Protonation and Reduction of the FeMo Cluster in Nitrogenase Studied by Quantum Mechanics/Molecular Mechanics (QM/MM) Calculations

Cao, Lili; Caldararu, Octav; Ryde, Ulf

Published in:
Journal of Chemical Theory and Computation

DOI:
10.1021/acs.jctc.8b00778

2018

Document Version:
Peer reviewed version (aka post-print)

Protonation and reduction of the FeMo cluster in nitrogenase studied by QM/MM calculations

Lili Cao, Octav Caldararu and Ulf Ryde *

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

Correspondence to Ulf Ryde, E-mail: Ulf.Ryde@teokem.lu.se, Tel: +46 – 46 2224502, Fax: +46 – 46 2228648

2018-10-15

Abstract

We have performed a systematic computational study of the relative energies of possible protonation states of the FeMo cluster in nitrogenase in the E_0–E_4 states, i.e. the resting state and states with 1–4 electrons and protons added but before N_2 binds. We use the combined quantum mechanics and molecular mechanics (QM/MM) approach, including the complete solvated heterotetrameric enzyme in the calculations. The QM system consisted of 112 atoms, i.e. the full FeMo cluster, as well all groups forming hydrogen bonds to it within 3.5 Å. It was treated with either the TPSS-D3 or B3LYP-D3 methods with the def2-SV(P) or def2-TZVPD basis sets. For each redox state, we calculated relative energies of at least 50 different possible positions for the proton, added to the most stable protonation state of the level with one electron less. We show quite conclusively that the resting E_0 state is not protonated using quantum refinement and by comparing geometries to the crystal structure. The E_1 state is protonated on S2B, in agreement with most previous computational studies. However, for the E_2–E_4 states, the two QM methods give diverging results, with relative energies that differ by over 300 kJ/mol for the most stable E_4 states. TPSS favours hydride ions binding to the Fe ions. The first bridges Fe2 and Fe6, whereas the next two bind terminally to either Fe4, Fe5 or Fe6 with nearly equal energies. On the other hand, B3LYP disfavours hydride ions and instead suggests that 1–3 protons bind to the central carbide ion.

Key Words: Nitrogenase, QM/MM, protonation, E_4 state, broken-symmetry DFT, quantum refinement.

Introduction

The N–N triple bond in N_2 is one of the strongest bonds in nature. For this reason, nitrogen is often a limiting element in plants, although the atmosphere is dominated by N_2.[1] Nitrogenase (EC 1.18/19.6.1) is the only enzyme that can cleave this bond, by performing the reaction

$$N_2 + 8\ e^- + 8\ H^+ + 16\ ATP \rightarrow 2\ NH_3 + H_2 + 16\ ADP + 16\ P_i \quad (1)$$

It is notable that H_2 seems to be a compulsory by-product of the reaction, although it is not stoichiometrically needed and it adds an extra cost of four ATP and two electrons. Owing to the importance of nitrogenase for biological systems, the enzyme has been extensively studied.[1–3]

Crystal structures show that nitrogenase consists of two proteins, the Fe protein, which delivers electrons, and the MoFe protein, which binds the substrate and performs the chemical reactions.[4–8] The latter protein contains two unusual and complicated iron–sulfur cofactors. The P cluster is a Fe_8S_7 complex with an electron-transfer function. The active site is the FeMo cluster, which is a $MoFe_7S_9C$(homocitrate) complex that is connected to the protein by a histidine and a cysteine residue.[6,9] There also exist alternative nitrogenases in which the Mo ion is replaced by V or Fe.[10] It is currently believed that the binding of the ATP-loaded Fe protein to the MoFe protein triggers a conformational change that leads to a slow transfer of an electron from the P cluster to the FeMo cluster. The P cluster is then rapidly re-reduced by the Fe protein. Finally, ATP is hydrolysed and the Fe protein dissociates again in a rate-limiting step.[1–3,11]

The reaction mechanism of nitrogenase is traditionally discussed in terms of the Lowe–Thorneley scheme, in which nine enzymic states are recognized, differing in the number of added electrons and protons, called E_0 to E_8.[12] Recent EPR and freeze-trapping experiments have significantly advanced our knowledge about the reaction mechanism.[1–3,9,13] According to these, four electrons and protons have to be added to FeMo cluster before N_2 can bind. In particular, it has been suggested that the binding of N_2 and probably the first two protonations of N_2 are facilitated by the formation and dissociation of H_2 from the FeMo cluster.[14,15] This provides an attractive explanation why H_2 is a obligate by-product in the reaction mechanism of the nitrogenases. However, the atomistic details of the mechanism are still not clear.

Nitrogenase has also been extensively studied by computational methods.[1,13,24–28,16–23] Earlier studies were hampered by the fact that the central carbide ion in the FeMo cluster was not identified until 2014[5–9] and that the redox state of the cluster was not settled until 2017.[29–31] Moreover, earlier studies did not involve the loading of the active site by four electrons and protons before the actual reaction could take place. However, even among the latest quantum mechanical (QM) studies, there are no agreement among the reaction mechanism. For example, Dance has suggested that N_2 binds side-on to one Fe ion and is protonated alternatively on the two N atoms,[32] whereas Nørskov and coworkers have proposed a mechanism in which N_2 binds with one N atom bridging two Fe ions, after the dissociation of one of the sulfide ions, and the non-coordinating N atom is fully protonated and dissociates before the other N atom is protonated.[25] On the other hand, Siegbahn has suggested that N_2 binds side-on, bridging between two Fe ions, after a triple protonation of the central carbide ion, which thereby moves to the periphery of the cluster and after the reduction of the cluster by six[26] or eight electrons.[33] Moreover, McKee has suggested that N_2 instead binds end-on to the central C atom and is protonated first on the distal N atom,[27] whereas Rao et al. has proposed that the central C is singly protonated and covalently interacts with N_2, which is protonated in a sequential manner.[28]

In fact, the various studies do not even agree on the most stable protonation states of the FeMo cluster in the various E_n states. On the contrary, there has recently been a dispute concerning the structure of the E_4 state: The Hoffman group suggested that it involves the

protonation of two sulfide ions and two hydride ions bridging two Fe ions each.[34] On the other hand, Siegbahn, has argued that it is energetically much more favourable (by 390 kJ/mol) to protonate the central carbide ion.[35]

Computational studies of nitrogenase are a formidable task. In the FeMo cluster, there are at least 21 sites that can be protonated (Mo, seven Fe, nine S, the carbide ion, as well as the cysteine, histidine and homocitrate ligands) and each of them can typically be protonated at two or three distinct positions. Moreover, for the metals, the hydride ion may either bind to a single metal ion or it can bridge two ions. As will be discussed in more detail below, this gives over 50 possible positions for each added proton. Thus, for the E_4 state, with four added protons, there are in principle $50^4 = 6.25 \cdot 10^6$ distinct structures and for each, there are 35 possible broken-symmetry states and 2–4 possible spin states, giving of the order of 10^9 possible states (and Siegbahn recently suggested that as much as eight protons and electrons need to be provided before N_2 binds,[33] giving ~10^{15} possible structures). Of course, it is currently impossible to systematically study all these states and it is therefore not unexpected that different scientists come to diverging conclusions regarding the most favourable protonation state, in particular as the energies depend on the QM approach employed, the density functional theory (DFT) method and the size of the QM system. Therefore, any computational scientist studying this system needs to apply some sort of heuristic approach to find the best protonation states.

The FeMo cofactor has a complicated electronic structure. All Fe ions are in the high-spin Fe(II) or Fe(III) state, whereas the Mo ion is in the +III state.[13] In DFT, this is normally treated by the broken-symmetry (BS) approach[20] and Noodleman and coworkers have shown that there are 35 possible BS states for the FeMo cluster in the asymmetric protein.[20,36] In a recent study, we suggested an approach to deal with these BS and spin states in studies of the nitrogenase mechanism.[37] We showed that the great majority of the states can be discarded by single-point energy calculations with medium-sized basis sets. Only states within 20 kJ/mol of the best one need to be studied by geometry optimisations and single-point energy calculations with big basis sets. However, both a pure and a hybrid DFT method should be employed, as they can give widely differing results.

In this study, we employ this approach to study the reduction and protonation of the FeMo cluster in nitrogenase from the resting E_0 state to the E_4 state. We employ the QM/MM approach,[38,39] which is less dependent on the size of the QM system than QM-cluster calculations.[40,41] Moreover, all atoms of the full (tetrameric) protein, as well a large amount of explicit water molecules, are included in the calculations, reducing the risk of making a biased choice of the QM system and avoiding the use of an ill-defined dielectric constant for the heterogeneous surroundings. Only a few QM/MM studies have previously been published for nitrogenase[21,31,37] and none of them has dealt with the protonation and reduction of the active site. We first address the question of the protonation state of the resting E_0 state by comparing to the crystal structure. Then, we study the E_1–E_4 states by systematically adding a proton to ~50 potential protonation sites, starting from the most stable structure of the state with one proton and electron less. All states are studied with both a pure and a hybrid DFT functional.

Methods

The protein

All calculations were based on the 1.0-Å crystal structure of nitrogenase from *Azotobacter vinelandii* (PDB code 3U7Q).[6] The setup of the protein is identical to that of our previous studies of the protein.[37,42] The entire heterotetramer was included in the calculations, because the various

subunits are entangled without any natural way to separate them. The QM calculations were concentrated on the FeMo clusters in the C subunit because there is a buried imidazole molecule from the solvent rather close to the active site (~11 Å) in the A subunit. The P cluster and the FeMo cluster in subunit A were modelled by MM in the fully reduced and resting states, respectively.[42]

The protonation states of all residues were the same as before:[42] All Arg, Lys, Asp, and Glu residues were assumed to be charged, except Glu-153, 440, and 231D (a letter "D" after the residue number indicates that it belongs to that subunit; if no letter is given, it belongs to subunit C; subunits A and B are identical to the C and D residues). Cys residues coordinating to Fe ions were assumed to be deprotonated. His-274, 451, 297D, 359D and 519D were assumed to be protonated on the ND1 atom, His-31, 196, 285, 383, 90D, 185D, 363D and 457D were presumed to be protonated on both the ND1 and NE2 atoms (and therefore positively charged), whereas the remaining 14 His residues were modelled with a proton on the NE2 atom. The homocitrate was modelled in the singly protonated state with a proton shared between the hydroxyl group (which coordinates to Mo) and the O1 carboxylate atom. This protonation state was found to be the most stable one in a recent extensive QM/MM, molecular dynamics and quantum-refinement study[42] and this protonation state is also supported by another recent study.[30]

The protein was solvated in a sphere with a radius of 70 Å around the geometrical centre of the protein. 160 Cl^- and 182 Na^+ ions were added at random positions (but not inside the protein[42]) to neutralise the protein and give an ionic strength of 0.2 M.[43] The final system contained 133 915 atoms. The added protons, counter ions and water molecules were optimised by a simulated annealing calculation (up to 370 K), followed by a minimisation, keeping the other atoms fixed at the crystal-structure positions.[42]

All MM calculations were performed with the Amber software.[44] For the protein, we used the Amber ff14SB force field[45] and water molecules were described by the TIP3P model.[46] For the metal sites, the MM parameters were the same as in our previous investigation.[42] The metal sites were treated by a non-bonded model[47] and charges were obtained with the restrained electrostatic potential method, obtained at the TPSS/def2-SV(P) level of theory[48,49] and sampled with the Merz–Kollman scheme.[50]

QM calculations

All QM calculations were performed with the Turbomole software (versions 7.1 and 7.2).[51] We employed two DFT methods, TPSS[48] and B3LYP,[52–54] and two different basis sets of increasing size, def2-SV(P)[49] and def2-TZVPD.[55] The calculations were sped up by expanding the Coulomb interactions in an auxiliary basis set, the resolution-of-identity (RI) approximation.[56,57] Empirical dispersion corrections were included with the DFT-D3 approach[58] and Becke–Johnson damping,[59] as implemented in Turbomole.

The FeMo cluster was modelled by $MoFe_7S_9C$(homocitrate)(CH_3S)(imidazole), where the two last groups are models of Cys-275 and His-442. In addition, all groups that form hydrogen bonds to the FeMo cluster within 3.5 Å in the crystal structure[6] were also included, viz. Arg-96, His-195 and Arg-359 (sidechains), Gly-356, Gly-357 (backbone), as well as two water molecules, in total 114 atoms (shown in Figure 1a). Following recent Mössbauer, anomalous dispersion and QM investigations,[13,26,29,30] we used the oxidation state-assignment $Mo^{III}Fe_3^{II}Fe_4^{III}$ of the metal ions in the resting state, giving a net charge of –5 for the QM system. In the present work, we studied five oxidation states of the FeMo cluster, obtained by adding 0–4 electrons to the resting states, denoted E_0, E_1, E_2, E_3 and E_4. One proton was added together with the electron.

However, for E_0, we also tried to add one proton without adding any electron.

Experiments have shown that the ground spin state of E_0 and E_2 are quartets with a surplus of three α electrons, whereas E_4 is a doublet.[1,13] Consequently, we used these spin states for these three states. However, before starting with a new oxidation state (using the most likely protonation state, based on our previous experience), we checked which the three lowest spin states had the lowest energy at the TPSS and B3LYP/def2-SV(P) levels of theory. This was also repeated after the completed study of the protonation states for the best structures of each oxidation state. For the E_1 and E_3 states (for which no experimental information is available), our calculations indicated that the ground spin states are the quintet and triplet, respectively.

The electronic structure of all QM calculations was obtained with the BS approach:[20] Each of the seven Fe ions were modelled in the high-spin state, with either a surplus of α (four Fe ions) or β (three Fe ions) spin. Such a state can be selected in 35 different ways ($\frac{7!}{3!\,4!}$).[37] A starting wavefunction was obtained by first optimising the all-high-spin state with 35 unpaired electrons and then changing the total α and β occupation numbers to the desired net spin. This gave one of the BS states. The other BS states were obtained by simply swapping the coordinates of the Fe ions.[60] In some cases, we instead used the fragment approach by Szilagyi and Winslow to obtain a proper BS state.[61]

We have thoroughly studied the 35 BS states for the resting state, the protonated resting state and the reduced state, and how their energies vary with the QM method, the size of the basis set, the geometry and the influence of the surroundings.[37] The conclusion was that the effect of the basis set and the surroundings were restricted (up to 7–11 kJ/mol), the effect of geometry intermediate (up to 37 kJ/mol, but the correlation, R^2, was 0.92–0.98) and that the effect of the DFT functional was large (up to 58 kJ/mol). Therefore, for each new oxidation (E_n) state, we here studied all BS states for the most likely structure. This study was then repeated again for the best protonation states at the end of the study and if the preference had changed, the protonation study was redone. In both cases, all 35 BS states were studied using single-point energy calculations. The energies were calculated with both the TPSS-D3 and B3LYP-D3 methods with the def2-SV(P) basis set. For all BS states within 20 kJ/mol of the lowest one at either level of theory, the geometry was optimised with that BS state and the TPSS/def2-SV(P) method. Finally, more accurate energies were calculated with TPSS-D3 and B3LYP-D3 and the larger def2-TZVPD basis set. The various spin states are defined in in Table S1 in the Supporting information (they are the same as in our previous study[37]). For the best structures of each E_n state, we also tested to optimise structures with the def2-SVP basis set (i.e. with polarising functions also on the added protons or hydride ions. However, as can be seen from Table S8, this did not change the bond length of these added H atoms by more than 0.02 Å and relative energies changed by less than 7 kJ/mol.

As have been discussed before,[37] TPSS-D3/def2-SV(P) calculations give geometries that reproduce the crystal structure of the resting state of nitrogenase excellently with average and maximum deviations of 0.05 and 0.09 Å for the metal–metal and 0.02 and 0.06 Å metal–ligand distances, respectively and a root-mean-squared-deviation (RMSD) of 0.06 Å for the metals and the first-sphere ligands. This is similar to the results obtained with the TPSSh approach[30] and appreciably better than with the B3LYP-D3/def2-SV(P) method, which gives average and maximum deviations of 0.08 and 0.12 Å for the metal–metal and 0.04 and 0.11 Å metal–ligand distances, respectively and a RMSD of 0.08 Å. Therefore, we used TPSS-D3/def2-SV(P) for all geometries, except in cases where it could be expected that the geometries obtained with B3LYP-D3/def2-SV(P) are significantly different (in particular when the central carbide is protonated).

QM/MM calculations

The QM/MM calculations were performed with the ComQum software.[62,63] In this approach, the protein and solvent are split into three subsystems: System 1 (the QM region) was relaxed by QM methods. System 2 contained all residues and water molecules with at least one atom within 6 Å of any atom in system 1 and it was optionally relaxed by MM. Finally, system 3 contained the remaining part of the protein and the solvent and it was kept fixed at the original coordinates (equilibrated crystal structure). The total system was spherical and non-periodic with 133915 atoms. An example file is given in the Supporting Information. In the great majority of the calculations, system 2 was kept fixed at the crystal geometry (to avoid the risk that different calculations end up in different local minima for system 2), but some test calculations with a relaxed system 2 are presented in Table S9 for the best E_n states. They show that the bond distances involving the added H atoms change by less than 0.02 Å (occasionally the longer distance of a bridging hydride change by up to 0.07 Å) and relative energies change by less than 13 kJ/mol.

In the QM calculations, system 1 was represented by a wavefunction, whereas all the other atoms were represented by an array of partial point charges, one for each atom, taken from the MM setup. Thereby, the polarisation of the QM system by the surroundings is included in a self-consistent manner (electrostatic embedding). When there is a bond between systems 1 and 2 (a junction), the hydrogen link-atom approach was employed: The QM system was capped with hydrogen atoms (hydrogen link atoms, HL), the positions of which are linearly related to the corresponding carbon atoms (carbon link atoms, CL) in the full system.[62,64] All atoms were included in the point-charge model, except the CL atoms.[41]

The total QM/MM energy in ComQum was calculated as[62,63]

$$E_{\mathrm{QM/MM}} = E^{\mathrm{HL}}_{\mathrm{QM1+ptch23}} + E^{\mathrm{CL}}_{\mathrm{MM123},q_1=0} - E^{\mathrm{HL}}_{\mathrm{MM1},q_1=0} \qquad (3)$$

where $E^{\mathrm{HL}}_{\mathrm{QM1+ptch23}}$ is the QM energy of the QM system truncated by HL atoms and embedded in the set of point charges modelling system 2 (but excluding the self-energy of the point charges). $E^{\mathrm{HL}}_{\mathrm{MM1},q_1=0}$ is the MM energy of the QM system, still truncated by HL atoms, but without any electrostatic interactions. Finally, $E^{\mathrm{CL}}_{\mathrm{MM123},q1=0}$ is the classical energy of all atoms in the system with CL atoms and with the charges of the QM region set to zero (to avoid double-counting of the electrostatic interactions). Thus, ComQum employs a subtractive scheme with electrostatic embedding and van der Waals link-atom corrections.[65] No cutoff is used for any of the interactions in the three energy terms in Eqn. 3.

The geometry optimisations were continued until the energy change between two iterations was less than 2.6 J/mol (10^{-6} a.u.) and the maximum norm of the Cartesian gradients was below 10^{-3} a.u. The QM/MM geometry optimisations were performed using the TPSS-D3 method[48,58] and the def2-SV(P)[49] basis set. For all structures, single-point QM/MM energy were also calculated with the B3LYP-D3 method and some structures were also optimised with this method. For structures within 20 kJ/mol of the most stable one, single-point QM/MM energies were calculated also at the TPSS-D3/def2-TZVPD and B3LYP-D3/def2-TZVPD levels of theory.

Quantum refinement

Quantum refinement is standard crystallographic refinement supplemented by quantum chemical calculations for a small, but interesting part of the protein.[66,67] Such calculations were performed with the ComQum-X software,[66] which is a combination of Turbomole[51] and the

crystallography and NMR system (CNS), [68,69] version 1.3. The quantum-refinement calculations were based on the same crystal structure as all the other calculations in this paper.[6] Coordinates, occupancies, *B* factors, and structure factors were obtained from the 3U7Q protein data bank files. From these files, we also obtained the space group, unit-cell parameters, resolution limits, *R* factors and the test set used for the evaluation of the R_{free} factor.

The calculations were performed the same way as in our previous study:[42] The full protein was used in all calculations, including all crystal water molecules. In each cycle of the geometry optimisation, the surrounding protein was allowed to relax by one cycle of crystallographic minimisation and one cycle of individual *B*-factor refinement. However, the new coordinates and *B* factors were accepted only if the *R* factor was reduced. For the protein, we used the standard CNS force field (protein_rep.param, water_rep.param, and ion.param). However, CNS does not support anisotropic *B* factors, so only isotropic *B* factors were used. The MM force field for non-standard residues were downloaded from the hetero-compound information centre Uppsala.[70] The w_A factor was determined by CNS to 0.0793. After quantum-refinement, anisotropic *B* factor and occupancy refinement was performed using *phenix.refine*[71]. Electron density maps were generated using *phenix.maps*. The QM calculations were performed at the TPSS/def2-SV(P) level of theory.

