Theoretical study of structure of catalytic copper site in nitrite reductase

Källrot, Niklas; Nilsson, Kristina; Rasmussen, Torben; Ryde, Ulf

Published in:
International Journal of Quantum Chemistry

DOI:
10.1002/qua.20386

2005

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

The structure of the catalytic copper site in nitrite reductase, studied by theoretical methods

Niklas Källrot, Kristina Nilsson, Torben Rasmussen, and 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

2017-04-02

The catalytic copper site in nitrite reductase contains a Cu^{2+} ion bound to three histidine (His) ligands and a solvent molecule. Sites from various sources show a conspicuous variation in the structure: In some proteins, it is close to to tetrahedral (even more than blue copper proteins), whereas in other proteins, it has a structure more similar to what can be expected for a type 2 copper site. We have studied this site with a number of theoretical methods, ranging from vacuum optimisations, combined quantum and molecular mechanics (QM/MM) optimisation, quantum refinement (X-ray crystallography supplemented by quantum chemical calculations), and accurate energy calculations. We show, that the difference in the structure arise from a movement of the solvent molecule and that this movement is determined by a compromise between its hydrogen bond interactions and the intrinsic preferences of the copper site. If the solvent molecule is deprotonated, the two structures have a similar energy, whereas if it is protonated, the more tetrahedral structure is energetically favourable. Neither of the structures involve a interaction as in the blue copper proteins; instead both are strongly distorted tetragonal structures with bonds to all four ligands. We have also examined the position of hydrogen atoms shared between second-sphere carboxylate groups and the first sphere solvent molecule and one of the histidine ligands. In the oxidised state, the structure with the solvent deprotonated but the histidine residue protonated seems to be most stable.

Keywords

density functional theory, quantum refinement, QM/MM calculations, copper, second-sphere ligands.

Introduction

Copper is one of the most important transition-metal ions in biology [1-3]. It is used for a great variety of functions, including electron transfer, oxygen activation and binding, nitrogen metabolism, and superoxide dismutase activity [4]. In biological systems, copper has two available oxidation states, Cu^{+} and Cu^{2+}. Cu^{+} is a d^{10} ion with closed electron shells and therefore prefers low coordination numbers (2–4) with a high symmetry (e.g. tetrahedral complexes with four ligands). On the other hand, Cu^{2+} is generally tetragonally coordinated in small inorganic compounds, being either square-planar, square pyramidal, or octahedral [5,6]. The reason for this is that Cu^{2+} has nine *d* electrons and therefore is Jahn–Teller unstable. For an octahedral complex, this instability can be lifted by a tetragonal distortion, elongating the bonds of two trans axial ligands, which ultimately may lead to a square-planar coordination. However, the Jahn–Teller instability can also be lifted by a trigonal distortion. This is illustrated by the fact that many Cu^{2+} complexes are trigonal pyramidal. In fact, such a geometry is often as stable as the square-pyramidal one. For example, for Cu(acetonitrile)$_5^{2+}$, the two structures are degenerate within 1 kJ/mole, according to a simple quantum chemical calculation in gas phase.

In biological systems, copper often forms sites with 1–4 copper ions [1-3]. The mononuclear sites are often tetragonal and trigonal bipyramidal structures. Such sites constitute the type 2 or normal copper sites, which have quite weak absorptions lines in the absorption spectrum ($< 1000\ M^{-1}cm^{-1}$), negative reduction potentials, and large hyperfine couplings in the electron paramagnetic resonance (EPR) spectra ($A_{\|} \sim 0.015\ cm^{-1}$).

However, there also exist numerous mononuclear copper sites with differing characteristics. The most well-known of these are the blue or type 1 copper proteins,

which are characterised by an intense blue colour (> 2000 $M^{-1}cm^{-1}$), positive reduction potentials (180–1000 mV), and narrow hyperfine couplings in the EPR spectrum ($A_{\|}$~0.007 cm^{-1}) [1-3]. In these sites, the copper ion is bound to three strong ligands (one cysteine and two histidine residues) in a trigonal plane, with one or two weaker axial ligands (methionine or amide oxygen). These sites have been extensively studied by experimental and theoretical methods, and it has been shown than the spectroscopic characteristics are mainly caused by the strong bond to the negatively charged and soft cysteine thiolate ligand [7-9]. In particular, the trigonal structure is stabilised by a strong bond between Cu and S_{Cys}.

We have used quantum chemical methods to study why some ligands prefer a tetragonal structure, whereas others prefer a trigonal structure of a four-coordinate copper complex [10]. Without any electronic effects, a complex with four equivalent ligands would prefer a tetrahedral coordination, because ligand–ligand repulsion is minimised in such an geometry. This is the reason why four-coordinate Cu^{+} complexes are tetrahedral. With electronic effects, we have seen that the Jahn–Teller distortion may give rise to either a square-planar structure or to a trigonal structure with one axial ligand. For most ligands, the former structure is most stable, because in such a structure, all four ligands can form favourable bonds to each of the four lobes of the singly-occupied Cu orbital. In the trigonal structure, one ligand forms a bond to two of the lobes, whereas the other two ligands in the trigonal plane form normal bonds to the singly-occupied orbital. However, the fourth ligand then has to interact with a fully occupied orbital and therefore becomes a weakly bound axial ligand.

We studied complexes of the type $Cu(NH_3)_3X^{2+}$, where *X* was varied in the periodic table around the simple Cys model SH^{-} [10]. It turned out that both the tetragonal and trigonal structures could be obtained for all complexes. For neutral ligands (NH_3, OH_2,

SH_2, SeH_2), the tetragonal structure was most stable by 32–42 kJ/mole. However, with negatively charged ligands, the stability of the trigonal structure increased when going down or to the left in the periodic table. Thus, F^- and OH^- gave stable tetragonal structures (by 75 and 36 kJ/mole, respectively), whereas, PH_2^- and SeH^- gave stable trigonal structures (by 5 kJ/mole). Interestingly, a Cys model alone was not enough to stabilise a trigonal structure – the tetragonal structure was 14 kJ/mole more stable than the trigonal one for $Cu(NH_3)_3(SH)^+$. However, if a model of the methionine ligand (which is an ideal weak axial ligand) was also added, i.e. $Cu(NH_3)_2(SH_2)(SH)^+$, the trigonal structure became most stable by 7 kJ/mole.

It also turned out that a trigonal structure is not necessary for the function of the blue copper proteins (electron transfer): The tetragonal structure with the blue-copper ligands is quite strongly tetrahedrally distorted (owing to a large transfer of charge from the thiolate to Cu, so that it becomes partly Cu^+). Such a structure is actually found in nature for some protein, viz. the electron-transfer site in green nitrite reductases [10-12]. Trigonal and (tetrahedrally distorted) tetragonal structures can be distinguished by the angle between the N–Cu–N and S_{Cys}–Cu–S_{Met} planes (it will be denoted in the following). It is 90 in an ideal trigonal (and also in a tetrahedral) structure, whereas it is 0 in an ideal square-planar structure. In practice, a survey of the available protein structures 1997 [10], indicated that trigonal sites had angles of 78–86 , whereas the tetragonal site in green nitrite reductases had an average angle of 61 . The tetragonal site is also characterised by two large (and two small) N–Cu–S angles of ~140 . These represent the angles of the two pairs of trans ligands, which in an ideal square-planar structure are 180 . In a trigonal structure, three angles in the trigonal plane are ~120 , whereas the other three angles are ~90 .

The survey of protein structure [10] showed that most available mononuclear copper

sites could be explained by these concepts. However, one site fell out: The *catalytic* type 2 copper site in nitrite reductase (not the electron-transfer site, discussed above) from *Achromobacter cycloclastes* had the largest angle in the survey, 88 , but it is not a blue-copper site (a recent atomic-resolution structure from *Alcaligenes xylosoxidans* shows a copper site [13]). As can be seen in Figure 1, it consists of a four-coordinate copper ion bound to one solvent molecule and three histidine (His) residues. The geometry is more tetrahedral than trigonal, with four copper angles of 102–108 and two N–Cu–O angles of 116–118 [14]. Based on the results from the $Cu(NH_3)_3X$ model, such a structure can be expected only if one of the ligands are deprotonated. Interestingly, crystal structures of a nitrite reductase from another source (*Alcaligenes faecalis*) shows a quite different structure, although the ligands are identical [15]. Thus, it has a angle of 69 and an N–Cu–O angle of 151 , indicating that the structure is tetragonal.

Nitrite reductases catalyses the one-electron reduction of nitrite to gaseous NO:

$$NO_2 + e^- + 2\ H^+ \quad NO + H_2O \qquad (1)$$

There are two types of nitrite reductases in nature, the haem-containing cytochrome cd_1 nitrite reductase (E.C. 1.9.3.2) and copper nitrite reductase (E.C. 1.7.99.3), which is found in about a third of the denitrifying bacteria. The latter enzymes contain two copper sites, a type 1 electron-transfer site, the and the catalytic type 2 site. The two sites are 12.6 Å apart, but two of their ligands, Cys-136 (type 1 site) and His-135 (type 2 site) are adjacent in sequence. The functional unit of the enzyme is a trimer, and the type 2 site is at the interface between the two subunits.

Copper nitrite reductases have two conserved carboxylate groups in the second coordination sphere of the copper ion: Glu-278 is hydrogen bonded to the (non-coordinating) N^{1} atom of one of the His ligands (His-100; the other two His ligands

form hydrogen bonds to back-bone carbonyl groups). Likewise, Asp-98 forms a hydrogen bond to the solvent molecule coordinated to the copper ion. These residues are also included in Figure 1.

The importance of charged second-sphere ligands for the function of biological metals sites has been increasingly realised [16-18]. It is frequently observed that metal-ligating His residues are hydrogen bonded to carboxylate groups. This can be seen as a way to neutralise the charge of the metal ion. However, it has in several cases been suggested that such an interaction may induce an imidazolate character onto the ligand, e.g. in haem peroxidases and vitamin B_{12} enzymes [16-19]. Recent theoretical investigations of these and other enzymes have indicated that the properties of the metal site are quite strongly modified by such a hydrogen bond, i.e. that this interaction need to be considered to understand the properties of the metal site [16-18,20]. It is likely that the same applies to metal–water–carboxylate interactions.

In this paper, we investigate the reason for the unusual structure of the catalytic copper ion in nitrite reductase and how it is related to these second-sphere interactions. We use three types of calculations to examine the structure and properties of this site: quantum chemical vacuum calculations, combined quantum mechanics and molecular mechanics calculations (QM/MM), and our recently developed quantum refinement method, which is standard crystallographic refinement, in which we have replaced the molecular mechanics potential (which normally is used in standard crystallography) for the metal site by more accurate quantum chemical calculations. Together, these calculations give a quite clear picture the reason of the unusual structure of this metal site and the positions of the protons in the carboxylate–metal-ligand links.

