Electronic Structure Studies of Iron-Sulfur Clusters: from Model Compounds to the Active Site of Nitrogenase

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Nico Spiller

aus

Hannover

Bonn 2022

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter: Prof. Frank Neese

2. Gutachter: Prof. Stefan Grimme

Tag der Promotion: 13. Juni 2022

Erscheinungsjahr: 2022

Summary

Iron-sulfur (FeS) clusters are omnipresent in nature, where they are involved in a variety of tasks such as electron transfer, DNA repair, Fe storage, and substrate activation. The enzyme nitrogenase contains one of the largest known biological FeS clusters. Nitrogenase is responsible for the conversion of inert N_2 gas to bioavailable ammonia, a reaction that is copied by the industrial Haber-Bosch process. However, despite decades of research, little is known about the molecular mechanism behind enzymatic N_2 reduction. The active site of nitrogenase is the iron-molybdenum cofactor (FeMoco, $[MoFe_7S_9C]$) in the MoFe protein, which contains a unique $\mu_6{-}C^{4-}$ center. The complex electronic structure of FeMoco pushes state-of-the-art quantum mechanical methods to their limits, therefore, smaller model compounds are often explored with high-level methods. The present work discusses the electronic structure of FeS clusters, ranging from monomeric and dimeric FeS model compounds up to a QM/MM model for FeMoco in the MoFe protein, and relates the results to experimental findings. The discussion of the results is separated into four chapters, each focusing on different electronic structure aspects of FeS clusters relevant to FeMoco.

Some S ligands in FeMoco can be selectively replaced with Se, which can be used as a probe for the electronic structure with element-specific spectroscopic techniques. However, it is unknown how much the electronic structure of FeMoco is altered by the replacement. The present work quantifies the perturbation of the S→Se substitution in $[Fe(XH)_4]^{1-,2-}$ and $[Fe_2X_2]^{2+,1+}$ (X = S, Se) model compounds. The analysis of the electronic structure focuses mostly on the multiconfigurational complete active space self-consistent field (CASSCF) wave function. The local electronic structure of the Fe−X bonds is characterized with *ab initio* ligand field theory (AILFT) and the angular overlap model (AOM), while the metal-metal interactions are related to spin Hamiltonian parameters. Se-based ligands show a smaller ligand field splitting and have a reduced donor strength compared to the S-based counterparts. For the homovalent $[Fe_2X_2]^{2+}$ clusters, S→Se substitution reduces the antiferromagnetic Heisenberg exchange coupling constant J by about 10 %. In the mixed-valent $[Fe_2X_2]^{1+}$ clusters, on the other hand, the coupling strength decreases by about 50 % upon S→Se substitution. The latter trend can be explained by increasing contributions from double exchange and vibronic coupling and is consistent with the mixture of spin states observed experimentally. S→Se substitution may therefore have a noticeable effect on the electronic structure of FeMoco, where antiferromagnetic coupling plays an important role.

Higher-nuclearity FeS clusters, such as FeMoco, often exhibit valence-delocalized $Fe^{2.5+}Fe^{2.5+}$ pairs as part of the electronic structure. In contrast, nearly all FeS dimers in the $[Fe_2S_2]^{1+}$ redox state have been reported to have a valence-localized $Fe^{2+}Fe^{3+}$ electronic structure. Cys→Ser variants of the $[Fe_2S_2]$ ferredoxins from *Clostridium pasteurianum* (*Cp*) and *Aquifex aeolicus* (*Aae*) are the only known examples for a valence-delocalized $[Fe_2S_2]^{1+}$ cluster. This work presents density functional theory (DFT) cluster model calculations for the $[Fe_2S_2]^{1+}$ ferredoxins from *Cp* and *Aae*. The electronic structure in the wild type model is localized, but delocalized in the Cys→Ser variant, consistent with experiment. Furthermore, protonation in the variant model leads to a localized electronic structure, which is consistent with the

experimentally observed pH dependence of the delocalization. The results suggest that the terminal ligands are central to valence delocalization.

The complex wave function of FeMoco is most often modeled using broken-symmetry (BS) DFT. Here, the local spins on each metal center can be aligned in a multitude of ways, but the most stable BS determinants are usually those that maximize the number of antiferromagnetically coupled Fe pairs. Furthermore, the localized orbital analysis yields the distribution of unpaired electrons across the metal centers and is helpful to rationalize structure-reactivity relations. The present work shows that calculating the coupling constants for different Fe pairs explicitly leads to an improved correlation between the energies and antiferromagnetic coupling. A comparison of localization algorithms suggests that Foster-Boys orbitals are the most robust in the context of the localized orbital analysis. Atomic charges, on the other hand, do not correlate with the localized orbitals analysis, but capture differences in the protein environment, such as in quantum mechanics/molecular mechanics (QM/MM) models of Mo nitrogenase compared to V nitrogenase. Furthermore, the atomic charges show that if the QM/MM boundary runs too close to the active site, the metal cluster becomes overpolarized by the MM charges. These findings help to interpret the results of DFT calculations for FeMoco and to choose a suitable QM region for QM/MM models.

The molecule CO is isoelectronic to N_2. It binds reversibly to Mo nitrogenase, where it acts as an inhibitor to N_2 reduction. V nitrogenase, on the other hand, reduces CO to hydrocarbons in a Fischer-Tropsch-like reaction. CO-inhibited nitrogenase gives rise to a number of experimentally well-characterized electron paramagnetic resonance (EPR) and infrared (IR) species, as well as X-ray diffraction (XRD) structures. However, details of the initial binding event are unknown, such as the redox state of the active site (E_n) or the role of the protein environment. In this work, the mechanism of CO binding to Mo and V nitrogenase is studied using QM/MM models. The models for the E_1 redox state feature a terminal and a bridging CO binding motif, where the calculated frequencies ($1922\,cm^{-1}$ and $1716\,cm^{-1}$) agree well with the experimentally observed IR bands ($1904\,cm^{-1}$ and $1715\,cm^{-1}$). Therefore, the QM/MM calculations are consistent with CO binding happening in the E_1 redox state. Alternatively, the calculated frequency for a semi-bridging CO in the E_2 QM/MM model ($1718\,cm^{-1}$) is also consistent with the latter IR band and the topology is the same as in the CO-bound XRD structure. Analogous models for V nitrogenase do not show significant differences, even though here CO binding has been reported without enzymatic turnover conditions. Furthermore, the careful analysis of the electronic structure reveals that CO coordination induces local spin pairing at the binding site. This, in turn, affects the magnetic interaction between the metal center and leads to an energy reordering of the BS determinants.

Acknowledgements

First and foremost, I want to thank my supervisor, Prof. Frank Neese, for welcoming me in his group, and for providing me with many fascinating research topics. I have learned so much during the past years and I am grateful for having been part of the diverse research environment at the MPI KOFO and CEC. Also, I want to thank Prof. Stefan Grimme for being my second supervisor at the University of Bonn. Furthermore, I would like to recognize the Max Planck Society for funding my PhD through the IMPRS-RECHARGE program.

I want to express my deep thanks to Vijay Chilkuri and Ragnar Björnsson, who have been invaluable mentors throughout my entire PhD. I am also grateful to Prof. Serena DeBeer for her support in the paper-writing process. Many thanks to Albert Thorhallsson, Barði Benediktsson, Zachary Mathe, Laure Decamps, Maxime Tarrago, and Corentin Poidevin for proof-reading my thesis. In addition to those mentioned above, I have had the great pleasure of working with Derek Rice, Aleksandra Wandzilak, Benjamin Cahier, and Romain Berraud-Pache. I want to thank all other the members of the Neese and the DeBeer groups who I did not mention explicitly for the many scientific, as well as non-scientific, discussions. They made my stay at the institute a unique and enriching experience. A special thanks goes to all the musicians at the MPI KOFO and CEC, who have shared their gear without hesitation with anyone wanting to play.

Finally, I would like to thank my parents Doris and Bernd, my sister Leonie, as well as all other family and friends, without whom this thesis would not have been possible. Last but not least, I have to thank Clara for her unwavering support.

Contents

1. Introduction

Nitrogen is one of the fundamental elements of life, occurring in biomolecules such as proteins or nucleic acids.[1] The largest nitrogen reservoir on the surface of the earth is the atmosphere, which consists of 78 % inert N_2 gas and is unavailable to most organisms. The chemical transformation of N_2 to bioavailable nitrogen sources, such as ammonia or amino acids, is called nitrogen fixation.[2] In nature, this process is performed almost exclusively by diazotrophic bacteria. Some of these bacteria live freely in soils and release nitrogen-containing compounds upon death. Others exist in symbiosis with higher plants, which supply them with carbohydrates in exchange for ammonia. Diazotrophs in symbiotic relationships are responsible for about 70 % of biological nitrogen fixation.[3] However, they are underrepresented in crop ecosystems, making the use of fertilizers necessary.

About 100 years ago, anthropogenic nitrogen fixation started becoming significant with the discovery of the Haber-Bosch process.[4,5] This industrial process uses heterogeneous, mostly Fe-based catalysts to synthesize ammonia from N_2 and H_2 gas. Estimates suggest that nowadays, anthropogenic nitrogen fixation has caught up with the biological fixation rate.[6] The Haber-Bosch process is typically run at high temperature (> 400 °C) and high pressure (> 200 bar) to acquire sufficient catalytic rates – in stark contrast to diazotrophic bacteria, which operate at ambient conditions.[7] Another demanding cost, however, comes with the H_2 gas required for the industrial process, which is synthesized via steam reforming from methane. As a consequence, the Haber-Bosch process is currently responsible for 1 %–2 % of the global energy consumption and about 1 % of the green house gas emission.[8,9] It has recently been shown that selective genetic engineering of bacteria enables the synthesis of small molecules through fermentation at industrial scale.[10] Therefore, understanding the details of biological nitrogen fixation could possibly open alternative pathways to the resource-intensive Haber-Bosch process. Furthermore, the enzyme responsible for nitrogen fixation in diazotrophic bacteria also reduces other substrates, *e.g.* protons to H_2 or CO and CO_2 to hydrocarbons, which is another highly relevant avenue for bio-inspired catalysis.[11,12]

The enzyme responsible for N_2 reduction in diazotrophic bacteria is called nitrogenase. Different forms of nitrogenase can be expressed depending on metal availability.[13] In the most common form, Mo nitrogenase, the active site is the so-called iron-molybdenum cofactor (FeMoco), which has the chemical formula $[MoFe_7S_9C]$. Enzymatic N_2 reduction is a complex, multistep process, where the active site cycles through eight one-electron reduction events.[14] Despite decades of research, little is known about the molecular mechanism and only a small fraction of the intermediate steps have been characterized experimentally (see Ref [15] for a recent review). The major problem is that enzymatic turnover typically yields a mixture of chemical species and the effective concentration of the active site is low due to the sheer size of the enzyme (about 40 000

atoms), which complicates the collection of spectroscopic data for well-defined reaction intermediates. Therefore, computational models of nitrogenase provide an opportunity to study selective redox states and structures, which help with the interpretation of experimental observations and allow for the proposition of mechanistic steps.

When designing computational models of nitrogenase, one faces two major challenges: (i) capturing the role of the protein environment during the mechanism and (ii) describing the complex electronic structure of the large metal cluster FeMoco. The protein creates a highly specific electrostatic environment around FeMoco and proximal amino acid residues may be directly involved in the reaction intermediates. Therefore, the protein has to be modeled, to some extent, explicitly. This is probably best realized in a QM/MM model, where the large part of the protein is modeled classically with molecular mechanics (MM), while the parts involved in the making and breaking of bonds are modeled with quantum mechanics (QM).[16] The QM part includes the active site and possibly important surrounding residues.

The number of open-shell metal centers in FeMoco strongly restricts the number of applicable QM methods. Routine QM/MM calculations therefore almost exclusively resort to density functional theory (DFT). DFT has been shown to yield accurate geometries for QM/MM models of nitrogenase.[17–19] However, DFT is also known to fail for other properties of bioinorganic systems, such as spin-state energetics or optical excitations, and it further suffers from the problem that the results cannot be systematically improved.[20] Alternatively, one can follow a bottom-up approach: For example, with recent advances in computational power and quantum mechanical algorithms, systems as large as $[Fe_2S_2]$ clusters can be modeled with multiconfigurational, wave function-based methods such as complete active space SCF (CASSCF).[21,22] The resulting multiconfigurational wave function reveals details about the magnetic interactions between two S-bridged Fe centers. Understanding the magnetic interactions in smaller FeS fragments is an important step towards modeling larger FeS clusters such as FeMoco.

The goal of this thesis is to promote the understanding of the relationship between the complex electronic structure, spectroscopic properties, and the reactivity of the active site of nitrogenase. To this end, computational models for smaller FeS clusters and FeMoco fragments are discussed, as well as large QM/MM models of the MoFe and the VFe protein. Chapter 2 investigates the S→Se substitution of the bridging ligands in synthetic $[Fe_2S_2]$ clusters, which have been shown to exhibit an unusual $S = \frac{3}{2}$ ground state in the reduced, Se-substituted form.[23] Chapter 3 constitutes a brief exploration the $[Fe_2S_2]$ ferredoxins from *Clostridium pasteurianum* (*Cp*) and *Aquifex aeolicus* (*Aae*), the only known examples of an $S = \frac{9}{2}$ ground state in $[Fe_2S_2]^{1+}$ clusters.[24] Chapter 4 is a thorough analysis of the computational models used for FeMoco and FeVco, the active sites of Mo and V nitrogenase, respectively. Chapter 5 uses the previously introduced models to study the binding of CO to the active site of nitrogenase. Large parts of Chapter 2 and Chapter 5 have been published in Ref [22] and Ref [25], respectively. The remainder of the introduction provides an overview of the computational methods used throughout this work (Section 1.1). Furthermore, it summarizes the current state of the relevant literature (Section 1.2).

1.1. Computational approaches

The following sections give an introduction to the quantum chemical concepts that are used throughout this work. Section 1.1.1 outlines the complete active space self-consistent field (CASSCF) ansatz for a multiconfigurational wave function as well as the n-electron valence state perturbation theory (NEVPT2). Section 1.1.2 discusses density functional theory (DFT) and its application for chemical systems. Section 1.1.3 contains a brief outline of molecular mechanics (MM) and how it is combined with quantum mechanics in QM/MM calculations. Section 1.1.4 covers different spin Hamiltonians that are commonly used to characterize magnetic interactions in molecules with multiple magnetic centers. Section 1.1.5 and 1.1.6 explain how population analysis and localized orbital analysis help to interpret the calculated electronic structure.

1.1.1. Wave function-based approaches

Time-independent Schrödinger equation

In quantum chemistry, stationary states are described by the time-independent Schrödinger equation

$$\hat{H}\Psi_n = E_n\Psi_n \tag{1.1}$$

where $\hat{H}$ is the molecular Hamiltonian, which contains all non-relativistic interactions involving electrons and nuclei. Eq 1.1 is an eigenvalue equation and solving it yields a set of eigenfunctions, the molecular wave functions Ψ_n, and a set of eigenvalues, the energies E_n of the n^{th} state.[26,27] The molecular Hamiltonian completely defines a chemical system in the non-relativistic limit and knowledge of the wave function allows for the extraction of all non-relativistic properties of that system. However, chemical systems present many-body problems and analytical solutions are not possible except for very simple cases. Also, the exact numerical solution is afflicted with a computational expense that render this approach inapplicable for all but the smallest systems. Therefore, solutions to Eq 1.1 can only be acquired in an approximate fashion. It is the goal of quantum chemistry to find the balance between accuracy and computational cost of these approximations.

The most common approximation to Eq 1.1 is the Born-Oppenheimer (BO) approximation. Here, the movement of the electrons is decoupled from the movement of the nuclei and molecular wave function becomes a product of the electronic wave function Ψ^{ele} and the nuclear wave function Ψ^{nuc}. The 10^3–10^5 larger mass for the nuclei justifies the use of the BO approximation, because the electrons respond nearly instantly to changes in the molecular geometry. This assumption allows for a separation of the molecular Hamiltonian $\hat{H}$ in an electronic term and a nuclear term – the coupling term is neglected. The electronic wave functions Ψ_n^{ele} are then the eigenfunctions of the electronic Hamiltonian

$$\hat{H}^{\text{ele}} = -\frac{1}{2}\sum_i \nabla_i^2 + \sum_{i<j}\frac{1}{r_{ij}} - \sum_{i,A}\frac{Z_A}{r_{iA}} \tag{1.2}$$

where the indices i, j run over all electrons and A over all nuclei in the system. ∇^2 is the Laplacian operator, r is the distance, and Z_A is the nuclear charge. The positions of the nuclei r_{iA} therefore enter as parameters in the electronic Hamiltonian. Solving the electronic part of the Schrödinger equation for different nuclear positions, *i.e.* molecular geometries, yields the potential energy surface of the system. The error introduced by the BO approximation is usually small compared to the approximations used for the electronic part. The coupling of the electron and nuclear motion is only strong in the proximity of a conical intersection, where the BO approximation breaks down.

Hartree-Fock wave function

Hartree-Fock (HF) theory describes an approximation to the electronic Schrödinger equation. Here, the complexity is reduced by restricting the form of the electronic wave function Ψ^{ele}. The ansatz for Ψ^{ele} is the Slater determinant

$$\Psi^{\text{SD}} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(1) & \cdots & \psi_1(N) \\ \cdots & \ddots & \cdots \\ \psi_N(1) & \cdots & \psi_N(N) \end{vmatrix} \tag{1.3}$$

where N is the number of electrons and ψ_i are one-electron functions, the spin orbital. A determinant is a linear combination of all possible matrix element products, which changes sign when two rows or columns are switched. The Slater determinant therefore conveniently includes two properties of the exact electronic wave function, namely, that the electrons are indistinguishable and that the wave function is antisymmetric with respect to the exchange of two electrons.

When the energy of a Slater determinant is minimized under the constraint that the spin orbitals remain orthonormal, one ends up with a set of N equations, the so-called HF equations

$$\hat{f}\psi_a = \varepsilon_a \psi_a \tag{1.4}$$

where the Fock operator $\hat{f}$ contains the core operator $\hat{h}$, the Coulomb operator $\hat{J}$, and the exchange operator $\hat{K}$

$$\hat{f}(1) = \hat{h}(1) + \sum_b \hat{J}_b(1) - \hat{K}_b(1) \tag{1.5}$$

$$\tag{1.6}$$

$$\hat{h}(1) = -\frac{1}{2}\nabla_1^2 - \sum_A \frac{Z_A}{r_{1A}}$$

$$\hat{J}_b(1)\psi_a(1) = [\int \frac{\psi_b^*(2)\psi_b(2)}{r_{12}}\,\mathrm{d}2]\psi_a(1)$$

$$\hat{K}_b(1)\psi_a(1) = [\int \frac{\psi_b^*(2)\psi_a(2)}{r_{12}}\,\mathrm{d}2]\psi_b(1)$$

Non-interacting in the sense of HF theory means that the Fock operator is a one-electron operator, which does not couple the movement of two electrons. However, the Coulomb and exchange operators implicitly depend on all orbitals: the electron in ψ_a feels the average effects of all other orbitals ψ_b. The Fock operator depends on the solution of the Fock equation, therefore, they can only be solved iteratively in a so-called self-consistent field (SCF). In practice, the spin orbitals are not varied freely, but are expanded within a basis

$$\psi_j = \sum_i c_{ij}\phi_i$$

The expansion coefficients c_{ij} are the parameters that are optimized during the SCF and are referred to as molecular orbital (MO) coefficients.

CASSCF wave function

For many chemical systems, a single determinant is an appropriate approximation for the exact ground state wave function, since it gives a qualitatively correct description. Therefore, HF theory provides a good starting point for more elaborate methods that recover dynamic electron correlation, such as configuration interaction (CI), many-body perturbation theory (MBPT), coupled cluster (CC) theory (see Ref [26] for introductions). However, a single determinant is not a good starting point for describing certain electronically excited states, the breaking of chemical bonds, or magnetic coupling. In these cases, it is advisable to approximate the exact wave function as a linear combination of multiple Slater determinants in a so-called multiconfigurational SCF (MCSCF) wave function

$$\Psi^{\mathrm{MCSCF}} = \sum_I C_I \Psi_I^{\mathrm{SD}} \tag{1.7}$$

where C_I are the expansion coefficients of Ψ_I^{SD}. Therefore, minimizing the energy for a Ψ^{MCSCF} wave function implies optimizing the expansion coefficients C_I alongside the MO coefficients c_{ij} for each Ψ_I^{SD}. In principle, the Ψ^{MCSCF} wave function can be a linear combination of an arbitrary number of Slater determinants, where each one has a unique set of MO coefficients. Even when using a single, finite atomic basis set (ϕ_i) to construct excited Slater determinants, and then expand the Ψ^{MCSCF} wave function, the dimension of the wave function becomes unmanageable even for small chemical systems. Therefore, the determinants included in the expansion need to be constrained.

A practical strategy to tame the MCSCF wave function is the complete active space SCF (CASSCF) approach, which is illustrated in Figure 1.1a. Here, a set of MOs is sorted into three categories: internal, active, and virtual orbitals. The internal orbitals are always doubly occupied, while the virtual orbitals are always empty. The CASSCF wave function expanded by all possible excited Slater determinants, where the active orbitals are either empty, singly, or doubly occupied. In other words, the dimension of the CASSCF problem is equivalent to the full CI problem in the subspace spanned by n electrons in o active orbitals.

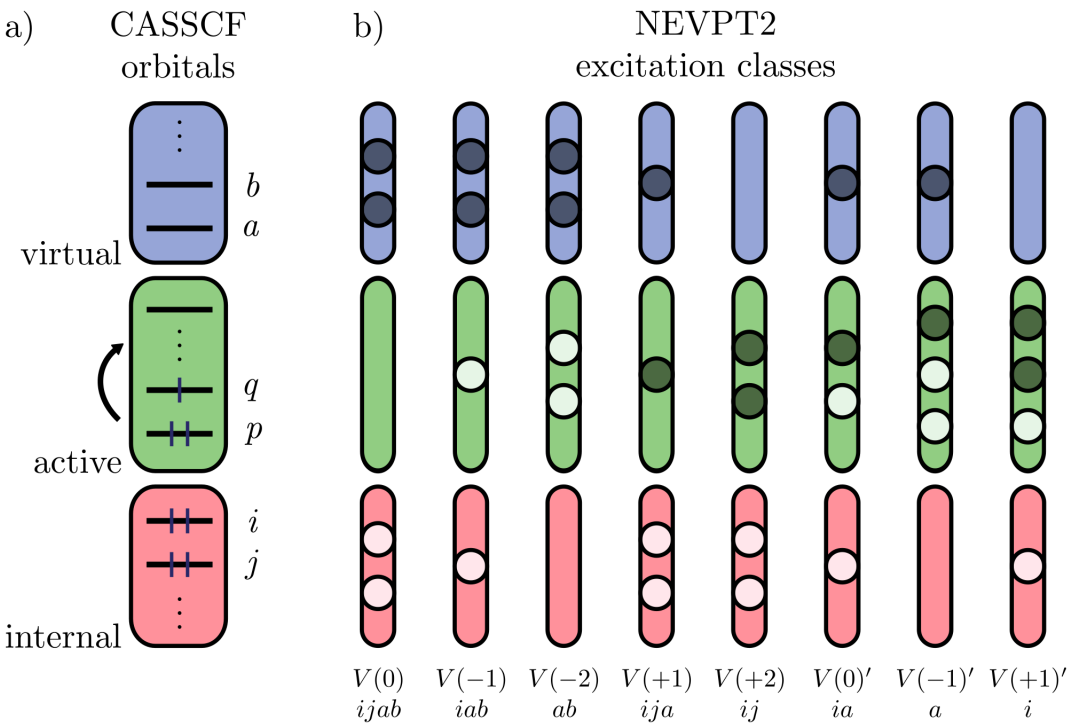

Figure 1.1. – (a) Schematic representation of orbital spaces in a CASSCF wave function and (b) the excitation classes in NEVPT2 (white: hole, black: particle). The n in $V(n)$ indicates the change in the number of electrons in the active space.

The CASSCF method requires its user to choose the number of electrons n and orbitals o for the active space, which is abbreviated as CAS(n,o). This choice is not unique, since it depends on the computational chemist to judge whether a CAS(n,o) yields a qualitatively correct wave function. Therefore, the CASSCF approach is not a black box method, but presumes a deep understanding of the electronic structure in question. In practice, even the CAS(n,o) is insufficient information to characterize the CASSCF wave function, because it is not always evident *which* o active orbitals give the lowest energy and the results of the orbital optimization depends heavily on the starting orbitals.

The MCSCF wave function, just like the exact non-relativistic electronic wave function, is an eigenfunction of the total spin operator $\hat{S}^2$.[28] Slater determinants, on the other hand, are generally not eigenfunctions of $\hat{S}^2$, but only of $\hat{S}_z$. Therefore, the C_I coefficients in Eq 1.7 can not be varied freely. A convenient way to make the Ψ^{MCSCF} wave function an eigenfunction of $\hat{S}^2$ is to summarize Slater determinants in spin-adapted linear combinations. In these so-called configuration state function (CSF), all Slater determinants have the same spatial part and differ only in their spin part. Therefore, a given CSF belongs to one and only one electron configuration (CFG), but not the other way around.

It is common practice to optimize one CASSCF wave function for multiple states simultaneously; the effects of this state averaging procedure is discussed in Section 2.2.3 and 2.3.3. The number of possible excited states is defined solely by the size of the

active space. A utility program to calculate the number of states for an arbitrary CAS(n,o) is presented in Section A1.2. Naturally, the number of states is equal to the number of CSFs in the expansion of the CASSCF wave function.

n-electron valence state perturbation theory (NEVPT2)

The correlation of electron movement, which means that an electron responds instantaneously to the position of another electron, is only partly possible within an approximate wave function. For example, the antisymmetry requirement introduces some correlation for electrons with the same spin, since they cannot occupy the same point in space. Furthermore, the electrons occupying the active space orbitals in a CASSCF wave function have the ability to avoid each other by "jumping" into an excited CSF (Figure 1.1a). The remaining electron correlation, which is contained in the excited CSFs beyond the active space, can be recovered via multireference perturbation theory (MRPT),[29] multireference configuration interaction (MRCI),[30,31] or multireference coupled cluster (MRCC).[32,33] The correlation included in the CASSCF wave function is typically referred to as static correlation, that recovered with post-CASSCF methods as dynamic correlation.

The MRPT approaches are the least computationally demanding and therefore applicable to the models in the present study. Electron correlation effects arise at least at the second order of perturbation theory, and two popular approaches are the n-electron valence state perturbation theory (NEVPT2)[34–37] and the complete active space second-order perturbation theory (CASPT2).[38] They differ in their choice of the zeroth order Hamiltonian, which includes only one-body terms for CASPT2, but one- and two-body terms for NEVPT2. The quality of the eigenstates of the zeroth order Hamiltonian is worse for CASPT2, which leads to intruder states and the necessity to apply an arbitrary level shift to the energy of the perturbation functions. Intruder states are mostly avoided in NEVPT2 by the choice of a better-suited zeroth order reference. NEVPT2 further has the advantage to be size consistent and invariant with respect to rotations of the active space orbitals; therefore, it is used throughout this work.

Since the Hamiltonian includes only one- and two-body terms (Eq 1.2), only singly and doubly excited CSFs need to be considered in the perturbational treatment. With the orbital subspaces in the CASSCF wave function (internal, active, virtual), the perturbation functions can be categorized into eight classes, which are shown in Figure 1.1b. Note that the double excitation within the active space do not generate perturbation functions, since they are already mixed variationally into the CASSCF wave function. Depending on the size of the active space, the number of CSFs used to expand the CASSCF wave function can be quite large.[†] If each CSF in this expansion is excited individually, even second order MRPT quickly becomes computationally

[†]The number of possible CSFs can quickly explode. For example in $[Fe_2S_2]^{2+}$, the antiferromagnetically coupled $S = 0$ ground state has to be described at least by a CAS(10,10) wave function (only Fe d orbitals in the active space). This state is expanded by 20 000 CSFs. When including the occupied ligand p orbitals in the active space, the active space grows to CAS(22,16) and the number of possible CSFs increases to 4 500 000. In both cases, however, only a single CSF dominates the wave function.

unmanageable. Therefore, NEVPT2 and CASPT2 make use of a concept called internal contraction. Here, a give excitation is only applied to the entire CASSCF wave function and not every single CSF. This has the advantage that the number of perturbation functions is kept small for larger active spaces.

1.1.2. Density functional theory

Two properties of the molecular ground state electron density $\rho(r)$ are that the integral over the entire physical space is equal to the number of electrons

$$N_{el} = \int \rho(r)\,\mathrm{d}r$$

and that the cusps, which are only located at the position of the nuclei, are related to the nuclear charge Z_A through

$$\lim_{r \to r_A} [\frac{\mathrm{d}}{\mathrm{d}r} + 2Z_A]\bar{\rho}(r) = 0$$

where r_A is the position of atom A and $\bar{\rho}(r)$ is the spherical average of $\rho(r)$.[39] Therefore, the electron density provides all information required to define a molecular system: the number of electrons and the geometry. This recognition forms the basis of density functional theory (DFT) and the goal is to determine the electronic energy and other properties that can be derived from the ground state density. A more rigorous foundation is given by the two Hohenberg-Kohn theorems, which prove that a functional exists that maps the ground state electron density onto the ground state energy (proof of existence) and that only the density yielding the lowest energy is the ground state density (variational principle).[40] However, the proof of existence remains silent about the explicit definition of the functional and the variational principle can also not guide the design of approximate functionals, since it is only valid for the exact functional.

Kohn-Sham approach to DFT

In analogy to the electronic Hamiltonian (Eq 1.2), the electronic energy can be written as

$$E(\rho) = T(\rho) + E_{ee}(\rho) + E_{Ne}(\rho)$$

where T is the kinetic energy, E_{ee} is the electron-electron repulsion energy, and E_{Ne} the electron-nuclei attraction energy. The electron-nuclei term is system-dependent and can be written exactly as a functional of the density as

$$E_{Ne}(\rho) = \int \rho(r) V_N\,\mathrm{d}r$$

where V_N is the electrostatic potential of the nuclei (and possibly other electric charges). The other two describe the behavior of electrons and are therefore valid for any system. However, their explicit form is unknown and the goal of DFT is to find accurate approximations. The following approach is used by most density functionals, but it is by no means unique.

The starting point for approximating the electron-electron interaction is the Coulomb repulsion $J(\rho)$ that would be present in a classical charge density:

$$\begin{aligned} E_{ee}(\rho) =& J(\rho) + \Delta E_{ee}(\rho) \\ =& \frac{1}{2} \int \frac{\rho(r_1)\rho(r_2)}{r_{12}} \, \mathrm{d}r_1 \, \mathrm{d}r_2 + \Delta E_{ee}(\rho) \end{aligned}$$

The correction term $\Delta E_{ee}(\rho)$ collects all remaining non-classical contributions, which includes the nonphysical electron self-interaction present in $J(\rho)$ and the exchange and Coulomb correlation missing in $J(\rho)$.

As a starting point for the kinetic energy, Kohn and Sham proposed an orbital-based approach, inspired by the success of methods such as HF.[41] They defined a fictitious system of N electrons, where the interactions are given by the Hamiltonian

$$\hat{H}_{\mathrm{KS}} = -\frac{1}{2} \sum_i \nabla_i^2 + \sum_i V_{\mathrm{eff}}(r_i)$$

The first term is the kinetic energy of the electrons and the second is an arbitrary local potential. The operator $\hat{H}_{\mathrm{KS}}$ is a sum of one-electron operators and therefore describes a system of non-interacting electrons.[26] The eigenfunction for such an operator is a Slater determinant (Eq 1.3). The key idea behind the Kohn-Sham approach is that the effective potential V_{eff} yields a determinant, which has the same density as the exact density $\rho(r)$. Naturally, V_{eff} depends on the molecular geometry, but also accounts for all electron correlation effects that are included in the exact density.

Solving the eigenvalue equation under the constraint that the eigenvectors are orthonormal yields a set of eigenvalues ϵ^{KS} and eigenvectors ψ^{KS}, which form the so-called Kohn-Sham determinant. It is important to realize that the Kohn-Sham determinant is *not* related to the exact wave function, in contrast to HF theory, where the Slater determinant approximates the exact wave function. To reiterate, the Kohn-Sham determinant describes a fictitious system of non-interacting electrons that happen to have the same electron density as the exact system. However, the kinetic energy in this system is easily accessible through

$$T_{\mathrm{KS}} = -\frac{1}{2} \sum_i \psi_i(x)^* \nabla_i^2 \psi_i x \, \mathrm{d}x$$

which is expected to capture a large part of the exact kinetic energy

$$T(\rho) = T_{\mathrm{KS}}(\rho) + \Delta T(\rho)$$

The remaining contributions to T are collected in the correction term ΔT. Note that in the electron density ρ does not explicitly appear in the current definition T_{KS}, but the Kohn-Sham orbitals depend on ρ through the effective potential V_{eff}.

With the conceptual separation of the kinetic energy $T(\rho)$ and the electron-electron energy $E_{ee}(\rho)$, the two correction terms make up the so-called exchange-correlation energy

$$E_{\mathrm{xc}}(\rho) = \Delta T(\rho) + \Delta E_{ee}(\rho) = (T(\rho) - T_{\mathrm{KS}}(\rho)) + (E_{ee}(\rho) - J(\rho))$$

The name "exchange correlation" is actually slightly inaccurate, since the quantity has to correct – in addition to the exchange and Coulomb correlation – the kinetic energy in the non-interacting Kohn-Sham determinant as well as the self-interaction error present in $J(\rho)$. The goal of Kohn-Sham DFT is to find approximate functionals for E_{xc}.

Approximate density functionals

In wave function theory, the approximate solution to the Schrödinger equation (Eq 1.1) can be gradually improved to approach the correct full CI wave function. In Kohn-Sham DFT, however, the approximations for the density functional E_{xc} cannot be systematically improved, since the exact form of E_{xc} is not known. The quality of an approximate density functional (in the following text simply: functional) is generally judged based on its performance in benchmarking studies. This has led to a plethora of functionals that predict certain properties of a limited class of compounds fairly well, but fail dramatically in other applications. To illustrate, focusing on the prediction of better electronic energies has led to worse electron densities, which clearly is not in the interest of an accurate computational model.[42] The application-oriented reader should refer Ref [43] for a recent overview of popular functionals.

The functionals categorized based on decisive quantities appearing in E_{xc}. The categories generally have a different accuracy-to-cost ratio and are therefore sorted on the so-called Jacob's ladder, where the lower rungs indicate less accurate but computationally efficient approximations.[43,44] The lowest rung is reserved for local density approximation (LDA) functionals, which are defined as

$$E_{\mathrm{xc}} = \int \rho(r)[\epsilon_x(\rho(r)) + \epsilon_c(\rho(r))]\,\mathrm{d}r$$

where ϵ_x and ϵ_c are the exchange and correlation functional for the uniform electron gas, respectively. An analytical expression exists for the term ϵ_x, but the term ϵ_c can only be solved numerically. Since ϵ_c can be approximated by different analytic expressions, different LDA functionals exist. An example is the VWN functional.[45]

The second rung of Jacob's ladder is occupied by generalized gradient approximation (GGA) functionals, which depend, in addition to the electron density, on its gradient

$$E_{\mathrm{xc}} = \int f(\rho(r), \nabla\rho(r))\,\mathrm{d}r$$

Explicit expressions are generally fairly involved, but the density gradient often enters in form of the scalar, dimensionless quantity $s(r) = \frac{|\nabla\rho(r)|}{\rho^{4/3}}$. A common practice is to mix the exchange and correlation terms from different authors. For example, the popular GGA functional BLYP uses the exchange term defined by Becke (B86)[46] and the correlation term defined by Lee, Yang, and Parr (LYP).[47] Unsurprisingly, the third rung on Jacob's ladder is reached by including the second derivative of the density. An example of these so-called meta-GGA (mGGA) functionals is the TPSS functional.[48]

The functionals of the three lowest rungs are purely local; they only depend on the value of the density and its derivatives at a given point in space. However, electron correlation is a non-local effect. For example, the HF wave function includes some

correlation for electrons with parallel spin: the integral defining the effect of the exchange operator acting on some spin orbital ($\hat{K}_b$ in Eq 1.5) depends on the value of that orbital everywhere in space. The fourth rung of Jacob's ladder is occupied by DFT/HF hybrid functionals, which use a mixture of local exchange and non-local HF exchange. The most popular example is the B3LYP functional,[49] which includes the exchange and correlation parts from the GGA functional BLYP, but also 20 % HF exchange. Another example is the TPSSh[50] functional, which includes 10 % HF exchange and is based on the mGGA functional TPSSh.[48] In the present work, TPSSh is used for all DFT calculations, because a low HF exchange has been shown to yield accurate geometries for FeS clusters in general[51–53] and for the active site of nitrogenase in particular.[54]

Non-local correlation effects can also be incorporated into E_{xc} by using the virtual Kohn-Sham orbitals.[55] These functionals constitute the fifth rung of Jacob's ladder. If perturbation theory is used to describe electron correlation, one speaks of double-hybrid functionals, which belong to the most accurate approximate functionals, but also come with a high computational cost. Other promising trends in functional development include the family of SCAN functionals,[56–58] which obey physical constraints known for the exact functional, or the DM21 functional,[59] which was trained by a neural network to correctly handle fictitious systems with fractional charge and spin.

Dispersion

Dispersion refers to attractive forces between neighboring molecules or parts of the same molecule due to fluctuations in the electron density.[39] The fluctuations generate transient electric multipole momenta, which, in turn, induce multipoles in the proximity. The attractive interaction decays as r^{-6}, r^{-8}, *etc.* for the interaction between dipoles, quadrupoles, *etc.* At short distances, the electron-electron repulsion dominates, resulting in an energy minimum at a specific distance. The strength of the dispersion interaction depends on the ease with which fluctuations in the electron density can occur and is therefore related to the polarizability of a chemical species.[60] Dispersion is often illustrated as an intermolecular effect, but it is important to realize that it is also an important interaction within molecules that has a decisive influence on the geometry.

The induction of multipoles is a long-range electron correlation effect and is therefore not captured in approximate local functionals.[39] A practical way to include dispersion effects in a functional is the DFT-D correction developed by Grimme *et al.*[61–65]. In the DFT-D3(BJ) method[64], which is used for DFT calculations in the present work, the dispersion energy is given by

$$E^{\text{disp}} = \sum_{AB}[s_6 \frac{C_{6,AB}}{r_{AB}^6 + f_{\text{damp},AB}^6} + s_8 \frac{C_{8,AB}}{r_{AB}^8 + f_{\text{damp},AB}^8}]$$

where the sum runs over all atom pairs. r_{AB} is the distance between atoms A and B, $C_{n,AB}$ is the pair-wise dispersion coefficient, s_n is a functional dependent scaling factor, and $f_{n,\text{damp}}$ is a damping function. Only the dipole-dipole (r^{-6}) and quadrupole-quadrupole (r^{-8}) interactions are considered, since higher-order terms have been found to introduce instabilities. The dispersion coefficient $C_{6,AB}$ is derived from the calculated

polarizability of atoms A and B. It also respects the chemical environment of each atom in the molecule: a higher coordination number diminishes the polarizability and therefore also the dispersion interaction. $C_{8,AB}$ is calculated recursively from $C_{6,AB}$. The damping function $f_{\text{damp},AB}$, Becke-Johnson (BJ) in this case, is necessary to avoid singularities at short interatomic distances.

The dispersion energy is added to the Kohn-Sham energy expression and therefore changes the potential energy surface, which leads to greatly improved geometries and thermodynamic properties.[43,66] More elaborate DFT-D methods incorporate three-body interactions (`ABC` keyword in `ORCA`)[61] or consider the atomic charge in a given molecule (`D4` keyword in `ORCA`).[65] An alternative approach is to included dispersion explicitly in a so-called van-der-Waals (vdW) functional,[67] where dispersion also affects the Kohn-Sham orbitals. In `ORCA`, appending `-NL` to a functional name includes the non-local part of the VV10 vdW functional.[68]

1.1.3. QM/MM approach

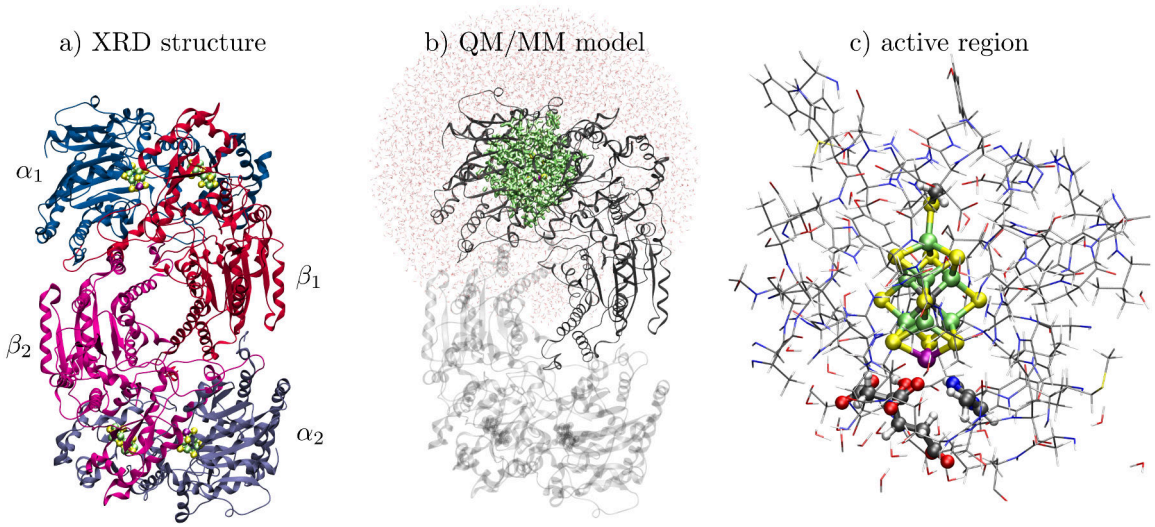

Figure 1.2. – The MoFe protein as an example for the QM/MM approach. (a) The XRD structure of the E_0 MoFe protein has four subdomains and a pseudo-C_{2v} symmetry.[69] (b) The QM/MM model developed by Benediktsson and Bjornsson includes only the α_1 and β_1 subdomains, which are enclosed in a water droplet.[17] The active region is shown in green; the shaded region is not included. (c) The active region is free to move during geometry optimizations (about 1000 atoms). The QM region is a subset of the active region (about 50 atoms, ball and stick).

The quantum mechanics/molecular mechanics (QM/MM) approach[16,70] is usually used to study a larger chemical system with a hybrid computational model: a small part of the system is described with a high-level QM method, called the QM region, while the remainder of the system is described with a low-level molecular force field, called the MM region. This is shown in Figure 1.2 for the MoFe protein of nitrogenase, which contains 40 000 atoms, including hydrogen. The advantage of this separation is to treat the region of interest with high accuracy while still capturing the relevant influence

outside the region of interest at low computational cost. In some contexts, the system is further partitioned into an active region, in which the atoms are unconstrained, and an inactive region, where the position of the atoms are fixed. This is done when structural changes are expected to affect only a certain part of the system, and to limit the number of degrees of freedom.

Important considerations for a QM/MM model are (i) the size and level of theory for the QM region, (ii) the choice of the MM force field, (iii) the interface between the QM and the MM region, and (iv) the size of the active region. If applied correctly, the QM/MM approach is a powerful tool to study metalloenzymes, such as the MoFe protein of nitrogenase, since their reactivity depends both on the unique electronic structure of the FeMoco active site and the distinct environment created by the protein. However, larger conformational changes cannot be captured in a simple QM/MM geometry optimization, but require molecular dynamics (MD) simulations in order to describe a thermodynamic ensemble. MD simulations are not easy to realize for systems such as the MoFe protein of nitrogenase, because the level of theory required for an accurate description of the QM region[17] makes the simulation of a reasonable timescale unfeasible. Promising approaches to treat these large systems are semi-empirical QM/MD simulations[71,72] or metadynamics.[73] The focus of the following section lies on the QM/MM setup used in this work, which uses a modified version of the `CHARMM36` force field for the MM part,[74] `ORCA` for the QM part,[75–77] and `chemshell` for the QM/MM interface.[78,79]

Molecular force fields

A molecular force field contains a set of interatomic potentials that capture the relevant interactions in a chemical system. One typically distinguishes between bonded and non-bonded potentials. Bonded potentials are used to describe the interaction between atoms connected via one, two, or three covalent bonds, therefore, they depend on the topology of the system. Non-bonded potentials describe the interaction between the remaining atom pairs. The following potentials are used in the `CHARMM` (CHemistry at HARvard Macromolecular Mechanics) force field.[74] Other popular force fields for biomolecules are `AMBER` (Assisted Model Building and Energy Refinement),[80] `GROMOS` (GROningen MOlecular Simulation) program package,[81] and `OPLS-AA`.[82]

The covalent *bonds* are described by the harmonic potential

$$V_b(b) = K_b(b - b_0)^2$$

where b is the distance between the two atoms. The force constant K_b, and the equilibrium distance b_0 are the force field parameters. The harmonic approximation is most accurate for small displacements around the equilibrium distance.[27] More accurate potentials, such as higher-order polynomials or the Morse potentials, capture anharmonicity effects, but they increase the number of parameters and may even decrease the accuracy of the predicted structures. When a covalent bond is stretched, the electronic structure becomes more and more complex, and a classical force field potential cannot easily capture the energy. Therefore, if the formation and breaking of chemical bonds is restricted to the QM region of QM/MM models, the harmonic approximation is well-suited in most applications.

The *angles* are described by the sum of two harmonic potentials

$$V_a(\theta, S) = K_\theta(\theta - \theta_0)^2 + K_{\mathrm{UB}}(S - S_0)^2$$

In the first term, θ is the angle defined by three covalently bound atoms i, j, and k and the force constant K_θ and the equilibrium angle θ_0 are the force field parameters. In the second term, S is the distance between the atoms i and j and the force constant K_{UB} and the equilibrium distance S_0 are the force field parameters. This is the so-called Urey-Bradley potential, which accounts for van-der-Waals interactions (*vide infra*) between two atoms bound to the same center, since these depend on the actual distance of the atoms and not simply the angle.

The *dihedral angles* are described by the periodic potential

$$V_{\mathrm{dih}}(\chi) = K_\chi(1 + \cos(n\chi - \delta))$$

where χ is the dihedral angle formed by atoms i, j, k, and l. The potential height K_χ and the offset δ are the force field parameters. The integer n defines the periodicity of the potential and is determined by the topology of the atoms.

The *improper torsion angles* are described by the harmonic potential

$$V_{\mathrm{imp}}(\psi) = K_\psi(\psi - \psi_0)^2$$

where ψ captures the out-of-plane bending of three planar atoms centered around a fourth. The force constant K_ψ and the equilibrium angle ψ_0 are the force field parameters. Even though ψ formally is a dihedral angle, V_{imp} is not a periodic potential, because it is used to maintain planarity or stereochemistry of certain atoms, for example an sp^2 carbon or an NR_3 amine. Improper torsion angles also fix the structure of transition metal complexes in a force field description.

The *van-der-Waals interaction* is described by the Lennard-Jones potential

$$V_{\mathrm{LJ}}(r) = \epsilon_{ij}[(\frac{R_{\mathrm{min}}}{r})^{12} - 2(\frac{R_{\mathrm{min}}}{r})^6]$$

where r is the distance between atoms i and j. The potential depth ϵ_{ij} and the equilibrium distance R_{min} are the force field parameters. The two terms, respectively, describe the repulsive part of the van-der-Waals interaction, which arises from the repulsion of the overlapping electron densities at short distances, and the attractive term, which is due to dispersion. The attractive $-r^{-6}$ term has the physically correct long range behavior of interacting dipoles, but the exponent of the repulsive r^{-12} term was chosen to be the attractive term squared in order to reduce the computational cost. In the `CHARMM` force field, the parameters are not fitted for every atom pair, but only for each atom. The individual and the pair-wise quantities through the geometric mean $\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}$ and the arithmetic mean $R_{\mathrm{min},ij} = R_{\mathrm{min},i}/2 + R_{\mathrm{min},j}/2$.

The *electrostatic interaction* is defined by the electrostatic potential

$$V_{\mathrm{E}}(r) = \frac{q_i q_j}{4\pi\epsilon_0\epsilon_R r}$$

where r is the distance between the atoms i and j. The charges q_i and q_j are the force field parameters and the constants ϵ_0 and ϵ_R are the vacuum permittivity and

the dielectric constant of the medium, respectively. Electrostatic interactions capture the attraction or repulsion between partial charges in molecules, which stem from an uneven electron distribution in a molecule due to polar bonds. Since the atomic charges are fixed at the position of the atoms, this force field cannot describe possible polarization effects. More sophisticated force fields allow for polarization, such as the Drude oscillator model,[83] where the atomic charge is linked to the atom via a harmonic potential. See Ref [84] for current developments in polarizable force fields.

The *dihedral cross terms* are a correction applied to the potentials of the two dihedral angles in the protein backbone involving the C_α carbon.[85,86] Using a strictly additive approach to calculate the energy of the protein backbone has been shown to result in systematic deviations in the prediction of the secondary structure. Therefore, an empirical grid-based potential is used in the `CHARMM36` force field, called cross-term map (CMAP), which couples the two dihedral angles in peptide bonds. The need for these cross terms in an otherwise additive force field reveals the limits of the force field approach, since it relies on the strict separability of molecular interactions.

Generally, the `CHARMM` parameters are optimized in an iterative scheme involving QM calculations and fitting against structural data (see Ref [87] for a concise outline of the procedure). It is important to understand that force field parameters are actually not defined for atoms, as for the sake of simplicity suggested above, but for atom types. The carbon in an alkane chain has a different atom type and therefore different parameters compared to an aromatic carbon, since they exhibit a different structural flexibility. However, it is not always obvious when atom types can be transferred from one molecule to another. For example, the atom types in DNA are not transferable to RNA, even though they differ by merely a hydroxyl group, because the parameters are biased towards the DNA structure.[88]

QM-MM interactions

The interaction between the QM and the MM region is the central point of the QM/MM approach, but it is not straightforward how it is best treated. The QM/MM models in the present work employ electrostatic embedding, which means that the QM region is polarized by the MM charges. Therefore, the MM charges appear explicitly as additional terms in the QM Hamiltonian, interacting with the electrons and the nuclear charges of the QM region. A more simplistic approach is the mechanical embedding, where the electrostatic interaction between the MM and the QM region is based only on the MM parameters defined for both atoms, the electron density in the QM region does not "see" any MM charges.[16] A more sophisticated approach involves bidirectional polarization of the QM and the MM region. This requires (i) the definition of charges in the QM region and (ii) the ability of the MM region to respond to the electrostatic influence, *i.e.* a polarizable force field.

Whenever the QM-MM interface cuts through a covalent bond, say between atoms A_{QM} and A_{MM}, a hydrogen atom is added to the QM region (see Ref [16] for other approaches to handle the QM-MM interface). This link atom is placed along the cut bond, and thus recreates the covalent bond for the QM region (A_{QM}−H). The position of the link atom is constrained by the orientation of the cut bond and a predefined scaling factor along that bond and the additional atom does therefore not

increase the degrees of freedom of the total system. Since the atom A_{MM} is fairly close to the link atom, the MM charge of A_{MM} may lead to an over-polarization of the electron density in the QM region. The MM charge is therefore removed from A_{MM} and distributed among the MM atoms bound to A_{MM} (`coupling=shift` in `chemshell`). The interaction between A_{QM} and all other atoms in the MM region still contributes to the MM energy expression. However, all interactions extending beyond A_{QM} into the QM region, including angles and dihedral involving the link atom, are disregarded in the MM energy expression, since they are already part of the QM energy.

The final interaction present between the QM and the MM region is the non-covalent van-der-Waals interactions. The calculation is straightforward and simply requires the definition of van-der-Waals parameters for all atoms in QM region as well.

QM/MM energy expression

There is no unique definition of the total energy of a QM/MM system,[16] but some definition is required, which can then be used to optimize the geometry. `chemshell` uses the additive scheme, where the total energy is given by

$$E_{\text{tot}}^{\text{add}} = E_{\text{MM}} + E_{\text{QM}} + E_{\text{QM-MM}}$$

where E_{MM} is the MM energy of the MM region, E_{QM} is the QM energy of the QM region including link atoms, and $E_{\text{QM-MM}}$ is the interaction energy between the QM and the MM region. The last term collects (i) the polarization of the QM electron density due to the MM charges including the shifted charges at the link atoms, (ii) the energy of the MM bonds crossing the QM/MM interface, and (iii) the non-bonded interaction between the QM and the MM atoms. Therefore, the atoms of the QM region only require non-bonded parameters plus bond parameters for the QM atoms directly at the QM/MM interface. The total QM/MM energy reported in the `chemshell` output file is split into a QM part (`ORCA`) and an MM part (`dl_poly`). The printed QM energy includes contribution (i), the printed MM energy includes contributions (ii) and (iii). Alternatively, the subtractive scheme for the total QM/MM energy does not include the coupling terms (i)-(iii), but it requires reliable MM parameters for the entire system including the QM region, which are particularly difficult to acquire for molecules with a complex electronic structure, such as FeS clusters.

1.1.4. Spin Hamiltonians

Spin Hamiltonians are phenomenological Hamiltonians that aim to describe magnetic interactions with minimal complexity.[89,90] Their success rests on the insight that characteristic patterns in certain spectra can be recreated with the energies in system expanded purely of spin functions. The interactions in this fictitious spin system are described by the spin Hamiltonian. The features observed experimentally are then

fitted to the parameters of the spin Hamiltonian.† Spin Hamiltonians can be used purely phenomenologically, that means to simply report a complex experimental spectrum in a concise way. In this case, no connection to the underlying molecular structure is made; the fictitious spin system exists only in spin space. However, spin Hamiltonian parameters account for relativistic effects, which are related to the underlying electronic structure. Quantitative predictions of the parameters can be achieved by constructing an effective Hamiltonian.[89,91] In a nutshell, in effective Hamiltonian theory a model space is defined that contains only a few low-energy states. Instead of diagonalizing the full electronic Hamiltonian matrix, the interaction between the low-energy and the higher-energy states is considered via, for example, second-order perturbation theory. The size of the model space is chosen to be equal to the spin Hamiltonian matrix. Therefore, the effective Hamiltonian spectrum can then be mapped onto the spin Hamiltonian spectrum. This allows for a decomposition of the spin Hamiltonian parameters into specific contributions in the wave function (see Ref [92] for a concise overview).

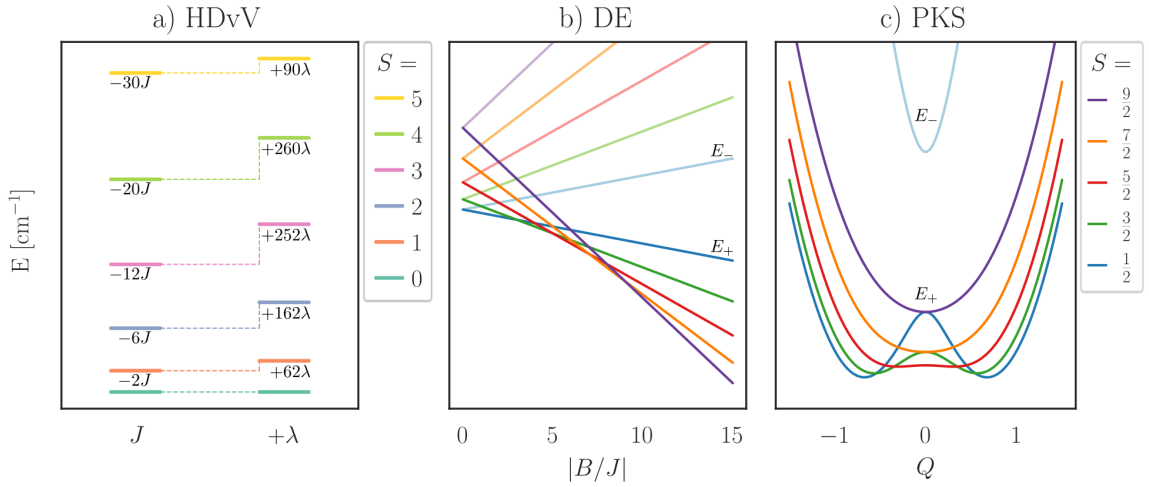

Figure 1.3. – Spin state energies of spin Hamiltonians often used for $[Fe_2S_2]^{2+,1+}$ clusters. (a) Heisenberg-Dirac-van-Vleck (HDvV) Hamiltonian for the homo-valent $[Fe_2S_2]^{2+}$ and the effect of the biquadratic term ($J = -200\,cm^{-1}$, $\lambda = +3\,cm^{-1}$, Eq 1.8 and Eq 1.10). (b) double exchange (DE) Hamiltonian in the symmetric mixed-valent $[Fe_2S_2]^{1+}$ (Eq 1.11). (c) Piepho-Krausz-Schatz (PKS) Hamiltonian, which considers the distortion of the symmetric structure in $[Fe_2S_2]^{1+}$ along the vibrational coordinate Q ($J = -200\,cm^{-1}$, $B = 1000\,cm^{-1}$, $\frac{\lambda^2}{k} = 8000\,cm^{-1}$, Eq 1.12).

†For example, the methyl radical $\cdot CH_3$ has a single unpaired electron. The EPR spectrum shows a typical $S = \frac{1}{2}$ signal with a hyperfine pattern due to the $I = \frac{1}{2}$ hydrogen nuclei. A fictitious spin system with an equivalent multiplet structure can be constructed from four abstract $S = \frac{1}{2}$ spin functions: one unique function, corresponding to the unpaired electron, and three degenerate ones, corresponding to the hydrogen nuclei. The interaction strength in the fictitious spin system is parameterized by the spin Hamiltonian and its eigenvalues can be therefore fitted to the actual spectrum. In the EPR spectrum of $\cdot CH_3$, the coupling between the electron spin and the nuclear spins is can be quantified by extracting the hyperfine tensor A.

Exchange Interactions

Magnetic centers in metal clusters can interact with each other either through space (dipole-dipole) or through chemical bonds (exchange).[90] Contributions to the exchange interaction are, for example, direct exchange, which arises from metal-metal bonding, or superexchange, which is relayed by intermediate atoms. Exchange interactions between two magnetic centers are most often described with the Heisenberg-Dirac-van-Vleck (HDvV) Hamiltonian[93–95]

$$\hat{H}^{\mathrm{HDvV}} = -2J\hat{S}_A\hat{S}_B \tag{1.8}$$

where $\hat{S}_A$ and $\hat{S}_B$ are the respective local spins and J is the coupling constant quantifying the Heisenberg exchange interaction. The spacing between the eigenstates of the HDvV Hamiltonian is defined by the Landé interval rule

$$\Delta E_{S-1,S} = 2JS \tag{1.9}$$

If $J > 0$, the two spins are parallel in the ground state, if $J < 0$ antiparallel. In Eq 1.8 only isotropic exchange is considered; asymmetric and anisotropic contributions are ignored.[96,97] Figure 1.3a shows the so-called Heisenberg ladder for a typical $[Fe_2S_2]^{2+}$ cluster.[98] The two Fe^{3+} centers in have local $S_{\mathrm{loc}} = \frac{5}{2}$ high spins. Since the coupling constant is on the order of $-200\,\mathrm{cm}^{-1}$ and the two magnetic centers coupled strongly to an $S = 0$ ground state.

The HDvV Hamiltonian can be extended to included higher-order contributions

$$\hat{H}^{\mathrm{biq}} = -2J\hat{S}_A\hat{S}_B - 4\lambda(\hat{S}_A\hat{S}_B)^2 \tag{1.10}$$

where λ is the parameter for the biquadratic term.[99] The effect of $\lambda \neq 0$ on the Heisenberg ladder is shown in Figure 1.3a. This leads to a deviation from the Landé interval rule. However, other contributions may also influence the energies of the Heisenberg ladder (*vide infra*). Experimentally, the coupling constant J is usually determined through the temperature dependence of magnetic susceptibility.[100] Due to the rather strong coupling in $[Fe_2S_2]^{2+}$ clusters, only the $S = 0, 1$ states are significantly populated at typical temperature ranges (up to 300 K). Therefore, only a single parameter can be extracted and other contributions, such as the biquadratic term, are folded into an effective J parameter.

Double exchange mechanism

Assuming a symmetric mixed-valent $[Fe_2S_2]^{1+}$ dimer, the two valence isomers $Fe_A^{2+}Fe_B^{3+}$ and $Fe_A^{3+}Fe_B^{2+}$ are degenerate. Zener proposed that the two valence isomers in a mixed-valent dimer may be in resonance.[101,102] In this double exchange (DE) mechanism (also called spin-dependent delocalization), the two metal centers are ferromagnetically aligned, which facilitates the travel of a minority spin electron between the two centers, since it does not need to flip its spin. For $[Fe_2S_2]^{1+}$, this would stabilize the $S = \frac{9}{2}$ state, where the electron would be valence-delocalized in an $Fe^{2.5+}Fe^{2.5+}$ pair. The energies in the spin Hamiltonian including DE are

$$E_{\pm}^{\mathrm{DE}} = -JS(S+1) \pm B(S+\tfrac{1}{2}) \tag{1.11}$$

where the DE parameter B is proportional to the transfer integral between the two orbitals holding the delocalized electron. The HDvV Hamiltonian, on the other hand, only applies to systems with localized spins. Note that the Landeé interval rule does not apply for the energies in Eq 1.11.

The influence of B on the spin states in a mixed-valent $[Fe_2S_2]^{1+}$ dimer is shown in Figure 1.3b. Without DE, *i.e.* for $B = 0$, the spin states form a typical Heisenberg spin ladder with an $S = \frac{1}{2}$ ground state. The electronic structure is localized and interconversion between the two degenerate valence isomers requires overcoming an energy barrier (Robin-Day class 2).[103] When DE dominates, *i.e.* at a ratio $|\frac{B}{J}| > 10$, the $S = \frac{9}{2}$ high spin state is the ground state. The electronic structure is delocalized, the Fe centers are indistinguishable (Robin-Day class 3). In an intermediate regime, *i.e.* $|\frac{B}{J}|$ 5, an intermediate spin state is lowest in energy. Therefore, Heisenberg exchange and double exchange are competing mechanisms, one favoring antiferromagnetic (in the case of $J < 0$), the other ferromagnetic coupling.

Vibronic coupling

The DE energy expression in Eq 1.11 is only valid if center A and B are identical.[104] An asymmetric vibrational mode involving the bridging ligands, for example, lifts the degeneracy and favors electron localization on one of the centers. Piepho-Krausz-Schatz (PKS) derived an energy expression including the vibrational coordinate x[104,105]

$$E_{\pm}^{\mathrm{PKS}}(S) = -JS(S+1) + \frac{1}{2}(\frac{\lambda^2}{k})Q^2 \pm \sqrt{\frac{1}{2}(\frac{\lambda^2}{k})^2Q^2 + B^2(S+\tfrac{1}{2})^2} \tag{1.12}$$

where λ is the vibronic coupling† term (not to be confused with the biquadratic term in Eq 1.10), k is the force constant of the harmonic vibration potential, and Q is the vibrational coordinate. Figure 1.3c illustrates the interplay between Heisenberg exchange, double exchange, and vibronic coupling. For a symmetric geometry, the $S = \frac{5}{2}$ state is lowest in energy ($Q = 0$, $|\frac{B}{J}| = 6$). The vibrational mode, however, lowers the energies of the lower spin states and the $S = \frac{1}{2}$ state becomes the ground state. Vibronic coupling therefore leads to an increase in antiferromagnetic coupling, in line with a localized electronic structure at the global minima.

In the case that vibronic coupling is much stronger than DE ($\lambda >> B$ in Eq 1.12), the vibrational coordinate Q_{min} is the same for all $E_{-}^{\mathrm{PKS}}(S)$.[104,107] The energies of these states at Q_{min} is given by

$$E_{-}^{\mathrm{loc}} = -JS(S+1) - \frac{B^2}{2\lambda}S(S+1) = -J^{\mathrm{eff}}S(S+1) \tag{1.13}$$

†The term vibronic is a composition of vibrational and electronic. Strictly speaking, vibronic coupling implies the coupling of nuclear and electron motion and therefore the breakdown of the Born-Oppenheimer (BO) approximation. However, in the present context of spin Hamiltonians, the literature widely uses the term vibronic coupling to refer to the effect of a geometric distortion on the spin state energies. Here, the BO approximation still holds. For a clarifying discussion of the term vibronic coupling the reader is referred to Ref [106].

which only depend on J, B, and λ and follow the Landé interval rule. The value of the effective coupling constant J^{eff} is now determined by the antiferromagnetic contributions from Heisenberg exchange (for $J < 0$) and vibronic coupling as well as the ferromagnetic contribution from DE.

The electronic structure of the mixed-valent $[Fe_2S_2]^{1+}$ dimers in Section 2.3.5 is localized with an $S = \frac{1}{2}$ ground state. Therefore, Eq 1.13 should be used to quantify the spin coupling in the ground state. As will be in more detail there, the minority electron on the high spin Fe^{2+} can also be excited locally in *d*-*d*-like transitions. The excitation energies are comparable to the spin-coupled states, but the spin Hamiltonians discussed so far do not consider any electronically excited states. As a consequence, Eq 1.13 should be extended to include electronically excited states

$$E_i^{\text{multi}} = -J_i^{\text{eff}} S(S+1) + \Delta_i \tag{1.14}$$

where Δ_i is the excitation energy and J_i^{eff} the state-specific coupling constant. The coupling constant may be different for each state, since the orientation of the excited electron has an effect on the metal-ligand overlap as well as on the transfer integral. This multi-state Hamiltonian has the correct number of states in the low-energy spectrum of mixed-valent $[Fe_2S_2]^{1+}$ clusters.[108]

Calculating the coupling constant J

The CASSCF wave function is an eigenfunction of the $\hat{S}^2$ operator, just like the exact non-relativistic electronic wave function. Furthermore, it has the flexibility to describe multiconfigurational states, such as the open-shell ground state singlet in the antiferromagnetically coupled homo-valent $[Fe_2S_2]^{2+}$ clusters. Therefore, the spin Hamiltonian parameters, can be extracted directly from the state energies. However, the single determinant approach, such as HF or Kohn-Sham DFT, cannot describe a multiconfigurational open-shell singlet state directly.[39]

In practice, a broken-symmetry (BS) determinant is constructed to access an open-shell singlet: the spin on center A is predominantly α, the spin on center B predominantly β; the spin symmetry is broken, because for a pure singlet, the spin density is zero everywhere in space. The BS determinant can be expressed as a mixture of different spin eigenfunctions (ignoring spin contamination for now).[109] In case of a small overlap between the magnetic orbitals on center A and B, the coupling constant can then be extracted from the energies of the high spin and the BS determinant via

$$J = -\frac{E_{\text{HS}} - E_{\text{BS}}}{S_{\text{max}}(S_{\text{max}}+1)}$$

where $S_{\text{max}} = S_A + S_B$.[109]

Spin contamination is the deviation of the spin expectation value $\langle \hat{S}^2 \rangle$ in an unrestricted determinant compared to the corresponding restricted determinant.[26] The deviation arises because unrestricted determinants are usually not spin eigenfunctions, but rather linear combinations of different spin states. For the unrestricted high spin determinant, spin contamination increases $\langle \hat{S}^2 \rangle_{\text{HS}}$ through spin polarization. For the unrestricted BS determinant, spin polarization increases $\langle \hat{S}^2 \rangle_{\text{BS}}$, but the mixing of

ionic CFGs, which are responsible for the magnetic coupling, decreases $\langle \hat{S}^2 \rangle_{\mathrm{BS}}$.[109] In the formalism proposed by Yamaguchi *et al.*,[110,111] which is used throughout this work, the coupling constant in the unrestricted case is given by

$$J = -\frac{E_{\mathrm{HS}} - E_{\mathrm{BS}}}{\left\langle \hat{S}^2{}_{\mathrm{HS}} \right\rangle - \left\langle \hat{S}^2{}_{\mathrm{BS}} \right\rangle} \tag{1.15}$$

For a comparison between the coupling constants extracted with wave function-based methods and BS-DFT, the reader is referred to Ref [112].

Calculating a BS determinant is straightforward in `ORCA`: Within the input file for a high spin determinant, the to-be-flipped atoms and the final M_S are specified: for example, a calculation with `mult 8` in the `%coords` block and `flipspin 1,4 finalms 1.5` in the `%scf` block (i) converges the $M_S = \frac{7}{2}$ high spin determinant, (ii) flips the spin on atoms 1 and 4, on which in this example one electron is localized, respectively, and (iii) reconverges the determinant with an $M_S = \frac{3}{2}$. Step (ii) uses the atomic basis functions to determine the affiliation with an atom and therefore may lead to unwanted results if the basis set contains too many delocalized atomic basis functions. Alternatively, the guess for the BS determinant can be generated by requesting the number of unpaired spin-up and spin-down electrons (`brokensym`). Here, localized orbitals are used to determine the to-be-flipped atom and the basis set dependency can be avoided. However, the approach is limited to flipping a single spin center and depends on the localization procedure, which can drastically change the shape of the magnetic orbitals in certain cases (see for example Figure 4.7).

1.1.5. Population analysis

The results of a wave function-based or DFT calculation yields the distribution of electrons within the molecule, *i.e.* the electron density. The electron density is closely related to the charge and integration over all space leads to the number of electrons in the molecule:

$$N_{el} = \int \rho(r)\,\mathrm{d}r \tag{1.16}$$

where $\rho(r)$ is the electron density at the position r. The idea behind population analysis is to divide a quantity, such as the electron density, among the atoms in the molecule. Integrating all the parts of the electron density associated with a particular atom leads to the atomic charge. In the unrestricted case, the integration of the spin density $\rho_S(r)$ leads to the excess number of α electrons

$$N_{el}^{\alpha} - N_{el}^{\beta} = \int \rho^{\alpha}(r)\,\mathrm{d}r - \int \rho^{\beta}(r)\,\mathrm{d}r = \int \rho_S(r)\,\mathrm{d}r$$

and dividing the spin density among the atoms leads to spin populations for each nucleus.

Attributing parts of the electronic structure to particular nuclei is a non-trivial task, because there is no obvious definition what an atom is within a molecule, and several fundamentally different approaches exist. In the following, several approaches will

be outlined for the atomic charges. Some approaches analyze a vector in the Hilbert space, the N_{el}-variate wave function Φ (or DFT determinant), which is related to the electron density by integrating over $N_{el} - 1$ electron coordinates

$$\int \Phi^*(r_1, r_2, \dots r_{N_{el}})\Phi(r_1, r_2, \dots r_{N_{el}})\,\mathrm{d}r_2 \dots \mathrm{d}N_{el} = \rho(r_1) = \rho(r)$$

where r_i indicates the position of the i^{th} electron. The most popular of these approaches were developed by Mulliken[113] and Loewdin,[114] and they further require an atom-centered basis to describe Φ. The other approaches analyze the physical, *i.e.* three-dimensional, space. Examples are the Voronoi[115] and Bader[116] charges, which use hard, discrete boundaries to separate atoms or Hirshfeld[117] charges, which use soft, fuzzy boundaries between atoms. While many authors refer to the approach developed by Bader as the theory of atoms in molecules (AIM), some authors use this term for any of the approaches listed above, because they attribute parts of a molecular property to atoms within that molecule.

Mulliken population

The Mulliken[113] partitioning can be used whenever a set of atom-centered basis functions $\{\varphi\}$ are used to expand the wave function or DFT determinant. For a single determinant ansatz, the electron density can be expressed in terms of molecular orbitals Φ and Eq 1.16 becomes

$$N_{el} = \int \left[\sum_i \Phi_i^*(r)\Phi_i(r)\right] \mathrm{d}r$$

where i runs over all occupied orbitals and Φ^* indicates the complex conjugate. Writing the molecular orbitals, in turn, as linear combinations of a set of basis functions $\{\varphi\}$ the above expression can be rewritten as

$$\begin{aligned} N_{el} &= \sum_{\mu\nu} \left[\sum_i c_{i\mu}^* c_{i\nu} \int \varphi_\mu^*(r)\varphi_\nu(r)\,\mathrm{d}r\right] \\ &= \sum_{\mu\nu} D_{\mu\nu} S_{\mu\nu} \\ &= \sum_{\mu\nu} (\mathbf{D} \odot \mathbf{S})_{\mu\nu} \end{aligned}$$

where μ, ν run over the entire set $\{\varphi\}$ and $\odot$ indicates the Hadamard (element-wise) matrix product. The elements of the density matrix $\mathbf{D}$ are defined as $D_{\mu\nu} = \sum_i c_{i\mu}^* c_{i\nu}$ and those of the overlap matrix $\mathbf{S}$ as $S_{\mu\nu} = \int \varphi_\mu^*(r)\varphi_\nu(r)\,\mathrm{d}r$. Since each φ is centered at an atom A, the matrices $\mathbf{D}$ and $\mathbf{S}$ as well as $\mathbf{D} \odot \mathbf{S}$ can be written in a block structure

$$\mathbf{D} \odot \mathbf{S} = \left[\begin{array}{c|c|c} AA & AB & \cdots \\ \hline BA & BB & \cdots \\ \hline \cdots & \cdots & \cdots \end{array}\right]$$

In the Mulliken partitioning scheme, the contributions to each atom consist of the diagonal block and $\frac{1}{2}$ of the off-diagonal blocks of the respective atoms. Adding the

charge of the nucleus Z_A, the Mulliken charge is defined as

$$q_A^{\mathrm{Mul}} = Z_A - \left[\sum_{\mu,\nu \in A} (\mathbf{D} \odot \mathbf{S})_{\mu\nu} + \frac{1}{2} \sum_{\mu \in A, \nu \notin A} (\mathbf{D} \odot \mathbf{S})_{\mu\nu} \right] \tag{1.17}$$

Loewdin population

The Loewdin[114] partitioning avoids the arbitrary $\frac{1}{2}$ separation of the off-diagonal blocks by eliminating off-diagonal blocks altogether. This can be achieved by orthogonalizing the basis, which makes the overlap matrix and therefore any Hadamard product diagonal. Generally, the basis of matrix $\mathbf{A}$ can be transformed using a transformation matrix $\mathbf{X}$ via

$$\mathbf{X}^\dagger \mathbf{A} \mathbf{X} = \mathbf{A}'$$

where A' denotes the matrix in the new basis.[26] Multiple procedures exist with which an orthogonal atomic basis $\{\varphi'\}$ can be generated, but the Loewdin partitioning uses the so-called symmetric orthogonalization ($\mathbf{X} = \mathbf{S}^{-\frac{\mathbf{1}}{\mathbf{2}}}$). The density matrix in the transformed basis is $\mathbf{D}' = \mathbf{S}^{-\frac{\mathbf{1}}{\mathbf{2}}\dagger} \mathbf{D} \mathbf{S}^{-\frac{\mathbf{1}}{\mathbf{2}}}$ and the overlap matrix becomes the identity matrix $\mathbf{S}' = \mathbf{I}$. Since the Hadamard product of the density and the overlap matrices is diagonal, no off-diagonal blocks are present

$$\mathbf{D}' \odot \mathbf{I} = \left[\begin{array}{c|c|c} AA & 0 & 0 \\ \hline 0 & BB & 0 \\ \hline 0 & 0 & \cdots \end{array}\right]$$

Therefore, the Loewdin charge can be expressed as

$$q_A^{\mathrm{Loew}} = Z_A - \sum_{\mu \in A} (\mathbf{D}' \odot \mathbf{I})_{\mu\mu} \tag{1.18}$$

The Loewdin approach does not actually solve the problem of how to deal with overlap between different atoms, but it simply outsources it to the orthogonalization procedure. The functions of the transformed basis $\{\varphi'\}$ are linear combinations of basis functions located on different atoms and there is no physical supremacy of the symmetric orthogonalization.

Voronoi cells

The simplest concept to partition the physical space in a molecule is to attribute each point in space to its closest atom.[115] This approach places each atom within an so-called Voronoi cell. Each cell is separated by discrete boundaries perpendicular to bonds from neighboring atoms, which is illustrated for guanine in Figure 1.4. The Voronoi cells are constructed only based on nuclear positions, not based on atom types. In analogy to Eq 1.17 the Voronoi atomic charge can be defined as

$$q_A^{\mathrm{Vor}} = Z_A - \int w_A^{\mathrm{Vor}}(r) \rho(r) \, \mathrm{d}r \tag{1.19}$$

where the discontinuous weight function $w_A^{\mathrm{Vor}}(r)$ is equal to 1 if A is the closest atom to r and otherwise 0. Because this simple approach treats all nuclei equally, an extension to the approach scales the Voronoi cells according to the van der Waals radii of the neutral atoms,[118] which results in similar charges as with the Hirshfeld method (*vide infra*).

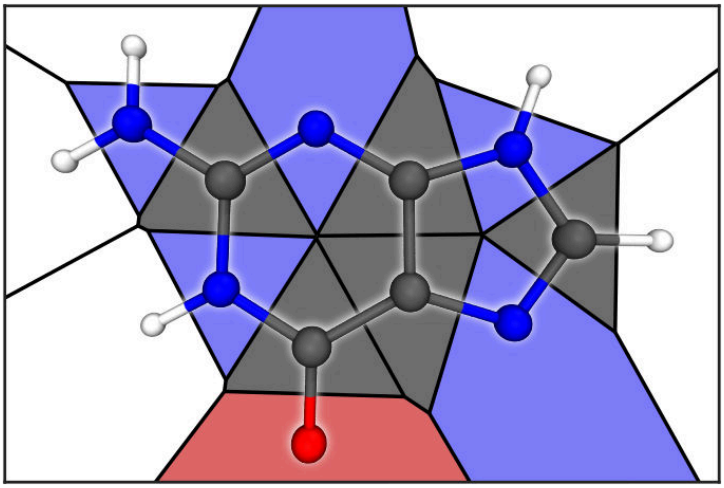

Figure 1.4. – 2D representation of Voronoi cells for the planar molecule guanine.

Bader volumes

The Bader[116] partitioning defines atomic volumes in a molecule based on the topology of the electron density.[119] If one follows the increase in electron density, one typically ends up extremely close to the position of a nucleus and, thus, each point in space can be assigned to a nucleus. More formally, the extreme values of the electron density satisfy the condition

$$\vec{\nabla}\rho(r) = 0$$

If the curvature at these points is negative, which can be analyzed by diagonalizing the Hessian matrix, the electron density is decreasing in all directions and the point is a maximum (also called attractor). Consequently, surfaces must exist between these maxima, where the electron density increases towards more than one maximum. This is equivalent to the requirement that the electron density flux through these surfaces vanishes, which can be expressed by

$$\vec{\nabla}\rho(r)\vec{n}(r) = 0$$

where $\vec{n}(r)$ is the surface normal. The volume confined by the zero-flux surfaces is the basis for the Bader partitioning of the electron density (also called basins). Examples for the volumes are given in Figure 1.5. In contrast to the Voronoi cells (see Figure 1.4) the surfaces are not planar and not necessarily halfway between to nuclei.

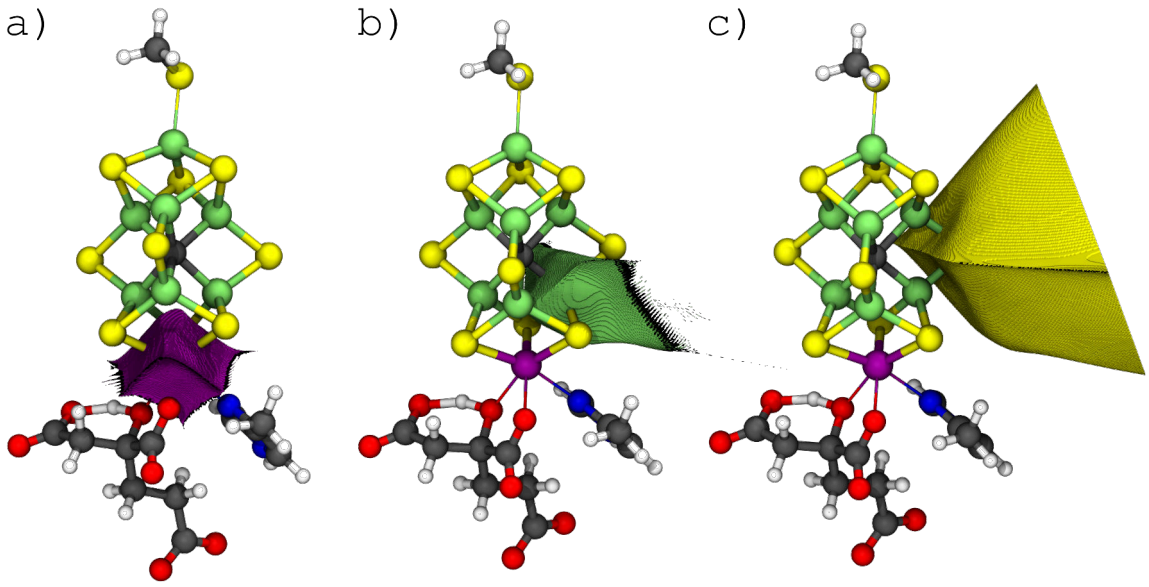

Figure 1.5. – Visualization of the Bader atomic volumes in FeMoco for (a) the Mo atom, (b) an Fe atom, and (c) a sulfur atom. The surface texture is due to the numerical grid used for the analysis of the electron density.

The Bader charges can be obtained via

$$q_A^{\text{Bader}} = Z_A - \int w_A^{\text{Bader}}(r)\rho(r)\,\mathrm{d}r$$

where w_A^{Bader} is 1 within the Bader volume associated with atom A and 0 elsewhere. Obviously, each electron density maximum needs to be mapped to a nucleus in order to account all electrons in the molecule. Even though fairly uncommon, electron densities have been calculated for symmetric and somewhat unusual molecules where a non-nuclear electron density maximum is stable.[120,121]

Hirshfeld partitioning

The atomic charges reported for the Hirshfeld[117] partitioning do not use the electron density directly, but measure the shift in electron density from a hypothetical, non-interacting system to actual system. The electron density of the non-interacting system, called the promolecule, is the sum of the free atom densities ρ_A^{free} located at each nucleus

$$\rho^{\text{pro}}(r) = \sum_A \rho_A^{\text{free}}(r - R_A)$$

The difference between the electron densities of the promolecule and the actual system is called the deformation density

$$\Delta\rho(r) = \rho(r) - \rho^{\text{pro}}(r)$$

which is a measure for the electrons being transferred during molecule formation.

Obtaining atomic charges from the deformation density requires the definition of an atomic domain and the Hirshfeld charge is defined as

$$q_A^{\mathrm{Hirsh}} = \int w_A^{\mathrm{Hirsh}}(r)\Delta\rho(r)\,\mathrm{d}r$$

where the weight function is closely related to promolecule density by

$$w_A^{\mathrm{Hirsh}}(r) = \frac{\rho_A^{\mathrm{free}}(r - R_A)}{\rho^{\mathrm{pro}}(r)}$$

Therefore, for Hirshfeld charges both the deformation density and the weight function depend on the definition of the promolecule. The charges obtained from the deformation density are not very sensitive to the definition of the weight function, and the using w_A^{Vor}, the discrete function used in the Voronoi partitioning has been shown to yield atomic charges similar to the Hirshfeld partitioning.[122] The atomic charges derived from the deformation density are typically smaller in magnitude than those derived from the direct integration of the electron density or analysis of atomic basis functions. Note that the Hirshfeld spin populations reported in this work are of similar magnitude to the other methods, because here the weight function $w_A^{\mathrm{Hirsh}}(r)$ is used to integrate the spin density directly.

In the iterative Hirshfeld (Hirshfeld-I)[123] scheme tries to account for polar bonds and the atom densities in the weight function $w_A^{\mathrm{Hirsh}}(r)$ are no longer those of the free atom ρ_A^{free}, but are expressed as linear combinations of differently charge states. The Charge model 5 (CM5)[124] approaches the problem that Hirshfeld charges are typically very small and do not recreate the molecular electrostatic potential. CM5 charges are the result of a parameterized mapping of Hirshfeld charges to recreate molecular dipole moments, making them applicable as QM-derived force field parameters.

Bond orders

Another metric that can be derived from the electronic structure and that depends on the definition of the atomic domain is the bond order, which reports the covalent bond strength between two atoms in a molecule. The most robust bond order that can be derived from the definition of the atomic basis set $\{\varphi\}$ is the Mayer[125] bond order, which is defined by

$$BO_{AB}^{\mathrm{Mayer}} = \sum_{\mu\in A,\nu\in B} (\mathbf{DS})_{\mu\nu}(\mathbf{DS})_{\nu\mu} \tag{1.20}$$

Note that here the sum runs over the elements of the matrix product **DS** and not the Hadamard product as in Eq 1.17 and Eq 1.18. As for the Mulliken and Loewdin charges, the Mayer bond order does not converge with an increasing basis set.

Using the discrete Bader volumes as atomic domains, the Bader bond order can be expressed analogous to Eq 1.20 as

$$BO_{AB}^{\mathrm{Bader}} = \sum_{\mu\nu} (\mathbf{DS}_A^{\mathrm{Bader}})_{\mu\nu}(\mathbf{DS}_B^{\mathrm{Bader}})_{\nu\mu} \tag{1.21}$$

where

$$S^{\text{Bader}}_{A,\mu\nu} = \int w^{\text{Bader}}_{A}(r)\varphi^{*}_{\mu}(r)\varphi_{\nu}(r)\,\mathrm{d}r$$

Here, the sum runs over all basis functions, because the overlap matrix is projected on the respective atomic domain ($\mathbf{S}^{\text{Bader}}_{A,B}$) using the weight function $w^{\text{Bader}}_{A,B}$. The Bader bond order is often referred to as the delocalization index (DI) and there is some debate whether the DI can be interpreted as a bond order (see Matito *et al.* [126] and references therein).

The so-called fuzzy bond order (FBO)[127] differs from the bond order definition in Eq 1.21 only in the weight function used to project the overlap matrix

$$S^{\text{FBO}}_{A,\mu\nu} = \int w^{\text{Hirsh}}_{A}(r)\varphi^{*}_{\mu}(r)\varphi_{\nu}(r)\,\mathrm{d}r$$

In this work, the fuzzy atomic domains in the FBO are defined by the weight function w^{Hirsh}_{A}, but this is by no means the only valid option. Matito *et al.* have discussed the connection between the Mayer bond order, the DI, and the FBO.[126]

1.1.6. Localized orbitals

The set of molecular orbitals $\{\Phi\}$ that result directly from an HF calculation are those that diagonalize the Fock matrix and are called canonical orbitals.[26] They are typically delocalized over the molecule, since they belong to an irreducible representation (irrep) of the molecule's point group. The set of molecular orbitals used to expand a wave function, however, is not unique and one obtain different sets via a unitary transformation of the molecular orbital basis without changing the overall wave function.[27] An infinite number of unitary basis transformations exists, but if the transformation is constrained to optimize a suitable orbital expectation value (*vide infra*) the resulting orbitals are more localized than the canonical ones. These localized orbitals paint a more intuitive picture of the electronic structure, since they often represent the electron pairs in the Lewis structure. Analogously, the canonical Kohn-Sham orbitals comprising the DFT determinant can be localized.

An example is given in Figure 1.6, which compares the canonical Kohn-Sham orbitals from a BS-DFT calculation of FeMoco, the active site of nitrogenase (see Section 4.2). In the charge state $MoFe_7S_9C^{1-}$, the total number of metal-based valence electrons is 41 assuming the oxidation states S^{2-} and C^{4-} for the ligands. Figure 1.6a shows those 41 canonical orbitals in which the sum over all metal contributions (*i.e.* Fe1 through Fe7 plus Mo) is the largest. Here, no structure is apparent. However, after localization, the 41 orbitals with the highest metal contributions are localized on at most two metal centers and clearly reveal the BS determinant. According to the localized orbitals, the oxidation states for the Fe centers are in between Fe^{2+} and Fe^{3+}, which is in line with the expectation for a tetrahedral Fe center in an FeS cluster.[98]

Localization algorithms

As mentioned above, the orbital transformation that results in localized orbitals is by no means unique.[27] The transformation proposed by Foster[128] yields a set of orbitals

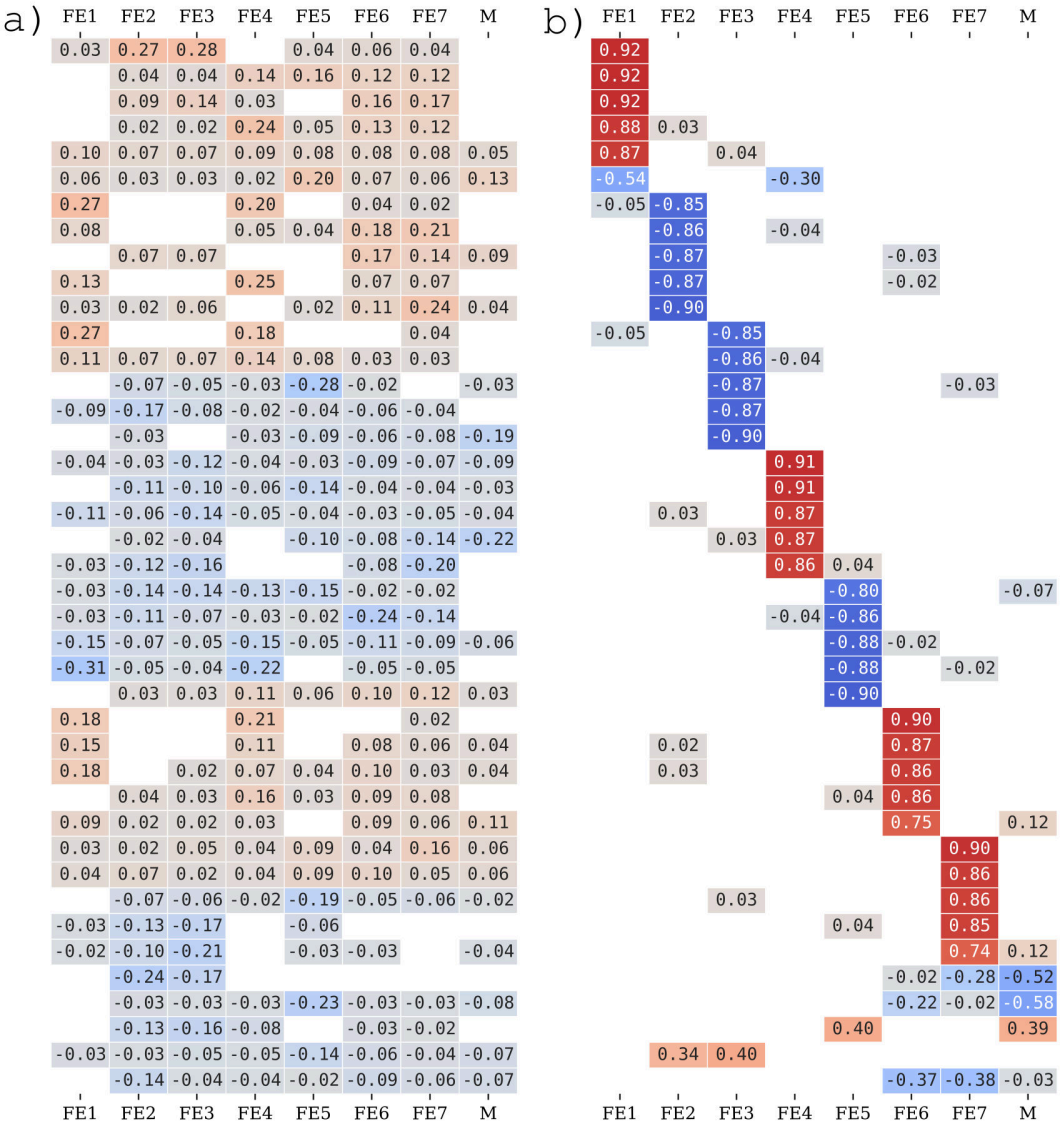

Figure 1.6. – Comparison between the Hirshfeld atomic contributions to the (a) canonical orbitals and (b) Foster-Boys localized orbitals in the resting state of FeMoco. Each row represents an orbital and the values show the respective metal contribution (α positive, β negative). In both cases, the 41 orbitals with the largest cumulated metal contribution are shown.

which are as compact as possible. This can be achieved by solving for the unitary transformation that minimizes the expectation value for the distance between two electrons in all orbitals

$$L_{FB} = \sum_i \int \Phi_i^*(r_1)\Phi_i^*(r_2)(r_1 - r_2)^2\Phi_i(r_1)\Phi_i(r_2)\,\mathrm{d}r_1\,\mathrm{d}r_2$$

Similar distance-based algorithms were proposed by Edmiston and Ruedenberg[129] and von Niessen[130].

Alternatively, Pipek and Mezey[131] proposed to simultaneously maximizes the atomic

charges for all orbitals. The corresponding unitary transformation maximizes the quantity

$$L_{PM} = \sum_A \sum_i |q_{A,i}^{\mathrm{Mul}}|$$

Here, $q_{A,i}^{Mul}$ is defined analogous to Eq 1.17 but for a single molecular orbital Φ_i. As discussed for the Mulliken charges, the Pipek-Mezey algorithm suffers from basis set dependence, but it has been shown that the localized orbitals obtained from maximizing the orbital-wise atomic charges do not significantly change with the definition of the atomic charges (*e.g.* those discussed above).[132]

An important difference between the Foster-Boys and the Pipek-Mezey algorithm is that the latter separates σ/π orbitals, *e.g.* the double bond of ethylene has one σ and one π orbital, while the first mixes them to three equivalent, banana-shaped bonding orbitals.[27] Maintaining σ/π separation appears tempting, since it agrees with the hybridization model of a covalent bond. However, it appears that Pipek-Mezey orbitals may lead to a more convoluted electronic structure compared to Foster-Boys for more complex systems, such as large FeS clusters (see Section 4.2.3).

1.2. Iron-sulfur (FeS) clusters

1.2.1. Biological FeS clusters

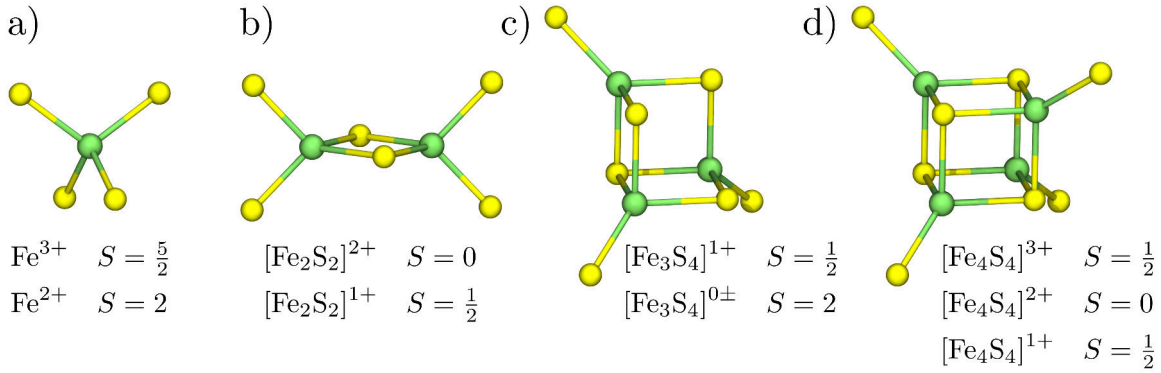

Figure 1.7. – Iron-sulfur (FeS) clusters and their most common charge and spin states encountered in biological systems: (a) monomeric $[FeS_4]$, (b) dimeric $[Fe_2S_2]$, (c) open cubanes $[Fe_3S_4]$, and (d) closed cubanes $[Fe_4S_4]$. Larger FeS clusters can be generated by fusing two cubanes, such as the P-cluster and FeMoco in nitrogenase (Figure 1.8).

Biological iron-sulfur (FeS) clusters perform a variety of functions such as electron transfer, DNA repair, Fe storage, and substrate activation.[98,133,134] This chemical flexibility results from a fortuitous combination of metal and ligand, which make a wide range of oxidation states and spin states accessible. An overview of common structural motifs is given in Figure 1.7. Much spectroscopic data is available for FeS clusters. However, the understanding of their electronic structure lags behind, mostly

because computational models that can describe the complex electronic wave function have only recently become applicable to larger FeS clusters.[21,108,135–138] The term *metal cluster* is used for molecules containing multiple metal centers, which exhibit some metal-metal bonding.[139] However, it in the context of biomolecules that contain an Fe−S bond, the term *FeS cluster* is used more loosely and often includes monomeric $[FeS_4]$ complexes, even though they contain only one metal center.

$[FeS_4]$ complexes

The simplest FeS complex contain a single Fe center coordinated by four S-based ligands (Figure 1.7a). A common example is the electron-transfer protein rubredoxin.[140] Here, the metal cycles between the two oxidation states Fe^{3+} and Fe^{2+}. As will be elaborated in Section 2.2, the tetrahedral coordination environment and the weak-field ligands favor high spin metal centers. The respective ground states therefore have the spins $S = \frac{5}{2}$ and $S = 2$. Prominent features in the electronic absorption spectra are ligand-to-metal charge transfer (LMCT) transitions at $> 10\,000\,\mathrm{cm}^{-1}$.[140] Here, an electron is transferred from a Cys−S^- ligand to the Fe center. In the Fe^{2+} redox state, additional *d-d* transitions are observed below $7000\,\mathrm{cm}^{-1}$.

Since the ligands in rubredoxin and similar monomeric complexes are usually cysteinates (Cys−S^-) and not inorganic sulfide (S^{2-}), some authors do not classify them as *FeS* complexes, but merely Fe complexes. However, this distinction will not be made in the present work, since cysteinate coordination is a common motif for higher-nuclearity complexes including S^{2-} as well. Therefore, the electronic structure in $[FeS_4]$ shares similarities with that of larges FeS complexes.

Valence-localized $[Fe_2S_2]$ clusters

In biological $[Fe_2S_2]$ clusters, the two Fe centers and the two S^{2-} ligands form a diamond core (Figure 1.7b). The capping ligands are either four Cys−S^-, as for example in ferredoxins, or two Cys−S^- on one Fe center and two His−N on the other, as in the Rieske protein.[98] For synthetic model compounds, the terminal ligands are sometimes replaced with β-ketiminato ligands ($NacNac^-$), which are rather soft for N-based ligands and therefore expected to mimic the ligand field of Cys−S^-.[141,142] In biological systems, they can cycle between the oxidized $[Fe_2S_2]^{2+}$ and the reduced $[Fe_2S_2]^{1+}$ form, for example, when acting as an electron transporter. As for the $[FeS_4]$ complexes, the Fe centers are in a tetrahedral coordination environment and the ligand field favors local high spin metal centers.

For the oxidized $[Fe_2S_2]^{2+}$ clusters, the diamagnetic $S = 0$ ground state is observed exclusively, which results from an antiferromagnetic coupling of the two local high spin Fe^{3+} centers ($S_{\mathrm{loc}} = \frac{5}{2}$). The two bridging S^{2-} ligands mediate strong antiferromagnetic coupling via the superexchange mechanism and the coupling constant J (Eq 1.8) is typically $-150\,\mathrm{cm}^{-1}$–$-200\,\mathrm{cm}^{-1}$.[143,144] Similar to the $[FeS_4]$ complexes, LMCT transitions are observed in an absorption spectrum above $10\,000\,\mathrm{cm}^{-1}$, which can originate either from the terminal Cys−S^- or the bridging S^{2-} ligands.[140] Absorptions below $1000\,\mathrm{cm}^{-1}$ correspond to vibrational modes of the $[Fe_2S_2]$ core.[144]

For the reduced $[Fe_2S_2]^{1+}$ clusters, a doublet $S = \frac{1}{2}$ ground state is almost exclusively

observed. The electronic structure is valence-localized, where the Fe_A^{2+} center ($S_{A,loc} = 2$) couples antiferromagnetically to the Fe_B^{3+} center ($S_{B,loc} = \frac{5}{2}$). Here, the coupling constant is on the order of $-100\,cm^{-1}$, consistent with a decreased Fe−S covalency in the reduced cluster.[143,144] Similar to the oxidized $[Fe_2S_2]^{2+}$ clusters, LMCT bands are also observed, but the inequivalence of the Fe centers is expected to broaden the LMCT features.[140] Transitions are also observed below $7000\,cm^{-1}$, which are assigned to local *d-d* transitions of the high spin Fe_A^{2+} center. Furthermore, an inter-valence charge transfer (IVCT) transition is possible, that leads to the inverted valence isomer $Fe_A^{3+}Fe_B^{2+}$.[104] For example, a band at $11\,000\,cm^{-1}$ has been proposed for this IVCT transition, but not been unambiguously assigned.[140] In principle, five IVCT transitions are possible, which would correspond to the local *d-d* excited states of the $Fe_A^{3+}Fe_B^{2+}$ valence isomer.

Valence delocalization in FeS clusters

A localized electronic structure can be favored by a vibrational mode of the FeS core, but also by an asymmetric electrostatic potential created by the environment. Most biological mixed-valent $[Fe_2S_2]^{1+}$ clusters are valence-localized and Heisenberg exchange therefore leads to an antiferromagnetically coupled $S = \frac{1}{2}$ ground state.[98,134] The double exchange mechanism, on the other hand, favors a valence-delocalized electronic structure and therefore the ferromagnetic alignment of metal centers. A ground state spin different from the typical $S = \frac{1}{2}$ has been observed in the $[Fe_2S_2]^{1+}$ ferredoxins of *Clostridium pasteurianum* (*Cp*) and *Aquifex aeolicus* (*Aae*).[24,145] While the respective wild type ferredoxins exhibit the usual $S = \frac{1}{2}$ spin, Cys→Ser variants[†] exist as a mixture of $S = \frac{9}{2}$ and $S = \frac{1}{2}$ spin states. These are the only known examples for $[Fe_2S_2]^{1+}$ cluster with a ground state different from a pure $S = \frac{1}{2}$ and their electronic structure is briefly explored in Chapter 3.

The $S = \frac{9}{2}$ spin state has been characterized by electron paramagnetic resonance (EPR) and variable-temperature magnetic circular dichroims (VTMCD) spectroscopy.[24,145] The valence-delocalized $Fe^{2.5+}Fe^{2.5+}$ electronic structure in state was supported by Mössbauer measurements (timescale $\propto 10^{-7}\,s$)[146] and resonance Raman spectra (timescale $\propto 10^{-12}\,s$).[147] The appearance of the $S = \frac{9}{2}$ spin state has further been shown to depend on the proton concentration: at pH ≤ 7 the $S = \frac{1}{2}$ species is exclusively detected, while at pH ≥ 10 the $S = \frac{9}{2}$ species is dominant ($> 80\,\%$).[24]

X-ray diffraction (XRD) structures of the wild type and the variants are only available for the ferredoxin from *Aae*, but the protein sequence in *Cp* is highly homologous.[148,149] The Cys→Ser substitution changes the coordination at one of the Fe centers from an S-based to an O-based terminal ligand. It was proposed that this counters an asymmetry in the electrostatic potential created by the protein environment and renders the two possible valence isomers, $Fe_A^{2+}Fe_B^{3+}$ and $Fe_A^{3+}Fe_B^{2+}$, accidentally degenerate. As a consequence, the electron delocalization would become feasible and double exchange would stabilize the high spin $S = \frac{9}{2}$ state (Eq 1.11). The pH-dependency further suggests that this delicate equilibrium can be easily disrupted through protonation/deprotonation in the proximity of the $[Fe_2S_2]$ cluster.

