}{2M_A}\nabla_A^2 - \sum_{i=1}^{N}\sum_{A=1}^{M} \frac{e^2 Z_A}{r_{iA}} + \\ &+ \frac{1}{2}\sum_{i=1}^{N}\sum_{j>1}^{N} \frac{Z_A}{r_{ij}} + \frac{1}{2}\sum_{A=1}^{M}\sum_{B>A}^{M} \frac{Z e_A^2 Z_B}{R_{AB}}\end{aligned} \tag{2.2}$$

where i and j run over the electrons and A and B over the nuclei. m_e and M_A are the mass of the electron and of nucleus and Z_A is the atomic number whereas r_{ij} and R_{AB} denote the interparticular distances. The term ∇^2 is the Laplacian operator, the sum of the second derivatives with respect to the coordinates, that in Cartesian coordinates takes the following form

$$\nabla^2 = \frac{\partial}{\partial x^2} + \frac{\partial}{\partial y^2} + \frac{\partial}{\partial z^2} \tag{2.3}$$

The five terms composing the Hamiltonian in equation 2.2 are operators representing, in order, the kinetic energy of the electrons ($\hat{T}_e$) and nuclei ($\hat{T}_n$), the attractive electron-nuclei potential ($\hat{V}_{en}$), the repulsive potential between the electrons ($\hat{V}_{ee}$) and between the nuclei ($\hat{V}_{nn}$).

The time-independent Schrödinger equation (2.1) can be solved exactly only in the case of hydrogen, of the H_2^+ molecule and of similar one-electron systems. Therefore, the introduction of approximations and simplifications is necessary to study more complex systems. Among the fundamental approximations that were employed in all calculations of this thesis, we recall the *adiabatic* approximation and the Born-Oppenheimer one. The first consists in neglecting the coupling between different electronic states so that the total wave function is restricted to one electronic surface. The Born-Oppenheimer approximation consists in the separation of the motion of the nuclei and the electrons, because of the large difference of electron and nuclear masses that reflects on the different velocities of the two. Consequently, the nuclei are much slower that the electrons which can therefore adjust instantaneously to a change in the nuclei position. Within this approximation, the nuclear kinetic energy ($\hat{T}_{en}$ in equation 2.2) can be omitted and the repulsive nuclear-nuclear potential energy term ($\hat{V}_{nn}$) is considered as a constant.

To solve the polyelectronic wave function, a further simplification consists in introducing the *orbital approximation* which assumes the independent motion of the single

electrons. The total wave function can therefore be factorised into a product of one-electron wave functions ($\phi_i(r_j, s_j)$), namely *molecular orbitals* (MOs) or *spin-orbitals*:

$$\Psi(r_1, s_1, r_2, s_2, \ldots, r_N, s_N) = \phi_i(r_1, s_1)\phi_j(r_2, s_2)\ldots\phi_k(r_N, s_N) \tag{2.4}$$

To each electron of the system a single function is assigned and it only depends on the spatial and spin coordinates of such electron. Electrons are indistinguishable particles with an associated spin of $\frac{1}{2}$. According to the *Pauli principle*, two electrons cannot be described by the same quantum numbers, therefore it requires the total wave function to be antisymmetric, i.e. it must change sign when interchanging two electrons. In the case of a polyelectronic wave function, the condition imposed by the principle of antisymmetry is achieved if the wave function is expressed in a determinant form, as a Slater determinant (SD), which in the case of N electrons and N spin-orbitals takes the form of:

$$\Phi_{SD} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(1) & \phi_2(1) & \ldots & \phi_N(1) \\ \phi_1(2) & \phi_2(2) & \ldots & \phi_N(2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(N) & \phi_2(N) & \ldots & \phi_N(N) \end{vmatrix} \tag{2.5}$$

MOs are given as the product of a spatial orbital ($\psi(r)$) and a spin function

$$\phi = \psi(r) \begin{cases} \alpha(\omega) \\ \beta(\omega) \end{cases} \tag{2.6}$$

and they satisfy the orthonormality condition

$$\langle \phi_i | \phi_j \rangle = \delta_{ij} \begin{cases} 1, i = j \\ 0, i \neq j \end{cases} \tag{2.7}$$

An additional simplification for the solution of the Schrödinger equation (2.1) consists in the description of the trial wave function by means of a single Slater determinant (2.5). Doing so, electron correlation is neglected because each electron moves in the average potential of all the other electrons, i.e the electron–electron repulsion is only included as an average effect.

2.1.1 The Hartree-Fock theory

In the case of the restricted Hartree-Fock closed-shell systems (even number of electrons in a singlet state), the spatial function $\psi(r)$ in equation 2.6 is assumed to be the same for the α and β electrons in the same orbital. The Hamiltonian operator is substituted by the Fock operator, which is given by

$$\hat{F}_i = \hat{h}_i + \sum_{j=1}^{N/2}(2\hat{J}_j - \hat{K}_j) \tag{2.8}$$

where the electron repulsion to all the other electrons is described via the Coulomb ($\hat{J}_j$) and the exchange ($\hat{K}_j$) operators, while the monoelectronic core operator ($\hat{h}_i$), defined as

$$\hat{h}_i = -\frac{1}{2}\nabla_i^2 - \sum_{A}^{N}\frac{Z_A}{r_{Ai}} \tag{2.9}$$

describes the kinetic energy of electron i and the attraction to all the nuclei.

By solving the eigenvalue equations

$$\hat{F}_i\phi_i = \epsilon_i\phi_i, \qquad i = 1, 2, \ldots, N \tag{2.10}$$

we obtain a set of one-electron wave functions (ϕ_i), called the Hartree–Fock molecular orbitals with associated one-electron energy (ϵ_i). The total energy of the system (E) does not correspond to the sum of the energies of the occupied orbitals because in that case the electronic repulsion would be counted twice. Instead, it can be expressed as

$$E = \sum_{i=1}^{N} 2h_{ii} + \sum_{i=1}^{N}\sum_{j=1}^{N}(2\hat{J}_{ij} - \hat{K}_{ij}) = \sum_{i=1}^{N}(\epsilon_i + h_{ii}) \tag{2.11}$$

The best set of monoelectronic wave functions is obtained by minimizing the energy with respect to the spin-orbitals, according to the *variational theorem*, which states that an approximate wave function has an energy that is higher or equal to the exact energy.

At this point, another approximation must be introduced which concerns the explicit analytical expression of the spatial part of the monoelectronic function. The molecular

orbitals are expanded in a basis set, a linear combination of functions whose analytical form are in general atomic orbitals χ_α.

$$\phi_i = \sum_{\alpha}^{\infty} c_{\alpha,i}\chi_\alpha \tag{2.12}$$

The best energy value resulting from the HF approximation, the *Hartree-Fock limit*, comes from considering a complete basis set that is an infinite number of basis functions in the MO linear combination. In the computational practise the summation must be truncated to a finite number M of basis functions, as reported in eq. 2.13. Evidently, as the size of the basis set is increased, the variational principle ensures that the energetic results become better.

$$\phi_i = \sum_{\alpha}^{M} c_{\alpha,i}\chi_\alpha \tag{2.13}$$

Thanks to this approximation, the HF equations can be collected in a matrix notation, the so-called *Roothan-Hall equation*

$$\mathbf{FC} = \mathbf{SC}\epsilon \tag{2.14}$$

where the $\mathbf{F}$ matrix contains the Fock matrix elements, the $\mathbf{C}$ matrix is composed of the coefficients of the molecular orbitals, the $\mathbf{S}$ matrix contains the overlap elements between basis functions and ϵ is a diagonal matrix formed by n eigenvalues corresponding to the energies of the n molecular orbitals.

2.1.2 Basis sets

Since only a limited number of basis functions are used, the size and quality of the basis set are important for the accuracy of QM calculations. There are two types of basis functions commonly used in electronic structure calculations: *Slater Type Orbital* (STO)

$$\chi_{\zeta,n,l,m}(r,\theta,\varphi) = NY_{l,m}(\theta,\varphi)r^{n-1}e^{-\zeta r} \tag{2.15}$$

and *Gaussian Type Orbitals* (GTO)

$$\chi_{\zeta,n,l,m}(r,\theta,\varphi) = NY_{l,m}(\theta,\varphi)r^{2n-2-l}e^{-\zeta r^2} \tag{2.16}$$

In both equations, n, l and m are the quantum numbers, N is a normalisation constant, $Y_{l,m}$ are the spherical harmonic functions, r is the distance between electron and nuclei and ζ is a predetermined constant determining the spread of the orbitals.

GTOs, in contrast to STOs, do not correctly represent the orbital shape. In particular, the difference is large when describing the proper behaviour near the nucleus, i.e. GTO has a zero slope, in contrast to a STO which has a cusp, and at large distances where GTO falls off too rapidly far from the nucleus compared with an STO. More GTOs are therefore necessary for achieving the same accuracy but the increase in the number of gaussian-type functions is compensated by the ease with which the required integrals can be calculated. Because of the achieved computational efficiency, GTOs are normally preferred in the great majority of QM softwares.

The smallest number of functions employable for a minimum description of the occupied orbitals, i.e. one basis function per electron pair of the neutral atom(s), is a *minimum basis set* or *single zeta* (SZ) basis set. An improvement of the basis sets is achieved by doubling of all basis functions, producing a *double zeta* (DZ) type basis. The next step up in basis set size is a *triple zeta* (TZ). In general, it is usually enough to describe the core orbitals by a minimum set of basis function while expanding only the number of basis functions of valence orbitals, producing a *split valence* basis, e.g. VDZ or VTZ. Additional basis functions with higher angular momentum, called *polarisation functions* can be added to treat charge polarisation and electron correlation.

2.1.3 Density Functional Theory

Density functional theory is a powerful tool in the computational study of chemical systems, based on the one-to-one correspondence between the electron density and the energy of a system. The electron density $\rho(\mathbf{r})$ is defined as the square of the wave function, integrated over $N-1$ electron coordinates and each spin coordinates.

$$\rho(\mathbf{r}) = N\int d\mathbf{r}_2 \cdots \int d\mathbf{r}_N |\Psi(\mathbf{r},\mathbf{r}_2,\ldots,\mathbf{r}_N)|^2 \tag{2.17}$$

The advantage of this approach derives from the simplification in the variables used to derive the energy of the polyelectronic system. In fact, while the electron function

depends on $4N$ coordinates, the electron density is a function of only the threes spatial coordinates.

The success of modern DFT methods is based on the *Kohn-Sham* (KS) formulation which suggests that the energy can be calculated from an auxiliary set of one-electron orbitals used for representing the electron density.

$$\rho = \sum_{i=0}^{N_{elec}} |\phi_i|^2 \tag{2.18}$$

The re-introduction of the concept of orbitals leads to an increase of the complexity of the problem, from 3 to $3N$ variables. The electron density is computed solving the N KS mono-electronic equations (2.19), whose operator is itself function of the electron density, therefore these equations must be solved in an iterative manner.

$$\hat{F}_{KS}(i)\phi_{u,KS}(i) = \epsilon_{u,KS}\phi_{u,KS}(i) \qquad u = 1, \ldots, n \tag{2.19}$$

The KS-SCF method is a variational method in the sense that the best obtained orbitals are those that generate an electron density that minimise the energy, $\epsilon[\rho]$.

The energy ground state E_0 is a function of an external potential that can be separated into various contributions, which are in turn functions of the electron density:

$$E_0[\rho] = E_v[\rho] = T_S[\rho] + J[\rho] + E_{ne}[\rho] + E_{XC}[\rho] \tag{2.20}$$

where $T_s[\rho]$ represents the kinetic energy of the non-interacting electrons, $J[\rho]$ the Coulomb electron-electron interaction, $E_{ne}[\rho]$ the nucleus-electron interaction energy and $E_{XC}[\rho]$ the exchange-correlation functional. The latter term considers both the exchange effects, due to the interplay between electrons with the same spin, and the correlation effects, related to interactions between same-spin electrons. The exact form of E_{XC} is not known exactly, therefore the main development in DFT methods consists in finding a proper approximation for this term.

In the so called *Local Density Approximation* methods (LDA), E_{XC} depends only on the local value of the electron density. Therefore, it is assumed that ρ can be treated as a uniform electron gas or a slightly varying function. The observation of the electron density trend of a molecule led to the formulation of functionals that depend not only on the electron density but also on derivatives of it. This approach is defined as the *Generalized Gradient Approximation* (GGA). An example of GGA fuctional empoyed in this thesis is BP86[37,38].

If the exchange-correlation functional includes a fraction of the Hartree-Fock exchange, we refer to a *Hybrid-GGA* method. For the hybrid functional B3LYP[39,40] used in this thesis, the E_{XC} term is defined as

$$E_{XC}^{B3LYP} = (1-a)E_x^{LSDA} + aE_x exact + b\Delta E_x^{B88} + (1-c)E_c^{LSDA} + cE_c^{LYP} \tag{2.21}$$

with a, b and c being parameters defined to fit experimental data with an associated value of 0.20, 0.72 and 0.81, respectively.

A newest class of DFT functionals named *Hybrid-Meta-GGA* methods, expresses the exchange-correlation functional as a function of the exact exchange, the electron density, its gradient and the kinetic energy density.

2.2 Solvent Effects

Evaluating the effect of the environment is an important aspect of computational chemistry as its presence may change the charge distribution or the conformational preference of a solute molecule. The long-range or polarisation effects induced by the protein environment in enzymatic catalysis can be modelled as a solvent. Computational treatment of solvent effect can be executed by considering all the individual solvent molecules, by describing the solvent as a continuous medium or by a combination of the two approaches.

In the case of *Continuum Solvation* models, the solvent is described as a uniform polarisable medium defined by a dielectric constant ϵ with the solute placed in a suitable cavity. The solvation model used in this thesis is the *COnductor-like Screening MOdel* (COSMO) that employs molecular shaped cavities that are typically assembled as a set of atom-centered spheres with radii approximately 20% larger than the Van der Waals radius.

2.3 Molecular Mechanics

Molecular Mechanics (MM) method describes molecules by classical mechanics allowing fast energy and properties computations of rather big molecular systems.
Within this description, electrons are not treated explicitly but each atom is simulated as a particle to which a radius is assigned, as well as a polarizability and a constant net

charge, whereas molecular bonds are described as springs having different length and stiffness.

Potential energy is calculated using a set of parameters, called *force field*, deriving both from experimental data and from high-level *ab initio* calculations. The term *force field* also refers to the functional form used to compute the energy. The quality of a force field calculation depends on the appropriateness of the mathematical functions used in the energy expression and on the goodness of the parameters. The general form of total energy is a sum of covalent and non-covalent contributions,

$$E_{FF} = E_{str} + E_{bend} + E_{tors} + E_{vdw} + E_{el} + E_{cross} \tag{2.22}$$

where E_{str} is the energy function for stretching a bond between two atoms, E_{bend} describes the energy involved in bending an angle, E_{tors} is the torsional energy for rotation around a bond, E_{vdw} and E_{el} represent the non-bonded energy, the attraction or repulsion between atoms that are not directly bonded. Finally, the E_{cross} term arises from coupling two or more of the bonded terms.

The most used bond energy (E_{str}) functional form derives from describing it as an harmonic oscillator, that written in a Taylor expansion around an "equilibrium" bond length r_0 becomes

$$E_{str} = \sum_{bonds} K_r(r - r_0)^2 \tag{2.23}$$

where K_r is the bond force constant. Likewise, the bending energy (E_{bend}) becomes

$$E_{bend} = \sum_{angles} k_0(\Theta - \Theta_0)^2 \tag{2.24}$$

where k_0 and Θ_0 represent the angle force constant and the reference angle, respectively. The torsional term (E_{tors}) describes part of the energy change associated with the rotation around a bond, therefore a periodic function is used, generally expressed as a Fourier series

$$E_{tors} = \sum_{n=1} V_n cos(n\omega) \tag{2.25}$$

The term $n = 1, 2, 3$ describes a rotation that is periodic around $360°$, $180°$ or $120°$, respectively; while the V_n constant shapes the size of the barrier for the rotation around

the bond. E_{vdw} describes the Van der Waals energy, deriving from the interaction of not-directly bonded atoms. This term is generally described as the sum of two contribution, a repulsive one reflecting the nuclear-nuclear repulsion, the electron-electron repulsion and the Pauli principle at very short distances, and an attractive one which arises from instantaneous correlation between the electronic motions in two non-bonded atoms, often referred to as *dispersion* or *London* force. The Lennard-Jones potential

$$E_{vdw} = \epsilon(\frac{r_0}{r^{12}} - 2\frac{r_0}{r^6}) \tag{2.26}$$

is one of the most popular function used to describe E_{vdw}. Here r_0 is the minimum energy distance and ϵ the depth of the minimum. The electrostatic interaction energy (E_{el}) is expressed by the Coulomb potential as

$$E_{el} = \sum_{i,j} \frac{q_i q_j}{\epsilon r_{ij}} \tag{2.27}$$

where q is the atomic charge, r the interatomic distance and ϵ the dielectric constant.

In this thesis, we employed the AMBER ff14SB force field[41] for the protein. Partial charges for new molecules were obtained with the restrained electrostatic potential (RESP) method[42], based on QM calculations.

2.4 Hybrid QM/MM Methods

Combined quantum mechanics and molecular mechanics (QM/MM) is a widely used method created to integrate the accuracy of QM calculations to the speed of MM calculations. This approach is of particular interest in the modelling of metalloproteins. In fact, in QM/MM, electrons belonging to metals can be treated explicitly but at the same time the entire macromolecule can be included in the calculation, allowing for a comprehensive study of the surrounding effects. In the case of enzymes, a small region of interest (among 20 to 300 atoms) which is commonly represented by the active-site is treated at the QM level whereas the remaining part of the protein, as well as an ample number of explicit solvent molecules, is described at MM level. A wide variety of QM/MM approaches exist which differ in the scheme used for the computation of the total energy, in the partition between the QM and MM systems and in the treatment of the interaction between QM and MM regions.

For sake of clarity, I will only describe the approach that was employed in the present thesis. Within the subtractive scheme approach, the total QM/MM energy can be expressed as

$$E_{QM/MM} = E_1^{QM} + E_{12}^{MM} - E_1^{MM} \tag{2.28}$$

where E_1^{QM} is the result of the QM calculation concerning the QM region (System 1) while the other terms derive form two MM calculations, one for the entire model (System 1 and System 2, E_{12}^{MM}) and one for System 1 (E_1^{MM}). According to the hydrogen-link atom approach, the QM system covalently connected to the MM region was truncated with hydrogen atoms (HL), whose position is linearly related to the corresponding carbon atoms (CL) in the whole system. The electrostatic QM–MM interaction was treated as defined in the electrostatic embedding procedure, including a point-charge model of System 2 in the QM calculations. The complete employed energy function is

$$E_{\mathrm{QM/MM}} = E^{\mathrm{HL}}_{\mathrm{QM1+ptch2}} + E^{\mathrm{CL}}_{\mathrm{MM12},q_1=0} - E^{\mathrm{HL}}_{\mathrm{MM1},q_1=0} \tag{2.29}$$

where $\mathrm{E}^{\mathrm{HL}}_{\mathrm{QM1+ptch2}}$ is the QM energy of System 1 truncated by HL atoms and embedded in the set of point charges modelling Systems 2. $\mathrm{E}^{\mathrm{HL}}_{\mathrm{MM1},q_1=0}$ is the MM energy of System 1, still truncated by HL atoms, without any electrostatic interactions. $\mathrm{E}^{\mathrm{CL}}_{\mathrm{MM12},q_1=0}$ is the classical energy of the whole system, with CL atoms and with the charges in System 1 set to zero, in order to avoid double counting of the electrostatic interactions.

2.5 Big-QM Approach

One of the limitations of both QM-cluster and QM/MM methods is that energies strictly depend on the size of the selected QM system. Based on the knowledge that QM/MM calculations converge faster than QM-only ones, but models of about 1000 QM atoms are needed to obtain convergence of the energies, the Big-QM approach was developed to obtain more accurate energies[43,44]. In the Big-QM calculations, the QM system (System 1) of the QM/MM optimizations is extended with all chemical groups with at least one atom within 6.0 Å of it and junctions are moved three amino-acids away from each residue in the minimal QM system. In addition, all buried charges inside the protein are included to create the Big-QM model. This typically gives a QM system of 600-1200 atoms.

3 | Summary of the Articles

This thesis is based on five publications, Papers I–V, concerning the MoCu CO dehydrogenase enzyme. Paper I is a short review focused on the analysis of the challenges posed by this system when describing it via a computational approach. Papers II and V focus the CO oxidation mechanism whereas Papers III and IV concern the catalysis of H_2 carried out by the enzyme.

3.1 Paper I

Previous knowledge concerning the CO oxidation catalytic cycle of MoCu CODH was compiled in paper I. In particular, we focused on the controversial aspects aroused from the previously published theoretical studies of the enzyme. We compared the suggested catalytic mechanisms, the models and the levels of theory chosen for the enzyme description of each work. We highlighted the challenges in the computational study of MoCu CODH, also in relation to other oxygen-atom transfer reactions accomplished by Mo or W complexes. Finally, we proposed different alternatives for the improvement of the *in silico* description of the enzyme.

3.2 Paper II

In the catalysis of CO carried out by MoCu CODH, a rather controversial aspect concerns the plausible formation of a thiocarbonate intermediate, first suggested by the crystallographer Dobbek and his group[22]. Such intermediate would be analogous to a thiocarbamate one evidenced during inhibition of the enzyme by means of *n*-butylisocyanide. Subsequent theoretical studies inspired by this proposal reported a surprisingly high stability, correlated with a low formation barrier, of the thiocarbonate intermediate[31,32,34].

The aim of paper II was to deepen the understanding of the role played by the thiocarbonate intermediate in the CO oxidation catalysis. We did so by using a hierarchy of QM-cluster models of the active site and QM/MM models of the whole active-site containing enzyme subunit.

Thanks to this study, we demonstrated that the size of the model and the chosen level of theory are fundamental aspects affecting the quality of the resulting energy of the thiocarbonate adduct. Our results indicate that the latter does not represent an energy sink on the catalytic energy profile when dispersion corrections are taken into account, and at the same time appropriate basis sets are used and an accurate modelling of the protein surroundings the active site is considered.

3.3 Paper III

Apart from CO oxidation to CO_2, MoCu CODH also expresses hydrogenase activity, as it is able to oxidise H_2 to protons and electrons. Only one previous computational study on a small QM-cluster model had been published before, and in such publication it was evidenced that changes in the MoCu CODH active site were necessary for the binding of H_2 to Cu[35], differently from what was previously hypothesised by the experimental group of Wilcoxen and Hille[27].

Paper III is the first published theoretical work concerning the H_2 oxidation catalytic mechanism of MoCu CODH. By means of the QM/MM approach, we were able to suggest two plausible mechanisms for the hydrogen molecule splitting. In the first case, the protein active site should be viewed as a Mo^{VI}=O/Cu^{I} Frustrated Lewis Pairs (FLP), where the Cu^{I} act as a Lewis acid and the equatorial oxo ligand of Mo act as a Lewis base toward the oxidation of H_2. The second mechanism involves the protonation of the sulphur atom of the Cu-bound cysteine residue as a favourable event for the binding of the substrate to the copper centre. We suggest that the two reactive channels, characterised by similar energy profiles, may be active in the enzyme thus enhancing the overall catalytic turnover. Our results indicate that the equatorial oxo ligand of Mo represents the strongest base of the active site and therefore it proved to be the most favourable site for the first deprotonation event of H_2 oxidation. This would lead to the formation of a Mo-OH–Cu(I)-H hydride with a corresponding intramolecular hydrogen bond that would stabilise the copper hydride, otherwise known to be very unstable.

3.4 Paper IV

Based on the results obtained in paper II, we knew that BigQM-based refinement of computed energies represents a valuable approach in the theoretical investigation of MoCu CODH. Therefore, in paper IV we refined the QM/MM energies for the intermediates involved in the dihydrogen oxidation catalysis reported in paper III by means of the BigQM approach.

An unexpected large difference in relative energies between the QM/MM and the BigQM results was found which lead to the observation of a sub-optimal description of the H_2–HN(backbone) interaction when the van der Waals parameters described in previous literature for H_2 were employed. Therefore, we developed a new set of van der Waals parameters in order to better describe the hydrogen–backbone interaction. We used a small [H_2 + $CH_3CONHCH_3$] model to determine the potential-energy surface for the H_2–HN interaction with QM methods (B3LYP-D3(BJ)/def2-TZVPD) and to subsequently determine van der Waals parameters to reproduce it.

The new van der Waals set we propose in this article turned out to be the best one in reproducing the H_2–HN(backbone) interaction, even with respect to parameters previously described in literature.

Thanks to this work, we were able to exclude the occurrence of side-on H_2 binding at the copper centre when the enzyme is in its resting state. Such reparameterization constitutes a useful tool also in the context of future QM/MM and MD modelling of enzymes which show hydrogenase activity.

Moreover, we evidenced that BigQM is a good approach to localise errors in the MM parameters and to ensure that the resulting energies are accurate.

3.5 Paper V

In analogy to Mo-containing metalloenzymes that catalyse oxygen-atom transfer reactions, MoCu CODH is believed to oxide CO by means of its Mo-oxo ligand. In this context, all previously proposed catalytic mechanisms for the oxidation of CO share the idea that the Mo-bound equatorial oxo ligand performs a nucleophilic attack on the activated CO substrate.

In paper V, we analyse the possibility, recently hypothesised by Hille and co-workers [15], that an activated water molecule present in the active site would play the role of nucleophile. We propose a new catalytic mechanism using a large QM-cluster model, previously validated by Rokhsana and coworkers [25]. Furthermore, the energetics involved in the activation barrier for the formation of the new C–O bond were evaluated by accounting for different protonation states of the active site as well as by explicit consideration of the whole protein environment by means of QM/MM followed by energy refinement via the BigQM technique.

Our results indicate that such alternative mechanism might constitute one of the reactive channels for the oxidation of CO by the enzyme even thought further QM/MM studies concerning all the states involved in the QM-cluster based catalytic cycle will be needed to confirm such a hypothesis.

4 Conclusions and Outlook

Soil bacteria contribute to control the atmosphere composition by processing or synthesising gases such as CH_4, CO, N_2, NO, NO_2, DMS and H_2. Gas transformations are performed thanks to the presence of metalloenzymes that catalyse specific redox-reactions. 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.

MoCu CO dehydrogenases, found in the aerobic soil bacteria *Oligotropha carboxidovorans*, exert the ability of oxidise CO to CO_2, enabling the microorganism to grow using CO as its sole source of carbon and energy. Oxidation of H_2 to protons and electrons can also constitute an alternative energy source. The crystal structure of MoCu CODH, characterised by a unique Mo/Cu bimetallic active site, was reported at the early of the 2000s. Nonetheless, an open debate still concerns key steps in the oxidation mechanisms of both CO and H_2. In this context, Density Functional Theory (DFT) based theoretical methods employed in this thesis proved to be powerful tools in providing structures and energetics of transition states and reaction intermediates and supplying insights into the factors that govern the enzymatic catalysis. Models of different sizes, ranging from the minimal metal clusters to very large systems, including the whole enzyme subunit containing the Mo/Cu bimetallic active site, have been developed, to study the essentials of the chemistry of this enzyme. Improvements of the enzyme *in silico* description with respect to previously published theoretical works turned out to be fundamental since such peculiar system evidenced a delicate dependence of the catalytic energy profile on the chosen level of theory. In light of these considerations, we have taken some steps towards the understanding of reaction mechanisms catalysed by MoCu CODH.

Concerning the CO oxidation catalysis, we showed that the possible occurrence of a thiocarbonate catalytic species does not represent an obstacle for catalysis as it was shown not to be a rate-limiting intermediate on the enzymatic energy landscape, differently from what it was reported in all previous computational studies. Moreover,

thanks to application of refined computational models, we were able to suggest also a different plausible catalytic reaction mechanism for the oxidation of CO which concerns the direct involvement of a water molecule, activated by the active site surroundings. In the same study, we also showed how the enzyme may allow the formation of an energetically favourable tetra-coordinated Cu^{I} species.

In the investigation of the MoCu CODH hydrogenase activity, a large hybrid quantum mechanical/molecular mechanical (QM/MM) model allowed us to define H_2 binding modes to the bimetallic active site and to propose two different mechanisms for the splitting of such a molecule. In such context, Mo=O_{eq} was showed to represents the most basic centre in the active site. BigQM-based energy refinement showed to be fundamental in elucidating the attainable H_2 coordination modes to the enzyme active site, evidencing the impossibility of binding H_2 to Cu^{I} in the protein resting state, and to locate inaccuracies in the application of MM parameters in QM/MM calculations.

In conclusion, the work of this thesis brought knowledge about the mechanisms of the reactions catalysed by the MoCu CODH enzyme, more specifically, by disclosing fundamental features of the active site. In such context, an accurate modelling of the protein surrounding around the active site, the employment of dispersion effect and of sufficiently large basis sets in the system description, proved to be crucial for the achievement of consistent results.

Further investigations are required to shed more light on the putative catalytic mechanisms, in particular as regards the choice of the computational description of the system. The latter in fact, similarly to homologous Mo-containing enzymes[45,46], showed to have rather pronounced influence on the computed energy differences. To tackle this issue, we are planning to apply high level *ab initio* methods –e.g, multiconfigurational post-Hartree-Fock approaches[47,48,49]– to treat the bimetallic system. Moreover, protein matrix phase space sampling by means of novel theoretical approaches –e.g, quantum mechanical/discrete molecular dynamics (QM/DMD) – may also provide insights into the metalloprotein catalysis by coupling the electronic description to the protein dynamics[50]. Finally, results reported in this thesis might also be helpful for the development of synthetic biomimetic catalysts for treating hazardous CO in daily life.

.

References

[1] Ralf Conrad. Soil microorganisms as controllers of atmospheric trace gases (H_2, CO, CH_4, OCS, N_2O, and NO). *Microbiological reviews*, 60(4):609–640, 1996.

[2] Daniel Kirk-Davidoff. The greenhouse effect, aerosols, and climate change. In *Green Chemistry*, pages 211–234. Elsevier, 2018.

[3] Hanqin Tian, Guangsheng Chen, Chaoqun Lu, Xiaofeng Xu, Wei Ren, Bowen Zhang, Kamaljit Banger, Bo Tao, Shufen Pan, Mingliang Liu, et al. Global methane and nitrous oxide emissions from terrestrial ecosystems due to multiple environmental changes. *Ecosystem Health and Sustainability*, 1(1):1–20, 2015.

[4] Licheng Liu, Qianlai Zhuang, Qing Zhu, Shaoqing Liu, Hella van Asperen, and Mari Pihlatie. Global soil consumption of atmospheric carbon monoxide: an analysis using a process-based biogeochemistry model. *Atmospheric Chemistry and Physics*, 18(11):7913, 2018.

[5] Edward I Stiefel. Molybdenum bolsters the bioinorganic brigade. *Science*, 272 (5268):1599–1600, 1996.

[6] Rafel Simó and Jordi Dachs. Global ocean emission of dimethylsulfide predicted from biogeophysical data. *Global Biogeochemical Cycles*, 16(4):26–1, 2002.

[7] GA Zavarzin and AN Nozhevnikova. Aerobic carboxydobacteria. *Microbial ecology*, 3(4):305–326, 1977.

[8] Gerhard Mörsdorf, Kurt Frunzke, Dilip Gadkari, and Ortwin Meyer. Microbial growth on carbon monoxide. *Biodegradation*, 3(1):61–82, 1992.

[9] Ralf Conrad. Metabolism of nitric oxide in soil and soil microorganisms and regulation of flux into the atmosphere. In *Microbiology of atmospheric trace gases*, pages 167–203. Springer, 1996.

[10] Surabi Menon, Kenneth L Denman, Guy Brasseur, Amnat Chidthaisong, Philippe Ciais, Peter M Cox, Robert E Dickinson, Didier Hauglustaine, Christoph Heinze, Elisabeth Holland, et al. Couplings between changes in the climate system and biogeochemistry. Technical report, Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), 2007.

[11] Russ Hille, James Hall, and Partha Basu. The mononuclear molybdenum enzymes. *Chemical reviews*, 114(7):3963–4038, 2014.

[12] Russ Hille. The molybdenum oxotransferases and related enzymes. *Dalton transactions*, 42(9):3029–3042, 2013.

[13] Axel Magalon, Justin G Fedor, Anne Walburger, and Joel H Weiner. Molybdenum enzymes in bacteria and their maturation. *Coordination chemistry reviews*, 255(9-10):1159–1178, 2011.

[14] Takeshi Nishino, Ken Okamoto, and Silke Leimkühler. Enzymes of the xanthine oxidase family. In *Molybdenum and Tungsten Enzymes*, pages 192–239. 2016.