Methods

Quantum chemical calculations

All quantum chemical calculations were performed with the density functional three-parameter hybrid functional B3LYP method, as implemented in the Turbomole package [21,22]. These calculations employed the 6-31G* basis set for all atoms [23], except for copper, for which we used the DZP basis sets of Schäfer et al. [24,25], enhanced with *p*. *d,* and *f*-type functions with exponents of 0.174, 0.132, and 0.39. In some test calculations, diffuse functions (6-31+G*) were added to the nitrogen and oxygen atoms. This did not change the structure significantly (the Cu–ligand distance changed by less than 2 pm). Accurate relative energies were estimated for some complexes with the 6-311+G(2d,2p) basis set [23]. The basis set for copper was then enhanced with *s*, *p*, and *f* function with exponents 0.0155, 0.046199, and 3.55. The structures were optimised until the change in energy between two iterations was below 2.6 J/mole (10^{-6} a.u.) and the norm of the maximum norm of the internal gradients was below 10^{-3} a.u.

Density functional methods have been shown to give excellent geometries for transition metal complexes, with errors in the bond distances of 0–7 pm [26,27]. In particular, the B3LYP method is known to give the most accurate energies of the widely used density functionals [28-30]. Calibrations on copper complexes have shown that the geometries and energies do not change significantly if the method or the basis sets are improved from the present level [31].

The His ligands were modelled by imidazole (Im), whereas Asp-98 and Glu-279 were modelled by acetate (Ac). Thus, the catalytic copper site in nitrite reductase was modelled by either $Cu(Im)_3(H_2O)$ or $Cu(Im)_3(H_2O)(Ac)_2$. In the former model, we considered different protonation states for one of the imidazole groups (corresponding to His-100; imidazole or imidazolate, Imm) and the water molecule (H_2O or OH^-). In

the latter case, we instead moved the proton shared by Glu-279 and His-100 or by H_2O and Asp-98, giving rise to protonated acetate (HAc) and Imm or Ac+Im, and H_2O+Ac or OH^-+HAc, respectively. Thus, in both cases, we have four possible states, depending on the protonation. In both cases we will denote the variants in the following way: Wat (H_2O and Im), Hyd (OH^- and Im), Imm (H_2O and Imm), and Both (OH^- and Imm).

Solvation effects were estimated by single-point calculations using the continuum conductor-like screening model (COSMO) [32,33]. These calculations were performed at the same level of theory as the geometry optimisation and with default values for all parameters (implying a water-like probe molecule) and a dielectric constant ( ) of 4 of 80. The latter value is close to that of pure water, whereas the former is closer to what is believed to be typical for a protein [34,35]. For the generation of the cavity, a set of atomic radii have to be defined. We used the optimised COSMO radii in Turbomole (130, 200, 183, 172, and 200 pm for H, C, N, O, and Cu, respectively [36]. Unrestricted open-shell theory was employed in all calculations. Several starting structures were tried for most complexes, but only the structure with the lowest energy of each type is reported. Calculations with both point-charges and a dielectric continuum were performed with the Gaussian-98 software [67].

QM/MM calculations

Combined quantum mechanical (QM) and molecular mechanical (MM) calculations (QM/MM) [37-45] where performed with the ComQum software [40,41], which is a combination of Turbomole [21] and Amber [42]. The model $Cu(Im)_3(H_2O)(Ac)_{0-2}$ was treated by quantum chemistry, using the B3LYP method and the 6-31+G* basis set, as described above. In the QM calculations, this model was represented by a wavefunction, whereas all the other atoms were represented by an array of point charges, one for each atom, taken from the Amber libraries. Thereby, the polarisation of the quantum chemical system by the surroundings is included in a self-consistent manner. In the MM force and

energy calculations, all atoms were represented by the Amber force field, but without any electrostatic interactions (which are already treated by quantum mechanics). Special action was taken when there is a bond between the classical and quantum chemical systems (a junction) [40]. The quantum chemical system was truncated by hydrogen atoms, the positions of which are linearly related to the corresponding carbon atoms in the full system, the hydrogen link-atom method [40,43].

The total energy is calculated as

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

Here, E_{QM} is the QM energy of the quantum system, truncated by hydrogen atoms, including all the electrostatic interactions. E_{MM1} is the classical energy of the quantum system, still truncated by hydrogen atoms, but without any electrostatic interactions. Finally, E_{MM123} is the MM energy of all atoms with normal atoms for the junctions and no electrostatics. The philosophy behind this energy 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 [44]. The calculated forces are the negative gradient of this energy, but owing to the different atoms for the junction atoms, the gradients have to be corrected using the chain rule. All calculations were performed with the surrounding protein fixed at the crystal positions (because we only want to reproduce the crystal structure.

The strain energy of the copper site in the protein (E_1) is calculated as the difference in QM energy in vacuum (i.e. without point charges) of the quantum system optimised in the protein and in vacuum. It should be noted that this energy may contain terms that are not normally considered as strain [45].

Quantum refinement calculations

Quantum refinement [38,46] is essentially standard crystallographic refinement

supplemented by QM calculations for a small part of the protein. Crystallographic refinement programs change the protein model (coordinates, occupancies, *B* factors, etc.) to improve the fit of the observed and calculated structure-factor amplitudes (usually estimated as the residual disagreement, the *R* factor). Owing to the limited resolution normally obtained for biomolecules, the experimental data are supplemented by chemical information, usually in the form of a MM force field [47]. Then, the refinement takes the form of a minimisation or simulated annealing calculation by molecular dynamics using an energy function of the form

$$E_{cryst} = w_A E_{Xref} + E_{MM} \quad (2)$$

where E_{Xref} is a penalty function, describing how well the model agrees with the experimental data (we used the maximum-likelihood refinement target using amplitudes, *MLF*) [48,49]), E_{MM} is a MM energy function with bond, angles, dihedral, and non-bonded terms, and w_A is a weight factor, which is necessary because E_{MM} is in energy units, whereas E_{Xref} is in arbitrary units [50].

Quantum chemistry can be introduced in this function by replacing the MM potential for a small (but interesting) part of the protein (system 1) by a QM calculation, yielding a QM energy for system 1, E_{QM1}. To avoid double counting we must then subtract the MM energy of system 1, E_{MM1}:

$$E_{tot} = E_{QM1} - E_{MM1} + E_{MM12} + w_A E_{Xref} \quad (3)$$

Thereby, we introduce an accurate energy function for the system of interest. Such a penalty function is implemented in the program ComQum-X [46], which is a combination of the software Turbomole [21] and Crystallography and NMR system (CNS) [51]. By comparing Eqns. 1 and 3, we see that ComQum and ComQum-X are closely related: Quantum refinement can be seen as a standard QM/MM program, where the structure is restrained to be close to the experimental electron density. Therefore, we

can use a similar program for both methods and we have also used the same hydrogen link-atom method for junctions in the quantum refinement calculations. Following crystallographic custom, hydrogen atoms and electrostatic interactions were ignored in all the refinements (hydrogen atoms are of course present in the quantum chemical calculations).

COMQUM-X has been tested by re-refining the structure of *N*-methylmesoporphyrin bound to ferrochelatase [46]. The results showed that we may improve the structure locally in terms of the R_{free} factor. Moreover, we have shown [52] that refinement with COMQUM-X of a medium-resolution (170 pm) crystal structure of cytochrome c_{553} brings the geometry of the haem group and its ligands closer to that observed in an atomic-resolution structure (97 pm) of the same protein [53]. For example, the errors in the Fe–ligand distances are reduced from 3–9, 12, and 32 pm to 1, 0, and 2 pm (for the porphyrin, histidine, and methionine ligands, respectively). We have also shown that we can decide the correct protonation status of metal-bound solvent molecules with this method [54].

Energy calculations with a continuum approach

Accurate relative energies of the various protein complexes were estimated by solving the Poisson–Boltzmann equation numerically using MEAD 2.2 software [55,56}. Charges for the quantum system were obtained by fitting to Merz–Kollman charges [57] using the Gaussian-98 software [67]. In the charge calculation, the wavefunction was perturbed by protein and solvent point charges, scaled by a factor of 4. For the rest of the protein, Amber [42] partial charges were used. PARSE radii [58] were used for all atoms in the MEAD calculations. A unity dielectric constant (= 1) was used for the quantum system, whereas = 4 in the remaining protein, and = 80 in

the surroundings (water). These values have previously been shown to provide reasonable agreement with experimental results [59-62]. To this solvation energies, the energy of the quantum system was added, calculated for a wavefunction, perturbed by scaled charges of the surroundings as above, but without the point charges in the energy calculation (i.e. the wavefunction was optimised with charges, then the charges were removed and the energy was evaluated without reoptimising the wavefunction).

The MEAD calculations were based on the quantum-refined geometries without any alternative configurations. All water molecules were removed except two, which directly interact with the quantum system in all structures. The MEAD calculations were performed with 351 grid points and a spacing of 30 pm, centring the grid on the catalytic copper ion. The reported energies are the average of seven calculations in which the grid origin was moved 15 pm in positive and negative direction along each Cartesian axis. The maximum difference among the seven calculations is 2.5 kJ/mole.

The proteins

The QM/MM and quantum-refinement calculations were based on the crystal structure of oxidised nitrite reductase at pH 6.0 (Brookhaven protein data bank file 1NIC, 1.9 Å resolution, *R* factor 0.168) [14]. The protein is composed of three identical subunits (although only one subunit is present in the crystallographic unit cell), each with 340 residues and two copper ions (one type 1 blue copper ion and the catalytic type 2 copper ion).

In the QM/MM calculations, the protein was truncated: all residues in the trimer with no atoms closer than 3.0 nm of the copper ion in subunit A were deleted. This resulted in 390 amino acids, two copper ions, and 173 crystal water molecules from all three subunits of the protein. A 2.4 nm water cap (720 water molecules) centred on the copper ion was included in the calculations, leading to a total of 8687 atoms. For disordered

residues, the conformation with the lowest occupancy was deleted (conformation A for Arg-54, and conformation B for Glu-197, Arg-250, and Arg-271). A sulphate ion and disordered crystal water molecules were also removed, except DIS-1250, 1275, and 1282. The positions of the solvent water molecules and the protein hydrogen atoms were optimised by a simulated-annealing calculation, as has been described before [12]. All Asp, Glu, Lys, and Arg residues were assumed to be ionised, whereas the protonation status of the His residues were determined from a study of the solvent exposure and hydrogen-bond interactions in the crystal (His-100, 135, and 306 were protonated on the N^{1} atom, His-60, 95, and 145 on the N^{2} atom, and the rest were protonated on both nitrogen atoms). The only Cys residue is a copper ligand and was therefore assumed to be deprotonated.