[†]C56S and C60S for *Cp*, C55S and C59S for *Aae*

All spectroscopic data for the $S = \frac{9}{2}$ state in the ferredoxin variants is consistent with a valence-delocalized $Fe^{2.5+}Fe^{2.5+}$ cluster. The characteristic EPR features are detectable between 4 K–60 K. According to the classification of mixed-valent metal dimers by Robin and Day,[103] this electronic structure corresponds to a class 3 dimer, where the delocalized electronic structure is the actual minimum of the potential energy surface. Several detailed studies on synthetic $[Fe_2S_2]^{1+}$ model compounds have shown that the Fe centers may be indistinguishable at higher temperatures of *e.g.* > 80 K, but become distinguishable at low temperatures.[141,142,150–153] For these compounds, only $S = \frac{1}{2}$ features have been observed in the EPR spectrum. Therefore, they full under the Robin-Day class 2, where each valence isomer corresponds to a separate minimum on the potential energy surface, but interconversion between the two happens easily.

Interestingly, substituting the bridging ligands in one specific synthetic $[Fe_2S_2]^{1+}$ model complex with Se or Te has been shown to stabilize an intermediate $S = \frac{3}{2}$ spin state.[23] This suggests that Heisenberg exchange, double exchange, and vibronic coupling are similarly important in these complexes and that the energy barrier for a transition from a valence-localized to a valence-delocalized is low. Valence-delocalized $Fe^{2.5+}Fe^{2.5+}$ pairs are also believed to be an essential aspect of the electronic structure of larger FeS clusters, such as open and closed FeS cubanes (Figure 1.7c and d)[98] and the active site of nitrogenase.[17,19] However, an isolated class 3 valence-delocalized $S = \frac{9}{2}$ Fe dimer has only been observed with three $\mu{-}OH^-$ bridging ligands.[154–156]

1.2.2. S→Se substitution in FeS clusters

The substitution of S with Se is a powerful tool in the study of FeS clusters. The two elements behave chemically similar, and the substitution constitutes only a small perturbation of the geometry and valence electronic structure. In other words, the S→Se substitution is assumed to be an *innocent* substitution. Intrinsic differences between S and Se can then be exploited spectroscopically, for examples, the different K-edge absorption energies or the distinct nuclear spins. The question to which extent the S→Se substitution is innocent, is dealt with of Chapter 2. Hereafter, experimental studies relevant in this context are reviewed.

Differences between S and Se

Some general properties of S and Se are compared in Table 1.1. Even though Se has 18 electrons more than S, the atomic as well as the dianionic radii of Se are only about 0.15 Å larger. This is a result of the *d* block contraction that happens in the fourth row of the periodic table.[159] The first ionization energy (IE) is 0.6 eV lower for Se, because the residual positive charge is distributed over a larger volume.[157] The about 10 keV higher K-edge energy in Se allows for a clear distinction between S and Se in X-ray absorption spectroscopy (XAS) experiments.[160] The only isotopes with a nuclear spin are ^{33}S ($I = \frac{3}{2}$) and ^{77}Se ($I = \frac{1}{2}$), which allows for a distinction between the hyperfine structures in EPR or nuclear magnetic resonance (NMR) spectra.[157] However, the natural abundance of these isotopes is low and enriching samples is a costly endeavor, especially in biological systems, where isotope loss is high and final concentrations are

Table 1.1. – Comparison between properties of the S and Se atoms and between the amino acids cysteine (Cys) and selenocysteine (Sec).

	S	Se
atomic number[157]	16	34
mass[157]	32.066 u	78.96 u
covalent radius[158]	1.05 Å	1.20 Å
dianionic radius[159]	1.84 Å	1.98 Å
ionization energy (IE)[157]	10.36 eV	9.75 eV
K-edge energy[160]	2.472 keV	12.658 keV
nuclear spin (abundance)[157]	$\frac{3}{2}$ (0.75 %)	$\frac{1}{2}$ (7.6 %)
SOC constant[157]	375 cm^{-1}	1540 cm^{-1}
pK_a(Cys/Sec)[161]	8.34	5.2
redox potential (Cys/Sec)[161]	−220 mV	−388 mV

low. The spin-orbit coupling (SOC) constant for Se is about four times higher than for S, which may lead to an increased mixing between spin states.[157]

Se is an essential trace element for mammals with an extremely narrow range between deficiency and toxicity.[162] In biological systems, it appears almost exclusively as the 21st amino acid selenocysteine (Sec) in enzymes such as coenzyme Q, glutathione peroxidase, and thioredoxin reductase.[163] Compared to cysteine (Cys), Sec is more easily deprotonated and a stronger reducing agent (see Table 1.1). The reason for the differences in p$K_a$is similar to the IE of the atoms: the residual negative charge after deprotonation can be stabilized more efficiently in Sec. The reduction potentials correlate with the stability of −S−S− and −Se−Se− bridges and the more favorable formation of −Se−Se bridges can be used to study protein folding mechanisms.[161] In line with the larger atomic radii, Se is more easily polarized than S and thus more susceptible to electrophilic or nucleophilic attacks, which makes Sec a powerful active handle in protein residue modification. These applications showcase that the reactivity of S and Se is indeed different, which may also reflect in the properties of Se-substituted FeS clusters.

Replacing S with Se in biological FeS clusters

Some biological FeS clusters may be removed from their protein under acidic conditions and spontaneously reassemble from a solvated Fe and S source.[164,165] Providing Se instead makes the corresponding FeSe cluster accessible, even though naturally occurring FeSe clusters have not yet been discovered. Orme-Johnson *et al.* used this approach already in 1968 to synthesize the Se-analog of the FeS protein putidaredoxin.[166] They selectively introduced the isotope ^{77}Se ($I = \frac{1}{2}$) and were able to resolve the hyperfine pattern in the $S = \frac{1}{2}$ EPR signal of the reduced $[Fe_2Se_2]^{1+}$ cluster, which had not been possible at that time using ^{33}S ($I = \frac{3}{2}$). The shape of the hyperfine pattern allowed them to deduce details about the topology long before an XRD structure at atomic resolution clarified the diamond core structure of $[Fe_2S_2]$ clusters.[167] This is one of the first examples of S→Se substitution in FeS clusters and illustrates its utility given

that the electronic structure remains qualitatively equivalent. An overview of studies characterizing FeSe analogs of FeS clusters is given in Ref [168].[143,169]

More recently, Se was introduced into FeMoco, the large FeS clusters forming the active site of the N_2-reducing enzyme nitrogenase, which will be discussed in Section 1.2.3.[170–172] Spatzal *et al.* replaced a single S with Se and followed the Se position during catalytic activity using spatially resolved anomalous dispersion (SpReAD).[170] Employing the same site-specific Se labeling, Henthorn *et al.* extracted accurate structural and electronic information about FeMoco through high-energy resolution fluorescence detected (HERFD) XAS measurements at the Se K-edge.[171] Their spectra suggest that the binding of the inhibitor molecule CO to FeMoco, which will be discussed in Section 1.2.4, triggers a redox redistribution between the metal centers, a change that is too small to be captured in typical XRD structures.

Spectroscopic differences between FeS and FeSe clusters

In the $S = \frac{1}{2}$ EPR signal of protein-bound and synthetic $[Fe_2S_2]^{1+}$ clusters, S→Se substitution has been shown to increase the isotropic g value as well as the anisotropy.†[151,166,173–175] At the same time, the Fe hyperfine coupling decreases by 3 MHz and the isomer shift by 0.01 mm/s according to Mössbauer and ^{57}Fe-electron nuclear double resonance (ENDOR) measurements.[176,177] The authors infer a higher Fe−Se covalency compared to Fe−S, which facilitates delocalization of the metal-based unpaired electrons.

The features in the infrared (IR) and UV/vis region typically appear at lower energies after S→Se substitution. In protein-bound $[Fe_2S_2]^{2+}$ clusters, UV/vis features between 20 000 cm^{-1} and 25 000 cm^{-1} are assigned to S→Fe LMCT transitions. In the corresponding $[Fe_2Se_2]^{2+}$ clusters, their energy is decreased by about 1000 cm^{-1}.[178,179] The supposedly LMCT bands are detected at lower energies in the reduced $[Fe_2S_2]^{1+}$ clusters, 17 500 cm^{-1}, but they undergo a similar red shift upon S→Se substitution.[175,179–181] Bands at 4500 cm^{-1} and 6000 cm^{-1} are assigned to *d*-*d* of the local high spin Fe^{2+} center in $[Fe_2S_2]^{1+}$, which appear at 4550 cm^{-1} and 5550 cm^{-1}, respectively, in $[Fe_2Se_2]^{1+}$.[140,173,175,179] Vibrational modes associated with the FeS core in protein-bound and synthetic $[Fe_2S_2]$ clusters are assigned to features from 200 cm^{-1}–500 cm^{-1} in resonance Raman spectra.[178,179,181,182] After S→Se substitution, they appear at lower energies, consistent with the higher atomic weight of Se,‡ but an unambiguous assignment is difficult.

The nature of the bridging ligand also affects the magnetic interactions between Fe centers. In protein-bound $[Fe_2S_2]$ ferredoxins, S→Se substitution has been shown to reduce the coupling constant J by about 10 % for both the oxidized $S = 0$ and the reduced $S = \frac{1}{2}$ form.[143] Similar trends have been reported for synthetic model compounds.[151,183] In a recent report, S→Se substitution has been shown to stabilize an intermediate spin state for synthetic mixed-valent clusters.[23] Here, the ground state

†For example, g = 2.02, 1.94, 1.89 $[Fe_2S_2]^{1+}$ and g = 2.04, 1.98, 1.95 in $[Fe_2Se_2]^{1+}$.[166] The values were converted from G using the formula $g = 0.714484\frac{\nu[\mathrm{MHz}]}{B[\mathrm{G}]}90$

‡In the harmonic oscillator approximation the vibrational frequency ω is proportional to $\frac{1}{\sqrt{\mu}}$, where μ is the reduced mass and $\sqrt{\frac{\mu_{\mathrm{FeSe}}}{\mu_{\mathrm{FeS}}}} = 1.27$

spin of the $[Fe_2S_2]^{1+}$ cluster is the typical $S = \frac{1}{2}$. In $[Fe_2Se_2]^{1+}$, both $S = \frac{1}{2}$ and the $S = \frac{3}{2}$ features are detectable in the EPR spectrum. Descending the group in the periodic table even further, the ground state of the S→Te substituted complex has a pure $S = \frac{3}{2}$ spin. The authors propose that the intermediate spin state is stabilized through a lower Heisenberg exchange, a lower vibronic coupling, and a higher double exchange in the Se- and Te-substituted clusters (see Eq 1.12). That S→Se has an effect on the ground state spin has also been observed for synthetic FeS cubanes, where the $S = \frac{3}{2}$ spin state has been shown to be stabilized over $S = \frac{1}{2}$ in the FeSe analogs.[184]

1.2.3. Nitrogenase

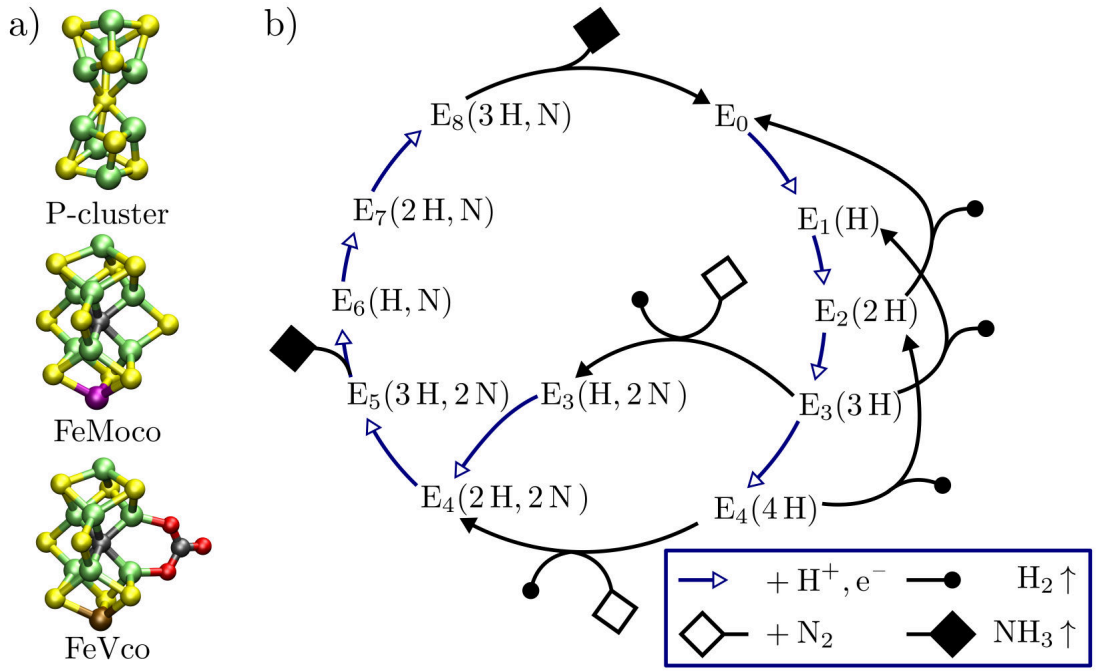

Figure 1.8. – (a) Structures of the P-cluster ($[Fe_8S_7]$), FeMoco ($[MoFe_7S_9C]$), and FeVco ($[VFe_7S_8CO_3C]$). (b) Lowe-Thorneley cycle for biocatalytic N_2 reduction in Mo nitrogenase.[14] The E_n state indicates the number of electrons transferred from the Fe protein to the MoFe protein, which are assumed to be accompanied by proton transfers.

Nitrogenases are responsible for reducing atmospheric N_2 to NH_3 and therefore complete the biological nitrogen cycle.[185] The most abundant and best studied variant is Mo nitrogenase, which consists of the Fe protein, a reductant containing an $[Fe_4S_4]$ cluster (Figure 1.7d), that transfers a single electron under ATP hydrolysis to the MoFe protein, where the reaction takes place. The MoFe protein contains two distinct FeS clusters, which are shown in Figure 1.8a. The P-cluster, $[Fe_8S_7]$, is the primary electron acceptor and relays the electrons to FeMoco, $[MoFe_7S_9C]$, the active site of nitrogenase. FeMoco has an unusual $\mu_6{-}C^{4-}$ carbide as the central atom, which is likely responsible for the unique reactivity of this enzyme.[69,186] The location of the FeS clusters within the MoFe protein is shown in Figure 1.2.

The stoichiometry of N_2 reduction, which has been established through kinetic measurements, is summarized in the Lowe-Thorneley cycle (Figure 1.8b).[14] The n in each E_n state indicates the number of reduction events, where an electron is transferred from the Fe protein to the MoFe protein. A proton transfer is assumed to accompany each reduction event. The complete chemical equation reads

$$N_2 + 8H^+ + 8e^- + 16\text{MgATP} \longrightarrow 2NH_3 + H_2 + 16\text{MgADP} + 16P_i$$

A total of eight electrons and protons are necessary to complete the cycle, since one equivalent of H_2 is produced for each conversion of N_2. H_2 evolution is not merely a byproduct, but appears to be an essential part of the mechanism: N_2 is believed to bind to the E_4 state, or possible the E_3 state, under reductive elimination of H_2.[187] Besides this, little is known about the mechanistic details due to the difficulty to trap a specific intermediate state for spectroscopic investigation (see Ref [15] for a recent review).

Alternative nitrogenases

Alternative nitrogenases contain V or only Fe instead of Mo.[188] The active sites are then called FeVco and FeFeco and the respective proteins VFe and FeFe protein. Mo, V, and Fe-only nitrogenase are believed to follow a similar N_2 reduction mechanism.[189] In FeVco (Figure 1.8a), one of the belt sulfides has been replaced with a carbonate ligand. V nitrogenase has received special attention because it was the first to catalyze Fischer-Tropsch-like chemistry, *i.e.* the reductive coupling of CO to hydrocarbons.[190] Later, similar reactivity was reported for Val70 variants of Mo nitrogenase[191] as well as for Fe-only nitrogenase.[192] In wild type Mo nitrogenase, CO merely acts as an inhibitor to N_2 reduction. V nitrogenase has also been reported to reduce CO_2 to CO.[193]

The FeMoco active site can be extracted from the MoFe protein into solution and later reconstituted into apo-MoFe protein, which completely restores the catalytic activity.[194,195] Reconstituting FeMoco in the apo-VFe protein instead yields the hybrid FeMoco@VFe nitrogenase.[196] Similarly, the hybrid FeVco@MoFe nitrogenase can be generated.[197] Interestingly, the hybrid nitrogenases are catalytically active, albeit with a lower turnover frequency compared to the wild-type nitrogenases. The g values of the characteristic $S = \frac{3}{2}$ EPR signals that correspond to the resting state FeXco (see below) are shifted in the hybrid nitrogenase, which shows that the MoFe and VFe proteins create a different environment for their respective active site.[197,198]

The electronic structure of the active site

In the resting state, the MoFe protein gives rise to a characteristic $S = \frac{3}{2}$ EPR signal ($g = [4.31, 3.77, 2.00]$). The signal arises from the paramagnetic FeMoco, while the P-cluster is EPR silent.[199] The charge state of FeMoco has been established as $[MoFe_7S_9C]^{1-}$ by a combination of Mössbauer measurements, SpReAD crystallography, and QM/MM modeling.[17,200,201] Assuming S^{2-} and C^{4-} oxidation states for the ligands, 41 valence electrons are left to be shared among the metal centers. According to BS-DFT calculations, the heterometal Mo^{3+} has an unusual non-Hund configuration, while the Fe centers have local high spins with oxidation states in between Fe^{2+} and

Fe^{3+}. The BS-DFT electronic structure of the resting state will be discussed in detail in Section 4.2.

It was long assumed that the resting state of the VFe protein is similar to the MoFe protein, since also features corresponding to an $S = \frac{3}{2}$ species were observed in the EPR under comparable conditions. However, the EPR spectra are not consistent for different methods of protein purification (see discussions in Ref [15] and [202]). In a recent study it has been suggested that the charge state of FeVco in the E_0 VFe protein is $[VFe_7S_8CO_3C]^{1-}$ and the valence electron count is therefore one less than for FeMoco in the E_0 MoFe protein.[202] However, when comparing different charge states in a VFe QM/MM model, $[VFe_7S_8CO_3C]^{2-}$ gives the best agreement with the VFe XRD structure.[19,203] In the computational models in the present study, E_0 refers to the charge states $[MoFe_7S_9C]^{1-}$ and $[VFe_7S_8CO_3C]^{2-}$, where both have 41 metal-based valence electrons, but it is important to keep in mind that the charge state of the E_0 VFe protein is still not well understood.

In the E_1 state, an additional electron and proton is added to FeMoco relative to E_0. In a combined extended X-ray absorption fine structure (EXAFS) and QM/MM study,[204,205] the electron has been found to reduce the Fe part of FeMoco, Mo is therefore not reduced. The proton most likely binds to one of the belt sulfides (S2B or S5A), which likely enhances the structural flexibility in the belt region.[204,206–209] FeMoco has an even number of electrons in E_1 state and is therefore EPR silent. Further reduction to the E_2 state yields two characteristic $S = \frac{3}{2}$ EPR signals.[210–213] Several authors have proposed hydride formation in the E_2 state based on DFT models,[206,208,213,214] which is consistent with the H_2 elimination from E_2.

1.2.4. Nitrogenase and CO

It has been known for over eighty years that the presence of carbon monoxide (CO) inhibits biological nitrogen fixation.[215] For wild type Mo nitrogenase, CO is a non-competitive inhibitor to N_2 reduction[216] and therefore either occupies a different binding site or binds to a different E_n redox state of the Lowe-Thorneley cycle (Figure 1.8b). However, in the alternative V and Fe nitrogenase as well as Val70 variants of Mo nitrogenase CO also acts as a substrate and can be coupled reductively to hydrocarbon chains.[191,192,217] See Ref [12] for a recent review on CO reduction by nitrogenase. The FeXco cofactor has been identified as a binding site,[218] but the mechanism of the binding event including the respective E_n states is unknown. In the following, those experimental studies are reviewed that characterizes the interaction of CO with nitrogenase. The focus lies on the initial binding event and Figure 1.9 gives an overview of the experimental species of CO-inhibited wild type Mo nitrogenase that are relevant to this work.

CO binding in the Lowe-Thorneley cycle

It is generally believed that CO does not interact with the resting state (E_0) of Mo nitrogenase, since the typical $S = \frac{3}{2}$ EPR signal is not affected by the presence of CO (Figure 1.9a). Rather, the observation of CO-bound species requires turnover conditions, *i.e.* a supply of electrons and protons to the cofactor, and binding is

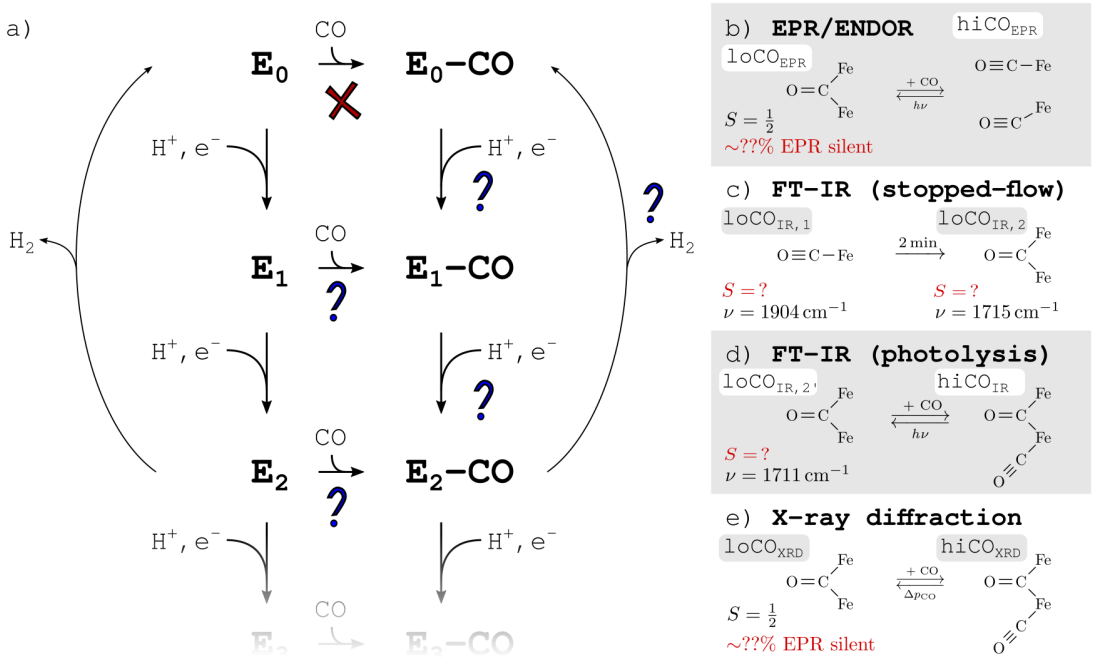

Figure 1.9. – Overview of CO inhibition in Mo nitrogenase. (a) Possible CO binding to the early redox states in the Lowe-Thorneley cycle (see Figure 1.8b). (b-e) Proposed binding modes in the experimental species that capture the initial CO binding event: EPR/ENDOR spectroscopy (b), SF-FT-IR spectroscopy (c) and (d), XRD structures (e). Lo and hi refer to the species under low and high CO partial pressure. Figure adapted from Ref [25].

discussed in the literature for the E_1 or E_2 state.[15] Interestingly, it has been reported for V nitrogenase that CO binds without turnover conditions, but merely in the presence of a strong reductant.[219,220] However, some controversy exists about this observation,[221] which is further complicated by the ambiguity of the resting state for V nitrogenase.[202]

EPR species

Under enzymatic turnover, two $S = \frac{1}{2}$ EPR signals arise that are characteristic for CO-bound FeMoco: loCO_{EPR} with $g = [2.09,\ 1.97,\ 1.93]$ and hiCO_{EPR} with $g = [2.17,\ 2.08, 2.05]$ (labeled lo-CO and hi-CO in the original work, Figure 1.9b).[222] The signals arise under low and high CO partial pressures, respectively. FeMoco has been confirmed as the binding site for both species through ^{57}Fe ENDOR measurements.[218] Furthermore, ^{13}C ENDOR measurements, which yield the hyperfine coupling tensor at the CO carbon, have led to the proposition that for loCO_{CO} a single bridging CO is bound and for hiCO_{CO} two terminal CO.[217] Both species can be interconverted through photolysis/annealing cycles, where the activation energy has been estimated at 1 kcal/mol.[223] Therefore, loCO_{CO} and hiCO_{CO} most likely correspond to the same E_n state. The detection of integer spin species is much more difficult compared to half-integer spin species in standard EPR measurements,[90] which is why the E_n states with an odd n are usually considered EPR silent.

It has been reported that for V nitrogenase the presence of CO and an Eu^{2+}-based reductant ($E^0 = -1.14\,\text{eV}$ at pH = 8) give rise to $S = \frac{1}{2}$ EPR signals without turnover conditions.[219,220] With g = [2.09, 1.99 1.91] and g = [2.13, 2.01, 1.97] at low and high CO partial pressure, respectively, these signals are highly similar to the Mo nitrogenase species loCO_{EPR} and hiCO_{EPR}. A different study states that CO binding to V nitrogenase requires turnover conditions.[221] EPR signals identical to loCO_{CO} and hiCO_{CO} have been measured for His195→Gln variants of Mo nitrogenase.[224] This study is the only one that reports spin quantitation of the EPR signals and it demonstrates that the CO-bound EPR species make up merely 10 %–25 % of the total enzyme. With the substrate-free $S = \frac{3}{2}$ E_0 signal constituting about 10 %, significantly more than half of the protein is in an EPR-silent E_n state and therefore unaccounted for in the EPR spectra.

IR species

In contrast to EPR spectroscopy, IR spectroscopy is generally not sensitive to the spin state. CO binding can be conveniently followed with IR measurements, since the stretching vibration is IR active due to the dipole moment changes during the CO stretching vibration.[60] Binding to a metal center reduces the frequency compared to unbound CO, $2143\,\text{cm}^{-1}$ for CO gas,[225] but the reported frequencies (*vide infra*) are still higher in energy than the amide I band of the protein, which would complicate the detection of peaks between $1600\,\text{cm}^{-1}$–$1700\,\text{cm}^{-1}$.[226] Time-resolved stopped-flow Fourier-transform IR (SF-FT-IR) measurement further reveal information about the mechanistic sequence of the binding event. However, since enzymatic turnover create a mixture of E_n states, the IR species can not be easily assigned to a specific redox state.o

SF-FT-IR measurements, which were performed under conditions comparable to the EPR studies, also show differences at low and high CO partial pressures (Figure 1.9c).[226–228] At low pressure, a single transient band appears within a few seconds, $\text{loCO}_{\text{IR,1}}$ at $1904\,\text{cm}^{-1}$, which has fully converted to $\text{loCO}_{\text{IR,2}}$ at $1715\,\text{cm}^{-1}$ after 2 min.[226] This measurement supports a mechanism in which CO binds initially as a terminal ligand and then assumes a bridging binding mode. As in EPR studies, turnover conditions are necessary to detect CO binding with IR spectroscopy. However, it has been reported that inhibition of azide (N_3^-) reduction due to the presence of CO happens within 400 ms of enzymatic turnover, which is quicker than the appearance of the first IR species.[229]

Under high CO pressure, a multitude of bands appear in the range of $1700\,\text{cm}^{-1}$–$2000\,\text{cm}^{-1}$, all of which have different time courses.[227,228,230] IR-monitored photolysis demonstrates that some bands correspond to species with two bound CO molecules.[231,232] However, two of the SF-FT-IR bands have time courses and frequencies ($1906\,\text{cm}^{-1}$ and $1715\,\text{cm}^{-1}$,) that are highly similar to the $\text{loCO}_{\text{IR,1}}$ and $\text{loCO}_{\text{IR,2}}$ species observed under low CO pressure and therefore most likely correspond to the initial binding of a single CO. Repeating the measurement with Val70 variants has been shown to shift the initial transient band from $1906\,\text{cm}^{-1}$ (wild type) to $1895\,\text{cm}^{-1}$ (Val70→Ile) or $1911\,\text{cm}^{-1}$ (Val70→Gly).[228] The Val70 residue is close to Fe6 in FeMoco, which suggests that Fe6 is the initial binding site.

One particular IR-monitored photolysis process of the high CO pressure species is the transformation of $hiCO_{IR}$ to $loCO_{IR,2'}$ (Hi-1 and Lo-1 in the original work, Figure 1.9d). It is characterized by the disappearance of bands 1973 cm^{-1} and 1679 cm^{-1} and the appearance of bands at about 2130 cm^{-1} and 1711 cm^{-1} and therefore likely corresponds to the conversion of a terminal plus bridging CO species to a bridging CO with the second CO molecule being trapped in some protein pocket. Subsequent annealing recovers the initial $hiCO_{IR}$ bands and the activation energy is with 1 kcal/mol in line with the EPR-monitored conversion between $loCO_{EPR}$ and $hiCO_{EPR}$. However, two terminal CO molecules were proposed for $hiCO_{EPR}$ based on the ^{13}C ENDOR measurements.[217] Furthermore, the frequency of $loCO_{IR,2'}$ coincides with $loCO_{IR,2}$ under low CO pressure, which may therefore correspond to the same species.

XRD structures

XRD structures have been solved for CO-inhibited MoFe protein at 1.5 Å and 1.33 Å resolution under low and high CO partial pressure, respectively ($loCO_{XRD}$ and $hiCO_{XRD}$ in Figure 1.9e).[233,234] In $loCO_{XRD}$, CO has replaced the S2B sulfide bridging Fe6 and Fe2, and is therefore consistent with the structural motif proposed for $loCO_{IR}$ and $loCO_{EPR}$. Exposing the $loCO_{XRD}$ crystals to high CO pressure yields a superposition of the $loCO_{XRD}$ and the $hiCO_{XRD}$ structures. In $hiCO_{XRD}$, an additional CO is bound terminally to Fe6, which has been suggested as the initial CO binding site based on the influence of Val70 substitutions on the $loCO_{IR}$ band[228] and may also be the binding site for the second CO molecule. EPR measurements of the $loCO_{XRD}$ and $hiCO_{XRD}$ crystals confirm the presence of the $loCO_{EPR}$ and $hiCO_{EPR}$ species, respectively. Selective isotope labeling of the central carbide combined with ^{13}C ENDOR measurements suggest that the central carbide is responsible for maintaining the structural integrity of FeMoco in the CO-bound structures.[235] However, as for the IR species, no spin quantitation was reported and a one-to-one correspondence can not be established. Furthermore, structures similar to $loCO_{XRD}$ and $hiCO_{XRD}$ have been reported for V nitrogenase.[221,236]

Solution-extracted FeMoco and model compounds

FeMoco can be extracted from the MoFe protein into N-methylformamide (NMF) solution. NMF-extracted FeMoco likely retains its structural integrity based EXAFS measurements and the observation of an $S = \frac{3}{2}$ EPR signal similar to that observed for the E_0 MoFe protein.[197] The solvent molecules likely coordinate to the Fe centers. CO was not found to interact with the redox state giving rise to the $S = \frac{3}{2}$ EPR signal, but required further reduction.[237–239] A single IR band appeared at 1835 cm^{-1} after one-electron reduction, which shifts to 1808 cm^{-1} after second one-electron reduction. Because these frequencies are more than 100 cm^{-1} lower compared to the MoFe protein, the authors proposed a bridging CO motif. Interestingly, the presence of cyanide allows CO to bind to solution-extracted FeMoco in a redox state that likely corresponds to the E_0 state.[239]

Propositions for binding motifs in CO-bound FeMoco based on IR frequencies are generally made by comparing to compounds with known structures. However, data

on suitable model compounds are scarce, since the binding of CO to a metal tends to increase the binding affinity for more CO ligands,[240] which appears to be prohibited for FeMoco in the MoFe protein.[226] The compounds that do feature a single CO bound to an Fe center[241–249] usually have a single Fe center coordinated by O-, N-, or P-based ligands, which makes the comparison to CO-bound FeMoco challenging. In a preprint manuscript, $[Fe_4S_4]$ clusters have been characterized where a single CO is bound terminally to one of the Fe centers.[250] The CO frequency was determined as $1902\,cm^{-1}$ for $[Fe_4S_4]^{1+}$ and $1832\,cm^{-1}$ for $[Fe_4S_4]^{0}$, which demonstrates that the CO bands observed in the solution-extracted FeMoco may also arise from a terminally bound CO.[237,238]

Computational models

Different aspects of CO binding to FeMoco have been studied with DFT-based computational models. In the context of experimental studies, Scott *et al.* calculated the CO frequencies for the $hiCO_{IR}$ species and proposed a terminal CO and a formyl species bound to an E_2-type cofactor.[251] In standalone computational studies, Rod and Nørskov reported a significantly increased CO binding affinity in the E_1 and E_2 redox state compared to E_0.[252] Dance discussed the coupling of the CO vibrational mode with hydrides present in more reduced E_n states.[253] Varley and Nørskov published a complete mechanism of CO reduction to CH_4.[254]

Except for the QM/MM model used in the quantum refinement of the $loCO_{XRD}$ structure by Bergmann *et al.*, none of the computational models investigating the CO−FeMoco interaction explicitly included the environment of the MoFe protein. However, multiple experiments using variant MoFe protein have demonstrated the importance of the protein environment on CO binding: substitutions can affect the frequency of the $loCO_{IR,1}$ band,[228] they generate different CO-bound species under high CO pressure,[231,232,251] they control the appearance of the $loCO_{EPR}$ and $hiCO_{EPR}$ signals,[223] and they even enable CO reduction for Mo nitrogenase.[191] Therefore, a QM/MM model is employed in the present work, which explicitly models the protein environment, and Chapter 5 focuses on the influence of the residue in the proximity of FeMoco during the CO binding event.

2. S→Se substitution in synthetic $[Fe_2S_2]$ clusters

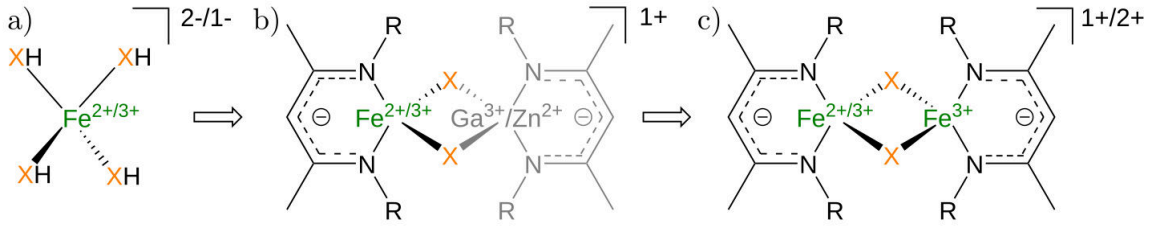

Figure 2.1. – Outline of the bottom-up approach used to study the effect of S→Se substitution on the electronic structure of FeS clusters: (a) monomeric $[Fe(XH)_4]^{2-,1-}$, (b) diamagnetically substituted $[FeGaX_2]^{1+}$ and $[FeZnX_2]^{1+}$ dimers, and (c) unsubstituted $[Fe_2X_2]^{2+,1+}$ dimers. Figure adapted from Ref [22].

The present chapter deals with the S→Se substitution in FeS clusters. This substitution is often believed to be innocent, which means that it does not or at most negligibly affects the electronic structure of the molecule. The selective substitution of S with Se is a powerful tool in understanding very complex clusters such FeMoco, the active site of nitrogenase, which will be the protagonist in Chapter 4 and 5. For example, Se has already been selectively introduced into FeMoco and the distinct X-ray absorption properties of S and Se have been exploited to study the electronic structure and the mechanism of substrate reduction.[170–172] For a more detailed account of the experimental data on S→Se substitution in FeS clusters, the reader is referred to Section 1.2.2.

This chapter addresses the question: How does S→Se substitution modulate the electronic structure of FeS clusters? The study follows a bottom-up approach, which is outlined in Figure 2.1. The computational details for all models are summarized in Section 2.1. Section 2.2 explores the S→Se substitution in the small and relatively simple monomeric $[Fe(XH)_4]^{1-,2-}$ (X=S,Se) complexes (Figure 2.1a). Here, the results of CASSCF calculations are discussed in detail, such as the composition of the wave function, the active orbitals, and the effect of the NEVPT2 correction. Furthermore, the electronic structure is further interpreted within the framework of *ab initio* ligand field theory (AILFT) and the angular overlap model (AOM). Finally, Section 2.3 investigates the dimers $[Fe_2X_2]^{2+,1+}$ (X=S,Se) and separates the analysis in two parts: the local electronic structure and the metal-metal interaction. The local electronic structure is studied in the diamagnetically substituted complexes $[FeGaX_2]$ and $[FeZnX_2]$ (Figure 2.1b), which is a commonly used technique in the study of system containing multiple magnetic centers.[21,256,257] For the metal-metal interaction, the CASSCF

wave function of the unsubstituted $[Fe_2X_2]$ complexes is analyzed (Figure 2.1c) and interpreted with phenomenological models such as the Heisenberg or double exchange model.

2.1. Computational details

Section 2.2.1 The geometries for the $[Fe(XH)_4]^{2-,1-}$ complexes (see Figure 2.2) were optimized using `ORCA` release version 4.1.[75–77] The hybrid DFT functional TPSSh (10 % HF exchange)[51–53] and the `ZORA-def2-TZVP` basis set[258] was used. A conductor-like polarizable continuum model (C-PCM) ($\epsilon = \infty$) was included to stabilize the charge of the anionic complexes.[259] Dispersion interactions were included via an atom-pairwise parameterization (`D3` keyword).[61,64] Scalar relativistic effects were included through the zeroth order regular approximation (ZORA).[260,261] The spin was assumed to be high spin $S = \frac{5}{2}, 2$ for $[Fe(XH)_4]^{2-,1-}$, respectively.

Section 2.2.2 The IEs were calculated with ionization potential equation of motion coupled cluster singles doubles (IP-EOM-CCSD),[262] which has been shown to give reliable results for IEs for valence electrons.[263] The evaluation of the Coulomb and exchange integrals was sped up with the resolution of identity (RI) approximation (`RI-JK` keyword) using the `def2/JK` auxiliary basis set. The remaining computational details are analogous to Section 2.2.1. The character of the ionization events, such as ligand-based or metal-based, was determined based on visual inspections of the major orbital contribution to the electron hole.

Section 2.2.3, 2.2.4, and 2.2.5 The environment of the $[Fe(XH)_4]$ complexes were modeled with a C-PCM to account for the negative charge. The minimal active space was used for the $[Fe(XH)_4]^{1-,2}$ complexes (CAS(5,5) and CAS(6,5), respectively). Unless specified otherwise, the wave function was averaged over all possible states in the given CAS (see Section A1.2 for details). NEVPT2 was used in the strongly-contracted formalism (SC-NEVTP2) to include dynamic correlation in the state energies.[34–37] The AILFT module was invoked via the `actorbs dorbs` keyword in the `casscf` block. The calculation of spin-orbit coupling was requested by setting `dosoc true` in the `rel` block within the `casscf` block. The basis set `ZORA-def2-TZVP`[258] was used in combination with the zeroth order regular approximation (ZORA).[260,261] As a general convergence strategy, the localized orbitals of a high spin HF calculation were used as starting orbitals for the CASSCF calculation. First, the wave function was converged using the small `def2-SVP` basis set and in a second step used as the starting orbitals for the calculation with the `ZORA-def2-TZVP` basis set.

Section 2.3.1 The geometry optimizations of $[Fe_2X_2]^{2+,1+}$ were performed using `ORCA` version 4.2.[75–77] The full structures were optimized with BS-DFT. The XRD structures of the diferric compounds were used as starting points for the geometry.[23] The spin of the high spin determinant was $M_S = 5, \frac{9}{2}$ for $[Fe_2X_2]^{2+,1+}$, respectively, and the keywords `brokensym 5,4` and `brokensym 5,5` in the `scf` block were used to invoke the BS calculations. The hybrid DFT functional TPSSh (10 % HF exchange)[51–53]

and the ZORA-def2-TZVP basis set[258] was used. The resolution of identity coulomb approximation and chain-of-sphere exchange (RIJCOSX) approximation[264,265] was used for the Coulomb and HF exchange integrals with SACR/J as the auxiliary basis set.[261] Dispersion interactions were included via an atom-pairwise parameterization (D3 keyword).[61,64] Scalar relativistic effects were included through the ZORA.[260,261] The molecules were embedded in a C-PCM ($\varepsilon = \infty$)[259] using the Gaussian charge scheme.[266] For the subsequent sections, the 2,6-diisopropylphenyl moieties were substituted with methyl groups and the position of the hydrogen atoms were reoptimized with the protocol explained above.

Section 2.3.2 To create the diamagnetically substituted complexes $[FeZnX_2]^{1+}$ and $[FeGaX_2]^{1+}$, the $Fe_A{}^{2+}$ center in $[Fe_2X_2]^{1+}$ was replaced with Zn^{2+} and the $Fe_B{}^{3+}$ center with Ga^{3+}. The CAS(5,5) and CAS(6,5) calculations followed the same protocol as for the $[Fe(XH)_4]^{1-,2-}$ complexes (Section 2.2.3 to 2.2.4) apart from the following two differences: (i) The integral evaluation was sped up with the RI approximation (RI-JK keyword)[264] in combination with an automatically created auxiliary basis sit (autoaux keyword). (ii) The C-PCM made use of the newly implemented Gaussian charge scheme.[266]

Section 2.3.3 The calculations on the optimized $[Fe_2X_2]^{2+}$ models were performed with ORCA release version 4.2.[75–77] The computational details are analogous to Section 2.3.2. The active space is discussed separately in the text. To ease convergence of the CAS(22,16) wave function, the $S = 5$ high spin state was converged first using a small def2-SVP basis set. Then the basis set was enlarged and the lower spin states $S = 0, \ldots 4$ were calculated.

The final active orbitals were transformed with the following protocol: (i) Orbitals were transformed to natural orbitals (actorbs natorbs keyword in %casscf block). (ii) The ten Fe-based active orbitals were localized (orca_loc utility, NewBoys algorithm,[128] Fe-based orbitals need to have consecutive orbital indices). (iii) The five orbitals localized on Fe_A and Fe_B, respectively, were transformed to diagonalize the Fock matrix (orca_blockf utility, apply twice on Fe_A and Fe_B orbitals, respectively, which need to have consecutive orbital indices, ! keepfock necessary to save Fock matrix to disk). By default, CFG contributions $< 0.1\,\%$ are not printed in the ORCA output file, but these are important contributions for the present models. Therefore, the printing threshold was lowered by adding ci tprintwf 1e-6 in the %casscf block.

Section 2.3.4 The calculations for symmetrized $[Fe_2X_2]^{2+}$ models were performed using ORCA release version 4.1.[75–77] The symmetrized model for $[Fe_2S_2]^{2+}$ is shown in Figure 2.22. The def2-TZVP basis set was used.[258] The evaluation of the Coulomb integrals and the HF exchange integrals were approximated using the RI approximation (RI-JK keyword) in conjunction with the def2/JK auxiliary basis set. The dynamic correlation was accounted for with NEVPT2 in the strongly-contracted formalism (SC-NEVPT2).[34–37] Symmetry was enforced with the usesym keyword. The molecular geometry needs to match an abelian point group with a threshold 10^{-4} Å. Symmetrization of the coordinates was performed with the molecular visualization

program `Chemcraft` version 1.8.[267] $[Fe_2X_2]^{2+}$ was symmetrized to the point group D_{2h} (C_{2v}) for X_2=S_2/Se_2 (X_2=SSe). The irrep for the roots alternated between $A_g, B_{1u}, \ldots (A_1, B_2, \ldots)$ for $S = 0, 1, \ldots$ in the point group D_{2h} (C_{2v}). The z-axis was oriented along the Fe−Fe vector and the x-axis along the X−X vector.

Section 2.3.5 All computational details are equivalent to Section 2.3.3. The convergence of the active orbitals and state-averaging is discussed separately in the text.

Section 2.3.6 The BS-DFT determinant properties were calculated with the same computational protocol as Section 2.3.1 using the full geometry (phenyl moieties not substituted), except for the frequency calculations, which used the methyl-substituted models. The g values were calculated using spin-orbit coupling as defined by the single keyword `! SOMF(1x)`.[268] The coupling constants J were extracted using the projection proposed by Yamaguchi *et al.* (Eq 1.15).[110,111] Frequencies were calculated for the full, numerical Hessian (`! numfreq`).

Visualization 3D Molecular representations were rendered with `VMD` version 1.9.3[269,270] using the workflow published online.[271] Volumetric data is shown with isosurfaces at ±0.05 (solid) and ±0.025 (transparent) unless otherwise noted. 2D molecular structures were created with `ChemDraw`. Nearly all other figures were created with the `python` libraries `matplotlib`[272] and `seaborn`[273]. The figures were post-processed with `Inkscape`,[274] using the extension `TexText` to embed LaTeX code.[275] Most figures are editable `.svg` files, which are made available along with this work, as well as the underlying data and the code used to generate the figures (see Section A1.1).

2.2. Monomeric FeS complexes: [Fe(XH)$_4$] (X=S,Se)

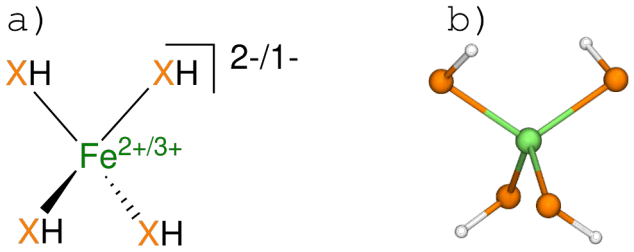

Figure 2.2. – (a) Schematic structure of the monomeric $[Fe(XH)_4]^{1-,2-}$ (X = S, Se) complexes and (b) 3D representation of the $[Fe(SeH)_4]^{1-}$ complex, showing the S_4 symmetry.

The structures of the $[Fe(XH)_4]$ complexes for which the S→Se is studied are shown in Figure 2.2. The oxidation states $[Fe(XH)_4]^{1-,2-}$ as well as the ligands SH^- and SeH^- are considered. The geometries of the complexes are compared in Section 2.2.1. In Section 2.2.2, the results of an IP-EOM-CCSD calculation is presented; the IEs between the unbound ligand and the ligand bound in the complex are compared, which allows for a quantification of the metal-ligand interaction. In Section 2.2.3, the CASSCF wave function is analyzed, discussing the active orbitals, the state compositions, state averaging, the NEVPT2 correction, and the size of the active space. In Section 2.2.4, the CASSCF wave function is related to ligand field theory (LFT) theory via AILFT, which yields d orbital energies, the Racah parameters B and C, and the spin-orbit coupling constant ζ. The LFT parameters are further interpreted in Section 2.2.5 with the AOM and the electronic structure can be summarized in the chemically intuitive AOM parameters such as e_σ, e_π, *etc.*

2.2.1. Geometry

Table 2.1. – Structural parameters of the DFT-optimized $[Fe(XH)_4]^{1-,2-}$ complexes.

	d(FeX)[Å]	d(XH)[Å]	$\angle(XFeX)$[°]	$\angle(FeXH)$[°]	$\chi(HXFeX)$[°]
$[Fe(SH)_4]^{1-}$	2.288	1.346	113.5	94.0	-60.0
$[Fe(SeH)_4]^{1-}$	2.411	1.471	113.5	94.0	-60.0
$[Fe(SH)_4]^{2-}$	2.353	1.347	116.0	100.5	-75.0
$[Fe(SeH)_4]^{2-}$	2.470	1.472	116.0	100.5	-75.0

The geometric parameters of the DFT-optimized structures of the complexes $[Fe(XH)_4]^{1-,2-}$ (X=S,Se) (see Figure 2.2) are listed in Table 2.1. The angles between XFeX and FeXH as well as the dihedral angles defined by HXFeX were averaged over X=S,Se for each oxidation state to make the electronic structures more comparable. As a consequence, the point group of the complexes is S_4.

Naturally, both the ligand exchange $SH^- \rightarrow SeH^-$ and the reduction event $[Fe(XH)_4]^{1-} \rightarrow [Fe(XH)_4]^{2-}$ affect the structure. The more bulky Se-based ligand increases the FeX and the XH bond length by about 0.12 Å. The FeX bond in the more reduced $[Fe(XH)_4]^{2-}$ is elongated by 0.06 Å compared to $[Fe(XH)_4]^{1-}$ while the XH bond is unaffected. The XFeX angle, which is cutting the improper rotation axis in half, and the FeXH angle are slightly increased upon reduction, therefore, the $[Fe(XH)_4]^{2-}$ complexes are slightly compressed along the improper rotation axis.

2.2.2. Ionization energies

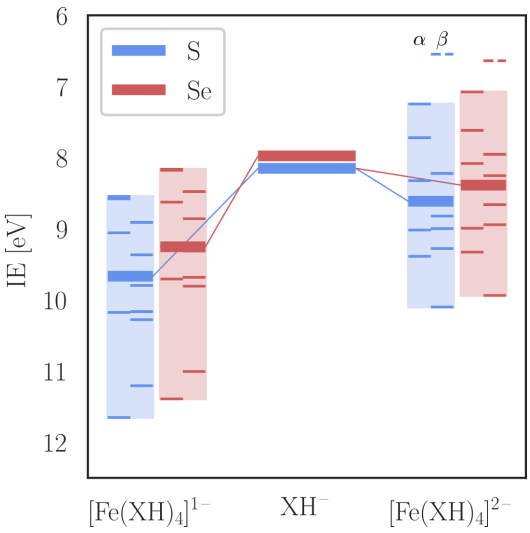

Figure 2.3. – ionization energies (IEs) for the valence electrons in the free ligand XH^- and in the complexes $[Fe(XH)_4]^{1-}$ and $[Fe(XH)_4]^{2-}$ calculated with IP-EOM-CCSD. For XH^-, the first IE is doubly degenerate, which splits into eight α and eight β IEs in $[Fe(XH)_4]^{1-,2-}$. The average is shown along the individual IEs in each box. The dashed lines for $[Fe(XH)_4]^{2-}$ indicate Fe-based IEs.

The lowest-energy IEs for the free anions SH^- and SeH^- are 8.15 eV and 7.98 eV, respectively. As for the atomic IEs of S and Se, 10.36 eV and 9.75 eV, respectively, electron is less strongly bound in the Se-based species.[157] When forming a metal complex, the presence of the positively charged metal stabilizes the negative charge on the anionic ligands and therefore increases the ligand-based IEs.

The ligand-based IEs are compared between the free anionic ligand XH^- and the complexes $[Fe(XH)_4]^{1-,2-}$ in Figure 2.3. It shows the lowest-energy IEs for XH^- in the middle, which is two-fold degenerate. The character of the two ionization events are dominated by the removal of an electron from the two p orbitals that are not involved in the X−H σ bond. When forming the complexes $[Fe(XH)_4]^{1-,2-}$, the interaction between the ligands lifts the degeneracy of the ligand-based IEs. As a consequence of the ligand interaction, an electron can be more easily removed from the ligands in the ferrous complex compared to the free ligand. Considering the two-fold degeneracy in the free ligands (neglecting spin for now), a total of $2 \cdot 4 = 8$ ligand-based ionization events have a similar character as the lowest-energy ionization events in the free ligand.

The barycenter of these ligand-based IEs in the complexes is shifted to higher energies compared to the free ligands in accord with the presence of the positively charged Fe center.

Furthermore, the presence of the open-shell metal centers in the $[Fe(XH)_4]^{1-,2-}$ complexes lifts the degeneracy of the α and β IEs through spin polarization, which is not present in the closed-shell free ligands. Therefore, a total of $2 \cdot 4 \cdot 2 = 16$ ligand-based ionization events in the complexes have the same character as the lowest-energy ionization from the free ligand (boxes in Figure 2.3). Interestingly, the lowest-energy ionization event in the ferrous complexes is not ligand-based, but describes the removal of an Fe-based β electron (dashed lines in Figure 2.3). This is a manifestation of the so-called inverted level scheme used in the discussion of high spin transition metal complexes, which is also facilitates the redox event $Fe^{3+} + e^- \rightleftharpoons Fe^{2+}$ in biological FeS clusters.[133]

The interaction between fragments of a chemical system is typically discussed in the basis of molecular orbitals. For example, the stabilization of ligand-based orbitals as a consequence of complex formation is expressed within the extended Hückel theory[276] as

$$\Delta E = \frac{H_{ML}^2}{e_M - e_L} \tag{2.1}$$

Here, e_M and e_L are the metal and ligand orbitals of the isolated fragments and H_{ML} is the effective resonance integral, which is, in turn, proportional to the overlap S_{ML} of the metal and ligand orbitals in the complex. Therefore, the stabilization of the ligand orbitals can be related to the mixing of S or Se with the Fe orbitals.

Table 2.2. – The average ionization energies (IEs) for isolated $Fe^{3+,2+}$ and XH^- as well as for the complexes $[Fe(XH)_4]^{1-,2-}$ calculated with IP-EOM-CCSD and estimated resonance integrals H_{ML} according to Eq 2.1. All values are given in eV.

	IP_{avg}	H_{ML}
SH^-	8.15	-
SeH^-	7.97	-
Fe^{3+}	34.73	-
Fe^{2+}	20.49	-
$[Fe(SH)_4]^{1-}$	9.67	6.35
$[Fe(SeH)_4]^{1-}$	9.25	5.84
$[Fe(SH)_4]^{2-}$	8.81	2.85
$[Fe(SeH)_4]^{2-}$	8.58	2.75

While orbitals can be easily plotted and interpreted in terms of atomic contributions, their energies are not observables, *i.e.* they cannot be measured experimentally. On the other hand, IEs are observables and further have the advantage that they respect electronic relaxation effects. Using the IEs instead of orbital energies in Eq 2.1, the resonance integral H_{ML} can be estimated and since H_{ML} is proportional to the overlap integral S_{ML}, the metal-ligand covalency quantified, as shown in Table 2.2. The

resonance integral estimates are around 6 eV for the ferric $[Fe(XH)_4]^{1-}$ and around 3 eV for the ferrous $[Fe(XH)_4]^{2-}$ complexes, consistent with the more positively charged metal center binding the ligands more strongly. According to the IEs, the metal-ligand covalency for Se-based ligands is 9 % and 4 % smaller in the ferric and ferrous complexes, respectively, compared to the S-based ligands. Since the metal-ligand bond is stronger in the ferric complex, differences between S and Se are larger in the ferric compared to the ferrous complexes.

2.2.3. CASSCF calculations

Active orbitals in the minimal CAS

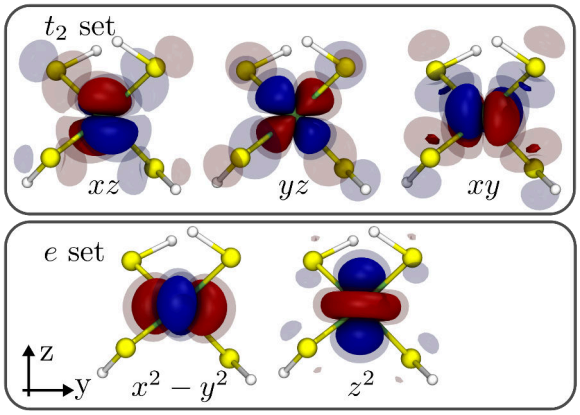

Figure 2.4. – Active orbitals for the CAS(6,5) wave function in $[Fe(SH)_4]^{1-}$. The active orbitals for all four $[Fe(XH)_4]^{1-,2-}$ have similar shapes. The labels t_2 and e correspond to pseudo-T_d symmetry. Isosurfaces are shown at ±0.05 and ±0.025.

The smallest active space necessary to describe the low-lying electronically excited states in $[Fe(XH)_4]^{1-,2-}$ complexes consists the five Fe-centered 3d orbitals. These active space orbitals are shown in Figure 2.4. The labels e and t_2 correspond the character of the $x^2 - y^2, z^2$ and xz, yz, xy orbital sets in T_d symmetry. In S_4 symmetry, only the xz and yz orbitals are degenerate. With the oxidation states $Fe^{3+,2+}$ the 3d orbitals are occupied by five and six electrons respectively, and the CASSCF calculations are labeled CAS(5,5) and CAS(6,5), respectively. Furthermore, the CASSCF wave function is invariant with respect to unitary transformations within the active space. Therefore, the active orbitals in Figure 2.4 and Table 2.3 were chosen to diagonalize the Fock matrix (`orca_blockf` utility), which allows for an easy labeling according to the real d orbitals of the hydrogen atom in a given point group.

The $x^2 - y^2$ and the z^2 active orbitals (e set, see Figure 2.4) have little ligand overlap, because the orbital lobes are oriented in between the ligands, and the metal character is 95.4 %–98.0 %. The lobes of the orbitals in the t_2 set, however, point towards the ligands and the metal character is reduced to 86.7 %–95.4 % as a consequence of larger ligand contributions. The xy orbital shows the largest metal-ligand overlap and the metal contribution is about 2 % lower compared to the xz and yz orbitals.

Table 2.3. – Loewdin atomic contribution of the Fe center ([%]) to the active orbitals of the CAS(5,5) and CAS(6,5) wave functions in $[Fe(XH)_4]^{1-,2-}$, respectively. The remainder stems from the XH^- ligand. See Figure 2.4 for orbital shapes.

	x^2-y^2	z^2	xz	yz	xy
$[Fe(SH)_4]^{1-}$	97.3	95.4	89.4	89.4	87.2
$[Fe(SeH)_4]^{1-}$	97.5	95.5	89.3	89.3	86.7
$[Fe(SH)_4]^{2-}$	97.9	97.2	95.2	95.2	93.7
$[Fe(SeH)_4]^{2-}$	98.0	97.3	95.4	95.4	94.0

The reduction $Fe^{3+}\rightarrow Fe^{2+}$ decreases the metal contribution to the active orbitals, which is in line with the increase in metal-ligand bond length. The metal character increases by up to 2 % in the e orbitals and by up to 6 % in the t_2 orbitals. Naturally, the composition of the t_2 orbitals is more sensitive to the reduction event due to the stronger metal-ligand mixing.

The ligand exchange $SH^-\rightarrow SeH^-$, on the other hand, increases the metal-ligand distance twice as much as the reduction event, but the composition of the active orbitals remains largely unchanged. The ligand exchange changes the metal character by 0.1 %–0.2 % in the e orbitals and by 0.1 %–0.5 % in the t_2 orbitals. Therefore, no clear change in the symmetry-restricted covalency is observed in the $[Fe(XH)_4]^{1-,2-}$ complexes upon S→Se substitution.

Ground state wave function

In addition to the orbital shapes, their occupation numbers are needed in order to fully characterize a CASSCF wave function. While the inactive orbitals are exclusively doubly occupied, the active orbitals can have occupation numbers between 0 and 2. The occupation numbers are defined by a linear combination of electron configurations in which an active orbital holds either 0, 1, or 2 electrons.

In their ground states, the complexes $[Fe(XH)_4]^{1-,2-}$ have high spin Fe centers with $S = \frac{5}{2}$ and 2, respectively. For the $S = \frac{5}{2}$ ground state, the only possible electron configuration is $(x^2-y^2)^1(z^2)^1(xz)^1(yz)^1(xy)^1$ (see Figure 2.4). Therefore, the occupation number of each active orbital is exactly 1 regardless of the choice of the active orbitals. For the $S = 2$ ground state the $(x^2-y^2)^2(z^2)^1(xz)^1(yz)^1(xy)^1$ dominates the in the wave function (> 99.9 %). Here, the compact representation of the wave function is only possible, because the active orbitals were transformed with the `orca_blockf` utility. A different set of active orbitals typically leads to a linear combination containing multiple electron configurations.

Excited states

Excited CAS states in $[Fe(XH)_4]^{1-,2-}$ with a spin other than the ground state are at least 15 000 cm^{-1} higher in energy (19 000 cm^{-1} after NEVPT2 correction). Within the given active space, the ground state is the only possible $S = \frac{5}{2}$ state for $[Fe(XH)_4]^{1-}$, but for $[Fe(XH)_4]^{2-}$ four excited $S = 2$ states exist in addition to the ground state.

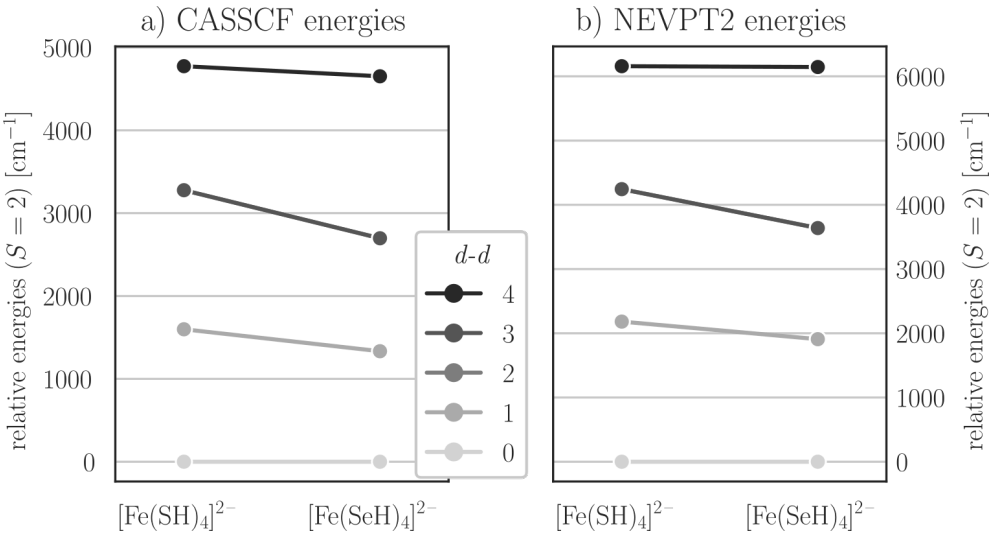

Figure 2.5. – Transition energies on the (a) CASSCF and (b) CASSCF/NEVPT2 level for the low-lying *d-d* states in [Fe(XH)$_4$]$^{2-}$ ($S = 2$, CAS(6,5)). The 2nd and 3rd transition energies are degenerate.

The first excited state is dominated by the single electron configuration $(x^2-y^2)^1(z^2)^2$ $(xz)^1(yz)^1(xy)^1$, the second with $(x^2-y^2)^1(z^2)^1(xz)^2(yz)^1(xy)^1$ *etc.* (see Figure 2.4 for orbitals). The second and the third excited state are degenerate, since the xz and yz orbitals are also degenerate in S_4.

The relative energies of five $S = 2$ states in [Fe(XH)$_4$]$^{2-}$ are shown in Figure 2.5. The CASSCF wave function was averaged only over the five $S = 2$ states (*vide infra*). All four excited CASSCF states lie within 5000 cm^{-1} of the ground state. The NEVPT2 correction increases the energy spacing between the state energies, placing the fourth excited state to more than 6000 cm^{-1} above the ground state. The increase in transition energies due to the NEVPT2 correction is larger for the first excited state (around 40 %) compared to the other three (around 30 %).

The S→Se substitution decreases the transition energies. The decrease in the transition energies based on the CASSCF energies is with 580 cm^{-1} most pronounced for the second and third excited state, respectively, followed by the first (265 cm^{-1}) and the fourth (122 cm^{-1}). In the ground state, the x^2-y^2 orbital is doubly occupied, which has the smallest metal-ligand overlap (see Table 2.3). Therefore, in any of the excited states the repulsion of metal-based electrons with the negatively charged ligands is increased. According to the difference in transition energies, this repulsion appears to be smaller for SeH$^-$ compared to SH$^-$ ligands. Indeed, the metal character of the active orbitals increases slightly in going from [Fe(SH)$_4$]$^{2-}$ to [Fe(SeH)$_4$]$^{2-}$ (see Table 2.3).

The NEVPT2 correction affects the transition energies in [Fe(SeH)$_4$]$^{2-}$ stronger than in [Fe(SH)$_4$]$^{2-}$. The NEVPT2 correction increases the four transition energies in [Fe(SH)$_4$]$^{2-}$ by 36 %, 30 %, 30 %, and 29 %, respectively and in [Fe(SeH)$_4$]$^{2-}$ by 43 %, 35 %, 35 %, and 32 %, respectively. Since both complexes have a different number of electrons, one cannot compare the absolute values of the NEVPT2 correction term, but the changes in the transition energies suggests that the CASSCF wave function is

a better reference for $[Fe(SH)_4]^{2-}$ compared to $[Fe(SeH)_4]^{2-}$.

State averaging in CASSCF calculations

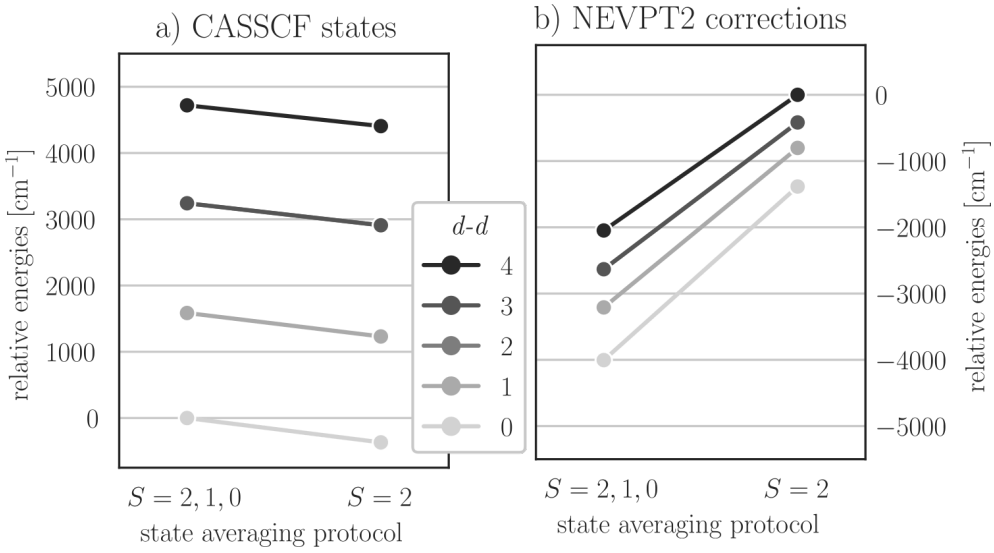

Figure 2.6. – (a) CASSCF state energies and (b) corresponding NEVPT2 corrections for the five lowest-lying $S = 2$ states in $[Fe(SH)_4]^{2-}$ (CAS(6,5)). The CAS(6,5) wave function was either averaged over all possible states ($S = 2, 1, 0$) or only over the five states shown ($S = 2$). The states d-d_2 and d-d_3 are degenerate.

The total number of states that can be generated with a CAS(5,5) wave function are 1, 24, and 75 for the spins $S = \frac{5}{2}$, $\frac{3}{2}$, and $\frac{1}{2}$, respectively, and 5, 45, and 50 for $S = 2$, 1, and 0, respectively, with a CAS(6,5) wave function. A computer code to calculate the number of states for arbitrary active spaces is presented in Section A1.2. Unless otherwise noted, the CASSCF wave function is optimized considering all possible CAS states. In `ORCA`, each set of states with the same spin has equal weight. For example of CAS(6,5), the sets of states with $S = 2$, 1, and 0 are all considered with a weight of $\frac{1}{3}$, respectively. Therefore, a state with $S = 2$ enters with a weight of $\frac{1}{3} \cdot \frac{1}{5} = \frac{1}{15}$, but a state with $S = 0$ with a weight of $\frac{1}{3} \cdot \frac{1}{75} = \frac{1}{225}$.

The approach of optimizing the CASSCF wave function for multiple states simultaneously is called state averaging. The goal is to obtain a wave function that describes all relevant states equally well and is not biased towards any state, *e.g.* the ground state, if one is interested in transition energies. However, the more states enter the state averaging, the more is the resulting wave function a compromise between the different states. Figure 2.6 compares the relative energies of states with $S = 2$ when averaging over all possible CAS states with averaging over only the five $S = 2$ states in $[Fe(SH)_4]^{2-}$. Including only the five $S = 2$ states in the state averaging lowers their energies on average by 340 cm^{-1}, because the wave function is not optimized simultaneously for the $S = 1$, and 0 states. On the other hand, the transition energies between the $S = 2$ CASSCF states change as well, but by less than 25 cm^{-1}. The NEVPT2 corrections for the respective states (Figure 2.6b) are larger when averaging over all $S = 2$, 1 and 0 states, which suggests that the perturbative correction attempts to

remedy the less optimal CASSCF wave function. In contrast to the CASSCF energies, however, the state averaging procedure strongly affects the NEVPT2-corrected $S = 2$ transition energies. These are as much as 200 cm^{-1} higher for the $S = 2$, 1, and 0 state averaged wave function compared to $S = 2$.

Beyond the minimal active space

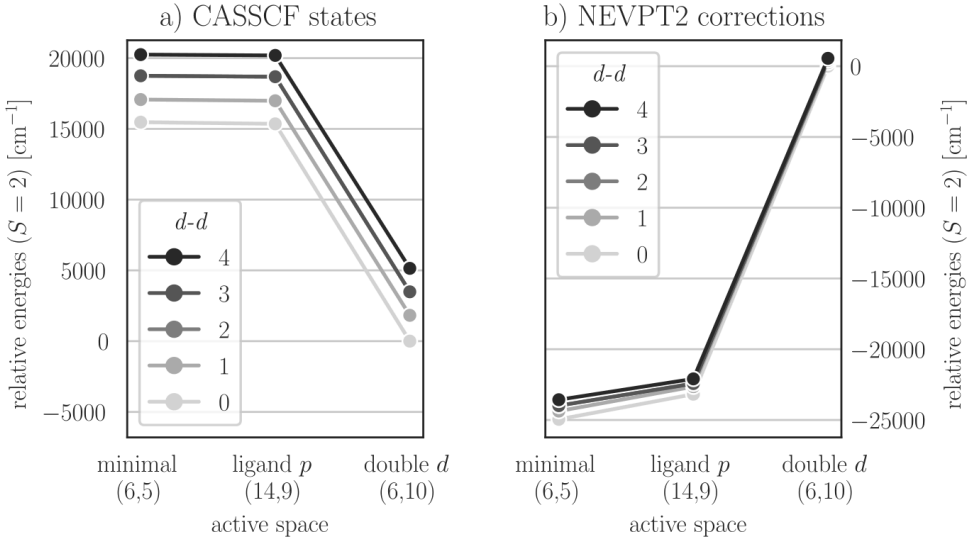

Figure 2.7. – (a) CASSCF state energies and (b) corresponding NEVPT2 corrections for the five lowest-lying $S = 2$ states in $[Fe(SH)_4]^{2-}$. The active space consist of the five Fe 3*d* orbitals (6,5), in addition to (6,5) one doubly occupied *p* orbital on each ligand (14,9), or in addition to (6,5) five unoccupied Fe 4*d* orbitals (6,10). The states *d*-d_2 and *d*-d_3 are degenerate.

While for the minimal active space in $[Fe(XH)_4]^{1-,2-}$ (CAS(5,5) and CAS(6,5), respectively) only *d*-*d* excited configurations contribute to the CASSCF wave function, extending the active space allows for different types of excited configurations to contribute as well. Including occupied ligand orbitals in the active space enables excitations from the ligand into the partly occupied metal orbitals, therefore, LMCT configurations are mixed into the wave function. Similarly, including unoccupied ligand orbitals in the active space mixes metal-to-ligand charge transfer (MLCT) configurations into the wave function. Also, one can include a second *d*-shell in the active space, which consists of the five unoccupied metal 4*d* orbitals. Both the 3*d* and the 4*d* set have the same angular momentum, but a different radial extent and 3*d*-4*d* excited configurations account for the so-called radial correlation.[27]

The different types of configurations are often categorized in the literature as contributing to either *static* or to *dynamic* electron correlation, where the first is to be handled by the flexibility of the reference, *i.e.* the CASSCF wave function, and the later by a suitable subsequent method, *e.g.* NEVPT2. The distinction between static and dynamic correlation is not clear-cut.[277] Malrieu *et al.* define dynamic correlation as arising from configurations beyond the minimal active space, therefore, the 3*d*-4*d*,

LMCT, and MLCT configurations mentioned above would all be considered to capture dynamic correlation effects.[109] However, the authors also highlight the observation that a minimal active space is often insufficient to quantitatively capture magnetic properties of transition metal complexes.

Figure 2.7 compares the five lowest CASSCF state energies and the respective NEVPT2 corrections for different active spaces in $[Fe(SH)_4]^{2-}$. All the following observations are qualitatively equivalent for the Se-substituted $[Fe(SeH)_4]^{2-}$. The five lowest states have a spin of $S = 2$ and the wave function was averaged over all five states. In the ligand p active space, each S contributes one doubly occupied predominantly $3p$ orbital. Including more ligand p orbitals in the active space, similar to Figure 2.23 in Section 2.3.4, was not possible due to SCF convergence issues. Including the second d-shell in the active space lowers the total CASSCF energies by about $15\,000\,\mathrm{cm}^{-1}$, while they are virtually unaffected by including ligand p orbitals. The NEVPT2 corrections for each state, in turn, are significantly smaller in magnitude for the active space including the second d-shell compared to the ligand p orbitals. Therefore, an active space including the second d-shell of the metal is a better starting point for the treatment of dynamic correlation with NEVPT2 than an active space including the ligand p orbitals

Furthermore, the spread of the NEVPT2 correction energies for the five states is significantly lower for the active space including the second d-shell (standard deviation $212\,\mathrm{cm}^{-1}$) compared to the minimal ($520\,\mathrm{cm}^{-1}$) and the ligand p active space ($401\,\mathrm{cm}^{-1}$). Given that the active space including the second d-shell is better suited for NEVPT2 calculations, it appears reasonable to include the second d-shell for reliable d-d excitation energies. The second d-shell is, however, not included for the remainder of this work, because it is (i) not compatible with the AILFT approach and (ii) not feasible for complexes with two magnetic centers due to computational limitations, both of which are the focus of the following analysis. Hence, it is important to keep a possible bias of a limited active space in mind in the following discussion.

2.2.4. *Ab Initio* ligand field theory

The experimental properties of transition metal complexes with a single magnetic center have been successfully interpreted with the help of ligand field theory for several decades. On the other hand, the development of efficient algorithms for *ab initio* quantum chemical methods and the increase in computational power have made the direct calculation of molecular properties with feasible for transition metal complexes as well. While the calculated properties become more and more accurate, the chemical interpretation of the calculated wave function is a non-trivial task. One approach to interpret the calculated results is the AILFT.[278,279] AILFT allows for the extraction of LFT parameters from an *ab initio* wave function.

The key idea behind AILFT lies in the observation that the space spanned by a CAS(n,5) wave function has the same dimensionality as the LFT model for a d^n transition metal complex.[278] Therefore, one can attempt to extract LFT parameters by fitting the CAS(n,5) wave function. The basis for LFT are the pure, real metal d orbitals. However, orbital mixing is not constrained in any way during the CASSCF and even though the active orbitals in a CAS(n,5) wave function for transition metal complexes

are typically dominated by the metal d orbitals, some metal-ligand mixing occurs (see Table 2.3). A unique mapping between the active orbitals and the LFT basis can be established by diagonalizing the $\hat{l}_z$ operator over the active orbitals. The resulting active orbitals have approximate magnetic quantum number of $m_l = -2, \ldots +2$, which can be easily combined to the real representation of the d orbitals. Now, the LFT parameters can be extracted from the CASSCF wave function much in the same way that they can be extracted from experimental data.

d orbital energies

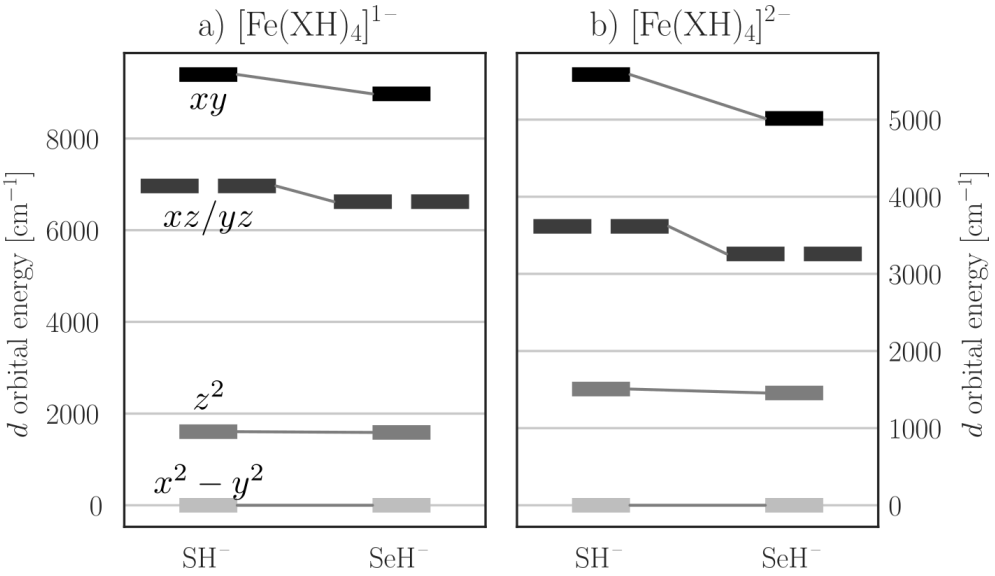

Figure 2.8. – AILFT d orbital energies for the (a) $[Fe(XH)_4]^{1-}$ and (b) $[Fe(XH)_4]^{2-}$ complexes (CAS(5,5) and CAS(6,5), respectively, including NEVPT2). See Figure 2.4 for orbital plots.

The d-orbital energies, *i.e.* the ligand field splitting, for the $[Fe(XH)_4]^{1-}$ ($[Fe(XH)_4]^{2-}$) complexes according to AILFT calculations are shown in Figure 2.8 as well as listed in Table 2.4. The lowest orbital in the e set is the $x^2 - y^2$ orbital, followed by the z^2 orbital at about 1600 cm^{-1} (1500 cm^{-1}). In the t_2 set, the degenerate orbitals xz and yz are about 2400 cm^{-1} (1850 cm^{-1}) lower than the xy orbital. Therefore, the orbitals with a z component are lowest in each set. The barycenters of the e and the t_2 set are about 6800 cm^{-1} and 3400 cm^{-1} apart, respectively, which is consistent with the ligand field splitting being larger for the ferric complexes.

The S→Se substitution in $[Fe(XH)_4]^{1-}$ ($[Fe(XH)_4]^{2-}$) only has a small influence on the e set. The energy gap between the $x^2 - y^2$ and the z^2 orbital is reduced by merely 16 cm^{-1} (51 cm^{-1}). The energy gap in the t_2 set, *i.e.* the difference between the xz/yz and the xy orbitals, increases with the substitution by 79 cm^{-1} (211 cm^{-1}). Alternatively, this increase can be interpreted as the barycenters of the e and the t_2 set approaching each other upon Se substitution: The gap between the barycenters decreases by 361 cm^{-1} (402 cm^{-1}). A similar decrease in the gap between the barycenters has been observed for S→Se substitution in Co^{3+} and Cr^{3+} complexes.[280,281]

Table 2.4. – AILFT parameters for $[Fe(XH)_4]^{1-,2-}$ with and without NEVPT2 correction. The d orbital energies, the Racah parameters B and C, and the spin orbit coupling constant ζ are given in [cm^{-1}]. The reduction relative to the free ions (rB, rC, and $r\zeta$) is given in [%]. See Table A1 for B, C, and ζ in the free $Fe^{3+,2+}$ ions; see Figure 2.8 for orbitals plots.

	CASSCF									
	x^2-y^2	z^2	xz/yz	xy	B	C	ζ	rB	rC	rζ
$[Fe(SH)_4]^{1-}$	0	1384	6267	8514	1066	4118	407	18.1	15.3	13.8
$[Fe(SeH)_4]^{1-}$	0	1329	5829	7952	1056	4098	372	18.8	15.7	21.1
$[Fe(SH)_4]^{2-}$	0	1404	3439	5275	1085	4055	391	8.4	7.3	5.8
$[Fe(SeH)_4]^{2-}$	0	1360	3094	4695	1089	4063	382	8.0	7.1	8.1
	NEVPT2									
	x^2-y^2	z^2	xz/yz	xy	B	C	ζ	rB	rC	rζ
$[Fe(SH)_4]^{1-}$	0	1605	6965	9395	847	3514	–	25.5	20.7	–
$[Fe(SeH)_4]^{1-}$	0	1589	6623	8974	826	3469	–	27.3	21.7	–
$[Fe(SH)_4]^{2-}$	0	1507	3616	5584	883	3592	–	13.4	9.4	–
$[Fe(SeH)_4]^{2-}$	0	1456	3258	5015	884	3606	–	13.3	9.1	–

Compared to plain CASSCF, the NEVPT2 correction increases the splitting between all d orbital energies (see Table 2.4). This is consistent with the Fe−X bond becoming more covalent when including dynamic correlation. Note that NEVPT2 does not change the orbitals, but enters the AILFT fit by changing the state energies. The increase for $[Fe(XH)_4]^{1-}$ ($[Fe(XH)_4]^{2-}$) is 10 %–20 % (5 %–8 %) and therefore more severe for the ferric complex. Similar to the excited state energies in the ferrous complex (see Figure 2.5), the d orbital energies in $[Fe(SeH)_4]^{2-}$ depend more strongly on the NEVPT2 correction (by 2 % on average) compared to $[Fe(SH)_4]^{2-}$. The d orbital energies in the ferric complexes follow the same trend.

Racah parameters

The Racah parameters B and C for the AILFT calculations on $[Fe(XH)_4]^{1-}$ ($[Fe(XH)_4]^{2-}$) are summarized in Table 2.4. The Racah parameters are a measure for the electron-electron repulsion at the Fe center. Focusing on the AILFT fit including NEVPT2-corrected energies, the reduction of B and C is 20.5 %–27.5 % (9.0 %–13.5 %) and therefore significantly stronger for the ferric complex. Despite the shorter FeX bond lengths in the ferric complexes (see Table 2.1), the electron-electron repulsion at the metal site is smaller at the ferric metal site, since the higher metal charge leads to a stronger mixing of the metal with the ligand orbitals.

The substitution S→Se only has a small influence on the respective Racah parameter. For the ferric complexes, the substitution increases the reduction of B and C by 1 %–2 %, but for the ferrous complexes, the reduction of B and C decreases by 0.1 %–0.3 %. Therefore, no clear trend in the change of the central field covalency is observed for the S→Se substitution.

The CASSCF wave function is known to be biased towards an overly ionic electronic structure, similar to HF. Hence, the electron-electron repulsion in the metal orbitals is expected to reduce by the inclusion of dynamic electron correlation via NEVPT2. The Racah parameters B and C from the AILFT fit without NEVPT2-corrected energies 22 %–28 % and 12 %–18 % higher, respectively, indicating stronger electron-electron repulsion in the metal orbitals. Similar to the d orbital energies (see Figure 2.8) and the transition energies in the ferrous complexes (see Figure 2.5) the contribution from the NEVPT2 correction to B and C is slightly larger (up to 2 %) for the Se complexes compared to S.

Spin-orbit coupling constant

The AILFT spin-orbit coupling constant for Fe is given in Table 2.4 for $[Fe(XH)_4]^{1-}$ ($[Fe(XH)_4]^{2-}$). Similar to the Racah parameters, ζ is reduced relative to the free Fe ion as a result of the metal-ligand covalency ($\zeta \propto \overline{r^{-3}}$, the average distance of the electron from the nucleus).[282] The reduction relative to the free ions is greater for the ferric compared to the ferrous complexes in line with the Racah parameters. In stark contrast to the Racah parameters, however, the reduction is with 21.1 % significantly stronger for $[Fe(SeH)_4]^{1-}$ compared to 13.8 % for $[Fe(SH)_4]^{1-}$. For the ferrous complexes as well, the reduction with 8.1 % stronger in $[Fe(SeH)_4]^{2-}$ compared to 5.8 % in $[Fe(SH)_4]^{2-}$. A similar trend for ζ has been calculated for Cr^{3+}-halide complexes: The reduction relative to the free ion is significantly stronger for Br^- and I^- compared to F^- and Cl^-, while the reduction in B only moderately increases over the halide group.[279] Note that ζ is not different after the NEVPT2 correction, since ζ does not depend on the state energies.

2.2.5. Angular overlap model

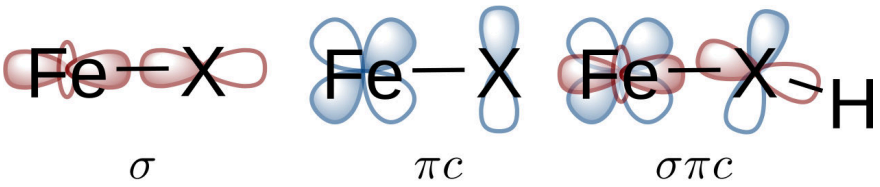

Figure 2.9. – Orbital interactions captured by the parameters in the angular overlap model (AOM). πs is analogous to π_c, but perpendicular to the FeXH plane. Figure adapted from Ref [22].

The relative energy of the d orbitals in Figure 2.8 is also known as the ligand field splitting. Given the geometry of the complex, the ligand field splitting can be rationalized in terms of the overlap of the metal with the ligand orbitals. Within the so-called angular overlap model (AOM) the ability of each ligand to affect the d orbital energies is quantified with parameters that capture different symmetries of the orbital interaction. For the XH^- ligands in $[Fe(XH)_4]^{1-,2-}$ these orbital interactions are visualized in Figure 2.9. The interaction of the XH^- p orbitals with the Fe d orbitals can have both σ and π symmetry, which is captured in the e_σ and e_π AOM parameters.

Table 2.5. – AOM parameters ([cm^{-1}]) for the XH$^-$ ligand in the complexes [Fe(XH)$_4$]$^{1-,2-}$ obtained by fitting the AILFT parameters, including the root-mean-square error (RMS).

	CASSCF				
	e_σ	$e_{\pi s}$	$e_{\pi c}$	$e_{\sigma\pi c}$	RMS
[Fe(SH)$_4$]$^{1-}$	5490	1233	6	187	98
[Fe(SeH)$_4$]$^{1-}$	5124	1232	5	163	120
[Fe(SH)$_4$]$^{2-}$	2960	662	31	304	53
[Fe(SeH)$_4$]$^{2-}$	2618	642	21	235	62
	NEVPT2				
	e_σ	$e_{\pi s}$	$e_{\pi c}$	$e_{\sigma\pi c}$	RMS
[Fe(SH)$_4$]$^{1-}$	6075	1401	6	189	98
[Fe(SeH)$_4$]$^{1-}$	5796	1435	5	168	123
[Fe(SH)$_4$]$^{2-}$	3124	712	36	335	55
[Fe(SeH)$_4$]$^{2-}$	2784	691	28	279	66

Since the atoms FeXH are not linear, one distinguishes between the in-plane $e_{\pi c}$ and the out-of-plane $e_{\pi s}$ parameter with respect to the plane formed by the atoms FeXH. Furthermore, the FeXH angle may lead to a perturbation of the p orbitals on X, which disrupts the mirror plane of the πc interaction. This disruption is captured by the misdirected valence parameter $e_{\sigma\pi c}$.[283] The misdirected valence parameter is subject to the constraint $e^2_{\sigma\pi c} = e_\sigma \cdot e_{\pi c}$, which increases the reliability of the fit.[284]

The AOM parameters for [Fe(XH)$_4$]$^{1-}$ ([Fe(XH)$_4$]$^{2-}$) are listed in Table 2.5. Including the NEVPT2 correction, the e_σ parameter is about 6000 cm^{-1} (3000 cm^{-1}) for [Fe(XH)$_4$]$^{1-}$ ([Fe(XH)$_4$]$^{2-}$). The e_π parameter is about 1400 cm^{-1} (700 cm^{-1}). Clearly, the σ and π donor strength of the XH$^-$ ligands is twofold as great in the ferric compared to the ferrous complexes. This ratio coincides with the ligand character in those active orbitals that have a large metal-ligand overlap, *i.e.* the t_2 set. However, the $e_{\sigma\pi c}$ parameter is with about 180 cm^{-1} (300 cm^{-1}) larger in the ferrous complexes. Apparently, the more covalent metal-ligand bond in the ferric complexes leads to a stronger ligand polarization, which reduces the role of the X−H bond in misdirecting the πc orbitals.

The S→Se substitution in [Fe(XH)$_4$]$^{1-}$ ([Fe(XH)$_4$]$^{2-}$) reduces the e_σ parameter by 4.7 % (10.9 %). Since the e_π parameters do not change significantly upon substitution, the reduction in the splitting between the e and the t_2 set (see Figure 2.8) is caused by the reduction in the e_σ parameters: In a tetrahedral ligand field, the splitting between the e and the t_2 set is given by $E_{t_2} - E_e = \frac{4}{3}e_\sigma - \frac{16}{9}e_\pi$. The $e_{\sigma\pi c}$ parameter decreases by about 15 % when substituting S with Se, which may indicate that the misdirected valence plays a less important role for larger, Se-based compared to smaller, S-based ligands. Similarly, the S→Se substitution in [CO(XPh)$_4$]$^{2-}$ also decreases the AOM parameters obtained via AILFT.[284]

The AOM parameters are fitted to the 1-electron ligand field matrix, which is

obtained from the AILFT calculations here. Considering the constraint $e^2_{\sigma\pi c} = e_\sigma \cdot e_{\pi c}$, three independent AOM parameters are fitted to the ligand field matrix, which contains six non-zero elements in S_4 symmetry, which reduces to five independent values, considering that the ligand field matrix can be chosen to be traceless. The ability of the AOM parameters to reproduce a particular ligand field splitting is measured by the root mean square (RMS). The RMS is twofold as great for the AOM fit in the ferric compared to the ferrous complexes. This could stem from the fact that LFT is designed for ionic complexes. The larger metal-ligand covalency in case of the ferric complexes could render the application of LFT less effective. Furthermore, the RMS is about 20 % higher for the Se-based complexes. LFT assumes point charges for the ligands and this approximation could be less applicable for the larger Se-based ligands.

Table 2.5 also compares the AOM parameters from a plain CASSCF with the NEVPT2-corrected calculation. Focusing on the e_σ and e_π parameters in $[Fe(XH)_4]^{1-}$ ($[Fe(XH)_4]^{2-}$), the NEVPT2 correction leads to an increase of 10 %–20 % (5 %–8 %), which is consistent with the differences between the d orbital energies (see Table 2.4).

2.3. Dimeric FeS clusters: [Fe_2X_2] (X=S,Se)

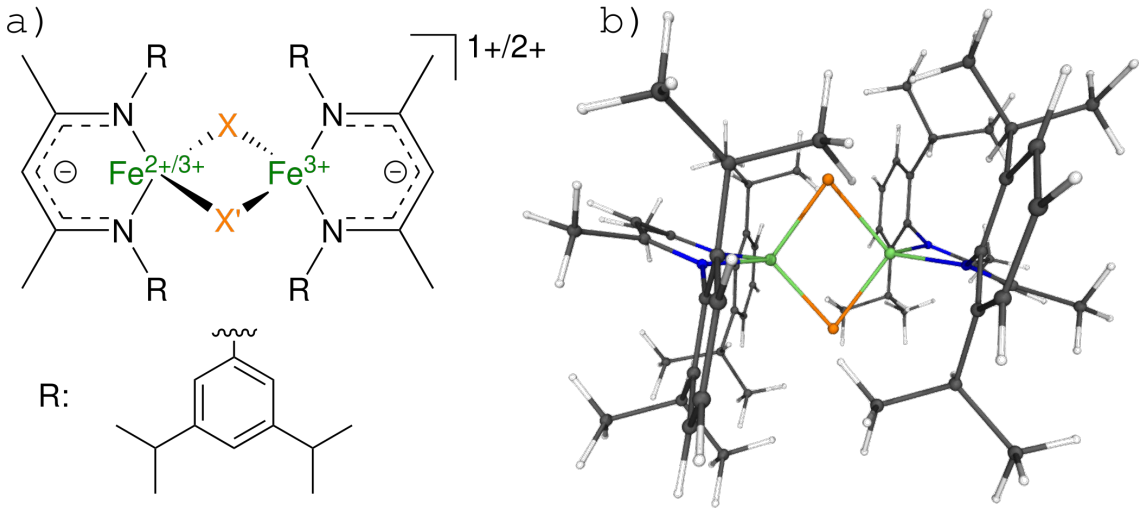

Figure 2.10. – (a) Schematic structure of the NacNac-capped [Fe_2X_2] clusters (X_2 = S_2, SSe, Se_2) and (b) 3D representation exemplary for $[Fe_2Se_2]^{2+}$. Except for the geometry optimization, the 2,6-diisopropylphenyl moieties (−R) were replaced with methyl groups ($-CH_3$).

The present Section studies S→Se substitution in [Fe_2X_2] dimers. Unless otherwise noted, the chemical formula [Fe_2X_2] refers to the β-ketiminato-capped ($NacNac^-$) FeS dimers in Figure 2.10. When substituting either one or both S in the bridge with Se, the clusters can be either labeled as homo-bridged, [Fe_2S_2] and [Fe_2Se_2], or hetero-bridged, [Fe_2SSe]. Because these ferredoxin-like FeS dimers typically occur as the oxidized $[Fe_2S_2]^{2+}$ and the reduced $[Fe_2S_2]^{1+}$, the discussion will focus on these two redox states. Both Fe centers are local high spins with Fe^{3+} oxidation states in the homo-valent $[Fe_2X_2]^{2+}$ and $Fe_A{}^{2+}/Fe_B{}^{3+}$ in the mixed-valent $[Fe_2X_2]^{1+}$, where the additional electron is localized on Fe_A in the ground state.

The geometries of the $[Fe_2X_2]^{2+,1+}$ series are compared in Section 2.3.1. The local electronic structure of either Fe center in the dimer is explored using the diamagnetically substituted clusters $[FeZnX_2]^{1+}$ and $[FeGaX_2]^{1+}$ in Section 2.3.2 and the discussion follows mostly that of the monomeric $[Fe(XH)_4]^{2-,1-}$ complexes in Section 2.2. In Section 2.3.3, the metal-metal interaction in the homo-valent $[Fe_2X_2]^{2+}$ dimers is explored. This includes an analysis of the CASSCF wave function for the antiferromagnetically coupled ground state ($S = 0$) as well as excited spin-coupled ($S = 1, \ldots 5$) and LMCT states. The calculations on the symmetrized $[Fe_2X_2]^{2+}$ clusters in Section 2.3.4 discusses the influence of symmetrizing the molecular geometry of the $[Fe_2X_2]^{2+}$ clusters. The effect of an additional electron on the electronic structure is studied for the mixed-valent $[Fe_2X_2]^{1+}$ clusters in Section 2.3.5. The additional electron is localized, yielding an antiferromagnetically coupled ground state ($S = \frac{1}{2}$), but it gives rise to a great number of low-lying excited states. Section 2.3.6 gives a brief account of the properties calculated with BS-DFT.

2.3.1. Geometry

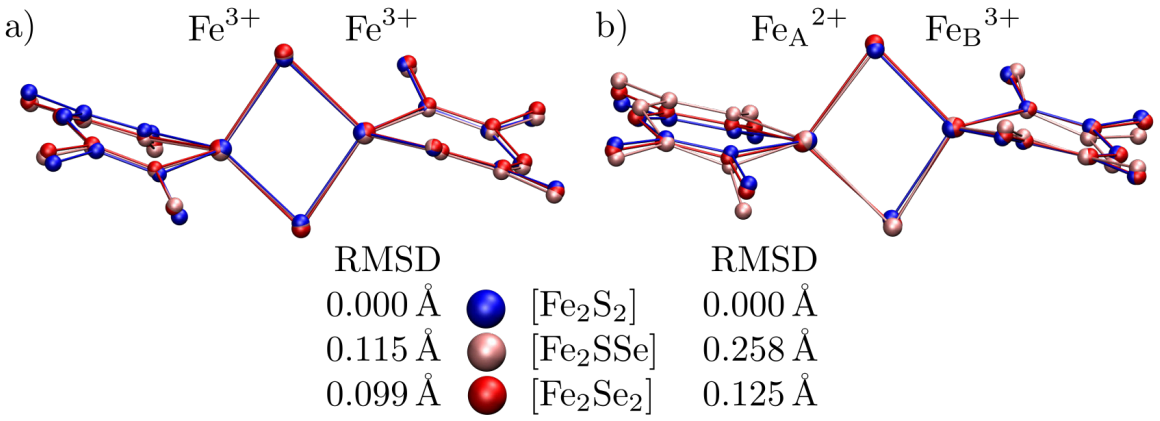

Figure 2.11. – Overlay of the BS-DFT-optimized structures for the NacNac-capped [Fe_2X_2] clusters in the redox states (a) $[Fe_2X_2]^{2+}$ ($M_S = 0$) and (b) $[Fe_2X_2]^{1+}$ ($M_S = \frac{1}{2}$). The root-mean-square deviation (RMSD) is shown relative to $[Fe_2S_2]^{2+,1+}$, respectively. The phenyl-based moieties and the hydrogen atoms are omitted for clarity (see Figure 2.10). Figure adapted from Ref [22].

The optimized structures for the $[Fe_2X_2]^{2+,1+}$ clusters are compared Figure 2.11. The overlaid structures all exhibit the same chair-like structure with the $NacNac^-$ ligands slightly bent out of the Fe_2N_4 plane. For the homo-valent $[Fe_2X_2]^{2+}$ clusters, both metal centers are in the same Fe^{3+} oxidation state, and they are related through symmetry in an approximate C_{2h} point group. For the mixed-valent $[Fe_2X_2]^{1+}$ clusters, the two metal centers are not equivalent. The metal centers are in the oxidation states $Fe_A{}^{2+}$ and $Fe_B{}^{3+}$, therefore, the approximate point group is C_s with the only symmetry element being a mirror plane through the Fe_2X_2 core. The localized electronic structure in $[Fe_2X_2]^{1+}$ is a direct consequence of the monodeterminantal ansatz in BS-DFT optimizations. A number of similar synthetic FeS dimers exhibit two distinct isomer shifts for the mixed-valent oxidation state in low-temperature Mössbauer spectra.[141,142,150–153] While this suggest a localized Fe^{2+} and Fe^{3+} oxidation states, one has to keep in mind that the DFT-optimized structure may be biased towards a localized geometry, which would affect all following calculations. The reader is referred to Section 3.2 for technical details using BS-DFT for geometry optimizations of mixed-valent dimers. The RMSD in Figure 2.11 measures the structural difference with respect to $[Fe_2S_2]^{2+,1+}$, respectively. It is evident, that the hetero bridge in $[Fe_2SSe]^{2+,1+}$ introduces a larger disorder than the homo bridge, especially for the mixed-valent clusters. This is expected to affect the interaction between the Fe centers.

The key bond lengths and angles for the $[Fe_2X_2]^{2+,1+}$ clusters are listed in Table 2.6. The metal oxidation state directly reflects in the geometric parameters. The FeFe distance is between 0.05 Å–0.11 Å longer in the more reduced $[Fe_2X_2]^{1+}$ clusters. The FeX distances are similar for $[Fe_2X_2]^{2+}$ and the $Fe_B{}^{3+}$ center in $[Fe_2X_2]^{1+}$ while the FeX distances involving $Fe_A{}^{2+}$ are about 0.1 Å longer. The dependence of the FeX bond length with the Fe oxidation state is stronger for $[Fe_2X_2]$ clusters compared to the $[Fe(XH)_4]$ complexes (around 0.06 Å, see Table 2.1) in line with higher ligand charge

Table 2.6. – Geometric parameters in the BS-DFT-optimized NacNac-capped $[Fe_2X_2]^{2+,1+}$ clusters. Distances are reported in Å, angles in °. The second value for the FeX distance in the $[Fe_2SSe]$ clusters refers to Se. See Table A3 for non-averaged values as well as for the $[Fe_2Te_2]^{2+,1+}$ clusters.

	homo-valent $[Fe_2X_2]^{2+}$				
	Fe−Fe	Fe−X	Fe−N	∠XFeX	∠NFeN
$[Fe_2S_2]^{2+}$	2.821	2.216	2.901	100.9	92.4
$[Fe_2SSe]^{2+}$	2.880	2.213 2.342	2.939	101.5	92.5
$[Fe_2Se_2]^{2+}$	2.932	2.338	2.971	102.3	93.1

	mixed-valent $[Fe_2X_2]^{1+}$						
	Fe−Fe	Fe_A−X	Fe_B−X	Fe−N	∠XFe_AX	∠XFe_BX	∠NFeN
$[Fe_2S_2]^{1+}$	2.875	2.301	2.204	2.964	97.8	103.7	90.9
$[Fe_2SSe]^{1+}$	2.933	2.298 2.438	2.188 2.340	2.998	98.2	104.5	92.0
$[Fe_2Se_2]^{1+}$	3.042	2.447	2.330	3.047	97.5	104.3	91.6

on X^{2-} compared to XH^-. The average FeN bond lengths is 0.05 Å–0.07 Å longer after the reduction from $[Fe_2X_2]^{2+}$ to $[Fe_2X_2]^{1+}$.

The substitution S→Se leads to an increase in all bond lengths. The FeFe distance in $[Fe_2X_2]^{2+}$ increases by 0.05 Å upon each substitution, which is consistent with the X-ray structures of similar $[Fe_2X_2]$ clusters.[180,285] For the one-electron reduced $[Fe_2X_2]^{1+}$, the FeFe distance increases by 0.06 Å for the first and by 0.11 Å for the second substitution. The FeS bonds (2.21 Å–2.30 Å) are consistently about 0.12 Å shorter than the FeSe bonds (2.33 Å–2.45 Å). This trend matches well with the $[Fe(XH)_4]$ complexes, but the FeX bond length in $[Fe(XH)_4]^{2-}$ are slightly longer compared to those observed in $[Fe_2X_2]^{1+}$ (+0.02 Å, see Table 2.1). The FeN bond lengths increase by about 0.04 Å upon S→Se substitution, which is about half of the change due to reduction and may therefore indicate more electron-rich Fe centers for the Se-based clusters. The XFeX angle increases by 0.5° $[Fe_2X_2]^{2+}$ when replacing S with Se, in line with the larger atomic radius of Se. However, for $[Fe_2X_2]^{1+}$ the angle increases upon the first, but decreases upon the second substitution. The NFeN angle follows the similar trends as the XFeX angle. The non-averaged geometric parameters as well as those for $[Fe_2Te_2]^{2+,1+}$ (not discussed) are give in Table A3.