The quality of the models was compared using the real-space difference density Z-score[72] (RSZD), calculated by EDSTATS, which measures the local accuracy of the model. The negative RZSD value for the protonated atoms was taken as the quality metric. In all calculations, the maximum value was positive. RSZD is typically < 3.0 in absolute values for a good model.

Protonations

To simplify the discussion, we will often say protonation of the various sites of the FeMo cluster, even if protonation of the metals gives rise to hydride ions bound to the metal with a significant covalent character. To facilitate comparisons with previous studies, the atoms in the FeMo cluster were named according to the crystal structure.[6] These names are depicted in Figure 1b.

We have tried to test all reasonable protonation sites for the entire FeMo cluster and also several possible directions of the proton for each site. For the carbide ion, we tested protonation from each of the three sides of the cluster, which we will call C(2367), C(3457) and C(2456) (indicating that the proton is on the Fe2–Fe3–Fe6–Fe7, Fe3–Fe4–Fe5–Fe7 and Fe2–Fe4–Fe5–Fe6 faces, respectively). For the Cys-275 ligand, we initially tested three different directions of the proton (towards the Fe2, Fe3 and Fe4 ions, respectively), but because these were always high in energy, we tested for most systems only one direction (towards Fe4). The His-195 ligand was always protonated on the NE2 atom, which is directed towards S2B. Therefore, we added a proton on ND1 when this residue was protonated. The homocitrate ligand (HCA) was always protonated on the alcohol O7 atom with a proton shared with the carboxylate O1 atom. We tested to protonate it also on the O2 carboxylate atom, which we have shown in our previous study to be the second most stable state.[42]

For the belt (μ_2) sulfide ions (S2B, S3A, S5A), we tested two directions of the proton, viz. towards the other two belt sulfide ions. These two directions will be called S2B(3) and S2B(5) in the following (and similar for the other two ions). Protonation in the orthogonal direction is typically prohibited by hydrogen bonds to the second-sphere residues in the protein model. Likewise, two directions were tested for the Fe-side (S1A, S2A and S4A) and Mo-side (S1B, S3B and S4B) μ_3 sulfide ions, directed either towards Fe1 or Mo and therefore called S1A(Fe1) and S1A(Mo) or similar. However, in some cases, the proton moved to point towards some of the

other nearby Fe ions instead of pointing to the opposite side of the cluster (e.g. S2A(Fe2)).

For the metal ions, we initially tested several different directions of the hydrogen atoms. However, we soon found that binding trans to the carbide ion was always most favourable and therefore we in general tested only that direction for Fe2–Fe7 (and this conformation is indicated by only giving the number of the Fe ion (e.g. Fe2). For Fe1, such a binding is not possible and we initially tested three different directions, which are named after closest atom (e.g. Fe1(S2A), Fe1(Fe2) and Fe1(Fe4)). However, these were often hard to find and high in energy. Therefore, we tested only one direction for most states. The same applied to protonation of Mo, which was even harder to find, because it is six-coordinated and therefore more crowded than Fe1.

In addition, we tried to add the hydrogen atom between all pairs of metal ion that are neighbours (cf. Table 1). They are named by giving the numbers of the two metal ions, e.g. Fe2/6 or Mo/Fe7. When Fe1 or Mo are involved, only a single direction is possible. The same applies also for Fe ions on the same side (Fe- or Mo-side) of the cluster. However, for hydrides bridging Fe ions on different sides of the cluster (Fe2/6, Fe3/7 and Fe4/5), two conformations are possible, depending on which side of the belt sulfide ions they are located. As for protonation of the belt sulfides, they are called after which belt sulfide atom they are directed towards, e.g. Fe2/6(3) or Fe2/6(5).

As can be seen in Table 1, we have tested in total about 50 different protonation states (68 in the initial studies) for each oxidation and protonation level of the FeMo cluster. In many cases, the optimisation failed to give the desired states. Then we tried to obtain them by using one or two H–X distance restraints and if the geometry optimisation now was successful, these restraints were removed and the structures were reoptimised (thus, all presented structures were obtained without any restraints). If this did not give the desired state, we made no further attempts to get that state. Therefore, each table or figure does not include all 50 states, only those that were actually found.

Finally, it should be noted that several states are close to each other. For example, protonation of the carbide ion always puts the proton close also to several (up to four) Fe ions. Likewise, there is an almost continuous transition between terminal binding to a metal and bridging between two metals. The latter binding is almost always asymmetric, with a difference between the two Fe–H distances. When this difference becomes large, the binding is essential terminal. A typical terminal Fe–H binding distance is 1.51–1.56 Å and we will consider the binding bridging if both Fe–H distances are shorter than 2.3 Å. In many cases, the binding became terminal, even when started from a bridging position.

Likewise, there are sometimes finer distinctions between various binding modes. For example, protonation of the carbide may sometimes lead to a proton symmetrically positioned with a similar distance to all four neighbouring Fe ions, whereas it sometimes is more tilted towards one of the sides with two shorter and two longer distances to the Fe ions. Likewise, a bridging binding may sometimes have almost the same Fe–H distances and sometimes more dissimilar Fe–H distances. However, we have not investigated the energetics of these fine details systematically, but only report the energy and the distances of the various states found.

Result and Discussion

In this paper, we study the reduction and protonation of the FeMo cluster in nitrogenase with QM/MM methods, starting from the resting state and adding up to four electrons and protons, to obtain the E_4 state, which is believed to be the state that binds N_2.[1] As discussed in the introduction, this is a formidable task, as there are billions of possibilities for the E_4 state. Therefore, we need to use some heuristic but systematic approach to study the various

protonation states. In a recent article, we examined the problem of the BS states and designed a procedure to deal with the BS states.[37] In this study, we follow that procedure: At the start and the end of the study of each oxidation state, we consider all possible BS states with two DFT functionals, TPSS and B3LYP. For the lowest states (within 20 kJ/mol), we optimise all structures and also use large basis sets (def2-TZVPD).

For the protonation states, we use a similar procedure. For each oxidation level, we start from the protonation state that was most stable with one proton and electron less and then add the proton to all possible sites (~50 different positions). When the TPSS and B3LYP functionals give different results (for E_2–E_4), we separately investigate the best structures obtained with the two functionals. In practice, it turned out to be hard to follow this procedure strictly as new best states were sometimes found based on experience gained from the more reduced states, and the problems with the DFT functionals and BS states were found to be more serious than expected. Moreover, several structures were sometimes found close in energy and then one of them had to be selected for further investigation.

Protonation of the resting state

To start with, we performed a thorough study of various protonations of the resting state (E_0) of the FeMo cluster. As discussed in the methods section, we tested almost 70 different possible protonation states of the cluster (cf. Table 1). All calculations were performed for the quartet spin state, which is the most stable state, in agreement with experiments,[1] and the BS7-3 BS state, which is the most stable BS state, as was shown in our previous study.[37] All structures were optimised with QM/MM, using the TPSS-D3/def2-SV(P) method, restraining the distance to the proton, followed by an unrestrained optimisation. For the successful optimisations (i.e. leading to the desired structure), we also calculated single-point TPSS-D3 and B3LYP-D3 energies with the large def2-TZVPD basis set. For optimisations that did not lead to the expected state, another set of optimisations was run, exploiting the experience gained from the successful optimisations.

The obtained structures are described in Table 2. It can be seen that we found 53 distinct structures. This included most of the structures in Table 1, except those with a hydride bridging Mo and a Fe ion, as well as two of the structures with the hydride bridging Fe1 and another Fe ion, and the bridging Fe5/6 and Fe6/7 structures (those structures changed to other states during the optimisation, typically terminal binding to Fe).

The geometries were mostly as expected. The protons bound to sulfides and to the SG atom of Cys-275 with S–H distances of 1.36–1.40 Å, except when it is also close to a Fe ion, in which cases, it increased to 1.45–1.50 Å. A hydride ion terminally bound to Fe had a Fe–H bond length of 1.52–1.55 Å, whereas the Mo–H bond was 1.69 Å. When the hydride bridges two Fe ion, the two Fe–H bonds were typically similar, 1.63–1.75 Å, but there were some cases with more dissimilar distances. The C–H distance was fairly variable, 1.22–1.29 Å, depending on the interaction with the surrounding four Fe ions.

Energies of the various protonated states are also shown in Table 2. Three energies are shown, viz. the QM/MM energies from the TPSS-D3/def2-SV(P) geometry optimisations, as well as single-point QM/MM energies with the larger def2-TZVPD basis set, obtained both at the TPSS-D3 and B3LYP-D3 levels of theory. It can be seen that the increase in the basis set changed the relative TPSS energies by less than 16 kJ/mol (4 kJ/mol on average); the correlation (R^2) between the two sets of relative energies is 0.98. Therefore, we will discuss only the result with the large basis set in the remainder of this section. On the other hand, the B3LYP energies are quite different, with differences in the relative energies of up to 123 kJ/mol and a correlation coefficient of only 0.55. The difference is largest for structures in which the proton binds to the

metals, whereas when it binds to sulfur atoms, the difference is up to 27 kJ/mol. The MM energy ($E^{\mathrm{CL}}_{\mathrm{MM12},q_1=0} - E^{\mathrm{HL}}_{\mathrm{MM1},q_1=0}$ in Eqn. 3) had only a small influence on the energies, less than 12 kJ/mol.

The most favourable protonation states were observed for the central belt (μ_2) sulfide ions, S2B, S3A and S5A. In particular, protonation of the S2B ion with the orientation shown in Figure 2 (S2B(3), i.e. with the proton directed towards S3) was found to be lowest in energy. This is the atom accepting a hydrogen bond from His-195 (Figure 1a), which is believed to deliver protons to the FeMo cluster. Thus, our results support such a suggestion and show that this is the most favourable protonation state of the FeMo cluster in the resting state. The most favourable conformation of the proton is perpendicular to the Fe2–Fe6 vector and pointing away from the HE2 atom of His-195 (His-195 is protonated on NE2 and not on ND1 in all our calculations). The structure with the proton pointing in the opposite direction is 17 kJ/mol higher in energy with both TPSS and B3LYP.

Protonation of S5A is only 6 kJ/mol less stable at the TPSS level (12 kJ/mol with B3LYP). This ion receives two hydrogen bonds from water molecules and from the sidechain of Arg-359. The other Arg group is also nearby. Protonation of this ion leads to some reorganisation of the hydrogen bonds. The opposite orientation of the proton is 14 (TPSS) or 4 kJ/mol (B3LYP) higher in energy.

Protonation of S3A is 23 kJ/mol less favourable than protonation of S2B at the TPSS level (14–17 kJ/mol with B3LYP). The reason for this is that it already receives two hydrogen bonds from the backbone of Ile-355 on one side and it is close to CG of Arg-96 on the other side.

Protonation of the three μ_3 sulfide ions on the Fe side of the FeMo cluster (S1A, S2A and S4A) is 33–65 kJ/mol less favourable than protonation of S2B at the TPSS level of theory. However, these structures are more stable at the B3LYP level of theory and protonation of S4A is actually only 2 kJ/mol less stable than protonation of S2B (14–31 kJ/mol for S1A and S2A). It is stabilised by a hydrogen bond from the added proton to one of the QM water molecules. The structure with the other direction of the proton, is 8 kJ/mol less stable, but it is still the third best structure at the B3LYP level of theory. Structures with S2A protonated are also among the low-lying structures at this level of theory.

Interestingly, protonation of the corresponding μ_3 sulfide ions on the Mo side of the cluster (S1B, S3B and S4B) is much less favourable, 58–111 kJ/mol less stable than the state with S2B protonated (71–126 kJ/mol with B3LYP). A reason may be that this side is more crowded, both by the extra Mo ligands (homocitrate and His-442) and by the second-sphere residues. However, it is likely that these sites are also intrinsically less favourable, perhaps owing to the Mo ion. The most stable of these structures has S4B protonated, although the proton also interacts with Fe5 (1.75 Å distance).

The Cys-275 SG atom can also be protonated. We obtained two different conformations, but they were 54–75 kJ/mol less favourable than protonation of S2B (55–79 kJ/mol with B3LYP). Likewise, protonation of the homocitrate ligand on a carboxylate atom not coordinating to Mo was 130 kJ/mol (120 kJ/mol with B3LYP) less stable than protonation of S2B. Protonation of the proton donor His-195 on the ND1 atom (i.e. the one not hydrogen-bonding to S2B) was more unfavourable, 148 (128) kJ/mol less stable than protonation of S2B. This shows that the FeMo cluster is readily protonated by His-195.

Three structures were studied with the central C^{4-} ion protonated, viz. with the proton facing each of the three sides of the cluster, each surrounded by four Fe ions. These states were strongly disfavoured, 90–95 kJ/mol less stable than the best state at the TPSS level. However, they are more favourable with B3LYP, 38–81 kJ/mol.

Next, we studied various protonations of the metal ions, which rather represents the binding of hydride ions to the metals, after the transfer of two electrons to the proton (and therefore oxidising the metal ions two steps). Protonation of Mo was unfavourable, 128 (166) kJ/mol less stable than protonation of S2B. No structures with the hydride bridging between Mo and Fe were found.

However, a great wealth of structures with the hydride bound to the Fe ions were obtained. First, it could coordinate terminally to any of the seven Fe ions. The structures with the lowest energy typically had the hydride ion trans to the central C^{4-} ion. However, other structures were also possible and they were typically only 10–15 kJ/mol less stable. The best structures had the hydride bound to Fe3, Fe4 or Fe5 and they were 26–30 kJ/mol less stable than protonation of S2B at the TPSS level. Protonation of Fe2, Fe6 and Fe7 were 44–50 kJ/mol less stable than protonation of S2B. On the other hand, all these states were strongly disfavoured by B3LYP, 99–128 kJ/mol less stable than protonation of S2B. The Fe1 ions is special, because it is at the end of the cluster, coordinating to Cys-275. Protonation of this ion, was the least stable at the TPSS level 69–76 kJ/mol less stable than protonation of S2B, but within the range of the other ions at the B3LYP level, 115–125 kJ/mol.

Finally, several structures with the hydride ion bridging between two Fe ions were obtained. These structures were quite unfavourable, 59–96 kJ/mol less stable than protonation of S2B, with the state with the hydride bridging Fe3 and Fe7 (Fe3/7) lowest in energy. At the B3LYP level, this increased to 125–186 kJ/mol. For the special Fe1 ion, the proton comes close to a sulfide ion, giving a S–H distance of 1.53 Å and two Fe–H distances of 1.70 and 1.95 Å. This structure was 77 kJ/mol less stable than protonation of S2B at the TPSS level, i.e. in the middle among the other bridging structures. However, at the B3LYP level, it became the most stable bridging structure, 90 kJ/mol less stable than protonation of S2B.

In some cases of hydride ions bridging between the two sides (Mo and Fe1 sides) of the FeMo cluster, structures with distinctly unequal bond lengths were obtained, 1.52 Å and 2.05–2.25 Å. The former distance indicates that these are better classified as terminal hydrides binding to one Fe ion, which happen to be rather close to another Fe ion. However, in all cases, these structures were appreciably less stable than the corresponding structures with the hydride ion terminally bound trans to the central C^{4-} ion, by ~35 (~15) kJ/mol. On the other hand, this is similar to the stability of the structures with the hydride bridging two Fe ions.

In conclusion, we get the following approximate trends for the various types of protonation at the TPSS level of theory: belt sulfide < Fe terminal < Fe-side sulfide < Cys < Mo-side sulfide ≈ Fe bridging < C^{4-} < homocitrate < Mo < His-195. With the B3LYP functional, protonation of all metals is strongly disfavoured and the protonation of C^{4-} is strongly favoured. Therefore, the trend is: belt sulfide < Fe-side sulfide < C^{4-} < Cys-275 < Mo-side sulfide < Fe terminal < homocitrate < Fe bridging < His-195 < Mo. However, we have seen that the local surroundings around the protonation site may change the energy by up to 54 kJ/mol, so there are extensive variations in the relative energies also for similar protonation sites. Therefore, the best conformation of each site needs to be investigated for each electronic state of the FeMo cluster.

All the data in Table 2 were obtained for the BS7-3 state. As is detailed in our previous article, this was found to be the most stable BS state for both the best state, protonated on S2B state and the state protonated on the S2A atom, although at the B3LYP level, some BS10 states also came close in energy.[37] In all cases, the quartet state was lowest in energy.

Is the resting state protonated?

It is normally (tacitly) assumed that the resting state of nitrogenase involves a $MoFe_7S_9C$ cluster that is not protonated (besides on the homocitrate ligand, which according to recent QM and quantum-refinement investigations is in the –3 charge state with one proton shared between the alcohol group and one of the carboxylate groups[31,42]). However, this is not evident. The resting state may very well be protonated, as has been assumed in some studies (for example, Nørskov and coworkers used a doubly protonated cluster for the resting E_0 state[25]). For the continued protonation studies, it is of course mandatory to first settle the protonation state of the resting state. Fortunately, this can be done by comparing geometries of various protonated states with the high-resolution crystal structure available for the resting state of the protein,[6] as has been recently done to settle the protonation state of homocitrate and the most stable BS state.[31,42]

To this aim, we compared the structures of all protonated states discussed in the previous section with the crystal structure of nitrogenase. In particular, we calculated the root-mean-squared deviation of the coordinates of all metal ions and the directly coordinating atoms (RMSD), as well as the average and maximum deviation of all 15 short metal–metal distances (below 3.0 Å) and all 34 metal–ligand distances (always the same for all structures). The results are collected in Table 3.

It can be seen that the RMSD is lower for the unprotonated structure (0.06 Å) than for any of the protonated structures (0.07–0.23 Å). The protonated structures with the lowest RMSD are those with His (0.07 Å) and HCA (0.08 Å) protonated, which are not among the low-energy structures. Likewise, the unprotonated structure has a lower average deviation for the metal–ligand distances (0.02 Å) than any of the protonated structures (0.03–0.05 Å). The same applies for the maximum deviation, which is 0.06 Å for the unprotonated structure and 0.06–0.59 Å protonated structures. Again, the best protonated structures were those with His (0.07 Å) or HCA (0.06 Å) protonated, illustrating that protonation inside the FeMo cluster lead to larger changes in the structure than outside the cluster (for HCA and His, the protonated atoms are not directly coordinated to the Mo or Fe ions).

However, for the metal–metal distances, the results are somewhat different. For these, the structure protonated on the central C atom (on the C(3457) side) gives the lowest average and maximum difference, 0.03 and 0.06 Å, respectively, whereas the unprotonated structure gives slightly worse results, 0.05 and 0.09 Å and this result is similar to the best of the other protonated structures, which give 0.05–0.10 and 0.09–0.44 Å, respectively. Still, it is clear that the best protonated state (in energy terms, S2B(3)) is worse than the unprotonated state for all five measures.

The reason for these results is that protonation of the sulfide ions typically leads to an elongation of the S–Fe bonds for the Fe ions to which it is bound. This can directly be seen from Table 3, which shows that the largest deviation from the crystal structure is always found for a Fe–S distance, involving the protonated S ion. The effect is relatively small for the belt sulfide ions (0.09–0.18 Å), but in general larger for the μ_3 sulfide ions, 0.17–0.59 Å. The effect on the Fe–Fe bonds is much smaller, 0.10–0.14 Å, i.e. only slightly larger than the maximum error in the unprotonated structure.

Terminal hydrides on the Fe ions have some effect on the Fe–ligand bond trans to the hydride (normally the central carbide), which changes by 0.10–0.16 Å). Likewise, bridging hydrides shorten the corresponding Fe–Fe bond by 0.14–0.19 Å.

Protonation of His, Cys, HCA and C(2367) has smaller effects on the structure and therefore these structures are most similar to the crystal among the protonated states. The good Fe–Fe distances in the C(3457) structure are caused by error cancellation: The TPSS-D3/def2-SV(P) calculations give systematically slightly too short Fe–Fe bonds (by 0.05 Å on average for the

unprotonated resting state). However, protonation of the central carbide leads to a slight expansion of the FeS cluster, leading to a few Fe–Fe distances that actually are too long, but a slightly improved result on average. However, this most likely do not represent any real improvement of the structure or an indication that the resting state is actually protonated on the carbide, as the other three measures and the relative energies indicate.

To further confirm this important conclusion, we have also re-refined the crystal structure with quantum refinement, using either the unprotonated or the structures protonated on S2B(3). Although the difference-density maps are noisy due to Fourier truncation ripples that occur around the heavy metal atoms in the FeMo cluster, one can clearly notice both positive and negative difference density peaks next to the S2B atom when it is protonated (marked with red arrows in Figure 3b). These difference density peaks are not present in the unprotonated structure (Figure 3a). This is further supported by the maximum negative RSZD value of the S2B atom going from –0.3 in the unprotonated structure to –2.6 in the protonated structure. Thus, the results conclusively show that the unprotonated structure fits the experimental data best.

This is a most important result for the remainder of this study. It shows that the resting E_0 state is not protonated and gives us a firm starting point for the protonation study. In the following we will successively add one to four electrons and protons to the FeMo cluster.

The E_1 state

To reach the E_1 state, we added one electron to the various singly-protonated resting states. Since we found only minor effects (4 kJ/mol on average) when the basis set was enlarged, these calculations were performed with the def2-SV(P) basis set (but still with both the TPSS-D3 and B3LYP-D3 functionals). Single-point energies with the def2-TZVPD basis set were calculated only for the most stable states (within 20 kJ/mol for each DFT method).

