Combined computational and crystallographic study of the oxidised states of [NiFe] hydrogenase

Söderhjelm, Pär; Ryde, Ulf

Published in:
Journal of molecular structure. Theochem

DOI:
10.1016/j.theochem.2006.06.008

2006

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

Combined computational and crystallographic study of the oxidised states of [NiFe] hydrogenase

Pär Söderhjelm & Ulf Ryde *

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

Correspondence to Ulf Ryde
E-mail: Ulf.Ryde@teokem.lu.se
Tel: +46 – 46 2224502
Fax: +46 – 46 2224543

2017-03-18

[NiFe] hydrogenases catalyse the reaction $H_2 \leftrightarrow 2H^+ + 2e^-$. Several states of the enzyme have been observed by spectroscopic methods. Among these, the two most oxidised states, called the unready Ni-A and Ni-SU states, have been especially intriguing, because they take a much longer time to activate than the corresponding ready Ni-B and Ni-SI states. It has recently been suggested that the unready states actually contain a (hydro)peroxide bridge between the Ni and Fe ions, in contrast to the hydroxide bridge in the ready states. In this paper, we use quantum refinement (crystallographic refinement, in which the molecular mechanics [MM] calculations, normally employed to supplement the crystallographic data, are replaced by more accurate quantum mechanics [QM] calculations), combined QM/MM calculations, and accurate energy estimates to study the nature of a recent oxidised crystal structure of [NiFe] hydrogenase from *Desulfovibrio fructosovorans*. We show that the structure contains a mixture of several states in the active site. The experimental data is best explained by structures with a hydroxide bridge but with two of the cysteine ligands (one bridging and one terminal) partly oxidised. When the terminal Cys-543 ligand is oxidised, the sulphur occupies an alternative position, observed in several crystal structures. The Glu-25 residue, that forms a hydrogen bond to this sulphur also changes position. A peroxide ligand may exist as a minor component in the crystal and the suggested structure is supported by the calculations. We suggest that oxidised states are slow-equilibrium mixtures of structures with a peroxide bound and structures with oxidised Cys residues, and that the former can be activated by replacement of the protonated peroxide with a H_2 or CO ligand, as has been observed in electrochemical experiments.

Keywords: [NiFe] hydrogenase, crystallographic refinement, QM/MM methods, solvation energy, peroxide bridge.

Introduction

Hydrogenases are enzymes that catalyse the reaction $H_2 \leftrightarrow 2\ H^+ + 2\ e^-$. They are found in a large number of bacteria and archaea as well as in some simple eukaryotes [1]. Both directions of reaction are encountered in nature: The forward reaction enables the organism to use H_2 as an energy source, whereas the reverse reaction provides a way to dispose excess electrons in the form of H_2. By studying the mechanism of these enzymes, we may hope to find new ways of effectively producing hydrogen to use as a clean fuel or generating cheaper fuel cell catalysts.

Hydrogenases can be classified into three groups by their metal components: nickel-iron [2], di-iron [3], and mono-iron [4] hydrogenases. Electron paramagnetic resonance (EPR) and infra-red spectroscopy (IR) experiments have shown that [NiFe] hydrogenase can exist in several different redox states [2]. Three of these, Ni-A, Ni-B, and Ni-C, are EPR active (i.e. paramagnetic), whereas Ni-SU, Ni-SI_a, Ni-SI_b, and Ni-R are EPR silent. Only Ni-R, Ni-C, and possibly Ni-SI_b are believed to participate in the catalytic cycle of the enzyme. The Ni-A and Ni-SU states are called unready, because they need long incubation with H_2 under anaerobic conditions to be activated, whereas the Ni-B state is called ready because it is rapidly activated. The assumed relationship between the states, as indicated by redox titrations, is shown in Figure 1 [2].

Crystal structures of [NiFe] hydrogenases from a number of different bacteria and in different redox states have been published [2,5]. The active site is highly conserved in all these enzymes so it is likely that the mechanism is the same for all of them. It contains a Ni ion bound to four cysteine ligands (cf. Figure 2). Two of these are terminal ligands, whereas the other two form bridges to the iron ion, which also coordinates one CO and two CN^- ligands. In oxidised structures (the Ni-A, Ni-B, and Ni-SU states), there is an additional bridging solvent species. Hydrogen atoms are usually not discernible in protein crystal structures, and therefore the nature of the bridging ligand is a subject of debate. Crystallography as well as ENDOR experiments [6] have shown that this bridging ligand is removed upon activation of the enzyme, probably when Ni-SI_a is formed. Recently, it has been suggested that it is a (hydro)peroxide ion in the Ni-A state and a hydroxide in Ni-B [7,8], which would explain the difference between the Ni-A and Ni-B states and their different mode of activation.

The oxidation and spin states of the metal ions in the [NiFe] hydrogenases have also been widely discussed. EPR, electron nuclear double resonance (ENDOR), and X-ray absorption spectroscopy (XAS) studies indicate that the iron ion is low-spin Fe^{2+} in all redox states, whereas nickel has been suggested to vary between the +III (EPR active) and +II oxidation states [2,9,10,11].

Several quantum mechanics (QM) studies of the mechanism of [NiFe] hydrogenases have been published [12,13,14,15,16,17,18,19,20,21]. Most of them considered only small model clusters, but Amara et al. [22] also included the surrounding protein by combined quantum and molecular mechanics (QM/MM) methods. The results differ in several aspects. For example,

Siegbahn et al. [13,14] as well as Hall et al. [18,19] have proposed that the H–H bond is cleaved at the Fe ion while Ni is in the +III oxidation state, but Amara et al. [22] suggested that H_2 is cleaved on Ni(II) and that a hydride bridge is present already before the H_2 cleavage, whereas all other groups considered that one of the hydrogen atoms from H_2 ends up as a hydride bridge. De Gioia et al. also suggested cleaving on Ni(II) [20,21] and Stein et al. had a similar view, but they suggested that a water molecule is bound to the Fe at the time of cleavage [16].

A major problem with the study of hydrogenase catalysis is that the substrates and products are invisible in standard protein crystal structures. Therefore, the binding site of H_2, hydride ions and protons remain uncertain. This applies also to the total number of protons in the active site (i.e. the total charge of the active site). Siegbahn et al. have argued that the active-site cluster should be neutral, because the protein is a low dielectric medium, but charged groups around it make it hard to decide where the neutral cluster ends [13]. They also suggested that the protonation state is of less importance and that different protonation states may be indistinguishable by QM methods. Other groups have tried to determine which protonation state agrees best with crystal structures, IR spectroscopy data, and EPR parameters, but with differing conclusions [16,18,21,22].

We have recently developed a method to combine QM methods with protein crystallography, quantum refinement [23]. In essence, we replace the molecular mechanics potential used to supplement the experimental data in standard protein structure refinement with more accurate QM calculations for a small, but interesting, part of the protein. We have shown that this method allows us to locally improve the structure [24] and that it can also be used to determine the protonation status of metal-bound solvent molecules in protein crystal structures [25,26]. In this paper, we use such an approach to re-refine an oxidised structure of [NiFe] hydrogenase from *Desulfovibrio fructosovorans*. We study the protonation of the active site and examine the recent suggestions that the bridging ligand is a (hydro)peroxide and that one or two of the Cys ligands are oxidised.

Methods

QM/MM calculations

QM/MM calculations were performed with the program ComQum [27,28]. It is an interface between the quantum mechanics (QM) software Turbomole 5.6 [29] and the molecular mechanics (MM) software AMBER 7 [30], using the Cornell force field [31] for the MM part. Like most other QM/MM methods, ComQum divides the protein and surrounding solvent into three subsystems. The central system 1 (the quantum system) is optimised by a QM method. System 2 consists of all amino acids and solvent molecules within a radius (6 Å in this study) of the quantum system. It is optimised with MM methods. Similarly, system 3 comprises the rest of the protein and a sphere of water molecules with a radius of 33 Å. System 3 is included in the calculations, but it is kept fixed at the crystal geometry.

In the QM calculations, system 1 is represented by a wavefunction, whereas systems 2 and 3 are modelled by an array of point charges, one for each atom (from the MM force field), which is included in the one-electron Hamiltonian. Therefore, the polarisation of the quantum system by the protein is included in a self-consistent manner. No cut-off is used for the electrostatic interactions (the system is non-periodic and finite). In the MM energy and force calculations, systems 1–3 are represented by the molecular mechanics force field, but without any electrostatic interactions (which are already treated by quantum mechanics).

In every step of the geometry optimisation of system 1, system 2 may be optimised with system 1 and 3 fixed. Thereby, we exploit the speed of a MM optimisation compared to that of the QM calculations. In these calculations, systems 2 and 3 are represented by standard MM parameters (including electrostatics), whereas system 1 is represented by MM parameters, but with charges from the QM calculation [28].

Special action is taken when there is a bond between systems 1 and 2 (a junction). The QM system is truncated by hydrogen atoms, the positions of which are linearly related to the corresponding heavy (typically carbon) atoms in the full system. The positions of the H junction atoms are optimised in the QM steps and the positions of the corresponding C junction atoms are updated from these positions. No constraints are imposed on any of the atoms in the boundary region, but the charges of the atoms bound to the C junction atoms are set to zero by spreading them out on the other atoms in the amino acid (a change of less than 0.05 e).

The total QM/MM energy is calculated as

$$E_{QM/MM} = E_{QM1} + E_{MM123} - E_{MM1} \qquad (1)$$

Here, E_{QM1} is the QM energy of system 1 with H junction atoms, including all the electrostatic interactions. Similarly, E_{MM1} is the classical energy of system 1, still with H junction atoms, but without any electrostatic interactions. Finally, E_{MM123} is the classical energy of systems 1–3 with C junction atoms and no electrostatics. The philosophy behind this is that the total energy should involve as much quantum chemistry as possible and that terms from the junction

truncations shall cancel out. This approach is similar to the one used in the ONIOM method [32]. The forces are the negative gradient of this energy, taking into account the relation between the H and C junction atoms using the chain rule [33].

Quantum refinement

The quantum refinement program, ComQum-X [23], can be seen as a modification of ComQum so that the structures are restrained by the crystallographic raw data. In ComQum-X, the MM program is replaced by the crystallographic refinement program CNS (Crystallography & NMR system) [34]. In crystallographic refinement, the model (coordinates, *B* factors, occupancies, etc.) is improved by optimising the fit of the observed and calculated structure-factor amplitudes, typically estimated by the residual disagreement, the *R* factor. Because of the limited resolution normally obtained with X-ray diffraction of biomolecules, a MM force field is used to supplement the data for the whole protein [35]. This force field ensures that the bond lengths and angles make chemical sense. In ComQum-X, this force field is replaced by more accurate QM calculations for system 1, in a manner completely analogous with the use of QM in QM/MM calculations. The junctions are handled in the same way as in ComQum.

Thus, the ComQum-X refinement takes the form of a minimisation using an energy function of the form

$$E_{ComQum\text{-}X} = E_{QM1} + E_{MM123} - E_{MM1} + w_A E_{Xray} \qquad (2)$$

Here, E_{Xray} is a penalty function, describing how well the model agrees with the experimental X-ray data. We have used the default maximum likelihood refinement target using amplitudes (MLF) in CNS [36]. w_A is a weight factor, which is necessary because E_{Xray} is in arbitrary units, whereas the other terms are in energy units. w_A was determined by CNS so that the root mean squares of the E_{MM123} gradients and E_{Xray} gradients are equal ($w_A = 0.43$ in all calculations).

In ComQum-X, all atoms are restrained towards the crystallographic data by E_{Xray}. Therefore, there is no need to fix the positions of any of the atoms (i.e. no system 3). In spite of the use of the R_{free} factor as the main quality factor, crystal structures are often heavily optimised with regards to the *R* factor. Therefore, the *R* factor often starts to increase if the structure is minimised without any QM data. In ComQum-X, this problem is avoided by running only one step of minimisation and then accept the structure only if the *R* factor is reduced [23]. As discussed below, this is never observed in this investigation. Therefore, the surrounding protein was effectively fixed. No hydrogen atoms were included in the MM region of the ComQum-X calculations (hydrogen atoms are not discernible in the crystal structure). Therefore, polarisation of the quantum system by the surrounding protein is not included in ComQum-X. However, in order to test the influence of polarisation, some calculations were carried out with hydrogen atoms in the MM region and point charges taken from the Amber force field in the QM calculations [37].

Finally, it should be noted that the MM force field used in CNS (protein_rep.param, water.param, and ion.param) is based on a statistical survey of crystal structures [35], rather than the energy-based force field in Amber and in the QM calculations. Therefore, the CNS energy has to be weighted by a factor of 1/3 to be comparable with the QM and Amber MM energies [23,35].

The protein

The calculations in this investigation are based on the $P2_12_12_1$ crystal structure of the Ser499Ala [NiFe] hydrogenase mutant from *D. fructosovorans* [38] (1.81 Å resolution, R = 0.149, and R_{free} = 0.187). Coordinates, occupancies, B factors and structure factors were generously provided by Profs. A. Volbeda and J. C. Fontecilla-Camps. We used the same test-set selection for the calculations of the R_{free} factors as in the original investigation. The hydroxide group of Ser-499 forms a hydrogen bond to one of the CN^- ligands of Fe, but in the mutant, a water molecule replaces this group, leading to unchanged enzymatic properties, vibrational frequencies, and structure, but much better diffracting crystals [38]. Over 80% of the purified enzyme was in the Ni-A state, as determined by IR spectroscopy [38] and this was probably also the dominant state in the crystals (it was originally suggested that the crystals contained mainly the Ni-SU state, because they were grown in the presence of H_2 [38], but it has recently been shown that the Ni-A state does not react with H_2 [39]). Still, during data collection, a partial reduction induced by the X-rays may have occurred, as will be discussed below.

In the quantum refinement calculations, alternative conformations were treated with standard procedures in CNS, except for the sulphur atom of Cys-543, whose less occupied position was deleted, because the atom belonged to the quantum region (but the occupancy of the remaining sulphur atom was still 0.8). In the QM/MM calculations, the alternative conformations with the lowest occupancy were deleted.

In most of the QM/MM calculations, all the surrounding protein was kept fixed, which corresponds to not having any system 2. The reason for this is that optimisations of system 2 may end up in different local minima for different protonation states of the quantum region, which makes the energies hard to compare. In the few optimisations without a fixed protein, system 2 consisted of all amino acids within 6 Å from any atom in the quantum region. All residues located more than 27 Å from the Ni atom were deleted and solvation water molecules were added to the protein, forming a sphere with radius 33 Å. As one of the three iron-sulphur clusters (the distal one) present in the enzyme is outside this cut-off radius, only four metal sites remained in the model, namely the proximal [4Fe-4S] cluster, the [3Fe-4S] cluster, a six-coordinated Mg ion at the C-terminus of the large subunit, and the [NiFe] site. For these sites, we used Merz–Kollman electrostatic potential (ESP) charges, taken from QM calculations for truncated models of each site. Other MM parameters were not needed for the metal sites, because they were not in system 2 and thus either had fixed positions or belonged to the

quantum region. Charges for the protein atoms and for water molecules were taken from the Amber force field.

Hydrogen atoms were added using the AMBER package. The protonation status of histidine residues was determined by inspection of the local surroundings and hydrogen-bond structure. This gave protonation of $N^{\delta 1}$ for residues S92 (an initial S indicates the small subunit, residue numbers without S refers to the the large subunit), S243, 481, and 549; protonation of $N^{\varepsilon 2}$ for S13, 27, 66, 79, 113, 118, 121, 123, 210, 228, 349, 367, and 419; and double protonation for 115, 204, 305, and 538. In some structures, His-79 (which forms a hydrogen bond to the Ni ligand Cys-546) was protonated on $N^{\delta 1}$ instead, as will be discussed below. All lysine, arginine, aspartate, and glutamate residues were charged, except Glu-25, which shares a hydrogen atom with the Ni ligand Cys-543. Before running the QM/MM calculations, the positions of hydrogen atoms and solvation water molecules were optimised by a simulated-annealing molecular dynamics calculation followed by a MM minimisation.

Three different sizes of quantum regions (system 1) were used. In the smallest quantum region (28–32 atoms), used in most quantum refinement and vacuum optimisations, only the first-sphere ligands of the metal ions were included (Figure 2). The cysteine thiol groups were modelled as CH_3S^-, with one hydrogen junction atom each, corresponding to the cysteine α-carbon in the MM region. It has been proposed by Stein et al. that a $C_2H_5S^-$ model for cysteines should be used in order to get reliable spin densities. When this was tested, a small effect was seen in the vacuum geometries (<0.04 Å in Ni–S distances, <0.08 Å in Fe–S distances, and <0.02 Å in other distances). However, this mainly reflects the tendency of this model to dissociate in vacuum (see below).

In the medium-sized quantum region (46–50 atoms), shown in Figure 3, the functional groups of two second-sphere residues (Glu-25 and His-79) were also included. They may have important influence on the structure and also play a role in the catalytic mechanism of the enzyme, because the crystal structures show that they share hydrogen atoms with Cys-543 and Cys-546, respectively. This system was used in all QM/MM calculations.

Finally, in the largest quantum region (113–117 atoms), shown in Figure 4, all polar interactions between the first-sphere ligands and the surrounding protein are included.

Quantum mechanical methods

If not otherwise stated, all QM calculations were performed at the density functional theory level with the Becke 1988–Perdew 1986 (BP86) functional [40,41] together with the 6-31G* basis set for H, C, N, O and S and the DZP basis set [42] for Fe and Ni. These calculations were sped up (by a factor of ~5) by expansion of the Coulomb interactions in auxiliary basis sets, the resolution-of-identity approximation [43,44]. However, to obtain more accurate energies, a single-point B3LYP [45] calculation was performed for the optimised structure. In these calculations, the 6-311+G(2d,2p) basis set was used for the light atoms, whereas the DZP basis set for the metals were enhanced with *s, p, d*, and two *f*-type functions with the following

exponents: 0.013772, 0.041843, 0.1244, 2.5, and 0.8 for Fe and 0.145763, 0.146588, 0.044447, 0.1458, 6.74, and 1.04 for Ni. We estimate that the error in relative B3LYP energies is ~25 kJ/mole [46,47].

Because experiments indicate that the active site is low-spin in all states, closed-shell singlets ($S = 0$) were studied for the EPR silent states and doublets ($S = 1/2$) were studied for Ni-A and Ni-B, if not otherwise stated. In some calculations, the triplet state ($S = 1$) of Ni-SU were also tried, because this state has been suggested by both experimental and QM investigations [18,19,48,49].