2.3.2. Diamagnetic substitution: $[FeGaX_2]$ and $[FeZnX_2]$ clusters

In going from the $[Fe(XH)_4]$ complexes in Section 2.2 to the $[Fe_2X_2]$ clusters, the complexity of the electronic structure increases sharply. One can ease this transition by taking an intermediate step via the diamagnetically substituted clusters $[FeGaX_2]^{1+}$ and $[FeZnX_2]^{1+}$. Here, either the $Fe_A{}^{2+}$ or the $Fe_B{}^{3+}$ center in the mixed-valent $[Fe_2X_2]^{1+}$ clusters is substituted by the diamagnetic ion Ga^{3+} or Zn^{2+}, respectively. Since the diamagnetic ions have a d^{10} electron configuration, they do not engage in spin-spin

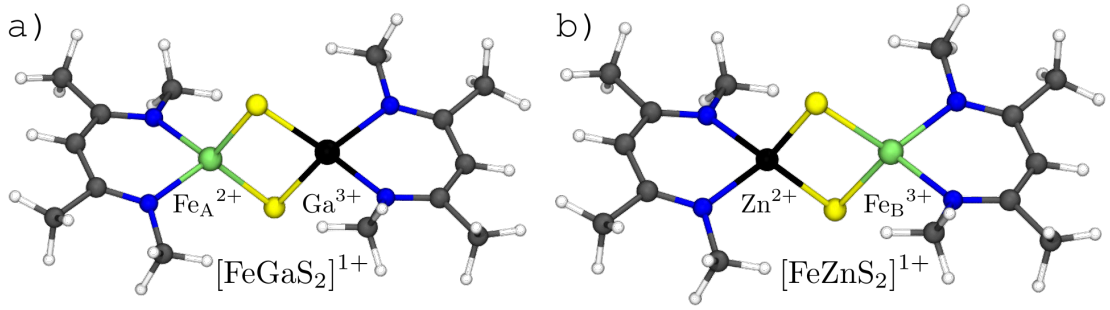

Figure 2.12. – Structures of the diamagnetically substituted $[Fe_2S_2]^{1+}$ clusters: (a) $[FeGaS_2]^{1+}$ and (b) $[FeZnS_2]^{1+}$. Ga^{3+} and Zn^{2+} are shown in black. The SSe- and Se_2-bridged clusters have equivalent structures.

interactions with the remaining open-shell Fe center and the same analysis used in the $[Fe(XH)_4]$ complexes can be performed. The crystal radii of the diamagnetic ions differ by less than 0.03 Å from those of the respective Fe ion.[159] Therefore, substitution is expected to maintain the electrostatic influence of the second metal center. The results of the diamagnetic substitution in the homo-valent clusters ($[Fe_2X_2]^{2+}$→$[FeGaX_2]^{2+}$) are similar to $[FeZnX_2]^{1+}$, since the local geometry of the Fe^{3+} centers in $[Fe_2X_2]^{2+}$ is very close to that of Fe_B^{3+} in $[Fe_2X_2]^{1+}$ (see Section 2.3.1). Therefore, only the diamagnetic substitution in the mixed-valent clusters is discussed here. As a side note, the reader is referred to Section 4.2.2 and Figure A12 for a more refined approach to choose the diamagnetic substituent, where the same technique is applied to study the electronic structure of FeMoco.

Table 2.7. – AILFT parameters for the $[FeZnX_2]^{1+}$ and $[FeGaX_2]^{1+}$ clusters and AOM parameters for the X^{2-} bridging ligands (CAS(5,5) and CAS(6,5), respectively, including NEVPT2). The d orbital energies and the AOM parameters are given in [cm^{-1}]; rB, rC, and $r\zeta$ are given in [%]. See Table A1 for B, C, and ζ in the free $Fe^{3+,2+}$ ions. AOM parameters for the $NacNac^-$ ligands are determined in a separate calculation (see Table A4).

	ferric $[FeZnX_2]^{1+}$										
	z^2	x^2-y^2	xy	xz	yz	rB	rC	rζ	e_σ	$e_{\pi s}$	$e_{\pi c}$
$[FeZnS_2]^{1+}$	0	1730	6086	9141	10784	21.8	22.4	12.6	7944	2504	1979
$[FeZnSSe]^{1+}$	0	1152	5508	8564	9238	23.0	22.6	15.0			
$[FeZnSe_2]^{1+}$	0	1371	5544	9229	9719	22.6	22.6	17.2	7582	1538	1350
	ferrous $[FeGaX_2]^{1+}$										
	z^2	x^2-y^2	xy	xz	yz	rB	rC	rζ	e_σ	$e_{\pi s}$	$e_{\pi c}$
$[FeGaS_2]^{1+}$	0	392	1998	4872	6779	12.6	10.0	5.4	5169	2228	1551
$[FeGaSSe]^{1+}$	0	505	1849	4112	7035	12.0	9.9	6.0			
$[FeGaSe_2]^{1+}$	0	88	1605	4159	6776	12.3	9.7	6.6	4382	1646	1257

Even though the Fe centers in $[Fe(XH)_4]$ and $[Fe_2X_2]$ are both coordinated by four ligands, the coordination environment is different in two major ways: First, the XFeXH

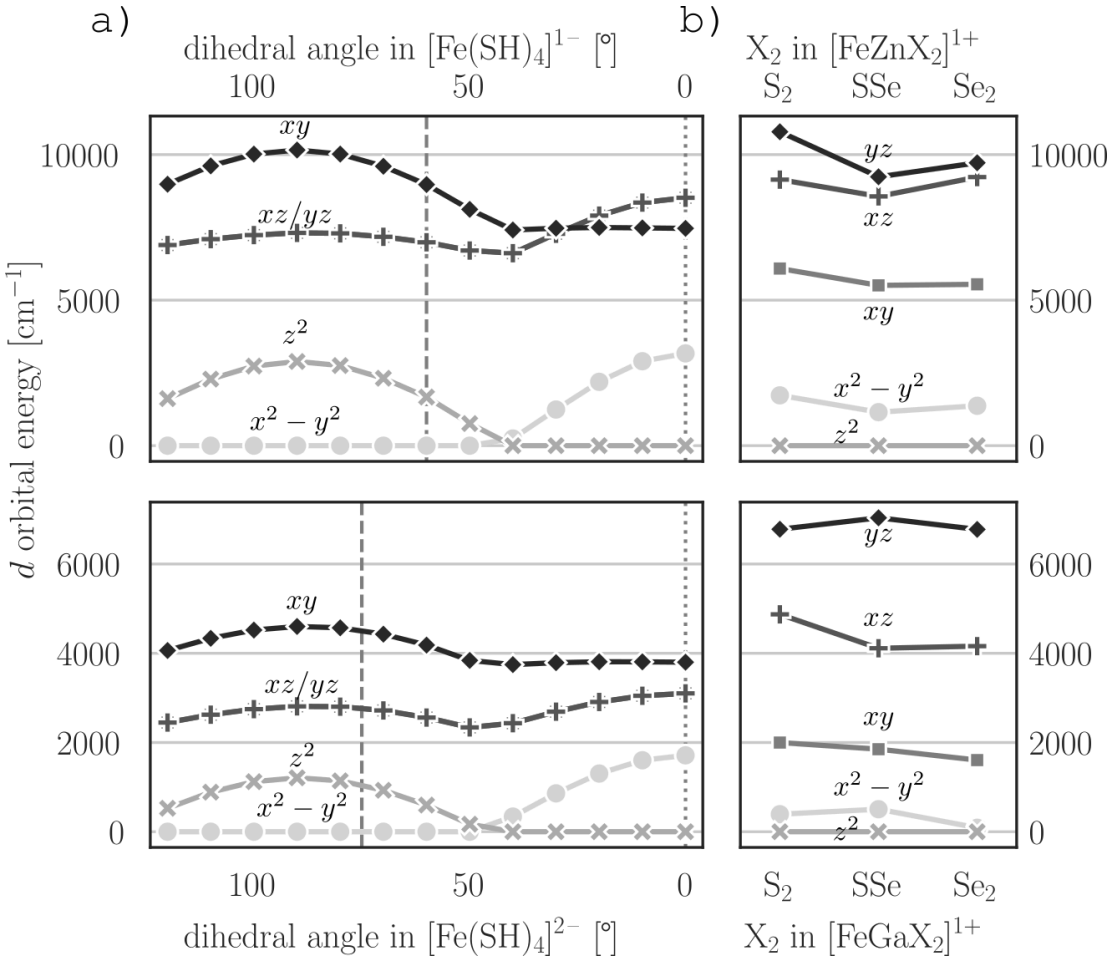

Figure 2.13. – (a) AILFT d orbital energies as a function of the SFeSH dihedral angle in $[Fe(SH)_4]^{1-}$ (top) and $[Fe(SH)_4]^{2-}$ (bottom). The dashed line marks the XFeXH dihedral angle in the DFT-optimized geometry of $[Fe(XH)_4]^{1-,2-}$. The dotted line marks the XFe_AXFe_B dihedral angle in the $[Fe_2X_2]$ clusters. (b) AILFT d orbital energies for different bridging ligands in $[FeGaX_2]^{1+}$ (top) and $[FeZnX_2]^{1+}$ (bottom). All values are based on CAS(5,5) calculations (top) or CAS(6,5) calculations (bottom) including NEVPT2. Figure adapted from Ref [22].

dihedral angle in the optimized $[Fe(XH)_4]$ geometries is 60°–75°, while the diamond core structure in $[Fe_2X_2]$ fixes the XFeXFe dihedral angle to 0°. Second, the capping $NacNac^-$ ligands are expected to create a different ligand field compared to the bridging X^{2-} ligands and therefore reducing the local symmetry at the Fe center from axial in $[Fe(XH)_4]$ to rhombic $[Fe_2X_2]$. In order to explore the effect of the dihedral angle, the AILFT d orbital energies in $[Fe(SH)_4]^{1-,2-}$ are plotted against the dihedral angle in Figure 2.13a. Analogous calculations for $[Fe(SeH)_4]^{1-,2-}$ result in similar plots and are therefore not discussed. At a dihedral angle of 0° the XH^- ligands point towards the z-axis and the complex assumes D_{2h} symmetry. At this geometry, the z^2 orbital becomes the lowest in energy and the energies of the xz and yz orbitals increase relative to xy.

The energies of the AILFT d orbitals in the diamagnetically substituted clusters $[FeZnX_2]^{1+}$ and $[FeGaX_2]^{1+}$ are shown in Figure 2.13b and also listed in Table 2.7. The active orbitals of the underlying CASSCF calculation are characterized in Table A2. The orbital energies cover about 9000 cm^{-1} in $[FeZnX_2]^{1+}$, which is comparable to $[Fe(XH)_4]^{1-}$, and about 7000 cm^{-1} in $[FeGaX_2]^{1+}$, which is 1500 cm^{-1} higher compared to $[Fe(XH)_4]^{2-}$. In the diamagnetically substituted dimers, the z-axis is oriented along the FeZn/FeGa bond, the bridging X^{2-} ligands lie in the xz plane, and the capping $NacNac^-$ ligands in the yz plane. Therefore, the local symmetry of Fe is approximately C_{2v}. As in the D_{2h} symmetric $[Fe(XH)_4]$ complexes, the z^2 orbital is lowest in energy. The energies of the xz and yz orbitals are more than 2000 cm^{-1} higher than the xy orbital. In the C_{2v} local symmetry, the degeneracy of the xz and yz orbitals is lifted. The xz and yz orbitals are separated by 500 cm^{-1}–1500 cm^{-1} in $[FeZnX_2]^{1+}$ and 1500 cm^{-1}–2000 cm^{-1} in $[FeGaX_2]^{1+}$. The energy of the xz orbital is lower than the yz orbital suggesting that the X^{2-} ligands are weaker donors than the $NacNac^-$ ligands. As shown in Table 2.7, the reduction of the Racah parameters B relative to the free ion is with 1 %–5 % weaker for $[FeZnX_2]^{1+}$ and $[FeGaX_2]^{1+}$ compared to $[Fe(XH)_4]^{1-,2-}$, while at the same time the reduction of C and ζ are 0 %–1 % and 1 %–6 % weaker, respectively.

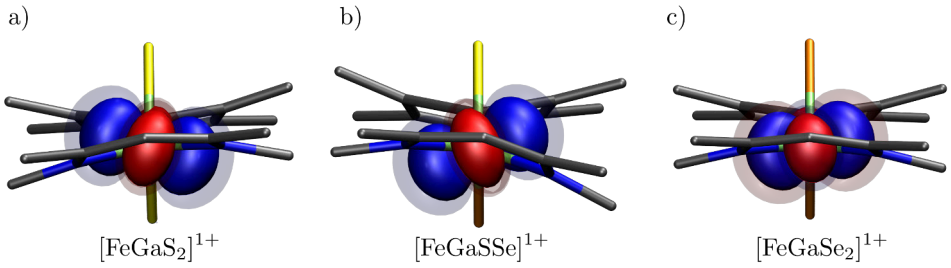

Figure 2.14. – Orientation of the z^2 orbital in the ground state of (a) $[FeGaS_2]^{1+}$, (b) $[FeGaSSe]^{1+}$, and (c) $[FeGaSe_2]^{1+}$. Isosurfaces are shown at ±0.05 and ±0.025. Figure adapted from Ref [22].

Most of the effects of the S→Se substitution in $[FeZnX_2]^{1+}$ and $[FeGaX_2]^{1+}$ are similar to $[Fe(XH)_4]^{1-,2-}$, respectively, though less pronounced, when disregarding the somewhat more distorted SSe-bridged molecule. Analogous to $[Fe(XH)_4]$, the S→Se substitution increases the reduction of B for the ferric $[FeZnX_2]^{1+}$, but leads to a slight decrease in the ferrous $[FeGaX_2]^{1+}$. The reduction of ζ is also significantly stronger in the Se-based complexes. The barycenter between the e set and the t_2 set decreases by about 300 cm^{-1} when substituting S with Se in $[FeZnS_2]^{1+}$ and by 200 cm^{-1} in $[FeGaS_2]^{1+}$. In contrast to $[Fe(XH)_4]$, the S→Se substitution decreases the energy gap between the z^2 and the $x^2 - y^2$ orbital substantially. In the ferrous case, the gap decreases from 392 cm^{-1} in $[FeGaS_2]^{1+}$ to merely 88 cm^{-1} in $[FeGaSe_2]^{1+}$. This decrease is partly a consequence of the lower donor strength of Se^{2-} compared to S^{2-} (*vide infra*), but may also be related to the geometry of the complexes. Figure 2.14 shows the z^2 orbital in the $[FeGaX_2]^{1+}$ series. The orientation of the z^2 orbital depends on the orientation of the $NacNac^-$ ligand and the parallel alignment of the two $NacNac^-$ ligands in $[FeGaSe_2]^{1+}$ (Figure 2.14c) may favor the near-degeneracy of the z^2 and the $x^2 - y^2$ orbital. Within LFT for a d^6 metal center, the lowest LFT d orbital is

doubly occupied in the ground state. Therefore, it may be possible to modulate the nature of the electronic structure in the ground state in mixed-valent $[Fe_2X_2]^{1+}$ by an appropriate choice of the bridging ligand and the geometry of the capping ligand. In biological $[Fe_2S_2]^{1+}$ clusters, the protein environment constrains the ligand geometry and the dihedral angle covers values from 0°–120°.[286]

The extraction of the AOM parameters from the AILFT calculation requires an additional step for the diamagnetically substituted complexes, because the fit is underdetermined with the two types of ligands present (bridging X^{2-} and capping $NacNac^-$). Therefore, the AOM parameters for the $NacNac^-$ ligands are extracted in a separate calculation, which is explained in Section A2.2. The AOM parameters for the bridging X^{2-} are given in Table 2.7. The misdirected valence parameters $e_{\sigma\pi}$ is not necessary here, because the XGa/XZn bonding interaction is weak compared to the XH bond in $[Fe(XH)_4]$. Compared to the monoanionic XH^- ligands (Table 2.5), the e_σ and $e_{\pi c}$ parameters are 1.5 to 3 times larger in the dianionic X^{2-} ligands. Even though the misdirected valence parameter is not included, the presence of the cationic Zn^{2+} and Ga^{3+} manifests in the inequality of the $e_{\pi s}$ (perpendicular to the FeX_2 plane) and $e_{\pi c}$ (within the FeX_2 plane) parameters. The $e_{\pi c}$ parameter is 10 %–45 % smaller compared to $e_{\pi s}$, because electron density is drawn to the positively charged second metal center. Similar to $[Fe(XH)_4]$, the e_σ parameter decreases by 5 %–15 % when substituting S with Se, but the e_π parameters also decrease by 20 %–40 %, which is not observed in $[Fe(XH)_4]$. The $e_{\pi s}/e_{\pi c}$ anisotropy is about 15 % less pronounced for Se compared to S, in line with the larger Se acting as a more potent buffer for the electron-withdrawing effects of the Zn^{2+}/Ga^{3+} centers.

2.3.3. Homo-valent dimers: $[Fe_2X_2]^{2+}$ clusters

The active space for the CAS(22,16) wave function of the $[Fe_2X_2]^{2+}$ clusters is shown in Figure 2.15a. The active orbitals were localized (see Section 2.1 for details) making it straightforward to assign them to either Fe center or the bridging ligands X. The CASSCF wave function can be expressed as a linear combination of multiple electron configurations (CFGs). In the so-called neutral CFG, each Fe-based orbital is doubly occupied and each X-based orbital doubly occupied, in line with the oxidation states Fe^{3+} and X^{2-}. A CFG where an electron is paired locally on either Fe center is called *d-d* non-Hund (nH). Transferring an electron from the X-based orbitals to either Fe center yields a ligand-to-metal charge transfer (LMCT) CFG, transferring it between the Fe centers yields a metal-to-metal charge transfer (MMCT) CFG. Note that CFGs only describe the occupation number and therefore do not distinguish between microstates with local high spins on the Fe centers and spin-flip *d-d* excitations.

The Fe-based orbitals are shown in Figure 2.15b. As for the $[Fe(XH)_4]$ complexes, the orbitals in the t_2 set have a stronger interaction with the ligands (see Figure 2.4). The z-axis lies along the Fe−Fe vector, the x-axis is parallel to the X−X vector, and the y-axis parallal to the $NacNac^-$ ligands. The full set of active orbitals is shown in Figure A5.

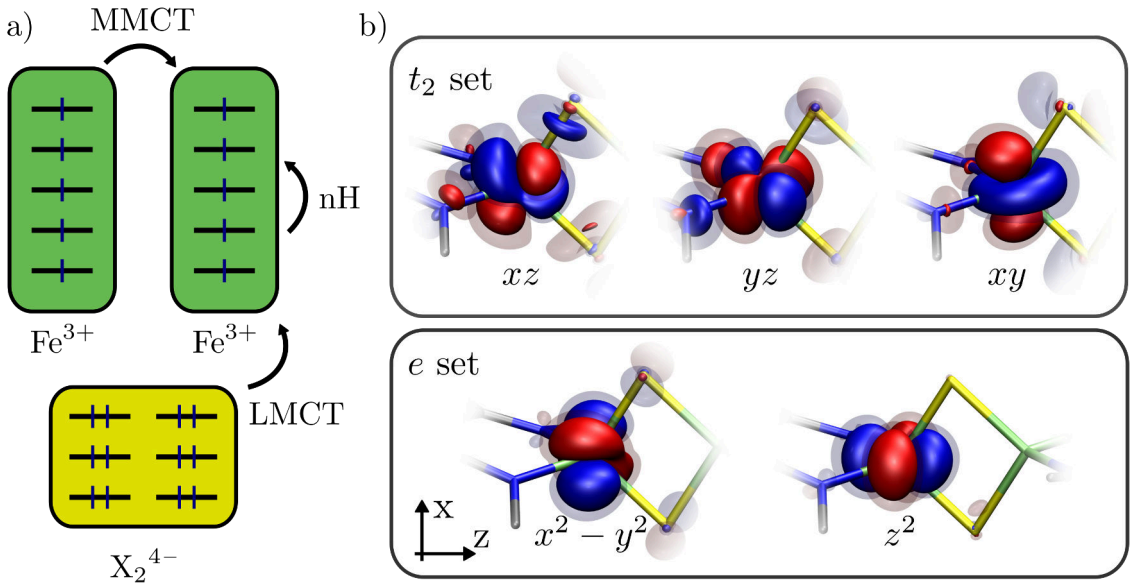

Figure 2.15. – (a) Schematic representations of the CFGs in the CAS(22,16) wave function of the homo-valent $[Fe_2X_2]^{2+}$ clusters. In the neutral CFG, all metal-based orbitals are singly occupied and all ligand-based orbital doubly occupied. (b) Plots of the five active orbitals on Fe_A with characters corresponding to pseudo-T_d symmetry. Orbitals on Fe_B are mirror images. The full set of all 16 active orbitals is shown in Figure A5. Isosurfaces are shown at ±0.05 and ±0.025.

Composition of the active orbitals in $[Fe_2X_2]^{2+}$

The S→Se substitution in $[Fe_2X_2]^{2+}$ does not change the overall shape of the active orbitals, but it does affect their composition slightly. The atomic contributions to the mostly Fe-based active orbitals are shown for the series $[Fe_2X_2]^{2+}$ in Figure 2.16. Compared to the diamagnetically substituted clusters $[FeZnX_2]^{1+}$, the weight of the ligand is 1 %–4 % smaller, *i.e.* the orbital is less covalent, which is in line with each bridging ligand forming two equally covalent FeX bonds in $[Fe_2X_2]^{2+}$ (see Table A2 for $[FeZnX_2]^{1+}$ and Table A5 for $[Fe_2X_2]^{2+}$). The largest differences are between the hetero-bridged $[Fe_2SSe]^{2+}$ cluster and the homo-bridged $[Fe_2S_2]^{2+}$ and $[Fe_2Se_2]^{2+}$ clusters. In the hetero-bridge cluster, the z^2 and the yz orbitals are 1 % and 0.3 % more covalent compared to the homo-bridged clusters, while at the same time the $x^2 - y^2$ and the xz orbitals are less covalent by the same amount, respectively. The $x^2 - y^2$ and the xz orbitals are the most covalent out of the e set and the t_2 set, respectively, but with two different bridging ligands, these orbitals tend to be more ionic. When directly comparing the homo-bridged $[Fe_2S_2]^{2+}$ with $[Fe_2Se_2]^{2+}$, the Se substitution does not lead to clear trends in the orbital composition and the Fe and the X contributions change by 0.0 %–0.3 %.

Composition of the wave function in $[Fe_2X_2]^{2+}$

The CASSCF wave function can be characterized based on the CFGs it contains. As explained in Figure 2.15, the CAS(22,16) wave function can have contributions from

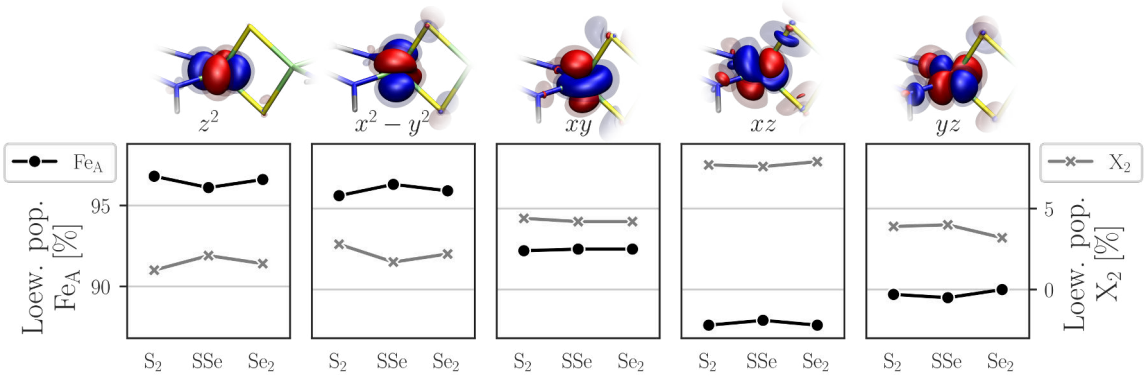

Figure 2.16. – Loewdin atomic contributions of Fe (left axis) and X_2 (right axis) to the metal-based active orbitals for $[Fe_2S_2]^{2+}$, $[Fe_2SSe]^{2+}$, and $[Fe_2Se_2]^{2+}$. Values for the Fe_A-based orbitals are shown; Fe_B-based orbitals are equivalent. See Table A5 for all active orbitals in the CAS(22,16) wave function, as well as $NacNac^-$ contributions. Isosurfaces are shown at ±0.05 and ±0.025.

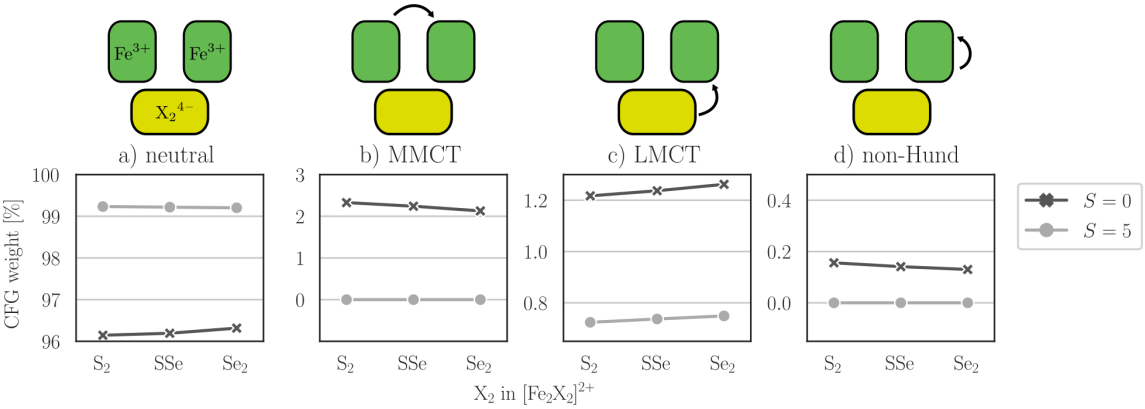

Figure 2.17. – (a-d) CFGs in the $S = 0$ and $S = 5$ CAS(22,16) wave functions for the clusters $[Fe_2S_2]^{2+}$, $[Fe_2SSe]^{2+}$, and $[Fe_2Se_2]^{2+}$. See Table A6 for values.

the neutral CFG, which corresponds to the oxidation states Fe^{3+} and X^{2-}, as well as MMCT, nH, and LMCT CFGs. In the ground state of the homo-valent $[Fe_2X_2]^{2+}$ clusters, the local high spin Fe^{3+} centers are coupled antiferromagnetically to an $S = 0$ singlet state. Figure 2.17 shows the composition of the ground state wave function as a function of the S→Se substitution. The neutral CFG has a weight of > 96 %, the MMCT CFGs contribute around 2 %, the LMCT CFGs around 1 %, and the nH CFG around 0.2 %. When substituting S with Se, the weight of the neutral CFG increases by about 0.2 % while the weight of the MMCT CFGs decreases similarly. The decrease in the MMCT CFGs correlates with the increase in the Fe−Fe distance by 0.1 Å (see Table 2.6), which makes the transfer of an electron between the Fe centers less favorable. The amount of MMCT CFGs is also related to the coupling strength between the Fe centers and is expected to affect the energy splitting between the spin states (*vide infra*).[109]

The weight of the LMCT CFGs is about 0.04 % higher in $[Fe_2Se_2]^{2+}$ compared

to $[Fe_2S_2]^{2+}$. This is consistent with the IEs of $[Fe(XH)_4]^{1-}$ (see Figure 2.3). Here, the removal of an SeH^--based electron requires 0.4 eV less energy compared to SH^- giving ready explanation for the higher LMCT CFGs weight in $[Fe_2Se_2]^{2+}$. The nH CFGs decrease by 0.03 % in going from the S- to the Se-based clusters. According to the AILFT analysis in the diamagnetically substituted clusters $[FeZnX_2]^{1+}$, the Se^{2-} ligands cause a smaller ligand field splitting than the S^{2-} ligands (see Figure 2.13), which would make any *d-d* excitations more favorable and actually suggesting an increase in the nH CFGs for the S→Se substitution. However, the decrease in nH CFGs also coincides with the increase in the LMCT CFGs, and the transfer of an electron from X to Fe could hinder any further local excitations.

Figure 2.17 also shows the composition of the wave function for the $S = 5$ spin state, where the Fe centers are ferromagnetically coupled. The energy of this spin state is several thousand cm^{-1} above the ground state (*vide infra*). The weight of the MMCT and nH CFGs is exactly zero for this state, since any electron pairing on the Fe centers would decrease the global spin, which is not possible in for an $S = 5$ spin state. Therefore, the change in the weight of the neutral CFG is directly coupled to the change in the weight of the LMCT CFGs, which is virtually identical to the $S = 0$ ground state.

State-averaged vs. state-specific energies in $[Fe_2X_2]^{2+}$

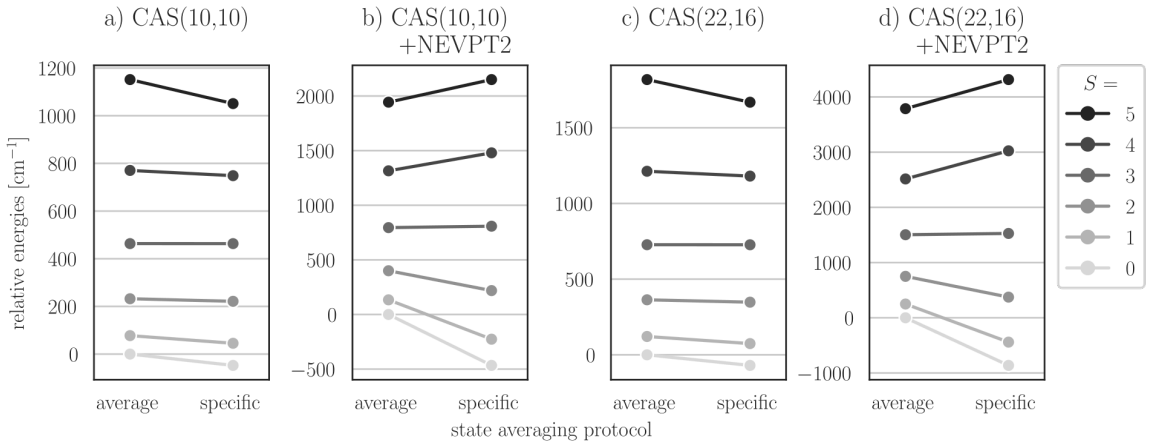

Figure 2.18. – State-averaged and state-specific energies of the spin-coupled states in $[Fe_2S_2]^{2+}$ relative to the state-averaged ground state: (a) CAS(10,10) energies, (b) NEVPT2-corrected CAS(10,10) energies, (c) CAS(22,16) energies, and (d) NEVPT2-corrected CAS(22,16) energies. The wave function was either optimized with an equal weight on all states (average) or for each state individually (specific). Values are listed in Table A7.

When calculating energy differences with CASSCF, one typically optimizes the wave function placing an equal weight on all states of interest. For example, the lowest-energy states in the $[Fe_2X_2]^{2+}$ clusters are the six spin-coupled states, $S = 0, \ldots 5$, and each state would enter the optimization of the CI coefficients and the molecular orbital coefficients (see Section 1.1.1) with a weight of $\frac{1}{6}$. In this way, one avoids bias of the

wave function towards any of the states and also limits the computational expense to a single CASSCF optimization. However, the quality of the CASSCF wave function may vary for different states. For example, the movement of electrons with parallel spin is correlated in CASSCF and different spin states therefore include different amounts of electrons correlation. Post-CASSCF corrections such as NEVPT2 may not be able to sufficiently recover the missing electron correlation leading to unbalanced state energies (*vide infra*). Alternatively, one can optimize each state of interest separately to obtain a distinct set of coefficients for each state. While this state specific protocol appears more rigorous, since it avoids any bias in the wave function, it may be problematic in the following cases. (i) The single optimized state may lead to a nearly doubly or nearly empty active orbital, which impedes the convergence of the CASSCF. For example, the occupation number of the $x^2 - y^2$ orbital in the ground state of $[Fe(XH)_4]^{2-}$ (see Figure 2.4) is virtually 2.0, but when averaging over the set of $S = 2$ states assures smooth convergence, since each active orbital has an occupation of 1.2 (Figure 2.6). (ii) When optimizing the wave function for an excited state, the orbitals become less suited for lower-lying states, therefore increasing their energies. This may lead to a reordering of states and the CASSCF algorithm may alternate between optimizing different states in two subsequent iterations, preventing convergence. This effect is further exploited in the $[Fe_2X_2]^{1+}$ clusters to converge a CASSCF wave function to CT excited states (see Figure 2.30).

The state-averaged and state-specific energies are compared for $[Fe_2S_2]^{2+}$ in Figure 2.18. Focusing on the CAS(10,10) energies (Figure 2.18a), the state-specific energies tend to be lower than the state-averaged ones. This is not surprising, because each state-specific wave function is more flexible than the state-averaged ones. However, the individual states are affected differently. The extreme states, *i.e.* $S = 0$ and 5, shift more strongly than $S = 1$ and 4, and the energies for $S = 2$ and 3 are virtually unchanged (see Table A7 for energy values). The orbitals in the state-averaged CASSCF wave function appear to be well-suited for the intermediate states $S = 2$ and 3. However, they are less well-suited for the states $S = 0$ and 5 Therefore, averaging over all $S = 0, \ldots 5$ states should be avoided when comparing to experimental observables, since typically only the lowest states are probed (the calculated coupling constants J is discussed later in the text).

Interestingly, the NEVPT2-corrected state energies (Figure 2.18b) show a different behavior than the CASSCF states. In going from state-averaged to state-specific energies, the states $S = 0$ and 1 still decrease in energy, but the decrease is more pronounced compared to the CASSCF states. The state-specific energies of the states $S = 4$ and 5, however, are higher compared to the state-averaged energies even though the reference states in Figure 2.18a show the opposite trend. In other words, the magnitude of the NEVPT2 correction increases for the $S = 0$ and 1 state in going from a state-averaged to a state-specific wave function, but it decreases at the same time for the $S = 4$ and 5 state. Ideally, a sufficiently accurate post-CASSCF correction would yield the same absolute energy for any state, regardless of using a state-averaged or state-specific reference wave function. Figure 2.18b shows that this is not the case for the NEVPT2-corrected energies. However, since NEVPT2 is a perturbative method, the lowest energy is not necessarily the best energy.

Figure 2.18c and d are the analogous plots to a and b using the larger active

space CAS(22,16). Apart from the total energy range, the trends are equivalent to the CAS(10,10) calculations. Therefore, the differences between state-averaged and state-specific energies do not depend on X-based orbitals in the active space.

Effect of S→Se substitution on the state energies in $[Fe_2X_2]^{2+}$

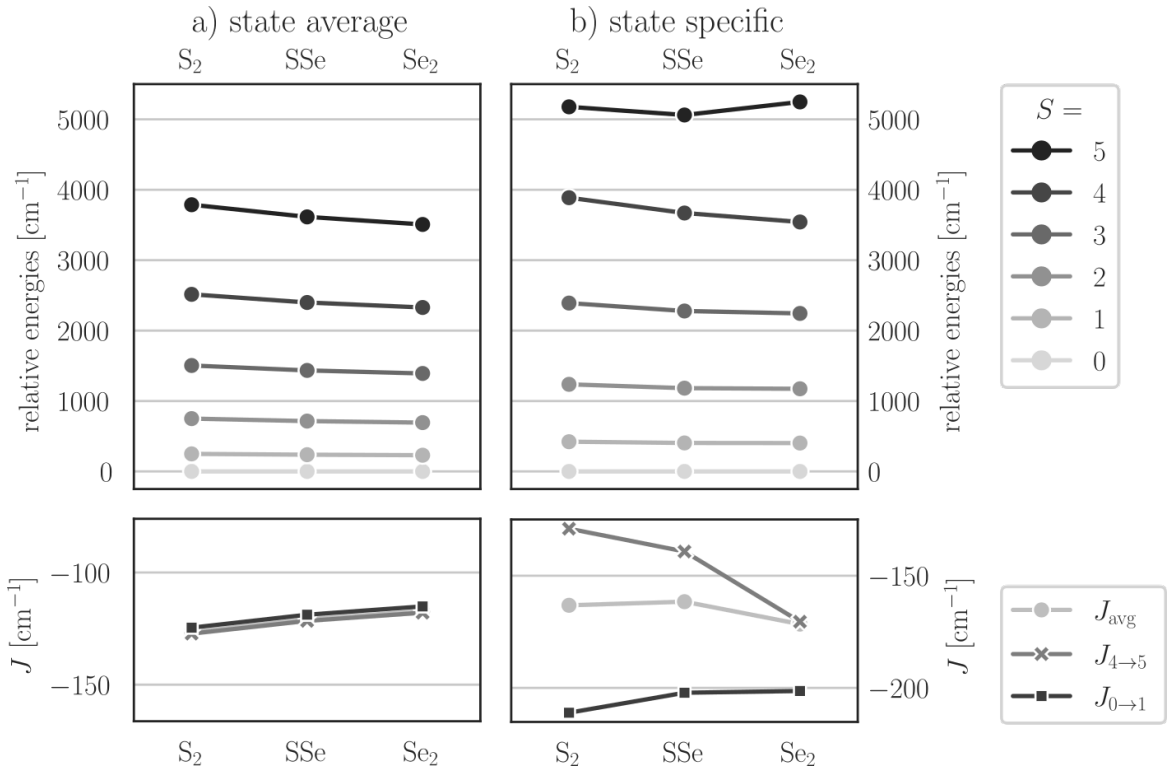

Figure 2.19. – Spin-coupled states in the $[Fe_2S_2]^{2+}$, $[Fe_2SSe]^{2+}$, and $[Fe_2Se_2]^{2+}$ (top) according to CAS(22,16)/NEVPT2 and coupling constants (bottom) fitted to one state pair ($J_{i\to j}$) or all state energies (J_{avg}). The energies are shown for (a) a state-averaged wave function and (b) a state-specific wave function. See Table A7 for values.

The dependence of the energy splitting between the states on the bridging ligand is shown in the top half of Figure 2.19 for $[Fe_2X_2]^{2+}$. With state-averaged energies (Figure 2.19a), the relative energies of all six spin states are reduced consistently by 3 %–5 % with each S→Se substitution, respectively, and about 8 % when comparing $[Fe_2S_2]^{2+}$ directly with $[Fe_2Se_2]^{2+}$. However, with state-specific energies (Figure 2.19b), the substitutions result in different trends for different spins. When comparing $[Fe_2S_2]^{2+}$ with $[Fe_2Se_2]^{2+}$, the $S = 4$ state is most strongly affected and its energy relative to the $S = 0$ ground state is reduced by 10 %. The reduction in the energies of the $S = 3, 2$, and 1 states decreases steadily in that order to 5 % for $S = 1$. In contrast, the energy of the $S = 5$ state relative to $S = 0$ is actually slightly higher (+1 %) in the Se-based cluster. Therefore, the effect of S→Se substitution on the state energies is fundamentally different when state-averaged or state-specific energies are considered.

Coupling constants J in $[Fe_2X_2]^{2+}$

The spin-coupled states that are a result of the two local high spin Fe^{3+} centers in the homo-valent $[Fe_2X_2]^{2+}$ clusters are typically modeled with the HDvV Hamiltonian (Eq 1.8). The spacing between the spin states is given by the simple Landé interval rule $\Delta E_{S-1,S} = 2JS$ (Eq 1.9). The bottom half of Figure 2.19 shows the coupling constants J extracted from the energy differences shown in the upper half via different protocols: using the energy difference between $S = 0$ and 1 ($J_{0\to1}$), between $S = 4$ and 5 ($J_{4\to5}$), and as the average from all five adjacent pairs (J_{avg}). If the HDvV Hamiltonian captured all the physics in the $[Fe_2X_2]^{2+}$ clusters, the Landé interval rule would yield the same J for any adjacent spin-coupled state pair.

The coupling constants extracted from state-averaged energies (Figure 2.19a) range from $-115\,\text{cm}^{-1}$–$-130\,\text{cm}^{-1}$ for the different bridging ligands in $[Fe_2X_2]^{2+}$. Within each each $[Fe_2X_2]^{2+}$ cluster, $J_{0\to1}$, $J_{4\to5}$, and J_{avg} differ by $< 3\,\text{cm}^{-1}$. With state-specific energies (Figure 2.19b), the coupling constants range from $-130\,\text{cm}^{-1}$–$-200\,\text{cm}^{-1}$ and therefore cover a significantly larger energy range compared to those extracted from state-averaged energies. For the $[Fe_2S_2]^{2+}$ cluster, the magnitude of J_{avg} is about $40\,\text{cm}^{-1}$ higher when using state-specific compared to state-averaged energies, which corresponds to an increase of about 30 %. The coupling constant J_{avg} ($-163\,\text{cm}^{-1}$) lies in between $J_{4\to5}$ ($-129\,\text{cm}^{-1}$) and $J_{0\to1}$ ($-211\,\text{cm}^{-1}$). It was shown in Figure 2.18a and c that the energies of the extreme spin states are reduced more strongly than the intermediate spins in going from state-averaged to state-specific energies. Since the state-specific energies for $S = 4$ and 5 move towards the energy of the average energy, but $S = 0$ and 1 move away, the coupling constant $J_{4\to5}$ decreases while $J_{0\to1}$ increases relative to J_{avg}. This is also true for the CASSCF energies (see Table A7) and hence not a result of the NEVPT2 correction.

With state-averaged energies (Figure 2.19a), all three coupling constants, J_{avg}, $J_{4\to5}$, and $J_{0\to1}$, decrease by about 5 % upon each S→Se substitution in $[Fe_2X_2]^{2+}$. However, with state-specific energies (Figure 2.19b), The absolute value of $J_{4\to5}$ increases by $10\,\text{cm}^{-1}$ and by $30\,\text{cm}^{-1}$ upon each Se substitution, respectively. The absolute value of $J_{0\to1}$, on the other hand, decreases by $10\,\text{cm}^{-1}$ and $1\,\text{cm}^{-1}$ at the same time. As discussed above, the weight of the MMCT CFGs decreases in the $S = 0$ wave function upon S→Se substitution. (see Figure 2.17), which decreases in the antiferromagnetic coupling strength.[109] The LMCT CFGs are also related to the coupling strength, but they enter at a higher order than the MMCT CFGs. Therefore, the changes in the MMCT outweigh the changes in the LMCT CFGs. J_{avg} shows an average behavior and increases slightly upon the first and decreases on the second substitution. Since $J_{4\to5}$ and $J_{0\to1}$ are more similar in the Se-based clusters, they appear to follow the HDvV Hamiltonian better compared to the S-based clusters.

The wide range coupling constants extracted via different protocols from one geometry raises the question which one is suited best for the comparison with experimental data. The experimental parameter J_{exp} is typically fitted to data collected at temperatures $< 300\,\text{K}$, *e.g.* in susceptibility measurements. For typical $[Fe_2X_2]^{2+}$ clusters, only the $S = 0$ and $S = 1$ states are decisively populated in these experiments, and therefore the calculated $J_{0\to1}$ should be used for the comparison to J_{exp}. Values for J_{exp} have been determined as $-195\,\text{cm}^{-1}$ and $-185\,\text{cm}^{-1}$ based on the temperature dependence

of ^{1}H-NMR in protein-bound $[Fe_2S_2]^{2+}$ and $[Fe_2Se_2]^{2+}$ clusters (values adapted for Hamiltonian in Eq 1.8).[143] The calculated $J_{0\to1}$ also decrease about 10 cm^{-1} upon Se-substitution for both the state-averaged and the state-specific energies, but the state-averaged energies yield a $J_{0\to1}$ that is about 70 cm^{-1} too small. This is to be expected, since a wave function, that is averaged over all six $S = 0, \dots 5$ states, is biased towards the intermediate spin states and therefore underestimates $J_{0\to1}$. While the state-specific energies yield a value for $J_{0\to1}$ that is in much better agreement with the experimental values, it is not clear whether a CAS(22,16) wave function combined with the NEVPT2 correction actually captures the relevant physics of the system or whether it is a result of fortunate error compensation.

Beyond the HDvV Hamiltonian

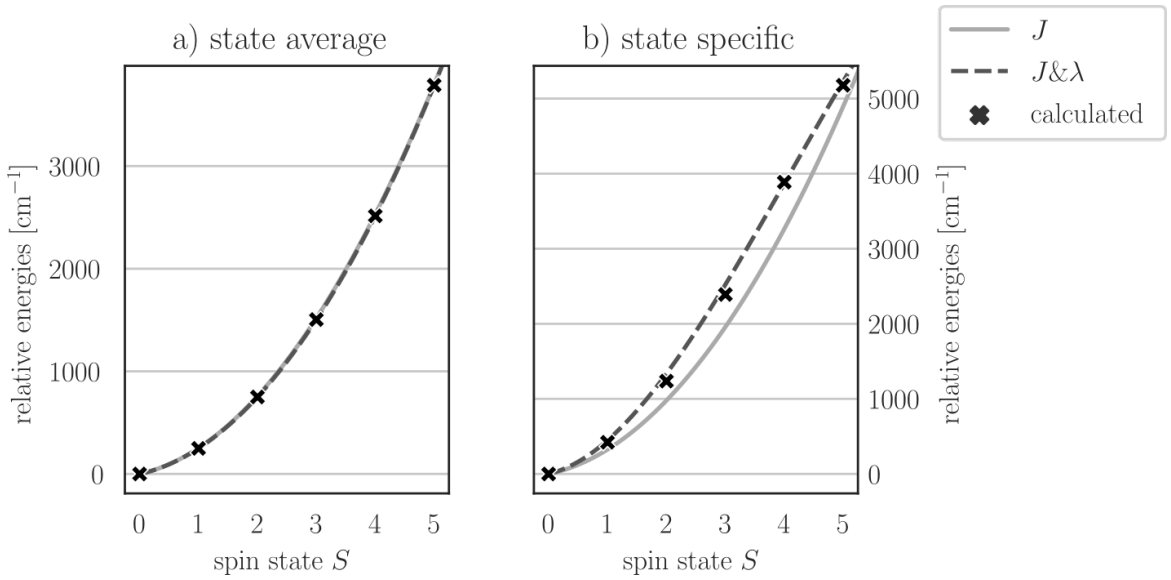

Figure 2.20. – Fits of the HDvV Hamiltonian and its biquadratic extension to the (a) state-averaged and (b) state-specific energies of the spin-coupled states in $[Fe_2S_2]^{2+}$ (CAS(22,16) including NEVPT2). Lines show third order polynomial fits to the spin Hamiltonian energies in Eq 1.8 and Eq 1.10, respectively. See Table A7 for values.

When two magnetic centers with more than one electron each are coupled magnetically, the resulting spin state energies can deviate from the Landé interval rule. The deviation can be modeled with a second-order, so-called biquadratic, term in the HDvV Hamiltonian, as suggested by Falk *et al.*[99] The biquadratic term attempts to capture nH contributions to the wave function, which are only possible for local spins of $S \geq 1$.[287] The extended spin Hamiltonian H_{biq} is given in Eq 1.10 and includes the parameter λ in addition to J.

The fit of the spin Hamiltonian H_{biq} to the state energies is shown in Figure 2.20. For state-averaged energies (Figure 2.20a), the H_{biq} Hamiltonian leads to a visually indistinguishable fit compared to the H_{HDvV} Hamiltonian. Focusing on $[Fe_2S_2]^{2+}$, H_{HDvV} yields $J_{\text{HDvV}} = -126.5\,\text{cm}^{-1}$ while H_{biq} yields $J_{\text{biq}} = -126.5\,\text{cm}^{-1}$ and $\lambda = -0.05\,\text{cm}^{-1}$

(see Table A7 for values). A closer look at the quality of the fit reveals that the standard deviation (SD) of the energy differences decreases from $5.2\,\mathrm{cm}^{-1}$ to $0.4\,\mathrm{cm}^{-1}$ when including λ. However, when fitting the state-specific energies (Figure 2.20b), H_{HDvV} is clearly insufficient to capture the energy differences correctly. Using H_{biq} instead changes the coupling constant from $J_{\mathrm{HDvV}} = -163.2\,\mathrm{cm}^{-1}$ to $J_{\mathrm{biq}} = -168.3\,\mathrm{cm}^{-1}$, while $\lambda = 1.97\,\mathrm{cm}^{-1}$ is considerably larger compared to the state-averaged energies. The quality of the fit for state-specific energies is SD $= 225.3\,\mathrm{cm}^{-1}$ for H_{HDvV} and SD $= 81.3\,\mathrm{cm}^{-1}$, which is two orders of magnitude higher compared to state-averaged energies. Curiously, the SD of H_{biq} drops to $7.7\,\mathrm{cm}^{-1}$ for a CAS(22,16) calculation without the NEVPT2 correction, but for the CAS(10,10) calculation it is of similar magnitude with and without the NEVPT2 correction ($1.7\,\mathrm{cm}^{-1}$–$2.0\,\mathrm{cm}^{-1}$). Therefore, the NEVPT2 correction may lead to a qualitatively different result when combined with the larger CAS(22,16) wave function.

When investigating the S→Se substitution in $[Fe_2X_2]^{2+}$, large changes in the λ parameter can only be observed when fitting to state-specific energies and only for CAS(22,16) calculations including the NEVPT2 correction (see Table A7). Here, λ drops from $1.97\,\mathrm{cm}^{-1}$ in $[Fe_2S_2]^{2+}$ to $0.35\,\mathrm{cm}^{-1}$ in $[Fe_2Se_2]^{2+}$. However, as discussed for Figure 2.19, since the coupling constant J_{avg} exhibits the opposite trend as $J_{0\to1}$ upon S→Se substitution, either the validity of the two spin Hamiltonians or the ability of the CAS(22,16)/NEVPT2 protocol to yield accurate energy differences is subject to further investigations.

The biquadratic term has been calculated for similar $[Fe_2S_2]^{2+}$ clusters with CASSCF and NEVPT2[21] and density matrix renormalization group (DMRG) calculations.[108] The computational protocol of the former is highly similar to the state-averaged approach in the present work and the reported $\lambda = -0.11\,\mathrm{cm}^{-1}$ is in good agreement with value extracted from the state-averaged energies in the present work ($\lambda = -0.05\,\mathrm{cm}^{-1}$). The DMRG calculations yield $\lambda = -1.25\,\mathrm{cm}^{-1}$, which surprisingly is the opposite sign compared to the present work. Since the experimental determination of J typically only involves the one energy difference between $S = 0$ and 1, higher-order contributions such as λ can not be easily extracted and the reported J_{exp} comprises an effective value.

LMCT excited state in $[Fe_2X_2]^{2+}$

The first excited state in $[Fe_2X_2]^{2+}$ with the same $S = 0$ spin as the ground state is an LMCT excited state. The LMCT state is an eigenstate of the system and not to be confused with the previously discussed LMCT CFGs in the CASSCF wave function. The energy of this state is shown in Figure 2.21a. Starting with $10\,174\,\mathrm{cm}^{-1}$ in $[Fe_2S_2]^{2+}$, S→Se substitution lowers the energy to $9272\,\mathrm{cm}^{-1}$ in $[Fe_2SSe]^{2+}$, and $8758\,\mathrm{cm}^{-1}$ in $[Fe_2Se_2]^{2+}$. In protein-bound $[Fe_2S_2]^{2+}$ clusters, a band around $10\,500\,\mathrm{cm}^{-1}$ has been observed experimentally, but its origin has not yet been clarified.[140] The calculated energies suggest that S→Se lowers the energy sufficiently to observe changes in the experimental vis/near-IR spectrum.

The electron redistribution during the LMCT excitation is shown in Figure 2.21b. The removal of electron density from the ground state is shown in orange, the addition to the excited state in purple. The substantial charge redistribution makes it obvious

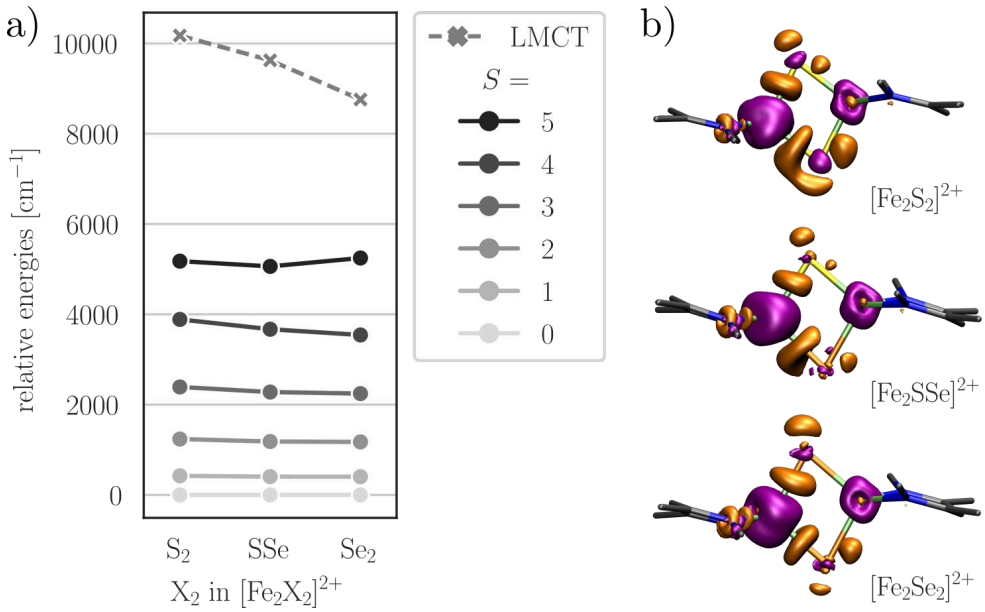

Figure 2.21. – (a) Energies of the $S = 0, \ldots 5$ spin-coupled states and the $S = 0$ ligand-to-metal charge transfer (LMCT) excited state in $[Fe_2S_2]^{2+}$, $[Fe_2SSe]^{2+}$, and $[Fe_2Se_2]^{1+}$ (CAS(22,16) including NEVPT2). (b) Difference electron density between the $S = 0$ ground state and LMCT excited state (attachment: purple, detachment: orange). Isosurfaces are drawn at ±0.005. Figure adapted from Ref [22].

that orbitals for the ground state are not the best for LMCT state and therefore the state-specific approach is applied here. Furthermore, it is evident that the excited electron, even though partly delocalized, favors the Fe_A center. Another LMCT excited state with the electron localized mostly on Fe_B is expected to be very close in energy due to the symmetry of the molecule, as well as four more LMCT states just for local excitations on Fe_A (d^6 Fe^{2+} center). Optimizing the wave function for the LMCT state on Fe_B can be a cumbersome endeavor, which follows to some extent the strategies discussed for the mixed-valent $[Fe_2X_2]^{1+}$ in Section 2.3.5.

2.3.4. CASSCF calculations and symmetry

Symmetry of the $[Fe_2X_2]^{2+}$ wave function

Symmetry-adapted calculations can be invoked in `ORCA` with the `!usesym` keyword. Even though the integral evaluation does not exploit symmetry properties, the size of the configuration space is reduced when using symmetry. In the `ORCA` output file, the irreps is printed along the respective molecular orbital and the electronic state. The structure for the symmetrized $[Fe_2S_2]^{1+}$ molecule is shown in Figure 2.22a. The point group of the symmetrized clusters is D_{2h} (C_{2v}), where the parentheses correspond to the hetero bridged dimer (X_2=SSe). The minimal active space for the $[Fe_2X_2]$ clusters consists of the five $3d$ orbitals on each Fe centers, that form bonding and antibonding linear combinations as illustrated in Figure 2.22b. In the $S = 5$ high spin state, only a single CFG enters the wave function; each active orbital is exactly singly

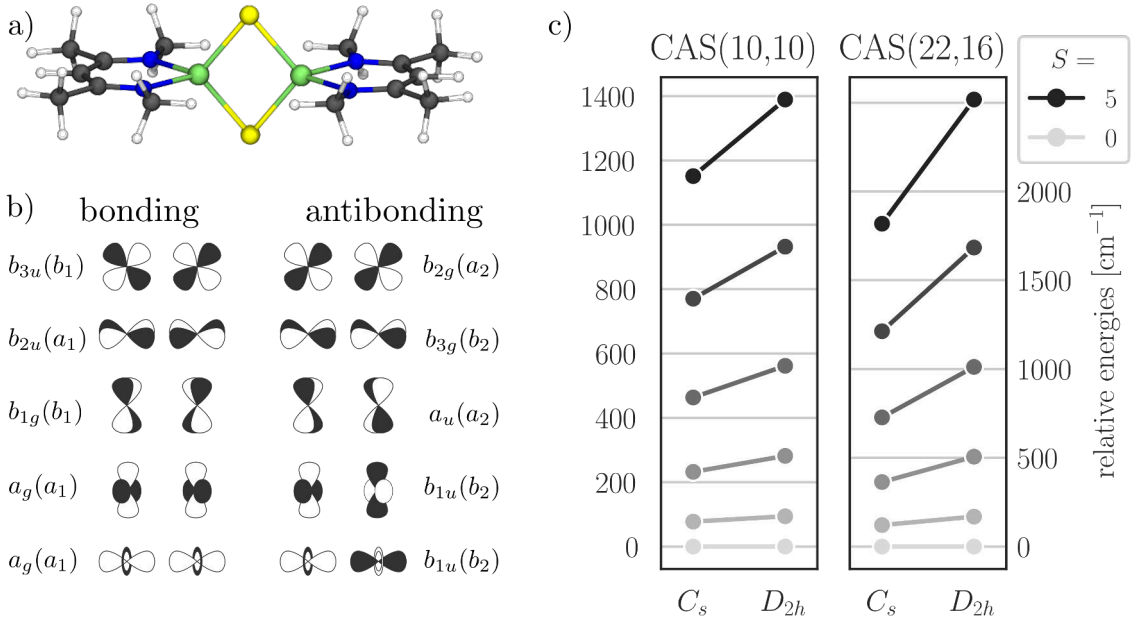

Figure 2.22. – (a) Structure of the symmetrized $[Fe_2S_2]^{2+}$ cluster. (b) Bonding and antibonding linear combinations of Fe 3*d* orbitals as well as the irreducible representation (irrep) in the point group D_{2h} (C_{2v}). Symmetrized $[Fe_2S_2]^{2+}$ and $[Fe_2Se_2]^{2+}$ belong to D_{2h}, and symmetrized $[Fe_2SSe]^{2+}$ to C_{2v}. (c) Energies of the spin-coupled states in the BS-DFT-optimized and symmetrized structures (CAS(10,10) and CAS(22,16) calculations include the NEVPT2 correction for $[Fe_2S_2]^{2+}$).

occupied and the irrep of the state is therefore B_{1u} (B_2). The $S = 0$ ground state belongs to the totally symmetric irrep A_g (A_1). The character of all $S = 5, 4, \ldots$ states alternate between the two irreps $B_{1u}, A_g, \ldots$ ($B_2, A_1, \ldots$). Although the configuration space is smaller when imposing symmetry, the representation of the wave function using the symmetry-adapted orbitals in Figure 2.22b is less compact compared to localizing the active orbitals. With localized active orbitals, the CFG where each orbital is singly occupied has the leading coefficient in all low-energy states (> 99 %, see Section 2.3.3), but with the symmetry-adapted active orbitals, numerous CFGs enter the wave function with small coefficients, especially for the states with lower spin.

The only symmetry element in the BS-DFT-optimized $[Fe_2S_2]^{2+}$ structure is a mirror plane containing the atoms of the rhomboid $[Fe_2S_2]$ core, resulting in the point group C_s. Symmetrizing the molecule to D_{2h} inevitable changes the geometric parameters, *e.g.* all four FeS bonds become equal. The FeS bond lengths decrease from 2.216 Å in C_s (Table 2.6) to 2.200 Å in D_{2h} and the FeFe bond length from 2.821 Å to 2.804 Å. Figure 2.22c shows the change in the low-energy states of $[Fe_2S_2]^{2+}$ when symmetrizing the as-optimized geometry. For the minimal active space (CAS(10,10)), the energy range covered by the states increases by 21 % as a direct consequence of the symmetrization. For the larger active space including orbitals on S^{2-} (CAS(22,16)) the change is even larger, and the energy range increases by 39 %. The energy differences

between the $S = 0, 1 \ldots$ states is proportional to the calculated coupling constant J (see Eq 1.8). Therefore, the effect of symmetrizing the molecular structure, which is common procedure in computational studies, should be investigated when comparing to experimental observables.

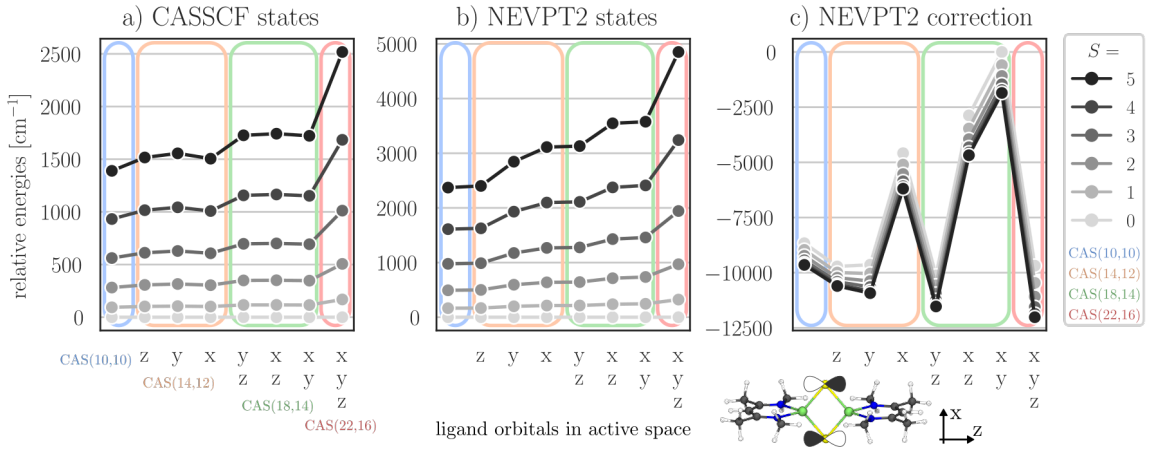

Figure 2.23. – (a) CASSCF and (b) NEVPT2-corrected energies of the spin-coupled states in $[Fe_2S_2]^{2+}$ as a function of the bridging ligand orbitals in the active space as well as (c) NEVPT2 correction energies. The active orbitals in the CAS(10,10) wave function (blue) are the 10 metal-based orbitals. Each x, y, or z indicates an increase of the active space by the ligand-based orbitals of the respective symmetry. For the CAS(22,16) wave function (red), all six doubly occupied $3p$ orbitals of the S_2^{4-} fragment are part of the active space.

Ligand-based active orbitals in symmetrized $[Fe_2S_2]^{2+}$

Up to three occupied p orbitals can be included in the active space for each X^{2-} ligand. When enforcing symmetry, these ligand orbitals transform as irreps of the molecular point group as well, and the p_x, p_y, and p_z orbitals on X do not mix. Because the orbital interaction between the Fe and X p depends on the orientation of the p orbital, the spin-coupling between the Fe centers is expected to depend on the specific ligand p orbital in the active space. Figure 2.23a shows how the low-energy CASSCF states in $[Fe_2S_2]^{2+}$ depend on the choice of the S $3p$ orbitals in the active space. Starting with the minimal active space (CAS(10,10)), the inclusion of a single occupied p orbital per X increases the relative state energies by 10 % (CAS(14,12)), two p orbitals per X by 25 % (CAS(18,14)), and all three per X by 80 % (22,16). Therefore, up to an active space of CAS(18,14) the contributions of the ligand p orbitals are roughly additive, but when including all of them, the coupling strength increases sharply. The orientation dependence is not additive: the p_y orbital leads to the largest splitting in the CAS(14,12) and the p_x and p_z orbitals in the CAS(18,14).

The NEVPT2-corrected state energies (Figure 2.23b) exhibit a different dependence on the orientation of the bridging ligand orbitals compared to the CASSCF energies. For example, including the p_z orbitals in the CAS(10,10) leaves the relative energies virtually unchanged (+1 %), but including the p_x orbitals increases the energy gaps as

much as including both the p_y and the p_z orbitals (+31 % and +32 %, respectively). The active spaces that include the p_x orbitals (oriented along the X−X vector) appear to produce to the largest splitting after applying the NEVPT2 correction, but not for the CASSCF energies. While the reason for this somewhat erratic behavior is not clear, it may be related to the absolute values of the NEVPT2 correction (Figure 2.23c). The absolute values of the NEVPT2 correction for the p_x CAS(14,12), the p_x/p_z CAS(18,14), and p_x/p_y CAS(18,14) calculations appear to follow a different trend. Therefore, the partial inclusion of ligand-based orbitals may lead to an unbalance active space and all the complete $3p$ shell should be included.

S→Se substitution in symmetrized $[Fe_2X_2]^{2+}$

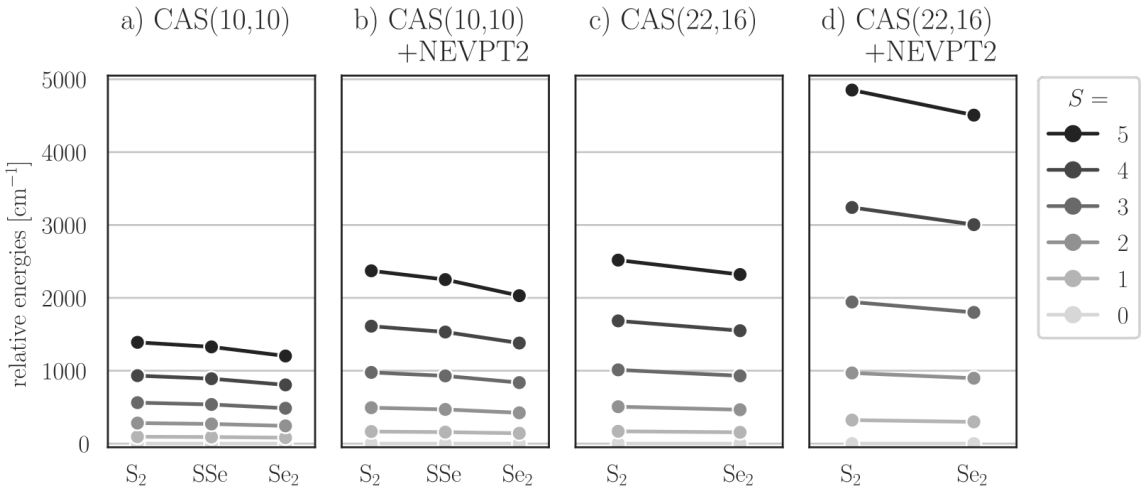

Figure 2.24. – Comparison of the energies of the spin-coupled states in symmetrized $[Fe_2S_2]^{2+}$ (D_{2h}), $[Fe_2SSe]^{2+}$ (C_{2v}), and $[Fe_2Se_2]^{2+}$ (D_{2h}) for (a) a CAS(10,10) calculation, (b) a CAS(10,10) calculation including NEVPT2, (c) a CAS(22,16) calculation, and (d) a CAS(22,16) calculation including NEVPT2.

The energy ordering of the low-energy spin states in the $[Fe_2X_2]^{2+}$ clusters is $S = 0, \ldots 5$, which is shown in Figure 2.24a for CAS(10,10) calculations. Focusing on $[Fe_2S_2]^{2+}$, the energy difference between $S = 0$ and $S = 5$ is 1390 cm^{-1}. Including dynamic correlation with NEVPT2 (Figure 2.24b) increases the energy difference to 2371 cm^{-1}. The splitting for CAS(10,10) including the NEVPT2 correction is of similar magnitude as the CAS(22,16) calculation without NEVPT2 correction (Figure 2.24c, 2518 cm^{-1}). The largest splitting between the spin states is achieved when using the large active space CAS(22,16) plus including the NEVPT2 correction (Figure 2.24d, 4852 cm^{-1}). The treatment of dynamic correlation and the expansion of the active space to include bridging ligand orbitals increases the energy differences between the spin states. The CAS(10,10) wave function is known to be biased towards an ionic wave function and therefore underestimates the spin-coupling between the two Fe centers leading to a smaller spacing between the energies of the spin states.

Also, the S→Se substitution reduces the energy differences regardless of the active space or the NEVPT2 correction (Figure 2.24a to d). When comparing $[Fe_2S_2]^{2+}$

with $[Fe_2Se_2]^{2+}$, the relative energies in Se-based clusters are reduced by 14 % for the CAS(10,10) calculations, but only by 7 % for the CAS(22,16) calculations. Assuming that the CAS(22,16) wave function gives more accurate energy differences, the spin-coupling mediated by S and Se is more similar than the overly ionic CAS(10,10) calculation suggests. A more detailed discussion about the S→Se substitution and the relationship between the electronic structure and the spin-coupling will be given in Section 2.3.3.

Mixed-valent $[Fe_2X_2]^{1+}$ clusters and symmetry

The two Fe centers in the mixed-valent $[Fe_2X_2]^{1+}$ clusters are inequivalent in the BS-DFT-optimized structures (see Section 2.3.1). Symmetrizing the geometry for the $[Fe_2X_2]^{1+}$ clusters in the same way as the $[Fe_2X_2]^{2+}$ clusters, so that both Fe centers are equivalent (D_{2h} for $[Fe_2S_2]^{2+}$), significantly changes the geometric structure. However, the CASSCF ground state wave function for the BS-DFT-optimized structure also features a predominantly localized electron with the oxidation states Fe_A^{2+} and Fe_B^{3+}, as will be discussed later in Section 2.3.5. Furthermore, by running a CASSCF calculation with the `! usesym` keyword, one imposes the same high symmetry on the wave function, therefore, leading to an $Fe^{2.5+}Fe^{2.5+}$ ground state, which is significantly higher in energy than the localized electronic structure. In order to allow for a localized electronic structure, the $[Fe_2S_2]^{1+}$ clusters should be symmetrized at most to the point group C_{2v}, where the rotational axis lies on through the Fe−Fe vector, or perform calculations on a symmetrized geometry without enforcing symmetry of the wave function.[†]

2.3.5. Mixed-valent Fe dimers: $[Fe_2X_2]^{1+}$ clusters

Figure 2.25a shows a schematic representation of the active space in the CAS(23,16) wave function of the mixed-valent $[Fe_2X_2]^{1+}$ clusters (analogous to Figure 2.15 for the homo-valent $[Fe_2X_2]^{2+}$). After transformation (see Section 2.1 for details), the 16 active orbitals can be assigned to the Fe_A center (5 orbitals), the Fe_B center (5 orbitals), and the X_2 fragment (6 orbitals). In the ground state, the high spin Fe_A^{2+} center ($S_{loc} = 2$) and the high spin Fe_B^{3+} center ($S_{loc} = \frac{5}{2}$) are antiferromagnetically coupled to an $S = \frac{1}{2}$ spin state. Here, the neutral CFG is the major contribution to the wave function: one orbital on Fe_A has an occupation number of 2, that of all other Fe-based orbitals is 1, and that on all X-based orbitals is 2, in line with the oxidation states Fe_A^{2+}, Fe_B^{3+}, and X_2^{4-}. Starting from the neutral CFG, excited CFGs can be classified as follows. (i) The extra electron on Fe_A can be excited locally (regular *d-d* CFG, number of doubly occupied orbitals remains the same). (ii) Any other

[†]When using state-averaging in mixed-valent clusters in combination with a symmetric geometry, the converged CASSCF wave function may be delocalized, even without enforcing a symmetric wave function. However, the localized wave function may be lower in energy. A practical approach is to first converge the localized wave function using an asymmetric structure, and then use the localized wave function as a starting guess for the symmetric geometry. If the electronic structure remains localized, and the final energy is lower than for the initially delocalized wave function, the localized electronic structure constitutes the ground state for the symmetric geometry.

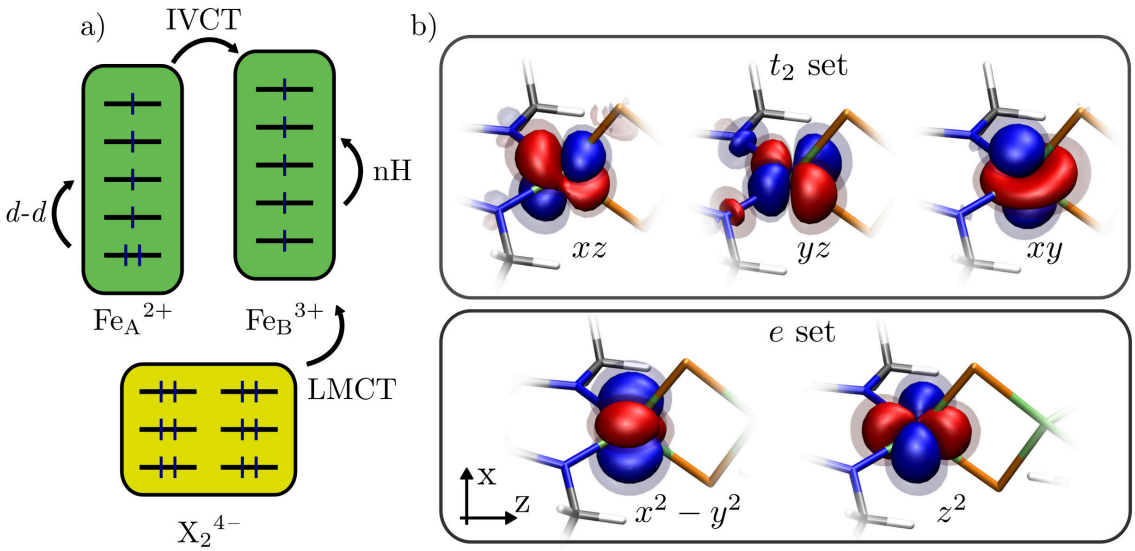

Figure 2.25. – (a) Schematic representations of the CFGs in the CAS(23,16) wave function of the mixed-valent $[Fe_2X_2]^{1+}$ clusters. In the neutral CFG, the z^2 orbital on Fe_A is doubly occupied. (b) Plots of the five active orbitals on Fe_A with characters corresponding to pseudo-T_d symmetry. Shapes for orbitals on Fe_B are similar; the full set of all 16 active orbitals is shown in Figure A7. Isosurfaces are shown at ±0.05 and ±0.025.

Fe-based electron can be excited locally (nH CFG, number of doubly occupied orbitals increases). (iii) An electron can be transferred between the Fe centers (inter-valence charge transfer (IVCT) CFG, analog of the MMCT CFG in the homo-valent $[Fe_2X_2]^{2+}$) (iv) An X-based electron can be transferred to either Fe center (LMCT CFG).

Figure 2.25b shows the active orbitals on Fe_A^{2+} in $[Fe_2X_2]^{1+}$. In contrast to the Fe^{3+} centers in $[Fe_2X_2]^{2+}$, the orbitals show a smaller FeX overlap. The z-axis lies along the Fe−Fe vector, the x-axis in the FeX_2 plane, and the y-axis in the FeN_2 plane. The full set of active orbitals is shown in Figure A7.

Composition of the active orbitals in $[Fe_2X_2]^{1+}$

The composition of the ten Fe-based active orbitals for the CAS(23,16) is shown in Figure 2.26. The composition of all 16 active orbitals is listed in Table A8 and corresponding plots are shown in Figure A7. The metal character of the orbitals localized on Fe_B is up to 5 % higher compared to Fe_A, since Fe_B is more oxidized in the ground state compared to Fe_A. The metal character of the orbitals in the t_2 set, which have a larger FeX overlap than those in the e set, is more sensitive to the metal oxidation state and the differences are similar to the active orbitals in the CAS(5,5)/CAS(6,5) wave functions of the $[Fe(XH)_4]^{1-,2-}$ clusters (cf. Table 2.3).

The active orbitals on Fe_B^{3+} have a 0.5 % larger FeX overlap compared to the Fe^{3+} centers in the homo-valent $[Fe_2X_2]^{2+}$ clusters. The Fe−X bond lengths are virtually identical for the Fe^{3+} centers in either the $[Fe_2X_2]^{2+}$ or the $[Fe_2X_2]^{1+}$ clusters. Therefore, the higher covalency of the Fe_BX bond is most likely due to the neighboring

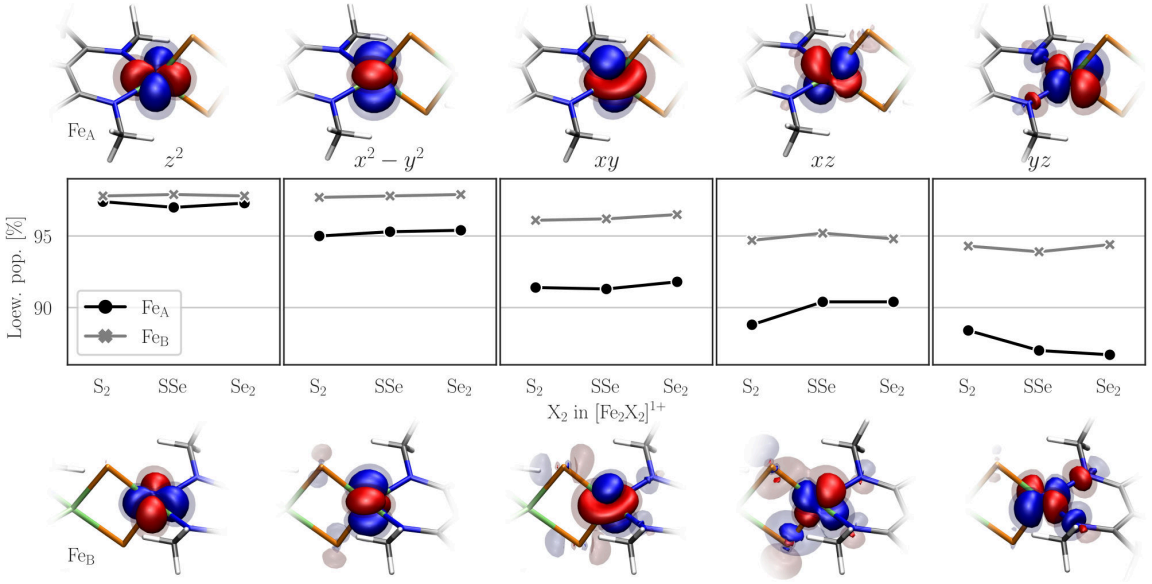

Figure 2.26. – Loewdin atomic contributions of Fe_A and Fe_B to the metal-based active orbitals for $[Fe_2S_2]^{1+}$, $[Fe_2SSe]^{1+}$, and $[Fe_2Se_2]^{1+}$. See Table A8 for all active orbitals in the $S = \frac{1}{2}$ CAS(23,16) ground state wave function as well as X^{2-} and $NacNac^-$ contributions.

$Fe_A{}^{2+}$ center. In the homo-valent $[Fe_2X_2]^{2+}$ clusters, both Fe^{3+} centers compete for the electrons of the bridging ligands, while the less positive charge on the $Fe_A{}^{2+}$, and hence reduced covalency of the Fe_AX bond, allows for a stronger Fe_BX bond.

The S→Se substitution in the $[Fe_2X_2]^{1+}$ clusters leads only to small changes in the composition of the metal-based active orbitals, similar to the homo-valent $[Fe_2X_2]^{2+}$ cluster (compare Figure 2.16). An increase of the Fe contribution correlates with the decrease the X_2 contribution and *vice versa*; the contributions from the $NacNac^-$ ligands remains constant. The magnitude of the substitution effect depends on the orientation of the orbital lobes and it is largest for xz and the yz orbital, which have the largest overlap with the bridging X ligands. Interestingly, this change has not been observed in the active orbitals of the diamagnetically substituted $[FeZnX_2]^{1+}$ and $[FeGaX_2]^{1+}$ clusters (see Table A2). Therefore, the mixing of the bridging ligands in the metal-based active orbitals may be related to the magnetic interaction between the Fe centers.

Composition of the wave function in $[Fe_2X_2]^{1+}$

The CFGs that contribute to the $S = \frac{1}{2}$ ground state of the $[Fe_2X_2]^{1+}$ clusters are shown in Figure 2.27a to e (see Table A8 for values). The neutral CFG dominates the wave function with a weight of 97 %–98 %, which is 1 %–2 % higher compared to the $[Fe_2X_2]^{2+}$ clusters (see Figure 2.17). When substituting S with Se, the weight of the neutral CFG increases by about 0.5 % for each substitution. In the $[Fe_2X_2]^{2+}$ clusters, the change in the neutral CFG appeared to be anticorrelated to the change in MMCT CFGs. However, the equivalent CFGs for the $[Fe_2X_2]^{1+}$, the IVCT CFGs, only decrease by 0.15 % for each substitution. Only together with the decrease in the

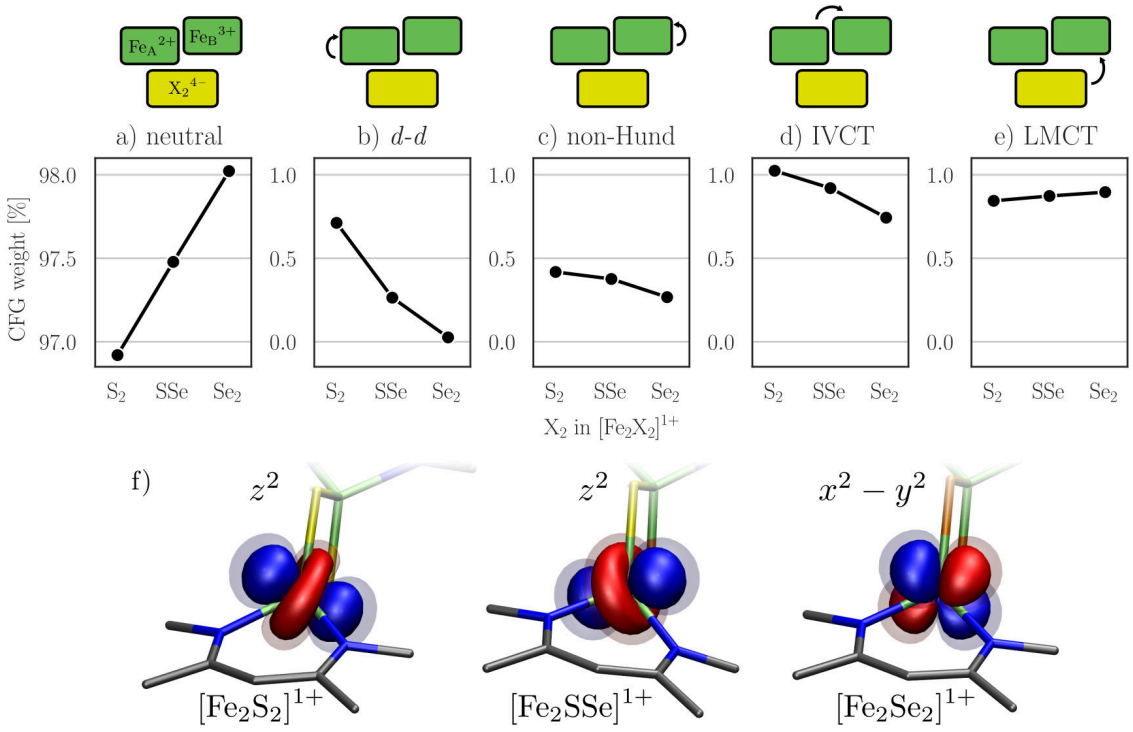

Figure 2.27. – (a-e) CFGs in the $S = \frac{1}{2}$ CAS(23,16) ground state wave functions for the clusters $[Fe_2S_2]^{1+}$, $[Fe_2SSe]^{1+}$, and $[Fe_2Se_2]^{1+}$. See Table A8 for values. (f) Plots of the respective active orbital that is doubly occupied in the neutral CFG. Isosurfaces are shown at ±0.05 and ±0.025. Figure (f) adapted from Ref [22].

d-d CFGs, 0.35 % for each substitution, is the increase in the neutral CFG accounted for. Interestingly, the weight of the *d-d* CFGs is nearly exactly zero in the $[Fe_2Se_2]^{1+}$ cluster, even though the ligand field of the Fe_A^{2+} is qualitatively the same according to the diamagnetically substituted $[FeGaX_2]^{1+}$ clusters.

The weight of the LMCT CFGs is with about 0.9 % slightly lower than in the $[Fe_2X_2]^{2+}$ clusters (about 1.2 %), which consistent with the lower metal charges for $[Fe_2X_2]^{1+}$. The slight increase of about 0.03 % upon each substitution is consistent with the $[Fe_2X_2]^{2+}$ clusters. The weight of the nH CFGs is two to three times as high compared to $[Fe_2X_2]^{2+}$. A higher weight of the nH CFGs is expected for the $[Fe_2X_2]^{1+}$ clusters, since the weaker ligand field around the ferrous Fe_A^{2+} center decreases the energy of promoting an electron locally. However, the weight of the nH CFGs decreases from 0.42 % in $[Fe_2S_2]^{1+}$ to 0.27 % in $[Fe_2Se]^{1+}$. According to the diamagnetically substituted $[FeGaX_2]^{1+}$ clusters, the ligand strength even decreases slightly in going from a bridging S^{2-} to Se^{2-}. Therefore, the ligand strength is not the reason for the trend in the nH CFGs.

The most striking difference in the $[Fe_2X_2]^{1+}$ series lies in the orbital occupation. Figure 2.27f shows the metal-based active orbital that is doubly occupied in the leading CFG of the $S = \frac{1}{2}$ ground state wave function. For $[Fe_2S_2]^{1+}$ and $[Fe_2SSe]^{1+}$, this orbital is the z^2 orbital, which is equivalent to the doubly occupied orbital in the diamagnetically substituted $[FeGaS_2]^{1+}$ and $[FeGaSSe]^{1+}$ cluster, respectively (see

Figure 2.14a and b). For $[Fe_2Se_2]^{1+}$, however, the doubly occupied orbital is the x^2-y^2 orbital. In the diamagnetically substituted $[FeGaSe_2]^{1+}$ cluster, the z^2 orbital is lowest in energy (Figure 2.14c), but according to the AILFT d orbital energies, the x^2-y^2 orbital is merely 88 cm^{-1} higher in energy. The metal-metal interaction in $[Fe_2Se_2]^{1+}$ therefore appears to be responsible for stabilizing an electronic structure with a doubly occupied x^2-y^2 orbital on Fe_A. This demonstrates that S→Se substitution in $[Fe_2S_2]^{1+}$ can change the nature of the ground state wave function.

Low-energy states landscape for $[Fe_2X_2]^{1+}$

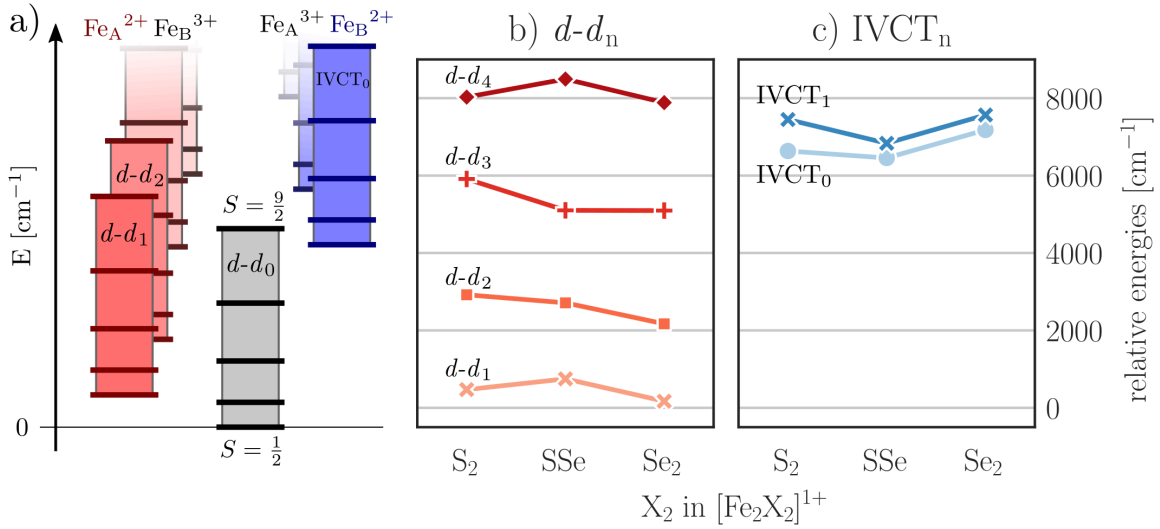

Figure 2.28. – (a) Schematic representation of possible low-lying states in the mixed-valent $[Fe_2X_2]^{1+}$ clusters. States with local intermediate or low spin Fe centers are not considered. Figure 2.29 shows all these states for the CAS(11,10) wave function. (b) Energies of the $S=\frac{1}{2}$ d-d_n excited states and (c) the lowest two $S=\frac{1}{2}$ IVCT$_n$ excited states for the $[Fe_2X_2]^{1+}$ clusters relative to the $S=\frac{1}{2}$ d-d_0 ground state (CAS(23,16) including NEVPT2). See Table A9 for values. Figure (a) adapted from Ref [22].

For the mixed-valent $[Fe_2X_2]^{1+}$ clusters, the low-energy spectrum contains many states, which is illustrated in Figure 2.28a. The Heisenberg spin ladder $S=\frac{1}{2},\ldots\frac{9}{2}$ arising for the ground state (d-d_0) is shown in gray. In the d-d_n excited states (red), an electron is excited locally on $Fe_A{}^{2+}$. When limiting to the local high spin $S_{loc}=2$ on Fe_A, four d-d_n excited states (n = 1,…4) are possible in addition to d-d_0. IVCT excited states (blue) comprise a charge transfer from Fe_A to Fe_B resulting in $Fe_A{}^{3+}Fe_B{}^{2+}$ oxidation states. Similar to the d-d_n excited states, locally excited states exist for $Fe_B{}^{2+}$ yielding a total of five IVCT$_n$ states (n = 0,…4). In each of the d-d_n and IVCT$_n$ states shown, the two Fe centers maintain their local high spins, which couple to give rise to their respective Heisenberg spin ladder with the $S=\frac{1}{2}$ spin being the lowest in energy. Therefore, the interaction between a high spin Fe^{2+} center ($S_{loc}=2$) and a high spin Fe^{3+} center ($S_{loc}=\frac{5}{2}$) generates a total number of $2\times5\times5=50$ states.

Figure 2.28b-d shows the influence of the S→Se substitution for some excited states. The d-d_n states (n = 0, ... 4, Figure 2.28b) span an energy range of about 8000 cm^{-1}. The d-d_n energies accurately follow the trend of the AILFT d orbital energies in the diamagnetically $[FeGaX_2]^{1+}$ clusters (see Figure 2.13). According to LFT, where the electron-electron repulsion is the same for each d orbital, the orbital energies are equal to the excitation energies within the $S = 2$ multiplet of a d^6 system. The correspondence between the two quantities emphasizes (i) that in the ground state of the $[Fe_2X_2]^{1+}$ clusters have a localized $S = \frac{1}{2}$ electronic structure and (ii) the usefulness of diamagnetic substitution to study the electronic structure of more complex systems. The fact that the $[FeGaSe_2]^{1+}$ cluster (doubly occupied z^2 orbital) predicts a different ground state for the $[Fe_2Se_2]^{1+}$ cluster (doubly occupied $x^2 - y^2$ orbital) is acceptable considering that the z^2 and $x^2 - y^2$ orbital energies in $[FeGaSe_2]^{1+}$ lie merely 88 cm^{-1} apart. In other words, the energies of the excited d-d_n states appear to be determined solely by the local environment around the Fe_A center and not by the metal-metal interaction.

The energies of the two lowest IVCT excited states is shown in Figure 2.28c, which lie between 6000 cm^{-1} and 8000 cm^{-1}. Therefore, the IVCT bands are expected to overlap with the d-d bands in experimental absorption spectra (near-IR). In going from $[Fe_2S_2]^{1+}$ to $[Fe_2SSe]^{1+}$, the energies of the $IVCT_0$ and the $IVCT_1$ state decrease by 200 cm^{-1} and 600 cm^{-1}, respectively. However, in going from $[Fe_2SSe]^{1+}$ to $[Fe_2Se_2]^{1+}$, the energies increase by 700 cm^{-1}, respectively. Therefore, the energy of the lowest IVCT state in $[Fe_2Se_2]^{1+}$ is about 500 cm^{-1} higher compared to $[Fe_2S_2]^{1+}$. Because the energies of the d-d excited states show the opposite trend, the S→Se substitution could be a valuable tool when assigning absorption bands in the near-IR region.

The energy gaps between the $S = \frac{1}{2}$ and the $S = \frac{3}{2}$ spin-coupled states decrease upon Se substitution from 370 cm^{-1} in $[Fe_2S_2]^{1+}$ to 200 cm^{-1} in $[Fe_2Se_2]^{1+}$ (Figure 2.28c). Each substitution therefore lowers the energy of the $S = \frac{3}{2}$ state by about 25 %. As a comparison, the energy gap between the $S = 0$ ground state and the $S = 1$ excited state in the homo-valent $[Fe_2X_2]^{2+}$ decreases by merely 5 %. The following text about spin Hamiltonian parameters includes a more detailed discussion about the influence of the bridging ligand on the spin-coupling in the $[Fe_2X_2]^{1+}$ clusters

Heisenberg and double exchange

As explained in Section 1.1.4, the double exchange mechanism can affect the spin states in mixed-valent metal dimers, such as $[Fe_2X_2]^{1+}$. Briefly, an electron can delocalize over the two metal centers favoring a ferromagnetic alignment and therefore lowering the energies of the higher spin states. Figure 1.3b shows that if double exchange dominates the antiferromagnetic coupling ($|B/J| > 5$), the ground state assumes a spin different from $S = \frac{1}{2}$. However, Figure 1.3b and the underlying Eq 1.11 describe the limiting case of a symmetric structure, where the metal centers are equivalent. Considering the localized electronic and geometric structure, this is clearly not the case for the $S = \frac{1}{2}$ ground state of the $[Fe_2X_2]^{1+}$ clusters (see bond lengths in Table 2.10 and energies in Table A9). That Eq 1.11 is not valid for the present models can be further demonstrated by attempting to extract the double exchange parameter: B is equal to the energy between a pair of states in which the extra electron occupies a bonding

Table 2.8. – Parameters of the multi-state spin Hamiltonian (Eq 1.14) for the $[Fe_2X_2]^{1+}$ clusters (CAS(23,16) including NEVPT2). State energies are given in Table A9.

	$J^{\text{eff}}_{d\text{-}d_0}$	$J^{\text{eff}}_{d\text{-}d_1}$	$J^{\text{eff}}_{d\text{-}d_2}$	$J^{\text{eff}}_{\text{IVCT}_0}$
$[Fe_2S_2]^{1+}$	-124	-141	-32	-184
$[Fe_2SSe]^{1+}$	-87	-95	-2	-224
$[Fe_2Se_2]^{1+}$	-66	-80	8	-194
	$\Delta(d\text{-}d_0)$	$\Delta(d\text{-}d_1)$	$\Delta(d\text{-}d_2)$	$\Delta(\text{IVCT}_0)$
$[Fe_2S_2]^{1+}$	0	470	2920	6635
$[Fe_2SSe]^{1+}$	0	748	2712	6454
$[Fe_2Se_2]^{1+}$	0	170	2170	7179

(more stable, −) and an antibonding (less stable, +) delocalized orbital, respectively, and it is proportional to the transfer integral β between the orbitals centered on A and B. For the $[Fe_2X_2]^{1+}$ clusters, these state pairs correspond to one of the d-d_n states (−) and the respective IVCT_n state (+). The energy difference between the − and the + state pair is $2B$ for $S = \frac{1}{2}$ and $4B$ for $S = \frac{3}{2}$. With the NEVPT2-corrected CAS(23,16) energies (Table A9), the double exchange parameter would amount to $B_{S=1/2} = 6635\,\text{cm}^{-1}$ in one case and to $B_{S=3/2} = 3654\,\text{cm}^{-1}$ in the other. Therefore, the double exchange Hamiltonian in the symmetric limit clearly does not capture the state energies in the mixed-valent $[Fe_2X_2]^{1+}$ clusters.

The effect of an asymmetric structural distortion on the spin states in a mixed-valent metal dimer can be modeled with the PKS spin Hamiltonian (Eq 1.12). As shown for some set of spin Hamiltonian parameters in Figure 1.3c, an asymmetric vibrational mode Q can make multiple spin states accessible, which is referred to as vibronic coupling in this context. However, if vibronic coupling is strong compared to double exchange, the vibrational mode Q does not change the ground state spin (Eq 1.13). The energies follow again the Landé interval rule and the single effective coupling constant J^{eff} incorporates all contributions from super exchange, double exchange, and vibronic coupling.

In the limit of Eq 1.13, the coupling constant is $-124\,\text{cm}^{-1}$, $-87\,\text{cm}^{-1}$, and $-66\,\text{cm}^{-1}$ for the $[Fe_2S_2]^{1+}$, $[Fe_2SSe]^{1+}$, and $[Fe_2Se_2]^{1+}$ cluster, respectively ($J^{\text{eff}}_{d\text{-}d_0}$ in Table 2.8). The coupling strength therefore decreases by 50 % when replacing the S bridge with Se. At the same time, the coupling strength in the homo-valent $[Fe_2X_2]^{2+}$ clusters decreased by merely 10 %. Therefore, the balance between Heisenberg exchange, double exchange, and vibronic coupling appears to be strongly affected by the S→Se substitution and Eq 1.13 may not be a valid approximation. The experimental EPR spectrum of $[Fe_2Se_2]^{1+}$ exhibits a mixture of both $S = \frac{1}{2}$ and $S = \frac{3}{2}$ signals.[23] It therefore seems as an appropriate step to investigate the effects of vibronic coupling on the spin state energies in future calculations.

Furthermore, the energy expression in Eq 1.13 does not have the same number of states as the low-energy spectrum shown in Figure 2.28. Each d-d_n and IVCT_n state gives rise to a spin ladder, therefore, a spin Hamiltonian that is consistent with the number of states is given by Eq 1.14, which contains the coupling constant J_i^{eff} for

each state (where i runs over all d-d_n and $IVCT_n$ states) plus the position of the spin ladder relative to the ground state Δ_i. The spin Hamiltonian parameters for the state energies shown in Figure 2.28 and using Eq 1.14 are listed in Table 2.8. Similar to the homo-valent $[Fe_2X_2]^{2+}$ clusters, the coupling constant J_i^{eff} is extracted using only the $S = \frac{1}{2} \rightarrow \frac{3}{2}$ energy gap, because it is the most relevant for any comparison with experimental values. The parameter Δ_i simply corresponds to the excitation energy of the respective d-d_n or $IVCT_n$ state.

The coupling constant J_i^{eff} strongly depends on the respective excited state. For the ground state (d-d_0) and the first locally excited state (d-d_1), J_i^{eff} is with $-120\,\text{cm}^{-1}$–$-140\,\text{cm}^{-1}$ of comparable magnitude. However, $J_{d\text{-}d_2}^{\text{eff}}$ is merely $-30\,\text{cm}^{-1}$, which indicates significantly reduced spin-coupling. The spin-coupling strength can be related to the orientation of the doubly occupied active orbital (*vide infra* CAS(11,10) calculations). S→Se substitution reduces the spin-coupling in the d-d_n states. When comparing $[Fe_2S_2]^{1+}$ with $[Fe_2Se_2]^{1+}$, the coupling constant of the ground state $J_{d\text{-}d_0}^{\text{eff}}$ decreases by 50 %, which is significantly stronger compared to 10 % decrease in the homo-valent $[Fe_2X_2]^{2+}$ clusters (see Figure 2.18). For protein-bound $[Fe_2X_2]^{1+}$ ferredoxins, the coupling constant was determined as $100\,\text{cm}^{-1}$ for the S-based and $92.5\,\text{cm}^{-1}$ for the Se-based cluster, which is a similarly moderate decrease.[143] The strong changes in spin-coupling, that S→Se substitution invokes in the $[Fe_2X_2]^{1+}$ clusters, may be a result of the Fe−Fe bond elongation, which is 50 % larger compared to homo-valent $[Fe_2X_2]^{2+}$ clusters (see Table 2.6). For the protein-bound $[Fe_2X_2]^{1+}$ ferredoxins, this structural change could be restrained by the protein environment. The elongation of the Fe−Fe bond length also makes a reduction of J^{eff} through an increase in double exchange less likely, since the transfer integral β is expected to decrease with the metal-metal distance.

CAS(11,10) minimal active space for $[Fe_2X_2]^{1+}$

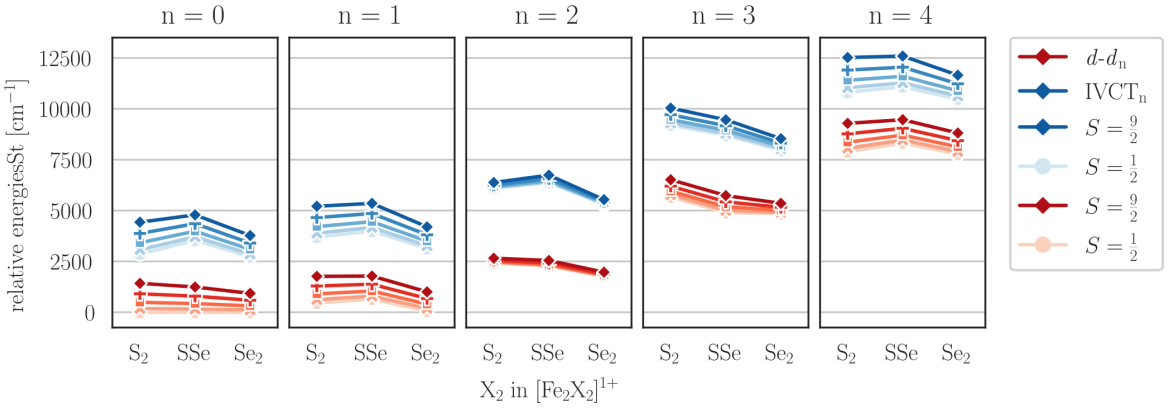

Figure 2.29. – Complete low-energy spectrum of the $[Fe_2X_2]^{1+}$ clusters based on NEVPT2-corrected CAS(11,10) energies. See Figure 2.28a for schematic representation. d-d_n states ($Fe_A^{2+}Fe_B^{3+}$) are shown in red and $IVCT_n$ states ($Fe_A^{3+}Fe_B^{2+}$) in blue. See Table A10 for values.

All state energies discussed previously for $[Fe_2X_2]^{1+}$ correspond to calculations with a CAS(23,16) active space, which includes occupied X_2 ligand orbitals. The calculations

for the homo-valent $[Fe_2X_2]^{2+}$ clusters have shown that the larger active space greatly improves the values for the calculated coupling constants J, probably due to a better description of the covalency in the FeX bond (see Table A7). Therefore, the energies in the mixed-valent $[Fe_2X_2]^{1+}$ clusters calculated with the CAS(11,10) minimal active space are not expected to be in quantitative agreement with experimental values. However, they can be acquired at relatively low computational cost, which makes them useful in discussing trends over all 50 low-energy states sketched in Figure 2.28a.