The results are presented in Table 4 and they are quite similar to those obtained for the resting state. The most favourable protonation site is still S2B, shown in Figure 4. The two directions of the proton differ by 7 kJ/mol with TPSS (4 kJ/mol with the large basis set), but 35 kJ/mol at the B3LYP level (34 kJ/mol with the large basis set). The most stable direction (S2B(3)) is the same as for the resting state. For TPSS, the third and fourth best states have hydrides on Fe4 and Fe5 (27–31 kJ/mol less stable than the best state), whereas the third best state is protonated on S2A for B3LYP (65 kJ/mol). When optimised at the B3LYP level, structures protonated on the central carbide were strongly stabilised (as will be further discussed in the coming sections), but they are still 9–29 kJ/mol less stable than the structure protonated on S2B.

At the TPSS level of theory, we get the following trends: belt S < Fe terminal < Fe-side S < Mo-side S < Fe bridging < Mo–Fe bridging < C < Cys < HCA < His. At the B3LYP level, metal binding is again strongly disfavoured and carbide binding is more favoured, giving the order: belt S < Fe-side S < C < Fe terminal < Mo-side S ≈ Mo–Fe bridging < Fe bridging < Cys < HCA < His.

For the best state (protonated on S2B), we checked that BS7-3 still was the most stable BS state. As can be seen in Figure S1, this is the case by 33 and 24 kJ/mol for the TPSS and B3LYP calculations. We also confirmed that the quintet was the most stable spin state, by 33 and 46 kJ/mol before the singlet and triplet states, respectively.

The E_2 state

Next, we added another electron and studied the various doubly-protonated states of the E_2 state. To reduce the number of possibilities, we always kept S2B(3) protonated (the best protonation site for E_1) and studied the energies of the states obtained by adding a second proton to the various positions (still with the def2-SV(P) basis set with both the TPSS-D3 and B3LYP-D3 functionals). The calculations were performed for BS7-3 in the quartet state, in accordance with experimental observations.[1] The results in Table 5 are somewhat different from those obtained for the E_0 and E_1 states, showing large differences between the two QM methods.

At the TPSS-D3 level, three structures are most stable and essentially degenerate in energy. Two of them has the second hydrogen atom bridging between Fe2 and Fe6, i.e. the two Fe ions that are bridged by S2B, which is already protonated. The two structures differ in which side of S2B they are: Fe2/6(3) or Fe2/6(5) (i.e. on the sides directed towards S3A or S5A), as is shown in Figures 5a and b. In the former case, the hydride ion is on the same side as the proton on S2B, with a H–H distance of only 2.33 Å. The hydride ion is much closer to Fe2 (1.53 Å) than to Fe5 (2.31 Å). In the other structure, the two hydrogen atoms are on different sides of S2B with a H–H distance of 3.49 Å. In this case, the two Fe–H distances are somewhat less different: 1.57 and 2.01 Å for Fe2–H and Fe6–H. With the def2-TZVPD basis set, the former structure is 3 kJ/mol more stable, whereas the opposite is true with the smaller basis set. The third structure, which is 3 kJ/mol less stable than the best one with both basis sets, has a hydride ion terminally bound to Fe5 (Figure 5c). The binding position is opposite to the carbide ion and with a Fe–H distance of 1.54 Å.

The other states are significantly less stable, with hydrides on Fe4 or Fe7 22 kJ/mol less stable and Fe4/5 26 kJ/mol less stable (Table 5). The general trend is Fe bridging ≈ Fe terminal < belt S < Fe-side S < Mo-side S < Mo–Fe bridging ≈ C < Cys < His ≈ HCA.

Unfortunately, the results are totally different at the B3LYP level, with differences in the relative energies of up to 291 kJ/mol. For this functional, protonation the Fe-side sulfides and the central carbide give the most stable structures. The best structure with S2B protonated and studied in the BS7-S state involves protonation of S2A(Mo). Protonation of the other Fe-side sulfides is 19–47 kJ/mol less stable. However, protonation of the central carbide on the Fe3–Fe4–Fe5–Fe7 side is only 11 kJ/mol less stable (23 kJ/mol with the def2-TZVPD basis set). As usual, B3LYP strongly destabilises all protonation of the metals. The best state has Fe5 protonated, but it is 52 kJ/mol less stable than the state protonated on S2A(Mo) (59 kJ/mol with the large basis set). The two Fe2/6 states that were most stable with TPSS were 63 and 74 kJ/mol less stable than the state with S2A protonated (70 and 85 kJ/mol with the large basis set). At the B3LYP level, the ordering of the states is approximately: Fe-side S < C < Mo-side S < belt S ≈ Fe terminal < Fe bridging < Cys < Mo–Fe bridging < His < homocitrate.

However, based on results from the more reduced states, we found out that the structures (especially those protonated on the carbide ion) are sensitive to the method used for the optimisation and also to the BS state used. Therefore, we reoptimised the best B3LYP structures at the B3LYP-D3/def2-SV(P) level. Interestingly, then a state doubly protonated on the carbide ion (on the C(2367) and C(3457) faces) in the BS8-6 state became most stable. It has two C–H bonds of 1.13 and 1.15 Å, although the C atom remains in the centre of the cluster, as can be seen in Figure 5d. This state is 194 kJ/mol less stable than the best Fe2/6 state at the TPSS level. The second-best structure had a proton on C(3457) (in addition to S2B(3); shown in Figure 5e) and was 31 kJ/mol less stable (optimised in the BS9-2 state; 39 kJ/mol with the def2-TZVPD basis set). The third structure was 53 kJ/mol less stable than the best one (41 kJ/mol with the larger basis set) and had a proton on S2A, but directed towards Fe1 (Figure 5f).

Finally, we studied the energies for the various BS for the best protonated states. As can be

seen in Table S2, BS7-3 is still the lowest state by 23–29 kJ/mol for the best TPSS structure (Fe2/6(5)). The same applies for the S2B(3) + S2A(Fe1) state (by at 13 kJ/mol). However, for the C(2367) + C(3457) state, BS8-6 was found to be most stable, as has already been mentioned, 50–68 kJ/mol more stable than the BS7 states. Likewise, for the S2B(3) + C(3457) state, BS9-2 was found to be lowest (43–50 kJ/mol lower than the BS7 states), although in both cases other BS states are close in energy (2–3 kJ/mol).

As mentioned above, experiments have indicated that E_2 is in the quartet state[1] and therefore all calculations were performed with this state. However, at the end of the investigation, we checked the relative energies of the doublet, quartet and sextet states for the most stable protonation states. As can be seen from Table S3, three protonation states (S2B(3) + Fe2/6(3), S2B(3) + Fe5 and C(2367) + C(3457)) gave the quartet lowest in energy, whereas for the other states, the doublet (S2B(3) + C(3457) or S2A(Fe1)) or the sextet (S2B(3) + Fe2/6(5)) was lower in energy, although by only 2–9 kJ/mol for both the TPSS and B3LYP functionals. Thus, the different spin states are too close in energy to allow a conclusive assignment of the ground spin states of the FeMo cluster.

The E_3 state

Next, we added a third proton and electron to the E2 state. Since the TPSS and B3LYP calculations gave differing predictions for the lowest E2 state, we first started three series of systematic protonation of all possible sites (all with S2B(3) protonated and TPSS structures): one with a hydride bridging Fe2 and Fe6 on the S3A side (Fe2/6(3)), one protonated on S2A(Mo) and one protonated on the central carbide on the C(3457) face. The first was studied primarily with TPSS-D3, the other two with B3LYP-D3. Different BS states were tested for the different sets.

We start to discuss the TPSS results of the first series. We obtained 35 different structures from these optimisations, listed in Table 6a. All of them retained the protonation on S2B (1.37 Å) and Fe2 (1.51–1.60 Å), and in most cases, the latter hydride ion bridged also to Fe6 (1.85–2.40 Å). The longer Fe2–H distances are connected with the shorter Fe6–H distances. As can be seen from Table 6a, the best structures were those protonated on a single Fe ion. Protonation of Fe5 gave the most stable structure, 9, 18 and 18 kJ/mol more stable than protonation of Fe4, Fe6 and Fe7, respectively (7, 7 and 17 kJ/mol with the def2-TZVPD basis set). A structure with a hydride ion bridging between Fe2 and Fe6, but on the other side of S2B was also low in energy, 19 kJ/mol above the best structure (28 kJ/mol with the large basis set). In general, the ordering of the protonation states was Fe < Fe bridging < belt-S < Fe-side S < Mo-side S < C < Fe–Mo bridging < Cys < His < HCA. At the B3LYP level, the energies and ordering were quite different: The best two were protonated on S2A(Mo) or C(3457), but they were 258–261 kJ/mol less stable than the best B3LYP structure.

As discussed in the previous section, two additional structures were low in energy for the E_2 state (Fe2/6(5) and Fe5). Based on the results in Tables 5a and 6a, we tested a number of combinations that were expected to be low in energy and they are presented at the end of Table 6a. The one with the lowest energy was protonated on S2B(3), Fe2/6(5) and Fe5(Mo). It was actually 1 kJ/mol lower than the corresponding Fe2/6(3) structure (isoenergetic with the def2-TZVPD basis set). It is shown in Figure 6a. A state with S2B(3), Fe5 and Fe4 protonated (Figure 6b) was only 5 kJ/mol higher in energy and it was actually the best structure with the large basis set by 8 kJ/mol. It was most stable in the BS2 state.

Next, we consider the second series with structures protonated on S2B(3), S2A(Mo) and a third atom (Table 6b), and we discuss primarily energies obtained with B3LYP (on TPSS structures). It can be seen that the best three structures have a proton on S3B (but the proton is

also close to Fe6, 1.91 Å), on C(3457) or on Fe2. In the BS7-3 state, the three structures have the same energy within 1 kJ/mol. However, for the C(3457) structure, we found that BS2 was actually lower in energy and with that state, this structure become the most stable one by 31 kJ/mol (however, there are other better structures outside this series, as the energies in Table 6b indicate). Structures with a hydride bridging Fe2 and Fe6 on either side of S2B were 43–45 kJ/mol higher in energy and a structure with a hydride on Fe5 was 48 kJ/mol less stable than the C(3457)structure. Other structures were less stable, following the trend Mo-side S < Fe-side S < belt S < Fe–Mo bridging < His < HCA < Cys. When evaluated with the TPSS functional, the energies were quite different and (as expected) the most stable state had Fe2/6(5) protonated, but it was 31 kJ/mol less stable than the best TPSS structure (25 kJ/mol with the large basis set).

We also run calculations on structures protonated on S2B, C(3457) and a third atom (Table 6c). However, they did not give any structures that were competitive compared to the structures in the other two series: The best structure at the B3LYP level, protonated on S2A(Mo), was 81 kJ/mol less stable than the best one from Table 6b. Moreover, many of these structures changed from protonation of the carbide to bridging or terminal interactions with the nearby Fe atoms. However, one structure was found to have a low energy at the TPSS level of theory. It had two protons on Fe5 (and the third on S2B) with Fe–H distances of 1.51 and 1.53 Å and the H–H distance is 1.99 Å. It was only 11 kJ/mol less stable than the best TPSS structure (9 kJ/mol with the def2-TZVPD basis set).

In the previous section we saw that some structures were very sensitive to the DFT method used in the optimisation. Therefore, we also optimised several low-energy structures with B3LYP-D3/def2-SV(P). It turned out that a structure with the carbide ion triply protonated (i.e. on all three faces) was lowest in energy (bottom of Table 6c). However, in variance to what was observed in QM-cluster calculations,[35,73] the carbide ion stayed in the centre of the cluster as can be seen Figure 6c, although it is strongly distorted and three Fe–C bonds are broken. The second-best structure had a doubly protonated carbide (on the (3457) and (2367) faces), in addition to a proton on S2B(3). It was 30 kJ/mol less stable (46 kJ/mol with the def2-TZVPD basis set) and it is shown in Figure 6d. The third best structure had protons on S2B(3), S2A(Mo) and C(3467), being 49 kJ/mol less stable than the best structure (50 kJ/mol with the large basis set; bottom of Table 6b). Geometry optimisation lowered the B3LYP energy of the TPSS-optimised structures by 108–267 kJ/mol and changed the relative energies by up to 159 kJ/mol, showing that it is risky to employ one functional for the structures and another for the energies.

These results show that there are extensive differences between the results obtained at the TPSS and B3LYP levels of theory. In fact, the best TPSS structures are 284–400 kJ/mol higher than the best structure at the B3LYP level (259–277 kJ/mol after geometry optimisation), and the best B3LYP structure is 240 kJ/mol less stable than the best structure at the TPSS level (177 kJ/mol after geometry optimisation). Of course, this is a major problem in the computational studies of nitrogenase.

For seven of the best structures, we examined the stability of the 35 BS states. As can be seen from Table S4, for three of the structures (S2B+Fe2/6(3 or 5)+Fe5 and S2B+Fe2/6(5)+S2A), BS7-3 is still the most stable state, by 11–26 kJ/mol. For a fourth protonation state (S2B+S2A+C(3457)), BS7-1 is most stable (by 11 kJ/mol). However, for the state triply protonated on the central carbide, BS10-4 is the lowest BS state, 3 kJ/mol lower than another BS10 state, but 44 kJ/mol lower than the BS7-3 state at the B3LYP level (all states were optimised). Likewise, BS8-6 is the lowest state when the carbide doubly protonated (by 13–14 kJ/mol), 36 kJ/mol lower than the BS7-3 state, and with a singly protonated carbide, BS2 is lowest, but only 5 kJ/mol lower than BS7-3 at the TPSS level, but by 35 kJ/mol with B3LYP. Thus, the BS-state ordering is more complicated with B3LYP than with TPSS, as was also noted

before.[37]

All calculation up to this point were performed on the triplet state. However, for six low-energy states, we also checked the stability of the (open-shell) singlet and quintet states. For three of the states, we found that the triplet was lowest by 12–45 kJ/mol, as can be seen in Table S5. However, for the other three states, the singlet state was lower in energy by 5–15 kJ/mol, especially when all BS states were considered. Yet, as the energy difference is small and a full study of all possible spin and BS states would be very demanding, we have not studied the spin states further, especially as the studies on the other E_n states do not indicate that the DFT spin-state ordering can be trusted.

The E_4 state

Finally, we added a fourth electron and proton to the E_3 state. In this case, we started from three different structures: Two of the best E_3 TPSS states (with S2B(3), Fe2/6(3) and Fe5 or Fe4 protonated), which differed by only 7–9 kJ/mol, and the B3LYP state based on the TPSS structures (with S2B(3), S2A(Mo) and C(3457) protonated). For each case, we tried as usual ~50 positions for the fourth proton, of which ~30 were successful. The results are shown in Table 7 and are described in separate paragraphs.

Starting from the state with protons on S2B(3), Fe2/6(3) and Fe5, the results (shown in Table 7a) were quite similar to those of the E_3 state. About half of the structures had the second proton mainly on Fe2 (with a Fe–H distance of 1.52–1.53 Å), whereas in the remainder of the structures, it bridged also to Fe6, typically with a longer Fe6–H distance of 2.03–2.28 Å. However, in the best structure, the opposite applies. It had Fe2–H = 1.91 Å and Fe6–H = 1.56 Å, and it had the fourth proton on Fe6 (Fe–H = 1.52 Å; Figure 7a). Thus, Fe6 binds two hydride ions, but they are 2.41 Å apart (one of them bridging to Fe2). This binding also affected the electronic structure so that the best BS state was BS2. The second-best state has a hydride on Fe4 but it was 17 kJ/mol less stable (23 kJ/mol with the def2-TZVPD basis set), followed by a second hydride bridging Fe2 and Fe6 on the other side or a hydride on Fe7 (both 32 kJ/mol less stable). The general order was Fe < Fe bridging < S Fe-side < S belt < S Mo-side < Cys < C < His < HCA.

The results for the structures with protons on S2B, Fe2/Fe6 and Fe4 were similar (Table 7b). The best structure had a hydride on Fe5 with Fe–H = 1.55 Å (Figure 7b). Like the best structure in the first series, the distance for the bridging hydride was shorter to Fe6 than to Fe2, 1.65 and 1.70 Å. Interestingly, it was found to be most stable in the BS6-2 state and it was 2 kJ/mol more stable than the best state in the other series (but 3 kJ/mol less stable with the def2-TZVDP basis set). Next came a state with the fourth proton on Fe6 with a Fe6–H distance of 1.51 Å. Again, it had two protons on Fe6 (H–H = 2.41 Å) and it was in the BS2 state. It was only 2 kJ/mol less stable than the previous state and 5 kJ/mol less stable than the best state with the large basis set. Other states with hydrides on Fe4, Fe7 or Fe2/6 were 39 kJ/mol less stable than the best state.

Based on the results for E_2 and E_3, we also tested some structures with a hydride in the Fe2/6(5) position, instead of Fe2/6(3), viz. with the third and fourth protons on Fe4+Fe5, Fe4+Fe6 and Fe5+Fe6. As can be seen at the end of Tables 7a and 7b, these structures were low in energy, as expected. The first was the second-best structure at the TPSS/def2-SV(P) level, only 1 kJ/mol less stable than the best one (but the fourth best with the def2-TZVPD basis set, 10 kJ/mol less stable than the best). The other two structures were 8 and 14 kJ/mol less stable than the best (14–18 kJ/mol with the large basis set). This shows that at the TPSS level, there are several structures close in energy with hydride ions on various combinations of the Fe2, Fe4, Fe5, and Fe6 ions.

Third, we added the fourth proton to the state with protons on S2B(3), S2A(Mo) and

C(3457). As for the corresponding states for E_3, this was somewhat problematic, because at the TPSS level the proton prefers to bind to Fe ions. Therefore, for more than half of the structures, the proton on C moved to one or two of the nearby Fe ions (Fe3, Fe5, Fe7, Fe3/7 or Fe4/5), even if the structure was first optimised with the proton restrained to the C atom (Table 7c). The best structure, according to the B3LYP energies, had two protons on the central carbide (on the (3457) and (2367) faces), with C–H distances of 1.19 and 1.21 Å. The carbide ion remained in the centre of the cluster, although it is quite distorted. It was most stable with the BS2 state. This structure was 91 kJ/mol more stable than a structure with the fourth proton bridging Fe2 and Fe6 (85 kJ/mol with the def2-TZVPD basis set). The general order of the structures was: Fe bridging < C < Fe < S Mo-side < S belt < S Fe-side < His ≈ HCA ≈ Cys. Structures that did not retain the C–H bond were high in energy at the B3LYP level (167 kJ/mol or more above the best structure).

However, if the structures were optimised at the B3LYP-D3/def2-SV(P) level of theory, the situation changed and additional structures were found as can be seen in Table 7d. In particular, a structure with a triply protonated carbide (on each of the three faces with all three C–H bonds = 1.10 Å) became most stable. It was quite distorted, as can be seen in Figure 7c. This structure was 51 kJ/mol more stable than the state doubly protonated on the carbide and also on S2A(Mo) (56 kJ/mol with the def2-TZVPD basis set), described in the previous paragraph (but after B3LYP optimisation). In fact, a similar structure, but with the S2A proton in the opposite direction (S2A(Fe1), shown in Figure 7d), was somewhat better, 38 kJ/mol less stable than the best structure (44 kJ/mol with the large basis set). The fourth-best structure (protonated on S2B(3), S2A(Mo), C(3457) and Fe6/7 was 168 kJ/mol less stable than the best one.

As before, there is a major discrepancy between relative energies calculated with TPSS and B3LYP. The four structures that are most stable at the TPSS level are 339–389 kJ/mol less stable than the best B3LYP structure (after optimisation). Likewise, the best B3LYP structure is 186 kJ/mol less stable than the best structure at the TPSS level.

Recently, there has been a controversy between Siegbahn and the Hoffman group regarding the best E_4 structures. The latter have assumed that all four protons bind on the Fe2–Fe3–Fe6–Fe7 face of the cluster and they have argued that the best structure is protonated on S2B, S5A, Fe2/6 and Fe2 (i.e. two hydride ions on Fe2, but without any H–H bond).[34] We can also obtain such a structure, but it is 86 kJ/mol higher than our best structure at the TPSS level (101 kJ/mol with the def2-TZVPD basis set) and 365 kJ/mol less stable than our best structure at the B3LYP level w (Table 7e, E(4H)b). Their second-best structure is protonated on S2B, S5A, Fe2/6 and Fe3/7 (Figure 7e), and in our calculations this structure is more stable, but it is still 54 kJ/mol less stable than our best TPSS structure (264 kJ/mol with B3LYP). Their third structure (with protons on S2B, S3A, Fe2/6 and Fe3/7) is 131 kJ/mol less stable than the best TPSS structure.

