A QM/MM study of the nitrogenase MoFe protein resting state: broken-symmetry states, protonation states and QM region convergence in the FeMoco active site

Bardi Benediktsson, Ragnar Bjornsson*
Science Institute, University of Iceland, Dunhagi 3, 107 Reykjavik, Iceland.
* E-mail: ragnar.bjornsson@gmail.com

Abstract

Nitrogenase is one of the most fascinating enzymes in nature, being responsible for all biological nitrogen reduction. Despite decades of research it is among the enzymes in bioinorganic chemistry whose mechanism is the most poorly understood. The MoFe protein of nitrogenase contains an iron-molybdenum-sulfur cluster, FeMoco, where N_2 reduction takes place. The resting state of FeMoco has been characterized by crystallography, multiple spectroscopic techniques and theory (broken-symmetry density functional theory) and all heavy atoms are now characterized. The cofactor charge, however, has been controversial, the electronic structure has proved enigmatic and little is known about the mechanism. While many computational studies have been performed on FeMoco, few have taken the protein environment properly into account. In this study, we put forward QM/MM models of the MoFe protein from *Azotobacter vinelandii*, centered on FeMoco. By a detailed analysis of the FeMoco geometry and comparing to the atomic resolution crystal structure we conclude that only the $[MoFe_7S_9C]^{1-}$ charge is a possible resting state charge. Further we find that, of the 3 lowest energy broken-symmetry solutions of FeMoco the BS7-235 spin isomer (where 235 refers to Fe atoms that are "spin-down") is the only one that can be reconciled with experiment. This is revealed by a comparison of the metal-metal distances in the experimental crystal structure, a rare case of spin-coupling phenomena being visible through the molecular structure. This could be interpreted as the enzyme deliberately stabilizing a specific electronic state of the cofactor, possibly for tuning specific reactivity on specific metal atoms. Finally, we show that the alkoxide group on the Mo-bound homocitrate must be protonated under resting state conditions; the presence of which has implications regarding the nature of FeMoco redox states as well as for potential substrate reduction mechanisms.

Introduction

Nitrogenases are nature's solution to making atmospheric nitrogen available to organisms.[1,2,3] These metalloproteins are the only enzymes that catalyze the reduction of dinitrogen to ammonia according to the reaction equation (in the case of molybdenum-dependent nitrogenase):

$$N_2 + 8e^- + 8H^+ + 16MgATP \rightarrow 2NH_3 + H_2 + 16MgADP + 16P_i$$

The mechanism for N_2 reduction (including binding site) as well as important questions such as why 8 electrons instead of 6 are required and why obligatory H_2 evolution is necessary for NH_3 formation, are all still unsolved problems, though reductive elimination of hydrides has been proposed for the latter[4].

Nitrogenases are multi-protein metalloenzymes that contain unusual and complex iron-sulfur clusters to deal with the difficult problem of N_2 activation. The most active and best characterized enzyme is the molybdenum-dependent nitrogenase which under turnover conditions is a protein complex consisting of the Fe protein (nifH) and MoFe protein (nifDK)[5]. The Fe protein is a homodimer and contains an iron-sulfur cubane cluster and it acts as a reductase. It binds to MoFe protein and reduces it one electron at a time, via a MgATP-dependent process. The MoFe protein is an $\alpha_2\beta_2$ heterotetramer and contains two metal-sulfur clusters in each heterodimer: the P-cluster (an Fe_8S_7 cluster believed to shuttle electrons to the active site) and the FeMo cofactor (FeMoco). FeMoco is an unusual iron-molybdenum-sulfur cluster ($MoFe_7S_9C$) that contains an interstitial atom that was revealed to be a carbon a few years ago[6,7]. It includes a homocitrate ligand and is anchored to the protein via histidine and cysteine residues. FeMoco is believed to be the catalytically active site of the MoFe protein, catalyzing the formation of ammonia from dinitrogen, electrons and protons and produces H_2 under N_2-free conditions.

While there have been many computational studies on FeMoco, only a few studies have considered the protein environment in detail and unfortunately many studies were performed prior to the interstitial carbon assignment. Noodleman and coworkers mainly used small models in their pioneering studies[8,9,10,11] but discussed the effect of the protein environment in a later study[12]. Studies by Norskov[13,14,15,16], Blöchl[17,18], Kästner[19], Szilagyi[20], Dance[21,22] and McKee[23] have used minimal cofactor models. Siegbahn has recently utilized relatively large cluster models to study the mechanism[24]. Cao et al.[25] were the first to perform a quantum mechanics/molecular mechanics (QM/MM) study of the MoFe protein, during the time when the interstitial atom of FeMoco was unknown and concluded that the interstitial atom was an oxygen. Recently Adamo[26] et al. used QM/QM' calculations to propose new mechanisms and Ryde et al. studied protonation states by QM/MM[27]. In previous articles by one of us (RB) we utilized large 225-atom cluster models in our studies of the spectroscopic properties of the FeMo cofactor[28,29,30,31,32]. Large cluster models become difficult to work with and require many constraints to keep the correct active site geometry and may not describe correctly the structural flexibility of the cofactor in the enzyme active site pocket. A better model of such a complex system should take into account both the unique electrostatic environment experienced by the cofactor as well as the flexible protein environment, including specific nearby amino acid sidechains that are known to be important for reactivity. A QM/MM model allows this, as in principle the full protein can be taken into account, while focusing the expensive quantum mechanical description on the important residues (FeMoco and nearby and/or charged residues) while the rest is described by simpler classical forcefields.

In this article, we describe QM/MM modelling of the MoFe protein, starting from the high-resolution crystal structure (PDB code: 3U7Q[6]). We utilize multiple-sized QM regions

and achieve convergence of the FeMoco geometric structure with respect to QM-region size and also demonstrate some convergence for redox properties. We consider all 3 charges of FeMoco that have been proposed in the literature and conclude that only the $[MoFe_7S_9C]^{1-}$ charge results in a geometry that is in acceptable agreement with the crystal structure, an assignment in agreement with recent spectroscopic and theoretical work. Importantly, we consider all 3 spin isomers of the broken-symmetry (BS) BS7 solution that is the accepted lowest energy BS solution of FeMoco. We find that the metal-metal distances of the BS7-235 solution are in much better agreement with the crystallographic metal-metal distances of FeMoco suggesting that the protein is stabilizing an electronic structure that is reasonably described by this particular broken-symmetry solution. This can be rationalized via an understanding of FeMoco in terms of localized and delocalized electrons on specific metal atoms. Finally, we show that the alkoxide group on the Mo-bound homocitrate is protonated under resting state conditions. This proton may play an important mechanistic role during N_2 reduction, possibly becoming a hydride in other redox states or even becoming a substrate during the H_2 evolution function of nitrogenase.

Computational details

Force Field Parameters

The CHARMM36 protein forcefield[33] was used in all MM and QM/MM calculations with minor modifications. As no forcefield parameters are available for complex metal clusters such as FeMoco and the P-cluster, we derived nonbonded parameters for these clusters and kept them frozen during all classical optimizations or MD simulations. Atomic charges for these residues were derived from BP86[34,35]/def2-TZVP[36] calculations of the clusters using natural population analysis[37]. Lennard-Jones (L-J) parameters for these clusters were similar to what has previously been applied to iron-sulfur clusters in MM simulations[38]: the CHARMM atom-type SM was used for inorganic sulfides and no Lennard-Jones sites were assigned to Fe or Mo. For homocitrate, atomic charges and L-J parameters were derived from parameters available for citrate.[39] Free imidazole molecules present in the crystal structure were not deleted from the structure and parameters for them were taken from CGenFF (Charmm general force field)[40]. All cysteines bound to FeMoco and P-cluster were modelled as deprotonated cysteines using parameters available in CHARMM36.

MM model preparation and solvation

The whole MoFe protein was modelled classically. The initial structure is based on the crystal structure of MoFe protein from *Azotobacter vinelandii* at an atomic resolution of 1.0 Å (PDB code: 3U7Q)[6]. Modifications were made to the basic structure: two calcium ions at the interface of subunits were changed to Fe(II) ions, according to the discovery of these sites as iron ions[41]. Only the amino acids present in the crystal structure were modelled, residues present in the peptide sequence but missing from the crystal structure were not added to the model. As the MoFe protein is a heterotetramer and containing two α and two β subunits, we will use the nomenclature used in the 3U7Q file when referring to a residue that belongs to a certain subunit, namely the A (α), B (β), C (α) and D (β) subunits. As an example, the residue number 153 refers to glutamate in subunits A and C. When referring to the residue in chain A, we will use "Glu153A" but when referring to both we will use "Glu153AC".

GROMACS, version 5.0.4[42,43,44] was used to set up the original MM model and add missing hydrogens. Protonation states of titrable residues were determined using manual inspection of hydrogen bonding patterns and also with the aid of PROPKA[45,46]. Due to being a heterotetramer, each residue appears twice in the protein (in each heterodimer). The symmetry was used in protonation state determination, that is, if a residue is protonated in the A subunit, it is also protonated in the C subunit. All aspartate and glutamate residues were modelled as deprotonated except Glu153AC. All lysine residues were determined to be protonated except Lys365BD. The protonation states of all histidine residues were checked and determined via careful manual inspection. The following histidine residues are protonated at the epsilon nitrogen His80AC, His83AC, His195AC, His362AC, His383AC, His442AC, His106BD, His193BD, His311BD, His363BD, His392BD, His429BD, His457BD, His477BD, His478BD and His480BD. The histidine residues protonated at delta nitrogen were His31AC, His196AC, His285AC, His451AC, His185BD, His297BD, His396BD and His519BD. Only His90BD was doubly protonated at both epsilon and delta nitrogen atoms. Water molecules present in the crystal structure as oxygen atoms were kept and hydrogens added to the oxygen atoms using GROMACS. The proton positions were manually inspected for the water pocket around the homocitrate residue in the active site.

After full hydrogenation of the MoFe protein, the number of atoms is 39566 atoms (before protonation there are 16295 atoms). The system was next solvated using GROMACS procedures by placing the protein inside a 90 x 90 x 90 Å box and filling it up with TIP3P waters. Water molecules present in the crystal structure were kept and modelled as TIP3P as well. The system has a total negative charge of -39 after the hydrogenation step. To balance the charge, we added 39 Na^+ counterions to the solvent. After the solvation and addition of counterions, the system consists of 320829 atoms and is shown in Figure 1a.