The CAS(11,10) energies of all d-d_n and $\mathrm{IVCT_n}$ states ($\mathrm{n} = 0, \ldots 4$) in the $[Fe_2X_2]^{1+}$ clusters, as well as the corresponding spin ladders ($S = \frac{1}{2}, \ldots \frac{9}{2}$) are shown in Figure 2.29. While the CAS(11,10) d-d_n excitation energies are consistently a few $100\,\mathrm{cm}^{-1}$ lower compared to the CAS(23,16) values, the $\mathrm{IVCT_n}$ excitation energies for $\mathrm{n} = 0, 1$ are underestimated by $3000\,\mathrm{cm}^{-1}$–$4000\,\mathrm{cm}^{-1}$ by with CAS(11,10). Similar to the $[Fe_2X_2]^{2+}$ clusters, the splitting between the lowest two spin states in the CAS(11,10) wave function is only 25 %–75 % compared to the larger active space. The differences increase strongly upon S→Se substitution. Therefore, the CAS(11,10) minimal active space significantly underestimates all energy differences. This is important to keep in mind when comparing to experimental observables as well as during the perturbational treatment of relativistic effects, which critically depend on the excitation energies.

As mentioned already for the CAS(23,16) calculations, the coupling constant J heavily depends on the respective state (see Table 2.8). However, the more complete state energy spectrum in Figure 2.29 makes it easier to identify possible trends. The spin ladders generated by the d-d_n and the $\mathrm{IVCT_n}$ states show a similar dependency for each n^{th} state. The spin-coupling appears to depend on the orbital that is predominantly doubly occupied in each excited state, which is equivalent for each d-d_n and $\mathrm{IVCT_n}$ state, respectively. For $\mathrm{n} = 0, 1$, and 4, the spin ladder spans a similar energy range of around $1100\,\mathrm{cm}^{-1}$. For $\mathrm{n} = 3$ and especially for $\mathrm{n} = 2$, the energy range decreases to around $700\,\mathrm{cm}^{-1}$ and $200\,\mathrm{cm}^{-1}$, respectively, indicating a significantly reduced spin-coupling for these states. For these states, the xz and xy orbital (see Figure 2.26) are predominantly doubly occupied, which are the two orbitals with the largest FeX overlap. This effect is not a mere artifact of the minimal active space, because the d-d_2 state for the CAS(23,16) wave function shows a similarly reduced spin-coupling. Rather, occupying orbitals pointing at the X_2 ligands appears to reduce the strength of the spin-coupling by decreasing in the FeX covalency.

CASSCF and charge transfer states

During an IVCT excitation, an electron is transferred from Fe_A to Fe_B. It was hinted already in the discussion of the LMCT excited state in the homo-valent $[Fe_2X_2]^{2+}$ clusters, that charge reorganization has a strong influence on the orbitals of a CASSCF wave function. Indeed, the orbitals that are suited well for an ${Fe_A}^{2+}{Fe_B}^{3+}$ state (d-d_n) are quite different to those for an ${Fe_A}^{3+}{Fe_B}^{2+}$ state ($\mathrm{IVCT_n}$). The dependence of the state energies on the orbitals optimized for either of the two state types is shown in Figure 2.30a. When the weight on the $\mathrm{IVCT_n}$ states (blue) is 0 %, conversely, the weight on the d-d_n states (red) is 100 %. Obviously, the orbitals are optimized for the d-d_n states and they are unsuited for the $\mathrm{IVCT_n}$ states, because these are about $1.0 \times 10^5\,\mathrm{cm}^{-1}$ higher in energy. When increasing the weight of the $\mathrm{IVCT_n}$ states, their

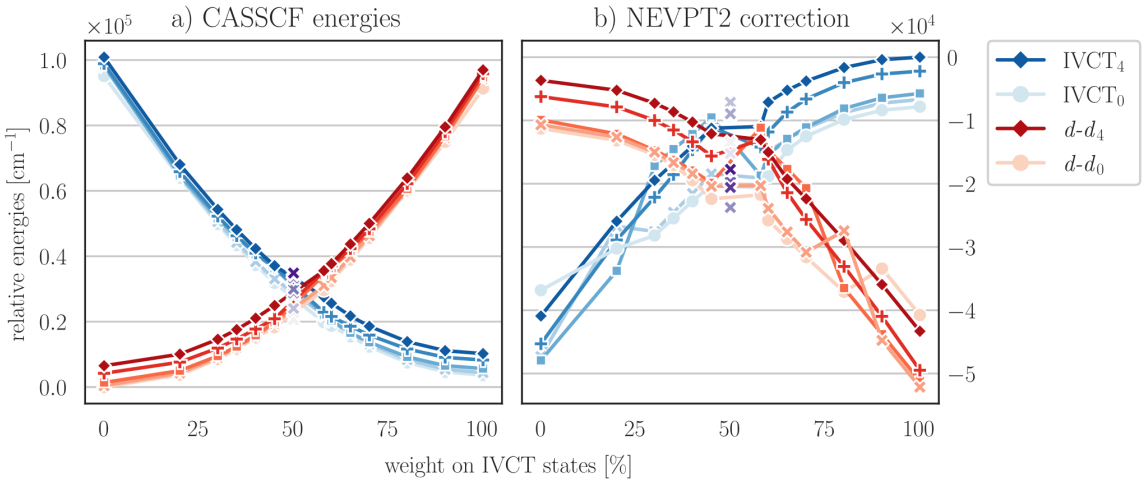

Figure 2.30. – (a) CASSCF energies of the d-d_n states (red) and $IVCT_n$ states (blue) in $[Fe_2S_2]^{1+}$ as a function of the weight placed on the $IVCT_n$ states during the CASSCF optimization and (b) the respective NEVPT2 correction. Energies are shown for a $S = \frac{9}{2}$ CAS(11,10) calculation relative to the $S = \frac{9}{2}$ d-d_0 state. At 0 %, all the weight is placed on the d-d_n states and the orbitals are well-suited for the d-d_n states, but ill-suited for the $IVCT_n$ states. The energy difference between the d-d_0 state at 0 % and the $IVCT_0$ state at 100 % is the IVCT excitation energy.

energy decreases while the energy of the d-d_n states increases – the orbitals attempt to adapt to both state types simultaneously. For a weight of 50 %, all 10 states are described equally (bad) and their energies are nearly identical. Reversely, when a weight of 100 % is on the $IVCT_n$ states, the energies of the d-d_n states are about $1.0 \times 10^5\,cm^{-1}$ higher. Note that the optimized $IVCT_n$ state energies (blue at 100 %) are still higher than the optimized d-d_n states energies (red at 0 %). This difference corresponds to the IVCT excitation energies discussed earlier.

Figure 2.30b shows the NEVPT2 correction for the CASSCF state energies as a function of the weight of the $IVCT_n$ states. When the orbitals are optimized for the d-d_n states (at 0 %), the NEVPT2 correction is small for the d-d_n states, but fairly large for the $IVCT_n$ states. This is not surprising, since in case of the $IVCT_n$ states, the NEVPT2 correction attempts to correct for the ill-suited orbitals in addition to dynamic correlation. When comparing the energy ranges between Figure 2.30a and b at 0 %, the NEVPT2 correction decreases the gap between the d-d_n and $IVCT_n$ states by about $0.5 \times 10^5\,cm^{-1}$, which is about half of the difference between the corresponding CASSCF energies. The absolute energies of the $IVCT_n$ states are therefore about $0.5 \times 10^5\,cm^{-1}$ higher compared to the d-d_0 ground state. Furthermore, the ordering of the $IVCT_n$ states is not maintained. A similar argument can be made for the d-d_n states at 100 %. Overall, this showcases the failure of the perturbational method NEVPT2 when using an insufficient zeroth order starting point. Therefore, care should be taken when averaging the wave function over states that include a charge transfer; with an equal weight on all 10 states (50 %, purple) the CASSCF plus NEVPT2 protocol leads to a random distribution of state energies.

Converging the orbitals for localized mixed-valent clusters such as $[Fe_2X_2]^{1+}$ is a non-trivial task, which has been shown by Domingo *et al.*[288] for $Fe^{2+}Fe^{3+}$ clusters. The strategy used in this work combines their approach with the findings regarding the state-averaging protocol discussed in Figure 2.18 for spin-coupled states. The goal is to acquire reliable values for (i) *d*-*d* excitation energies, (ii) IVCT excitation energies, and coupling constants J (*e.g.* $S = \frac{1}{2} \rightarrow \frac{3}{2}$).

The following protocol is written using the `ORCA` program's syntax, but it should be transferable to other quantum chemistry programs. The orbitals produced in one step are used as starting orbitals in the next (`%moinp some.gbw`). In order to converge the states with an $Fe_A^{2+}Fe_B^{3+}$ electron distribution:

- Starting with a reasonable guess for the Fe-based orbitals (*e.g.* HF orbitals), partly converge the $S = \frac{9}{2}$ high spin CAS(11,10) wave function averaged over 10 roots (`mult 10 nroots 10`). This corresponds to 50 % in Figure 2.30a. Localize active orbitals (`actorbs locorbs`).

- Identify how many of the lowest roots have oxidation states $Fe_A^{2+}Fe_B^{3+}$ for more than half of the CFGs. Select that many roots, interrupt after a few macro-iterations, and repeat until the 5 lowest roots are all dominated by CFGs with $Fe_A^{2+}Fe_B^{3+}$. Fully converge orbitals.

- Expand active space by the X_2-based occupied orbitals. Converge high spin CAS(23,16) wave function averaged over 5 roots (`mult 10 nroots 5`).

- Converge $S = \frac{1}{2}$ and $\frac{3}{2}$ separately (`mult 2` and `mult 4`, respectively). Keep averaging over 5 roots in order to avoid occupation numbers close to 2 for the active orbitals, since this may cause convergence problems. If one is specifically interested in just one *d*-d_n excitation energy, one could average only over the ground and n^{th} excited state (*e.g.* `mult 2 nroots 5 weights[0] 1,0,0,1,0` for $n = 3$).

For the $IVCT_n$ excited states ($Fe_A^{3+}Fe_B^{2+}$), the above protocol requires an extra step after the first one:

- Select a root with a high contribution from CFGs corresponding to an $Fe_A^{3+}Fe_B^{2+}$ oxidation state. Optimize only this root (*e.g.* `nroots 10` and `weights[0] = 0,0,0,0,0,0,1,0,0,0`), interrupt when root flipping occurs. The intermediate orbitals will lower the energies for $IVCT_n$ states. Repeat until at least the lowest two roots are dominated by CFGs corresponding to the $Fe_A^{3+}Fe_B^{2+}$ oxidation state.

The basis set for the above protocol can be kept small, *e.g.* `def2-SVP`, and only after the last step increased to *e.g.* `def2-TZVP`. Note that expanding the basis set from one using effective core potentials (`def2-XVP`) to an all-electron basis set (`ZORA-def2-XVP`) may cause slow convergence. Therefore, it may be economical to reduce the number of SCF cycles for the large basis set and use the small, but all-electron, basis set (`ZORA-def2-SVP`) from the very beginning.

2.3.6. Properties calculated with DFT

DFT usually provides a computationally inexpensive protocol to obtain accurate molecular structures, once a suitable functional is validated against structural data for related compounds. While the TPSSh functional has been shown to yield accurate structures for FeS clusters,[51–53] it may not necessarily yield accurate results for other properties.[20] The calculated properties discussed below were obtained with the same level of theory used to optimize the geometries (Section 2.1). No proper benchmark against experimental values was performed. The observables calculated with DFT are included to serve as a starting point for future studies and to discuss trends in properties that are not easily accessible via the CASSCF wave function.

Table 2.9. – Properties for the $[Fe_2X_2]^{2+,1+}$ clusters calculated with BS-DFT: the coupling constant J (Eq 1.8), the Mössbauer isomer shift δ and quadrupole splitting ΔE_q, and the g values for the $M_S = \frac{1}{2}$ BS determinant. The calibration for the isomer shifts is shown in Figure A4. Values for J were obtained with the Yamaguchi projection.

	J_{DFT} [cm^{-1}]	δ_A [mm/s]	δ_B	ΔEq_A [MHz]	ΔEq_B	g_1	g_2	g_3	g_{iso}
				$[Fe_2X_2]^{2+}$					
$[Fe_2S_2]^{2+}$	-226	0.25	0.25	1.03	1.03				
$[Fe_2SSe]^{2+}$	-218	0.26	0.26	-0.99	0.99				
$[Fe_2Se_2]^{2+}$	-214	0.27	0.27	-0.89	-0.90				
	J_{DFT}	δ_A	δ_B	ΔEq_A	ΔEq_B	g_1	g_2	g_3	g_{iso}
				$[Fe_2X_2]^{1+}$					
$[Fe_2S_2]^{1+}$	-102	0.62	0.28	-2.39	1.32	1.94	1.99	2.04	1.99
$[Fe_2SSe]^{1+}$	-130	0.63	0.29	-2.20	1.10	1.95	2.03	2.08	2.02
$[Fe_2Se_2]^{1+}$	-116	0.61	0.43	-1.04	0.88	2.04	2.07	2.10	2.07

Coupling constants

The coupling constants for $[Fe_2X_2]^{2+,1+}$ calculated with BS-DFT (J_{DFT}) are shown in Table 2.9. As discussed earlier, they were also extracted directly from the CASSCF state energies (see Table A7 and 2.8). For the homo-valent clusters, the values for extracted from the $S = 0$ and $S = 1$ CASSCF state energies ($J_{0\to1} = 211\,cm^{-1}$, $202\,cm^{-1}$, and $201\,cm^{-1}$ for $[Fe_2S_2]^{2+}$, $[Fe_2SSe]^{2+}$, and $[Fe_2Se_2]^{2+}$, respectively) are only about $15\,cm^{-1}$ lower than the BS-DFT values. However, since the BS-DFT values were obtained from the difference of the high spin determinant ($M_S = 5$) and the low spin BS determinant ($M_S = 0$), a better comparison would be with the energy difference between the $S = 0$ and $S = 5$ CASSCF states ($J_{0\to5} = 173\,cm^{-1}$, $169\,cm^{-1}$, and $174\,cm^{-1}$, respectively). Here, the values are about $50\,cm^{-1}$ lower compared to BS-DFT and do not show the same decrease upon S→Se substitution. Therefore, the good agreement between J_{DFT} and $J_{0\to1}$ probably does not imply that both methods calculate the coupling constant similarly well.

For the mixed-valent clusters, J_{DFT} and the CASSCF-extracted $J_{1/2\to3/2}$ ($-124\,\mathrm{cm}^{-1}$, $-87\,\mathrm{cm}^{-1}$, and $-66\,\mathrm{cm}^{-1}$ for $[Fe_2S_2]^{1+}$, $[Fe_2SSe]^{1+}$, and $[Fe_2Se_2]^{1+}$, respectively) do not agree as well as the for the homo-valent clusters. For $[Fe_2S_2]^{1+}$, the absolute value of $J_{1/2\to3/2}$ is merely $20\,\mathrm{cm}^{-1}$ higher compared to J_{DFT}. However, the absolute value of $J_{0\to1}$ decreases by 50 % in going from $[Fe_2S_2]^{1+}$ to $[Fe_2Se_2]^{1+}$, while that of J_{DFT} increases by 10 %. The $S = \frac{9}{2}$ state was not calculated with CASSCF on the same level of theory, therefore, a comparison with $J_{1/2\to9/2}$ is not possible.

Mössbauer parameters

The ^{57}Fe Mössbauer parameters for the $[Fe_2X_2]^{2+,1+}$ clusters listed in Table 2.9. The isomer shift requires a calibration of the electron density at the nucleus against experimental values, which is shown in Figure A4. The isomer shifts δ and the quadrupole splitting for the homo-valent $[Fe_2S_2]^{2+}$ cluster are in very good agreement with the experimentally determined values for a similar NacNac$^-$capped $[Fe_2S_2]^{2+}$ cluster ($\delta_{\mathrm{exp}} = 0.29\,\mathrm{mm/s}$ and $|\Delta E_{\mathrm{Q,exp}}| = 1.06\,\mathrm{mm/s}$).[141] When comparing $[Fe_2Se_2]^{2+}$ with $[Fe_2S_2]^{2+}$, the isomer shift increases by 0.02 mm/s and the absolute value of the quadrupole splitting by 0.1 mm/s, but not conclusive experiment data exists for $[Fe_2Se_2]^{2+}$ clusters.[168,176]

For a similar NacNac-capped $[Fe_2S_2]^{1+}$ cluster, Mössbauer measurements at 80 K exhibit a single doublet[141], however, other model clusters show two clearly distinguishable doublets in line with a localized $Fe^{2+}Fe^{3+}$ electronic structure.[142] The calculated parameters for $[Fe_2S_2]^{1+}$ are also consistent with a localized electronic structure with the $Fe_A{}^{2+}$ center having a higher isomer shift and a higher absolute quadrupole splitting than the Fe^{3+} center, similar to the values extracted for both model compounds and protein-bound $[Fe_2S_2]^{1+}$ clusters.[142] S→Se substitution in $[Fe_2S_2]^{1+}$ leads to large changes in the calculated parameters except for δ_{A}, whereas experimental measurements suggest similar parameters for protein-bound $[Fe_2S_2]^{1+}$ and $[Fe_2Se_2]^{1+}$.

g values

Table 2.9 shows the calculated g values for the $[Fe_2X_2]^{1+}$ clusters based on the $M_S = \frac{1}{2}$ BS determinant. The g values for $[Fe_2S_2]^{1+}$ span a range of 1.94–2.04 with an average of $g_{\mathrm{iso}} = 1.99$. Compared to the $S = \frac{1}{2}$ EPR spectrum of a similar NacNac$^-$-capped $[Fe_2S_2]^{1+}$ cluster ($g = 2.06$, 1.95, and 1.73), the calculated g_{iso} value is higher, but the g anisotropy is lower. Upon S→Se substitution, the g_{iso} value increase to 2.07 for $[Fe_2Se_2]^{1+}$. An increase was also observed experimentally in protein-bound $[Fe_2X_2]^{1+}$ clusters.[166,173–175]

Vibrational frequencies

The calculated IR spectra in the energy range covering vibrations of the FeX core are shown in Figure 2.31 for $[Fe_2S_2]^{2+}$ (a), $[Fe_2Se_2]^{2+}$ (b), $[Fe_2S_2]^{1+}$ (c), and $[Fe_2Se_2]^{1+}$ (d). Also, vibrations involving the entirety of the NacNac$^-$ ligands fall in this energy range. The four vibrations with the largest changes in the Fe−X bond lengths are indicated in the spectra. Consistent with the $[Fe_2X_2]^{2+}$ cluster belonging to the point group D_{2h} (a and b), *i.e.* an inversion center exists, the vibrational modes that do not

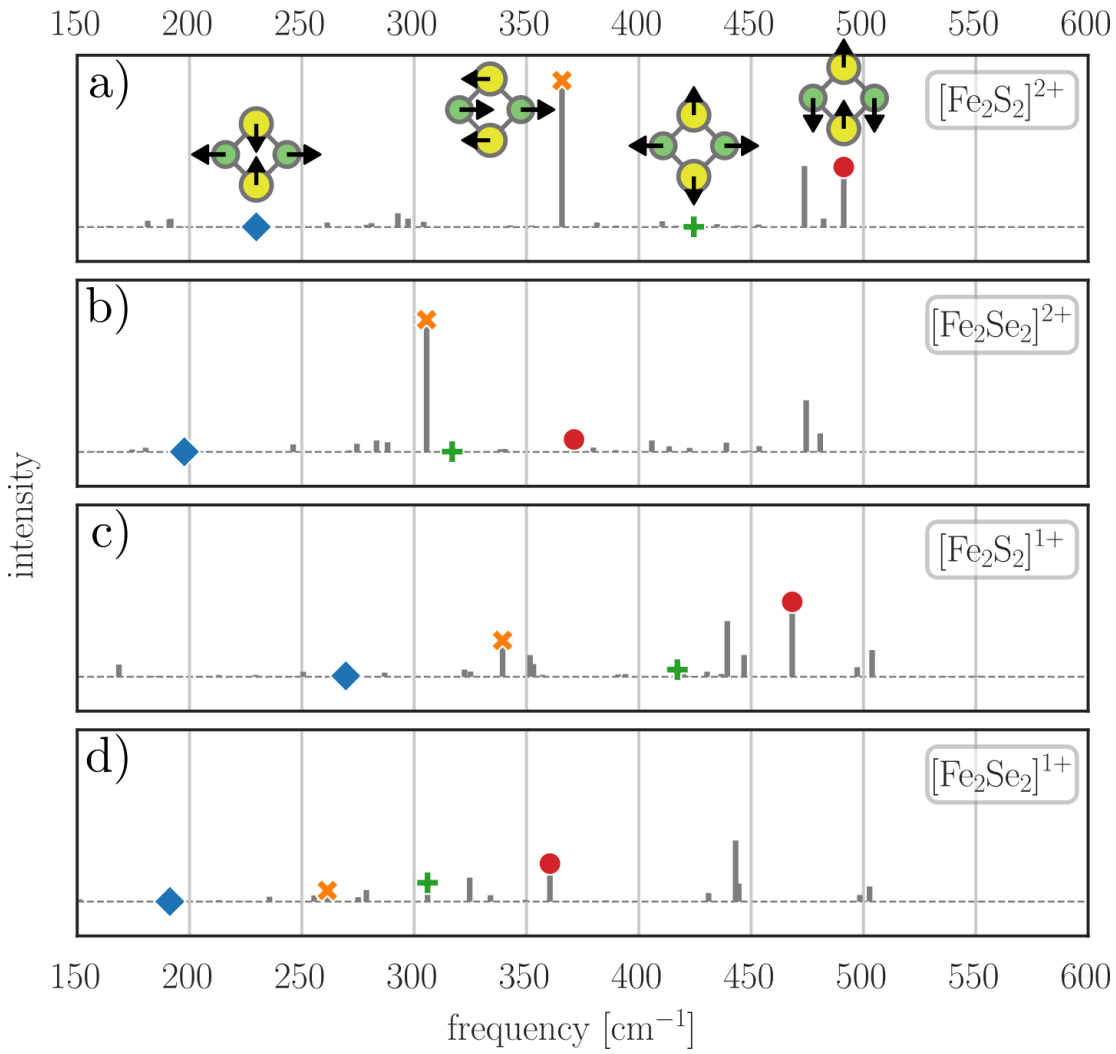

Figure 2.31. – Frequencies and intensities (au) of the low-energy IR vibrations calculated with BS-DFT for the clusters (a) $[Fe_2S_2]^{2+}$, (b) $[Fe_2Se_2]^{2+}$, (c) $[Fe_2S_2]^{1+}$, and (d) $[Fe_2Se_2]^{1+}$. Four of the normal modes involving the FeX core are depicted.

change the dipole moment (◆ and ✚) have near-zero IR intensity. On the other hand, those vibrations that do change the dipole moment (✖ and ●), have large intensities. According to the rule of mutual exclusion,[289] vibration without IR activity are Raman active and *vice versa*. The rule applies to some extent to the $[Fe_2X_2]^{1+}$ clusters (c and d), even though they do not have an inversion center due to the inequivalence of the two Fe centers. Reducing the $[Fe_2X_2]^{2+}$ clusters to $[Fe_2X_2]^{1+}$ tends to lower the energy of the four vibrational modes by about $10\,cm^{-1}$–$40\,cm^{-1}$, consistent with the weaker FeX bond.

Substituting S with Se leads to a red shift, between $10\,cm^{-1}$ and $120\,cm^{-1}$; for all four highlighted vibrational energies. This is to be expected from the higher atomic mass for Se compared to S while forming a similarly strong FeX bond (see Table 1.1). The shifts are of similar magnitude as those observed in Raman spectra of protein-bound $[Fe_2X_2]^{2+}$ clusters, however, since individual peak assignments were not reported, one cannot comment on the quantitative agreement.[178,179] Furthermore, the low-energy

vibrational frequencies are known to be inaccurate within the harmonic approximation, which make them unsuitable to calculate properties such as the vibrational contribution to the entropy.[66]

2.4. Conclusion

The present chapter explored the effect of the S→Se substitution on the geometry and the electronic structure of monomeric and dimeric FeS complexes. While the geometric structures were optimized with DFT, the discussion of the electronic structure focused, for the most part, on the CASSCF wave function. The analysis followed a bottom-up approach, where the complexity of the systems increases gradually. As a starting point, monomeric $[Fe(XH)_4]^{1-,2-}$ complexes (X = S, Se) were studied in Section 2.2. The purpose of these simple molecules was to gauge the magnitude of the perturbation that the S→Se substitution introduces. The frameworks used to analyze the metal-ligand interactions, namely AILFT and AOM, were introduced along with this simple system. Section 2.3 discussed dimeric, $NacNac^-$-capped $[Fe_2X_2]^{2+,1+}$ clusters, which are commonly used as synthetic model compounds for the biological $[Fe_2S_2]$ ferredoxins.[141,142] First, the local electronic structure was explored in the diamagnetically substituted clusters $[FeGaX_2]^{1+}$ and $[FeZnX_2]^{1+}$. Because these molecules contain only one magnetic center, the concepts that were introduced for $[Fe(XH)_4]$ can be conveniently applied. Finally, the electronic structures of the homo-valent $[Fe_2X_2]^{2+}$ and the mixed-valent $[Fe_2X_2]^{1+}$ clusters were analyzed in detail, which included the composition of the CASSCF wave function, the active orbitals, and the energies of low-lying excited states.

Geometry

In comparison to sulfur, a selenium atom has 18 more electrons and the atomic radius is about 0.15 Å larger, making it spatially more demanding. The Fe−Se bond is therefore consistently longer compared to an analogous Fe−S bond and the elongation is 0.1 Å–0.2 Å for both the monomeric $[Fe(XH)_4]^{1-,2-}$ complexes and the dimeric $[Fe_2X_2]^{2+,1+}$ clusters (Table 2.1 and 2.6). For the $[Fe_2X_2]^{2+,1+}$ clusters, each S→Se increases Fe−Fe distance by 0.05 Å, consistent with the XRD structures reported for these and similar compounds.[23,180,285] Changes in the Fe−Fe distance are expected to have an immediate effect on the magnetic coupling, which is related to the electron hopping between the Fe centers.[104,109]

The BS-DFT optimized geometries of the mixed-valent $[Fe_2X_2]^{1+}$ clusters show that the orientation of the $NacNac^-$ ligand on Fe_A is sensitive to the S→Se substitution (Figure 2.11). Due to the dispersion interactions between the phenyl moieties bound to the $NacNac^-$ ligands (Figure 2.10), the larger Se bridging ligands results in a slightly different minimum on the potential energy surface. Therefore, the perturbation introduced by the S→Se substitution may be sufficient to induce geometric changes, given a suitable flexibility in the chemical system. The local electronic structure on Fe_A, in turn, is susceptible to the orientation of the $NacNac^-$ ligand (Figure 2.14).

Donor strength of S- and Se-based ligands

The AOM parameters have been determined for the S- and Se-based ligands in the $[Fe(XH)_4]^{1-,2-}$ monomers (Table 2.5) and in the diamagnetically substituted $[FeGaX_2]^{1+}$ and $[FeZnX_2]^{1+}$ clusters (Table 2.7). S→Se substitution for XH^- reduces e_σ, and therefore the σ donor strength, and is responsible for the decrease between e and the t_2 set in the LFT orbitals. The same trend has been observed both experimentally and computationally for S- and Se-based ligands in Co^{3+} and Cr^{3+} complexes.[280,281,284] Furthermore, the $e_{\sigma\pi c}$ parameter, which captures the misalignment of the ligand orbitals relative to the metal center (Figure 2.9), is smaller for SeH^-. Therefore, the covalent X−H bond has a smaller influence on the donor strength in Se-based ligand compared to S, which is probably a result of the larger size of Se. The e_σ and e_π parameters for the bridging X^{2-} ligands also decrease upon S→Se substitution, which is consistent with the AOM parameters for XH^- ligands.

FeS complexes exhibit a particularly covalent metal-ligand bond. It is important to keep in mind that the CASSCF wave function is biased towards an ionic electronic structure and potential metal-ligand covalency can be only partly recovered by including electron correlation with NEVPT2.[279] Therefore, the quantification of the metal-ligand covalency according to the AILFT and AOM parameters is most likely underestimated.

Ground state wave function

The ground state CASSCF wave function is dominated by the neutral electron configuration (CFG), where the oxidation states are $[Fe^{3+}Fe^{3+}]$ for the homo-valent cluster and $[Fe_A{}^{2+}Fe_B{}^{3+}]$ in the mixed-valent cluster (Figure 2.17 and 2.27, respectively). The S→Se substitution affects the amount of excited CFGs in the ground state wave function similarly for both redox states. For instance, electron hopping between Fe_A and Fe_B decreases, *i.e.* MMCT and IVCT CFGs, which is in line with the elongation of the Fe−Fe bond. At the same time, the electron transfer from the ligands to the metal increases, *i.e.* the LMCT CFGs, which is consistent with an easier reduction of Se (Table 1.1). The MMCT and LMCT CFGs facilitate antiferromagnetic coupling via kinetic exchange or superexchange, respectively.[109] Therefore, it is not immediately obvious from the ground state wave function how the magnetic coupling is affected by the S→Se substitution.

While the electronic structure in the homo-valent dimers is qualitatively equivalent for all bridging ligands, going from S to Se changes the nature of the ground state in the mixed-valent dimers (Figure 2.27f). For the $[Fe_2S_2]^{1+}$ and $[Fe_2SSe]^{1+}$ clusters, the z^2 orbital on Fe_A is doubly occupied in the leading CFG of the ground state wave function. However, for the $[Fe_2Se_2]^{1+}$ cluster, the x^2-y^2 orbital is doubly occupied. The AILFT analysis of the diamagnetically substituted models shows that the z^2 and x^2-y^2 orbitals for the Se_2-bridged complex are close in energy ($[FeGaSe_2]^{1+}$ in Figure 2.13). The change in the electronic structure in going from $[Fe_2S_2]^{1+}$ to $[Fe_2Se_2]^{1+}$ is therefore a combined effect of changes in the local ligand field and changes in the metal-metal interactions due to the substitution of the bridging X ligands.

Spin-coupled states

The lowest-energy states in the homo-valent $[Fe_2X_2]^{2+}$ dimers form a spin ladder ranging from the diamagnetic $S = 0$ ground state to the $S = 5$ high spin state. If the HDvV spin Hamiltonian is valid (Eq 1.8), the coupling constant J can be conveniently extracted from any state pair via the Landé interval rule (Eq 1.9). However, it becomes quickly apparent that the relative state energies depend strongly on the state-averaging method used to converge the CASSCF wave function (Figure 2.18). For state-averaged energies, the Landé interval applies fairly well, as can be seen from the small standard deviation of fitting J to all energy differences (Table A7). For state-specific energies, however, the highest and lowest spin states are stabilized relative to the state-averaged energies, while the intermediate spin states remain virtually unchanged. As a consequence, the Landé interval rule breaks down: the coupling constant extracted from the $S = 0, 1$ state pair is significantly larger compared to that extracted from the $S = 4, 5$ pair (Figure 2.19). The fit of the state-specific energies can be improved by including a biquadratic term in the spin Hamiltonian (Figure 2.20), which decreases the standard deviation by an order of magnitude.

The experimentally determined coupling constants are on the order of $J = -200\,\mathrm{cm}^{-1}$ for $[Fe_2S_2]^{2+}$ clusters similar to the ones studied in the present work.[141] Therefore, only the $S = 0$ and the $S = 1$ state are significantly populated at room temperature and contribute to J.[†] Therefore, the focus should lie on the energy difference between the low spin states when comparing the calculated coupling constants to experimental values. Furthermore, the bias towards higher spin states in the state-averaged wave function artificially decreases the energy differences between the lower spin states. Therefore, the state-specific energies of the $S = 0$ and $S = 1$ should be used to compare the calculated J values with the experiment ($J_{0\to1}$ in Figure 2.19b).

The substitution of the bridging ligands reduces the calculated coupling constant by about $10\,\mathrm{cm}^{-1}$ in going from the $[Fe_2S_2]^{2+}$ to the $[Fe_2Se_2]^{2+}$ cluster (Figure 2.19). A similar decrease has been observed experimentally for protein-bound and similar synthetic $[Fe_2X_2]^{2+}$ clusters.[143,180,183] The trend is also consistent with the decrease of MMCT CFGs in the $S = 0$ ground state wave function (Figure 2.17), which contribute to antiferromagnetic coupling through kinetic exchange.[109] The LMCT CFGs, which contribute to antiferromagnetic coupling via superexchange, also increase upon S→Se substitution, but their influence appears to be outweighed by kinetic exchange.

The mixed-valent clusters exhibit a valence-localized $S = \frac{1}{2}$ ground state. The effective coupling constant J^{eff} (Eq 1.13) determined from the energy difference between the $S = \frac{1}{2}, \frac{3}{2}$ state pair is $-124\,\mathrm{cm}^{-1}$, $-87\,\mathrm{cm}^{-1}$, and $-66\,\mathrm{cm}^{-1}$ for $[Fe_2S_2]^{1+}$, $[Fe_2SSe]^{1+}$, and $[Fe_2Se_2]^{1+}$, respectively (Table 2.8). This is in stark contrast to experimental measurements for similar FeS dimers, where the magnitude of the coupling constant decreases by merely 10 % in going from $[Fe_2S_2]^{1+}$ to $[Fe_2Se_2]^{1+}$.[143,151,180] The electron hopping between the metal centers as well as from the ligands to the metal shows a similar dependency on the bridging ligands as in the $[Fe_2X_2]^{2+}$ clusters (Figure 2.27). However, the calculated energy difference between the $S = \frac{1}{2}$ and the $S = \frac{3}{2}$ spin states is compatible with the recently reported EPR spectrum of $[Fe_2Se_2]^{1+}$, which exhibits both $S = \frac{1}{2}$ and $S = \frac{3}{2}$ features.[23] Here, the authors suggest that a combined decrease

[†] At $T = 300\,\mathrm{K}$, the available energy to thermally populate excited states is on the order of $k_\mathrm{B}T = 209\,\mathrm{cm}^{-1}$, where k_B is the Boltzmann constant.

in Heisenberg exchange, increase in double exchange, and decrease in vibronic coupling relative to $[Fe_2S_2]^{1+}$ is responsible for the $S = \frac{3}{2}$ spin state observed for $[Fe_2Se_2]^{1+}$. It is therefore well possible that the 50 % decrease in J^{eff} in going from $[Fe_2S_2]^{1+}$ to $[Fe_2Se_2]^{1+}$ arises from a growing contribution of the double exchange mechanism, which favors ferromagnetic alignment.

Electronically excited states

The first $S = 0$ excited state in the homo-valent $[Fe_2X_2]^{2+}$ dimers is an LMCT state from the bridging ligand to Fe (Figure 2.21). The excitation energy of the LMCT state depends strongly on the nature of the bridging ligand and decreases from about $10\,200\,\text{cm}^{-1}$ in $[Fe_2S_2]^{2+}$ to about $8800\,\text{cm}^{-1}$ in $[Fe_2Se_2]^{2+}$. The lowest-energy LMCT state in the experimental spectra of $[Fe_2S_2]^{2+,1+}$ clusters has not been unambiguously assigned, but has been proposed to appear in the near-IR/vis region,[140] which is in good agreement with the calculated energy ($10\,200\,\text{cm}^{-1} = 925\,\text{nm}$). Furthermore, two bands between $20\,000\,\text{cm}^{-1}$ and $25\,000\,\text{cm}^{-1}$ in the UV/vis spectra of protein-bound $[Fe_2X_2]^{2+}$ clusters have been proposed to arise due to LMCT excitations.[178,179] These bands are shifted by around $1000\,\text{cm}^{-1}$ to lower energies upon S→Se substitution, which is in excellent agreement with the calculated shift.

The low-energy spectrum for the mixed-valent $[Fe_2X_2]^{1+}$ clusters contains significantly more states than the homo-valent $[Fe_2X_2]^{2+}$ clusters due to d-d_n and $IVCT_n$ excited states, which furthermore give rise to an individual spin ladder (Figure 2.28). The four d-d_n states, where the minority spin electron on $Fe_A{}^{2+}$ is excited locally, occur within $8500\,\text{cm}^{-1}$ of the ground state. Their energies are lowered by the S→Se substitution, but the decrease varies, between $100\,\text{cm}^{-1}$ and $1000\,\text{cm}^{-1}$, depending on the respective d-d_n state. The energy trend follows accurately AILFT orbitals in the diamagnetically substituted $[FeGaX_2]^{1+}$ (Figure 2.13). This highlights that the d-d_n excitations are mostly local excitations. The LFT picture supplies an easy explanation for the different sensitivity of the d-d_n state energies on the bridging ligands: if the electron is excited into an orbital which has a large Fe−X overlap, this corresponding state will be more strongly affected by the S→Se substitution.

The calculated shifts are consistent with IR measurements in protein-bound $[Fe_2S_2]^{1+}$ clusters, where S→Se substitution has been shown to shift the bands of the d-$d_{4,5}$ transitions at $4500\,\text{cm}^{-1}$ and $6000\,\text{cm}^{-1}$ to $4550\,\text{cm}^{-1}$ and $5550\,\text{cm}^{-1}$, respectively.[140,173,175,179] Furthermore, the d-d_1 excitation energy in $[Fe_2S_2]^{1+}$ is $470\,\text{cm}^{-1}$, consistent with the energy determined through the temperature dependence of the quadrupole splitting in Mössbauer measurements of protein-bound $[Fe_2S_2]^{1+}$ clusters.[290] A similar measurement for $[Fe_2Se_2]^{1+}$ could verify the lower calculated d-d_1 energy ($170\,\text{cm}^{-1}$) and orbital occupation of ground state wave (Figure 2.27f).

In an IVCT excitation, an electron is transferred from Fe_A to Fe_B, resulting in the $Fe_A{}^{3+}Fe_B{}^{2+}$ valence isomer. Similar to the local states on Fe_A, a total number of five $IVCT_n$ states are expected, where the excited electron occupies different orbitals on Fe_B. In stark contrast to the d-d_n and LMCT excited states, S→Se substitution leads to an increase in the excitation energy of lowest IVCT state from about $6600\,\text{cm}^{-1}$ in $[Fe_2S_2]^{1+}$ to about $7200\,\text{cm}^{-1}$ in $[Fe_2Se_2]^{1+}$. This is most likely a result of the Fe−Fe bond elongation. The IVCT transition has been tentatively assigned to a feature around

11 000 cm^{-1} for protein-bound $[Fe_2S_2]^{1+}$ clusters, where the electrostatic environment of the protein would have a pronounced influence on the excitation energy.[140] Since S→Se substitution raises the energy of the IVCT state, this behavior could be used to assign the experimental features in the near-IR/vis/UV spectra.

For the electronic ground state d-d_0, the splitting between the $S = \frac{1}{2}$ and the $S = \frac{3}{2}$ state decreased by about 50 % in going from $[Fe_2S_2]^{1+}$ to $[Fe_2Se_2]^{1+}$ ($J^{\text{eff}}_{d\text{-}d_0}$ in Table 2.8). The magnitude of J^{eff} is different in the d-d_n and $IVCT_n$ excited states, since the excited electron changes its orientation with respect to the bridging ligands. For the $IVCT_0$ excited state, the coupling constant is actually slightly higher for the Se-based cluster. Therefore, how strongly and in which direction S→Se substitution affects the magnetic interactions in the mixed-valent $[Fe_2X_2]^{1+}$ clusters depends critically on the orientation of the doubly occupied orbitals, which could be controlled by the capping ligands or the protein environment.

Outlook

The computational study of the synthetic $[Fe_2X_2]^{2+,1+}$ clusters has shown that S→Se substitution can have a considerable effect on the magnetic interactions between Fe centers. This has also been observed experimentally for FeS dimers and cubanes.[23,143,151,180,183] In FeMoco, a large FeS cluster in nitrogenase, antiferromagnetic coupling is an important contribution to the stability of the electronic structure (Section 4.2.2).[291] Selective S→Se substitution has already been reported for FeMoco and no qualitative changes in the electronic structure have been observed based on XAS, EXAFS, and EPR spectroscopy.[170,171] However, it is not obvious if the replacement affects the energies of intermediates during substrate reduction due to changes in the antiferromagnetic coupling and thus alters reactivity in a subtle way. As a starting point, the energies of the BS-DFT determinants could be calculated for substrate-free Se-substituted FeMoco and compared with unsubstituted FeMoco (Figure 4.9).

Furthermore, the binding of the strong field ligand CO to FeMoco favors local spin pairing and therefore competes with antiferromagnetic coupling between the Fe centers (Chapter 5). S→Se substitution in FeS clusters reduces the spin-coupling, which would favor CO binding, but also decreases the local ligand field splitting, which would disfavor CO binding. It would therefore be interesting to determine how the frequency of CO-bound FeMoco would change upon Se-substitution.

Many excited CFGs that are important for the magnetic Fe−Fe interaction (MMCT and LMCT) are included variationally in the CAS(22,16) and CAS(23,16) wave functions for the $[Fe_2X_2]^{2+}$ and $[Fe_2X_2]^{1+}$ clusters. The perturbative NEVPT2 correction incorporates the effect of the excited CFGs beyond the active space only to a certain extent. These are, however, crucial for an accurate description of the magnetic interactions in metal clusters,[109] but a further extension of the active space with the CASSCF/NEVPT2 strategy is currently computationally too demanding. Promising developments to approximate CASSCF wave functions with larger active spaces are, for example, based on the density matrix renormalization group (DMRG),[292,293] the iterative configuration expansion (ICE),[294,295] or full CI quantum Monte Carlo (FCIQMC).[296–298] An accurate calculation of the magnetic interactions is vital to observables that depend on relativistic effects, such g values of $[Fe_2X_2]^{1+}$ clusters, which shift upon S→Se

substitution,[166,173–175] since they depend strongly on accurate energies for the low-lying states.[92]

The dramatic decrease in the effective coupling constant in going from $[Fe_2S_2]^{1+}$ to $[Fe_2Se_2]^{1+}$ is likely linked to an increase in double exchange and a decrease in vibronic coupling. If these contributions are of comparable magnitude, Eq 1.13 is not applicable and the spin state energies can not be fitted with an effective coupling constant. Even though the ground state electronic structure is localized, the interplay between Heisenberg exchange, double exchange, and vibronic coupling is not obvious from the calculations at a fixed geometry. Therefore, it would be advisable to track the spin state energies along an asymmetric vibrational mode of the $[Fe_2X_2]$ core, similar to Figure 1.3c. In doing so, one could determine the Heisenberg exchange parameter J, the double exchange parameter B, and the vibronic coupling parameter $\frac{\lambda^2}{k}$ for the computational $[Fe_2X_2]^{1+}$ models. A natural extension would be to include the $[Fe_2Te_2]^{1+}$ cluster in the analysis, which has recently been shown to exhibit a pure $S = \frac{3}{2}$ ground state.[23]

3. Valence-delocalized $[Fe_2S_2]$ ferredoxins

Variants of ferredoxins from *Clostridium pasteurianum* (*Cp*) and *Aquifex aeolicus* (*Aae*) are the only known examples for valence-delocalized $S = \frac{9}{2}$ FeS dimers.[24,98] The respective wild type ferredoxins exhibits a valence-localized $S = \frac{1}{2}$ ground state, but substitution of a terminal Cys ligand with Ser results in a mixture of the $S = \frac{1}{2}$ and $S = \frac{9}{2}$ ($> 80\,\%$) spin state. Furthermore, the spin states depend on the proton concentration: at low pH, only the $S = \frac{1}{2}$ state is observed, but at high pH, the mixture of spin states is observed (see Section 1.2.1 for more details).

The present chapter is a brief exploration of the electronic structure of the $[Fe_2S_2]^{1+}$ ferredoxins from *Cp* and *Aae*. The computational models, which are shown in Figure 3.1, are based on the XRD structure reported for the wild type and variant ferredoxins from *Aae*.[149] The influence of the Cys→Ser substitution as well as the proton concentration on the stability of the two spin states is analyzed with the help of BS-DFT calculations. Valence delocalization is also a characteristic feature of the electronic structure in higher-nuclearity FeS clusters and will be discussed in the context of FeMoco, the active site of nitrogenase, in Chapter 4 and 5.

3.1. Computational details

The calculations were performed with `ORCA` release version 4.2.[75–77] The models for the $[Fe_2S_2]^{1+}$ ferredoxins from *Cp* and *Aae* were created from the XRD structures solved for *Aae*.[149] The $[Fe_2S_2]$ core was cut out from the XRD structures of the wild type (PDB `1M2A`) and the C55S variant (PDB `1M2B`). For the *Cp* ferredoxin, no XRD structure has been reported, but the structure is expected to be very similar to *Aae*, since both protein sequences are highly homologous.[148,149] The terminal Cys and Ser residues were substituted by CH_3S^- and CH_3O^- (see Figure 3.1). The geometries were optimized with BS-DFT (`flipspin` keyword, $M_S = \frac{1}{2}$) with and without a proton on Cys55/Ser55 while constraining the Fe−S/O bond lengths to the XRD structure. Furthermore, the protein environment is crudely modeled with a C-PCM ($\varepsilon = 4$) using the Gaussian charge scheme.[266] All other details of the level of theory are equivalent to the geometry optimizations of the $[Fe_2X_2]$ model in Chapter 2 (see Section 2.1).

3.2. Electronic structure of $[Fe_2S_2]^{1+}$ ferredoxins

The computational models for the ferredoxins from *Cp* and *Aae* are shown in Figure 3.1. The wild type model (wt) has four terminal $CH_3{-}S^-$ ligands. In the Cys→Ser

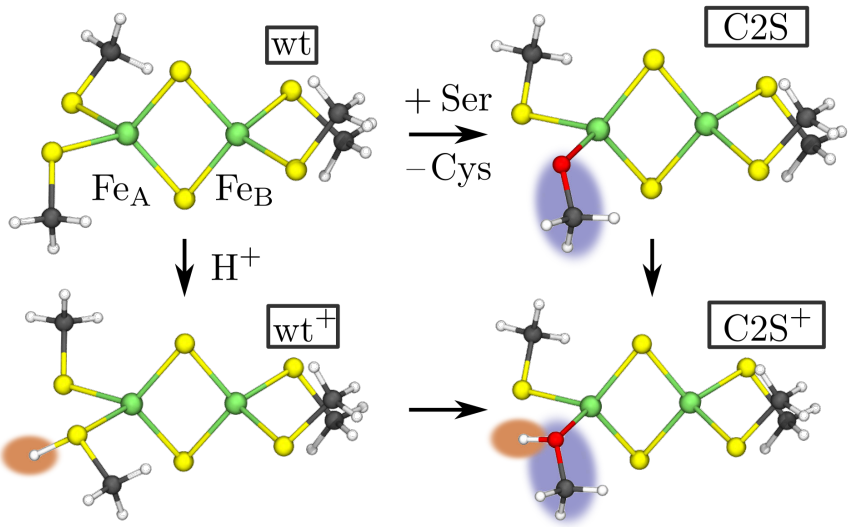

Figure 3.1. – Computational models for the $[Fe_2S_2]$ ferredoxins from *Cp* and *Aae*. The wild type (wt) and the Cys→Ser variant (C2S) models correspond to high pH conditions and the respective protonated models (wt^+ and $C2S^+$) to low pH conditions.

substituted model (C2S), one of the terminal $CH_3{-}S^-$ is replaced with $CH_3{-}O^-$. The C2S model corresponds to the C55S variant from *Aae* and the C56S variant from *Cp*. Furthermore, to model high pH conditions, all four terminal ligands are unprotonated (wt and C2S) and to model low pH conditions, one terminal ligand is protonated to $CH_3{-}SH/CH_3{-}OH$ ($wt^+/C2S^+$). Only the reduced $[Fe_2S_2]^{1+}$ redox state is discussed.

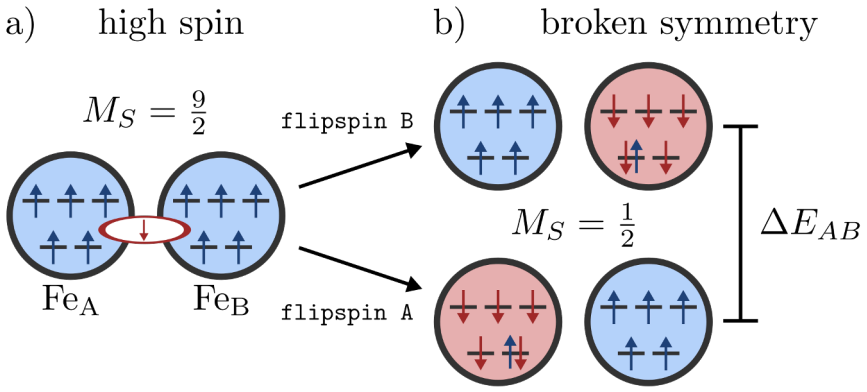

Figure 3.2. – Schematic representation of (a) the high spin determinant in a mixed-valent $[Fe_2S_2]^{1+}$ cluster and (b) the two possible BS determinants. ΔE_{AB} is the energy difference between the BS determinants, if the local redox potential of the two Fe centers is different.

In the high spin $M_S = \frac{9}{2}$ DFT determinant of all models, the magnetic orbitals have predominantly Fe 3*d* character. The localized orbital analysis, which has been introduced in Section 1.1.6, reveals that the both Fe are local high spin centers and share one minority spin electron (Figure 3.2a). The high spin determinants therefore

Table 3.1. – Summary of the DFT calculations for the models shown in Figure 3.1. The Fe−S/O bond lengths were constrained to the XRD structure.[149] ΔE_{AB} is the difference between the two possible $M_S = \frac{1}{2}$ BS determinants (Figure 3.2). The delocalized orbital (deloc. orb.) in the high spin $M_S = \frac{9}{2}$ determinant is shown in Figure 3.3. Orbitals were localized with the Foster-Boys algorithm; atomic contributions are based on Hirshfeld populations. The values for J were obtained with the Yamaguchi projection.

	d(Fe$_A$−S) [Å]	d(Fe$_B$−S) [Å]	ΔE_{AB} [cm^{-1}]	deloc. orb. Fe$_A$ [%]	deloc. orb. Fe$_B$ [%]	J [cm^{-1}]
wt	2.229/2.250	2.236/2.212	2150	45.8	35.4	-110.5
wt$^+$			2692	65.4	20.0	-100.6
C2S	2.202/2.305	2.215/2.262	3	41.3	41.7	+9.3
C2S$^+$			3616	67.8	19.8	-92.8

exhibit a valence-delocalized electronic structure with the oxidation states $Fe^{2.5+}Fe^{2.5+}$.

In order to obtain broken-symmetry (BS) determinants, the spin parts of the orbitals on one of the Fe centers are flipped, and the determinant is converged with $M_S = \frac{1}{2}$. This causes the formerly delocalized electron to localize on the spin down Fe center (Figure 3.2b). Since the models are not symmetric, flipping Fe$_A$ or Fe$_B$ results in distinct final energies, which are separated by ΔE_{AB}. The two BS determinants correspond to the two possible valence isomers $Fe_A^{2+}Fe_B^{3+}$ and $Fe_A^{3+}Fe_B^{2+}$.† ΔE_{AB} is therefore the BS-DFT equivalent for the IVCT excitation energies based on CASSCF states discussed for the synthetic $[Fe_2X_2]^{1+}$ clusters in Section 2.3.5.

As can be seen for ΔE_{AB} in Table 3.1, the $Fe_A^{2+}Fe_B^{3+}$ valence isomer is clearly favored for the wt, wt$^+$, and C2S$^+$ models by 2000 cm^{-1}–4000 cm^{-1}. Even for the wt model, where the type of ligands is the same for the two Fe centers, the geometry of the terminal ligands makes Fe$_A$ the preferred center of reduction. However, for the C2S model, the two valence isomers are virtually degenerate. Since the Fe−S/O bond lengths are exactly the same in C2S and C2S$^+$, the change in ΔE_{AB} can be attributed entirely to the proton on CH_3−OH. Therefore, the local redox potentials are nearly identical for Fe$_A$ and Fe$_B$, – they are equipotent – even though their chemical environment appear rather different.

Differences in the local redox potential are also observable in the high spin $M_S = \frac{9}{2}$ determinant. The delocalized minority spin orbital is partly localized in the wt, wt$^+$, and C2S$^+$ model with an at least 10 % higher contribution from Fe$_A$. In the C2S model, however, the electron is evenly shared between Fe$_A$ and Fe$_B$, 41.3 % and 41.7 %, respectively, indicating that both centers are equipotent. The uneven distribution is not immediately obvious from visual inspection, as can be seen in the orbital plots for

†Alternatively, the center of localization can be controlled with the sign of the M_S value of the BS determinant: setting $M_S = -\frac{1}{2}$ forces the electron onto the spin up center. Since `ORCA` only allows $M_S \geq 0$, the localization can only be controlled via the choice of the to-be-flipped center. Extracting the coupling constant from the higher-energy BS determinant will artificially shift J to more positive values. Therefore, if the lower-energy BS determinant is not obvious, one has to test for both options. Recognizing the most stable M_S value for a given BS determinant is even more complicated for the complex spin coupling patterns in FeMoco.

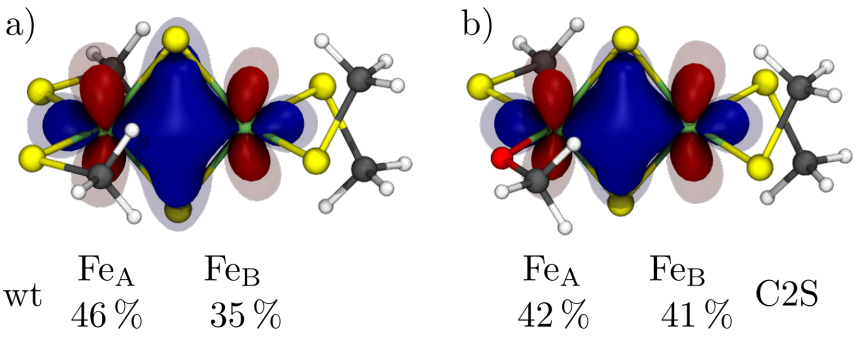

Figure 3.3. – Delocalized orbital in the high spin $M_S = \frac{9}{2}$ determinant for (a) the wt model and (b) the C2S model. Orbitals were localized with the Foster-Boys algorithm; atomic contributions are based on Hirshfeld populations. Isosurfaces are shown at ±0.05 and ±0.025.

the wt and the C2S model in Figure 3.3.

The coupling constants J calculated with BS-DFT are on the order of $-100\,\mathrm{cm}^{-1}$ for the wt, wt^+, and $\mathrm{C2S}^+$ model (Table 3.1). This fairly strong antiferromagnetic coupling is consistent with a valence-localized $S = \frac{1}{2}$ ground state observed experimentally for the wild type ferredoxins and the Cys→Ser variant under low pH conditions.[24,145–149] Interestingly, the coupling constant for the C2S model is $+9.3\,\mathrm{cm}^{-1}$ and therefore close to zero; the high spin and the BS determinant are nearly degenerate. This is in agreement with the experimentally observed mixture of the $S = \frac{1}{2}$ and $S = \frac{9}{2}$ spin states in the Cys→Ser variants under high pH conditions.

3.3. Conclusion

This chapter presented DFT calculations for the *Cp* and *Aae* $[Fe_2S_2]^{1+}$ ferredoxins. Cys→Ser variants of these ferredoxins are the only known $[Fe_2S_2]^{1+}$ clusters that exhibit an $S = \frac{9}{2}$ valence-delocalized ground state.[24,145–149] At high pH, they exist as a mixture of an $S = \frac{9}{2}$ and an $S = \frac{1}{2}$ spin state, but at low pH, only the $S = \frac{1}{2}$ is observed. The computational models covered the wild type and Cys→Ser variant at high pH (wt and C2S, respectively) and at low pH (wt^+ and $\mathrm{C2S}^+$, respectively, Figure 3.1). The coupling constants J for the wt, wt^+, and $\mathrm{C2S}^+$ models are on the order of $-100\,\mathrm{cm}^{-1}$, which indicates strong antiferromagnetic coupling in the $M_S = \frac{1}{2}$ BS determinant (Table 3.1). For the C2S model, on the other hand, $J = +9.3\,\mathrm{cm}^{-1}$, the $M_S = \frac{9}{2}$ high spin determinant and the $M_S = \frac{1}{2}$ determinant are therefore nearly degenerate. This is in good agreement with the experimentally observed mixture of the $S = \frac{9}{2}$ and $S = \frac{1}{2}$ spin states for the Cys→Ser variant at high pH.

The two localized valence isomers $Fe_A^{2+}Fe_B^{3+}$ and $Fe_A^{3+}Fe_B^{2+}$ corresponding to the two possible $M_S = \frac{1}{2}$ BS determinants nearly have the same energy in the C2S model (Table 3.1). This is not the case for the other models, where electron localization on Fe_A is favored by more than $2000\,\mathrm{cm}^{-1}$. Furthermore, the delocalized electron in the high spin $M_S = \frac{9}{2}$ determinant is shared evenly between Fe_A and Fe_B in the C2S model. In the other models, it is partly localized on Fe_A. Therefore, both the high spin

and the BS determinant suggest that the two Fe centers accommodate the additional electron equally well in the C2S model, but not in the wt, wt^+, and $C2S^+$ models. This supports the hypothesis that the Cys→Ser substitution leads to an accidental degeneracy of the local redox potentials for Fe_A and Fe_B, making them equipotent.[24] Equipotency facilitates electron delocalization and therefore the stabilization of the higher spin states due to the double exchange mechanism (Section 1.1.4).

Outlook

The computational models in this chapter are of rather exploratory character: the terminal Cys/Ser ligands are replaced with CH_3−S/O ligands and the surrounding protein is crudely modeled with an implicit solvation model. It remains to be shown if similar results are obtained for a more realistic QM/MM model, which recreate the electrostatic influence of the protein environment more accurately. Nevertheless, the present calculations demonstrate nicely that valence delocalization can happen between two chemically distinct metal centers and that the local redox potential plays an important role. Ferromagnetically aligned valence-delocalized $Fe^{2.5+}Fe^{2.5+}$ pairs are an essential aspect of the electronic structure in higher-nuclearity FeS clusters,[98] where they are stabilized through antiferromagnetic coupling with neighboring Fe centers. In the DFT models of FeMoco, the active site of nitrogenase, valence-delocalized pairs are also observed,[17] and they will therefore be omnipresent in the coming two chapters.

The projection method for the BS determinant as well as the HDvV Hamiltonian (Eq 1.8) are not valid for valence-delocalized dimers. Therefore, the near-degeneracy of the $M_S = \frac{9}{2}$ high spin and the $M_S = \frac{1}{2}$ BS determinant may not represent the energies of the $S = \frac{9}{2}$ and $S = \frac{1}{2}$ spin states observed experimentally and the small coupling constant calculated for the C2S model may not be accurate. The energies of the high and low spin states could be verified with computational methods that yield spin eigenfunctions, such as CASSCF. Preliminary CAS(11,10) calculations including NEVPT2 predict that the $S = \frac{1}{2}$ spin state is at least $1000\,cm^{-1}$ lower in energy compared to $S = \frac{9}{2}$ for all four models discussed here. This minimal active space underestimates the metal-ligand covalency and at least the occupied valence p orbitals of the bridging S^{2-} ligands should be included in the active space in order to acquire better relative spin state energies (see discussion also in Section 2.3.5).

4. The resting state of nitrogenase

The best-studied state of Mo nitrogenase is the resting state, mostly because it has proven difficult to purify other intermediates of the catalytic cycle and therefore to characterize them spectroscopically.[15] In the resting state, the FeMoco active site gives rise to an $S = \frac{3}{2}$ EPR signal[199] and a similar signal has been observed for solution-extracted FeMoco suggesting that active site is likely stable in solution.[238] The charge of FeMoco in this odd-electron state has been established as $[MoFe_7S_9C]^{1-}$ through a combination of experimental and computational studies.[17,200,201] Assuming the anionic oxidation states S^{2-} and C^{4-}, this leaves 41 valence electrons to be distributed among the metal centers. According to BS-DFT calculations, the Fe centers have local high spin states with oxidation states between Fe^{2+} and Fe^{3+} and the Mo^{3+} center has an unusual non-Hund local spin state.[256] The nature of the resting state in V nitrogenase is still debated,[202] but assuming an odd-electron $S = \frac{3}{2}$ resting state the charge of FeVco is most likely $[VFe_7S_8CO_3C]^{2-}$ with a V^{3+} oxidation state and a one-electron reduced Fe part compared to FeMoco.[19,299]

The present chapter analyzes DFT-based resting state models of nitrogenase. The models were adapted from the QM/MM models of either Mo nitrogenase[17] or V nitrogenase[19] by Benediktsson and Bjornsson. Section 4.1 contains the details for the different models used in the present chapter as well as to a large extent in Chapter 5. Section 4.2 focuses on the intrinsic properties of the electronic structure of the active site. Here, different descriptors with which to interpret the electronic structure are compared, contributions to the stability of the various BS determinants are analyzed, and differences between FeMoco and FeVco are highlighted. Section 4.3 investigates how the protein environment affects the electronic structure. First, different models for the protein environment are compared. Then, the roles of individual amino acid residues are explored by deleting residue charges, creating a QM/MM model for a hybrid nitrogenase, and by introducing residue substitutions. The following Chapter 5 focuses on the interaction of CO with the active sites.

4.1. Computational details

Model sizes

Throughout this chapter and Chapter 5 different models will be used to study certain aspects of nitrogenase. An overview of the models is given in Figure 4.1. The Figure shows the structures for Mo nitrogenase, but V nitrogenase models are analogous given some differences in the protein environment, as explained below. The following list specifies each model in more detail, where the residues in parentheses refer to V nitrogenase. The protein backbone was omitted for all residues, except for Ser278

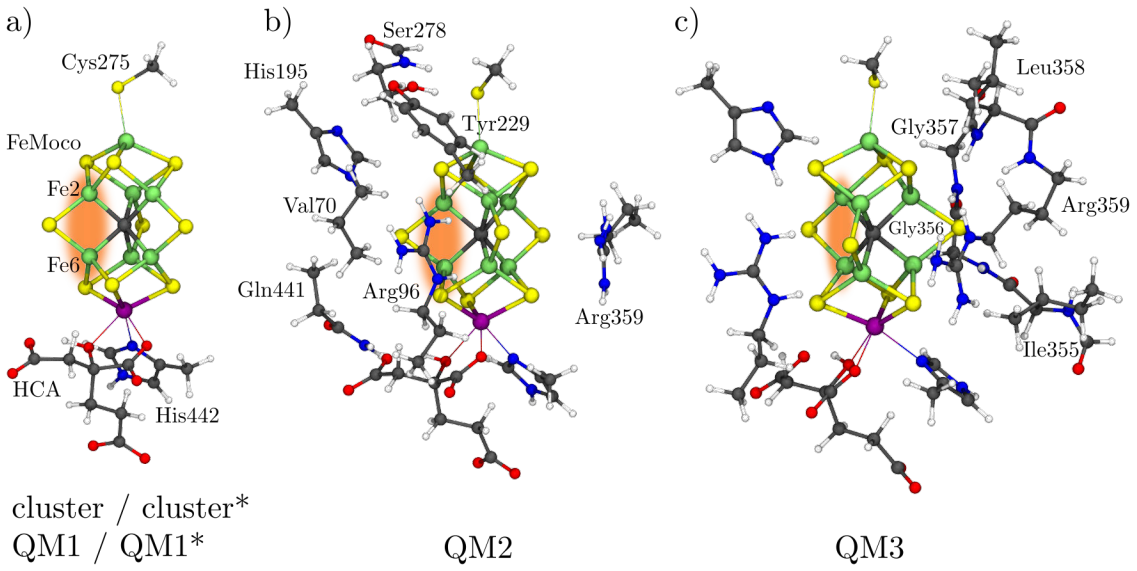

Figure 4.1. – Overview of the computational models for Mo nitrogenase used in this work. V nitrogenase models are analogous, considering the differences in the cofactor and the protein environment. Only the QM part of the respective model is shown: the cluster model is embedded in an implicit solvation model, while the other models are QM/MM models, where the protein environment is described explicitly with a molecular force field. (a) The cluster model and QM1/QM1* model only include FeMoco plus residues with direct coordination to Fe1 and Mo. (b) The QM2 model includes residues that are expected to affect substrate binding at Fe2 and Fe6. (c) The QM3 model includes all residues that form hydrogen bonds with the cofactor. Fe2 and Fe6 are highlighted to guide the eye of the reader.

(Ser260), where the amide hydrogen forms a weak bond with the sulfide of Cys275 (Cys257). The cut bonds were terminated with hydrogen atoms (*vide infra*).

a. The **cluster** model contains FeMoco (FeVco), HCA, His442 (His423), and Cys275 (Cys257). This cluster is not part of a QM/MM embedding, but the surrounding protein is modeled by C-PCM ($\varepsilon = 4$). The initial structures were taken from the published models[17,19] (BS7-235 $M_S = \frac{3}{2}$). Following the initial publication, HCA is a trianion. Only the FeXco (X=Mo,V) coordinates were optimized; HCA, His, and Cys residues were kept fixed in order to avoid conformations that would differ qualitatively from the cofactor cavity. The cluster model can help to separate the explicit effects of amino acid residues from intrinsic properties of the FeXco. The **QM1** model has a QM region that contains the same atoms as the cluster model, while the surrounding protein is modeled with MM. Both models can be used for a fast, but inaccurate screening of different conformations, binding sites, BS determinants, and M_S values. Substrate binding is especially sensitive to the cofactor-environment interaction and a QM/MM interface between a substrate (QM) and a close-lying protein residue (MM) should be avoided.

   The **cluster*** model and **QM1*** model are slimmed-down versions of the cluster model and the QM1 model, respectively. They use a smaller basis set and, in

case of QM1*, a smaller active region (see below for details).

b. The **QM2** region contains QM1 plus His195 (His180), Gln441 (Gln176), Val70 (Val57), Arg96 (Lys83), Arg359 (Arg339), Tyr229 (Phe211), and Ser278 (Ser260). QM2 was designed for studying substrate interactions at Fe2 and Fe6.

c. The **QM3** region contains QM1 plus His442 (His423), His195 (His180), Arg96 (Lys83), Arg359 (Arg339), Leu358 (Pro338), Gly357 (Gly337), Gly356 (Gly336), and Ile355 (Thr335). Therefore, it includes all charged residues in the first residue shell of the cofactor and also those that for hydrogen bonds with FeXco sulfides and the carbonate in FeVco. QM3 was designed to study how differences in the protein environment between Mo and V nitrogenase affect the electronic structure of FeXco.

QM/MM setup

The QM/MM setup follows closely the previously published protocols[17–19,300]. The parameters for the MM region are defined by the `CHARMM36` force field[301] and handled by the `DL_POLY` package.[302] The force field was modified to include the non-bonding parameters for the metal clusters in the XFe protein.[17,19] The `chemshell` setup is made available online.[303] The active region contains roughly 1000 atoms and the total model size is about 37 000 (32 500) atoms. The interaction within the MM region and the interface between the MM and the QM region was handled by a custom version of `chemshell` based on version 3.7[78,79] (`coupling=shift`, see Section 1.1.3 for details). The QM region is calculated with `ORCA` version 4.2.[75–77]

QM level of theory

The electronic structure is described with DFT with the hybrid functional TPSSh (10 % HF exchange).[48,50] The BS determinant was obtained in two steps: 1) For the resting state, a determinant with a multiplicity of 36 (34) was converged. The multiplicity was reduced by 2 for each electron added or CO bound to the cofactor. 2) The `flipspin` keyword was used to invert the spin on 3 Fe centers and the V if present. An all-electron basis set of triple-ζ quality (ZORA-def2-TZVP) was used for all Fe, Mo, V, S, the central carbide, HCA, and additional protons/substrates. A double-ζ basis set (ZORA-def2-SVP) was used for the remaining atoms in the QM region. In the cluster model, the triple-ζ basis was used for all atoms, but in the cluster* model the double-ζ basis was used for all atoms. Relativistic effects were included by using the scalar relativistic Hamiltonian ZORA (ZORA in the basis set name indicates that they were recontracted for ZORA).[260,261] For the evaluation of the Coulomb and HF exchange integrals, the RIJCOSX approximation was used in combination with the auxiliary basis set SARC/J.[264,265] Dispersion forces were approximated by the atom-pairwise parameterization including Becke-Johnson damping (`D3BJ` keyword) proposed by Grimme *et al.*[61,64] The analysis of the electronic structure is carried out with the wave function analyzer program `Multiwfn` version 3.7.[304] This includes Hirshfeld charges and spin populations,[117] fuzzy bond orders (based on Hirshfeld),[127] Bader charges and spin populations,[116] and delocalization indices (based on Bader volumes).[305] Localized

orbitals are obtained with the Foster-Boys algorithm, unless otherwise noted.[128] Orbital contributions are reported as Hirshfeld populations (workflow is made available online, see Ref [303]).

Geometry optimizations

Geometry optimizations are performed with `chemshell`[78,79] for the QM/MM models (`L-BFGS` within `dl_find`[306] and by `ORCA` for the cluster model (`opt` keyword). A modification exists for QM1, called QM1*, in which entire QM region is described with the def2-SVP basis set (without ZORA) and the active region consists only of QM2 (not 1000 atoms). The advantage of this model lies in its performance, and it can be used to screen many BS determinants and M_S values. The most favorable of these can be further refined with a more accurate model.

Diamagnetic substitution

Diamagnetically substituted cofactors were created by substituting all metal centers, except for the centers of interest, by a diamagnetically equivalent ion (see also Section 2.1). The Mo^{3+} center was substituted with In^{3+}. Typically, one would substitute Fe^{3+} with Ga^{3+} and Fe^{2+} with Zn^{2+}, but the average oxidation state of Fe in FeMoco is $Fe^{2.5+}$.[200] Therefore, the electrostatic influence of $Fe^{2.5+}$ may be modeled better with a diamagnetic ion *in between* Fe^{3+} and Fe^{2+} *i.e.* with a fractional oxidation state. Such an ion can be emulated by adjusting the nuclear charge Z in the quantum chemical calculation. To test this approach, all Fe centers have been replaced by either Ga^{3+} ($Z = 31$), $Ga^{2.5+}$ ($Z = 30.5$), $Zn^{2.5+}$ ($Z = 30.5$), and Zn^{2+} ($Z = 30$). Figure A12 shows the difference in the S atomic charges between unsubstituted FeMoco and the diamagnetically substituted FeMoco for the different diamagnetic options. $Ga^{2.5+}$ is found to lead to the lowest average deviation in the S atomic charges and is therefore used as the diamagnetic equivalent for Fe throughout this work. Note that adjusting the nuclear charge while monitoring the ligand atomic charges opens a way to systematically improve diamagnetically substituted models.

Visualization

Throughout this work, atomic or atom pair properties are visualized with schematic illustrations of the cofactor as shown in Figure 4.2. The schematics illustrate how properties are distributed across the cofactor. This way atom-centered properties can be visualized, such as charges and spin populations, as well as atom-pair properties, such as bond orders and bond distances. Molecular structures and orbitals were rendered with `VMD` version 1.9.3.[269,270] Orbitals are shown with isosurfaces at 0.05 and 0.025 (workflow available online[271]).

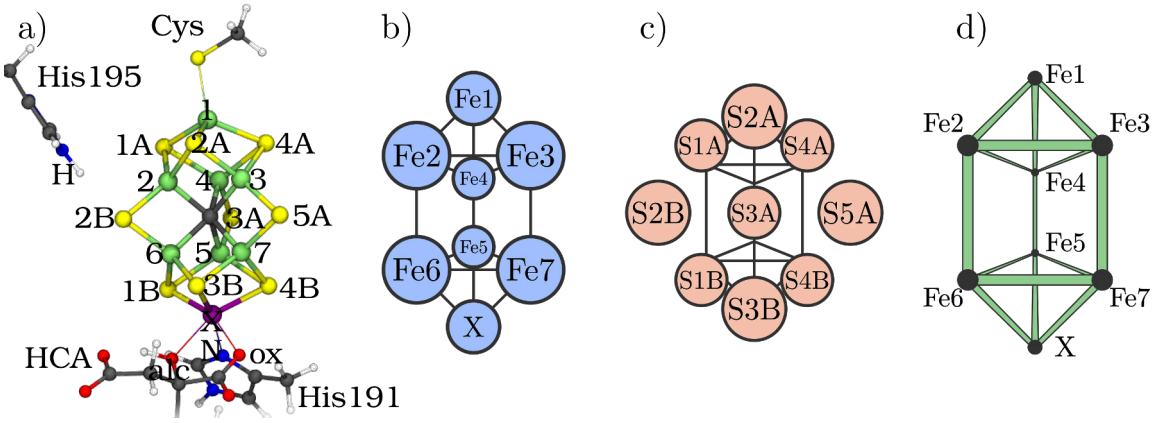

Figure 4.2. – Schematic representations of the cofactor used to visualize atomic properties. (a) Orientation of the atomic labels with respect to His195, HCA, and His191. (b) Schematic used for metal-based properties. (c) Schematic used for ligand-based properties. (d) Schematic used for properties based on metal pairs.

4.2. Electronic structure

In the following, the electronic structure of resting state FeMoco and FeVco is analyzed within the framework of BS-DFT. Section 4.2.1 compares concepts commonly used to interpret the electronic structure, such as localized orbitals, atomic charges, spin populations, and bond orders. Section 4.2.2 considers the full set of BS determinants, that are possible for FeXco, and their stability is rationalized in terms of antiferromagnetic coupling. While the previous sections focus on FeMoco, Section 4.2.3 highlights differences in the electronic structure that arise from the Mo→V and the S3A→CO_3 substitution.

4.2.1. Interpretation of the electronic structure

Localized orbitals

In previous works, the electronic structure of FeMoco according to BS-DFT has been interpreted with the help of metal-based localized orbitals, which have major contributions from most often one, but sometimes two metal centers.[17] Following closely the original work by Benediktsson and Bjornsson, Figure 4.3 shows the localized orbitals for two BS7 determinants in the resting state FeMoco. Contributions from β electrons are shown as negative/blue, therefore, the precise spin-coupling within the BS determinants (BS7-235 and BS7-346) becomes immediately obvious, along with the location of the delocalized electrons. Thus, localized orbitals provide an easy protocol to associate all 41 unpaired electrons in FeMoco with their respective metal center(s).

However, some differences can be found between the values reported for BS7-235 in the original work by Benediktsson and Bjornsson and Figure 4.3. For example, the electron shared between Fe2/Fe3 is only partly delocalized in the original work (35/56 %), but nearly fully delocalized here (34/40 %). Also, the Mo center in the original work was reported as a localized d^3 Mo^{3+}, while Figure 4.3 suggests a more delocalized electronic structure involving Mo with two partly delocalized Mo/Fe orbitals

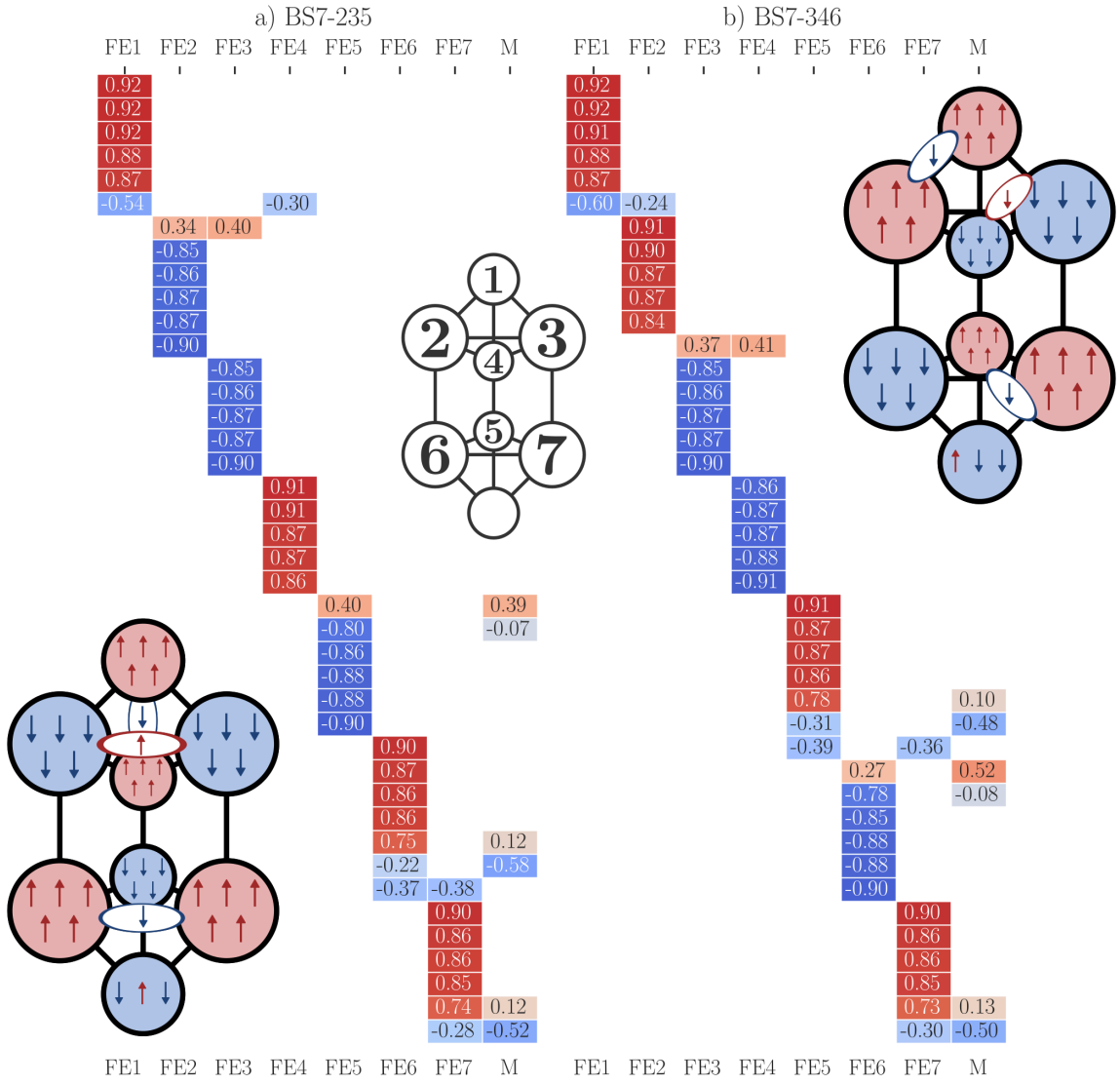

Figure 4.3. – Hirshfeld atomic contributions to the Foster-Boys localized orbitals of (a) the BS7-235 determinant and (b) the BS7-346 determinant in FeMoco (E_0 QM2 model, $M_S = \frac{3}{2}$). Red/positive values corresponds to α electrons and blue/negative values to β electrons. Only the metal-based valence orbitals are shown. The insets show the simplified electronic structure that can be derived from the localized orbital analysis (recreated after Benediktsson and Bjornsson,[17] see Figure 4.2 for atom labels).

(58/22 % and 52/28 %), one delocalized Mo/Fe orbital (39/40 %), and additionally two partly delocalized Fe/Mo orbitals (75/12 % and 74/12 %). Because the authors demonstrated that the localized orbitals are fairly insensitive to the computational details, such as the size of the QM region, the differences most likely stem from either the localization method (original work: PM; here: FB; see Section 1.1.6) or the quantification scheme to count the atomic contributions in each orbital (original work: Mullliken; here: Hirshfeld; see Section 1.1.5). Different localization methods will be discussed in Section 4.2.3 for V nitrogenase, because there they lead to more

pronounced differences. Quantifying atomic contributions based on atomic orbitals can be problematic for large basis sets, as will be discussed for atomic charges in Section 4.2.1. For highly localized orbitals, this is not expected to be significant, but for orbitals involving multiple centers this may lead to different results compared to quantification schemes based on the physical space.

Atomic charges

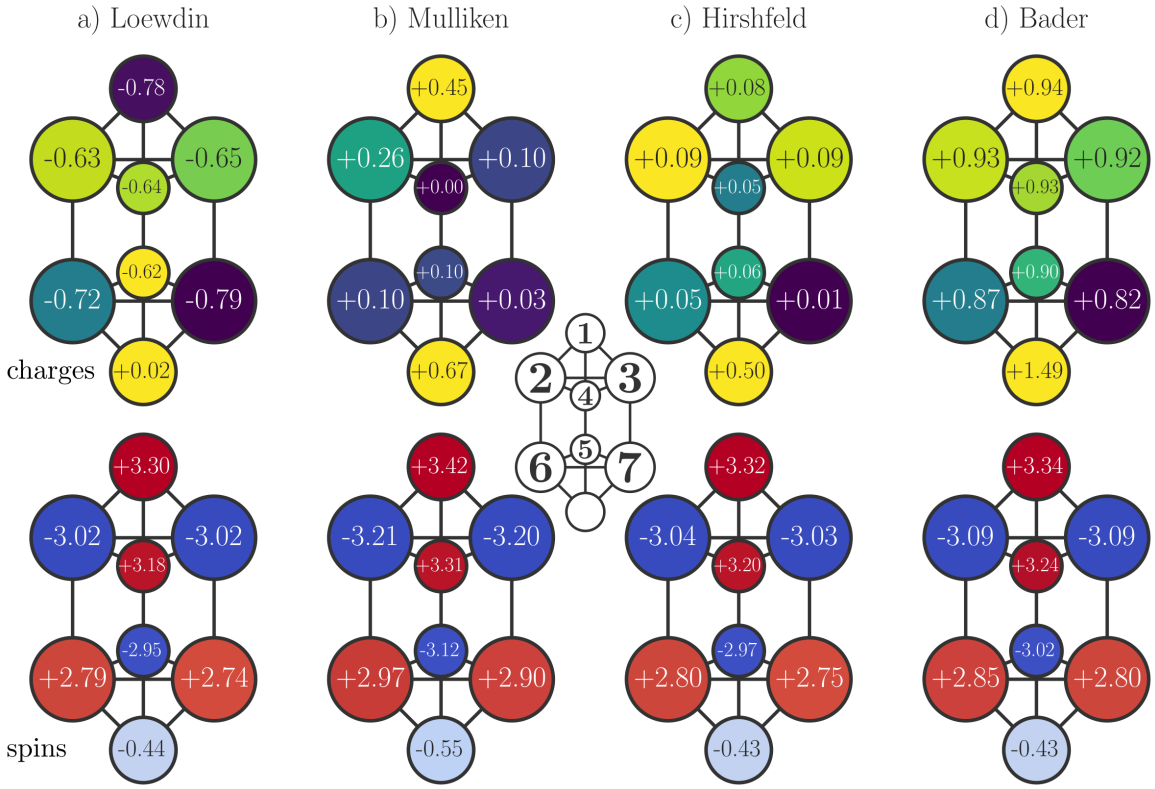

Figure 4.4. – Comparison of different definitions for the atomic charges (top) and spin populations (bottom) for the metal centers in FeMoco (E_0 QM2 model, BS7-235 $M_S = \frac{3}{2}$). See Figure 4.2 for atom labels.

Atomic charges draw an intuitive picture of the charge distribution within a molecule. Different schemes can be used to attribute electron density to a particular atom and they can be divided into whether they partition the atomic basis space (*e.g.* Loewdin or Mulliken) or the physical space (*e.g.* Hirshfeld or Bader). More details on the theoretical background can be found in Section 1.1.5.

The atomic charges for the metals in FeMoco are shown in Figure 4.4. The Fe Loewdin charges range from −0.79–−0.62, while the Fe Mulliken charges range from 0.00–+0.26. Fe1 is one of the most reduced Fe centers according to the Loewdin charge, but the most oxidized according to the Mulliken charge. The Hirshfeld and Bader charges paint a more consistent picture of the relative charges compared to the Loewdin and the Mulliken scheme. The Fe Hirshfeld charges range from +0.01–+0.09 and the Fe Bader charges from +0.82–+0.94. According to both schemes, the Fe centers in

the Fe-only cubane are more oxidized than those in the Mo cubane. The dependence of the Fe charges on the protein environment is discussed in detail in Section 4.3.1. The Mo charge is consistently more positive than the Fe charges for all models, likely due to the less covalent Mo−S compared to Fe−S bonds. The S atomic charges are compared in Figure A9.

The seemingly random distribution of atomic charges according to the Loewdin and Mulliken scheme are not surprising considering the relatively large basis set used to build the atomic basis. In particular, diffuse basis functions centered on a given atom often describe electrons far away from it. As a consequence, the atomic charges are not comparable for the same atom in different chemical environments, for example, for Fe1, which is coordinated by one terminal Cys−S and three μ_3−S^{2-}, and for Fe2, which is coordinated by one μ_2−S^{2-}, two μ_3−S^{2-}, and one μ_6−C^{4-}.

The absolute values for the Hirshfeld charges are smaller than for the other methods, because they describe the shift of the electron density relative to a promolecule, that is constructed with neutral atom densities. The iterative Hirshfeld-I scheme is designed with polar bonds in mind and scales the atomic volume iteratively with the atomic charge, but the high negative charge on the central carbide causes convergence failure. An extension to the Hirshfeld-I scheme is the Hirshfeld-E[307] scheme, which claims to overcome the problems associated with using densities for isolated anions, but it was not tested, because it is not implemented in `Multiwfn` at this time.

It is immediately obvious that none of the charge schemes allows for an easy determination of oxidation states. As suggested by the localized orbital analysis in Figure 4.3, both the Fe2/Fe3 pair and the Fe6/Fe7 pair share a delocalized electron in the BS7-235 determinant, leading to an $Fe^{2.5+}Fe^{2.5+}$ oxidation state. However, while the Fe2/Fe3 pair has similar Hirshfeld charges ($+0.09/+0.09$), the charges for the Fe6/Fe7 pair are quite different ($+0.05/+0.01$). Also, even though the electron shared between Fe1/Fe4 is localized more on Fe1 than on Fe4, Fe1 is more positive than Fe4 according to the Hirshfeld charges. Therefore, the metal atomic charges do not correspond to the electron distribution of the localized orbital analysis.

Spin populations

Similar to the atomic charges, the different partitioning schemes of the electron density can also be used to attribute spin density to a particular atom. Different spin populations are compared in Figure 4.4 for the metal centers in FeMoco. The absolute Fe spin populations range from 2.7–3.5. The Mo spin population is merely -0.5, consistent with a non-Hund local spin (see Figure 4.3).[17,256] The S spin populations are shown in Figure A10.

The magnitude of the Mulliken spins are consistently about +0.2 larger compared to the Loewdin spins. On the other hand, Loewdin, Hirshfeld, and Bader spins differ on average by less than 0.05. In contrast to the atomic charges, the spin populations paint a consistent picture. Because the spin is localized mostly on the magnetic centers, *i.e.* the metal centers, it is less susceptible to the partitioning scheme.

The sign of the spin nicely shows the spin coupling of the BS7-235 determinant, but comparing the absolute magnitudes reveals more subtle details in the electronic structure. For example, the absolute spin on Fe5 is about +0.2 higher compared to the

Fe6/Fe7 pair, which are interpreted as Fe^{3+} and $Fe^{2.5+}Fe^{2.5+}$, respectively, according to a QM/MM model.[17] However, the spin on Fe2/Fe3 is about +0.2 higher compared to Fe6/Fe7, even though both are interpreted as an $Fe^{2.5+}Fe^{2.5+}$ pair. Therefore, the differences in the chemical environment between the Fe-only cubane and the Mo cubane also show in the spin population. Also, the spin on Fe1 being larger compared to Fe4 is not consistent with the localized orbital interpretation (see Figure 4.3), because the shared electron being localized more on Fe1 should lead to more spin cancellation on Fe1.

Bond orders

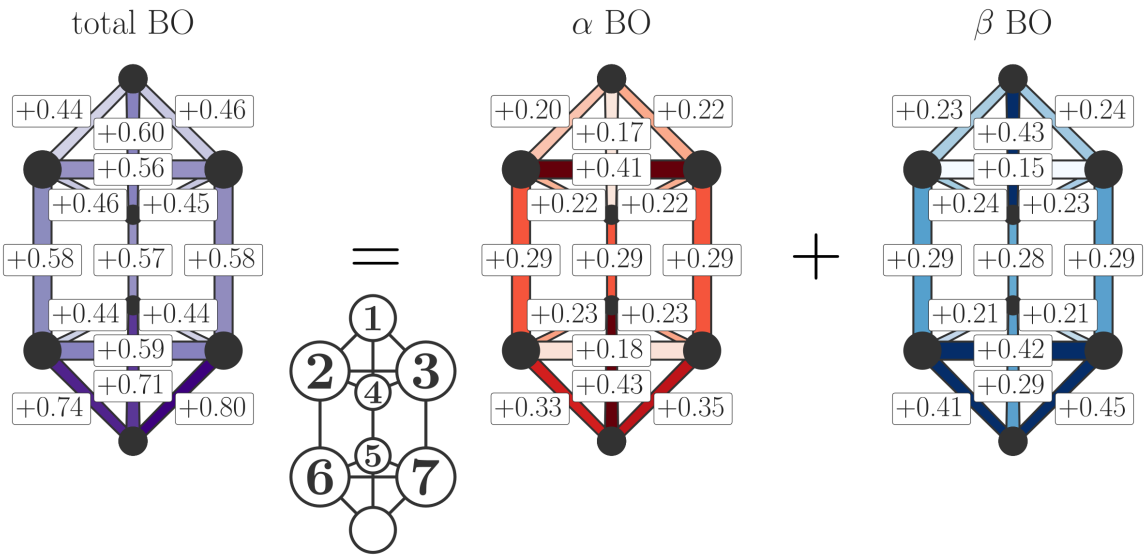

Figure 4.5. – Decomposition of the total fuzzy bond order between metal centers in FeMoco into α and β contributions (E_0 QM2 model, BS7-235 $M_S = \frac{3}{2}$). See Figure 4.2 for atom labels.

The partitioning schemes discussed for the atomic charges can also be used to determine bond orders between atoms from the electron density (see Section 1.1.5 for details). Figure 4.5 shows the fuzzy bond order between the metal centers in FeMoco, which uses the Hirshfeld partitioning. Other definitions of the bond order are compared in Figure A11. The Fe−Fe bond orders range from 0.4–0.6, while the Mo−Fe bond orders are higher with 0.7–0.8, despite the longer Mo−Fe compared to Fe−Fe distances in FeMoco. The bond order between ferromagnetically coupled Fe pairs is 0.10–0.15 higher compared to the respective symmetry-related antiferromagnetically coupled pairs of. For example, Fe2/Fe3 are ferromagnetically coupled, Fe2/Fe4 and Fe3/Fe4 are antiferromagnetically coupled, even though all three pairs are related through a pseudo C_3 rotation. This is consistent with the delocalized pairs in the localized orbital analysis (Figure 4.3).

The spin-resolved bond orders corroborate the relation to the delocalized electrons. For instance, the α part of the electron density is with 0.41 the main contributor to the bond order between Fe2/Fe3, while the β part contributes only 0.15. The bond

orders for the antiferromagnetically coupled belt Fe pairs, *i.e.* Fe2/Fe6, Fe3/Fe7, and Fe4/Fe5, is as high as for the Fe pairs sharing an electron, but the contributions are equal part α and β density. The metal-metal bond order is therefore a convenient metric to describe the location of delocalized electrons in FeMoco.

In the mixed-valent delocalized pairs, *i.e.* Fe1/Fe4, Fe2/Fe3, and Fe6/Fe7, the minority spin contribution is with 0.41 and 0.43 very similar. However, the localized orbitals suggest a weaker bond between Fe1/Fe4 compared to Fe2/Fe3, because the electron is only partly localized in the former, but the minority spin bond order is actually higher. Therefore, a one-to-one correspondence between the bond order and the localized orbitals does not exist.

4.2.2. Stability of BS determinants

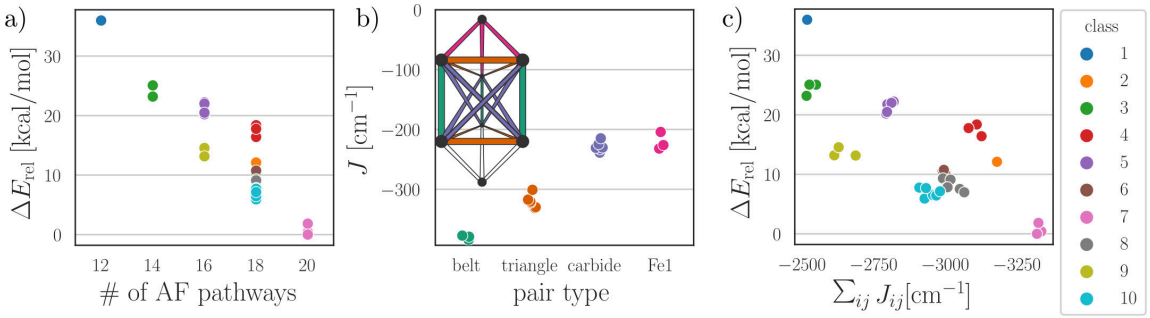

Figure 4.6. – Connection between the energies of the BS determinants in FeMoco and antiferromagnetic coupling between Fe pairs. (a) Dependence of the BS determinant energies on the number of antiferromagnetic pathways present in the respective determinant (adapted from Harris and Szilagyi[291] with the energies for the E_0 $M_S = \frac{3}{2}$ FeMoco cluster* model). (b) Calculated coupling constants for different $Fe^{3+}Fe^{3+}$ pairs in the diamagnetically substituted FeMoco (structures based on the E_0 $M_S = \frac{3}{2}$ FeMoco QM1 model). (c) Dependence of the BS determinant energies on the summed J_{ij} metric (energies for the E_0 $M_S = \frac{3}{2}$ FeMoco cluster* model). See text for further details.

The idea that antiferromagnetic coupling is crucial to the stability of a particular BS determinant has been suggested by Noodleman and coworkers, when they classified the BS determinants based on their symmetry (class BS1 through BS10).[308] They found that those spin arrangements lead to the lower-energy BS determinants that included the largest number of antiferromagnetically coupled Fe pairs. This thought was further refined by Harris and Szilagyi, who weighted each antiferromagnetically coupled pair by the number of bridging ligands involved. In the present work, the number of bridging ligands between two metal centers will be referred to as pathways.[291] For example, the Fe2/Fe3 pair is connected by two pathways, because it is bridged by a μ_3−S^{2-} and the central μ_6−C^{4-}. In contrast to Noodleman's model, the pathway model also counts contributions from Fe pairs that are connected only by the central μ_6−C^{4-}, such as Fe2/Fe7. Figure 4.6a shows the plot of the pathway model using the BS determinant energies of the cluster* model. The plot nicely shows that the BS7 class, which has the

maximum number of 20 pathways connecting antiferromagnetically coupled Fe pairs, constitutes the lowest-energy determinants. However, when comparing the classes BS2, BS4, BS6, BS8, and BS10 – all of which have 18 pathways – it becomes evident that the pathway model does not resolve their relative energies.

The key idea behind the pathway model is that the number of the bridging ligands directly affects the antiferromagnetic coupling strength through superexchange (Eq 1.8). Naturally, the question arises whether the types of the bridging ligands play an important role as well, *i.e.* how different is the $\mu_6{-}C^{4-}$ ligand compared to a $\mu{-}S^{2-}$ or a $\mu_3{-}S^{2-}$ ligand in facilitating antiferromagnetic coupling? Figure 4.6b shows the coupling constants J_{ij} for the Fe pairs in FeMoco. They were calculated for the respective $Fe^{3+}Fe^{3+}$ pair in the otherwise diamagnetically substituted FeMoco (see Section 4.1 for details on diamagnetic substitution). The Fe pairs in each category have similar coupling constants, with the belt pairs exhibiting the strongest coupling (around $-380\,\mathrm{cm}^{-1}$), followed by those in the triangles formed by Fe2/Fe3/Fe4 and Fe5/Fe6/Fe7, respectively (around $-325\,\mathrm{cm}^{-1}$). The Fe pairs bridged only via the central $\mu_6{-}C^{4-}$ and the Fe pairs involving Fe1 have similar coupling constants (about $-220\,\mathrm{cm}^{-1}$). This suggests that the number of bridging ligands is indeed insufficient to describe the antiferromagnetic coupling strength. For example, the Fe pairs in the Fe1 category are bridged by two $\mu_3{-}S^{2-}$ and those in the carbide category by a single $\mu_6{-}C^{4-}$, but both categories have have similar coupling constants.

Figure 4.6c shows the relative energy of the BS determinants as a function of the summed J_{ij} for the respective determinant. Similar to the pathway model in Figure 4.6a, only the antiferromagnetically aligned Fe pairs contribute to the stability metric, but their contributions are weighted according to the coupling constants in Figure 4.6 b. The energies in the summed J_{ij} model follow a similar general trend as in the pathway model. However, the energies of the classes BS8 and BS10 (or BS2 and BS4) decrease with the summed J_{ij} metric, whereas these determinants all have an equivalent weight (18) in the pathway model. Therefore, the summed J_{ij} model constitutes an improved description of stability of different BS determinants.

While calculating the antiferromagnetic coupling constants clearly is a step towards understanding the stability of BS determinants in FeMoco, it is also obvious from Figure 4.6c that the summed J_{ij} model do not capture all the trends in the energies. It appears that the BS classes form at least two energetically separated sets: classes BS1 through BS5 follow a higher-energy trend, while classes BS6 through BS10 follow a lower-energy trend. The spin populations (see Figure A13) suggest some significant differences in the electronic structure for most determinants in the higher-energy set compared to the lower-energy set. For example, the Mo center in the BS1, BS2 and BS3 classes is in a local high spin state as opposed to the non-Hund local spin state in the lower-energy determinants. Also, the determinants of the BS5 class show spin-pairing to a local intermediate spin Fe center in the Fe-only cubane, while the lower-energy determinants all exhibit local high spin Fe centers. Furthermore, BS determinants with an unequal number of α and β Fe centers in the Fe-only cubane (classes BS2 through BS5) do not exhibit two delocalized electrons, which may affect their energies. Altogether, the factors contributing to the stability of a BS determinant is not well understood. Contributions not considered in the summed J_{ij} model could be related to local spin pairing or the delocalization of electrons. However, a good understanding

is necessary, because the nature of the bridging ligand (*e.g.* in FeVco, see Section 4.2.3) or substrate binding (*e.g.* CO, see Chapter 5) further affects the energy ordering, and it is not clear which BS determinants have to be considered.

4.2.3. V nitrogenase

Resting state of the VFe protein

In contrast to the MoFe protein, the charge of the cofactor in the resting state of the VFe protein is not firmly established. A quantitation of the $S = \frac{3}{2}$ EPR features accounts for only $< 10\,\%$ of the sample under resting state conditions (*i.e.* dithionite-reduced) and recent study has proposed an integer spin for the VFe resting state.[202] A high resolution XRD structure (1.35 Å) has been solved in the presence of dithionite, however, no EPR spectrum was reported for the XRD structure sample.[309] An extensive comparison of the XRD structure with a QM/MM model featuring different cofactor charges favors a $[\mathrm{FeVco}]^{2-}$ oxidation state, but it is impossible to discriminate the charge state based on RMSD alone .[19] All QM/MM models for the VFe protein used in this thesis build on this QM/MM model by Benediktsson and Bjornsson.[19] Similar to the $[\mathrm{FeMoco}]^{1-}$ resting state of the MoFe protein, $[\mathrm{FeVco}]^{2-}$ holds 41 metal-based valence electrons. For simplicity the $[\mathrm{FeVco}]^{2-}$ charge state will be referred to as the resting state E_0 (more reduced E_n states follow accordingly). However, one has to keep in mind that the resting state of the VFe protein is still debated.[202]

Localized orbitals have helped with the interpretation of the electronic structure of FeVco as much as with FeMoco.[17,19,299] They suggest a d^2 V^{3+} as opposed to the d^3 Mo^{3+} leading to one-electron reduced Fe part for FeVco (see Figure A14). The additional electron localizes on Fe5 in the BS7-235 determinant of FeVco, which is significantly lower in energy compared to BS7-346 and BS7247. Important features, such as evenly delocalized orbitals between the Fe2/Fe3 and the Fe6/Fe7 pair in the BS7-235 determinant (compare FeMoco in Figure 4.3) are equivalent for FeMoco and FeVco.

Breakdown of the localized orbital analysis

Unfortunately, the definition of localized orbitals is not unique and several approaches have been proposed to obtain the unitary transformation that converts a set of orbitals, *e.g.* the canonical orbitals, to localized ones (see Section 1.1.6). Typically, the different localization approaches lead to the same qualitative picture, *i.e.* how many electrons are localized on the metal centers and where the delocalized electrons are located. Figure 4.7 shows one of the rare examples where this is not the case. The orbitals belong to the BS7-235 $M_S = \frac{3}{2}$ determinant for the FeVco cluster model. In analogy to the localized orbitals reported for FeVco and FeMoco in QM/MM models[17,19] one would expect a delocalized orbital between Fe2/Fe3 and a delocalized orbital between Fe6/Fe7 (see Figure 4.3 and A14). When comparing the localization methods implemented in `ORCA` (`orca_loc` utility), they all lead to a consistent picture for the orbital evenly shared in the Fe6/Fe7 pair. As shown in Figure 4.7top, the contribution of the central carbide to the delocalized orbital highlight the σ/π mixing in case of Foster-Boys

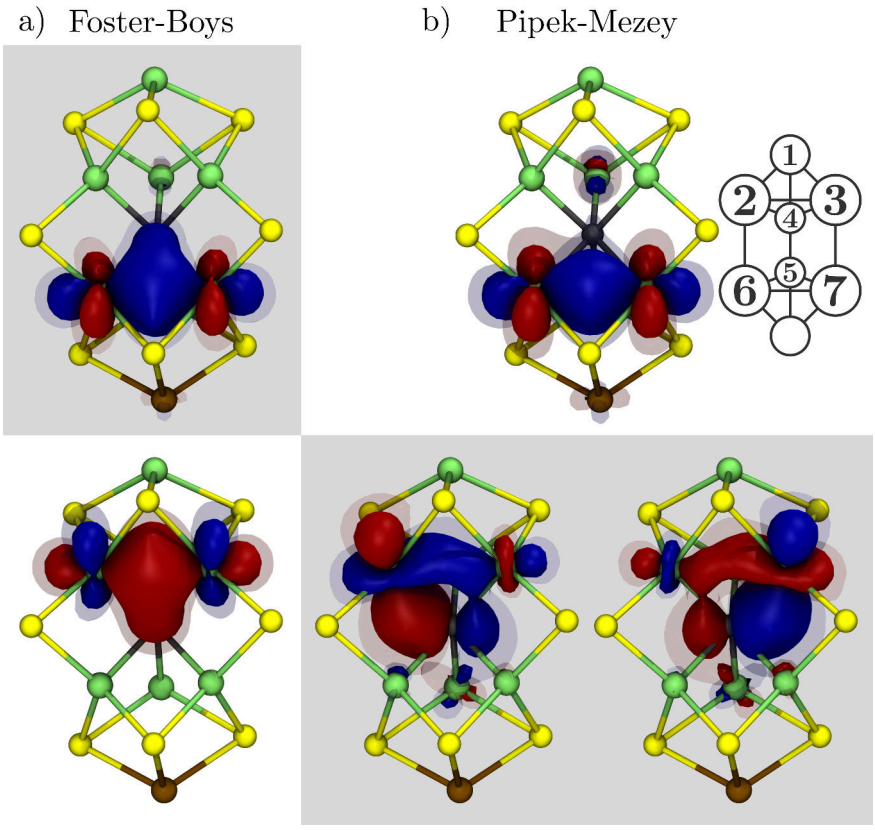

Figure 4.7. – Comparison between (a) Foster-Boys and (b) Pipek-Mezey localized orbitals in FeVco (E_0 cluster model, BS7-235 $M_S = \frac{3}{2}$). For the orbital shared between the Fe6/Fe7 pair (top), both algorithms yield an evenly delocalized orbital. However, for the Fe2/Fe3 (bottom), only the Foster-Boys orbital is evenly delocalized, while Pipek-Mezey yields two asymmetric orbitals. IAO-Boys and new-Boys orbitals were virtually indistinguishable to Foster-Boys, while the IAO-IBO orbitals showed an asymmetry somewhat similar to Pipek-Mezey. Isosurfaces are shown at ±0.05 and ±0.025.

and the σ/π separation in case of Pipek-Mezey. However, for the Fe2/Fe3 pair these two localization methods lead to qualitatively different results. According to the Foster-Boys orbitals the Fe2/Fe3 pair is a mirror image of the Fe6/Fe7 pair. According to the Pipek-Mezey orbitals the Fe2/Fe3 pair harbors two delocalized orbitals, that also involve the central carbide. The central carbide, in turn, carries only 7 localized orbitals instead of 8, which would be expected for the formal C^{4-}). Other descriptors of the electronic structure, such as spin population, atomic charges, bond orders *etc.*, are in line with a single delocalized orbital shown in a) and not with an extraordinary involvement of the central carbide, as may be inferred from b).