[15] Russ Hille, Stephanie Dingwall, and Jarett Wilcoxen. The aerobic CO dehydrogenase from Oligotropha carboxidovorans. *JBIC Journal of Biological Inorganic Chemistry*, 20(2):243–251, 2015.

[16] Lance C Seefeldt, Brian M Hoffman, and Dennis R Dean. Mechanism of Mo-dependent nitrogenase. *Annual review of biochemistry*, 78:701–722, 2009.

[17] K Alef and D Kleiner. Rapid and sensitive determination of microbial activity in soils and in soil aggregates by dimethylsulfoxide reduction. *Biology and fertility of soils*, 8(4):349–355, 1989.

[18] Frank Schneider, Jan Löwe, Robert Huber, Hermann Schindelin, Caroline Kisker, and Jörg Knäblein. Crystal Structure of Dimethyl Sulfoxide Reductase fromRhodobacter capsulatusat 1.88 Å Resolution. *Journal of molecular biology*, 263(1):53–69, 1996.

[19] Gary M King and Carolyn F Weber. Distribution, diversity and ecology of aerobic CO-oxidizing bacteria. *Nature Reviews Microbiology*, 5(2):107–118, 2007.

[20] Liliana Quiza, Isabelle Lalonde, Claude Guertin, and Philippe Constant. Land-use influences the distribution and activity of high affinity CO-oxidizing bacteria associated to type I-coxL genotype in soil. *Frontiers in microbiology*, 5:271, 2014.

[21] Joana C Xavier, Martina Preiner, and William F Martin. Something special about CO-dependent CO_2 fixation. *The FEBS journal*, 285(22):4181–4195, 2018.

[22] Holger Dobbek, Lothar Gremer, Reiner Kiefersauer, Robert Huber, and Ortwin Meyer. Catalysis at a dinuclear [CuSMo(=O)OH] cluster in a CO dehydrogenase resolved at 1.1-Å resolution. *Proceedings of the National Academy of Sciences*, 99(25):15971–15976, 2002.

[23] Manuel Gnida, Reinhold Ferner, Lothar Gremer, Ortwin Meyer, and Wolfram Meyer-Klaucke. A novel binuclear [CuSMo] cluster at the active site of carbon monoxide dehydrogenase: characterization by X-ray absorption spectroscopy. *Biochemistry*, 42(1):222–230, 2003.

[24] Bo Zhang, Craig F Hemann, and Russ Hille. Kinetic and spectroscopic studies of the molybdenum-copper CO dehydrogenase from Oligotropha carboxidovorans. *Journal of Biological Chemistry*, 285(17):12571–12578, 2010.

[25] Dalia Rokhsana, Tao AG Large, Morgan C Dienst, Marius Retegan, and Frank Neese. A realistic in silico model for structure/function studies of molybdenum–copper CO dehydrogenase. *JBIC Journal of Biological Inorganic Chemistry*, 21(4): 491–499, 2016.

[26] Beatrix Santiago and Ortwin Meyer. Characterization of hydrogenase activities associated with the molybdenum CO dehydrogenase from Oligotropha carboxidovorans. *FEMS microbiology letters*, 136(2):157–162, 1996.

[27] Jarett Wilcoxen and Russ Hille. The Hydrogenase Activity of the Molybdenum/Copper-containing Carbon Monoxide Dehydrogenase of Oligotropha carboxidovorans. *Journal of Biological Chemistry*, 288(50):36052–36060, 2013.

[28] Muralidharan Shanmugam, Jarett Wilcoxen, Diana Habel-Rodriguez, George E Cutsail III, Martin L Kirk, Brian M Hoffman, and Russ Hille. ^{13}C and ^{63,65}Cu ENDOR studies of CO Dehydrogenase from Oligotropha carboxidovorans. Experimental Evidence in Support of a Copper–Carbonyl Intermediate. *Journal of the American Chemical Society*, 135(47):17775–17782, 2013.

[29] Craig Gourlay, David J Nielsen, Jonathan M White, Sushilla Z Knottenbelt, Martin L Kirk, and Charles G Young. Paramagnetic active site models for the molybdenum-copper carbon monoxide dehydrogenase. *Journal of the American Chemical Society*, 128(7):2164–2165, 2006.

[30] Oliver Kreß, Manuel Gnida, Astrid M Pelzmann, Christian Marx, Wolfram Meyer-Klaucke, and Ortwin Meyer. Reversible inactivation of CO dehydrogenase with thiol compounds. *Biochemical and biophysical research communications*, 447(3):413–418, 2014.

[31] Per E. M. Siegbahn and Alexander F. Shestakov. Quantum chemical modeling of CO oxidation by the active site of molybdenum CO dehydrogenase. *Journal of Computational Chemistry*, 26(9):888–898, 2005.

[32] Matthias Hofmann, Jutta K Kassube, and Tobias Graf. The mechanism of Mo/Cu dependent CO dehydrogenase. *JBIC Journal of Biological Inorganic Chemistry*, 10(5):490–495, 2005.

[33] Benjamin W Stein and Martin L Kirk. Orbital contributions to CO oxidation in Mo–Cu carbon monoxide dehydrogenase. *Chemical communications*, 50(9): 1104–1106, 2014.

[34] Kai Xu and Hajime Hirao. Revisiting the catalytic mechanism of Mo–Cu carbon monoxide dehydrogenase using QM/MM and DFT calculations. *Physical Chemistry Chemical Physics*, 20(28):18938–18948, 2018.

[35] Raffaella Breglia, Maurizio Bruschi, Ugo Cosentino, Luca De Gioia, Claudio Greco, Toshiko Miyake, and Giorgio Moro. A theoretical study on the reactivity of the Mo/Cu-containing carbon monoxide dehydrogenase with dihydrogen. *Protein Engineering, Design and Selection*, 30(3):169–174, 2017.

[36] Stephanie Dingwall, Jarett Wilcoxen, Dimitri Niks, and Russ Hille. Studies of carbon monoxide dehydrogenase from Oligotropha carboxidovorans. *Journal of Molecular Catalysis B: Enzymatic*, 134:317–322, 2016.

[37] Axel D Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical review A*, 38(6):3098, 1988.

[38] John P Perdew. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Physical Review B*, 33(12):8822, 1986.

[39] Axel D Becke. A new mixing of Hartree–Fock and local density-functional theories. *The Journal of chemical physics*, 98(2):1372–1377, 1993.

[40] Chengteh Lee, Weitao Yang, and Robert G Parr. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical review B*, 37(2):785, 1988.

[41] James A Maier, Carmenza Martinez, Koushik Kasavajhala, Lauren Wickstrom, Kevin E Hauser, and Carlos Simmerling. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *Journal of chemical theory and computation*, 11(8):3696–3713, 2015.

[42] C. I. Bayly, P. Cieplak, W. D. Cornell, and P. A. Kollman. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.*, 97:10269–10280, 1993.

[43] Li Hong Hu, Par Söderhjelm, and Ulf Ryde. Accurate reaction energies in proteins obtained by combining QM/MM and large QM calculations. *J Chem Theory Comput*, 9(1):640–649, 2012.

[44] Sophie Sumner, Par Söderhjelm, and Ulf Ryde. Effect of geometry optimizations on QM-cluster and QM/MM studies of reaction energies in proteins. *J Chem Theory Comput*, 9(9):4205–4214, 2013.

[45] Ji-Lai Li, Ricardo A Mata, and Ulf Ryde. Large density-functional and basis-set effects for the DMSO reductase catalyzed oxo-transfer reaction. *Journal of chemical theory and computation*, 9(3):1799–1807, 2013.

[46] Jilai Li, Milica Andrejić, Ricardo A Mata, and Ulf Ryde. A Computational Comparison of Oxygen Atom Transfer Catalyzed by Dimethyl Sulfoxide Reductase with Mo and W. *European Journal of Inorganic Chemistry*, 2015(21):3580–3589, 2015.

[47] Quan Manh Phung, Sebastian Wouters, and Kristine Pierloot. Cumulant Approximated Second-Order Perturbation Theory Based on the Density Matrix Renormalization Group for Transition Metal Complexes: A Benchmark Study. *Journal of chemical theory and computation*, 12(9):4352–4361, 2016.

[48] Quan Manh Phung, Alex Domingo, and Kristine Pierloot. Dinuclear Iron(II) Spin-Crossover Compounds: A Theoretical Study. *Chemistry–A European Journal*, 24(20):5183–5190, 2018.

[49] Geng Dong, Quan Manh Phung, Simon D Hallaert, Kristine Pierloot, and Ulf Ryde. H_2 binding to the active site of [NiFe] hydrogenase studied by multiconfigurational and coupled-cluster methods. *Physical Chemistry Chemical Physics*, 19(16):10590–10601, 2017.

[50] Manuel Sparta, David Shirvanyants, Feng Ding, Nikolay V Dokholyan, and Anastassia N Alexandrova. Hybrid dynamics simulation engine for metalloproteins. *Biophysical journal*, 103(4):767–776, 2012.

Scientific publications

Paper I

A. Rovaletti, M. Bruschi, G. Moro, U. Cosentino and C. Greco
The challenging in silico description of carbon monoxide oxidation as catalyzed by molybdenum-copper CO dehydrogenase
Frontiers in chemistry, 2019, 6, pp. 630

frontiers in Chemistry

MINI REVIEW
published: 09 January 2019
doi: 10.3389/fchem.2018.00630

The Challenging *in silico* Description of Carbon Monoxide Oxidation as Catalyzed by Molybdenum-Copper CO Dehydrogenase

Anna Rovaletti[1], *Maurizio Bruschi*[1], *Giorgio Moro*[2], *Ugo Cosentino*[1] *and Claudio Greco*[1]*

[1] *Dipartimento di Scienze dell'Ambiente e della Terra, Università Degli Studi di Milano-Bicocca, Milan, Italy,* [2] *Dipartimento di Biotecnologie e Bioscienze, Università Degli Studi di Milano-Bicocca, Milan, Italy*

Carbon monoxide (CO) is a highly toxic gas to many living organisms. However, some microorganisms are able to use this molecule as the sole source of carbon and energy. Soil bacteria such as the aerobic *Oligotropha carboxidovorans* are responsible for the annual removal of about $2x10^8$ tons of CO from the atmosphere. Detoxification through oxidation of CO to CO_2 is enabled by the MoCu-dependent CO-dehydrogenase enzyme (MoCu-CODH) which—differently from other enzyme classes with similar function—retains its catalytic activity in the presence of atmospheric O_2. In the last few years, targeted advancements have been described in the field of bioengineering and biomimetics, which is functional for future technological exploitation of the catalytic properties of MoCu-CODH and for the reproduction of its reactivity in synthetic complexes. Notably, a growing interest for the quantum chemical investigation of this enzyme has recently also emerged. This mini-review compiles the current knowledge of the MoCu-CODH catalytic cycle, with a specific focus on the outcomes of theoretical studies on this enzyme class. Rather controversial aspects from different theoretical studies will be highlighted, thus illustrating the challenges posed by this system as far as the application of density functional theory and hybrid quantum-classical methods are concerned.

Keywords: molybdenum, copper, dehydrogenase, DFT, carbon monoxide

OPEN ACCESS

Edited by:
Fahmi Himo,
Stockholm University, Sweden

Reviewed by:
Lubomir Rulisek,
Institute of Organic Chemistry and Biochemistry (ASCR), Czechia
Per E. M. Siegbahn,
Stockholm University, Sweden

***Correspondence:**
Claudio Greco
claudio.greco@unimib.it

Specialty section:
This article was submitted to Theoretical and Computational Chemistry, a section of the journal Frontiers in Chemistry

Received: *29 October 2018*
Accepted: *03 December 2018*
Published: *09 January 2019*

Citation:
Rovaletti A, Bruschi M, Moro G, Cosentino U and Greco C (2019) The Challenging in silico Description of Carbon Monoxide Oxidation as Catalyzed by Molybdenum-Copper CO Dehydrogenase. Front. Chem. 6:630. doi: 10.3389/fchem.2018.00630

Carbon monoxide (CO) is a fatal gas to many living organisms as well as an indirect greenhouse gas in the atmosphere (Liu et al., 2018). Global CO emissions derive both from anthropogenic and natural sources (Choi et al., 2017). One of the main sinks of atmospheric CO is constituted by the soil, in which it is consumed in large amounts by microbial oxidation (Liu et al., 2018). One example of these important soil microorganisms is represented by the aerobic bacteria *Oligotropha carboxidovorans*. The latter is able to grow using CO as its sole source of carbon and energy (Hille et al., 2015). This metabolism is ascribed to the air-stable Mo/Cu-dependent CO dehydrogenase (MoCu-CODH) enzyme that catalyzes the oxidation of CO to CO_2 (Zhang et al., 2010).

This enzyme contains a unique active site composed by a molybdenum/copper bimetallic center (see **Figure 1A**). The molybdenum ion is found in a square-pyramidal geometry with one apical oxo ligand, a dithiolene ligand from the molybdopterin cytosine dinucleotide (MCD) cofactor, an equatorial oxo ligand and a sulfido ligand. The latter bridges to the copper center, which links the active site to the protein matrix by coordinating the sulfur atom of Cys388. Moreover, Cu is also

coordinated to a weakly bound water molecule (Gnida et al., 2003; Rokhsana et al., 2016), and can coordinate not only CO (i.e., the physiologic substrate) but also H_2. In fact, the MoCu-CODH enzyme has the ability to catalyze dihydrogen oxidation, even though such hydrogenase activity is rather low (Santiago and Meyer, 1996; Wilcoxen and Hille, 2013).

The protonation state of the active site has been object of debate. In fact, experimental X-ray diffraction (XRD) results were interpreted as indicative of a Mo(=O)OH state both for the oxidized and the reduced forms of the enzyme (Dobbek et al., 2002). Differently, extended X-ray absorption fine structure (EXAFS) spectroscopy suggested the presence of a Mo^{VI}(=O)O unit and a Mo^{IV}(=O)OH_2 unit for the oxidized and reduced states, respectively (Gnida et al., 2003). Recent data in support of the EXAFS-based hypothesis came from electron paramagnetic resonance (EPR) analysis of the Mo^{V} analog and from theoretical

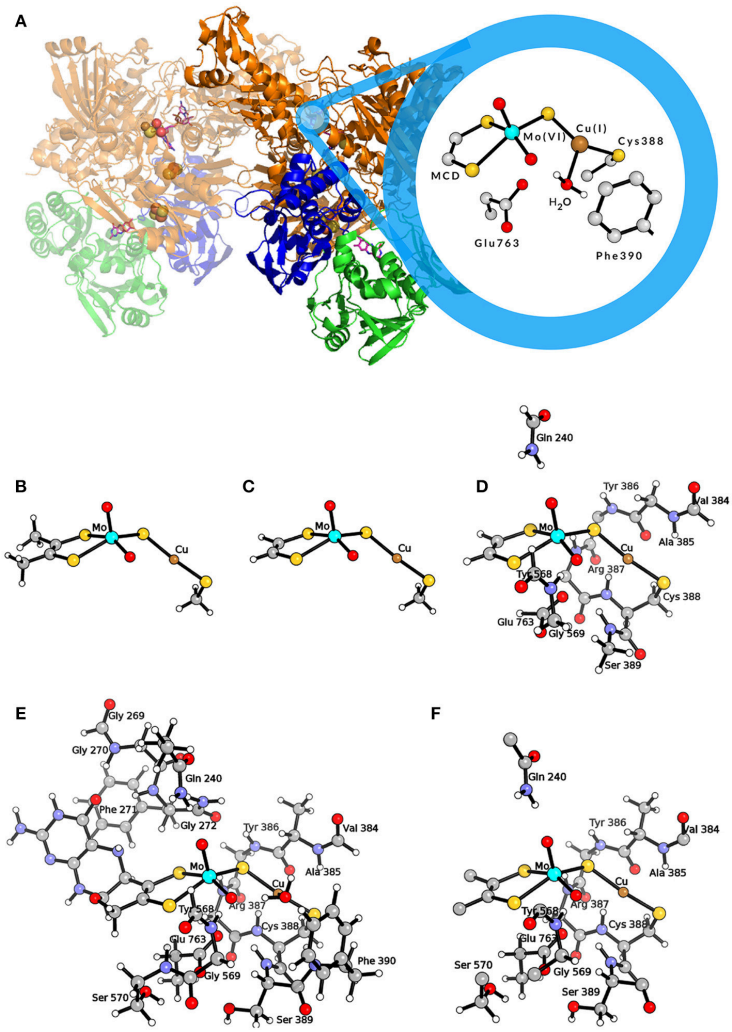


FIGURE 1 | (A) Representation of the MoCu-CODH enzyme and of the active site in the Mo(VI)Cu(I) resting state. **(B)** QM model used by Hofmann et al. (2005) and by Stein and Kirk (2014); **(C)** QM model used by Siegbahn and Shestakov (2005) and by Breglia et al. (2017); **(D)** larger QM model used by Siegbahn and Shestakov (2005); **(E)** QM model used by Rokhsana et al. (2016); **(F)** QM region of the hybrid QM/MM model used by Xu and Hirao (2018). Color code of atoms in the active site: cyan, molybdenum; orange, copper; yellow, sulfur; red, oxygen; gray, carbon; white, hydrogen.

calculations (Zhang et al., 2010; Rokhsana et al., 2016). As far as the oxidation state of the Cu ion is concerned, it maintains the +1 state throughout the enzymatic catalytic cycle (Dobbek et al., 2002; Gnida et al., 2003). In fact, CO oxidation occurs directly at the Cu center (Shanmugam et al., 2013), and the two-electron transfer to Mo at each catalytic cycle is allowed by the highly delocalized nature of the Mo(μ-S)Cu unit (Gourlay et al., 2006). Key second-sphere amino acid residues are conserved in MoCu-CODH enzymes and in homologues with different activity, i.e., xanthine oxidases (Hille, 2013). In particular, a conserved glutamate residue (Glu763) is found in proximity to the equatorial oxo ligand of Mo and is considered to act as a base to facilitate deprotonation events (Wilcoxen and Hille, 2013; Hille et al., 2015). Moreover, the aromatic ring of a phenylalanine residue (Phe390), located in front of the Cu^{I} ion, is thought to have relevance for the correct positioning of the substrate within the active site (Rokhsana et al., 2016).

Carbon monoxide (|C≡O) is expected to show similar reactivity with respect to the isoelectronic isocyanide (|C≡N-R) species. They share the presence of non-bonding electron pairs in the *sp* orbital of the terminal carbon atom and a triple bond between the latter and a more electronegative atom. Indeed, Dobbek et al. reported the inhibitory activity of *n*-butylisocyanide, $|C{\equiv}N\text{-}(CH_2)_3CH_3$, toward the oxidized MoCu-CODH (Dobbek et al., 2002). In the same study, the corresponding crystallographic structure of the inhibited enzyme was determined (PDB ID: 1N62). The resulting inactive complex is characterized by a thiocarbamate geometry in which the isocyanide group forms covalent bonds with the μ-sulfido ligand, the equatorial oxygen of Mo and the Cu atom, while the alkyl chain of *n*-butylisocyanide extends into the hydrophobic interior of the substrate channel.

The features of the crystal structure of the *n*-butylisocyanide-bound state prompted Dobbek and coworkers to advance the first hypothesis ever proposed for the MoCu-CODH catalytic mechanism (see **Figure 2A**) (Dobbek et al., 2002). The latter involves the formation of a thiocarbonate-like intermediate—analogous to the thiocarbamate derivative formed during the aforementioned inhibition—after the CO substrate accesses the oxidized active site. Such thiocarbonate species would be characterized by the insertion of CO between copper, the μ-sulfido ligand and the equatorial oxo ligand of the Mo atom. Taking inspiration from such a mechanistic hypothesis for CO-oxidation catalysis, the enzymatic mechanism has been subsequently studied with computational methods by several groups, thus giving birth to a debate that is still ongoing (*vide infra*).

Recently, it has been reported that rather bulky thiol molecules—e.g., L-cysteine, coenzyme A or glutathione—can reach the bimetallic active site (Kreß et al., 2014). They cause a reversible inactivation of the enzymatic activity, competing with the substrates for the same position at the Cu^{I} center.

In the following two sections of this manuscript we aim at reviewing the theoretical studies that have been published on MoCu-CODH using quantum mechanical (QM) and hybrid quantum mechanical/molecular mechanical (QM/MM) approaches. In doing so, we will pay special attention to the controversies and the challenges that have emerged. Moreover, promising future developments in the theoretical description of this system will be proposed in the concluding section.

1. SMALL- AND MEDIUM-SIZED QM-CLUSTER MODELS

The first theoretical investigations of the catalytic activity of MoCu-CODH were carried out independently by two theoretical groups in 2005 (Hofmann et al., 2005; Siegbahn and Shestakov, 2005). To model the enzymatic active site, Hofmann et al. used a small cluster of 24 atoms representing the two transition metals and their first coordination spheres (see **Figure 1B**) (Hofmann et al., 2005), whereas Siegbahn and Shestakov performed calculations with two different models, a small one

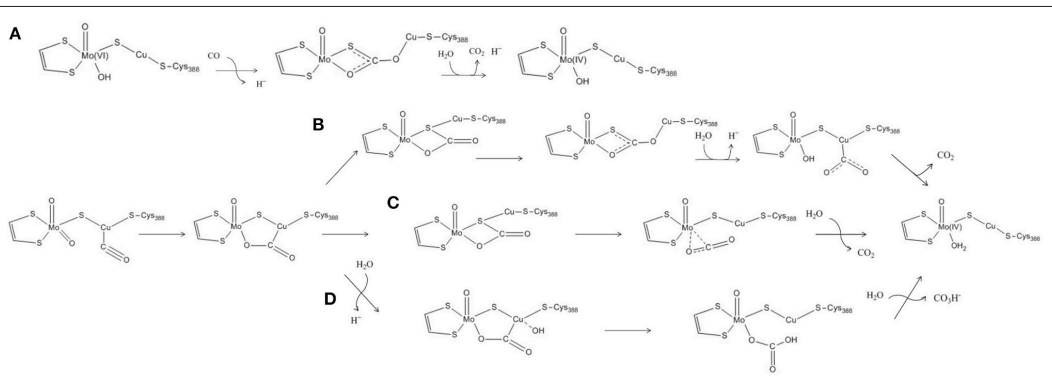


FIGURE 2 | Proposed reaction mechanisms for the oxidation of CO by means of MoCu-CODH. **(A)** Reaction mechanism proposed by Dobbek's experimental group. **(B–D)** Reaction mechanisms proposed by Siegbahn and Shestakov, by Hofmann et al. and by Stein and Kirk, on the basis of their respective computational studies.

of about 20 atoms and a bigger one composed of about 70 atoms (see **Figures 1C,D**). In the latter, some residues belonging to the second coordination spheres of the metals were explicitly included (Siegbahn and Shestakov, 2005). The hybrid density functional B3LYP (Lee et al., 1988; Becke, 1993) was employed in both cases to optimize the geometries and compute relative energies of intermediates along the putative catalytic cycles. For geometry optimizations, Hofmann and coworkers adopted the Lanl2DZ (Dunning and Hay, 1976; Hay and Wadt, 1985a,b; Wadt and Hay, 1985) effective core basis set with additional d-type functions on S atoms. This was followed by single point energy calculations using the SDD (Dunning, 1970; Dunning and Hay, 1977; Dolg et al., 1987) basis set augmented by d-type polarization functions on all non-hydrogen, non-metal atoms. The basis sets employed by Siegbahn and Shestakov were lacvp and lacv3p*—with ECP for Mo, Cu, and S atoms—for geometry optimizations and energy calculations, respectively. In both studies, the protein matrix surrounding the active site was modeled by a continuum dielectric with $\epsilon = 4$ (Eckert and Klamt, 2002; Cossi et al., 2003).

A comparative analysis of the results coming from two such early investigations evidences a significant variability of the results obtained as a function of the adopted level of theory, within models of the same size. In particular, the calculated energy differences showed significant dependency on the basis sets used, with deviations up to 21 kJ/mol in the case of intermediates involved in formation of the new C–O bond, a key step in the catalytic process. Interestingly, previous studies of oxygen-atom transfer (OAT) reactions involving Mo-complexes (Li et al., 2013) and other transition-metal-containing systems (Hu and Chen, 2015; Li et al., 2015) also evidenced the strong basis set effect on computed energy differences along the reactive paths.

Notwithstanding the shortcomings deriving from the choice of basis sets, both Siegbahn, Shestakov and their respective coworkers evidenced a surprisingly high stability for the thiocarbonate intermediate. The presence of such a deep minimum on the energy landscape pertaining to the reaction mechanism was interpreted differently by the two groups. Siegbahn and Shestakov in particular proposed that the thiocarbonate derivative represents an intermediate of the CO-oxidation mechanism. However, in the proposed mechanism the barrier for the release of the CO_2 product was estimated to be rather high, as it would require the insertion of a water molecule which was reported not to be a facile step (see **Figure 2B**). Differently, Hofmann and coworkers raised the possibility that the thiocarbonate adduct lies outside the catalytic cycle, in a deep potential energy well that would effectively slow down enzymatic activity (see **Figure 2C**). These authors further proposed that the constrains imposed by the protein matrix could prevent formation of such a stable off-path adduct, a hypothesis that— however—was later discarded as a result of a theoretical study focused on this topic (Siegbahn, 2011).

In a more recent theoretical study, a different mechanism for the oxidation of CO by the MoCu-CODH enzyme was proposed (see **Figure 2D**) (Stein and Kirk, 2014). Using a cluster model analogous to the one previously employed by Hofmann and coworkers (see **Figure 1B**), at the PBE/TZP (Perdew et al., 1996; Ernzerhof and Scuseria, 1999) level of theory and including continuum dielectric contributions (ϵ = hexane) (Klamt and Schüürmann, 1993), Stein and Kirk proposed that the stable thiocarbonate intermediate formation could be bypassed by evolving bicarbonate as a final product rather than CO_2. Bicarbonate formation would proceed via nucleophilic attack of a copper-activated water molecule on the C atom of the metal-bound CO_2. However, such a picture is at odds with recent experimental studies, which appear to exclude the possibility of forming a bicarbonate complex during catalysis (Dingwall et al., 2016).

Breglia and coworkers published the most recent theoretical study of MoCu-CODH, in which only the first shell coordination spheres were included in a QM model (see **Figure 1C**) (Breglia et al., 2017). Such a study mainly regards the hydrogenase activity of the enzyme and includes a comparative analysis of the binding reactions of the physiologic substrate—i.e., CO—and of dihydrogen to the Cu ion. Similarly to previous studies on H_2- and CO-binding enzyme models in which a pure functional was used in conjuction with triple-zeta bases (Greco et al., 2015; Rovaletti and Greco, 2018), geometry optimizations and energy calculations were carried out in vacuo at BP86/def2-TZVP level (Perdew, 1986; Becke, 1988; Weigend and Ahlrichs, 2005). As far as the energetics of CO binding is concerned, the computed ΔE was as negative as −64 kJ/mol. A comparison of the latter value with the results of corresponding calculations previously performed at different levels of theory for the same reaction by Siegbahn, Hofmann and their respective coworkers, evidences significant discrepancies ($\Delta\Delta E$ = 50 and 31 kJ/mol, respectively). Actually, the occurrence of such large differences comes with little surprise, given the well-known shortcomings in binding energy calculations using quantum chemical models of coordination complexes and their ligands (Husch et al., 2018).

2. LARGE QM-CLUSTER MODELS AND HYBRID MODELS

The importance of extending the dimension of the bimetallic active site model and systematically accounting for the effects of the second-sphere residues on energetics was evidenced in a recent theoretical work by Rokhsana and coworkers (Rokhsana et al., 2016). In fact, a large-size cluster model of around 180 atoms (see **Figure 1E**) turned out to be required for a fully satisfactory reproduction of experimentally determined structural parameters. The same held true for the evaluation of the most plausible protonation states of the Mo/Cu core, which was done taking into account key geometric features of the enzyme crystal structure and redox potential measurements available in literature. The authors were able to assess in particular the protonation state of the equatorial oxo ligand of Mo, at the different redox states attained during catalysis. As for the adopted level of theory, Rokhsana et al. employed the def2-TZVP basis set for all elements, apart from H and C atoms, for which the def2-SVP basis set was used (Weigend and Ahlrichs, 2005). The protein environment was modeled by a continuum dielectric with $\epsilon = 4$

(Klamt and Schüürmann, 1993). The BP86 and B3LYP density functionals were used for geometry optimizations and energy evaluation, respectively.

In a subsequent study, explicit consideration of the whole protein environment was achieved by means of a hybrid quantum mechanics/molecular mechanics (QM/MM) approach (Xu and Hirao, 2018). In the work of Xu and Hirao, the active site was described by a QM region of 89 atoms (see **Figure 1F**) using the B3LYP functional. During geometry optimization, the SDD effective core potential basis set was employed to represent the transition metal ions, whereas the 6-31G* basis set was adopted for all the other atoms (Dolg et al., 1987; Andrae et al., 1990). Single-point energy calculations were carried out using the larger def2-TZVP basis set. For the molecular mechanics calculations, the AMBER03 force field was employed (Duan et al., 2003). Moreover, the Grimme's D3-correction with Becke−Johnson damping [D3(BJ)] was taken into account in the calculations (Grimme et al., 2011). The previously proposed catalytic mechanisms that involve the formation of the thiocarbonate species were re-investigated at such a level of theory. According to Xu and Hirao's study, the S-C-bound adduct would be formed along the reaction pathway as previously suggested (Dobbek et al., 2002; Siegbahn and Shestakov, 2005). However, based on the novel QM/MM results, the thiocarbonate species thus formed would not be as stable as previously proposed. It has to be remarked that the thiocarbonate intermediate was not found to be directly linked to the transition state for CO_2 releasing. In fact, it was proposed that— after thiocarbonate formation—the reaction needs to follow a reverse process to productively proceed toward CO_2 evolution. Notably, the overall barrier for the proposed catalytic mechanism was found to be low (in the order of 50 kJ/mol). In the same study, Xu and Hirao also carried out purely QM calculations with a QM-cluster of the same size of the quantum-mechanical region of their hybrid model, and compared the obtained results with those coming from QM/MM modeling. Such a comparison evidenced that the protein environment is not involved in modulating the kinetic barrier associated with the investigated catalytic mechanism. However, it was found that the protein matrix plays an important role in the stabilization of the CO_2-released state. Finally, it was also reported that the inclusion of dispersive corrections lowers by 15 kJ/mol the activation barrier of the product-releasing step, in line with what was expected for the modeling of a bimolecular reaction step.

CONCLUDING REMARKS

Over the last fifteen years, the theoretical investigation of the CO oxidation mechanism by MoCu-CODH has given rise to a debate, the essentials of which are centered on the possible occurrence and on the role of a thiocarbonate catalytic intermediate. In the above sections, we have reported key details of the various computational studies published to date, and we are now in the condition to present a more general outlook on the state of the art regarding MoCu-CODH.

The early studies by Siegbahn, Hofmann, and their respective coworkers evidenced that the thiocarbonate intermediate would occupy a deep well in the energy profiles pertaining to the investigated reaction mechanisms. However, the kinetic barriers they computed for CO_2 evolution were at least 30 kJ/mol higher than the recently determined experimental counterpart (Zhang et al., 2010). In part, this picture depends on the neglect of dispersive corrections: their inclusion became a standard possibility only after the publication of the mentioned study (Siegbahn, 2011).

Results more compatible with the experimental evidence of a kinetic barrier of around 50 kJ/mol were obtained by Xu and Hirao, who exploited a larger QM-cluster model with the explicit inclusion of most of the second-sphere coordination environment, along with employment of large basis sets and dispersion corrections (Xu and Hirao, 2018). It is noticeable that, according to Xu and Hirao's results, the thiocarbonate species still appears to behave as a thermodynamic sink. Even though not a very deep one, such a sink would effectively hamper the advancement along the proposed path toward products, a rather unusual role for a species formed during an enzymatic process.

All the catalytic mechanisms proposed in the theoretical studies reviewed here focus on the possibility that the Mo-bound equatorial oxo ligand performs the nucleophilic attack on the activated CO substrate bound to the Cu ion. This is in line with what has been suggested in the case of the catalytic mechanism of the homologous xanthine oxidases enzymes. However, a variant with respect to such a picture has been recently proposed, in which an activated water molecule would play the role of nucleophile (Hille et al., 2015). Notably, to the best of our knowledge this alternative mechanism has not yet been investigated by QM or hybrid QM/MM studies.