Charges for the quantum-refinement method was determined in a similar way, but in this case, all atoms were included and a ball of solvent molecules with a radius of 34 Å was added, entered on the catalytic copper ion. The positions of the solvent water molecules and the protein hydrogen atoms were determined by a 100 ps simulated annealing molecular dynamics simulation. The temperature was kept at 370 K during the first 30 ps and was then decreased linearly with time to 0 K at the end of the simulation. The simulated annealing was concluded by a 10 000 step molecular mechanics minimisation. In all simulations, a non-bounded cut-off of 1.5 nm was applied, the neighbour list was updated every 25 fs, the time step was 0.5 fs, a dielectric constant = 1 was used, and no bond-length constraints were included. All heavy atoms present in the PDB files were kept fixed and no bonds between copper and the ligands were defined. The charges on the copper ions and their ligands were taken from Merz–Kollman calculations on $Cu(Im)_3(H_2O)$ or $Cu(Im)_2(SCH_3)(S(CH_3)_2)$. The software Amber 7.0 [42] with the Cornell et al. 1995 force field [63] was used.

In the quantum-refinement calculations, the full protein was treated, including all water molecules and the SO_4^{2-} ion. Coordinates, occupancies, *B* factors, and structure factors were downloaded from the Brookhaven protein data bank. From these files, we

also obtained the space group, unit-cell parameters, resolution limits, and *R* factors. One set of calculations was performed with and one without the alternatively conformations of nine residues, present in the original files.

The full geometry of the proteins was optimised, using the same convergence criteria is in the vacuum QM calculations. In each cycle of the geometry optimisation, the surrounding protein was allowed to relax by one cycle of crystallographic minimisation and one cycle of individual *B*-factor refinement. However, the new coordinates and *B* factors were accepted only if the R_{free} factor was reduced. For the protein, we used the standard CNS force field (protein_rep.param, water.param, and ion.param). For the other program parameters, we used data form the PDB files or the default choices. Residue (real-space) *R* factors [64] were calculated from $_A$-weighted maps using CNS. For the w_A factor, we used the default choice of CNS, 0.40 without and 0.21 with alternative conformations.

A set of quantum refinement calculations were also performed on the crystal structure of nitrite reductase from *A. faecalis* (PDB file 2afn, 200 pm resolution)[15]. All atoms were included in the calculations (there are no alternative conformations) and the QM calculations were performed on the copper ion in subunit C. Otherwise, all calculations were similar to those of the protein from *A. cycloclastes*.

Results and Discussion

Vacuum calculations without hydrogen bonding

In order to get a feeling about the preferred geometry of the copper site in various protonation states, we first performed a series of QM vacuum calculations of the small $Cu(Im)_3(H_2O)$ model. The results of these calculations are collected in Table 1. As for the $Cu(NH_3)_3X$ models [10], we were able to find both a tetragonal and a trigonal structure (at least with constraints to C_s symmetry) for all protonation states. These vacuum results confirm our earlier results [10]: The tetragonal structures have angles

of 9–19 , whereas the angles of the trigonal structures are 83–90 . Moreover, all the tetragonal complexes have two large (164–175) and four small angles (86-94), whereas the trigonal complexes have more similar angles (90–135). Typical examples of the two structures are shown on the left-hand side of Figures 2a–b.

The right-hand side of the same figures shows the singly-occupied orbital for the structures. In the tetragonal state, each lobe of the Cu 3*d* orbital forms a anti-bond with a *p* orbital on each of the four ligands, explaining the square-planar structure of the complex. In the trigonal Imm (and Wat) complex (Figure 2b), on the other hand, there is still two anti-bonds between the Cu 3*d* orbital and two ligand. However, the other two lobes of the Cu 3*d* orbital forms a anti-bond with two lobes of a orbital of the ligand. The fourth ligand (water) does not interact with the singly occupied Cu 3*d*, explaining why it becomes a weakly bound axial ligand. This is also completely analogous to what was found for the blue copper models [10].

This is also reflected in the spin densities of the complexes, listed in Table 2. In the tetragonal structure (Tet1), the spin density on Cu is high (0.55–0.62 *e*) and there is significant spin on all the four ligands. Negatively charged ligands have a higher density than the other ligands, and water has the lowest spin density. In the trigonal Imm complex, Cu has a significantly lower spin (0.35 *e*), the Imm ligand has a high spin (0.58 *e*), whereas water has no spin at all.

However, the trigonal structures of the Hyd and Both models are different. In these complexes (Figure 2c), the singly occupied orbital has Cu d_{z^2} character and forms a strong anti bond with OH^-. This orbital has only minor overlap with the other three ligands. Consequently, the copper ion in these states has about the same spin density as in the corresponding Tet1 states, whereas OH^- has a large spin density (0.3 *e*) and the ligand in the plane has zero spin density (Table 2).

For all models, except Imm, the tetragonal structure is most stable, by 43 (Wat) to 80 (Both) kJ/mole. For the Imm state, the trigonal structure is more stable by 11 kJ/mole. This is in line with the results for the $Cu(NH_3)_3X$ models (46, 32, and –1 kJ/mole for X = H_2O, OH^-, and NH_2^-, respectively) [10]. However, the trigonal Imm model is 82 kJ/mole less stable than the tetragonal Hyd model (they are comparable, because they contain the same number of protons). Therefore, the vacuum results indicate that all trigonal states are quite unstable. The energy does not change appreciably when a larger basis set is used (less than 9 kJ/mole), but in water solution, the tetragonal state of Imm is 22 kJ/mole more stable than the trigonal structure. All the Imm complexes are also stabilised relative to the Hyd complexes.

Interestingly, we actually were able to find *two* different tetragonal states for all structures, one nearly planar with a low angle of 9–19 (Tet1; described above and in Figure 2a), and one that was appreciably more tetrahedrally distorted, with a angle of 45–71 (Tet2; Figures 2d and e). This difference is also reflected in larger N–Cu–N/O angles in the Tet1 structure (164–173 compared to 123–156). For all protonation states, except Imm, the planar Tet1 structure was more stable by 3–18 kJ/mole (in solution, the Tet1 structure is most stable also for Imm).

The Tet2 complexes have a more complicated electronic structure. For the Imm model (Figure 2d), the singly-occupied Cu 3*d* orbital, forms anti-bonds to the two Im ligands and to the water ligand. However, only for one of the Im ligands is the geometry nearly ideal, whereas for the other two ligands, the bonds are bent (especially for water). For the fourth ligand (Imm), one lobe of a orbital overlaps with the fourth lobe of the Cu 3*d* orbital (in a -like manner), but there is also a tendency to a interaction with the other lobe of the orbital and the Cu 3*d* lobe that also interacted (non-ideally) with an Im ligand, explaining the distorted bonds to this Im and also to the

water ligand. Thus, the Imm–Cu interaction is a mixture between and . This is reflected by a low spin density on Cu and a high density on Imm, as in the Trig complex. Thus, the electronic structure of the Imm Tet2 complex is analogous to that of the rhombic blue copper proteins, e.g. in the electron-transfer (type 1) copper site in green nitrite reductase [10,12].

For the other three models (Figure 2e), the singly occupied orbital and the spin densities are quite similar to those of the Tet1 complexes. For these models, the Tet2 structure seems mainly to be caused by ligand–ligand repulsion.

The Cu–N_{Im} bond lengths in all complexes are 196–207 pm (except in the trigonal Hyd complex, where it is an axial ligand at a distance of 219 pm; cf. Table 1). They increase when the total charge of the complex decreases and they are often longer in the tetragonal than in the trigonal complexes. The Cu–O_{Wat} distances are 211–246 pm and they are longer in the trigonal complexes, where water always is an axial ligand. The Cu–N_{Imm} bond length is 191–197 pm, except in the trigonal Both structure, where it is actually an axial ligand. The Cu–O_{OH} distance is 183–191 pm and always longer in the tetragonal structures (the bond in the trigonal structures is improved when the Cu–O bond is shortened.

If we compare these structures to crystal structures of nitrite reductase from *A. faecalis* and *A. cycloclastes*, we see that the Cu–O distance in the former structure is closest to that of water, whereas that in the latter structure is closer to that of OH–, although it is 3–15 pm too long. However, the angles in the crystal structures are not reproduced by any vacuum structure.

Finally, we have also optimised the corresponding reduced (Cu^+) structures, because it is possible that the copper site is reduced during the X-ray data collection. It turned out that one or more ligands dissociated from the copper ion in all structures with a

negatively charged ligand. Thus, the only stable reduced structure was Wat and it ended up in a trigonal structure with an axial water molecule at a long distance of 247 pm. The Cu–N bonds were similar to those in the oxidised structures, 199–201 pm. Thus, this structure is not very similar to the crystal structures.

Vacuum calculations with hydrogen bonding

We have also optimised the structures of $Cu(Im)_3(H_2O)(Ac)_2$ in the various protonation states. In vacuum, these models are quite hard to optimise to reasonable structures because the carboxylate oxygen atom that is not involved in the hydrogen bond has a strong tendency to interact with the other atoms in the structure, giving rise to structures with little relevance to the protein (e.g. with the Ac group coordinating to the Cu ion). For this reason, we have not been able to obtain any Trig structures of these complexes and all Tet1 structures involve a syn coordination of the Glu-297 model and a second hydrogen bond to HN of another imidazole ligand (cf. Figure 3a). The optimised structures are described in Table 3 and the Tet2 structures of the four protonation states are shown in Figures 3b-e.

It can be seen that the hydrogen bonds have quite extensive effects on the geometry. In particular, all geometries become more similar and the differences between protonated and deprotonated ligands become less pronounced. Thus, the Cu–O distance is 195–198 pm for water (shorter than in the small models) and 188–191 for OH^- (longer than for the small models). Likewise, the Cu–N_{Imm} distance is 193–200 pm, whereas the Cu–N distance of the same ligand when protonated is 196–203 pm and the Cu–N distances of the other Im ligands are 203–210 pm. The angles around the copper ion have also changed, but the trends are less clear. The largest angles are 164–175 in the Tet1 complexes and 122–144 in the Tet2 complexes. The angles are 7–35 for

Tet1 and 46–71 for Tet2. Thus, the Cu–ligand distances of the Tet2 complexes are quite similar to those observed in the crystal structures, especially that from *A. faecalis*. However, no vacuum structure has angles similar to those of the nearly tetrahedral structure from *A. cycloclastes.*

The electronic structure of these complexes is similar to that of complexes without any acetate groups. As can be seen in Table 2, there is no spin density on the Asp model, but significant spin on the Glu model (0.02–0.04 *e*), especially for the Wat complex (0.21 *e*). The singly-occupied orbital is similar to that of the Tet2 models without any acetate models (i.e. somewhat distorted anti-bonds to all four ligands) as can be seen in Figure 3f. However, the spin densities and orbitals of the Imm model is much more similar to the other complexes, than for the small models without acetate (i.e. the Imm ligand does not form any bond to Cu). We have tried hard to obtain any bonded structure of the Imm model, but in vain.