The inconsistencies in the localized orbitals in Figure 4.7 may to be related to σ/π separation, which is maintained during the Pipek-Mezey localization, but not during Foster-Boys. The separation constrains the shape of the localized orbitals and could therefore lead to artifacts such as those seen here. Since significant inconsistencies were observed only for the FeVco cluster model, the problem may be specific to the electronic structure of FeVco and/or the isotropic environment of the C-PCM. To which

extent the belt carbonate – being on the backside of the central carbide relative to the Fe2/Fe3 pair – influences the localization procedure in FeVco remains to be determined. However, the obvious failure of the Pipek-Mezey algorithm for the FeVco cluster model may also manifest itself more subtly in other cases, such as the FeMoco QM/MM model, for which an asymmetric orbital between Fe2/Fe3 was reported (34/56 % according to the Mulliken population).[17] In any case, the relationship between the localized orbitals and the electronic structure should be explored in more detailed in light of the inconsistencies shown in Figure 4.7.

Effect of the carbonate on the electronic structure

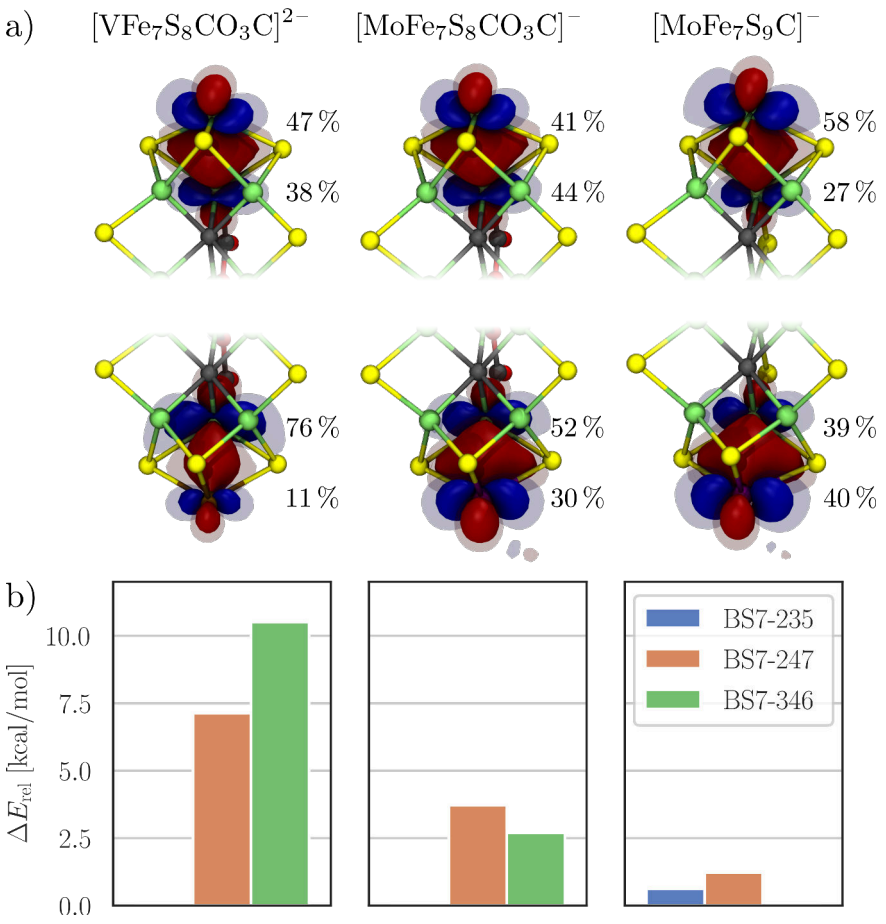

Figure 4.8. – Stepwise transformation from FeVco (left) to FeMoco (right) via an intermediate cofactor that contains both Mo and carbonate (E_0 cluster model, BS7235 $M_S = \frac{3}{2}$). (a) Foster-Boys localized orbitals shared by the Fe1/Fe4 and Fe5/X pairs (X = Mo, V). Fe atomic contributions are based on Hirshfeld populations. Isosurfaces are shown at ±0.05 and ±0.025. (b) Energies in the BS7 determinants relative to the most stable (BS7-235, BS7-235, and BS7-346, respectively).

The largest difference between the resting states of FeMoco and FeVco is the electron configuration on the heterometal, which is reported as a d^3 Mo^{3+} and a d^2 V^{3+}, respectively, according to QM/MM models.[17,19] While the cluster model suggests a

more delocalized Mo−Fe interaction rather than a localized d^3 Mo^{3+}, the localized orbitals still indicate a shift of an electron from the heterometal towards Fe5 in FeVco (76/11 %) compared to FeMoco (39/40 %). Because the difference between the structures of FeMoco and FeVco is both the heterometal and the ligand bridging the Fe4/Fe5 pair, it is not obvious which chemical modification is responsible for the change in electron configuration. To distinguish between the effects of the heterometal and the bridging ligand on the electronic structure, the V in FeVco was replaced by Mo while keeping the carbonate ligand, shown as FeVco(V→Mo) in Figure 4.8.

Figure 4.8a shows the localized orbitals shared by the Fe1/Fe4 pair and by the Fe5/X pair (X=Mo,V) in the BS7-235 determinant. The comparison of FeVco(V→Mo) with FeMoco, which differ only by the Fe4/Fe5 bridging ligand, shows that the carbonate ligand causes a shift in the delocalized orbitals between Fe1/Fe4 and Fe5/Mo by 10 %–20 % towards Fe4 and Fe5, respectively. However, the comparison between FeVco(V→Mo) with FeVco, which differ only in the heterometal, shows that the orbital shared between Fe5/X is only fairly localized on Fe5 in case of V. At the same time, the Fe1/Fe4 orbital appears to respond to the changes in the Fe5/X orbital by shifting from Fe4 to Fe1. In summary, the location of the electron in the Fe5/X pair appears to be affected by both the heterometal and the ligand bridging the Fe4/Fe5 pair.

The structural differences between FeVco and FeMoco also have a pronounced influence on the relative energies of the determinants in the BS7 class. As shown in Figure 4.8b, the energies are spread over 10 kcal/mol in FeVco, but only over 2 kcal/mol in FeMoco. For FeVco, the BS7-235 determinant is clearly more stable (> 7 kcal/mol), whereas the BS7-346 is slightly more stable for FeMoco. Substituting the V in FeVco with Mo (FeVco(V→Mo)) shows that the preference for the BS7-235 determinant remains, but is reduced (> 3 kcal/mol) compared to unsubstituted FeMoco. Therefore, the strong preference for the BS7-235 determinant in FeVco results from both the V and from the carbonate.

Interestingly, the heterometal also locks the relative energies of the BS7-247 and the BS7-346 determinants, which are mirror images with respect to the plane defined by the carbonate ligand. While V as the heterometal favors the BS7-247 determinant by 3.5 kcal/mol, Mo as the heterometal favors the BS7-346 by 1 kcal/mol. How the heterometal determines the relative stability of these two symmetry-related determinants remains to be elucidated.

Stability of BS determinants in FeVco

Figure 4.9 compares the average energy and the standard deviation of all 10 BS classes for FeMoco and FeVco. The energies within each BS class are fairly similar for FeMoco, but cover a greater range for FeVco. This is consistent with the BS7 determinants in the cluster model, where the larger spread of energies in FeVco has been shown to be a result of both the bridging carbonate and the heterometal V (see Figure 4.8b). The values are reported for the QM1* model, in which the protein is modeled explicitly, however, the average energies and standard deviations exhibit only small differences between the cluster* and the QM1* model for FeMoco (see Figure A15). Therefore, the spread in energies for all BS classes probably stems from the carbonate and the V in FeVco. Since the lowest-energy determinant (BS7-235) is energetically favored in

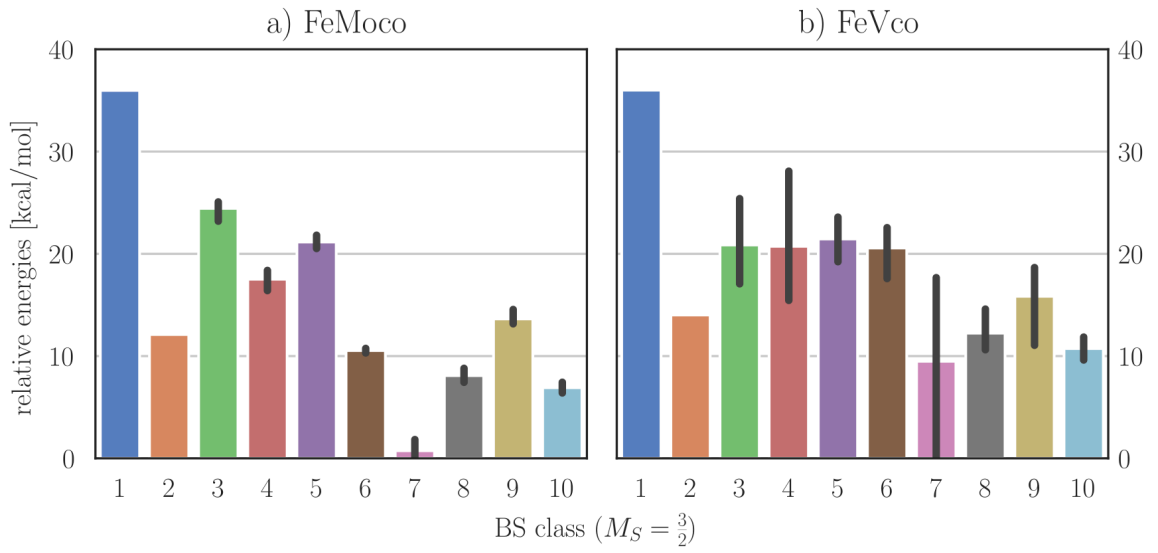

Figure 4.9. – Average energies of the 10 BS determinant classes in (a) FeMoco and (b) FeVco, including the standard deviation for each class. (E_0 QM1* model, $M_S = \frac{3}{2}$). The lowest-energy models are BS7-346 for FeMoco and BS7-235 for FeVco.

FeVco, it may be possible that this BS determinant is a better approximation of the true wave function, unlike in FeMoco, where the three BS7 determinants are nearly degenerate. This further implies that in BS-DFT studies of FeVco a smaller number of determinants may be sampled due to the clearer energy separation.

Since the μ–CO_3^{2-} ligand has been shown to directly affect the relative energies of the BS7 class (see Figure 4.8), one may wonder if this is a result from a change in antiferromagnetic coupling strength. Figure 4.10 is the FeVco-analog of Figure 4.6 and shows the relative energies sorted with the pathway model (a), the calculated coupling constants J_{ij} (b), and the relative energies sorted with the summed J_{ij} model (c). As demonstrated already in Figure 4.9, the energies within all BS classes cover a greater range in FeVco compared to FeMoco. As expected from the lower symmetry in FeVco, the calculated coupling constants for different Fe pairs form less compact clusters. In particular, the μ–CO_3^{2-}-bridged belt Fe pair has a coupling constant of merely $-200\,\mathrm{cm}^{-1}$ while the μ–S^{2-}-bridged belt Fe pairs lie around $-340\,\mathrm{cm}^{-1}$. The triangle and carbide Fe pairs also have lower average absolute coupling constants and as a result, the absolute sums of J_{ij} are smaller in FeVco compared to FeMoco (*e.g.* around $-2700\,\mathrm{cm}^{-1}$ and $-3300\,\mathrm{cm}^{-1}$ for the BS7 class, respectively). Therefore, the antiferromagnetic alignment of Fe pairs is expected to be less important to the stability of BS determinants in FeVco. Nevertheless, the summed J_{ij} model clearly fails to describe anything but the general trend in the relative energies and a more refined model is needed. As discussed for FeMoco, the current models clearly fail to capture the trends in the relative energies of the BS determinants (Section 4.2.2). The stepwise substitution from FeMoco to FeVco (S→CO_3 and Mo→V, Figure 4.8) further supports that the interactions involving the heterometal have a pronounced effect on the energy.

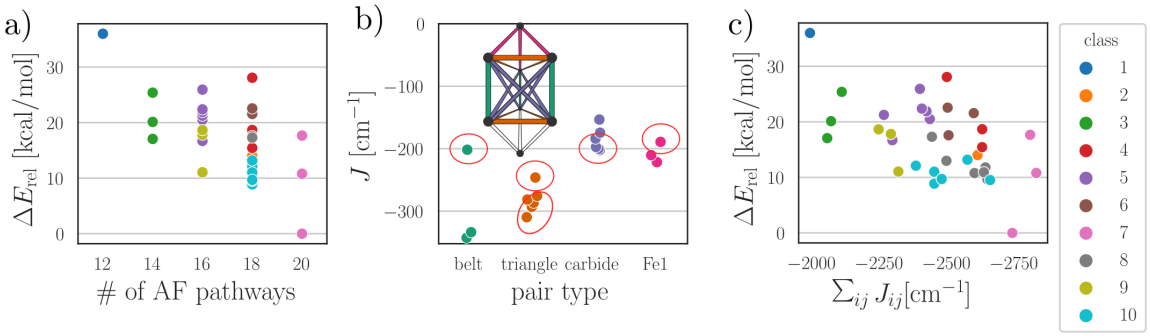

Figure 4.10. – Connection between the energies of the BS determinants in FeMoco and antiferromagnetic coupling between Fe pairs. (a) Dependence of the BS determinant energies on the number of antiferromagnetic pathways present in the respective determinant (adapted from Harris and Szilagyi[291] with the energies for the E_0 $M_S = \frac{3}{2}$ FeMoco cluster* model). (b) Calculated coupling constants for different $Fe^{3+}Fe^{3+}$ pairs in the diamagnetically substituted FeMoco (structures based on the E_0 $M_S = \frac{3}{2}$ FeMoco QM1 model). (c) Dependence of the BS determinant energies on the summed J_{ij} metric (energies for the E_0 $M_S = \frac{3}{2}$ FeMoco cluster* model). See text for further details. The figure is analogous to Figure 4.6 for FeMoco, red circles in b) highlight the Fe pairs involving the carbonate-bound Fe4 and Fe5.

4.3. Influence of the protein environment on the electronic structure

While Section 4.2 dealt with the intrinsic properties of the cofactors, the following analysis explores the influence of the protein environment. Section 4.3.1 compares a cluster model, where the cofactor is placed within an implicit solvent continuum, with the explicit QM/MM model. Furthermore, the influence of the QM/MM interface on the cofactor is quantified by expanding the QM region. Section 4.3.2 focuses on the charged Arg residues in the MoFe protein that are close to FeMoco. Section 4.3.3 works out the differences between the MoFe and the VFe protein. This is done by analyzing hybrid models, where the parent XFe proteins are switched, as well as via specific substitutions of residues.

4.3.1. Different models for the protein environment

The computational models studied in this work (see Figure 4.1) use varying degrees of sophistication to describe the cofactor-protein interaction. In the cluster model, the protein environment is modeled crudely by a polarizable continuum (C-PCM). Here, charges can be induced on a surface around the molecule, which models a dielectric bulk material. In the QM/MM models, the surrounding residues are included explicitly. These residues are either part of the QM region and therefore described alongside FeMoco with DFT, or part of the MM region and therefore interact with the QM region via the atom-centered charges defined in the MM force field. One can therefore compare different QM regions to understand how the QM/MM interface differs from the DFT-only description.

Cluster vs. QM/MM model

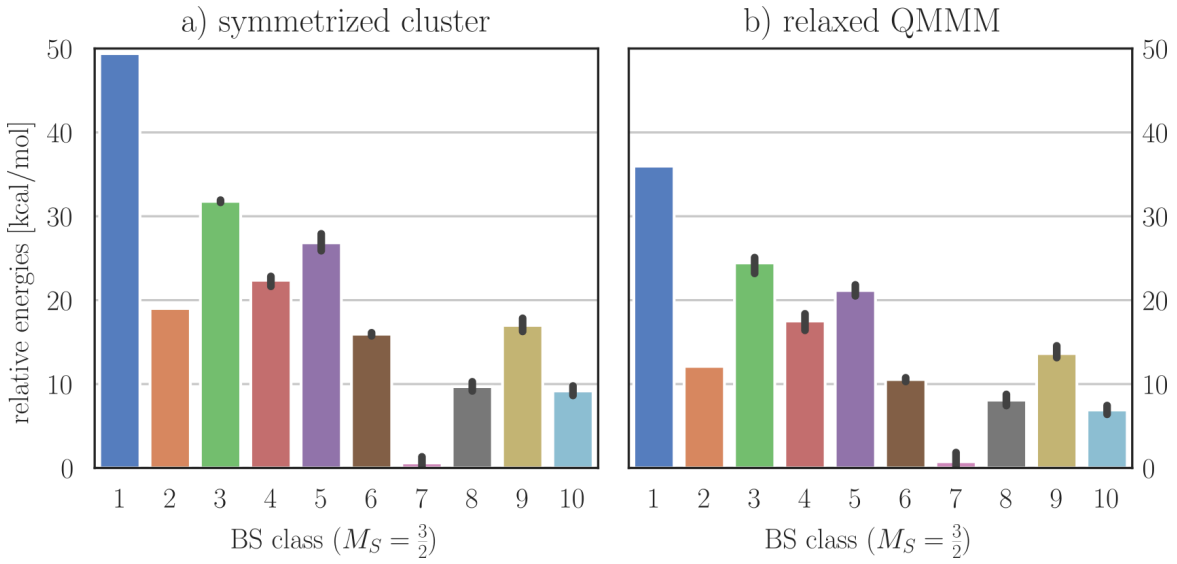

Figure 4.11. – Average energies of the 10 BS determinant classes in (a) the cluster* model and (b) the QM1* model for FeMoco including the standard deviation for each class. (E_0 models, $M_S = \frac{3}{2}$, lowest energy: BS7-235). The cluster* model was symmetrized to C_{3v}, while the QM1* model was optimized for each BS determinant.

Figure 4.11 compares the average energies of the 10 BS classes for both models. The cluster* model and QM1* model use the same low level of theory (double ζ basis set, Figure 4.1a). To calculate the energies for the cluster* model, the geometry optimized with the BS7-235 $M_S = \frac{3}{2}$ cluster model was used and symmetrized to C_{3v} symmetry using `chemcraft`.[267] The standard deviation of the energies in each BS class is of similar magnitude for both models, despite the more asymmetric electrostatic potential in the QM/MM model. This shows that the spin coupling between the metal centers is a more decisive aspect for the stability of a given BS determinant in comparison to the explicit protein environment. However, the energy range spanned by all 35 BS determinants is smaller in the QM1* model. This observation is most likely the result of the geometry relaxation in the QM1* model, that allows the cofactor to adapt to the changes in the spin-coupling pattern. It is therefore important to keep in mind that geometry relaxation can change the relative energies between BS determinants substantially (*e.g.* 15 kcal/mol for BS1-123) when estimating energies based on single point calculations.

The atomic charges of FeMoco according to the different models is shown in Figure 4.12. The charge distribution is evidently most symmetric in the cluster model: the sulfur charges are similar for the three respective symmetry-related sulfides (i: Fe-only cubane, ii: cofactor belt, iii: Mo cubane). For the metal charges, the symmetry is broken and the charges on Fe4 and Fe5 are higher than those on the respective symmetry-related Fe centers. Here, the spin-coupling pattern in the BS determinant becomes evident: Fe4 and Fe5 lie within the single mirror plane of the BS7-235 determinant.

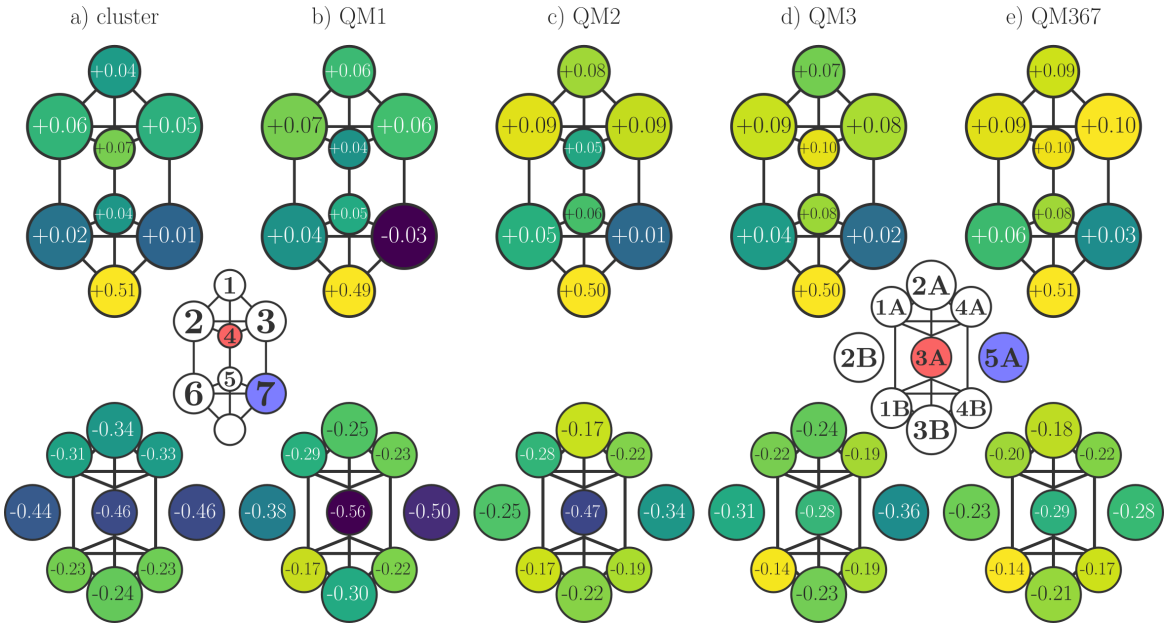

Figure 4.12. – Comparison of the Hirshfeld atomic charges for the metal centers (top) and the sulfide ligands (bottom) in different FeMoco models (E_0, $M_S = \frac{3}{2}$). The size of the QM region increases from (a) to (e), as shown in Figure 4.1 (except for the QM367 model, which was kindly supplied by Benediktsson and Bjornsson[17]). The highlighted atoms are discussed in the text (Fe4: red, Fe7: blue, S3A: red, S5A: blue).

In the QM1 model, the proximal residues affect the QM density via the MM point charges defined in the force field. In comparison with the cluster model, the explicit description of the residues leads to an asymmetry in the sulfur charges. In particular, the sulfides S3A and S5A (red and blue in Figure 4.12, respectively) are significantly more negative in the QM1 model, both of which form hydrogen bonds with surrounding residues. Describing the protein environment explicitly is furthermore noticeable in the metal charges. In the QM1 model, Fe4 and Fe7 (red and blue in Figure 4.12, respectively) are significantly more negative compared to the symmetry-related Fe centers. Note that the apparent shift of electron density towards Fe4 and Fe7, respectively, is not reflected in the delocalized orbitals involving the two Fe centers. For example, the delocalized orbital of both the Fe2/Fe3 pair and of the Fe6/Fe7 pair shifts by 3 % towards Fe3 and Fe7, respectively, in going from the cluster to the QM1 model, however, the atomic charges indicate large changes only for the Fe6/Fe7 pair. It is therefore not obvious whether the atomic charges indicate an actual redistribution of the valence electrons within the cofactor that may, in turn, affect the relative reactivity of Fe6 and Fe7 as potential binding sites.

Extending the QM region

The charges of the S3A and the S5A sulfide (red and blue in Figure 4.12, respectively) are affected further when including the respective hydrogen-bonded residues in the QM region. In the QM2 model, the Arg359 residue is part of the QM region, which

forms a hydrogen bond with the S5A. Compared to the QM1 model, where Arg359 is part of the MM region, the S5A charge is significantly less negative in the QM2 model. The atomic charges therefore indicate that the negative charge on sulfide is reduced if the hydrogen-bonded residue is described by DFT as opposed to the MM force field. Similarly, the S3A charge becomes less negative in the QM3 model compared to the QM1 and QM2 model. In the QM3 model, the residues Leu358 through Gly356 are part of the QM region and their backbone forms hydrogen bonds with S3A. The Hirshfeld charges do not rely on the partitioning of the physical space and are therefore not expected to change significantly as a consequence of basis set superposition (see Section 1.1.5). Furthermore, the QM/MM interface also affects the charges on those metal centers that are coordinated by the respective sulfides. For example, the Fe4 and the Fe7 charges (red and blue in Figure 4.12, respectively) correlate with the S3A and S5A charges and become more positive, if the hydrogen-bound residues become part of the QM region. However, the localized orbitals do not show any significant differences (not shown), and it remains to be determined whether the changes in the atomic charges also reflect in the reactivity of FeMoco.

While the QM1 region includes only residues with covalent bonds to the cofactor, the QM367 region of Benediktsson and Bjornsson has residues surrounding the cofactor in all directions. Following the trend of the QM2 and QM3 model, the QM367 model results in even less negative sulfur atomic charges compared to the QM1 model. Therefore, the atomic charges also change when extending the QM region with residues that do not form hydrogen bonds with FeMoco. Still, S3A and S5A (red and blue in Figure 4.12, respectively) carry the most negative compared to all other sulfurs in the QM367 model, which is consistent with the hydrogen bonding. Overall, the sulfur charges in the QM367 model span a smaller range of values than in any other model discussed here. A more uniform charge distribution in the QM367 model is also apparent from the metal charges. The metal charges in the Fe-only cubane are nearly identical, which links the unequal distribution in the QM1 model to the description of the protein environment. In direct comparison, the Fe charges in the Mo cubane are less equivalent compared to the Fe-only cubane: Fe7 (blue in Figure 4.12) is the least positive Fe center in the Mo cubane, which is probably related to the positively charged Arg359 in the proximity of Fe7. While this may indicate a different reactivity of the Fe centers in the Mo cubane, the interaction of a possible substrate with the surrounding residues will probably affect the reaction mechanism more strongly. Finally, whether an ever larger QM region leads to a more realistic representation of MoFe protein is not immediately obvious. The force field parameters are optimized for protein structures and may capture important cooperative effects. The current DFT protocol for the QM region, on the other hand, uses a small basis set for the surrounding residues, which may not yield reliable protein structures.

4.3.2. Charged residues in the proximity of FeMoco

In addition to Cys275 and HCA, which directly coordinate Fe1 and Mo, the residues Arg96 and Arg339 are the charge carriers next closest to FeMoco. Arg96 forms hydrogen bonds with S5A and S3B and Arg339 with S4A and S4B (see also Figure 4.16). The proximity of the positively charged residues is expected to have an effect on substrate-bound geometries and therefore the reactivity of FeMoco.

Deleting charges on Arg residues

One approach to assess the influence of the charged residues on FeMoco is to simply delete the charge from the model, which is straightforward if the residue is part of the MM region. Here, the atom-centered MM charges, which are responsible for the electrostatic interactions between MM atoms but also polarize the QM region, were set to zero for the charged group of Arg96 and Arg339. All other parameters of the force field were left untouched (see Section A1.2 for code) and the modified model was reoptimized.

The changes in the structure and the atomic charges upon charge deletion are shown in Figure 4.13. The center of the charged groups moved < 0.5 Å after the charge deletion (Figure 4.13a). This indicates that the Arg residues are not only held in their position by electrostatic interactions, but to a large extent by the protein backbone. The atomic charges exhibit a shift within the cofactor (Figure 4.13b and c). In particular, the belt sulfide S5A, which is wedged in between the two Arg residues, becomes significantly more positive (+0.16). The shift of charge happens from the atoms close to the S5A position towards those close to S2B. The charge redistribution is clearly related to the removal of the MM charges on the Arg residues. Simultaneously, the metal-metal distances respond to the charge migration (Figure 4.13d). The distances involving Fe1, Fe2, and Fe6 increase, while those involving Fe1, Fe3, and Fe7 decrease. Naturally, the bond lengths correlate with the electron density in a particular region of FeMoco.

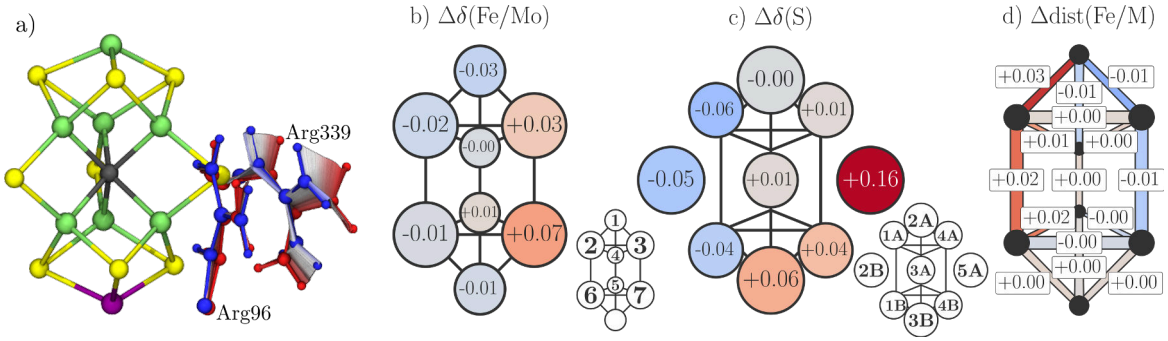

Figure 4.13. – Effect of deleting the MM charges on the cationic Arg96 and Arg339 residues on the geometric and electronic structure of FeMoco (E_0 QM1 model, BS7-235 $M_S = \frac{3}{2}$). (a) Displacement of the Arg96 and Arg339 side chains upon deleting their MM charges (red to blue). (b) Changes in the Hirshfeld atomic charges for the metal centers relative to the original QM1 model. (c) Same as b) but for the sulfide ligands. (d) Changes in the metal-metal distances in [Å] relative to the original QM1 model.

The atomic charges in Figure 4.13 show that the positively charged Arg residues draw electrons towards the Fe7 center in the QM1 model. The discussion of different protein models has shown that modeling residues close to FeMoco with MM parameters tends to overestimate charge redistribution (see Figure 4.12). This should be kept in mind when trying to interpret the corresponding geometric changes. Nevertheless, the polarizing influence of the Arg residues appears to be also present to some extent in those models where the Arg residues are part of the QM region (QM2, QM3, and QM367 in Figure 4.12). Still, Fe7 being the most negative Fe center is inconsistent

with experimental data. The SpReAD-refined XRD structure at the Fe K-edge energy suggests that Fe3 and Fe7 are actually more oxidized compared to Fe2, Fe6, Fe5, and Fe6.[201] This contradiction warrants more detailed theoretical and experimental analysis.

4.3.3. Comparison of the MoFe and the VFe protein

In the following, the influences of the MoFe and the VFe protein on the electronic structure of the respective cofactor are compared. Intrinsic differences between FeMoco and FeVco were analyzed in Section 4.2.3 and the differences during the CO binding event are studied in Section 5.3.3 and 5.4.4.

Hydrogen bonds in MoFe and VFe

Differences in the protein-cofactor interactions between the MoFe and the VFe protein are compared in Figure 4.14. These are predominantly differences in the hydrogen bonds involving the bridging sulfides/carbonate. Figure 4.14a shows that the Arg359(Mo) residue forms two similarly weak hydrogen bonds (3.2 Å–3.7 Å) with the S4A and S4B sulfides of FeMoco, respectively. The Arg339(V) residue, on the other hand, forms a shorter hydrogen bond with S4A (2.2 Å), a longer one with S4B (3.7 Å), and a third with the carbonate. The VFe protein is therefore expected to stabilize negative charge on the S4A more strongly than the MoFe protein. As shown in Figure 4.14b, the Arg96(Mo) residue forms two relatively short hydrogen bonds with the sulfides S5A (2.1 Å) and S3B (2.5 Å). The corresponding Lys83(V) residue also forms hydrogen bonds to S5A (2.7 Å) and S3B (2.0 Å), but here the bond with S3A is shorter. Finally, the interaction involving the peptide coiling around the S3A(Mo)/carbonate(V) is shown in Figure 4.14c. The S3A(Mo) sulfide is loosely bound by the amide hydrogens of four residue Arg359(Mo), Leu358(Mo), Gly357(Mo), and Gly356(Mo). In the VFe protein, the Leu358(Mo) residue is replaced by the Pro338(V) residue, but since Pro is a cyclic residue, it can not provide an amide hydrogen. Therefore, only the three residues Arg339(V), Gly337(V), and Gly336(V) interact with the carbonate(V) via the amide hydrogen. However, the guanidinium side chain of Arg339(V) also forms a hydrogen bond with the carbonate(V).

The sulfur charges for protein-bound FeMoco and FeVco are compared in Figure 4.15a (the carbonate atomic charges for FeVco are not shown). The charges are of similar magnitude and the S2B and the S5A belt positions are the most negative. This is consistent with the S2B sulfide forming a hydrogen bond with the His195(Mo)/His180(V) residue and the S5A sulfide being close to both Arg359(Mo)/Arg339(V) and Arg96(Mo)/Lys83(V). The $\mu_3{-}S^{2-}$ in the Fe-only cubane tend to be more negative in FeMoco, but more positive in FeVco, which suggest differences in the electrostatic potentials created by the explicit protein environment. The hydrogen bonds involving Arg359(Mo) and Arg339(V) (Figure 4.14a) also reflect in the charges of the S4A and S4B sulfide. For FeMoco, the two charges are virtually identical, but for FeVco, S4A carries a more negative charge compared to S4B, in line with the shorter Arg339(Mo)−S4A hydrogen bond in the VFe protein. The differences between Arg96(Mo) and Lys83(V) are smaller and (Figure 4.14b) do not lead to any clear trend in the S5A and S3B charges.

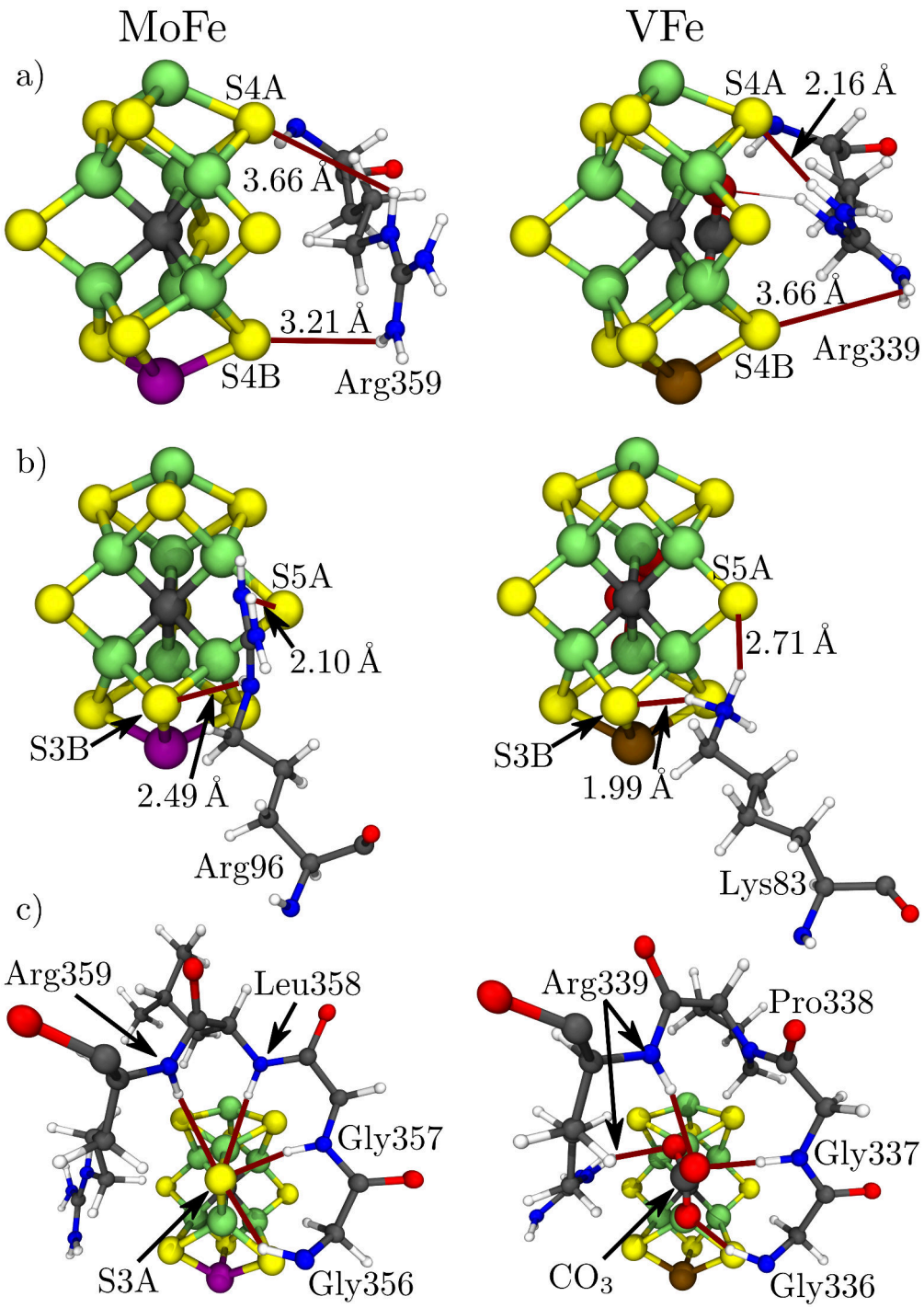

Figure 4.14. – Comparison between the MoFe protein (left) and the VFe protein (right) in terms of the interactions between the protein environment and the cofactor (E_0 QM3 model, BS7-235 $M_S = \frac{3}{2}$). Differences are apparent for the interactions involving (a) the Arg359(Mo) and Arg339(V) residues, (b) the Arg96(Mo) and Lys83(V) residues, and (c) the protein backbone interacting with S3B in FeMoco and the carbonate in FeVco.

Also, potential differences between the amide hydrogen bonds (Figure 4.14c) can not be tracked via the atomic charges, since the S3A sulfide in FeMoco is replaced by a carbonate in FeVco. In order to circumvent these problems and observe changes solely based on the protein environment, one can analyze hybrid nitrogenase models. Here, either FeMoco or FeVco is placed within the respective other XFe protein, thus, canceling all effects that may arise from intrinsic differences between FeMoco and FeVco.

Switching proteins: hybrid nitrogenases

In hybrid nitrogenases, FeVco is placed in the MoFe protein (FeVco@MoFe) or FeMoco in the VFe protein (FeMoco@VFe). The easiest way to create a hybrid QM/MM models is to start with, for example, the wild type MoFe model and replace the heterometal Mo→V and the belt position S3A→CO_3, as shown for the cluster model in Figure 4.8. The FeMoco@VFe hybrid model can be generated analogously. Since each substitution involves only atoms that are already defined in the MM force field, and the cofactors are part of the QM region, no additional groups need to be defined for QM/MM calculations. However, in this approach the partial charges for the other atoms of the cofactor were not adapted, which are different for FeMoco and FeVco, and therefore the electrostatic interaction in the MM step is treated inconsistently (compare `RESID ICS` and `8P8` in the CHARMM parameter file, respectively). Furthermore, since the conformations in the protein environment are not altered during the substitution, the amino acid residues are biased towards the conformations of the wild type protein, which are not necessarily the most favorable in the hybrids.

Switching the XFe proteins around cofactors has an evident effect on the atomic charges, which are shown in Figure 4.15b. The sulfur charges nicely demonstrate that changes in FeMoco for the MoFe→VFe substitution (top) are the inverse of the changes in FeVco for the VFe→MoFe substitution (bottom). As discussed previously, the shortArg339(V)−S4A hydrogen bond (Figure 4.14a) leads to a stabilization of negative charge on S4A in the VFe protein, which is also observed for the FeMoco@VFe hybrid. However, the differences in the hydrogen bonds involving Arg96(Mo) and Lys83(V) (Figure 4.14b) are too small to clearly affect the charges on S3B and S4B hybrids as well. The interaction of the Arg339(V) side chain with the S3A position (Figure 4.14c) is also present in the FeMoco@VFe hybrid. The S3A charge, however, is not affected by the MoFe→VFe replacement, which suggests that the hydrogen bonds with the S3A position are of similar strength in both the MoFe and the VFe protein, even though they exhibit different patterns. Additionally, large changes occur for the S2A and S1B sulfides. There is no obvious reason for these changes in the immediate protein environment around the cofactor. Therefore, proximal charged residues and hydrogen bonds with the cofactors are only one contribution to the charge distribution within the active site. The residues beyond the first coordination sphere in the MoFe and the VFe protein appear to create distinct electrostatic potentials, which may fine-tune the reactivity at the active site.

The changes in the metal-metal bond lengths when creating the hybrids are compared in Figure 4.15c. As for the sulfur charges, the protein substitutions have the respective inverse effect on the bond lengths. Some correlation between the bond lengths and the

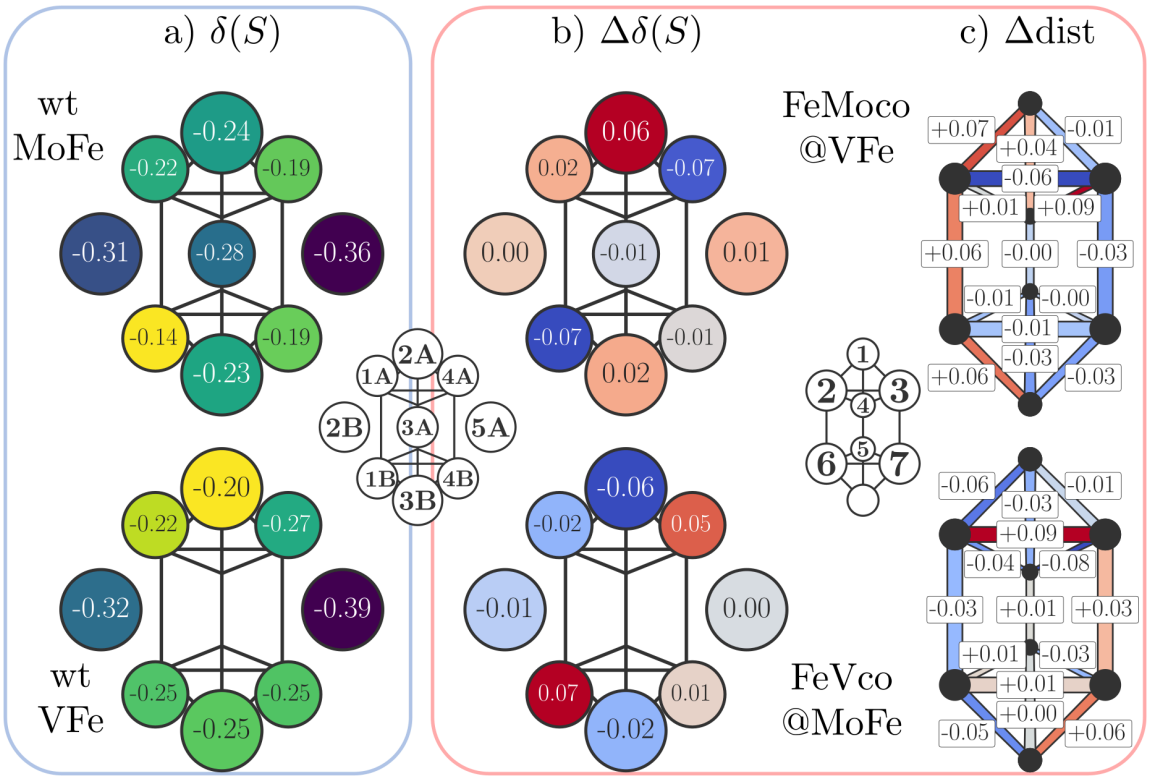

Figure 4.15. – Comparison between the wild type MoFe and VFe protein and the FeMoco@VFe and FeVco@MoFe hybrids (E_0 QM3 model, BS7-235 $M_S = \frac{3}{2}$). (a) Hirshfeld atomic charges for the sulfide ligands FeMoco and FeVco in the wild type MoFe and VFe protein, respectively. Note that the S3A position in FeVco is occupied by a carbonate and the atomic charge is therefore omitted. (b) Changes in the Hirshfeld atomic charges for the sulfide ligands in FeMoco and FeVco upon MoFe→VFe and VFe→MoFe substitution, respectively. (c) Changes in the metal-metal distances ([Å]) in FeMoco and FeVco upon MoFe→VFe and VFe→MoFe substitution, respectively.

sulfur charges is apparent: focusing on the FeMoco@MoFe hybrid, the Fe2−Fe3 bond becomes shorter, while the charge on S2A becomes more positive, and the Fe3−Fe4 bond becomes longer, while the charge on S4A becomes more negative. However, the similarly large elongation of the Fe2−Fe6 bond does not correlate with the change in the charge of the corresponding S2B sulfide. In the MoFe protein Fe1/Fe2/Fe6/X edge of the cofactor is elongated, but the Fe1/Fe3/Fe7/X edge is compressed. Interestingly, the volume[†] of FeMoco expands by 1.5 % upon MoFe→VFe substitution, which would be consistent with the differences in the Leu358(Mo)/Pro338(V) residues, leaving more space for the cofactor in the VFe protein. However, the volume of FeVco remains virtually unchanged upon VFe→MoFe substitution.

The energies of the three BS7 determinants are compared in Figure 4.16 for the wild type and the hybrid models. In the FeMoco@VFe hybrid, the BS7-247 determinant is

[†]Measured as the volume confined by the convex, irregular polyhedron defined by the coordinates of the metal nuclei.

the lowest in energy followed by the BS7-235 determinant, while in the wild type the BS7-346 determinant is lowest. Interestingly, in the FeVco@MoFe hybrid, the BS7-247 determinant is also stabilized relative to the BS7-235 determinant when comparing to the wild type VFe model. This shows that both the MoFe→VFe and the VFe→MoFe substitution stabilizes the BS7-247 relative to the BS7-235 determinant. This is in contrast to the changes in the atomic charges and the bond lengths (Figure 4.15b and c), which largely exhibited inverse trends. The energy changes cannot be easily connected to the changes in atomic charges or bond lengths of the cofactor and a deeper understanding of the electronic structure may be necessary to explain the influence of the protein environment on the relative stability of the BS determinants.

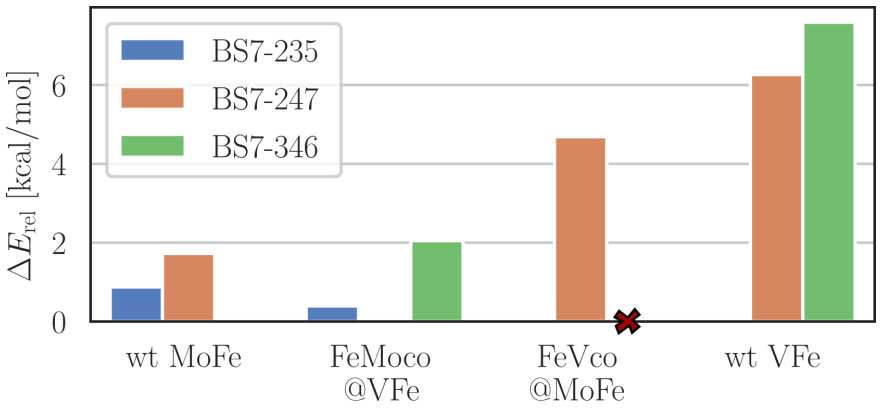

Figure 4.16. – Energies of the BS7 determinants in wild type (wt) MoFe and VFe and in the respective hybrids (E_0 QM3 model, $M_S = \frac{3}{2}$). The lowest-energy determinants are BS7-346 (wt MoFe), BS7-247 (FeMoco@VFe), BS7-235 (FeVco@MoFe), and BS7-235 (wt VFe). The value for BS7-346 in the FeVco@MoFe hybrid is missing due to an error in the model setup.

Arg96→Lys substitution in MoFe

As shown in Figure 4.16b, an immediately obvious difference in the MoFe and the VFe protein lies in the Arg96(Mo)/Lys83(V) residues. It is therefore instructive to introduce the Arg96→Lys substitution in the MoFe QM/MM model and compare its effects with switching the entire protein, *i.e.* with the hybrid models in Figure 4.15. The workflow to introduce substitutions to an existing QM/MM model is explained in Section A1.2.

The comparison between the wild type MoFe model and the Arg96→Lys substituted MoFe model is given in Figure 4.17. The substitution affects the relative energies of the BS7 determinants (Figure 4.17a) and the BS7-235 determinant is stabilized relative to the other two determinants, which has not been observed in the FeMoco@VFe hybrid model. The change in the sulfur atomic charges (Figure 4.17b) indicate that the S3B sulfide becomes more negative and the S5A more positive. This would be consistent with the hydrogen bond patterns in the wild type MoFe and VFe protein (see Figure 4.16b), where Arg96(Mo) forms a shorter hydrogen bond with S5A and a longer one with S3B and *vice versa* for Lys96(V). However, the Lys96(Mo) in the Arg96Lys MoFe variant also forms a shorter hydrogen bond with the S5A sulfide (2.12 Å) and a

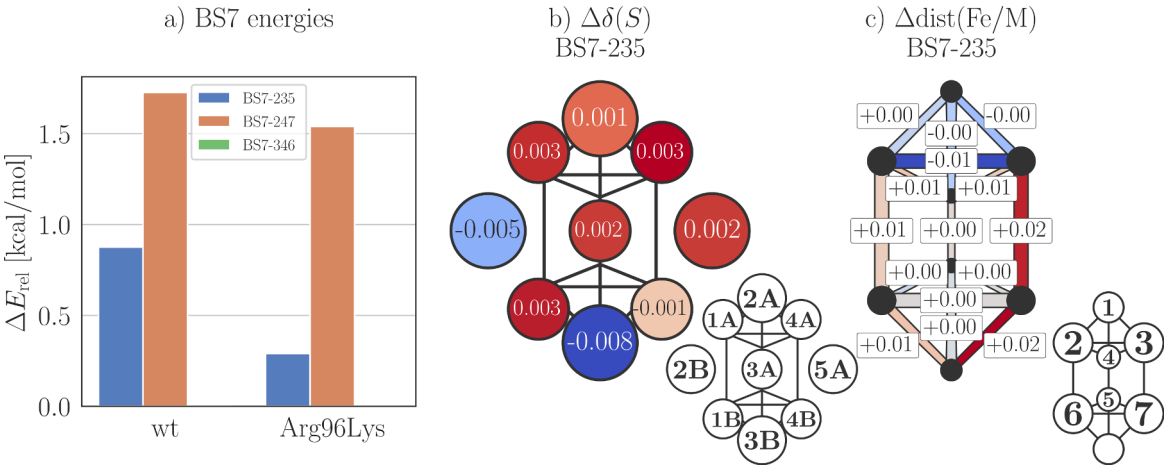

Figure 4.17. – Comparison between FeMoco in the wild type MoFe protein and in the Arg96Lys variant. (a) Energies of the BS7 determinants relative to the BS7-346 determinant (E_0 QM3 model, $M_S = \frac{3}{2}$). (b) Changes in the Hirshfeld atomic charges for the FeMoco sulfide ligands in the Arg96Lys variant relative to wild type MoFe. (c) Changes in the FeMoco metal-metal distances in the Arg96Lys variant relative to wild type MoFe.

shorter one with S3B (2.34 Å). Therefore, the hydrogen bonds in the Arg96Lys variant resemble more closely those in the wild type MoFe protein than the wild type VFe protein. It appears that the position of the Lys96(Mo) residue is constrained by the anchor in the protein backbone, which aligns with the understanding that the protein creates a highly specific electrostatic potential. However, it is also possible that the actual variant MoFe protein undergoes more complex conformational changes than those accessible via simple potential energy surface scans. Finally, the changes in the metal-metal bond lengths (Figure 4.17c) show an expansion of the cofactor close to the substitution. This is consistent with the side chain of the Lys residue being spatially less demanding (chain of 5 non-hydrogen atoms) compared to the Arg residue (chain of 6 non-hydrogen atoms).

4.4. Conclusion

The present chapter explored the electronic structure of the FeMoco and FeVco resting state on the basis of BS-DFT calculations. Section 4.1 introduced the different computational models used throughout this work and specified the level of theory. Section 4.2 focused on intrinsic properties of the cofactors and discussed different descriptors that aid the interpretation of the electronic structure (Section 4.2.1), the relation between antiferromagnetic coupling and the stability of BS determinants (Section 4.2.2), and the differences between FeMoco and FeVco (Section 4.2.3). Section 4.3 quantified the influence of the protein environment on the electronic structure. Different protein models and QM regions were compared (Section 4.3.1), the influence of charged residues was explored (Section 4.3.2), and differences in the protein environment between MoFe and VFe were highlighted (Section 4.3.3).

Antiferromagnetic coupling and the energy of BS determinants

The stability of a spin-coupling pattern in a BS determinant is related to antiferromagnetic coupling of the local high spin Fe centers. This has been modeled by simply counting the antiferromagnetically aligned Fe pairs in a given BS determinant[308] or by further respecting the number of bridging ligands.[291] To refine these models, the coupling constants J_{ij} were explicitly calculated for all Fe pairs in FeMoco using diamagnetic substitution (Figure 4.6). This illustrates that the coupling constants depend strongly on the respective bridging ligands. Using the calculated coupling constants to sort the energies of the BS determinants in FeMoco constitutes an improvement over the previous models.

The energies of the BS determinants are split into multiple, apparently independent, sets, which follow different energy trends. Preliminary analysis of the spin populations suggests that the higher-energy determinants exhibit local spin pairing, while the Fe centers are all local high spin centers in the lower-energy determinants (Figure A13). Furthermore, the same approach for FeVco shows that the spin-coupling mediated through the CO_3^{2-} ligand is significantly lower compared to μ–S^{2-} ligands (Figure 4.10). Surprisingly, modeling the energies of the BS determinants in FeVco with the calculated coupling constants results in a significantly worse prediction compared to FeMoco. This clearly shows that considering only antiferromagnetic coupling is insufficient to rationalize the relative energies of the BS determinants.

Differences between FeMoco and FeVco

Two of the major differences between E_0 FeMoco and FeVco are: (i) In FeVco, the BS7-235 determinant is clearly lower in energy than BS7-247 and BS7-346 (> 7.5 kcal/mol), whereas the three BS7 determinants are nearly degenerate in FeMoco (< 2 kcal/mol, Figure 4.8b). (ii) Previously reported[19,299] XAS measurements and localized orbital analyses yield a more reduced Fe part for FeVeco with oxidation states Mo^{3+}–$Fe5^{3+}$ and V^{3+}–$Fe5^{2+}$ for FeMoco and FeVco, respectively. Both observations appear initially related to the carbonate ligand in FeVco: (i) The S→CO_3 replacement breaks the C_3 symmetry of the cofactor and singles out the BS7-235 determinant, which is symmetric

with respect to the carbonate. (ii) The reduced Fe5 center is coordinated by the carbonate. However, the stepwise substitution of the heterometal (Mo→V) and the bridging ligand (S→CO_3) shows that the nature of the heterometal is relevant for the energies of the BS7 determinants as well as for the oxidation state of Fe5 (Figure 4.8). The reason for this is not obvious, since the Mo→V replacement does not affect the symmetry of the system.

The immediate surrounding around the cofactor is highly conserved for the MoFe and the VFe protein, and differences involve mostly hydrogen bond lengths (Figure 4.1). At a first glance, the most drastic difference lies in the Arg96(Mo)/Lys83(V) residues, because both are positively charged and form distinct hydrogen bonds with the respective cofactor (Figure 4.14). However, when performing the Arg96→Lys substitution in the FeMoco QM/MM model (Figure 4.17), the changes in the electronic structure are rather small compared to substituting the entire protein MoFe→VFe (Figure 4.15). Comparing the hybrid and wild type models shows that the effect of the protein environment on the electronic structure of the cofactor is not dominated by the proximal residues. Instead, the charge redistribution upon switching the protein environment is more likely the result of distinct long-range electrostatic influences in the MoFe and the VFe protein. This difference may prove important when explaining the specific reactivity of the wild type and the hybrid nitrogenases.[196–198]

Localized orbital analysis

The localized orbitals for the E_0 FeMoco model (Figure 4.8a) yield a shared electron between Mo and Fe5 in the BS7-235 determinant. This electron was previously assigned as predominantly Mo-based, yielding the oxidation state $Fe5^{3+}$ based on Mulliken atomic populations,[17] while Hirshfeld partitioning suggests an evenly delocalized electron. The Mulliken population works in the atomic orbital basis, while the Hirshfeld approach works in physical space. Neither approach is unique, because they depend either on the choice of the basis set or on the choice of the promolecular density (see Section 1.1.5 for details). Therefore, while localized orbitals are a powerful tool to compress and visualize the complexity of the electronic structure, the resulting oxidation state assignments should be taken with a grain of salt.

Different localization algorithms can lead to a qualitatively different electronic structure interpretation. For example, in the BS7-235 determinant of the E_0 FeVco model, the Foster-Boys orbitals suggest an evenly delocalized electron for the Fe2−Fe3 pair (Figure 4.7, IAO-Boys and new-Boys orbitals are virtually identical). The Pipek-Mezey orbitals, on the other hand, exhibit two asymmetric orbitals involving Fe2, Fe3, and the central carbide (IAO-IBO orbitals also exhibit some asymmetry). The differences appear to be related to the σ/π separation maintained during Pipek-Mezey localization, but not during Foster-Boys (see Section 1.1.6 for details) and the latter allows for an arbitrary mixing of the central carbide in the predominantly metal-based orbitals. It is not obvious whether σ/π separation is a desirable property for the localized orbital analysis, but since all other substrate-free E_0 cofactor models show evenly delocalized orbitals in the BS7 determinants, one could argue that the Foster-Boys algorithm leads to the most consistent picture.

In contrast to localized orbitals, which are fairly insensitive to environmental in-

fluences, atomic charges have proven as a useful descriptor for the interaction of the protein environment with the cofactor. They can aid the understanding of the electrostatic potential created by the protein (Figure 4.15 and 4.13) and demonstrate the convergence of the QM region size (Figure 4.12). Spin populations, on the other hand, allow for a quick identification of the spin coupling within a given BS determinant and even help to spot local spin pairing (Figure A13). However, neither atomic charges nor spin populations reveal the distribution of the unpaired electrons across the metal centers. This makes the localized orbital analysis a powerful complementary tool when studying BS-DFT determinants for larger FeS clusters.

According to the localized orbital analysis, electron delocalization is a characteristic property of BS-DFT electronic structure of the resting state FeMoco and FeVco (Figure 4.3 and A14).[17,19] Here, most Fe centers have an oxidation state close to $Fe^{2.5+}$. A similarly delocalized electronic structure has been reported for other computational models.[256,308] On the other hand, experimental X-ray data suggest an electronic structure with localized Fe^{2+} and Fe^{3+} centers.[171,201,310] The XRD structure of the resting state MoFe protein recorded with SpReAD refinement of the Fe K-edge XAS indicates, that Fe1, Fe3, and Fe7 are more oxidized compared to Fe2, Fe4, Fe5, and Fe6.[201] Therefore, either the localized orbitals may not correspond to physical oxidation states or the XAS absorption energy may not correlate with the d^n configuration of a given metal center in FeMoco.

Se K-edge HERFD XAS measurements of Se-substituted FeMoco also indicate a localized electronic structure, where Fe2 and Fe6 are more oxidized and Fe3, Fe4, Fe5, and Fe7 are more reduced.[171] However, the study of the $[Fe_2X_2]$ dimers in Chapter 2 has shown that Se compared to S bridges lower the coupling strength between the Fe centers and the coupling strength is imperative to the stability of a given BS determinant (Figure 4.6). It is therefore not obvious if S→Se substitution in FeMoco stabilizes a different electron distribution across the metal centers compared to unsubstituted FeMoco. Admittedly, since the difference in the coupling strength is rather small (5 % for the substitution of one μ–S^{2-} ligand), introducing a single Se into the cofactor is not expected to have a large influence on the electronic structure. On the other hand, the calculations for FeVco show that CO_3^{2-} reduces the coupling strength by nearly 50 % compared to S^{2-} (Figure 4.10). Also, S→CO_3 substitution has visible effects on electron localization and the energies of the BS determinants (Figure 4.8). Therefore, while the S→CO_3 substitution introduces pronounced changes to the electronic structure of the cofactor, the S→Se substitution is expected to be an innocent one.

Outlook

The BS7 class of determinants being the most stable for resting state FeMoco and FeVco has been verified in several computational studies (Figure A15).[311–313] Also, QM/MM optimized structures have been shown to align well with the high-resolution XRD structures of the MoFe and the VFe proteins when using the BS7 determinants.[17,19] While their exceptional energetic position has been rationalized with maximal antiferromagnetic coupling, none of the proposed models describes the energies of the other BS classes correctly, nor can they explain the significantly lower energy of the BS7-235 determinant to BS7-247 and BS7-346 in FeVco (Figure 4.6 and 4.10).[291,308]

Furthermore, in more reduced E_n redox states, or when binding substrates to the cofactor, determinants other than BS7 become lowest in energy,[25,206,214,300,312,314,315] and it is not *a priori* obvious, which BS determinant yields the lowest energies. Combined with the uncertainty of the most stable M_S value, any rigorous exploration of mechanistic details with DFT models is soon overwhelmed by the many possible BS/M_S combinations for each structure. It is therefore of utmost importance to understand the factors that contribute to the stability of a given BS determinant in addition to antiferromagnetic coupling. A more complete model could include contributions from ferromagnetically coupled pairs, or delocalized electrons, which have been shown to stabilize ferromagnetic alignment through double exchange (see also Chapter 3).[24,98] Also, the electronic structure of the heterometal should be considered, since Mo→V substitution in FeMoco has been shown to have a pronounced influence on the BS7 energies (Figure 4.8).

The resting state of FeVco is still debated,[202] but computational models suggest that for a valence electron count equal to E_0 FeMoco, $[MoFe_7S_9C]^-$ and $[VFe_7S_8CO_3C]^{2-}$, respectively, the Fe part in FeVco is more reduced.[19,299] According to the BS7-235 determinant of FeVco, which is clearly the most stable BS determinant in FeVco, Fe5 is the center of reduction (Figure 4.8). SpReAD refinement at the Fe K-edge for FeVco could help clarify if the reduction of the Fe part in FeVco relative to FeMoco is indeed localized.[201] In this context, it could prove insightful to compare measurements on the wild type MoFe and VFe protein, as well as on the respective hybrid proteins. This would allow for a subtraction of the electrostatic potential created by the respective protein environment, which has been shown to stabilize certain charge distributions within the cofactors (Figure 4.16).

5. CO binding to nitrogenase

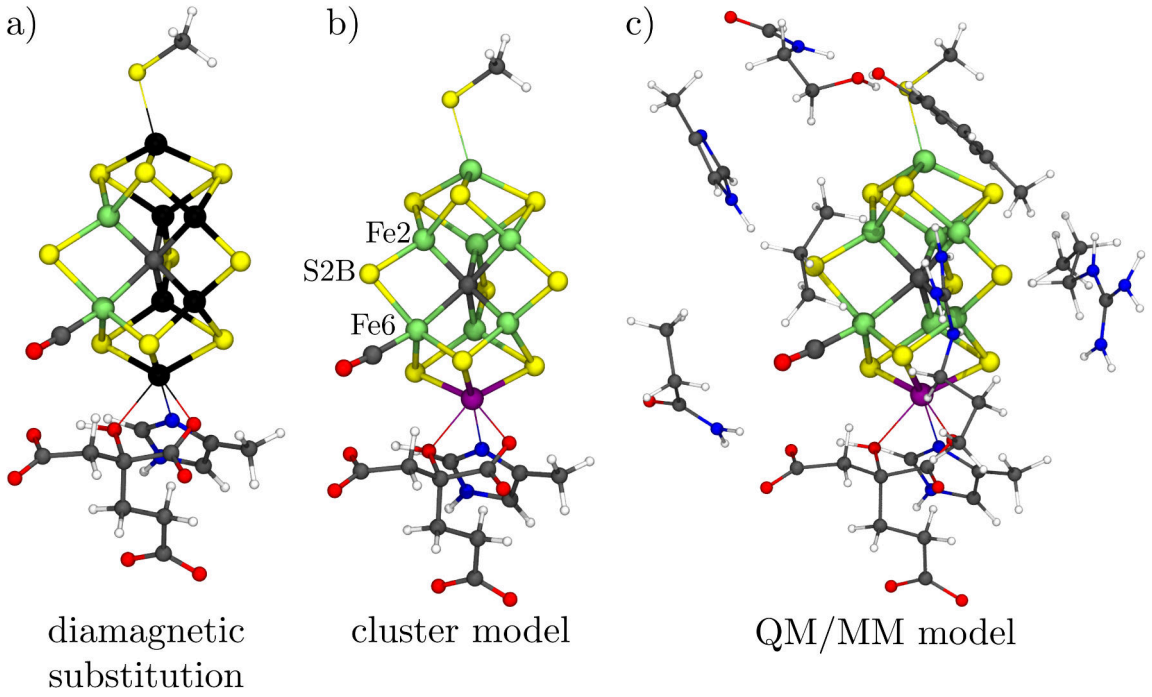

Figure 5.1. – Illustration of the computational models used to study CO binding to nitrogenase: (a) a diamagnetically substituted cluster model, (b) an unsubstituted cluster model, and (c) a QM/MM model. The diamagnetic ions Ga^{3+} and In^{3+} are shown in black. The present chapter uses the cluster and the QM2 model introduced in Chapter 4.

Carbon monoxide (CO) is isoelectronic to N_2. It can reversibly bind to the nitrogenase cofactor, where it can act either as an inhibitor to N_2 reduction, such as in wild type Mo nitrogenase,[204] or can be reduced to hydrocarbons in Fischer-Tropsch-like chemistry, such as in wild type V nitrogenase[190] or Val70 variants of Mo nitrogenase.[191] The binding of CO to FeMoco generally requires enzymatic turnover and the resulting CO-bound species have been characterized with IR spectroscopy,[226–228] X-ray crystallography,[233,234] and EPR spectroscopy.[217,218,222] The experimental species relevant to the discussion are summarized in Figure 1.9. For a more complete account of the current literature, see Section 1.2.4. Briefly, SF-FT-IR measurements suggest that CO binds initially as a terminal ligand to FeMoco ($\mathrm{loCO}_{\mathrm{IR},1}$, 1904 cm^{-1}), which is then transformed to a bridging CO ($\mathrm{loCO}_{\mathrm{IR},2}$, 1715 cm^{-1}). The X-ray structure of the CO-bound MoFe protein also shows a bridging binding mode ($\mathrm{loCO}_{\mathrm{XRD}}$). The EPR spectrum exhibits a characteristic $S = \frac{1}{2}$ signal in the presence of CO ($\mathrm{loCO}_{\mathrm{EPR}}$), which corresponds

to an odd-electron redox state of the CO-bound cofactor. However, quantification of the EPR signal accounts only for a small fraction of the sample,[224] which implies that the majority of FeMoco – CO-bound or not – does not correspond to the $\mathrm{loCO_{EPR}}$ species. This raises the question which redox state initially binds CO and whether further enzymatic reduction is required for the conversion of terminal to bridging CO.

Even though V nitrogenase reduces CO, the experimental species bear close resemblance to those of Mo nitrogenase, such as the highly similar EPR signals[219,220] and the X-ray structures.[221,236] It has been reported that CO binding does not require enzymatic turnover for V nitrogenase,[219–221] but this observation is still under debate.[221] Furthermore, the redox state of FeVco in the resting state is not unambiguously defined.[17,202] To clarify the context of the following discussion, $\mathrm{E_0}$ refers to the resting state as previously reported for the QM/MM models,[17,19] where both FeMoco and FeVco have an equal valence electron count of 41 (see Figure 4.3 and A14). But the reader should keep the uncertainty of the resting state for V nitrogenase in mind.

The present chapter studies the binding of CO to the active site of nitrogenase with the models shown in Figure 5.1. These models were already introduced in the previous chapter, therefore, only the computational details specific to CO binding are given in Section 5.1. First, the local electronic structure of the CO−Fe bond is discussed in Section 5.2 with the help of the diamagnetically substituted cofactor models (Figure 5.1a), similar to Section 2.3.2 for the Fe dimers. Then, in Section 5.3, CO is bound to a cluster model (Figure 5.1b), which neglects the explicit protein environment for the time being and focuses on the interplay between the CO−Fe and the metal-metal interactions. Section 5.4 uses a QM/MM model (Figure 5.1c) to investigate the influence of the explicit protein environment on the CO-bound cofactor. The conclusion in Section 5.5 summarizes the previous results and puts them in context of the current understanding of CO binding to nitrogenase in the literature.

5.1. Computational details

The models used to study the binding of CO were already introduced in the previous chapter and the greater part of the setup also applies here, which is explained in-depth in Section 4.1. Therefore, only the computational details specific to CO binding are given in the following.

CO binding was mainly explored for the $\mathrm{E_0}$ and the $\mathrm{E_1}$ redox states (see Figure 1.9). The CO-bound XRD structures show that CO eventually replaces the S2B sulfide in Mo and V nitrogenase[233,236] and according to the quantum-refinement of the CO-inhibited XRD structure of Mo nitrogenase no other sulfide is protonated.[255] An analysis of EXAFS structural data combined with QM/MM modeling suggests that the belt sulfides S2B or S5A (see Figure 4.2) are protonated in the $\mathrm{E_1}$ redox states.[204] Since S2B protonation appears as a reasonable first step in CO-binding, the $\mathrm{E_1}$ models were created from the $\mathrm{E_0}$ models by adding an electron and protonating the S2B sulfide. However, the reader should keep in mind that it is also possible for a different or no sulfide to be protonated in the $\mathrm{E_1}$ redox state. For the brief exploration of the $\mathrm{E_2}$ model (Mo nitrogenase and CO-bound only) it was assumed that the S2B sulfide had dissociated from the cofactor in the form of $\mathrm{H_2S}$. Therefore, the cofactor in the $\mathrm{E_2}$

model has the same number of protons as in the E_0 model, but is reduced by two electrons.

Since the XRD structures[233,236] show that CO bridges Fe6 and Fe2, those two centers are considered as potential binding sites. In particular, the QM2 model (Figure 4.1) was designed with substrate binding to Fe6 and Fe2 in mind. Therefore, the QM region of the QM2 model includes all amino acids in the proximity of Fe6 and Fe2 in order to avoid artifacts that may arise at the QM/MM interface, as demonstrated in Figure 4.12, which are expected to heavily influence substrate-cofactor interactions.

In the initial structures for the geometry optimization, CO was coordinated to either Fe6 or Fe2 *trans* to the carbide with an initial bond length of 1.8 Å. The substrate-free E_0 and E_1 models were also optimized in order to use their energies for the estimation of the CO binding energy. The binding energy is calculated as $\Delta E = E_{\mathrm{AB}} - (E_{\mathrm{A}} + E_{\mathrm{B}})$, where AB is the CO-bound system, A is the free CO molecule, and B is the substrate-free model. The electronic binding energies ΔE are expected to be a good approximation to the binding enthalpy ΔH.[†] Generally, only the lowest-energy models are discussed, but all other tested BS/M_S combinations are given in Section A4.3. For the E_2 state, only the CO-bound model was calculated and therefore the binding energy cannot be calculated.

The multiplicity of the high-spin determinant, which provides the guess orbitals for the BS determinant, was reduced by 2 for each added electron, because the reduction happens on a local high spin Fe center[204]. It was further reduced by 2 when CO was bound, because CO has been shown to introduce local spin pairing (*vide infra*). The final BS determinant is usually not dependent on the high spin multiplicity, but adjusting for expected electronic structure change improves the SCF convergence. For the added proton and for CO the higher-quality triple-ζ basis set was used except in the cluster* and QM1* model.

The CO vibrational frequencies in the cluster models were calculated numerically using the `! numfreq` keyword in `ORCA`, while those in the QM/MM models by a numeric three-point harmonic approximation using `chemshell`. In both cases, a partial Hessian was diagonalized, which included only the CO and the coordinated Fe atoms, because the inclusion of further atoms of the cofactor has proven to leave the CO frequencies unaffected (see Table A17). The calculated frequencies were scaled so that the calculated frequency of free CO matches the experimentally determined frequency of gaseous CO.[225] The calculated frequencies in the text represent exclusively scaled frequencies, but all unscaled frequencies are listed in Table A18.

Since CO binding significantly changes the geometry of the cofactor, the geometry for the diamagnetically substituted models was taken from the CO-bound models and not from the substrate-free models. For terminally bound CO, the lowest-energy CO-bound E_0 cluster model was used to create the diamagnetically substituted models CO−[$FeGa_6InS_9C$] and CO−[$Fe_2Ga_5InS_9C$]. Afterwards, only the CO coordinates were optimized in order to capture CO activation. Since the bridging CO motif was found to be stable only within an explicit protein environment, the respective

[†]In the E_1 cluster model with CO bound to Fe6 $\Delta E = -15.0$ kcal/mol and $\Delta H = -14.1$ kcal/mol. To calculate ΔH, all vibrational modes of FeMoco and CO were considered. The Cys, His, and HCA residues were not included in the partial Hessian, since they were constrained during the geometry optimization (Section 4.1).

diamagnetically substituted model was created from the geometry of the μ−CO-bound E_1 QM/MM model. In this model, μ−CO−[$Fe_2Ga_5InHS_9C$], the CO coordinates were not further optimized and therefore no CO frequency is calculated. The general procedure of diamagnetic substitution, in particular the choice of the diamagnetic ions, is explained in Section 4.1.

In the ΔHis195 QM2 model, the steric hindrance introduced by the His195 residue was explored (Mo nitrogenase only). In this model, the side chain of His195 was simply deleted from the QM region and only the FeMoco and CO coordinates were optimized in the field of the MM point charges and the remaining, but frozen, QM region. The energy values for the ΔHis195 QM2 model refer to the QM energy, which includes the polarization of the MM point charges. The comparison of QM energies (cluster and ΔHis195 QM2 model) and total QM/MM energies (QM2 model) can be justified, since only energy differences within the respective model are discussed in this context.

5.2. Diamagnetic substitution: the local electronic structure

Substituting a subset of the magnetic transition metal centers with diamagnetic ions allows for a separation between local properties of the electronic structure and those that involve the interaction between multiple magnetic centers. Briefly, out of the eight magnetic centers in FeMoco all but one or two Fe centers are substituted with diamagnetic main group ions that have a similar ionic radius as the transition metal ions (see Section 4.1 for detailed explanation). Afterwards, by comparing the diamagnetically substituted models with the original model, one can categorize aspects of the electronic structure as being a local property or a collective property. This is a bottom-up approach, where complexity of the system is increased step-by-step. In terms of CO binding, it eventually helps to understand the interplay of CO binding and the magnetic interactions between the metal centers.

The outline of the diamagnetic substitution study is as follows: In Section 5.2.1, a single Fe center is left unsubstituted, to which a terminal CO is bound (Fe6 in [$FeGa_6InS_9C$]). In Section 5.2.2, two Fe centers are left unsubstituted in order to study the effect of the Fe−Fe interaction on the Fe−CO bond (Fe6 and Fe2 in [$Fe_2Ga_5InS_9C$]). The model Section 5.2.3 also contains two Fe centers, but CO assumes a bridging binding mode between them (μ−CO−[$Fe_2Ga_5InHS_9C$]). Diamagnetic substitution has been used previously in Section 2.3.2 to study the ligand field strength of bridging ligands in [Fe_2X_2] dimers and in Section 4.2.2 and 4.2.3 to decompose the energy in a given BS determinant into pair-wise contributions.

5.2.1. Terminal CO binding to a single Fe center

Local spin states in FeS clusters

As explained in Section 1.2.1, the Fe centers in biological FeS clusters are typically coordinated by four sulfide ligands and have oxidation states ranging from Fe^{3+} to Fe^{2+}.[98] Consistent with the ideas of LFT, the Fe centers exhibit local high spin states,

because the weak field ligands and the low coordination number result in a small ligand field splitting.[316] Mössbauer data and computational models suggest, that these properties of the local electronic structure also apply to the μ_6−C^{4-}-containing FeMoco (see Figure 4.3).[17,19,200] Local high spins, *i.e.* five majority spin electrons per center, are also lowest in energy for the single Fe6 center and the Fe6Fe2 dimer in the diamagnetically substituted substrate-free cofactors using the same level of theory (data not shown). Therefore, a local high spin is also energetically favored for an Fe center in FeMoco without or with significantly reduced metal-metal interactions.

In contrast to the sulfide ligands, coordination of CO typically leads to low spin transition metal complexes.[240] This can be rationalized with the π back bonding from occupied metal d orbitals to the unoccupied π^* orbitals on the CO ligand. The π back bond becomes stronger the more electrons occupy the metal orbitals of suitable symmetry and CO binding is therefore particularly strong for electron-rich low-spin metal centers. A CO bound terminally to an Fe center of FeMoco would add a strong field ligand and increase the coordination number to five, therefore increasing the ligand field splitting. Since few studies have experimentally characterized Fe complexes with a comparable coordination environment,[243,244] it is not obvious how the electronic structure of FeMoco responds to a CO ligand. Therefore, *in silico* diamagnetic substitution is a convenient procedure to generate the relevant model compound and study the electronic structure of a CO-bound Fe center in the nearly exact ligand environment provided by FeMoco.

Electronic structure of CO−[$FeGa_6InS_9C$]

The geometry for CO bound to Fe6 in the diamagnetically substituted [$FeGa_6InS_9C$] is shown in Figure 5.2a. The optimization of the CO coordinates yields stable geometries for all three tested oxidation states $Fe^{3+,2+,1+}$ and all possible M_S. While the high spin M_S value is the most stable for Fe6 in the substrate-free [$FeGa_6InS_9C$], either an intermediate or a low spin M_S value is stabilized upon CO binding (Figure 5.2d). Repeating the calculations for Fe2 instead of Fe6 or with FeVco instead of FeMoco leads to qualitatively equivalent results, which is not surprising, since the ligand environment of the Fe center is similar in those cases (see Figure A17).

The orbital occupation scheme for the Fe center can be derived from the localized orbitals, which helps to visualize interactions between metal and ligand (Figure 5.2b and c). Only the xz and yz orbitals have the correct symmetry for the π system of CO and therefore have significant metal-ligand overlap. The series $Fe^{3+\rightarrow 2+\rightarrow 1+}$ demonstrates nicely that these two orbitals are preferentially doubly occupied, if a sufficient number of electrons is available, *i.e.* not in Fe^{3+}. The z^2 orbital, which is oriented along the Fe−CO bond, is unoccupied in all cases. The intermediate spin of Fe6 in [$FeGa_6InS_9C$] is therefore the result of an interplay between the σ donor and π acceptor ligand CO and the weak field ligands S^{2-}.

CO activation

The strength of the π back bond also reflects in the parameters often used to discuss CO activation (see Table 5.1). For the series $Fe^{3+\rightarrow 2+\rightarrow 1+}$, where the M_S values decrease as

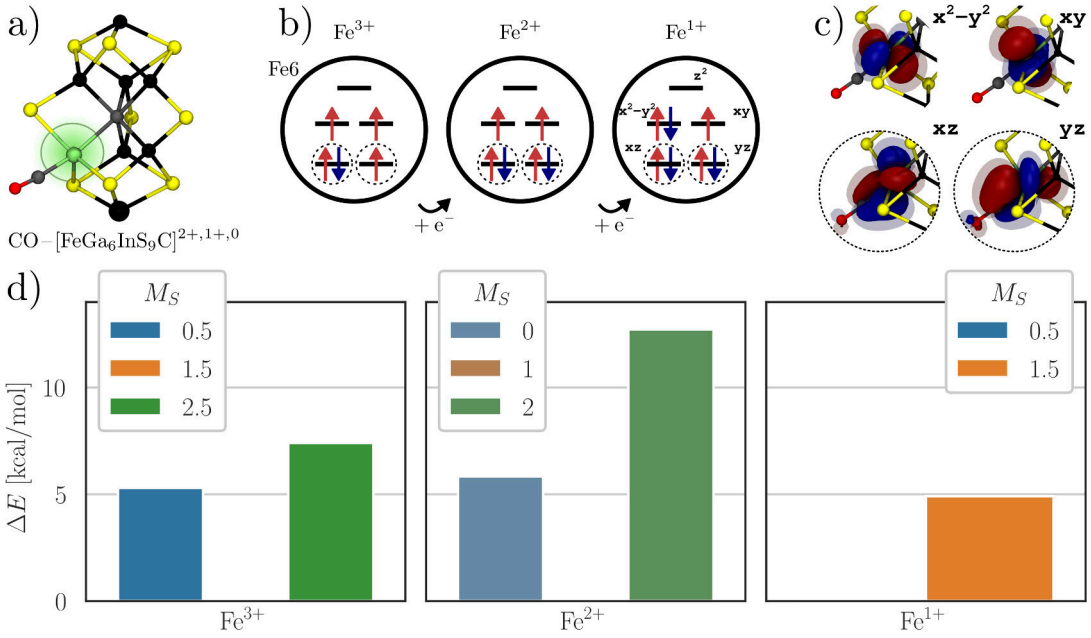

Figure 5.2. – Summary of the diamagnetically substituted cofactor $[FeGa_6InS_9C]$ with CO bound to Fe6. (a) Structure of CO–$[FeGa_6InS_9C]$. Diamagnetic ions (Fe→Ga, Mo→In) are shown in black. (b) Orbital occupation scheme for the most stable M_S value in the series $Fe^{3+,2+,1+}$ based on Foster-Boys localized orbitals. (c) Shape of orbitals in (b). Dashed lines highlight orbitals with a significant Fe–CO overlap. Isosurfaces are shown at ±0.05 and ±0.025. (d) Energies of different M_S values relative to the lowest one (M_S = 1.5, 1, 0.5 for $Fe^{3+,2+,1+}$, respectively). Figure (a-c) adapted from Ref [25].

$\frac{3}{2}$, 1, $\frac{1}{2}$, the C–O bond length increases from 1.129 Å (free CO) to 1.146 Å, 1.161 Å, and 1.183 Å, respectively, while the Fe–CO bond length decreases from 1.850 Å to 1.756 Å to 1.710 Å, respectively. At the same time, the CO frequency, which is also often used as a measure for CO activation, decreases from 2143 cm^{-1} (free CO) to 1978 cm^{-1}, 1884 cm^{-1}, and 1772 cm^{-1}. Therefore, a more electron-rich Fe center clearly results in a longer C–O bond, a shorter Fe–CO bond, and a lower CO frequency. However, while the changes in C–O bond length and CO frequency are strongly correlated ($r = -0.998$),[†] the Fe–CO bond length is more strongly affected by the $Fe^{3+\rightarrow2+}$ reduction, but the C–O bond length by the $Fe^{2+\rightarrow1+}$ reduction ($r = -0.954$). This showcases that the C–O and the Fe–CO bond length are not equally well suited as a measure for CO activation.

As it is obvious from the orbital occupation scheme, the Fe–CO bond favors a strong π back bond and therefore induces spin-pairing. According to the calculated CO frequency and the C–O bond length, CO activation is increased for the $Fe^{2+\rightarrow1+}$ reduction compared to $Fe^{3+\rightarrow2+}$, even though all four spin orbitals involved in the π back bond are already occupied in Fe^{2+}: the additional electron in Fe^{1+} occupies the $x^2 - y^2$ orbital, which is not directly involved in the π back bond. However, the

[†]The quantity Person's r measures the correlation between two variables. If $r = \pm1$, the variables are linearly dependent and their functional dependence is a straight line.

Table 5.1. – Characterization of the CO−Fe interaction in the diamagnetically substituted cofactors models CO−[$FeGa_6InS_9C$], CO−[$Fe_2Ga_5InS_9C$], and μ−CO−[$Fe_2Ga_5InHS_9C$]. The respective structures are shown in Figure 5.2, 5.3, and 5.4. Note that the CO coordinates were not optimized for μ−CO−[$Fe_2Ga_5InHS_9C$]. S_{pop} denote Hirshfeld spin populations. ν_{CO} are scaled frequencies (see Table A18).

binding mode	M_S	S_{pop} Fe6	S_{pop} Fe2	ν_{CO} [cm^{-1}]	$d_{C–O}$ [Å]	$d_{Fe–CO}$ [Å]
			$[FeGa_6InS_9C]^{2+,1+,0}$			
CO−Fe^{3+}	1.5	2.50	-	1978	1.146	1.850
CO−Fe^{2+}	1.0	1.84	-	1884	1.161	1.756
CO−Fe^{1+}	0.5	1.04	-	1772	1.183	1.710
			$[Fe_2Ga_5InS_9C]^{1+}$			
CO−$Fe^{2+}\uparrow Fe^{3+}\uparrow$	3.5	2.05	3.58	1908	1.157	1.768
CO−$Fe^{2+}\downarrow Fe^{3+}\uparrow$	1.5	−1.69	3.59	1917	1.157	1.757
			$[Fe_2Ga_5InHS_9C]^{2+}$			
μ−CO−$Fe^{2+}Fe^{3+}$	2.5	2.29	2.50	-	1.193	1.837/1.960

metal-ligand overlap of the xz and yz orbitals are increased by the $Fe^{2+\rightarrow 1+}$ reduction. Therefore, the occupation of the metal-based π-donor orbitals has an equally large effect on CO-activation as enhancing the donor strength of the metal through further reduction.

The experimental IR bands for terminally bound CO are reported between 1900 cm^{-1} and 1980 cm^{-1} for binding to FeMoco in the MoFe protein.[226,227,231,232] The calculated frequencies for an Fe^{3+} or an Fe^{2+} center in [$FeGa_6InS_9C$], 1978 Å and 1884 Å, respectively, are in good agreement with the experimental values and consistent with the commonly discussed average $Fe^{2.5+}$ oxidation state in FeMoco.[17,98,200] In particular, the transient experimental band at 1904 cm^{-1} ($loCO_{IR,1}$), which captures the initial CO binding event, is in good agreement with the calculated frequency for CO bound to the Fe^{2+} center. The conversion of the $loCO_{IR,1}$ band to $loCO_{IR,2}$ (1715 cm^{-1}) is usually interpreted as a change in binding mode from terminal to bridging. The calculated frequency for the Fe^{1+} center (1772 cm^{-1}) suggests that a frequency below 1800 cm^{-1} may be also possible for a terminally bound CO. However, the accumulation of charge at the binding site to form an Fe^{1+} center appears rather unlikely considering the low barrier for redox reorganization within large FeS clusters such as FeMoco.[98] Furthermore, the frequency for Fe^{1+} is still more than 50 cm^{-1} higher than the experimentally observed band.

5.2.2. Terminal CO binding to an Fe dimer

As discussed in Section 4.2.2, the high spin Fe centers in FeMoco preferentially couple antiferromagnetically. The study of the Fe−CO bond in [$FeGa_6InS_9C$] shows that π back bonding results in local spin-pairing and that an intermediate spin is energetically preferred at the binding site. Decreasing the local spin would decrease

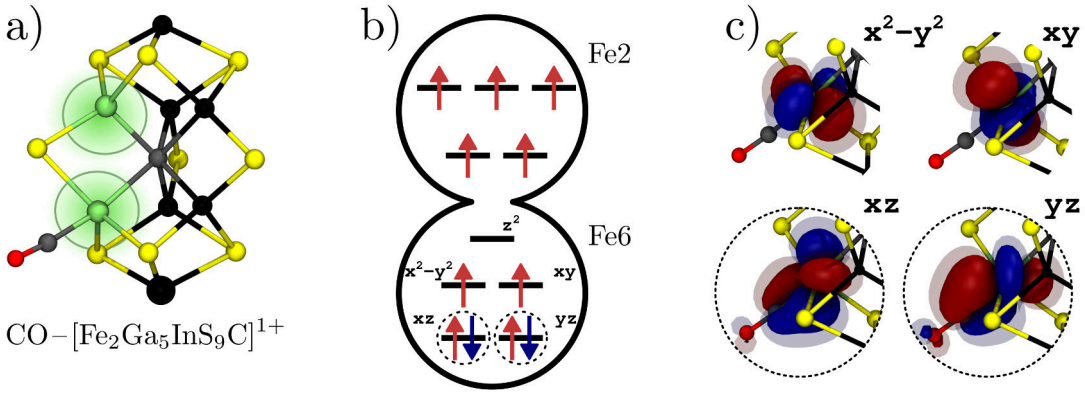

Figure 5.3. – Geometric and electronic structure of the diamagnetically substituted cofactor $[Fe_2Ga_5InS_9C]$ with CO bound to Fe6. (a) Structure of CO−$[Fe_2Ga_5InS_9C]$. Diamagnetic ions (Fe→Ga, Mo→In) are shown in black. (b) Orbital occupation scheme based on Foster-Boys localized orbitals for the high spin $M_S = \frac{7}{2}$ determinant, which is the most stable for an $Fe^{2+}Fe^{3+}$ redox state. (c) Shape of orbitals in (b). Dashed lines highlight orbitals with a significant Fe−CO overlap. Isosurfaces are shown at ±0.05 and ±0.025. Figure adapted from Ref [25].

the antiferromagnetic coupling strength, and it is therefore interesting to assess the relative importance of spin-pairing and spin-coupling. This can be done with the diamagnetically substituted cofactor $[Fe_2Ga_5InS_9C]$, where two instead of one Fe center are left unsubstituted: Fe6, which binds CO, and Fe2, which is magnetically coupled to Fe6. The structure of CO-bound $[Fe_2Ga_5InS_9C]$ is shown in Figure 5.3a.

Substrate-free $[Fe_2Ga_5InS_9C]$

In the substrate-free FeMoco, the Fe6 center exhibits strong antiferromagnetic coupling to all neighboring metal centers, which has been demonstrated in Figure 4.6b. More precisely, according to BS-DFT, the mixed-valent $Fe^{2+}Fe^{3+}$ pair in $[Fe_2Ga_5InS_9C]$ is antiferromagnetically coupled in an $M_S = \frac{1}{2}$ determinant, where the electron is localized on one of the Fe centers ($J = -117\,cm^{-1}$, also Table A11). In the corresponding high spin determinant ($M_S = \frac{9}{2}$), which is obviously higher in energy, one electron is shared between Fe6 and Fe2 to form a valence-delocalized $Fe^{2.5+}Fe^{2.5+}$ pair. Fe6 and Fe2 maintain their local high spin states in both the high spin and the BS determinant. Valence delocalization is frequently observed in BS-DFT determinants of unsubstituted FeMoco between ferromagnetically aligned Fe centers (Figure 4.3).

Spin-pairing vs. spin-coupling

As can be seen from the orbital occupations in Figure 5.3b, The CO-bound Fe6 center is ferromagnetically coupled to Fe2 in the $[Fe_2Ga_5InS_9C]$ model. The coupling constant J is $+34\,cm^{-1}$, which is in stark contrast to the substrate-free $[Fe_2Ga_5InS_9C]$ ($-117\,cm^{-1}$). However, the lowest-energy determinant does not have an $M_S = \frac{9}{2}$, where all α orbitals are occupied, but an $M_S = \frac{7}{2}$. The localized orbitals show that Fe6 is an intermediate spin Fe^{2+} center with the same orbital occupation as the single

CO-bound Fe center in [$FeGa_6InS_9C$] (Figure 5.2b). That is, the xz and yz orbitals are doubly occupied and involved in the π back bond while the z^2 orbital, which is oriented along the Fe−CO bond, is unoccupied. Fe2 is a high spin Fe^{3+} center.

The change from a high to an intermediate spin on Fe6 delivers a ready explanation for the reduced coupling. Furthermore, CO coordination disturbs the geometry of the cofactor (*vide infra*) and the Fe6−Fe2 bond length increases from 2.59 Å to 2.78 Å, which mitigates the magnetic interactions. Since antiferromagnetic coupling plays a decisive role in the stability of BS determinants (see Figure 4.6), terminal CO binding is expected to affect their relative energies. In other words, the antiferromagnetic alignment of the CO-bound Fe center may be less important than between the unbound Fe centers.

CO activation and metal-metal interactions

The parameters of the Fe−CO bond in [$Fe_2Ga_5InS_9C$] are summarized in Table 5.1. Here, the CO coordinates were optimized for both the ferromagnetic and the antiferromagnetic alignment of Fe6 and Fe2 (CO−$Fe^{2+}\uparrow Fe^{3+}\uparrow$ and CO−$Fe^{2+}\downarrow Fe^{3+}\uparrow$, respectively). The calculated vibrational frequency and the C−O bond length are similar to the CO-bound Fe^{2+} center in [$FeGa_6InS_9C$], indicating a rather small influence of the second Fe center. However, one has to keep in mind that the diamagnetic substitution of Fe2 still introduces an error in the electrostatic influence and potential differences between [$Fe_2Ga_5InS_9C$] and the [$FeGa_6InS_9C$] may not only be due to magnetic interactions.

The CO vibrational frequency for the ferromagnetically aligned Fe dimer is with 1908 cm^{-1} slightly lower than for the antiferromagnetic alignment (1917 cm^{-1}). The difference in CO activation can be explained with the Pauli exclusion principle: when both Fe centers are ferromagnetically aligned, the majority spin electrons on Fe6 cannot delocalize towards Fe2, which already holds the maximum number of α electrons. The small difference in frequency does not reflect in the C−O bond length (1.157 Å in both cases). This is not surprising, since when interpolating the changes for [$FeGa_6InS_9C$], one would expect the C−O bond length to change by about 0.001 Å for a frequency change of 10 cm^{-1}. On the other hand, the Fe−CO bond length is more susceptible to CO activation and it decreases by 0.01 Å along with the 10 cm^{-1} increase in frequency. All in all, magnetic interactions between neighboring metal centers have a small, but measurable effect on CO activation.

5.2.3. CO-bridged Fe dimer

That CO eventually replaces the S2B sulfide and assumes the bridging position between Fe6 and Fe2, has been established through X-ray crystallography ($loCO_{XRD}$, see Figure 1.9).[233,234] The discussion in the coming sections will show that the bridging CO motif requires both the protonation of the S2B sulfide and the interaction with the His195 residue. Therefore, the previous approach to construct the diamagnetically substituted cofactors from the CO-bound cluster model is not applicable for the bridging CO motif. Instead, the diamagnetic substitution is performed on the geometry of the E_1 QM/MM model featuring a bridging CO, resulting in the μ−CO−[$Fe_2Ga_5InHS_9C$]

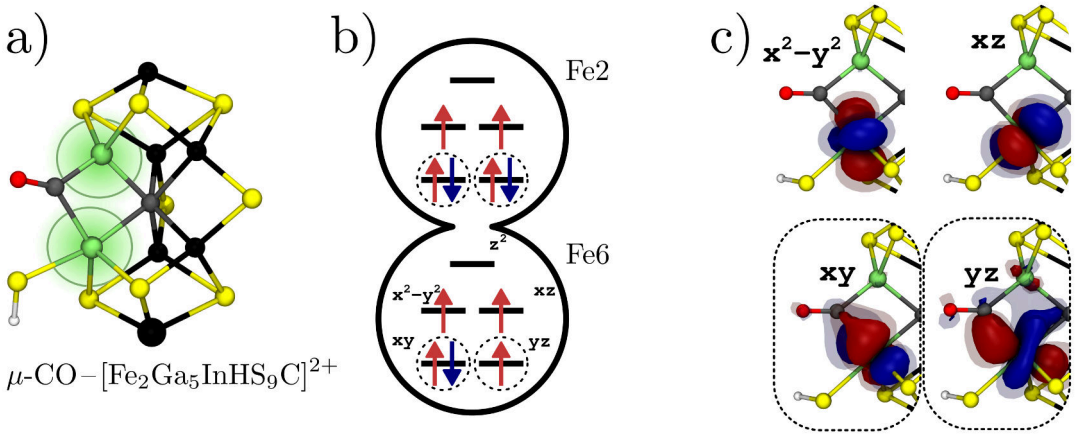

Figure 5.4. – Geometric and electronic structure of the diamagnetically substituted cofactor $[Fe_2Ga_5InHS_9C]$ with CO bridging Fe6 and Fe2. (a) Structure of μ−CO−$[Fe_2Ga_5InHS_9C]$. Diamagnetic ions (Fe→Ga, Mo→In) are shown in black. (b) Orbital occupation scheme based on Foster-Boys localized orbitals for the high spin $M_S = \frac{5}{2}$ determinant, which is the most stable for an $Fe^{2+}Fe^{3+}$ redox state. (c) Shape of orbitals in (b). Dashed lines highlight orbitals with a significant Fe−CO overlap. Isosurfaces are shown at ±0.05 and ±0.025. Figure adapted from Ref [25].

model (Figure 5.4a). This structure is not further optimized, since the bridging CO motif is not stable, and therefore no reliable CO frequency can be calculated. The geometric parameters are given in Table 5.1 for completion, but note that they correspond to the optimized geometry of the QM/MM model. Nevertheless, it is instructive to study the electronic structure of the bridging CO motif here, since the replacement of the S2B belt sulfide constitutes a dramatic change to ligand environment of both Fe6 and Fe2.

The localized orbitals for μ−CO−$[Fe_2Ga_5InHS_9C]$ are shown in Figure 5.4b and c. The oxidation states are $Fe6^{3+}$ and $Fe2^{2+}$, seemingly in response to the SH^- ligand bound to Fe6. Both Fe centers have a local intermediate spin, where the orbital oriented along the Fe−carbide bond is unoccupied (z^2). As for the terminal binding mode, spin-pairing appears to be a consequence of CO coordination. For the bridging motif, the maximum number of orbitals that can participate in the π back bond is eight, out of which seven are occupied. Due to anisotropy in the Fe−CO bond, one has to distinguish between in-plane (yz) and out-of-plane interactions (xy). The occupation of the orbitals on Fe6 suggests that the out-of-plane interaction is favored.

The bridging CO favors the ferromagnetic alignment of Fe6 and Fe2 ($J = +47\,\mathrm{cm}^{-1}$, Table A11). This may again be related to local spin-pairing, however, the yz orbital shows significant delocalization over Fe6 and Fe2. Valence delocalization is an important feature of the double exchange mechanism, which stabilizes the ferromagnetic state (see Chapter 3). All in all, both the terminal and the bridging CO motif have shown to change the preferred alignment of Fe6 and Fe2 from anti- to ferromagnetic, which is expected to disrupt the energy ordering of the BS determinants in the full cofactor models.

5.3. Cluster model: metal-metal interactions

The previous section discussed possible changes that CO binding can cause locally at the binding site. Spin-pairing and π back bonding was found to be a characteristic feature of the Fe−CO bond in the diamagnetically substituted cofactors. The present section focuses on the electronic structure within the unsubstituted, full cofactor model and the balance between the local aspects of the Fe−CO bond and the influence of the magnetic interactions between all metal centers. For now, the asymmetry of the protein environment is neglected and the discussion is limited to the cluster model, where the protein environment is modeled to the most part by a uniform polarizable continuum (Figure 4.1). A dielectric constant of $\epsilon = 4$ is used, as is commonly done to mimic the electrostatic influence of the bulk protein.[317] Section 5.4 will discuss the QM/MM models of nitrogenase, where all amino acid residues are modeled explicitly.

The interaction of CO with FeMoco in the resting state of the MoFe protein has not been observed experimentally, which is most likely true as well for solution-extracted FeMoco.[15,237] However, the resting state is by far the best-understood redox state, both experimentally and computationally. It is therefore instructive to study how an Fe−CO bond alters the electronic structure of the E_0 FeMoco cluster model, which is done in Section 5.3.1. CO binding to the experimentally more relevant E_1 FeMoco cluster model, which is reduced and protonated relative to E_0, is explored in Section 5.3.2. Section 5.3.3 highlights differences between FeMoco and FeVco in the previous analysis. The structures for these models are shown in Figure 5.5.

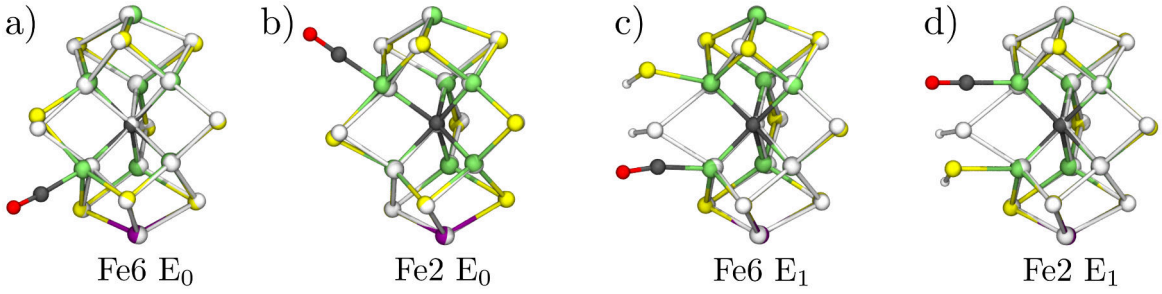

Figure 5.5. – CO binding motifs for the FeMoco cluster models in (a,b) the E_0 and (c,d) the E_1 redox state. The substrate-free structure is overlaid in black and white. The binding motifs for the FeVco cluster model are equivalent.