Future theoretical studies on this (and other) putative catalytic mechanism will possibly face the challenge associated with rather pronounced fluctuations of computed energy differences as a function of the adopted level of theory. As far as the application of density functional theory is concerned, extensive benchmarking could in principle help to improve the theoretical predictions. In this regard, the available experimental data on enzyme inhibition— by thiols in particular—could represent a useful dataset, keeping in mind that the reproduction of binding energies in the case of bulky thiols might require extensive protein matrix phase space sampling in the case of QM/MM studies. High level *ab initio* methods are a viable—though still challenging—alternative for providing reliable results. In fact, thanks to recent methodological developments, the treatment of relatively large bimetallic systems has been shown to be computationally affordable with multiconfigurational post-Hartree-Fock approaches (Phung et al., 2016, 2018; Dong et al., 2017).

The relevance of MoCu-CODH as an inspiring system for future biomimetic and bioengineering applications is currently growing. This is due not only to the relevance of the reactions it catalyzes, but also to its resistance to atmospheric O_2 exposure—a rare feature in the case of enzymes expressing carbon monoxide dehydrogenase and hydrogenase activities (Choi et al., 2017; Gourlay et al., 2018; Groysman et al., 2018). Notably, the recent establishment of a functional heterologous expression system for

the MoCu-CODH enzyme (Kaufmann et al., 2018) together with developments in the computational chemistry field will hopefully boost the positive feedback among biochemical, biomimetic and quantum chemical studies, opening new perspectives for a deeper understanding of this interesting metalloenzyme.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

REFERENCES

Andrae, D., Haeussermann, U., Dolg, M., Stoll, H., and Preuss, H. (1990). Energy-adjusted *ab initio* pseudopotentials for the second and third row transition elements. *Theor. Chim. Acta* 77, 123–141.

Becke, A. D. (1988). Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* 38:3098.

Becke, A. D. (1993). A new mixing of Hartree–Fock and local density-functional theories. *J. Chem. Phys.* 98, 1372–1377.

Breglia, R., Bruschi, M., Cosentino, U., De Gioia, L., Greco, C., Miyake, T., et al. (2017). A theoretical study on the reactivity of the Mo/Cu-containing carbon monoxide dehydrogenase with dihydrogen. *Protein Eng. Design Select.* 30, 169–174. doi: 10.1093/protein/gzw071

Choi, E. S., Min, K., Kim, G.-J., Kwon, I., and Kim, Y. H. (2017). Expression and characterization of *Pantoea* CO dehydrogenase to utilize CO-containing industrial waste gas for expanding the versatility of CO dehydrogenase. *Sci. Rep.* 7:44323. doi: 10.1038/srep44323

Cossi, M., Rega, N., Scalmani, G., and Barone, V. (2003). Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. *J. Comput. Chem.* 24, 669–681. doi: 10.1002/jcc.10189

Dingwall, S., Wilcoxen, J., Niks, D., and Hille, R. (2016). Studies of carbon monoxide dehydrogenase from *Oligotropha carboxidovorans*. *J. Mol. Catal. B Enzymat.* 134, 317–322. doi: 10.1016/j.molcatb.2016.10.007

Dobbek, H., Gremer, L., Kiefersauer, R., Huber, R., and Meyer, O. (2002). Catalysis at a dinuclear [CuSMo(=O)OH] cluster in a CO dehydrogenase resolved at 1.1-Å resolution. *Proc. Natl. Acad. Sci. U.S.A.* 99, 15971–15976. doi: 10.1073/pnas.212640899

Dolg, M., Wedig, U., Stoll, H., and Preuss, H. (1987). Energy-adjusted *ab initio* pseudopotentials for the first row transition elements. *J. Chem. Phys.* 86, 866–872.

Dong, G., Phung, Q. M., Hallaert, S. D., Pierloot, K., and Ryde, U. (2017). H_2 binding to the active site of [NiFe] hydrogenase studied by multiconfigurational and coupled-cluster methods. *Phys. Chem. Chem. Phys.* 19, 10590–10601. doi: 10.1039/C7CP01331K

Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G., Zhang, W., et al. (2003). A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* 24, 1999–2012. doi: 10.1002/jcc.10349

Dunning, T. and Hay, P. (1976). *Modern Theoretical Chemistry*. Plenum, NY: HF Schaefer III.

Dunning, T. H., and Hay, P. J. (1977). "Gaussian basis sets for molecular calculations," in *Methods of Electronic Structure Theory*, ed H. F. Schaefer (Boston, MA: Springer), 1–27.

Dunning, T. H. Jr. (1970). Gaussian basis functions for use in molecular calculations. I. contraction of (9s5p) atomic basis sets for the first-row atoms. *J. Chem. Phys.* 53, 2823–2833.

Eckert, F., and Klamt, A. (2002). Fast solvent screening via quantum chemistry: COSMO-RS approach. *AIChE J.* 48, 369–385. doi: 10.1002/aic.690480220

Ernzerhof, M., and Scuseria, G. E. (1999). Assessment of the Perdew–Burke–Ernzerhof exchange-correlation functional. *J. Chem. Phys.* 110, 5029–5036.

Gnida, M., Ferner, R., Gremer, L., Meyer, O., and Meyer-Klaucke, W. (2003). A novel binuclear [CuSMo] cluster at the active site of carbon monoxide dehydrogenase: characterization by X-ray absorption spectroscopy. *Biochemistry* 42, 222–230. doi: 10.1021/bi026514n

Gourlay, C., Nielsen, D. J., Evans, D. J., White, J. M., and Young, C. G. (2018). Models for aerobic carbon monoxide dehydrogenase: synthesis, characterization and reactivity of paramagnetic $Mo^{V}O(\mu\text{-}S)Cu^{I}$ complexes. *Chem. Sci.* 9, 876–888. doi: 10.1039/c7sc04239f

Gourlay, C., Nielsen, D. J., White, J. M., Knottenbelt, S. Z., Kirk, M. L., and Young, C. G. (2006). Paramagnetic active site models for the molybdenum-copper carbon monoxide dehydrogenase. *J. Am. Chem. Soc.* 128, 2164–2165. doi: 10.1021/ja056500f

Greco, C., Ciancetta, A., Bruschi, M., Kulesza, A., Moro, G., and Cosentino, U. (2015). Influence of key amino acid mutation on the active site structure and on folding in acetyl-CoA synthase: a theoretical perspective. *Chem. Commun.* 51, 8551–8554. doi: 10.1039/c5cc01575h

Grimme, S., Ehrlich, S., and Goerigk, L. (2011). Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* 32, 1456–1465. doi: 10.1002/jcc.21759

Groysman, S., Hollingsworth, T., Hollingsworth, R. L., and Lord, R. L. (2018). Cooperative bimetallic reactivity of a heterodinuclear molybdenum-copper model of Mo-Cu CODH. *Dalton Trans.* 47, 10017–10024. doi: 10.1039/c8dt02323a

Hay, P. J., and Wadt, W. R. (1985a). *Ab initio* effective core potentials for molecular calculations. Potentials for K to Au including the outermost core orbitals. *J. Chem. Phys.* 82, 299–310. doi: 10.1063/1.448975

Hay, P. J., and Wadt, W. R. (1985b). *Ab initio* effective core potentials for molecular calculations. Potentials for the transition metal atoms Sc to Hg. *J. Chem. Phys.* 82, 270–283. doi: 10.1063/1.448799

Hille, R. (2013). The molybdenum oxotransferases and related enzymes. *Dalton Trans.* 42, 3029–3042. doi: 10.1039/c2dt32376a

Hille, R., Dingwall, S., and Wilcoxen, J. (2015). The aerobic CO dehydrogenase from Oligotropha carboxidovorans. *J. Biol. Inorgan. Chem.* 20, 243–251. doi: 10.1007/s00775-014-1188-4

Hofmann, M., Kassube, J. K., and Graf, T. (2005). The mechanism of Mo/Cu dependent CO dehydrogenase. *J. Biol. Inorgan. Chem.* 10, 490–495. doi: 10.1007/s00775-005-0661-5

Hu, L., and Chen, H. (2015). Assessment of DFT methods for computing activation energies of Mo/W-mediated reactions. *J. Chem. Theor. Comput.* 11, 4601–4614. doi: 10.1021/acs.jctc.5b00373

Husch, T., Freitag, L., and Reiher, M. (2018). Calculation of ligand dissociation energies in large transition-metal complexes. *J. Chem. Theor. Comput.* 14, 2456–2468. doi: 10.1021/acs.jctc.8b00061

Kaufmann, P., Duffus, B. R., Teutloff, C., and Leimkuhler, S. (2018). Functional studies on oligotropha carboxidovorans molybdenum–copper CO dehydrogenase produced in *Escherichia coli*. *Biochemistry* 57, 2889–2901. doi: 10.1021/acs.biochem.8b00128

Klamt, A., and Schüürmann, G. (1993). COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc. Perkin Trans.* 2, 799–805. doi: 10.1039/P29930000799

Kreß, O., Gnida, M., Pelzmann, A. M., Marx, C., Meyer-Klaucke, W., and Meyer, O. (2014). Reversible inactivation of CO dehydrogenase with thiol compounds. *Biochem. Biophys. Res. Commun.* 447, 413–418. doi: 10.1016/j.bbrc.2014.03.147

Lee, C., Yang, W., and Parr, R. G. (1988). Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* 37:785. doi: 10.1103/PhysRevB.37.785

Li, J., Andrejić, M., Mata, R. A., and Ryde, U. (2015). A computational comparison of oxygen atom transfer catalyzed by dimethyl sulfoxide reductase with Mo and W. *Eur. J. Inorgan. Chem.* 2015, 3580–3589. doi: 10.1002/ejic.201500209

Li, J.-L., Mata, R. A., and Ryde, U. (2013). Large density-functional and basis-set effects for the DMSO reductase catalyzed oxo-transfer reaction. *J. Chem. Theor. Comput.* 9, 1799–1807. doi: 10.1021/ct301094r

Liu, L., Zhuang, Q., Zhu, Q., Liu, S., van Asperen, H., and Pihlatie, M. (2018). Global soil consumption of atmospheric carbon monoxide: an analysis

using a process-based biogeochemistry model. *Atmos. Chem. Phys.* 18:7913. doi: 10.5194/acp-18-7913-2018

Perdew, J. P. (1986). Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B* 33:8822. doi: 10.1103/PhysRevB.33.8822

Perdew, J. P., Burke, K., and Ernzerhof, M. (1996). Generalized gradient approximation made simple. *Phys. Rev. Lett.* 77:3865. doi: 10.1103/PhysRevLett.77.3865

Phung, Q. M., Domingo, A., and Pierloot, K. (2018). Dinuclear iron(II) spin-crossover Compounds: a theoretical study. *Chem. A Eur. J.* 24, 5183–5190. doi: 10.1002/chem.201704441

Phung, Q. M., Wouters, S., and Pierloot, K. (2016). Cumulant approximated second-Order perturbation theory based on the density matrix renormalization group for transition metal complexes: a benchmark study. *J. Chem. Theor. Comput.* 12, 4352–4361. doi: 10.1021/acs.jctc.6b00714

Rokhsana, D., Large, T. A., Dienst, M. C., Retegan, M., and Neese, F. (2016). A realistic *in silico* model for structure/function studies of molybdenum–copper CO dehydrogenase. *J. Biol. Inorgan. Chem.* 21, 491–499. doi: 10.1007/s00775-016-1359-6

Rovaletti, A., and Greco, C. (2018). Organophosphorous ligands in hydrogenase-inspired iron-based catalysts: a DFT study on the energetics of metal protonation as a function of P-atom substitution. *J. Phys. Org. Chem.* 31:e3748. doi: 10.1002/poc.3748

Santiago, B., and Meyer, O. (1996). Characterization of hydrogenase activities associated with the molybdenum CO dehydrogenase from Oligotropha carboxidovorans. *FEMS Microbiol. Lett.* 136, 157–162. doi: 10.1111/j.1574-6968.1996.tb08042.x

Shanmugam, M., Wilcoxen, J., Habel-Rodriguez, D., Cutsail III, G. E., Kirk, M. L., Hoffman, B. M., et al. (2013). 13C and 63,65Cu ENDOR studies of CO Dehydrogenase from Oligotropha carboxidovorans. Experimental evidence in support of a copper–carbonyl intermediate. *J. Am. Chem. Soc.* 135, 17775–17782. doi: 10.1021/ja406136f

Siegbahn, P. E. (2011). The effect of backbone constraints: the case of water oxidation by the oxygen-evolving complex in PSII. *ChemPhysChem* 12, 3274–3280. doi: 10.1002/cphc.201100475

Siegbahn, P. E. M., and Shestakov, A. F. (2005). Quantum chemical modeling of CO oxidation by the active site of molybdenum CO dehydrogenase. *J. Comput. Chem.* 26, 888–898. doi: 10.1002/jcc.20230

Stein, B. W., and Kirk, M. L. (2014). Orbital contributions to CO oxidation in Mo–Cu carbon monoxide dehydrogenase. *Chem. Commun.* 50, 1104–1106. doi: 10.1039/C3CC47705C

Wadt, W. R., and Hay, P. J. (1985). *Ab initio* effective core potentials for molecular calculations. Potentials for main group elements Na to Bi. *J. Chem. Phys.* 82, 284–298. doi: 10.1063/1.448800

Weigend, F., and Ahlrichs, R. (2005). Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for h to rn: design and assessment of accuracy. *Phys. Chem. Chem. Phys.* 7, 3297–3305. doi: 10.1039/b508541a

Wilcoxen, J., and Hille, R. (2013). The hydrogenase activity of the molybdenum/copper-containing carbon monoxide dehydrogenase of oligotropha carboxidovorans. *J. Biol. Chem.* 288, 36052–36060. doi: 10.1074/jbc.M113.522441

Xu, K., and Hirao, H. (2018). Revisiting the catalytic mechanism of Mo–Cu carbon monoxide dehydrogenase using QM/MM and DFT calculations. *Phys. Chem. Chem. Phys.* 20, 18938–18948. doi: 10.1039/C8CP00858B

Zhang, B., Hemann, C. F., and Hille, R. (2010). Kinetic and spectroscopic studies of the molybdenum-copper CO dehydrogenase from Oligotropha carboxidovorans. *J. Biol. Chem.* 285, 12571–12578. doi: 10.1074/jbc.M109.076851

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Paper II

A. Rovaletti, M. Bruschi, G. Moro, U. Cosentino, U. Ryde, and C. Greco
A thiocarbonate sink on the enzymatic energy landscape of aerobic CO oxidation? Answers from DFT and QM/MM models of MoCu CO-dehydrogenases
Journal of catalysis, 2019, 372, pp. 201–205

Journal of Catalysis 372 (2019) 201–205

Contents lists available at ScienceDirect

Journal of Catalysis

journal homepage: www.elsevier.com/locate/jcat

Priority Communication

A thiocarbonate sink on the enzymatic energy landscape of aerobic CO oxidation? Answers from DFT and QM/MM models of Mo—Cu CO-dehydrogenases

Anna Rovaletti [a], Maurizio Bruschi [a], Giorgio Moro [b], Ugo Cosentino [a], Ulf Ryde [c,*], Claudio Greco [a,*]

[a] Department of Earth and Environmental Sciences, Milano – Bicocca University, Piazza della Scienza 1, 20126 Milan, Italy
[b] Department of Biotechnology and Biosciences, Milano – Bicocca University, Piazza della Scienza 2, 20126 Milan, Italy
[c] Department of Theoretical Chemistry, Lund University, Chemical Centre, P.O. Box 124, SE-221 00 Lund, Sweden

ARTICLE INFO

Article history:
Received 2 January 2019
Revised 26 February 2019
Accepted 26 February 2019
Available online 18 March 2019

Keywords:
CO-dehydrogenase
Molybdenum
Copper
DFT
Carbon monoxide
Thiocarbonate intermediate

ABSTRACT

We present a theoretical investigation providing key insights on a long-standing controversial issue that dominated the debate on carbon monoxide oxidation by Mo—Cu CO-dehydrogenases. Previous investigations gravitate around the possible occurrence of a thiocarbonate intermediate, that was repeatedly reported to behave as a thermodynamic sink on the catalytic energy landscape. By using a hierarchy of quantum mechanical and hybrid quantum/classical models of the enzyme, we show that no such energy sink is present on the catalytic energy profile. Consequent perspectives for the definition of a novel mechanistic proposal for the enzyme-catalyzed CO-oxidation are discussed in light of the recent literature.

1. Introduction

The MoCu-dependent CO-dehydrogenase (MoCu-CODH), expressed by the aerobic soil bacterium *Oligotropha carboxidovorans*, is an enzyme with key environmental relevance, as it detoxifies 2×10^8 tons of atmospheric CO annually [1]. Its catalytic mechanism has been the subject of extensive research efforts, which have led to controversies regarding the nature of the intermediates along the CO-oxidizing cycle giving rise to CO_2 evolution [2–5]. The MoCu-CODH active site presents a unique architecture in which the molybdenum ion shows a distorted square-pyramidal geometry (Fig. 1). An oxo (O^{2-}) ligand is found at the apex and the base is formed by a dithiolene ligand from a molybdopterin–cytosine dinucleotide (MCD) cofactor, another oxo ligand and a sulfido ligand. The latter bridges to the Cu ion, which is also coordinated by the sulfur atom of Cys388 and a weakly bound water molecule [6–8]. The copper ion remains in the +1 oxidation state throughout the catalytic cycle [2,6]. Substrate binding and its oxidation are believed to occur at the Cu(I) center [9], owing to the high degree of delocalization within the Mo—S—Cu unit, allowing for facile electron transfer toward the redox-active Mo ion during catalysis [10].

The first mechanistic hypothesis for the catalytic cycle of MoCu-CODH was based on information coming from the crystal structure of the enzyme complexed with the *n*-butylisocyanide inhibitor [2]. According to this proposal, after CO binding to Cu(I), the equatorial oxo ligand of molybdenum (O_{eq}) may act as a nucleophile. Such a reaction would yield an intermediate with a five-membered ring, involving both the Mo and Cu ions. In the next step, CO may insert between the sulfido ligand and copper to yield a thiocarbonate intermediate (D in Fig. 1), which is analogous to the thiocarbamate derivative formed during inhibition of the enzyme by *n*-butylisocyanide.

Inspired by this proposal, quantum mechanical (QM) studies based on density functional theory (DFT) were carried out on cluster models of the enzyme active site [3,4]. These studies discovered an unexpected stability of the thiocarbonate intermediate which prevented an easy progress along the CO-oxidation mechanism. Siegbahn et al. suggested that the release of CO_2 can be facilitated by the binding of a water molecule to molybdenum. In another theoretical study, Hofmann and coworkers proposed that the thiocarbonate species is not part of the catalytic cycle. Based on their results, the thiocarbonate would represent a deep thermodynamic

* Corresponding authors.
E-mail addresses: Ulf.Ryde@teokem.lu.se (U. Ryde), Claudio.greco@unimib.it (C. Greco).

https://doi.org/10.1016/j.jcat.2019.02.032
0021-9517/

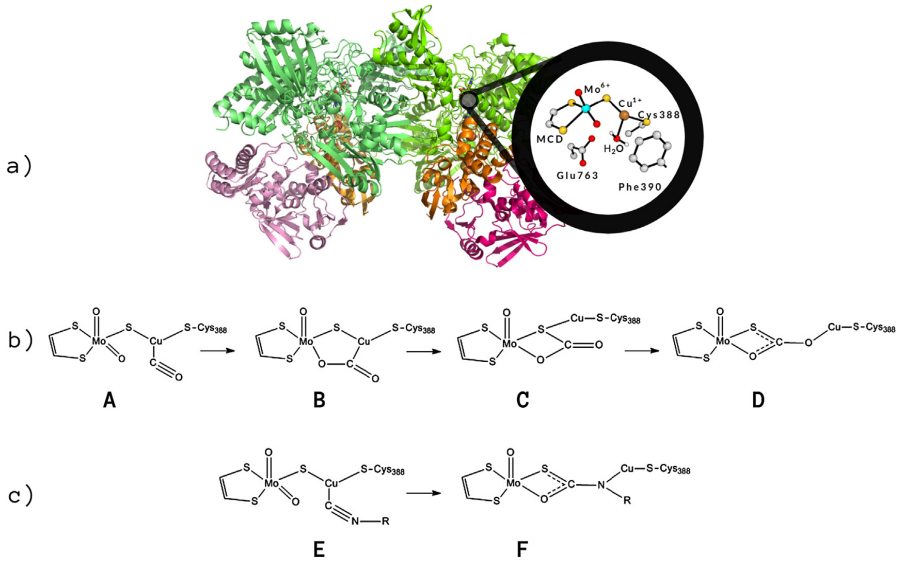


Fig. 1. (a) Representation of the MoCu-CODH enzyme and of the active site in the Mo(VI)Cu(I) resting state; (b) Intermediates involved in the thiocarbonate-formation mechanism; (c) Intermediates involved during inhibition of MoCu-CODH by *n*-butylisocyanide. R = *n*-butyl.

sink located outside the CO_2-evolving route, the presence of which would significantly slow down catalysis.

Both these QM studies were performed on small cluster models of the active site, which neglect most effects of the surrounding protein matrix. However, a more recent study showed that several residues surrounding the Mo—Cu binuclear center are crucial to maintain the structure of the enzyme core, arguing that a 179-atom QM model is needed for accurate results [8]. Unfortunately, this model has not yet been used to investigate the catalytic mechanism of the enzyme. Recently, the previously reported reaction mechanism for the oxidation of CO has been revisited by hybrid quantum mechanics/molecular mechanics (QM/MM) methods [11]. The study supported the formation of a relatively stable thiocarbonate adduct, lying somehow off-path the catalytic cycle. In fact, the authors suggested that the reaction has to follow a reverse process in order to release the CO_2 product from the thiocarbonate species.

By this study, we want to deepen the understanding of the controversial role of the thiocarbonate adduct in MoCu-CODH. We do so by using small and large-size QM cluster models of the active site, as well as QM/MM models of the whole enzyme. We show that the size of the model, as well as the level of theory, have a strong influence on the computed stability of the investigated intermediates. Most importantly, the thiocarbonate species turned out not represent a potential-energy well when the enzyme is described using large QM-cluster or all-atom QM/MM models, together with sufficiently large basis sets.

2. Computational methods

All calculations were based on the crystal structure of the enzyme in its oxidized form (PDB ID: 1N5W) [2]. MoCu-CODH is a dimer of heterotrimers. In this study, only the large subunit of one dimer, containing the active site, was considered.

2.1. QM-cluster calculations (*Models 1 & 2*)

The smallest **Model 1** is the model used by Hofmann and coworkers and it is constituted by 24 atoms, representing the two metal centers and their first coordination spheres (shown in Fig. 2). **Model 2** is appreciably larger and it was created and validated by Rokhsana et al. [8]. It is composed of 179 atoms with CO substrate and 189 atoms when considering the *n*-butylisocyanide inhibitor. In addition to the atoms in **Model 1**, it includes also models of residues Gln240, Gly269–Gly272, Val384–Phe390, Tyr568–Ser570, and Glu763 (also shown in Fig. 2). All QM calculations were carried out with DFT. Geometry optimizations were performed *in vacuo* using the TURBOMOLE 7.1 software with the BP86 method [12–14]. For **Model 1**, the calculations employed the def2-TZVP basis set, whereas for **Model 2**, the def2-TZVP basis set was used for the first coordination shell (31 or 41 atoms), whereas the rest of the model was described by the smaller def2-SVP basis set [12–15]. All calculations included Grimme's dispersion correction with Becke–Johnson damping (D3(BJ)) [16]. The resolution-of-identity technique was employed to accelerate the calculations [17]. For **Model 2**, all Cα atoms were kept frozen at the crystallographic positions together with one attached proton each. Finally, for both models, single-point energy calculations at the BP86 level were performed on the optimized geometries using the COSMO continuum solvent model [18], to evaluate solvation effects on the relative energies of the intermediates ($\varepsilon = 4$). Moreover, the hybrid GGA functional B3LYP-D3(BJ) [19,20], together with the large def2-TZVPD basis set, was employed to obtain more accurate energies.

2.2. QM/MM calculations (*Model 3*)

The QM/MM calculations required a detailed analysis of all protonable residues in the protein, based on calculations with PROPKA [21], studies of the hydrogen-bond pattern, the solvent accessibil-

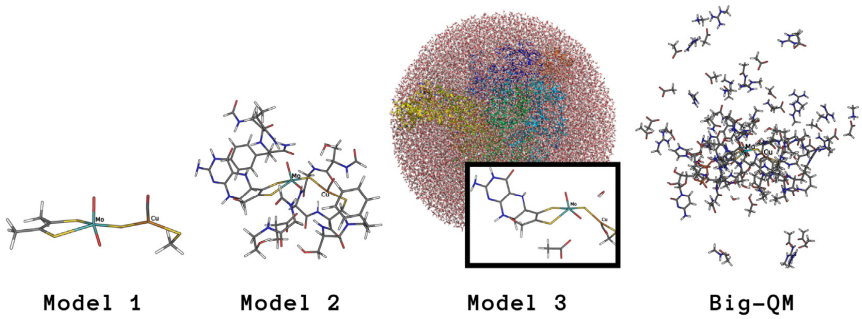


Fig. 2. Models used for the description of the MoCu-dependent CO-dehydrogenase. Color code of atoms: cyan, molybdenum; brown, copper; yellow, sulphur; red, oxygen; grey, carbon; white, hydrogen.

ity and the possible formation of ionic pairs. Based on this analysis, we decided to let all Arg, Lys, Asp, and Glu residues be charged, with exception of Glu29 and Glu488 that were protonated on OE2, whereas Asp684 was protonated on OD1. Cysteine ligands coordinating to metals were deprotonated. Among the His residues, His61, 339, 766 and 793 were protonated on the ND1 atom, His177, 178, 210, 213, 243, 700, 753, 754 and 788 were assumed to be protonated on NE2 atom, whereas the other His residues were assumed to be doubly protonated. The protein was solvated with water molecules forming a sphere with a radius of 60 Å around the geometric center of the protein. Added protons and water molecules were then optimized by a 1-ns simulated-annealing molecular-dynamics simulation, followed by a minimization.

The QM/MM calculations were performed with the ComQum software [22,23]. In this approach, the protein and the solvent are split into three subsystems: System 1 is the QM subsystem (containing 49 or 65 atoms considering CO or *n*-butylisocyanide, respectively, shown in the inset in Fig. 2 and in Fig. S1 in the Supporting Information), System 2 consists of all residues with any atom within 6 Å of any atom in System 1, whereas System 3 contains the remaining part of the protein and the water molecules. During the QM/MM geometry optimizations, System 1 was relaxed by QM methods, whereas Systems 2 and 3 were kept fixed at the crystallographic coordinates. In the QM calculations, System 1 was represented by a wavefunction whereas all the other atoms were represented by an array of partial point charges.

Covalent bonds between the QM and MM systems were truncated using the hydrogen link-atom approach [24]. The QM system is capped with hydrogen atoms (hydrogen link atoms, HL), the position of which are linearly related to the corresponding carbon atoms (carbon link atoms, CL) in the full system [22]. The CL atoms are not included in the point-charge model. The total QM/MM energy in ComQum is calculated as

$$E_{QM/MM} = E^{HL}_{QM1+ptch23} + E^{CL}_{MM123,q1=0} - E^{HL}_{MM1,q1=0} \quad (1)$$

where $E^{HL}_{QM1+ptch23}$ is the QM energy of System 1 truncated by HL atoms and embedded in the set of point charges modeling Systems 2 and 3. $E^{HL}_{MM1,q1=0}$ is the MM energy of System 1, truncated by HL atoms, without any electrostatic interactions. $E^{CL}_{MM123,q1=0}$ is the classical energy of the whole system, with CL atoms and with the charges present in System 1 set to zero, in order to avoid double counting of the electrostatic interactions. Thus, ComQum employs a subtractive scheme with electrostatic embedding and van der Waals link-atom corrections [25].

The QM calculations were carried out using TURBOMOLE 7.1 software [12]. Geometry optimizations were performed using the BP86 functional [13,14] in combination with def2-TZVP basis set [15], including empirical dispersion correction D3(BJ) [16] and the resolution-of-identity technique to accelerate the calculations [17]. The MM calculations were carried out by means of the Amber software [26], using the Amber ff14SB force field for the protein [27], and the general Amber force field [28] with restrained electrostatic potential (RESP) charges [29] for carbon monoxide, *n*-butylisocyanide and MCD. The two Fe_2S_2 clusters were described with RESP charges and a non-bonded model (they are kept fixed in the calculations). Single-point calculations at B3LYP-D3(BJ)/def2-TZVPD [19,20,15] level were also run on the optimized geometries.

2.3. Big-QM calculation

More accurate QM/MM energies were also obtained by means of the Big-QM technique [30,31]. This approach was developed to reduce the dependence of QM/MM energies on the size of the QM system and on the arbitrariness of the choice of the latter [32], based on the knowledge that QM/MM calculations converge faster than QM-only ones, but models of about 1000 QM atoms are needed to obtain convergence of the energies [33–36]. The minimal QM system (System 1) was extended with all chemical groups with at least one atom within 6.0 Å and junctions were moved two amino-acids away from each residue in the minimal QM system. In addition, all buried charges inside the protein were included, with exception of the two iron–sulfur clusters (see Supporting Information Table S2). The resulting Big-QM model is constituted by 1060 atoms when the *n*-butylisocyanide ligand is present in the system. It is constituted by 990 atoms when the CO ligand is present instead (see Fig. 2). The Big-QM calculations were performed on coordinates from the QM/MM optimizations of **Model 3**, with a surrounding point-charge model, at the BP86-D3(BJ)/def2-SVP level. The multipole-accelerated resolution-of-identity *J* approach (marij keyword) was employed to accelerate the calculations [37]. The resulting energies were corrected with a QM/MM MM term (cf. Eq. (1)) for the Big-QM region:

$$E_{MM} = E^{CL}_{MM123,q1=0} - E^{HL}_{MM1,q1=0} \quad (2)$$

Finally, the energies were also corrected by taking into consideration the B3LYP-D3(BJ)/def2-TZVPD functional and basis set effects, using calculations with the standard QM/MM QM system with a point-charge model of the surroundings:

$$E_{corr} = E^{B3LYP/TZVPD}_{QM1,ptch23} - E^{BP86/SVP}_{QM1,ptch23} \quad (3)$$

3. Results

In this work, we focus on the first steps of the CO-oxidation mechanism and on the associated side reaction described by Hof-

mann and coworkers [4]. In particular, we consider the thiocarbonate adduct (**D** in Fig. 1) and the intermediates that give rise to it (**A**, **B** and **C** in Fig. 1).

Structural comparison of each of these species, optimized using models **1**–**3**, showed great structural similarity (see Supporting Information, Table S3). The mean absolute deviation for the bonded distances in that table between **Model 1** or **Model 2** and **Model 3** is only 0.03 and 0.02 Å, respectively. The maximum deviation is 0.10 Å, observed both for the Cu—C(CO) bond in intermediate **A** and for the Mo—O_{eq} bond in intermediate **D** for **Model 1**.

Larger differences are observed when comparing distances between non-bonded atoms. In particular, the position of the CO substrate with respect to the putative nucleophile O_{eq} in **A** depends strongly on the model size. As the description of the active site environment becomes more and more complete, the C···O_{eq} distance becomes considerably shorter (3.54, 2.62 and 2.26 Å for **Models 1**, **2** and **3**, respectively). The latter result suggests that the protein surrounding the active site has a role to assist the correct positioning of the substrate with respect to the incipient nucleophile. The presence of the aromatic ring of Phe390 in front of the Cu ion might be responsible of this effect, as it was proposed to influence the nature of the interaction between the copper ion and a weakly-coordinated water molecule in the crystallographic experiments [8]. Moreover, mutation of Phe390 to Tyr was reported to partially impair CO-oxidation catalysis [38]. This illustrates the importance of going beyond the first coordination sphere when details on CO binding and processing need to be modeled.

The relevance of the second coordination sphere and of the long-range effects of the protein matrix becomes evident also when the relative energies of the various intermediates are considered (see Table 1). We focus first on the stability of **D**, given the key relevance of this thiocarbonate intermediate in previous theoretical investigations of MoCu-CODH. The computed ΔE_{DA} between **D** and the parent adduct **A** is large in the case of **Model 1**, which is not surprising given the results of previous studies based on small models of the active site ($\Delta E_{DA}^{Model1} = -47$ kJ/mol). However, the energy difference becomes much smaller when the size of the model increases: $\Delta E_{DA}^{Model2} = -25$ kJ/mol, $\Delta E_{DA}^{Model3} = -8$ kJ/mol. The latter energy difference shrinks further when Big-QM/MM corrections are adopted ($\Delta E_{DA}^{BigQM} = -5$ kJ/mol).