At this point, 2 different MM models were created: i) a constrained X-ray model ii) an unconstrained MM model. The constrained X-ray model kept all protein heavy-atom (non-hydrogen atoms) positions frozen (not waters) during relaxation and molecular dynamics steps while the unconstrained MM model kept the following residues frozen in space: FeMoco, homocitrate, P-cluster, 7492-Fe(II), 6492-Fe(II) and had additionally the following atomic constraints: delta-nitrogen of the the Mo-bound His442, sulfur atoms of all deprotonated cysteine and all oxygen atoms of residues connected to residue 6492 and residue 7492 (the two Fe ions). The constrained X-ray model is thus deliberately biased towards the experimental structure while the unconstrained model has no such bias but can suffer from artifacts due to dynamical events in the active region resulting from the approximate forcefield used for the complex metalloprotein. MD simulations were performed for both the constrained X-ray model and the unconstrained MM model, in the canonical ensemble using the velocity Verlet algorithm. The constrained model was simulated for 1 ns while the unconstrained model was simulated for 5 ns. In addition to the above-mentioned constraints, bond constraints using the LINCS algorithm[47,48] were applied to all X-H bonds in order to maintain a 1 fs timestep during the simulation. A Nosé-Hoover thermostat[49,50,51,52] with chain number of 4 with coupling to the whole system, was used for heating and maintaining a simulation temperature of 300 K. The system was gradually heated from 50 K to 300 K in 0-500 ps for the unconstrained model but in 0-100 ps for the constrained model. For the unconstrained model, the average RMSD of heavy protein atoms converged quickly

to a value of 0.31 Å after approximately 1 ns and maintained this value for the rest of the 5 ns simulation, demonstrating that our classical modelling is reasonable.

QM/MM preparation and calculations

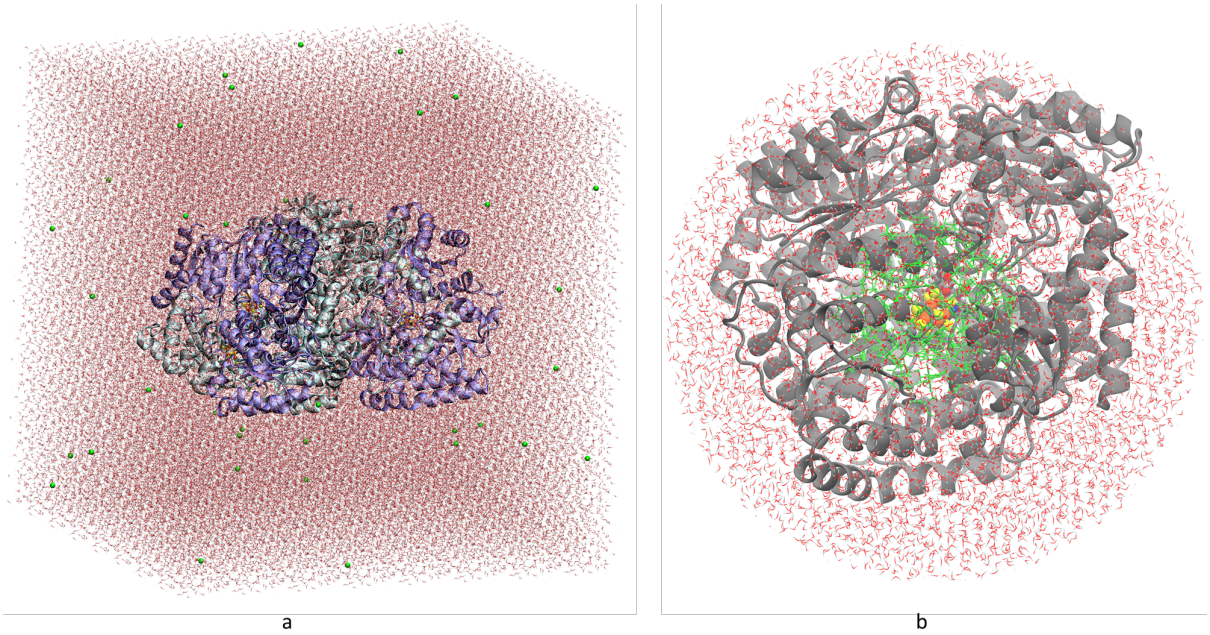


Figure 1: a) The 320 K atom MM model (in a periodic box) showing the whole heterotetrametic MoFe protein dissolved in 93754 solvent molecules and 39 Na^+ ions. b) The spherical QM/MM model centered on one FeMoco unit (in one of the heterodimers) showing a minimal QM region (ball-and-stick) and the MM region divided up into frozen (black and red/white) and active regions (green). The spherical model contains whole chains A and B and a part of chain C.

Spherical QM/MM models (as shown in Figure 1b) were next cut from the 320K atom MM box from the two separate MD simulations on the constrained and unconstrained MM models. The constrained model used a snapshot after 940 ps while the unconstrained after 1862 ps. This spherical droplet model is used to facilitate QM/MM geometry optimizations of large proteins as popularized by Thiel and coworkers for years[53,54,55] and has been confirmed to be a highly economical yet accurate QM/MM modelling strategy according to a recent comparison to periodic QM/MM calculations[56]. The models are centered on the carbide atom of FeMoco present in the A,B-chain half of the MoFe protein and all whole residues (amino acids and waters) within 42 Å are included. Models were cut so that they remained neutral (24 Na^+ ions included in each spherical model). The entire A and B chains were included; as some part of the C chain was inside the 42 Å radius, a longer chain of it was included for a more natural cut. That includes residues 242, 320-326, 342-369, 378-392, 410-419, 437-443, 458-469, 476 and 468-523. The constrained model contains one amino acid residue less, residue Ile443C. While the MM preparation steps were the same for constrained and

unconstrained MM models (MD step differed as mentioned), the constrained and unconstrained QM/MM models thus ended up being slightly different: the unconstrained MM model contains 37060 atoms and the QM/MM version of the constrained MM model containing 36989 atoms. Aside from the slightly different spherical cut, there were also some differences due to differing amount of water molecules. Most calculations described in this article were performed using the constrained QM/MM model. The reason is that due to water dynamics in the active site around FeMoco during the MD simulation of the unconstrained model, some water molecules pushed residues like His195 away from its crystallographic positions, resulting in geometries that seems incompatible with all the crystal structures available for MoFe protein. These water molecules had to be deleted prior to QM/MM geometry optimizations. We confirmed, however, that similar FeMoco geometries are obtained for the unconstrained model as shown in the supporting information.

Chemshell version 3.7[57,58] was used for all QM/MM calculations. Scripts in Python and Tcl to convert the GROMACS MM model format to Chemshell format were written (available as supporting information). Additional Chemshell code was written to facilitate QM-region and active regions definitions. The ORCA quantum chemistry code (version 3.0.3)[59] was interfaced to Chemshell via a modified Chemshell-ORCA interface that allows for electrostatic embedding. The modified interface allows easier use of any QM protocol available in ORCA and allows geometry optimizations on the broken-symmetry surface via the spin-flipping procedure in ORCA. Broken-symmetry solutions were found by converging first to a high-spin ferromagnetic solution of (e.g. M_S=35/2 for the $[MoFe_7S_9C]^{1-}$ charge), then flipping the spin on specific atoms and converging to the broken-symmetry state with M_S=3/2 (for resting state FeMoco). All QM/MM calculations used electrostatic embedding and link atoms were used to terminate the QM-MM border together with the charge-shift procedure as implemented in Chemshell[60]. In electrostatic (a.k.a. electronic) embedding the MM charges are included directly in the QM calculation, polarizing the QM electron density. For the QM part, the TPSSh hybrid density functional[61,62] with D3BJ dispersion correction[63,64] and the ZORA relativistic approximation[65,66] was used and employing the relativistically contracted def2 Ahlrichs basis set[36,67]. A triple-zeta ZORA-def2-TZVP basis set was used for all Mo, Fe and S atoms (and interstitial carbide of FeMoco) while a double-zeta ZORA-def2-SVP basis set was used for all other atoms. The RIJCOSX approximation[68,69] was used to speed up Coulomb and Exchange integrals. An example ORCA inputfile used in the QM/MM and QM calculations is shown in the supporting information. Different QM-regions for the active site were defined and used as shown in Figure 2 (and the supporting information). The MM part was calculated using DL_POLY[70] as implemented in Chemshell using the previously defined modified CHARMM36 forcefield for the MoFe protein. The QM/MM geometry optimizations were done using the DL-FIND[71] program inside Chemshell. The VMD program[72] was used to visualize results and was used to render the molecular images shown in this article.

In all QM/MM calculations a large approximately spherical active atoms region is defined. All atoms within this sphere (either described at the QM or MM level) are free to move during the geometry optimization while the rest is kept frozen. When defining the active atoms region, all protein residues (or water molecules) that contain an atom within a certain radius are included in this selection. A 11 Å active region, containing all residues with an atom within 11 Å of the central carbide of FeMoco, was used for most calculations

(corresponds to 1002 active atoms in the constrained model but 978 atoms in the unconstrained model) but large QM-region calculations (larger than 154 atoms) used a larger 13 Å radius active region (corresponds to 1576 active atoms in the constrained model but 1596 atoms in the unconstrained model). Active regions are used out of convenience, due to the large degrees of freedom present in the geometry optimization problem and to avoid falling into new local minima during each new optimization[54]. The QM/MM optimization results were overall not very sensitive to the size of the active region.

Results and Discussion

The QM/MM calculations performed allow a better modelling of the molecular structure of FeMoco than simplified cluster models and as we will show, even allow insight into the electronic structure. Previous cluster models have kept residues surrounding the cofactor either frozen or partially frozen in order to keep the cluster model consistent with the geometry imposed by the protein. As our QM/MM model incorporates a much larger part of the whole protein, no amino acids near FeMoco need to be frozen (the active atoms region is chosen to be large, ~1000 atoms or more). Furthermore, the QM/MM model can be systematically improved by increasing the QM region. In this study, multiple QM regions of increasing size were explored; Figure 2 shows the most important QM region models while the supporting information shows other defined QM regions and additional details. The effect of QM-region size on the structure as well as the redox properties of FeMoco are discussed in the last section. We start off discussing the charge of the cofactor, this is followed by a discussion of broken-symmetry solutions and spin isomers and their effect on the FeMoco geometry, we then discuss the protonation state of Mo-bound homocitrate and end on QM region convergence effects in our QM/MM calculations.