The relative energies of the complexes have also changed. Since all complexes have the same atoms, their energies are comparable. It can be seen from Table 3 that the Tet1 complexes are 86–117 kJ/mole more stable than the corresponding Tet2 complexes. Of course this is mainly an effect of the change in the binding of the Glu-279 model and therefore has little relevance for the protein (when the carboxylate group coordinates to Cu, the structure becomes even more stable). Of more interest is the relative energy of the various protonation states. For the Tet2 structures, the Both state is 14–56 kJ/mole more stable than the other structures. The larger basis set changes these energies by less than 7 kJ/mole, but in a solvent, the energies become much closer, with all four states within 12 kJ/mole and with the Hyd state only 2 kJ/mole less stable than the Both state.

We have studied the reduced state also of these hydrogen-bonded systems. As for the small models, an imidazole ligand dissociated from the Both and Hyd states. However,

the Wat and Imm states gave converged and quite similar structures. They had longer Cu–N and Cu–O distances than the oxidised complexes (201–217 and 230–235 pm, respectively and 198 pm for the deprotonated Imm ligand). The N–Cu–N and N–Cu–O angles are 110–129 and 88–109 , respectively, and the angle is 83–85 . Thus, these structures are quite similar to the crystal structure from *A. cycloclastes*, although the Cu–O bond is much too long details of the angles are somewhat different. The reduced Wat structure is shown in Figure 3g.

QM/MM calculations

The vacuum calculations with the small $Cu(Im)_3(H_2O)$ model indicate that the catalytic copper site in nitrite reductase behaves in the same way as copper sites, i.e. that there are trigonal and tetragonal structures, the relative stability of which depends on the protonation of the ligands. However, they give no clear explanation to why a nearly tetrahedral structure is observed in nitrite reductase from *A. cycloclastes*. Instead, they indicate that the site should involve a hydroxide ion, although such an ion would stabilise the tetragonal structure by 52–80 kJ/mole, depending on whether His-100 is deprotonated or not. Yet, it is well-known that a protein may change the relative stability of various structures [65]. In particular, it can be expected that the second-sphere carboxylate groups may change the protonation as well as the geometric preferences of the copper site. However, we saw that these charged groups are hard to study in vacuum, because they interact with the copper site with only one of the carboxylate atoms and are free to form non-native interactions with the second atom. A solution to this problem is to perform the calculations with the $Cu(Im)_3(H_2O)(Ac)_2$ model using QM/MM methods, where the full protein is included in the calculations and the carboxylate groups are held in place by their link to the protein backbone and by

hydrogen bonds to other parts of the protein. The results of the QM/MM calculations are collected in Table 4.

We first performed two calculations with only $Cu(Im)_3(H_2O)^{2+}$ or $Cu(Im)_3(OH)^{+}$ in the quantum system. Interestingly, it was not possible to converge any structure with the former system (Wat). On the other hand, the Hyd model converged nicely to a tetrahedrally distorted tetragonal structure with = 74 . It had an appreciably larger variation in the Cu–N distances (197–210 pm) than the vacuum structure and a shorter Cu–O bond (183 pm). The largest angle (N–Cu–O) was 136 , compared to 146 in the vacuum structure.

Likewise, if a model of Asp-98 or Glu-279 was added, only structures with a OH^- ion converged. With only the Asp-98 model in the quantum system, an interesting structure was obtained with = 83 and one weakly bond His ligand (245 pm). However, the largest N–Cu–O angle is still 139 . Interestingly, nearly the same structure was also obtained with the full $Cu(Im)_3(H_2O)(Ac)_2$ model, which also ended up in the Hyd state, $Cu(Im)_3(OH)(Ac)(HAc)$, shown in Figure 4.

The electronic structure of this model (Figure 4b) is quite complicated and closer to the Tet2 structure of the Imm model with acetate groups than to the other purely bonded structures. Thus, Cu forms a close to ideal bond to His-306, whereas it forms a mixed bond and anti-bond to the OH^- group. His-100 interacts with a orbital and forms a mixture of a and bond to Cu. His-135 shows a very weak interaction with Cu, as is also indicated by a vanishing spin density (Table 2). Likewise, no spin is present on the two carboxylate ligands.

Thus, these two Hyd structures are good candidates for the geometry observed in nitrite reductase from *A. cycloclastes*. However, compared to this structure, the optimised structures have a too long Cu–N bond (~244 compared to 220 pm), a too

short Cu–O bond (184 compared to 194 pm), and a too large N–Cu–O angle (139 compared to 118).

Quantum refinements

In order to explain the discrepancies between the QM/MM calculations and the crystal structure and to obtain an optimum compromise between quantum chemistry and crystallography, we performed a set of re-refinements of the crystal structure of nitrite reductase from *A. cycloclastes* [14], using our recently developed quantum refinement program COMQUM-X [46]. This is essentially the standard crystallographic refinement software CNS, in which we have replaced the MM potential of the copper site with QM calculations. Thereby, we obtain an accurate description of the ideal structure of the copper site and we can therefore gauge the effect of various protonations of the copper ligands. Thus, we have re-refined the structure using the four different protonation states (Wat, Hyd, Imm, and Both) of the active-site model $Cu(Im)_3(H_2O)(Ac)_2$, and we then evaluate which of these four state fits the experimental raw data best. The calculations were performed both with and without the alternative configurations of some residues, present in the original crystal structure. The results of the quantum refinements are collected in Table 5. It can directly be seen that the calculations with alternative configurations always gave a lower R_{free} factor, showing that they improve the structure.

It turned out that we could obtain two different structures for each state, depending on the orientation of the hydrogen atom(s) on the copper-bound solvent molecule. In one of these, a hydrogen atom points towards Glu-279 (see Figure 5a; we call this Hconf2). In the other conformation (Hconf1; Figure 5c-f), the hydrogen atom instead points in the opposite direction. We also obtained a few complexes, with the hydrogen atom in an intermediate position (Hconf3, Figure 5b). Hconf2 always had a ~20

kJ/mole lower strain energy (E_1, i.e. the QM energy difference of the structure optimised in vacuum and in the protein) than Hconf1, but it also consistently gave higher *R* factors (Table 5; type 3 gave the highest *R* factors). We will therefore restrict our discussion to Hconf1, which probably is stabilised by the enzyme (the results for Hconf2 are similar and the two structures differ mainly in the position of the solvent oxygen and hydrogen atoms).

From the results in Table 5, it can be seen that all structures are quite similar. For example, the N–Cu–N angles differ by less than 3 between the various calculations. The N–Cu–O angles are also similar, but for the largest one, the variation is up to 9 . Likewise, the angles show a very minor variation, 79–84 . It is notable that this is appreciably lower than in the original crystal structure (89). This illustrates that the structure is quite well determined, especially for angles and dihedrals, the variation of which involves quite large movement of the atoms involved.

For the two His residues that are not hydrogen bonded to Glu, the Cu–N bond lengths also show a minor variation, but the bond to His-135 is always ~10 pm shorter (212–216 pm) than that to His-306 (220–225 pm). This asymmetry was present also in the original crystal structure (210 and 220 pm), but not in the recent atomic-resolution crystal structure of nitrite reductase from *A. xylosoxidans* [13]. The bond to His-100 is even shorter, as could be expected from its hydrogen bond to Glu-279. However, this bond length show a larger variation, which of course depends on the location of the proton. If it resides on the His residue, the bond length is 201–203 pm, whereas if it resides on Glu-279, the Cu–N bond becomes slightly shorter196–197 pm.

The Cu–O bond shows similar, but more even more pronounced, trends. When the proton resides on water, the bond is 190–195 pm (longer for Wat than for Imm), whereas when it resides on Asp-98, it is 184–186 pm (longer for Hyd than for Both).

Thus, Cu–O bonds are shorter than in vacuum (196–198 or 189 pm for the Tet2 complexes).

The electronic structures of the four complexes are also similar. As can be seen in Figure 5g, the OH^- group forms a distorted bond of mainly character to one of the lobes of the Cu 3*d* orbital. A orbital of the His-100 group forms a bond of mixed and (mainly) character to two lobes of the singly occupied Cu orbital. Finally, the other two His residues form distorted bonds to the last lobe of the 3*d* orbital. In the Wat and Imm complexes, the bond is stronger to His-306, as can be seen from the spin densities in Table 2, whereas in the other two complexes, the bond is slightly stronger to His-135. From this table, it can also be seen that the Wat and Hyd complexes have much spin on the Glu model, whereas in the Imm complex, the spin on His-100 is unusually high (0.36 *e*).

From Table 5, it can be seen that the Both complexes give the lowest R_{free} factors, both with and without alternative configurations. For example, with alternative configurations, the R_{free} factor is 0.1510 for the Both state, whereas it is 0.1512, 0.1514, and 0.1516 for the Hyd, Imm, and Wat complexes. Admittedly, these differences are so small that it is uncertain if they are significant (remember that R_{free} is a global factor of the whole protein, quite insensitive to changes in a single group). Therefore, we have also calculated the residue (real-space) *R* factor for the seven residues in the quantum system. Encouragingly, this factor shows the same trend, but with ~10 times larger differences, 0.038, 0.042, 0.046, and 0.047 for the Both, Hyd, Imm, and Wat states, respectively. Almost all of the difference comes from the solvent molecule alone. It can also be seen that the best ComQum-X structures also give better *R* factors than the original crystal structure, showing that we actually improve the crystal structure locally.

The strain energies (E_1) give a similar result: They are always lowest for the Both

state and highest for the Wat state, but they are lower for Imm than for the Hyd state. It can also be seen that they are consistently ~20 kJ/mole higher for the calculations with alternative configurations than for those without. This is caused by the use of different values of the w_A factor (the strain energy is very sensitive to this factor [46]).

Finally, we have also studied the reduced site for all four protonations states (only with alternative conformations). From the results in Table 5, we can see that this did not lead to any improvement in the structure: The R_{free} factors are always higher than in Hconf1 and the residue R factor (0.050–0.052) is larger than all the calculations with oxidised sites. Likewise, the strain energies are higher than for the corresponding oxidised state, except for the Wat state, which is the most stable reduced state (6 kJ/mole more stable than the Imm state). The angle is 1–3 larger for the reduced than for the oxidised state, but still 5 less than in the crystal structure. Thus, we can conclude that it is unlikely that the crystal structure of *A. cycloclastes* represents a reduced catalytic copper site.