As for intermediate **B**, the ΔE_{BA} energy differences appear to depend much less on the size of the adopted model, with a variation of up to 23 kJ/mol. More specifically, ΔE_{BA} is essentially independent of the second coordination sphere (**Models 1** and **2** give nearly identical energies), whereas inclusion of long-range interactions influences the energy difference somewhat more (cf. the QM/MM results in Table 1).

The effect of the second coordination sphere of the metals turned out to be more pronounced when the stability of intermediate **C** is considered. In fact, for **Model 1**, this intermediate is predicted to be very stable ($\Delta E_{CA}^{Model1} = -60$ kJ/mol). However, the computed energy difference becomes much smaller when the active site model includes the second coordination sphere ($\Delta E_{CA}^{Model2} = -22$ kJ/mol), and it shrinks further when the whole protein is considered ($\Delta E_{CA} = -2$ kJ/mol and -12 kJ/mol in the case of QM/MM calculations without or with Big-QM corrections, respectively).

Table 1
Relative energies (kJ/mol) of the intermediates **A**, **B**, **C** and **D**, calculated using the QM-cluster models, the QM/MM model and the Big-QM approach at B3LYP-D3(BJ)/def2-TZVPD level.

	A	B	C	D
Model 1	0.0	−5.0	−60.2	−46.9
Model 2	0.0	−2.1	−21.8	−25.1
Model 3	0.0	−24.7	−2.1	−8.4
Big-QM	0.0	−15.5	−11.7	−5.0

Table 2
Relative energies (kJ/mol) of the intermediates **E** and **F**, calculated using the QM-cluster models, the QM/MM model and the Big-QM approach at B3LYP-D3(BJ)/def2-TZVPD level.

	E	F
Model 1	0.0	−89.9
Model 2	0.0	−88.3
Model 3	0.0	−97.1
Big-QM	0.0	−65.7

In analogy with some of the previous studies on the catalytic mechanism of the MoCu-CODH [4,11], we also carried out ancillary calculations on the inhibition of the enzyme by *n*-butylisocyanide (|C≡N—R, a ligand that is isoelectronic with the C≡O substrate). This was done to assess the capability of our larger models to reproduce the deep potential well associated with the formation of a thiocarbamate adduct, which is at the basis of the experimentally observed inhibition. Analogously to previous reports, the stability of the thiocarbamate adduct (**F** in Fig. 1) was evaluated by considering the parent encounter complex as a reference (**E** in Fig. 1). As expected, our calculations show for all four models a high stability of the thiocarbamate product **F** with respect to the reference **E**: $\Delta E_{FE} = -66$ to -97 kJ/mol, as can be seen in Table 2. Thus, our large-sized models – and in particular the Big-QM calculations – provide energy landscapes that are compatible both with the inhibitory role of the thiocarbamate adduct and with a non-inhibitory role of the thiocarbonate intermediate.

4. Discussion

In this work, we evaluated the relative stability of the thiocarbonate intermediate in MoCu-CODH, whose formation is thought to represent a key intermediate during catalytic CO oxidation. To this end, a hierarchy of QM-cluster models of different sizes, as well as QM/MM models were adopted to describe the enzyme.

Our results from a small QM-cluster model including only the metal ions and their first coordination spheres suggest that the thiocarbonate adduct corresponds to a deep well on the potential-energy surface. Notably, the early studies by Hofmann, Siegbahn, and coworkers [3,4], reported energies even more negative for this intermediate (up to −93 kJ/mol as compared to the encounter complex). The latter models not only bear the shortcomings from complete or partial neglect of the second coordination sphere. In fact, the lack of dispersion energy corrections in those studies was later reported to contribute to inaccuracies in the calculation of energy differences involving the thiocarbonate intermediate [39]. In a recent QM/MM study by Xu and Hirao, the thiocarbonate conformer was predicted to be less stable. However, according to those authors, this species would still represent a thermodynamic sink, with a QM/MM ΔE_{DA} energy difference of −43 kJ/mol. This would drive the enzyme somehow off-path when the **C** intermediate is reached, as the rearrangement of the latter to give the thiocarbonate adduct was reported to be energetically favorable. Notably, the basis set used in the latter QM/MM study for geometry optimizations was of double-zeta quality both for metal atoms and for light atoms. Previous studies on similar Mo-containing systems have suggested that the choice of the basis set is critical to obtain accurate energies [40,41]. An evaluation of the influence of the quality of the basis set used during the optimization on our system, evidenced an effect of 19 kJ/mol on the final energies even when the involved intermediates retain a high degree of structural similarity (see Supporting Information, Tables S4 and S5).

In conclusion, accurate modeling of the protein surroundings around the active site, coupled with consideration of dispersion effects and employment of sufficiently large basis sets allowed us to reveal that the thiocarbonate adduct does not play the role of an unusual rate-limiting intermediate in MoCu-CODH. Instead, our results are in line with a recent proposal [9] that intermediates **B**, **C** and **D** may be readily interconverted during biocatalysis.

Acknowledgements

This investigation has been supported by grants from the Swedish research council (project 2018-05003). The computations were performed on computer resources provided by the Swedish National Infrastructure for Computing (SNIC) at Lunarc at Lund University.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jcat.2019.02.032.

References

[1] G. Mörsdorf, K. Frunzke, D. Gadkari, O. Meyer, Microbial growth on carbon monoxide, Biodegradation 3 (1992) 61–82.
[2] H. Dobbek, L. Gremer, R. Kiefersauer, R. Huber, O. Meyer, Catalysis at a dinuclear [CuSMo(O)OH] cluster in a CO dehydrogenase resolved at 1.1-Å resolution, Proc. Natl. Acad. Sci. USA 99 (2002) 15971–15976.
[3] P.E.M. Siegbahn, A.F. Shestakov, Quantum chemical modeling of CO oxidation by the active site of molybdenum CO dehydrogenase, J. Comput. Chem. 26 (2005) 888–898.
[4] M. Hofmann, J.K. Kassube, T. Graf, The mechanism of Mo-/Cu-dependent CO dehydrogenase, J. Biol. Inorg. Chem. JBIC 10 (2005) 490–495.
[5] R. Hille, S. Dingwall, J. Wilcoxen, The aerobic CO dehydrogenase from *Oligotropha carboxidovorans*, J. Biol. Inorg. Chem. JBIC 20 (2015) 243–251.
[6] M. Gnida, R. Ferner, L. Gremer, O. Meyer, W. Meyer-Klaucke, A novel binuclear [CuSMo] cluster at the active site of carbon monoxide dehydrogenase: characterization by X-ray absorption spectroscopy, Biochemistry 42 (2003) 222–230.
[7] J. Wilcoxen, R. Hille, The Hydrogenase activity of the molybdenum/copper-containing carbon monoxide dehydrogenase of *Oligotropha carboxidovorans*, J. Biol. Chem. 288 (2013) 36052–36060.
[8] D. Rokhsana, T.A. Large, M.C. Dienst, M. Retegan, F. Neese, A realistic *in silico* model for structure/function studies of molybdenum–copper CO dehydrogenase, J. Biol. Inorg. Chem. JBIC 21 (2016) 491–499.
[9] M. Shanmugam, J. Wilcoxen, D. Habel-Rodriguez, G.E. Cutsail III, M.L. Kirk, B.M. Hoffman, R. Hille, ^{13}C and $^{63,65}Cu$ ENDOR studies of CO dehydrogenase from *Oligotropha carboxidovorans*. Experimental evidence in support of a copper–carbonyl intermediate, J. Am. Chem. Soc. 135 (2013) 17775–17782.
[10] C. Gourlay, D.J. Nielsen, J.M. White, S.Z. Knottenbelt, M.L. Kirk, C.G. Young, Paramagnetic active site models for the molybdenum-copper carbon monoxide dehydrogenase, J. Am. Chem. Soc. 128 (2006) 2164–2165.
[11] K. Xu, H. Hirao, Revisiting the catalytic mechanism of Mo–Cu carbon monoxide dehydrogenase using QM/MM and DFT calculations, Phys. Chem. Chem. Phys. 20 (2018) 18938–18948.
[12] TURBOMOLE V7.1 2016, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from http://www.turbomole.com.
[13] A.D. Becke, Density-functional exchange-energy approximation with correct asymptotic behavior, Phys. Rev. A 38 (1988) 3098–3100.
[14] J.P. Perdew, Density-functional approximation for the correlation energy of the inhomogeneous electron gas, Phys. Rev. B 33 (1986) 8822.
[15] F. Weigend, R. Ahlrichs, Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy, Phys. Chem. Chem. Phys. 7 (2005) 3297–3305.
[16] S. Grimme, S. Ehrlich, L. Goerigk, Effect of the damping function in dispersion corrected density functional theory, J. Comput. Chem. 32 (2011) 1456–1465.
[17] K. Eichkorn, F. Weigend, O. Treutler, R. Ahlrichs, Auxiliary basis sets for main row atoms and transition metals and their use to approximate Coulomb potentials, Theor. Chem. Acc. 97 (1997) 119–124.
[18] A. Klamt, G. Schüürmann, COSMO a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient, J. Chem. Soc., Perkin Trans. 2 (1993) 799–805.
[19] A.D. Becke, A new mixing of Hartree-Fock and local density-functional theories, J. Phys. Chem. 98 (1993) 1372–1377.
[20] C. Lee, W. Yang, R.G. Parr, Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density, Phys. Rev. B 37 (1988) 785–789.
[21] M.H.M. Olsson, C.R. Søndergaard, M. Rostkowski, J.H. Jensen, PROPKA3: consistent treatment of internal and surface residues in empirical pK_a predictions, J. Chem. Theory Comput. 7 (2011) 525–537.
[22] U. Ryde, The coordination of the catalytic zinc ion in alcohol dehydrogenase studied by combined quantum chemical and molecular mechanical calculations, J. Comput. Aid. Mol. Des. 10 (1996) 153–164.
[23] U. Ryde, M.H.M. Olsson, Structure, strain, and reorganization energy of blue copper models in the protein, Int. J. Quantum Chem. 81 (2001) 335–347.
[24] N. Reuter, A. Dejaegere, B. Maigret, M. Karplus, Frontier bonds in QM/MM methods: a comparison of different approaches, J. Phys. Chem. A 104 (2000) 1720–1735.
[25] L. Cao, U. Ryde, On the difference between additive and subtractive QM/MM calculations, Front. Chem. 6 (2018) 89.
[26] D. Case, V. Babin, J. Berryman, R. Betz, Q. Cai, D. Cerutti, T. Cheatham III, T. Darden, R. Duke, H. Gohlke, et al., "Amber 14", 2014. http://www.ambermd.org.
[27] J.A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K.E. Hauser, C. Simmerling, ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB, J. Chem. Theory Comput. 11 (2015) 3696–3713.
[28] J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, D.A. Case, Development and testing of a general amber force field, J. Comput. Chem. 25 (2004) 1157–1174.
[29] C.I. Bayly, P. Cieplak, W.D. Cornell, P.A. Kollman, A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the resp model, J. Phys. Chem. 97 (1993) 10269–10280.
[30] L. Hu, P. Söderhjelm, U. Ryde, Accurate reaction energies in proteins obtained by combining QM/MM and large QM calculations, J. Chem. Theory Comput. 9 (2013) 640–649.
[31] S. Sumner, P. Söderhjelm, U. Ryde, Effect of geometry optimizations on QM-cluster and QM/MM studies of reaction energies in proteins, J. Chem. Theory Comput. 9 (2013) 4205–4214.
[32] L. Hu, J. Eliasson, J. Heimdal, U. Ryde, Do quantum mechanical energies calculated for small models of protein-active sites converge?, J Phys. Chem. A 113 (2009) 11793–11800.
[33] C.V. Sumowski, C. Ochsenfeld, A convergence study of QM/MM isomerization energies with the selected size of the QM region for peptidic systems, J. Phys. Chem. A 113 (2009) 11734–11741.
[34] D. Flaig, M. Beer, C. Ochsenfeld, Convergence of electronic structure with the size of the QM region: example of QM/MM NMR shieldings, J. Chem. Theory Comput. 8 (2012) 2260–2271.
[35] R.-Z. Liao, W. Thiel, Convergence in the QM-only and QM/MM modeling of enzymatic reactions: A case study for acetylene hydratase, J. Comput. Chem. 34 (2013) 2389–2397.
[36] S. Roßbach, C. Ochsenfeld, Influence of coupling and embedding schemes on QM size convergence in QM/MM approaches for the example of a proton transfer in DNA, J. Chem. Theory Comput. 13 (2017) 1102–1107.
[37] M. Sierka, A. Hogekamp, R. Ahlrichs, Fast evaluation of the Coulomb potential for electron densities using multipole accelerated resolution of identity approximation, J. Chem. Phys. 118 (2003) 9136–9148.
[38] P. Kaufmann, B.R. Duffus, C. Teutloff, S. Leimkühler, Functional studies on *Oligotropha carboxidovorans* molybdenum–copper CO dehydrogenase produced in *Escherichia coli*, Biochemistry 57 (2018) 2889–2901.
[39] P.E.M. Siegbahn, The effect of backbone constraints: the case of water oxidation by the oxygen-evolving complex in PSII, ChemPhysChem 12 (2011) 3274–3280.
[40] J.-L. Li, R.A. Mata, U. Ryde, Large density-functional and basis-set effects for the DMSO reductase catalyzed oxo-transfer reaction, J. Chem. Theory Comput. 9 (2013) 1799–1807.
[41] J. Li, M. Andrejic, R.A. Mata, U. Ryde, A computational comparison of oxygen atom transfer catalyzed by dimethyl sulfoxide reductase with Mo and W, Eur. J. Inorg. Chem. 2015 (2015) 3580–3589.

Paper III

A. Rovaletti, M. Bruschi, G. Moro, U. Cosentino, C. Greco and U. Ryde
Theoretical Insights into the Aerobic Hydrogenase Activity of Molybdenum–Copper CO Dehydrogenase
Inorganics, 2019, 7.11, pp. 135

inorganics

MDPI

Article

Theoretical Insights into the Aerobic Hydrogenase Activity of Molybdenum–Copper CO Dehydrogenase

Anna Rovaletti [1], Maurizio Bruschi [1], Giorgio Moro [2], Ugo Cosentino [1] and Claudio Greco [1,*] and Ulf Ryde [3,*]

1 Department of Earth and Environmental Sciences, University of Milano-Bicocca, Piazza della Scienza 1, 20126 Milan, Italy; a.rovaletti@campus.unimib.it (A.R.); maurizio.bruschi@unimib.it (M.B.); ugo.cosentino@unimib.it (U.C.)

2 Department of Biotechnology and Biosciences, University of Milano-Bicocca, Piazza della Scienza 2, 20126 Milan, Italy; giorgio.moro@unimib.it

3 Department of Theoretical Chemistry, Lund University, Chemical Centre, P.O. Box 124, SE-221 00 Lund, Sweden

* Correspondence: claudio.greco@unimib.it (C.G.); ulf.ryde@teokem.lu.se (U.R.)

Received: 14 September 2019; Accepted: 4 November 2019; Published: 9 November 2019

Abstract: The Mo/Cu-dependent CO dehydrogenase from *O. carboxidovorans* is an enzyme that is able to catalyse CO oxidation to CO_2; moreover, it also expresses hydrogenase activity, as it is able to oxidize H_2. Here, we have studied the dihydrogen oxidation catalysis by this enzyme using QM/MM calculations. Our results indicate that the equatorial oxo ligand of Mo is the best suited base for catalysis. Moreover, extraction of the first proton from H_2 by means of this basic centre leads to the formation of a Mo–OH–Cu^{I}H hydride that allows for the stabilization of the copper hydride, otherwise known to be very unstable. In light of our results, two mechanisms for the hydrogenase activity of the enzyme are proposed. The first reactive channel depends on protonation of the sulphur atom of a Cu-bound cysteine residues, which appears to favour the binding and activation of the substrate. The second reactive channel involves a frustrated Lewis pair, formed by the equatorial oxo group bound to Mo and by the copper centre. In this case, no binding of the hydrogen molecule to the Cu center is observed but once H_2 enters into the active site, it can be split following a low-energy path.

Keywords: CO dehydrogenase; dihydrogen; hydrogenase; quantum/classical modeling; density functional theory

1. Introduction

Carbon monoxide dehydrogenases (CODHs) have proved to be an essential component for the biogeochemical carbon monoxide (CO) consumption. They contribute to maintenance of sub-toxic concentration of CO in the lower atmosphere by processing approximately 2×10^8 tons of it annually [1]. To date, only two enzymes have been found to be able to use CO as carbon and energy source [2]. The first type is represented by the oxygen-sensitive Ni,Fe-dependent CO dehydrogenase enzyme (NiFe-CODH), found in anaerobic bacteria and archea, whereas in aerobic carboxido-bacteria, such as *Oligotropha carboxidovorans*, the same function is performed by the oxygen-tolerant MoCu-dependent CO dehydrogenase (MoCu-CODH) enzyme [3]. These bacterial enzymes catalyze the oxidation of CO to CO_2 following the reaction:

$$CO + H_2O \longrightarrow CO_2 + 2H^+ + 2e^-$$

In addition to this reaction, both enzymes show the ability to catalyze other reactions, although at lower activity. Reduction of CO_2, following the reverse mechanism, is explicated by NiFe-CODH, while the ability of oxidizing H_2 to protons has been reported for MoCu-CODH [3,4].

The latter belongs to the xanthine oxidase family of enzymes, whose members are usually characterised by a mononuclear redox-active site, $LMo^{VI}OS(OH)$, with a square-pyramidal coordination geometry [5]. The oxo (O^{2-}) group is an apical ligand and the equatorial plane has a dithiolene ligand from a molybdopterin–cytosine dinucleotide (MCD) cofactor, a catalytically labile hydroxyl group (OH^-) and a sulphide ion (S^{2-}) [6]. MoCu-CODH presents a noncanonical binuclear active site, in which the sulphido ligand bridges to a second metal, –S–Cu–S-Cys and the equatorial hydroxyl group is deprotonated to form another oxo group. The copper ion is found in a +1 oxidation state and it is also coordinated by a weakly-bound water molecule [7].

The protein is a $(\alpha\beta\gamma)_2$ heterodimer. The active site is located in the large subunit (CoxL; 89 kDa), the medium subunit (CoxM; 30 kDa) contains a FAD cofactor, whereas the small one (CoxS; 18 kDa) possesses two [2Fe–2S] iron–sulfur clusters [7].

The protein resting state is characterised by a $Mo^{VI}(=O)SCu^{I}$ core. Binding of either CO or H_2 is believed to occur at the Cu^{I} centre, allowing the proper placement and activation of the substrate [8,9]. The subsequent steps lead to the transfer of two electrons to the Mo ion. This process is favoured by the presence of a highly delocalised redox-active orbital over the Mo–μS–Cu moiety [10]. The catalytic activity of MoCu-CODH is also promoted by the presence of highly conserved amino acids in the metal second coordination sphere [11]. A glutamate residue (Glu763) positioned in proximity of the equatorial oxygen ligand of molybdenum is believed to promote deprotonation events [9,12]. In addition, a phenylalanine (Phe390), situated in a flexible loop in front of the copper centre, has been proposed to contribute to the proper orientation of the Cu-bound ligand [13].

Notwithstanding the fact that the general features of the catalytic reaction cycle of the enzyme has been clarified thanks to previous experimental and theoretical efforts, a detailed understanding of the reaction mechanisms for the oxidation of both CO and H_2 has not yet been reached. As for H_2, a mechanism for the dihydrogen molecule splitting was proposed by Wilcoxen and Hille (see Figure 1) [9]. According to electron paramagnetic resonance (EPR) spectroscopy studies, H_2 is believed to coordinate Cu^{I} in a side-on fashion. Once the hydrogen is bound, it would be activated by the metal and could easily be split by extraction of a proton thanks to a nearby base. Subsequently, a second proton would leave the active site and the two electrons would be transferred to the Mo centre. To date, no theoretical data have been reported in support of the experimental proposal. Only one previous theoretical work investigating the hydrogenase activity has been published so far [14]. In this study, it was unexpectedly found that the copper ion could not bind the H_2 substrate. Therefore, the authors suggested that changes in the active site are needed, for example, by protonation events, before the catalytic cycle can start. In such context, it is interesting to notice that, to the best of our knowledge, side-on binding of H_2 to molecular species containing a Cu(I) ion has been previously reported only in a couple of gas-phase case studies, that is, H_2–CuCl [15] and H_2–CuF [16]. Somehow, the interaction between H_2 and the enzyme appears to find a closer parallel with the case of porous materials containing electron-rich copper ions, such as certain zeolites proposed for H_2 storage applications [17–19] and Prussian-blue analogues [20].

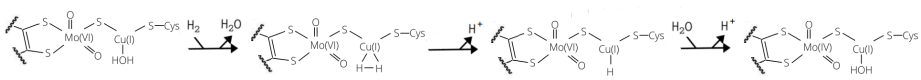


Figure 1. Reaction mechanism for H_2 activation proposed by Wilcoxen and Hille on the basis of experimental results [9].

In the present work we present a plausible mechanistic picture for the MoCu-CODH hydrogenase activity, taking into account the effects of the protein matrix by means of hybrid quantum mechanical/molecular mechanical (QM/MM) enzyme models.

2. Results and Discussion

2.1. Study of H_2 Binding Modes to the Copper Centre

We started with an investigation of the most plausible binding sites for the H_2 substrate. Considering the resting state of the protein, two types of coordination to the copper centre were considered (**1HR** and **2HR** in Figure 2). We find that H_2 can bind to the Cu ion both when the latter is bi-coordinated (μS–Cu–S-Cys388) and also when the Cu coordination sphere includes a weakly bound water molecule, as proposed in previous investigations of the enzyme resting state [7,13]. The hydrogen molecule is found to be slightly activated, with a H–H distance of 0.77 Å in both cases (the equilibrium distance of H_2 at BP86/def2-TZVP level is 0.75 Å).

For both binding modes, we estimated the binding energy of H_2 to Cu^I by comparing the energy of the coordinated structure to the corresponding one in which H_2 is not linked to copper but it is already present in the active site (**1HNoB** and **2HNoB** in Figure 3). The resulting QM/MM total energy differences were found to be +46.9 and +46.4 kJ/mol, in the absence (**1H**) or presence of the coordinated water molecule (**2H**), respectively (see Table 1).

Table 1. Relative energies (kJ/mol) calculated using the quantum mechanical/molecular mechanical (QM/MM) approach at B3LYP-D3(BJ)/def2-TZVPD level.

	NoB	R	P1	TS	P	TS2	P2
1H	−46.9	0.0	57.7		−2.9	31.8	−130.1
2H	−46.4	0.0	83.3				
3H	−10.9	0.0		29.3	28.0	56.4	−90.8
1H-FLP	0.8	0.0		45.2	44.4	79.1	−82.8

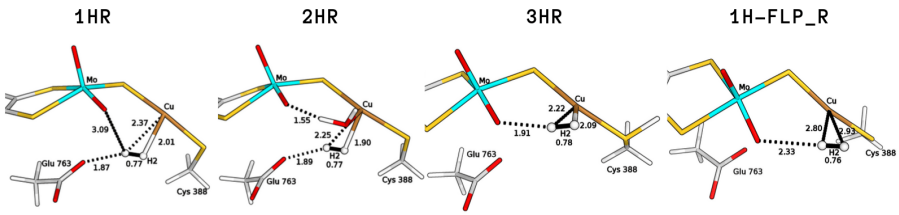


Figure 2. Binding modes of H_2 to Cu^I characterised by means of QM/MM calculations considering different copper coordination modes (**1HR**, bi-coordinated; **2HR**, tri-coordinated) and active site protonation states (**3HR**, protonated Cys388) and the stable intermediate **1H-FLP_R**. All distances are in Å.

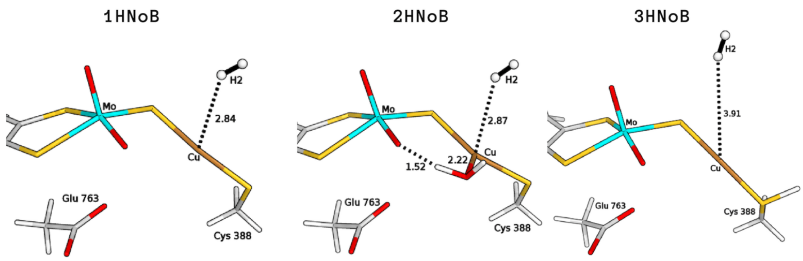


Figure 3. QM/MM optimised structures of intermediates where the H_2 molecule is present in the active site but not bound to the Cu ion, considering different copper coordination modes (**1HNoB**, bi-coordinated; **2HNoB**, tri-coordinated), active-site protonation states (**3HNoB**, protonated Cys388). All distances are in Å.

Thus, formation of these reactants was found to be strongly unfavourable from an energetic point of view. Therefore, we contemplated the possibility that a sterically hindered Lewis acid–Lewis base pair (a so-called frustrated Lewis pair, FLP), may form in the active site, which would eventually lead to heterolytic H_2 cleavage. Indeed, combination of Lewis acid and bases have been demonstrated to activate a wide range of small molecules [21]. Moreover, the plausible involvement of an FLPs in enzymatic catalysis has been recently proposed for the H_2 splitting by [NiFe]-hydrogenase [22] and an FLP might actually be active also in the case of [FeFe]-hydrogenases [23–25].

We found a variant of **1HR** (**1H-FLP_R** in Figure 2), in which H_2 assumes a different position in the active site, such that the Mo–O_{eq} atom and the copper centre appear to be able to act as a FLP. The **1H-FLP_R** reactant is characterised by a H_2 molecule positioned in front of the copper atom, with long Cu–H_{H_2} distances (2.80 and 2.93 Å) and a O_{eq}–H_{H_2} distance of 2.33 Å. **1H-FLP_R** was found to be 0.8 kJ/mol lower in energy than the **1HNoB** species in which the hydrogen molecule is separated from the bimetallic centre.

Based on previous literature, which proposed that the active site could be protonated prior to substrate binding [14,26], we also evaluated the influence of protonation on the coordination of H_2 to the metal. Protonation of the bridging sulphide ion turned out not to favour dihydrogen binding, as no H_2-bound geometry was found on the QM/MM potential energy surface. On the other hand, protonation of the sulphur atom of Cys388 favours coordination of H_2 to Cu. The latter process would lead to the formation of a H_2-coordinated geometry (**3HR** in Figure 2), in which dihydrogen is activated to a H–H bond of 0.78 Å and the computed binding energy is +10.9 kJ/mol.

It is noticeable that energy minimizations of the one-electron reduced counterparts of **1HR**, **2HR** and **3HR** lead to the detachment of the hydrogen molecule from the copper centre. This result is not compatible with the experimental outcomes that identified a H_2-bound $Mo^{V}Cu^{I}$ species by means of EPR experiments [9]. However, since one-electron reduced counterparts do not represent plausible intermediates in the catalytic cycle—there are no evidences that H_2 binding is followed by Mo reduction, prior to substrate oxidation—we have not investigated this aspect any further. The aim of the work is focused on the mechanistic characteristics of the H_2 oxidation reaction, the details of which are described in the subsequent sections.

2.2. Exploring Basic Residues in the Active Site

Next, we investigated which functional group may represent the base for the abstraction of the first proton from the activated hydrogen molecule. We considered the Glu763 residue, which has been suggested to take part into deprotonation events (*vide infra*) and the equatorial oxo ligand of molybdenum. As for the reactants, we considered not only **1H-FLP_R** and **3HR** but also **1HR** and **2HR**. In fact, even though H_2 binding in the latter two cases was computed to be disfavoured, one cannot completely rule out the hypothesis that such adducts have some minor role in dihydrogen oxidation catalysis, a possibility that would be enabled in case the subsequent reaction steps are associated with an overall smooth energy profile.

The two bases would lead to the formation of two copper hydrides (see Figure 4), namely either GluOOH–$Cu^{I}H$ (**P1**) or Mo–OH–$Cu^{I}H$ (**P**). For the **1HR** and **2HR** states (see Figure 2), the oxygen of the glutamate residue is closest (the O_{Glu}–H_{H_2} distance is 1.87 Å and 1.89 Å for **1HR** and **2HR**), while the Mo–oxo ligand is 3.09 and 3.28 Å from the closest H atom of H_2, respectively. However, for **3HR**, O_{eq} is the closest base (the O–H distance is 1.91 Å) while the oxygen of Glu763 is 3.83 Å away.

The search for the GluOOH–Cu^{I}H products was found to be challenging, as the glutamic acid sidechain in the guess geometries turned out to release a proton to reform a Cu-bound H_2 along optimizations. Therefore, to optimize the intermediates **1HP1** and **2HP1**, (see Figure 5) we had to fix the O_{Glu}–H distance; this was done in the hypothesis of the existence of a barrier for proton transfer, although so small that convergence on the desired minimum could be hampered by the choices made on the starting input structures to be optimized. The total energy differences of these structures, with respect to the corresponding reagents, were found to be 57.7 and 83.3 kJ/mol for **1H** and **2H**, respectively. In the case of **3H**, no GluOOH–Cu^{I}H (**3HP1**) product was found.

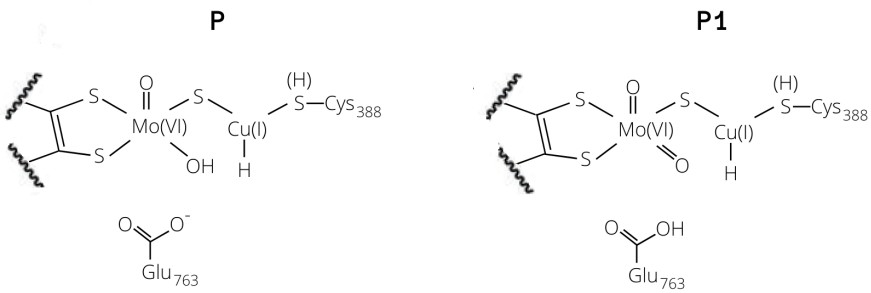


Figure 4. Copper–hydride complexes formed by proton extraction by means of Glu763 (**P1**) or Mo–O_{eq} (**P**).

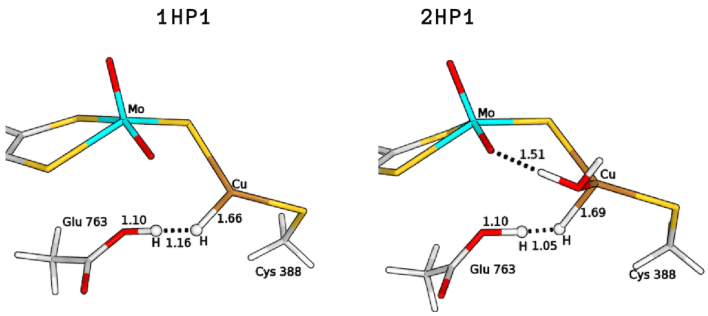


Figure 5. QM/MM optimised structures of the copper hydride complexes formed by proton abstraction by means of Glu763 (**P1**). All distances are in Å.

The product formed by extraction of the proton by the equatorial ligand of Mo (**P**) was not found in the case of **2H**, whereas it corresponds to structures with a total energy difference of −2.9 and 28.0 kJ/mol with respect to the reagents for **1H** and **3H**, respectively (**1H-FLP_P** in Figure 6 and **3HP** in Figure 7). **1H-FLP_P** can also represent the product for the H_2 splitting reaction starting from the **1H-FLP_R** structure. In this case, the resulting Cu–hydride was found to be 44.4 kJ/mol less stable than the reagent.

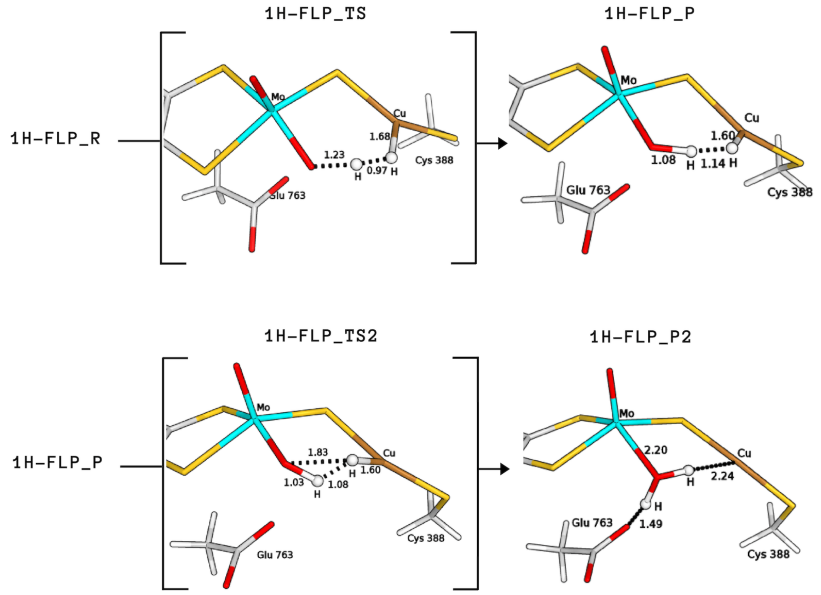


Figure 6. QM/MM structures of intermediates and transition states (in square brackets) of species involved in the catalytic cycle **B** (see Figure 8). All distances are in Å.