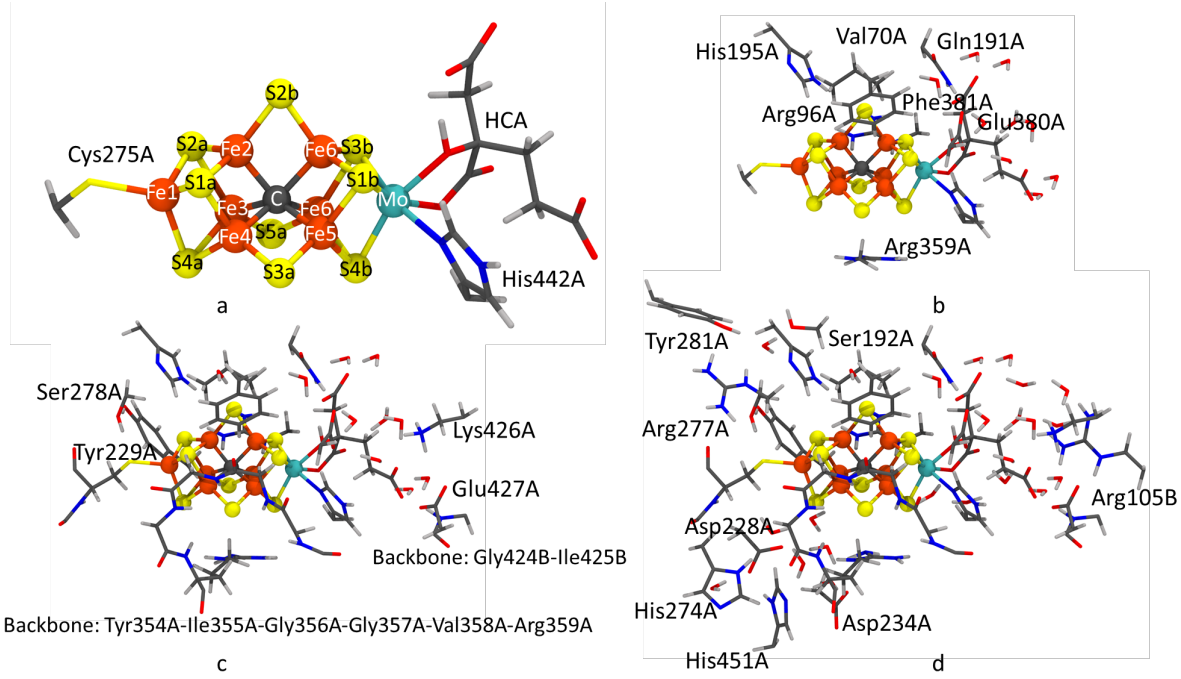


Figure 2. Examples of the QM regions used in the QM/MM calculations (link atoms not shown). Additional details about the QM regions are available in the SI. a) 54 QM atom region consisting of the [$MoFe_7S_9$] cluster, homocitrate (HCA), His442A truncated as methylimidazole and Cys275A truncated as methylthiolate. b) 154 QM atom region that additionally includes the sidechains of

Val70A, Arg96A, His195A Arg359A, Glu380A, Gln191A, Phe381A and 9 water molecules. c) 247 QM atom region that additionally contains Tyr229A, Ser278A, Lys426A, Glu427A, the carbonyl group of the Tyr354A peptide bond, amide cage around S3B (S8) consisting of the peptide backbone of residues Tyr354A-Ile355A-Gly356A-Gly357A-Val358A-Arg359A, the whole residue of Arg359A, larger part of residue Cys275A, peptide bond between residues Gly424B and Ile425B and an additional water molecule above S2A d) 368 QM atom region that includes additionally the sidechains of Ser192A, Asp228A, Asp234A, His274A, Arg277A, Tyr281A, His451A, Arg105B and 12 additional water molecules.

A. QM/MM FeMoco geometry: effect of cofactor charge

The charge on FeMoco has been controversial in nitrogenase research as discussed in a recent article by one of us[32]. Early spectroscopic work[73,74,75,76] suggested metal oxidation states of Mo(IV)6Fe(II)1Fe(III), corresponding to $[MoFe_7S_9C]^{3-}$ (sulfides and carbides taken in their closed-shell forms, S^{2-} and C^{4-}) while later work[77] suggested Mo(IV)4Fe(II)3Fe(III), corresponding to $[MoFe_7S_9C]^{-}$, to be more likely. A more recent computational study[20] suggested a more oxidized charge: $[MoFe_7S_9C]^{+}$. The total spin of resting state FeMoco is experimentally known to be S=3/2[1] making these 3 odd-electron charges the only reasonable possibilities. In a recent study by one of us, the Mössbauer properties of the cofactor were calculated[32] and it was demonstrated that only the $[MoFe_7S_9C]^{-}$ charge fits the experimental Mössbauer parameters well. This is also in line with a recent spatially resolved anomalous dispersion refinement (SpReAD study)[78] that shows 3 Fe ions to be more reduced than the other Fe ions (according to the Fe XAS edge positions). This would be in agreement with a formal oxidation state assignment of Mo(III)3Fe(II)4Fe(III) that corresponds to a charge of $[MoFe_7S_9C]^{-}$.

It is nonetheless of interest to use all available experimental data and structural data is both fundamental, accessible from computations and a vital test for a computational model. In cluster modelling studies it is always an open question of how well the environment is being described while a QM/MM model should be describing the local FeMoco environment better than cluster models due to the inclusion of the protein environment in the model. Previous attempts by us and others[7] to deduce the cofactor charge by a geometric analysis have proved futile. There is also the question of the reliability of the experimental crystal structures (particularly when heavy elements with bonded light atoms are involved). With better modelling of the FeMoco environment being performed in this study, it is therefore interesting to compare the optimized QM/MM structures for the 3 proposed charges of FeMoco and compare to the atomic resolution crystal structure available and see whether clear differences are obvious and what insights can be drawn from the molecular structure. Table 1 shows root-mean-square-deviations (RMSD) for QM/MM optimized structures of FeMoco with the 3 different charges, using a QM region of 154 atoms (see Figure 2b). RMSDs are shown for both the pure metal cluster (metal-sulfur-carbide part) and also when including in the RMSD calculation the homocitrate, the methylimidazole of Mo-bound His442 and the methylthiolate group of the Fe1-bound Cys275 residue (hydrogens not included). The results show that when considering the $[MoFe_7S_9C]$ part, the -1 charge of the cofactor is considerably more likely as the RMSD is approximately half of the RMSD of the +1 charge and more than half of the RMSD of the -3 charge. When considering also the

ligands bound to the metals, the RMSD values still favor strongly the -1 charge as the likely resting state charge.

Table 1: Root-mean square-deviations (RMSD) in Å of different QM/MM-optimized geometries (154-atom QM models were used, see Figure 2b) with different FeMoco charges compared to the experimental crystal structure (cofactor bound to chains A and B; PDB ID: 3U7Q[6]). The broken-symmetry solution BS7-235 was used in all calculations (see later discussion) and homocitrate was in its singly protonated form (see later). Table shows RMSD values for both the $[MoFe_7S_9C]$ part of the cofactor and the $[MoFe_7S_9C]$+homocitrate(HCA)+His442-sidechain+Cys275-SCH_2 group where only heavy atoms were included in the calculation of RMSD.

Charge on cofactor	$[MoFe_7S_9C]$	$[MoFe_7S_9C]$+HCA+His442+Cys275
$[MoFe_7S_9C]^{+}$	0.082	0.137
$[MoFe_7S_9C]^{-}$	0.044	0.104
$[MoFe_7S_9C]^{3-}$	0.097	0.139

Table 2: Various bond lengths and atom-atom distances (in Å) in the cofactor from the experimental crystal structure (average of both cofactors in protein), QM/MM calculations (154 QM atoms) of the 3 different charges and QM/MM calculations of the 3 different BS7 spin isomers for the $[MoFe_7S_9C]^{1-}$ charge.

	3U7Q-ave	**$[MoFe_7S_9C]^{+}$**	**$[MoFe_7S_9C]^{-}$**	**$[MoFe_7S_9C]^{3-}$**	**BS7-235 $[MoFe_7S_9C]^{-}$**	**BS7-247 $[MoFe_7S_9C]^{-}$**	**BS7-346 $[MoFe_7S_9C]^{-}$**
Fe1-Fe2	2.67	2.68	2.67	2.66	2.65	2.63	2.67
Fe1-Fe3	2.67	2.65	2.63	2.67	2.64	2.67	2.64
Fe1-Fe4	2.66	2.80	2.64	2.75	2.65	2.65	2.66
Mo-Fe5	2.73	2.70	2.72	2.77	2.71	2.64	2.64
Mo-Fe6	2.67	2.64	2.66	2.84	2.65	2.64	2.71
Mo-Fe7	2.68	2.62	2.61	2.62	2.62	2.69	2.61
Fe2-Fe3	2.67	2.82	2.66	2.64	2.65	2.64	2.64
Fe2-Fe4	2.65	2.64	2.62	2.64	2.63	2.64	2.63
Fe3-Fe4	2.64	2.64	2.62	2.64	2.62	2.61	2.59
Fe5-Fe6	2.63	2.62	2.63	2.66	2.63	2.57	2.63
Fe5-Fe7	2.63	2.64	2.64	2.66	2.61	2.62	2.56
Fe6-Fe7	2.60	2.53	2.56	2.55	2.56	2.61	2.61
Fe1-Mo	7.00	6.96	6.96	7.10	6.96	6.96	6.97
Fe1-C	3.54	3.44	3.46	3.52	3.46	3.46	3.49
Mo-C	3.46	3.52	3.50	3.58	3.50	3.50	3.49
Fe2-Fe6	2.58	2.56	2.59	2.60	2.59	2.58	2.57
Fe3-Fe7	2.58	2.57	2.58	2.62	2.60	2.58	2.59
Fe4-Fe5	2.61	2.57	2.60	2.59	2.60	2.62	2.63
C-Fe2	2.01	2.00	1.99	2.01	1.99	1.99	1.98
C-Fe3	1.99	2.01	2.01	2.02	2.01	1.99	2.00
C-Fe4	2.00	1.95	1.99	1.98	1.98	2.00	2.00
C-Fe5	2.01	2.02	1.98	1.99	1.97	1.99	1.99
C-Fe6	2.01	2.01	2.00	1.99	2.00	2.00	1.98
C-Fe7	1.99	2.00	2.00	2.01	1.99	1.98	1.98
Mo-O_{alk}	2.17	2.09	2.17	2.20	2.17	2.16	2.17
Mo-O_{carbox}	2.21	2.13	2.20	2.25	2.20	2.22	2.21
Mo-N_{His}	2.34	2.26	2.27	2.25	2.28	2.28	2.27
Fe1-S_{Cys}	2.27	2.21	2.27	2.25	2.28	2.26	2.27

It is even more useful to consider specifically the metal-metal distances within the cofactor. Table 2 contains a list of most of the relevant distances for the cofactor for all 3 charges while Figure 3 shows more convenient diagrams of the basic metal-skeleton of the cofactor and the important metal-metal distances. Due to resolution limitations and potential disorder in the crystal structure as well as systematic errors in the DFT calculations (plus lack of vibrational motion), we emphasize that trends in the distances should be more relevant than comparison

of absolute distances when comparing to the crystal structure data. The experimental distances are shown in Figure 3d where it is notable that the Mo-Fe5 distances is longer than the Mo-Fe6 and Mo-Fe7 distances by 0.05-0.06 Å, the Fe6-Fe7 distances being slightly shorter (by 0.03 Å) than the Fe5-Fe6 and Fe5-Fe7 distances in the triangle near Mo while the Fe2-Fe3 distance is longer than Fe2-Fe4 and Fe3-Fe4 (0.02-0.03 Å) in the triangle farther from Mo. When comparing the same distances for the calculated geometries with different charges it is immediately noticeable that only the $[MoFe_7S_9C]^{1-}$ charge reproduces most of these trends in distances. The $[MoFe_7S_9C]^{3-}$ charge geometry (Figure 3c) has a noticeably large Mo-Fe6 distance (2.84 Å or 0.22 Å longer than the shorter Mo-Fe7) in sharp contrast with the experimental structure (0.06 Å difference between the shortest and longest Mo-Fe distance) and does not show a longer Fe2-Fe3 distance; these deviations seem large enough to rule out $[MoFe_7S_9C]^{3-}$ as the cofactor resting state charge. The $[MoFe_7S_9C]^{1+}$ charge (Figure 3a) has Fe1-Fe4 and Fe2-Fe3 distances that are much too large compared to the crystal structure. On the other hand, $[MoFe_7S_9C]^{1-}$ (Figure 3b) gives a structure in relatively good agreement with the distance trends seen in the crystal structure: The Mo-Fe5 distance is slightly larger than the Mo-Fe6 and Mo-Fe7 lengths in agreement with the experimental geometry (although the Mo-Fe6 and Mo-Fe7 are not as similar as in the experimental structure). The Fe6-Fe7 distance is notably smaller than the Fe5-Fe6 and Fe5-Fe7 distances as in the experimental structure and the Fe2-Fe3 distance is slightly longer than the Fe2-Fe4 and Fe3-Fe4 distances, also in good agreement. No unusually large or small metal-metal distances are seen for the $[MoFe_7S_9C]^{1-}$ geometry.