Solvation energy calculations

Energy calculations in proteins are problematic, because long-range interactions, screening of charges, solvation, and dynamic effects must be treated in a proper way [37,50]. We have compared two methods for estimating the influence of the surrounding protein on the quantum energy. The standard method in ComQum is to include the point charges of the MM region in the one-electron part of the Hamiltonian for the quantum region. Moreover, the difference in MM energy of system 2 is added ($E_{MM123} - E_{MM1}$). If system 2 and 3 are kept fixed (as in most of our calculations), this energy difference reduces to bond and angle energy for the junction atoms, as well as Van der Waals interaction between the quantum and MM region. The point charges have much larger influence on the relative energies of different structures than the MM energy, and their effect sometimes seems to be exaggerated.

By using a solvation model, it is possible to screen the surrounding point charges in a proper way, in addition to include the difference in the free energy of solvation of the various systems. We used the program Mead 2.2 [51], which solves the Poisson–Boltzmann equation for the quantum region solvated in the protein and water (both surrounding the protein and in empty spaces inside the protein). The quantum region was in these calculations modelled as an array of point charges, taken as Merz–Kollman electrostatic potential charges obtained from the wavefunction. The dielectric constant, ε_r, was 1 in the quantum region, 80 in the solvent, whereas we varied ε_r in the protein between 1 and 16. Since no significant difference in relative energies were observed, we decided to use $\varepsilon_r = 4$ in the protein. We also tried removing the crystal waters, but this only affected the absolute energies. Therefore, we decided to treat the crystal waters (but not the solvation water molecules) explicitly.

The Mead calculations were performed with 351^3 grid points and a spacing of 0.25 Å, centring the grid on the catalytic Ni ion. The reported energies are the average of seven calculations in which the grid origin was moved 0.1 Å in positive and negative direction along each Cartesian axis. The maximum difference among the seven calculations is 2 kJ/mol. Parse radii [52] were used for all atoms. These values have previously been shown to provide reasonable agreement with experimental results [53,54,55].

Nomenclature

In this paper, we have studied a large number of models of the [NiFe] active site. In order to effectively describe these models, we have adopted the following notation. Sulphur atoms of the four cysteines are referred to by their order in the sequence (to avoid using the specific numbering of *D. fructosovorans*). Thus, S1, S2, S3, and S4 are the S^{γ} atoms of Cys-72, 75, 543, and 546, respectively. The state of the active site is given by a number of symbols in brackets, followed by the total charge of the quantum region. The first symbol indicates the oxidation number of nickel (typically II or III). Then follows positions of added hydrogen atoms (H) or, in the case of oxidised cysteines, oxygen atoms (O). The symbol H1 indicates that the sulphur of Cys-72 is protonated whereas SO2 indicates that Cys-75 is oxidised. Finally, the last part of the bracket shows the chemical formula of the bridging ligand (HOH, OH, O, OO, or HOO, omitting the charge; the symbol "/" indicates no bridging ligand). In absence of added protons, it is assumed that all four cysteines of the active site are deprotonated, whereas Glu-25 is protonated (with a proton that forms a hydrogen bond to S3) and His-79 is singly protonated on the $N^{\varepsilon 2}$ atom that is directed towards S4, except for calculations with S4 protonated, in which the protonation of His-79 was simultaneously changed to the $N^{\delta 1}$ atom. The oxidation state of Fe is always +II.

Results and discussion

Calibration of the method on a model compound

The quantum refinement method strongly depends on the accuracy of the QM method employed. Previous investigations have shown an excellent performance for the distances around iron in haem complexes (errors of less than 0.03 Å for Fe–N and Fe–S distances) [24] and of Cu and Mn complexes (errors of up to 0.07 Å) [56,57]. However, we have not studied Ni complexes before. Therefore, we started this investigation with a test of the performance of various methods and basis sets for a model compound, which has been characterised by X-ray diffraction [58]. The complex is shown in Figure 5. The optimised geometry for various basis sets and density functional methods are listed in Table 1, together with the experimental values, which have an accuracy of 0.01 Å or better [58]. As can be seen, the optimised parameters agree very well with experiment. For the selected method (BP86/DZP,6-31G*), the Ni–S bond lengths are overestimated by 0.01–0.04 Å and the Ni–N bond lengths by 0.04–0.05 Å. These distances are not very basis set-dependent, but the errors are somewhat larger with the B3LYP functional than with the BP86 functional. The same conclusion was recently obtained with a different Ni complex [48]. The Ni–Ni distance is more sensitive to the basis set, but the error is only 0.05 Å with the selected method.

Next, we tested the effect of protonation of the thiolate S_4 group. Such a protonation has been studied experimentally by extended X-ray absorption fine structure (EXAFS) measurements, indicating that the average Ni–S distance increased by 0.02 Å [59]. Our results

give a similar, but more detailed picture: All Ni–S bond lengths change by less than 0.02 Å. The bonds to S_4 and the nearby S_1 are elongated, whereas the others are shortened. The changes are essentially independent of the basis set. Such small changes are unexpected: Protonation of Cys ligand of the structural zinc site in alcohol dehydrogenase ($ZnCys_4$) gave an increase in the Zn–S distance of 0.45 Å or more [60]. Evidently, the changes are much smaller in the binuclear sites. This is quite unfortunate, because it will make it much harder to determine protonation states using quantum refinement.

Finally, we also investigated the effect of alkylating the same atom (S_4) with CH_2Phe. The structure of the alkylated complex is also known from a crystal structure and shows changes in the bond distances of less than 0.04 Å [59]. The corresponding QM structure reproduce all distances in the X-ray structure within 0.04 Å, but the differences compared to the unalkylated structure are slightly smaller (up to 0.03 Å).

Thus, we can conclude that the theoretical method performs as well on Ni-containing systems as on other metal complexes [24,57,61,62]. Therefore, we expect that the optimised metal–ligand distances will be within 0.05 Å of the crystallographic ones and that we can accurately model the effects of ligand protonations.

Quantum refinements

The primary goal of the quantum refinement calculations was to establish the protonation state of the active site in the examined crystal structure [38], especially whether the bridging ligand is H_2O, OH^-, or O^{2-}. As described earlier, the protein in the crystal is expected to be mainly in the Ni-A state. Most experimental and theoretical studies agree that the Ni oxidation number varies between +II (e.g. for Ni-SU) and +III (e.g. for Ni-A). Furthermore, it is often assumed that only the terminal cysteines (S1 and S3) may be protonated, because they should have higher pK_a values than the bridging cysteines, which interact with two positively charged metals. Therefore, we made a more systematic investigation with those structures. However, several non-conventional structures were also tested, including protonated bridging cysteines, no bridging ligand, Ni(I) and Ni(IV), and high-spin Ni(II).

As in our earlier investigations, we re-refine the protein structure with all possible protonation states in the QM system and then try to determine which structure fits the experimental data best using a number of criteria [23,25,26]: The R and R_{free} factors give a direct estimate as to how well the structure fits the structure factors. R_{free} has the advantage of being less biased to the current model and it is less sensitive to overfitting. On the other hand, it is based on fewer reflections and can therefore be more noisy. Second, we compare how much the quantum-refined structures differ from the corresponding vacuum structure (i.e. how well the vacuum structure fits the electron density). As we saw in the previous section, QM vacuum calculations are expected to reproduce metal-ligand bond lengths within 0.05 Å. Therefore, we look at the average and maximum difference in the bond lengths between the vacuum and quantum-refined structures. We also calculate the strain energy (ΔE_{QM1}), which is the vacuum

energy difference between the structures optimised in vacuum and in the protein. It gives a global estimate of all changes in geometry and it also weights them in energy terms. On the other hand, it is sensitive to polar interactions between the QM system and the surroundings and therefore to the total charge of the complex [63].

Unfortunately, it turned out that for several structures, especially those with a high negative charge, some of the ligands dissociated in the vacuum optimisations. Typically, the Ni–S4 bond was elongated (Ni–S = 2.8–3.0 Å) or fully broken (>3Å) during the optimisation, because the bridging oxygen and the other three Cys ligands formed an approximate plane with S4 as an axial ligand, which is a much weaker ligand for the Jahn–Teller active low-spin Ni(II) ion. In order to calculate strain energies for the corresponding quantum-refined structures, we therefore performed vacuum calculations with the Ni–S4 distance constrained to the value obtained in the quantum refinement.

The results of the ComQum-X calculations are summarised in Table 2. It is clear that calculations without any bridging ligand (giving a putative Ni-SI_a state) give appreciably larger values for both R and R_{free} all the other structures. Interestingly, only half of the increase in R is caused by the actual removal of the electron density of the oxygen atom, whereas the rest comes from other changes in the geometry, caused by the removal of the bridge. This clearly shows that there is indeed a bridge present in the crystal and that the method works properly.

Moreover, it can be seen that the complexes with a bridging OH^- group give consistently lower R factors than those with either H_2O or O^{2-} as the bridging ligand (ΔR = 0.00008–0.00012 compared to 0.00012–0.00019 and 0.00012–0.00018, respectively). This is mostly confirmed by the ΔE_{QM1} values, which are 74–101 kJ/mole for the OH^- complexes (except for three complexes that dissociate in vacuum), but 82–158 kJ/mole for the other complexes. The average difference in the bond lengths show a similar trend. However, the R_{free} factors show a somewhat differing view, giving the smallest values for some of the O^{2-} complexes. On the other hand, the R_{free} factors are only slightly correlated to the other quality criteria, giving almost random results for the various calculations. Considering the small differences in R_{free} and the rather small number of reflections used for the calculation of these factors, we believe that the other quality criteria give the more reliable picture.

Furthermore, the picture is complicated by the fact that various oxidation, protonation, and spin states give different results. Thus, the low-spin Ni(II) states give consistently higher R factors and strain energies than the Ni(III) states. However, the high-spin Ni(II) complexes are clearly best among the H_2O complexes, whereas they seem to be slightly worse than the Ni(III) complexes if the bridging ligand is OH^- (for the O^{2-}-bridged structures, no reliable triplet geometries in the vacuum calculations could be obtained, because these complexes do not contain any weakened axial ligand so any of the Ni–S bonds may break). Finally, double protonation give consistently higher strain energies than the complexes with one or no extra proton and this is often also supported by the R factors. On the other hand, it seems impossible to say anything further about the actual site of protonation, based on the quantum refinement

calculations.

Consequently, the three Ni(III) states with a bridging OH^- ion and with no or one extra proton seem to fit the crystallographic data best, with *R* factors, strain energies and average and maximum bond length deviations among the best observed for all complexes. However, this conclusion is quite weak because the differences between the various complexes are small and various criteria sometimes give different results.

Finally, we make a note about the spin state of the Ni(II) complexes. Most previous calculations have assumed that Ni(II) is in the singlet state [13,15,21]. However, recent experimental and theoretical results have suggested that it may actually rather be in the triplet state [18,19,48,49]. Therefore, we performed also a set of calculations with Ni(II) in the triplet state. With the BP86 functional, the two states have a similar energy, with the triplet most stable in the protein by 1–20 kJ/mole, except for the states with a O^{2-} bridge, for which the singlet is more stable by 0–14 kJ/mole. In vacuum, the O^{2-} states dissociate, the OH^-states are even more stable in the triplet state (by 13–38 kJ/mole), whereas the H_2O states are most stable in the singlet state (by 43–55 kJ/mole). Unfortunately, the normally more accurate [46,47] B3LYP functional gives quite different results, increasing the relative stability of the singlet state by ~60 kJ/mole, compared to the BP86 results. Therefore, the most stable spin state of Ni(II) in the hydrogenases cannot yet be settled, but it is clear that the protein does not always stabilise the triplet state, as was recently suggested [19].

Geometries

Calculated bond distances for vacuum and quantum refinement optimisations are collected in Table 3 together with the crystal data. It can be seen that most of the distances in the quantum refinement and vacuum optimisations agree quite well with those in the original crystal structure. The most evident difference is the long Ni–S_{bridge} distances (2.52 Å) in the crystal structure. In all vacuum calculations, at least one of these distances is significantly shorter, around 2.35 Å, and in most structures, both distances are shorter than 2.4 Å. The Fe–O distance is also problematic: In the vacuum calculations, it is typically 2.05 Å when the bridge is hydroxide and 1.90 when it is oxo, but in the crystal structure it is 2.22 Å. In the H_2O vacuum structures, it is even longer, 2.51–3.08 Å. Similarly, one of the Fe–CN distances is 2.09 Å in the crystal structure but around 1.90 Å in the calculations, which is equal to the other Fe–CN distance.

In the quantum-refined structures, these distances respond in different ways: The Fe–O and Fe–CN distances become almost equal to the vacuum-optimised distances, which indicates that the QM potential is steep and the crystallographic electron-density landscape is flat (owing to the limited resolution of X-ray diffraction). Therefore, we can conclude that the quantum-refined Fe–O and Fe–CN distances are more reliable than those in the original crystal structure.

On the other hand, the Ni–S_{bridge} distances change in the quantum refinement calculations to values close to those in the original crystal structure. This shows that they are well-determined

in the crystal geometry, although they are not fully consistent with the QM calculations (the *R* factor increases by 0.0004, i.e. more than in any other structure, if a Ni–S distance is shortened to the vacuum value). It is notable that a recent EXAFS study gave two Ni–S distances of 2.17 Å and two Ni–S distances of 2.35 Å [64], i.e. in good agreement with the vacuum QM calculations.

Furthermore, it can be seen from Table 2 that the maximum difference in the metal–ligand distances between the vacuum and quantum-refined structures is larger than 0.13 Å in all structures. This is more than twice as large as the expected uncertainty in the QM method (Table 1), and it is also much larger than in any previous application of ComQum-X [23,24,25,26,37,56], indicating a serious problem in this application. This is supported by the fact that the R_{free} factor is not improved significantly, in contrary to what is usually seen, and that the various quality criteria give differing results.

A conceivable reason for these problems is that we have ignored electrostatic interactions in the optimisation. We therefore performed a set of calculations in which the surrounding protein is modelled by a set of point charges, taken from the Amber force field [31]. The resulting metal–ligand distances are also included in Table 3 (in bold face). It can be seen that the effect of the point charges is quite small. They lengthen the Ni–S4 bond (so much that it becomes even longer than in the crystal structure) owing to the strong attraction from His-79, which in these calculations was modelled as doubly protonated (positively charged). However, the *R* factor was always larger in these calculations than in the corresponding quantum refinement without point charges.

Another possibility is that the quantum region is not large enough. Therefore, we also performed some vacuum and quantum refinements with the largest quantum region (Figure 4). The results are collected in Tables 4 (quality criteria) and 5 (geometries). It can be seen that the geometries are quite similar to those obtained with the small quantum region. In fact, both *R* factors increased significantly compared to the calculations with the smaller system, suggesting that the QM calculations pulled the system away from the experimental geometry. Thus, the size of the QM system cannot explain the large geometric discrepancies.

Peroxide structures

A straight-forward explanation for the unclear quantum refinement results may be that we have still not found the correct quantum system. Recently, it has been suggested that the bridging ligand actually is not a solvent-derived hydroxide, but instead a (hydro)peroxide ion [7,8]. We therefore included nine ComQum-X calculations with such a bridging ligand. From Table 2, it can be seen that those complexes had quite small *R* factors (ΔR = 0.00010–0.00017), only slightly higher than those obtained for the complexes with a OH^- bridge (and one actually better with ΔR = 0.00007) and also the lowest R_{free} factors obtained (ΔR_{free} = –0.00007 to –0.00015, with one exception). However, these low factors were only obtained after an optimisation of the occupancy of the bridging ligand, giving an occupancy of only 0.17–0.32 (it

was 0.6–0.7 in the recent peroxide crystal structures [7,8]). If the occupancy of the OH^- bridge is optimised for the [III OH]$^{2-}$ state (one of the best previous structures), the occupancy becomes 0.98 and the *R* factors do not change.

The strain energies of the peroxide complexes are also as low as or even lower than the complexes with a OH^- bridge, 41–80 kJ/mole (compared to ~75 kJ/mole for the best OH^- models). Unfortunately, neither of the peroxide models are stable in vacuum and even with one constrained distance, five of the complexes still dissociate. Therefore, average and maximum deviations could only be calculated for four of the peroxide complexes, which gave values that are similar to those of the best OH^- models.

This indicates that the peroxide structures are strong candidates for the structure observed in the crystal. Therefore, we tried to discriminate between the structures with an OH^- or peroxide bridge by detailed comparison between one structure of each type, viz. the [II OH]$^{3-}$ and [III H3 OO]$^{2-}$ structures (the selection of these two structures is somewhat arbitrary, but they are clearly among the best of each type and the difference between structures differing in the protonation status is quite small). According to the residue (real-space) *R* factors, as calculated by the CNS program [65], the OH^- structure fitted the experimental data best. The main difference comes from the fully occupied bridging O atom, for which the residue *R* factor was 0.260 in the OH^- structure, compared to 0.378 for the O_2^{2-} structure (using the omit map). Moreover, the $F_o - F_c$ difference map of the peroxide structure displays a large negative density for the second O atom (white volume near the OO and OH atoms in Figure 6). Thus, both the residue *R* factor and the electron-density maps strongly indicate that a hydroxide bridge fits the experimental electron density better than a peroxide structure. The maps in Figure 6 also show that there is no significant residual density around the Fe–O and Fe–CN bonds that are modified in the quantum-refined structures. This strengthens our conclusion that these distances are better described in the quantum-refined structures than in the original crystal structure.

Thus, we can conclude that a (partly occupied) peroxide model is consistent with the crystallographic and QM data, but the best OH^- models fit the experimental data slightly better. Moreover, the peroxide structures do not resolve the discrepancies in the metal–ligand distances: One of the two bridging Ni–S_{bridge} distances were always shorter than in the crystal structure, as can be seen in Table 3. In particular, the maximum difference between the quantum-refined and vacuum structures were 0.16–0.19 Å, i.e. at least three times larger than expected.

Structural disorder

There remain two possible reasons for the failure of the quantum refinement for this particular structure: either we have not used the correct structure for the quantum region, or there is disorder in the QM system. Although we cannot guarantee that we have tested all conceivable structures, it is more likely that the crystal is disordered.

For example, it is possible that in some of the proteins, one of the Ni–S_{bridge} bonds is partially

broken (as is often seen in optimisations without the surrounding protein). If both Ni–S_{bridge} bonds may be elongated, but in different molecules, this could give the long average bond lengths observed in the crystal. This was tested by keeping one of the distances fixed at various values while letting the rest of the quantum system relax in vacuum. As can be seen in Figure 7 it costs only ~7 kJ/mole to elongate the Ni–S bond to the distance observed in the crystal structure and that this leads to a small shortening (~0.05 Å) in the other Ni–S_{bridge} bond length. This indicates that the suggestion is possible.