5.3.1. Binding to the E_0 redox state

Screening BS/M_S combinations

CO binding is expected to favor spin-pairing at the binding site, which (i) may affect the relative energies of the BS determinants and (ii) may change the most stable M_S value. Prior to CO binding the determinants of the class BS7 with an $M_S = \frac{3}{2}$ are the most stable, followed by the classes BS8 and BS10 (Figure 4.6). Therefore, these three classes and M_S values of $\frac{1}{2}$, $\frac{3}{2}$, and $\frac{5}{2}$ are likely combinations for the CO-bound E_0 FeMoco model. The energies are estimated using the cluster* model (Figure 4.1), which employs a small basis set and is therefore suitable to run many calculations for

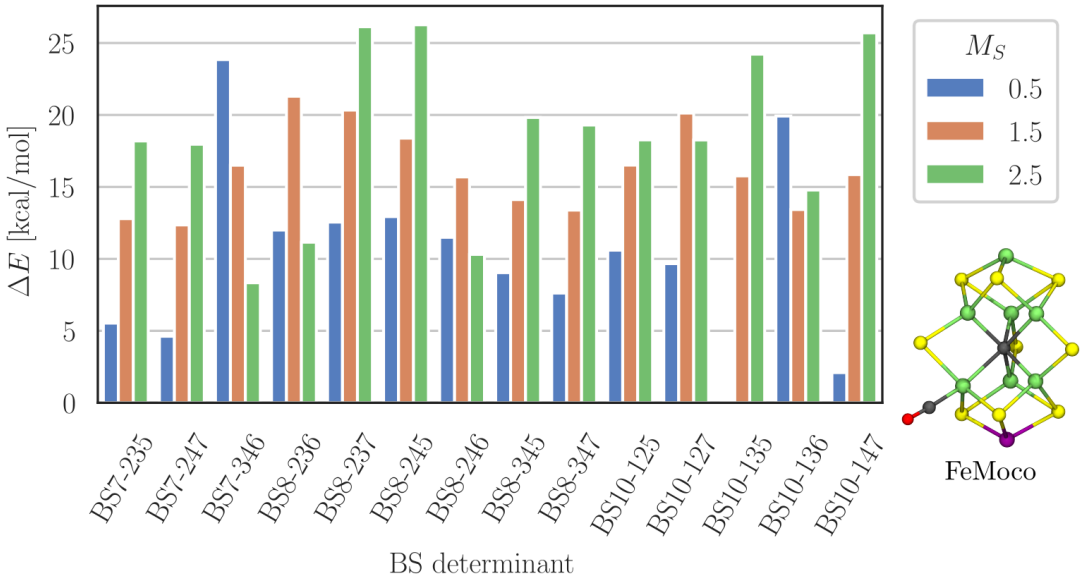

Figure 5.6. – Energies for the E_0 FeMoco cluster* model with CO bound to Fe6 relative to BS10-135 $M_S = \frac{1}{2}$. The plot shows all BS determinants in the classes BS7, BS8, and BS10.

screening purposes. The error due to the smaller basis is on the order of 1 kcal/mol,[†] which makes subsequent verification with a larger basis set necessary. The cluster* calculations allow for the identification of important trends in the energies and help to rule out many BS/M_S combinations.

The energies for CO bound to Fe6 in the cluster* model are shown in Figure 5.6. In contrast to the substrate-free cofactor, the respective $M_S = \frac{3}{2}$ determinant is not among the lowest in energy (except for BS10-136). The energy of the $M_S = \frac{1}{2}$ determinant is lower than the $M_S = \frac{5}{2}$ determinant whenever Fe6 is not flipped in a given BS determinant and *vice versa* whenever the spin on Fe6 is flipped. This observation is consistent with the orbital occupation schemes of the diamagnetically substituted cofactors, where spin-pairing decreases the local M_S value at Fe6. This decreases the global M_S value if Fe6 is spin up, but increases it if Fe6 is spin down.

Interestingly, the BS7 class does not constitute the lowest-energy determinant for the CO-bound models. Rather, the BS10-135 and BS10-147 determinants are the most stable, the two of which are related through a pseudo-mirror plane including CO and Fe6. To summarize, CO binding disrupts the energy ordering of the BS determinants and stabilizes other M_S values than $\frac{3}{2}$. Therefore, in the subsequent analysis, the determinants BS10-135 and BS10-147 are tested in addition to the BS7 class and the M_S value is adjusted either up or down to account for spin-pairing at the binding site.

[†]The energy gap between the $M_S = \frac{3}{2}$ and the $M_S = \frac{5}{2}$ BS7-235 determinant is 1.36 kcal/mol for the cluster* model (def2-SVP), but 0.06 kcal/mol for the cluster model (def2-TZVP).

Geometry

The E_0 FeMoco cluster model converges to stable structures with CO bound to Fe6 and Fe2, which are shown in Figure 5.5a and b. CO binding increases the coordination number of the respective Fe center from four ($3 \times S^{2-}$ and $1 \times C^{4-}$) to five (+CO). With the change in the ligand environment from approximately tetrahedral to trigonal bipyramidal, the CO-bound Fe center is pulled out of the cofactor, which increases the bond lengths to the neighboring metal centers by 0.1 Å–0.2 Å relative to the substrate-free model. At the same time, the Fe−carbide distance increases from 2.00 Å (substrate-free) to 2.18 Å and 2.15 Å when binding to Fe6 and Fe2 respectively. The remaining bond lengths change by less than 0.05 Å.

Electronic structure

The lowest-energy BS determinants for the CO-bound E_0 FeMoco cluster model are characterized in Table 5.2. As predicted by the cluster* model, the BS10-135 determinant is the most stable when CO is bound to Fe6 (see Table A13 for all tested BS/M_S combinations). For CO bound to Fe2, the BS10-147 determinant the most stable, which is the pseudo-mirror image of BS10-135 with respect to Fe6/Fe2. In both cases, the $M_S = \frac{1}{2}$ determinant is more stable than $M_S = \frac{3}{2}$ or $M_S = \frac{5}{2}$. As indicated by the spin populations, the reduction in the global M_S value is directly related to the decrease in the local spin at the respective CO-bound Fe center, which is spin up in both determinants.

The localized orbitals corroborate that the electronic structure changes are limited to the binding Fe center (Figure A18). The orbital occupation scheme, which can be derived from the localized orbitals, is qualitatively equivalent to that of the Fe^{2+} center in the diamagnetically substituted CO-bound [$FeGa_6InS_9C$] (Figure 5.2). The electron localization due to CO binding ($Fe^{2.5+}Fe^{2.5+} \rightarrow CO-Fe^{2+}Fe^{3+}$), which was observed in the [$Fe_2Ga_5InS_9C$] models (Figure 5.3) is also happening unsubstituted models. Here, the localization happens between Fe pairs within the same cubane. For example, in the substrate-free BS10-147 determinant, an electron is shared between Fe2 and Fe3. CO binding to Fe2 causes this electron to localize, yielding the oxidation states $Fe2^{2+}Fe3^{3+}$. This showcases the utility of diamagnetic substitution to decompose the complex electronic structure of larger metal clusters. However, some more subtle effects cannot be captured by the diamagnetically substituted models: the orbitals on CO-bound Fe center have some overlap with multiple neighboring metal centers, especially if they are ferromagnetically aligned. This may indicate the involvement of several metal centers on the Fe−CO bond.

CO activation

The Fe−CO bond parameters for the E_0 FeMoco cluster model are listed in Table 5.2. The calculated CO frequency for the Fe6-bound model is 1910 cm^{-1}, which is within 10 cm^{-1} of the frequency in CO−[$Fe_2Ga_5InS_9C$] (Table 5.1). Again, this suggests that the diamagnetically substituted model already includes a majority of the effects contributing to CO-activation in FeMoco. The calculated CO frequency for Fe2 is about 15 cm^{-1} lower, indicating a slightly more activated CO for the Fe-only cubane

Table 5.2. – Summary of the lowest-energy CO-bound FeMoco and FeVco cluster models in the redox states E_0 and E_1. S_{pop} denote Hirshfeld spin populations. ν_{CO} are scaled frequencies (see Table A18). The substrate-free models used as reference for the binding energy are listed in Table A12.

binding site	S_{pop} Fe6	S_{pop} Fe2	ν_{CO} [cm^{-1}]	$d_{C\text{-}O}$ [Å]	$d_{Fe\text{-}CO}$ [Å]	BS	M_S	ΔE_{bnd} [kcal/mol]
				E_0 FeMoco				
Fe6	1.56	3.19	1910	1.158	1.760	BS10-135	0.5	-6.0
Fe2	2.93	1.95	1895	1.159	1.756	BS10-147	0.5	-9.1
				E_1 FeMoco				
Fe6	1.07	3.22	1822	1.171	1.735	BS10-147	0	-15.0
Fe2	3.00	1.61	1856	1.166	1.745	BS10-147	0	-11.0
				E_0 FeVco				
Fe6	1.96	3.22	1887	1.162	1.748	BS8-345	0.5	-5.3
Fe2	3.28	-1.62	1894	1.162	1.735	BS7-235	2.5	-5.1
				E_1 FeVco				
Fe6	1.82	3.11	1799	1.176	1.749	BS10-135	0	-11.8
Fe2	3.03	1.87	1810	1.173	1.749	BS10-135	0	-8.9

compared to the Mo cubane. The lower frequency correlates again with a small, albeit consistent, elongation of the C−O and a contraction of the Fe−CO bond.

Furthermore, the total energy for the Fe2-bound model is about 3 kcal/mol lower compared to Fe6. The reason for this binding site preference is not obvious. Since the CO activation is rather similar in both cases, the energy difference may be related to the electronic structure of each cubane: In the substrate-free BS7 determinants, Fe2 couples antiferromagnetically to Fe1, whereas Fe6 couples antiferromagnetically to Mo. If the loss of spin-coupling were responsible for the energy difference, the reduction of the local spin on Fe6 and hence breaking the coupling between Fe and Mo would be less favorable than breaking the coupling between Fe and Fe.

The contribution of Mo to the reactivity of FeMoco is not well understood. For the diamagnetically substituted models, CO activation was clearly related to the orbital occupation schemes derived from the localized orbitals (Figure 5.2). For the unsubstituted models, the electron count on CO-bound Fe6 is higher compared to CO-bound Fe2 due to an orbital shared between Mo and Fe6 (Figure A18), even though CO is more activated for Fe2. Since the localized orbital analysis cannot explain the small differences between the Fe-only and the Mo cubane, it may not be applicable when different metals are involved.

5.3.2. Binding to the E_1 redox state

Geometry

In going from the E_0 to the E_1 cluster model, FeMoco is reduced by one electron and the S2B sulfide is protonated. As shown in Figure 5.5, CO binding to either Fe6 or Fe2 leads spontaneously to an opening of the S2B bridge, resulting in a terminal CO and a terminal SH^- on the neighboring Fe centers. The Fe6−Fe2 distance is with about 2.56 Å close to that in the substrate-free E_0 cluster model (2.59 Å) and the bond lengths not involving Fe6 or Fe2 change by less than 0.03 Å. Therefore, the metal-metal distances in the CO-bound E_1 model are close to the substrate-free E_0 model, whereas CO pulls one Fe center out of the cofactor in the E_0 model. Nevertheless, the opening of the S2B bridge is expected to have pronounced effect on the electronic structure.

Electronic structure

For the substrate-free E_1 cluster model, the BS7-346 determinant with an $M_S = 2$ is lowest in energy, which is consistent with the E_1 QM/MM model reported previously.[204] One of the Fe centers in the Mo cubane is reduced relative to the E_0 redox state (Fe6 in case of BS7-346). The relative energies of the BS determinants in the substrate-free E_1 cluster model are compared in Figure A24.

The metal spin populations for the most stable BS determinants are shown along with the structures in Figure 5.7. The spin populations allow for a quick identification of the BS determinant and possible local spin-pairing. CO binding and the accompanied SH^- bridge opening leads to two major changes in the electronic structure: (i) The BS10-147 $M_S = 0$ determinant is the lowest in energy for both binding sites. (ii) The local spin at the CO-bound Fe center is reduced, which indicates spin-pairing. The localized orbitals (Figure A19) further support an intermediate local spin. The respective Fe center has only four occupied majority spin orbitals and the orbital oriented along the Fe−CO bond is unoccupied. However, the minority spin orbitals cannot be unequivocally assigned as either localized or delocalized. Therefore, it is not possible to draw a simple orbital occupation scheme as for the E_0 cluster model and the diamagnetically substituted models. Some orbitals having contributions from both Fe6 and Fe2 may also indicate that these exposed centers are rather reactive. Nevertheless, when neglecting the small delocalization tail between Fe6 and Fe2, the localized orbitals allow for determining the redox states of the cubanes. The additional electron in the E_1 models relative to E_0 follows the CO molecule and resides in the Mo cubane, when CO is bound to Fe6, and in the Fe-only cubane, when CO is bound to Fe2.

CO activation

The CO bond in the E_1 models is significantly more activated compared to the E_0 models (Table 5.2). The calculated CO frequencies, 1822 cm^{-1} for Fe6 and 1856 cm^{-1} for Fe2, are 90 cm^{-1} and 90 cm^{-1} lower compared to the E_0 model, respectively. Again, a lower CO frequency correlates with a longer C−O and a shorter Fe−CO bond length. As discussed above, the localized orbitals are not a concise description of the electronic structure at the binding site and therefore cannot be used to explain the differences

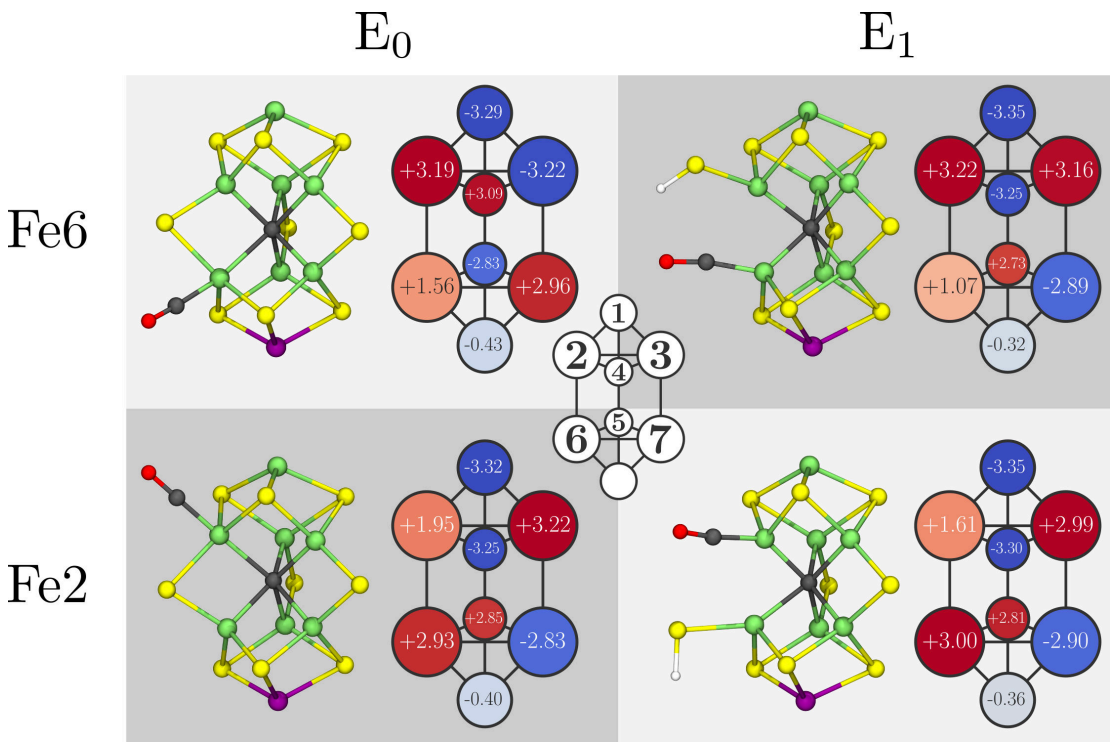

Figure 5.7. – Structures and Hirshfeld spin populations of the metal centers for the CO-bound FeMoco cluster model in the redox states E_0 and E_1. The lowest-energy models are BS10-135 $M_S = \frac{1}{2}$ (Fe6/E_0), BS10-147 $M_S = \frac{1}{2}$ (Fe2/E_0), and BS10-147 $M_S = 0$ (Fe6/E_1 and Fe2/E_1).

in CO activation. The spin population (Figure 5.7) as well as the atomic charge (not shown) of the CO-bound Fe center is lower compared to the E_0 model, which indicates a higher degree of spin-pairing and a more electron rich Fe center in E_1.

In addition to the more reduced cofactor, the structural changes may also contribute to CO activation. The concomitant opening of the S2B bridge due to protonation leaves the coordination number of all Fe centers unchanged, resulting in a less convoluted binding site compared to the E_0 models. The Fe−CO bond lengths is about 1.74 Å in the E_1 models, which is 0.02 Å shorter compared to the E_0 models. At the same time, the Fe−carbide bond is 0.1 Å–0.2 Å shorter, which increases the electron density at the Fe center and therefore enhances π back bonding to CO.

Binding energies

The binding energies in the E_1 cluster model are −15.0 kcal/mol when CO is bound to Fe6 and −11.0 kcal/mol when bound to Fe2 (Table 5.2, all other tested BS/M_S combinations are listed in Table A14). This is significantly stronger than in the E_0 cluster model and the release of energy would be sufficient to overcome the entropic penalty associated with the binding event, which is estimated to be around 10 kcal/mol at 300 K.[318] However, it is important to keep in mind that this estimate corresponds to gas phase reactions and the protein environment may have a critical impact on thermodynamic properties. Furthermore, the binding site preferences is inverted

relative to E_0, because Fe6 is the preferred binding site. This can be rationalized with the electronic structure of the two cubanes: When CO binds to the E_0 models, it merely leads to charge localization within the respective cubane and an intermediate spin at the binding site. The intermediate spin, in turn, affects the spin-coupling, which is different in the Fe-only and the Mo cubane due to the Mo center. In the E_1 models, however, the additional electron exhibits some flexibility and follows CO into the respective cubane. Since the Mo cubane is the preferred residence of the electron according to the substrate-free E_1 model, CO binding to Fe2 comes with the cost of the electron transfer into the Fe-only cubane.

5.3.3. Differences between FeMoco and FeVco

Electronic structure in the E_0 redox state

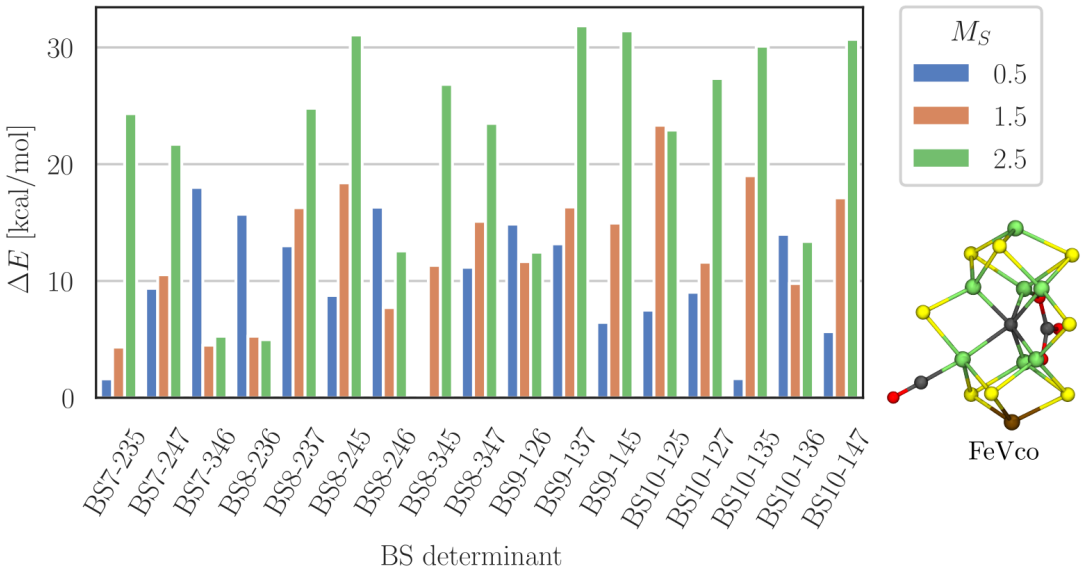

Figure 5.8. – Energies for the E_0 FeVco cluster* model with CO bound to Fe6, relative to BS8-347 $M_S = \frac{1}{2}$. The plot show all BS determinants in the classes BS7, BS8, and BS10. The equivalent plot for FeMoco is Figure 5.6.

As discussed in Section 5.3.1, a larger number of BS/M_S combinations can be tested using the less accurate but computationally inexpensive cluster* model. Similar to FeMoco, the classes BS7, BS8, and BS10 are lowest in energy for the substrate-free FeMoco (Figure 4.9). Therefore, these determinants are tested in combination with $M_S = \frac{5}{2}$, $\frac{3}{2}$, and $\frac{1}{2}$ for CO binding to Fe6 in the E_0 FeVco cluster* model and the resulting energies are shown in Figure 5.8. The lowest-energy determinant is BS8-345 followed by BS10-135 and BS7-235. This is in stark contrast to FeMoco, where BS8-345 is about 10 kcal/mol higher in energy.

The near-degeneracy of the BS determinants in a given class is lifted due to the presence of the carbonate ligand and the heterometal V, as discussed for the substrate-free cofactors (Figure 4.8). This also holds true for the CO-bound cofactor. For example, the energy of BS10-147 is about 5 kcal/mol higher than BS10-135. These two

determinants are pseudo-mirror images with respect to Fe6 in FeMoco, where they lie merely 2 kcal/mol apart, demonstrating the symmetry breaking due to the carbonate ligand in FeVco. It is not obvious why the BS8-345 determinant becomes lowest in energy when binding CO to Fe6, and it is therefore also possible that other BS classes not tested here would be even lower in energy. When going beyond the resting state, it is of utmost importance to either screen many BS/M_S combinations or to understand what contributes to the stability of a BS determinant, as discussed in Section 4.2.2 and 4.2.3. For FeVco, the first step would be to understand the effect of the carbonate ligand on the energies in the substrate-free E_0 BS determinants.

Consistent with the cluster* model, the BS8-345 determinant with an $M_S = \frac{1}{2}$ is also the most stable when binding CO to Fe6 in the cluster model (see Table A13 for all tested BS/M_S combinations). When binding to Fe2, the BS7-235 determinant is the most stable and since Fe2 is spin down in this determinant, the reduction of the local spin at Fe2 favors a global $M_S = \frac{5}{2}$. Even though completely different BS determinants are stabilized in FeVco, the localized orbitals reveal that the local electronic structure at the CO-bound Fe center is the same as in FeMoco (Figure A20). To summarize, the CO-bond Fe center has the same intermediate local spin and an Fe^{2+} oxidation state in all E_0 cluster models, but the BS determinant specifies the location of the ferromagnetically aligned pairs, and therefore the origin of the to-be-localized electron.

Electronic structure in the E_1 redox state

The main difference between FeMoco and FeVco in the substrate-free E_0 redox state lies in the heterometal and the oxidation state of Fe5 (green in Figure 5.9). For the BS7-235 determinant, the oxidation states are $Mo^{3+}Fe5^{3+}$ for FeMoco and $V^{3+}Fe5^{2+}$ for FeVco (see also discussion of Figure 4.8). This directly affects the location of the additional electron in the E_1 redox state (purple in Figure 5.9). For FeMoco, the $Fe5^{3+}$ center gets reduced to $Fe5^{2+}$, but in FeVco, the valence-delocalized $Fe1^{2.5+}Fe4^{2.5+}$ pair gets reduced to a localized $Fe1^{2+}Fe4^{2+}$ pair. Since the additional electron is added to a β orbital in the latter case, the E_1 FeVco model has an $M_S = 1$. The BS7-235 determinant maintains the isolated energetic position in the E_1 FeVco model as a result of the carbonate ligand (see Figure A24).

CO binding to either Fe6 or Fe2 in the E_1 FeVco model is accompanied by the same opening of the protonated S2B sulfide bridge discussed for FeMoco (Figure 5.7) Similarly, the CO-bound Fe centers have a reduced local spin. According to the localized orbitals (not shown), the orbital along the Fe−CO bond is unoccupied. The localized orbitals further show that the additional electron in E_1 FeVco follows CO into the respective cubane, as already discussed for FeMoco.

The BS10-135 determinant is the most stable when binding CO to either Fe6 or Fe2 in the E_1 FeVco model, in contrast to FeMoco, where BS10-147 is most stable. The preference of FeVco for BS10-135 over BS10-147 has also been apparent for the CO-bound E_0 cluster* model (Figure 5.8). For the CO-bound E_1 FeVco model, the energy of the BS10-135 determinant is about 6 kcal/mol lower compared to BS10-147 (Table A14), which is likely related to the carbonate ligand breaking the symmetry and therefore represents a property specific to FeVco.

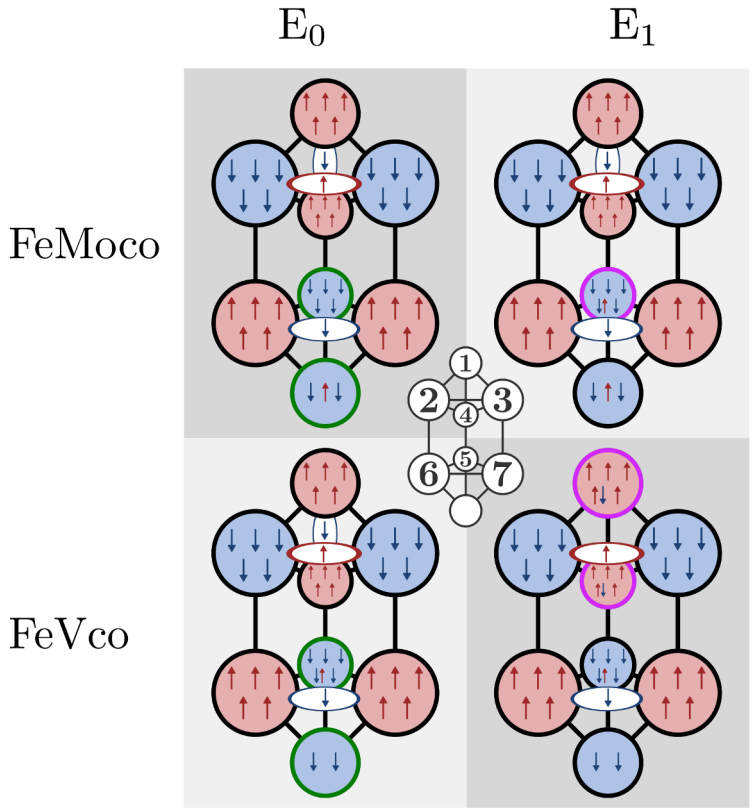

Figure 5.9. – Comparison between the orbital occupation schemes derived from Foster-Boys localized orbitals for the substrate-free FeMoco and FeVco cluster models in the E_0 and E_1 redox states (BS7-235, E_0 FeMoco/FeVco: $M_S = \frac{3}{2}$, E_1 FeMoco: $M_S = 1$, E_1 FeVco: $M_S = 2$). The differences between FeMoco and FeVco are highlighted in green for the E_0 redox state and in purple for the E_1 redox state.

Binding energy

The CO binding energies are consistently more negative for FeVco than for FeMoco (Table 5.2). Unlike E_0 FeMoco, where Fe2 is the preferred binding site, the E_0 FeVco cluster model shows no binding site preference. The reason for this may be related to the more reduced Fe centers in the V cubane (Figure 5.9), which stabilizes the electron relocation towards Fe6. Alternatively, the spin-coupling between Fe and the heterometal Mo or V could be an important contribution. A comparative study of FeMoco and FeVco through XAS measurements suggests that the Mo−Fe orbital overlap is larger than that for V−Fe.[299] A strong interaction between Fe and Mo in FeMoco could disfavor an intermediate local spin on Fe6, because it would compete with the spin-coupling to Mo, leading to a reduced binding affinity for Fe6 in E_0 FeMoco. If the spin-coupling between V and Fe6 is weaker in comparison and coincides with that between Fe2 and Fe1, this could also remove the binding site preference in the E_0 FeVco cluster model.

For the E_1 cluster models, binding to Fe6 is energetically favored in both cofactors. For FeMoco this was rationalized with the electronic structure in the substrate-free E_1 model, where the additional electron reduces the Mo cubane (Figure 5.9). However,

applying the same reasoning to FeVco would predict that Fe2 is the preferred binding site, since the additional electron reduces the Fe-only cubane in the substrate-free E_1 FeVco. Therefore, the slight preference for Fe6 in the E_1 FeVco cluster model still lacks an explanation.

CO frequency

Table 5.2 compares the CO activation parameters for FeMoco and FeVco. With Fe2 as the binding site, the calculated CO frequencies in the E_0 cluster models are virtually identical for FeMoco (1895 cm^{-1}) and FeVco (1894 cm^{-1}). With Fe6 as the binding site, CO is less activated for FeMoco (1910 cm^{-1}) compared to FeVco (1887 cm^{-1}). The local electronic structure at the binding site is equivalent in all cases, but the redox level of the respective cubane delivers a ready explanation for the difference in CO activation (Figure 5.9). The Fe centers in the V cubane are reduced by one electron relative to the Fe centers in the Mo cubane and Fe6 in FeVco is therefore part of a more electron rich cubane. The Fe-only cubane has the same number of electrons for both cofactors, which leads to similar CO activation.

The calculated CO frequency in the E_1 FeVco model is about 20 cm^{-1} lower for Fe6 and 50 cm^{-1} lower for Fe2 in comparison to E_1 FeMoco model. The first case is not surprising considering the more reduced V cubane. However, difference for Fe2 is less obvious, since the number of electrons in the Fe-only cubane is the equivalent for both models. However, the Fe2−Fe6 distance is 0.1 Å shorter for the CO-bound E_1 FeVco, which may increase the overlap between both metal centers and therefore the CO activation. The reason for the shorter Fe2−Fe6 distance in FeVco may be related to the less strong V−Fe interaction compared to Mo−Fe.[299]

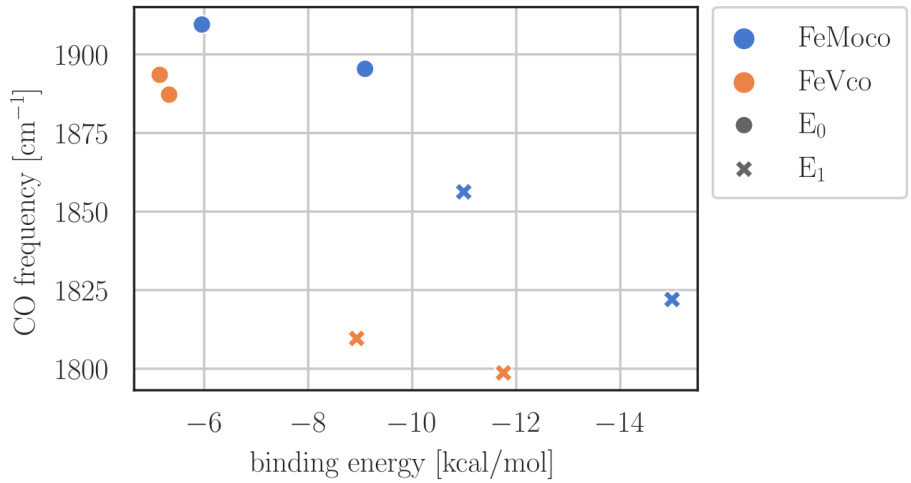

Figure 5.10. – Relationship between scaled CO frequencies and CO binding energies for the FeMoco and FeVco cluster models in the E_0 and the E_1 redox states. See Table 5.2 for values and model details.

The decrease in the CO frequency correlates with a more negative binding energy for both the FeMoco or the FeVco cluster model, as shown in Figure 5.10. FeVco yields lower frequencies at comparable binding energies, which is consistent with the more

reduced Fe centers. However, it is also possible that the carbonate ligand allows for a stronger binding of CO in FeVco. Since the binding sites Fe6 and Fe2 are equally affected, the difference between Mo and V appears to play a minor role in the binding energy.

5.4. QM/MM model: interactions with the protein environment

The following discussion will focus on the role of the protein environment during the CO binding event. The protein environment creates a highly specific electrostatic environment that is key to enzymatic catalysis. Instead of the implicit solvation model used in the clusters models, the amino acid residues are modeled explicitly in the QM/MM models, where they are either part of the QM or the MM region. See Section 4.3.1 for a detailed comparison between the cluster and QM/MM model in the substrate-free cofactor. The QM2 model (Figure 4.1) is used, which was specifically designed to study CO binding and those residues that are expected to affect binding to Fe6 or Fe2 are modeled with high accuracy as part of the QM region. The discussion builds on the analysis in the CO-bound cluster models (Section 5.3) and therefore mainly focuses on the *differences* in geometric and electronic structure introduced by the explicit protein environment in the QM/MM models.

Section 5.4.1 discusses CO binding to Fe6 and Fe2 of the cofactor in the E_0 resting state. Section 5.4.2 focuses on the one-electron reduced and S2B-protonated E_1 redox state. The question will be addressed whether the QM/MM model favors a different binding mode and stabilizes a different electronic structure compared to the cluster model. Section 5.4.3 presents tentative results on potential sulfide dissociation in the even further reduced and protonated E_2 redox state. In this context, the compliance of the E_1 and E_2 models with the available experimentally observed species will be discussed. Finally, Section 5.4.4 compares how the results for V nitrogenase differ from those of Mo nitrogenase.

5.4.1. Binding to the E_0 redox state

Geometry

The structures for CO-bound E_0 FeMoco in the QM/MM model are shown in Figure 5.11a and b. As for the cluster model, CO binding to either Fe6 or Fe2 increases the coordination number of the respective Fe center from four to five, but does not affect the topology of the system otherwise. Also, the perturbation of the cofactor geometry is of comparable magnitude to the cluster model, which comprises an elongation of the bonds involving either Fe6 or Fe2, respectively, leaving the remaining bonds largely unaffected (see Section 5.3.1 for discussion of the cluster model geometry).

The most prominent geometric change happens at the His195 residue, which forms a hydrogen bond with the S2B belt sulfide bridging Fe6 and Fe2. The hydrogen bond remains intact in the CO-bound structures, but changes from 2.25 Å (substrate-free) to 2.20 Å and 2.41 Å when CO is bound to Fe6 and Fe2, respectively. For Fe6, the His195

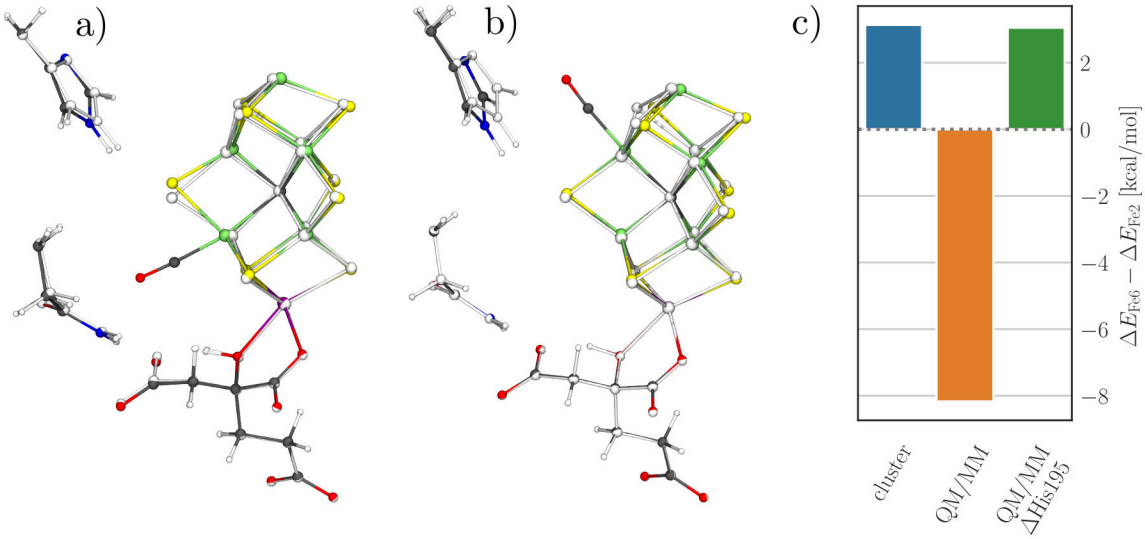

Figure 5.11. – Structures of FeMoco and surrounding residues with CO bound to (a) Fe6 and (b) Fe2 in the E_0 QM2 model. The substrate-free structure is overlaid in black and white. (c) Binding site preference for Fe6 over Fe2. For more details on the cluster and QM2 models, see Table 5.2 and 5.3; the ΔHis195 QM2 model is explained in the text. Figure adapted from Ref [25].

residue appears to simply follow the S2B sulfide, which is affected by the Fe6−S2B bond elongation. For Fe2, the displacement of the His195 residue is significantly more pronounced. The slight elongation of the His195−S2B hydrogen bond suggests a competition between the hydrogen bond and the steric requirement of the CO ligand, which clashes with His195. On the other hand, the positions of the residues Gln441 and HCA remain fairly unperturbed for either binding site. Note that in contrast to the cluster model, where for example the HCA residue was fixed, none of the atoms in proximity of the cofactor are constrained during the geometry optimization of the QM/MM models shown in Figure 5.11a and b.

To assess how the His195 residue affects the binding event, the modified ΔHis195 QM2 model was created from the QM2 model by removing the His195 residue (see Section 5.1 for details). Now, the influence of the His195 residue on the binding site preference can be quantified, which is shown in Figure 5.11c. In the cluster model, binding to Fe2 was favored over Fe6 by 3 kcal/mol. The QM2 model exhibits the opposite preference and the Fe6-bound structure is 8 kcal/mol more stable. Finally, in the ΔHis195 QM2 model Fe2 is again favored by 3 kcal/mol and the binding site preference is therefore virtually identical to the cluster model. This agreement between the cluster and the ΔHis19 QM2 model is a strong indication that Fe2 preference is an intrinsic property of E_0 FeMoco, but the His195 in the MoFe protein environment shields Fe2 by forming a hydrogen bond with the S2B sulfide. Therefore, substrate binding to Fe2 is unlikely, unless larger rearrangements happen in the protein environment.

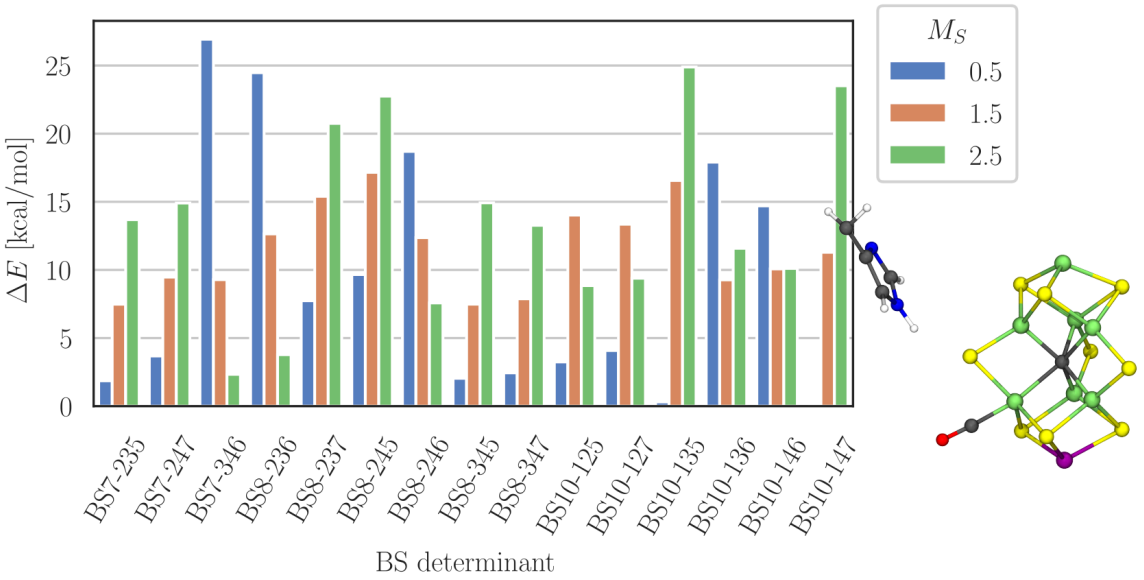

Figure 5.12. – Relative energies for the determinant classes BS7, BS8, and BS10 with CO bound to Fe6. Values correspond to the E_0 FeMoco QM1* model.

Electronic structure

To test if the protein environment has an influence on the relative energies of the BS determinants, multiple BS/M_S combinations were calculated using the QM1* model, in which CO was bound to Fe6, which is shown in Figure 5.12. The QM1* model uses the small QM1 region (Figure 4.1) as well as a small basis set in order to screen many BS/M_S combinations. This is the equivalent approach employed for the cluster* model (Figure 5.6). As in the cluster* model, the BS10-135 and BS10-147 determinants are the most stable. However, in the QM1* model multiple BS8 determinants, such as BS8-345, are within 5 kcal/mol of the most stable determinant, whereas BS8-345 was about 10 kcal/mol higher in the cluster* model. At the same time, the energies of the BS7 determinants are about 5 kcal/mol lower in the QM1* model. The observation that the protein environment stabilizes certain BS couplings may be related to hydrogen bonds between residues and sulfide ligands or to charged residues in the proximity of the cofactor. These have been shown to affect the relative energies of the BS determinants in the substrate-free cofactor (Section 4.3.1). In any case, the explicit protein environment in the QM/MM model increases the number of candidates for low-energy BS determinants when CO is bound to the cofactor.

Turning to the CO-bound QM2 model, which has a larger and more accurate QM region (Figure 4.1), the most stable CO-bound structures are characterized in Table 5.3. The energies for all other tested BS/M_S combinations are listed in Table A15. The BS10-147 determinant is the most stable when CO is bound to Fe6, consistent with the QM1* model, showcasing the usefulness of the rudimentary model to screen the BS/M_S landscape. For Fe2, the BS7-346 determinant is the most stable, but the energy of the BS10-147 determinant is virtually identical. In contrast, the BS10-147 determinant in the cluster model is here about 4 kcal/mol more stable than BS7-346 (Table A13) confirming a strong influence of the protein environment for the QM2

Table 5.3. – Summary of the CO-bound MoFe and VFe QM2 models. The substrate-free QM2 models used as a reference to calculate the binding energies are given in Table A12. The same parameters for the cluster models are shown in Table 5.2.

binding site	S_{pop} Fe6	S_{pop} Fe2	ν_{CO} [cm^{-1}]	$d_{C\text{-}O}$ [Å]	$d_{Fe\text{-}CO}$ [Å]	BS	M_S	ΔE_{bnd} [kcal/mol]
				E_0 MoFe				
Fe6	1.51	3.06	1966	1.152	1.763	BS10-147	0.5	-7.8
Fe2	-2.96	1.81	1957	1.153	1.775	BS7-346	0.5	0.4
				E_1 MoFe				
Fe6	-1.88	3.16	1922	1.157	1.783	BS7-346	2	-8.3
Fe2/Fe6	2.26	2.11	1716	1.193	1.795/ 2.022	BS10-147	0	-5.1
				E_0 VFe				
Fe6	1.71	-3.15	1924	1.160	1.730	BS7-235	0.5	-8.1
Fe2	3.22	-1.73	1957	1.154	1.762	BS7-235	2.5	6.9
				E_1 VFe				
Fe6	1.73	-3.19	1921	1.160	1.747	BS7-235	0	-8.5
Fe2/Fe6	2.50	2.29	1721	1.194	1.828/ 1.969	BS10-147	0	-7.4

model.

As hinted by the magnitude of the Fe spin population (Table 5.3) and confirmed with the localized orbital analysis (not shown), the CO-bound Fe center exhibits an intermediate local spin state. This particular feature of the electronic structure is equivalent to the E_0 cluster model and has already been discussed in Section 5.3.1. To summarize the discussion of the electronic structure in the CO-bound E_0 models, the relative energies of the BS determinants are not easily predicable, since they are defined by a combined contribution of electronic structure changes and environmental influences. However, the changes in the local electronic structure at the binding site follow the same set of rules, which comprises spin-pairing and electron localization. Since the remaining metal centers maintain their local spin state, the M_S value for a given CO-bound BS determinant changes by ± 1 relative to the substrate-free resting state, depending on whether the binding Fe center is flipped.

CO activation

As shown in Table 5.3, the calculated CO frequencies differ slightly for Fe6 (1966 cm^{-1}) and Fe2 (1957 cm^{-1}). The lower CO frequency correlates with a longer C−O bond length, but not with a shorter Fe−CO bond length. The longer Fe2−CO bond length is probably a result of the His195 residue fixing the position of the S2B sulfide in the coordination sphere of Fe2 with the hydrogen bond (Figure 5.11b). The experimentally observed transient SF-FT-IR band loCO$_{IR,1}$ at 1904 cm^{-1} is more than 50 cm^{-1} lower compared to the calculated frequencies for the E_0 QM2 model. This is in line with the

observation that CO does not bind to the resting state protein and further suggests that a more reduced cofactor may be required.

A slightly higher CO frequency for Fe6 compared to Fe2 has also been observed in the cluster model (Table 5.2) consistent with the analogous local electronic structure at the binding site. However, the CO frequencies for both Fe6 and Fe2 are about 50 cm^{-1} higher in the QM2 model. The difference between the QM/MM and the cluster model is likely related to the implicit solvation model in the cluster model and the choice of the dielectric constant. While $\varepsilon = 4$ may describe the average for many proteins well,[317] it may not be suitable for immediate environment of CO bound to the active site of nitrogenase. The immediate protein environment should be therefore described explicitly to capture its highly specific electrostatic.

Binding energies

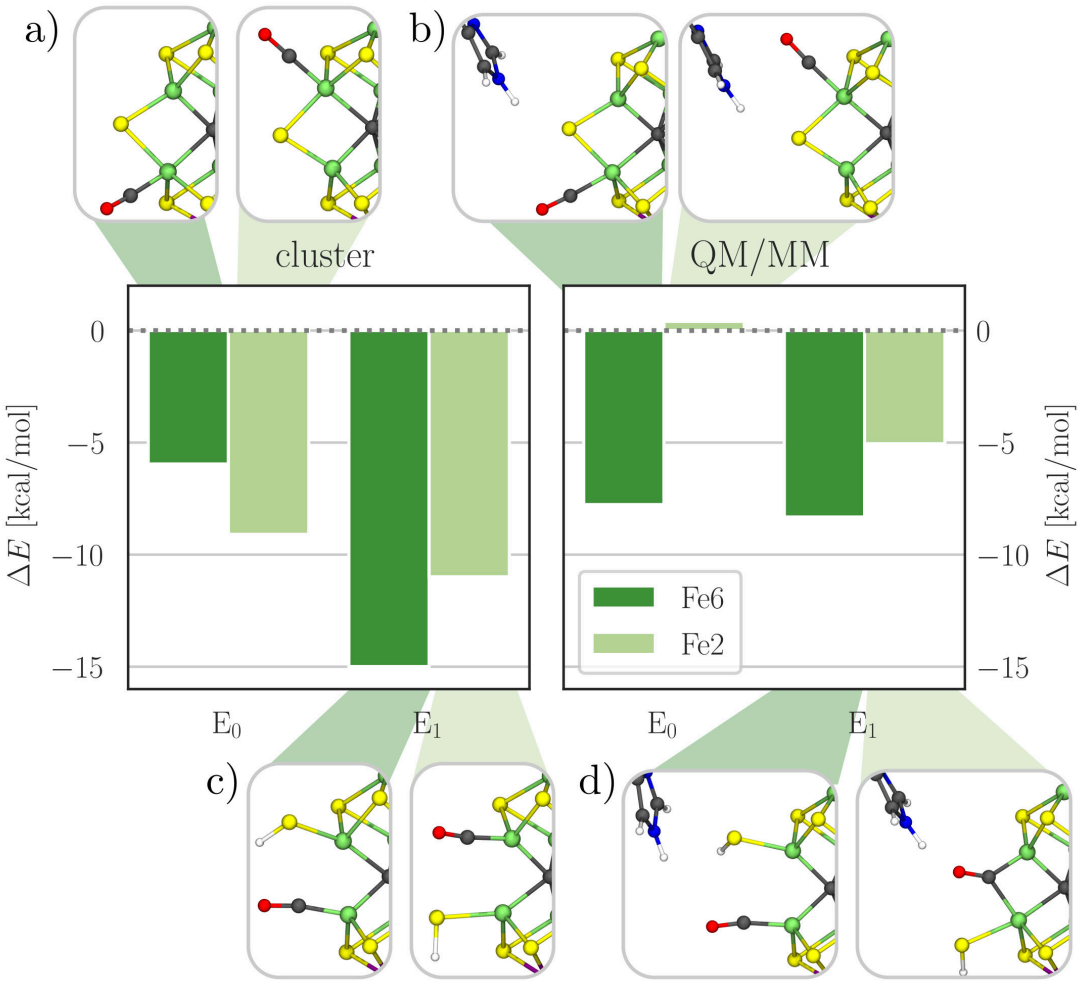

Figure 5.13. – CO binding energies for the FeMoco cluster model (left) and the QM2 model (right) in (a,b) the E_0 and (c,d) the E_1 redox state. Insets show corresponding binding motifs. Further details of the models are given in Table 5.2 and 5.3. Figure adapted from Ref [25].

The CO binding energies for the QM2 model are shown as part of Figure 5.13, which summarizes the binding energies and structures of all E_0 and E_1 models. When focusing on the E_0 QM2 model (Figure 5.13b), the binding site preference for Fe6 is immediately evident, which has been already been shown to be a consequence of the His195 residue. When binding COto Fe6, where the protein environment does not result in any steric hindrance, the binding energy in the QM2 model is with −7.8 kcal/mol about 2 kcal/mol more negative compared to the cluster model (Figure 5.13). Nevertheless, the energy release for the binding event is well below 10 kcal/mol, which is likely insufficient to surmount the reduction in entropy during the associative process.[318] Therefore, the calculated binding energies are consistent with the observation that CO does not interact with the resting state MoFe protein.

5.4.2. Binding to the E_1 redox state

Geometry

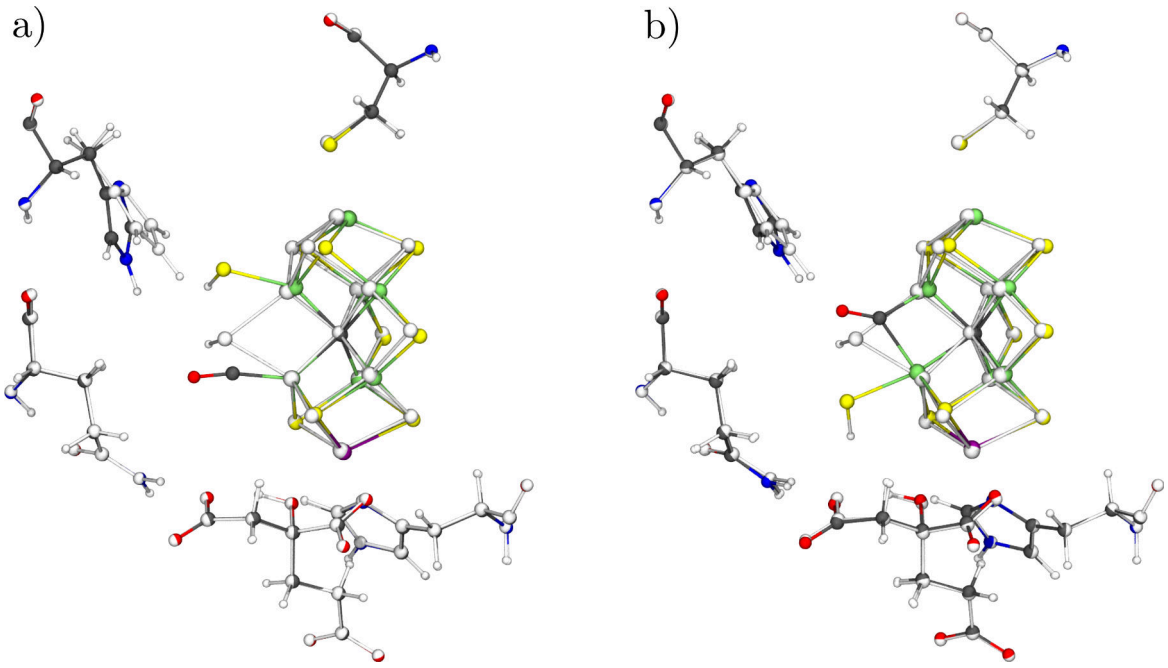

Figure 5.14. – Structures of FeMoco and surrounding residues with CO bound to (a) Fe6 and (b) Fe2 in the E_1 QM2 model. The substrate-free structure is overlaid in black and white. For more details on the models, see Table 5.3.

In the substrate-free E_1 model, the protonated S2B belt sulfide bridges Fe6 and Fe2 and further interacts with the His195 residue via a hydrogen bond. As shown in Figure 5.14, the S2B bridge opens when CO binds to either Fe6 or Fe2 in the QM2 model. For Fe6, the optimized structure features a terminal CO at Fe6 and a terminal SH^- at Fe2. This is the same binding mode observed for the cluster model (Figure 5.5). However, the protein environment in the QM2 model interferes with this structural rearrangement, in particular, the His195 residue clashes with the terminal SH^- on Fe2.

The interference of His195 with a terminal ligand on Fe2 has been shown to greatly increase the energy in the E_0 QM2 model (Figure 5.11).

When binding CO initially to Fe2, the S2B sulfide bridge opens as well, but in the other direction to form a terminal SH^- on Fe6. Simultaneously, CO assumes the bridging position between Fe6 and Fe2. In the slightly asymmetric bridge, the Fe6−CO bond (2.022 Å) is longer than the Fe2−CO bond (1.795 Å). Since no terminal ligand is bound to Fe2, no interference with the His195 residue is observed. On the contrary, His195 now forms a 1.75 Å hydrogen bond with the oxygen of CO, which is shorter than the hydrogen bond with S2B in the substrate-free model (2.30 Å). The bridging CO motif is not stable in the E_1 cluster model and therefore most likely stabilized explicitly by the hydrogen bond with His195. The cluster model suggests that the bridging CO motif is less favorable in the E_1 models compared to an opened bridge with two terminal ligands.

Electronic structure for terminal CO

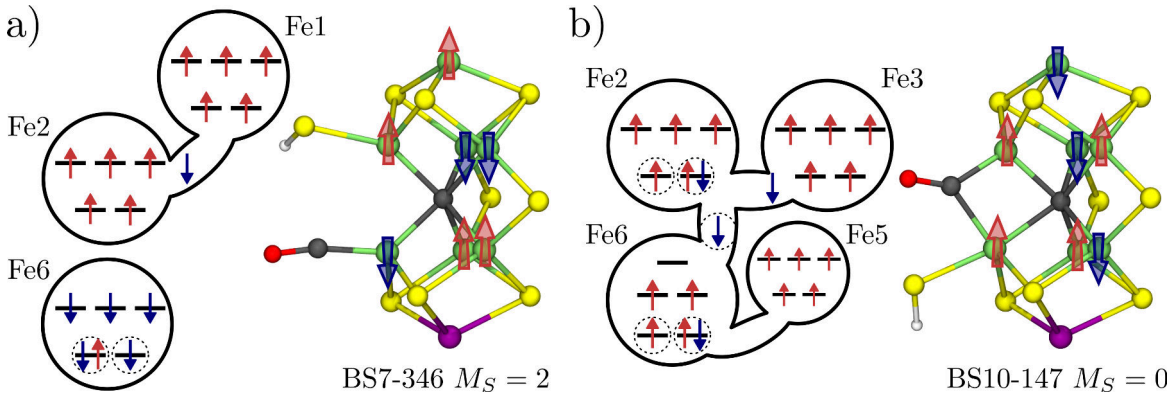

Figure 5.15. – Orbital occupations based on Foster-Boys localized orbitals, structure, and spin-coupling pattern in the BS determinant for (a) CO bound to Fe6 and for (b) CO bridging Fe6 and Fe2 in the E_1 FeMoco QM2 model. Dashed circles indicate orbitals with a large CO overlap. For more details on the models, see Table 5.3. Figure adapted from Ref [25].

Figure 5.15a shows thee spin-coupling between the metal centers as well as the orbital occupation scheme for the E_1 QM2 model with a terminally bound CO. Here, the BS7-346 determinant is lowest in energy, whereas BS10-147 was the most stable in the cluster model. Since the binding mode is equivalent in the QM2 and the cluster model, the change in the BS determinant is likely a consequence of the electrostatic influence of the protein environment. In the E_1 cluster model, the BS7-346 and BS10-147 determinants lie merely 1 kcal/mol apart (Table A14), therefore different descriptions of the environment may be sufficient to swap their relative energies. A stabilization of BS7-346 over BS10-147 has also been observed for the CO-bound E_0 models (Table 5.2 and 5.3). Antiferromagnetic coupling between Fe centers is central to the stability of a given BS determinant, and it has been shown that the explicit protein environment lifts the degeneracy in the calculated J coupling constants of pseudo-symmetry related Fe pairs (Figure 4.6).

As in the E_1 cluster model, some localized orbitals exhibit partial delocalization across multiple metal centers. Therefore, the orbital occupation schemes in Figure 5.15 represent a simplified picture, while the more complete localized orbitals compositions are shown in Figure A21. For the terminally bound CO, the additional electron in the E_1 redox state localizes on Fe6 and therefore participates directly in the π back bond. A key difference between the cluster and the QM2 model is the local spin at the binding site. In the cluster model, Fe6 exhibits a local intermediate spin, where the orbital along the Fe−CO bond is unoccupied. In the QM2 model, however, Fe6 maintains a local high spin, which is the first example of a CO-bound high spin Fe center in the models discussed so far. Consequently, only three Fe-based orbitals participate in the π back bonding. The BS7-346 determinant is the most stable for both the substrate-free E_1 QM2 model and for that with CO bound terminally to Fe6. The $M_S = 2$ is found most stable in both cases and CO binding does not induce any large changes in the electronic structure, even though the geometric changes at Fe6 and Fe2 are quite large.

Electronic structure for bridging CO

As shown in Figure 5.15b, the bridging CO motif in the E_1 QM2 model features the BS10-147 determinant with an $M_S = 0$. Here, Fe6 and Fe2 are ferromagnetically aligned, which is not surprising, since the calculated coupling constant in the diamagnetically substituted μ−CO−[$Fe_2Ga_5InHS_9C$] is $J = +47\,cm^{-1}$ (Table A11). This is the same BS/M_S combination as for the cluster model, even though here the bridge between Fe6 and Fe2 is open.

In the orbital occupation scheme for the bridging CO motif, Fe2 has a local high spin and Fe6 a local intermediate spin. Three spin orbitals localized on each Fe center participate in the π back bond; the respective doubly occupied orbitals contribute to the in-plane interaction (yz in Figure 5.4). Additionally, an orbital delocalized over Fe6 and Fe2 has a large overlap with CO. This orbital is a linear combination of the xy orbitals on Fe6 and Fe2, respectively, and is therefore responsible for out-of-plane interaction.

Even though the CO-bridged Fe6 and Fe2 centers are ferromagnetically aligned in both the QM2 model (Figure 5.14b) and in the diamagnetically substituted [$Fe_2Ga_5InHS_9C$] model (Figure 5.14b and 5.4, respectively), the orbital occupation schemes exhibit some significant differences. The Fe2 center has a local high spin in the QM2 model, but an intermediate spin in the [$Fe_2Ga_5InHS_9C$] model. This is most likely an effect of the unsubstituted Fe centers in the QM2 model, which favor antiferromagnetic coupling and therefore counteract spin-pairing on Fe2. Furthermore, the out-of-plane xy orbital is evenly delocalized over Fe6, Fe2, and CO in the QM2 model, while the orbitals in [$Fe_2Ga_5InHS_9C$] are all fairly localized on either Fe center. This may be a consequence of the local high spin on Fe2 in the QM2 model, which would facilitate electron delocalization through double exchange (see Section 1.1.4). In contrast to the E_0 models, where the changes in the electronic structure due to CO binding are well described by the [$FeGa_6InS_9C$] and the [$Fe_2Ga_5InS_9C$] model, the bridging CO affects more metal centers than just Fe6 and Fe2. In other words, the μ−CO−[$Fe_2Ga_5InHS_9C$] model does not capture the essential aspects of

the μ−CO−FeMoco electronic structure. Therefore, a suitable diamagnetically substituted model would probably have to additionally include at least Fe3, since electron delocalization between Fe2 and Fe3 is apparent in the orbital occupation scheme (Figure 5.15b).

Binding energies

Figure 5.13 compares the CO binding energies for all Mo nitrogenase models discussed so far. The binding energies for the terminal and the bridging motif in the E_1 QM2 model are −8.3 kcal/mol and −5.1 kcal/mol, respectively, which implies a preference for the terminal binding mode. The binding strength for the terminal CO in E_1 is virtually identical to the Fe6-bound E_0 QM2 model. In the cluster model, however, the binding strength significantly increased in going from the E_0 to the E_1 redox state, which has been argued to stem from a more reduced cofactor and the additional flexibility of the protonated S2B sulfide. In the QM2 model, this conformational freedom has disappeared owing to the steric interference of a terminal ligand on Fe2 with the His195 residue. Since CO binding in the E_1 cluster model is clearly favored, the question arises if simple QM/MM potential energy surface explorations possibly fail to capture important conformational rearrangements in the protein environment upon substrate binding. For example, it has been reported for V nitrogenase that the His180(V) residue can switch its hydrogen bonding partner from S2B to Gln176(V) (change from "out" to "in" conformation).[236] However, when this particular conformation was tested for the CO-bound QM2 models, it did not lower the energy of the system.

CO activation

The scaled calculated frequencies for terminal and bridging CO in the E_1 QM2 models are 1922 cm^{-1} and 1716 cm^{-1}, respectively (Table 5.3). The exceptionally low frequency for the bridging CO is consistent with the orbital occupation scheme (Figure 5.15), where seven spin orbitals have a large Fe−CO overlap. The frequency for the terminal CO is about 50 cm^{-1} lower compared to the E_0 QM2 model. This is somewhat surprising, since only three instead of four spin orbitals contribute to the π back bond and the Fe−CO bond is longer. The stronger activation of terminal CO is therefore most likely related to the open S2B bridge, which exposes a more reactive Fe6 center.

The two calculated CO frequencies in the E_1 QM2 model agree very well with the experimentally observed SF-FT-IR bands $loCO_{IR,1}$ and $loCO_{IR,2}$ (1904 cm^{-1} and 1715 , respectively). Therefore, the QM/MM calculations suggest that the initial CO binding may already happen in the E_1 redox state. Also, the calculated CO frequencies in the E_0 QM2 model are about 50 cm^{-1} higher than the experimentally observed ones, further supporting the hypothesis that CO does not interact with resting state Mo nitrogenase. In the E_0 cluster model, on the other hand, the calculated frequencies would agree well with the experimental $loCO_{IR,1}$ band. This agreement could lead one to propose that an E_0 redox state is responsible for the $loCO_{IR,1}$ species. While the calculated CO frequency alone should be taken as a reliable quantity to assign an experimental species, the discrepancies between the cluster and the QM2 model strongly objects the use of a cluster model to study mechanistic details in nitrogenase.

CO activation in the Val70→Ile variant

Substitutions of the Val70 residue (Figure 4.1) have been shown to modulate the frequency of the $\mathrm{loCO_{IR,1}}$ band.[228] Replacement with the more bulky Ile residue shifted the observed frequency from 1906 $\mathrm{cm^{-1}}$ (wild type) to 1895 $\mathrm{cm^{-1}}$, whereas the less bulky Gly residue resulted in a higher frequency of 1911 $\mathrm{cm^{-1}}$. Using the work flow described in Section A1.2, the Val70Ile substitution was carried out in the $\mathrm{E_1}$ QM2 model in which CO was bound terminally to Fe6. Then, the structure was optimized using the same BS/M_S combination as for the wild type model. The calculated CO frequency in the Val2Ile $\mathrm{E_1}$ QM2 model is 1905 $\mathrm{cm^{-1}}$ and therefore 17 $\mathrm{cm^{-1}}$ lower compared to the wild type model. This is in good agreement with the 11 $\mathrm{cm^{-1}}$ decrease in the experimental $\mathrm{loCO_{IR,1}}$ band. This agreement lends strong support to the hypothesis that $\mathrm{loCO_{IR,1}}$ features a terminally bound CO at Fe6.

Proton transfer to the S2B sulfide

The CO-bound X-ray structure of the MoFe protein clearly shows that sulfide dissociation is eventually part of the CO inhibition mechanism ($\mathrm{loCO_{XRD}}$ in Figure 1.9e), though the mechanism for this is unknown. When starting with an elongated Fe−S2B bond ($> 3.3\,\text{Å}$) in the bridging CO $\mathrm{E_1}$ QM2 model, $\mathrm{SH^-}$ does not dissociate, but the structure converges back to the bound structure. While $\mathrm{SH^-}$ dissociation may require an even more reduced and protonated cofactor, it may also be possible that an essential conformational change in the protein environment is not captured by a simple potential energy surface explorations performed in this work.

To explore the conformational space of the protein environment, one can employ QM/MD simulations. However, using DFT for the QM theory is not feasible, since each QM step takes up to 1 h computational time and conformational changes happen on the timescale of about 50 ps. With 1 fs time steps, this would require about 10^4 to 10^5 individual calculations. As an alternative, the semi-empirical method extended tight-binding (xTB) can be used to screen the conformational space of systems with up to 1000 atoms. Here, the GFN2-xTB method[319] was used for the QM part. The QM region is identical to the QM2 region, however, the atoms of the cofactor including the bridging CO and the $\mathrm{SH^-}$ ligand were constrained. The active region was the same as in the QM2 model. The geometric constraints were chosen because the structure of the μ−CO-bound FeMoco is not expected to be stable using the restricted open-shell description of the xTB method. The goal of the xTB/MD simulation is to explore possible responses of the protein environment to the structure of ligand-bound FeMoco. Note that the S2B sulfide is also part of the constraints and sulfide dissociation cannot be captured here.

The results of the 50 ps xTB/MD simulation for the bridging CO structure are shown in Figure 5.16. The time evolution of the xTB/MM energy is stable over the length of the simulation. However, the xTB energy decreases over the first 20 ps, which implies that the MM energy increases similarly. It is well possible that structural change in the QM part would lower the energy of the entire system, given the fairly small active region of about 1000 atoms, which may not sufficiently respond to the changes in xTB part. However, the goal of the xTB/MD simulation is not to accurately describe the

time evolution of the MoFe protein, but to uncover larger structural changes that are connected to CO binding.

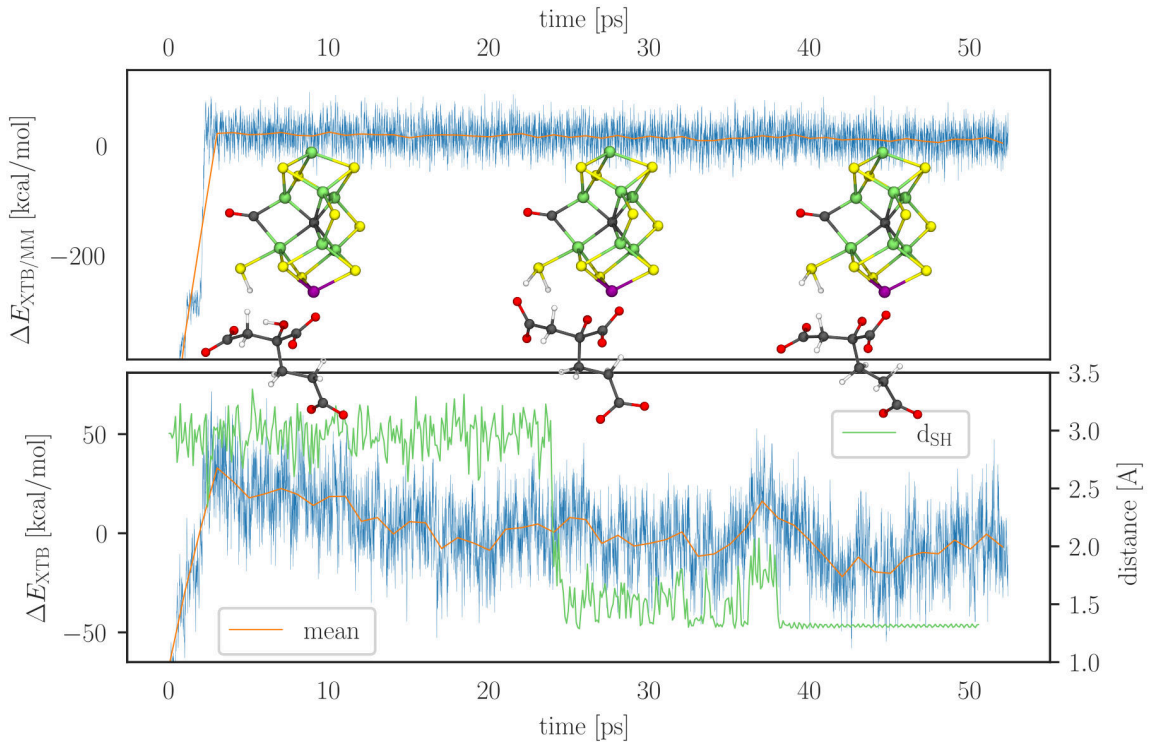

Figure 5.16. – xTB/MD simulation for the bridging CO motif in the E_1 FeMoco QM2 model. The structure of FeMoco, including CO and the terminal SH^- ligand (μ−CO−$[MoFe_7S_9H]^-$), was constrained during the simulation. Top: Time course of the xTB/MM energy (blue) and 1 ps average (orange). Bottom: Time course of the xTB energy (blue) and 1 ps average (orange). The distance d_{SH} between the S2B sulfide and the HCA proton is plotted on the secondary axis (green). Inset structures are for time points 0 ps, 25 ps, and 50 ps, respectively.

As can be seen in the inset structures, a proton is transferred from the HCA residue to S2B after about 25 ps. The proton was formerly bound to the alkoxide of HCA in the vicinity of Mo. The carboxylate pointing in the direction of S2B changes its conformation to transfer the proton. The HCA residue acting as a protonation agent aligns well with a proposed proton channel terminating at HCA.[320] Here, a chain of water molecules can transfer protons from the surface of the MoFe protein to FeMoco via the Grotthuss mechanism.[321] Testing for the proton transfer from HCA to S2B in the E_1 QM2 model increases the energy by 2.3 kcal/mol. Elongating the Fe−S2B bond does not lead to the dissociation of H_2S. Nevertheless, the xTB/MD simulation suggested a mechanistically plausible proton transfer from HCA to S2B and subsequent verification with the QM2 model has shown that this step requires little energy given a terminal SH^- on Fe6. Assuming the uncomplicated supply of protons to HCA via a proton channel, the proton transfer from HCA to S2B may well be a viable mechanistic step during CO binding and sulfide dissociation. However, another electron may also be required.

5.4.3. Binding to the E_2 Redox state

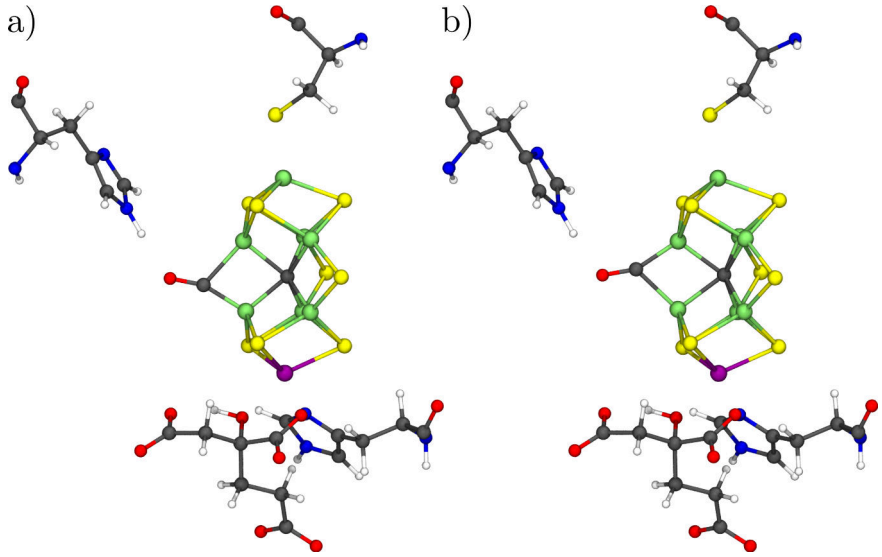

Figure 5.17. – Structures of a semi-bridging CO motif in FeMoco leaning either towards (a) Fe6 or towards (b) Fe2 in the E_2 QM2 model (BS10-147 $M_S = \frac{1}{2}$). Structure (b) is 2.7 kcal/mol higher in energy.

While the redox states of the experimental CO-bound FeMoco species are not known, the X-ray structure clearly shows that CO eventually replaces the S2B belt sulfide (loCO_{XRD} in Figure 1.9e). To describe sulfide dissociation from the cofactor with a computational model would require an extensive exploration of the conformational space of the protein environment, which cannot be achieved within the QM/MM models used here (see discussion of Figure 5.16). Therefore, the assumption for the following model is that the S2B sulfide has dissociated from the cofactor and is stored in some distal protein pocket – it is simply removed from the QM2 model. It is further assumed that dissociation happened in the form of H_2S and that adding another proton to the system is coupled to an additional reduction event. The quantum refinement study of the loCO_{XRD} structure does not support any protonation of the cofactor.[255] The resulting QM2 model therefore corresponds to an E_2 redox state and differs from the E_1 model by an additional electron and a missing H_2S. However, for the E_2 model, only two different BS/M_S combinations were explored and the results should be therefore viewed as preliminary.

The converged CO-bound structures for the E_2 QM2 model are shown in Figure 5.17. CO forms an asymmetric semi-bridge between Fe6 and Fe2, which is either leaning towards Fe6 (a) or Fe2 (b), depending on the starting structure. The loCO_{XRD} structure, on the other hand, features a symmetric CO bridge,[233] but a semi-bridging CO has also been reported for the computational model used for the quantum refinement of loCO_{XRD}.[255] The two structures are merely 2.7 kcal/mol apart and may well be virtually degenerate, given the accuracy of the computational model. This raises the question if the XRD structure, which has been recorded at 100 K,[233] is actually an average of two asymmetric structures. The localized orbitals exhibit strong delocalization over Fe6, Fe2, and CO (Figure A23), making a concise description of the electronic

structure difficult. The sum of the contributions to the atoms of CO, however, is nearly identical to the bridging CO in the E_1 QM2 model (Figure A21b), indicating a similar amount of CO activation. The diamagnetically substituted model corresponding to the CO-bound E_2 structure, where Fe6 and Fe2 are left unsubstituted, predicts a coupling constant of $+18\,cm^{-1}$ (Table A11). This is lower than for the E_1 structure ($+47\,cm^{-1}$), which suggests that ferromagnetic coupling is strengthened by a more symmetric CO bridge. Consistent with this hypothesis, the energies of the two tested determinants BS10-147 and BS7-235, where Fe6 and Fe2 are ferromagnetically and antiferromagnetically aligned, respectively, lie merely 1.2 kcal/mol apart.

For the lowest-energy E_2 model (BS7-235 $M_S = \frac{1}{2}$), the CO bond is 1.199 Å and the hydrogen bond between His195 and CO is 1.88 Å long. The calculated CO frequency is $1718\,cm^{-1}$, which is virtually identical to the bridging CO E_1 model (Table 5.3). The similar CO activation in the E_1 and the E_2 model is consistent with the geometric parameters and the contributions of the localized orbitals. It is important to note that the calculated CO frequencies for both redox states are consistent with the experimentally observed SF-FT-IR species $loCO_{IR,2}$, lending strong support to the proposed bridging motif, but also illustrating that the frequency may not be unambiguously related to the redox state of the cofactor.

5.4.4. Differences between Mo and V nitrogenase

Binding to the E_0 redox state

The CO binding modes in the E_0 redox state are qualitatively identical for the Mo and V nitrogenase according to the QM2 models (Figure 5.11a and b). For the V model, the BS7-235 is the most stable determinant for both Fe6 and Fe2 (see Table A15 for all tested determinants). Since in case of Fe6 the most stable determinant changes from BS8-345 to BS7-235, the VFe protein environment appears to stabilize the BS7 determinants, similar to the MoFe protein. The changes in the local electronic structure upon CO binding, *i.e.* the electron localization and the intermediate local spin at the binding site, are the same as discussed for the Mo QM2 model and are therefore characteristic to all CO-bound E_0 models.

The CO binding energies for Fe6 are virtually identical in the Mo and the V QM2 model (Table 5.3). However, the binding to Fe2 is much more destabilized in the V model: the Fe2 binding energy is 7 kcal/mol more positive compared to the Mo model. Here, the hydrogen bond between His180(V) and S2B is 2.32 Å and therefore identical to the substrate-free structure. Because CO binding to Fe2 resulted in an elongation of the hydrogen bond between His195(Mo) and S2B, the hydrogen bond in the V model appears to be stronger, which would explain the high energy of the Fe2-bound structure.

The calculated CO frequencies for Fe2 are virtually identical for the Mo and the V models ($1957\,cm^{-1}$), which was also the case in the cluster models. Therefore, the significantly higher binding energy for Fe2 in the V model does not affect CO activation. The CO frequency for Fe6, however, is about $40\,cm^{-1}$ lower in the V model, consistent with the more reduced Fe centers in the V cubane (Figure 5.9). If the initial CO binding were to happen in the E_0 redox state, Fe6 is the most likely binding site.

Here, the CO frequency is expected to differ between Mo and V nitrogenase, assuming that the E_0 redox state of the V model is correct.[19,202] In any case, binding to the E_0 redox state is equally unlikely for either Mo or V nitrogenase according to the binding energies in the QM2 models.

Binding to the E_1 redox state

Also, for the E_1 redox state, the binding modes are equivalent for the Mo and the V QM2 model, which comprises a terminal CO bound to Fe2 and a bridging CO between Fe6 and Fe2 (Figure 5.14). Furthermore, the calculated CO frequencies and the C−O bond lengths are virtually identical (Table 5.3). The CO bridge is slightly more symmetric in the V model; the two Fe−CO bond lengths differ by less than 0.1 Å (about 0.2 Å for Mo). The binding energy of the bridging CO is 2 kcal/mol more negative in the V model, making the terminal and the bridging motif nearly degenerate (see Table A16 for all tested BS/M_S combinations).

Even though the binding motifs appear very similar in the Mo and the V model, the electronic structures exhibit some significant differences. For structure with a terminal CO at Fe6, the localized orbitals show that the additional electron in the E_1 redox state localizes in the Fe-only cubane of FeVco (Figure A22), but in the Mo cubane for FeMoco. In the V E_1 cluster model, the additional electron followed the CO into the respective cubane, therefore, it reduced the V cubane, when CO was bound to Fe6. The reduction of the Fe-only cubane in the V QM2 model, even though CO binds to the V cubane, is a direct result of the protein environment. In the E_0 redox state, the Fe centers in the V cubane are already reduced by one electron compared to the Mo cubane (Figure 5.9).

Bulk electrostatic effects in the VFe model

In the original report of the QM/MM model for E_0 VFe protein, Benediktsson and Bjornsson reported on the long-range effect of bulk water.[19] They found that a large number of explicit water molecules around the VFe protein is necessary to stabilize the electronic energies (unbound electron problem). However, the effect on the cofactor geometry was negligible.

When using the same bulk water extension for the QM2 model, the binding energy for CO bound to Fe6 changes from −8.1 kcal/mol to −8.4 kcal/mol. This, admittedly small, difference can be related to the electrostatic influence of the MM part. The electrostatic potential created by the MM charges at the position of the metal nuclei is shown in Figure 5.18a. Note that these values do not represent the actual electrostatic potential, since those residues that are part of the QM region do not hold any MM charges, such as the charged Arg339(V) or Lys83(V) residues (Figure 4.1b). When including the bulk water, the electrostatic potential undergoes a fairly uniform shift of about +0.15 au (Figure 5.18b). The positive shift in electrostatic potential nicely illustrates the stabilization of the electronic energies discussed by Benediktsson and Bjornsson.[19] However, due to a small anisotropy introduced by the bulk water extension, the electrostatic potential at the Fe2 and Fe6 nuclei increases above average (red) while that at the Fe3 and Fe7 nuclei increases below average (blue). Considering

that CO binding causes electron localization with a change in the oxidation from $Fe^{2.5+}Fe^{2.5+}$ to $CO{-}Fe^{2+}Fe^{3+}$, the more positive electrostatic potential on Fe6 yields a ready explanation for the slight increase in binding strength of 0.3 kcal/mol. Since including the bulk water adds about 5×10^5 point charges to the system and therefore significantly increases the computational cost, future models could make use of the approximate treatment of the long-range electrostatic interactions discussed in Ref [16].

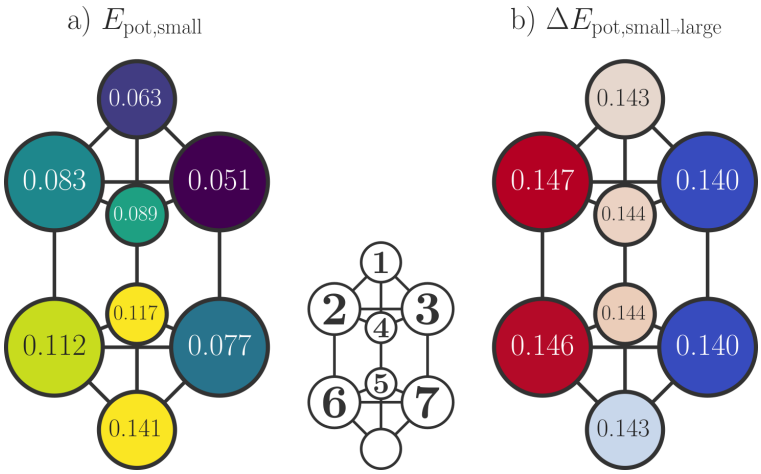

Figure 5.18. – (a) Electrostatic potential created by the MM charges at the positions of the metal centers in FeVco (E_0 QM2 model, BS7-235 $M_S = \frac{3}{2}$). (b) Changes in the electrostatic potential in (a) due to the addition of MM bulk water as reported by Benediktsson and Bjornsson.[19] The electrostatic potential is reported in atomic units, *i.e.* $\frac{e}{a_0}$.

5.5. Conclusion

The aim of this chapter was to analyze in detail how CO binding affects the electronic and geometric structure of the active site of nitrogenase, how the protein environment modulates the binding event, and what protonation and redox state of the protein is required to bind CO. To this end, the local electronic structure at the binding site was studied in a diamagnetically substituted cofactor (Section 5.2). Then, the interplay between the local electronic structure and the metal-metal interactions were analyzed in a cluster model (Section 5.3). Finally, the influence of the protein environment was explored within a QM/MM model (Section 5.4). The main focus was put on the E_0 and E_1 redox states, but the E_2 redox state was also briefly explored. An overview of the experimental species relevant in this context is given in Figure 1.9.

Binding energies

Since CO binding is an associative chemical reaction, the energy release needs to overcompensate the entropy reduction.[†] For an associative gas phase reaction, the entropy contribution is estimated to be on the order of 10 kcal/mol at room temperature.[318] It is not obvious if this approximation applies to enzymatic reactions, where the protein environment may result in significant pre-organization of the substrate and therefore mitigate the entropic contribution.

To bind CO to the cofactor of Mo nitrogenase, turnover conditions are generally required and CO−cofactor interaction is not observed for the resting state.[222,226–228,233] Similarly, for solution-extracted FeMoco reduction to an equivalent of the E_1 redox state is required to bind CO.[237–239] All computational E_0 models presented here converge to stable, CO-bound structures and the energy released is at most about −9 kcal/mol, which therefore appears to be insufficient to overcome the entropy penalty (Figure 5.13). In the E_1 cluster model, the binding energy is with −15 kcal/mol significantly more negative and therefore consistent with the experimental observed binding to a more reduced solution-extracted FeMoco. In the QM/MM model, the binding energy is of similar magnitude for the E_0 and the E_1 redox state, which would indicate equally unlikely binding in either redox state. However, this is likely due to the protein environment in the QM/MM model limiting the structural changes that accompany CO binding. Therefore, the cluster model suggests a significantly increased binding affinity in the E_1 redox state relative to E_0, but the binding energies of the QM/MM model are probably not reliable quantities for the binding affinity.

Binding sites

Fe6 and Fe2 were tested as potential binding sites, since the X-ray structure loCO_{XRD} shows that CO eventually replaces the S2B belt sulfide bridging Fe6 and Fe2.[233] In

[†]The binding free energy ΔG depends on the binding enthalpy ΔH, the binding entropy ΔS and the temperature T:

$$\Delta G = \Delta H - T\Delta S$$

Only if $\Delta G < 0$, *i.e.* an exergonic process, does the reaction happen. Here, the binding enthalpy is approximated by the binding energy ΔE.

the E_0 cluster model, binding to Fe2 is more stable than binding to Fe6, which implies that the Fe-only cubane is more reactive than the Mo cubane (Figure 5.13). In the E_1 cluster model, on the other hand, binding to Fe6 is preferred. This aligns well with the additional electron in the substrate-free E_1 models reducing the Fe centers in the Mo cubane, which, in turn, facilitates CO binding to Fe6. Therefore, the reactivity of the two cubanes depends on the respective redox state.

In the QM/MM model, the protein environment dominates the binding site selectivity. In the E_0 QM/MM model, the His195 residue shields binding to Fe2 and therefore reverses the binding site selectivity relative to the E_0 cluster model (Figure 5.11). The influence of the His195 residue is also evident in the E_1 QM/MM model, where it either raises the energy of the structure with a terminal SH^- at Fe2 or guides CO into a bridging position between Fe6 and Fe2. Unless the protein environment undergoes drastic conformational changes prior to CO binding, the His195 residue therefore prevents binding to Fe2, making Fe6 the most likely center for the initial ligand attack.

CO frequencies

The binding of a single CO molecule to the MoFe protein under turnover conditions has been followed with SF-FT-IR spectroscopy and the conversion of the $loCO_{IR,1}$ band at $1904\,cm^{-1}$ to the $loCO_{IR,2}$ band at $1715\,cm^{-1}$ is interpreted as an initially terminal CO that eventually assumes a bridging binding mode.[226] The calculated frequencies in the E_0 QM/MM model are about $60\,cm^{-1}$ higher than the band observed for $loCO_{IR,1}$ (Table 5.3). The E_1 QM/MM model features both a terminal and a bridging binding CO motif and the calculated frequencies, $1922\,cm^{-1}$ and $1716\,cm^{-1}$, respectively, are in good agreement with $loCO_{IR,1}$ and $loCO_{IR,2}$ bands. Furthermore, the calculated frequency of the semi-bridging CO motif in the E_2 QM/MM model is with $1718\,cm^{-1}$ virtually indistinguishable from the bridging CO in the E_1 model. Therefore, both the E_1 and the E_2 models are viable candidates for the $loCO_{IR,2}$ species. This also implies that the CO frequency may not be sensitive to the redox state of the cofactor.

Substitutions of the Val70 residue have been shown to modulate the frequency of the $loCO_{IR,1}$ band, where the Val70→Ile replacement shifts the band by $-11\,cm^{-1}$.[228] The same substitution in the E_1 QM/MM model produced a comparable shift of $-17\,cm^{-1}$ when a CO was bound to Fe6 (Table A18). This further supports Fe6 being the center of initial, terminal CO binding.

Solution-extracted FeMoco binds CO in a redox state that most likely corresponds to the E_1 state of the MoFe protein.[238] The experimental frequency, $1835\,cm^{-1}$, agrees well with the calculated CO frequencies for the E_1 cluster model, $1822\,cm^{-1}$ and $1856\,cm^{-1}$ for Fe6 and Fe2, respectively. However, the solvent NMF used in the experiment has a dielectric constant of $\varepsilon = 170$,[322] while the cluster model uses $\varepsilon = 4$ to approximate the polarizability of the protein environment.[317] Furthermore, it is not known if the solvated FeMoco is protonated or interacts with another Lewis acid, which is most likely necessary in order for the S2B bridge to open. Therefore, the agreement between the calculated and experimentally measured CO frequency for solution-extracted FeMoco may well be accidental.

Comparing the cluster with the QM/MM model illustrates the significance of an accurate description of the protein environment when calculating the CO frequency.

For the E_0 redox state, the calculated frequencies are around 1900 cm^{-1} in the cluster model and around 1960 cm^{-1} in the QM/MM model. In light of the experimentally observed 1904 cm^{-1} band for $loCO_{IR,1}$, one may come to the conclusion that this species corresponds to an E_0 redox state based on the cluster calculations only. However, when using QM/MM modeling, the agreement with $loCO_{IR,1}$ is better for the E_1 redox state. Therefore, the error introduced by the environmental model can be quite large for the CO frequency, even though the explicit residues in the QM/MM models do not directly interact with CO.

Sulfide dissociation

That the S2B sulfide eventually dissociates from the cofactor, is obvious from the X-ray structures of the CO-bound MoFe and VFe protein.[233,234,236] However, it is not obvious if it happens spontaneously during the binding event or if it is a prerequisite to CO binding. The displacement of the S2B sulfide and potential binding pockets in the protein have been discussed in the context of enzymatic catalysis in Mo nitrogenase[170] and V nitrogenase.[309]

While CO induces the spontaneous opening of the protonated S2B bridge in the E_1 QM/MM model (Figure 5.14), the resulting terminal SH^- does not dissociate from the cofactor, even if the Fe−S bond is forcefully elongated. Some X-ray structures, in which the S2B sulfide has dissociated, show conformational changes involving the His and Gln residue in proximity of Fe6 and Fe2.[236,309] This suggests a flexibility in the protein environment that is insufficiently captured by the simple potential energy surface explorations in this work. Attempts to recreate these conformational changes in the QM/MM model also did not lead to sulfide dissociation.

An xTB/MD simulation of the CO-bound E_1 model structure indicates a highly flexible homocitrate and that a proton transfer from the homocitrate to SH^- can occur (Figure 5.16). The homocitrate acting as a protonation agent is consistent with a proposed proton channel that connects the homocitrate with the surface of the protein.[320] The same proton transfer in the QM/MM model raises the energy by merely 2 kcal/mol. Therefore, the xTB/MD simulation uncovered a step in a potential binding mechanism that would not have been considered using only QM/MM geometry optimizations.

Differences between Mo and V Nitrogenase

V nitrogenase has been reported to bind CO without the need to put the system under enzymatic turnover, merely in the presence of a reductant.[219,220] However, some controversy exists around this observation.[221] Furthermore, the redox state of FeVco in the resting state is not unambiguously defined, which makes V nitrogenase a difficult system to model computationally.[202] In the present work, the V nitrogenase model is adopted from Benediktsson and Bjornsson, which accurately captures the geometry of the VFe X-ray structure assumed to be in the E_0 redox state. Therefore, E_0 FeVco refers to that redox state, which holds the same number of metal-based valence electrons as E_0 FeMoco, *i.e.* 41.[19] According to this redox state assignment, the major difference between FeMoco and FeVco in the resting state is the more reduced Fe part in the V

cubane compared to the Mo cubane (Figure 5.9).[17,19,299]

The E_0 Mo and V QM/MM models do not show significant differences in the CO binding energies (Table 5.3). The experimentally observed binding of CO in V nitrogenase without turnover conditions has been proposed based on the occurrence of an $S = \frac{1}{2}$ EPR signal similar to $loCO_{EPR,1}$ in Mo nitrogenase. If binding to E_0 can be excluded based on the similar binding energies in the Mo and V QM/MM models, the CO-bound FeVco would have to be in an odd-electron redox state with 43 valence electrons in order to give rise to the $S = \frac{1}{2}$ EPR signal. In light of the recent study that suggests even a more oxidized FeVco resting state, *i.e.* with 40 metal-based valence electrons,[202] a highly reduced CO-bound FeVco may not be relevant for enzymatic catalysis and only accessible via electrochemical reduction.

Arguably the most interesting property of V nitrogenase lies in its ability to reduce CO to hydrocarbon chains, demonstrating the ability to perform Fischer-Tropsch-like chemistry,[190,192] whereas CO acts only as an inhibitor to N_2 reduction in wild type Mo nitrogenase.[15] However, this feature cannot be solely attributed to the cofactors, since a similar catalytic potential has been observed in the Val70Ala and Val70Gly substituted Mo nitrogenase.[191] The CO-bound Mo and V QM/MM models in the E_1 redox state do not show any differences that would explain their distinct reactivity. Also, the Val70 residue does not show any strong interaction with CO or with the cofactor. Therefore, CO reduction is likely related to the ability of the system to stabilize crucial intermediates after the formation of a bridging CO structure. The similarities for Mo and V nitrogenase in the experimentally observed CO-bound species, such as the EPR species $loCO_{EPR}$ and $hiCO_{EPR}$[219,220] and the X-ray structures $loCO_{XRD}$ and $hiCO_{XRD}$,[221,236] further support a similar mechanism for the CO binding event.

Proposed binding mechanism

To summarize the results from the QM/MM models, Figure 5.19 proposes a mechanism for the initial binding of CO to the active site in Mo nitrogenase. During experiments, CO was not found to affect the cofactor in the resting state of the MoFe protein, but turnover conditions are required to bind CO.[222,226–228,233] Consistent with this observation, the calculated binding energy in the E_0 model (-7.8 kcal/mol, Table 5.3) is not negative enough to overcome the entropy penalty associated with the binding event, which is estimated around 10 kcal/mol. Furthermore, the calculated CO frequency is with 1966 cm^{-1} more than 60 cm^{-1} higher than the initially observed SF-FT-IR band $loCO_{IR,1}$ at 1904 cm^{-1}.

The further reduced and protonated E_1 model exhibits two distinct binding motifs: (i) a terminal CO at Fe6 in combination with a terminal SH^- at Fe2 and (ii) a bridging CO between Fe6 and Fe2 in combination with a terminal SH^- at Fe6. The calculated CO frequencies, 1922 cm^{-1} and 1716 cm^{-1}, respectively, agree well with the experimentally observed bands for $loCO_{IR,1}$ and $loCO_{IR,2}$ at 1904 cm^{-1} and 1715 cm^{-1}, respectively However, the binding energies are with -8.3 kcal/mol and -5.1 kcal/mol comparable or even less negative than in the E_0 model (Figure 5.13). Because the cluster model suggests a significantly stronger binding for E_1, the reduced binding strength in the E_1 QM/MM model likely stems from an interference with the protein environment. Here, the simple potential energy surface explorations may not consider necessary

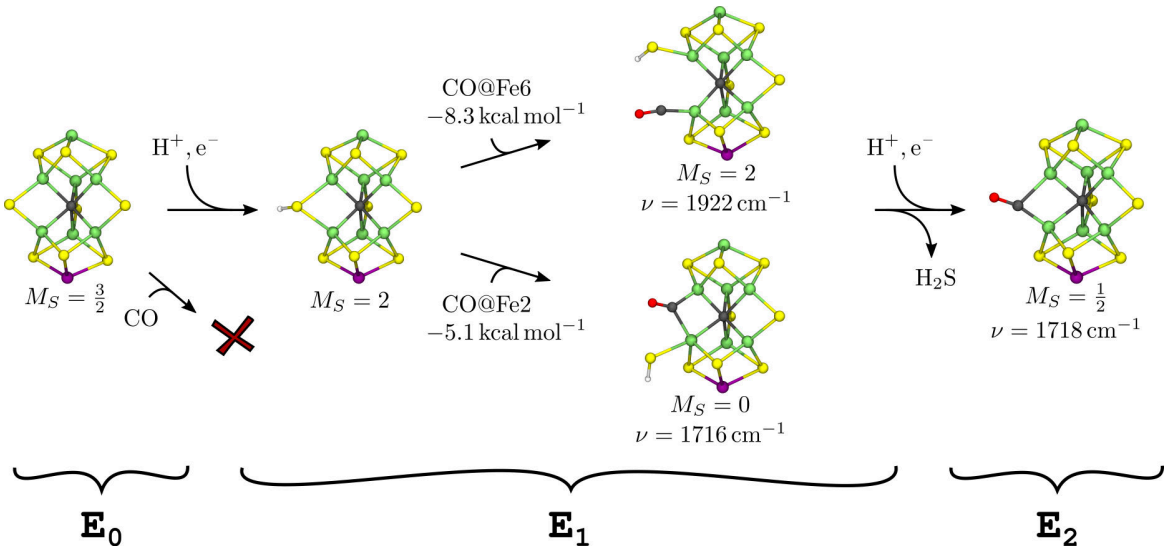

Figure 5.19. – Mechanism for the initial binding of CO to FeMoco based on QM/MM modeling. CO binding to the E_0 redox state is found unlikely in the QM/MM model, consistent with experimental observations. In the E_1 redox state, the CO-bound FeMoco features a terminal CO and a bridging CO binding motif, where the protonated S2B sulfide bridge has spontaneously opened. The calculated CO frequencies agree well with the experimentally observed bands at $1904\,\mathrm{cm^{-1}}$ and $1715\,\mathrm{cm^{-1}}$ ($\mathrm{loCO_{IR,1}}$ and $\mathrm{loCO_{IR,2}}$, Figure 1.9). Alternatively, further protonation and reduction could lead to the dissociation of H_2S in the E_2 redox state, which is observed in the X-ray structure ($\mathrm{loCO_{XRD}}$). The E_2 QM/MM model shows a semi-bridging CO motif. The calculated frequency is virtually identical to that of the E_1 model, therefore consistent with the $\mathrm{loCO_{IR,2}}$ band, and the $M_S = \frac{1}{2}$ agrees with the experimentally observed $S = \frac{1}{2}$ EPR signal ($\mathrm{loCO_{EPR}}$). Figure adapted from Ref [25].

conformational changes in the protein environment that accommodate structural rearrangements at Fe6 and Fe2.

The X-ray structure of CO-inhibited Mo nitrogenase $\mathrm{loCO_{XRD}}$ suggests that the removal of the S2B sulfide is related to the mechanism of CO binding. Since sulfide dissociation is accompanied by conformational changes in the protein environment and therefore cannot be easily described in a QM/MM model, the S2B sulfide was simply removed in order to explore a CO-bound E_2 redox state. The E_2 model has an odd number of electrons, and would be therefore consistent with the $S = \frac{1}{2}$ EPR species $\mathrm{loCO_{EPR}}$. CO forms a semi-bridge between Fe6 and Fe2 in the E_2 QM/MM model, while the $\mathrm{loCO_{XRD}}$ structure exhibits a symmetric bridge. The calculated CO frequency, $1718\,\mathrm{cm^{-1}}$, is virtually identical to the bridging CO in the E_1 model, therefore, the IR species $\mathrm{loCO_{IR,2}}$ may correspond to an E_1 as well as an E_2 redox state. However, the E_2 model was only explored briefly and a more thorough treatment is necessary to corroborate the preliminary findings.

Outlook

The binding of CO to nitrogenase is accompanied initially by the opening and eventually the dissociation of the S2B sulfide bridge. It is unclear whether these structural changes can be reliably captured with the potential energy surface explorations performed in this work. Also, the structure of the QM/MM model is biased towards the substrate-free E_0 X-ray structure.[17] Especially sulfide dissociation is expected to be accompanied by larger rearrangements in the protein environment. These would require molecular dynamics simulations, which are not computationally feasible with the present QM/MM level of theory. Promising developments are related to semi-empirical QM methods, such as the xTB/MD simulation shown in Figure 5.16, or metadynamics.[73] Once the location of the dissociated sulfide has been elucidated, the nudge elastic band (NEB) method could prove useful to compute the dissociation mechanism.[323,324]

The Lowe-Thorneley cycle implies that the next E_n redox state in the catalytic cycle is reached by electron transfer from the Fe protein and a simultaneous proton transfer. (Figure 1.8b). On the atomic scale, however, the supply of protons to the cofactor may be more or less decoupled from the electron transfer, which is rarely rigorously explored in computational models. Understanding the details of proton transport within the MoFe protein is essential for mechanistic studies, because protonation may stabilize reaction intermediates. For example, the proton transfer from the homocitrate to the S2B sulfide suggested by the xTB/MD simulations is a plausible step in the sulfide dissociation mechanism (Figure 5.16). The deprotonation of the homocitrate is limited by a fixed number of protons in the present computational model. However, a proton channel within the MoFe protein has been proposed, that connects the protein surface with the homocitrate,[320] which may enable a constant, low-barrier proton supply to the cofactor. Currently, this can only be accounted for by testing multiple protonation states in the computational models.

IR spectroscopy reports on all CO-bound species, regardless of the E_n state of the cofactor. The EPR spectroscopy selectively reports on odd-electron, and therefore even-numbered, E_n state, but only accounts for a fraction of the species present. Therefore, a combined SF-FT-IR and EPR experimental study would shed more light on the CO binding mechanism. This would reveal, for example, if the $loCO_{IR,1}$ or $loCO_{IR,2}$ bands, which capture the initial CO binding, correlate with the appearance of the $loCO_{EPR,1}$ signal, which arises at the earliest after two reduction events, thus clarifying the E_n state necessary to bind CO. Furthermore, including V nitrogenase in this experimental study would disclose potential differences to Mo nitrogenase prior to CO reduction. For example, the $loCO_{IR,1}$ and $loCO_{IR,2}$ bands, if they are observed, could show a different frequency or time evolution, which could aid the understanding of why V nitrogenase can reduce CO.

A. Appendix

A1. Additional data

A1.1. Reproducibility repository

The original data reported here is published alongside this work in an online reproducibility repository (repro repo), which is linked in Ref [325]. This repository makes the more complete output of the computational models accessible and also all files necessary to set up and rerun the calculations. Therefore, the exact details of the computational models are unambiguously defined. In addition, the code used to generate most plots in this work is also included in the repository. This is done to (i) make the data processing more transparent and (ii) allow the reader to easily access and further explore the raw data.

The repro repo is organized around a `Jupyter` notebook containing `python` code (formerly known as `IPython` notebook).[326] This notebook contains all plots and tables in order of appearance. The respective notebook cell includes (i) the path to the computational model files and (ii) the `python` code to plot figures and LaTeX tables. The data is intermediately stored in `pandas` data frames in the `pickle` format. Therefore, by navigating to the respective notebook cell for a plot or table, the reader gets access to the underlying raw computational model files, the easy-to-handle `pandas` data frames, and the `python` code used to generate the plots. The `python` environment was managed with `Anaconda`,[327] and the package list is stored in `conda_env.yml`. This file can be used to recreate the exact package versions used to process the data in this work (`conda env create -f conda_env.yml`).

A1.2. Standalone programs

Some of the code that has been written along this work has been cast into standalone programs in order to share it with other researches. The programs are located in `./progs/`. Some have further become part of the `orca_helpers` repository.[328] This repository was initiated by Benedikt Floeser and myself in order to openly share and improve programs written for and by the `ORCA` community.