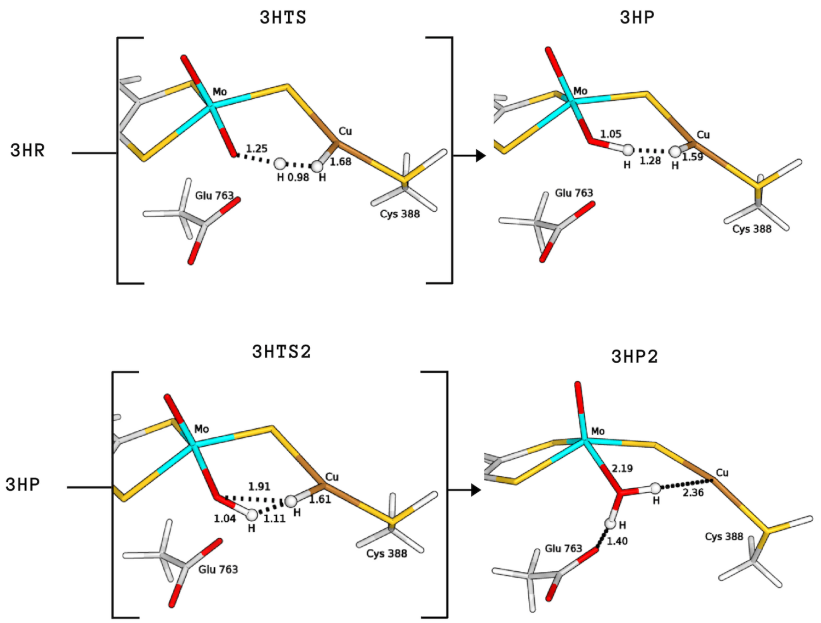


Figure 7. QM/MM structures of intermediates and transition states (in square brackets) of species involved in the catalytic cycle **A** (see Figure 8). All distances are in Å.

2.3. Plausible Activation Mechanisms for H_2 Splitting

Based on the above results on the relative energies of the intermediates potentially involved in the catalysis, we proceeded with the computation of the activation barriers in order to estimate which reaction mechanism would represent the most energetically favourable path for the hydrogenase activity of MoCu-CODH.

Starting from the H_2-bound conformation **1HR**, we have just shown that the equatorial Mo–oxo ligand represents the most favourable basic group for the first proton abstraction, leading to a rather stable product (**1H-FLP_P**). However, despite extensive efforts, we were not able to locate any transition state for the reaction **1HR** → **1H-FLP_P**. This issue is probably due to the large distance between the labile proton and the base. We also considered the **1HR** → **1HP1** reaction because, even if the GluOOH–Cu^{I}H product is not as stable as the Mo–OH–Cu^{I}H product, the whole catalysis may be driven by subsequent exoenergetic reactions. However, again, no transition state was found.

On the other hand, the **1H-FLP_R** species can be linked to the **1H-FLP_P** product through a low-lying transition state (**1H-FLP_TS**), reported in Figure 6. The reaction **1H-FLP_R** → **1H-FLP_P** was found to be endergonic (ΔE = 44.4 kJ/mol) with an associated activation barrier of 45.2 kJ/mol. The latter results support the hypothesis that Cu and the totally oxidised equatorial ligand of molybdenum could play the role of a FLP.

For the complex with a water molecule at the copper centre (**2HR**), we could not locate any transition state geometry that would allow the deprotonation of the hydrogen molecule by the anionic Glu763 residue.

Finally, we considered protonation of Cys388 and found that such protonation can be functional for the hydrogenase activity. In fact, in this case the activation energy for H_2 cleavage involving the equatorial oxo-ligand is only 29.3 kJ/mol (**3HTS** in Figure 7), that is, 15.9 kJ/mol lower than for **1H-FLP**.

Independently of the protonation state of Cys388, our results support a mechanism in which the equatorial oxo group of Mo extracts the first proton. Such process is characterised by a rather low activation barrier that may allow a relatively rapid proton exchange, in accordance with experimental evidence [9]. However, differently from what was hypothesised in the literature [9], we propose that this rapid proton exchange is promoted by the oxo ligand on Mo, rather than by Glu763. Moreover, the equilibrium of this step favours the H_2 reactant complex over copper hydride in both cases, as discussed by Wilcoxen and Hille in light of the pH independence of the reaction. Finally, the intermediates **3HP** or **1H-FLP_P** should be converted to the fully reduced form of the enzyme by the transfer of the Cu-bound hydride to the (now protonated) O_{eq}, giving $Mo^{IV}O(OH_2)Cu^{I}$ [8,13,27] (**P2**; see Figure 8). In both cases, the resulting reduced intermediates were found to be very stable, −90.8 and −82.8 kJ/mol for **3HP2** and **1H-FLP_P2**, respectively (see Table 1 and Figures 6 and 7). Moreover, the reaction barriers for the above two reactions are relatively low (28.4 and 34.7 kJ/mol, respectively). Finally, the reverse reactions are energetically impeded since the reverse activation barriers exceed 140 kJ/mol in both cases (TSs energies are and 147.2 and 161.9 kJ/mol respectively). Again, this is in good agreement with experiments, showing no production of H_2 by the enzyme [9].

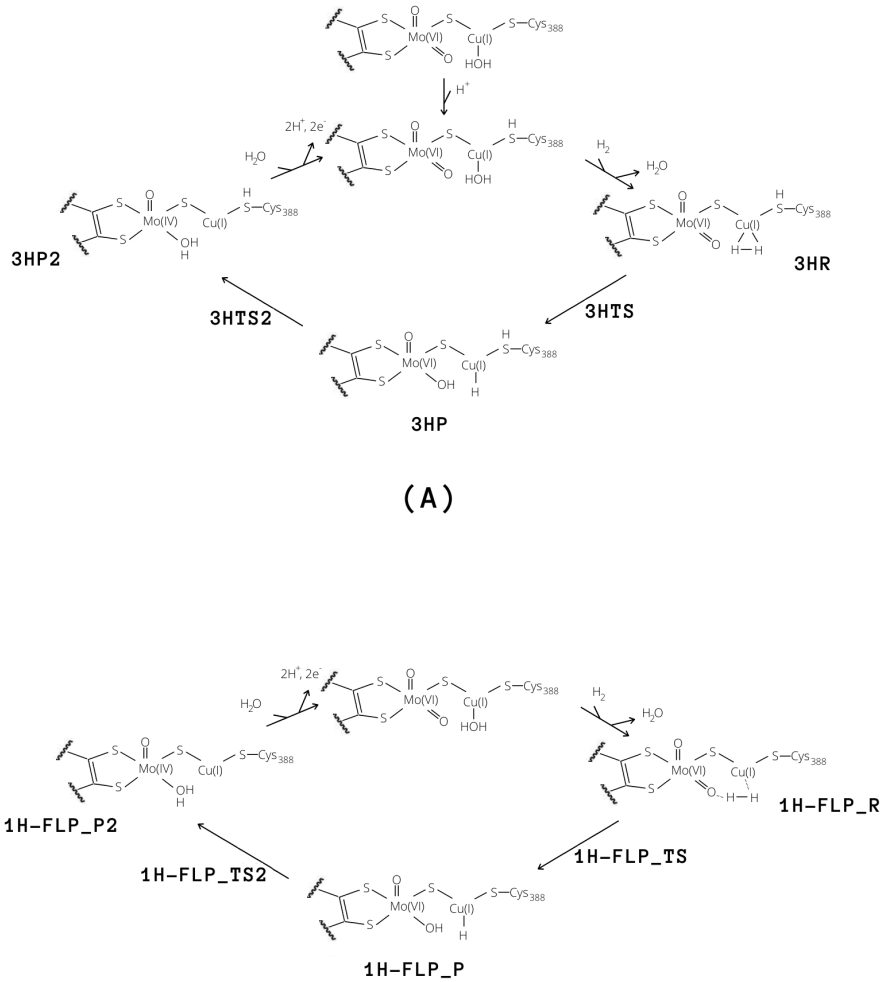


Figure 8. Proposed catalytic cycles for the oxidation of H_2 by means of MoCu-CODH considering either the presence of protonated (**A**) and deprotonated (**B**) Cys388.

3. Methods

3.1. The Protein

All calculations were based on the crystal structure of MoCu-CODH hydrogenase in its oxidised form (PDB ID: 1N5W) [7]. Only the large subunit (CoxL) of one monomer, containing the active site, was considered in this study. The enzyme was set up in the same way as in our previous study [28]. The protonation state of all the residues was determined based on calculations with PROPKA [29] and on studies of the hydrogen-bond pattern, of the solvent accessibility and of the possible formation of ionic pairs. All Arg, Lys, Asp and Glu residues were assumed to be charged, with exception of Glu29 and Glu488 that were protonated on OE2, whereas Asp684 was protonated on OD1. Cysteine ligands coordinating to metals were deprotonated. Among the His residues, His61, 339, 766 and 793 were

protonated on the ND1 atom, His177, 178, 210, 213, 243, 700, 753, 754 and 788 were assumed to be protonated on NE2 atom, whereas the other His residues were assumed to be doubly protonated. The protein was solvated with water molecules, forming a sphere with a radius of 60 Å around the geometric centre of the protein. The added protons and water molecules were then optimised by a 1-ns simulated-annealing molecular-dynamics simulation, followed by energy minimization [28] (for details on the adopted force field, see the next subsection).

3.2. QM/MM Calculations

The QM/MM calculations were performed with the ComQum software [30,31]. According to this approach, the protein and the solvent are split into three subsystems—System 1 corresponds to the QM region and it was relaxed by QM methods. It consisted of the molybdenum ion and its first coordination sphere (two O^{2-}, the bridging sulphide and the MCD cofactor, truncated to exclude the phosphate and cytosine moieties), the copper ion, the ligand molecules H_2, H_2O, as well as the sidechains of residues Cys388 and Glu763 (see Figure 9).

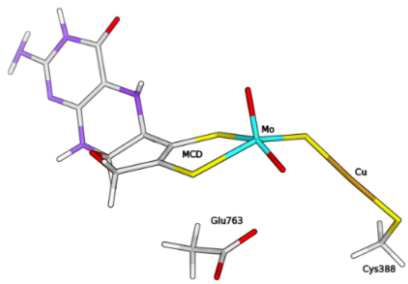


Figure 9. Composition of the QM system within the hybrid QM/MM model of the enzyme, in the form free from H_2/H_2O ligands at the Cu centre.

System 2 consists of all residues with any atom within 6 Å of any atom in System 1, whereas System 3 contains the remaining part of the protein and the water molecules. The latter two systems were kept fixed at the crystallographic coordinates in the QM/MM calculations.

In the QM calculations, System 1 was represented by a wavefunction whereas all the other atoms were represented by an array of partial point charges.

When there is a covalent bond between the QM and MM systems, the QM system was truncated using the hydrogen link-atom approach [32]. The QM system was capped with hydrogen atoms (hydrogen link atoms, HL), the position of which are linearly related to the corresponding carbon atoms (carbon link atoms, CL) in the full system [30]. All atoms were included in the point-charge model, except the CL atoms. ComQum employs a subtractive scheme with electrostatic embedding and van der Waals link-atom corrections [33]. The total QM/MM energy is calculated as

$$E_{\mathrm{QM/MM}} = E^{\mathrm{HL}}_{\mathrm{QM1+ptch23}} + E^{\mathrm{CL}}_{\mathrm{MM123},q_1=0} - E^{\mathrm{HL}}_{\mathrm{MM1},q_1=0} \quad (1)$$

where $E^{\mathrm{HL}}_{\mathrm{QM1+ptch23}}$ is the QM energy of System 1 truncated by HL atoms and embedded in the set of point charges modeling Systems 2 and 3. $E^{\mathrm{HL}}_{\mathrm{MM1},q_1=0}$ is the MM energy of System 1, truncated by HL atoms, without any electrostatic interactions. $E^{\mathrm{CL}}_{\mathrm{MM123},q_1=0}$ is the classical energy of the whole system, with CL atoms and with the charges in System 1 set to zero, to avoid double counting of the electrostatic interactions. The QM calculations were carried out using TURBOMOLE 7.1 software [34]. Geometry optimizations and TS searches (the latter based on minimizations with geometric constraints imposed step-wise to selected atoms, along the putative reactive path) were performed using the BP86 functional [35,36] in combination with def2-TZVP basis set [37]. All calculations included Grimme's

dispersion correction with Becke–Johnson damping (D3(BJ)) [38]. The resolution-of-identity technique was employed to accelerate the calculations [39].

The MM calculations were performed with the Amber software [40], using the Amber ff14SB force field for the protein [41] and the general Amber force field [42] with restrained electrostatic potential (RESP) charges [43] for H_2 and MCD. The two Fe_2S_2 clusters were described with RESP charges and a non-bonded model (they are kept fixed in the calculations).

4. Conclusions

We have studied dihydrogen oxidation by MoCu-CODH using QM/MM calculations. Our results indicate that the equatorial oxo ligand of Mo is a better base than Glu763 during catalysis. Moreover, extraction of the first proton from H_2 by means of this basic centre leads to the formation of a Mo-OH–$Cu^{I}H$ hydride that allows for the stabilization of the copper hydride, otherwise known to be very unstable [44]. It is intriguing to notice that our proposal of a direct involvement of the Mo-bound oxo group in H_2 splitting finds a conceptually similar case in a recently published mechanistic study on the [Fe]-hydrogenase. In the latter enzyme, a deprotonared OH group that belongs to the iron–guanylylpyridinol cofactor and that is in γ-position with respect to the metal-activated H_2, was suggested to be involved in substrate splitting [45]. Such mechanistic picture on the [Fe]-hydrogenase was the result of crystallographic studies that took advantage of the fact that the latter enzyme exists in two forms—an open substrate-accessible form and a closed catalytically active form. This is not the case for MoCu-CODH; however, key details on the proton transfer events following H_2 binding might still be gained at experimental level, most likely by means of application of high-sensitivity infrared spectroscopy techniques (see for example, Reference [46]).

In light of our results, two plausible mechanisms for the hydrogenase activity of the enzyme can be proposed, as reported in Figure 8. The first reactive channel (cycle **A** in Figure 8) depends on protonation of the sulphur atom of Cys388, which appears to favour the binding and activation of the hydrogen molecule. The second reactive channel involves a frustrated Lewis pair, formed by the Mo–O_{eq} oxo group and the copper centre (cycle **B** in Figure 8). In this case, no binding of the hydrogen molecule to the Cu center is observed but once H_2 enters into the active site, it can be split following a low-energy path. All in all, we cannot exclude that both pathways may be active in MoCu-CODH, thus enhancing the overall catalytic turnover. The two mechanisms for H_2 oxidation are characterized by energy profiles that were found to be quite similar (see Figure 10) and although the FLP-based mechanism features a ~20 kJ/mol higher rate-determining barrier, it does not require any preliminary active-site protonation. However, the H^+ source for protonation of Cys388 might be the H_2 molecule itself, as a result of a possible proton transfer following initial H_2 splitting at the active site by means of the FLP mechanism. Alternatively, protonation of Cys388 may occur as a parallel, concerted process at the onset of H_2 oxidation catalysis. Notably, the existence of a pair of distinct proton pathways has been recently suggested for [FeFe]-hydrogenases. There, a "regulatory" protonation of the [4Fe–4S] portion of the hexanuclear active site was proposed to occur in concomitance with a "catalytic" protonation step occurring at the di-iron portion of the same active site [47]. The results presented in the latter study and—hopefully—those discussed in the present paper are likely to provide useful information for future efforts to deepen insights into proton transfer mechanisms from and to the MoCu-CODH active site.

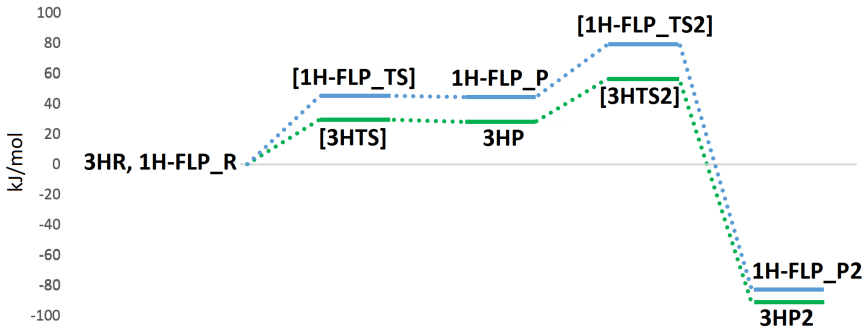


Figure 10. QM/MM energy profiles (in kJ/mol) for the oxidation of H_2 in MoCu-CODH, considering either the presence of protonated (profile traced in green) or deprotonated (profile traced in blue) Cys388.

Author Contributions: Conceptualization, C.G. and U.R.; investigation and formal analysis, A.R., C.G. and U.R.; data curation, A.R.; writing—original draft preparation, A.R.; writing—review and editing, U.C., M.B., G.M., C.G. and U.R.

Funding: This investigation has been supported by grants from the Swedish research council (project 2018-05003). The computations were performed on computer resources provided by the Swedish National Infrastructure for Computing (SNIC) at Lunarc at Lund University.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mörsdorf, G.; Frunzke, K.; Gadkari, D.; Meyer, O. Microbial growth on carbon monoxide. *Biodegradation* **1992**, *3*, 61–82. [CrossRef]
2. Xavier, J.C.; Preiner, M.; Martin, W.F. Microbial growth on carbon monoxide. *FEBS J.* **2018**, *285*, 4181–4195. [CrossRef] [PubMed]
3. Jeoung, J.-H.; Martins, B.M.; Dobbek, H. X-ray Crystallography of Carbon Monoxide Dehydrogenases. In *Metalloproteins*; Hu, Y., Ed.; Humana Press: New York, NY, USA, 2019; pp. 167–178.
4. Santiago, B.; Meyer, O. Characterization of hydrogenase activities associated with the molybdenum CO dehydrogenase from *Oligotropha carboxidovorans*. *FEMS Microbiol. Lett.* **1996**, *136*, 157–162. [CrossRef]
5. Nishino, T.; Okamoto, K.; Leimkühler, S. Enzymes of the Xanthine Oxidase Family. In *Molybdenum and Tungsten Enzymes: Biochemistry*; Hille, R., Schulzke, C., Kirk, M.L., Eds.; The Royal Society of Chemistry: London, UK, 2016; pp. 192–239.
6. Hille, R.; Hall, J.; Basu, P. The mononuclear molybdenum enzymes. *Chem. Rev.* **2014**, *114*, 3963–4038. [CrossRef] [PubMed]
7. Dobbek, H.; Gremer, L.; Kiefersauer, R.; Huber, R.; Meyer, O. Catalysis at a dinuclear [CuSMo(=O)OH] cluster in a CO dehydrogenase resolved at 1.1-Å resolution. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 15971–15976. [CrossRef] [PubMed]
8. Zhang, B.; Hemann, C.F.; Hille, R. Kinetic and spectroscopic studies of the molybdenum-copper CO dehydrogenase from Oligotropha carboxidovorans. *J. Biol. Chem.* **2010**, *285*, 12571–12578. [CrossRef] [PubMed]
9. Wilcoxen, J.; Hille, R. The Hydrogenase Activity of the Molybdenum/Copper-containing Carbon Monoxide Dehydrogenase of *Oligotropha carboxidovorans*. *J. Biol. Chem.* **2013**, *288*, 36052–36060. [CrossRef] [PubMed]
10. Gourlay, C; Nielsen, D.J.; White, J.M.; Knottenbelt, S.Z.; Kirk, M.L.; Young, C.G. Paramagnetic active site models for the molybdenum–copper carbon monoxide dehydrogenase. *J. Am. Chem. Soc.* **2006**, *128*, 2164–2165. [CrossRef] [PubMed]
11. Kaufmann, P.; Duffus, B.R.; Teutloff, C.; Leimkühler, S. Functional Studies on *Oligotropha carboxidovorans* Molybdenum-Copper CO Dehydrogenase Produced in *Escherichia coli*. *Biochemistry* **2018**, *57*, 2889–2901. [CrossRef] [PubMed]

12. Hille, R.; Dingwall, S.; Wilcoxen, J. The aerobic CO dehydrogenase from *Oligotropha carboxidovorans*. *J. Biol. Inorg. Chem. JBIC* **2015**, *20*, 243–251. [CrossRef] [PubMed]
13. Rokhsana, D.; Large, T.A.G.; Dienst, M.C.; Retegan, M.; Neese, F. A realistic in silico model for structure/function studies of molybdenum–copper CO dehydrogenase. *J. Biol. Inorg. Chem. JBIC* **2016**, *21*, 491–499. [CrossRef] [PubMed]
14. Breglia, R.; Bruschi, M.; Cosentino, U.; De Gioia, L.; Greco, C.; Miyake, T.; Moro, G. A theoretical study on the reactivity of the Mo/Cu-containing carbon monoxide dehydrogenase with dihydrogen. *Protein Eng. Des. Sel.* **2017**, *30*, 169–174. [CrossRef] [PubMed]
15. Plitt, H.S.; Bär, M.R.; Ahlrichs, R.; Schnöckel, H. [Cu(η^2-H_2)Cl], a model compound for H_2 complexes. Ab initio calculations and identification by IR spectroscopy. *Angew. Chem. Int. Ed.* **1991**, *30*, 832–834. [CrossRef]
16. Frohman, D.J; Grubbs, G.S., II; Yu, Z.; Novick, S.E. Probing the Chemical Nature of Dihydrogen Complexation to Transition Metals, a Gas Phase Case Study: H_2–CuF. *Inorg. Chem.* **2013**, *52*, 816–822. [CrossRef] [PubMed]
17. Serykh, A.I.; Kazansky, V.B. Unusually strong adsorption of molecular hydrogen on Cu^+ sites in copper-modified ZSM-5. *Phys. Chem. Chem. Phys.* **2004**, *6*, 5250–5255. [CrossRef]
18. Spoto, G.; Gribov, E.; Bordiga, S.; Lamberti, C.; Ricchiardi, G.; Scarano, D.; Zecchina, A. $Cu^+(H_2)$ and $Na^+(H_2)$ adducts in exchanged ZSM-5 zeolites. *Chem. Commun.* **2004**, 2768–2769. [CrossRef] [PubMed]
19. Georgiev, P.A.; Albinati, A.; Mojet, B.L.; Ollivier, J.; Eckert, J. Observation of Exceptionally Strong Binding of Molecular Hydrogen in a Porous Material: Formation of an η^2-H_2 Complex in a Cu-Exchanged ZSM-5 Zeolite. *J. Am. Chem. Soc.* **2007**, *129*, 8086–8087. [CrossRef] [PubMed]
20. Reguera, L.; Krap, C.P.; Balmaseda, J.; Reguera, E. Hydrogen storage in copper prussian blue analogues: Evidence of H_2 coordination to the copper atom. *J. Phys. Chem. C* **2008**, *112*, 15893–15899. [CrossRef]
21. Stephan, D.W. "Frustrated Lewis pairs": A concept for new reactivity and catalysis. *Org. Biomol. Chem.* **2008**, *6*, 1535–1539. [CrossRef] [PubMed]
22. Carr, S.B.; Evans, R.M.; Brooke, E.J.; Wehlin, S.A.M.; Nomerotskaia, E.; Sargent, F.; Armstrong, F.A.; Phillips, S.E.V. Hydrogen activation by [NiFe]-hydrogenases. *Biochem. Soc. Trans.* **2016**, *44*, 863–868. [CrossRef] [PubMed]
23. Miyake, T.; Bruschi, M.; Cosentino, U.; Baffert, C.; Fourmond, V.; Legér, C.; Moro, G.; De Gioia, L.; Greco, C. Does the environment around the H-cluster allow coordination of the pendant amine to the catalytic iron center in [FeFe] hydrogenases? Answers from theory. *J. Biol. Inorg. Chem. JBIC* **2013**, *18*, 693–700. [CrossRef] [PubMed]
24. Greco, C.; Silakov, A.; Bruschi, M.; Ryde, U.; De Gioia, L.; Lubitz, W. Magnetic Properties of [FeFe]-Hydrogenases: A Theoretical Investigation Based on Extended QM and QM/MM Models of the H-Cluster and Its Surroundings. *Eur. J. Inorg. Chem.* **2011**, 1043–1049. [CrossRef]
25. Greco, C. H_2 Binding and Splitting on a New-Generation [FeFe]-Hydrogenase Model Featuring a Redox-Active Decamethylferrocenyl Phosphine Ligand: A Theoretical Investigation. *Inorg. Chem.* **2013**, *52*, 1901–1908. [CrossRef] [PubMed]
26. Siegbahn, P.E.M.; Shestakov, A.F. Quantum chemical modeling of CO oxidation by the active site of molybdenum CO dehydrogenase. *J. Comput. Chem.* **2005**, *26*, 888–898. [CrossRef] [PubMed]
27. Gnida, M.; Ferner, R.; Gremer, L.; Meyer, O.; Meyer-Klaucke, W. A novel binuclear [CuSMo] cluster at the active site of carbon monoxide dehydrogenase: Characterization by X-ray absorption spectroscopy. *Biochemistry* **2003**, *42*, 222–230. [CrossRef] [PubMed]
28. Rovaletti, A.; Bruschi, M.; Moro, G.; Cosentino, U.; Ryde, U.; Greco, C. A thiocarbonate sink on the enzymatic energy landscape of aerobic CO oxidation? Answers from DFT and QM/MM models of MoCu CO-dehydrogenases. *J. Catal.* **2019**, *372*, 201–205. [CrossRef]
29. Olsson, M.H.M.; Søndergaard, C.R.; Rostkowski, M.; Jensen, J.H. PROPKA3: Consistent treatment of internal and surface residues in empirical pK_a Predictions. *J. Chem. Theory Comput.* **2011**, *7*, 525–537. [CrossRef] [PubMed]
30. Ryde, U. The coordination of the catalytic zinc ion in alcohol dehydrogenase studied by combined quantum-chemical and molecular mechanics calculations. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 153–164. [CrossRef] [PubMed]
31. Ryde, U.; Olsson, M.H.M. Structure, strain, and reorganization energy of blue copper models in the protein. *Int. J. Quantum Chem.* **2001**, *81*, 335–347. [CrossRef]

32. Reuter, N.; Dejaegere, A.; Maigret, B.; Karplus, M. Frontier bonds in QM/MM methods: A comparison of different approaches. *J. Phys. Chem. A* **2000**, *104*, 1720–1735. [CrossRef]
33. Cao, L.; Ryde, U. On the difference between additive and subtractive QM/MM calculations. *Front. Chem.* **2018**, *6*, 89. [CrossRef] [PubMed]
34. TURBOMOLE V6.3 2011, a Development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989–2007; TURBOMOLE GmbH, Since 2007. Available online: http://www.turbomole.com (accessed on 8 November 2019)
35. Becke, A. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100. [CrossRef] [PubMed]
36. Perdew, J.P. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B* **1986**, *33*, 8822–8824. [CrossRef] [PubMed]
37. Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305. [CrossRef] [PubMed]
38. Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465. [CrossRef] [PubMed]
39. Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. Auxiliary basis sets for main row atoms and transition metals and their use to approximate Coulomb potentials. *Theor. Chem. Acc.* **1997**, *97*, 119–124. [CrossRef]
40. Case, D.A.; Babin, V.; Berryman, J.; Betz, R.M.; Cai, Q.; Cerutti, D.S.; Cheatham, T.E., III; Darden, T.A.; Duke, R.E.; Gohlke, H.; et al. *AMBER 14*; University of California: San Francisco, CA, USA, 2014.
41. Maier, J.A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K.E.; Simmerling, C. ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713. [CrossRef] [PubMed]
42. Wang, J.; Wolf, R.M.; Caldwell, J.W.; Kollman, P.A.; Case, D.A. Development and testing of a general Amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174. [CrossRef] [PubMed]
43. Bayly, C.I.; Cieplak, P.; Cornell, W.D.; Kollman, P.A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993**, *97*, 10269–10280. [CrossRef]
44. Romero, E.A.; Olsen, P.M.; Jazzar, R.; Soleilhavoup, M.; Gembicky, M.; Bertrand, G. Spectroscopic Evidence for a Monomeric Copper(I) Hydride and Crystallographic Characterization of a Monomeric Silver(I) Hydride. *Angew. Chem. Int. Ed.* **2017**, *56*, 4024–4027. [CrossRef] [PubMed]
45. Huang, G.; Wagner, T.; Wodrich, M. D.; Ataka, K.; Bill, E.; Ermler, U.; Hu, X.; Shima, S. The atomic-resolution crystal structure of activated [Fe]-hydrogenase. *Nat. Catal.* **2019**, *2*, 537–543 [CrossRef]
46. Tai, H.; Nishikawa, K.; Higuchi, Y.; Mao, Z.-W.; Hirota, S. Cysteine SH and Glutamate COOH Contributions to [NiFe] Hydrogenase Proton Transfer Revealed by Highly Sensitive FTIR Spectroscopy. *Angew. Chem. Int. Ed.* **2019**, *58*, 13285–13290. [CrossRef] [PubMed]
47. Senger, M.; Mebs, S.; Duan, J.; Shulenina, O.; Laun, K.; Kertess, L.; Wittkamp, F.; Apfel, U.P.; Happe, T.; Winkler, M. Protonation/reduction dynamics at the [4Fe–4S] cluster of the hydrogen-forming cofactor in [FeFe]-hydrogenases. *Phys. Chem. Chem. Phys.* **2018**, *20*, 3128–3140. [CrossRef] [PubMed]

Paper IV

A. Rovaletti, C. Greco and U. Ryde
QM/MM study of the binding of H_2 to MoCu CO dehydrogenase: development and application of improved H_2 parameters for a correct treatment of gas-protein interaction
Submitted to the Journal of Molecular Modelling

Noname manuscript No.
(will be inserted by the editor)

QM/MM study of the binding of H_2 to MoCu CO dehydrogenase: development and applications of improved H_2 van der Waals parameters

Anna Rovaletti · Claudio Greco · Ulf Ryde

Received: date / Accepted: date

Abstract MoCu CO dehydrogenase can oxidise molecular hydrogen to protons and electrons. Even if its hydrogenase activity has been known for decades, a debate is ongoing on the most plausible mode for the binding of H_2 to the enzyme active site and the hydrogen oxidation mechanism. In the present work, we provide a new perspective on the MoCu-CODH hydrogenase activity by improving the *in silico* description of the enzyme. Energy refinement – by means of the BigQM approach – was performed on the intermediates involved in the dihydrogen oxidation catalysis reported in our previously published work (Rovaletti, et al. "Theoretical Insights into the Aerobic Hydrogenase Activity of Molybdenum–Copper CO Dehydrogenase." Inorganics 7 (2019) 135). A sub-optimal description of the H_2–HN(backbone) interaction was observed when the van der Waals parameters described in previous literature for H_2 were employed. Therefore, a new set of van der Waals parameters is developed here in order to better describe the hydrogen–backbone interaction. Implications of the resulting outcomes for a better understanding of hydrogen oxidation catalysis mechanisms are proposed and discussed.

Keywords Hydrogenases · MoCu CO dehydrogenase · H_2 oxidation · BigQM approach · QM/MM · force field parametrization

A. Rovaletti
Department of Earth and Environmental Sciences, Milano-Bicocca University,
Piazza della Scienza 1, 20126 Milan, Italy
E-mail: a.rovaletti@campus.unimib.it

C. Greco
Department of Earth and Environmental Sciences, Milano-Bicocca University,
Piazza della Scienza 1, 20126 Milan, Italy
E-mail: claudio.greco@unimib.it

U. Ryde
Department of Theoretical Chemistry, Lund University, Chemical Centre, P.O. Box 124, SE-221 00 Lund, Sweden.
E-mail: ulf.ryde@teokem.lu.se

1 Introduction

Hydrogenases represent a large enzyme family that catalyse the conversion of H_2 to protons and electrons, as well as the reverse reaction [26]. They are all metalloenzymes and are subdivided into three main classes depending on the type of metal ions that compose their active site: [NiFe], [FeFe] and [Fe] hydrogenases. Still, also other metalloenzymes, such as the MoCu CO dehydrogenase (MoCu-CODH, featuring a MoCu binuclear active site) [11] have the ability to oxidise H_2 to protons and electrons [37, 47].