On the other hand, Siegbahn has argued that structures protonated on the central carbide are much more stable than the Hoffman structures by 175–231 kJ/mol at the B3LYP level and 104–221 kJ/mol with BP86.[35] Unfortunately, these energies are based on a very strange structure employed for the Hoffman-type state (according to the structure in the supporting material). It is correctly protonated on S3B, S5A and Fe2/6, but the fourth proton is placed in the centre of the Fe1, Fe2, Fe3, Fe4 cubane cluster, giving a strongly distorted structure and unfavourable energies (202 kJ/mol higher than our best TPSS structure). On the other hand, our results confirm that when the energies are evaluated with B3LYP and especially when structures are optimised with B3LYP, protonation of the central carbide gives the most stable structures. However, our QM/MM structures are quite different from those presented by Siegbahn in that the carbide is still in the centre of the cluster, whereas in Siegbahn's structure, it has moved to the surface of the cluster, interacting with only three of the six central Fe ions and prone to dissociation as CH_4. Such structures have been hard to find in our QM/MM calculations, owing to the constraints

enforced by the surrounding protein.

Several groups have studied the reaction mechanism of nitrogenase and suggest other protonation states for E_4. For example, Adamo and coworkers used the M06-2X functional and suggested a structure with protons on S2B, S2A, S5A and the central carbide protonated.[28] When optimised with B3LYP with the BS3-2 state, it is 146 kJ/mol less stable than the best B3LYP structure and therefore our fourth best B3LYP structure (Table 7d). On the other hand, Dance used the pure BLYP functional and suggested a structure with hydride ions terminally bound on Fe2 and Fe6, in addition to protons on S2B and S3B.[74,75] In our hands, such a structure is 72 kJ/mol less stable than our best structure at the TPSS level and 426 kJ/mol less stable than the best structures as the B3LYP level (Table 8e). Likewise, McKee showed that with the B3LYP functional, a structure with the central carbide is 108 kJ/mol more stable than Dance's structure,[27] which is in agreement with our B3LYP results.

For eight low-energy states, we have studied the relative stability of all 35 BS states. All results are collected in Table S6. It can be seen that the results are quite confusing: For none of the structures is the BS7-3 lowest, although it is rather low for all TPSS structures, as well as the Hoffman structure and two of the B3LYP structures (4–18 kJ/mol higher than the best BS state). However, for one of the B3LYP structures, it 157 kJ/mol less stable than the best state. Unfortunately, there is no other state that is always most stable. Instead, the most stable state is BS6-2 for two of the TPSS structures, BS2 and BS9-2 for one of the TPSS and one of the B3LYP structures each, whereas BS10-1 and BS10-5 are most stable for the remaining two structures. However, BS6-2 is always low in energy (within 8 kJ/mol of the best) for the TPSS and Hoffman structures (evaluated with the TPSS functional), whereas BS2 is always low in energy for the B3LYP structures (within 9 kJ/mol). Even worse, the relative energies of the various structures change much when the geometry is optimised. Therefore, it seems to be necessary to determine the best BS state for each intermediate with a full investigation of the BS states, including full geometry optimisation.

Although experimental data show that the E_4 state should be in the doublet state[1,13] (the state employed in all calculations up to now), we also studied the quartet state (we did not find any reasonable sextet state) for the best structures. As can be seen in Table S7, for the best TPSS and B3LYP structures (protonated on S2B, Fe2/6, Fe4 and Fe5 or triply protonated on the central carbide), the doublet state was most stable by 14 kJ/mol or 65 kJ/mol, respectively. However, for the second-best TPSS state, the quartet was found to be 3 kJ/mol lower in energy.

Conclusions

We have performed a systematic study of the protonation of the FeMo cluster in nitrogenase in the E_0–E_4 states. We use QM/MM calculations with the entire FeMo cluster, as well as all groups that form hydrogen bonds to it within 3.5 Å in the crystal structure in the QM system and keeping the surrounding protein fixed (to avoid the local-minima problem). By employing quantum refinement and comparing geometries with that of the crystal structure, we show quite conclusively that the resting (E_0) state of the enzyme does not have any additional proton. This has been assumed in most previous studies, except by Nørskov and coworkers,[25] but it is not at all evident.

In the E_1 state, we find that the most favourable protonation site is on S2B and this also agrees with most previous theoretical studies.[1,13,24–28,16–23] However, for the E_2–E_4 states, we find new preferred protonation states that have not been suggested before and moreover, the predictions strongly depend on which DFT method is used to calculate the relative energies. With the pure TPSS functional, we obtain structures with hydride ions on the Fe ions, typically with

several structures close in energies, indicating that the hydride ions may move quite freely within the cluster. The first hydride ion bridges the Fe2 and Fe6 ions, with approximately the same energy for states with the hydride on either side of S2B. The next proton binds terminally either to Fe4 or Fe5 and the E_4 state is predicted to have two terminal hydrides on two of Fe4, Fe5 or Fe6 ions. On the other hand, the hybrid functional B3LYP disfavours metal binding and instead strongly prefers protonation of the central carbide ion. In fact, the most stable structures of the E_2, E_3 and E_4 states have the carbide ion doubly or triply protonated. This is in agreement with the suggestions by Siegbahn,[35,73] although the carbide ion stays in the centre of the (quite distorted) FeMo cluster in our QM/MM calculations. In agreement with our previous study, the energies are insensitive to the size of the basis set (changing by less than 16 kJ/mol going from def2-SV(P) to def2-TZVPD).

As discussed in the introduction, there is a very large number of possible protonation states of the FeMo cluster in nitrogenase. Therefore, a heuristic approach needs to be applied. In this study, we have made the following assumptions:

a) The best protonation state at one oxidation level can be found by adding a proton and an electron to the best state at the previous oxidation/protonation level.
b) Reliable energies can be obtained by optimising structures at the TPSS-D3/def2-SV(P) level, followed by a single-point energy calculation at the B3LYP-D3/def2-SV(P) level, and for the best states, also single-point energies by TPSS-D3 and B3LYP-D3 with the def2-TZVPD basis set.
c) All structures at the same oxidation level can be studied with the same BS and spin state. For E_0, E_2 and E_4, we have used the experimentally determined spin states, quartet, quartet and doublet, respectively.
d) It is assumed that all BS states can be reached by swapping the Fe ions.

Of course, all these assumptions can be discussed and criticised – we in no way claim that the present calculations finally solve the protonation-state problem of nitrogenase and we currently ignore several potentially important effects, e.g. entropy or conformational changes. On the contrary, our results show that the problem is much more complicated than assumed before in several aspects.

In particular, assumption c) is clearly not valid: The various protonation states often have varying preferred BS states. For example, for the eight protonation states studied for E_4, five different optimum BS states were found (cf. Table S6). Even worse, the optimum BS state is sensitive to the geometry (by > 50 kJ/mol), so studies of the BS state for the more reduced and protonated FeMo cluster require geometry optimisation of each studied BS state. Moreover, both TPSS and B3LYP seem to have problems to find the correct spin state, especially as these often are close in energy.

Likewise, assumption d) also becomes increasingly problematic when the number of protonated metals increase. Swapping of the metals assume that the metals are equivalent, but when they are protonated, they become distinct from the others, typically with a lower spin density. Then, it becomes more complicated to automatically generate all possible BS states.

Furthermore, we have also seen that even assumption a) is not always valid: The E_1 state is most stable with S2B protonated, even with B3LYP, but for E_2, double protonation of the carbide was found to be the most favourable state. Moreover, we have seen that our approach has been hard to follow in practice as new low-energy states are continuously found as the experience grows by the studies of the more reduced states. To reduce the time consumption, we decided to start the study of the next E_n level when all the def2-SV(P) calculations were finished. However, it was sometimes found that the ordering changed slightly when the larger basis set was used (although the effect is still quite restricted) and in particular when better BS states were found or

when structures were optimised by B3LYP. Therefore, the selection of protonation series can sometimes seem suboptimal in retrospect.

Throughout this study, we have found that it is very unfavourable to protonate the His, Cys and HCA ligands, typically by >100 kJ/mol (i.e. >17 pK_a units). This is somewhat alarming, because it indicates that the calculated pK_as of these groups are far from their experimental values (7–10 for free His or Cys).[76] This may indicate that the model of the FeMo cluster has a too high negative charge, i.e. that it actually should be more protonated. However, the comparison of the resting state to the crystal structure quite conclusively shows that the latter is not further protonated.

We have concentrated on protonated states of the cluster. Occasionally, we have obtained structures with H_2 bound to the metals or dissociated from the cluster, but we have not systematically studied such structures. However, we find that they typically have favourable energies, especially at the TPSS level (e.g. –48 kJ/mol for E_4). Considering that previous studies have shown that protons in the FeMo cluster seem to be able to move rather freely between the various possible protonation positions,[26,74,75] it seems to be a big challenge for any computational study of nitrogenase to explain how the enzyme avoids that the added protons and electrons leave the cluster in the form of H_2 before N_2 binds. Thus, it is not enough to show that N_2 can bind to the cluster and may be protonated – it must also be shown that H_2 formation and dissociation from any of the E_n, $n \geq 2$, state is not thermodynamically favourable and kinetically feasible.

However, the most important problem with the study of nitrogenase is that different DFT functionals give so varying results for energies, BS states, as well as geometries. In general, TPSS favours hydride ions bound to or bridging metals, whereas B3LYP favours protonation of the sulfides and especially the central carbide ion. Predictions of the relative energies of the best protonation states for E_4 differ by over 300 kJ/mol. This is the reason for the recent controversy between Siegbahn and the Hoffman group about the best protonation of the E_4 state[34,35] and probably also an important reason for the many differing mechanisms suggested for the enzyme.[1,13,24–28,16–23] Undoubtedly, we urgently need more accurate QM methods that can decide which DFT approaches give the more accurate energies. At present, we can only point out that pure functionals give more accurate structures of the resting state (compared to the crystal structure[6]) and a protonation of the E_4 state that is closer to the interpretation of the experiments (with two bridging hydride ions),[34] whereas B3LYP seems to give more reasonable energies of states involving H_2. Unfortunately, neither TPSS nor B3LYP give a structure of the E_4 state with two bridging hydride ions, as suggested by experiments,[1] but the best TPSS structure of E_2 has one protonated sulfide ion and one hydride ion, in accordance with experimental suggestions.[1]

Acknowledgements

This investigation has been supported by grants from the Swedish research council (project 2014-5540), from COST through Action CM1305 (ECOSTBio), by the Royal Physiographic Society in Lund and by a scholarship to LC from the China Scholarship Council. The computations were performed on computer resources provided by the Swedish National Infrastructure for Computing (SNIC) at Lunarc at Lund University and HPC2N at Umeå University.

Supporting Information

Definition of the 35 BS states, relative energies of the 35 BS states for selected E_1–E_4 states, relative energies of the lowest two or three spin states for selected E_2–E_4 states, relative energies and geometries for the best E_0–E_4 states obtained with the def2-SVP basis set or with system 2

optimised with MM, as well as coordinates and absolute energies for the best E_0–E_4 states. In addition, a pdb file of the entire QM/MM system is provided. This information is available free of charge via the Internet at http://pubs.acs.org.

References

(1) 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.

(2) Burgess, B. K.; Lowe, D. J. Mechanism of Molybdenum Nitrogenase. *Chem. Rev.* **1996**, *96* (7), 2983–3012.

(3) Schmid, B.; Chiu, H.-J.; Ramakrishnan, V.; Howard, J. B.; Rees, D. C. Nitrogenase. In *Handbook of Metalloproteins*; John Wiley & Sons, Ltd, 2006; pp 1025–1036.

(4) Kim, J.; Rees, D. C. Structural Models for the Metal Centers in the Nitrogenase Molybdenum-Iron Protein. *Science* **1992**, *257* (5077), 1677–1682.

(5) 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* (5587), 1696.

(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) 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* (6204), 1620–1623.

(8) Einsle, O. Nitrogenase FeMo Cofactor: An Atomic Structure in Three Simple Steps. *J. Biol. Inorg. Chem.* **2014**, *19* (6), 737–745.

(9) 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* (6058), 974–977.

(10) Eady, R. R. Structure−Function Relationships of Alternative Nitrogenases. *Chem. Rev.* **1996**, *96*, 3013–3030.

(11) Danyal, K.; Dean, D. R.; Hoffman, B. M.; Seefeldt, L. C. Electron Transfer within Nitrogenase: Evidence for a Deficit-Spending Mechanism. *Biochemistry* **2011**, *50*, 9255–9263.

(12) Thorneley, R. N. F.; Lowe, D. J. No Title. In *Molybdenum Enzymes*; Spiro, T. G., Ed.; Wiley: New York, 1985; pp 221–284.

(13) 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* (8), 3096–3103.

(14) Lukoyanov, D.; Yang, Z.-Y.; Khadka, N.; Dean, D. R.; Seefeldt, L. C.; Hoffman, B. M. Identification of a Key Catalytic Intermediate Demonstrates That Nitrogenase Is Activated by the Reversible Exchange of N2 for H2. *J. Am. Chem. Soc.* **2015**, *137*, 3610–3615.

(15) Lukoyanov, D.; Khadka, N.; Yang, Z.-Y.; Dean, D. R.; Seefeldt, L. C.; Hoffman, B. M. Reductive Elimination of H2 Activates Nitrogenase to Reduce the N≡N Triple Bond: Characterization of the E4(4H) Janus Intermediate in Wild-Type Enzyme. *J. Am. Chem. Soc.* **2016**, *138*, 10674–10683.

(16) Tuczek, F. Nitrogen Fixation in Nitrogenase and Related Small-Molecule Models: Results of DFT Calculations. In *RSC Metallobiology Series 7*; Hille, R., Schulzke, C., Kirk, M. L., Eds.; Royal Society of Chemistry: Cambridge, 2017; pp 223–274.

(17) Dance, I. Theoretical Investigations of the Mechanism of Biological Nitrogen Fixation at the FeMo Cluster Site. *J. Biol. Inorg. Chem.* **1996**, *1* (6), 581–586.
(18) Stavrev, K. K.; Zerner, M. C. Studies on the Hydrogenation Steps of the Nitrogen Molecule at the Azotobacter Vinelandii Nitrogenase Site. *Int. J. Quantum Chem.* **1998**, *70* (6), 1159–1168.
(19) Siegbahn, P. E. M.; Westerberg, J.; Svensson, M.; Crabtree, R. H. Nitrogen Fixation by Nitrogenases: A Quantum Chemical Study. *J. Phys. Chem. B* **1998**, *102* (9), 1615–1623.
(20) 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.
(21) 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* (36), 11435–11439.
(22) 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.
(23) Hallmen, P. P.; Kästner, J. N2 Binding to the FeMo-Cofactor of Nitrogenase. *Zeitschrift fur Anorg. und Allg. Chemie* **2015**, *641* (1), 118–122.
(24) Dance, I. Activation of N2, the Enzymatic Way. *Zeitschrift für Anorg. und Allg. Chemie* **2015**, *641*, 91–99.
(25) 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* (44), 29541–29547.
(26) Siegbahn, P. E. M. 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* (33), 10485–10495.
(27) 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* (5), 754–764.
(28) 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 Catal.* **2016**, *6* (3), 1567–1577.
(29) 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* (Iii), 10902.
(30) 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* (3), 1470–1477.
(31) Benediktsson, B.; Bjornsson, R. QM/MM 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.
(32) Dance, I. Activation of N2, the Enzymatic Way. *Zeitschrift für Anorg. und Allg. Chemie* **2015**, *641*, 91–99.
(33) Siegbahn, P. E. M. A Major Structural Change of the Homocitrate Ligand of Probable Importance for the Nitrogenase Mechanism. *Inorg. Chem.* **2018**, *57*, 1090–1095.
(34) Lukoyanov, D.; Khadka, N.; Dean, D. R.; Raugei, S.; Seefeldt, L. C.; Hoffman, B. M. Photoinduced Reductive Elimination of H2 from the Nitrogenase Dihydride (Janus) State Involves a FeMo-Cofactor-H2 Intermediate. *Inorg. Chem.* **2017**, *56*, 2233–2240.
(35) Siegbahn, P. E. M. Is There Computational Support for an Unprotonated Carbon in the E4

State of Nitrogenase? *J. Comput. Chem.* **2018**, *39*, 743–747.
(36) 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* (26), 11437–11449.
(37) 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 (16 pages).
(38) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chemie - Int. Ed.* **2009**, *48* (7), 1198–1229.
(39) Ryde, U. *QM/MM Calculations on Proteins*; 2016.
(40) 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.
(41) Hu, L.; Söderhjelm, P.; Ryde, U. On the Convergence of QM/MM Energies. *J. Chem. Theory Comput.* **2011**, *7* (3), 761–777.
(42) 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**, *121*, 8242–8262.
(43) Barney, B. M.; Mcclead, J.; Lukoyanov, D.; Laryukhin, M.; Yang, T.; Dean, D. R.; Hoffman, B. M.; Seefeldt, L. C. Diazene (HN=NH) Is a Substrate for Nitrogenase : Insights into the Pathway Of. *Biochemistry* **2007**, *46*, 6784–6794.
(44) Case, D. A.; Berryman, J. T.; Betz, R. M.; Cerutti, D. S.; Cheatham, T. E.; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; et al. AMBER 14. University of California: San Francisco 2014.
(45) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.
(46) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79* (2), 926–935.
(47) Hu, L.; Ryde, U. Comparison of Methods to Obtain Force-Field Parameters for Metal Sites. *J. Chem. Theory Comput.* **2011**, *7* (8), 2452–2463.
(48) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the Density Functional Ladder: Non-Empirical Meta-Generalized Gradient Approximation Designed for Molecules and Solids. *Phys. Rev. Lett.* **2003**, *91* (14), 146401.
(49) Schäfer, A.; Horn, H.; Ahlrichs, R. Fully Optimized Contracted Gaussian Basis Sets for Atoms Li to Kr. *J. Chem. Phys.* **1992**, *97* (4), 2571–2577.
(50) Besler, B. H.; Merz, K. M.; Kollman, P. A. Atomic Charges Derived from Semiempirical Methods. *J. Comput. Chem.* **1990**, *11* (4), 431–439.
(51) Furche, F.; Ahlrichs, R.; Hättig, C.; Klopper, W.; Sierka, M.; Weigend, F. Turbomole. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4* (2), 91–100.
(52) Becke, A. D. Density-Functional Exchange-Energy Approximation With Correct Asymptotic-Behavior. *Phys. Rev. A* **1988**, *38* (6), 3098–3100.
(53) 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* (2), 785–789.
(54) Becke, A. D. A New Mixing of Hartree–Fock and Local Density-Functional Theories. *J. Chem. Phys.* **1993**, *98* (2), 1372.

(55) 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* (18), 3297–3305.
(56) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. Auxiliary Basis-Sets to Approximate Coulomb Potentials. *Chem. Phys. Lett.* **1995**, *240* (4), 283–289.
(57) Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. Auxiliary Basis Sets for Main Row Atoms and Transition Metals and Their Use to Approximate Coulomb Potentials. *Theor. Chem. Acc.* **1997**, *97* (1–4), 119–124.
(58) 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* (15), 154104 (19 pages).
(59) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32* (7), 1456–1465.
(60) Greco, C.; Fantucci, P.; Ryde, U.; Gioia, L. de. Fast Generation of Broken-Symmetry States in a Large System Including Multiple Iron–Sulfur Assemblies: Investigation of QM/MM Energies, Clusters Charges, and Spin Populations. *Int. J. Quantum Chem.* **2011**, *111* (14), 3949–3960.
(61) Szilagyi, R. K.; Winslow, M. A. On the Accuracy of Density Functional Theory for Iron—Sulfur Clusters. *J. Comput. Chem.* **2006**, *27* (12), 1385–1397.
(62) Ryde, U. The Coordination of the Catalytic Zinc in Alcohol Dehydrogenase Studied by Combined Quantum-Chemical and Molecular Mechanics Calculations. *J. Comput. Aided. Mol. Des.* **1996**, *10*, 153–164.
(63) Ryde, U.; Olsson, M. H. M. Structure, Strain, and Reorganization Energy of Blue Copper Models in the Protein. *Int. J. Quantum Chem.* **2001**, *81*, 335–347.
(64) Reuter, N.; Dejaegere, A.; Maigret, B.; Karplus, M. Frontier Bonds in QM/MM Methods: A Comparison of Different Approaches. *J. Phys. Chem. A* **2000**, *104* (8), 1720–1735.
(65) Cao, L.; Ryde, U. On the Difference between Additive and Subtractive QM/MM Calculations. *Front. Chem.* **2018**, *6*, 89.
(66) Ryde, U.; Olsen, L.; Nilsson, K. Quantum Chemical Geometry Optimizations in Proteins Using Crystallographic Raw Data. *J. Comput. Chem.* **2002**, *23* (11), 1058–1070.
(67) Ryde, U.; Nilsson, K. Quantum Chemistry Can Locally Improve Protein Crystal Structures. *J. Am. Chem. Soc.* **2003**, *125* (47), 14232–14233.
(68) Brunger, A. T.; Adams, P. D.; Clore, G. M.; Delano, W. L.; Gros, P.; Grosse-kunstleve, R. W.; Jiang, J.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; et al. Crystallography & NMR System : A New Software Suite for Macromolecular Structure Determination. *Acta Crystallogr. D* **1998**, *54*, 905–921.
(69) Brunger, A. T. Version 1.2 of the Crystallography and NMR System. *Nat. Protoc.* **2007**, *2* (11), 2728–2733.
(70) Kleywegt, G. J. Crystallographic Refinement of Ligand Complexes. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2006**, *63* (1), 94–100.
(71) Afonine, P. V; Ralf, W.; Headd, J. J.; Thomas, C. Towards Automated Crystallographic Structure Refinement with Phenix.Refine. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2012**, *68*, 352–367.
(72) Tickle, I. J. Statistical Quality Indicators for Electron-Density Maps. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2012**, *68*, 454–467.
(73) Siegbahn, P. E. M. 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* (33), 10485–10495.