We note that the $[MoFe_7S_9C]^{1-}$ model does not reproduce all distances (but most of the trends) in the crystal structure but the geometric differences between FeMoco charges (which is reflected in the bond lengths due to metal oxidation state differences) are large enough in our opinion to completely rule out the $[MoFe_7S_9C]^{1+}$ and $[MoFe_7S_9C]^{3-}$ charges as viable models for the resting state of FeMoco. That this can be established so clearly, demonstrates the utility of our new QM/MM models of MoFe protein.

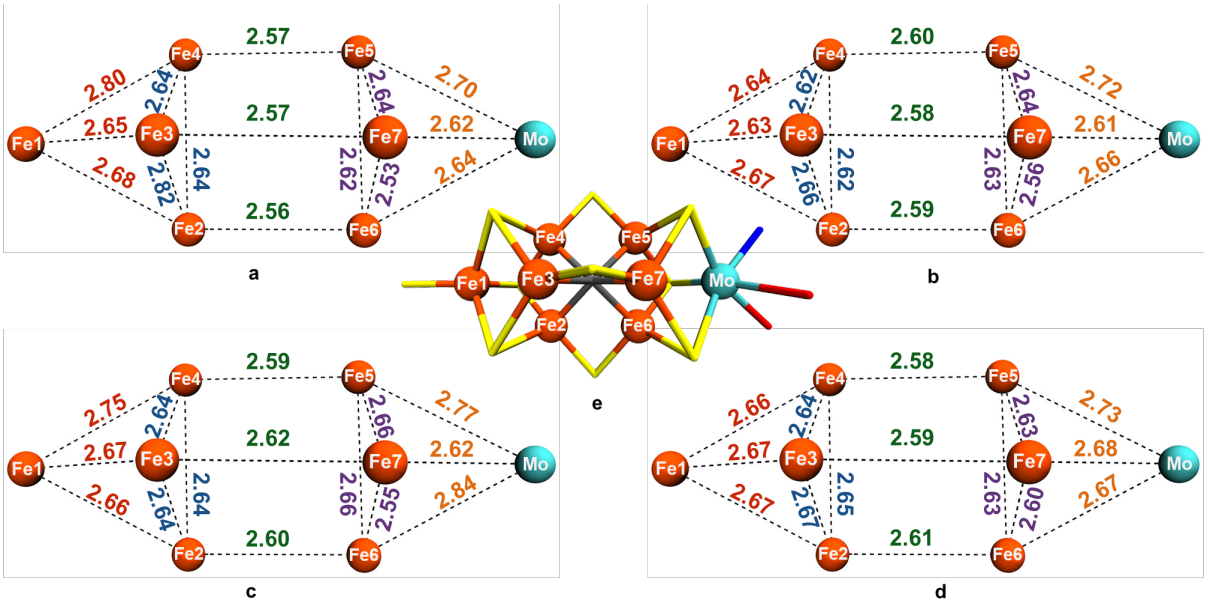


Figure 3: Metal-metal distances (in Å) in FeMoco structures with different charges and the experimental crystal structure. 154 QM-atom models were used and the BS7-235 solution was used for all charges. a) $[MoFe_7S_9]^{1+}$, b) $[MoFe_7S_9]^{1-}$, c) $[MoFe_7S_9]^{3-}$, d) Average of the two cofactors of the 3U7Q FeMoco crystal structure. e) The full FeMoco structure

B. QM/MM FeMoco geometry: effect of broken-symmetry spin isomers

As concluded in the last section, only the $[MoFe_7S_9C]^{1-}$charge is likely to be the resting state based on this new geometric analysis that is in agreement with the most recent analysis of the Mössbauer properties[32] as well as the recent spatially resolved anomalous dispersion refinement (SpReAD) study[78]. All of our QM/MM calculations have used broken-symmetry SCF solutions (with M_S=3/2) using the broken-symmetry solution known as BS7 as proposed by Noodleman[11]. This broken-symmetry solution has been found by many research groups[11,18,20,21,7,28] to be the lowest energy one which can be explained as being mainly due to presence of more favorable antiferromagnetic interactions between Fe atoms than other BS solutions. As pointed out in an article by one of us[28], the molybdenum atom (a Mo(III) ion according to XAS and theory[28,29]), also takes part in the spin coupling (and is even part of weak metal-metal bonding with Fe5, Fe6 and Fe7, however, its electrons are flipped on their own during the SCF procedure and usually does not need to be considered when considering the spin-flipping problem (in rare cases though, the Mo atom ends up in a high-energy electron configuration). We have confirmed this multiple times that the BS7 solution originally proposed by Noodleman is the lowest energy solution that can be found, despite the Mo atom not being part of the spin-flipping problem. Curiously in the lowest energy BS7 state the Mo(III) ion is in a non-Hund electronic configuration (due to the interactions with Fe atoms) as previously discussed[28] and seems to be a general feature of Mo-Fe-S cubane-like clusters[79]. What is seldom pointed out (although by Noodleman[9,8,12]) and we made a point of in a recent article[32] is that each broken-symmetry solution, including BS7, actually has 3 different spin isomers that all need to be considered when calculating a cofactor in an environment where approximate C_3 symmetry is no longer present. These 3 spin isomers

should be energetically equivalent in a perfect symmetric environment but might be expected to differ energetically in a non-symmetric protein environment and due to the presence of mixed-valence delocalized pairs[80,81] as previously discussed[32]. The 3 different spin isomers lead to the mixed-valence delocalized pairs as well as more localized metal spin sites swapping positions. These 3 spin isomers of the BS7 solution have here been labelled as: BS7-235, BS7-346 and BS7-247 where the latter 3 numbers refer to which Fe atoms are flipped to find the energetically favored BS7 class of solutions. Previous analysis[32] using a cluster model showed that these spin isomers are energetically very similar, within ~1 kcal/mol of each other at the TPSSh level but it was not possible to to distinguish between them by energy or Mössbauer properties.

We were curious whether considering these 3 spin isomers within a QM/MM model might energetically stabilize one of these solutions or perhaps whether the geometric differences caused by different locations of localized/delocalized Fe atoms would reveal evidence of one of these electronic states being stabilized within the protein. The crystal structure might even be an average of all 3 of them.

As the previous analysis of the FeMoco charges suggested, the crystal structure does show some clear differences in certain Fe-Fe distances, possibly pointing to specific locations of delocalized Fe-Fe pairs for example. It should be pointed out that this also presents an important test for broken-symmetry DFT methodology as it remains unclear how capable this approximate methodology is at describing the electronic structure for such an exotic spin-coupled cofactor as FeMoco. Broken-symmetry solutions are after all not eigenfunctions of the total spin operator and the approximate exchange-correlation functionals used in DFT are far from perfect.

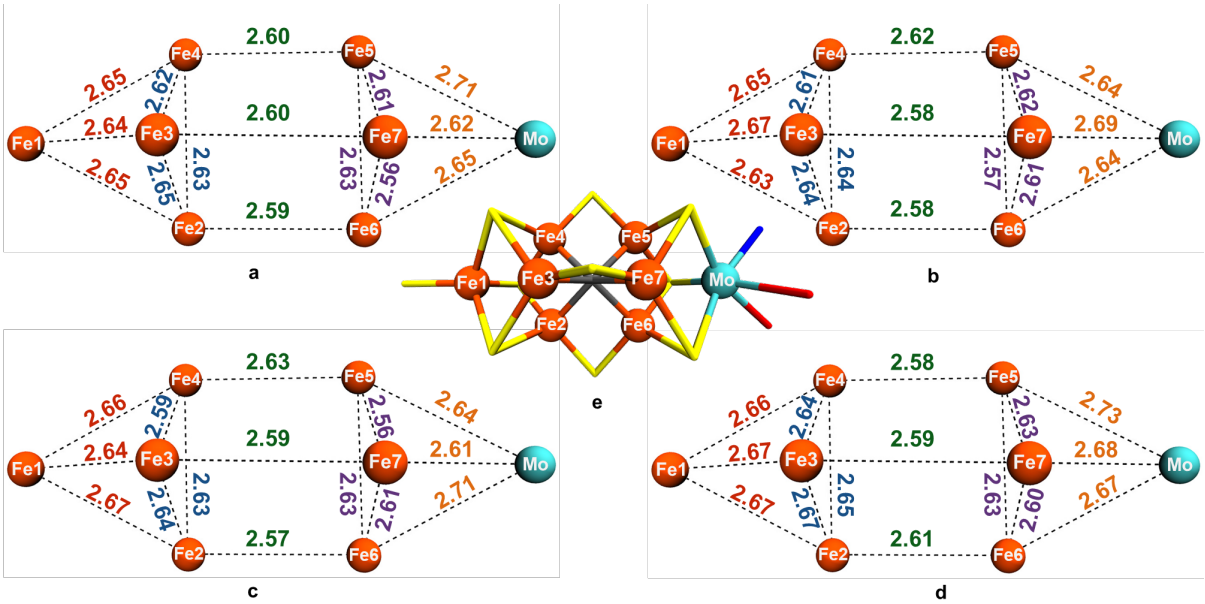


Figure 4: Comparison of the metal-metal distances of FeMoco for the $[MoFe_7S_9C]^{1-}$ charge when comparing different spin isomers of the BS7 solution. The QM/MM model with 247

QM atoms as seen in Figure 2c was used in all calculations. a) BS7-235, b) BS7-247 c) BS7-346, d) Average of the two cofactors of the 3U7Q FeMoco crystal structure. e) The full FeMoco structure

For these spin isomer comparisons we performed QM/MM calculations with an even larger QM region than previously used, 247 atoms (see Figure 2c). RMS deviations with respect to the crystal structure interestingly reveal that the BS7-235 solution is slightly closer to the crystal structure than the other solutions with RMSD values of 0.044 Å (BS7-235), 0.050 Å (BS7-247) and 0.074 Å (BS7-346) for the [$MoFe_7S_9C$] part) but the differences are probably too small to differentiate between BS solutions.