Many theoretical studies of hydrogenase reactivity have been published, based either on quantum mechanical (QM) cluster or on hybrid quantum mechanics/molecular mechanics (QM/MM) approaches [38, 30, 4, 6, 16, 12, 33]. Moreover, H_2 and O_2 gas diffusion inside the various protein matrices has been investigated via molecular dynamics (MD) simulations [9, 10, 42, 43].

The accuracy of calculations at the molecular mechanics (MM) level critically depends on the accuracy of the employed force field (ff). Suboptimal ff parameters may give rise to incorrect conformations or misleading energies. This problem may also arise at the interface between the QM and MM subsystems in hybrid QM/MM calculations.

In a previous study, we proposed a mechanistic picture for the MoCu-CODH hydrogenase activity on the basis of QM/MM results [33]. As far as computed energies are concerned, it has been pointed out that both QM-cluster and QM/MM results may depend critically on the size of the QM system [41, 21, 25, 14, 22, 40, 32]. To tackle this issue, one of us has developed the BigQM approach within which stable energies are normally obtained using a large quantum model (∼1000 atoms). The latter is composed by all groups within 4.5–6 Å from the QM system of the QM/MM calculation, all buried charged groups in the protein, and by moving truncated groups three residues away from the active site [22]. A previous QM/MM study of the first steps of the CO oxidation reaction in MoCu-CODH [34] indicated that BigQM-based refinement of computed energies is a valuable approach also in the theoretical investigation of the latter enzyme.

In the present study, we have refined the QM/MM energies associated with the hydrogenase activity of MoCu-CODH. As described in the following, we found unexpected large differences in relative energies between the QM/MM and the BigQM results. We show that these large deviations are caused by poor van der Waals (VdW) parameters used for H_2 which lead to much too short interactions with nearby backbone HN groups. Therefore, we have developed improved H_2 VdW parameters. A comparison with the results obtained with parameters previously described in literature shows that our developments represent a key step to obtain reliable energies and structures in the modelling of H_2–protein interactions.

2 Methods

2.1 The MoCu-CODH protein

The setup of the protein was the same as in our previous QM/MM studies of MoCu-CODH [33]. All calculations were based on the crystal structure of CODH in its oxidised form (PDB ID: 1N5W) [11]. Only the active site-containing large subunit (L) of the MoCu-CODH dimer of LMS heterotrimers was considered. The setup of the protein involved a detailed analysis of all the protonable residues, based on calculations with PROPKA [28], studies of the hydrogen-bond pattern, the solvent accessibility and the possible formation of ionic pairs. Based on this, all Arg, Lys, Asp, and Glu residues were considered in their charged form, with exception of Glu29 and Glu488 that were protonated on OE2 and Asp684 that was protonated on OD1. Cysteine ligands coordinating to metals were deprotonated. Histidine residues were assumed to be doubly protonated with exception of His177, 178, 210, 213, 243, 700, 753, and 788 that were protonated only on the NE2 atom and of His61, 339, 766 and 793 that were protonated on the ND1 atom. The protein was solvated with water molecules, forming a sphere with a 60 Å radius around the geometric centre. A 1-ns simulated-annealing molecular-dynamics simulation, followed by a minimisation, was run to optimise added protons and water molecules.

2.1.1 QM/MM calculations

The QM/MM calculations were performed using the ComQum software [35, 36]. The protein and the solvent were split into two subsystems: System 1 is the QM system (see Supporting Information) while System 2 contains the remaining part of the protein and the water molecules. During the QM/MM geometry minimisation, System 1 is described by a wavefunction and is relaxed by QM methods, whereas Systems 2 is represented by an array of partial point charges (electrostatic embedding) and is kept fixed at the crystallographic coordinates. Covalent bonds between the QM and MM systems were truncated using the hydrogen link-atom approach [31]. The QM system is capped with hydrogen atoms (hydrogen link atoms, HL), the position of which are linearly related to the corresponding carbon atoms (carbon link atoms, CL) in the full system. The latter are not included in the point-charge model. The total QM/MM energy in ComQum is calculated as [35, 36]

$$E_{\mathrm{QM/MM}} = E^{\mathrm{HL}}_{\mathrm{QM1+ptch2}} + E^{\mathrm{CL}}_{\mathrm{MM12},q_1=0} - E^{\mathrm{HL}}_{\mathrm{MM1},q_1=0} \quad (1)$$

where $\mathrm{E}^{\mathrm{HL}}_{\mathrm{QM1+ptch2}}$ is the QM energy of System 1 truncated by HL atoms and embedded in the set of point charges modelling Systems 2. $\mathrm{E}^{\mathrm{HL}}_{\mathrm{MM1},q_1=0}$ is the MM energy of System 1, still truncated by HL atoms, without any electrostatic interactions. $\mathrm{E}^{\mathrm{CL}}_{\mathrm{MM12},q_1=0}$ is the classical energy of the whole system, with CL atoms and with the charges in System 1 set to zero, in order to avoid double counting of the electrostatic interactions. Thus, ComQum employs a

subtractive scheme with electrostatic embedding and van der Waals link-atom corrections [7].

The QM calculations were carried out at the BP86-D3(BJ)/def2-TZVP [2, 29, 17, 46] level of theory using TURBOMOLE 7.2 software [44]. The resolution-of-identity technique was used to accelerate the calculations [13]. The MM calculations were carried out by means of the Amber software [8], using the Amber FF14SB force field for the protein [27], and the general Amber force field [45] with restrained electrostatic potential (RESP) charges [1] for the molybdopterin-cytosine dinucleotide (MCD) cofactor. Parameters used for H_2 are discussed below. The two Fe_2S_2 clusters – fixed during the optimisation process – were described with RESP charges and a non-bonded model [20]. Single-point calculations at the B3LYP-D3(BJ)/def2-TZVPD [3, 24, 46] level were also run on the optimised geometries.

2.1.2 BigQM calculations

The BigQM technique was applied with the aim of obtaining more accurate QM/MM energies [22, 40], based on the knowledge that QM/MM calculations converge faster than QM-only ones, but models of about 1000 QM atoms are needed to obtain convergence of the energies [41, 14, 25, 32]. The minimal QM system (System 1) was extended with all chemical groups with at least one atom within 6.0 Å of it and junctions were moved three amino-acids away from each residue in the minimal QM system. In addition, all buried charges inside the protein were included, with exception of the two iron–sulfur clusters (see Supporting Information). The resulting BigQM model consisted of 983 (most calculations), 991 (when Cys388 is protonated) or 1019 atoms (when a water molecule is present in the active site). The BigQM calculations were performed on coordinates from the QM/MM optimisations, with a surrounding point-charge model, at the BP86-D3(BJ)/def2-SVP level. The multipole accelerated resolution-of-identity J approach (marij keyword) was employed to accelerate the calculations [39]. The resulting energies were corrected with a QM/MM MM term for the BigQM region:

$$E_{\mathrm{MM}} = E^{\mathrm{CL}}_{\mathrm{MM12},q_1=0} - E^{\mathrm{HL}}_{\mathrm{MM1},q_1=0} \tag{2}$$

Finally, the energies were also corrected by taking into consideration the B3LYP-D3(BJ)/def2-TZVPD functional and basis set effects, using calculations with the standard QM/MM QM system with a point-charge model of the surroundings:

$$E_{\mathrm{corr}} = E^{\mathrm{B3LYP/TZVPD}}_{\mathrm{QM1,ptch2}} - E^{\mathrm{BP86/SVP}}_{\mathrm{QM1,ptch2}} \tag{3}$$

2.2 The H_2–$CH_3CONHCH_3$ Model

A small [H_2 + $CH_3CONHCH_3$] model (see Figure 1) was used to determine the potential-energy surface for the H_2–HN interaction with QM methods and to subsequently determine van der Waals parameters to reproduce it.

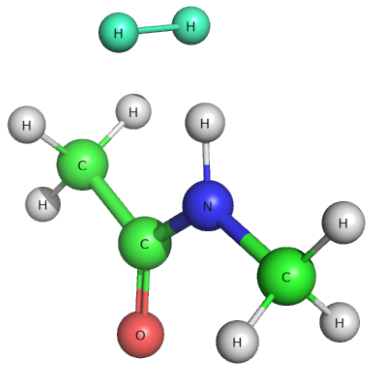


Fig. 1 The small [H_2 + $CH_3CONHCH_3$] model used for parametrization of the VdW parameters of H_2. Colour code: blue, N; red, O; green, C; gray, H; cyan, H_2 molecule.

2.2.1 DFT calculations

The calculations were performed at the B3LYP-D3(BJ)/def2-TZVPD level using TURBOMOLE 7.2 software. The resolution-of-identity technique was employed to accelerate the calculations. The potential-energy surface for the H_2–HN distance was scanned from a distance of 3.50 Å to 1.02 Å with steps of 0.25 Å, through optimisation of all degree of freedom with exception of the analysed H–H distance that was kept fixed at the predetermined distance.

2.2.2 Parameter fitting

The Amber software was used to carry out all MM calculations in conjunction with Amber FF14SB force field for the $CH_3CONHCH_3$ protein backbone model. Restrained electrostatic potential (RESP) charges were obtained separately for H_2 and $CH_3CONHCH_3$ running a single point energy calculation at HF/6-31G* [18, 15, 19] on the QM-optimised structure, followed by antechamber charge generation. Naturally, the charges on the two H atoms in H_2 were zero. The H–H optimum bond length was set to 0.7443 Å with a force constant of 416.038 kcal/mol/Å^2. Improved non-bonded parameters were obtained for the H_2 molecule, as described in Section 3.2.

3 Results and Discussion

3.1 Evidence of inaccuracies in the QM/MM description of the H_2–HN(backbone) interaction

In a previous study [33], we employed the hybrid QM/MM technique to investigate plausible mechanisms for the oxidation of H_2 by MoCu-CODH. Binding of H_2 to the active site was explored considering the enzyme resting state (**1H**), the presence of a water molecule in proximity to the copper ion (**2H**) and the hypothetical protonation of the Cu-bound cysteine residue (**3H**). Moreover, a variant of (**1H**) was considered in which Cu and the Mo–O_{eq} atom act as a frustrated Lewis pair (**1H-FLP**). For each of these variants we studied possible H_2 binding modes to the Cu centre, comparing the energy of the coordinated structures (hereafter indicated with an **R** tag) to the corresponding structure with the hydrogen molecule present in the active site but not linked to any atom (**NoB**). Deprotonation of the metal-activated H_2 by means of the equatorial Mo-bound oxygen (referred to with **P** tag) or by an active-site glutamate residue (Glu763) (**P1**) was investigated. A second protonation of the equatorial Mo–O ligand was considered (**P2**) to form the fully reduced enzyme, $Mo^{IV}O(OH_2)Cu^{I}$. Finally, the transition state for the proton-transfer reaction to the Mo-oxo ligand was located (**TS**) as well as the one involved in the transfer of the Cu-bound hydride to the Mo–OH_{eq} group (**TS2**).

With the aim of refining the QM/MM energies obtained for the intermediates, we performed BigQM calculations on the resulting optimised QM/MM geometries. Relative energies with respect to the **R** adducts are reported for the identified intermediates in Table 1.

Table 1 Relative energies (kcal/mol) calculated using the QM/MM and BigQM approaches at the B3LYP-D3(BJ)/def2-TZVPD level.

		NoB	**R**	**P1**	**TS**	**P**	**TS2**	**P2**
QM/MM	1H	-11.2	0.0	13.8		-0.7	7.6	-31.1
BigQM		-38.3	0.0	0.2		-42.0	-32.3	-64.2
QM/MM	2H	-11.1	0.0	19.9				
BigQM		-42.0	0.0	21.3				
QM/MM	3H	-2.6	0.0		7.0	6.7	13.5	-21.7
BigQM		15.0	0.0		4.5	3.6	12.4	-10.9
QM/MM	1H-FLP	0.2	0.0		10.8	10.6	18.9	-19.8
BigQM		10.0	0.0		7.0	6.3	16.0	-15.9

Comparing the energy profiles obtained by the QM/MM and BigQM methods, large deviations are observed for the binding energies of H_2 to Cu ($\Delta\Delta E^{NoB/R}_{BigQM/QMMM}$) for all of the tested active site states (**1H**, **2H** and **3H**). The differences in relative energies were found to be 27, 31, 18 and 10 kcal/mol for **1H**, **2H**, **3H** and **1H-FLP**, respectively. Moreover, the **1H** species gave very large differences for all studied states, up 41 kcal/mol for the **1HR**/**1HP**

pair. $\Delta\Delta E^{3HP2/3HR}_{BigQM/QMMM}$ is also rather large (11 kcal/mol), whereas all other intermediates exhibit a difference in relative energy that is less than 5 kcal/mol, i.e. similar to what is observed for other systems [34].

To understand the origin of these large differences between the QM/MM and BigQM relative energies we focused on the **1HR** → **1HP** step and calculated single-point BigQM energies at the BP86-D3(BJ)/def2-SV(P) level (i.e. without the additional corrections in equation 2) for QM systems of increasing size. Thus the original BigQM system – composed of 983 atoms (see Section 2) – was gradually reduced until the QM system of the QM/MM calculations was reached (46 atoms). The shrinking of the QM model was obtained by modulating three parameters for setting up the BigQM system (see Section 2), i.e. the cutoff radius, the number of residues junctions are moved away, and whether charged buried residues are included. To reach the smallest models, the MCD cofactor was also truncated before or after the phosphate groups. The resulting energy differences between the **1HP** and **1HR** intermediates are reported in Table 2.

Table 2 Single point BigQM calculations with QM systems of different sizes. The models are obtained with varying cutoff radii (r in Å), varying number of residues the junctions are moved away (n_{aa}), and whether buried charged residues are included (ch aa). The resulting QM systems are described by the total number of atoms (n_{atoms}), the total charge (charge) and the relative energy of **1HP** with respect to **1HR** expressed in kcal/mol obtained at the BP86-D3(BJ)/def2-SV(P) level. The last line shows the same relative energy obtained in the standard QM/MM calculation. Models C–M include the entire MCD cofactor, whereas in B, MCD is modelled by the molybdopterin ligand plus the –PO_2–O–PO_2–OH group and in A, it is modelled by the molybdopterin ligand only, as in the original QM system of QM/MM calculations.

BigQM	r	n_{aa}	ch aa	n_{atoms}	charge	ΔE
M	6	3	yes	983	-5	-49.2
L	6	3	no	590	-2	-49.2
K	4	3	yes	764	-5	-51.3
J	4	3	no	351	-3	-49.5
I	2	3	yes	621	-5	-44.4
H	2	3	no	154	-3	-45.1
G	1	3	yes	600	-5	-44.2
F	1	3	no	126	-4	-44.2
E	1	1	no	95	-4	-42.8
D	0	1	no	95	-4	-42.8
C	0	0	no	81	-4	-7.9
B	0	0	no	54	-4	-7.9
A	0	0	no	42	-2	-7.3
QM/MM	-	-	-	46	-3	-7.9

It can be seen that BigQM energies change by less than 2 kcal/mol if the buried charged residues are included or excluded. Somewhat larger variations are observed for the various cutoff radii (43–51 kcal/mol), but the difference between the two largest radii is less than 2 kcal/mol. Truncating the MCD cofactor also has a minimal effect (less than 1 kcal/mol). Instead, the large

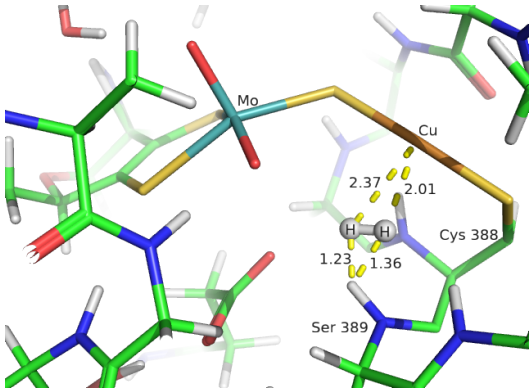


Fig. 2 QM/MM optimised geometry of **1HR** focusing on the active site. Distances are in Å. Colour code: cyan, Mo; copper, Cu; yellow, S; blue, N; red, O; green, C; gray, H.

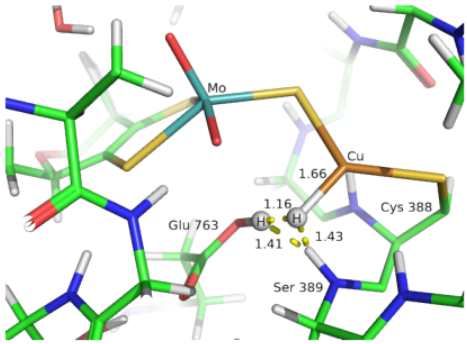
difference comes from the number of residues the junctions are moved away. In the present case, this involves only the neighbours of the Cys388 residue. Reducing the number from 3 (in standard BigQM) to 1 has small effect (1.4 kcal/mol between models E and F). However, excluding also the last residue (from model D to C), ΔE suddenly jumps by 35 kcal/mol. By a careful analysis of intermediates **1HR** and **1HP** in the two models, we realised that the deviating energies arise from the very short H_2–HN(Ser389) distances (1.23 and 1.36 Å) observed in the **1HR** geometry, as shown in Figure 2.

Based on this observation, we further analysed all the QM/MM optimised geometries and we found short H_2–HN(backbone) interactions also for the **1HP1** (1.41, 1.43 Å), **2HR** (1.35, 1.25 Å) and **2HP1** (1.10, 1.28 Å) as shown in Figure 3.

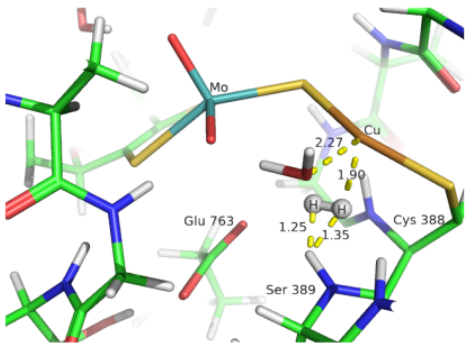
Such short H_2–backbone distances indicates that the van der Waals repulsion between the peptide NH group and the hydrogen molecule is underestimated at the MM level (in these QM/MM calculations, H_2 is in the QM system but Ser389 is in the MM system, so that their interactions are determined by the MM parameters, whereas in the BigQM calculation, both groups are in the QM system). This was examined in the following sections, by using a small [H_2 + $CH_3NHCOCH_3$] model.

3.2 Improving the H_2–backbone non-bonded MM parameters

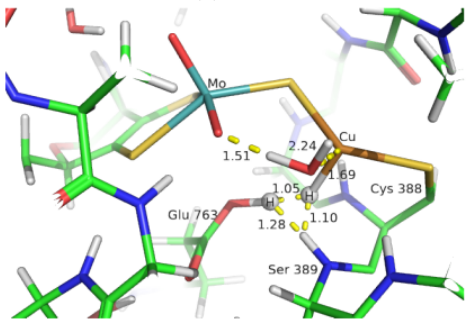
To confirm that the problem lies in the VdW parameters of H_2, we optimised the structure of the [H_2 + $CH_3NHCOCH_3$] model as described in Section 2. The optimised geometry, reported in Figure 4, shows H_2–HN distances of 2.52 and 2.73 Å, i.e. much longer than the corresponding ones in the QM/MM



(a) **1HP1**



(b) **2HR**



(c) **2HP1**

Fig. 3 QM/MM optimised geometries for **1HP1**, **2HR** and **2HP1**. All distances are in Å. The colour code is the same as Figure 2.

structures of MoCu-CODH (see above). On the other hand, a full optimisation of the same [H_2 + $CH_3NHCOCH_3$] model with the original MM parameters gave H_2–HN distances of 1.49 and 1.50 Å. This shows that the original VdW parameters are suboptimal.

Starting from the QM structure, we performed a series of constrained optimisations in order to scan the potential-energy surface (PES) for the H_2–HN interaction. The resulting QM PES is shown in Figure 5a (black thick line).

The non-bonded interaction between two atoms at a distance of r is described in the Amber MM force-field as a sum of an attractive r^{-6} (dispersion) and a repulsive r^{-12} (exchange repulsion) term, according to the Lennard-Jones equation:

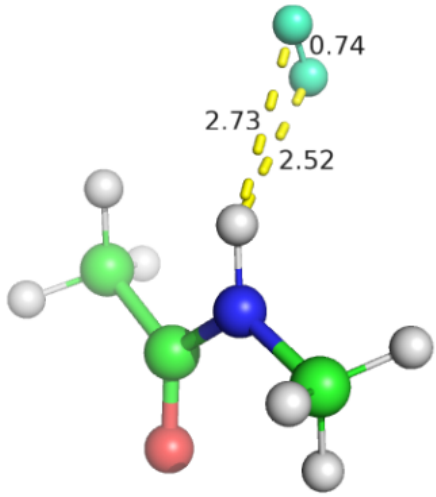
$$E_{\mathrm{VdW}} = \epsilon \left[\left(\frac{r^*}{r} \right)^{12} - 2 \left(\frac{r^*}{r} \right)^{6} \right] \tag{4}$$

where r^* represents the distance at the energy minimum and ϵ defines the well depth.

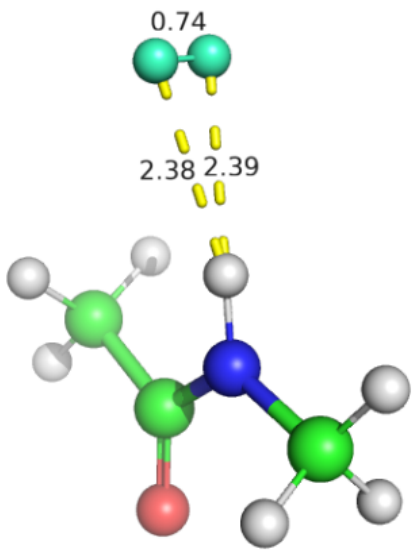
The original VdW parameters for H in H_2 were taken from the Amber FF14SB atom type H (H bound to N; the same parameters are used for H bound to S), with $r^* = 0.60$ Å and $\epsilon = 0.0157$ kcal/mol. They clearly give too short equilibrium distances and a much too little repulsive PES at short distances as is also shown in Figure 5a (red line marked H; the MM PESs were calculated by single-point calculations to avoid any distortion of the structure or the influence from other suboptimal MM parameters).

To obtain a better description of the PES, we tested different sets of VdW parameters (r^*= 1.50, 1.60, 1.70 and 1.85 Å, as well as ϵ= 0.01, 0.0157, 0.05, 0.055, 0.06, 1, 10 kcal/mol). The results are also included in Figure 5a (lines marked with the values of r^*/ϵ) and they are described in Table 3. It can be seen that the optimum H(H_2)–HN distance depends mainly on r^*, whereas ϵ is less important. It is also clear that it is not possible to reproduce both the long- and short-distance part of the PES (the MM potential is more repulsive than the QM one). We decided that it was most important to describe the minimum and longer distances, where the potential energy is close to the minimum energy and therefore often encountered in simulations. Figure 5b concentrates on that region. Employing the mean absolute deviation (MAD) in the interval d = 2.02—3.50 Å as the quality measure, we obtained optimum VdW parameters for $r^* = 1.7$ Å and $\epsilon = 0.055$ kcal/mol. With these, the MM curve reproduces the QM relative energies with a MAD of 0.06 kcal/mol (0.01 kcal/mol for distances 2.27–3.50 Å). At shorter distances, the curve is much too steep but this is acceptable because this part of the PES will rarely be visited in real simulations. A comparison of the optimum structure obtained with QM and the new MM parameters is shown in Figure 4. It does not fully reproduce the distances observed in the QM structure because H_2 moves to a more symmetric position in the MM structure and the methyl groups show different rotations.

The Amber FF14SB contains 13 sets of VdW parameters for different H atom types. Two of them (for water and OH groups) have zeroed parameters

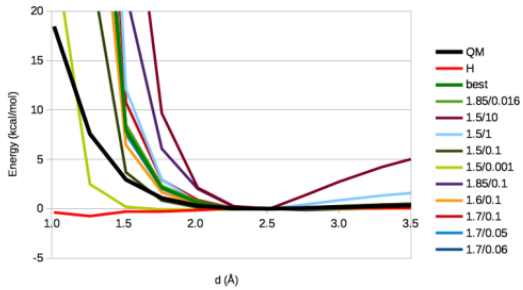


(a) **QM**

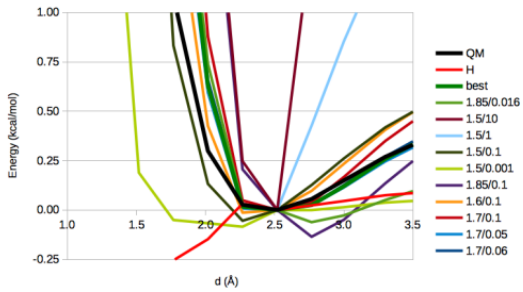


(b) **MM**

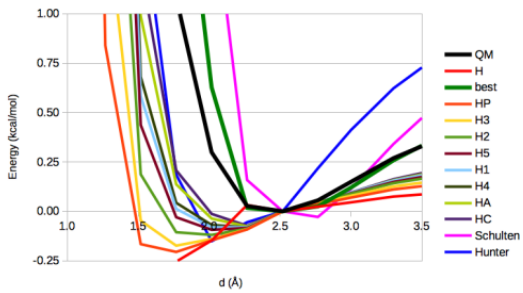
Fig. 4 The [H_2 + $CH_3NHCOCH_3$] model optimised at the B3LYP-D3(BJ)/def2-TZVPD level (QM) and at MM level with the new VdW parameters (MM). All distances in Å. The colour code is the same as in Figure 1.



(a)



(b)



(c)

Fig. 5 PES for the [H_2 + $CH_3NHCOCH_3$] model obtained with QM and different MM force fields (those starting with H are the various AMBER atom types). The pairs of numbers in the legends indicate the r^* and ϵ parameters in our calibration. "Schulten" and "Hunter" indicate the force fields suggested in refs. [10] and [23], respectively. "best" is our suggested force field.

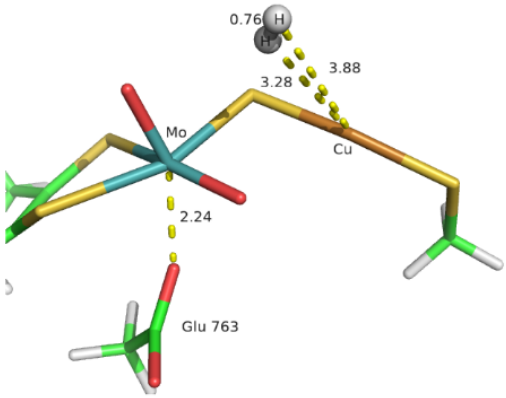
(the interaction is determined instead by VdW interactions by the heavy atom to which the H atom is bound), whereas the other have similar $\epsilon = 0.0157$ or 0.0150 kcal/mol, but a varying $r^* = 0.6 - 1.487$ Å (low values for H bound to polar atoms and high values for H bound to C). Our suggested parameters are both significantly larger than the AMBER VdW parameters. Figure 5c shows that all AMBER parameters give too short H_2–HN distances (but the H atom type is by far the worst).

In MD simulations of the diffusion of H_2 in hydrogenases [9, 10, 42, 43], two different sets of VdW parameters have been used. One was suggested by Schulten and coworkers [10] and employs $r^* = 1.7682$ Å and $\epsilon = 0.1521$ kcal/mol. From Figure 5c and Table 3 it can be seen that it gives a slightly too long H_2–HN optimum distance and a too repulsive potential at longer distances. The other is a more sophisticated model by Hunter *et al.* [23]. In variance to all the other models it employs an extra point at the H–H bond centre with a charge of -0.95 e and charges of 0.475 e on the H atoms (in order to give a proper quadrupole moment). Only the bond centre has non-zero VdW parameters, $r^* = 1.501$ Å and $\epsilon = 0.4769$ kcal/mol. In Figure 5c and Table 3 it can be seen that it gives a too short H_2–HN optimum distance and a too repulsive potential at longer distances.

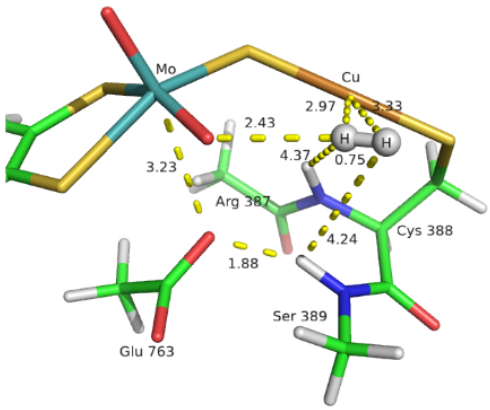
Thus, we can conclude that the H_2 interaction with the protein backbone NH group is sensitive to the VdW parameters and that only our suggested parameters give an accurate description of this interaction.

3.3 Test case: H_2 interaction with the MoCu-CODH active site

In order to validate the new set of VdW parameters, we run new QM/MM optimisations of the **1HR** intermediate. We tested two approaches. In the first (**1HR-A**), we used the QM system described in section 2, but employed our new VdW parameters for H_2 ($r^* = 1.7$ Å and $\epsilon = 0.055$ kcal/mol). In the second approach (**1HR-B**), the backbone neighbouring Cys388 was included in the QM system (see Figure S1), so that the H_2–HN interaction was treated at a QM level. Both optimisations led to a complete detachment of H_2 from the copper ion (see Figure 6), with a Cu–H distance of 3.28, 3.88 Å and 2.97, 3.33 Å for **1HR-A** and **1HR-B** , respectively. The main difference between the two structures is that Glu763 coordinates to Mo in **1HR-A** with a Mo–O distance of 2.26 Å, whereas the Mo–O distance is 3.23 Å in the **1HR-B** structure, because it instead forms a hydrogen bond to the backbone HN group of Ser389, which is also introduced in the QM region. This leads to quite extensive changes in the local structure around the Mo ion. The new QM/MM optimisations show that the VdW parameters of H_2 have a significant influence on the structure and that the **1HR** minimum previously obtained by means of the old VdW parameters for the H_2 (Figure 2) was an artefact caused by a bad description of the hydrogen–backbone interaction. Moreover, it clears that it is crucial to include the backbone around Cys388 in the QM system, with the two NH groups pointing into in the active site, providing potential



(a) **1HR-A**



(b) **1HR-B**

Fig. 6 QM/MM optimised **1HR** at the BP86-D3(BJ)/def2-TZVP level obtained by employing the new VdW parameters (**1HR-A**) or by including the backbone neighbouring Cys388 in the QM region of the QM/MM calculation (**1HR-B**). All distances in Å. The colour code is the same as in Figure 2.

hydrogen bonds to both Glu763 or various reaction intermediates. Apparently, the MM parameters are not accurate enough to model the subtle competition between Mo and the HN group of Ser389 for the carboxylate group of Glu763.

Table 3 Performance of the various tested sets of H_2 VdW parameters for the [H_2 + $CH_3NHCOCH_3$] model. The table lists the r^* and ϵ parameters, as well as the mean absolute deviation (MAD in kcal/mol) of the resulting MM energy compared to the results obtained at the BP86-D3(BJ)/def2-TZVPD level in the range $d = 2.02 - 3.50$ Å, the error in the minimum distance (Δd in Å) and the error in the minimum energy (ΔE in kcal/mol).

Set	r^*	ϵ	MAD	Δd	ΔE
1.85/0.016	1.85	0.0157	0.17	0.25	0.06
1.5/10	1.5	10	2.08	0.00	0.00
1.5/1	1.5	1	0.53	0.00	0.00
1.5/0.1	1.5	0.1	0.11	-0.25	0.05
1.5/0.001	1.5	0.001	0.17	-2.39	0.08
1.85/0.1	1.85	0.1	0.36	-0.25	0.14
1.6/0.1	1.6	0.1	0.08	0.00	0.00
1.7/0.1	1.7	0.1	0.12	0.00	0.00
1.7/0.05	1.7	0.05	0.06	0.00	0.00
best	1.7	0.055	0.06	0.00	0.00
1.7/0.06	1.7	0.06	0.06	0.00	0.00
H	0.6	0.0157	0.15	-1.25	0.75
HP	1.1	0.0157	0.15	-0.75	0.21
H3	1.187	0.0157	0.14	-0.75	0.17
H2	1.287	0.0157	0.13	-0.50	0.12
H5	1.359	0.015	0.12	-0.50	0.09
H1	1.387	0.0157	0.12	-0.25	0.08
H4	1.409	0.015	0.12	-0.25	0.08
HA	1.459	0.015	0.11	-0.25	0.08
HC	1.487	0.0157	0.10	-0.25	0.07
Schulten	1.7682	0.1521	0.26	0.25	0.03
Hunter	1.5010	0.4769	0.24	-0.50	0.15

4 Conclusions

In this paper, we have deepened our insights about the energetics of the putative initial steps for dihydrogen oxidation by MoCu CO dehydrogenase, keeping as a reference the QM/MM results from our previous investigation of the hydrogenase activity [33]. The energies were refined by using the BigQM approach on the QM/MM optimised geometries, analogously to our previous work on the CO-oxidation catalysis by this enzyme [34]. Unexpectedly, we observed large deviations in the relative energies obtained by the two approaches. Therefore, a detailed investigation was carried out to identify the cause of the deviation. We found that the reason was a suboptimal description of the H_2–backbone interaction, caused by poor van der Waals parameters for H_2. Using the simple H_2 + $CH_3NHCOCH_3$ model, we show that none of the previously suggested VdW parameters for H_2 are able to accurately reproduce the H_2–HN(backbone) interaction, which has key relevance for a correct description of the binding of H_2 in the active site pocket. Therefore, we have developed a new set of VdW parameters, with the aim of providing an accurate description of this interaction.
Based on the energy refinement work here presented, it is possible to exclude

the previously proposed occurrence [47] of side-on H_2 binding at the copper centre when the enzyme is in its resting state (**1HR**). Other routes for the initial interaction between substrate and enzyme active site need to be considered to account for the hydrogenase activity of Mo/Cu CO dehydrogenases, in line with previous theoretical investigations [5, 33].