(74) Dance, I. The Chemical Mechanism of Nitrogenase: Calculated Details of the Intramolecular Mechanism for Hydrogenation of H2-N2 on FeMo-Co to NH3. *Dalt. Trans.* **2008**, *2*, 5977–5991.
(75) Dance, I. Ramifications of C-Centering Rather than N-Centering of the Active Site FeMo-Co of the Enzyme Nitrogenase. *Dalt. Trans.* **2012**, *41* (16), 4859.
(76) Ullmann, G. M.; Knapp, E. W. Electrostatic Models for Computing Protonation and Redox Equilibria in Proteins. *Eur. Biophys. J.* **1999**, *28* (7), 533–551.

Table 1. A list of the tested protonation sites, their grouping and the number of tested conformations in the initial (E_0 state; in brackets) and bulk (E_1–E_4) studies.

Site	Group	# directions
Cys-275		1 (3)
His-195		1
HCA		1
C		3
S2B	Belt sulfide	2
S3A		2
S5A		2
S1A	Fe-side sulfide	2
S2A		2
S4A		2
S1B	Mo-side sulfide	2
S3B		2
S4B		2
Fe1	Fe1	1 (3)
Fe2	Terminal Fe	1 (3)
Fe3		1 (3)
Fe4		1 (3)
Fe5		1 (3)
Fe6		1 (3)
Fe7		1 (3)
Mo	Mo	1 (3)
Fe1/2	Fe bridging	1
Fe1/3		1
Fe1/4		1
Fe2/3		1
Fe2/4		1
Fe3/4		1
Fe2/6		2
Fe3/7		2
Fe4/5		2
Fe5/6		1
Fe5/7		1
Fe6/7		1
Mo/Fe5	Mo–Fe bridging	1
Mo/Fe6		1
Mo/Fe7		1
Total		**50 (68)**

Table 2. Results for the protonated resting state (E_0), giving the protonation site, the X–H distance, as well as other short distances to the added proton (Å; S2B=2.02 means that the S2B–H distance is 2.02 Å and so on) and three sets of energies (kJ/mol; SV and TZ are the def2-SV(P) and def2-TZVPD basis sets).

Site	X–H	Other	TPSS		B3LYP
			SV	TZ	TZ
Cys(Fe2)	1.39		87	75	77
Cys(Fe4)	1.36		61	54	55
His	1.09	S2B=2.02	151	148	128
HCA	0.99		128	130	120
C(2367)	1.24	Fe3=1.74, Fe7=1.76	87	95	51
C(3457)	1.29	Fe3=1.96, Fe4=1.81, Fe5=1.90	88	95	81
C(2456)	1.22	Fe4=1.77, Fe5=1.78	81	90	38
S2B(3)	1.37		0	0	0
S2B(5)	1.37		19	17	17
S3A(2)	1.37		20	23	14
S3A(5)	1.36		25	23	16
S5A(3)	1.37		4	6	12
S5A(2)	1.38		22	19	16
S1A(Fe1)	1.39	Fe1=2.21	31	33	22
S1A(Mo)	1.39		44	49	31
S1A(Cys)	1.53		113	114	99
S2A(Mo)	1.37	S_{Cys}=1.85, Fe1=2.00	35	38	14
S2A(Fe1)	1.40		35	36	30
S2A(Fe2)	1.48	Fe2=1.84	77	72	78
S4A(Fe1)	1.38		46	47	10
S4A(Fe4)	1.36		57	65	2
S1B(Mo)	1.38		92	97	108
S1B(Fe5)	1.48	Fe5=1.78	76	77	103
S3B(Mo)	1.38		107	111	126
S3B(Fe6)	1.45	Fe6=1.87	71	75	90
S4B(Mo)	1.40		77	87	81
S4B(Fe5)	1.50	Fe5=1.75	57	58	71
Fe1(S2A)	1.52		67	69	123
Fe1(Fe2)	1.52		78	76	125
Fe1(Fe4)	1.54		76	76	115
Fe2	1.53		45	50	101
Fe2(S1A)	1.53		62	52	103
Fe2(Fe3)	1.53		59	62	109
Fe3	1.54		28	27	106
Fe3(S4A)	1.52	Fe7=2.16	65	72	129
Fe4	1.53		27	26	110
Fe5	1.53		29	30	117
Fe5(S4B)	1.52	Fe4=2.25	61	71	126
Fe6	1.55		43	45	99
Fe7	1.54		37	44	127
Fe7(S4B)	1.52	Fe3=2.05	61	73	136
Mo(Fe7)	1.69		114	128	166
Fe1/3	1.70	1.95, S2A=1.53	79	77	90
Fe2/3	1.70	1.71, S2A=1.99	91	96	149
Fe2/4	1.69	1.71	84	91	127
Fe3/4	1.67	1.71	76	92	125
Fe2/6(3)	1.75	1.75	70	72	183
Fe2/6(5)	1.74	1.63	74	79	175
Fe4/5(2)	1.66	1.71	60	63	186
Fe4/5(5)	1.68	1.69	86	95	166
Fe3/7(3)	1.64	1.74	86	94	157
Fe3/7(2)	1.68	1.64	47	59	178
Fe5/7	1.77	1.63	88	96	145

Table 3. RMSD, as well as average and maximum metal–metal and meta–ligand distances for the unprotonated (No) and the various protonated states of the resting state, compared to the crystal structure.[6] The best result is highlighted in bold face in each column.

Site	RMSD	Metal–metal distances			Metal–ligand distances		
		Av.	Max	Bond	Av	Max	Bond
No	**0.063**	0.054	0.092	Fe6–Fe7	**0.023**	**0.058**	Fe2–C
Cys(Fe2)	0.101	0.070	0.142	Fe1–Fe3	0.028	0.081	Mo–O1
Cys(Fe4)	0.093	0.075	0.168	Fe1–Fe3	0.031	0.113	Fe1–S_{Cys}
His	0.074	0.052	0.113	Fe5–Fe7	0.026	0.068	Mo–O1
HCA	0.079	0.060	0.090	Fe5–Mo	0.027	0.061	Fe5–C
C(2367)	0.096	0.054	0.124	Fe2–Fe3	0.036	0.149	Fe3–C
C(3457)	0.081	**0.029**	**0.060**	Fe1–Fe3	0.027	0.090	Mo–O1
C(2456)	0.104	0.066	0.206	Fe4–Fe5	0.033	0.146	Fe4–C
S2B(3)	0.094	0.063	0.099	Fe2–Fe4	0.036	0.174	Fe2–S2B
S2B(5)	0.091	0.062	0.099	Fe2–Fe4	0.036	0.183	Fe2–S2B
S3A(2)	0.093	0.061	0.122	Fe3–Fe4	0.033	0.119	Fe5–S3A
S3A(5)	0.090	0.063	0.119	Fe5–Fe7	0.033	0.111	Fe5–S3A
S5A(3)	0.121	0.064	0.114	Fe3–Fe4	0.035	0.090	Fe7–S5A
S5A(2)	0.110	0.062	0.105	Fe5–Fe7	0.036	0.112	Fe3–S5A
S1A(Fe1)	0.125	0.064	0.123	Fe1–Fe3	0.038	0.340	Fe1–S1A
S1A(Mo)	0.123	0.063	0.131	Fe1–Fe3	0.039	0.369	Fe1–S1A
S2A(Mo)	0.177	0.061	0.127	Fe1–Fe4	0.049	0.588	Fe1–S2A
S2A(Fe1)	0.145	0.057	0.111	Fe1–Fe4	0.044	0.399	Fe1–S2A
S4A(Fe1)	0.123	0.057	0.109	Fe5–Fe7	0.050	0.461	Fe1–S4A
S4A(Fe4)	0.122	0.069	0.137	Fe5–Fe7	0.049	0.270	Fe4–S4A
S1B(Mo)	0.106	0.060	0.138	Fe5–Fe7	0.044	0.345	Fe5–S1B
S1B(Fe5)	0.100	0.054	0.103	Fe5–Fe7	0.036	0.206	Fe5–S1B
S3B(Mo)	0.119	0.053	0.116	Fe5–Fe7	0.039	0.174	Mo–S3B
S3B(Fe6)	0.104	0.058	0.104	Fe5–Fe6	0.040	0.221	Fe6–S3B
S4B(Mo)	0.100	0.060	0.117	Fe5–Fe6	0.042	0.361	Fe5–S4B
S4B(Fe5)	0.087	0.055	0.110	Fe3–Fe4	0.035	0.192	Fe5–S4B
Fe1(S2A)	0.122	0.066	0.168	Fe1–Fe2	0.032	0.144	Fe1–S2A
Fe2	0.102	0.052	0.108	Fe3–Fe4	0.034	0.097	Fe2–C
Fe3	0.119	0.050	0.147	Fe5–Fe7	0.035	0.159	Fe3–C
Fe4	0.113	0.056	0.157	Fe5–Fe7	0.035	0.126	Fe4–C
Fe5	0.105	0.057	0.143	Fe3–Fe4	0.036	0.128	Fe5–C
Fe5(S4B)	0.136	0.065	0.143	Fe5–Fe7	0.029	0.139	Fe5–S8
Fe6	0.102	0.053	0.136	Fe5–Fe7	0.035	0.164	Fe6–C
Fe7	0.109	0.058	0.141	Fe3–Fe4	0.036	0.145	Fe7–C
Mo	0.226	0.096	0.440	Fe5–Mo	0.040	0.201	Mo–S4B
Fe1/3	0.113	0.055	0.114	Fe5–Fe7	0.037	0.294	Fe1–S2A
Fe2/3	0.089	0.056	0.136	Fe2–Fe3	0.032	0.130	Fe2–S2A
Fe2/4	0.107	0.056	0.141	Fe2–Fe4	0.032	0.156	Fe2–S1A
Fe3/4	0.138	0.045	0.148	Fe3–Fe4	0.033	0.126	Fe4–S4A
Fe2/6(3)	0.137	0.060	0.186	Fe2–Fe6	0.028	0.079	Mo–O1
Fe2/6(5)	0.108	0.060	0.188	Fe2–Fe6	0.029	0.090	Fe5–C
Fe4/5(2)	0.146	0.065	0.191	Fe4–Fe5	0.028	0.094	Mo–O1
Fe4/5(5)	0.110	0.049	0.126	Fe4–Fe5	0.027	0.088	Fe2–C
Fe3/7(3)	0.108	0.046	0.136	Fe3–Fe7	0.027	0.090	Fe2–C
Fe3/7(2)	0.179	0.067	0.193	Fe3–Fe7	0.027	0.097	Mo–O1
Fe5/7	0.090	0.047	0.144	Fe5–Fe7	0.031	0.096	Mo–O1

Table 4. Results for the various protonated states of the E_1 state, showing the protonated site, the X–H bond length of the added proton, as well as other short distances to this proton (Å) and three sets of relative energies (kJ/mol; energies in brackets are obtained with def2-TZVPD basis set). In the B3Opt column, structures were optimised at the B3LYP-D3/def2-SV(P) level of theory. The reference structure was protonated on S2B(3).

Site	X–H	Other	TPSS	B3LYP	B3Opt
Cys	1.41		140	168	
His	1.02		206	203	
Hca	0.99		181	175	
C(2367)	1.26	Fe3=1.72, Fe7=1.74	88	85	29
C(3457)	1.29	Fe3=2.04, Fe5=1.82 Fe4=1.81, Fe7=2.08	90	105	9
S2B(3)	1.37		0 (0)	0 (0)	0
S2B(5)	1.38		7 (4)	35 (34)	
S3A(5)	1.37		57	87	
S3A(2)	1.37		55	69	
S5A(3)	1.37		32	72	
S5A(2)	1.38		53	79	
S1A(Mo)	1.37		69	91	
S1A(Fe1)	1.39		53	80	
S2A(Mo)	1.37		58	65	
S2A(Fe1)	1.40		59	80	
S4A(Fe1)	1.38		74	81	
S4A(Fe4)	1.36		75	72	
S1B(Fe5)	1.50	Fe5=1.78	82	130	
S1B(Mo)	1.39		92	119	
S3B(Mo)	1.41		106	138	
S3B(Fe6)	1.56	Fe6=1.67	71	109	
S4B(Fe5)	1.50	Fe5=1.76	71	116	
S4B(Mo)	1.39		91	116	
Fe1	1.52		76	168	
Fe2	1.53		51	131	
Fe3	1.52		43	129	
Fe4	1.54		27	96	
Fe5	1.53		31	141	
Fe6	1.51		37	120	
Fe7	1.51		40	158	
Fe3/4	1.73	1.66	85	155	
Fe3/7(3)	1.70	1.65	85	180	
Fe4/5(5)	1.79	1.60	83	181	
Fe4/5(2)	1.97	1.53	141	248	
Mo/Fe5	1.78	1.68	93	120	
Mo/Fe7	1.74	1.80	86	111	

Table 5. Results for the various protonated E_2 states of the state, giving the second protonated site (the first proton is always on S2B(3), except for the C_2 structures), the two X–H bond lengths, as well as other short distances to the protons (Å) and five sets of relative energies (kJ/mol). In the upper part, the structures are optimised with TPSS and in the lower part by B3LYP. The reference state is protonated on Fe2/6(5) for TPSS, and on C(2367) + C(3457) (= C_2) for B3LYP and B3Opt.

Site	X–H distances (Å)			def2-SV(P)			def2-TZVPD	
	S2B	2nd	Other	TPSS	B3LYP	B3Opt	TPSS	B3LYP
Cys	1.37	1.43		113	263			
His	1.37			172	313			
HCA	1.37			170	318			
C(2367)	1.38	1.38	2.12,1.71,2.11,1.70	95	237			
C(3457)	1.38	1.18	1.81,2.27,2.22,2.02	104	157	75	102	149
C(2456)	1.38	1.26	2.17,1.73,1.74,2.21	92	205			
S3A(2)	1.38	1.37		57	198	128	42	201
S3A(5)	1.38	1.39		52	241			
S5A(2)	1.38	1.38		30	216	129	18	200
S5A(3)	1.38	1.37		37	200	112	19	188
S1A(Fe1)	1.38	1.38	Fe1=2.29	47	177	104	25	158
S1A(Mo)	1.38	1.38		55	193		42	180
S2A(Mo)	1.38	1.39		47	146	90	34	126
S2A(Fe1)	1.38	1.39	Fe1=2.22	47	191	53	31	179
S4A(Fe4)	1.37	1.38		50	174	99	53	168
S4A(Fe1)	1.37	1.42	Fe1=1.93	53	165	97	49	156
S1B(Mo)	1.37	1.42		80	225			
S3B(Mo)	1.37	1.42		84	216			
S3B(Fe6)	1.37	1.42	Fe6=1.91	66	186			
S4B(Mo)	1.37	1.44		90	230			
S4B(Fe5)	1.37	1.45		93	244			
Fe1	1.38	1.52		33	228			
Fe2	1.37	1.53		27	212		20	197
Fe3	1.38	1.52		53	336			
Fe4	1.38	1.54		22	250			
Fe5	1.37	1.54		3	198	159	0	185
Fe6	1.38	1.51		35	255			
Fe7	1.37	1.51		22	216	178	21	206
Fe3/4	1.38	1.65	1.74	81	287			
Fe2/6(5)	1.38	1.57	2.01	0	220	156	0	212
Fe2/6(3)	1.37	1.53		2	209	153	-3	196
Fe3/7(2)	1.38	1.52		53	344			
Fe3/7	1.37	2.28	1.51	28	233			
Fe4/5(2)	1.38	2.21	1.52	26	236			
Fe4/5(5)	1.37	1.69	1.78	64	274			
Mo/Fe5	1.37	1.80	1.67	111	271			
MoFe7	1.38	1.74	1.79	95	326			
C_2[b]	1.13	1.16	1.87,2.20,2.15,2.01 1.83,2.12,1.81,2.23	155	0	0	170	0
Structures optimised with B3LYP-D3/def2-SV(P)								
S2A(Fe1)	1.36	1.37		133		53	124	41[c]
C(3457)[a]	1.36	1.11	2.21,2.12,2.09,2.34	127		31	122	39[c]
C_2[b]	1.13	1.15	2.39,2.13,2.24,2.40 2.20,2.17,2.22	194		0	228	0[c]

[a] Studied in the BS9-2 state.
[b] Protons on both C(2367) and C(3457) and not on S2B(3). Studied in the BS8-6 state.
[c] With the C(2367)+C(3457) structure optimised with B3LYP as the reference

Table 6. Results for the various E_3 states, listing the protonated site, the X–H bond lengths, as well as other short distances to the added protons (Å) and up to five sets of relative energies (kJ/mol; B3Opt means optimised with B3LYP-D3/def2-SV(P); values in brackets in (c) are for the def2-TZVPD basis set). All structures are protonated on S2B(3) (except C_3). In (a) the second proton (hydride) is on Fe2/6(3). In (b), it is on S2A(Mo). In (c) it is on central carbide on the C(3457) face. Each energy column has a reference value (0 kJ/mol) that is common for the three tables (S2B(3) + Fe2/6(5) + Fe5(Mo) for TPSS and C(2367) + C(3457) + C(2456) (= C_3) for B3LYP and B3Opt).

(a) Protonated on S2B(3), Fe2/6(3) and a third site

Site	X–H distance					def2-SV(P)			def2-TZVPD	
	S2B	Fe2	Fe6	3rd	Other	TPSS	B3LYP	B3Opt	TPSS	B3LYP
Cys	1.37	1.53	2.27	1.42		93	345			
His	1.37	1.53				176	430			
HCA	1.37	1.53	2.30			178	433			
C(2367)	1.37	1.52		1.38	2.27,1.68,2.10,1.70	79	333	125		
C(3457)	1.37	1.53	2.26	1.20	1.76,2.26,2.28,2.00	89	261	196	87	271
C(2456)	1.37	1.56		1.12	Fe2=2.14	109	265	214	119	268
S3A(2)	1.37	1.53		1.37		46	290			
S3A(5)	1.37	1.53		1.37		51	298			
S5A(2)	1.37	1.53		1.38		38	331			
S5A(3)	1.37	1.52	2.27	1.37		30	291			
S1A(Fe1)	1.37	1.54	2.08	1.40	Fe1=2.13	32	275	214	26	263
S1A(Mo)	1.37	1.55	1.95	1.37		43	273			
S2A(Mo)	1.37	1.52		1.39		44	258	207	31	241
S2A(Fe1)	1.37	1.52	2.08	1.39	Fe1=2.23	39	298			
S4A(Fe4)	1.37	1.53		1.38		48	282			
S4A(Fe1)	1.37	1.52		1.38	Fe1=2.27	53	295			
S1B(Mo)	1.37	1.54	2.21	1.41		79	336			
S3B(Mo)	1.37	1.60	1.85	1.41		73	323			
S3B(Fe6)	1.37	1.57	1.96	1.39	Fe6=2.17	60	302			
S4B(Mo)	1.37	1.53		1.41		87	336			
S4B(Fe5)	1.37	1.53		1.45		98	358			
Fe1	1.37	1.53		1.52		36	348			
Fe3	1.37	1.53	2.21	1.53		56	369			
Fe4	1.37	1.52		1.54		10	341		7	330
Fe5	1.37	1.53	2.27	1.54		**1**	284	**259**	**0**	319
Fe6	1.37	1.54	2.11	1.52		19	336	288	7	323
Fe7	1.37	1.53		1.51		19	334	301	17	325
Fe2/6(5)	1.37	1.51	2.14	1.61	1.75	20	321	284	28	321
Fe3/7(2)	1.37	1.52		1.52	2.16	39	362			
Fe3/7(3)	1.37	1.52		2.37	1.51	27	342			
Fe4/5(2)	1.37	1.52		2.34	1.51	39	355			
Fe4/5(5)	1.37	1.55	2.03	2.27	1.52	35	364			
MoFe5	1.37	1.53	2.26	1.80	1.68	93	358			
MoFe7	1.37	1.53	2.26	1.75	1.78	85	348			
With Fe2/6(5) instead of Fe2/6(3)										
Fe4	1.38	1.56	2.04	1.53		13	345	284	19	343
Fe5	1.37	1.56	2.04	1.54		7	324	272	9	318
Fe5(Mo)	1.38	1.55	2.13	1.54		**0**	314	264	**0**	307
Fe6	1.38	1.58	1.90	1.51		24	355	296	23	347
Fe7	1.38	1.56	2.04	1.51		21	344	288	25	339
With Fe5 instead of Fe2/6(3)										
Fe4[a]	1.38	1.56		1.55		5	400	277	-8	385
Fe6	1.38	1.55		1.51		35	360			
Fe7	1.37	1.54		1.50		20	324			

[a] Studied in the BS2 state.

(b) Protonated on S2B(3), S2A(Mo) and a third site.