Figure 4 shows the metal-metal distances for all 3 spin isomers compared to the crystal structure. It is remarkable that the Mo-Fe5 distance in the crystal structure, that is notably longer than the Mo-Fe7 and Mo-Fe6 distances, is also the longest for the BS7-235 geometry while for the BS7-247 and the BS7-346 solutions the Mo-Fe7 and Mo-Fe6 respectively are the longer molybdenum-iron bond distances. This is unlikely to be a coincidence and can be explained quite easily by an understanding of the electronic structure in terms of delocalized Fe-Fe pairs and localized Fe sites (as has been discussed in a recent article[32]). For the BS7-235 solution the Fe5 atom is down-spin, effectively creating a localized Fe(III) site (since no other Fe atom has spin parallel to it) while Fe6 and Fe7 become involved in a delocalized mixed-valence Fe(2.5)-Fe(2.5) pair (we assume complete delocalization at this point). This relatively simple picture (that comes from analysis of the localized orbitals that is shown in the supporting information) explains well why the Fe6-Fe7 distance is so short both in the experimental structure and the BS7-235 structure, why the Mo-Fe6 and Mo-Fe7 distances are so similar in the experimental crystal structure (BS7-235 shows some asymmetry here) and finally why the Fe5-Fe6 distance is short in the BS7-247 solution and the Fe5-Fe7 distance in the BS7-346 solution. It is all a question of where the Fe(2.5)-Fe(2.5) pair ends up (and the localized Fe(III) site).

A similar albeit different comparison can be seen in the other Fe triangle. There the distances are dependent on the BS solution and clearly where the delocalized vs. localized Fe sites end up but curiously the delocalized pair seems to result in a longer distance for the BS7-235 solution instead of a shorter one. This is not easily understood at present (mixed-valence delocalization in iron-sulfur clusters is still not well understood[82]) but importantly this asymmetry is seen in both the experimental structure and in the BS7-235 solution but not the other solutions and BS7-235 is the only one that reproduces the slightly longer Fe2-Fe3 distance. We think this difference between Fe triangles in the cofactor may reflect some competition between all the spin-couplings, weak metal-metal bonding and other interactions within the cofactor.

The relative QM/MM energies of BS7 spin isomers show that they fall in a range of 1.1 kcal/mol with BS7-346 being the lowest in energy, BS7-235 being 0.7 kcal/mol higher in energy and BS7-247 1.1 kcal/mol higher in energy. These relative energies (all at the TPSSh level) are similar to what has been found before[32] but due to the small differences we do not think these energies help us deduce the correct BS solution. After all, there is nothing to indicate that our calculations come close to chemical accuracy (~ 1 kcal/mol).

To conclude, we have been able to show that the FeMo cofactor appears to be in a specific electronic state according to a comparison of the metal-metal distances in the cofactor when considering 3 practically equivalent electronic structure solutions and the experimental crystal structure. The relative energies of the solutions are within ~1 kcal/mol of each other and are too close to confidently deduce anything while the geometric analysis paints a very clear picture of the FeMo cofactor being in a specific electronic state, approximately described by the BS7-235 broken-symmetry solution. We cannot put much faith in the relative energies due to the approximations in use (approximate exchange-correlation functionals and the use of determinants that are not eigenfunctions of the total spin operator) but it is conceivable that 3 genuine electronic states of the cofactor are this close in energy, all of which could be populated at experimental conditions. Perhaps more likely, there is a larger energy difference between these states at experimental conditions and that the state as characterized by broken-symmetry DFT, BS7-235, is the most populated as suggested strongly by the metal-metal distance analysis. With progress being made in multiconfigurational and multireference wavefunction theory it is our hope that such calculations (that were e.g. used to calculate the electronic structure of Fe_4S_4 cubane clusters[82]) can shed further light on this issue in the future. It would be very important to find out if the picture suggested by broken-symmetry DFT of delocalized pairs and localized Fe(II)/Fe(III) sites holds within a multiconfigurational wavefunction picture.

Finally, it is important to note that the BS7-235 broken-symmetry solution and our picture of FeMoco involving specific delocalized pairs/localized Fe(II)/Fe(III) sites (see Figure 5) is not inconsistent with the results that came from spatially resolved anomalous dispersion refinement[78] (SpReAD) where 3 Fe atoms (no. 1, 3 and 7) were found to be more reduced (similar to the P-cluster) than the others. This suggests an oxidation state distribution of Mo(III)3Fe(II)4Fe(III) while the BS-DFT calculations suggests something closer to Mo(III)1Fe(II)4Fe(2.5)2Fe(III) according to localized orbital analysis. However, as we pointed out[32], it is entirely possible that spin localization could occur in these mixed-valence pairs that we suggest are present in Fe2-Fe3 and Fe6-Fe7. Should the spin localize somewhat, e.g. by vibronic coupling, it is entirely possible that Fe3 and Fe7 would have more Fe(II) character than Fe(2.5) character. We also note that the minority-spin electrons in these mixed-valence pairs are not always completely delocalized as can be seen in the localized orbitals (see supporting information) and are furthermore dependent on the QM-region and density functional used. Finally we note that XAS edges of mixed-valence Fe dimers are not fully understood as discussed by DeBeer et al.[83]

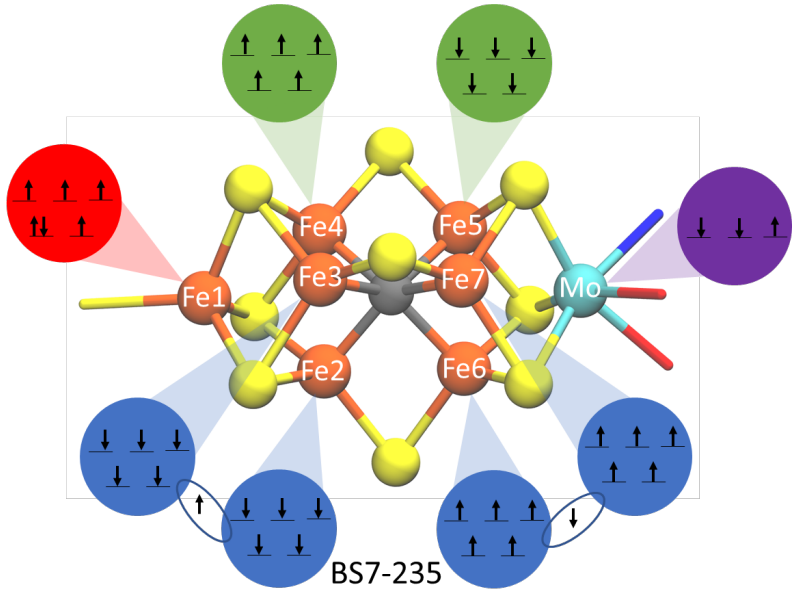


Figure 5. Our current understanding of the electronic structure of FeMoco, featuring the $[MoFe_7S_9C]^{1-}$ charge and the BS7-235 broken-symmetry solution based on an analysis of the localized orbitals (see supporting information). Fe6 and Fe7 as well as Fe2 and Fe3 are found to be mixed-valence delocalized pairs. A partially delocalized electron between Fe1 and Fe4 is also found but this electron ends up being more localized on Fe1, making the latter's oxidation state closer to Fe(II). Fe4 and Fe5 can be interpreted as having more Fe(III) character (leading to strong antiferromagnetic coupling) while the pairs Fe6-Fe7 and Fe2-Fe3 could be interpreted as all having an oxidation state of Fe(2.5). Alternatively, localization of the minority spin electron in these delocalized pairs could result in e.g. Fe7 becoming Fe(II) (and Fe6 being Fe(III)) and Fe3 becoming Fe(II) (and Fe2 being Fe(III)). The localized orbital analysis does show that there is sometimes partial localization of these mixed-valence pairs and e.g. the delocalized electron is slightly more on Fe3 than Fe2 while the the electron in the Fe6-Fe7 is more delocalized. The magnitude of localization/delocalization is found to be rather dependent on the QM-region size and is also very dependent on the density functional so we prefer not taking these arguments too far.

C. Protonation State of homocitrate

While all heavy atoms on FeMoco are now confidently characterized, the same is not necessarily the case for the protonation state. Protons are almost always invisible in limited-resolution crystal structures. The homocitrate ligand bound to molybdenum contains 3 carboxylate groups and 1 alkoxide group. Szilagyi et al.[20] were the first to propose that the alkoxide group could be protonated based on a comparison of the molybdenum-oxygen bond lengths in their DFT models to the lower resolution (1.16 Å) crystal structure (PDB code:1M1N)[84]. Due to resolution limitations and the fact that homocitrate was severely truncated and approximated as a glycolate group in their study, the validity of this assignment was unclear. In previous studies[28,29,30,31,32] by one of us (RB) we have assigned the alkoxide group of the homocitrate as protonated but have not until now given a complete justification for this as the question of how well the protein environment is described surrounding the homocitrate remained unclear.

Figure 6a shows a close-up of the molybdenum environment of the cofactor from the 3U7Q crystal structure. Notably the O-O distance between the Mo-bound alkoxide group and one of the oxygens of the shorter carboxylate-arm is rather small (2.48 Å) suggesting a hydrogen bond between the oxygens. Figure 6b shows the result of the QM/MM model with a 154-atom QM region showing how the O-O distance can be well reproduced when a proton is present between the two O atoms. Removing the proton and reoptimizing as shown in Figure 6c, however, results in a dramatically elongated O-O distance that cannot be reconciled with the experimental crystal structure distance. We note that other resting state crystal structures of *A. vinelandii* at close to physiological pH (pH 8.0) with resolutions between 1.0 - 2.3 Å show this short alkoxide-carboxylate O-O distance to always be present (range of 2.4- 2.6 Å). The only exception to this is the 4ND8 crystal structure, recently published by Rees & Howard et al.[85] which is a crystal structure solved under basic (pH 9.5) conditions, where the O-O distance is on average 2.77 Å as shown in Figure 6d. This is in very good agreement with a QM/MM model where the proton is missing (Figure 6d). It should be noted that the resolution of the latter crystal structure is 2.0 Å (compared to 1.0 Å for the resting state 3U7Q structure) but in view of the large change in this O-O-distance we think this must suggest a missing proton at the pH 9.5 conditions (especially since this proposed proton should be rather acidic). Recently a new 2.3 Å crystal structure (PDB code: 5VQ4)[86] of the MoFe protein at pH 5 conditions became available. This crystal structure is in line with previous crystal structures for the homocitrate where an average O-O distance of 2.41 Å is found, indicating that there is a proton present on the alkoxide group. The possibility of protonated carboxylate groups was briefly explored by us but a geometrical analysis was inconclusive and the carboxylate groups have been modelled as unprotonated in all calculations in this study.