From a methodological point of view, the BigQM approach confirmed to be a highly valuable approach for the exploration of the complex potential energy surface associated with metalloenzymes activity. Moreover, the reparameterization that turned out to be necessary in this work will hopefully be relevant in the context of future efforts towards the theoretical description of the interaction between dihydrogen and metalloenzymes for which the former is a reagent, a product, or both.

Acknowledgements This work was supported by grants from the Swedish research council (project 2018-05003). The computations were performed on computer sources provided by the Swedish National Infrastructure for Computing (SNIC) at Lunarc at Lund University.

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. Bayly CI, Cieplak P, Cornell WD, Kollman PA (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. J Phys Chem 97:10269–10280
2. Becke AD (1988) Density-functional exchange-energy approximation with correct asymptotic behavior. Physical review A 38(6):3098
3. Becke AD (1993) A new mixing of Hartree–Fock and local density-functional theories. The Journal of chemical physics 98(2):1372–1377
4. Bertini L, Greco C, Bruschi M, Fantucci P, De Gioia L (2010) CO affinity and bonding properties of [FeFe] hydrogenase active site models. A DFT study. Organometallics 29(9):2013–2025
5. Breglia R, Bruschi M, Cosentino U, De Gioia L, Greco C, Miyake T, Moro G (2017) A theoretical study on the reactivity of the Mo/Cu-containing carbon monoxide dehydrogenase with dihydrogen. Protein Engineering, Design and Selection 30(3):169–174
6. Breglia R, Greco C, Fantucci P, De Gioia L, Bruschi M (2018) Reactivation of the ready and unready oxidized states of [NiFe]-hydrogenases: mechanistic insights from DFT calculations. Inorganic chemistry 58(1):279–293
7. Cao L, Ryde U (2018) On the difference between additive and subtractive QM/MM calculations. Frontiers in chemistry 6:89
8. Case D, Babin V, Berryman J, Betz R, Cai Q, Cerutti D, Cheatham Iii T, Darden T, Duke R, Gohlke H, et al. (2014) Amber 14

9. Cohen J, Kim K, King P, Seibert M, Schulten K (2005) Finding gas diffusion pathways in proteins: application to O_2 and H_2 transport in CpI [FeFe]-hydrogenase and the role of packing defects. Structure 13(9):1321–1329
10. Cohen J, Kim K, Posewitz M, Ghirardi ML, Schulten K, Seibert M, King P (2005) Molecular dynamics and experimental investigation of H_2 and O_2 diffusion in [Fe]-hydrogenase. Biochem Soc Trans 33(1):80–82
11. Dobbek H, Gremer L, Kiefersauer R, Huber R, Meyer O (2002) Catalysis at a dinuclear [CuSMo(=O)OH] cluster in a CO dehydrogenase resolved at 1.1-Å resolution. Proceedings of the National Academy of Sciences 99(25):15971–15976
12. Dong G, Phung QM, Pierloot K, Ryde U (2018) Reaction mechanism of [NiFe] hydrogenase studied by computational methods. Inorganic chemistry 57(24):15289–15298
13. Eichkorn K, Weigend F, Treutler O, Ahlrichs R (1997) Auxiliary basis sets for main row atoms and transition metals and their use to approximate Coulomb potentials. Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta) 97(1):119–124
14. Flaig D, Beer M, Ochsenfeld C (2012) Convergence of electronic structure with the size of the QM region: example of QM/MM NMR shieldings. Journal of chemical theory and computation 8(7):2260–2271
15. Fock V (1930) Näherungsmethode zur lösung des quantenmechanischen mehrkörperproblems. Zeitschrift für Physik 62:795–805
16. Greco C, Bruschi M, Fantucci P, Ryde U, De Gioia L (2011) Mechanistic and physiological implications of the interplay among iron–sulfur clusters in [FeFe]-hydrogenases. A QM/MM perspective. Journal of the American Chemical Society 133(46):18742–18749
17. Grimme S, Ehrlich S, Goerigk L (2011) Effect of the damping function in dispersion corrected density functional theory. Journal of computational chemistry 32(7):1456–1465
18. Hartree D (1928) Proc. cambridge soc. vol 24:89–111
19. Hehre W, Radom L, Pople J, Schleyer Pv (1986) Ab initio molecular orbital theory
20. Hu L, Ryde U (2011) Comparison of methods to obtain force-field parameters for metal sites. J Chem Theory Comput y:2452–2463
21. Hu LH, Eliasson J, Heimdal J, Ryde U (2009) Do quantum mechanical energies calculated for small models of protein-active sites converge? J Phys Chem A 113(43):11793–11800
22. Hu LH, Söderhjelm P, Ryde U (2012) Accurate reaction energies in proteins obtained by combining QM/MM and large QM calculations. J Chem Theory Comput 9(1):640–649
23. Hunter JE, Taylor DG, Strauss HL (1992) Calculation of the rotational Raman spectrum of H_2 dissolved in water. J Chem Phys 97:50–59
24. Lee C, Yang W, Parr RG (1988) Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Physical review B 37(2):785

25. Liao RZ, Thiel W (2013) Convergence in the QM-only and QM/MM modeling of enzymatic reactions: A case study for acetylene hydratase. Journal of computational chemistry 34(27):2389–2397
26. Lubitz, Wolfgang and Ogata, Hideaki and Rüdiger, Olaf and Reijerse, Edward (2014) Hydrogenases. Chemical reviews 114(8):4081–4148
27. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C (2015) ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. Journal of chemical theory and computation 11(8):3696–3713
28. Olsson MH, Søndergaard CR, Rostkowski M, Jensen JH (2011) PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions. Journal of chemical theory and computation 7(2):525–537
29. Perdew JP (1986) Density-functional approximation for the correlation energy of the inhomogeneous electron gas. Physical Review B 33(12):8822
30. Qui S, Xu Y, Shen S, Sun C (2019) Learning from nature: Understanding hydrogenase enzyme using computational approach. WIRE's Comput Mol Sci p e1422
31. Reuter N, Dejaegere A, Maigret B, Karplus M (2000) Frontier bonds in QM/MM methods: A comparison of different approaches. The Journal of Physical Chemistry A 104(8):1720–1735
32. Roßbach S, Ochsenfeld C (2017) Influence of coupling and embedding schemes on QM size convergence in QM/MM approaches for the example of a proton transfer in DNA. Journal of chemical theory and computation 13(3):1102–1107
33. Rovaletti A, Bruschi M, Moro G, Cosentino U, Greco C, Ryde U (2019) Theoretical Insights into the Aerobic Hydrogenase Activity of Molybdenum–Copper CO Dehydrogenase. Inorganics 7(11):135
34. Rovaletti A, Bruschi M, Moro G, Cosentino U, Ryde U, Greco C (2019) A thiocarbonate sink on the enzymatic energy landscape of aerobic CO oxidation? Answers from DFT and QM/MM models of MoCu CO-dehydrogenases. Journal of Catalysis 372:201–205
35. Ryde U (1996) The coordination of the catalytic zinc ion in alcohol dehydrogenase studied by combined quantum chemical and molecular mechanical calculations. J Comput Aid Mol Des 10:153–164
36. Ryde U, Olsson MH (2001) Structure, strain, and reorganization energy of blue copper models in the protein. International Journal of Quantum Chemistry 81(5):335–347
37. Santiago B, Meyer O (1996) Characterization of hydrogenase activities associated with the molybdenum CO dehydrogenase from Oligotropha carboxidovorans. FEMS microbiology letters 136(2):157–162
38. Siegbahn PEM, Tye JW, Hall MB (2007) Computational Studies of [NiFe] and [FeFe] Hydrogenases. Chem Rev 107:4414–4435
39. Sierka M, Hogekamp A, Ahlrichs R (2003) Fast evaluation of the Coulomb potential for electron densities using multipole accelerated resolution of identity approximation. The Journal of chemical physics 118(20):9136–

9148
40. Sumner S, Söderhjelm P, Ryde U (2013) Effect of geometry optimizations on QM-cluster and QM/MM studies of reaction energies in proteins. J Chem Theory Comput 9(9):4205–4214
41. Sumowski CV, Ochsenfeld C (2009) A convergence study of QM/MM isomerization energies with the selected size of the QM region for peptidic systems. The Journal of Physical Chemistry A 113(43):11734–11741
42. Teixeira VH, Baptista AM, Soares CM (2006) Pathways of H_2 toward the active site of [NiFe]-hydrogenase. Biophysical journal 91(6):2035–2045
43. Topin J, Rousset M, Antonczak S, Golebiowski J (2012) Kinetics and thermodynamics of gas diffusion in a NiFe hydrogenase. Proteins: Structure, Function, and Bioinformatics 80(3):677–682
44. TURBOMOLE, "V7-2" (2017) a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989–2007, TURBOMOLE GmbH, since 2007. `http://wwwturbomolecom`
45. Wang J, Wolf RM, Caldwell JW, Kollman PA, A CD (2004) Development and testing of a general Amber force field. J Comput Chem 25:1157–1174
46. Weigend F, Ahlrichs R (2005) Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. Physical Chemistry Chemical Physics 7(18):3297–3305
47. Wilcoxen J, Hille R (2013) The Hydrogenase Activity of the Molybdenum/Copper-containing Carbon Monoxide Dehydrogenase of Oligotropha carboxidovorans. Journal of Biological Chemistry 288(50):36052–36060

Paper v

A. Rovaletti, M. Bruschi, U. Cosentino, G. Moro, U. Ryde and C. Greco
The direct role of water in CO oxidation catalysed by Mo-Cu CO-dehydrogenase. Answers from theory.
In preparation

The direct role of water in CO oxidation catalysed by Mo-Cu CO-dehydrogenase. Answers from theory.

Anna Rovaletti, Maurizio Bruschi, Ugo Cosentino,
Giorgio Moro, Ulf Ryde, Claudio Greco

October 22, 2020

1 Introduction

Carbon monoxide dehydrogenase (MoCu-CODH) from the aerobic soil bacterium *Oligotropha carboxidovorans* has an important role in the conversion of an estimated 2 x 10^8 tons of CO from the lower atmosphere and soil per year [27], following the reaction:

$$CO + H_2O \longrightarrow CO_2 + 2\,H^+ + 2\,e^-$$

The active site architecture of *O. carboxidovorans* CO-dehydrogenase is very peculiar. It is constituted of a unique $MoO_2(\mu S)Cu$ bimetallic centre in the oxidised form of the enzyme. The molybdenum ion shows a distorted square pyramidal geometry, with one axial oxo ligand and four equatorial ligands: a dithiolene ligand from a molybdopterin cytosine dinucleotide (MCD) cofactor, one oxo (O^{2-}) ligand and a μ-sulfido ligand. The latter links the Mo-centre to the Cu-centre, which is also coordinated by the sulphur atom of Cys388 and by a weakly coordinated water molecule [13, 44, 30]. Although the protonation state of the Mo(VI)(=O)O unit in the oxidised active site has been subject of

debate, due to controversial experimental results obtained by X-ray diffraction (XRD) techniques and by extended X-ray absorption fine structure (EXAFS) spectroscopy [9, 13], a consensus has now been reached on the presence of a MoO_2 core. In fact, recent data in support of this hypothesis has come from Electron Paramagnetic Resonance (EPR) analysis of the Mo(V) analogue and from theoretical calculations [47, 30]. The molybdenum ion can also attain a +4 redox state in which the equatorial oxo ligand is found to be di-protonated [30, 13, 36]. As far as the copper ion is concerned, the +1 oxidation state is maintained during the enzymatic catalytic cycle [9, 13]. The CO molecule binds to the Cu(I) centre [36] and its oxidation takes place at the same site, thanks to the high degree of delocalisation within the Mo(μ-S)Cu unit [14]. Such delocalisation allows for facile electron transfers toward the redox-active Mo ion during catalysis [14]. Notably, a glutamate residue (Glu763), which belongs to the second coordination shell of the metals, is conserved within the xanthine oxidase family of enzymes, to which MoCu-CODH belongs [17] and occupies a position such that it could facilitate the deprotonation events of protic ligands at the equatorial positions of the Mo ion.

The first mechanistic hypothesis for the catalytic cycle of the MoCu-CODH was published in 2002 by Dobbek *et al.*, based on structural evidences obtained from X-ray diffraction experiments [9]. According to their proposal, coordination of CO to the resting state of the enzyme is followed by formation of a thiocarbonate intermediate, obtained through a nucleophilic attack of the equatorial Mo-oxo ligand on the CO carbon atom. The structure of the latter is analogous to the thiocarbamate derivative formed during inhibition of the enzyme by *n*-butylisocyanide. Finally, a water molecule could regenerate the oxygenic molybdenum ligand with the consequent release of CO_2. The completion of the reaction cycle occurs upon reoxidation of Mo(IV) to Mo(VI) via the transfer of two electrons.

Inspired by such proposal, two research groups (Hofmann *et al.* and Siegbahn *et al.* [20, 37]) performed quantum mechanical calculations on cluster models of the enzyme active site. Hofmann *et al.* noted that the thiocarbonate intermediate

corresponds to a very stable side-product, the occurrence of which seems not to be functional for catalysis. However, Siegbahn and co-workers found that the SCO-bond of the thiocarbonate can be broken upon insertion of a water molecule in the first coordination shell of the molybdenum ion but the barriers were estimated to be high. More recent theoretical studies based on much larger models showed that the thiocarbonate species does not represent a energy sink on the catalytic energy profile [46, 33] but its formation still remains problematic for the CO-oxidation catalysis. In fact, no transition state for CO_2 release could be directly linked to the thiocarbonate intermediate by means of the hybrid quantum mechanics/molecular mechanics approach [46].

A different mechanism was suggested by Stein and Kirk [39], which avoids formation of a stable C–S bonded intermediate based on the hypothesis that a bicarbonate product can be formed after a nucleophilic attack of a water molecule on μ_2-η^2-CO_2 intermediate. The rate limiting step for such mechanism is lower than the highest barriers occurring in the previously proposed mechanism. Unfortunately, experimental EPR studies of the partially reduced binuclear centre in complex with bicarbonate showed that bicarbonate coordination to Mo is unlikely to happen during catalysis [8].

Finally, in a recent theoretical study on the MoCu-CODH hydrogenase activity, as well as in an experimental and theoretical investigation concerning a synthetic model of the enzyme active site, a hypothesis was formulated, according to which the protein active site should be viewed as a Mo^{VI}=O/Cu^{I} Frustrated Lewis Pairs (FLP), where the Cu^{I} act as a Lewis acid and the equatorial oxo ligand of Mo act as a Lewis base toward the oxidation of both H_2 and CO [32, 12]

The previously proposed mechanisms share the idea that the equatorial oxo-ligand of Mo is responsible of a direct oxygen transfer to CO. However, a consensus on this key issue has not been reached, due to the lack of experimental evidences in support of such scenario. In this regard, Hille and coworkers have recently hypothesised that water might play the role of nucleophile, in place of the Mo-bound oxide [18]. Such a role for water would be favoured by the presence of basic groups that might facilitate its deprotonation upon CO-oxidation.

3

This inspired us to investigate a new catalytic cycle using density functional theory (DFT) by considering the possibility of a direct role of water in the C–O bond formation.

2 Methods

2.1 QM-cluster models

Calculations were carried out in the framework of DFT. The realistic *in silico* model created and validated by Rokhsana *et al.* was used to model the structure of the catalytic site of MoCu-CODH [30]. The quantum mechanical (QM) cluster, composed of 179 atoms (180 in the case of protonated Glu763 residue), was derived from the crystal structure of the enzyme in its oxidised form (PDB ID: 1N5W) [9]. Geometry optimizations were performed *in vacuo* using the TURBOMOLE v7.2 program at the BP86/def2-SVP/def2-TZVP level [1, 3, 28, 43]. During optimizations, all αC atoms were kept frozen at the crystallographic positions along with one attached proton each. The def2-TZVP basis set was used for the first coordination shell (31 or 32 atoms considering Glu763 as anionic or neutral, respectively), while the rest of the model was described by the smaller def2-SVP basis set (see Supplementary Information Figure S1 for details). The resolution of identity technique was employed to accelerate the self-consistent field (SCF) calculations [10]. All calculations included Grimme's dispersion correction with Becke–Johnson damping (D3(BJ)) [15].

Transition states (TS) were localised by eigenvector-following approaches as previously described [5], by taking advantage of the trust-region image minimisation (TRIM) method [16]. In the optimised TSs, the presence of constrained atoms resulted in the occurrence of several imaginary frequencies upon vibrational modes calculations. However, the eigenvector corresponding to the reaction coordinate was easily identified as the one with an eigenvalue much lower than the other negative eigenvalues.

Unless otherwise stated, all energy differences discussed in the following are

based on electronic energies obtained by means of single point calculations on the geometries optimised as reported above, carried out by using the large def2-TZVPD basis set to all atoms, and by employing the hybrid functional B3LYP [4, 24]. The effects of the protein environment was taken into account by performing single-point energy calculations at the BP86 level using the COSMO continuum solvent model [23], and assuming that a general solvent with dielectric constant equal to 4 exerts solvation effects similar to the ones of the protein matrix and the default optimised radii for all atoms (and 2 Å for the metals).

2.2 QM/MM calculations

Hybrid QM/MM model was performed in the same way as in our previous works [33, 32] [JMolMod2020]. The system was split into two subsystems: the QM region or System 1 (see Supplementary Information Figure S2), which was described by a wavefunction and is relaxed by QM methods, and Systems 2, which was represented by an array of partial point charges (electrostatic embedding) and was kept fixed at the crystallographic coordinates. Covalent bonds between the QM and MM systems were treated by the hydrogen link-atom approach [29]. The calculations were performed using the ComQum software [34, 35]. Total QM/MM energy in ComQum is calculated as [34, 35]

$$E_{\mathrm{QM/MM}} = E^{\mathrm{HL}}_{\mathrm{QM1+ptch2}} + E^{\mathrm{CL}}_{\mathrm{MM12},q_1=0} - E^{\mathrm{HL}}_{\mathrm{MM1},q_1=0} \quad (1)$$

where $\mathrm{E}^{\mathrm{HL}}_{\mathrm{QM1+ptch2}}$ is the QM energy of System 1 truncated by hydrogen-link (HL) atoms and embedded in the set of point charges modelling Systems 2. $\mathrm{E}^{\mathrm{HL}}_{\mathrm{MM1},q_1=0}$ is the MM energy of System 1, still truncated by HL atoms, without any electrostatic interactions. $\mathrm{E}^{\mathrm{CL}}_{\mathrm{MM12},q_1=0}$ is the classical energy of the whole system, with carbon-link (CL) atoms and with the charges in System 1 set to zero, in order to avoid double-counting of the electrostatic interactions. Thus, ComQum uses a subtractive scheme with electrostatic embedding and van der Waals link-atom corrections [6].

The QM calculations were carried out using TURBOMOLE 7.2 software [1] at the BP86-D3(BJ)/def2-TZVP [3, 28, 15, 43] level of theory. The resolution-

of-identity technique was used to accelerate the calculations [10]. The MM calculations were carried out by means of the Amber software [7], using the Amber FF14SB force field for the protein [26], and the general Amber force field [42] with restrained electrostatic potential (RESP) charges [2] for CO and the MCD cofactor. The two Fe_2S_2 clusters were described with RESP charges and a non-bonded model [21].

2.3 BigQM energy calculation

The BigQM approach was employed to obtain refined energies of key steps in the CO-oxidation catalysis. This technique [22, 40] has been shown to be a valuable tool in theoretical investigation of this enzyme [33][JMolMod2020]. It had been demonstrated that models of about 1000 QM atoms are needed to obtain convergence of the energies [41, 11, 25, 31]. Since, QM/MM calculations converge faster than QM cluster calculations, we started from optimised QM/MM geometries to build the large BigQM model of 1035 atoms. We included in the calculations all chemical groups with at least one atom within 6.0 Å of the QM system in the QM/MM calculations and junctions were moved three amino-acids away from each residue in the minimal QM system. In addition, all buried charges inside the protein were included, with exception of the two iron–sulfur clusters (see Supplementary Information List S3). The BigQM calculations were performed on coordinates from the QM/MM optimisations, with a surrounding point-charge model, at the BP86-D3(BJ)/def2-SVP level. The multipole accelerated resolution-of-identity J approach (marij keyword) was employed to accelerate the calculations [38]. The resulting energies were corrected with a QM/MM MM term for the BigQM region:

$$E_{\mathrm{MM}} = E^{\mathrm{CL}}_{\mathrm{MM12},q_1=0} - E^{\mathrm{HL}}_{\mathrm{MM1},q_1=0} \tag{2}$$

Finally, the energies were also corrected by taking into consideration the B3LYP-D3(BJ)/def2-TZVPD functional and basis set effects, using calculations with the

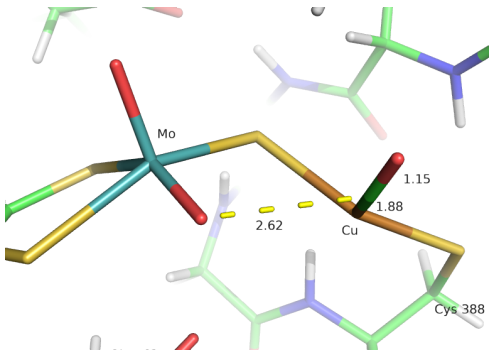


Figure 1: Optimised CO-bound intermediate. All distances in angstrom. Colour code: cyan, Mo; copper, Cu; yellow, S; blue, N; red, O; green, C; gray, H.

standard QM/MM QM system and a point-charge model of the surroundings:

$$E_{\text{corr}} = E^{\text{B3LYP/TZVPD}}_{\text{QM1,ptch2}} - E^{\text{BP86/SVP}}_{\text{QM1,ptch2}} \tag{3}$$

3 Results and Discussion

As explained in the Introduction, we focus here on the possible role that water may play in the catalysis of CO oxidation. In the first part of this section, we performed a study of plausible poses of H_2O within the active site cavity, functional to nucleophilic attack of water to the metal-bound CO ligand. Based on this premise, we were able to propose a new catalytic path for MoCu-CODH, which is discussed in the second part of this section.

3.1 Relevant H_2O poses in the CO-bound active-site cavity

The CO-bound form of the enzyme features a trigonal planar coordination geometry of the Cu ion (Fig. 1), as previously reported [37, 20]. Given the features of the second coordination spheres of the metals, a limited number of water poses

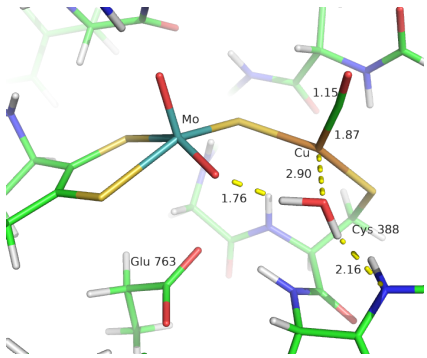
can be devised, which place the O atom of water in proximity of the C atom of the CO ligand. In fact, water may form hydrogen bonds either with the the Mo-bound oxo ligand in the proximal position with respect to the CO substrate, or with the sidechain of Glu763; the corresponding structures are tagged **B1** and **B2** in Figure 2. Moreover, we explored the possibility of coordination of water to the CO-bound Cu center: a tetracoordinated structure was actually obtained upon geometry optimisation (**B'** in Fig. 2), which turned out to be 4.3 and 9.1 kcal/mol lower in energy than **B1** and **B2**, respectively. In intermediate **B'**, the water molecule is H-bonded to the two nearest strong bases, Glu763 and the equatorial Mo-oxo ligand, at 1.67 Å and 1.68 Å, respectively. However, the involvement of **B'** in the nucleophilic attack of water to the C atom of CO is questionable. In fact, it was not possible to find any relevant transition state for this reaction, although eigenvector-following calculations were started from guess structures with an imaginary mode associated with Cu–O(H_2O) bond stretching.

As mentioned, **B1** features hydrogen bonding between water and one of the Mo-bound oxo groups with a O(oxo)–H(H_2O) distance of 1.76 Å, whereas the other H atom of H_2O is in proximity of an amide nitrogen of the protein backbone (N(HN-Phe390)-H($_2$O) distance: 2.16 Å). The distance between the O atom of water and the C atom of the CO ligand is relatively short (2.66 Å). Therefore, it is not surprising that we were able to find a TS structure for C–O bond formation connected to this structure (see below, sect. 3.2). The latter does not apply to **B2**, which features a much larger distance between the CO ligand and H_2O (C(CO)–O(H_2O) distance: 4.23 Å).

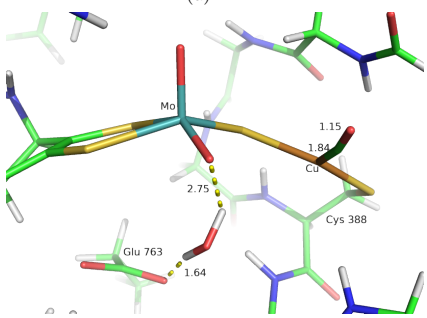
3.2 Proposed catalytic cycle for CO oxidation mediated by a metal-activate water molecule

The catalytic cycle starts with CO (substrate) binding to the resting state of the enzyme (Mo^{VI} and Cu^{I}), state **A** (Figure 3).

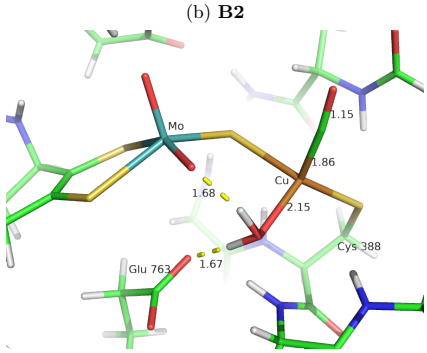
The latter features a tricoordinated Cu centre (Fig. 4), with one of the coordina-



(a) **B1**



(b) **B2**



(c) **B'**

Figure 2: Optimised-intermediates **B1**, **B2**, **B'**. All distances in Å.

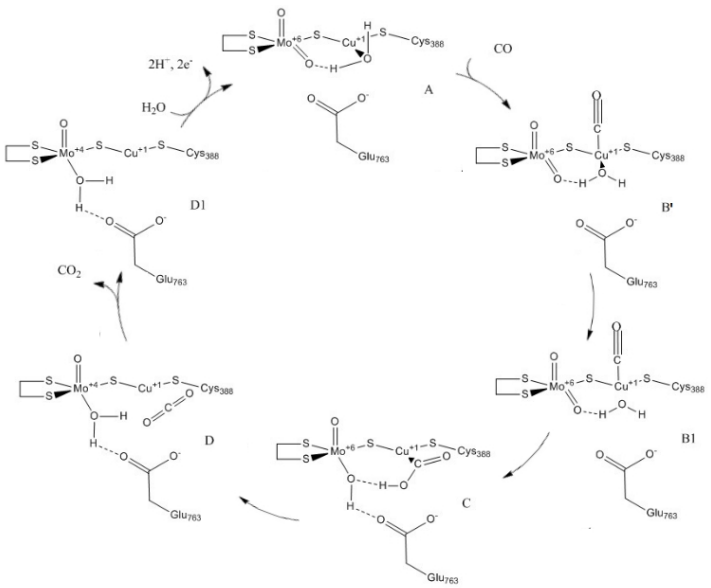


Figure 3: Proposed catalytic cycle for CO-oxidation in MoCu-CODH.

tion positions occupied by a water molecule. A comparison of the Cu–O(H_2O) distances between the crystal structure and our model evidences underestimation of the bond length (2.49 vs. 2.12 Å, respectively).

A similar outcome was discussed previously in the theoretical investigation performed by Rokhsana *et al.* (Cu-O(H_2O) distance: 2.14 Å) [30]. These authors argued that the weakly coordinated H_2O molecule might occupy slightly different positions within the crystal; however, they also showed that experimental EPR hyperfine couplings are very well reproduced by their model. The difference can also be partially explained by dynamic effects at ambient temperatures for this weak Cu^I–OH_2 bond.

CO binding to **A** gives rise to the tetracoordinated adduct **B'**, which was de-

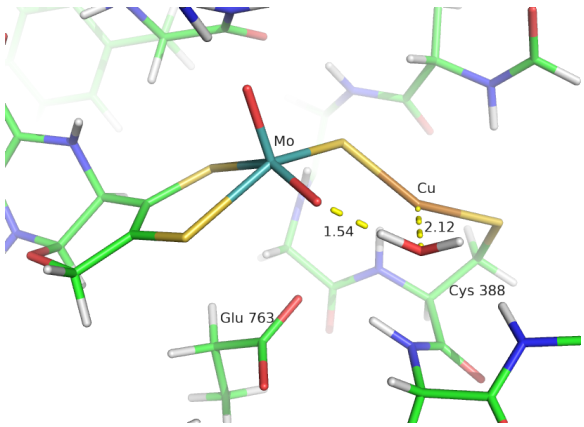


Figure 4: Optimised-intermediates **A**. All distances in Å.

scribed above (See Section 3.1).

As already mentioned earlier, even if a migratory-insertion reaction mechanism on the Cu atom was first hypothesised, it was not possible to localise a reactive channel for C–O bond formation starting from **B'**. Therefore, decoordination of water from Cu was considered as a next step in the catalytic cycle, which leads to formation of **B1**.

At this point, a new C–O bond can be formed as a result of a nucleophilic addition reaction of the water molecule to the carbon monoxide coordinated to copper. Formation of the new bond is assisted by the equatorial oxo ligand of molybdenum which augments the water nucleophilicity by formation of a strong H-bond interaction. **TS1** (Fig. 8), which links the reagent **B1** to the product **C**, is characterised by a seven-membered ring structure. The C(CO)–O(H_2O) distance is as long as 1.71 Å in **TS1**; concomitantly, one of the O–H bonds distances elongates up to 1.30 Å. This occurs thanks to the presence of the Mo=O_{eq} ligand, acting as a base during the reaction (O(oxo)-H(H_2O) distance in **TS1**: 1.13 Å). A vibrational frequency calculations evidences the presence

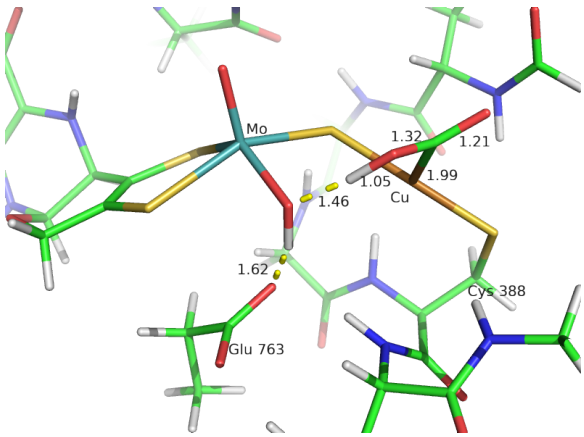


Figure 5: Optimised-intermediates **C**. All distances in Å.

of an imaginary mode (–563 cm^{-1}) that highlights how the nucleophilic attack of the water oxygen to the CO-carbon is coupled with the deprotonation of water by the Mo oxo ligand. The computed activation barrier associated to this reaction is quite high, ΔE_a = 23.5 kcal/mol.