Site	X–H distance				def2-SV(P)		def2-TZVPD	
	S2B	S2A	Third	Other	TPSS	B3LYP	TPSS	B3LYP
Cys	1.38	1.38	1.44		155	369		
His	1.37	1.39			213	359		
HCA	1.38	1.39			215	363		
C(3457)[a]	1.37	1.38	1.16	1.89,2.29, 2.28,2.01	108	**237** **205**[a]	109	
S3A(2)	1.38	1.38	1.37		103	302		
S3A(5)	1.38	1.38	1.39		95	274		
S5A(2)	1.38	1.39	1.38		82	288		
S5A(3)	1.38	1.37	1.37		65	292		
S1A(Fe1)	1.38	1.38	1.40	Fe1=2.05	95	289		
S1A(Mo)	1.38	1.38	1.39		104	290		
S4A(Fe4)	1.38	1.38	1.38		85	270		
S4A(Fe1)	1.38	1.38	1.38		104	283		
S1B(Mo)	1.37	1.39	1.42		124	277		
S3B(Mo)	1.37	1.39	1.43		126	265		
S3B(Fe6)	1.37	1.39	1.42	Fe6=1.91	111	**236**	101	229
S4B(Mo)	1.38	1.38	1.44		132	269		
S4B(Fe5)	1.37	1.39	1.45		135	285		
Fe1	1.38	1.38	1.52		48	321		
Fe2	1.38	1.38	1.56		61	**237**	43	235
Fe3	1.38	1.38	1.51		80	378		
Fe4	1.38	1.38	1.54		47	309	34	290
Fe5	1.38	1.38	1.53		41	253	25	234
Fe6	1.38	1.38	1.51		60	315		
Fe7	1.38	1.39	1.51		68	282		
Fe1/3	1.38	1.38	1.62	1.86	35	323	21	306
Fe2/6(5)	1.37	1.38	1.57	1.91	31	248	25	233
Fe2/6(3)	1.37	1.39	1.52	2.35	37	384	24	237
Fe3/7(2)	1.38	1.39	1.53	2.08	83	277		
Fe4/5(2)	1.38	1.39	2.21	1.52	69	298		
Fe4/5(5)	1.37	1.39	2.22	1.52	76	322		
MoFe5	1.38	1.38	1.82	1.66	133	374		
MoFe7	1.38	1.38	1.75	1.77	133	218		
Optimised with B3LYP/def2-SV(P)[b]								
C(3457)	1.36	1.36	1.12	2.02,2.28, 2.36,2.04	212	49	211	50
S3B(Fe6)[a]	1.36	1.36	1.37	Fe6=2.30	216	186	218	185
Fe2	1.36	1.36	1.57		153	195		
Fe2/6(3)	1.36	1.37	1.52		133	200		
Fe2/6(5)	1.37	1.36	1.54	2.25	131	204		

[a] Studied in the BS2 state.

[b]With the C(2367)+C(3457)+C(2456) structure optimised with B3LYP-D3/def2-SV(P) rather than TPSS as the reference for the B3LYP energies.

(c) Protonated on S2B(3), C(3457) and a third site.

Site	X–H distance (Å)								Energy (kJ/mol)	
	S2B	C	Fe3	Fe4	Fe5	Fe7	Third	Other	TPSS	B3LYP
Cys	1.37	1.19	1.64	2.07	2.12	1.73	1.43		192	398
His	1.38	1.26	1.77	2.14	2.14	1.81			252	435
HCA	1.38	1.26	1.77	2.15	2.16	1.81			254	438
S3A(2)	1.38			2.05	1.56		1.37		39	293
S3A(3)	1.38			1.94	1.57		1.37		41	312
S5A(2)	1.37		2.02			1.55	1.37		43	310
S5A(3)	1.37		1.92			1.55	1.37		32	294
S1A(Fe1)	1.38	1.31	1.76	2.21	2.23	1.71	1.38	Fe1=2.34	111	309
S1A(Mo)	1.38	1.29	1.73	2.10	2.19	1.83	1.38		118	305
S2A(Fe1)	1.38		2.25			1.52	1.39	Fe1=2.23	73	335
S2A(Mo)	1.38	1.25	1.77	2.10	2.21	1.81	1.38		116	**286**
S4A(Fe1)	1.38			1.54			1.40		71	348
S3B(Mo)	1.37			2.30	1.52		1.42		111	352
S4B(Mo)	1.38	1.26	1.86	2.22	2.14	1.75	1.41		167	335
S4B(Fe5)	1.38	1.28	2.17	1.83	1.75	2.14	1.43		173	360
Fe1	1.38	1.27	1.75	2.16	2.22	1.79	1.52		114	348
Fe2	1.37	1.30	1.76	2.05	2.15	1.80	1.55		108	322
Fe4	1.38			2.39	1.51		1.54		44	391
Fe5	1.37		2.30			1.51	1.54		23	329
$Fe5_2$	1.37				1.51		1.57		**11 (1)**	325
Fe6	1.38		2.29			1.51	1.51		55	376
Fe7	1.38	1.21	1.82	2.32	2.21	1.80	1.50		91	315
Fe2/6(5)	1.38	1.27	1.72	2.23	2.19	1.76	1.55	2.04	73	328
Fe2/6(3)	1.37	1.25	1.74	2.15	2.24	1.80	1.52	2.29	73	309
Fe3/7(2)[a]	1.38	1.24	1.75	2.21	2.32	1.76	1.53	2.15	104	368
Fe4/5(2)	1.38		2.31			1.51	2.30	1.52	47	360
Mo/Fe7	1.38	1.26	1.74	2.22	2.36	1.74	1.75	1.82	138	368
C_3[c]	1.13[d]	1.15		2.02	2.03	2.11	1.11		177(189)	**0 (0)**
Optimised with B3LYP/def2-SV(P)[e]										
C(2367)[b]	1.36	1.12	2.28	2.29	2.07		1.12		231(252)	30 (46)
C(2456)	1.36	1.13	2.11		2.27	2.19	1.11		223(218)	71 (69)
S3A(5)	1.36	1.13	2.04	2.15	2.08	2.21	1.35		254	195
S5A(3)	1.36	1.13	2.21	2.05	2.16	2.13	1.36		207	128
Fe2	1.36	1.14	2.13	2.06	2.03	2.19	1.54		231	188
C_3[c]	1.10[d]	1.15	2.11		2.35		1.10		240	**0 (0)**

[a]Studied in the BS2 state. [b]Studied in the BS8-6 state.

[c]C_3 = C(2367)+C(3457)+C(2456); studied in the BS10-4 state.

[d]C(2367)

[e]With the C(2367)+C(3457)+C(2456) structure optimised with B3LYP-D3/def2-SV(P) rather than TPSS as the reference for the B3LYP energies.

Table 7. Results for the various E_4 states, listing the fourth protonated site, the four X–H bond lengths, as well as other short distances to the added protons (Å) and three sets of relative energies (kJ/mol; values in brackets are calculated with the def2-TZVPD basis set; B3Opt means that the structure was optimised by B3LYP-D3/def2-SV(P)). Nearly all structures are protonated on S2B(3). In (a) a hydride ion is bridging Fe2 and Fe6 on the S3A side and another is terminally bound to Fe5. In (b), the third proton is terminally bound to Fe4. In (c) there are protons on S2B, S2A and on the central carbide ion (Fe3–F4–Fe5–Fe7 face). In (d), structures are optimised with B3LYP-D3/def2-SV(P). In (e), three Hoffman-type and one Dance structure are examined. In all cases, the reference structure for the energies is protonated on S2B(3), Fe2/6(5), Fe5 and Fe6 for TPSS and on S2B(3), C(2367), C(3457) and C(2456) for B3LYP.

(a) Protonated on S2B(3), Fe2/6(3), Fe5 and a fourth site.

Site	H–X distance						def2-SV(P)		
	S2B	Fe2	Fe6	Fe5	4th	Other	TPSS	B3LYP	B3Opt
Cys	1.37	1.53	2.23	1.52	1.42		105	384	
His	1.37	1.52	2.23	1.53			210	501	
HCA	1.37	1.52	2.24	1.52			211	505	
C(3457)	1.37	1.52		1.52	1.20	1.83,2.20,2.11,2.11	109 (113)	292	
S3A(2)	1.37	1.53		1.52	1.37		56	319	
S3A(5)	1.37	1.53		1.52	1.36		56 (85)	321	
S5A(2)	1.37	1.52		1.52	1.38		53	351	
S5A(3)	1.37	1.52		1.52	1.37		50	341	
S1A(Fe1)	1.37	1.54	2.10	1.52	1.39		42	326	
S1A(Mo)	1.37	1.55	1.90	1.52	1.38		53	309	
S2A(Mo)	1.37	1.53	2.22	1.51	1.39		52	299	
S2A(Fe1)	1.37	1.53	1.99	1.52	1.39	Fe1=2.23	60	347	
S4A(Fe4)	1.37	1.52		1.51	1.38		62 (73)	316	
S4A(Fe1)	1.37	1.53		1.52	1.38		69	322	
S1B(Mo)	1.37	1.54	2.15	1.53	1.40		82 (84)	356	
S3B(Mo)	1.37	1.59	1.84	1.52	1.41		89	364	
S3B(Fe6)	1.37	1.56	1.96	1.52	1.39	Fe6=2.20	72	339	
S4B(Mo)	1.37	1.53		1.53	1.40		96	363	
S4B(Fe5)	1.37	1.53		1.51	1.45		113	396	
Fe1	1.37	1.53		1.51	1.52		49	383	
Fe3	1.37	1.53	2.26	1.50	1.52		82	430	
Fe4	1.37	1.52	2.28	1.51	1.55		17 (23)	366	
Fe5	1.37	1.52		1.50	1.51		38	374	
Fe6[a]	1.37	1.91	1.56	1.51	1.52		**0 (0)**	362 (362)	389
Fe7	1.37	1.52	2.30	1.52	1.51		32	389	
Fe7(Fe3)	1.37	1.52		1.52	1.51		38	378	
Fe2/6(5)	1.37	1.52	2.04	1.52	1.60	1.75	32	362	
Fe3/7(2)[a]	1.37	1.75	1.61	1.51	1.52	2.14	44	422	
With Fe2/6(5) instead of Fe2/6(3)									
Fe4[b]	1.37	1.75	1.62	1.55	1.55		-1 (10)	343 (379)	
Fe6[a]	1.37	1.91	1.57	1.56	1.52		12 (18)	369 (375)	

[a] Studied in the BS2 state.
[b] Studied in the BS6-2 state.

(b) Protonated on S2B(3), Fe2/6(3), Fe4 and a fourth site.

Site	H–X distance (Å)						def2-SV(P)		
	S2B	Fe2	Fe6	Fe4	4th	Other	TPSS	B3LYP	B3Opt
Cys	2.11	1.52		1.54	1.42		108	413	
His	1.52	1.52		1.54			190	485	
HCA	1.37	1.52		1.54			191	490	
C(3457)	1.37	1.52		1.55	1.19	1.79,2.16, 2.20, 2.17	113(120)	323 (332)	
S3A(2)	1.37	1.53		1.56	1.38		64	343	
S3A(5)	1.37	1.53		1.58	1.38		64	354	
S5A(2)	1.37	1.52		1.53	1.38		52	371	
S5A(3)	1.37	1.52		1.54	1.37		42	360	
S1A(Fe1)	1.37	1.53	2.28	1.55	1.39		44 (45)	323 (320)	
S1A(Mo)	1.37	1.52	2.19	1.54	1.38		62 (61)	318 (321)	
S2A(Mo)	1.37	1.52		1.54	1.38		47 (41)	319 (315)	
S2A(Fe1)	1.37	1.51	2.16	1.54	1.39	Fe1=2.27	55 (52)	349	
S4A(Fe4)	1.37	1.52	2.28	1.54	1.39	Fe1=2.23	53	331	
S4A(Fe1)	1.37	1.52		1.54	1.39	Fe1=2.21	53	330	
S1B(Mo)	1.37	1.53	2.19	1.54	1.41		85	383	
S3B(Mo)	1.37	1.55	2.04	1.54	1.42		98	392	
S3B(Fe6)	1.37	1.54	2.13	1.54	1.40	Fe6=2.04	83	358	
S4B(Mo)	1.38	1.62	Fe5=1.82	1.56	1.40		91	410	
S4B(Fe5)	1.37	1.52		1.53	1.44		106	414	
Fe1	1.37	1.53		1.54	1.53		50	398	
Fe2	1.37	1.54		1.55	1.52		55 (10)	315	
Fe4[a]	1.37	1.76	1.61	1.51	1.54	H–H=2.06	37	392	
Fe5[b]	1.37	1.70	2.65	1.55	1.55		-2 (3)	341 (347)	350
Fe6[b]	1.37	1.89	1.57	1.56	1.51		0 (5)	322 (327)	339
Fe7	1.37	1.52		1.55	1.51		37	407	
Fe7(Fe3)	1.37	1.52		1.54	1.51	Fe3=2.42	45	394	
Fe2/6(5)	1.37	1.52	2.06	1.54	1.63	1.72	37	380	
Fe3/7(2)	1.37	1.52		1.55	1.54	2.11	73	444	
Fe4/5(5)	1.37	1.53	2.15	1.53	2.24	1.52	49	415	
With Fe2/6(5) instead of Fe2/6(3)									
Fe6[b]	1.37	2.01	1.56	1.55	1.51		6 (14)	331 (339)	

[a] Studied in the BS2 state.
[b] Studied in the BS6-2 state.

(c) Protonated on S2B(3), S2A(Mo), C(3457) (although the C–H bond is often broken) and a fourth site.

4th	H–X distance (Å)									def2-SV(P)	
Site	S2B	S2A	C	Fe3	Fe4	Fe5	Fe7	4th	Other	TPSS	B3LYP
Cys	1.38	1.38		2.11			1.54	1.44		193 (172)	450
His	1.38	1.38	1.26	1.76	2.09	2.20	1.82			304	447
HCA	1.38	1.38	1.25	1.76	2.11	2.21	1.82			306	448
C(2367)[a]	1.38	1.38	1.21		1.89	1.76		1.19	1.81,2.21, 1.80	188 (207)	**145 (163)**
S3A(2)	1.38	1.38			1.73	1.66		1.37		99	339
S3A(2)	1.38	1.38			2.05	1.56		1.37		113 (106)	333
S3A(5)	1.38	1.38			2.05	1.56		1.37		103 (90)	323
S5A(2)	1.38	1.39		1.79			1.65	1.38		113	346
S5A(3)	1.38	1.39		2.10			1.53	1.37		107	327
S5A(3)	1.38	1.39		1.99			1.55	1.37		109	358
S1A(Mo)	1.38	1.38		2.24			1.52	1.39		136	348
S1A(Fe1)	1.38	1.38		2.29			1.52	1.40	Fe1=2.04	138	365
S4A(Fe1)	1.38	1.38					1.52	1.39		145	351
S4A(Fe4)	1.38	1.38				1.52		1.38		135 (140)	350
S4A(Fe4)	1.38	1.37		2.20			1.52	1.38		130	348
S1B(Mo)	1.38	1.38		1.51				1.39		150	372
S1B(Mo)	1.38	1.38				1.51		1.40		150	369
S3B(Mo)	1.37	1.38	1.32	2.16	1.92	1.78	2.19	1.43		211	359
S3B(Fe6)	1.38	1.38	1.26	1.80	1.71	2.16	1.97	1.43	Fe6=1.90	205 (201)	325
S3B(Fe6)	1.38	1.38	1.30	2.18	2.20	1.80	2.25	1.42	Fe6=1.92	201 (197)	343
S3B(Mo)	1.37	1.37					1.51	1.39		139	312
S3B(Mo)	1.38	1.38					1.51	1.39		149	372
S4B(Mo)	1.38	1.38	1.23	1.85	2.21	2.27	1.78	1.41		223	324
S4B(Mo)	1.38	1.38				1.54		1.40		133	326
S4B(Fe5)	1.38	1.38					1.54	1.40		155	360
S4B(Fe5)	1.38	1.38			2.05	1.53		1.45		185	370
Fe1	1.38	1.38	1.29	1.81	2.07	2.26	1.69	1.51		146	382
Fe1	1.38	1.38					1.52	1.52		91	398
Fe2	1.38	1.38	1.26	1.81	1.76	2.16	1.79	1.54		150 (139)	317
Fe3[a]	1.38	1.38	1.20	1.65	2.16	2.23	1.92	1.52		147	390
Fe4	1.38	1.37	1.27	2.18	1.76	1.78		1.54		128	382
Fe4	1.38	1.38					1.52	1.55		94 (92)	386
Fe5	1.38	1.38	1.23		1.77	1.77	2.24	1.53		134	340
Fe5 other	1.38	1.38	1.23		2.22	1.77	2.24	1.53		136	341
Fe5[b]	1.38	1.38				1.51		1.53		74 (79)	327
Fe6[a]	1.38	1.38	1.19	1.78			1.79	1.52		145	284
Fe7	1.38	1.38	1.22	1.79	2.29	2.27	1.78	1.51		139	336
Fe7	1.38	1.37					1.51	1.52	Fe3=2.37	76 (76)	340 (344)
Fe1/3	1.38	1.38	1.23	1.75			1.78	1.66	1.77	112	372
Fe2/6(5)	1.37	1.38	1.18	1.78			1.95	1.56	1.90	135 (144)	235 (248)
Fe2/6(5a)	1.38	1.38	1.26	1.80	2.06	2.18	1.82	1.55	1.98	130 (137)	286 (297)
Fe2/6(3)[a]	1.37	1.38	1.26	1.73	2.14	2.27	1.79	1.52	2.26	119 (116)	304
Fe3/7(2)[a]	1.38	1.38	1.24	1.76	2.16	2.31	1.77	1.54	1.77	152	323
Fe4/5(2)	1.38	1.38	1.28	1.74	2.01		1.84	2.35	1.51	161	369
Fe4/5(2)	1.38	1.38			1.81	1.57		1.83	1.55	128	405
Mo/Fe5	1.37	1.38				1.53		1.83	1.67	157	333
Mo/Fe7	1.38	1.38					1.51	1.76	1.82	165	385
C_3[d]	1.37	1.12[e]	1.13	2.01		2.15	2.26	1.11		186 (197)	**0 (0)**

[a] Studied in the BS2 state. [b] Studied in the BS10-5 state.

[c]Studied in the BS9-8 state. Fe3–H=2.15, Fe6–H=2.12 and Fe7–H=2.25.

[d]C_3 = S2B(3) + C(2367)+C(3457)+C(2456); studied in the BS2 state.

[e]C(2367)

(d) Various structures optimised with B3LYP-D3/def2-SV(P). The first six are structures related to those in Table 7c.