Energy calculations

It is somewhat surprising that the QM/MM results indicate that Hyd is the only stable state in the protein, whereas the quantum refinements indicate that the Both state better fit to the crystal structure and gives lower strain energies. A possible reason for this discrepancy is that the electrostatic influence of the surrounding enzyme is included in the QM/MM calculations, but not the the quantum refinements (the electrostatics is determined mainly by the polar hydrogen atoms, but these are not seen in the crystal structure, so they could not be included in the calculations). Thus, it is possible that the Hyd structure is favoured by electrostatic interactions in the protein. In order to check this possibility, we made a number of estimates of the relative energy for the four

protonation states in the quantum refined structures. These calculations were based on the Hconf1 structures without alternative conformations (because the programs used cannot use alternative conformations). The energies are collected in Table 6.

All these energies quite clearly show that the Hyd structure is strongly stabilised by electrostatic interactions and solvation inside the protein. Thus, the Both state is ~30 kJ/mole more stable than the Hyd and Imm state and ~50 kJ/mole more stable than the Wat state both in the ComQum-X and single-point vacuum calculations based on the ComQum-X structures (Vac. in Table 6). However, if the active sites (quantum systems) are inserted in a continuum solvent with a dielectric constant of 80 (like water), all structures are stabilised relative to the Both state, and the Hyd state actually becomes most stable (by 12 kJ/mole; = 80 column). The size of the dielectric constant in a protein has been much discussed. Clearly, it is not well defined for such an inhomogeneous system such as a protein, but most authors suggest that the effective dielectric constant is 2–16, with 4 as the most commonly used value [34,35]. With this value for, the effect is smaller and the Both state becomes most stable again (by 10 kJ/mole; = 4 column).

A continuum model of the surroundings is very crude and ignore all details (e.g. hydrogen bonds) of the surrounding protein. A more satisfactorily way to model the effect of the protein is to include it as a point-charge model (i.e. one partial charge for each protein atom, including a sphere of water molecules, taken from the Amber force field). If this is done, it can be seen from Table 6 (column Ptch) that the Wat and especially the Hyd states are more stable than the Both state (by 53 kJ/mole for the latter). However, if the point charges are scaled by a factor of 4 (once again modelling a dielectric screening of the protein, the result becomes close to the original one (column Ptch/4). If we finally combine both these methods so that we model the protein by and array of point charges, scaled by a dielectric constant of 4, and immerse this in a continuum solvent, also with = 4, the Hyd state again becomes most stable by 29

kJ/mole (column Ptch+ =4).

Thus, the relative energies of the protonation states depend strongly on the treatment of the surrounding protein. We have recently developed a more stable method to estimate the effect of the protein on the energy of a geometry obtained by QM/MM methods [61]. In this, the protein is described by an array of point charges, which is immersed in a dielectric continuum with a dielectric constant that varies between the quantum system (= 1), the protein, (= 4), and the solvent (= 80). The effective solvation energy of the quantum system in the protein and surrounding solvent is found by solving Poisson's equation numerically. The charges of the quantum system are obtained from a wavefunction that is polarised by the point charges, scaled down by a dielectric constant of 4. Similar methods have successfully been used to estimate acid constants and reduction potentials of various residues and metal sites in proteins [59,66].

With such a method, we obtain solvation energies (column Mead in Table 6) that strongly favour the Hyd state, whereas the other states have similar energies. This energy should finally be added to the quantum chemical vacuum energy (preferably for a wavefunction polarised by the scaled point charges, column SCF1) to give the total energy in the last column in Table 6. It shows that the Hyd state is slightly (8 kJ/mole) more stable than the Both state, whereas the other two states are 26–36 kJ/mole less stable. This is our best estimate of the relative energies of the various complexes.

The differential stabilisation of the various complexes can be quite well understood by calculating the contribution from each residue from the interaction of an unscaled point-charge model of the protein with the quantum chemical charges of the quantum system, employed in the Mead calculations. Such a calculation reproduces the Mead solvation energy within 6 kJ/mole. It shows that the main difference between the Hyd

and Both complexes comes from the protein (41 out of 48 kJ/mole), in particular from Lys-262 (38 kJ/mole), the side-chain of which forms a hydrogen bond to Glu-279. For the other two complexes, the Both state is favoured by the protein (by 19–26 kJ/mole), but this is more than compensated by the solvent (32–33 kJ/mole).

Finally, we have made an attempt to run ComQum-X calculations calculations, including electrostatics. The simplest way to do this is to include in all QM calculations an array of point charges, one for each atom of the surrounding protein. We employed the scaled Amber charges used above (Table 6, column Ptch/4). If the protein is not allowed to relax, this array can be fixed and the same in all calculations, making its implementation very simple. However, such an approach did not change the results much. The results in Table 5 (no alternative configurations, Hconf1, footnote [c]) show that the R_{free} factor is in general improved, whereas the strain energy increases by ~4 kJ/mole. The metal–ligand bonds change by less than 2 pm and the angles by less than 3 . Most importantly, the relative energy of the various model did not change significantly from the calculations without point charges, but this could be anticipated from the results in Table 6 (column Ptch/4). We also tried to run calculations with unscaled point charges, but in all cases except one, the calculations crashed owing to exaggerated electrostatic interactions. Work is in progress in our laboratory to design and test a stable method to include the electrostatic interactions with the surrounding protein in the quantum refinement calculations.

Comparison with other crystal structures

Since the survey of crystal structures containing copper 1997 [10], many new structures of nitrite reductase have been published: In January 2004, there were 40 structures in the PDB databank. Seventeen of these had a fully occupied catalytic site, without any mutations of the ligands and with a solvent molecule bound to the copper

ion. If we plot the angle of these versus the maximum N–Cu–N angle, two groups of structures can clearly be seen (Figure 6): one with large angles (82–89) and small maximum N–Cu–N angles (111–125 ; we will call this the 1nic type in the following) and one with smaller angles (65–77) and larger maximum N–Cu–N angles (134–162). We will call these two types of sites 1nic and 2afn, respectively, after the typical crystal structures, cited above and included in the Tables [14,15]. Interestingly, the type of structure is not fully correlated to the bacterial source of the protein. Although all structures from *A. cycloclastes* are of the 1nic type, there are structures of both types from the other two bacteria *(A. faecalis* and *A. xylosoxidans*). This indicates that the structure of the site is more related to experimental conditions, e.g. pH, than to the architecture of the protein.

In fact, the difference between the two configurations is almost entirely restricted to the position of the solvent molecule, whereas the copper ion and the amino-acid ligands are at essentially the same positions, as can be seen in Figure 7. This movement of the solvent molecule strongly weakens its hydrogen bond to Asp-98: the O–O distance becomes 308–351 pm, compared to 261 pm in the other crystal structure. Interestingly, the electron-density difference map of the 2afn crystal structure from *A. faecalis*, also included in Figure 7, shows a strong peak of residual positive density close to the position of the solvent molecule in the 1nic *A. cycloclastes* crystal structure. This indicates that this molecule is disordered in the structure and that it actually may occupy two different positions.

Therefore, we have also performed a series of quantum refinement calculations based on the 2nic *A. faecalis* structure [15]. The results of these calculations are shown in Table 7. Interestingly, we were able to obtain the 2afn-type structure, as well as the 1nic-type structure. This illustrates the disorder shown in Figure 7; the only difference

between the two structures is the position of the water molecule. For all states, the 2afn conformation gave the lowest value of R_{free} as well as the lowest residue R factor (the main difference is in the R factor of the solvent molecule, which is 0.09–0.10 for the 2afn conformation, but 0.22–0.23 for the 1nic conformation). This shows that the 2afn conformation is clearly dominating in the crystal structure.

For the Hyd and Both states, both conformations could be obtained without any constraints and they had similar (strain) energies, the 1nic being more stable by 8–10 kJ/mole. However, for the Wat and Imm states, only the 1nic was obtained without any constraints and it was 79–90 kJ/mole more stable than the 2afn conformation. This indicates that the HOH–Ac hydrogen bond (which is essentially broken in the 2afn structure but nearly ideal in the 1nic structure) is appreciably stronger than the HO–HAc bond. Thus, the preferred structure seems to be determined by the competition between this hydrogen bond and the intrinsic stability the copper ligand sphere (where the 2afn structure is quite close to the ideal vacuum Tet2 structure).

Finally, we have also tested a reduced structure (only for the Wat state), but as for the other protein, this structure fits the crystal data worse than the oxidised 2afn structure in terms of the R_{free} and residue R factors. Moreover, it gives a low value of the angle (71).

Conclusions

We have studied the structure of the catalytic (type 2) copper site in nitrite reductase with a series of different quantum chemical methods, viz. vacuum optimisations, QM/MM calculations, quantum refinement, and estimates of the relative energies of the various protonation states. These have given us a quite complete description of this site and some explanation of its unusual structure.

First, we have seen that the isolated site behaves the same way as blue copper proteins and related sites. Thus, all models could be found in both a trigonal and a tetragonal structure. In fact, we could even obtain two different tetragonal structure, one almost perfectly square planar and one quite strongly tetrahedrally distorted. The three structures differed in the angle and also in the largest N–Cu–N and O–Cu–N angles.

However, when the second-sphere carboxylate groups were added to the models, the structures became more similar and we could not obtain the trigonal structure. The Both protonation state is most stable, but the Hyd state is almost as stable in solution. Yet, none of the geometries were very similar to that found in the protein.

The QM/MM results are very clear, showing that stable structures are found only if the Cu-bound solvent molecule is deprotonated. However, if the structures are optimised with restraints to the crystallographic data (quantum refinement), all four protonation states can be found again, all with similar geometries. If no electrostatic effects are included, the Both state once again seems to be most stable by ~30 kJ/mole. However, it is clear that the Hyd state is strongly stabilised by electrostatic and solvation effects in the protein. Our best calculations indicate that it is actually ~10 kJ/mole more stable than the Both state.

Finally, we saw that there are two different states observed in crystal structures, one with a large angle and a relatively small largest N–Cu–N angle and one with a smaller angle and a larger N–Cu–N angle. These two structures are related by a 160 pm movement of the solvent molecule, with the other atoms almost fixed. Both structures are obtained with proteins from various sources and the solvent molecule seems to be disordered in some structures (both conformations are occupied). The movement is connected with a weakening of the hydrogen bond to Asp-88.

The results in this paper allow us to address three important questions regarding the

structure of nitrite reductase. The first is why the catalytic type-2 copper site have two qualitatively different structures in various crystal structures, illustrated in Figure 6. We have shown that both sites appear to be present in a conformational equilibrium in some structures (Figure 7) and that they are related to each other by a simple movement of only the solvent molecule. Moreover, we have seen that the two structures have similar energy for some protonation states (Hyd and Both), even when calculated in vacuum, indicating that small changes in the protein may stabilise either of the structures. Therefore, it is possible that the two types of structures reflect movements of the protons shared by the first- and second-sphere ligands of the copper ion (the 2afn-type structures do not have a deprotonated solvent molecule). Apparently, the position of the solvent molecule is determined as a compromise of the hydrogen bond to the Asp-98 residue and the ideal structure of the copper site.