Another reason for the disorder is that the structure changes during data collection. The high-intensity X-ray beam produced at a synchrotron can ionise atoms and thus generate radicals and photoelectrons that may modify the structure that is studied, even if diffraction data are collected at cryogenic temperatures [66]. X-ray radiation damage may manifest itself through the reduction (and thereby cleavage) of disulphide bonds and through the decarboxylation of glutamate and aspartate side-chains [67]. The extent of such damage depends on the X-ray dose to which the studied crystal has been subjected. When redox-active metal sites are present, these may be reduced as well, making it difficult to obtain structures of oxidized states [56,66].

The studied structure shows several signs of radiation damage: A disulphide bond on the surface is cleaved and several Asp and Glu side chains are partially decarboxylated [38]. However, these changes are restricted to or close to the protein surface, where the largest effects of radiation damage is expected. The active [NiFe] site lies close to the centre of the enzyme, almost 30 Å below the surface. Therefore, it should be more protected against radiation damage.

Still, there is clear evidence of disorder of the active site in this crystal: S3 shows two different conformations, one coordinating to Ni with an occupancy of 0.8 and a second conformation ~3.0 Å from Ni (cf. Figure 6, S3 alt.). Likewise, significant features in the difference maps are seen around S2, S4, and Glu-25 in Figure 6. These will be discussed more below, but it is clear that the active site is not in a single geometry. This is most likely the cause of the problems encountered in the quantum refinement. Similar disorder is observed in essentially all reported structures of oxidised [NiFe] hydrogenases [7,8,38], so it is unlikely that better results would be obtained with other structures.

QM/MM calculations

Apparently, it is not possible to determine the protonation state of the hydrogenase active site from the present crystal structure using quantum refinement calculations. We therefore decided to try another approach, viz. to discriminate between the various structures using their relative energies. The previous section showed that the structures obtained with quantum refinement are probably not reliable, owing to disorder in the active site. Therefore, we used structures obtained by standard QM/MM methods instead. We also used a larger quantum system, viz. the medium-sized system in Figure 4, which includes also models of Glu-25 and

His-79.

Of course, we can only compare energies for structures that have the same number of atoms and electrons. Therefore, we focused on states with one extra proton for Ni-A and two extra protons for Ni-SU. The motivation for this choice is that the [NiFe] active site should be close to neutral if it is to be stable in the centre of a protein [12]. There is only one positively charged residue in the neighbourhood, Arg-476 (His-79 can also be positively charged, but when it was protonated in our calculations, one of the protons always moved to Cys-546 instead). Therefore, we have assumed that the total charge of this large system should be neutral when Glu-25 is assumed to share a proton with Cys-543. Consequently, the medium-sized system should have a single negative charge, because Arg-476 is missing.

The results of the QM/MM calculations are shown in Table 6 (bond lengths) and Table 7 (energies). It can be seen that the geometry does not change much between the various protonation states. The largest structural differences are observed when the bridging atom group is protonated. Depending on whether it is an O^{2-}, OH^{-}, OH_2, or peroxide bridge, the Ni–O and Fe–O distances vary quite appreciably (Ni–O 1.73–1.74, Ni–OH 1.82–1.87, Ni–OH_2 1.93, Ni–OO 1.83–1.95, Fe–O 1.87–1.92, Fe–OH 2.01–2.02, Fe–OH_2 2.15, Fe–OO 1.99–2.12 Å). This is in good agreement with earlier calculations. In accordance with the quantum refinements, the observed Ni–O distance, 1.86 Å, is consistent with both a hydroxide or peroxide bridge (but not with an O^{2-} bridge), but the observed Fe–O distance, 2.22 Å is longer than in any of the optimised structures, although it is closer to the distances observed in some of the calculated peroxide structures (as well as the structure with a water bridge). Likewise, the Ni–Fe distance in the crystal structure (2.93 Å) is somewhat more consistent with the peroxide and OH^{-} structures (2.85–3.01 and 2.87–2.97 Å, respectively) than with the O^{2-} or water bridged structures (2.79–2.88 and 3.07 Å, respectively).

In the published crystal structures of [NiFe] hydrogenases with a putative bridging peroxide ligand, the is 1.44-1.45 [7] and 1.57 Å [8], respectively. In our calculations, this bond-length varies widely, ranging from 1.41-1.43 Å for a deprotonated O_2^{2-} ligand to 1.48-1.59 Å for a protonated HO_2^{-} ligand (longer in the reduced than in the oxidised structures, but shorter when the proton formed a hydrogen bond to the Cys ligands). Thus, there is a clear difference in the O–O bond length between a peroxide and a hydroperoxide, but the available crystal structures are probably not accurate enough to decide the protonation status of the peroxide from these data.

In general, the differences between the QM/MM and vacuum structures are small, supporting the view that the influence of the protein on the geometries is small. A general trend is that the Ni–S4 bond is longer in the QM/MM calculations, owing to interaction with His-79, which was not present in the vacuum optimisations.

As mentioned above, two of the Cys ligands share a proton with a surrounding amino acid: Cys-543 with Glu-25 and Cys-546 with His-79. We started to test where these protons prefer to reside by optimising the two possibilities. It turned out that for the first pair, the proton prefers

to reside on Glu-25 (giving a neutral glutamic acid residue) in all structures. Moving it to the sulphur atom resulted in an increase in energy by at least 150 kJ/mole. Likewise, it turned out to be at least 200 kJ/mole more favourable to have the proton on $N^{\varepsilon 2}$ of His-79 than on S^{γ} of Cys-546, as long as the His residue was not doubly protonated. In the latter case, the proton preferred to reside on the sulphur atom, but the energy difference was small.

Considering that Cys-543 and Glu-25 share a hydrogen bond, it seems quite unlikely that Cys-543 could be further protonated. However, test calculations showed that this was actually possible and gave structures with not significantly increased energies. Yet, the protonation destroyed the O_{Glu}–H···S3 hydrogen bond and increased the S3–O_{Glu} distance from 3.1 to 3.7 Å, which is inconsistent with the crystal structures, 3.3 Å. This indicates that Cys-543 should not be additionally protonated. On the other hand it supports a possible role of Cys-543 in proton transfer.

Relative energies of the various protonation states are presented in Table 7. As discussed above, the surrounding protein and solvent have a quite strong influence on the calculated energies. Therefore, three different energies are given, viz. energies obtained in a vacuum (first column), energies obtained with a point-charge model of the surrounding protein (last column), and energies including the solvation effect of the surrounding protein and solvent, as well as the difference in the MM energy (column "Total"). The latter energy is considered most accurate, but the estimated accuracy is not better than ~25 kJ/mole.

The results for the Ni^{III} states without any peroxide show that an OH^- bridge is much more stable than an O^{2-} bridge and protonation of any of the Cys residues (by at least 113 kJ/mole; the same result is obtained also in vacuum or with the point-charge model). This is in accordance with the previous quantum refinement results and indicates that these structures are in the $[III\ OH]^{2-}$ state. As can be seen in Table 6, it gives Ni–S distances of 2.22 and 2.19 Å for the terminal cysteines, and 2.37 and 2.42 Å for the bridging cysteines. This is in reasonable agreement with a recent EXAFS study that gave two Ni–S distances of 2.17 Å and another two of 2.35 Å for the oxidised enzyme [64]. Of course, we cannot rule out the possibility that the state has an O^{2-} bridge and no protons on the basis of energies (the energies are not comparable), but such a complex has significantly longer Ni–S4 distance (2.8–2.9 Å), which is not consistent with the EXAFS or crystal data [38,64].

For the corresponding reduced state, the relative energies indicate that the extra proton binds to the Cys residues, rather than to the OH^- ion. This is an effect mainly of the surrounding protein and the energy difference is smaller in vacuum (9 kJ/mole, compared to 54 kJ/mole). However, the results do not allow us to decide which of the Cys residues is most likely to be protonated, because protonation on Cys-72 and Cys-546 (S1 and S4) give the same energy within 12 kJ/mole, and protonation of Cys-543 (S3) is favoured by the point-charge model.

Almost the same ordering was also observed for the oxidised state. Thus, it is notable that the QM/MM energies do not confirm the consensus that it is much harder to protonate the bridging Cys ligands. On the contrary, the total energies indicate that the bridging Cys-546 (S4)

ligand is among the two most easily protonated ligands in both the reduced and oxidised state. This is most likely connected to the fact that this residue forms a hydrogen bond to His-79. However, even the other bridging Cys ligand (S2) is intrinsically (i.e. in vacuum) the second most easily protonated residue in the reduced state (however, it has the highest total energy in both states, mainly owing to the surrounding protein).

We have also optimised a number of structures with a bridging (hydro)peroxide ion with QM/MM. As can be seen from Figure 8, the QM/MM structures (with a varying protonation of the peroxide and oxidation state of Ni) are quite similar to the two crystal structures with peroxide ligands (Ni-A/Ni-SU structures from *D. fructosovorans* and *D. gigas* [7]). The orientation of the peroxide group is slightly different in the QM/MM and crystal structures, but the largest deviation between the same atom in the overlayed structures is 0.47 Å for the atom closest to Fe and 0.64 Å for the other O atom. This is not much larger than the largest difference between the two crystal structures (0.32 Å) or the various QM/MM structures (0.46 Å). This shows that the suggested position of the peroxide in the crystal structures makes chemical sense.

Structures with an oxidised cysteine ligand

We have also performed a series of tests of the recent suggestions that some of the cysteine ligands may be oxidised [7,8,38,68]. First, we tried to oxidise the bridging Cys-75 ligand (S2) to a Cys-SO group, as was suggested in two recent oxidised structures [7,8]. As can be seen from Figure 9, the QM/MM structures (both with Ni^{II} or Ni^{III}) are quite similar to the crystal structure [7], despite the facts that no X-ray data have been used in the calculations and that the crystal contains partly a peroxide bridge whereas a OH^- bridge was employed in the calculations. In particular, S2 of the oxidised Cys residue is essentially on the same position as before the oxidation (the deviation in the overlayed structures is only 0.39 Å, explaining why only one position of this atom is discerned in the crystal structure. The deviation for the O2 atom between the crystal and QM/MM structures is somewhat larger, 0.65 Å, but it still shows that the suggestion of an oxidised Cys residue is chemically reasonable, as is the position suggested by the electron density. Thus, we can conclude that it is likely that Cys-75 (S2) is partly oxidised in the oxidised enzyme. This is also confirmed by quantum refinement calculations, which gave an optimised occupancy of 0.2 for an oxidised S2 (with OH^- as the bridging ligand).

In several crystal structures, an alternative configuration of the terminal S3 atom has been reported [7,38,68] (cf. Figures 8 and 9). This is a quite strange position, because it cannot make any favourable interactions with the surrounding protein in that position; the closest polar atom is N^{ε} of Arg-476, which is 3.1 Å away, in an unfavourable geometry for a hydrogen bond. However, it is closer to the bridging peroxide/hydroxide group of the [NiFe] cluster, 2.6 Å S–O distance. In fact, the distance between this alternative position for S2 and the hydrogen atom of the bridging OH^- group is typically 2.2–2.3 Å in our structures, which gives a reasonable

hydrogen bond. However, any attempt to optimise a structure with a normal Cys ligand in this alternative position failed (independent of the bridging ligand), even if a favourable hydrogen bond with the bridging ligand was present from the beginning – in all structures, S2 moved back to the normal position, showing that this is appreciably more stable than the alternative position.

Therefore, we made some further investigations of this alternative position. First, we tested if both the normal and the alternative position may be occupied simultaneously if Cys-543 (S3) is oxidised to SO^- or SO_2^-. If the S atom was placed in the alternative position, such structures were possible, but the O atom ended up in positions not observed in crystal structures (irrespectively if the bridging ligand was OH^- or $(H)O_2$). It is interesting to observe that in these structures, Glu-25 followed the movement of S3 to form a hydrogen bond to O3 of the oxidised Cys-543 (Figure 10).

We also tested to put S3 in the crystal position and O3 close to the alternative position. However, O3 ended up 1.1 Å from the alternative position and the S atom moved 0.6 Å away from the crystal position. On the other hand, Glu-25 stayed close to the crystal position. In such a structure, there was no hydrogen bond between the bridging hydroxide and the oxidised cysteine residue (the OH–O3 distance was 3.0 Å, with an O–H–O3 angle of 84°). All the other residues stayed close to their crystal positions. The structure was 80 kJ/mole less stable than the structure with S3 in the alternative position.

A similar structure was obtained also with peroxide as the bridging ligand (O3 1.1 Å from the alternative position and S3 0.4 Å from the crystal position), but in this case an excellent hydrogen bond could be formed between the peroxide and the oxidised cysteine (H–O3 distance 1.5 Å with an O–H–O3 angle of 157°). In fact, the hydrogen atom preferred to reside on the oxidised cysteine (giving CH_3SOH), rather than on the peroxide, although the optimisation was started with an HO_2^- ion. However, this structure was also ~30 kJ/mole less stable than the structures with S3 in the alternative position.

Finally, we looked at the energies of the various oxidised structures, collected in Table 7. Interestingly, it turned out to be much more favourable to oxidise the Cys residues than the OH^- bridge: irrespectively of the oxidation state of the Ni ion, a O_2^{2-} or HO_2^- bridge is over 180 kJ/mole less stable than the two states with S2 or S3 oxidised. The latter two states have the same energy within 30 kJ/mole and calculations systems with Ni in the +II or +III oxidation states gave opposite results. Therefore, we suggest that in the present structure both Cys-75 and 543 are partly oxidised. Oxidised Cys-75 stays close to the crystal position, whereas for oxidised Cys-543, it becomes favourable to move to the alternative position, forming a favourable hydrogen bond to the hydroxide ion. As observed in the QM/MM optimisations, Glu-25 follows this movement to an alternative position, where it still forms a hydrogen bond to Cys-543.

We tested this suggestion by running a quantum refinement with the QM system consisting of an oxidised Cys-543, with S3 in the alternative position, and with Glu-25 in an alternative

configuration (S3, O, and COO of Glu-25 with an occupancy of 0.2). The normal positions of S3 and Glu-25 were in the MM system with 0.8 occupancy. Interestingly, this gave an excellent structure that is completely in agreement with the electron density map. Figure 11 shows the difference map before and after the oxidation of S3 and the movement of Glu-25. It can be seen that the oxidation removes all positive density around S3 (green) and most of it also around S2. Furthermore, no additional differential density is observed close to the added oxygen atom, showing that it is actually compatible with the electron density, although it is located on the edge of the density. Finally, a strong improvement can be observed around Glu-25 (yellow density), strongly indicating that it is disordered in the crystal structure in a way that probably is connected to the oxidation of S3.

Interestingly, in the high-resolution structure of the Ni-A state of hydrogenase from *D. vulgaris Miyazaki*, S3 is also oxidised, but the oxygen atom (O3) is in a different positions: It points towards S4, with an O3–S4 distance of only 2.40 Å (Figure 12, green). The S3–O3 distance is 1.76 Å, significantly longer than the S2–O2 bond (1.58 Å), and the corresponding bonds in the QM/MM structures (1.54–1.66 Å). In such a configuration, both the S3 and O3 atoms may coordinate to Ni, with distances of 2.12 and 2.26 Å, respectively. In the Ni-B state (Figure 12, magenta) and in reduced structures, the O3 atom has moved ~1 Å away from S4, giving a Ni–O3 distance of 2.74 Å [8]. The distance between the O3 position in our [III SO3 OH]$^{2-}$ structure (Figure 12, atomic colours) and this crystal structure of the Ni-A state is 2.71 Å. We have tried to obtain such structures also by QM/MM optimisations. However, only the structure observed for the Ni-B state was obtained in the optimisations (structures started from the Ni-A structure ended up in the same structure as those started from the Ni-B state). They have comparable energies to that of the other structures. This shows that there are two possible positions of the O3 atom of an oxidised Cys-543 (Figure 12, magenta and atomic colours), but the structure observed in the Ni-A state of *D. vulgaris Miyazaki* (Figure 12, green) is most likely an artefact, perhaps representing a mixture with the S4 atom also oxidised, as the short O3–S4 distance also indicates.

Conclusions

In this investigation, we have used a series of different QM methods, some of them combined with experimental data, to elucidate the nature of an oxidised crystal structure of [NiFe] hydrogenase. Our original aim was to use quantum refinement to decide the protonation status of the active [NiFe] site. However, this turned out to be hard, because of the minimal change in the geometry when the Cys ligands are protonated. Moreover, we found an unusual disagreement between the crystallographic data and the QM calculations, with much larger discrepancies between the quantum-refined and vacuum structures than have been seen in any earlier investigation. The reason for this is most likely disorder in the crystal structure: it contains a mixture of different oxidation (and protonation) states of the Ni ion and its ligands.

Therefore, the crystallographic data are incompatible with a single structure and we turned instead to QM/MM optimisations and accurate energy calculations to estimate the intrinsic stability of various protonation states. Using such a procedure, we could clearly show that a $Ni^{III}OH^-$ structure is appreciably more stable than states with a bridging O^{2-} group. However, for the corresponding reduced state, the energy differences are smaller (even if they still point towards a OH^- bridge). Moreover, the accuracy does not allow us to decide which of the four Cys ligands is most easily protonated. In fact, the results indicate that it cannot even be assumed that the terminal Cys ligands are more easily protonated than the bridging ones, as is normally assumed.

Still, the calculations have allowed us to solve several important issues for the [NiFe] hydrogenases. First, both the quantum refinement and QM/MM results clearly show that the major structure contains a OH^- bridge, rather than a water or O^{2-} bridge. A similar conclusion was reached by Hall et al. by vacuum calculations on a smaller model [18,69], whereas Stein et al. proposed a O^{2-} bridge for the Ni-A state [70].

Second, we have seen that quantum refinement is sensitive to disorder in the crystal structure. Of course, this is a problem in the application of this method. On the other hand, it also shows that quantum refinement can be used to identify disorder in crystal structures. Naturally, it is important to know if disorder is present, because if so is the case, the details of the structure cannot be trusted, because it is a mixture. The disorder in the crystal structure is partly caused by the fact that the original protein preparation already contained a mixture of spectroscopic states (~80% Ni-A). However more importantly, the crystal structure shows clear signs of X-ray induced radiation damage (breakage of the Cys–Cys link on the surface and decarboxylation of some Asp and Glu residues), so it is likely that also the active site has been partly reduced during data collection.

Third, we have identified several errors in the original structure. In particular, one of the Fe–CN distances is 2.09 Å in the crystal structure, but it should be ~1.90 Å, like the other Fe–CN distance. Likewise, the Fe–O distance should be ~2.05 Å rather than 2.22 Å, reported in the crystal structure. Finally, the two Ni–S distances of the two bridging Cys ligands are appreciably longer in the crystal structure (2.52 Å) than in the calculations and recent EXAFS measurements (~2.35 Å) [64]. This is a clear feature of the electron density, but is most likely caused by the disorder of the active site, rather than representing the true distances in any single state of the active site.