	H–X distance (Å)							def2-SV(P)	
Structure	S2B	2nd proton		3rd proton		4th proton		TPSS	B3LYP
C(2367)[a]	1.36	S2A(Mo)	1.37	C(3457)	1.10	C(2367)	1.11	303	51 (56)
S5A(2)	1.36	S2A(Mo)	1.36	C(3457)	1.18	S5A(2)	1.36	287	283
S3B(Mo)	1.36	S2A(Mo)	1.36	C(3457)	1.13	S3B(Mo)	1.39	317	183
S3B(Mo)	1.36	S2A(Mo)	1.36	Fe7	1.56	S3B(Mo)	1.37	250	276
Fe7	1.36	S2A(Mo)	1.36	Fe7	1.50	Fe7	1.51	181	316
Fe2/6[b]	1.36	S2A(Mo)	1.36	C(3457)	1.11	Fe2/6(5)	2.33/1.52	212	177
Fe6/7[c]	1.36	S2A(Mo)	1.36	C(3457)	1.12	Fe6/7	1.55/2.04	237	168
C_2[d]	1.36	S2A(Fe1)	1.36	C(3457)	1.10	C(2367)	1.11	333(359)	38 (44)
C_3[e]	1.36	C(2456)	1.10	C(3457)	1.10	C(2367)	1.10	308(329)	0 (0)
Adamo[f]	1.36	S2A(Mo)	1.36	S5A	1.36	C(3457)	1.13	288	146

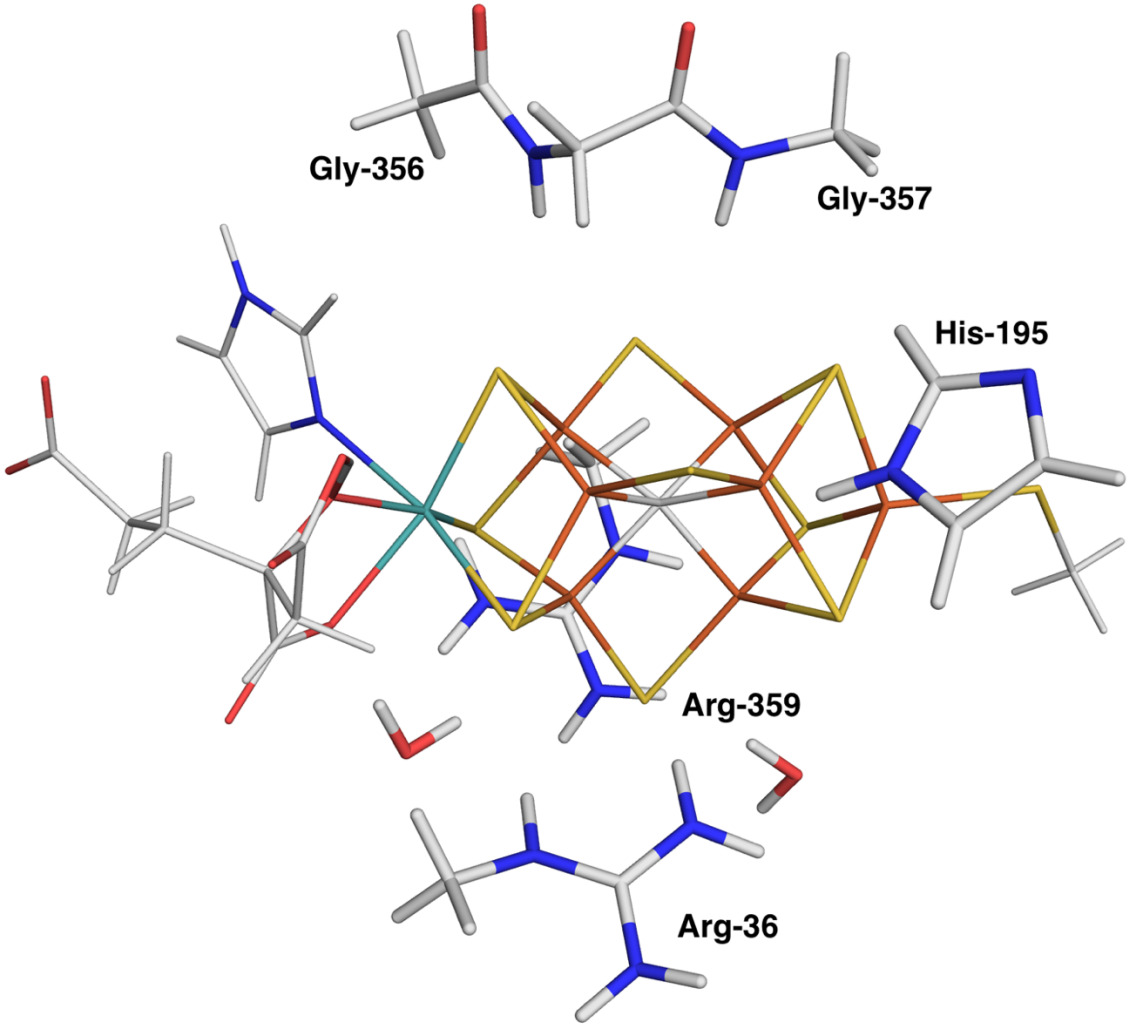
[a]Studied in the BS9-2 state. [b]Studied in the BS10-1 state. [c]Studied in the BS2 state
[d]C_2 = S2B(3) + S2A(Fe1) + C(2367) + C(3457); studied in the BS10-5 state.
[e]C_3 = S2B(3) + C(2367)+C(3457)+C(2456); studied in the BS10-4 state.
[f]S2B(3) + S2A(Mo) + S5A(3) + C(3457); studied in the BS3-2 state.

(e) Hoffman-type (S2B(5)+Fe2/6(5)+two more on the same Fe2–Fe3–Fe6–Fe7 face) and Dance-type (S2B(3)+S3B(Fe6)+Fe2+Fe6) structures.

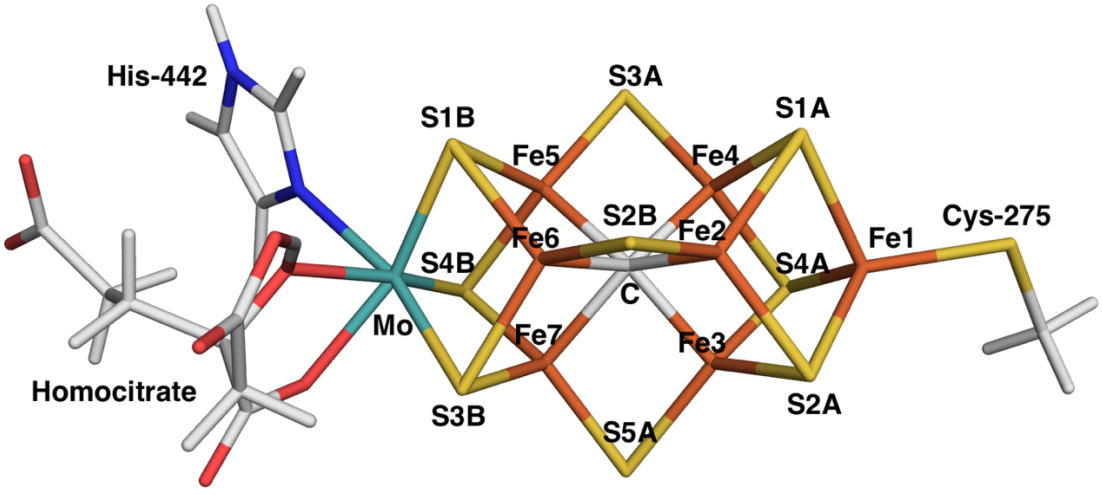
Structure	H–X distance (Å)									def2-SV(P) (kJ/mol)		
	S2B	Fe2	Fe6	3rd proton		4th proton		Other		TPSS	B3LYP	B3Opt
E(4H)[a]	1.36	1.53	1.80	S5A	1.36	Fe3	1.86	Fe7	1.57	52 (60)	370 (378)	264
E(4H)b	1.37	1.58	1.68	S5A	1.37	Fe2	1.55			84 (101)	377	365
E(4H)c	1.36	1.53	1.79	S3A	1.37	Fe3	2.00	Fe7	1.52	131	433	412
Dance	1.37	1.57		Fe6	1.56	S3B	1.56	Fe6	2.13	72	517	426

[a]Studied in the BS9-2 state.

Figure 1. FeMo cluster, illustrating (a) the QM system used in the QM/MM calculations with the hydrogen-bonding residues marked, and (b) the FeMo cluster with atom names indicated.



a



b

Figure 2. The best protonated E_0 structure, with the proton on the S2B(3) atom (green sphere), obtained at the TPSS-D3/def2-SV(P) level of theory.

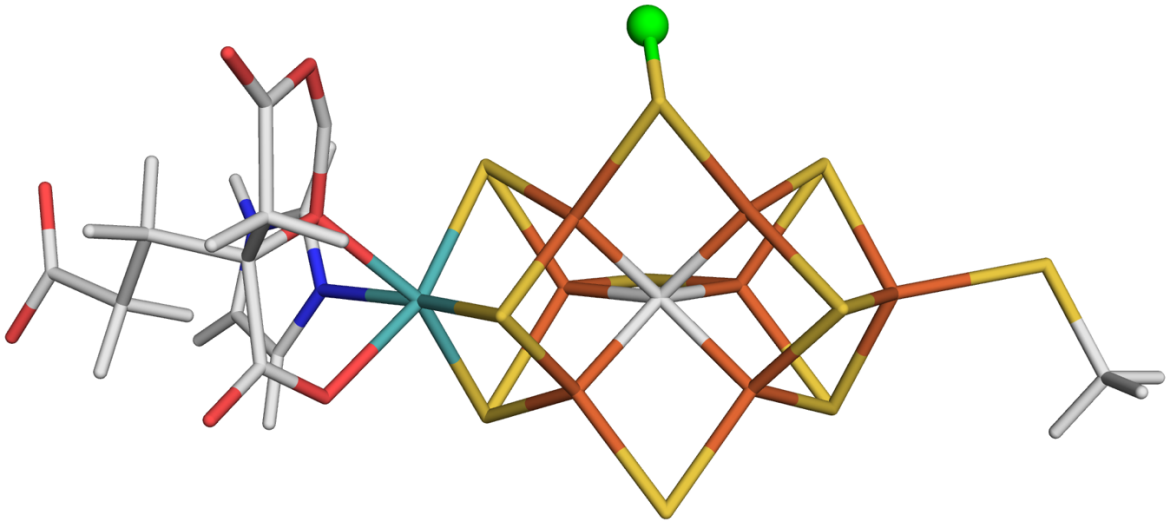
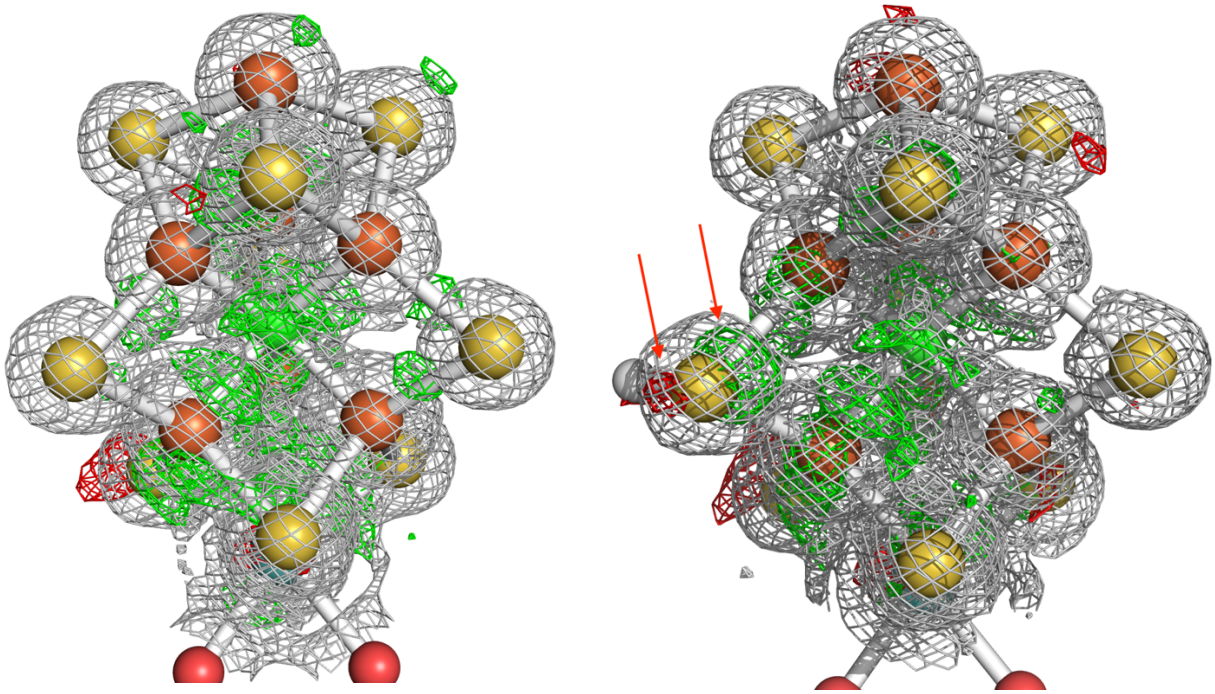


Figure 3. Electron-density maps of the resting E_0 state (a) without or (b) with a proton on the S2B(3) atom (middle left side) in the quantum refinement of the nitrogenase 3U7Q structure.[6] The $2mF_o–DF_c$ maps are contoured at 1.0 σ and the $mF_o – DF_c$ maps are contoured at +3.0 σ (green) and –3.0 σ (red). The red arrows point out the extra difference density for the protonated structure.



a b

Figure 4. The best E_1 structure, protonated on the S2B(3) atom, obtained at the TPSS-D3/def2-SV(P) level of theory.

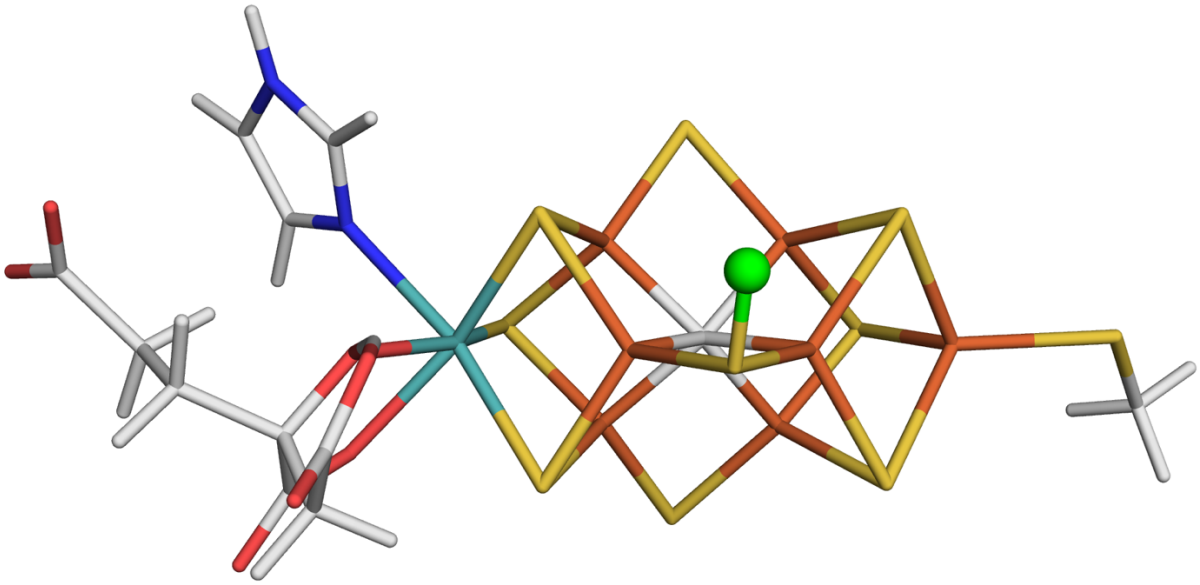
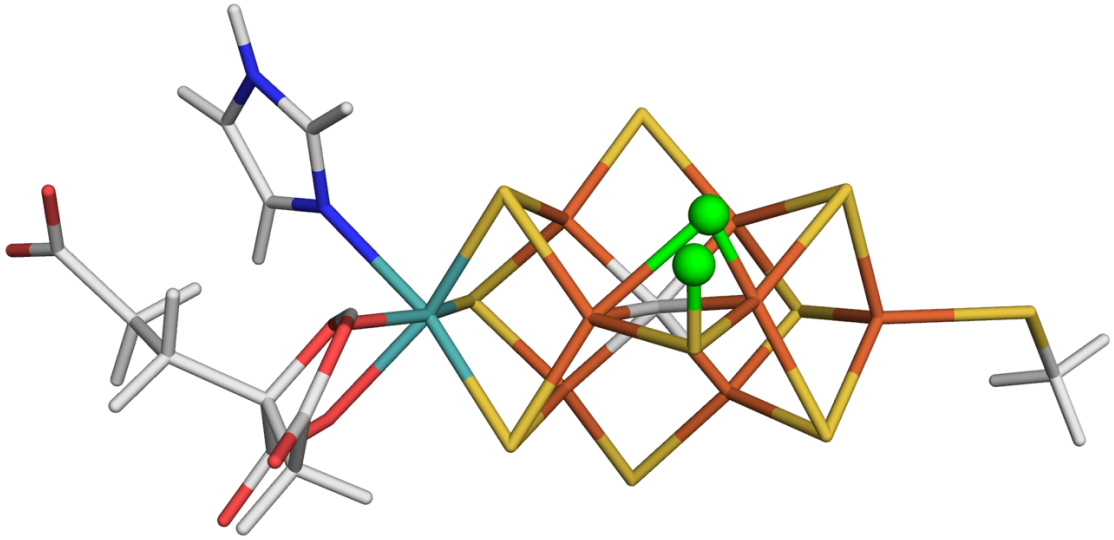
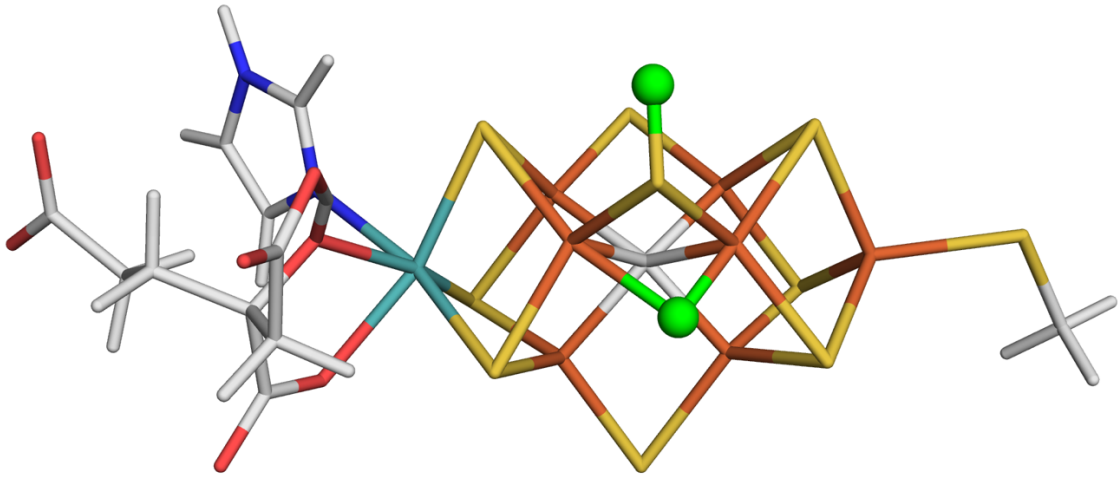


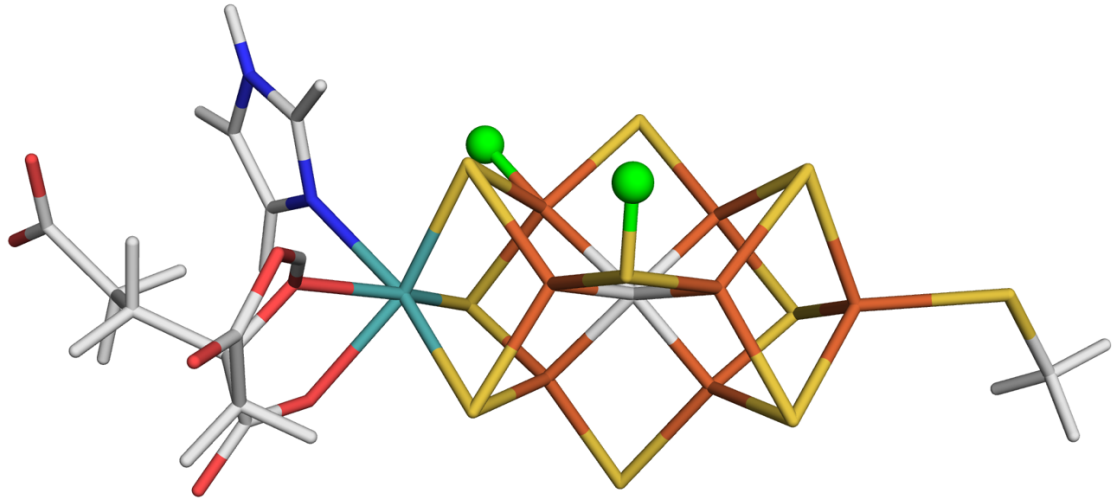
Figure 5. The best E_2 structures protonated on the following sites: (a) S2B(3) + Fe2/6(3), (b) S2B(3) + Fe2/6(5), (c) S2B(3) + Fe5, (d) C(2367) + C(3457), (e) S2B(3) + C(3457), (f) S2B(3) + S2A(Fe1). Structures (a)–(c) were obtained at the TPSS-D3/def2-SV(P) level of theory, whereas structures (d)–(f) were obtained with B3LYP-D3/def2-SV(P).