Second, we have investigated the nature of these two structures. The 2afn structure is similar to the Tet2 structure, obtained also in vacuum. Thus, it is a tetrahedrally distorted -bonded tetragonal structure. However, the 1nic structure was not obtained in vacuum. We have shown that this structure is not caused by a to the copper ion as in the blue copper proteins (i.e. it is not trigonal in that sense). Instead, this site clearly has a strongly distorted tetragonal structure with bonds to all ligands, although some of the bonds are distorted and have both a bonding and anti-bonding character. It seems to be stabilised by the strong hydrogen bond between Asp-98 and the solvent molecule. Thus, the second-sphere ligands are important to obtain this structure. The reason why it was not found in vacuum is that the carboxylate groups form very strong hydrogen bond within the structure with both oxygen atom and therefore tend to distort the structure from the one obtained in the protein. However, we have shown that the two structures are close in energy, even when compared in a vacuum, showing that no significant strain

need to be built up in the protein.

Finally, we have looked upon the actual protonation of the catalytic copper ion in vivo. We have shown that for the oxidised state, the Hyd protonation is most stable in the protein, although the Both protonation state is more stable in vacuum. Thus, the Hyd state should be used in future theoretical investigations of the reaction mechanism of this protein. For the reduced state, the Wat state seems to be most stable.

In conclusion, we have seen that the protonation of metal ligands are important for the properties of the metal site and that second-sphere ligands can strongly modify these. Therefore, these second-sphere ligands have to be taken into account in the accurate modelling of metal sites. We have also seen that the combination of various theoretical methods can give us some insight into the structures and energies of the active sites of proteins.

Acknowledgements

This investigation has been supported by grants from the Crafoord Foundation and the Swedish research council, and by computer resources of Lunarc at Lund University. TR gratefully acknowledges financial support from the Carlsberg Foundation.

References

1. R. H. Holm, P. Kennepohl, E. I. Solomon (1996) Chem. Rev. 96, 2239-2314.
2. A. G. Sykes (1990) Adv. Inorg. Chem. 36, 377-408.
3. E. T. Adman (1991) Advan. Protein. Chem. 42, 145-197.
4. W. Kaim, B. Schwederski, Bioinorganic chemistry: Inorganic elements in the chemistry of life., John Wiley & Sons, Chichester, 1994, p. 27.
5. F. A. Cotton, G. Wilkinson (1988) Advanced Inorganic Chemistry, Wiley, New York.
6. J. J. R. Fraústo da Silvia, R. J. P. Williams (1991) The biological chemistry of the elements.
7. D. W. Randall, D. R. Gamelin, L. B. LaCroix, E. I. Solomon (2000) J. Biol. Inorg. Chem. 5, 16-32.
8. H. B. Gray, B. G. Malmström, R. P. J. Williams (2000), J. Biol. Inorg. Chem. 5, 551-559.
9. U. Ryde, M. H. M. Olsson, B. O. Roos, J. O. A. De Kerpel, K. Pierloot (2002), J. Biol. Inorg. Chem. 5, 565-574.
10. M. H. M. Olsson, U. Ryde, B. O. Roos & K. Pierloot (1998) "On the relative stability of tetragonal and trigonal Cu(II) complexes with relevance to the blue copper proteins". *J. Biol. Inorg. Chem.* , **3**, 109-125.
11. L. B. LaCroix, S. E. Shadle, Y. Wang, B. A. Averill, B. Hedman, K. O. Hodgson, E. I Solomon (1996) J. Am. Chem. Soc. 118, 7755-7768.
12. K. Pierloot, J. O. A. De Kerpel, U. Ryde, M. H. M. Olsson, B. O. Roos (1998) J. Am. Chem. Soc. 120, 13156-13166
13. Ellis, M. J.; Dodd, F. E.; Sawers, G.; Eady, R. R.; Hasnain, S. S. J. Mol. Biol. 2003, 328, 429-440.

14. E. T. Adman, J. W. Godden, S. Turley (1995) J. Biol. Chem. 270, 27458-27474.

15. Murphy, M. E.; Turley, S.; Kukimoto, M; Nishiyama, M.; Horinouchi, S.; Sasaki, H.; Tanokura, M.; Adman, E. T. Biochem. 1995, 34, 12107-12119.

16. L. De Santis, P. Carloni (1999), Proteins, Struct. Funct. Gen. 37, 611.

17. K. P. Jensen, U. Ryde (2003) Mol. Phys. 101, 2003-2018.

18 K. P. Jensen, U. Ryde (2002) J. Mol. Struct. (Theochem), 585, 239-255.

19. Gajhede M (2001) In: Messerschmidt A, Huber R, Poulos T, Wieghart K (eds) Handbook of Metalloproteins. J. Wiley & Sons, Chichester, pp 195-209

20. P. Rydberg, E. Sigfridsson & U. Ryde (2004) J. Biol. Inorg. Chem., 9, 203-223.

21. R. Alrichs, M. Bär, M. Häser, H. Horn, and C. Kölmel, Chem. Phys. Lett., 162 (1989) 165.

22. R. H. Hertwig and W. Koch, Chem. Phys. Lett., 268 (1997) 345.

23. Hehre, W. J; Radom L.; Schleyer P. v. R.; Pople J. A. In Ab initio molecular orbital theory, Wiley-Interscience, New York, **1986**.

24. A. Schäfer, C. Huber & R. Ahlrichs, J. Chem. Phys. 100 (1994) 5829.

25. A. Schäfer, H. Horn, and R. Alrichs, J. Chem. Phys., 97 (1992) 2571.

26. E. Sigfridsson, M. H. M. Olsson and U. Ryde, J. Phys. Chem. B, 105 (2001) 5546-5552.

27. M. H. M. Olsson and U. Ryde, J. Am. Chem. Soc., 123 (2001) 7866-7876.

28. P. E. M. Siegbahn and M. R. A. Blomberg, Annu. Rev. Phys. Chem., 50 (1999) 221–249.

29. P. E. M. Siegbahn and M. R. A. Blomberg, Chem. Rev., 100, (2000) 421–437.

30. C. W. Bauschlicher, Chem. Phys. Lett., 246 (1995) 40.

31. U. Ryde and M. H. M. Olsson, Intern. J. Quant. Chem., 81 (2001) 335-347.

32. Klamt A, Schüürmann J (1993) J Chem Soc Perkin Trans 2 5:799-805

33. Schäfer A, Klamt A, Sattel D, Lohrenz JCW, Eckert F (2000) Phys Chem Chem Phys 2:2187-2193

34. Sharp, K. A., 1990, Annu. Rev. Biophys. Biophys. Chem., **19**, 301.

35. Honig, B., 1995, Science, **268**, 1144.

36. Klamt A, Jonas V, Bürger T, Lohrenz JCW (1998) J Phys Chem 102:5074-5085

37. Mulholland, A. J. In Theoretical biochemistry – processes and properties of biological systems (Theoretical and computational chemistry, Vol. 9). Eriksson, L. A., ed., Elsevier, Amsterdam; **2001**, 597.

38. U. Ryde & K. Nilsson (2003) J. Mol. Struct. (Theochem) 632, 259-275.

39. G. Monard, K. M. Merz (1999) Acc. Chem. Res. 32, 904—911.

40. Ryde U (1996) J Comp-Aided Mol Design 10: 153–164

41. Ryde U, Olsson MHM, (2001) Intern J Quant Chem, 81: 335–347

42. Case DA, Pearlman DA, Caldwell JW, Cheatham TE, Ross WS, Simmerling CL, Darden TA, Merz KM, Stanton RV, Cheng AL, Vincent JJ, Crowley M, Ferguson DM, Radmer RJ, Seibel GL, Singh UC, Weiner PK, Kollman PA (1997), Amber 5, University of California, San Francisco

43. Reuter N, Dejaegere A, Maigret B, Karplus M (2000) J Phys Chem A 104, 1720-1735.

44. Svensson M, Humbel S, Froese RDJ, Matsubara T, Sieber S, Morokuma K (1996) J Phys Chem 100: 19357

45. Ryde U (2002) "On the role of covalent strain in protein function", in Recent Research Developments in Protein Engineering, Research Signpost, Trivandrum, pp. 65-91.

46. U. Ryde, L. Olsen & K. Nilsson (2002) J. Comp. Chem., 23, 1058-1070.

47. G. J. Kleywegt, T. A. Jones" (1998) Acta Cryst. D54, 1119-1131.

48. N. S. Pannu, R. J. Read (1996) Acta Cryst. A52, 659-668.

49. P. D. Adams, N. S. Pannu, R. J. Read, A. T. Brunger (1997) Proc. Natnl. Ac. Sci. US 94, 5018-5023.

50. A. T. Brünger (1997) Meth. Enzym. 243, 269-277.

51. Brünger, 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**.

52. U. Ryde & K. Nilsson (2003) J. Am. Chem. Soc., 125, 14232-14233.

53. Benini, S.; González, A.; Rypniewski, W. R.; Wilson, K. S.; Van Beeumen, J. J.; Ciurli, S.; *Biochemistry*, **2000**, *39*, 13115.

54. K. Nilsson & U. Ryde (2004) "Protonation status of protein ligands can be determined by quantum refinement", J. Inorg. Biochem, accepted.

55. Bashford, D. In Y. Ishikawa, R. R. Oldehoeft, J. V. W. Reynders, M. Tholburn, eds., Scientific Computing in Object-Oriented Parallel Environments, Lecture Notes in Computer Science, vol. 1343, pp. 233-240, Berlin, 1997, Springer.

56. Tishmack, P. A., Bashford, D., Harms, E., Van Etten, R. L. (1997) Biochem. 36, 11984.

57. Besler BH, Merz KM, Kollman PA (1990) J Comput Chem 11: 431.

58. Sitkoff, T. Sharp, B. Honig, (1994) J. Phys. Chem. 98:1978—1988.

59. G. M. Ullmann, L. Noodleman, D. A. Case (2002) J. Biol. Inorg. Chem., 7, 632-639

60. D. Bashford, D. A. Case, C. Dalvit, L. Tennant, P. E. Wright, (1993) Biochem., 32, 8045-8056.

61. T. Rasmussen, K. Nilsson, U. Ryde, submitted to J. Am. Chem. Soc.

62. L. Olsen, T. Rasmussen, L. Hemmingsen, U. Ryde "Binding of benzylpenicillin to

metallo- -lactamase studied by theoretical methods", submitted to J. Phys. Chem.