Fourth, our QM/MM calculations show that the suggested structures with a peroxide bridge in the unready Ni-A and Ni-SU [7,8] are chemically reasonable and in accordance with QM calculations. Therefore, it is likely that the reported structures of Ni-A and Ni-SU [7,8] contain a peroxide ligand. However, the quantum refinements indicate that a hydroxide bridge fits better the electron density map of our studied structure [38], although we cannot exclude that peroxide structures are present as minor conformations in the crystal. The reason for this is probably that the active site is photoreduced by the X-ray radiation in this structure. In fact, our

results indicate that the studied crystal structure most likely contains at least three conformations: The major conformation contains a bridging hydroxide ligand between the two metal ions. The first minor conformation (~20 % occupancy) contains an oxidised Cys-543 (S3), but still a OH^- bridge. S3 has moved to an alternative position and the oxygen atom resides in a new position 1.5 Å from the original position of S3. The nearby Glu-25 moves ~1.8 Å to form a strong hydrogen bond to this oxygen. This provides a novel and attractive explanation to the otherwise strange and isolated alternative configuration of S3, observed in several crystal structures [7,8,38,68]. The second minor conformation (~5 % occupancy) contains an oxidised Cys-75, in which S2 resides essentially in its original position. Such a combination of three conformations explain most of the features in the crystal structure. The oxidation state of the three structures is quite uncertain and it is possible that there are more minor conformations with the same atoms but a different oxidation state of Ni.

Fifth, we have shown that there is also another possible positions for the atoms in an oxidised Cys-543, which are observed in the recent structure of oxidised [NiFe] hydrogenase from *D. vulgaris* [8], although only one of the two suggested positions is compatible with the QM data (Figure 12, magenta).

This gives a quite complicated picture of the oxidised states of [Ni,Fe] hydrogenases, consisting of mixtures of peroxide bridges and oxidations of two or three of the Cys ligands, with the atoms in different positions. However, the results are in accordance with the observation of such mixtures in all recent high-resolution crystal structures of oxidised states (also those with less radiation damage than the present structure) [7,8,38,68]. It is also consistent with the recent electrochemical suggestion that there is at least one unobserved intermediate between Ni-SU and Ni-SI [71]. An investigation of the C–O and C–N bond lengths (strongly correlated to the corresponding vibrational frequencies) show that the difference in these distances between the various peroxide and oxidised Cys structures are small, indicating that it may be hard to discern the various structures by IR spectroscopy.

Recent electrochemical experiments [71] have shown that Ni-A can be activated by a three-step mechanism, involving a reversible reduction (to Ni-SU), a slow reversible step to an isomer of Ni-SU, followed by a final irreversible step induced by H_2 or CO. The latter step indicates the exchange of a dissociable ligand to the solvent. It is attractive to assign the latter to the dissociation of a hydrogen peroxide or hydroperoxide [71], in accordance with the recent crystal structures [7,8]. A possible interpretation of the previous step, in view of the present results, would be an isomerisation between oxidised Cys residues and a peroxide structure. The active site might need to be reduced to stabilise a protonated hydroperoxide or even a hydrogen peroxide, which should dissociate more easily than a fully deprotonated O_2^{2-} ion.

Such a mechanism would nicely explain why an oxygen ligand to Ni in the Ni-A state may derive both from O_2 (via an unstable peroxide bridge) and H_2O (the OH^- ligand) [6]. Moreover, it explains also why it takes longer time to activate the unready states than the ready B and SI states (which have a OH^- bridge or no bridge at all, respectively, and no oxidised Cys ligand)

[7]. Finally, it could explain why the *Desulfomicrobium baculatum* enzyme with a selenocysteine residue in the position corresponding to Cys-543 is resistant to oxygen and does not form the unready states [72]. However, test calculations (Table 7) indicate that Se is as easily oxidised as S, at least compared to HO_2^-.

Moreover, the energies in Table 4 indicate that all peroxide structures are strongly disfavoured (by more than 183 kJ/mole) compared to structures with oxidised Cys ligands. This is in accordance with S-O bonds being stronger than O-O bonds [73], but it indicates that no peroxide structures would be expected for [Ni,Fe] hydrogenase, in variance to the crystallographic observations [7,8]. There are at least two possible solutions to this dilemma. The first is that the diatomic bridge is not a peroxide, but another species, e.g. HOS^-, which would have a lower tendency to oxidise the ligands. The second explanation is that the activation barrier for the oxidation of a Cys ligand by the peroxide bridge is so high that the reaction is slow at ambient temperature. This would also explain why both a diatomic bridge and oxidised Cys residues with fractional occupancies are always observed in the same crystal structure [7,8].

In conclusion, we have obtained a number of interesting results for the [NiFe] hydrogenases, using a combination of crystallographic data, QM/MM, and accurate energy calculations. A few questions regarding the oxidised states still remain. An important step towards their solution would be crystal structures of pure spectroscopic states and information about the spectroscopic state of the crystals both before and after data collection.

Acknowledgements

This investigation has been supported by grants from the Swedish research council, and by computer resources of Lunarc at Lund University. We thank Profs. A. Volbeda and J. C. Fontecilla-Camps for kindly providing us the crystallographic structure factors for orthorombic structure of [NiFe] hydrogenase from *D. fructosovorans* and for fruitful discussions, and Dr. T. Rasmussen for help with the MEAD calculations.

References

1. P.M. Vignais, B. Billoud and J. Meyer, FEMS Microbiol Rev. 25 (2001) 455-501.
2. M. Frey, J.C. Fontecilla-Camps, A. and Volbeda, in A. Messerschmidt, R. Huber, T. Poulus, and K. Wieghardt (Eds.) Handbook of metalloproteins. Wiley: New York, 2001, pp. 880-896.
3. Lemon, B. J.; Peters, J.W.; *Iron-only hydrogenases*. In: Messerschmidt, A.; Huber, R; Poulus, T; Wieghardt, K; editors. *Handbook of metalloproteins*. Wiley: New York, 2001. p. 738-751.
4. E.J. Lyon, S. Shima, G. Buurman, S. Chowdhuri, A. Gatschauer, K. Steinbach, R.K.

Thauer, Eur. J. Biochem. 271 (2004) 195-204.
5. A. Volbeda & JC Fontecilla-Camps, *Coord Chem Rev* (2005), **249**, 1609-1619.
6. Carepo, M.; Tierney, D. L.; Brondino, C. D.; Yang, T. C.; Pamplona, A.; Telser, J.; Moura, I.; Moura, J. J. G.; Hoffman, B. M. *J. Am. Chem. Soc.* **2002**, *124*, 281-286.
7. Volbeda A, Martin L, Cavazza C, Matho M, Faber BW, Rosenboom W, Albracht SPJ, Garcin E, Rousset M, Fontecilla-Camps JC (2005) J Biol Inorg Chem, 10:239-249; erratum in J Biol Inorg Chem 10:591
8. H. Ogata, S. Hirota, A. Nakahara, H. Komori, N. Shibata, T. Kato, K. Kano, Y. Higuchi (2005) Structure 13, 1635-1642.
9. Davidson G, Choudhury SB, Gu Z, Bose K, Roseboom W, Albracht SJ, Maroney MJ (2000) Biochemistry 39:7468-7479
10. Dole F, Fournel A, Magro V, Hatchikian EC, Bertrand P, Guigliarelli B (1997) Biochem. 36:7847-7854
11. Huyett JE, Carepo M, Pamplona A, Franco R, Moura I, Moura JJG, Hoffman BM (1997) J Am Chem Soc 119:9291-9292
12. Pavlov, M.; Blomberg, M. R. A.; Siegbahn, P. E. M. *Int. J. Quantum Chem.* **1999**, *73*, 197-207.
13. Siegbahn, P. E. M.; Blomberg, M. R. A.; Wirstam, M.; Crabtree, R. H. *J. Biol. Inorg. Chem.* **2001**, *6*, 460-466.
14. Siegbahn PEM (2004) Adv Inorg Chem 56:101-125
15. Stein, M.; Lubitz, W. *Curr. Opin. Chem. Biol.* **2002**, *6*, 243-249.
16. Stein M, Lubitz W (2004) J Inorg Biochem 98:862-877
17. Fan, H.-J.; Hall, M. B. *J. Biol. Inorg. Chem.* **2001**, *6*, 467-473.
18. Fan H-J, Hall MB (2002) J Am Chem Soc 124:394-395
19. A. Pardo, A.L De Lacey, V.M. Fernández, H.-J. Fan, Y. Fan, M.B. Hall, J. BIol. Inorg. Chem. 11 (2006) 286-306.
20. De Gioia, L.; Fantucci, P.; Guigliarelli, B.; Bertrand, P. *Inorg. Chem.* **1999**, *38*, 2658-2662.
21. M. Bruschi, G. Zampella, P. Fantucci, L. De Gioia (2005) Coord. Chem. Rev. 249, 1620-1640.
22. Amara, P.; Volbeda, A.; Fontecilla-Camps, J. C.; Field, M. J. *J. Am. Chem. Soc.* **1999**, *121*, 4468-4477.
23. Ryde, U., Olsen, L, Nilsson, K. *J. Comp. Chem.* **2002**, *23*, 1058-70.
24. Ryde U, Nilsson K (2003), J Am Chem Soc 125:14232-14233
25. Nilsson K, Ryde U (2004) J Inorg Biochem 98:1539-1546
26. Nilsson K, Hersleth HP, Rod TH, Andersson KK, Ryde U (2004) Biophys J 87:3437-3447
27. Ryde, U. *J. Comput.-Aided mole. Design* **1996**, *10*, 153-164.
28. Ryde, U.; Olsson, M. H. M. *Int. J. Quantum Chem.* **2001**, *81*, 335-347.

29. Treutler, O.; Ahlrichs, R. *J. Chem. Phys.* **1995**, 102, 346-354.
30. Case, D. A.; Pearlman, D. A.; Caldwell, K. W.; Cheatham III, T. E.; Wang, J.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Gohlke, H.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; SIngh, U. C.; Werner, P. K.; Kolman, P. A. AMBER 7, University of California, San Francisco, 2002.
31. Cornell, W. D.; Cieplak, P.; Bayly, C. I.;Gould, I. R.; Merz, K.M.; Ferguson, D. M.; Spellmeyer, D. C.;Fox, T.; Caldwell, J. W.;Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179-5197.
32. Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357-19363.
33. Maseras, F.; Morokuma, K. *J. Comput. Chem.*, **1995**, *16*, 1170-1179.
34. Brunger, A. T.; Adams, P. D.; Clore, G. M.; Delano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J.-S.; Kuszewsk, J. I.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L.; Crystallography & NMR System CNS, Version 1.0, Yale University, 2000.
35. Engh, R. A.; Huber, R. *Acta Cryst. A*, **1991**, *47*, 392-400.
36. Pannu, N. S.; Read, R. J. *Acta Cryst. A*, **1996**, *52*, 659-668.
37. Källrot N, Nilsson K, Rasmussen T, Ryde U (2005) Intern J Quant Chem, 102:520-541
38. Volbeda A.; Montet, Y.; Vernède, X.; Hatchikian, E. C.; Fontecilla-Camps, J. C. *Intern. J. Hydrogen Energy* **2002**, *27*, 1449-1461.
39. Kurkin S, George SJ, Thorneley RN, Albracht SP (2004), *Biochemistry* **43**, 6820-6831.
40. Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098-3100.
41. Perdew J. P. *Phys. Rev. B* **1986**, *33*, 8822-8824.
42. Schäfer, A.; Huber, C.; Ahlrichs, R. *J. Chem. Phys.* **1994**, *100*, 5829-5835.
43. Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. *Chem. Phys. Lett.* **1995**, *240*, 283-290.
44. Eichkorn K , Weigend F, TreutlerO , Ahlrichs R (1997) Theor Chem Acc 97:119-126
45. Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372-1377.
46. Siegbahn, P. E. M.; Blomberg, M. R. A. *Annu. Rev. Phys. Chem.* **1999**, *50*, 221-249.
47. Siegbahn, P. E. M.; Blomberg, M. R. A. *Chem. Rev.* **2000**, *100*, 421-437.
48. Bruschi M, De Gioia, L, Zampella G, Reiher M, Fantucci P, Stein M (2004) J Biol Inorg Chem 9:873-884
49. Wang H, Ralston CY, Patil DS, Jones RM, Gu W, Verhagen M, Adams M, Ge P, Riordan C, Marganian CA, Mascharak P, Kovacs J, Miller CG, Collins TJ, Brooker S, Croucher PD, Wang K, Stiefel EI, Cramer SP (2000) J Am Chem Soc 122:10544–10552
50. Klähn M, Braun-Sand S, Rosta E, Warshel A (2005) J Phys Chem B, DOI: 10.1021/jp0521757, ASAP article.
51. Bashford, D.; Gerwert, K. *J. Mol. Biol.* **1992** *224*, 473-486.

52. Sitkoff T, Sharp KA, Honig B (1994) J Phys Chem 98:1978-1988
53. Ullmann GM, Noodleman L, Case DA (2002) J Biol Inorg Chem 7:632-639
54. Bashford D, Case DA, Dalvit C, Tennant, L, Wright PE (1993) Biochem 32:8045-8056
55. Olsen L, Rasmussen T, Hemmingsen L, Ryde U (2004) J Phys Chem B 108:17639-17648
56. L. Rulíšek & U. Ryde (2006) J. Phys. Chem. B, in press.
57. Shen, Y.; Ryde, U. *J. Inorg. Biochem.*, **2004**, *98*, 878-895..
58. Choudhury, S. B.; Pressler, M. A.; Mirza, S. A.; Day, R. O.; Maroney, M. J. *Inorg. Chem.* **1994**, *33*, 4831-4839
59. Allan, C. B.; Davidson, G.; Choudhury, S. B.; Gu, Z.; Bose, K.; Day, R. O.; Maroney, M. J. *Inorg. Chem* **1998**, *37*, 4166-4167
60. Ryde U (1996) Eur Biophys J 24:213-221.
61. Sigfridsson, E.; Olsson, M. H. M.; Ryde, U. *J. Phys. Chem. B* **2001**, *105*, 5546-5552.
62. Olsson, M. H. M.; Ryde, U. *J. Am. Chem. Soc.* **2001**, *123*, 7866-7876.
63. Ryde U (2002) in Recent Research Developments in Protein Engineering, 2, pp 65-91; Research Signpost, Trivandrum
64. Gu, W.; Jacquamet, L.; Patil, D. S.; Wang, H.-X.; Evans, D. J.; Smith, M. C.; Millar, M.; Koch, S.; Eichhorn, D. M.; Latimer, M.; Cramer, S. P. *J. Inorg. Biochem.* **2003**, *93*, 41-51.
65. Jones TA, Zou JY, Cowand WE, Kjeldgaard M (1991) Acta Cryst A47:110-119
66. C.M. Wilmot, T. Sjögren, G.H. Carlsson, G.I. Berglund, J. Hajdu, Methods in Enzymology 353 (2002) 301-318.
67. Weik M, Ravelli RB, Kryger G, McSweeney S, Raves ML, Harel M, Gros P, Silman I, Kroon J, Sussman JL (2000) Proc Natl Acad Sci USA 97:623-628.
68. Matias PM, Soares CM, Saraiva LM, Coelho R, Morais J, Legall J, Carrondo MA (2001) J Biol Inorg Chem 6:63-81
69. Niu, S. Q.; Thomson, L. M.; Hall, M. B. *J. Am. Chem. Soc.* **1999**, *121*, 4000-4007.
70. M. Stein, W. Lubitz, PhysChemChemPhys 3 (2001) 2668-2675.
71. SE Lamle, SPJ Albracht & FA Armstrong (2005), JACS 127, 6595-6604
72. Garcin, E.; Vernède, X.; Hatchikian, E. C.; Volbeda, A.; Frey, M.; Fontecilla-Camps, J. C. *Structure* **1999**, *7*, 557-566.
73. W.E. Dascent, Inorganic Energetics, Cambridge University Press, New York, 1982.

Table 1. Geometries for the model compound in Figure 5, as calculated with various methods and from experiment. The basis set indicated is for Ni only. DZpdf differs from DZP by inclusion of one *p*, one *d*, and one *f*-type function. For other atoms, 6-31G* was used together with DZP and TVZP, whereas 6-31+G** was used together with DZpdf. AAD is the average absolute differences in the nine listed bond distances between the calculation and the corresponding crystal structure.

Method	Bond length to Ni_1 (Å)					Bond length to Ni_2 (Å)				Angles (°)		AAD
	Ni_2	S_1	S_2	S_3	N_1	S_1	S_3	S_4	N_2	$S_1Ni_1S_3$	$S_1Ni_2S_3$	(Å)
Unprotonated												
Experiment	2.68	2.22	2.17	2.18	1.94	2.18	2.21	2.14	1.92	82.4	82.7	
BP86 / DZP	2.63	2.23	2.19	2.19	1.99	2.19	2.24	2.18	1.96	83.4	83.3	0.030
BP86 / TZVP	2.67	2.24	2.19	2.20	2.00	2.19	2.25	2.19	1.98	82.7	82.6	0.033
BP86 / DZpdf	2.69	2.24	2.19	2.20	2.00	2.20	2.24	2.19	1.98	82.0	82.0	0.033
SVWN / DZP	2.52	2.18	2.14	2.14	1.93	2.13	2.18	2.13	1.91	83.1	83.2	0.041
SVWN / DZpdf	2.57	2.19	2.15	2.15	1.94	2.14	2.19	2.14	1.92	82.0	82.1	0.030
B3LYP / DZP	2.68	2.27	2.20	2.23	2.00	2.23	2.27	2.20	1.98	82.9	83.0	0.048
B3LYP / DZpdf	2.75	2.28	2.21	2.23	2.01	2.23	2.27	2.20	1.99	82.2	82.2	0.059
Protonated												
BP86 / DZP	2.66	2.22	2.17	2.18	1.98	2.21	2.22	2.19	1.98	83.6	82.9	
BP86 / DZpdf	2.71	2.23	2.18	2.18	1.99	2.22	2.22	2.21	2.00	82.3	81.7	
Alkylated												
Experiment	2.64	2.21	2.16	2.18	1.94	2.19	2.18	2.17	1.95	80.8	81.0	
BP86 / DZP	2.66	2.22	2.17	2.18	1.98	2.21	2.22	2.19	1.98	83.7	83.0	0.034
BP86 / DZpdf	2.72	2.23	2.18	2.18	1.99	2.22	2.22	2.20	1.99	82.4	81.8	0.048

Table 2. Results of the quantum refinement calculations with the small quantum region. Av. and Max |Δ*r*| denotes the average and maximum differences in the bond lengths listed in Table 3 between the quantum-refined and vacuum geometries. The strain energy, ΔE_{QM1}, is defined as the difference in vacuum energy between the quantum-refined and the vacuum-optimised geometry. In all the quantum refinement calculations involving a peroxide ligand, the occupancy of the second oxygen was optimised. This had no effect on the structure, but reduced Δ*R* by ~70%. The optimised occupancy is shown in the last column.