This small study nicely demonstrates that the alkoxide proton on homocitrate is likely present during resting state conditions and is possibly a proton that plays a role in other FeMoco redox states (E_1-E_8), possibly becoming a substrate for H_2 formation or possibly is used to protonate N_2 for ammonia formation. Future studies will reveal the importance of this proton in conjunction with other protonation pathways near FeMoco. Finally, we note that Ryde et al.[27] have recently come to the same conclusions regarding the protonation state of

homocitrate using a combination of QM/MM calculations, pK_a calculations via thermochemical cycles and quantum crystal refinement.

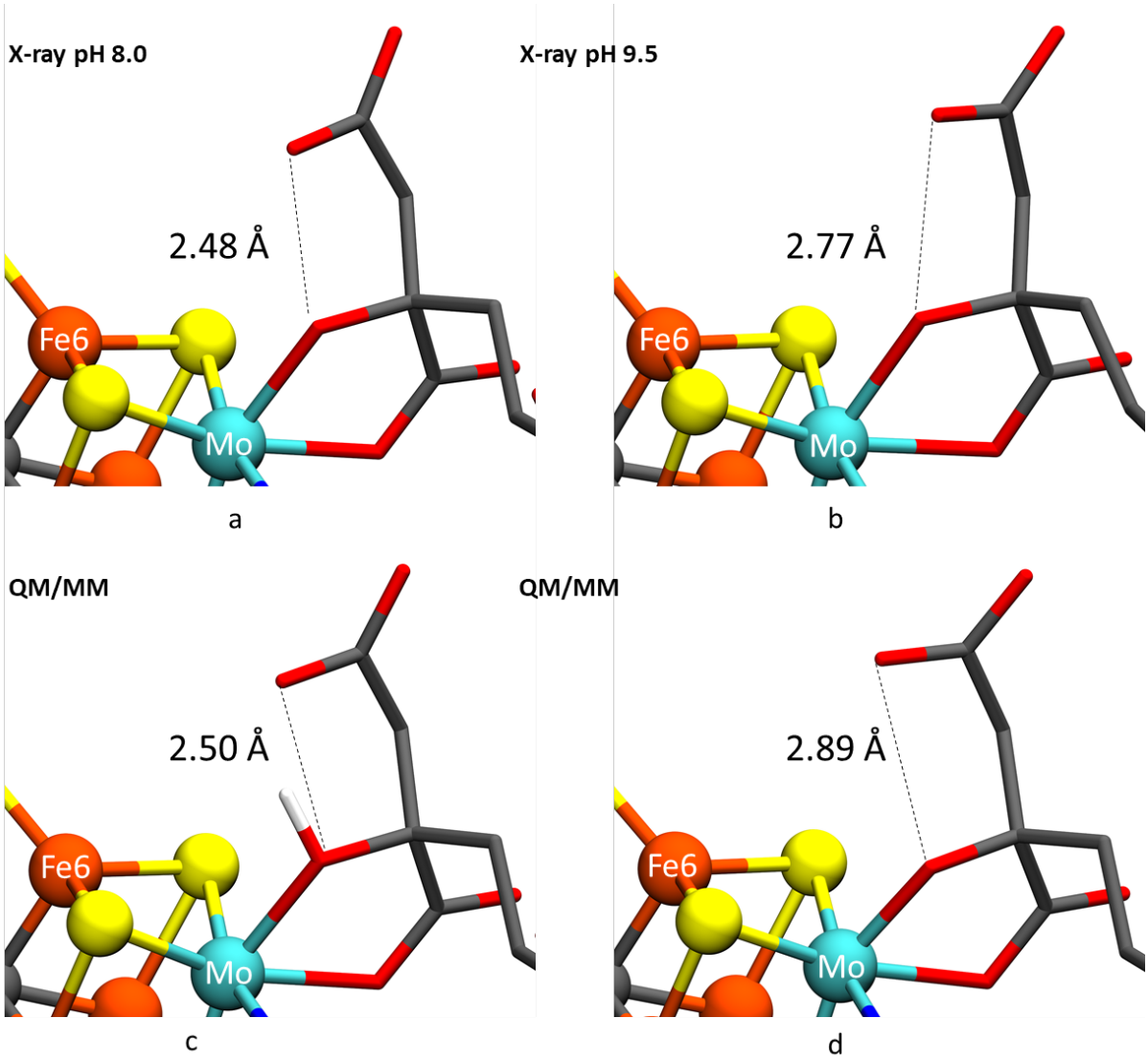


Figure 6. Close-up of the molybdenum coordination environment on FeMoco showing the Mo-coordinated alkoxide group and the O-O distance between the alkoxide group and the carboxylate group of the shorter carboxylate arm. a) the average O-O distance in the 1.0 Å crystal structure (3U7Q)[6] (at pH 8.0). b) the average O-O distance in the 2.0 Å crystal structure (4ND8) at pH 9.5[85]. c) the O-O distance in the protonated QM/MM model (154-atom QM-region). d) the O-O distance in the unprotonated QM/MM model (153-atom QM-region).

D. QM region size dependence: structures, vertical redox reactions and deprotonations

Large QM regions have been used in the geometric analysis of FeMoco in this study. The sometimes slow convergence of QM/MM properties with respect to QM region size is an issue that has been discussed in recent articles[87,88,89,90,91]. It has been suggested that QM/MM properties such as chemical reaction barriers in proteins do not converge until the QM region is several hundred atoms in size. As the knowledge of chemical barriers or the reaction mechanism in nitrogenase is in its infancy, this is a difficult question to answer at present. However, it is possible to systematically explore the sensitivity of the QM/MM calculations with respect to QM region size for properties such as geometries and simple reactions such as redox reactions.

Table 3 shows a comparison of the RMSD from the crystal structure for the FeMoco geometry (both with and without the homocitrate, His and Cys ligands) in QM/MM-optimized geometries with QM region sizes ranging from 54 atoms to 367 atoms. Four of these QM-region sizes were introduced in Figure 2, the other QM regions are introduced in the supporting information. The RMSD values clearly reveal that even a minimal QM region of 54 atoms gives small deviations with respect to the crystal structure and the deviation changes very little with larger QM regions and appears completely converged (within 0.001 Å for the bare cofactor atoms) with a size of 334 or 367 atoms. This suggests that one can get away with small QM regions in QM/MM calculations if one only intends to describe the FeMoco structure well and demonstrates that the QM regions employed in the previous geometric analysis should have been large enough. This may not be the case for redox properties or chemical barriers, however. Also shown in Table 3 is the RMSD when a 54 atom QM cluster in a dielectric continuum described by the COSMO[92] solvation model (and without any explicit protein environment) is calculated at the same level of theory. The structure deviates considerably from the 54 QM-atom QM/MM model. This demonstrates well how incorporation of structural flexibility and natural constraints within the QM/MM model can have a considerable impact on the local geometry of the cofactor. Such large structural deviations as seen in the 54-atom COSMO model (due to model defects caused by the lack of an explicit protein environment) may well adversely affect the energy landscape in calculations of reaction mechanisms.

Table 3: Root-mean-square deviations (RMSD in Å) for the QM/MM-optimized geometries for the increasing QM region from one of the cofactors (connected to A and B chains) of the crystal structure (3U7Q). RMSDs shown for the [$MoFe_7S_9C$] part alone and also including the heavy atoms of homocitrate(HCA)+His442-sidechain+Cys275-SCH_2.

QM region size (atoms)	[$MoFe_7S_9C$]	$MoFe_7S_9C$-HCA-His-Cys
54	0.040	0.105
103	0.039	0.099
127	0.043	0.102
154	0.044	0.104
198	0.050	0.103
247	0.044	0.100

282	0.045	0.101
334	0.046	0.106
367	0.046	0.100
54 (COSMO)	0.081	0.698

Additionally, we explored the redox properties and deprotonation energy of FeMoco by computation of vertical 1-electron oxidation and reduction energies as well as the vertical deprotonation energy of removing the alkoxide proton on homocitrate. The purpose here is only to understand the dependence of these properties with respect to QM region size and not too compare to experimental data which we will explore in future studies. For the oxidized and reduced species we removed/added one electron to the QM/MM optimized resting state geometry, kept the same broken-symmetry solution (BS7-235) and converged to a final M_S= 0 for the oxidized state and M_S=1 for reduced state. We note that little is known about the oxidized and reduced FeMoco states (spin states are controversial and protonation states are unknown) and the current calculations are only performed to study QM region convergence for these redox energies and to test the methodology for future calculations of redox potentials. Point charges from the MM region were included in all of these calculations.

The results are shown in Table 4 and reveals a much larger dependence of these properties on the size of the QM region. While there is only a 1 kcal/mol jump in ΔE_{Ox} and ΔE_{Red} in going from 54 QM atoms to 103 atoms, there is a larger jump when increasing the size to 198 atoms and subsequently there are relatively small changes (1-2 kcal/mol) as we approach the largest QM region of 367 atoms. The deprotonation energies on the other hand decrease rapidly with increasing QM region size and shows little sign of convergence. It should be pointed out, however, that these results are obtained with the TPSSh functional and that self-interaction error in the density functional could affect these results as shown by Ryde for QM/MM calculations on proteins[91]. Future studies will look into these properties with range-separated hybrid functionals but the current results suggest that QM cluster sizes, whether used in cut-out cluster-continuum models of MoFe protein or in QM/MM models, are an important aspect to consider in future mechanistic calculations of redox catalysis on FeMoco.

Table 4: Vertical oxidation, reduction and deprotonation QM/MM energies in kcal/mol for FeMoco for different sized QM regions.

QM region size (atoms)	ΔE_{Ox}	ΔE_{red}	ΔE_{deprot}
54	56.14	52.49	432.71
103	57.06	53.34	426.73
127	57.35	46.15	425.25
154	59.08	48.82	420.11
198	64.16	33.00	416.12
247	66.65	33.29	412.37
282	67.70	31.88	411.15
334	65.77	33.67	412.45
367	66.82	31.62	408.84

Conclusions

We have presented a detailed QM/MM study of the resting state of the MoFe protein of nitrogenase. We have focused exclusively on the active site containing the FeMo cofactor and the S=3/2 resting state and have shown based on a careful analysis of the metal-metal distances of the cofactor that only the $[MoFe_7S_9C]^{1-}$ charge can be the resting state of the cofactor. This fits well with recent spectroscopic and computational work. Furthermore we show that when considering all 3 spin isomers of the lowest-energy class of broken-symmetry solutions (BS7) that only the solution BS7-235 (where Fe atoms Fe2, Fe3 and Fe5 are spin-down) fits well when comparing to the trends in the crystallographic Fe-Fe and Mo-Fe distances. We further show that this can be understood well by an understanding of the electronic structure in terms of specific locations of delocalized mixed-valence ferromagnetic pairs and more localized Fe sites. This electronic structure of the cofactor fits reasonably well with the picture from spatially resolved anomalous dispersion refinement where 3 Fe atoms (Fe1,Fe3, Fe7) are found to be more reduced according to their XAS edges and we propose partial localization (possibly by vibronic coupling) of the minority-spin electrons in these mixed-valence ferromagnetic Fe-Fe pairs to be the reason for why Fe3 and Fe7 appear as more reduced in the SpReAD. We also demonstrate the presence of a proton on the alkoxide group under resting state conditions of the homocitrate group by comparison to multiple crystal structures. Finally we have performed a convergence study of our QM/MM model with respect to QM region size. The geometry of the cofactor converges very quickly with QM region size and we achieve some convergence of the 1-electron vertical redox properties as well. Deprotonation energies converge slowly.