Determine number of roots

The number of eigenstates N that can be formed with n electrons in as many singly occupied orbitals is

$$N = \frac{(2S+1)N!}{(\frac{N}{2}+S+1)!(\frac{N}{2}-S)!}$$

for a given spin S. However, in state-averaged CASSCF calculations one wants to include *all* possible states that can be generated by distributing n electrons over o orbitals, allowing singly, doubly, and unoccupied orbitals. The CASSCF block in `ORCA` expects the number of states (or roots: `nroots`) for each spin (or multiplicity: `mult`). The program `roots.py` outputs the maximum number of roots for each possible spin multiplicity for any given combination of number of electrons (`nel`) and orbitals (`norb`). For a CAS(6,5) calculation, *i.e.* 6 electrons in 5 orbitals, the program prints

```
Ms>=0 nα nβ microstates |     S mult states/roots
    2  5  1           5 |     2    5           5
    1  4  2          50 |     1    3          45
    0  3  3         100 |     0    1          50
total              155 | total              100
```

Here, the left block shows all possible microstates with $M_S \geq 0$; the number of microstates with $M_S < 0$ is equal to those with $M_S > 0$. The right block lists the number of roots for each multiplicity. The procedure to obtain the right from the left block is also outlined in text books explaining term symbols for transition metal complexes.[139,329]

Sort `xyz` files

The program `xyzSorter.py` matches the order of atom coordinates in an `xyz` file to a given template `xyz` file. Both files need to contain molecules with the same chemical formula. The intended use is to unify the ordering for the same molecular fragment from different XRD structures.

Heat map for Loewdin populations

The program `plotLoewdin.py` reads in the data from the `LOEWDIN REDUCED ORBITAL POPULATIONS` block in an `ORCA` output file and prints it as a heat map as shown in Figure A1. The user has a fine control over the selection of atoms, subshells, and orbital ranges. In case of output files from CASSCF calculations, the program suggests orbital rotations, if the active orbitals do not have the maximum weight for a given selection. The code is optimized to handle large output files generated by molecules with several hundred atoms.

Automated `cube` file creation

The program `create_cube.py` is a simple wrapper for the `orca_plot` utility. It creates `cube` files for an arbitrary number of α and β orbitals via a single command.

Extracting AOM parameters from AILFT calculations

The python program `aom.py` contains the code used to obtain the AOM parameters for the $[Fe(XH)_4]$ complexes in Section 2.2. It is limited to complexes similar to $[Fe(XH)_4]$, because all ligand is fitted with the same set of AOM parameters. Therefore, it also

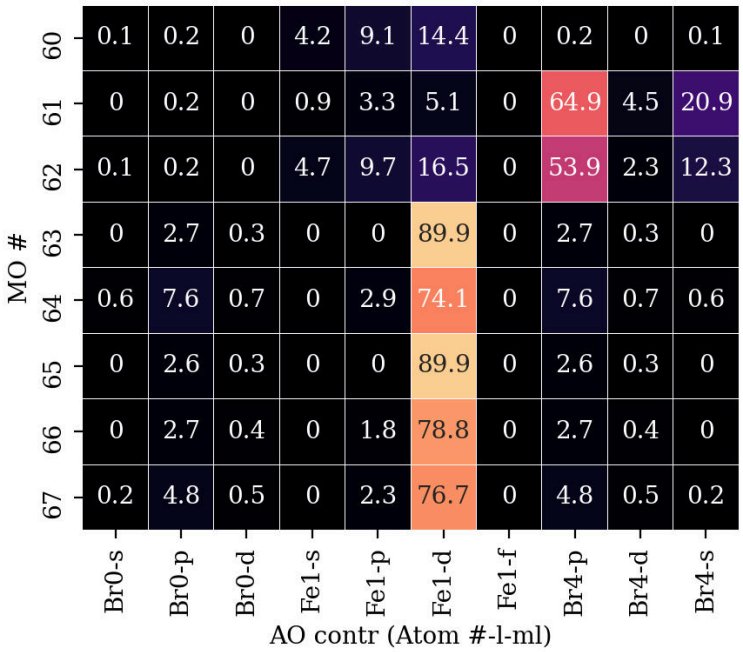

Figure A1. – Plot created with `plotLoewdin.py` for the orbitals in a CAS(5,5) wave function of $FeBr_4^{1-}$. Here, Loewdin atomic populations were contracted over all m_l (*e.g.* p_x, p_y, p_z),

cannot read in AOM parameters for a subset of ligands, which is necessary for the diamagnetically substituted dimers in Section 2.3.2.

The topology of the molecule is hard coded in the `connectivity` list at the beginning of the `aom.py` file. This is necessary to extract the ligand angles, *i.e.* Θ, Φ, and Ψ, from the Cartesian coordinates. The topology definition includes the central atom and the coordinating atom (necessary for Θ and Φ) as well as the atom bound to the coordinating atom (necessary for Ψ). The latter atom still needs to be defined for the current implementation, even when it is not required for the AOM parameters, such as for a monoatomic ligand with only e_σ and e_π parameter. The AOM parameters implement are e_σ, e_π, $e_{\pi s}$, $e_{\pi c}$, $e_{\sigma\pi s}$, and $e_{\sigma\pi c}$. The AOM parameters are defined together with the initial guess value and the limits for the fitting procedure in the `aom_params` list at the beginning of the script. Here, one can also place additional constraints on the AOM parameters (`constraints` list) and select if the CASSCF or NEVPT2 AILFT values are to be used (`correlation` list).

The `aom.py` program takes the output file of an AILFT calculation as an input (`ORCA` version 4.2). The topology and details of the AOM fit are defined in the `USER INPUT` part at the beginning of the script as explained above. The optimized AOM parameters together with the parameter-specific AOM matrices are printed as output.

`vim` syntax highlighting for `ORCA` input files

The folder `.vim` contains files to enable basic syntax highlighting for `ORCA` input files when using the `vim` editor, as shown in Figure A2. Single keywords, blocks, and the

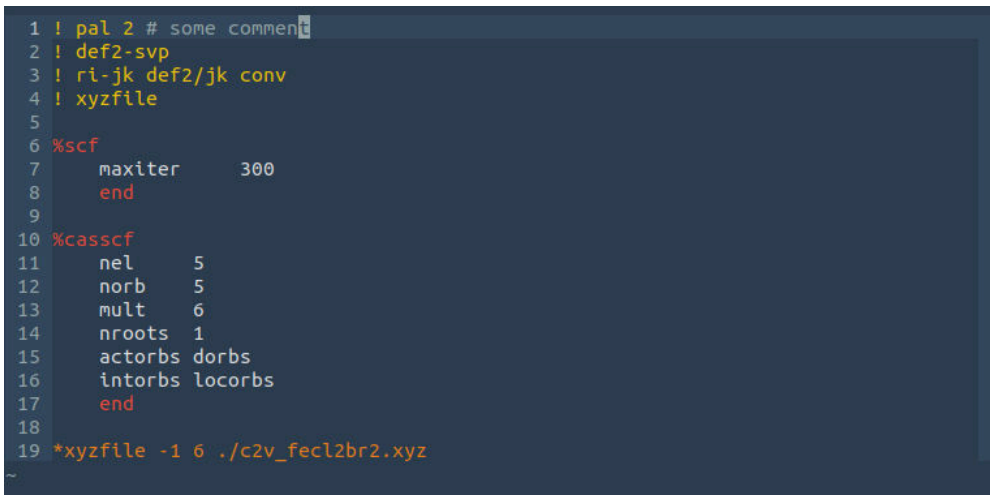

Figure A2. – Syntax highlighting in an `ORCA` input file using the `vim` editor.

coordinate section are visually distinguished, and the colors may make the input files more fun to work with.

Fit parameters for calculated Mössbauer parameters

In order to calculate the ^{57}Fe Mössbauer isomer shift, one typically correlates the calculated electron density at the Fe nuclei to the experimentally determined isomer shifts in a benchmark set of molecules.[330] This linear fit is specific to a particular level of theory, such as basis set or DFT functional. The folder `autoFitMoessbauer` contains a `bash`/`python` workflow that facilitates the generation of `ORCA` input files required for the benchmark calculations, if one desires to obtain fit parameters for a different level of theory. An example for the linear fit is shown in Figure A4. The code is also made available online in the `autoFitMoessbauer` repository.

Residue substitutions in QM/MM models

The folder `regenerate_psf` contains a workflow written in `python`, `bash`, and `tcl` that allows one to easily introduce residue substitutions to an existing `chemshell` QM/MM model. Starting with an existing QM/MM model, the workflow generates all necessary files to run the `chemshell` calculation once the desired residues have been substituted. The code is also made available online here.

Deleting MM charges

The `python` program `deleteMMcharges.py` sets the charges in the `save-new.chm` file to 0 for an arbitrary number of atoms. The code is also made available online in the `chemshell-QMMM-protein-setup` repository.[303]

Localized orbital analysis with `Multiwfn`

The folder `MultiWFN_analysis` contains an automated workflow to perform the localized orbital analysis for large systems. It includes an orbital localization step via the `orca_loc` utility and a subsequent analysis with `Multiwfn`.[304] The workflow was written for nitrogenase QM/MM models, where it facilitates the metal-based electronic structure analysis of BS-DFT determinants, but the code works for any `ORCA` calculation (modify `orca2name` dictionary in `parse_multiwfn.py`). The localized orbitals are summarized in a heat map similar to Figure A14. The atomic charges as well as localization and delocalization indices are printed as well for the selected atoms. The code is also made available online in the `multiWFN_analysis` directory of the `chemshell-QMMM-protein-setup` repository.[303]

A2. S→Se substitution in synthetic [Fe_2S_2] clusters

A2.1. Monomeric FeS complexes: [$Fe(XH)_4$] (X=S,Se)

Table A1. – AILFT parameters for the free ions Fe^{3+} and Fe^{2+} extracted from CAS(5,5) and CAS(6,5) calculations, respectively. All values in [cm^{-1}].

	CASSCF		
	B	C	ζ
Fe^{3+}	1302	4860	472
Fe^{2+}	1184	4373	416
	NEVPT2		
Fe^{3+}	1136	4432	–
Fe^{2+}	1020	3965	–

A2.2. Dimeric FeS clusters: $[Fe_2X_2]$ (X=S,Se)

Table A2. – Loewdin atomic contributions to the active orbitals for the CAS(5,5) and CAS(6,5) wave functions of $[FeZnX_2]^{1+}$ and $[FeGaX_2]^{1+}$, respectively (averaged over all states, see Figure 2.12 for structures). The active orbitals diagonalize the Fock matrix (`orca_blockf` utility).

	Fe contribution [%]				
	z^2	x^2-y^2	xy	xz	yz
$[FeZnS_2]^{1+}$	97.0	94.6	90.9	89.3	85.2
$[FeZnSSe]^{1+}$	96.8	94.8	90.8	89.1	85.6
$[FeZnSe_2]^{1+}$	97.2	95.1	91.1	89.1	85.2
$[FeGaS_2]^{1+}$	97.6	97.4	95.6	93.7	93.2
$[FeGaSSe]^{1+}$	97.7	97.4	95.8	94.2	92.8
$[FeGaSe_2]^{1+}$	97.5	97.5	96.1	93.7	93.4
	X^{2-} contribution [%]				
	z^2	x^2-y^2	xy	xz	yz
$[FeZnS_2]^{1+}$	1.3	4.1	6.5	3.1	11.5
$[FeZnSSe]^{1+}$	1.4	4.0	6.7	3.4	11.3
$[FeZnSe_2]^{1+}$	1.3	3.7	6.3	2.9	11.4
$[FeGaS_2]^{1+}$	0.6	1.0	2.8	3.5	2.1
$[FeGaSSe]^{1+}$	0.7	0.9	2.7	3.0	2.3
$[FeGaSe_2]^{1+}$	1.2	0.5	2.2	4.0	1.5
	$NacNac^-$ contribution [%]				
	z^2	x^2-y^2	xy	xz	yz
$[FeZnS_2]^{1+}$	0.7	0.4	1.4	6.3	2.2
$[FeZnSSe]^{1+}$	0.9	0.6	1.5	5.4	2.1
$[FeZnSe_2]^{1+}$	0.7	0.4	1.4	6.6	2.2
$[FeGaS_2]^{1+}$	0.5	0.4	0.4	1.5	2.6
$[FeGaSSe]^{1+}$	0.4	0.5	0.6	1.2	2.5
$[FeGaSe_2]^{1+}$	0.8	0.6	0.4	1.1	3.2

Table A3. – Geometry of the BS-DFT-optimized $[Fe_2X_2]^{2+,1+}$ clusters (see Figure 2.10). All distances are reported in Å and all angles in °. For $[Fe_2SSe]$, X1 corresponds to S, X2 to Se. See Section 2.1 for computational details.

	$[Fe_2X_2]^{2+}$ $M_S = 0$										
	Fe−Fe	Fe_A−X1	Fe_A−X2	Fe_B−X1	Fe_B−X2	Fe_A−N	Fe_B−N	∠XFe_AX	∠XFe_BX	∠NFe_AN	∠NFe_BN
$[Fe_2S_2]^{2+}$	2.821	2.213	2.219	2.219	2.213	2.901	2.901	100.9	100.9	92.4	92.4
$[Fe_2SSe]^{2+}$	2.880	2.215	2.337	2.212	2.348	2.939	2.939	101.6	101.3	92.8	92.2
$[Fe_2Se_2]^{2+}$	2.932	2.333	2.345	2.341	2.333	2.971	2.971	102.3	102.4	93.1	93.1
$[Fe_2Te_2]^{2+}$	3.121	2.540	2.548	2.549	2.539	3.362	3.361	104.4	104.3	94.2	94.2
	$[Fe_2X_2]^{1+}$ $M_S = \frac{1}{2}$										
	Fe−Fe	Fe_A−X1	Fe_A−X2	Fe_B−X1	Fe_B−X2	Fe_A−N	Fe_B−N	∠XFe_AX	∠XFe_BX	∠NFe_AN	∠NFe_BN
$[Fe_2S_2]^{1+}$	2.875	2.299	2.302	2.202	2.207	2.978	2.950	97.8	103.7	89.0	90.9
$[Fe_2SSe]^{1+}$	2.933	2.298	2.438	2.188	2.340	3.013	2.983	98.2	104.5	89.4	92.0
$[Fe_2Se_2]^{1+}$	3.042	2.451	2.443	2.325	2.336	3.066	3.028	97.5	104.3	90.8	91.6
$[Fe_2Te_2]^{1+}$	3.245	2.649	2.650	2.531	2.544	3.456	3.457	99.5	105.7	92.0	92.7

AOM parameters for the NacNac$^-$ ligand

The extraction of the AOM parameters for the diamagnetically substituted complexes $[FeZnX_2]^{1+}$ and $[FeGaX_2]^{1+}$ in Section 2.3.2 requires two steps. First, parameters for the capping NacNac$^-$ ligands are determined. These parameters enter the parameter extraction of the bridging X^{2-} ligands as fixed values. The two-step procedure is necessary, since fitting the AILFT matrix with the six AOM parameters (e_σ, $e_{\pi s}$, and $e_{\pi c}$ for both X^{2-} and NacNac$^-$) is an underdetermined optimization problem in the approximate C_{2v} symmetry of the Fe center.

To extract the NacNac$^-$ AOM parameters, the complexes $[Fe(NacNac)_2]^{1+,0}$ are constructed from the Fe(NacNac) fragment of the $[Fe_2X_2]^{1+}$ complexes. The NacNac$^-$ ligand is duplicated to yield $[Fe(NacNac)_2]$, as shown in Figure A3. The AOM parameters are extracted for each Fe center in $[Fe_2X_2]^{1+}$ individually in order to account for the slightly different geometric parameters. The resulting parameters are listed in Table A4. The NacNac$^-$ ligands are stronger donors than the X^{2-} ligands, which is consistent with the yz orbital (lobes pointing at NacNac$^-$ ligand) being higher in energy than the xz orbital (lobes pointing at X^{2-} ligands).

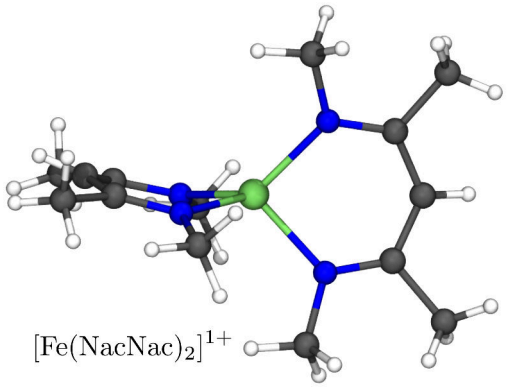

Figure A3. – Structure of the $[Fe(NacNac)_2]^{1+}$ complex constructed from the Fe–NacNac fragment in $[FeZnS_2]^{1+}$. The structures are similar for all six $[FeZnX_2]^{1+}$ and $[FeGaX_2]^{1+}$ combinations.

Table A4. – AOM parameters for the NacNac$^-$ ligand in $[Fe(NacNac)_2]^{1+,0}$ extracted via AILFT (CAS(5,5) or CAS(6,5) including NEVPT2). Structures were created from the Fe–NacNac fragment in $[FeZnX_2]^{a}$nd $[FeGaX_2]^{(}$Figure A3). All values in $[cm^{-1}]$.

parent structure	e_σ	$e_{\pi s}$	$e_{\pi c}$
$[FeZnS_2]^{1+}$	10681	4514	2289
$[FeZnSSe]^{1+}$	9603	3833	1606
$[FeZnSe_2]^{1+}$	10810	4589	2320
$[FeGaS_2]^{1+}$	6835	2712	1404
$[FeGaSSe]^{1+}$	6533	2541	1189
$[FeGaSe_2]^{1+}$	7502	3126	1713

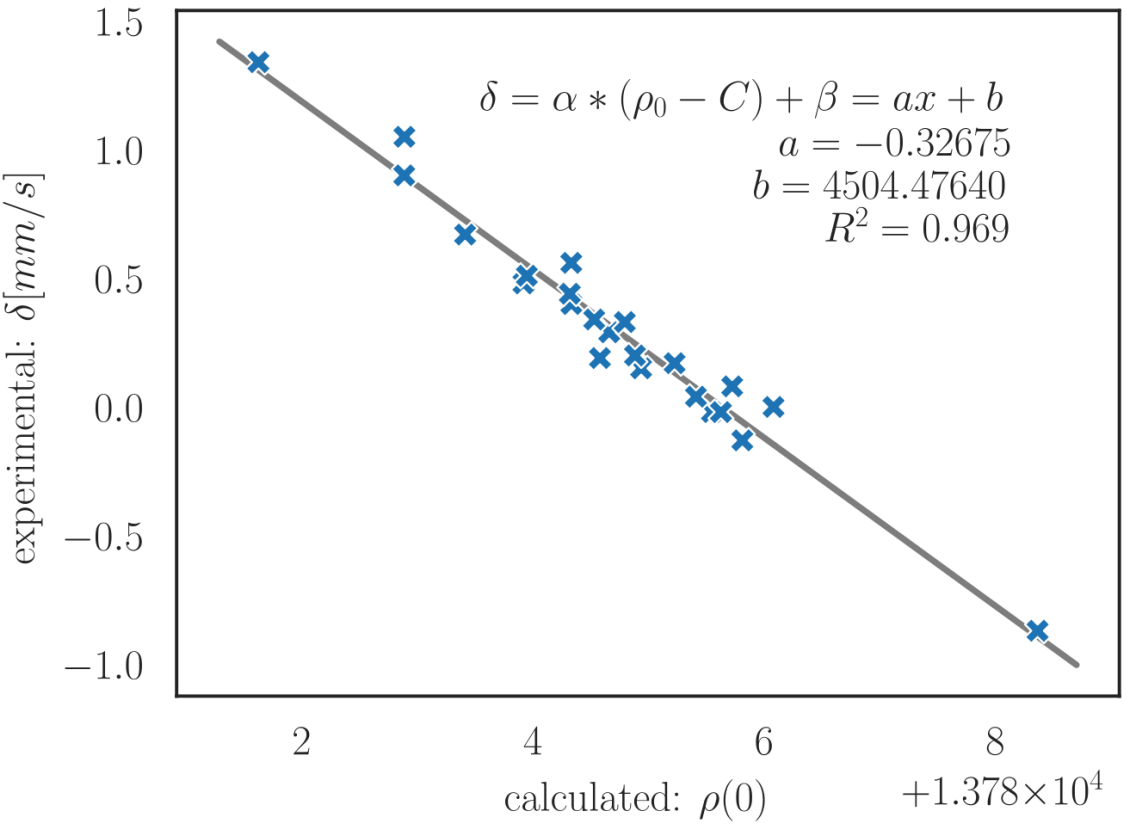

Figure A4. – Calibration curve for calculated Mössbauer isomer shifts against calculated values using the TPSSh functional. Details on the level of theory are given in Section 2.1 for the geometry optimizations of [Fe_2X_2] with BS-DFT. The set of geometries and experimental values were taken from Ref [330]. The code to automate the generation of fit curves for other levels of theory is published along with this work (see Section A1.2).

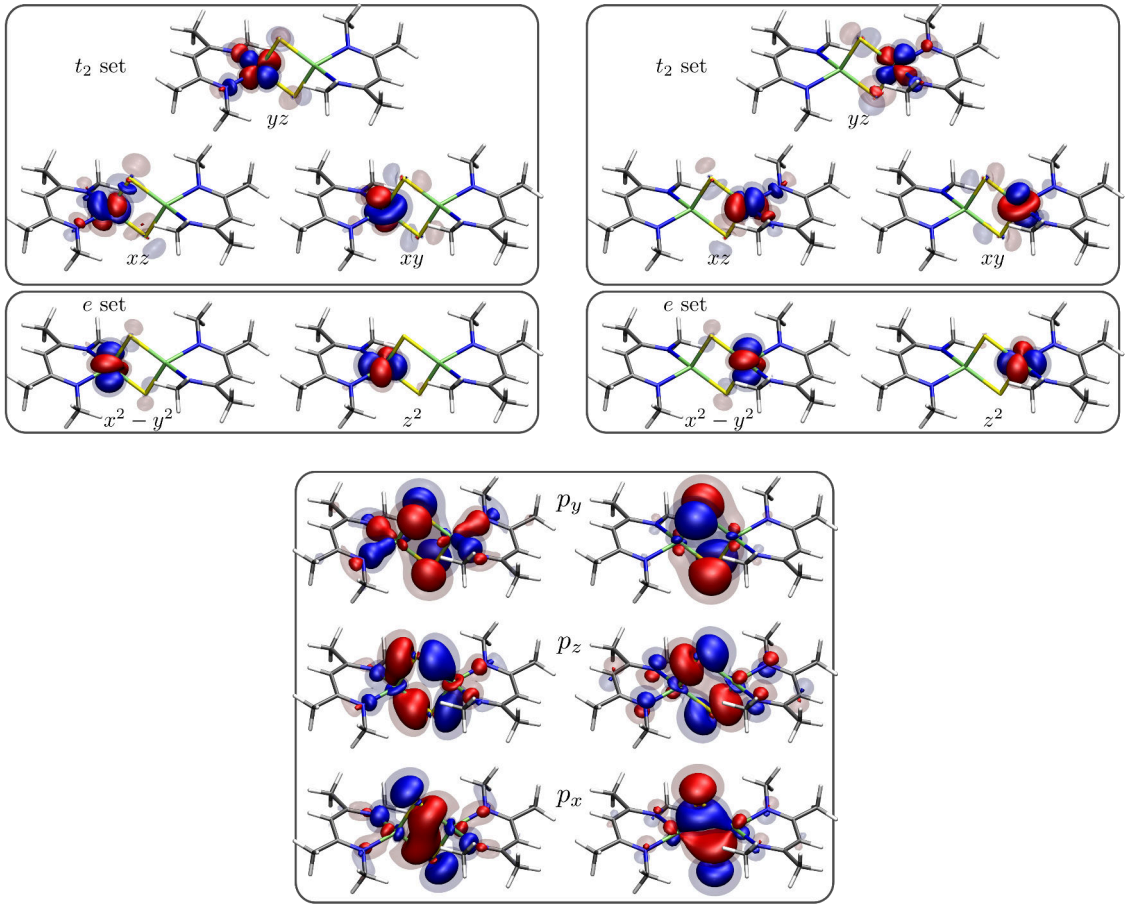

Figure A5. – Plots of the active orbitals in the $S = 0$ CAS(22,16) wave function for the NacNac-capped $[Fe_2S_2]^{2+}$ clusters (see Figure 2.15). Orbital shapes are similar for $[Fe_2SSe]^{2+}$ and $[Fe_2Se_2]^{2+}$. Isosurfaces are shown at ± 0.05 and ± 0.025.

Figure A6. – Contributions of CFGs to the $S = 0$ and $S = 5$ CAS(22,16) wave function for the NacNac-capped $[Fe_2X_2]^{2+}$ clusters (see also Figure 2.15). All values in [%]. The CASSCF wave function optimized for each spin state separately.

	$S = 0$			
	neutral	LMCT	MMCT	non-Hund
$[Fe_2S_2]^{2+}$	96.14	1.22	2.33	0.16
$[Fe_2SSe]^{2+}$	96.19	1.24	2.24	0.14
$[Fe_2Se_2]^{2+}$	96.32	1.26	2.13	0.13
	$S = 5$			
	neutral	LMCT	MMCT	non-Hund
$[Fe_2S_2]^{2+}$	99.23	0.72	0.00	0.00
$[Fe_2SSe]^{2+}$	99.22	0.74	0.00	0.00
$[Fe_2Se_2]^{2+}$	99.20	0.75	0.00	0.00

Table A5. – Loewdin atomic contributions to the active orbitals in the $S = 0$ CAS(22,16) wave function for the NacNac-capped $[Fe_2S_2]^{2+}$ clusters (see Figure A5 for plots). The first six rows correspond to orbitals based manly on the bridging ligands X, the next five on Fe_A, and the last five on Fe_B. The CASSCF wave function optimized just for the $S = 0$ spin state.

	p_x	p_{z^*}	p_{y^*}	p_y	p_z	p_{x^*}	z^2	x^2-y^2	xy	xz	yz	z^2	x^2-y^2	xy	xz	yz
									Fe_A [%]							
$[Fe_2S_2]^{2+}$	13.2	10.8	11.2	12.8	16.0	14.4	97.0	95.8	92.4	87.8	89.7	0.2	0.3	0.9	0.2	0.3
$[Fe_2SSe]^{2+}$	14.1	10.5	11.0	12.9	15.9	15.1	96.3	96.5	92.5	88.1	89.5	0.1	0.3	0.9	0.3	0.1
$[Fe_2Se_2]^{2+}$	13.9	10.2	11.2	12.6	16.5	15.2	96.8	96.1	92.5	87.8	90.0	0.2	0.3	1.0	0.4	0.2
									Fe_B [%]							
$[Fe_2S_2]^{2+}$	13.2	10.8	11.2	12.8	16.0	14.4	0.2	0.3	0.9	0.4	0.2	97.0	95.6	92.5	89.3	88.4
$[Fe_2SSe]^{2+}$	13.7	10.9	11.3	12.5	16.3	14.5	0.2	0.1	0.9	0.4	0.2	97.3	95.3	92.4	88.0	89.6
$[Fe_2Se_2]^{2+}$	13.8	10.5	11.0	12.3	16.6	14.9	0.1	0.2	1.0	0.4	0.2	97.0	95.8	92.5	87.8	89.9
									X_2 [%]							
$[Fe_2S_2]^{2+}$	58.0	56.0	74.8	44.2	57.2	64.2	1.2	2.8	4.4	7.7	3.9	1.1	2.9	4.4	5.1	6.8
$[Fe_2SSe]^{2+}$	65.3	55.2	73.9	44.8	56.9	63.1	2.1	1.7	4.2	7.6	4.0	1.2	2.7	4.4	7.3	4.0
$[Fe_2Se_2]^{2+}$	64.2	53.8	74.3	43.3	56.2	62.0	1.6	2.2	4.2	7.9	3.2	1.2	2.6	4.2	8.2	3.3
									NacNac [%]							
$[Fe_2S_2]^{2+}$	14.0	21.4	1.2	25.2	9.0	4.6	0.6	0.2	1.1	2.1	4.2	0.5	0.2	1.1	4.1	3.3
$[Fe_2SSe]^{2+}$	5.2	22.0	2.4	26.4	8.7	5.0	0.3	0.5	1.2	2.1	4.2	0.6	1.0	1.1	2.8	4.4
$[Fe_2Se_2]^{2+}$	6.6	24.0	1.7	27.3	9.1	5.4	0.4	0.4	1.1	2.0	4.7	0.5	0.3	1.2	2.4	5.1

Table A6. – Loewdin atomic contributions to the active orbitals in the $S = 5$ CAS(22,16) wave function for the NacNac-capped $[Fe_2S_2]^{2+}$ clusters (see Figure A5 for plots). The first six rows correspond to orbitals based manly on the bridging ligands X, the next five on Fe_A, and the last five on Fe_B. The CASSCF wave function optimized just for the $S = 5$ spin state.

	p_x	p_{z^*}	p_{y^*}	p_y	p_z	p_{x^*}	z^2	x^2-y^2	xy	xz	yz	z^2	x^2-y^2	xy	xz	yz
									Fe_A [%]							
$[Fe_2S_2]^{2+}$	13.5	11.2	11.0	12.8	15.0	13.8	97.0	95.9	93.6	89.2	89.2	0.2	0.3	0.7	0.1	0.1
$[Fe_2SSe]^{2+}$	14.4	10.5	10.9	13.0	15.1	14.6	96.7	96.4	93.6	89.1	89.3	0.1	0.2	0.7	0.2	0.1
$[Fe_2Se_2]^{2+}$	14.5	9.3	10.9	13.4	15.8	16.2	97.0	96.2	93.5	88.5	89.8	0.2	0.2	0.7	0.2	0.2
									Fe_B [%]							
$[Fe_2S_2]^{2+}$	13.5	11.2	11.0	12.8	14.9	13.8	0.2	0.3	0.7	0.1	0.1	96.9	95.9	93.6	89.2	89.1
$[Fe_2SSe]^{2+}$	13.8	11.6	11.0	12.4	15.0	13.6	0.1	0.2	0.7	0.2	0.1	97.3	96.0	93.6	88.4	90.1
$[Fe_2Se_2]^{2+}$	13.8	12.8	10.7	11.9	14.7	12.7	0.2	0.2	0.7	0.2	0.0	96.8	96.1	93.6	88.6	90.1
									X_2 [%]							
$[Fe_2S_2]^{2+}$	57.1	53.6	74.4	42.9	60.2	63.6	1.2	2.6	3.7	5.8	5.3	1.2	2.6	3.7	5.9	5.5
$[Fe_2SSe]^{2+}$	64.8	52.9	74.3	42.9	59.9	62.4	1.6	2.0	3.5	6.6	4.1	1.3	2.6	3.7	7.7	3.1
$[Fe_2Se_2]^{2+}$	63.6	51.7	73.8	41.8	59.4	60.8	1.4	2.3	3.7	7.3	3.3	1.2	2.3	3.7	7.8	3.1
									NacNac [%]							
$[Fe_2S_2]^{2+}$	13.8	22.6	1.6	26.9	8.0	6.4	0.6	0.2	1.0	3.3	3.5	0.5	0.2	1.0	3.7	3.8
$[Fe_2SSe]^{2+}$	5.5	23.3	2.5	28.8	8.0	6.9	0.4	0.3	1.1	2.4	4.2	0.6	0.5	1.0	2.3	4.8
$[Fe_2Se_2]^{2+}$	6.5	24.8	2.4	28.1	8.0	7.9	0.8	0.5	1.0	2.0	4.7	0.4	0.4	1.1	2.3	5.2

Table A7. – CASSCF and NEVPT2 energies for the $S = 0, \ldots 5$ states in the NacNac-capped $[Fe_2X_2]^{2+}$ clusters relative to the respective $S = 0$ all-states averaged energy (see also Figure 2.19) and parameters extracted via spin Hamiltonians in Eq 1.8 and Eq 1.10 including the standard deviation (SD). The CASSCF wave function was either averaged over all $S = 0, \ldots 5$ states (state average) or optimized for each state separately (state specific). All values in [cm^{-1}].

$S =$	0	1	2	3	4	5	J	SD_E	J_{biq}	λ	SD_E
						CAS(10,10)					
						state average					
$[Fe_2S_2]^{2+}$	0	77	232	463	770	1152	-38.3	1.4	-38.3	0.01	0.1
$[Fe_2SSe]^{2+}$	0	71	214	428	711	1063	-35.4	1.2	-35.4	0.01	0.1
$[Fe_2Se_2]^{2+}$	0	67	202	403	670	1002	-33.4	0.9	-33.4	0.01	0.0
						state specific					
$[Fe_2S_2]^{2+}$	-48	45	221	463	748	1051	-34.8	33.6	-35.6	0.32	1.7
$[Fe_2SSe]^{2+}$	-44	42	204	428	691	971	-32.1	30.8	-32.9	0.29	1.6
$[Fe_2Se_2]^{2+}$	-41	40	192	403	651	916	-30.3	28.6	-31.0	0.27	1.4
						CAS(10,10)+NEVPT2					
						state average					
$[Fe_2S_2]^{2+}$	0	121	363	727	1213	1818	-60.6	0.3	-60.6	0.00	0.3
$[Fe_2SSe]^{2+}$	0	114	341	682	1137	1706	-56.9	0.2	-56.9	-0.00	0.2
$[Fe_2Se_2]^{2+}$	0	109	326	653	1089	1636	-54.6	0.7	-54.5	-0.01	0.2
						state specific					
$[Fe_2S_2]^{2+}$	-70	74	348	727	1181	1669	-55.3	48.0	-56.5	0.46	2.0
$[Fe_2SSe]^{2+}$	-67	68	326	681	1106	1563	-51.8	45.5	-52.9	0.44	2.0
$[Fe_2Se_2]^{2+}$	-66	64	311	652	1059	1496	-49.7	43.9	-50.8	0.42	1.7
$S =$	0	1	2	3	4	5	J	SD_E	J_{biq}	λ	SD_E
						CAS(22,16)					
						state average					
$[Fe_2S_2]^{2+}$	0	134	400	795	1315	1944	-64.3	9.9	-64.5	0.09	1.7
$[Fe_2SSe]^{2+}$	0	122	366	728	1205	1781	-58.9	9.3	-59.1	0.09	1.9
$[Fe_2Se_2]^{2+}$	0	115	344	683	1129	1668	-55.2	8.5	-55.4	0.08	1.5
						state specific					
$[Fe_2S_2]^{2+}$	-466	-226	219	808	1480	2150	-81.3	106.1	-83.9	1.02	7.7
$[Fe_2SSe]^{2+}$	-459	-233	183	738	1369	2003	-76.5	98.3	-79.0	0.94	7.5
$[Fe_2Se_2]^{2+}$	-459	-244	158	692	1300	1912	-73.8	93.8	-76.1	0.90	7.2
						CAS(22,16)+NEVPT2					
						state average					
$[Fe_2S_2]^{2+}$	0	249	750	1504	2515	3787	-126.5	5.2	-126.4	-0.05	0.4
$[Fe_2SSe]^{2+}$	0	238	715	1434	2399	3615	-120.8	5.6	-120.7	-0.05	0.2
$[Fe_2Se_2]^{2+}$	0	230	693	1391	2328	3507	-117.2	5.3	-117.1	-0.05	0.5
						state specific					
$[Fe_2S_2]^{2+}$	-863	-441	374	1527	3024	4314	-163.2	225.3	-168.3	1.97	81.3
$[Fe_2SSe]^{2+}$	-828	-423	355	1451	2842	4234	-161.6	150.4	-165.2	1.38	35.2
$[Fe_2Se_2]^{2+}$	-842	-439	332	1402	2701	4405	-171.6	63.2	-172.5	0.35	54.1

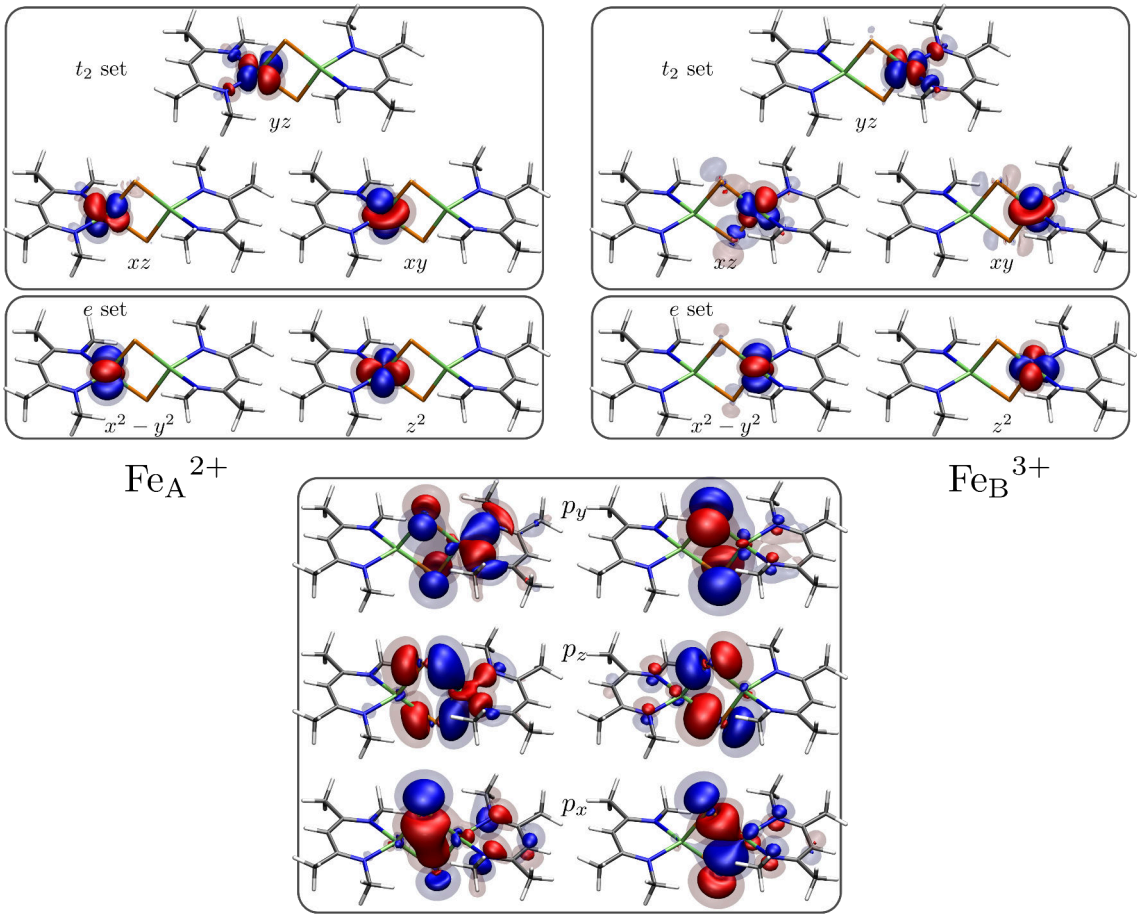

Figure A7. – Plots of the active orbitals in the CAS(23,16) wave function for the NacNac-capped $[Fe_2S_2]^{1+}$ clusters (see Figure 2.25). Orbital shapes are similar for $[Fe_2SSe]^{1+}$ and $[Fe_2Se_2]^{1+}$. Isosurfaces are shown at ± 0.05 and ± 0.025.

Figure A8. – Contributions of CFGs to the $S = \frac{1}{2}$ CAS(23,16) wave function for the NacNac-capped $[Fe_2X_2]^{1+}$ clusters (see also Figure 2.25). All values in [%].

	neutral	*d*-*d*	non-Hund	IVCT	LMCT
$[Fe_2S_2]^{1+}$	96.92	0.71	0.42	1.02	0.84
$[Fe_2SSe]^{1+}$	97.48	0.26	0.38	0.92	0.87
$[Fe_2Se_2]^{1+}$	98.02	0.03	0.27	0.74	0.90

Table A8. – Loewdin atomic contributions to the active orbitals in the $S = \frac{1}{2}$ CAS(23,16) wave function for the NacNac-capped $[Fe_2S_2]^{1+}$ clusters (see Figure A7 for plots). The first six rows correspond to orbitals based manly on the bridging ligands X, the next five on Fe_A, and the last five on Fe_B. The CASSCF wave function was averaged over the 5 d-d_n states with $S = \frac{1}{2}$.

	p_x	p_{z^*}	p_{y^*}	p_y	p_z	p_{x^*}	z^2	x^2-y^2	xy	xz	yz	z^2	x^2-y^2	xy	xz	yz
									Fe_A [%]							
$[Fe_2S_2]^{2+}$	3.7	13.9	14.3	18.6	20.9	23.7	97.4	95.0	91.4	88.8	88.4	0.1	0.1	0.5	0.2	0.1
$[Fe_2SSe]^{2+}$	5.3	12.9	13.8	18.9	21.3	24.0	97.0	95.3	91.3	90.4	87.0	0.1	0.1	0.4	0.1	0.1
$[Fe_2Se_2]^{2+}$	4.0	13.5	14.0	18.8	21.3	24.5	97.3	95.4	91.8	90.4	86.7	0.0	0.1	0.5	0.2	0.0
									Fe_B [%]							
$[Fe_2S_2]^{2+}$	14.7	8.9	5.9	2.8	8.6	2.2	0.1	0.3	0.8	0.2	0.2	97.8	97.7	96.1	94.7	94.3
$[Fe_2SSe]^{2+}$	13.3	7.0	5.4	2.9	8.4	2.3	0.1	0.3	0.8	0.0	0.4	97.9	97.8	96.2	95.2	93.9
$[Fe_2Se_2]^{2+}$	13.7	8.1	5.3	2.3	8.2	1.8	0.0	0.2	0.8	0.0	0.3	97.8	97.9	96.5	94.8	94.4
									X_2 [%]							
$[Fe_2S_2]^{2+}$	4.8	18.3	4.6	43.6	11.7	8.3	0.6	0.4	1.1	3.1	2.9	0.1	0.1	0.1	1.1	2.1
$[Fe_2SSe]^{2+}$	8.4	30.9	7.3	41.1	12.9	6.7	0.8	0.2	1.3	4.8	1.4	0.1	0.3	0.5	0.9	2.0
$[Fe_2Se_2]^{2+}$	6.9	22.6	6.1	46.1	13.0	9.5	0.5	0.2	1.2	5.6	1.9	0.6	0.2	0.6	0.7	2.4
									NacNac [%]							
$[Fe_2S_2]^{2+}$	75.3	57.3	74.2	32.1	57.1	64.8	1.1	3.5	5.5	5.8	6.6	0.6	0.7	2.1	2.8	1.9
$[Fe_2SSe]^{2+}$	71.7	47.0	71.9	34.3	56.4	65.3	1.2	3.3	5.6	2.9	9.7	0.6	0.7	2.0	2.3	2.2
$[Fe_2Se_2]^{2+}$	74.3	53.8	73.5	29.8	55.8	63.2	1.0	3.6	5.3	2.7	10.1	0.4	0.9	1.7	2.9	1.3

Table A9. – Energies for select low-lying excited states in the CAS(23,16) wave function including NEVPT2 for NacNac-capped $[Fe_2X_2]^{1+}$ clusters (see also Figure 2.28). d-d_n states correspond to the $Fe_A{}^{2+}Fe_B{}^{3+}$ valence isomer, $IVCT_n$ to $Fe_A{}^{3+}Fe_B{}^{2+}$. The CASSCF wave function was averaged either over the 5 d-d_n states or the 5 $IVCT_n$ states for each spin state separately. All values are given in $[cm^{-1}]$.

	d-d_0	d-d_1	d-d_2	d-d_3	d-d_4	$IVCT_0$	$IVCT_1$
				$S = \frac{1}{2}$			
$[Fe_2S_2]^{1+}$	0	470	2920	5912	8028	6635	7448
$[Fe_2SSe]^{1+}$	0	748	2712	5101	8494	6454	6835
$[Fe_2Se_2]^{1+}$	0	170	2170	5096	7884	7179	7566
				$S = \frac{3}{2}$			
$[Fe_2S_2]^{1+}$	372	893	3016			7186	9718
$[Fe_2SSe]^{1+}$	261	1032	2717			7127	9805
$[Fe_2Se_2]^{1+}$	199	411	2147			7762	10183

Table A10. – Energies of all low-lying excited states in the CAS(11,10) wave function for the NacNac-capped $[Fe_2X_2]^{1+}$ clusters, including NEVPT2. d-d_n and $IVCT_n$ states correspond to the $Fe_A^{2+}Fe_B^{3+}$ and $Fe_A^{3+}Fe_B^{2+}$ valence isomer, respectively (see also Figure 2.29). The CASSCF wave function was averaged either over all 25 d-d_n states or over all 25 $IVCT_n$ states. All values in [cm^{-1}].

	$S=\frac{1}{2}$	$S=\frac{3}{2}$	$S=\frac{5}{2}$	$S=\frac{7}{2}$	$S=\frac{9}{2}$
			d-d_0		
$[Fe_2S_2]^{1+}$	0	185	491	909	1423
$[Fe_2SSe]^{1+}$	0	161	427	793	1243
$[Fe_2Se_2]^{1+}$	0	119	317	589	929
			d-d_1		
$[Fe_2S_2]^{1+}$	454	623	904	1291	1766
$[Fe_2SSe]^{1+}$	648	799	1046	1381	1782
$[Fe_2Se_2]^{1+}$	67	193	399	676	1004
			d-d_2		
$[Fe_2S_2]^{1+}$	2441	2460	2496	2558	2658
$[Fe_2SSe]^{1+}$	2274	2298	2344	2420	2546
$[Fe_2Se_2]^{1+}$	1779	1796	1829	1883	1970
			d-d_3		
$[Fe_2S_2]^{1+}$	5675	5780	5955	6200	6516
$[Fe_2SSe]^{1+}$	4906	5011	5185	5428	5736
$[Fe_2Se_2]^{1+}$	4848	4910	5014	5161	5355
			d-d_4		
$[Fe_2S_2]^{1+}$	7886	8063	8359	8771	9286
$[Fe_2SSe]^{1+}$	8327	8469	8708	9043	9466
$[Fe_2Se_2]^{1+}$	7775	7910	8132	8436	8810
			$IVCT_0$		
$[Fe_2S_2]^{1+}$	2871	3078	3416	3875	4429
$[Fe_2SSe]^{1+}$	3540	3710	3986	4355	4788
$[Fe_2Se_2]^{1+}$	2727	2867	3095	3403	3773
			$IVCT_1$		
$[Fe_2S_2]^{1+}$	3701	3892	4211	4656	5212
$[Fe_2SSe]^{1+}$	3994	4167	4454	4854	5354
$[Fe_2Se_2]^{1+}$	3134	3270	3496	3808	4197
			$IVCT_2$		
$[Fe_2S_2]^{1+}$	6102	6125	6170	6248	6380
$[Fe_2SSe]^{1+}$	6377	6409	6469	6572	6740
$[Fe_2Se_2]^{1+}$	5311	5333	5373	5438	5537
			$IVCT_3$		
$[Fe_2S_2]^{1+}$	9202	9300	9468	9711	10036
$[Fe_2SSe]^{1+}$	8702	8792	8946	9167	9461
$[Fe_2Se_2]^{1+}$	7967	8035	8150	8316	8533
			$IVCT_4$		
$[Fe_2S_2]^{1+}$	10817	11038	11402	11906	12523
$[Fe_2SSe]^{1+}$	11083	11278	11602	12047	12593
$[Fe_2Se_2]^{1+}$	10481	10634	10885	11230	11654

A3. The resting state of nitrogenase

A3.1. Electronic structure

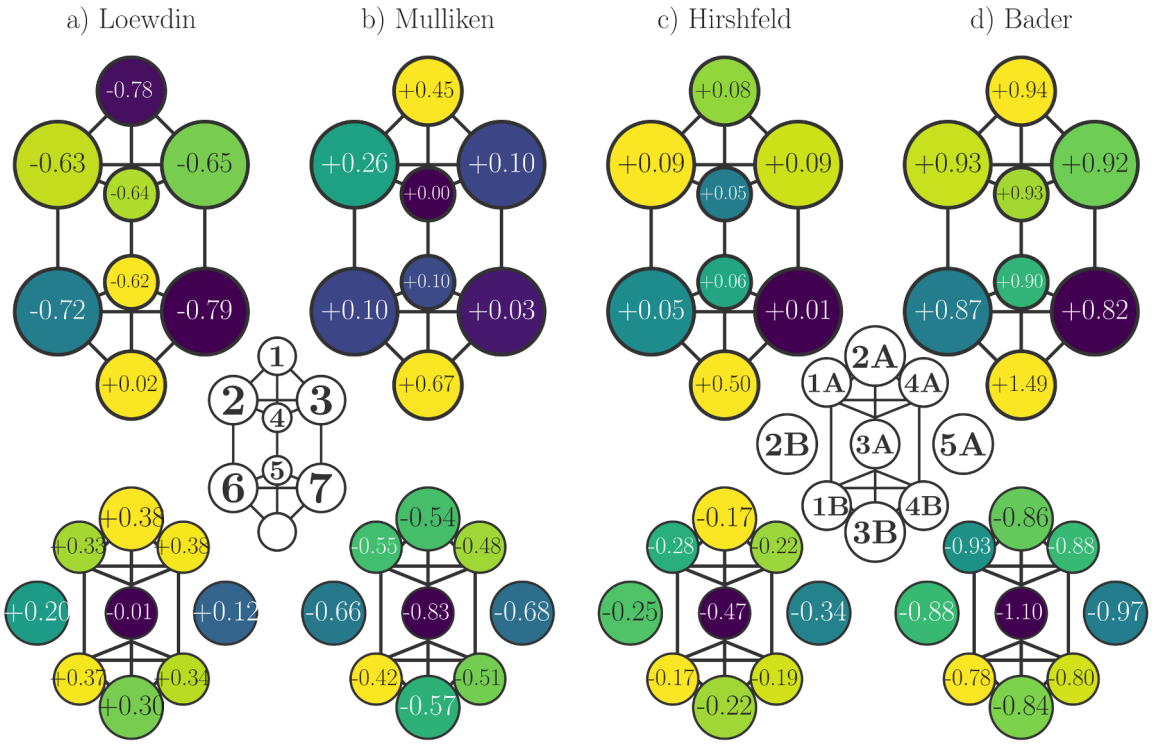

Figure A9. – Comparison of atomic charges for the metal centers (top) and the sulfide ligands (bottom) in FeMoco (E_0 QM2 model, BS7-235 $M_S = \frac{3}{2}$). See Figure 4.2 for atom labels.

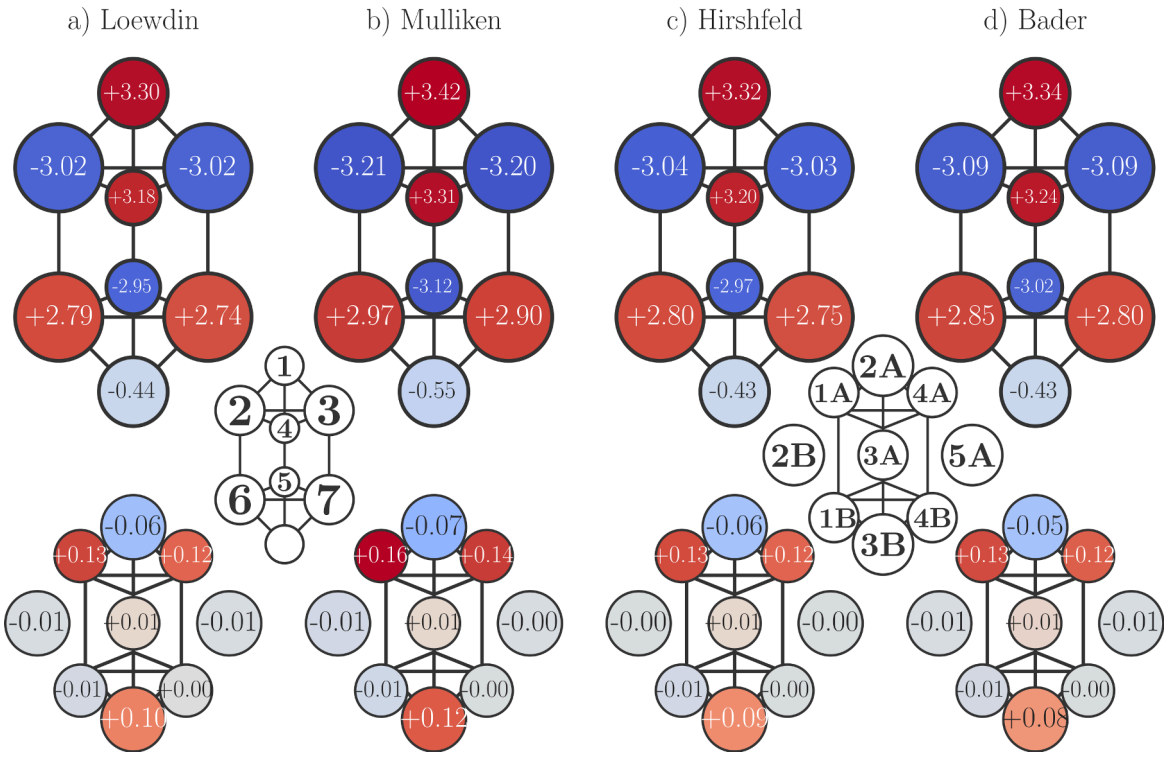

Figure A10. – Comparison of spin populations for the metal centers (top) and the sulfide ligands (bottom) in FeMoco (E_0 QM2 model, BS7-235 $M_S = \frac{3}{2}$). See Figure 4.2 for atom labels.

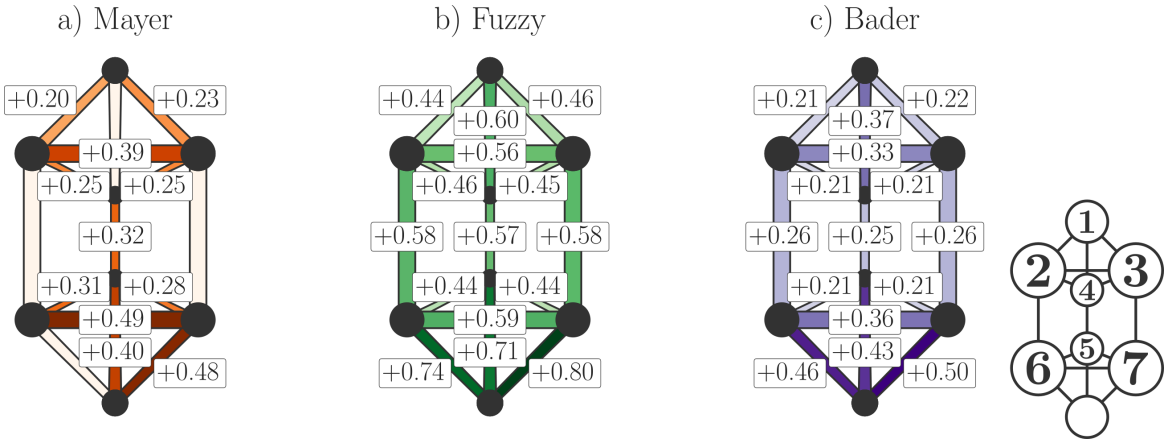

Figure A11. – Comparison of metal-metal bond orders in FeMoco (E_0 QM2 model, BS7-235 $M_S = \frac{3}{2}$). Values below 0.1 are not shown for the Mayer bond order. See Figure 4.2 for atom labels.

Choice of the diamagnetic ions

With diamagnetic substitution, one attempts to emulate the electrostatic influence of the magnetic centers with diamagnetic ions as well as possible. With an average oxidation state of $Fe^{2.5+}$ the choice for the diamagnetic ion is not obvious for FeMoco. A plausible indicator is the charge on the ligand, which should not change much upon diamagnetic substitution. Figure A12 shows the sulfide atomic charges in FeMoco for the substitution Fe→Ga^{3+} and Fe→Zn^{2+} as well as custom diamagnetic ions that emulate intermediate oxidation states. The custom ions are generated by adjusting the charge of the nuclei, therefore expanding or contracting the Ga^{3+} and Zn^{2+} ions, respectively. The smallest deviation of the sulfide ligands relative to the unsubstituted FeMoco was achieved for $Ga^{2.5+}$.

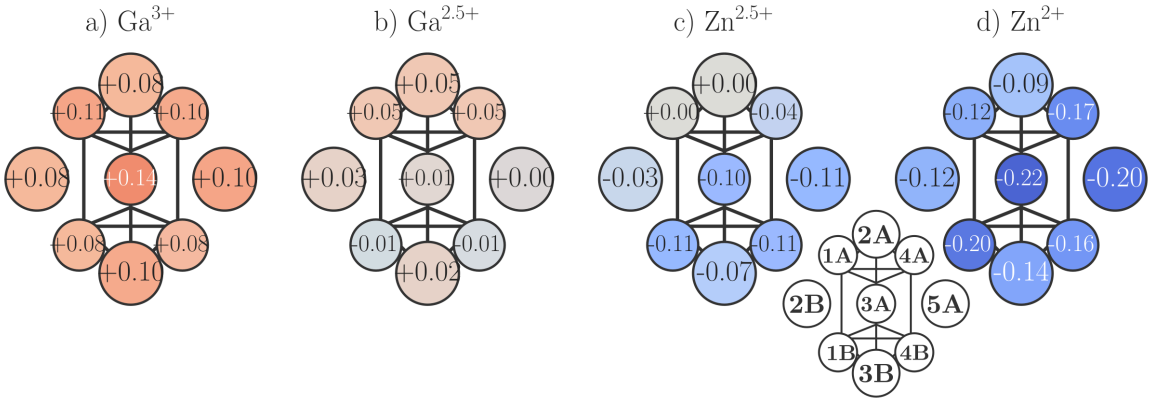

Figure A12. – Differences in the Hirshfeld atomic charges of the sulfide ligands in fully diamagnetically substituted FeMoco relative to the unsubstituted FeMoco (E_0 QM2 model, BS7-235 $M_S = \frac{3}{2}$). The Fe centers are substituted with (a) Ga^{3+} $Z = 31$, (b) $Ga^{2.5+}$ $Z = 30.5$, (c) $Zn^{2.5+}$ $Z = 30.5$, and (d) Zn^{3+} $Z = 30$. Mo was substituted with In^{3+} in all cases. See Figure 4.2 for atom labels.

	FE1	FE2	FE3	FE4	FE5	FE6	FE7	M
BS1 567	3.52	0.50	3.39	3.28	-3.27	-3.09	-3.12	1.30
BS2 234	3.57	-3.13	-3.23	-2.99	3.23	3.16	3.14	-1.36
BS3 123	-3.39	-2.34	-3.34	3.37	3.40	3.23	3.15	-1.43
BS3 124	-3.17	-2.51	3.38	-3.30	3.31	3.24	3.15	-1.40
BS3 134	-3.57	3.51	-3.21	-3.12	3.45	3.40	3.26	-1.36
BS4 257	3.32	-3.33	3.26	2.35	-3.13	3.13	-3.00	0.30
BS4 356	2.51	3.11	-3.36	3.24	-3.13	-3.03	3.06	0.57
BS4 467	3.33	2.28	3.25	-3.33	3.14	-3.10	-3.04	0.59
BS5 256	3.46	-3.36	2.47	3.30	-3.18	-3.09	2.93	0.45
BS5 267	3.48	-3.38	3.27	2.52	3.00	-3.12	-3.11	0.49
BS5 357	3.41	2.54	-3.39	3.27	-3.19	2.97	-3.07	0.30
BS5 367	2.55	3.20	-3.36	3.28	3.12	-3.16	-3.08	0.54
BS5 456	3.45	3.34	2.31	-3.36	-3.18	-3.12	2.96	0.51
BS5 457	3.44	2.34	3.35	-3.35	-3.16	2.99	-3.08	0.29
BS6 156	-3.58	3.10	3.25	3.12	-3.08	-3.04	2.97	0.53
BS6 157	-3.58	3.15	3.16	3.11	-3.06	2.94	-2.98	0.43
BS6 167	-3.59	3.08	3.15	3.17	3.03	-3.03	-2.99	0.55
BS7 235	3.52	-3.23	-3.26	3.36	-3.18	2.99	3.00	-0.54
BS7 247	3.51	-3.24	3.36	-3.24	3.06	2.96	-3.13	-0.47
BS7 346	3.51	3.35	-3.27	-3.27	3.06	-3.17	2.95	-0.25
BS8 236	3.50	-3.37	-3.38	3.39	3.16	-3.14	3.03	-0.24
BS8 237	3.51	-3.39	-3.39	3.39	3.18	3.08	-3.09	-0.46
BS8 245	3.50	-3.36	3.41	-3.34	-3.15	3.07	3.07	-0.47
BS8 246	3.49	-3.36	3.40	-3.35	3.11	-3.13	3.05	-0.24
BS8 345	3.49	3.37	-3.37	-3.37	-3.15	3.09	3.06	-0.45
BS8 347	3.49	3.35	-3.38	-3.38	3.13	3.07	-3.06	-0.41
BS9 126	-3.47	-3.35	3.36	3.35	3.16	-3.00	3.07	-0.22
BS9 137	-3.46	3.32	-3.38	3.30	3.16	3.11	-2.93	-0.37
BS9 145	-3.48	3.32	3.36	-3.34	-3.03	3.14	3.10	-0.41
BS10 125	-3.49	-3.34	3.45	3.33	-3.03	3.10	3.18	-0.55
BS10 127	-3.49	-3.33	3.31	3.44	3.27	3.09	-2.97	-0.53
BS10 135	-3.50	3.42	-3.36	3.29	-3.01	3.23	3.08	-0.54
BS10 136	-3.50	3.29	-3.35	3.42	3.25	-3.02	3.07	-0.32
BS10 146	-3.51	3.33	3.45	-3.34	3.14	-3.04	3.16	-0.31
BS10 147	-3.51	3.43	3.33	-3.35	3.14	3.19	-2.96	-0.51

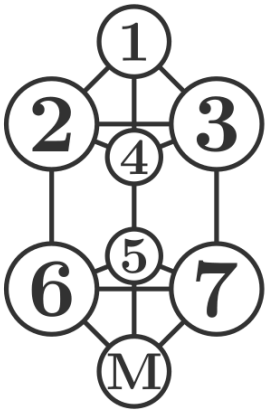

Figure A13. – Mulliken spin populations for the metal centers in FeMoco for all 35 possible BS determinants (E_0 cluster* model, $M_S = \frac{3}{2}$). See Figure 4.2 for atom labels.

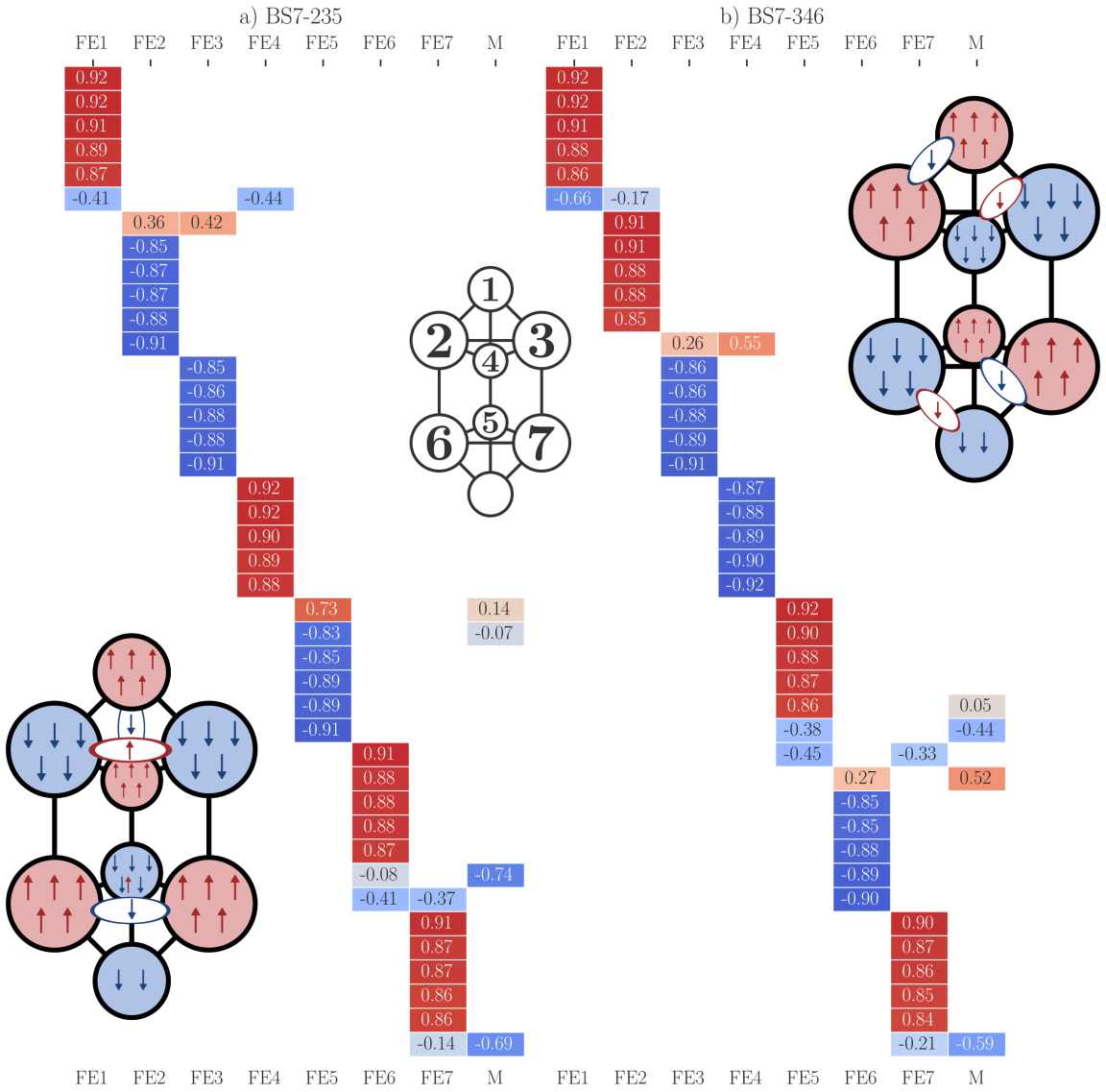

Figure A14. – Hirshfeld atomic contributions to the Foster-Boys localized orbitals of (a) the BS7-235 determinant and (b) the BS7-346 determinant in FeVco (E_0 QM2 model, $M_S = \frac{3}{2}$). Red/positive values corresponds to α electrons and blue/negative values to β electrons. Only the metal-based valence orbitals are shown. The insets show the simplified electronic structure that can be derived from the localized orbital analysis (recreated after Benediktsson and Bjornsson,[19] see Figure 4.2 for atom labels). The analogous plot for FeMoco is shown in Figure 4.3.

A3.2. Influence of the protein environment

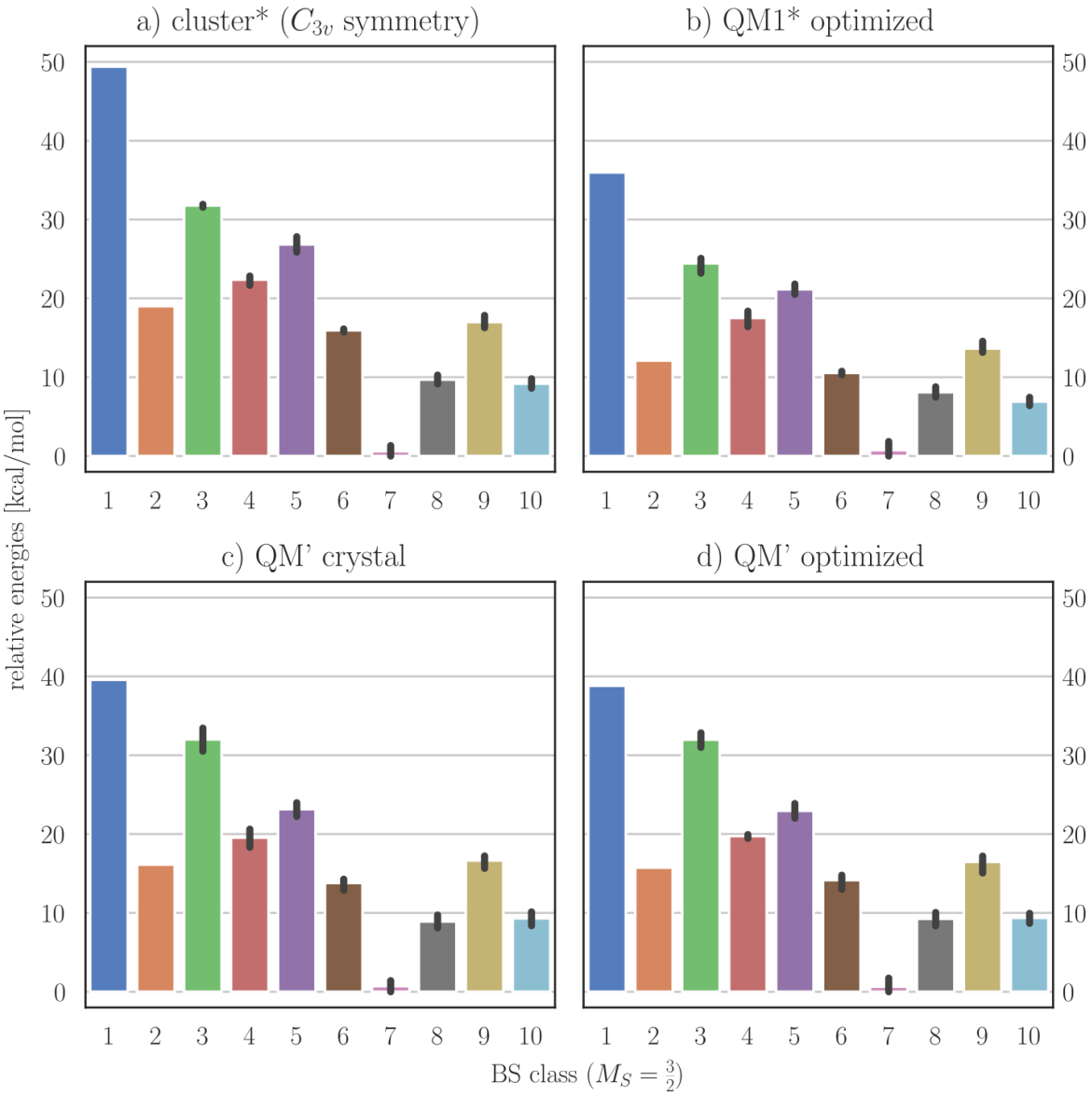

Figure A15. – Average energies of the 10 BS determinant classes in FeMoco for (a) the cluster* model, (b) the QM1* model, (c) single point QM/MM calculations on the E_0 MoFe XRD structure (QM' model),[69] (d) same as (c), but for relaxed geometries. Energies for c) and d) were kindly supplied by Thorhallsson and the QM' region is of comparable size to QM2. $M_S = \frac{3}{2}$ is used for all BS determinants and BS7-235 is lowest in energy in all models.

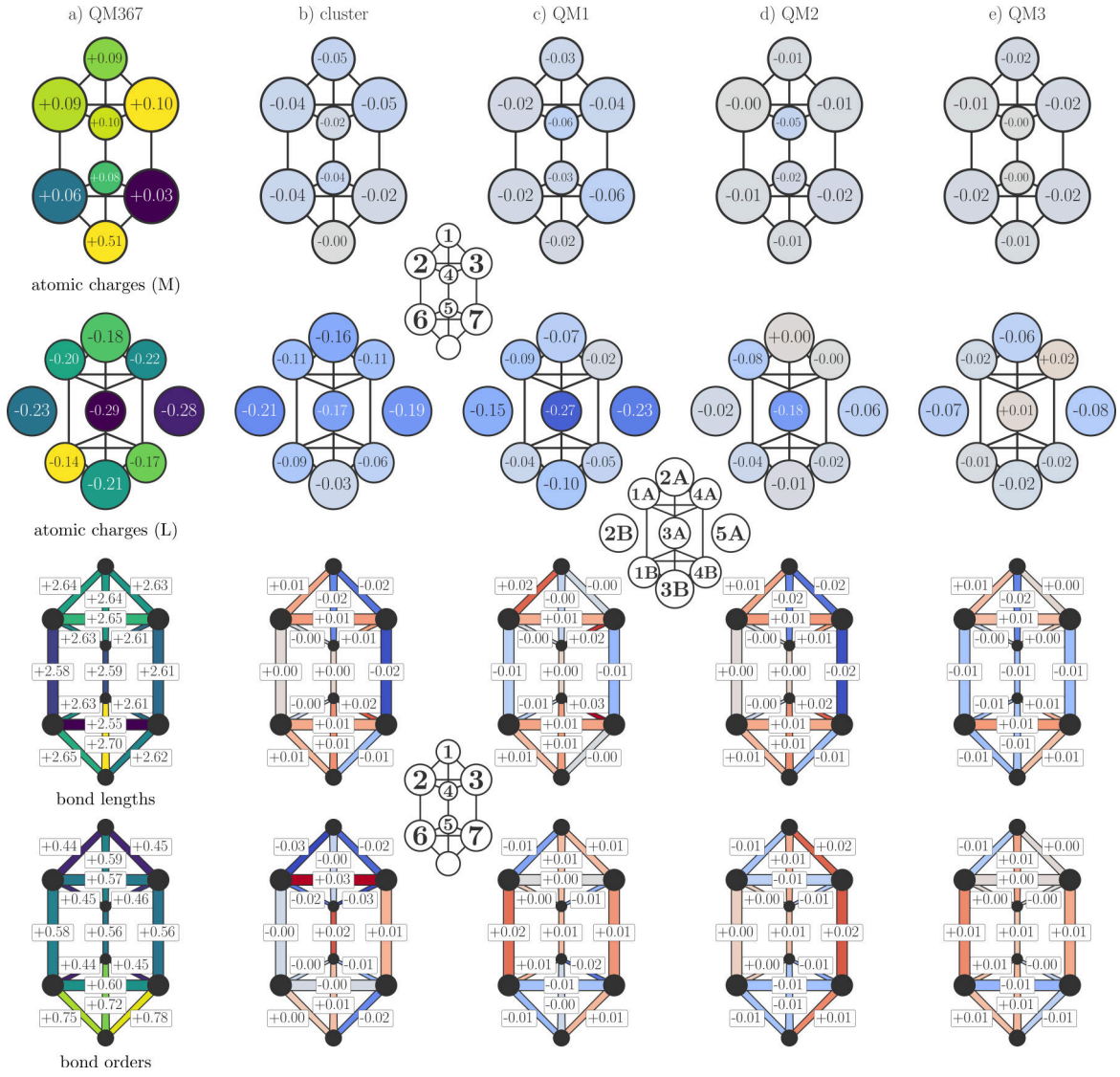

Figure A16. – Comparison of FeMoco models used in this work with the QM367 model provided by Benediktsson (E_0 BS7-235 $M_S = \frac{3}{2}$). (a) Hirshfeld atomic charges for the metal centers and sulfide ligands in FeMoco (top) as well and the metal-metal bond lengths [Å] and bond orders (bottom). (b-e) Differences in the cluster, QM1, QM2, and QM3 model relative to the QM367 model. See Figure 4.1 for model and Figure 4.2 for atom labels.

A4. CO binding to nitrogenase

A4.1. Diamagnetically substituted cofactors

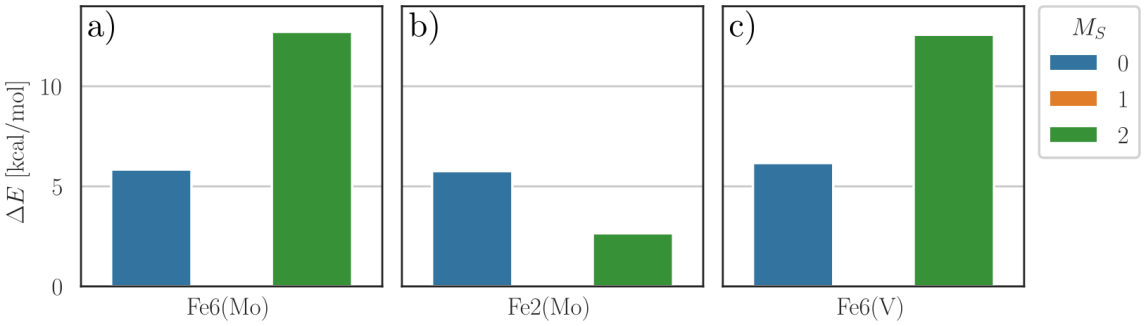

Figure A17. – Energies of different M_S values for the CO-bound Fe^{2+} center in (a) FeMoco-based [$FeGa_6InS_9C$] with Fe6 unsubstituted, (b) FeMoco-based [$FeGa_6InS_9C$] with Fe2 unsubstituted, and (c) FeVco-based [$FeGa_7S_8CO_3C$] Fe6 unsubstituted (E_0 cluster model). $M_S = 0$ is lowest in energy in all cases.

Table A11. – Calculated coupling constants for the Fe6/Fe2 dimer in diamagnetically FeMoco with and without CO bound. The redox state is $Fe^{2+}Fe^{3+}$. Values for J were obtained with the Yamaguchi projection from the $M_S = \frac{9}{2}$ high spin determinant and the $M_S = \frac{1}{2}$ BS determinant.

CO	model	J [cm^{-1}]
no CO	cluster	-117.10
CO−Fe6	cluster (ferro opt)	33.62
	cluster (anti opt)	25.81
μ−CO	E_1 QM/MM (single point)	46.70
	E_2 QM/MM (single point)	18.13

A4.2. Localized orbitals

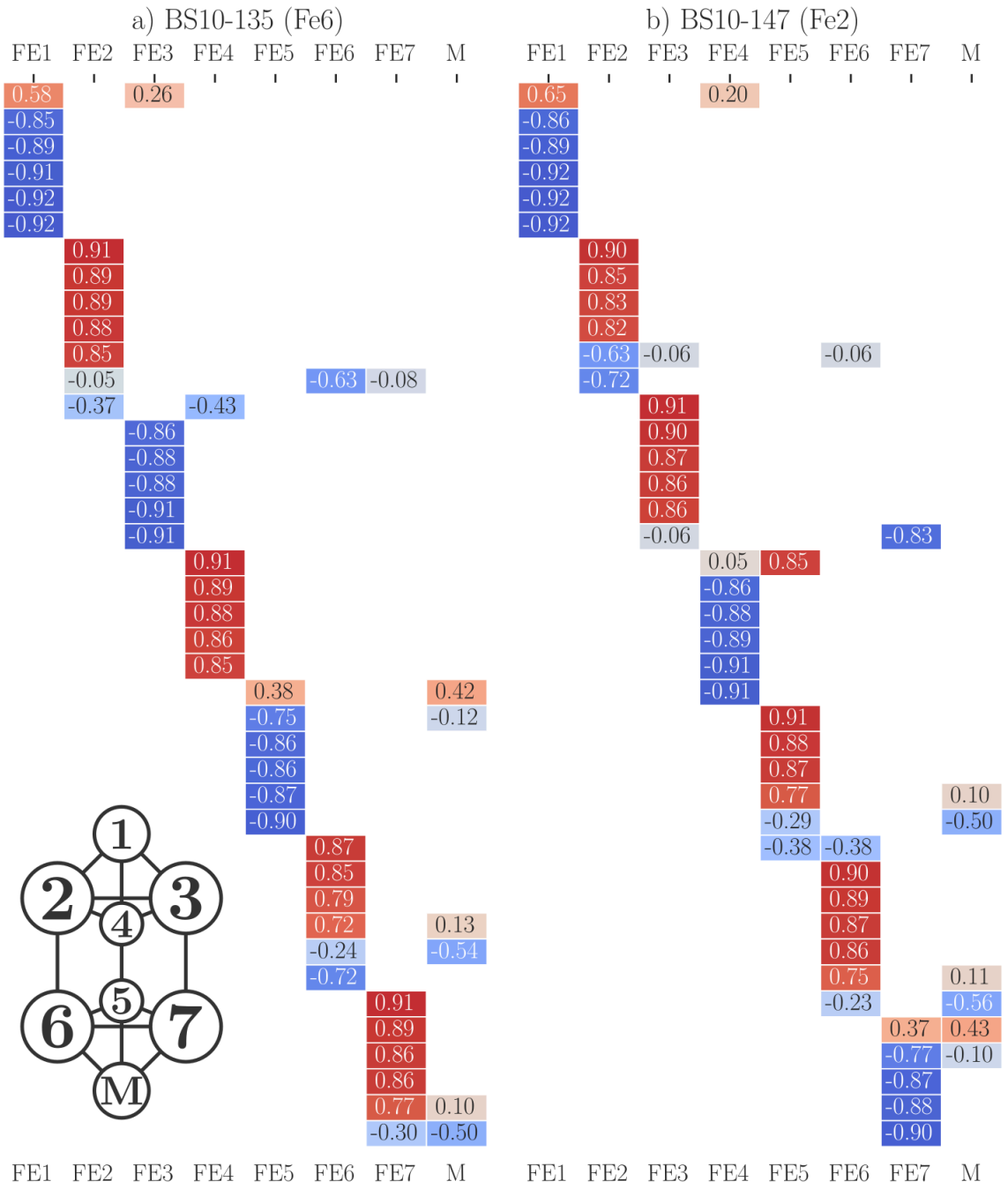

Figure A18. – Hirshfeld atomic contributions to the Foster-Boys localized orbitals in the E_0 FeMoco cluster model with CO bound to (a) Fe6 or (b) Fe2 (BS10-135 $M_S = \frac{1}{2}$ and BS10-147 $M_S = \frac{1}{2}$, respectively). Red/positive values corresponds to α electrons and blue/negative values to β electrons. Only the metal-based valence orbitals are shown. See Figure 4.2 for atom labels.

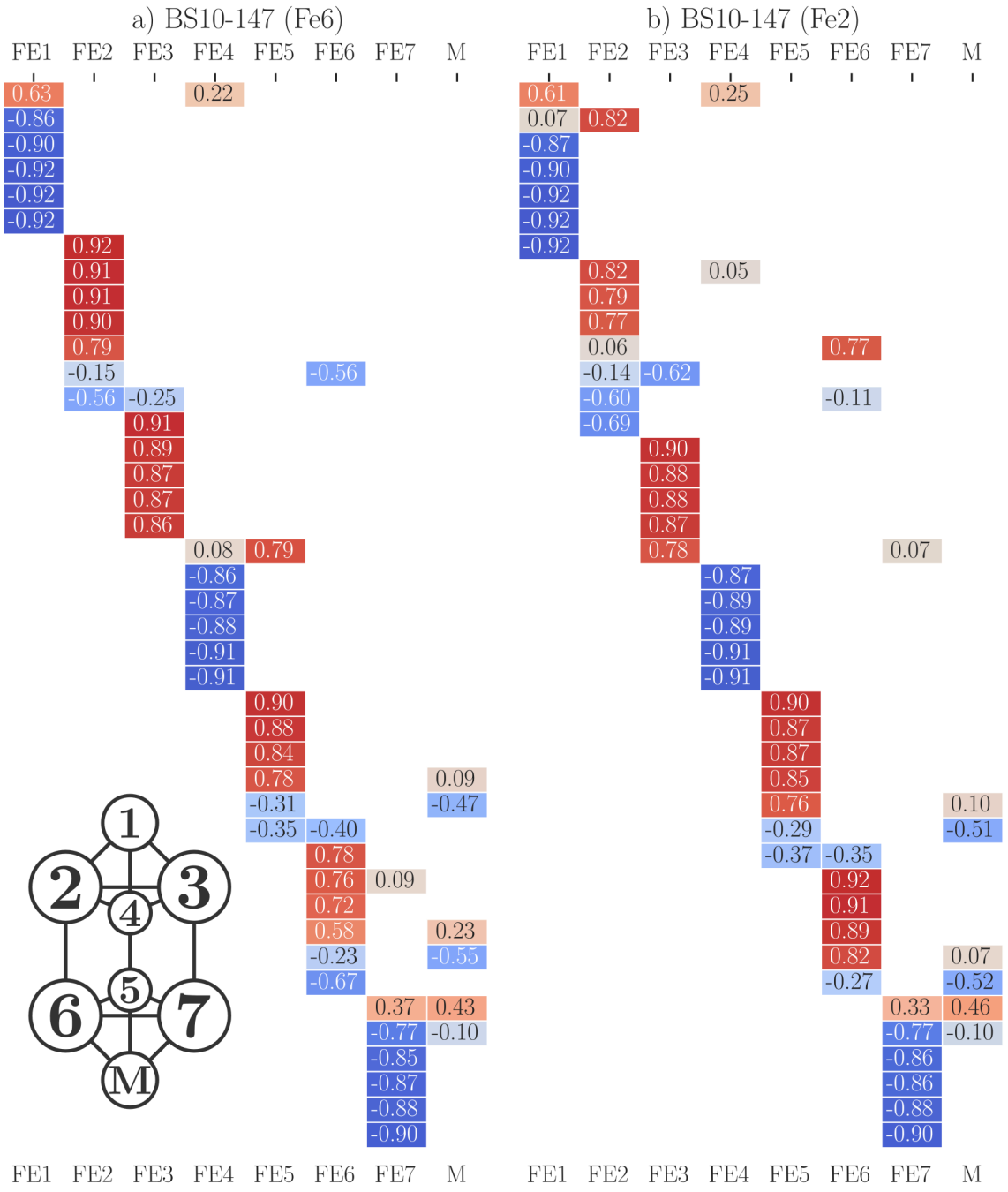

Figure A19. – Hirshfeld atomic contributions to the Foster-Boys localized orbitals in the E_1 FeMoco cluster model with CO bound to (a) Fe6 or (b) Fe2 (BS10-147 $M_S = 2$, respectively). Red/positive values corresponds to α electrons and blue/negative values to β electrons. Only the metal-based valence orbitals are shown. See Figure 4.2 for atom labels.

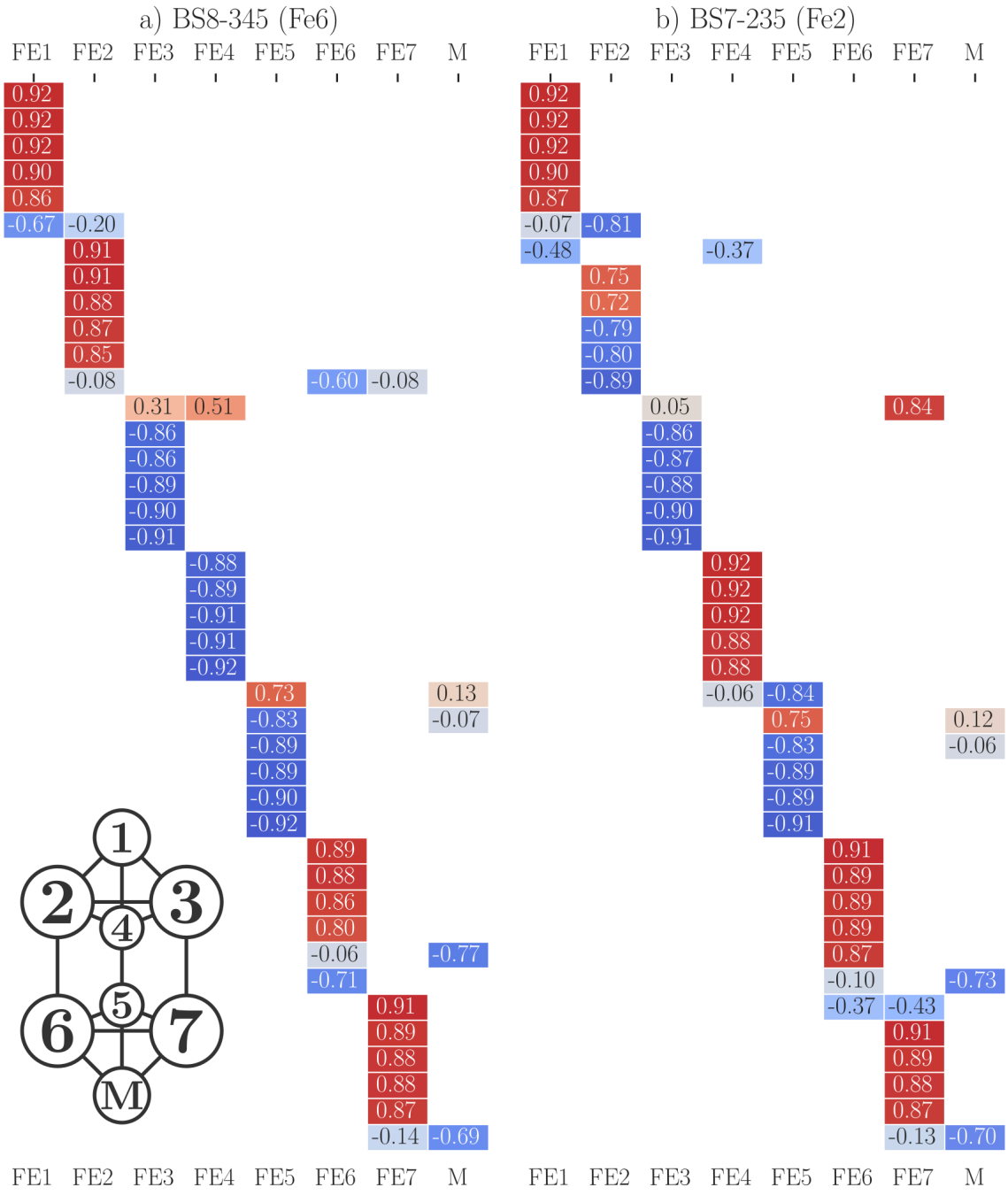

Figure A20. – Hirshfeld atomic contributions to the Foster-Boys localized orbitals in the E_0 FeVco cluster model with CO bound to (a) Fe6 or (b) Fe2 (BS8-345 $M_S = \frac{1}{2}$ and BS7-235 $M_S = \frac{5}{2}$, respectively). Red/positive values corresponds to α electrons and blue/negative values to β electrons. Only the metal-based valence orbitals are shown. See Figure 4.2 for atom labels.

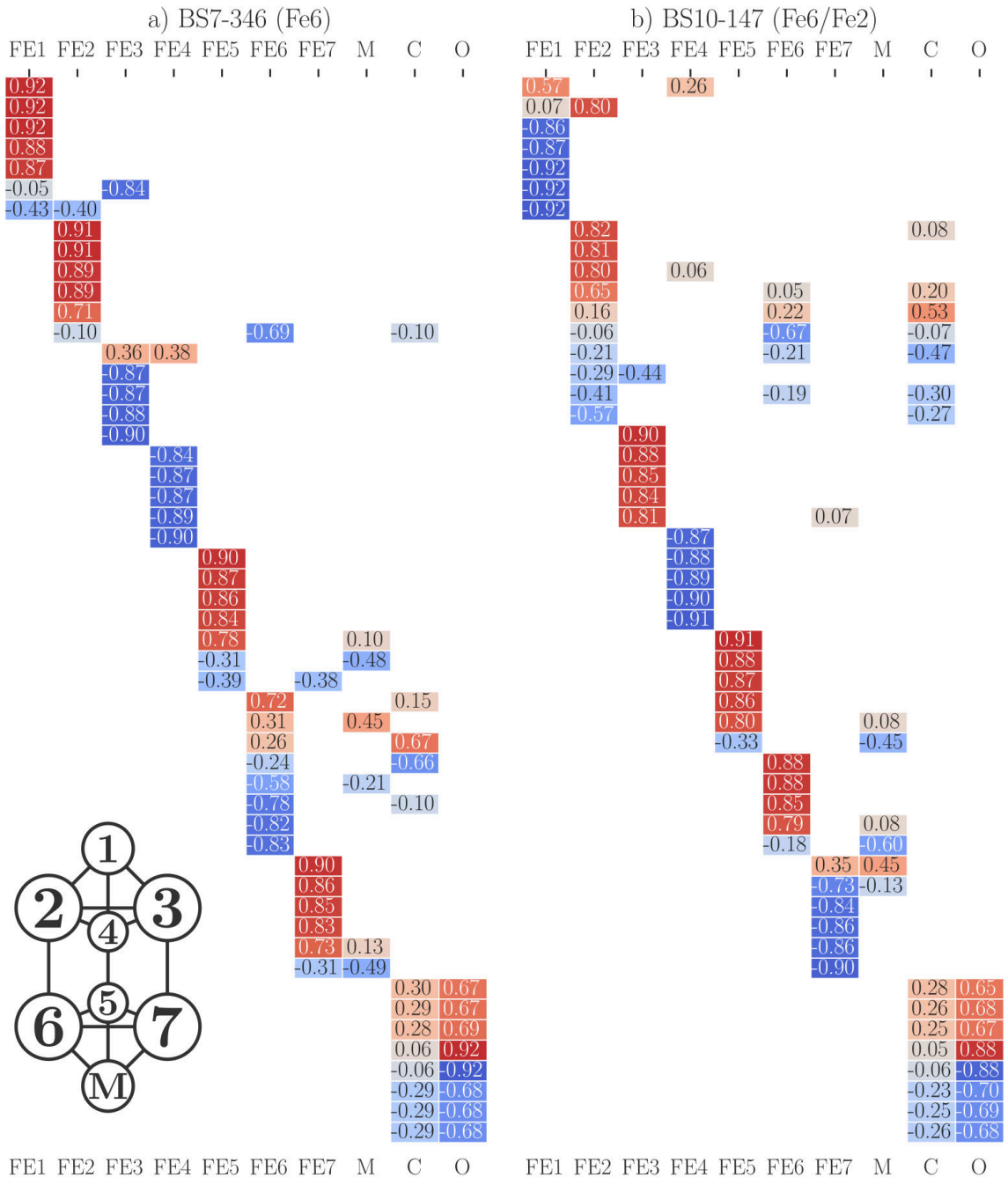

Figure A21. – Hirshfeld atomic contributions to the Foster-Boys localized orbitals in the E_1 FeMoco QM2 model with CO bound to (a) Fe6 or (b) bridging Fe6/Fe2 (BS7-346 $M_S = 2$ and BS10-147 $M_S = 0$, respectively). Red/positive values corresponds to α electrons and blue/negative values to β electrons. Only the metal-based and CO-based valence orbitals are shown. See Figure 4.2 for atom labels.

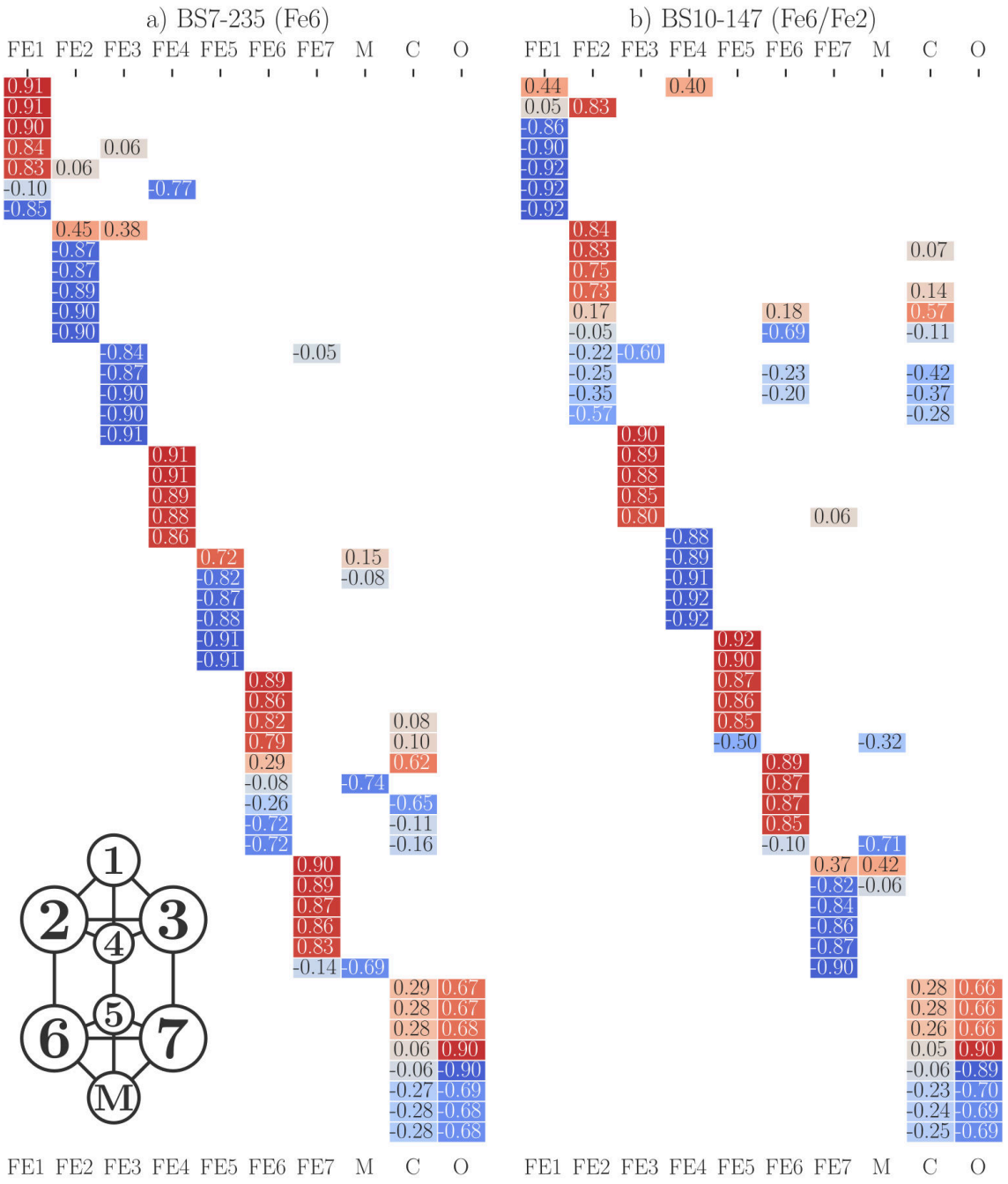

Figure A22. – Hirshfeld atomic contributions to the Foster-Boys localized orbitals in the E_1 FeVco QM2 model with CO bound to (a) Fe6 or (b) bridging Fe6/Fe2 (BS7-235 $M_S = 0$ and BS10-147 $M_S = 0$, respectively). Red/positive values corresponds to α electrons and blue/negative values to β electrons. Only the metal-based and CO-based valence orbitals are shown. See Figure 4.2 for atom labels.

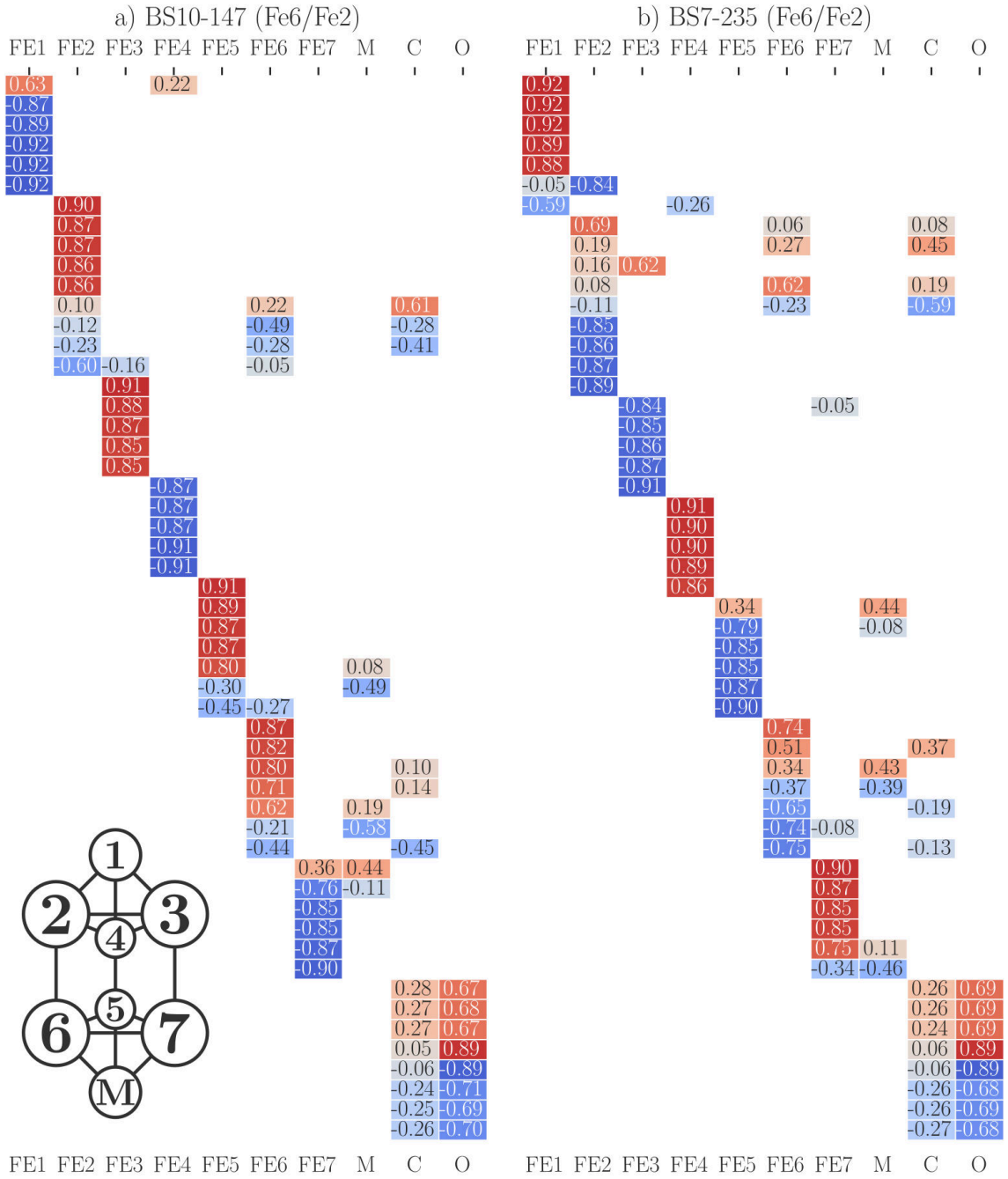

Figure A23. – Hirshfeld atomic contributions to the Foster-Boys localized orbitals in the E_2 FeMoco QM2 model with CO bound bridging Fe6/Fe2 for the (a) BS7-235 $M_S = \frac{1}{2}$ determinant and the (b) BS10-147 $M_S = \frac{1}{2}$ determinant. Red/positive values corresponds to α electrons and blue/negative values to β electrons. Only the metal-based and CO-based valence orbitals are shown. See Figure 4.2 for atom labels.

A4.3. Binding energies

Table A12. – Lowest-energy BS/M_S combination of for each substrate-free model, which serve as reference for the CO binding energy estimate (see Section 5.1).