	Δ*R**1000	ΔR_{free}*1000	Av. |Δ*r*| (Å)	Max |Δ*r*| (Å)	ΔE_{QM1} (kJ/mole)	Occupancy of $(H)O_2$
[II /]$^{2-}$	0.38	0.19	0.04	0.18	148	
[III /]$^{-}$	0.37	0.20	0.04	0.19	137	
[^{3}II /]$^{2-}$	0.31	0.13	0.04	0.13	111	
[I /]$^{3-}$	0.32	0.15	0.06	0.16	115	
[II H1 HOH]$^{-}$	0.19	0.02	0.20	1.02	154	
[II H3 HOH]$^{-}$	0.18	0.09	0.15	0.93	158	
[III H1 HOH]^{0}	0.15	0.05	0.13	0.88	104	
[III H3 HOH]^{0}	0.15	0.06	0.12	0.79	122	
[^{3}II H1 HOH]$^{-}$	0.12	0.01	0.08	0.41	89	
[^{3}II H3 HOH]$^{-}$	0.13	0.10	0.11	0.56	97	
[II O]$^{4-}$	0.15	0.04	0.09	0.20	111[a]	
[II H1 O]$^{3-}$	0.14	-0.06	0.08	0.21	108[a]	
[II H3 O]$^{3-}$	0.12	0.10	0.09	0.28	104	
[II H1 H3 O]$^{2-}$	0.12	-0.01	0.10	0.21	123	
[III O]$^{3-}$	0.12	-0.05	0.08	0.20	97[a]	
[III H1 O]$^{2-}$	0.13	-0.10	0.10	0.23	105	
[III H3 O]$^{2-}$	0.14	-0.01	0.08	0.22	82[a]	
[III H1 H3 O]$^{-}$	0.15	-0.06	0.11	0.43	118	
[^{4}III O]$^{3-}$	0.14	-0.01	Ni–S_4 breaks in vacuum		104	
[^{3}II O]$^{4-}$	0.18	0.02	Dissociates in vacuum			
[^{3}II H1 O]$^{3-}$	0.14	-0.09	Dissociates in vacuum			
[^{3}II H3 O]$^{3-}$	0.12	0.03	Dissociates in vacuum			
[^{3}II H1 H3 O]$^{2-}$	0.12	-0.09	Dissociates in vacuum			
[I H1 H3 O]$^{3-}$		Dissociates into [II H_3 O]$^{3-}$ and H				
[II OH]$^{3-}$	0.11	0.02	0.05	0.19	79[a]	
[II H1 OH]$^{2-}$	0.11	-0.02	0.09	0.46	86	
[II H3 OH]$^{2-}$	0.11	0.05	0.06	0.29	80	
[II H1 H3 OH]$^{-}$	0.12	-0.01	0.05	0.21	89	
[II H2 H4 OH]$^{-}$	0.10	0.04	Ni–S_4 breaks in vacuum		120	93
[III OH]$^{2-}$	0.08	0.01	0.04	0.15	76	
[III H1 OH]$^{-}$	0.09	-0.02	0.04	0.13	75[a]	
[III H3 OH]$^{-}$	0.09	0.04	0.06	0.17	72	
[III H1 H3 OH]^{0}	0.10	0.00	0.10	0.40	98	
[III H2 H4 OH^{0}	0.10	0.02	Ni–S_4 breaks in vacuum		118	
[^{3}II OH]$^{3-}$	0.11	0.05	0.08	0.22	93	
[^{3}II H1 OH]$^{2-}$	0.09	-0.02	0.09	0.47		
[^{3}II H3 OH]$^{2-}$	0.09	0.09	0.17	1.44	107[a]	
[^{3}II H1 H3 OH]$^{-}$	0.08	-0.02	0.13	0.37	101	
[III H1 OO]$^{2-}$	0.10	-0.15	Dissociates in vacuum			0.29
[III H3 OO]$^{2-}$	0.10	-0.06	0.05	0.19	41[b]	0.22
[III OOH]$^{2-}$	0.10	-0.06	0.04	0.20	60[b]	0.23
[II H1 OO]$^{3-}$	0.10	-0.10	Ni–S_4 breaks in vacuum		76[b]	0.24
[II H3 OO]$^{3-}$	0.17	-0.06	Ni–S_4 breaks in vacuum		75[b]	0.21
[II OOH]$^{3-}$	0.11	-0.07	Ni–S_4 breaks in vacuum		80[b]	0.20
[^{3}II H3 OO]$^{3-}$	0.12	0.05	0.05	0.16	52[b]	0.32
[IV H3 OO]$^{-}$	0.11	-0.06	0.06	0.17	55[b]	0.20
[^{3}IV H3 OO]$^{-}$	0.07	-0.07	Ni–S_4 breaks in vacuum		57[b]	0.17

[a]Vacuum calculation done with Ni–S4 distance constrained to the quantum-refined value
[b]Vacuum calculation done with Ni–S3 distance constrained to the quantum-refined value.

Table 3. Bond distances as optimised by ComQum-X (first line), by ComQum-X with point charges (bold) and in vacuum (within brackets).

	Ni–S3	Ni–S1	Ni–S2	Ni–S4	Ni–O	Ni–Fe	Fe–O	Fe–S2	Fe–S4	Fe–CO	Fe–CN	Fe–CN
Crystal [38]	2.19	2.19	2.52	2.52	1.86	2.93	2.22	2.33	2.32	1.72	1.95	2.09
[II /]$^{2-}$	2.19	2.25	2.42	2.34		2.89		2.30	2.25	1.70	1.89	1.91
	(2.25)	(2.23)	(2.24)	(2.22)		(2.84)		(2.33)	(2.27)	(1.70)	(1.89)	(1.89)
[III /]$^{-}$	2.19	2.22	2.42	2.34		2.91		2.28	2.25	1.70	1.88	1.89
	(2.21)	(2.19)	(2.22)	(2.22)		(2.82)		(2.30)	(2.26)	(1.71)	(1.88)	(1.87)
[^{3}II /]$^{2-}$	2.20	2.24	2.44	2.44		2.92		2.28	2.30	1.69	1.89	1.90
	(2.23)	(2.28)	(2.32)	(2.31)		(2.88)		(2.32)	(2.32)	(1.69)	(1.89)	(1.88)
[I /]$^{3-}$	2.22	2.29	2.43	2.43		2.90		2.29	2.30	1.71	1.91	1.92
	(2.27)	(2.41)	(2.30)	(2.31)		(2.74)		(2.34)	(2.35)	(1.72)	(1.90)	(1.90)
[II H1 HOH]$^{-}$	2.22	2.22	2.42	2.43	2.18	2.89	2.64	2.29	2.27	1.70	1.88	1.91
	(2.26)	(2.56)	(2.27)	(2.24)	(2.03)	(2.51)	(3.66)	(2.32)	(2.30)	(1.72)	(1.87)	(1.84)
[II H3 HOH]$^{-}$	2.18	2.23	2.40	2.45	2.20	2.89	2.54	2.30	2.28	1.70	1.88	1.90
	(2.20)	(2.25)	(2.23)	(2.19)	(3.14)	(3.08)	(2.43)	(2.32)	(2.32)	(1.70)	(1.88)	(1.89)
[III H1 HOH]^{0}	2.21	2.23	2.44	2.39	2.12	2.88	2.30	2.30	2.27	1.73	1.88	1.90
	(2.20)	(2.25)	(2.27)	(2.17)	(3.00)	(2.92)	(2.19)	(2.32)	(2.32)	(1.74)	(1.90)	(1.89)
[III H3 HOH]^{0}	2.23	2.20	2.41	2.42	2.15	2.88	2.32	2.29	2.28	1.73	1.88	1.89
	(2.30)	(2.16)	(2.20)	(2.22)	(2.94)	(2.87)	(2.23)	(2.29)	(2.34)	(1.74)	(1.89)	(1.89)
[^{3}II H1 HOH]$^{-}$	2.23	2.26	2.48	2.44	2.18	2.90	2.38	2.31	2.30	1.70	1.89	1.90
	(2.24)	(2.53)	(2.42)	(2.34)	(2.23)	(2.87)	(2.79)	(2.30)	(2.30)	(1.70)	(1.87)	(1.88)
[^{3}II H3 HOH]$^{-}$	2.25	2.24	2.47	2.42	2.22	2.90	2.45	2.30	2.29	1.70	1.89	1.90
	(2.58)	(2.23)	(2.40)	(2.38)	(2.16)	(2.70)	(3.01)	(2.29)	(2.29)	(1.71)	(1.86)	(1.88)
[II O]$^{4-}$	2.23	2.26	2.54	2.65	1.82	2.95	1.96	2.38	2.39	1.74	1.91	1.94
	2.24	**2.23**	**2.53**	**2.72**	**1.82**	**2.96**	**1.93**	**2.39**	**2.42**	**1.75**	**1.93**	**1.95**
	(2.42)	(2.45)	(2.34)	(2.65)[a]	(1.83)	(2.92)	(1.90)	(2.57)	(2.59)	(1.74)	(1.91)	(1.91)
[II H1 O]$^{3-}$	2.24	2.17	2.52	2.62	1.81	2.93	1.97	2.37	2.37	1.74	1.90	1.92
	2.25	**2.18**	**2.50**	**2.70**	**1.81**	**2.95**	**1.94**	**2.37**	**2.40**	**1.74**	**1.91**	**1.93**
	(2.36)	(2.20)	(2.38)	(2.62)[a]	(1.80)	(2.78	(1.88	(2.57	(2.58)	(1.75)	(1.88)	(1.89)
[II H3 O]$^{3-}$	2.15	2.27	2.50	2.58	1.81	2.92	1.97	2.36	2.37	1.73	1.90	1.92
	(2.18)	(2.40)	(2.62)	(2.30)	(1.84)	(2.87)	(1.91)	(2.54)	(2.50)	(1.74)	(1.90)	(1.90)
[II H1 H3 O]$^{2-}$	2.19	2.21	2.49	2.55	1.80	2.91	1.96	2.35	2.37	1.74	1.89	1.92
	(2.21)	(2.30)	(2.41)	(2.36)	(1.82)	(2.70)	(1.86)	(2.56)	(2.57)	(1.75)	(1.89)	(1.88)
[III O]$^{3-}$	2.24	2.26	2.50	2.52	1.79	2.91	1.99	2.35	2.36	1.73	1.90	1.92
	2.24	**2.21**	**2.51**	**2.69**	**1.77**	**2.94**	**1.96**	**2.34**	**2.39**	**1.73**	**1.92**	**1.94**
	(2.34)	(2.37)	(2.33)	(2.52)[a]	(1.81)	(2.84)	(1.91)	(2.46)	(2.56)	(1.74)	(1.89)	(1.89)
[III H1 O]$^{2-}$	2.24	2.24	2.46	2.49	1.79	2.88	1.97	2.34	2.35	1.74	1.89	1.91
	2.25	**2.22**	**2.46**	**2.63**	**1.77**	**2.91**	**1.94**	**2.34**	**2.38**	**1.74**	**1.91**	**1.93**
	(2.32)	(2.33)	(2.35)	(2.39)	(1.77)	(2.68)	(1.85)	(2.50)	(2.57)	(1.76)	(1.88)	(1.88)
[III H3 O]$^{2-}$	2.22	2.26	2.44	2.51	1.78	2.89	1.98	2.33	2.36	1.73	1.90	1.91
	(2.32)	(2.33)	(2.26)	(2.51)[a]	(1.78)	(2.78)	(1.87)	(2.44)	(2.58)	(1.75)	(1.89)	(1.89)
[III H1 H3 O]$^{-}$	2.24	2.25	2.43	2.49	1.78	2.87	1.93	2.33	2.36	1.75	1.89	1.91
	(2.32)	(2.30)	(2.32)	(2.27)	(1.82)	(2.45)	(1.87)	(2.45)	(2.53)	(1.77)	(1.90)	(1.90)
[^{4}III O]$^{3-}$	2.28	2.29	2.62	2.47	1.81	2.93	1.96	2.36	2.35	1.73	1.91	1.92
	(2.36)	(2.32)	*(3.86)*	(2.33)	(1.78)	(2.93)	(1.92)	(2.50)	(2.49)	(1.75)	(1.90)	(1.90)
[^{3}II O]$^{4-}$	2.31	2.31	2.63	2.54	1.85	2.93	1.98	2.38	2.39	1.73	1.91	1.93
[^{3}II H1 O]$^{3-}$	2.32	2.23	2.59	2.51	1.84	2.91	1.99	2.37	2.37	1.73	1.90	1.92
[^{3}II H3 O]$^{3-}$	2.25	2.30	2.57	2.52	1.82	2.91	2.01	2.35	2.37	1.73	1.91	1.92
[^{3}II H1 H3 O]$^{2-}$	2.28	2.24	2.53	2.49	1.82	2.90	2.00	2.35	2.36	1.73	1.90	1.91
[II OH]$^{3-}$	2.21	2.21	2.49	2.63	1.90	2.96	2.04	2.34	2.36	1.71	1.90	1.93
	2.22	**2.19**	**2.50**	**2.70**	**1.90**	**2.98**	**2.02**	**2.34**	**2.38**	**1.71**	**1.92**	**1.94**
	(2.30)	(2.26)	(2.30)	(2.63)[a]	(1.93)	(2.98)	(2.05)	(2.40)	(2.45)	(1.71)	(1.90)	(1.91)
[II H1 OH]$^{2-}$	2.22	2.17	2.46	2.58	1.90	2.93	2.06	2.33	2.35	1.72	1.89	1.92
	2.23	**2.16**	**2.47**	**2.66**	**1.88**	**2.96**	**2.03**	**2.33**	**2.37**	**1.72**	**1.90**	**1.93**
	(2.26)	(2.16)	(2.92)	(2.28)	(1.88)	(2.95)	(2.06)	(2.48)	(2.38)	(1.72)	(1.90)	(1.89)
[II H3 OH]$^{2-}$	2.17	2.22	2.44	2.56	1.90	2.93	2.08	2.33	2.34	1.71	1.90	1.91
	(2.19)	(2.25)	(2.55)	(2.27)	(1.95)	(2.92)	(2.07)	(2.46)	(2.36)	(1.72)	(1.90)	(1.89)
[II H1 H3 OH]$^{-}$	2.20	2.21	2.42	2.52	1.90	2.91	2.08	2.31	2.33	1.73	1.89	1.91
	(2.24)	(2.21)	(2.40)	(2.31)	(1.91)	(2.82)	(2.06)	(2.48)	(2.35)	(1.72)	(1.89)	(1.88)
[II H2 H4 OH]$^{-}$	2.16	2.17	2.63	2.64	1.75	2.98	1.95	2.32	2.33	1.75	1.88	1.90
	(2.16)	(2.25)	(2.16)	*(4.15)*	(1.82)	(3.03)	(1.97)	(2.32)	(2.36)	(1.75)	(1.88)	(1.89)
[III OH]$^{2-}$	2.22	2.22	2.49	2.47	1.90	2.92	2.06	2.33	2.32	1.72	1.89	1.91
	2.22	**2.20**	**2.49**	**2.52**	**1.90**	**2.94**	**2.04**	**2.33**	**2.34**	**1.71**	**1.91**	**1.92**
	(2.29)	(2.23)	(2.40)	(2.32)	(1.93)	(2.91)	(2.06)	(2.41)	(2.36)	(1.72)	(1.90)	(1.89)
[III H1 OH]$^{-}$	2.21	2.22	2.46	2.46	1.89	2.89	2.07	2.32	2.30	1.73	1.89	1.91
	2.22	**2.22**	**2.47**	**2.51**	**1.88**	**2.92**	**2.05**	**2.32**	**2.32**	**1.73**	**1.89**	**1.91**
	(2.23)	(2.25)	(2.33)	(2.38)	(1.89)	(2.81)	(2.06)	(2.37)	(2.37)	(1.73)	(1.88)	(1.89)

[III H3 OH]$^{-}$	2.21	2.22	2.44	2.46	1.91	2.89	2.08	2.31	2.30	1.73	1.89	1.91
	(2.34)	(2.22)	(2.27)	(2.36)	(1.94)	(2.81)	(2.07)	(2.30)	(2.42)	(1.73)	(1.89)	(1.89)
[III H1 H3 OH]^{0}	2.25	2.26	2.48	2.43	1.91	2.89	2.06	2.31	2.32	1.75	1.89	1.89
	(2.38)	(2.24)	(2.24)	(2.27)	(1.91)	(2.49)	(2.02)	(2.47)	(2.38)	(1.75)	(1.89)	(1.89)
[III H2 H4 OH]^{0}	2.16	2.21	2.57	2.63	1.76	2.94	1.90	2.32	2.33	1.79	1.88	1.90
	(2.22)	(2.19)	(2.18)	*(3.68)*	(1.79)	(3.05)	(1.86)	(2.37)	(2.40)	(1.79)	(1.87)	(1.89)
[^{3}II OH]$^{3-}$	2.27	2.28	2.59	2.47	1.97	2.95	2.06	2.35	2.35	1.71	1.90	1.92
	(2.49)	(2.38)	(2.39)	(2.52)	(2.07)	(3.04)	(2.04)	(2.45)	(2.40)	(1.72)	(1.90)	(1.90)
[^{3}II H1 OH]$^{2-}$	2.27	2.25	2.54	2.45	1.94	2.91	2.08	2.34	2.33	1.72	1.89	1.91
	(2.28)	(2.72)	(2.34)	(2.39)	(2.01)	(2.87)	(2.04)	(2.48)	(2.41)	(1.72)	(1.89)	(1.88)
[^{3}II H3 OH]$^{2-}$	2.26	2.27	2.53	2.44	1.95	2.91	2.10	2.33	2.33	1.71	1.89	1.91
	(3.70)	(2.22)	(2.33)	(2.44)[a]	(1.97)	(2.81)	(2.08)	(2.36)	(2.49)	(1.72)	(1.89)	(1.89)
[^{3}II H1 H3 OH]$^{-}$	2.27	2.25	2.51	2.44	1.93	2.89	2.11	2.33	2.31	1.73	1.89	1.90
	(2.26)	(2.56)	(2.27)	(2.29)	(2.06)	(2.52)	(2.01)	(2.47)	(2.40)	(1.74)	(1.90)	(1.90)
[III H1 OO]$^{2-}$	2.32	2.25	2.55	2.53	1.85	2.92	2.04	2.37	2.38	1.72	1.9	1.91
[III H3 OO]$^{2-}$	2.32	2.27	2.54	2.48	1.88	2.92	2.07	2.35	2.38	1.71	1.90	1.91
	(2.32)[a]	(2.29)	(2.64)	(2.29)	(1.90)	(2.84)	(2.13)	(2.33)	(2.50)	(1.71)	(1.90)	(1.88)
[III OOH]$^{2-}$	2.26	2.24	2.50	2.50	1.87	2.94	2.02	2.37	2.34	1.72	1.90	1.90
	(2.25)[a]	(2.25)	(2.30)	(2.41)	(1.94)	(2.94)	(2.07)	(2.38)	(2.36)	(1.72)	(1.89)	(1.88)
[II H1 OO]$^{3-}$	2.32	2.18	2.54	2.63	1.84	2.96	2.06	2.39	2.39	1.72	1.91	1.92
	(2.31)[a]	(2.20)	(2.30)	*(2.94)*	(1.89)	(2.99)	(2.09)	(2.41)	(2.49)	(1.72)	(1.91)	(1.91)
[II H3 OO]$^{3-}$	2.42	2.28	2.65	2.46	1.88	2.93	2.08	2.38	2.38	1.71	1.91	1.92
	(2.42)[a]	(2.29)	*(3.40)*	(2.26)	(1.85)	(3.01)	(2.16)	(2.41)	(2.47)	(1.71)	(1.91)	(1.90)
[II OOH]$^{3-}$	2.26	2.24	2.51	2.64	1.83	2.97	2.02	2.38	2.38	1.72	1.91	1.92
	(2.26)[a]	(2.28)	(2.24)	*(3.48)*	(1.89)	(3.20)	(2.11)	(2.37)	(2.43)	(1.71)	(1.91)	(1.92)
[^{3}II H3 OO]$^{3-}$	2.29	2.33	2.56	2.55	2.04	2.94	2.05	2.38	2.40	1.72	1.91	1.92
	(2.29)[a]	(2.49)	(2.46)	(2.53)	(2.10)	(2.96)	(2.04)	(2.39)	(2.55)	(1.72)	(1.91)	(1.89)
[IV H3 OO]$^{-}$	2.29	2.26	2.47	2.44	1.91	2.89	2.07	2.31	2.36	1.74	1.90	1.90
	(2.29)[a]	(2.25)	(2.30)	(2.28)	(1.93)	(2.77)	(2.12)	(2.24)	(2.50)	(1.73)	(1.90)	(1.88)
[^{3}IV H3 OO]$^{-}$	2.32	2.24	2.52	2.50	1.88	2.91	2.07	2.36	2.37	1.74	1.90	1.90
	(2.32)[a]	(2.22)	*(2.98)*	(2.24)	(1.90)	(2.79)	(2.10)	(2.26)	(2.48)	(1.75)	(1.92)	(1.89)

[a] This distance was constrained to the ComQum-X distance.