While the electronic structure of the resting state remains exotic and is not completely understood we believe that we are reaching the limits of what can be accomplished within a broken-symmetry DFT picture. It is our hope that affordable multiconfigurational wavefunction theory will soon be able to present a more detailed picture. This study has nonetheless revealed remarkable agreement that can be achieved between broken-symmetry DFT and experiment, when an explicit description of the protein environment is included. Future QM/MM studies will focus on characterizing the redox states of FeMoco and substrate interactions as well as unraveling the specific role that amino acid residues around FeMoco play in the mechanism for substrate catalysis.

Supporting Information Available:

Details about QM regions, RMSD data on the unconstrained model, localized orbitals, more QM and QM/MM technical details and Cartesian coordinates for all QM regions.

Acknowledgements

We thank Albert Þ. Þórhallsson for useful discussions. RB acknowledges support from the Icelandic Research Fund, Grants No. 141218051 and 141218052 and the University of Iceland Research Fund. RB thanks Hannes Jónsson and Egill Skúlason for support.

References

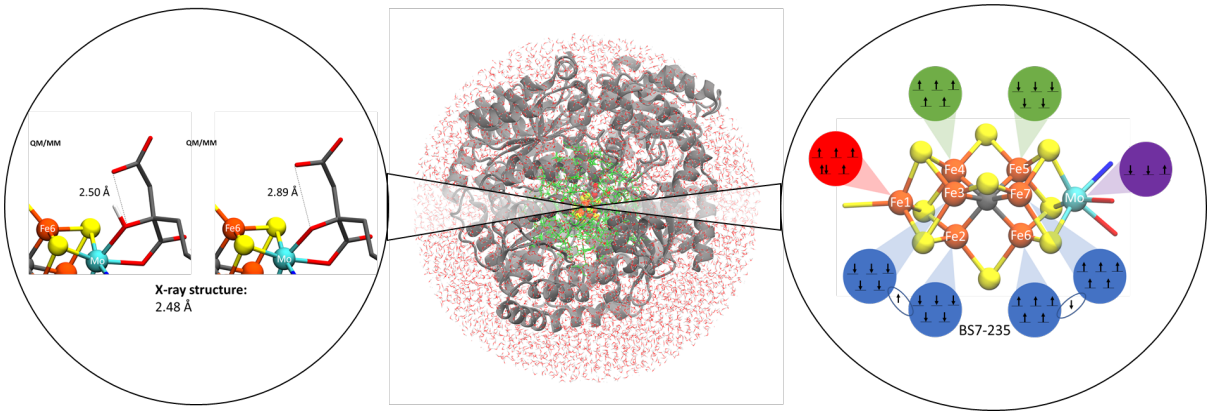
(1) Burgess, B. K. The iron-molybdenum cofactor of nitrogenase. *Chem. Rev.* **1990**, *90,* 1377-1406.
(2) Burgess, B. K.; Lowe, D. J. Mechanism of Molybdenum Nitrogenase. *Chem. Rev.* **1996**, *96,* 2983-3012.
(3) Howard, J. B.; Rees, D. C. Structural Basis of Biological Nitrogen Fixation. *Chem. Rev.* **1996**, *96,* 2965-2982.
(4) 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.
(5) Ribbe, M. W.; Hu, Y.; Hodgson, K. O.; Hedman, B. Biosynthesis of nitrogenase metalloclusters. *Chem. Rev.* **2014**, *114,* 4063-80.
(6) 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.
(7) 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.
(8) Lovell, T.; Li, J.; Liu, T.; Case, D. A.; Noodleman, L. FeMo Cofactor of Nitrogenase: A Density Functional Study of States M N, M OX, M R, and M I. *J. Am. Chem. Soc.* **2001**, *123,* 12392-12410.
(9) Lovell, T.; Torres, R. A.; Han, W.; Liu, T.; Case, D. A.; Noodleman, L. Metal Substitution in the Active Site of Nitrogenase MFe 7S 9(M = Mo 4+, V 3+, Fe 3+). *Inorg. Chem.* **2002**, *41,* 5744-5753.
(10) Lovell, T. T.; Liu, T. T.; Case, D. A.; Noodleman, L. L. Structural, spectroscopic, and redox consequences of a central ligand in the FeMoco of nitrogenase: a density functional theoretical study. *J. Am. Chem. Soc.* **2003**, *125,* 8377-8383.
(11) Lukoyanov, D.; Pelmenschikov, V.; Maeser, N.; Laryukhin, M.; Yang, T. C.; Noodleman, L.; Dean, D. R.; Case, D. A.; Seefeldt, L. C.; Hoffman, B. M. Testing if the interstitial atom, X, of the nitrogenase molybdenum-iron cofactor is N or C: ENDOR, ESEEM, and DFT studies of the S = 3/2 resting state in multiple environments. *Inorg. Chem.* **2007**, *46,* 11437-11449.
(12) Lovell, T.; Li, J.; Case, D. A.; Noodleman, L. FeMo cofactor of nitrogenase: energetics and local interactions in the protein environment. *J. Biol. Inorg. Chem.* **2002**, *7,* 735-49.
(13) Rod, T. H.; Nørskov, J. K. Modeling the nitrogenase FeMo cofactor. *J. Am. Chem. Soc.* **2000**, *122,* 12751-12763.
(14) Hinnemann, B.; Nørskov, J. K. Modeling a central ligand in the nitrogenase FeMo cofactor. *J. Am. Chem. Soc.* **2003**, *125,* 1466-7.
(15) Hinnemann, B.; Nørskov, J. K. Chemical Activity of the Nitrogenase FeMo Cofactor with a Central Nitrogen Ligand: Density Functional Study. *J. Am. Chem. Soc.* **2004**, *126,* 3920-3927.
(16) Varley, J. B.; Wang, Y.; Chan, K.; Studt, F.; Nørskov, J. K. Mechanistic insights into nitrogen fixation by nitrogenase enzymes. *Phys. Chem. Chem. Phys.* **2015**, *17,* 29541-7.
(17) Kästner, J.; Blöchl, P. E. Towards an understanding of the workings of nitrogenase from DFT calculations. *ChemPhysChem* **2005**, *6,* 1724-6.
(18) Kästner, J.; Blöchl, P. E. Ammonia production at the FeMo cofactor of nitrogenase: results from density functional theory. *J. Am. Chem. Soc.* **2007**, *129,* 2998-3006.
(19) Hallmen, P. P.; Kästner, J. N2 Binding to the FeMo-Cofactor of Nitrogenase. *Z. Anorg. Allg. Chem.* **2015**, *641,* 118-122.

(20) Harris, T. V.; Szilagyi, R. K. Comparative assessment of the composition and charge state of nitrogenase FeMo-cofactor. *Inorg. Chem.* **2011**, *50,* 4811-4824.
(21) Dance, I. Electronic dimensions of FeMo-co, the active site of nitrogenase, and its catalytic intermediates. *Inorg. Chem.* **2011**, *50,* 178-192.
(22) Dance, I. Nitrogenase: a general hydrogenator of small molecules. *Chem. Comm.* **2013**, *49,* 10893-907.
(23) McKee, M. L. A New Nitrogenase Mechanism Using a CFe8S9 Model: Does H2 Elimination Activate the Complex to N2 Addition to the Central Carbon Atom?. *J. Phys. Chem. A* **2016**, *120,* 754-64.
(24) Siegbahn, P. E. Model calculations suggest that the central carbon in the FeMo-cofactor of nitrogenase becomes protonated in the process of nitrogen fixation. *J. Am. Chem. Soc.* **2016**, *138,* 10485-10495.
(25) Xie, H.; Wu, R.; Zhou, Z.; Cao, Z. Exploring the interstitial atom in the FeMo cofactor of nitrogenase: insights from QM and QM/MM calculations. *J. Phys. Chem. B* **2008**, *112,* 11435-9.
(26) Rao, L.; Xu, X.; Adamo, C. Theoretical investigation on the role of the central carbon atom and close protein environment on the nitrogen reduction in Mo nitrogenase. *ACS. Catalysis* **2016**, *6,* 1567-1577.
(27) Cao, L.; Caldararu, O.; Ryde, U. Protonation States of Homocitrate and Nearby Residues in Nitrogenase Studied by Computational Methods and Quantum Refinement. *J. Phys. Chem. B.* **2017** DOI: 10.1021/acs.jpcb.7b0271
(28) Bjornsson, R.; Lima, F. A.; Spatzal, T.; Weyhermueller, T.; Glatzel, P.; 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.
(29) Bjornsson, R.; Delgado-Jaime, M. U.; Lima, F. A.; Sippel, D.; Schlesier, J.; Weyhermüller, T.; Einsle, O.; Neese, F.; DeBeer, S. Molybdenum L-Edge XAS Spectra of MoFe Nitrogenase. *Z. Anorg. Allg. Chem.* **2015**, *641,* 65-71.
(30) Rees, J. A.; Bjornsson, R.; Schlesier, J.; Sippel, D.; Einsle, O.; DeBeer, S. The Fe-V Cofactor of Vanadium Nitrogenase Contains an Interstitial Carbon Atom. *Angew. Chem. Int. Ed. Engl.* **2015**, *54,* 13249-52.
(31) Rees, J. A.; Bjornsson, R.; Kowalska, J. K.; Lima, F. A.; Schlesier, J.; Sippel, D.; Weyhermüller, T.; Einsle, O.; Kovacs, J. A.; DeBeer, S. Comparative electronic structures of nitrogenase FeMoco and FeVco. *Dalton Trans.* **2017**, *46,* 2445-2455.
(32) 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.
(33) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E.; Mittal, J.; Feig, M.; Mackerell, A. D. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain $\chi(1)$ and $\chi(2)$ dihedral angles. *J. Chem. Theory Comput.* **2012**, *8,* 3257-3273.
(34) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38,* 3098-3100.
(35) Perdew, J. P. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B* **1986**, *33,* 8822-8824.
(36) 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.
(37) Reed, A. E.; Weinstock, R. B.; Weinhold, F. Natural population analysis. *J. Chem. Phys.* **1985**, *83,* 735-746.
(38) Chang, C. H.; Kim, K. Density Functional Theory Calculation of Bonding and Charge Parameters for Molecular Dynamics Studies on [FeFe] Hydrogenases. *J. Chem. Theory*