63. W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman (1995) J. Am. Chem. Soc. 117, 5179-5197.

64. T. A. Jones, J.-Y. Zou , W. E. Cowand, M. Kjeldgaard (1991) ActaCryst. A47, 110119.

65. U. Ryde (2003) Curr. Opin. Chem. Biol., 7, 136-142.

66. Li J, Fisher CL, Konecny R, Lovell T, Bashford D, Noodleman L (1999) Inorg. Chem. 38, 929-939.

67. M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, V.G. Zakrzewski, J.A. Montgomery, R.E. Stratmann, J.C. Burant, S. Dapprich, J.M. Millam, A.D. Daniels, K.N. Knudin, M.C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski,G.A. Petersson, P.Y. Ayala, Q. Cui, K. Morokuma, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J. Cioslowski, J.V. Ortiz, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P.M.W. Gill, B.G. Johnson, W. Chen, M.W. Wong, J.L. Andres, M. Head-Gordon, E.S. Replogle, J.A. Pople 1998. Gaussian 98, Revision A.9, Gaussian, Inc. Pittsburgh PA.

Table 1. Result of vacuum calculations with the $Cu(Im)_3(H_2O)$ model in various protonation states. For all structures, two tetragonal geometries were obtained, one with a small and one with a larger angle, Tet1 and Tet2.

State	Geo	Rel. energy[b]	Distances to Cu (pm)		Angles around Cu ( )		
	metry	kJ/mole	N	O	N–N	N–O	( )
Wat	Tet1	0.0	199,199,199	207	93,93,173	86,88,171	10
	Tet2	2.9, 2.7, 8.1	197,199,199	211	97,98,153	90,91,142	45
	Trig[c]	42.7, 37.6, 55.6	196,196,198	226	111,122,122	95,100,100	86
	Red	–	199,199,201	247	116,118,123	91,101,105	86
Hyd	Tet1	0.0	204,204,206	187	92,93,164	88,90,170	19
	Tet2	7.5, 9.8, 4.5	205,205,205	187	97,97,147	91,93,146	46
	Tet2[a]	–	203,204,207	186	99,99,136	89,98,144	51
	Trig[c]	51.8, 48.8, 68.8	202,202,219	183	105,105,109	96,119,119	83
Imm	Tet1	105.9, 96.9, 38.4	193,199,201	216	93,94,172	86,89,173	9
	Tet2	93.1, 86.8, 70.1	196,202,206	226	109,110,128	93,94,123	71
	Trig	81.8, 74.1, 60.6	191,201,203	241	105,116,135	90,95,102	86
Both	Tet1	0.0	197,206,207	191	92,92,167	88,89,175	13
	Tet2	18.3, 19.5, 1.9	195,208,210	188	97,98,137	94,91,156	48
	Trig[c]	80.1, 77.2, 72.8	206,206,209	183	98,98,101	111,122,122	90
Crystal	2afn[d]		199–223	208–226	97–116	91–156	66–73
	1nic[e]		198,210,220	194	102,105,106	108,116,118	89
	1oe1[f]		196,200,200	198	103,109,111	110,111,112	88

[a] Optimised with diffuse functions in the basis set (6-31+G*).

[b] The three reported energies (kJ/mole) are calculated with the 6-31G* basis set, with the larger 6-311+G(2d,2p) basis set, and in water solution (= 80) and the small basis set.

[c] Constrained to have C_s symmetry.

[d] Nitrite reductase from *A. faecalis* [15].

[e] Nitrite reductase from *A. cycloclastes* [14].

[f] Nitrite reductase from *A. xylosoxidans* at atomic resolution (104 pm) [13].

Table 2. Spin densities of the various optimised complexes.

Calculation	State	Geo	Spin density *(e)* on						
			Cu	O-ligand	His1	His2	His3	Asp	Glu
Small, Vacuum	Wat	Tet1	0.62	0.05	0.14	0.10	0.09		
		Tet2	0.66	0.03	0.14	0.10	0.08		
		Trig	0.69	**0.00**	0.12	0.12	0.08		
	Hyd	Tet1	0.57	**0.22**	0.08	0.08	0.05		
		Tet2	0.57	**0.26**	0.07	0.06	0.04		
		Trig	0.57	**0.33**	0.05	0.05	**0.00**		
	Imm	Tet1	0.55	0.03	**0.27**	0.08	0.07		
		Tet2	**0.31**	0.00	**0.65**	0.03	0.01		
		Trig	**0.35**	**0.00**	**0.58**	0.04	0.03		
	Both	Tet1	0.60	0.14	0.10	0.09	0.08		
		Tet2	0.62	0.18	0.08	0.07	0.06		
		Trig	0.60	**0.32**	0.04	0.04	**0.00**		
Large, Vaccum	Wat	Tet1	0.60	0.11	0.11	0.08	0.09	0.00	0.00
		Tet2	0.55	0.08	0.09	0.03	0.04	0.00	**0.21**
	Hyd	Tet2	0.61	0.18	0.11	0.05	0.05	0.00	0.02
	Imm	Tet2	0.65	0.10	0.16	0.00	0.05	0.00	0.04
	Both	Tet2	0.61	0.17	0.13	0.00	0.05	0.00	0.04
ComQum	Hyd1		0.64	**0.20**	0.07	0.03	0.06	0.00	0.00
	Hyd2		0.62	**0.23**	0.11	0.00	0.03	0.00	0.00
	Hyd		0.62	**0.23**	0.11	0.00	0.03	0.00	0.00
ComQum-X	Wat		0.43	0.03	0.08	0.00	0.01	0.01	**0.44**
	Hyd		0.53	**0.15**	0.04	0.02	0.00	0.00	**0.26**
	Imm		0.55	0.06	**0.36**	0.00	0.01	0.01	0.01
	Both		0.63	**0.18**	0.16	0.00	0.00	0.00	0.02

Table 3. Result of vacuum calculations with the $Cu(Im)_3(H_2O)(Ac)_2$ model in various protonation states. Relative energies are in kJ/mole.

State	Geo	Rel.	Distances to Cu (pm)		Angles around Cu ( )		
	metry	energy	N	O	N–N	N–O	( )
Wat	Tet1[ab]	30.6	203,203,203	195	89,92,164	91,94,162	24
	Tet2[b]	142.1	197,205,208	198	101,101,132	101,102,122	71
	Red	5.3	201,206,212	230	114,114,122	89,100,107	83
Hyd	Tet1	13.1	203,204,205	190	89,92,175	88,92,174	7
	Tet2	129.8	197,205,209	188	96,96,141	88,103,144	46
Imm	Tet1	9.6	199,203,204	195	92,93,155	92,94,154	35
	Tet2	101.1	193,204,207	196	99,104,137	99,102,122	68
	Red	0.0	198,206,217	235	110,114,129	88,98,109	85
Both	Tet1	0.0	200,203,205	191	90,92,172	87,92,172	10
	Tet2	86.5	195,206,210	189	95,96,142	88,104,143	50
Crystal	2afn[c]		199–223	208–226	97–116	91–156	66–73
	1nic[d]		198,210,220	194	102,105,106	108,116,118	89
	1oe1[e]		196,200,200	198	103,109,111	110,111,112	88

[a] Fixed O–H bond length (110 pm).

[b] Fixed N–H bond length (107 pm).

[c] Nitrite reductase from *A. faecalis* [15].

[d] Nitrite reductase from *A. cycloclastes* [14].

[e] Nitrite reductase from *A. xylosoxidans* at atomic resolution (104 pm) [13].

Table 4. Result of the QM/MM calculations with three different models, 1: $Cu(Im)_3(OH)^+$, 2: $Cu(Im)_3(OH)(HAc)$ (HAc hydrogen bonded to the OH^- ion), and 3:$Cu(Im)_3(H_2O)(Ac)_2$. The other protonation states did not give any converged QM/MM structures.

State	Model	E_1	Distances to Cu (pm)		Angles around Cu		
			N	O	N–N	N–O	
Hyd	1	109.8	197,201,210	183	93,98,112	103,112,136	74
Hyd	2	113.5	197,202,245	184	101,102,103	102,106,139	83
Hyd	3	210.1	197,203,244	184	101,102,103	102,107,139	83
Crystal	1afn[a]		199–223	208–226	97–116	91–156	66–73
	1nic[b]		198,210,220	194	102,105,106	108,116,118	89

[a] Nitrite reductase from *A. faecalis* [15].

[b] Nitrite reductase from *A. cycloclastes* [14].

Table 5. Result of the COMQUM-X re-refinements of the crystal structure of nitrite reductase from *A. cycloclastes* [14]. The calculations were run with the $Cu(Im)_3(H_2O)(Ac)_2$ model in various protonation states. E_1 is in kJ/mole. The Hconf variants are described in the text and in Figure 5.

State	Alt. conf	H conf	R_{free}	residue R	E_1	Distances to Cu (pm) N	Distances to Cu (pm) O	Angles around Cu N–N	Angles around Cu N–O	
Wat[ab]	No	1	0.15483	0.046	155.0	201,213,221	194	100,108,110	105,108,125	82
		1[c]	0.15460	0.046	159.0	201,213,221	192	99,108,112	106,106,126	81
		2	0.15484	0.047	135.1	201,213,221	193	99,108,110	105,107,127	81
	Yes	1	0.15156	0.047	173.1	200,214,222	195	100,108,110	105,107,125	82
		2	0.15160	0.049	151.5	201,214,222	193	99,108,111	104,106,128	80
		Red	0.15158	0.050	138.6	203,212,224	204	100,109,110	108,113,118	84
Hyd	No	1	0.15469	0.042	136.5	203,213,223	186	99,107,111	108,109,122	83
		1[c]	0.15457	0.038	139.2	203,212,223	184	99,108,112	107,109,122	82
		2	0.15471	0.043	112.4	201,214,222	185	99,107,110	107,109,124	82
	Yes	1	0.15120	0.042	158.0	203,212,225	186	99,108,111	107,110,122	83
		2	0.15123	0.045	130.4	201,215,222	185	99,107,111	106,108,126	81
		Red	0.15133	0.000	172.7	205,215,226	197	98,106,107	107,115,122	84
Imm[b]	No	1	0.15468	0.046	133.2	197,214,221	192	102,108,111	105,106,125	81
		1[c]	0.15450	0.045	137.1	197,213,221	192	101,107,112	104,107,125	81
		2	0.15474	0.048	105.7	196,214,220	190	101,107,111	103,106,128	80
	Yes	1	0.15141	0.046	150.3	197,213,223	193	102,108,113	105,105,125	81
		2	0.15156	0.051	120.6	196,215,221	190	101,108,112	102,105,129	79
		Red	0.15145	0.052	145.1	203,213,225	205	101,108,110	107,114,117	84
Both	No	1	0.15450	0.041	105.6	197,214,224	184	102,107,112	108,108,120	83
		1[c]	0.15440	0.042	106.7	199,215,223	185	101,107,112	107,107,123	82
		2	0.15461	0.043	75.6	196,215,222	184	101,107,110	105,109,124	82
	Yes	1	0.15100	0.038	126.4	197,214,225	185	102,107,113	107,109,120	82
		2	0.15109	0.044	93.6	196,216,223	184	101,107,111	104,107,125	81
		3	0.15176	0.046	106.6	196,211,225	184	99,105,109	101,114,128	84
		Red	0.15121	0.050	182.1	205,216,227	197	99,106,108	106,117,121	83
Cryst.	2afn[d]					199–223	208–226	97–116	91–156	66–73
	1nic[e]		0.15140	0.045		198,210,220	194	102,105,106	108,116,118	89
	1oe1[f]					196,200,200	198	103,109,111	110,111,112	88

[a] Fixed O–H bond length (110 pm).
[b] Fixed N–H bond length (107 pm).
[c] Run with an array of scaled (by a factor of 4) point charges.
[d] Nitrite reductase from *A. faecalis* [15].
[e] Nitrite reductase from *A. cycloclastes* [14].