The product of the nucleophilic attack (**C**) involves a $COOH^-$ ligand coordinated to the Cu(I) centre at a distance of 1.99 Å (Fig. 5). The conversion of intermediate **B1** to product **C** only involves small reorganisations of the molybdenum coordination sphere. The equatorial oxo ligand of molybdenum was protonated to an hydroxo ligand in **C**, is involved in two hydrogen bonds. It donates one to the Glu763 residue (1.62 Å) and accepts the other from the $COOH^-$ molecule (1.46 Å). The intermediate **C** is computed to be 2.4 kcal/mol less stable than **B1**.

The next step in the catalytic cycle is the release of carbon dioxide. This requires oxidation of the formate and a second proton transfer to the equatorial Mo-ligand. It takes place through a seven-membered ring transition state structure (**TS2**, Fig. 8b) to allow the passage of the $COOH^-$ proton to the Mo–OH

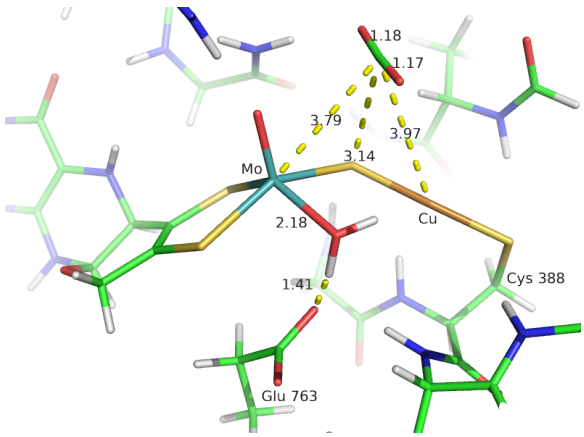


Figure 6: Optimised-intermediates **D**. All distances in Å.

ligand. While going from intermediate **C** to **TS2**, elongations of the O–H and Cu–C bonds of 0.08 and 0.01 Å are observed while the O(O_{eq}H-Mo)–H(COOH) distance decreases by 0.15 Å and the OCO angle is increased by 3.5°. The energy barrier of the deprotonation reaction is very low, ΔE_a =0.4 kcal/mol.

Finally, the product of the whole CO-oxidation reaction is represented by state **D**. It features a reduced $Mo^{IV}Cu^{I}$ centre in which the molybdenum ion is coordinated by a water molecule in the equatorial position at a 2.18 Å distance from the metal (Fig. 6). The water molecule forms a hydrogen bond with the oxygen of Glu763 (1.41 Å), while CO_2 has completely dissociated from the Mo(μS)Cu core (Mo–C = 3.79 Å, S–C = 3.14 Å, Cu–C = 3.97 Å). The reaction product **D** is very stable with respect to intermediate **C**, ΔE = –20.9 kcal/mol, and the activation barrier of the reverse reaction is 21.3 kcal/mol.

On the basis of the previous observations, we suggest a new alternative catalytic cycle for the oxidation of carbon monoxide by MoCu-CODH, based on the properties of its unique bimetallic active site and the presence of a water molecule. The catalytic cycle is summarised in Figure 3 and the associated energy profile

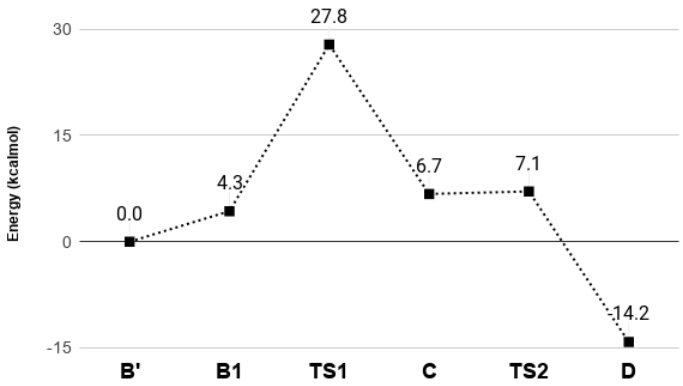
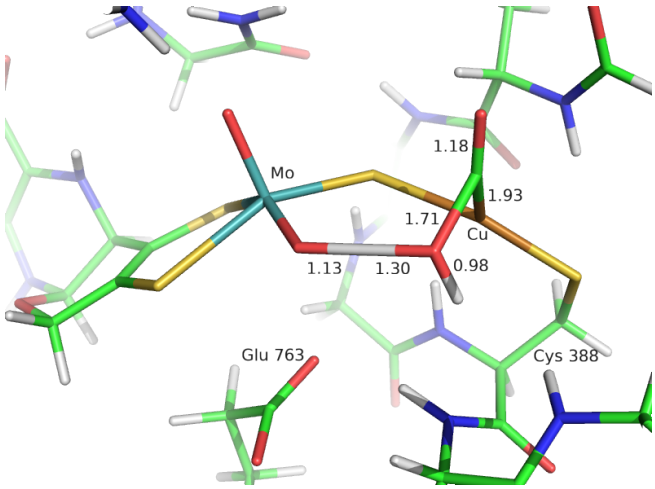


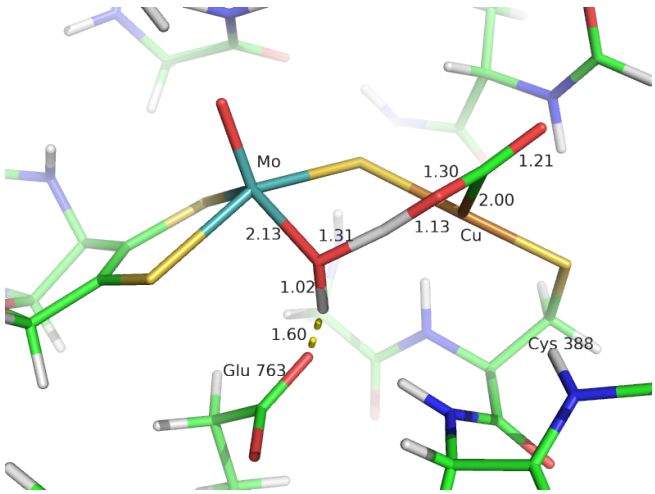
Figure 7: Potential energy surface of the new CO-oxidation mechanism, computed at the B3LYP-D3(BJ)/def2-TZVPD level. Relative energies in kcal/mol.

is shown in Figure 7.

Starting from intermediate **A**, the entrance of the CO substrate would allow the formation of the tricoordinated intermediate **B1**. Activation of the water molecule by means of the equatorial Mo–oxo ligand enables the formation of a new C–O bond in intermediate **C**. Deprotonation of the HCOO-ligand leads to the CO_2 product which dissociates from the copper atom, to generate intermediate **D**. Finally, CO_2 may leave the active site and the $Mo^{IV}(=O)_2Cu^{I}$ species **A** can be regenerated by the release of two protons and two electrons to a quinone pool of the electron transport chain [45]. The concomitant entrance of a new water molecule in the active site would close the enzymatic catalytic cycle.



(a) **TS1**



(b) **TS2**

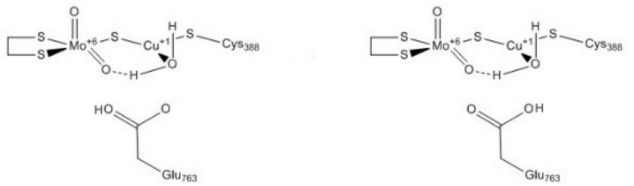
Figure 8: Transition state structures **TS1** and **TS2**. All distances in angstrom.

3.3 Further investigation of the rate limiting B1 to C reaction

In the present study, we tested the hypothesis [18] that water might act as a nucleophile upon C-O bond formation, during the CO oxidation reaction catalyzed by the MoCu-CODH. The rate limiting step computed for the catalytic cycle is 27.8 kcal/mol. Therefore, we investigated the role of the protonation state of Glu763, as well as by the protein matrix on the energetics of the activation barrier of the **B1** $\rightarrow$ **C** reaction.

First, we considered the possibility that Glu763 may be protonated at this step of catalysis. The protonation state of this residue during catalysis has not been settled yet. However, it may exerts the role of a proton acceptor of a protonated equatorial ligand of Mo. Glutamic acid was described in its neutral form by adding a proton to the QM model previously used in order to represent two possible protonated forms – **GluOH1** and **GluOH2** – as shown in Figure 9. Intermediates **B1** and **C** were found for both protonation states (see Supplementary Information Figure S4) but we successfully identified a transition state for the **B1** $\rightarrow$ **C** reaction only in the case of **GluOH1**. The transition state showed an imaginary frequency of –673.3 cm^{-1} analogous to the one identified in the case of the anionic Glu763. The activation barrier associated with this reaction was calculated to be 30.9 kcal/mol, significantly higher than if Glu763 is deprotonated.

Next, we studied the effects that the protein environment might have on this section of the catalytic cycle. This was done by application of the QM/MM approaches in order to obtain the optimised geometries and by BigQM calculation for a better estimate of the resulting energies (see sections 2.2 and 2.3). **B1** optimised with QM/MM shows a C(CO)–O(H_2O) distance of 2.95 Å and (N(HN-Phe390)-H($_2$O) distance of 3.04 Å (see Supplementary Information Figure S5). Compared to the **B1** minima previously obtained using the QM cluster model, we observe a slight elongation of those distances by 0.29 Å and 0.88 Å, as well as a partial modification of the orientation of the water molecule with



(a) **GluOH1** (b) **GluOH2**

Figure 9: MoCu-CODH active site representing the two possible neutral forms of the Glu763 residue.

respect to CO. The latter was found to be less perpendicular in the QM/MM geometry (computed H(H_2O)–O(H_2O)–C(CO) angle of 50.4° and 92.3° in the QM/MM model and QM-cluster model, respectively). Despite of these structural variations, a **TS1** was located on the PES, by scanning the C(CO)–O(H_2O) distance. The resulting optimised transition state geometry is characterised by a C(CO)–O(H_2O) distance of 1.75 Å (see Supplementary Information Figure S6). By using these optimised geometries, the BigQM approach was applied to refine the activation energy barrier at the B3LYP-D3(BJ)/def2-TZVPD level on a model of 1035 atoms. This gave an activation energy of 18.2 kcal/mol, 5 kcal/mol lower than for the QM-cluster model.

4 Conclusions

In the present study, we used DFT calculations to test the hypothesis that water might act as a nucleophile upon C–O bond formation, during the CO oxidation reaction catalysed by MoCu-CODH. Notably, the most plausible mechanistic alternatives, are the catalytic cycles proposed by Siegbahn and Shestakov [37] and by Xu and Hirao [46]. In both mechanisms, formation of a thiocarbonate intermediate in the course of catalysis is contemplated and it was shown to act as a potential energy well, impeding an easy release of the CO_2 molecule. In order

to complete the catalytic cycle, Siegbahn and Shestakov suggested the insertion of a water molecule in the first coordination sphere of Mo but found it to be a arduous step. Difficulties in the involvement of the S–C intermediate in the CO oxidation catalysis were also faced by Xu and Hirao. They suggested that after thiocarbonate formation, the catalytic cycle should follow a reverse process and reform the precursor of the thiocarbonate species to productively favour CO_2 evolution. However, in a recent work we found that the thiocarbonate intermediate was evidenced not to represent an energy sink along the catalytic energy profile when an accurate description of the enzyme is included – e.g. by modelling the active site surroundings and the dispersion effects as well as by employing sufficiently large basis sets for the description of the system even during optimisation calculations [33].

Here we show that an alternative mechanism may be followed in the oxidation catalysis of CO that naturally avoids formation of the thiocarbonate intermediate, since the proximal oxo ligand of Mo works as a base and not as a nucleophile. Such alternative mechanism, already proposed in the literature [18], might constitute one of the reactive channel for the oxidation of CO by the enzyme. The role of the oxo ligand shown here is atypical for the members of the xanthine oxidase family, in which oxo-transfer reactions is generally mediated by the catalytically labile equatorial ligand of Mo. Therefore, evidences in support of the reported hypothesis might derive from future experimental studies by ^{18}O-isotopic labelling in order to determine if water constitutes the primary source of the oxygen atom incorporated into product in the course of catalysis. Such technique turned out to be a successful tool in previous mechanistic studies of xanthine oxidase enzymes [19].

Future QM/MM studies concerning all the states involved in the QM-cluster based catalytic cycle will be needed to confirm the possibility that this mechanism may constitute a reactive channel in the CO oxidation by MoCu CODH.

References

[1] TURBOMOLE V7.1 2016, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from `http://www.turbomole.com`.

[2] C. I. Bayly, P. Cieplak, W. D. Cornell, and P. A. Kollman. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.*, 97:10269–10280, 1993.

[3] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical review A*, 38(6):3098, 1988.

[4] A. D. Becke. A new mixing of Hartree–Fock and local density-functional theories. *The Journal of chemical physics*, 98(2):1372–1377, 1993.

[5] M. Bruschi, M. Tiberti, A. Guerra, and L. De Gioia. Disclosure of key stereoelectronic factors for efficient h2 binding and cleavage in the active site of [nife]-hydrogenases. *Journal of the American Chemical Society*, 136(5):1803–1814, 2014.

[6] L. Cao and U. Ryde. On the difference between additive and subtractive QM/MM calculations. *Frontiers in chemistry*, 6:89, 2018.

[7] D. Case, V. Babin, J. Berryman, R. Betz, Q. Cai, D. Cerutti, T. Cheatham Iii, T. Darden, R. Duke, H. Gohlke, et al. Amber 14. 2014.

[8] S. Dingwall, J. Wilcoxen, D. Niks, and R. Hille. Studies of carbon monoxide dehydrogenase from Oligotropha carboxidovorans. *Journal of Molecular Catalysis B: Enzymatic*, 134:317–322, 2016.

[9] H. Dobbek, L. Gremer, R. Kiefersauer, R. Huber, and O. Meyer. Catalysis at a dinuclear [CuSMo(=O)OH] cluster in a CO dehydrogenase resolved at 1.1-Å resolution. *Proceedings of the National Academy of Sciences*, 99(25):15971–15976, 2002.

[10] K. Eichkorn, F. Weigend, O. Treutler, and R. Ahlrichs. Auxiliary basis sets for main row atoms and transition metals and their use to approximate Coulomb potentials. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)*, 97(1):119–124, 1997.

[11] D. Flaig, M. Beer, and C. Ochsenfeld. Convergence of electronic structure with the size of the QM region: example of QM/MM NMR shieldings. *Journal of chemical theory and computation*, 8(7):2260–2271, 2012.

[12] D. Ghosh, S. Sinhababu, B. D. Santarsiero, and N. P. Mankad. A W/Cu synthetic model for the Mo/Cu cofactor of aerobic CODH indicates that biochemical CO oxidation requires a frustrated Lewis acid/base pair. *Journal of the American Chemical Society*, 142(29):12635–12642, 2020.

[13] M. Gnida, R. Ferner, L. Gremer, O. Meyer, and W. Meyer-Klaucke. A novel binuclear [CuSMo] cluster at the active site of carbon monoxide dehydrogenase: characterization by X-ray absorption spectroscopy. *Biochemistry*, 42(1):222–230, 2003.

[14] C. Gourlay, D. J. Nielsen, J. M. White, S. Z. Knottenbelt, M. L. Kirk, and C. G. Young. Paramagnetic active site models for the molybdenum-copper carbon monoxide dehydrogenase. *Journal of the American Chemical Society*, 128(7):2164–2165, 2006.

[15] S. Grimme, S. Ehrlich, and L. Goerigk. Effect of the damping function in dispersion corrected density functional theory. *Journal of computational chemistry*, 32(7):1456–1465, 2011.

[16] T. Helgaker. Transition-state optimizations by trust-region image minimization. *Chemical Physics Letters*, 182(5):503–510, 1991.

[17] R. Hille. The molybdenum oxotransferases and related enzymes. *Dalton transactions*, 42(9):3029–3042, 2013.

[18] R. Hille, S. Dingwall, and J. Wilcoxen. The aerobic CO dehydrogenase from Oligotropha carboxidovorans. *JBIC Journal of Biological Inorganic Chemistry*, 20(2):243–251, 2015.

[19] R. Hille and H. Sprecher. On the mechanism of action of xanthine oxidase. evidence in support of an oxo transfer mechanism in the molybdenum-containing hydroxylases. *Journal of Biological Chemistry*, 262(23):10914–10917, 1987.

[20] M. Hofmann, J. K. Kassube, and T. Graf. The mechanism of Mo/Cu dependent CO dehydrogenase. *JBIC Journal of Biological Inorganic Chemistry*, 10(5):490–495, 2005.

[21] L. Hu and U. Ryde. Comparison of methods to obtain force-field parameters for metal sites. *J. Chem. Theory Comput.*, y:2452–2463, 2011.

[22] L. H. Hu, P. Söderhjelm, and U. Ryde. Accurate reaction energies in proteins obtained by combining QM/MM and large QM calculations. *J Chem Theory Comput*, 9(1):640–649, 2012.

[23] A. Klamt and G. Schüürmann. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *Journal of the Chemical Society, Perkin Transactions 2*, (5):799–805, 1993.

[24] C. Lee, W. Yang, and R. G. Parr. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical review B*, 37(2):785, 1988.

[25] R.-Z. Liao and W. Thiel. Convergence in the QM-only and QM/MM modeling of enzymatic reactions: A case study for acetylene hydratase. *Journal of computational chemistry*, 34(27):2389–2397, 2013.

[26] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, and C. Simmerling. ff14SB: improving the accuracy of protein side chain

and backbone parameters from ff99SB. *Journal of chemical theory and computation*, 11(8):3696–3713, 2015.

[27] G. Mörsdorf, K. Frunzke, D. Gadkari, and O. Meyer. Microbial growth on carbon monoxide. *Biodegradation*, 3(1):61–82, 1992.

[28] J. P. Perdew. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Physical Review B*, 33(12):8822, 1986.

[29] N. Reuter, A. Dejaegere, B. Maigret, and M. Karplus. Frontier bonds in QM/MM methods: A comparison of different approaches. *The Journal of Physical Chemistry A*, 104(8):1720–1735, 2000.

[30] D. Rokhsana, T. A. Large, M. C. Dienst, M. Retegan, and F. Neese. A realistic in silico model for structure/function studies of molybdenum–copper CO dehydrogenase. *JBIC Journal of Biological Inorganic Chemistry*, 21(4):491–499, 2016.

[31] S. Roßbach and C. Ochsenfeld. Influence of coupling and embedding schemes on QM size convergence in QM/MM approaches for the example of a proton transfer in DNA. *Journal of chemical theory and computation*, 13(3):1102–1107, 2017.

[32] A. Rovaletti, M. Bruschi, G. Moro, U. Cosentino, C. Greco, and U. Ryde. Theoretical Insights into the Aerobic Hydrogenase Activity of Molybdenum–Copper CO Dehydrogenase. *Inorganics*, 7(11):135, 2019.

[33] A. Rovaletti, M. Bruschi, G. Moro, U. Cosentino, U. Ryde, and C. Greco. A thiocarbonate sink on the enzymatic energy landscape of aerobic CO oxidation? Answers from DFT and QM/MM models of MoCu CO-dehydrogenases. *Journal of Catalysis*, 372:201–205, 2019.

[34] U. Ryde. The coordination of the catalytic zinc ion in alcohol dehydrogenase studied by combined quantum chemical and molecular mechanical calculations. *J Comput Aid Mol Des*, 10:153–164, 1996.

[35] U. Ryde and M. H. Olsson. Structure, strain, and reorganization energy of blue copper models in the protein. *International Journal of Quantum Chemistry*, 81(5):335–347, 2001.

[36] M. Shanmugam, J. Wilcoxen, D. Habel-Rodriguez, G. E. Cutsail III, M. L. Kirk, B. M. Hoffman, and R. Hille. ^{13}C and ^{63,65}Cu ENDOR studies of CO Dehydrogenase from Oligotropha carboxidovorans. Experimental Evidence in Support of a Copper–Carbonyl Intermediate. *Journal of the American Chemical Society*, 135(47):17775–17782, 2013.

[37] P. E. M. Siegbahn and A. F. Shestakov. Quantum chemical modeling of CO oxidation by the active site of molybdenum CO dehydrogenase. *Journal of Computational Chemistry*, 26(9):888–898, 2005.

[38] M. Sierka, A. Hogekamp, and R. Ahlrichs. Fast evaluation of the Coulomb potential for electron densities using multipole accelerated resolution of identity approximation. *The Journal of chemical physics*, 118(20):9136–9148, 2003.

[39] B. W. Stein and M. L. Kirk. Orbital contributions to CO oxidation in Mo–Cu carbon monoxide dehydrogenase. *Chemical communications*, 50(9):1104–1106, 2014.

[40] S. Sumner, P. Söderhjelm, and U. Ryde. Effect of geometry optimizations on QM-cluster and QM/MM studies of reaction energies in proteins. *J Chem Theory Comput*, 9(9):4205–4214, 2013.

[41] C. V. Sumowski and C. Ochsenfeld. A convergence study of QM/MM isomerization energies with the selected size of the QM region for peptidic systems. *The Journal of Physical Chemistry A*, 113(43):11734–11741, 2009.

[42] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and C. D. A. Development and testing of a general Amber force field. *J. Comput. Chem.*, 25:1157–1174, 2004.

[43] F. Weigend and R. Ahlrichs. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Physical Chemistry Chemical Physics*, 7(18):3297–3305, 2005.

[44] J. Wilcoxen and R. Hille. The Hydrogenase Activity of the Molybdenum/Copper-containing Carbon Monoxide Dehydrogenase of Oligotropha carboxidovorans. *Journal of Biological Chemistry*, 288(50):36052–36060, 2013.

[45] J. Wilcoxen, B. Zhang, and R. Hille. Reaction of the molybdenum-and copper-containing carbon monoxide dehydrogenase from Oligotropha carboxydovorans with quinones. *Biochemistry*, 50(11):1910–1916, 2011.

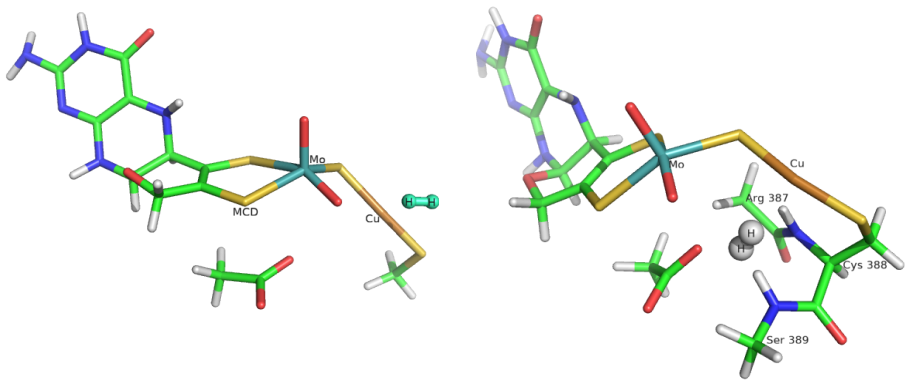
[46] K. Xu and H. Hirao. Revisiting the catalytic mechanism of Mo–Cu carbon monoxide dehydrogenase using QM/MM and DFT calculations. *Physical Chemistry Chemical Physics*, 20(28):18938–18948, 2018.

[47] B. Zhang, C. F. Hemann, and R. Hille. Kinetic and spectroscopic studies of the molybdenum-copper CO dehydrogenase from Oligotropha carboxidovorans. *Journal of Biological Chemistry*, 285(17):12571–12578, 2010.

Appendix A

QM/MM study of the binding of H_2 to MoCu CO dehydrogenase: development and application of improved H_2 parameters for a correct treatment of gas-protein interaction
Supplementary Information

Figure S1. System 1 used in the standard QM/MM calculations (a) and in QM/MM calculations including the backbone neighbouring Cys388 (b). Colour code: cyan, Mo; copper, Cu; yellow, S; blue, N; red, O; green, C; gray, H.



(a) (b)

List S2. List of all the buried charges included in the Big-QM calculations:

Arg27, 30, 56, 112, 126, 138, 188, 211, 261, 301, 312, 387, 391, 418, 460, 571, 797
Asp30, 42, 43, 103, 124, 192, 266, 313, 338, 439, 507, 779, 794
Glu117, 218, 299, 394, 400, 404, 533, 555, 705, 714, 763
Lys60, 185, 274, 296, 379, 759
Hip532

Appendix B

The direct role of water in CO oxidation catalysed by Mo-Cu CO-dehydrogenase. Answers from theory.
Supplementary Information

Figure S1 - The QM cluster model used in the calculations. Atoms shows as sphere were described using the def2-TZVP basis set during geometry optimization calculations.
Atoms labelled as "*" were kept fixed at the crystallographic coordinates during geometry optimization calculations.
Colour code: cyan, Mo; copper, Cu; yellow, S; blue, N; red, O; green, C; gray, H.

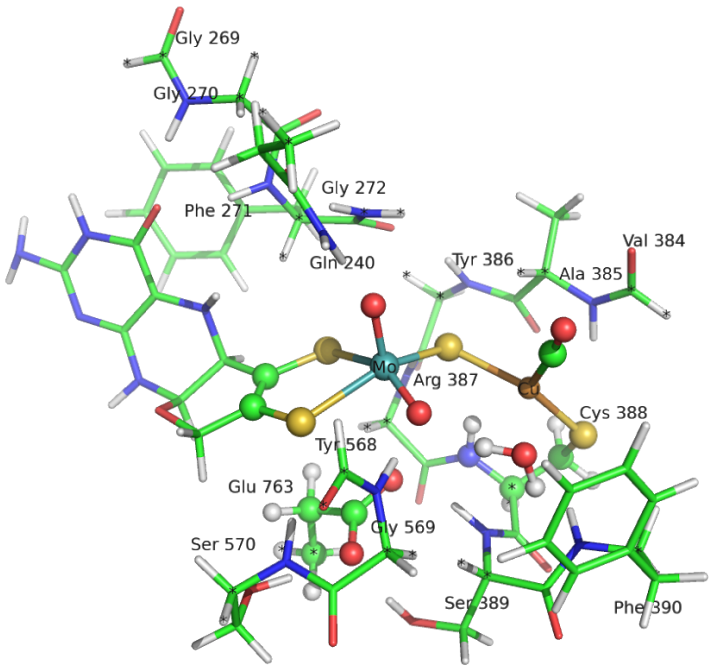
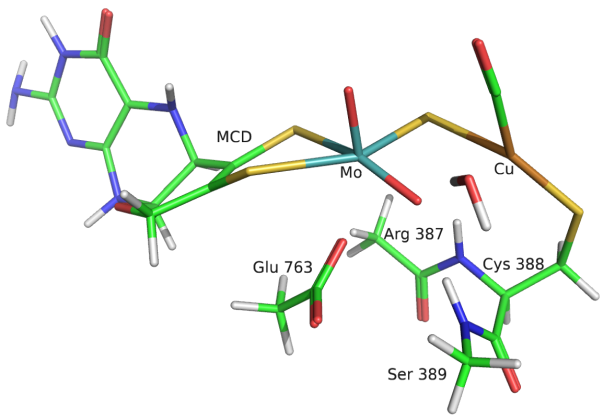


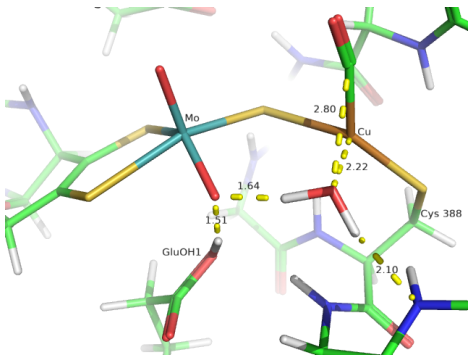
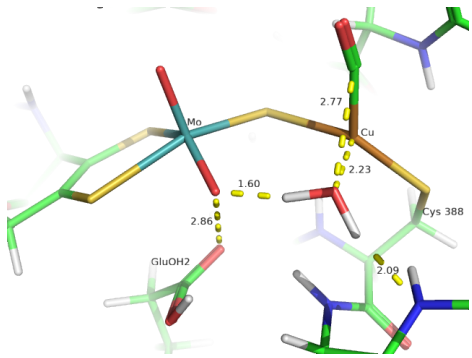
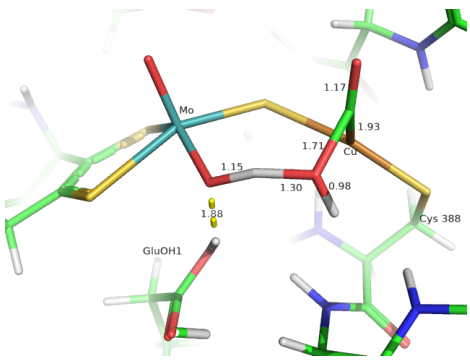
Figure S2 - System 1 used in QM/MM calculations. Colour code: the same as in Fig. S1.



List S3 - List of all the buried charges included in the Big-QM calculations:

Arg27, 30, 56, 112, 126, 138, 188, 211, 261, 301, 312, 387, 391, 418, 460, 571, 797
Asp30, 42, 43, 103, 124, 192, 266, 313, 338, 439, 507, 779, 794
Glu117, 218, 299, 394, 400, 404, 533, 555, 705, 714, 763
Lys60, 185, 274, 296, 379, 759
Hip532

Figure S4 - Intermediates **B1, C, TS1** GluOH1 and GluOH2

GluOH1	**GluOH2**
B1	
	
TS1	
	Not found.

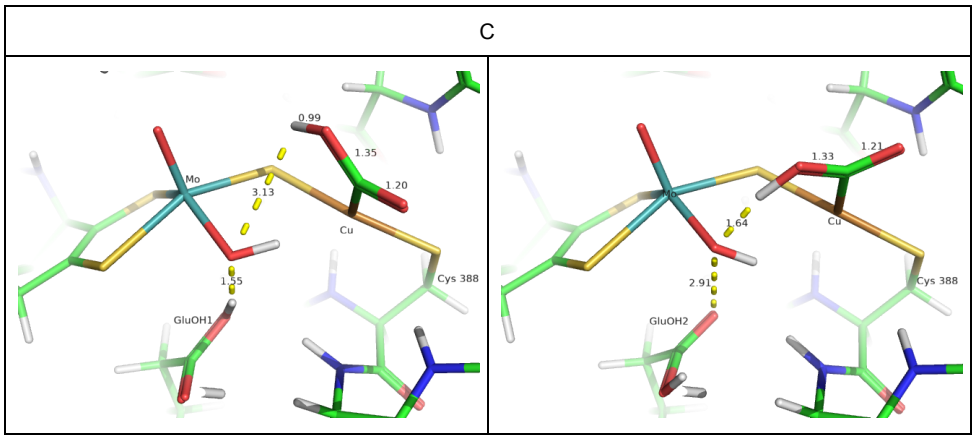


Figure S5 - Intermediate **B1** obtained with the QM/MM model. All distances in Å.

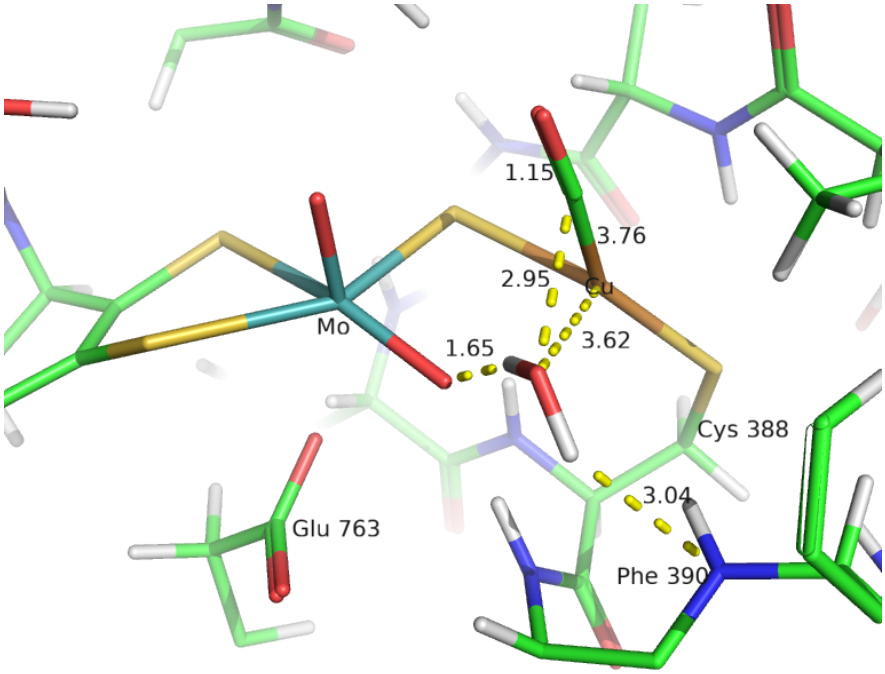


Figure S6 - Intermediate **TS1** obtained with the QM/MM model. All distances in Å.