Comput. **2009**, *5*, 1137-45.
(39) Wright, L. B.; Rodger, P. M.; Walsh, T. R. Aqueous citrate: a first-principles and force-field molecular dynamics study. *RSC*. *Adv*. **2013**, *3*, 16399-16409.
(40) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J*. *Comput*. *Chem*. **2010**, *31*, 671-90.
(41) Zhang, L.; Kaiser, J. T.; Meloni, G.; Yang, K. Y.; Spatzal, T.; Andrade, S. L.; Einsle, O.; Howard, J. B.; Rees, D. C. The sixteenth iron in the nitrogenase MoFe protein. *Angew*. *Chem*. *Int*. *Ed*. *Engl*. **2013**, *52*, 10529-32.
(42) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J*. *Chem*. *Theory Comput*. **2008**, *4*, 435-47.
(43) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013**, *29*, 845-54.
(44) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*. **2015**, *1*, 19-25.
(45) Li, H.; Robertson, A. D.; Jensen, J. H. Very fast empirical prediction and rationalization of protein pKa values. *Proteins Struct*. *Funct*. *Bioinf*. **2005**, *61*, 704-721.
(46) Olsson, M. H.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *J*. *Chem*. *Theory Comput*. **2011**, *7*, 525-37.
(47) Hess, B.; Bekker, H.; Berendsen, H. J.; Fraaije, J. G. LINCS: a linear constraint solver for molecular simulations. *J*. *Comput*. *Chem*. **1997**, *18*, 1463-1472.
(48) Hess, B. P-LINCS: A parallel linear constraint solver for molecular simulation. *J*. *Chem*. *Theory Comput*. **2008**, *4*, 116-122.
(49) Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. *J*. *Chem*. *Phys*. **1984**, *81*, 511-519.
(50) Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol*. *Phys*. **1984**, *52*, 255-268.
(51) Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys*. *Rev*. *A* **1985**, *31*, 1695-1697.
(52) Martyna, G. J. ; Klein, M. L.; Tuckerman, M. Nosé-Hoover chains: The canonical ensemble via continuous dynamics. *J*. *Chem*. *Phys*. **1992**, *97*, 2635-2643.
(53) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew*. *Chem*. *Int*. *Ed*. **2009**, *48*, 1198-1229.
(54) Thiel, W. In *Multiscale Simulation Methods in Molecular Sciences*; Grotendorst, J.; Attig, N.; Blügel, S.; and Marx, D., Editor.; John von Neuman Institute for Computing: Jülich, 2009; p. . 203-214.
(55) Senn, H. M., Thiel, W. In *Topics in Current Chemistry*, Reiher, M., Ed.; Springer Berlin Heidelberg, **2007**; p. 173-290.
(56) Vasilevskaya, T.; Thiel, W. Periodic Boundary Conditions in QM/MM Calculations: Implementation and Tests. *J*. *Chem*. *Theory Comput*. **2016**, *12*, 3561-70.
(57) Sherwood, P.; de Vries, A. H.; Guest, M. F.; Schreckenbach, G.; Catlow, C. R. A.; French, S. A.; Sokol, A. A.; Bromley, S. T.; Thiel, W.; Turner, A. J.; Billeter, S.; Terstegen, F.; Thiel, S.; Kendrick, J.; Rogers, S. C.; Casci, J.; Watson, M.; King, F.; Karlsen, E.; Sjøvoll, M.; Fahmi, A.; Schäfer, A.; Lennartz, C. QUASI: A general purpose implementation of the

QM/MM approach and its application to problems in catalysis. *J. Mol. Struct.: THEOCHEM* **2003**, *632*, 1-28.
(58) Metz, S.; Kästner, J.; Sokol, A. A.; Keal, T. W.; Sherwood, P. ChemShell modular software package for QM/MM simulations. *WIREs. Comput. Mol. Sci.* **2014**, *4*, 101-110.
(59) Neese, F. The ORCA program system. *WIREs. Comput. Mol. Sci.* **2011**, *2*, 73-78.
(60) Sherwood, P. In *Modern Methods and Algorithms of Quantum Chemistry Proceedings*; Grotendorst, J., Ed.; John von Neumann Institute for Computing : Jülich, Germany; p. 285-305.
(61) Tao, J.; Perdew, J.; Staroverov, V.; Scuseria, G. Climbing the density functional ladder: Nonempirical meta--generalized gradient approximation designed for molecules and solids. *Phys. Rev. Lett.* **2003**, *91*, 146401.
(62) 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.
(63) 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.
(64) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456-1465.
(65) van Lenthe, E.; Baerends, E. J.; Snijders, J. G. Relativistic regular two-component Hamiltonians. *J. Chem. Phys.* **1993**, *99*, 4597-4610.
(66) 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.
(67) Pantazis, D. A.; Chen, X.; Landis, C. R.; Neese, F. All-Electron Scalar Relativistic Basis Sets for Third-Row Transition Metal Atoms. *J. Chem. Theory Comput.* **2008**, *4*, 908-919.
(68) 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. *Chem. Phys.* **2009**, *356*, 98-109.
(69) Izsák, R.; Neese, F. An overlap fitted chain of spheres exchange method. *J. Chem. Phys.* **2011**, *135*, 144105.
(70) Smith, W.; Forester, T. DL_POLY_2.0: A general-purpose parallel molecular dynamics simulation package. *J. Mol. Graphics.* **1996**, *14*, 136-141.
(71) 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.
(72) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **1996**, *14*, 33-8, 27-8.
(73) Hoffman, B. M.; Roberts, J. E.; Orme-Johnson, W. H. Molybdenum-95 and proton ENDOR spectroscopy of the nitrogenase molybdenum-iron protein. *J. Am. Chem. Soc.* **1982**, *104*, 860-862.
(74) Venters, R. A.; Nelson, M. J.; McLean, P. A.; True, A. E.; Levy, M. A.; Hoffman, B. M.; Orme-Johnson, W. H. ENDOR of the resting state of nitrogenase molybdenum-iron proteins from Azotobacter vinelandii, Klebsiella pneumoniae, and Clostridium pasteurianum. Proton, iron-57, molybdenum-95, and sulfur-33 studies. *J. Am. Chem. Soc.* **1986**, *108*, 3487-3498.
(75) Lee, H.; Hales, B. J.; Hoffman, B. M. Metal-Ion Valencies of the FeMo Cofactor in CO-Inhibited and Resting State Nitrogenase by 57Fe Q-Band ENDOR. *J. Am. Chem. Soc.* **1997**, *119*, 11395-11400.
(76) True, A. E.; McLean, P.; Nelson, M. J.; Orme-Johnson, W. H.; Hoffman, B. M.

Comparison of wild-type and nifV mutant molybdenum-iron proteins of nitrogenase from Klebsiella pneumoniae by ENDOR spectroscopy. *J. Am. Chem. Soc.* **1990**, *112,* 651-657.
(77) Yoo, S. J.; Angove, H. C.; Papaefthymiou, V.; Burgess, B. K.; Münck, E. Mössbauer Study of the MoFe Protein of Nitrogenase from Azotobacter vinelandiiUsing Selective 57Fe Enrichment of the M-Centers. *J. Am. Chem. Soc.* **2000**, *122,* 4926-4936.
(78) Spatzal, T.; Schlesier, J.; Burger, E. M.; Sippel, D.; Zhang, L.; Andrade, S. L.; Rees, D. C.; Einsle, O. Nitrogenase FeMoco investigated by spatially resolved anomalous dispersion refinement. *Nat. Commun.* **2016**, *7,* 10902.
(79) Bjornsson, R.; Neese, F.; Schrock, R. R.; Einsle, O.; DeBeer, S. The discovery of Mo(III) in FeMoco: reuniting enzyme and model chemistry. *J. Biol. Inorg. Chem.* **2015**, *20,* 447-60.
(80) L. Noodleman, C. Y. Peng, D. A. Case, and J. M. Mouesca, *Coord. Chem. Rev.*, 1995, **144**, 199–244.
(81) H. Beinert, *Science*, 1997, **277**, 653–659.
(82) Sharma, S.; Sivalingam, K.; Neese, F.; Chan, G. K. Low-energy spectrum of iron-sulfur clusters directly from many-particle quantum mechanics. *Nat. Chem.* **2014**, *6,* 927-33.
(83) 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,* 4485-97.
(84) Einsle, O.; Tezcan, F. A.; Andrade, S. L. A.; Schmid, B.; Yoshida, M.; Howard, J. B.; Rees, D. C. Nitrogenase MoFe-Protein at 1.16 Å Resolution: A Central Ligand in the FeMo-Cofactor. *Science* **2002**, *297,* 1696-1700.
(85) Yang, K. Y.; Haynes, C. A.; Spatzal, T.; Rees, D. C.; Howard, J. B. Turnover-dependent inactivation of the nitrogenase MoFe-protein at high pH. *Biochemistry.* **2014**, *53,* 333-43.
(86) Morrison, C. N.; Spatzal, T.; Rees, D. C. Reversible Protonated Resting State of the Nitrogenase Active Site. *J. Am. Chem. Soc.* **2017**, *139,* 10856-10862.
(87) Kulik, H. J.; Zhang, J.; Klinman, J. P.; Martínez, T. J. How Large Should the QM Region Be in QM/MM Calculations? The Case of Catechol O-Methyltransferase. *J. Phys. Chem. B* **2016**, *120,* 11381-11394.
(88) Karelina, M.; Kulik, H. J. Systematic Quantum Mechanical Region Determination in QM/MM Simulation. *J. Chem. Theory Comput.* **2017**, *13,* 563-576.
(89) Liao, R. Z.; Thiel, W. Convergence in the QM-only and QM/MM modeling of enzymatic reactions: A case study for acetylene hydratase. *J. Comput. Chem.* **2013**, *34,* 2389-97.
(90) Roßbach, S.; Ochsenfeld, C. Influence of Coupling and Embedding Schemes on QM Size Convergence in QM/MM Approaches for the Example of a Proton Transfer in DNA. *J. Chem. Theory. Comput.* **2017**, *13,* 1102-1107.
(91) Fouda, A.; Ryde, U. Does the DFT Self-Interaction Error Affect Energies Calculated in Proteins with Large QM Systems?. *J. Chem. Theory. Comput.* **2016**, *12,* 5667-5679.
(92) Klamt, A.; Schüürmann, G. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc. Perkin Trans. 2* **1993**, 0, 799-805.



For Table of Contents Only

A QM/MM model of the MoFe protein of nitrogenase, accounting explicitly for the protein environment, has been used to study the geometry and properties of the FeMo cofactor. Analysis of the metal-metal distances reveals that the cofactor resting state charge is $[MoFe_7S_9C]^{1-}$. Furthermore we demonstrate that a specific spin isomer of FeMoco, BS7-235, is favored over others. Mo-bound homocitrate is furthermore found to be protonated under resting state conditions.