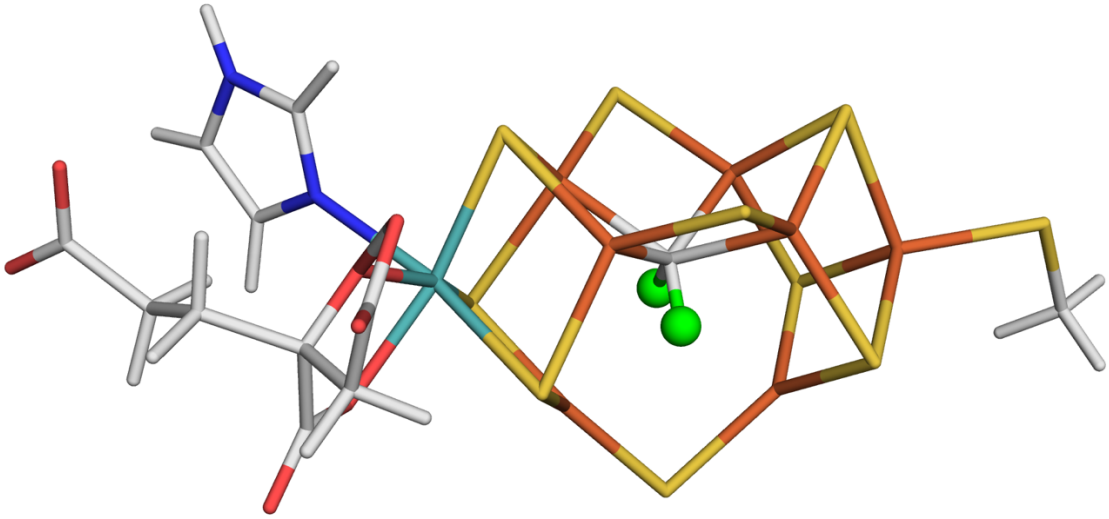
a



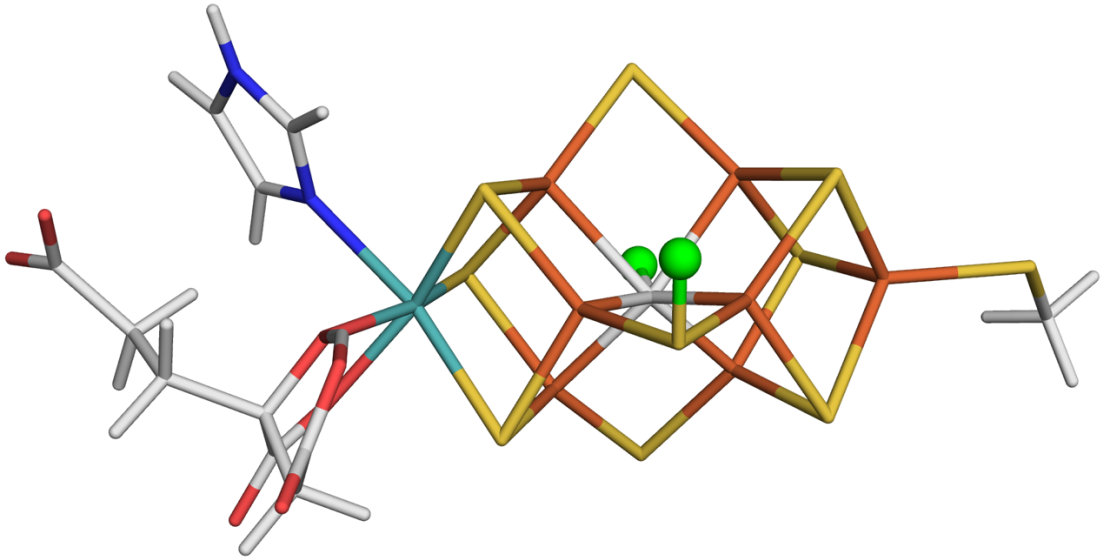
b



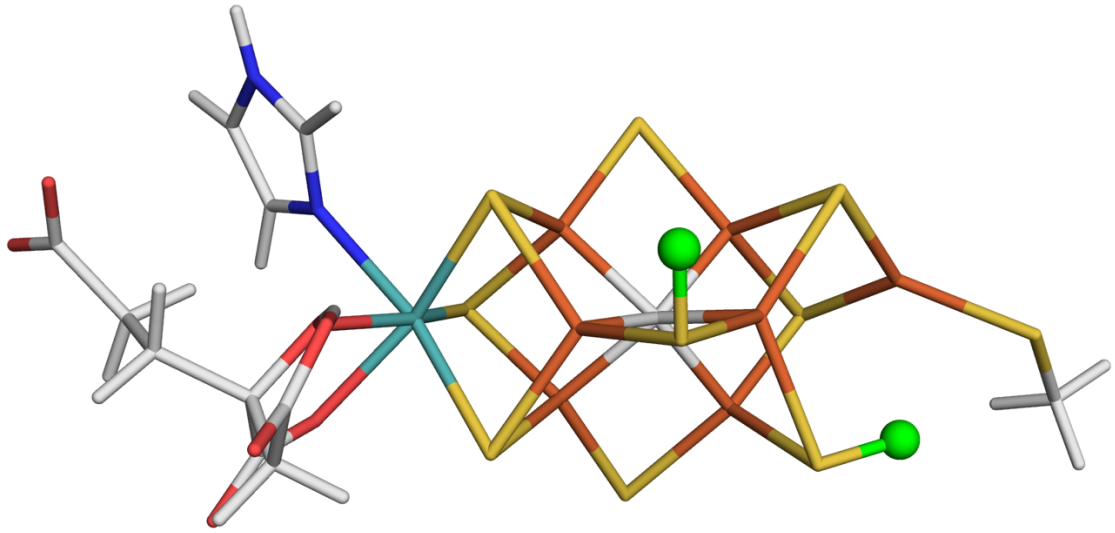
c



d

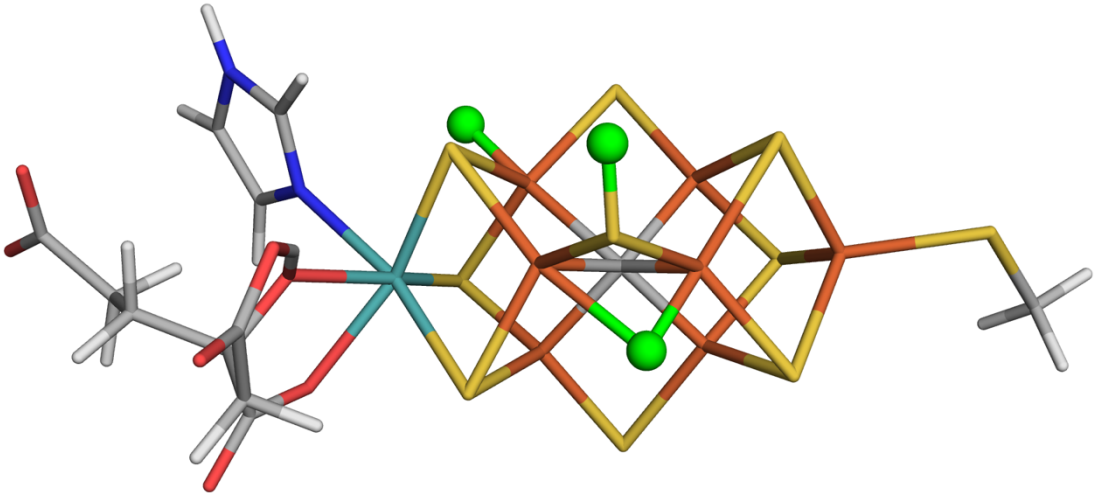


e

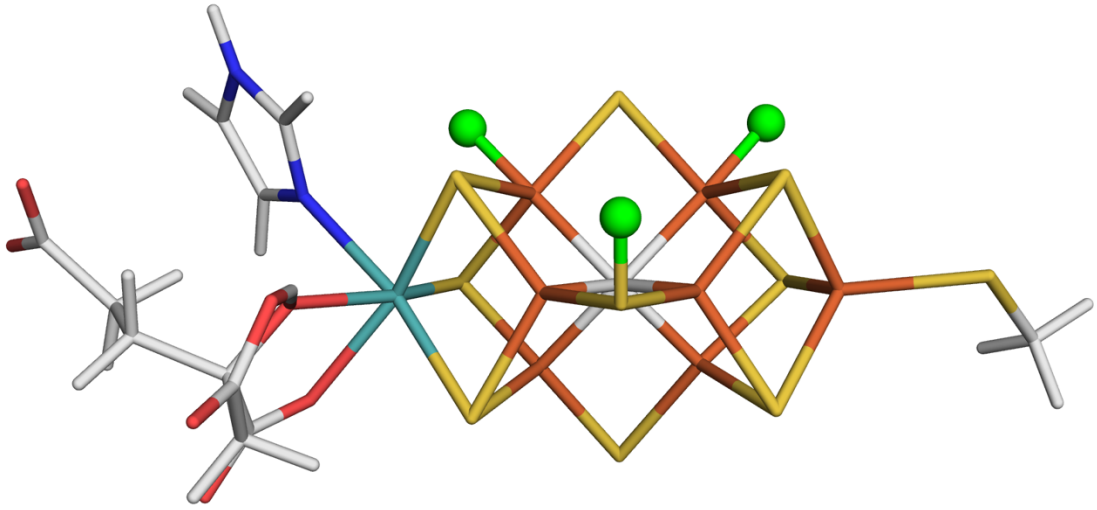


f

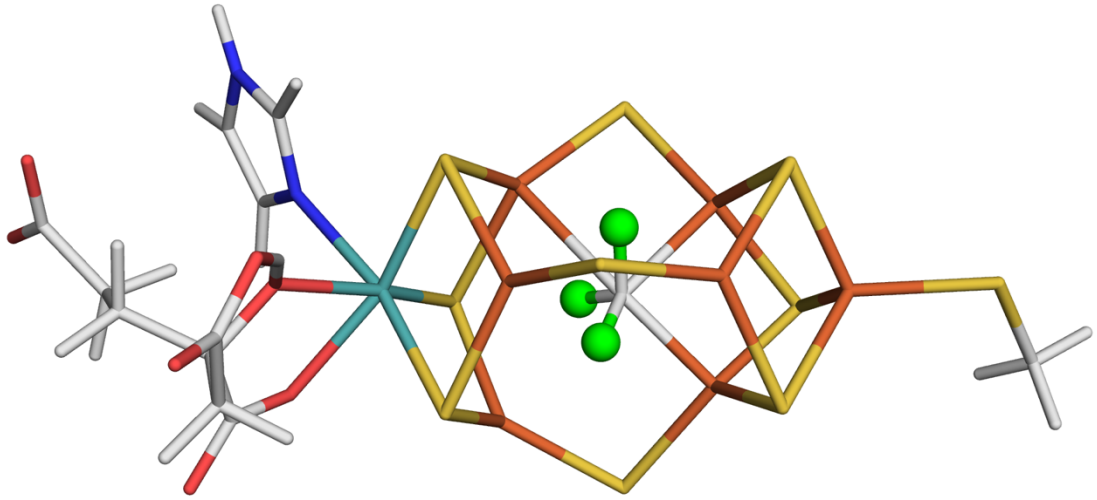
Figure 6. The best E_3 structures, protonated on (a) S2B(3), Fe2/6(5) and Fe5, (b) S2B(3), Fe4 and Fe5, (c) C(2367), C(3457) and C(2456) or on (d) S2B(3), C(3457) and C(2456). Structures (a) and (b) were obtained at the TPSS-D3/def2-SV(P) level of theory, whereas structures (c) and (d) were obtained at the B3LYP-D3/def2-SV(P) level of theory.



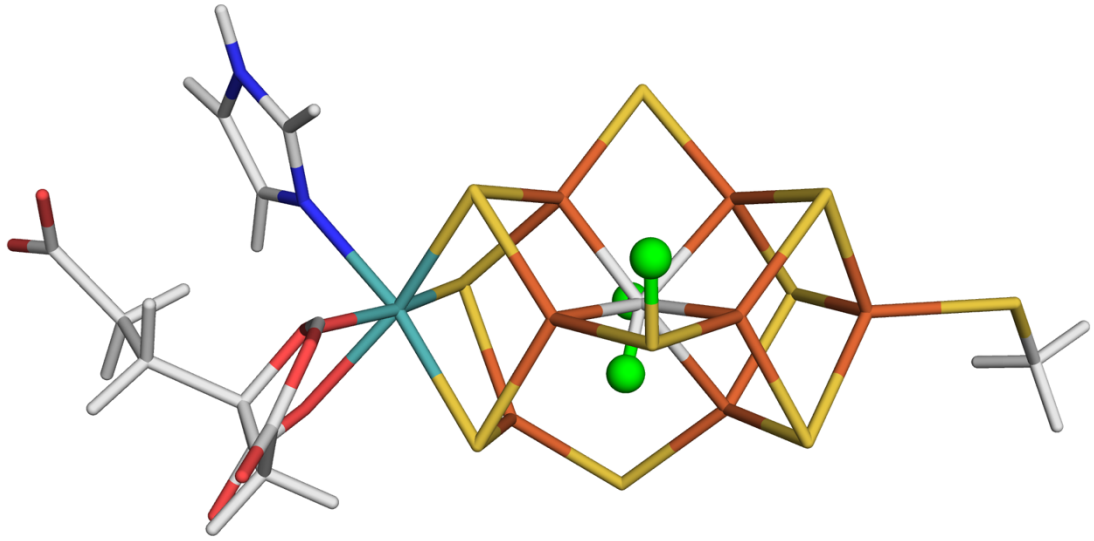
a



b

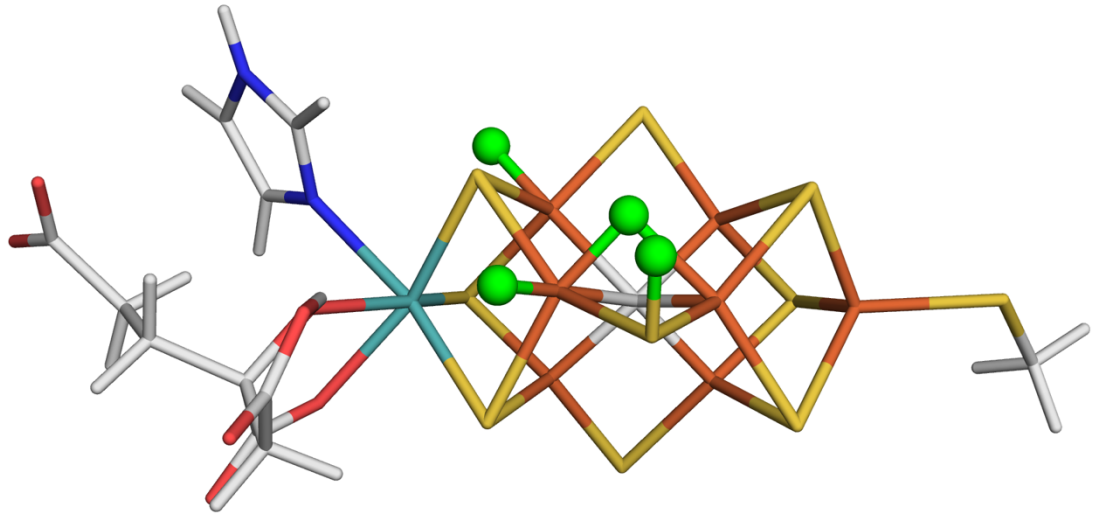


c

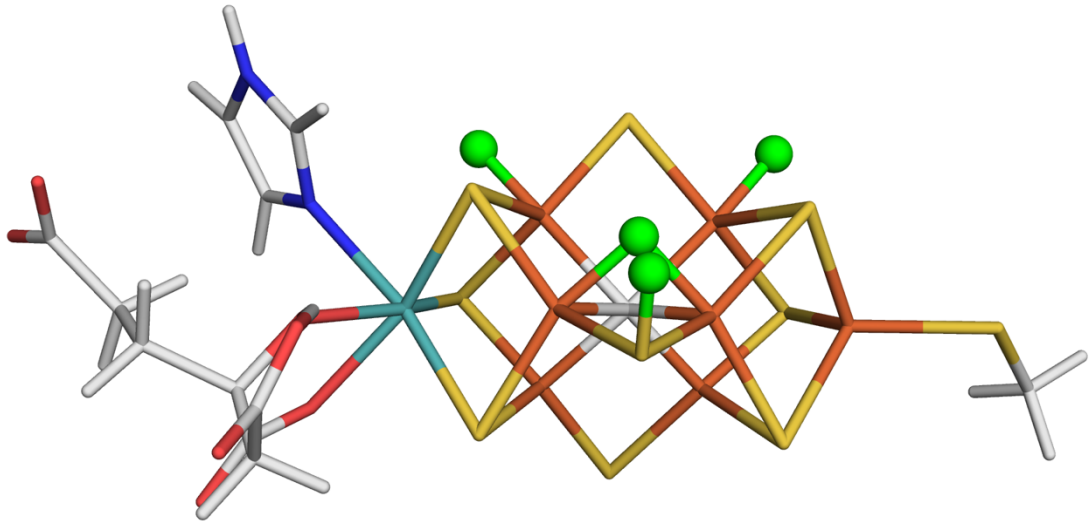


d

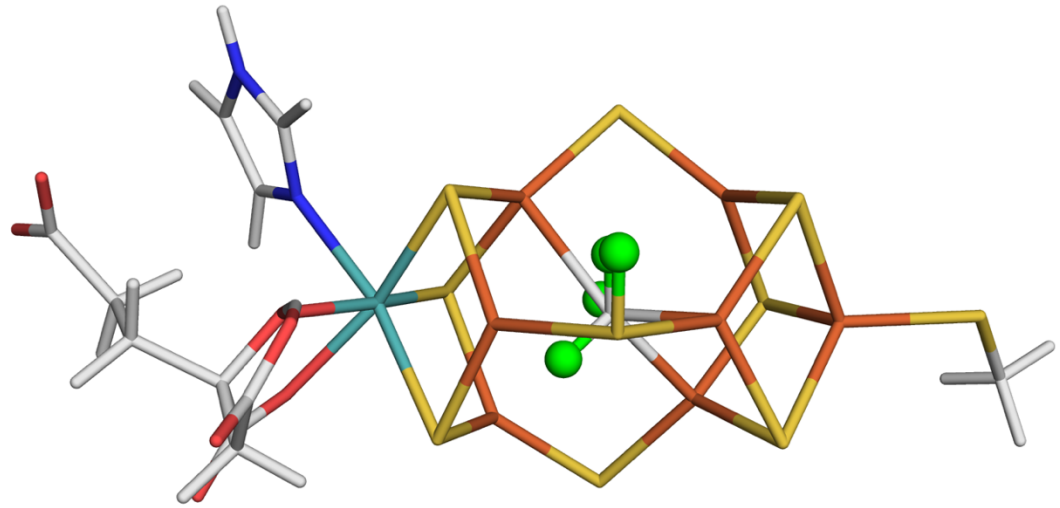
Figure 7. The best E_4 structures, all protonated on S2B and in addition on (a) Fe2/6(3), Fe5 and Fe6 or (b) Fe2/6(3), Fe4 and Fe5, (c) C(2367), C(3457) and C(2456) (d) S2A(Fe1), C(2367) and C(3457) or (e) S2B(5), Fe2/6(5), S5A(2) and Fe3/7(2). Structures (a), (b) and (e) were obtained at the TPSS-D3/def2-SV(P) level of theory, whereas structures (c) and (d) were obtained at the B3LYP-D3/def2-SV(P) level of theory.



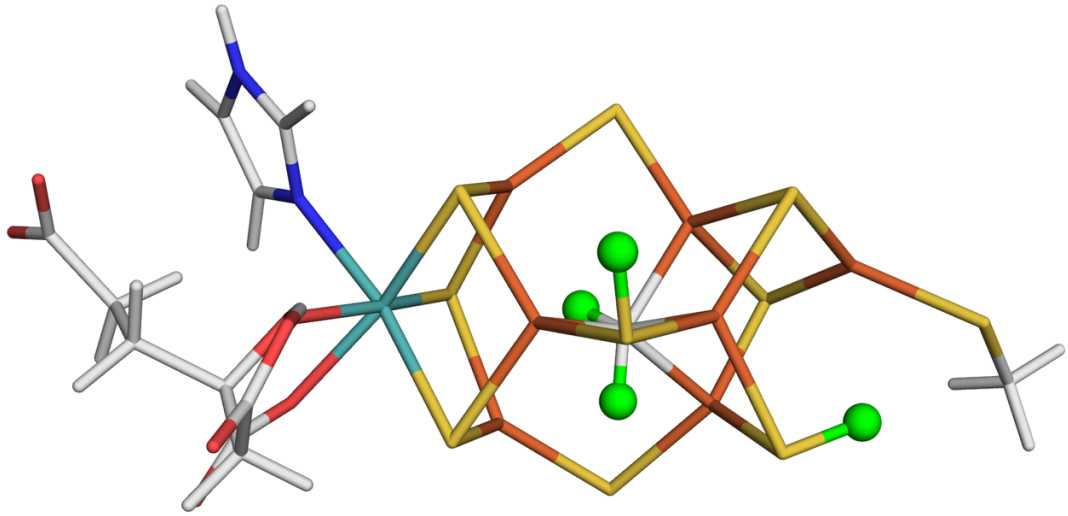
a



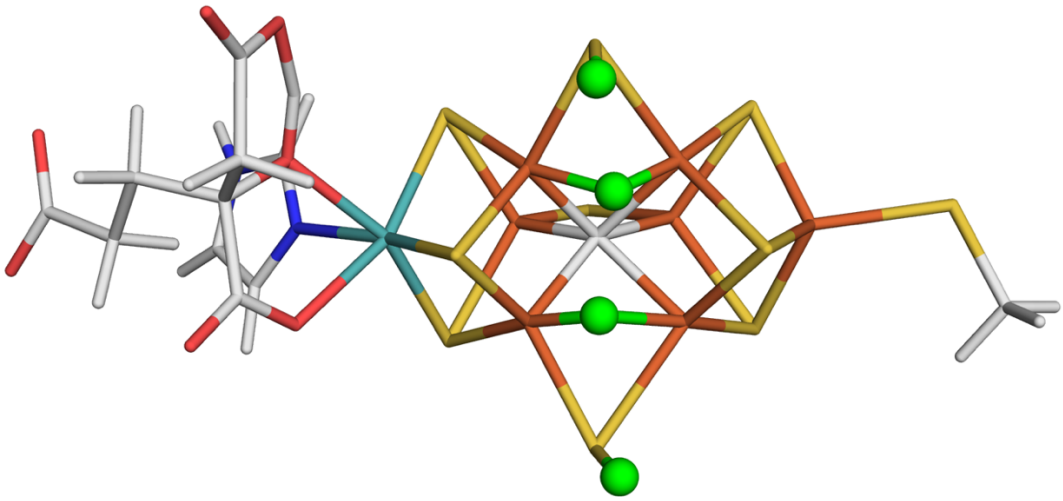
b



c



d



e

TOC graphics