E_n	model	nitrogenase	BS	M_S
E_0	cluster	Mo	BS7-346	1.5
E_0	cluster	V	BS7-235	1.5
E_0	QM2	Mo	BS7-346	1.5
E_0	QM2	V	BS7-235	1.5
E_1	cluster	Mo	BS7-346	2
E_1	cluster	V	BS7-235	1
E_1	QM2	Mo	BS7-235	2
E_1	QM2	V	BS7-235	1

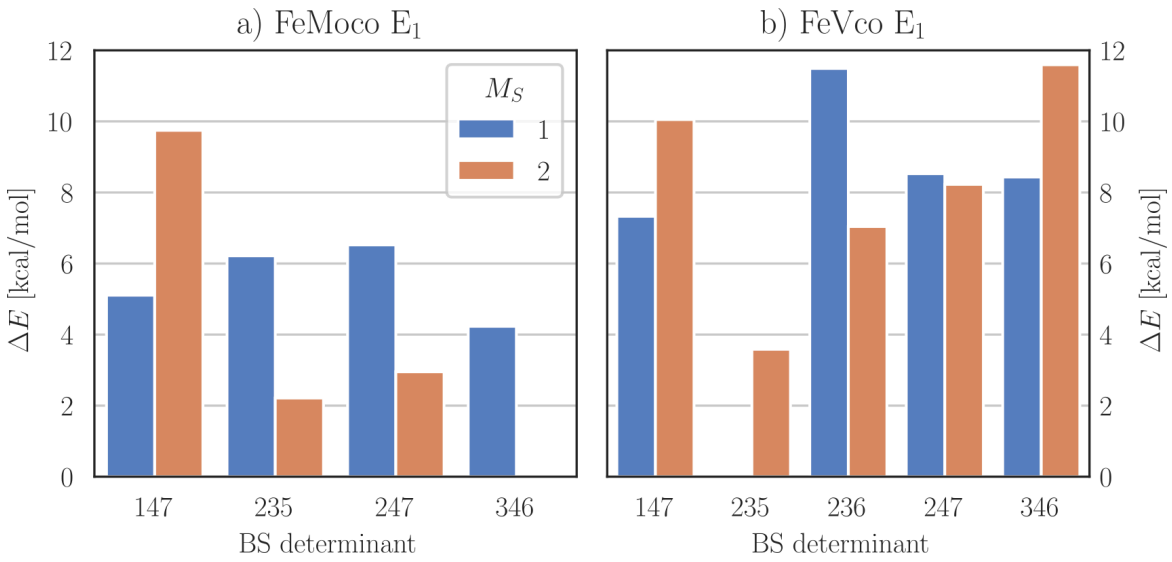

Figure A24. – Energies for the substrate-free (a) FeMoco and (b) FeVco cluster model in the E_1 redox state relative to BS7-346 $M_S = 2$ and BS7-235 $M_S = 1$, respectively.

Table A13. – CO binding energies [kcal/mol] for the E_0 FeMoco (left) and FeVco (right) cluster model.

BS	M_S	binding site	ΔE_{bnd}
135	0.5	Fe6	-5.96
147	0.5	Fe6	-5.26
235	0.5	Fe6	-4.12
345	0.5	Fe6	-3.18
346	2.5	Fe6	-2.61
347	0.5	Fe6	-2.37
125	0.5	Fe6	-1.42
247	0.5	Fe6	-1.28
127	0.5	Fe6	-0.38
236	2.5	Fe6	0.86
246	2.5	Fe6	2.23
245	0.5	Fe6	4.19
237	0.5	Fe6	4.82
146	2.5	Fe6	5.33
136	2.5	Fe6	5.35
147	0.5	Fe2	-9.09
135	0.5	Fe2	-7.47
235	2.5	Fe2	-7.42
247	2.5	Fe2	-6.56
346	0.5	Fe2	-5.27
236	2.5	Fe2	-4.51
246	2.5	Fe2	-3.01
347	0.5	Fe2	-2.70
146	0.5	Fe2	-2.57
345	0.5	Fe2	-2.37
136	0.5	Fe2	-1.60
237	2.5	Fe2	1.53
127	2.5	Fe2	1.86
125	2.5	Fe2	2.20
245	2.5	Fe2	3.35

BS	M_S	binding site	ΔE_{bnd}
345	0.5	Fe6	-5.33
235	0.5	Fe6	-4.29
147	0.5	Fe6	-2.91
135	0.5	Fe6	-1.82
125	0.5	Fe6	1.20
236	2.5	Fe6	2.16
127	0.5	Fe6	2.16
346	2.5	Fe6	3.05
347	0.5	Fe6	4.38
247	0.5	Fe6	4.60
245	0.5	Fe6	5.56
237	0.5	Fe6	5.76
136	2.5	Fe6	7.65
146	2.5	Fe6	7.69
246	2.5	Fe6	8.20
235	2.5	Fe2	-5.15
147	0.5	Fe2	-4.12
135	0.5	Fe2	-2.53
345	2.5	Fe2	-2.37
236	2.5	Fe2	-0.17
247	2.5	Fe2	0.90
146	0.5	Fe2	0.94
346	0.5	Fe2	1.24
245	2.5	Fe2	2.31
136	0.5	Fe2	2.87
347	0.5	Fe2	3.85
237	2.5	Fe2	4.31
246	2.5	Fe2	5.91
127	2.5	Fe2	6.08
125	2.5	Fe2	7.31

Table A14. – CO binding energies [kcal/mol] for the E_1 FeMoco (left) and FeVco (right) cluster model.

BS	M_S	binding site	ΔE_{bnd}
147	0	Fe6	-15.00
346	2	Fe6	-13.46
346	3	Fe6	-11.50
135	0	Fe6	-11.07
235	1	Fe6	-10.90
147	0	Fe2	-10.99
135	0	Fe2	-9.46
346	3	Fe2	-3.73
346	2	Fe2	-2.65
235	1	Fe2	-2.17

BS	M_S	binding site	ΔE_{bnd}
135	0	Fe6	-11.75
235	1	Fe6	-11.42
235	0	Fe6	-9.54
147	1	Fe6	-7.68
147	0	Fe6	-6.54
136	2	Fe6	-5.85
135	0	Fe2	-8.93
235	0	Fe2	-8.90
147	1	Fe2	-3.73
147	0	Fe2	-2.68
235	1	Fe2	-2.07

Table A15. – CO binding energies [kcal/mol] for the E_0 FeMoco (left) and FeVco (right) QM2 model.

BS	M_S	binding site	ΔE_{bnd}
147	0.5	Fe6	-7.77
135	0.5	Fe6	-7.10
235	0.5	Fe6	-5.69
347	0.5	Fe6	-5.55
345	0.5	Fe6	-5.11
247	0.5	Fe6	-3.98
125	0.5	Fe6	-3.84
236	0.5	Fe6	-3.30
127	0.5	Fe6	-3.07
346	2.5	Fe6	-3.02
235	1.5	Fe6	-1.24
346	0.5	Fe6	-0.41
346	1.5	Fe6	0.49
247	1.5	Fe6	1.19
146	2.5	Fe6	2.15
147	1.5	Fe6	4.05
346	0.5	Fe2	0.41
147	0.5	Fe2	0.48
235	2.5	Fe2	2.25
346	1.5	Fe2	6.49
235	1.5	Fe2	8.80
147	1.5	Fe2	9.85
247	1.5	Fe2	10.83
247	0.5	Fe2	17.08
235	0.5	Fe2	18.87

BS	M_S	binding site	ΔE_{bnd}
235	0.5	Fe6	-8.14
147	0.5	Fe6	-6.63
345	0.5	Fe6	-5.73
145	0.5	Fe6	-5.16
135	0.5	Fe6	-4.74
236	2.5	Fe6	-2.05
235	1.5	Fe6	0.85
147	1.5	Fe6	4.15
247	0.5	Fe6	9.00
346	0.5	Fe6	11.15
346	1.5	Fe6	11.97
235	2.5	Fe2	6.88
247	2.5	Fe2	7.56
247	0.5	Fe2	8.62
346	0.5	Fe2	10.50
147	0.5	Fe2	11.12
235	1.5	Fe2	13.73
346	1.5	Fe2	17.93
247	1.5	Fe2	18.59
235	0.5	Fe2	19.53
147	1.5	Fe2	21.26

Table A16. – CO binding energies [kcal/mol] for the E_1 FeMoco (left) and FeVco (right) QM2 model.

BS	M_S	binding site	ΔE_{bnd}
346	2	Fe6	-8.32
147	0	Fe6	-7.31
346	3	Fe6	-6.01
235	1	Fe6	-6.00
135	0	Fe6	-5.83
247	1	Fe6	-5.47
127	0	Fe6	-5.43
125	0	Fe6	-5.41
235	2	Fe6	-0.84
247	2	Fe6	-0.50
147	1	Fe6	7.54
346	1	Fe6	7.96
147	2	Fe6	12.90
147	0	Fe2	-5.05
235	2	Fe2	-4.91
235	3	Fe2	-2.54
147	1	Fe2	-1.43
235	1	Fe2	0.56
346	1	Fe2	0.90
247	2	Fe2	1.41
346	2	Fe2	1.69
247	1	Fe2	3.18
147	2	Fe2	11.27

BS	M_S	binding site	ΔE_{bnd}
235	0	Fe6	-8.47
235	1	Fe6	-7.55
147	0	Fe6	21.62
135	0	Fe6	22.09
346	2	Fe6	22.11
147	1	Fe6	23.02
127	0	Fe6	23.25
346	3	Fe6	24.02
125	0	Fe6	24.60
135	1	Fe6	27.52
147	0	Fe2	-7.36
135	0	Fe2	-5.49
235	2	Fe2	21.02
235	3	Fe2	23.74
147	1	Fe2	27.36
235	1	Fe2	28.39
247	2	Fe2	28.90
346	1	Fe2	29.85
247	3	Fe2	31.02
346	2	Fe2	33.84
346	0	Fe2	33.92
147	2	Fe2	38.49
247	1	Fe2	40.32

A4.4. CO frequencies

Table A17. – Calculated CO frequency for different partial Hessians for the E_1 QM/MM FeMoco model (CO bound to Fe6, BS7-247, $M_S = 1$).

unfrozen atoms	$\nu_{\text{calculated}}$	ν_{scaled}
CO+Fe	1978.1	1943.6
CO+Fe+3S+carbide	1978.1	1943.6
CO+6Fe+Mo+3S+carbide	1977.1	1943.4

Table A18. – As-calculated, unscaled CO frequencies and scaled frequencies [cm^{-1}] for all CO-bound models. The scaling factor (0.9825) is equal to the ratio of the as-calculated frequency of free CO (2143 cm^{-1}) and the experimentally determined frequency CO gas (2181 cm^{-1}).[225]

binding site	BS	M_S	$\nu_{calculated}$	ν_{scaled}
diamagnetically substituted FeMoco				
CO−Fe^{3+}		1.5	2013	1978
CO−Fe^{2+}		1.0	1918	1884
CO−Fe^{1+}		0.5	1803	1772
CO−Fe^{2+}()Fe^{3+}()		3.5	1942	1908
CO−Fe^{2+}(v)Fe^{3+}()		1.5	1951	1917
E_0 FeMoco cluster				
Fe6	BS10-135	0.5	1943	1910
Fe2	BS10-147	0.5	1929	1895
E_1 FeMoco cluster				
Fe6	BS10-147	0	1854	1822
Fe2	BS10-147	0	1889	1856
E_0 FeVco cluster				
Fe6	BS8-345	0.5	1921	1887
Fe2	BS7-235	2.5	1927	1894
E_1 FeVco cluster				
Fe6	BS10-135	0	1831	1799
Fe2	BS10-135	0	1842	1810
E_0 FeMoco QM/MM				
Fe6	BS10-147	0.5	2001	1966
Fe2	BS7-346	0.5	1992	1957
E_1 FeMoco QM/MM				
Fe6	BS7-346	2	1956	1922
Fe2/Fe6	BS10-147	0	1746	1716
E_1 FeMoco QM/MM Val2Ile				
Fe6	BS7-346	2	1939	1905
E_0 FeVco QM/MM				
Fe6	BS7-235	0.5	1958	1924
Fe2	BS7-235	2.5	1992	1957
E_1 FeVco QM/MM				
Fe6	BS7-235	0	1955	1921
Fe2/Fe6	BS10-147	0	1752	1721
E_2 FeMoco QM/MM				
Fe2/Fe6	BS7-235	0.5	1748	1718
Fe2/Fe6	BS7-147	0.5	1767	1738

Bibliography

(1) Rehder, D., *Bioinorganic Chemistry: An Introduction*; Oxford University Press: Oxford, 2014; 230 pp., ISBN: 978-0-19-965519-9.

(2) Zhang, X.; Ward, B. B.; Sigman, D. M. Global Nitrogen Cycle: Critical Enzymes, Organisms, and Processes for Nitrogen Budgets and Dynamics. *Chem. Rev.* **2020**, *120*, 5308–5351, DOI: 10.1021/acs.chemrev.9b00613.

(3) Rong Li; Chen, H.; Yang, Z.; Yuan, S.; Zhou, X. Research status of soybean symbiosis nitrogen fixation. *Oil Crop Science* **2020**, *5*, 6–10, DOI: 10.1016/j.ocsci.2020.03.005.

(4) Haber, F. Über die Darstellung des Ammoniaks aus Stickstoff und Wasserstoff. *Naturwissenschaften* **1922**, *10*, 1041–1049, DOI: 10.1007/BF01565394.

(5) Haber, F. Neue Arbeitsweisen. *Naturwissenschaften* **1923**, *11*, 753–756, DOI: 10.1007/BF01551097.

(6) Fowler, D. *et al.* Effects of global change during the 21st century on the nitrogen cycle. *Atmospheric Chemistry and Physics* **2015**, *15*, 13849–13893, DOI: 10.5194/acp-15-13849-2015.

(7) Humphreys, J.; Lan, R.; Tao, S. Development and Recent Progress on Ammonia Synthesis Catalysts for Haber–Bosch Process. *Advanced Energy and Sustainability Research* **2021**, *2*, 2000043, DOI: 10.1002/aesr.202000043.

(8) Smith, C.; K. Hill, A.; Torrente-Murciano, L. Current and future role of Haber–Bosch ammonia in a carbon-free energy landscape. *Energy & Environmental Science* **2020**, *13*, 331–344, DOI: 10.1039/C9EE02873K.

(9) Kyriakou, V.; Garagounis, I.; Vourros, A.; Vasileiou, E.; Stoukides, M. An Electrochemical Haber-Bosch Process. *Joule* **2020**, *4*, 142–158, DOI: 10.1016/j.joule.2019.10.006.

(10) Liew, F. E. *et al.* Carbon-negative production of acetone and isopropanol by gas fermentation at industrial pilot scale. *Nat Biotechnol* **2022**, *40*, 335–344, DOI: 10.1038/s41587-021-01195-w.

(11) Seefeldt, L. C.; Yang, Z.-Y.; Lukoyanov, D. A.; Harris, D. F.; Dean, D. R.; Raugei, S.; Hoffman, B. M. Reduction of Substrates by Nitrogenases. *Chem. Rev.* **2020**, *120*, 5082–5106, DOI: 10.1021/acs.chemrev.9b00556.

(12) Oehlmann, N. N.; Rebelein, J. G. The Conversion of Carbon Monoxide and Carbon Dioxide by Nitrogenases. *ChemBioChem* **2021**, DOI: 10.1002/cbic.202100453.

(13) Jasniewski, A. J.; Lee, C. C.; Ribbe, M. W.; Hu, Y. Reactivity, Mechanism, and Assembly of the Alternative Nitrogenases. *Chem. Rev.* **2020**, *120*, 5107–5157, DOI: 10.1021/acs.chemrev.9b00704.

(14) Thorneley, R. N. F.; Lowe, D. J. The mechanism of Klebsiella pneumoniae nitrogenase action. Pre-steady-state kinetics of an enzyme-bound intermediate in N2 reduction and of NH3 formation. *Biochemical Journal* **1984**, *224*, 887–894, DOI: 10.1042/bj2240887.

(15) Van Stappen, C.; Decamps, L.; Cutsail, G. E.; Bjornsson, R.; Henthorn, J. T.; Birrell, J. A.; DeBeer, S. The Spectroscopy of Nitrogenases. *Chem. Rev.* **2020**, *120*, 5005–5081, DOI: 10.1021/acs.chemrev.9b00650.

(16) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angewandte Chemie International Edition* **2009**, *48*, 1198–1229, DOI: 10.1002/anie.200802019.

(17) Benediktsson, B.; Bjornsson, R. QMMM Study of the Nitrogenase MoFe Protein Resting State: Broken-Symmetry States, Protonation States, and QM Region Convergence in the FeMoco Active Site. *Inorg. Chem.* **2017**, *56*, 13417–13429, DOI: 10.1021/acs.inorgchem.7b02158.

(18) Benediktsson, B.; Thorhallsson, A. T.; Bjornsson, R. QMMM calculations reveal a bridging hydroxo group in a vanadium nitrogenase crystal structure. *Chem. Commun.* **2018**, DOI: 10.1039/C8CC03793K.

(19) Benediktsson, B.; Bjornsson, R. Quantum Mechanics/Molecular Mechanics Study of Resting-State Vanadium Nitrogenase: Molecular and Electronic Structure of the Iron–Vanadium Cofactor. *Inorg. Chem.* **2020**, *59*, 11514–11527, DOI: 10.1021/acs.inorgchem.0c01320.

(20) Neese, F. A critical evaluation of DFT, including time-dependent DFT, applied to bioinorganic chemistry. *J Biol Inorg Chem* **2006**, *11*, 702–711, DOI: 10.1007/s00775-006-0138-1.

(21) Chilkuri, V. G.; DeBeer, S.; Neese, F. Ligand Field Theory and Angular Overlap Model Based Analysis of the Electronic Structure of Homovalent Iron–Sulfur Dimers. *Inorg. Chem.* **2020**, *59*, 984–995, DOI: 10.1021/acs.inorgchem.9b00974.

(22) Spiller, N.; Chilkuri, V. G.; DeBeer, S.; Neese, F. Sulfur vs. Selenium as Bridging Ligand in Di-Iron Complexes: A Theoretical Analysis. *European Journal of Inorganic Chemistry* **2020**, *2020*, 1525–1538, DOI: https://doi.org/10.1002/ejic.202000033.

(23) Henthorn, J. T.; Cutsail, G. E.; Weyhermüller, T.; DeBeer, S. Stabilization of intermediate spin states in mixed-valent diiron dichalcogenide complexes. *Nat. Chem.* **2022**, 1–6, DOI: 10.1038/s41557-021-00853-5.

(24) Subramanian, S.; Duin, E. C.; Fawcett, S. E. J.; Armstrong, F. A.; Meyer, J.; Johnson, M. K. Spectroscopic and Redox Studies of Valence-Delocalized [Fe2S2]+ Centers in Thioredoxin-like Ferredoxins. *J. Am. Chem. Soc.* **2015**, *137*, 4567–4580, DOI: 10.1021/jacs.5b01869.

(25) Spiller, N.; Bjornsson, R.; DeBeer, S.; Neese, F. Carbon Monoxide Binding to the Iron–Molybdenum Cofactor of Nitrogenase: a Detailed Quantum Mechanics/Molecular Mechanics Investigation. *Inorg. Chem.* **2021**, *60*, 18031–18047, DOI: 10.1021/acs.inorgchem.1c02649.

(26) Szabo, A.; Ostlund, N. S., *Modern quantum chemistry : introduction to advanced electronic structure theory*; Dover Publications: 1996; 466 pp., ISBN: 978-0-486-69186-2.

(27) Jensen, F., *Introduction to Computational Chemistry*, 2nd edition; JW: Chichester, England ; Hoboken, NJ, 2011; 620 pp., ISBN: 978-0-470-01187-4.

(28) Helgaker, T.; Jørgensen, P.; Olsen, J., *Molecular Electronic-Structure Theory: Helgaker/Molecular Electronic-Structure Theory*; John Wiley & Sons, Ltd: Chichester, UK, 2000, ISBN: 978-1-119-01957-2, DOI: 10.1002/9781119019572.

(29) Chattopadhyay, S.; Chaudhuri, R. K.; Mahapatra, U. S.; Ghosh, A.; Ray, S. S. State-specific multireference perturbation theory: development and present status. *WIREs Computational Molecular Science* **2016**, *6*, 266–291, DOI: 10.1002/wcms.1248.

(30) Buenker, R. J.; Peyerimhoff, S. D. Individualized configuration selection in CI calculations with subsequent energy extrapolation. *Theoret. Chim. Acta* **1974**, *35*, 33–58, DOI: 10.1007/PL00020553.

(31) Buenker, R. J.; Peyerimhoff, S. D. Energy extrapolation in CI calculations. *Theoret. Chim. Acta* **1975**, *39*, 217–228, DOI: 10.1007/BF00555301.

(32) Bartlett, R. J.; Musiał, M. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* **2007**, *79*, 291–352, DOI: 10.1103/RevModPhys.79.291.

(33) Köhn, A.; Hanauer, M.; Mück, L. A.; Jagau, T.-C.; Gauss, J. State-specific multireference coupled-cluster theory. *WIREs Computational Molecular Science* **2013**, *3*, 176–197, DOI: 10.1002/wcms.1120.

(34) Angeli, C.; Cimiraglia, R.; Malrieu, J.-P. N-electron valence state perturbation theory: a fast implementation of the strongly contracted variant. *Chemical Physics Letters* **2001**, *350*, 297–305, DOI: 10.1016/S0009-2614(01)01303-3.

(35) Angeli, C.; Cimiraglia, R.; Evangelisti, S.; Leininger, T.; Malrieu, J.-P. Introduction of n-electron valence states for multireference perturbation theory. *J. Chem. Phys.* **2001**, *114*, 10252–10264, DOI: 10.1063/1.1361246.

(36) Angeli, C.; Cimiraglia, R.; Malrieu, J.-P. n-electron valence state perturbation theory: A spinless formulation and an efficient implementation of the strongly contracted and of the partially contracted variants. *J. Chem. Phys.* **2002**, *117*, 9138–9153, DOI: 10.1063/1.1515317.

(37) Angeli, C.; Cimiraglia, R. Multireference perturbation CI IV. Selection procedure for one-electron properties. *Theor Chem Acc* **2001**, *105*, 259–264, DOI: 10.1007/s002140000212.

(38) Andersson, K.; Malmqvist, P. A.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. Second-order perturbation theory with a CASSCF reference function. *J. Phys. Chem.* **1990**, *94*, 5483–5488, DOI: 10.1021/j100377a012.

(39) Koch, W.; Holthausen, M. C., *A Chemist's Guide to Density Functional Theory*, 1st ed.; Wiley: 2001, ISBN: 978-3-527-60004-5, DOI: 10.1002/3527600043.

(40) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136*, B864–B871, DOI: 10.1103/PhysRev.136.B864.

(41) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, A1133–A1138, DOI: 10.1103/PhysRev.140.A1133.

(42) Medvedev, M. G.; Bushmarinov, I. S.; Sun, J.; Perdew, J. P.; Lyssenko, K. A. Density functional theory is straying from the path toward the exact functional. *Science* **2017**, *355*, 49–52, DOI: 10.1126/science.aah5975.

(43) Goerigk, L.; Mehta, N. A Trip to the Density Functional Theory Zoo: Warnings and Recommendations for the User. *Aust. J. Chem.* **2019**, *72*, 563–573, DOI: 10.1071/CH19023.

(44) Perdew, J. P.; Schmidt, K. Jacob's ladder of density functional approximations for the exchange-correlation energy. *AIP Conference Proceedings* **2001**, *577*, 1–20, DOI: 10.1063/1.1390175.

(45) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys.* **1980**, *58*, 1200–1211, DOI: 10.1139/p80-159.

(46) Becke, A. D. Density functional calculations of molecular bond energies. *J. Chem. Phys.* **1986**, *84*, 4524–4529, DOI: 10.1063/1.450025.

(47) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785–789, DOI: 10.1103/PhysRevB.37.785.

(48) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the Density Functional Ladder: Nonempirical Meta–Generalized Gradient Approximation Designed for Molecules and Solids. *Phys. Rev. Lett.* **2003**, *91*, 146401, DOI: 10.1103/PhysRevLett.91.146401.

(49) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, *98*, 11623–11627, DOI: 10.1021/j100096a001.

(50) Staroverov, V. N.; Scuseria, G. E.; Tao, J.; Perdew, J. P. Comparative assessment of a new nonempirical density functional: Molecules and hydrogen-bonded complexes. *J. Chem. Phys.* **2003**, *119*, 12129–12137, DOI: 10.1063/1.1626543.

(51) Szilagyi, R. K.; Winslow, M. A. On the accuracy of density functional theory for iron—sulfur clusters. *Journal of Computational Chemistry* **2006**, *27*, 1385–1397, DOI: 10.1002/jcc.20449.

(52) Sandala, G. M.; Hopmann, K. H.; Ghosh, A.; Noodleman, L. Calibration of DFT Functionals for the Prediction of 57Fe Mössbauer Spectral Parameters in Iron–Nitrosyl and Iron–Sulfur Complexes: Accurate Geometries Prove Essential. *J. Chem. Theory Comput.* **2011**, *7*, 3232–3247, DOI: 10.1021/ct200187d.

(53) Harris, T. V.; Szilagyi, R. K. Iron–sulfur bond covalency from electronic structure calculations for classical iron–sulfur clusters. *Journal of Computational Chemistry* **2014**, *35*, 540–552, DOI: 10.1002/jcc.23518.

(54) Cao, L.; Ryde, U. Extremely large differences in DFT energies for nitrogenase models. *Physical Chemistry Chemical Physics* **2019**, *21*, 2480–2488, DOI: 10.1039/C8CP06930A.

(55) Goerigk, L.; Grimme, S. Double-hybrid density functionals. *WIREs Computational Molecular Science* **2014**, *4*, 576–600, DOI: 10.1002/wcms.1193.

(56) Sun, J.; Ruzsinszky, A.; Perdew, J. P. Strongly Constrained and Appropriately Normed Semilocal Density Functional. *Phys. Rev. Lett.* **2015**, *115*, 036402, DOI: 10.1103/PhysRevLett.115.036402.

(57) Bartók, A. P.; Yates, J. R. Regularized SCAN functional. *J. Chem. Phys.* **2019**, *150*, 161101, DOI: 10.1063/1.5094646.

(58) Furness, J. W.; Kaplan, A. D.; Ning, J.; Perdew, J. P.; Sun, J. Accurate and Numerically Efficient r2SCAN Meta-Generalized Gradient Approximation. *J. Phys. Chem. Lett.* **2020**, *11*, 8208–8215, DOI: 10.1021/acs.jpclett.0c02405.

(59) Kirkpatrick, J. *et al.* Pushing the frontiers of density functionals by solving the fractional electron problem. *Science* **2021**, *374*, 1385–1389, DOI: 10.1126/science.abj6511.

(60) Atkins, P. W.; Friedman, R. S., *Molecular Quantum Mechanics*, 5th edition; Oxford University Press: Oxford ; New York, 2010; 560 pp., ISBN: 978-0-19-954142-3.

(61) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104, DOI: 10.1063/1.3382344.

(62) Grimme, S. Accurate description of van der Waals complexes by density functional theory including empirical corrections. *Journal of Computational Chemistry* **2004**, *25*, 1463–1473, DOI: 10.1002/jcc.20078.

(63) Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *Journal of Computational Chemistry* **2006**, *27*, 1787–1799, DOI: 10.1002/jcc.20495.

(64) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *Journal of Computational Chemistry* **2011**, *32*, 1456–1465, DOI: https://doi.org/10.1002/jcc.21759.

(65) Caldeweyher, E.; Bannwarth, C.; Grimme, S. Extension of the D3 dispersion coefficient model. *J. Chem. Phys.* **2017**, *147*, 034112, DOI: 10.1063/1.4993215.

(66) Grimme, S. Supramolecular Binding Thermodynamics by Dispersion-Corrected Density Functional Theory. *Chemistry – A European Journal* **2012**, *18*, 9955–9964, DOI: 10.1002/chem.201200497.

(67) Lee, K.; Murray, É. D.; Kong, L.; Lundqvist, B. I.; Langreth, D. C. Higher-accuracy van der Waals density functional. *Phys. Rev. B* **2010**, *82*, 081101, DOI: 10.1103/PhysRevB.82.081101.

(68) Vydrov, O. A.; Van Voorhis, T. Nonlocal van der Waals density functional: The simpler the better. *J. Chem. Phys.* **2010**, *133*, 244103, DOI: 10.1063/1.3521275.

(69) Spatzal, T.; Aksoyoglu, M.; Zhang, L.; Andrade, S. L. A.; Schleicher, E.; Weber, S.; Rees, D. C.; Einsle, O. Evidence for Interstitial Carbon in Nitrogenase FeMo Cofactor. *Science* **2011**, *334*, 940–940, DOI: 10.1126/science.1214025.

(70) Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *Journal of Molecular Biology* **1976**, *103*, 227–249, DOI: 10.1016/0022-2836(76)90311-9.

(71) Senn, H. M.; Thiel, S.; Thiel, W. Enzymatic Hydroxylation in p-Hydroxybenzoate Hydroxylase: A Case Study for QM/MM Molecular Dynamics. *J. Chem. Theory Comput.* **2005**, *1*, 494–505, DOI: 10.1021/ct049844p.

(72) Zou, Y.; Houk, K. N. Mechanisms and Dynamics of Synthetic and Biosynthetic Formation of Delitschiapyrones: Solvent Control of Ambimodal Periselectivity. *J. Am. Chem. Soc.* **2021**, *143*, 11734–11740, DOI: 10.1021/jacs.1c05293.

(73) Bussi, G.; Laio, A. Using metadynamics to explore complex free-energy landscapes. *Nat Rev Phys* **2020**, *2*, 200–212, DOI: 10.1038/s42254-020-0153-0.

(74) Huang, J.; MacKerell, A. D. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J. Comput. Chem.* **2013**, *34*, 2135–2145, DOI: 10.1002/jcc.23354.

(75) Neese, F. The ORCA program system. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2012**, *2*, 73–78, DOI: 10.1002/wcms.81.

(76) Neese, F. Software update: the ORCA program system, version 4.0. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2018**, *8*, e1327, DOI: 10.1002/wcms.1327.

(77) Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The ORCA quantum chemistry program package. *J. Chem. Phys.* **2020**, *152*, 224108, DOI: 10.1063/5.0004608.

(78) Metz, S.; Kästner, J.; Sokol, A. A.; Keal, T. W.; Sherwood, P. ChemShell—a modular software package for QM/MM simulations. *WIREs Computational Molecular Science* **2014**, *4*, 101–110, DOI: https://doi.org/10.1002/wcms.1163.

(79) Sherwood, P. *et al.* QUASI: A general purpose implementation of the QM/MM approach and its application to problems in catalysis. *Journal of Molecular Structure: THEOCHEM* **2003**, *632*, 1–28, DOI: 10.1016/S0166-1280(03)00285-9.

(80) Case, D. A.; Cheatham III, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz Jr., K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *Journal of Computational Chemistry* **2005**, *26*, 1668–1688, DOI: 10.1002/jcc.20290.

(81) Scott, W. R. P.; Hünenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Krüger, P.; van Gunsteren, W. F. The GROMOS Biomolecular Simulation Program Package. *J. Phys. Chem. A* **1999**, *103*, 3596–3607, DOI: 10.1021/jp984217f.

(82) Robertson, M. J.; Tirado-Rives, J.; Jorgensen, W. L. Improved Peptide and Protein Torsional Energetics with the OPLS-AA Force Field. *J. Chem. Theory Comput.* **2015**, *11*, 3499–3509, DOI: 10.1021/acs.jctc.5b00356.

(83) Lamoureux, G.; Roux, B. Modeling induced polarization with classical Drude oscillators: Theory and molecular dynamics simulation algorithm. *J. Chem. Phys.* **2003**, *119*, 3025–3039, DOI: 10.1063/1.1589749.

(84) Jing, Z.; Liu, C.; Cheng, S. Y.; Qi, R.; Walker, B. D.; Piquemal, J.-P.; Ren, P. Polarizable Force Fields for Biomolecular Simulations: Recent Advances and Applications. *Annu. Rev. Biophys.* **2019**, *48*, 371–394, DOI: 10.1146/annurev-biophys-070317-033349.

(85) MacKerell, A. D.; Feig, M.; Brooks, C. L. Improved Treatment of the Protein Backbone in Empirical Force Fields. *J. Am. Chem. Soc.* **2004**, *126*, 698–699, DOI: 10.1021/ja036959e.

(86) Mackerell Jr., A. D.; Feig, M.; Brooks III, C. L. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *Journal of Computational Chemistry* **2004**, *25*, 1400–1415, DOI: 10.1002/jcc.20065.

(87) Vangaveti, S.; Ranganathan, S. V.; Chen, A. A. Advances in RNA molecular dynamics: a simulator's guide to RNA force fields. *WIREs RNA* **2017**, *8*, e1396, DOI: 10.1002/wrna.1396.

(88) Yildirim, I. In *Biomolecular Simulations: Methods and Protocols*, Bonomi, M., Camilloni, C., Eds.; Methods in Molecular Biology; Springer: New York, NY, 2019, pp 55–74, DOI: 10.1007/978-1-4939-9608-7_3.

(89) McWeeny, R., *Spins in Chemistry*, 2004 edition; Dover Publications: Mineola, N.Y, 2004; 176 pp., ISBN: 978-0-486-43486-5.

(90) Hagen, W. R., *Biomolecular EPR Spectroscopy*; CRC Press: 2008; 264 pp., ISBN: 978-1-4200-5958-8.

(91) Mostafanejad, M. Basics of the spin Hamiltonian formalism. *International Journal of Quantum Chemistry* **2014**, *114*, 1495–1512, DOI: 10.1002/qua.24721.

(92) Neese, F. In *eMagRes*; American Cancer Society: 2017, pp 1–22, DOI: 10.1002/9780470034590.emrstm1505.

(93) Heisenberg, W. In *Zur Theorie des Ferromagnetismus*, Blum, W., Rechenberg, H., Dürr, H.-P., Eds.; Werner Heisenberg Gesammelte Werke Collected Works; Springer: Berlin, Heidelberg, 1985, pp 580–597, DOI: 10.1007/978-3-642-61659-4_35.

(94) Dirac, P. A. M.; Fowler, R. H. On the theory of quantum mechanics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **1926**, *112*, 661–677, DOI: 10.1098/rspa.1926.0133.

(95) Van Vleck, J. H. The Theory of Electric and Magnetic Susceptibilities. *Clarendon Press* **1932**, *130*, 490–491, DOI: 10.1038/130490a0.

(96) Moriya, T. Anisotropic Superexchange Interaction and Weak Ferromagnetism. *Phys. Rev.* **1960**, *120*, 91–98, DOI: 10.1103/PhysRev.120.91.

(97) Anderson, P. W. In *Solid State Physics*, Seitz, F., Turnbull, D., Eds.; Academic Press: 1963; Vol. 14, pp 99–214, DOI: 10.1016/S0081-1947(08)60260-X.

(98) Beinert, H.; Holm, R. H.; Münck, E. Iron-Sulfur Clusters: Nature's Modular, Multipurpose Structures. *Science* **1997**, *277*, 653–659, DOI: 10.1126/science.277.5326.653.

(99) Falk, U.; Furrer, A.; Kjems, J. K.; Güdel, H. U. Biquadratic Exchange in CsMnxMg1-xBr3. *Phys. Rev. Lett.* **1984**, *52*, 1336–1339, DOI: 10.1103/PhysRevLett.52.1336.

(100) Landee, C. P.; Turnbull, M. M. Review: A gentle introduction to magnetism: units, fields, theory, and experiment. *Journal of Coordination Chemistry* **2014**, *67*, 375–439, DOI: 10.1080/00958972.2014.889294.

(101) Zener, C. Interaction between the d-Shells in the Transition Metals. II. Ferromagnetic Compounds of Manganese with Perovskite Structure. *Phys. Rev.* **1951**, *82*, 403–405, DOI: 10.1103/PhysRev.82.403.

(102) Anderson, P. W.; Hasegawa, H. Considerations on Double Exchange. *Phys. Rev.* **1955**, *100*, 675–681, DOI: 10.1103/PhysRev.100.675.

(103) Robin, M. B.; Day, P. In *Molecules into Materials*; WORLD SCIENTIFIC: 2007, pp 142–317, DOI: 10.1142/9789812706836_0016.

(104) Kahn, O., *Molecular magnetism*; VCH: New York, NY, 1993; 380 pp., ISBN: 978-1-56081-566-2.

(105) Piepho, S. B.; Krausz, E. R.; Schatz, P. N. Vibronic coupling model for calculation of mixed valence absorption profiles. *J. Am. Chem. Soc.* **1978**, *100*, 2996–3005, DOI: 10.1021/ja00478a011.

(106) Azumi, T.; Matsuzaki, K. What Does the Term "Vibronic Coupling" Mean? *Photochemistry and Photobiology* **1977**, *25*, 315–326, DOI: 10.1111/j.1751-1097.1977.tb06918.x.

(107) Blondin, G.; Girerd, J. J. Interplay of electron exchange and electron transfer in metal polynuclear complexes in proteins or chemical models. *Chem. Rev.* **1990**, *90*, 1359–1376, DOI: 10.1021/cr00106a001.

(108) Sharma, S.; Sivalingam, K.; Neese, F.; Chan, G. K.-L. Low-energy spectrum of iron–sulfur clusters directly from many-particle quantum mechanics. *Nature Chemistry* **2014**, *6*, 927–933, DOI: 10.1038/nchem.2041.

(109) Malrieu, J. P.; Caballol, R.; Calzado, C. J.; de Graaf, C.; Guihéry, N. Magnetic Interactions in Molecules and Highly Correlated Materials: Physical Content, Analytical Derivation, and Rigorous Extraction of Magnetic Hamiltonians. *Chem. Rev.* **2014**, *114*, 429–492, DOI: 10.1021/cr300500z.

(110) Yamaguchi, K.; Takahara, Y.; Fueno, T. In *Applied Quantum Chemistry*, ed. by Smith, V. H.; Schaefer, H. F.; Morokuma, K., Springer Netherlands: Dordrecht, 1986, pp 155–184, DOI: 10.1007/978-94-009-4746-7_11.

(111) Soda, T.; Kitagawa, Y.; Onishi, T.; Takano, Y.; Shigeta, Y.; Nagao, H.; Yoshioka, Y.; Yamaguchi, K. Ab initio computations of effective exchange integrals for H–H, H–He–H and Mn2O2 complex: comparison of broken-symmetry approaches. *Chemical Physics Letters* **2000**, *319*, 223–230, DOI: 10.1016/S0009-2614(00)00166-4.

(112) Costa, R.; Valero, R.; Mañeru, D. R.; Moreira, I. d. P. R.; Illas, F. Spin Adapted versus Broken Symmetry Approaches in the Description of Magnetic Coupling in Heterodinuclear Complexes. *J. Chem. Theory Comput.* **2015**, *11*, 1006–1019, DOI: 10.1021/ct5011483.

(113) Mulliken, R. S. Criteria for the Construction of Good Self-Consistent-Field Molecular Orbital Wave Functions, and the Significance of LCAO-MO Population Analysis. *J. Chem. Phys.* **1962**, *36*, 3428–3439, DOI: 10.1063/1.1732476.

(114) Löwdin, P.-O. In *Advances in Quantum Chemistry*, Löwdin, P.-O., Ed.; Academic Press: 1970; Vol. 5, pp 185–199, DOI: 10.1016/S0065-3276(08)60339-1.

(115) Politzer, P.; Harris, R. R. Properties of atoms in molecules. I. Proposed definition of the charge on an atom in a molecule. *J. Am. Chem. Soc.* **1970**, *92*, 6451–6454, DOI: 10.1021/ja00725a011.

(116) Bader, R. F. W.; Carroll, M. T.; Cheeseman, J. R.; Chang, C. Properties of atoms in molecules: atomic volumes. *J. Am. Chem. Soc.* **1987**, *109*, 7968–7979, DOI: 10.1021/ja00260a006.

(117) Hirshfeld, F. L. Bonded-atom fragments for describing molecular charge densities. *Theoret. Chim. Acta* **1977**, *44*, 129–138, DOI: 10.1007/BF00549096.

(118) Rousseau, B.; Peeters, A.; Van Alsenoy, C. Atomic charges from modified Voronoi polyhedra. *Journal of Molecular Structure: THEOCHEM* **2001**, *538*, 235–238, DOI: 10.1016/S0166-1280(00)00692-8.

(119) Matta, C. F.; Boyd, R. J. In *The Quantum Theory of Atoms in Molecules*, Matta, C. F., Boyd, R. J., Eds.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2007, pp 1–34, DOI: 10.1002/9783527610709.ch1.

(120) Bader, R. F. W.; Platts, J. A. Characterization of an F-center in an alkali halide cluster. *J. Chem. Phys.* **1997**, *107*, 9, DOI: 10.1063/1.3393036.

(121) Timerghazin, Q. K.; Peslherbe, G. H. Non-nuclear attractor of electron density as a manifestation of the solvated electron. *The Journal of Chemical Physics* **2007**, *127*, 064108, DOI: 10.1063/1.2747250.

(122) Guerra, C. F.; Handgraaf, J.-W.; Baerends, E. J.; Bickelhaupt, F. M. Voronoi deformation density (VDD) charges: Assessment of the Mulliken, Bader, Hirshfeld, Weinhold, and VDD methods for charge analysis. *Journal of Computational Chemistry* **2004**, *25*, 189–210, DOI: https://doi.org/10.1002/jcc.10351.

(123) Bultinck, P.; Van Alsenoy, C.; Ayers, P. W.; Carbó-Dorca, R. Critical analysis and extension of the Hirshfeld atoms in molecules. *The Journal of Chemical Physics* **2007**, *126*, 144111, DOI: 10.1063/1.2715563.

(124) Marenich, A. V.; Jerome, S. V.; Cramer, C. J.; Truhlar, D. G. Charge Model 5: An Extension of Hirshfeld Population Analysis for the Accurate Description of Molecular Interactions in Gaseous and Condensed Phases. *J. Chem. Theory Comput.* **2012**, *8*, 527–541, DOI: 10.1021/ct200866d.

(125) Mayer, I. Charge, bond order and valence in the AB initio SCF theory. *Chemical Physics Letters* **1983**, *97*, 270–274, DOI: 10.1016/0009-2614(83)80005-0.

(126) Matito, E.; Poater, J.; Solà, M.; Duran, M.; Salvador, P. Comparison of the AIM Delocalization Index and the Mayer and Fuzzy Atom Bond Orders. *J. Phys. Chem. A* **2005**, *109*, 9904–9910, DOI: 10.1021/jp0538464.

(127) Mayer, I.; Salvador, P. Overlap populations, bond orders and valences for 'fuzzy' atoms. *Chemical Physics Letters* **2004**, *383*, 368–375, DOI: 10.1016/j.cplett.2003.11.048.

(128) Foster, J. M. Canonical Configurational Interaction Procedure. *Rev. Mod. Phys.* **1960**, *32*, 300–302, DOI: 10.1103/RevModPhys.32.300.

(129) Edmiston, C.; Ruedenberg, K. Localized Atomic and Molecular Orbitals. II. *J. Chem. Phys.* **1965**, *43*, S97–S116, DOI: 10.1063/1.1701520.

(130) Von Niessen, W. Density Localization of Atomic and Molecular Orbitals. I. *J. Chem. Phys.* **1972**, *56*, 4290–4297, DOI: 10.1063/1.1677859.

(131) Pipek, J.; Mezey, P. G. A fast intrinsic localization procedure applicable for ab initio and semiempirical linear combination of atomic orbital wave functions. *J. Chem. Phys.* **1989**, *90*, 4916–4926, DOI: 10.1063/1.456588.

(132) Lehtola, S.; Jónsson, H. Pipek–Mezey Orbital Localization Using Various Partial Charge Estimates. *J. Chem. Theory Comput.* **2014**, *10*, 642–649, DOI: 10.1021/ct401016x.

(133) Noodleman, L.; Lovell, T.; Liu, T.; Himo, F.; Torres, R. A. Insights into properties and energetics of iron–sulfur proteins from simple clusters to nitrogenase. *Current Opinion in Chemical Biology* **2002**, *6*, 259–273, DOI: 10.1016/S1367-5931(02)00309-5.

(134) Johnson, D. C.; Dean, D. R.; Smith, A. D.; Johnson, M. K. STRUCTURE, FUNCTION, AND FORMATION OF BIOLOGICAL IRON-SULFUR CLUSTERS. *Annu. Rev. Biochem.* **2005**, *74*, 247–281, DOI: 10.1146/annurev.biochem.74.082803.133518.

(135) Lunghi, A.; Totti, F. The Role of Anisotropic Exchange in Single Molecule Magnets: A CASSCF/NEVPT2 Study of the Fe4 SMM Building Block [Fe2(OCH3)2(dbm)4] Dimer. *Inorganics* **2016**, *4*, 28, DOI: 10.3390/inorganics4040028.

(136) Li, Z.; Guo, S.; Sun, Q.; Chan, G. K.-L. Electronic landscape of the P-cluster of nitrogenase as revealed through many-electron quantum wavefunction simulations. *Nat. Chem.* **2019**, *11*, 1026–1033, DOI: 10.1038/s41557-019-0337-3.

(137) Li, Z.; Li, J.; Dattani, N. S.; Umrigar, C. J.; Chan, G. K.-L. The electronic complexity of the ground-state of the FeMo cofactor of nitrogenase as relevant to quantum simulations. *J. Chem. Phys.* **2019**, *150*, 024302, DOI: 10.1063/1.5063376.

(138) Brabec, J.; Brandejs, J.; Kowalski, K.; Xantheas, S.; Legeza, Ö.; Veis, L. Massively parallel quantum chemical density matrix renormalization group method. *arXiv:2001.04890 [physics.chem-ph] (Preprint)* **2020**, DOI: 10.48550/arXiv.2001.04890.

(139) Gade, L. H., *Koordinationschemie*, 1.; Wiley: Weinheim, 2010; 582 pp., ISBN: 978-3-527-29503-6.

(140) Eaton, W. A.; Palmer, G.; Fee, J. A.; Kimura, T.; Lovenberg, W. Tetrahedral Iron in the Active Center of Plant Ferredoxins and Beef Adrenodoxin. *Proceedings of the National Academy of Sciences* **1971**, *68*, 3015–3020, DOI: 10.1073/pnas.68.12.3015.

(141) Yao, S.; Meier, F.; Lindenmaier, N.; Rudolph, R.; Blom, B.; Adelhardt, M.; Sutter, J.; Mebs, S.; Haumann, M.; Meyer, K.; Kaupp, M.; Driess, M. Biomimetic [2Fe-2S] Clusters with Extensively Delocalized Mixed-Valence Iron Centers. *Angew. Chem. Int. Ed.* **2015**, *54*, 12506–12510, DOI: 10.1002/anie.201506788.

(142) Reesbeck, M. E.; Rodriguez, M. M.; Brennessel, W. W.; Mercado, B. Q.; Vinyard, D.; Holland, P. L. Oxidized and reduced [2Fe–2S] clusters from an iron(I) synthon. *J Biol Inorg Chem* **2015**, *20*, 875–883, DOI: 10.1007/s00775-015-1272-4.

(143) Benini, S.; Ciurli, S.; Luchinat, C. Oxidized and Reduced [Fe2Q2] (Q = S, Se) Cores of Spinach Ferredoxin: a Comparative Study Using 1H NMR Spectroscopy. *Inorganic Chemistry* **1995**, *34*, 417–420, DOI: 10.1021/ic00105a068.

(144) Bertini, I.; Ciurli, S.; Luchinat, C. In *Iron-Sulfur Proteins Perovskites*; Springer Berlin Heidelberg: Berlin, Heidelberg, 1995; Vol. 83, pp 1–53, DOI: 10.1007/3-540-59105-2_1.

(145) Crouse, B. R.; Meyer, J.; Johnson, M. K. Spectroscopic Evidence for a Reduced Fe2S2 Cluster with a S = 9/2 Ground State in Mutant Forms of Clostridium pasteurianum 2Fe Ferredoxin. *J. Am. Chem. Soc.* **1995**, *117*, 9612–9613, DOI: 10.1021/ja00142a049.

(146) Achim, C.; Golinelli, M.-P.; Bominaar, E. L.; Meyer, J.; Münck, E. Mössbauer Study of Cys56Ser Mutant 2Fe Ferredoxin from Clostridium Pasteurianum: Evidence for Double Exchange in an [Fe2S2]+ Cluster. *J. Am. Chem. Soc.* **1996**, *118*, 8168–8169, DOI: 10.1021/ja9617698.

(147) Johnson, M. K.; Duin, E. C.; Crouse, B. R.; Golinelli, M.-P.; Meyer, J. In *Spectroscopic Methods in Bioinorganic Chemistry*; ACS Symposium Series 692, Vol. 692; American Chemical Society: 1998, pp 286–301, DOI: 10.1021/bk-1998-0692.ch016.

(148) Yeh, A. P.; Chatelet, C.; Soltis, S. M.; Kuhn, P.; Meyer, J.; Rees, D. C. Structure of a thioredoxin-like [2Fe-2S] ferredoxin from Aquifex aeolicus. *Journal of Molecular Biology* **2000**, *300*, 587–595, DOI: 10.1006/jmbi.2000.3871.

(149) Yeh, A. P.; Ambroggio, X. I.; Andrade, S. L. A.; Einsle, O.; Chatelet, C.; Meyer, J.; Rees, D. C. High Resolution Crystal Structures of the Wild Type and Cys-55→Ser and Cys-59→Ser Variants of the Thioredoxin-like [2Fe-2S] Ferredoxin from Aquifex aeolicus. *J. Biol. Chem.* **2002**, *277*, 34499–34507, DOI: 10.1074/jbc.M205096200.

(150) Mascharak, P. K.; Papaefthymiou, G. C.; Frankel, R. B.; Holm, R. H. Evidence for the localized iron(III)/iron(II) oxidation state configuration as an intrinsic property of [Fe2S2(SR)4]3- clusters. *J. Am. Chem. Soc.* **1981**, *103*, 6110–6116, DOI: 10.1021/ja00410a021.

(151) Beardwood, P.; F. Gibson, J. Electron spin resonance spectra of reduced [Fe2S2 (SC6H4Y-p)4]2- (Y = Cl, H, or Me) complexes and their selenium-ligated homologues. *Journal of the Chemical Society, Dalton Transactions* **1983**, *0*, 737–748, DOI: 10.1039/DT9830000737.

(152) Ding, X.-Q.; Bill, E.; Trautwein, A. X.; Winkler, H.; Kostikas, A.; Papaefthymiou, V.; Simopoulos, A.; Beardwood, P.; Gibson, J. F. Exchange interactions, charge delocalization, and spin relaxation in a mixed-valence di-iron complex studied by Mössbauer spectroscopy. *J. Chem. Phys.* **1993**, *99*, 6421–6428, DOI: 10.1063/1.465881.

(153) Albers, A.; Demeshko, S.; Dechert, S.; Bill, E.; Bothe, E.; Meyer, F. The Complete Characterization of a Reduced Biomimetic [2Fe-2S] Cluster. *Angewandte Chemie International Edition* **2011**, *50*, 9191–9194, DOI: 10.1002/anie.201100727.

(154) Drüeke, S.; Chaudhuri, P.; Pohl, K.; Wieghardt, K.; Ding, X.-Q.; Bill, E.; Sawaryn, A.; Trautwein, A. X.; Winkler, H.; Gurman, S. J. The novel mixed-valence, exchange-coupled, class III dimer [L2Fe2(μ-OH)3]2+(L =N,N',N"-trimethyl-1,4,7-triazacyclononane). *J. Chem. Soc., Chem. Commun.* **1989**, *0*, 59–62, DOI: 10.1039/C39890000059.

(155) Ding, X.-Q.; Bominaar, E. L.; Bill, E.; Winkler, H.; Trautwein, A. X.; Drüeke, S.; Chaudhuri, P.; Wieghardt, K. Mössbauer and electron paramagnetic resonance study of the double-exchange and Heisenberg-exchange interactions in a novel binuclear Fe(II/III) delocalized-valence compound. *The Journal of Chemical Physics* **1990**, *92*, 178–186, DOI: 10.1063/1.458460.

(156) Gamelin, D. R.; Bominaar, E. L.; Kirk, M. L.; Wieghardt, K.; Solomon, E. I. Excited-State Contributions to Ground-State Properties of Mixed-Valence Dimers: Spectral and Electronic-Structural Studies of [Fe2(OH)3(tmtacn)2]2+ Related to the [Fe2S2]+ Active Sites of Plant-Type Ferredoxins. *J. Am. Chem. Soc.* **1996**, *118*, 8085–8097, DOI: 10.1021/ja9536633.

(157) Kramida, A.; Yu. Ralchenko; Reader, J.; and NIST ASD Team NIST Atomic Spectra Database (ver. 5.7.1). **2019**, DOI: 10.18434/T4W30F.

(158) Cordero, B.; Gómez, V.; Platero-Prats, A. E.; Revés, M.; Echeverría, J.; Cremades, E.; Barragán, F.; Alvarez, S. Covalent radii revisited. *Dalton Trans.* **2008**, 2832, DOI: 10.1039/b801115j.

(159) Shannon, R. D. Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides. *Acta Cryst A* **1976**, *32*, 751–767, DOI: 10.1107/S0567739476001551.

(160) Thompson, A.; Kirz, J.; Attwood, D.; Gullikson, E.; Howells, M.; Kortright, J.; Liu, Y.; Robinson, A.; Kim, K.-J.; Lindau, I.; Pianetta, P.; Scofield, J.; Underwood, J.; Williams, G.; Winick, *X-Ray Data Booklet*, 3rd ed.; Lawrence Berkeley National Laboratory, University of California: Berkeley, California 94720, 2009.

(161) Mousa, R.; Notis Dardashti, R.; Metanis, N. Selenium and Selenocysteine in Protein Chemistry. *Angewandte Chemie International Edition* **2017**, *56*, 15818–15827, DOI: 10.1002/anie.201706876.

(162) Fordyce, F. M. In *Essentials of Medical Geology: Revised Edition*, Selinus, O., Ed.; Springer Netherlands: Dordrecht, 2013, pp 375–416, DOI: 10.1007/978-94-007-4375-5_16.

(163) Stadtman, T. C. Selenium-Dependent Enzymes. *Annu. Rev. Biochem.* **1980**, *49*, 93–110, DOI: 10.1146/annurev.bi.49.070180.000521.

(164) Malkin, R.; Rabinowitz, J. C. The reconstitution of clostridial ferredoxin. *Biochemical and Biophysical Research Communications* **1966**, *23*, 822–827, DOI: 10.1016/0006-291X(66)90561-4.

(165) Ogasawara, Y.; Isoda, S.; Tanabe, S. Reconstitution of an iron-sulfur cluster with bound sulfur: a possible source of acid-labile sulfur in biological systems. *Biol. Pharm. Bull.* **1995**, *18*, 1045–1048, DOI: 10.1248/bpb.18.1045.

(166) Orme-Johnson, W. H.; Hansen, R. E.; Beinert, H.; Tsibris, J. C.; Bartholomaus, R. C.; Gunsalus, I. C. On the sulfur components of iron-sulfur proteins. I. The number of acid-labile sulfur groups sharing an unpaired electron with iron. *PNAS* **1968**, *60*, 368–372.

(167) Tsukihara, T.; Fukuyama, K.; Nakamura, M.; Katsube, Y.; Tanaka, N.; Kakudo, M.; Wada, K.; Hase, T.; Matsubara, H. X-Ray Analysis of a [2Fe-2S] Ferredoxin from 'Spirulina platensis. Main Chain Fold and Location of Side Chains at 2.5 Å Resolution. *The Journal of Biochemistry* **1981**, *90*, 1763–1773, DOI: 10.1093/oxfordjournals.jbchem.a133654.

(168) Meyer, J.; Moulis, J.-M.; Gaillard, J.; Lutz, M. In *Advances in Inorganic Chemistry*, Cammack, R., Ed.; Academic Press: 1992; Vol. 38, pp 73–115, DOI: 10.1016/S0898-8838(08)60062-8.

(169) Hallenbeck, P. C.; George, G. N.; Prince, R. C.; Thorneley, R. N. F. Characterization of a modified nitrogenase Fe protein from Klebsiella pneumoniae in which the 4Fe4S cluster has been replaced by a 4Fe4Se cluster. *J Biol Inorg Chem* **2009**, *14*, 673–682, DOI: 10.1007/s00775-009-0480-1.

(170) Spatzal, T.; Perez, K. A.; Howard, J. B.; Rees, D. C. Catalysis-dependent selenium incorporation and migration in the nitrogenase active site iron-molybdenum cofactor. *eLife* **2015**, *4*, DOI: 10.7554/eLife.11620.

(171) Henthorn, J. T.; Arias, R. J.; Koroidov, S.; Kroll, T.; Sokaras, D.; Bergmann, U.; Rees, D. C.; DeBeer, S. Localized Electronic Structure of Nitrogenase FeMoco Revealed by Selenium K-Edge High Resolution X-ray Absorption Spectroscopy. *J. Am. Chem. Soc.* **2019**, *141*, 13676–13688, DOI: 10.1021/jacs.9b06988.

(172) Tanifuji, K. *et al.* Tracing the incorporation of the "ninth sulfur" into the nitrogenase cofactor precursor with selenite and tellurite. *Nat. Chem.* **2021**, DOI: 10.1038/s41557-021-00799-8.

(173) Tsibris, J. C. M.; Namtvedt, M. J.; Gunsalus, I. C. Selenium as an acid labile sulfur replacement in putidaredoxin. *Biochemical and Biophysical Research Communications* **1968**, *30*, 323–327, DOI: 10.1016/0006-291X(68)90454-3.

(174) Mukai, K.; Huang, J. J.; Kimura, T. Electron paramagnetic resonance studies on the selenium-replaced derivatives of adrenodoxin: The presence of the one selenium-one sulfur compound. *Biochemical and Biophysical Research Communications* **1973**, *50*, 105–110, DOI: 10.1016/0006-291X(73)91070-X.

(175) Fee, J. A.; Palmer, G. The properties of parsley ferredoxin and its selenium-containing homolog. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1971**, *245*, 175–195, DOI: 10.1016/0005-2728(71)90020-X.

(176) Muenck, E.; Debrunner, P. G.; Tsibris, J. C. M.; Gunsalus, I. C. Moessbauer parameters of putidaredoxin and its selenium analog. *Biochemistry* **1972**, *11*, 855–863, DOI: 10.1021/bi00755a027.

(177) Bowman, M.; Kevan, L.; Mukai, K.; Kimura, T. Comparative studies on adrenal iron-sulfur protein and its selenium derivatives by electron nuclear double resonance. *Biochimica et Biophysica Acta (BBA) - Protein Structure* **1973**, *328*, 244–251, DOI: 10.1016/0005-2795(73)90258-4.

(178) Meyer, J.; Moulis, J.-M.; Lutz, M. High-yield chemical assembly of [2Fe-2X] (X = S, Se) clusters into spinach apoferredoxin: product characterization by resonance Raman spectroscopy. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* **1986**, *871*, 243–249, DOI: 10.1016/0167-4838(86)90205-0.

(179) Tang, S.-P. W.; Spiro, T. G.; Mukai, K.; Kimura, T. Resonance Raman scattering and optical absorption of adrenodoxin and selena-adrenodoxin. *Biochemical and Biophysical Research Communications* **1973**, *53*, 869–874, DOI: 10.1016/0006-291X(73)90173-3.

(180) Yu, S. B.; Papaefthymiou, G. C.; Holm, R. H. Comprehensive iron-selenium-thiolate cluster chemistry. *Inorg. Chem.* **1991**, *30*, 3476–3485, DOI: 10.1021/ic00018a018.

(181) Strasdeit, H.; Krebs, B.; Henkel, G. Synthetic routes to [Fe2X2(X5)2]2- anions (X = S, Se). Structure and properties of [Fe2Se2(Se5)2]2-, a complex with selenido and pentaselenido ligands. *Inorganica Chimica Acta* **1984**, *89*, L11–L13, DOI: 10.1016/S0020-1693(00)82418-4.

(182) Beardwood, P.; Gibson, J. F. Resonance-Raman and infrared spectra of the dimeric iron–chalcogenide complexes [Fe2X2(YR)4]2– (X and Y = S or Se). *J. Chem. Soc., Dalton Trans.* **1984**, 1507–1516, DOI: 10.1039/DT9840001507.

(183) Fohlmeister, L.; Vignesh, K. R.; Winter, F.; Moubaraki, B.; Rajaraman, G.; Pöttgen, R.; Murray, K. S.; Jones, C. Neutral diiron(III) complexes with Fe2(μ-E)2 (E = O, S, Se) core structures: reactivity of an iron(I) dimer towards chalcogens. *Dalton Trans.* **2014**, *44*, 1700–1708, DOI: 10.1039/C4DT03081H.

(184) Carney, M. J.; Papaefthymiou, G. C.; Spartalian, K.; Frankel, R. B.; Holm, R. H. Ground spin state variability in [Fe4S4(SR)4]3-. Synthetic analogs of the reduced clusters in ferredoxins and other iron-sulfur proteins: cases of extreme sensitivity of electronic state and structure to extrinsic factors. *J. Am. Chem. Soc.* **1988**, *110*, 6084–6095, DOI: 10.1021/ja00226a025.

(185) Burgess, B. K.; Lowe, D. J. Mechanism of Molybdenum Nitrogenase. *Chem. Rev.* **1996**, *96*, 2983–3012, DOI: 10.1021/cr950055x.

(186) Lancaster, K. M.; Roemelt, M.; Ettenhuber, P.; Hu, Y.; Ribbe, M. W.; Neese, F.; Bergmann, U.; DeBeer, S. X-ray Emission Spectroscopy Evidences a Central Carbon in the Nitrogenase Iron-Molybdenum Cofactor. *Science* **2011**, *334*, 974–977, DOI: 10.1126/science.1206445.

(187) Hoffman, B. M.; Lukoyanov, D.; Yang, Z.-Y.; Dean, D. R.; Seefeldt, L. C. Mechanism of Nitrogen Fixation by Nitrogenase: The Next Stage. *Chem Rev* **2014**, *114*, 4041–4062, DOI: 10.1021/cr400641x.

(188) Eady, R. R. Structure-Function Relationships of Alternative Nitrogenases. *Chem. Rev.* **1996**, *96*, 3013–3030, DOI: 10.1021/cr950057h.

(189) Harris, D. F.; Lukoyanov, D. A.; Kallas, H.; Trncik, C.; Yang, Z.-Y.; Compton, P.; Kelleher, N.; Einsle, O.; Dean, D. R.; Hoffman, B. M.; Seefeldt, L. C. Mo-, V-, and Fe-Nitrogenases Use a Universal Eight-Electron Reductive-Elimination Mechanism To Achieve N2 Reduction. *Biochemistry* **2019**, *58*, 3293–3301, DOI: 10.1021/acs.biochem.9b00468.

(190) Lee, C. C.; Hu, Y.; Ribbe, M. W. Vanadium Nitrogenase Reduces CO. *Science* **2010**, *329*, 642–642, DOI: 10.1126/science.1191455.

(191) Yang, Z.-Y.; Dean, D. R.; Seefeldt, L. C. Molybdenum Nitrogenase Catalyzes the Reduction and Coupling of CO to Form Hydrocarbons. *J. Biol. Chem.* **2011**, *286*, 19417–19421, DOI: 10.1074/jbc.M111.229344.

(192) Harris, D. F.; Jimenez-Vicente, E.; Yang, Z.-Y.; Hoffman, B. M.; Dean, D. R.; Seefeldt, L. C. CO as a substrate and inhibitor of H+ reduction for the Mo-, V-, and Fe-nitrogenase isozymes. *Journal of Inorganic Biochemistry* **2020**, *213*, 111278, DOI: 10.1016/j.jinorgbio.2020.111278.

(193) Sickerman, N. S.; Hu, Y.; Ribbe, M. W. Activation of CO2 by Vanadium Nitrogenase. *Chemistry – An Asian Journal* **2017**, *12*, 1985–1996, DOI: 10.1002/asia.201700624.

(194) Shah, V. K.; Brill, W. J. Isolation of an iron-molybdenum cofactor from nitrogenase. *PNAS* **1977**, *74*, 3249–3253, DOI: 10.1073/pnas.74.8.3249.

(195) Robinson, A. C.; Dean, D. R.; Burgess, B. K. Iron-molybdenum cofactor biosynthesis in Azotobacter vinelandii requires the iron protein of nitrogenase. *J Biol Chem* **1987**, *262*, 14327–14332.

(196) Moore, V. G.; Tittsworth, R. C.; Hales, B. J. Construction and Characterization of Hybrid Component 1 from V-Nitrogenase Containing FeMo Cofactor. *Journal of the American Chemical Society* **1994**, *116*, 12101–12102, DOI: 10.1021/ja00105a080.

(197) Fay, A. W.; Blank, M. A.; Lee, C. C.; Hu, Y.; Hodgson, K. O.; Hedman, B.; Ribbe, M. W. Characterization of Isolated Nitrogenase FeVco. *J. Am. Chem. Soc.* **2010**, *132*, 12612–12618, DOI: 10.1021/ja1019657.

(198) Rebelein, J. G.; Lee, C. C.; Newcomb, M.; Hu, Y.; Ribbe, M. W. Characterization of an M-Cluster-Substituted Nitrogenase VFe Protein. *mBio* **2018**, *9*, DOI: 10.1128/mBio.00310-18.

(199) Münck, E.; Rhodes, H.; Orme-Johnson, W. H.; Davis, L. C.; Brill, W. J.; Shah, V. K. Nitrogenase. VIII. Mössbauer and EPR spectroscopy. The MoFe protein component from Azotobacter vinelandii OP. *Biochimica et Biophysica Acta (BBA) - Protein Structure* **1975**, *400*, 32–53, DOI: 10.1016/0005-2795(75)90124-5.

(200) Bjornsson, R.; Neese, F.; DeBeer, S. Revisiting the Mössbauer Isomer Shifts of the FeMoco Cluster of Nitrogenase and the Cofactor Charge. *Inorg. Chem.* **2017**, *56*, 1470–1477, DOI: 10.1021/acs.inorgchem.6b02540.

(201) Spatzal, T.; Schlesier, J.; Burger, E.-M.; Sippel, D.; Zhang, L.; Andrade, S. L. A.; Rees, D. C.; Einsle, O. Nitrogenase FeMoco investigated by spatially resolved anomalous dispersion refinement. *Nat Commun* **2016**, *7*, 1–7, DOI: 10.1038/ncomms10902.

(202) Yang, Z.-Y.; Jimenez-Vicente, E.; Kallas, H.; Lukoyanov, D. A.; Yang, H.; Campo, J. S. M. d.; Dean, D. R.; Hoffman, B. M.; Seefeldt, L. C. The electronic structure of FeV-cofactor in vanadium-dependent nitrogenase. *Chem. Sci.* **2021**, *12*, 6913–6922, DOI: 10.1039/D0SC06561G.

(203) Sippel, D.; Einsle, O. The structure of vanadium nitrogenase reveals an unusual bridging ligand. *Nat Chem Biol* **2017**, *13*, 956–960, DOI: 10.1038/nchembio.2428.

(204) Van Stappen, C.; Thorhallsson, A. T.; Decamps, L.; Bjornsson, R.; DeBeer, S. Resolving the structure of the E1 state of Mo nitrogenase through Mo and Fe K-edge EXAFS and QM/MM calculations. *Chem. Sci.* **2019**, *10*, 9807–9821, DOI: 10.1039/C9SC02187F.

(205) Van Stappen, C.; Davydov, R.; Yang, Z.-Y.; Fan, R.; Guo, Y.; Bill, E.; Seefeldt, L. C.; Hoffman, B. M.; DeBeer, S. Spectroscopic Description of the E1 State of Mo Nitrogenase Based on Mo and Fe X-ray Absorption and Mössbauer Studies. *Inorg. Chem.* **2019**, *58*, 12365–12376, DOI: 10.1021/acs.inorgchem.9b01951.

(206) Cao, L.; Caldararu, O.; Ryde, U. Protonation and Reduction of the FeMo Cluster in Nitrogenase Studied by Quantum Mechanics/Molecular Mechanics (QM/MM) Calculations. *J. Chem. Theory Comput.* **2018**, *14*, 6653–6678, DOI: 10.1021/acs.jctc.8b00778.

(207) Varley, J.; Wang, Y.; Chan, K.; Studt, F.; Nørskov, J. K. Mechanistic insights into nitrogen fixation by nitrogenase enzymes. *Physical Chemistry Chemical Physics* **2015**, *17*, 29541–29547, DOI: 10.1039/C5CP04034E.

(208) Dance, I. Survey of the Geometric and Electronic Structures of the Key Hydrogenated Forms of FeMo-co, the Active Site of the Enzyme Nitrogenase: Principles of the Mechanistically Significant Coordination Chemistry. *Inorganics* **2019**, *7*, 8, DOI: 10.3390/inorganics7010008.

(209) Pelmenschikov, V.; Case, D. A.; Noodleman, L. Ligand-Bound S = 1/2 FeMo-Cofactor of Nitrogenase: Hyperfine Interaction Analysis and Implication for the Central Ligand X Identity. *Inorg. Chem.* **2008**, *47*, 6162–6172, DOI: 10.1021/ic7022743.

(210) Fisher, K.; Dilworth, M. J.; Kim, C.-H.; Newton, W. E. Azotobacter vinelandii Nitrogenases Containing Altered MoFe Proteins with Substitutions in the FeMo-Cofactor Environment: Effects on the Catalyzed Reduction of Acetylene and Ethylene. *Biochemistry* **2000**, *39*, 2970–2979, DOI: 10.1021/bi992092e.

(211) Fisher, K.; Lowe, D. J.; Tavares, P.; Pereira, A. S.; Huynh, B. H.; Edmondson, D.; Newton, W. E. Conformations generated during turnover of the Azotobacter vinelandii nitrogenase MoFe protein and their relationship to physiological function. *Journal of Inorganic Biochemistry* **2007**, *101*, 1649–1656, DOI: 10.1016/j.jinorgbio.2007.07.037.

(212) Lukoyanov, D.; Barney, B. M.; Dean, D. R.; Seefeldt, L. C.; Hoffman, B. M. Connecting nitrogenase intermediates with the kinetic scheme for N2 reduction by a relaxation protocol and identification of the N2 binding state. *PNAS* **2007**, *104*, 1451–1455, DOI: 10.1073/pnas.0610975104.

(213) Lukoyanov, D. A.; Khadka, N.; Yang, Z.-Y.; Dean, D. R.; Seefeldt, L. C.; Hoffman, B. M. Hydride Conformers of the Nitrogenase FeMo-cofactor Two-Electron Reduced State E2 (2H), Assigned Using Cryogenic Intra Electron Paramagnetic Resonance Cavity Photolysis. *Inorg. Chem.* **2018**, *57*, 6847–6852, DOI: 10.1021/acs.inorgchem.8b00271.

(214) Thorhallsson, A. T.; Bjornsson, R. The E2 state of FeMoco: hydride formation vs. Fe reduction and a mechanism for H2 evolution. *Chemistry – A European Journal* **2021**, *n/a*, DOI: 10.1002/chem.202102730.

(215) Lind, C. J.; Wilson, P. W. Mechanism of Biological Nitrogen Fixation. VIII. Carbon Monoxide as an Inhibitor for Nitrogen Fixation by Red Clover1. *J. Am. Chem. Soc.* **1941**, *63*, 3511–3514, DOI: 10.1021/ja01857a073.

(216) Hwang, J. C.; Chen, C. H.; Burris, R. H. Inhibition of nitrogenase-catalyzed reductions. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1973**, *292*, 256–270, DOI: 10.1016/0005-2728(73)90270-3.

(217) Lee, H.-I.; Cameron, L. M.; Hales, B. J.; Hoffman, B. M. CO Binding to the FeMo Cofactor of CO-Inhibited Nitrogenase: 13CO and 1H Q-Band ENDOR Investigation. *J. Am. Chem. Soc.* **1997**, *119*, 10121–10126, DOI: 10.1021/ja9715096.

(218) Christie, P. D.; Lee, H.-I.; Cameron, L. M.; Hales, B. J.; Orme-Johnson, W. H.; Hoffman, B. M. Identification of the CO-Binding Cluster in Nitrogenase MoFe Protein by ENDOR of 57Fe Isotopomers. *J. Am. Chem. Soc.* **1996**, *118*, 8707–8709, DOI: 10.1021/ja961289g.

(219) Lee, C. C.; Fay, A. W.; Weng, T.-C.; Krest, C. M.; Hedman, B.; Hodgson, K. O.; Hu, Y.; Ribbe, M. W. Uncoupling binding of substrate CO from turnover by vanadium nitrogenase. *PNAS* **2015**, *112*, 13845–13849, DOI: 10.1073/pnas.1519696112.

(220) Lee, C. C.; Wilcoxen, J.; Hiller, C. J.; Britt, R. D.; Hu, Y. Evaluation of the Catalytic Relevance of the CO-Bound States of V-Nitrogenase. *Angewandte Chemie International Edition* **2018**, *57*, 3411–3414, DOI: 10.1002/anie.201800189.

(221) Rohde, M.; Laun, K.; Zebger, I.; Stripp, S. T.; Einsle, O. Two ligand-binding sites in CO-reducing V nitrogenase reveal a general mechanistic principle. *Science Advances* **2021**, *7*, eabg4474, DOI: 10.1126/sciadv.abg4474.

(222) Davis, L. C.; Henzl, M. T.; Burris, R. H.; Orme-Johnson, W. H. Iron-sulfur clusters in the molybdenum-iron protein component of nitrogenase. Electron paramagnetic resonance of the carbon monoxide inhibited state. *Biochemistry* **1979**, *18*, 4860–4869, DOI: 10.1021/bi00589a014.

(223) Maskos, Z.; Hales, B. J. Photo-lability of CO bound to Mo-nitrogenase from Azotobacter vinelandii. *Journal of Inorganic Biochemistry* **2003**, *93*, 11–17, DOI: 10.1016/S0162-0134(02)00480-4.

(224) Sørlie, M.; Christiansen, J.; Lemon, B. J.; Peters, J. W.; Dean, D. R.; Hales, B. J. Mechanistic Features and Structure of the Nitrogenase α-Gln195 MoFe Protein, *Biochemistry* **2001**, *40*, 1540–1549, DOI: 10.1021/bi0013997.

(225) Mina-Camilde, N.; Manzanares I., C.; Caballero, J. F. Molecular Constants of Carbon Monoxide at v = 0, 1, 2, and 3: A Vibrational Spectroscopy Experiment in Physical Chemistry. *J. Chem. Educ.* **1996**, *73*, 804, DOI: 10.1021/ed073p804.

(226) *Prokaryotic Nitrogen Fixation: A Model System for the Analysis of a Biological Process*, 1st edition; Triplett, E. W., Ed.; Horizon Press: Wymondham, 2000, 820 pp.

(227) George, S. J.; Ashby, G. A.; Wharton, C. W.; Thorneley, R. N. F. Time-Resolved Binding of Carbon Monoxide to Nitrogenase Monitored by Stopped-Flow Infrared Spectroscopy. *J. Am. Chem. Soc.* **1997**, *119*, 6450–6451, DOI: 10.1021/ja971088s.

(228) Yang, Z.-Y.; Seefeldt, L. C.; Dean, D. R.; Cramer, S. P.; George, S. J. Steric Control of the Hi-CO MoFe Nitrogenase Complex Revealed by Stopped-Flow Infrared Spectroscopy. *Angewandte Chemie* **2011**, *123*, 286–289, DOI: https://doi.org/10.1002/ange.201005145.

(229) Tolland, J. D.; Thorneley, R. N. F. Stopped-Flow Fourier Transform Infrared Spectroscopy Allows Continuous Monitoring of Azide Reduction, Carbon Monoxide Inhibition, and ATP Hydrolysis by Nitrogenase. *Biochemistry* **2005**, *44*, 9520–9527, DOI: 10.1021/bi050453m.

(230) Gee, L. B.; Scott, A. D.; Dapper, C. H.; Newton, W. E.; Cramer, S. P. Carbon monoxide binding to α-R277H Mo-nitrogenase – Evidence for multiple pH-dependent species from IR-monitored photolysis. *Journal of Inorganic Biochemistry* **2022**, 111806, DOI: 10.1016/j.jinorgbio.2022.111806.

(231) Yan, L.; Dapper, C. H.; George, S. J.; Wang, H.; Mitra, D.; Dong, W.; Newton, W. E.; Cramer, S. P. Photolysis of Hi-CO Nitrogenase – Observation of a Plethora of Distinct CO Species Using Infrared Spectroscopy. *European Journal of Inorganic Chemistry* **2011**, *2011*, 2064–2074, DOI: 10.1002/ejic.201100029.

(232) Yan, L.; Pelmenschikov, V.; Dapper, C. H.; Scott, A. D.; Newton, W. E.; Cramer, S. P. IR-Monitored Photolysis of CO-Inhibited Nitrogenase: A Major EPR-Silent Species with Coupled Terminal CO Ligands. *Chem. Eur. J.* **2012**, *18*, 16349–16357, DOI: 10.1002/chem.201202072.

(233) Spatzal, T.; Perez, K. A.; Einsle, O.; Howard, J. B.; Rees, D. C. Ligand binding to the FeMo-cofactor: Structures of CO-bound and reactivated nitrogenase. *Science* **2014**, *345*, 1620–1623, DOI: 10.1126/science.1256679.

(234) Buscagan, T. M.; Perez, K. A.; Maggiolo, A. O.; Rees, D. C.; Spatzal, T. Structural Characterization of Two CO Molecules Bound to the Nitrogenase Active Site. *Angew. Chem. Int. Ed.* **2021**, *60*, 5704–5707, DOI: 10.1002/anie.202015751.

(235) Pérez-González, A.; Yang, Z.-Y.; Lukoyanov, D. A.; Dean, D. R.; Seefeldt, L. C.; Hoffman, B. M. Exploring the Role of the Central Carbide of the Nitrogenase Active-Site FeMo-cofactor through Targeted 13C Labeling and ENDOR Spectroscopy. *J. Am. Chem. Soc.* **2021**, *143*, 9183–9190, DOI: 10.1021/jacs.1c04152.

(236) Rohde, M.; Grunau, K.; Einsle, O. CO Binding to the FeV Cofactor of CO-Reducing Vanadium Nitrogenase at Atomic Resolution. *Angewandte Chemie International Edition* **2020**, *59*, 23626–23630, DOI: 10.1002/anie.202010790.

(237) Ibrahim, S. K.; Gormal, C. A.; Smith, B. E.; Pickett, C. J.; Vincent, K.; Best, S. P. The isolated iron–molybdenum cofactor of nitrogenase binds carbon monoxide upon electrochemically accessing reduced states. *Chem. Commun.* **1999**, 1019–1020, DOI: 10.1039/A902371B.

(238) Pickett, C. J.; Vincent, K. A.; Ibrahim, S. K.; Gormal, C. A.; Smith, B. E.; Best, S. P. Electron-Transfer Chemistry of the Iron–Molybdenum Cofactor of Nitrogenase: Delocalized and Localized Reduced States of FeMoco which Allow Binding of Carbon Monoxide to Iron and Molybdenum. *Chem. Eur. J.* **2003**, *9*, 76–87, DOI: 10.1002/chem.200390033.

(239) Pickett, C. J.; Vincent, K. A.; Ibrahim, S. K.; Gormal, C. A.; Smith, B. E.; Fairhurst, S. A.; Best, S. P. Synergic Binding of Carbon Monoxide and Cyanide to the FeMo Cofactor of Nitrogenase: Relic Chemistry of an Ancient Enzyme? *Chemistry – A European Journal* **2004**, *10*, 4770–4776, DOI: 10.1002/chem.200400382.

(240) King, R. B.; Bitterwolf, T. E. Metal carbonyl analogues of iron–sulfur clusters found in metalloenzyme chemistry. *Coordination Chemistry Reviews* **2000**, *206-207*, 563–579, DOI: 10.1016/S0010-8545(99)00251-9.

(241) Mascharak, P. K.; Armstrong, W. H.; Mizobe, Y.; Holm, R. H. Single cubane-type MFe3S4 clusters (M = Mo, W): synthesis and properties of oxidized and reduced forms. *J. Am. Chem. Soc.* **1983**, *105*, 475–483, DOI: 10.1021/ja00341a029.

(242) Nguyen, D. H.; Hsu, H.-F.; Millar, M.; Koch, S. A.; Achim, C.; Bominaar, E. L.; Münck, E. Nickel(II) Thiolate Complex with Carbon Monoxide and Its Fe(II) Analog: Synthetic Models for CO Adducts of Nickel–Iron-Containing Enzymes. *J. Am. Chem. Soc.* **1996**, *118*, 8963–8964, DOI: 10.1021/ja961968r.

(243) Davies, S. C.; Hughes, D. L.; Richards, R. L.; Roger Sanders, J. Iron complexes of the N(CH2CH2S)33– ligand; a paramagnetic, trigonal bipyramidal Fe(II) CO complex as a chelate ligand. *Chem. Commun.* **1998**, 2699–2700, DOI: 10.1039/a807655c.

(244) Sellmann, D.; Blum, D. C. F.; Heinemann, F. W.; Sutter, J. Synthesis, Structure and Reactivity of Sulfur-Rich [Fe(L1)(L2)('tpS4')] Complexes with Rigid [Fe('tpS4'] Cores and σ–π Co-Ligands (L1, L2 = CO, PR3, NO). *European Journal of Inorganic Chemistry* **2003**, *2003*, 418–426, DOI: 10.1002/ejic.200390058.

(245) Moret, M.-E.; Peters, J. C. Terminal Iron Dinitrogen and Iron Imide Complexes Supported by a Tris(phosphino)borane Ligand. *Angewandte Chemie International Edition* **2011**, *50*, 2063–2067, DOI: 10.1002/anie.201006918.

(246) Lee, Y.; Peters, J. C. Silylation of Iron-Bound Carbon Monoxide Affords a Terminal Fe Carbyne. *J. Am. Chem. Soc.* **2011**, *133*, 4438–4446, DOI: 10.1021/ja109678y.

(247) Rittle, J.; Peters, J. C. Fe–N2/CO complexes that model a possible role for the interstitial C atom of FeMo-cofactor (FeMoco). *Proc Natl Acad Sci U S A* **2013**, *110*, 15898–15903, DOI: 10.1073/pnas.1310153110.

(248) Arnett, C. H.; Chalkley, M. J.; Agapie, T. A Thermodynamic Model for Redox-Dependent Binding of Carbon Monoxide at Site-Differentiated, High Spin Iron Clusters. *J. Am. Chem. Soc.* **2018**, *140*, 5569–5578, DOI: 10.1021/jacs.8b01825.

(249) Bloch, E. D. *et al.* Reversible CO Binding Enables Tunable CO/H2 and CO/N2 Separations in Metal–Organic Frameworks with Exposed Divalent Metal Cations. *J. Am. Chem. Soc.* **2014**, *136*, 10752–10761, DOI: 10.1021/ja505318p.

(250) Brown, A.; Thompson, N.; Suess, D. Evidence for Low-valent Electronic Configurations in Iron–Sulfur Clusters. *ChemRxiv (Preprint)* **2022**, DOI: 10.26434/chemrxiv-2022-x058p.

(251) Scott, A. D.; Pelmenschikov, V.; Guo, Y.; Yan, L.; Wang, H.; George, S. J.; Dapper, C. H.; Newton, W. E.; Yoda, Y.; Tanaka, Y.; Cramer, S. P. Structural Characterization of CO-Inhibited Mo-Nitrogenase by Combined Application of Nuclear Resonance Vibrational Spectroscopy, Extended X-ray Absorption Fine Structure, and Density Functional Theory: New Insights into the Effects of CO Binding and the Role of the Interstitial Atom. *J. Am. Chem. Soc.* **2014**, *136*, 15942–15954, DOI: 10.1021/ja505720m.

(252) Rod, T. H.; Nørskov, J. K. Modeling the Nitrogenase FeMo Cofactor. *J. Am. Chem. Soc.* **2000**, *122*, 12751–12763, DOI: 10.1021/ja001163q.

(253) Dance, I. Calculated vibrational frequencies for FeMo-co, the active site of nitrogenase, bearing hydrogen atoms and carbon monoxide. *Dalton Trans.* **2011**, *40*, 6480–6489, DOI: 10.1039/C1DT10505A.

(254) Varley, J. B.; Nørskov, J. K. First-Principles Calculations of Fischer–Tropsch Processes Catalyzed by Nitrogenase Enzymes. *ChemCatChem* **2013**, *5*, 732–736, DOI: 10.1002/cctc.201200635.

(255) Bergmann, J.; Oksanen, E.; Ryde, U. Quantum-refinement studies of the bidentate ligand of V nitrogenase and the protonation state of CO-inhibited Mo nitrogenase. *Journal of Inorganic Biochemistry* **2021**, *219*, 111426, DOI: 10.1016/j.jinorgbio.2021.111426.

(256) Bjornsson, R.; Lima, F. A.; Spatzal, T.; Weyhermüller, T.; Glatzel, P.; Bill, E.; Einsle, O.; Neese, F.; DeBeer, S. Identification of a spin-coupled Mo(III) in the nitrogenase iron–molybdenum cofactor. *Chem. Sci.* **2014**, *5*, 3096–3103, DOI: 10.1039/C4SC00337C.

(257) Krewald, V.; Neese, F.; Pantazis, D. A. On the Magnetic and Spectroscopic Properties of High-Valent Mn3CaO4 Cubanes as Structural Units of Natural and Artificial Water-Oxidizing Catalysts. *J. Am. Chem. Soc.* **2013**, *135*, 5726–5739, DOI: 10.1021/ja312552f.

(258) 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, DOI: 10.1039/B508541A.

(259) Barone, V.; Cossi, M. Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model. *J. Phys. Chem. A* **1998**, *102*, 1995–2001, DOI: 10.1021/jp9716997.

(260) Lenthe, E. v.; Baerends, E. J.; Snijders, J. G. Relativistic regular two-component Hamiltonians. *J. Chem. Phys.* **1993**, *99*, 4597–4610, DOI: 10.1063/1.466059.

(261) Van Wüllen, C. Molecular density functional calculations in the regular relativistic approximation: Method, application to coinage metal diatomics, hydrides, fluorides and chlorides, and comparison with first-order relativistic calculations. *J. Chem. Phys.* **1998**, *109*, 392–399, DOI: 10.1063/1.476576.

(262) Stanton, J. F.; Bartlett, R. J. The equation of motion coupled-cluster method. A systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties. *The Journal of Chemical Physics* **1998**, *98*, 7029, DOI: 10.1063/1.464746.

(263) Dutta, A. K.; Vaval, N.; Pal, S. Lower scaling approximation to EOM-CCSD: A critical assessment of the ionization problem. *International Journal of Quantum Chemistry* **2018**, *118*, e25594, DOI: 10.1002/qua.25594.

(264) Neese, F. An improvement of the resolution of the identity approximation for the formation of the Coulomb matrix. *Journal of Computational Chemistry* **2003**, *24*, 1740–1747, DOI: 10.1002/jcc.10318.

(265) Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. Efficient, approximate and parallel Hartree–Fock and hybrid DFT calculations. A 'chain-of-spheres' algorithm for the Hartree–Fock exchange. *Chemical Physics* **2009**, *356*, 98–109, DOI: 10.1016/j.chemphys.2008.10.036.

(266) York, D. M.; Karplus, M. A Smooth Solvation Potential Based on the Conductor-Like Screening Model. *J. Phys. Chem. A* **1999**, *103*, 11060–11079, DOI: 10.1021/jp9920971.

(267) Chemcraft - Graphical program for visualization of quantum chemistry computations, https://www.chemcraftprog.com/ (accessed 04/23/2021).

(268) Neese, F. Efficient and accurate approximations to the molecular spin-orbit coupling operator and their use in molecular g-tensor calculations. *J. Chem. Phys.* **2005**, *122*, 034107, DOI: 10.1063/1.1829047.

(269) Humphrey, W.; Dalke, A.; Schulten, K. VMD – Visual Molecular Dynamics. *Journal of Molecular Graphics* **1996**, *14*, 33–38, DOI: 10.1016/0263-7855(96)00018-5.

(270) Stone, J. An Efficient Library for Parallel Ray Tracing and Animation, Master's Thesis, Computer Science Department, University of Missouri-Rolla, 1998.

(271) Plasser, F.; Mewes, J. M. vmd-orb.bash, http://workadayqc.blogspot.de/2014/01/shiny-pictures-mo-porn.html (accessed 05/17/2021).

(272) Hunter, J. D. Matplotlib: A 2D graphics environment. *Computing in Science & Engineering* **2007**, *9*, 90–95, DOI: 10.1109/MCSE.2007.55.

(273) Waskom, M. L. seaborn: statistical data visualization. *Journal of Open Source Software* **2021**, *6*, 3021, DOI: 10.21105/joss.03021.

(274) Inkscape Project Inkscape, version 0.92.5, 2020, https://inkscape.org.

(275) Virtanen, P. TexText - A TeX extension for Inkscape, 2021, https://github.com/textext/textext (accessed 11/02/2021).

(276) Reinhold, J., *Quantentheorie der Moleküle*; Springer Fachmedien Wiesbaden: Wiesbaden, 2015, ISBN: 978-3-658-09409-6.

(277) Roos, B. O.; Lindh, R.; Malmqvist, P. Å.; Veryazov, V.; Widmark, P.-O., *Multiconfigurational Quantum Chemistry*, 1.; Wiley: Hoboken, New Jersey, 2016; 240 pp., ISBN: 978-0-470-63346-5.

(278) Atanasov, M.; Ganyushin, D.; Sivalingam, K.; Neese, F. In *Molecular Electronic Structures of Transition Metal Complexes II*; Structure and Bonding; Springer, Berlin, Heidelberg: 2011, pp 149–220, DOI: 10.1007/430_2011_57.

(279) Singh, S. K.; Eng, J.; Atanasov, M.; Neese, F. Covalency and chemical bonding in transition metal complexes: An ab initio based ligand field perspective. *Coordination Chemistry Reviews* **2017**, *344*, 2–25, DOI: 10.1016/j.ccr.2017.03.018.

(280) Furlani, C.; Cervone, E.; Diomedi Camassei, F. Transition metal N,N-diethyldiselenocarbamates. *Inorg. Chem.* **1968**, *7*, 265–268, DOI: 10.1021/ic50060a019.

(281) Jørgensen, C. K. Spectra and electronic structure of complexes with sulphur-containing ligands. *Inorganica Chimica Acta Reviews* **1968**, *2*, 65–88, DOI: 10.1016/0073-8085(68)80015-4.

(282) Figgis, B. N.; Hitchman, M. A., *Ligand Field Theory and Its Applications*, 1 edition; Wiley-VCH: New York, 1999; 376 pp., ISBN: 978-0-471-31776-0.

(283) Deeth, R. J.; Duer, M. J.; Gerloch, M. Ligand fields from misdirected valency. 1. Lone-pair contributions in planar cobalt(II) Schiff-base complexes. *Inorg. Chem.* **1987**, *26*, 2573–2578, DOI: 10.1021/ic00263a004.

(284) Suturina, E. A.; Maganas, D.; Bill, E.; Atanasov, M.; Neese, F. Magneto-Structural Correlations in a Series of Pseudotetrahedral [CoII(XR)4]2– Single Molecule Magnets: An ab Initio Ligand Field Study. *Inorg. Chem.* **2015**, *54*, 9948–9961, DOI: 10.1021/acs.inorgchem.5b01706.

(285) Mayerle, J. J.; Denmark, S. E.; DePamphilis, B. V.; Ibers, J. A.; Holm, R. H. Synthetic analogs of the active sites of iron-sulfur proteins. XI. Synthesis and properties of complexes containing the iron sulfide (Fe2S2) core and the structures of Bis[o-xylyl-a,a'-dithiolato-mu-sulfido-ferrate(III)] and Bis[p-tolylthiolato-mu-sulfido-ferrate(III)] dianions. *J. Am. Chem. Soc.* **1975**, *97*, 1032–1045, DOI: 10.1021/ja00838a015.

(286) De Oliveira, F. T.; Münck, E.; Bominaar, E. L. DFT study of the intrinsic conformations of [2Fe–2S–4(SCH3)]2– clusters and their influence on exchange coupling. *Inorganica Chimica Acta* **2008**, *361*, 1192–1196, DOI: 10.1016/j.ica.2007.07.002.

(287) Bastardis, R.; Guihéry, N.; de Graaf, C. Isotropic non-Heisenberg terms in the magnetic coupling of transition metal complexes. *J. Chem. Phys.* **2008**, *129*, 104102, DOI: 10.1063/1.2975336.

(288) Domingo, A.; Angeli, C.; de Graaf, C.; Robert, V. Electronic reorganization triggered by electron transfer: The intervalence charge transfer of a Fe3+/Fe2+ bimetallic complex. *Journal of Computational Chemistry* **2015**, *36*, 861–869, DOI: 10.1002/jcc.23871.

(289) Nakamoto, K., *Infrared and Raman spectra of inorganic and coordination compounds*; Wiley: 1986; 504 pp., ISBN: 978-0-471-01066-1.

(290) Dunham, W. R.; Bearden, A. J.; Salmeen, I. T.; Palmer, G.; Sands, R. H.; Orme-Johnson, W. H.; Beinert, H. The two-iron ferredoxins in spinach, parsley, pig adrenal cortex, Azotobacter vinelandii, and Clostridium pasteurianum: Studies by magnetic field Mössbauer spectroscopy. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1971**, *253*, 134–152, DOI: 10.1016/0005-2728(71)90240-4.

(291) Harris, T. V.; Szilagyi, R. K. Comparative Assessment of the Composition and Charge State of Nitrogenase FeMo-Cofactor. *Inorg. Chem.* **2011**, *50*, 4811–4824, DOI: 10.1021/ic102446n.

(292) Chan, G. K.-L.; Head-Gordon, M. Highly correlated calculations with a polynomial cost algorithm: A study of the density matrix renormalization group. *J. Chem. Phys.* **2002**, *116*, 4462–4476, DOI: 10.1063/1.1449459.

(293) Chan, G. K.-L. An algorithm for large scale density matrix renormalization group calculations. *J. Chem. Phys.* **2004**, *120*, 3172–3178, DOI: 10.1063/1.1638734.

(294) Chilkuri, V. G.; Neese, F. Comparison of many-particle representations for selected-CI I: A tree based approach. *Journal of Computational Chemistry* **2021**, *42*, 982–1005, DOI: 10.1002/jcc.26518.

(295) Chilkuri, V. G.; Neese, F. Comparison of Many-Particle Representations for Selected Configuration Interaction: II. Numerical Benchmark Calculations. *J. Chem. Theory Comput.* **2021**, *17*, 2868–2885, DOI: 10.1021/acs.jctc.1c00081.

(296) Booth, G. H.; Thom, A. J. W.; Alavi, A. Fermion Monte Carlo without fixed nodes: A game of life, death, and annihilation in Slater determinant space. *J. Chem. Phys.* **2009**, *131*, 054106, DOI: 10.1063/1.3193710.

(297) Cleland, D.; Booth, G. H.; Alavi, A. Communications: Survival of the fittest: Accelerating convergence in full configuration-interaction quantum Monte Carlo. *J. Chem. Phys.* **2010**, *132*, 041103, DOI: 10.1063/1.3302277.

(298) Guther, K. *et al.* NECI: N-Electron Configuration Interaction with an emphasis on state-of-the-art stochastic methods. *J. Chem. Phys.* **2020**, *153*, 034107, DOI: 10.1063/5.0005754.

(299) Rees, J. A.; Bjornsson, R.; Kowalska, J. K.; Lima, F.; Schlesier, J.; Sippel, D.; Weyhermüller, T.; Einsle, O.; A. Kovacs, J.; DeBeer, S. Comparative electronic structures of nitrogenase FeMoco and FeVco. *Dalton Transactions* **2017**, *46*, 2445–2455, DOI: 10.1039/C7DT00128B.

(300) Thorhallsson, A. T.; Benediktsson, B.; Bjornsson, R. A model for dinitrogen binding in the E4 state of nitrogenase. *Chem. Sci.* **2019**, *10*, 11110–11124, DOI: 10.1039/C9SC03610E.

(301) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone Phi, Psi and Side-Chain Chi1 and Chi2 Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273, DOI: 10.1021/ct300400x.

(302) Smith, W.; Forester, T. R. DL_POLY_2.0: A general-purpose parallel molecular dynamics simulation package. *Journal of Molecular Graphics* **1996**, *14*, 136–141, DOI: 10.1016/S0263-7855(96)00043-4.

(303) Spiller, N. QM/MM setup and workflows for ORCA and chemshell, 2021, https://github.com/nspiller/chemshell-QMMM-protein-setup (accessed 03/17/2022).

(304) Lu, T.; Chen, F. Multiwfn: A multifunctional wavefunction analyzer. *J. Comput. Chem.* **2012**, *33*, 580–592, DOI: 10.1002/jcc.22885.

(305) Fradera, X.; Austen, M. A.; Bader, R. F. W. The Lewis Model and Beyond. *J. Phys. Chem. A* **1999**, *103*, 304–314, DOI: 10.1021/jp983362q.

(306) Kästner, J.; Carr, J. M.; Keal, T. W.; Thiel, W.; Wander, A.; Sherwood, P. DL-FIND: An Open-Source Geometry Optimizer for Atomistic Simulations. *J. Phys. Chem. A* **2009**, *113*, 11856–11865, DOI: 10.1021/jp9028968.

(307) Verstraelen, T.; Ayers, P. W.; Van Speybroeck, V.; Waroquier, M. Hirshfeld-E Partitioning: AIM Charges with an Improved Trade-off between Robustness and Accurate Electrostatics. *J. Chem. Theory Comput.* **2013**, *9*, 2221–2225, DOI: 10.1021/ct4000923.

(308) Lovell, T.; Li, J.; Liu, T.; Case, D. A.; Noodleman, L. FeMo Cofactor of Nitrogenase: A Density Functional Study of States MN,MOX,MR, and MI. *J. Am. Chem. Soc.* **2001**, *123*, 12392–12410, DOI: 10.1021/ja011860y.

(309) Sippel, D.; Rohde, M.; Netzer, J.; Trncik, C.; Gies, J.; Grunau, K.; Djurdjevic, I.; Decamps, L.; Andrade, S. L. A.; Einsle, O. A bound reaction intermediate sheds light on the mechanism of nitrogenase. *Science* **2018**, *359*, 1484–1489, DOI: 10.1126/science.aar2765.

(310) Kowalska, J. K.; Hahn, A. W.; Albers, A.; Schiewer, C. E.; Bjornsson, R.; Lima, F. A.; Meyer, F.; DeBeer, S. X-ray Absorption and Emission Spectroscopic Studies of [L2Fe2S2]n Model Complexes: Implications for the Experimental Evaluation of Redox States in Iron–Sulfur Clusters. *Inorg. Chem.* **2016**, *55*, 4485–4497, DOI: 10.1021/acs.inorgchem.6b00295.

(311) Cao, L.; Ryde, U. Influence of the protein and DFT method on the broken-symmetry and spin states in nitrogenase. *Int J Quantum Chem* **2018**, *118*, e25627, DOI: 10.1002/qua.25627.

(312) Raugei, S.; Seefeldt, L. C.; Hoffman, B. M. Critical computational analysis illuminates the reductive-elimination mechanism that activates nitrogenase for N2 reduction. *PNAS* **2018**, *115*, E10521–E10530, DOI: 10.1073/pnas.1810211115.

(313) Grunenberg, J. The Interstitial Carbon of the Nitrogenase FeMo Cofactor is Far Better Stabilized than Previously Assumed. *Angewandte Chemie International Edition* **2017**, *56*, 7288–7291, DOI: 10.1002/anie.201701790.

(314) Cao, L.; Ryde, U. What Is the Structure of the E4 Intermediate in Nitrogenase? *J. Chem. Theory Comput.* **2020**, *16*, 1936–1952, DOI: 10.1021/acs.jctc.9b01254.

(315) Cao, L.; Ryde, U. Putative reaction mechanism of nitrogenase after dissociation of a sulfide ligand. *Journal of Catalysis* **2020**, *391*, 247–259, DOI: 10.1016/j.jcat.2020.08.028.

(316) Sugano, S., *Multiplets of Transition-Metal Ions in Crystals*; Academic Press: 1970; 348 pp., ISBN: 978-0-12-431675-1.

(317) Li, L.; Li, C.; Zhang, Z.; Alexov, E. On the Dielectric "Constant" of Proteins: Smooth Dielectric Function for Macromolecular Modeling and Its Implementation in DelPhi. *J. Chem. Theory Comput.* **2013**, *9*, 2126–2136, DOI: 10.1021/ct400065j.

(318) Watson, L.; Eisenstein, O. Entropy Explained: The Origin of Some Simple Trends. *J. Chem. Educ.* **2002**, *79*, 1269, DOI: 10.1021/ed079p1269.

(319) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB - an Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *ChemRxiv (Preprint)* **2018**, DOI: 10.26434/chemrxiv.7246238.v1.

(320) Durrant, M. C. Controlled protonation of iron–molybdenum cofactor by nitrogenase: a structural and theoretical analysis. *Biochemical Journal* **2001**, *355*, 569–576, DOI: 10.1042/bj3550569.

(321) Pomès, R. Theoretical Studies of the Grotthuss Mechanism in Biological Proton Wires. *Israel Journal of Chemistry* **1999**, *39*, 387–395, DOI: 10.1002/ijch.199900044.

(322) Sengwa, R. J.; Sankhla, S.; Khatri, V. Static dielectric constants of the binary mixtures of N-methylformamide with water, ethyl alcohol, ethylene glycol, dimethylsulphoxide, acetone and 1,4-dioxane. *Philosophical Magazine Letters* **2010**, DOI: 10.1080/09500831003757782.

(323) Mills, G.; Jónsson, H.; Schenter, G. K. Reversible work transition state theory: application to dissociative adsorption of hydrogen. *Surface Science* **1995**, *324*, 305–337, DOI: 10.1016/0039-6028(94)00731-4.

(324) Henkelman, G.; Jónsson, H. Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points. *J. Chem. Phys.* **2000**, *113*, 9978–9985, DOI: 10.1063/1.1323224.

(325) Spiller, N. PhD thesis Jupyter notebook, 2022, https://github.com/nspiller/phd_notebook (accessed 04/05/2022).

(326) Kluyver, T.; Ragan-Kelley, B.; Pérez, F.; Granger, B.; Bussonnier, M.; Frederic, J.; Kelley, K.; Hamrick, J.; Grout, J.; Corlay, S.; Ivanov, P.; Avila, D.; Abdalla, S.; Willing, C. In *Positioning and Power in Academic Publishing: Players, Agents and Agendas*, ed. by Loizides, F.; Schmidt, B., IOS Press: 2016, pp 87–90.

(327) Anaconda Software Distribution, version 2-2.4.0, 2020, https://docs.anaconda.com/.

(328) Flöser, B.; Spiller, N. ORCA Helpers: Utility programs for ORCA, 2021, https://gitlab.gwdg.de/orca-helpers/orca-helpers (accessed 03/17/2022).

(329) Drago, R. S., *Physical Methods for Chemists*, 2nd edition; Saunders College Publishing: Ft. Worth, 1992; 750 pp., ISBN: 978-0-03-075176-9.

(330) Römelt, M.; Ye, S.; Neese, F. Calibration of Modern Density Functional Theory Methods for the Prediction of 57Fe Mössbauer Isomer Shifts: Meta-GGA and Double-Hybrid Functionals. *Inorg. Chem.* **2009**, *48*, 784–785, DOI: 10.1021/ic801535v.