Table 4. Results of the quantum refinement calculations with the large quantum region.

	ΔR*1000	ΔR_{free}*1000	Av |Δr| (Å)	Max |Δr| (Å)	ΔE_{QM1} (kJ/mole)
[II H4 O]$^{2-}$	0.36	0.31	0.07	0.24	185
[^{3}II H4 O]$^{2-}$	0.35	0.32			
[III H4 O]$^{-}$	0.32	0.18	0.05	0.22	125
[II H4 OH]$^{-}$	0.31	0.18			
[II H2 H4 OH]^{0}	0.32	0.21	Ni–S4 breaks in vacuum		135
[III OH]$^{-}$	0.31	0.10	0.04	0.12	118
[III OO]$^{2-}$	0.32	0.13	0.05	0.17	61

Table 5. Geometries from the quantum refinement calculations with the large quantum region (bold) and the corresponding vacuum geometries calculated by keeping the heavy atoms outside the small quantum system (Figure 2) fixed at the crystal geometry but allowing hydrogen atoms of the large system as well as all atoms in the small quantum system to relax.

	Ni–S3	Ni–S1	Ni–S2	Ni–S4	Ni–O	Ni–Fe	Fe–O	Fe–S2	Fe–S4	Fe–CO	Fe–CN1	Fe–CN2
Crystal [38]	2.19	2.19	2.52	2.52	1.86	2.93	2.22	2.33	2.32	1.72	1.95	2.09
												1.8
												7
	2.22	**2.23**	**2.51**	**2.58**	**1.83**	**2.92**	**1.97**	**2.38**	**2.38**	**1.75**	**1.87**	1.8
[II H4 O]$^{2-}$	2.33	2.27	2.29	2.82	1.82	2.96	1.95	2.37	2.43	1.76	1.84	4
												1.8
												7
	2.22	**2.20**	**2.51**	**2.59**	**1.78**	**2.92**	**1.96**	**2.36**	**2.36**	**1.76**	**1.88**	1.8
[III H4 O]$^{-}$	2.30	2.21	2.34	2.81	1.76	2.93	1.93	2.35	2.37	1.77	1.86	6
												1.8
[^{3}II H4 O]$^{2-}$	**2.25**	**2.25**	**2.51**	**2.49**	**1.83**	**2.90**	**1.99**	**2.37**	**2.37**	**1.75**	**1.87**	**7**
												1.8
[II H4 OH]$^{-}$	**2.21**	**2.18**	**2.51**	**2.57**	**1.91**	**2.95**	**2.05**	**2.36**	**2.35**	**1.74**	**1.88**	**7**
												1.8
												9
	2.19	**2.15**	**2.56**	**2.61**	**1.89**	**2.99**	**2.02**	**2.33**	**2.37**	**1.74**	**1.86**	1.8
[II H2 H4 OH]^{0}	2.13	2.09	3.49	2.84	1.89	3.35	2.04	2.36	2.38	1.73	1.87	8
	2.21	**2.20**	**2.49**	**2.46**	**1.92**	**2.92**	**2.09**	**2.35**	**2.33**	**1.74**	**1.87**	**1.88**
[III OH]$^{-}$	2.28	2.21	2.36	2.38	1.91	2.89	2.04	2.36	2.37	1.74	1.86	1.86
	2.29	**2.23**	**2.57**	**2.51**	**1.88**	**2.94**	**2.12**	**2.35**	**2.36**	**1.73**	**1.88**	**1.88**
[III OO]$^{2-}$	2.37	2.23	2.69	2.34	1.87	2.89	2.10	2.34	2.43	1.73	1.87	1.85
												1.8
[II O]$^{3-}$	2.34	2.30	2.27	2.97	1.82	2.95	1.96	2.39	2.47	1.76	1.85	4
												1.8
[II H1 O]$^{2-}$	2.33	2.20	2.28	2.85	1.79	2.89	1.95	2.39	2.45	1.76	1.85	4
												1.8
[II H2 O]$^{2-}$	2.15	2.21	3.23	2.91	1.77	3.18	1.98	2.40	2.41	1.76	1.85	7
												1.8
[II H1 OH]$^{-}$	2.25	2.14	2.32	2.69	1.87	2.94	2.07	2.38	2.37	1.74	1.86	6
												1.8
[II OH]$^{2-}$	2.25	2.20	2.33	2.86	1.91	3.03	2.06	2.37	2.39	1.74	1.86	5
												1.8
[II H1 H4 OH]^{0}	2.24	2.14	2.33	2.80	1.86	2.97	2.06	2.37	2.37	1.73	1.87	6
												1.8
[III OH]$^{-}$	2.28	2.21	2.36	2.38	1.91	2.89	2.04	2.36	2.37	1.74	1.86	6
												1.8
[IV O]$^{-}$	2.28	2.27	2.35	2.22	1.80	2.80	1.89	2.42	2.46	1.77	1.85	5
												1.8
[^{4}III OH]$^{-}$	2.30	2.20	2.79	2.27	1.98	2.88	2.06	2.35	2.36	1.75	1.88	6

												1.8
$[^{3}\mathrm{II\ OH}]^{2-}$	2.37	2.26	2.76	2.32	2.04	3.05	2.05	2.40	2.38	1.74	1.87	5
												1.8
$[\mathrm{I\ OH}]^{3-}$	2.38	2.24	2.87	2.45	2.06	3.15	2.04	2.41	2.38	1.73	1.87	7
												1.8
$[^{4}\mathrm{I\ OH}]^{3-}$	2.44	2.27	2.71	2.33	2.01	3.05	2.04	2.42	2.37	1.73	1.87	8

Table 6. Metal–ligand distances (Å) in the QM/MM optimisations with the medium-sized quantum region.

	Distance to Ni (Å)							Distance to Fe (Å)					
	S3	S1	S2	S4	O1	O2	Fe	O1	S2	S4	CO	CN1	CN2
Crystal structures													
oxidised [38]	2.19	2.19	2.52	2.52	1.86		2.93	2.22	2.33	2.32	1.72	1.95	2.09
Ni-B [7]	2.17	2.15	2.30	2.55	1.88		2.81	1.89	2.24	2.28	1.72	1.90	1.90
Ni-A [7]	2.26	2.15	2.59	2.55	1.90	2.11	2.93	1.90	2.26	2.29	1.73	1.90	1.90
Ni-B [8]	2.07	2.23	2.38	2.55	1.67		2.69	2.14	2.26	2.34	1.82	1.98	2.08
Ni-A [8]	2.12	2.21	2.53	2.48	1.70	2.04	2.80	2.20	2.25	2.32	1.74	1.88	1.87
[III OH]$^{2-}$	2.22	2.19	2.37	2.42	1.87		2.87	2.02	2.36	2.38	1.72	1.87	1.88
[III H1 O]$^{2-}$	2.22	2.22	2.30	2.81	1.73		2.79	1.87	2.38	2.46	1.76	1.87	1.88
[III H2 O]$^{2-}$	2.16	2.19	2.33	2.98	1.76		2.97	1.93	2.27	2.43	1.75	1.87	1.89
[III H3 O]$^{2-}$	2.26	2.22	2.26	2.91	1.74		2.86	1.89	2.37	2.44	1.76	1.87	1.88
[III H4 O]$^{2-}$	2.22	2.21	2.33	2.88	1.74		2.88	1.92	2.36	2.41	1.75	1.88	1.87
[III O]$^{3-}$	2.22	2.24	2.31	2.98	1.74		2.89	1.91	2.35	2.45	1.74	1.87	1.88
[II H1 OH]$^{2-}$	2.22	2.13	2.30	2.89	1.82		2.94	2.02	2.35	2.42	1.72	1.87	1.88
[II H2 OH]$^{2-}$	2.16	2.18	2.23	3.06	1.87		3.06	2.02	2.28	2.41	1.72	1.86	1.88
[II H3 OH]$^{2-}$	2.18	2.19	2.28	2.87	1.84		2.96	2.02	2.36	2.42	1.72	1.87	1.88
[II H4 OH]$^{2-}$	2.20	2.18	2.32	2.88	1.84		2.97	2.01	2.36	2.40	1.72	1.88	1.87
[II HOH]$^{2-}$	2.20	2.14	2.33	2.93	1.93		3.07	2.15	2.34	2.37	1.70	1.87	1.88
[III OOH]$^{2-}$	2.25	2.18	2.42	2.56	1.95	2.00	2.96	2.12	2.31	2.38	1.69	1.86	1.88
[III HOO]$^{2-}$	2.24	2.19	2.42	2.39	1.87	2.20	2.85	2.02	2.32	2.41	1.71	1.87	1.88
[III H4 OO]$^{2-}$	2.21	2.19	2.47	2.84	1.86	1.94	3.00	2.04	2.32	2.38	1.72	1.88	1.88
[III SO2 OH]$^{2-}$	2.19	2.20	3.00	2.43	1.91	1.92	3.09	2.01	2.24	2.44	1.72	1.87	1.87
[III SO3 OH]$^{2-}$	3.13	2.16	2.30	2.53	1.84	1.92	2.92	1.99	2.33	2.41	1.72	1.86	1.87
[III SO3 OH]$^{2-}$ [a]	2.21	2.20	2.35	2.42	1.85	3.04	2.86	2.01	2.34	2.40	1.71	1.86	1.88
[III $SO_2$3 H_4]$^{2-}$	2.65	2.12	2.20	3.24	2.08	1.93	3.11	2.52	2.28	2.34	1.68	1.86	1.89
[III $SO_2$3 OH]$^{2-}$	3.03	2.16	2.29	2.51	1.86	1.91	2.92	1.96	2.34	2.40	1.71	1.86	1.87
[III SO3 HOO]$^{2-}$	3.11	2.12	2.27	3.35	1.94	1.89	3.19	1.94	2.34	2.37	1.72	1.87	1.90
[II OOH]$^{3-}$	2.20	2.16	2.46	3.44	2.45	1.76	3.49	1.99	2.37	2.39	1.71	1.86	1.89
[II HOO]$^{3-}$	2.20	2.17	2.41	2.77	1.83	2.34	2.96	2.00	2.33	2.42	1.71	1.88	1.89
[II H_4 OO]$^{3-}$	2.22	2.19	2.51	2.89	1.83	2.03	3.01	2.05	2.33	2.39	1.72	1.88	1.88
[II SO2 OH]$^{3-}$	2.16	2.19	3.01	3.09	1.88	1.90	3.27	2.02	2.25	2.46	1.72	1.89	1.87
[III Se3 HOO]$^{2-}$	2.33	2.20	2.35	2.38	1.87	2.22	2.80	2.04	2.31	2.40	1.71	1.87	1.87
[III SeO3 OH]$^{2-}$	3.20	2.16	2.32	2.52	1.85	1.93	2.94	2.02	2.32	2.36	1.72	1.86	1.87
[II Se3 HOO]$^{3-}$	2.29	2.17	2.35	2.75	1.84	2.37	2.94	2.02	2.35	2.41	1.71	1.88	1.88
[II SeO3 OH]$^{3-}$	3.19	2.17	2.26	3.21	1.84	1.91	3.10	2.03	2.35	2.42	1.71	1.87	1.88

[a] Alternative structure (with S3 and not O3 coordinating to Ni), started from the crystal structures of the Ni-A and Ni-B states from *D. vulgaris Miyazaki* [8].

Table 7. Relative energies of various protonation states for structures optimised by QM/MM using the medium-sized quantum region. The various energies are the B3LYP/6-311+G(2d,2p) vacuum energy, the Mead solvation energy, and the MM energy. The total energy is the sum of the latter three components. The B3LYP/6-311+G(2d,2p) energy including protein point charges is given for comparison in the last column.

	Energy (kJ/mole)				
	Vacuum	Solvation	MM	Total	Point charges
[III OH]$^{2-}$	0	0	0	**0**	0
[III H1 O]$^{2-}$	181	-68	0	**113**	165
[III H2 O]$^{2-}$	227	-19	-4	**204**	233
[III H3 O]$^{2-}$	225	-77	4	**152**	134
[III H4 O]$^{2-}$	235	-106	2	**132**	178
[II H1 OH]$^{2-}$	0	0	0	**0**	0
[II H2 OH]$^{2-}$	34	21	3	**58**	74
[II H3 OH]$^{2-}$	44	-31	8	**21**	-7
[II H4 OH]$^{2-}$	49	-66	5	**-12**	23
[II HOH]$^{2-}$	9	31	1	**42**	66
[III OOH]$^{2-}$	281	4	-20	**266**	292
[III HOO]$^{2-}$	226	-28	-14	**183**	212
[III H4 OO]$^{2-}$	379	-106	-13	**259**	383
[III SO2 OH]$^{2-}$	0	0	0	**0**	0
[III SO3 OH]$^{2-}$	23	-22	29	**27**	-31
[III $SO_2$3 H4]$^{2-}$	313	-66	43	**290**	269
[II HOO]$^{3-}$	250	-14	7	**243**	216
[II SO2 OH]$^{3-}$	16	3	11	**30**	8
[II SO3 OH]$^{3-}$	0	0	0	**0**	0
[III Se3 HOO]$^{2-}$	210	-2	-8	**200**	205
[III SeO3 OH]$^{2-}$	0	0	0	**0**	0
[II Se3 HOO]$^{3-}$	249	-9	-2	**237**	245
[II SeO3 OH]$^{3-}$	0	0	0	**0**	0

Figure 1. Relationship between the various redox states of [NiFe] hydrogenase. All reactions are more or less reversible, but simple arrows are used for clarity. The exact difference between Ni-SU and Ni-SI_a is not known and thus no species is shown at that reaction arrow.

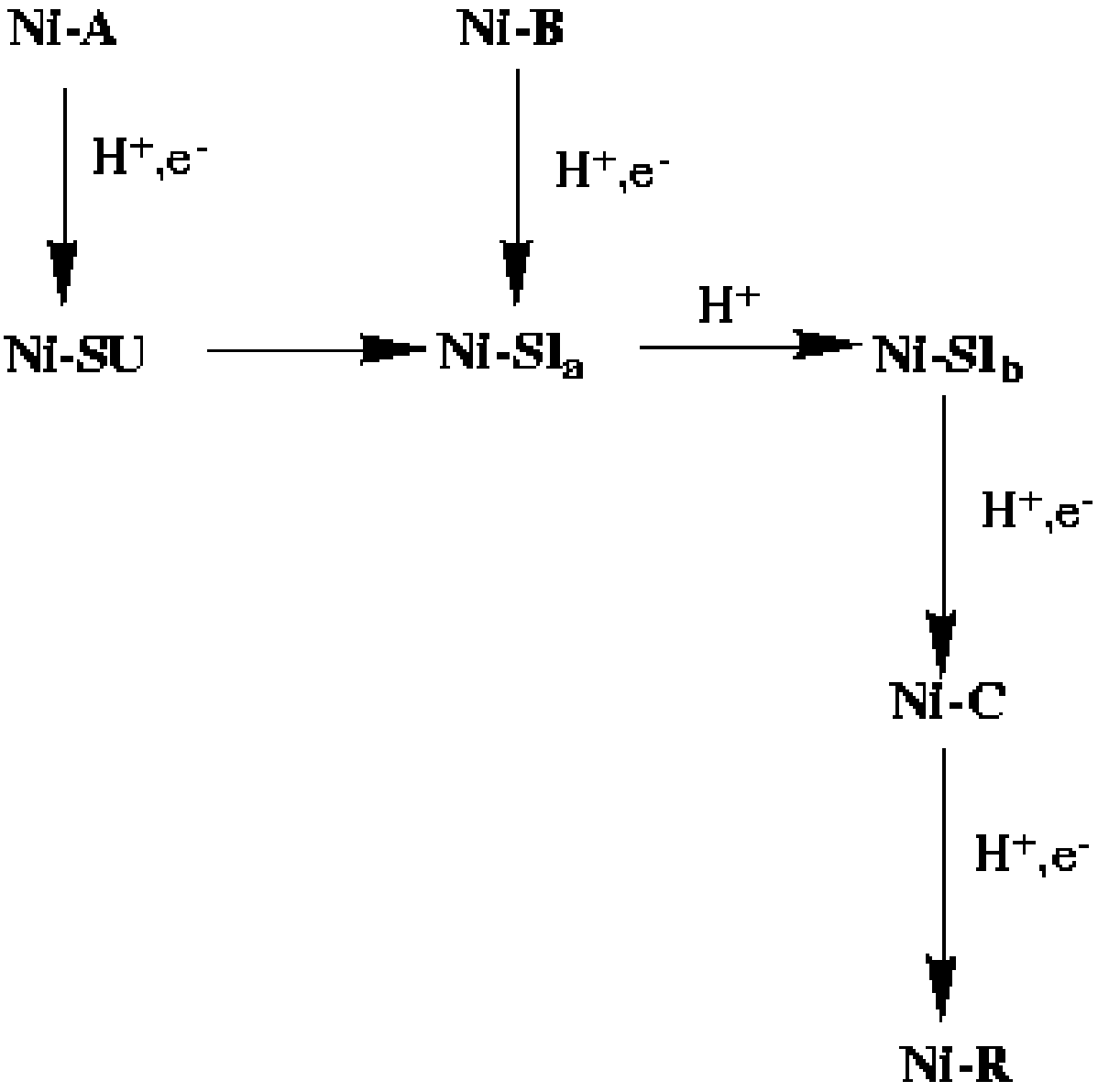

Figure 2. Active site of [NiFe] hydrogenase. The numbering of amino-acid residues is for the large subunit of *D. fructosovorans*. The corresponding numbers in *D. gigas* are 65, 68, 530, and 533.

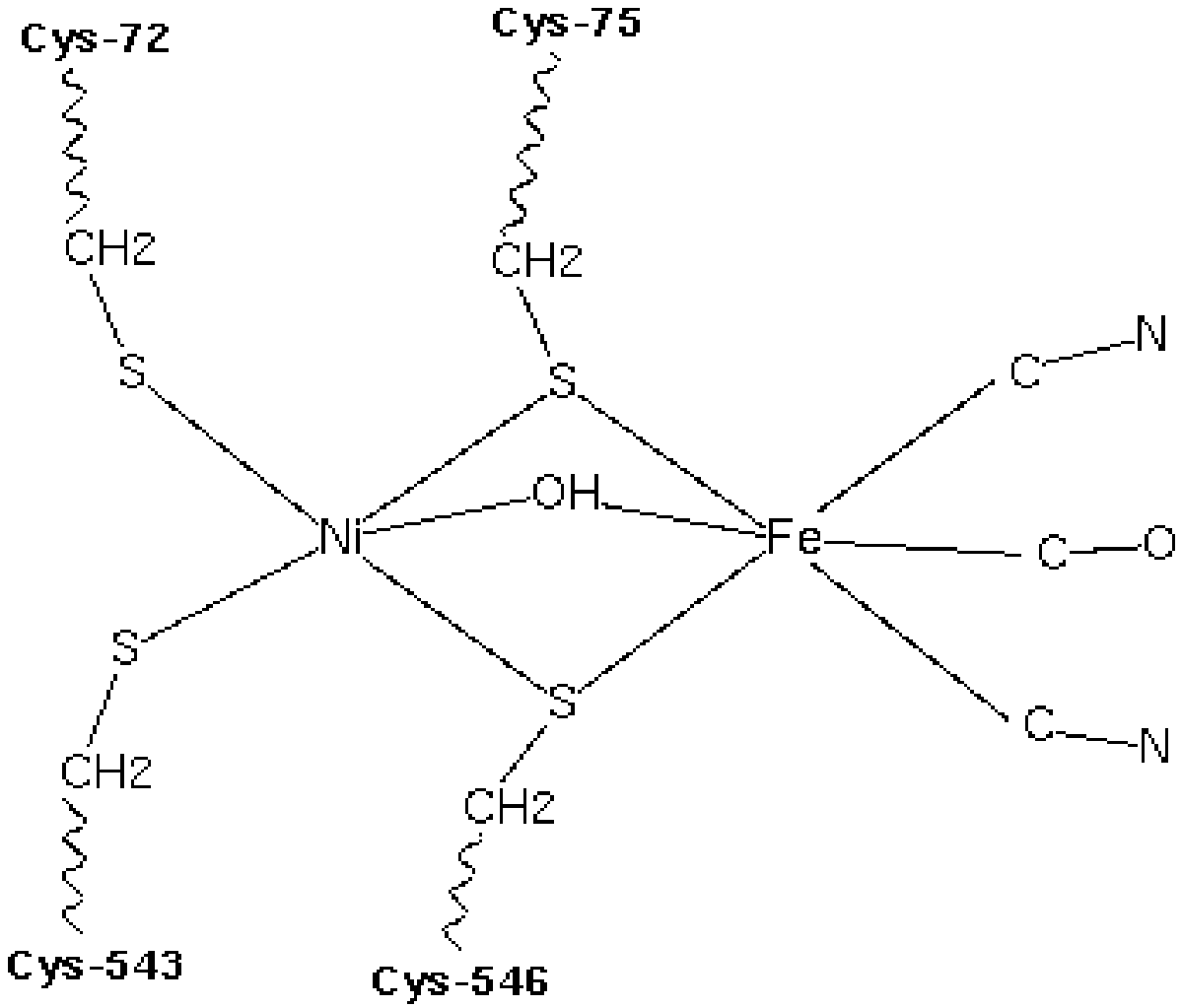

Figure 3. Medium quantum region used for QM/MM calculations. The example shown is [III OH]$^{2-}$. Polar interaction is shown as dashed lines in the structure. Waved lines indicate junctions, with the actual protein atom replaced by a hydrogen atom in the QM calculations.

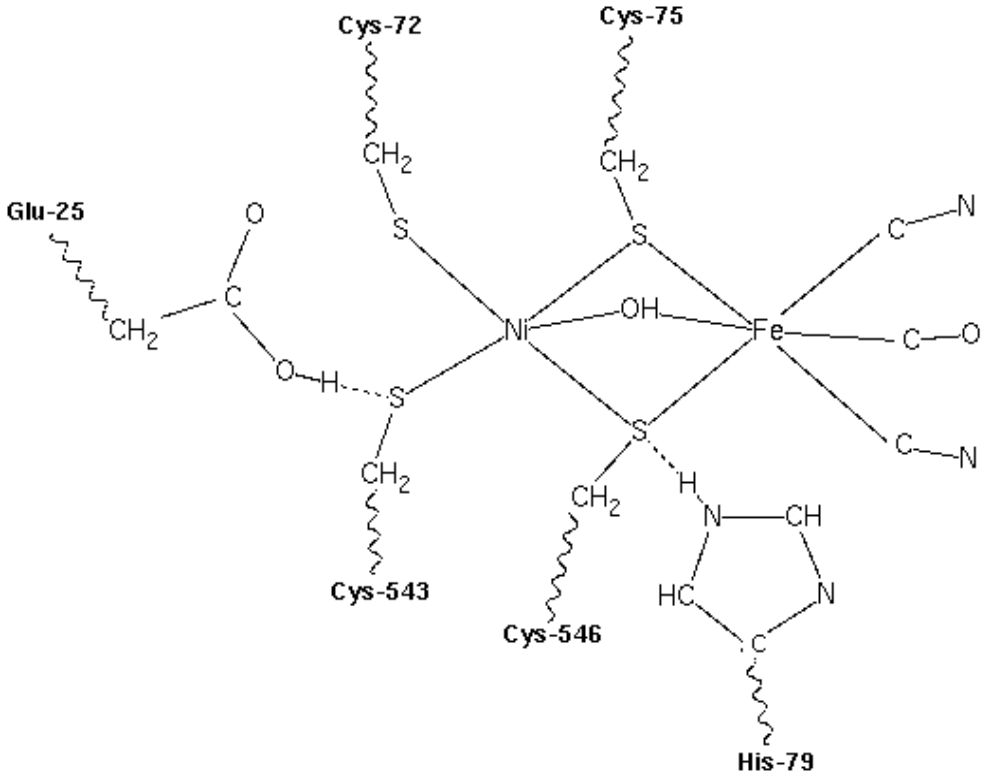

Figure 4. Large quantum region used for some of the quantum refinement calculations and vacuum calculations with fixed surroundings. The example shown is [III OH]$^{2-}$. Polar interaction is shown as dashed lines in the structure. Waved lines indicate junctions, with the actual protein atom replaced by a hydrogen atom in the QM calculations.

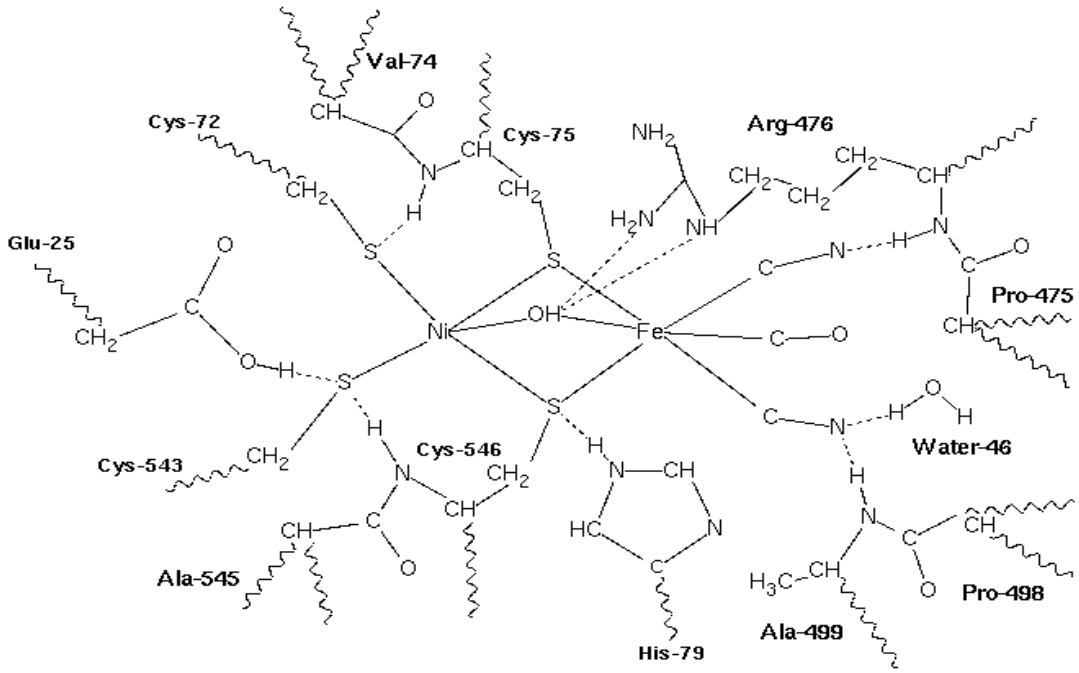

Figure 5. Model compound. Hydrogen atoms are omitted for simplicity.

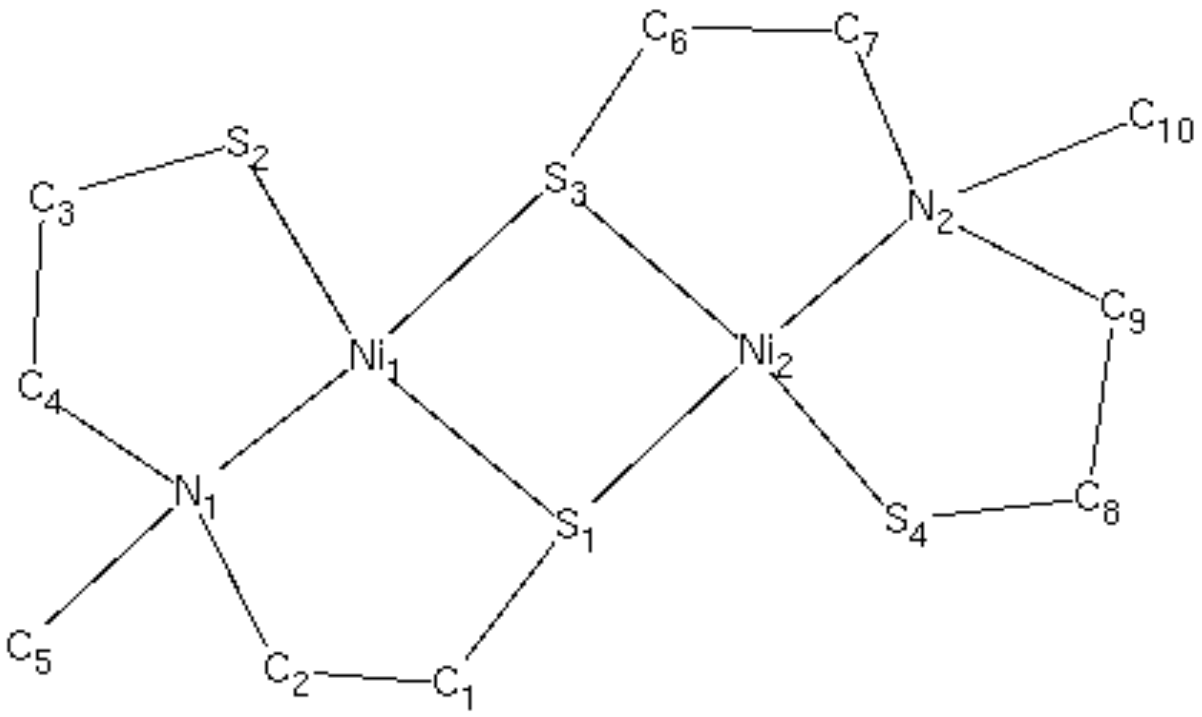

Figure 6. A comparison of the quantum-refined structures of [II OH]$^{3-}$ and [III H3 OO]$^{2-}$. Also included are the $F_o - F_c$ difference maps of the two structures at the 2.8 σ level (blue and red for the [II OH]$^{3-}$ structure and green and white for [III H3 OO]$^{2-}$). Positive densities are blue and green, negative read and white.

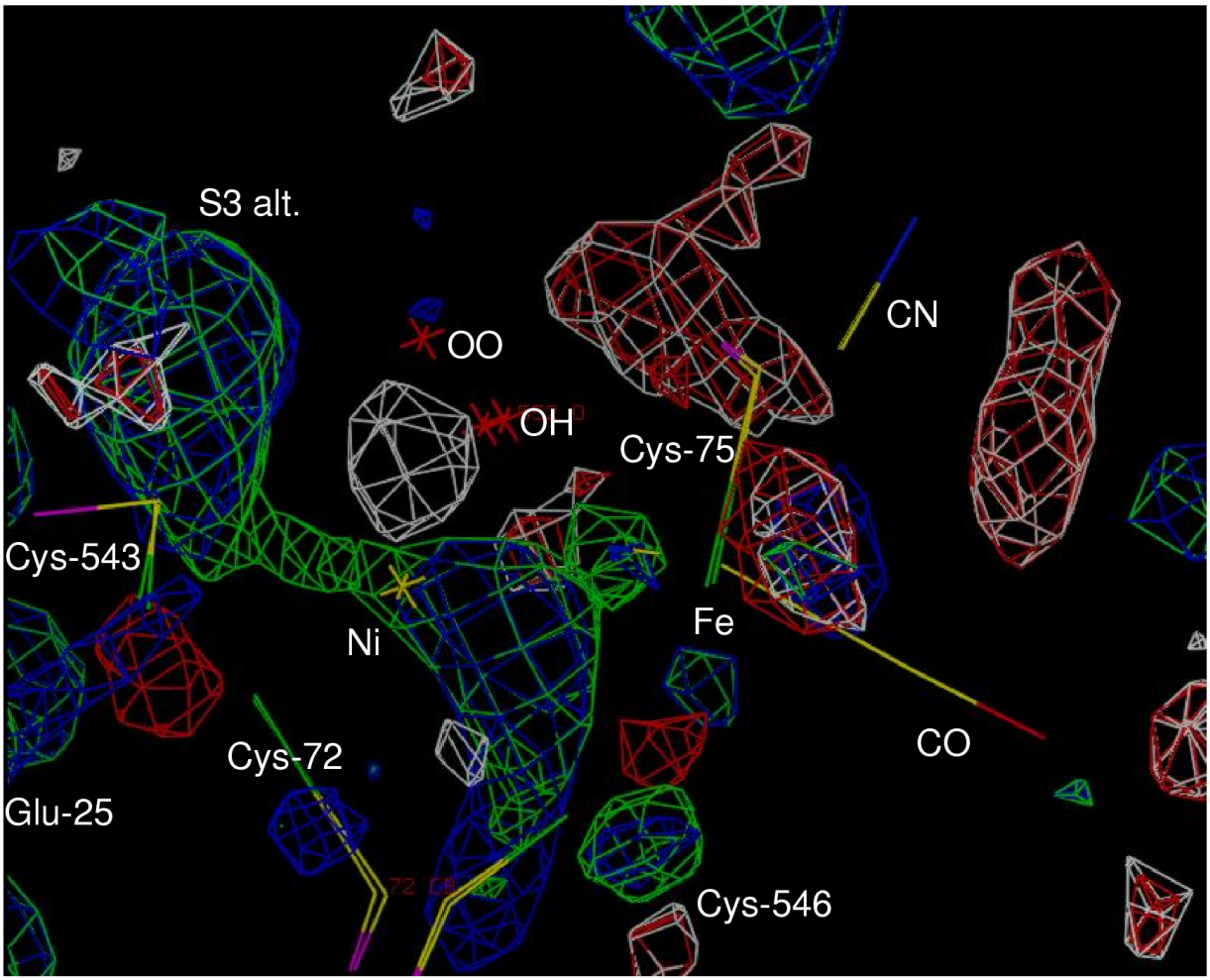

Figure 7. Optimised Ni–S2 distance and the corresponding energy for fixed values of the Ni–S4 distance. Energies are calculated at the BP/DZP level.

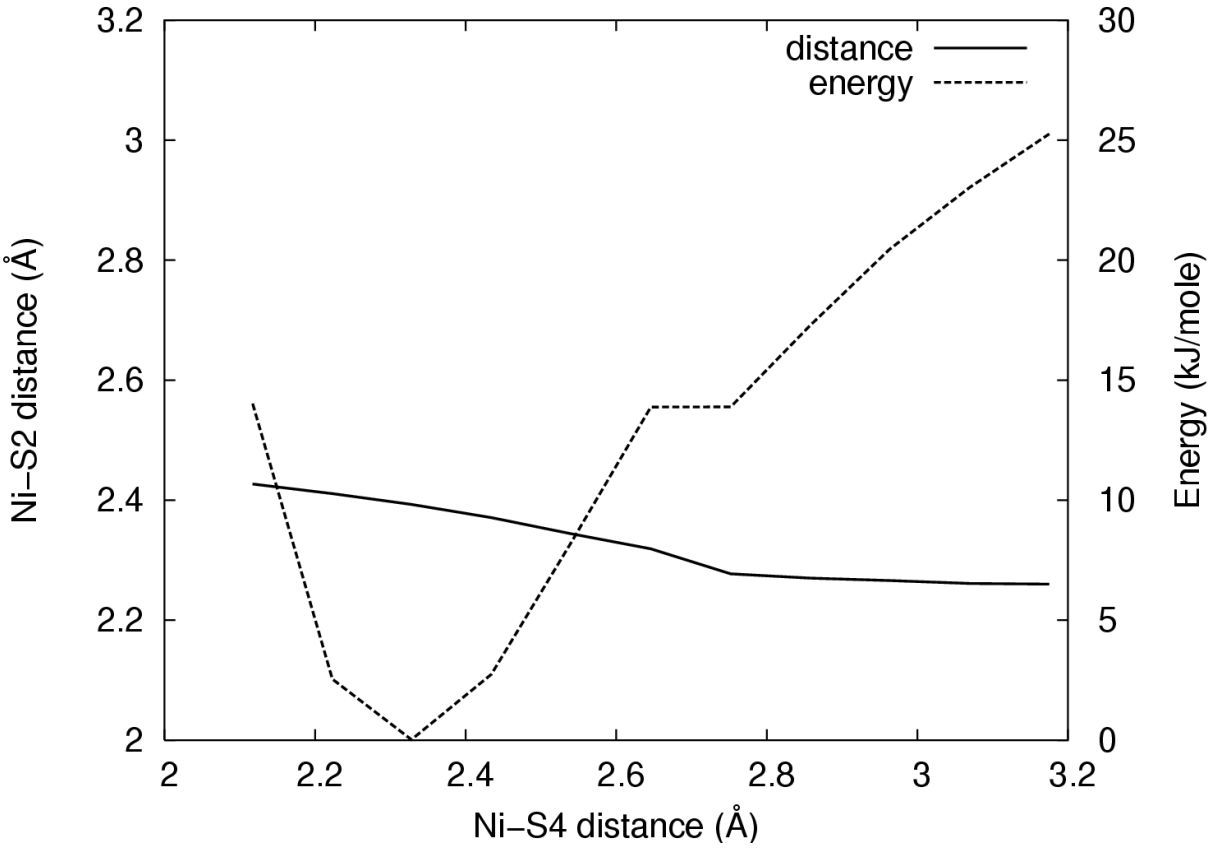

Figure 8. Five optimised structures with a bridging $(H)O_2$ ion and either Ni^{II} or Ni^{III}, compared to the crystal structures of Ni-SU from *D. fructosovorans* (green) and Ni-A from *D. gigas* (cyan) [7]. Hydrogen atoms are omitted for clarity.

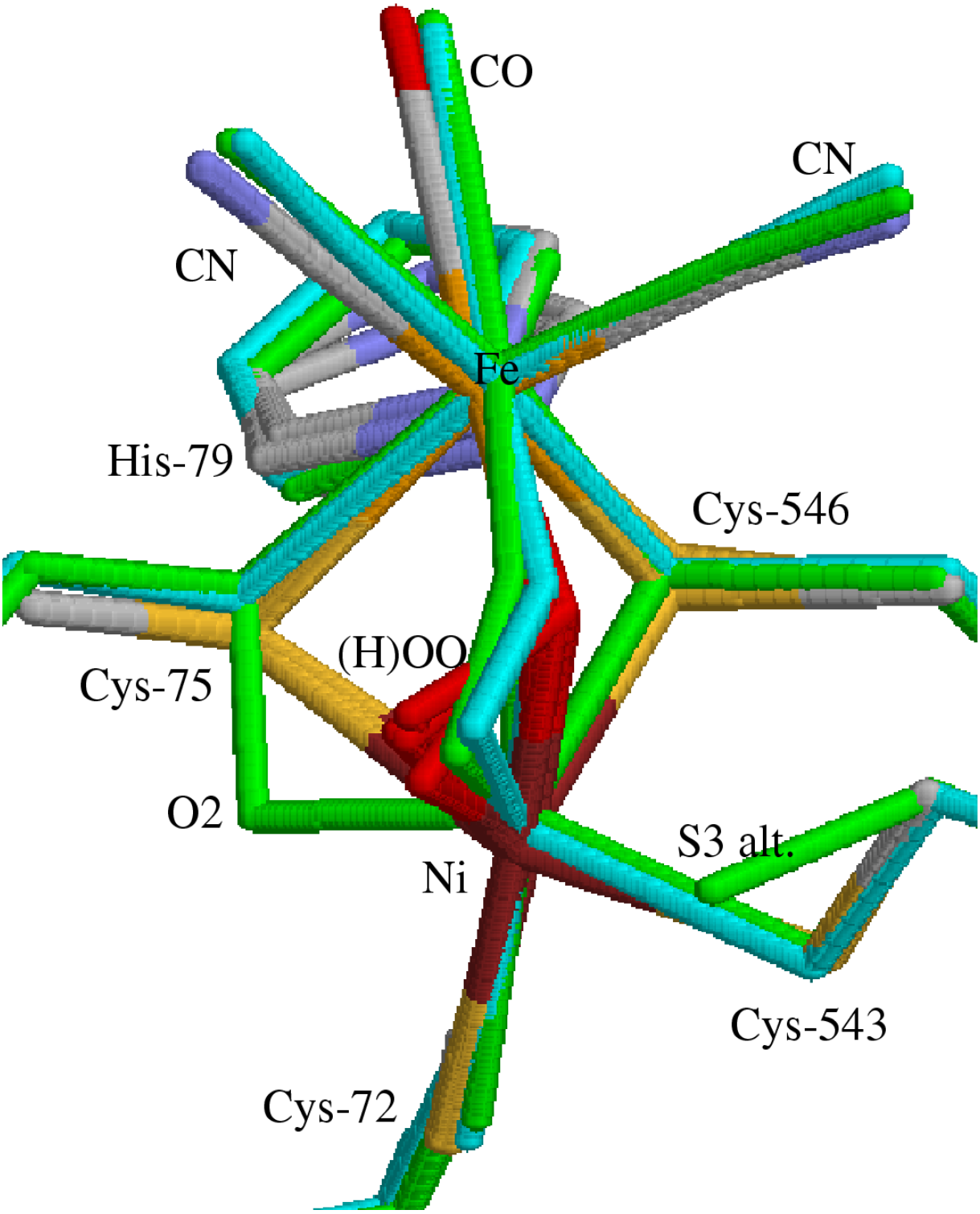

Figure 9. Two optimised structures with Cys-75 (S2) oxidised (the extra oxygen atom is marked by an arrow) and either Ni^{II} or Ni^{III}, compared to the crystal structures of Ni-SU from *D. fructosovorans* (green) [7].

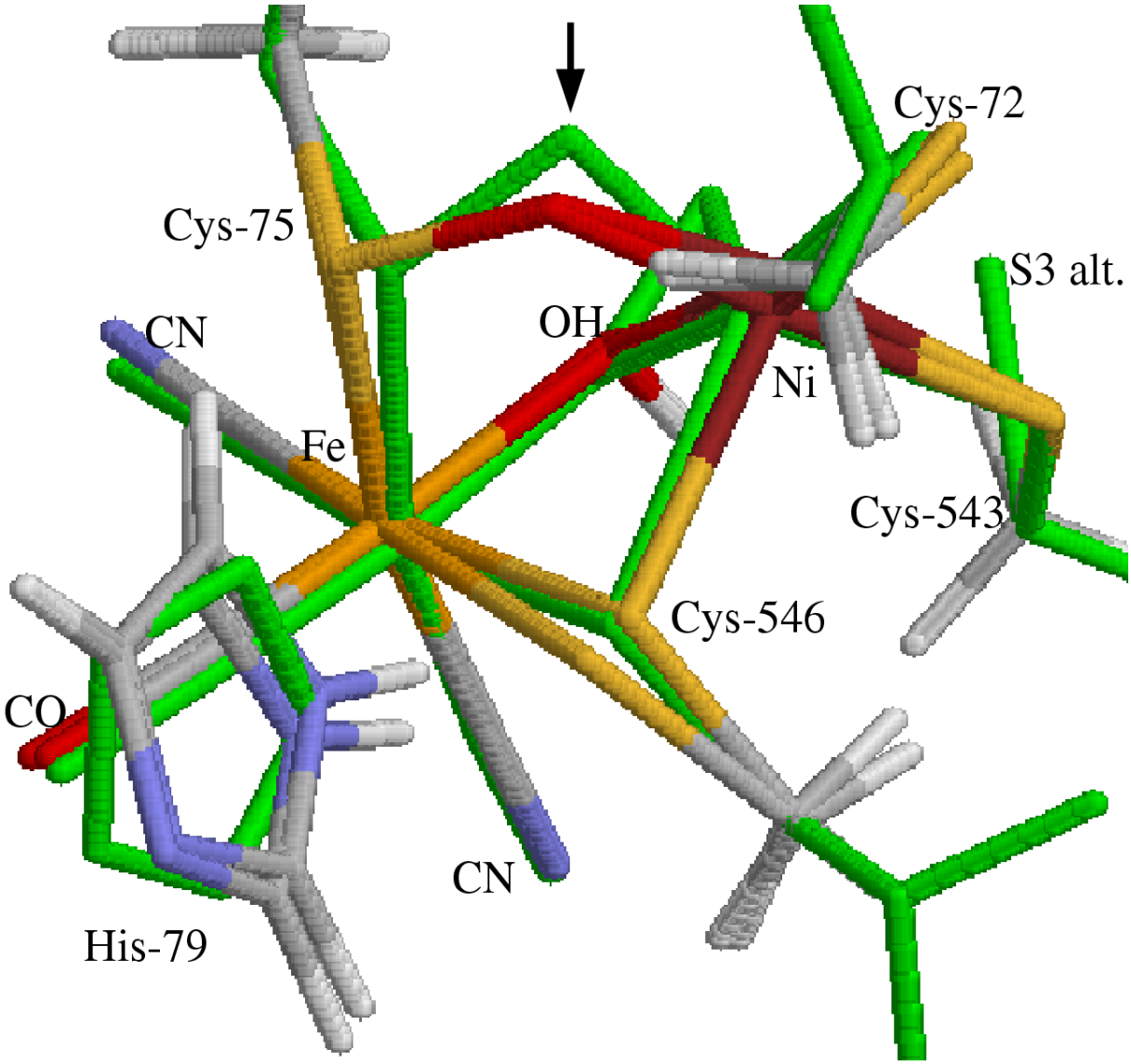

Figure 10. The optimised structure of Ni^{III} with Cys-543 (S3) oxidised and placed in the alternative position, compared to the crystal structure of Ni-SU from *D. fructosovorans* (green) [7].

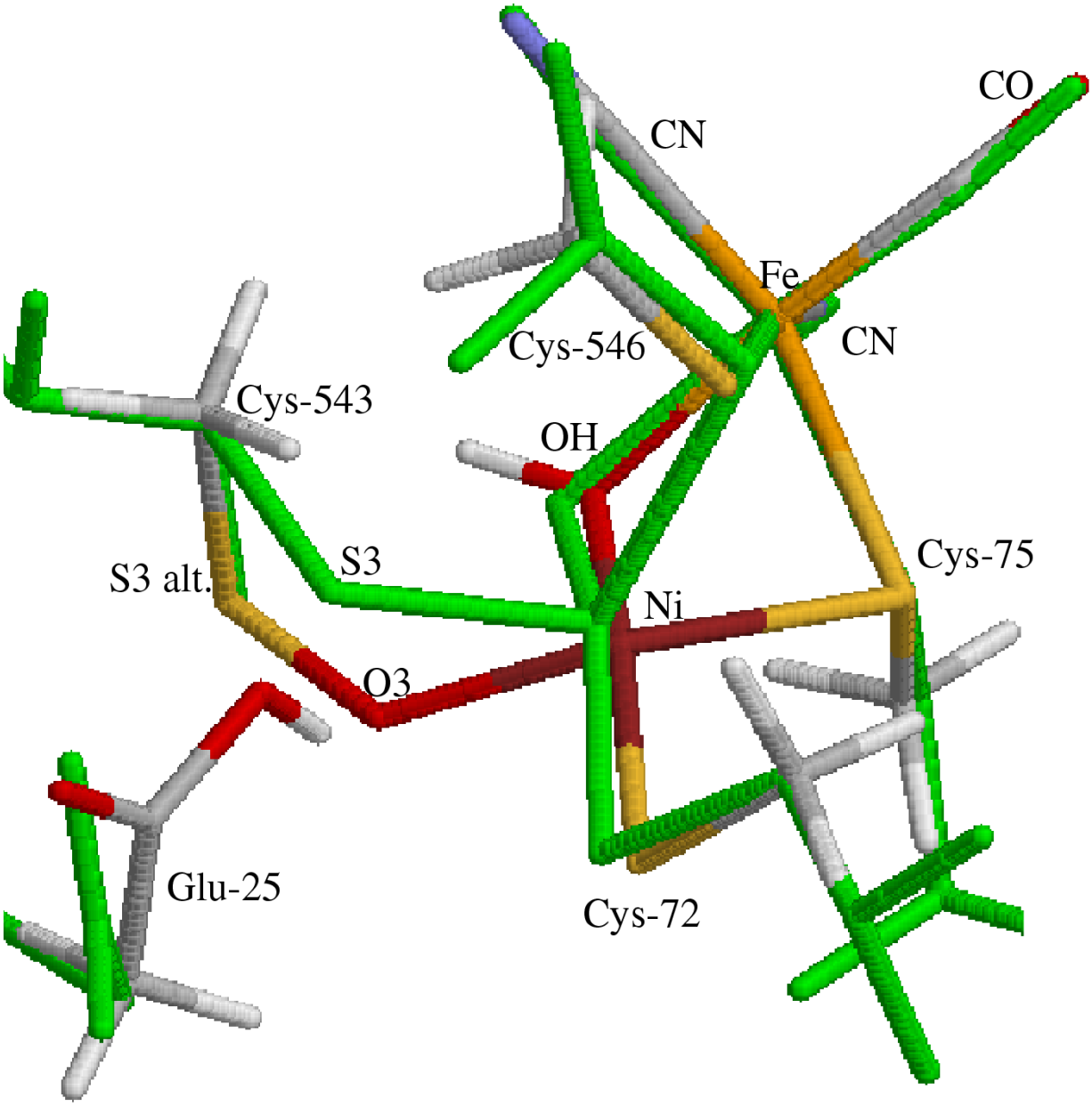

Figure 11. A comparison of the quantum-refined structures of [II OH]$^{3-}$ and [III SO3 OH]$^{2-}$ (thick sticks). Also included are the $F_o - F_c$ difference maps of the two structures at the 3.0 σ level (blue and red for the [III SO3 OH]$^{2-}$ structure and green and yellow for [II OH]$^{3-}$).

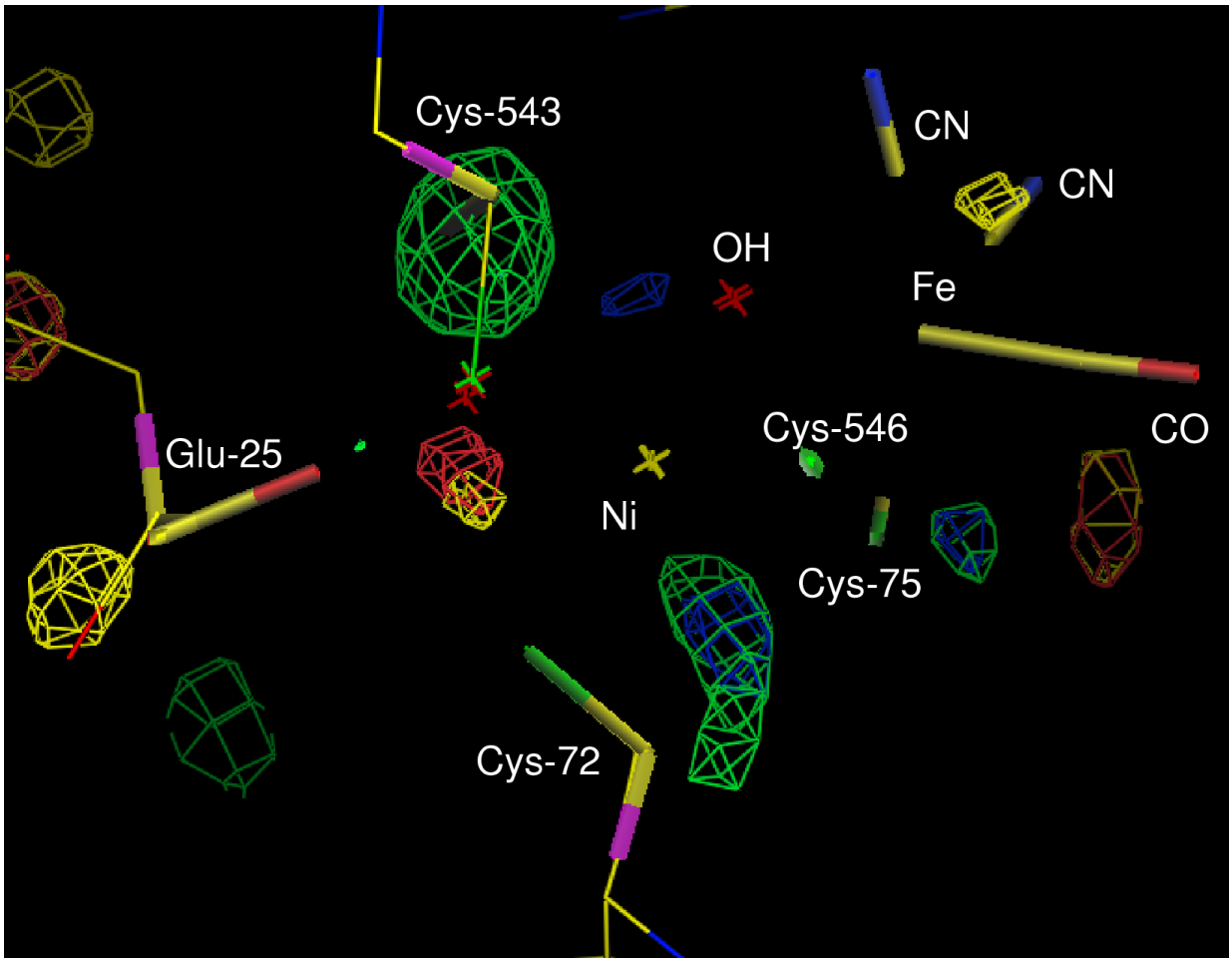

Figure 12. An overlay of the crystal structures of the Ni-A (green, 1wui) [8], Ni-B (magenta, 1wuj) [8], and Ni-SU states (cyan, 1yqw) [7] and the QM/MM structure of the [III SO3 OH]$^{2-}$ state. The figure shows that the structures agree in the positions of the oxidised S2 atom and the peroxide ligand (not present in the Ni- B and the QM/MM structures), but not in the position of the oxidised S3 atom.