[f] Nitrite reductase from *A. xylosoxidans* at atomic resolution (104 pm) [13].

Table 6. Relative energies of the four protonation states in the ComQum-X structures in the Hconf1 state and no alternative conformations. The various energy terms are explained in the text.

State	Relative energies (kJ/mole)									
	ComQum-X	Vac.	= 80	= 4	Ptch	Ptch/4	Ptch+ = 4	Mead	SCF1	Total
Wat	51.9	49.3	12.8	37.0	-0.2	51.4	11.1	-19.9	54.1	34.2
Hyd	28.2	30.8	-11.6	9.7	-52.8	24.6	-29.3	-48.0	40.4	-7.6
Imm	32.4	27.5	12.4	18.7	39.2	30.6	36.6	-1.6	27.5	25.9
Both	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table 7. Result of the COMQUM-X calculations with the $Cu(Im)_3(H_2O)(Ac)_2$ model in various protonation states in the crystal structure of *A. faecalis* [15]. E_1 is in kJ/mole.

State	Conf	R_{free}	residue	E_1	Distances to Cu (pm)		Angles around Cu		
			R	kJ/mol	N	O	N–N	N–O	
Wat[a]	2afn	0.15521	0.044	224.6	203,211,220	211	98,103,114	94,100,147	71
	1nic	0.15529	0.062	145.9	205,215,222	191	98,105,112	102,105,132	83
	Red	0.15525	0.046	138.6	205,220,224	217	100,104,114	91,102,146	71
Hyd[b]	2afn	0.15530	0.044	136.2	203,218,219	184	97,101,115	95,101,149	68
	1nic	0.15530	0.060	126.6	205,216,222	186	98,105,111	104,106,130	84
Imm[b]	2afn	0.15529	0.042	203.7	199,219,220	207	100,104,114	94,99,146	71
	1nic	0.15536	0.060	113.8	199,214,222	188	100,105,111	99,104,134	81
Both	2afn	0.15533	0.042	91.2	199,217,219	183	98,101,115	94,98,151	69
	1nic	0.15537	0.060	82.9	199,215,223	185	100,105,111	101,104,132	82
Cryst.	2afn[c]				199–223	208–226	97–116	91–156	66–73
		0.15508	0.041						
	1nic[d]				198,210,220	194	102,105,106	108,116,118	89
	1oe1[e]				196,200,200	198	103,109,111	110,111,112	88

[a] Fixed O–H bond length (110 pm).

[b] Fixed N–H bond length (107 pm).

[c] Nitrite reductase from *A. faecalis* [15].

[d] Nitrite reductase from *A. cycloclastes* [14].

[e] Nitrite reductase from *A. xylosoxidans* at atomic resolution (104 pm) [13].

Figure 1. The active site of nitrite reductase from *A. cycloclases* [14], showing the tetrahedral structure and second-sphere interactions with carboxylate groups.

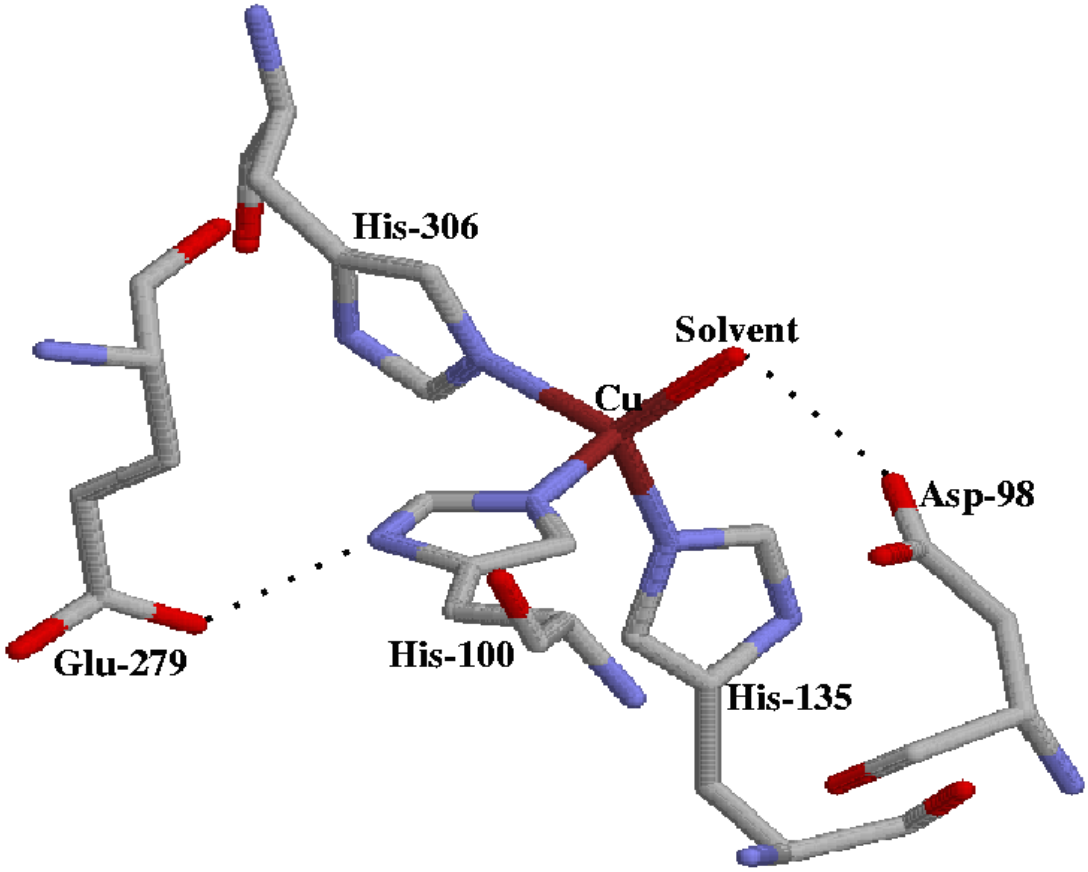
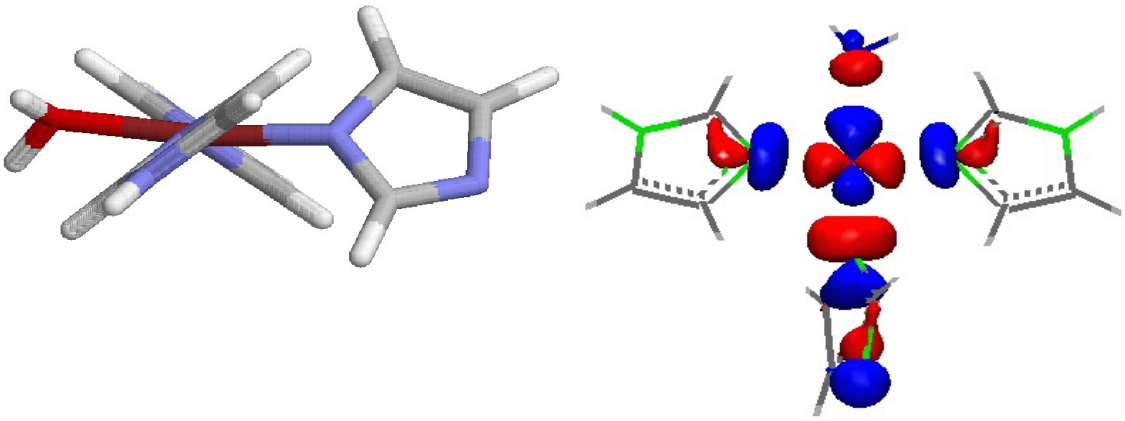
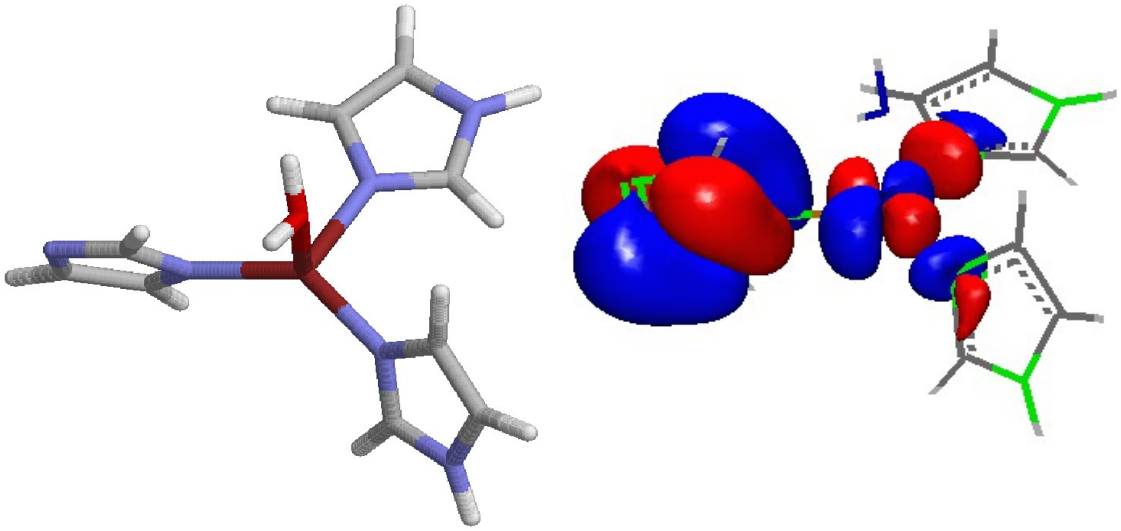


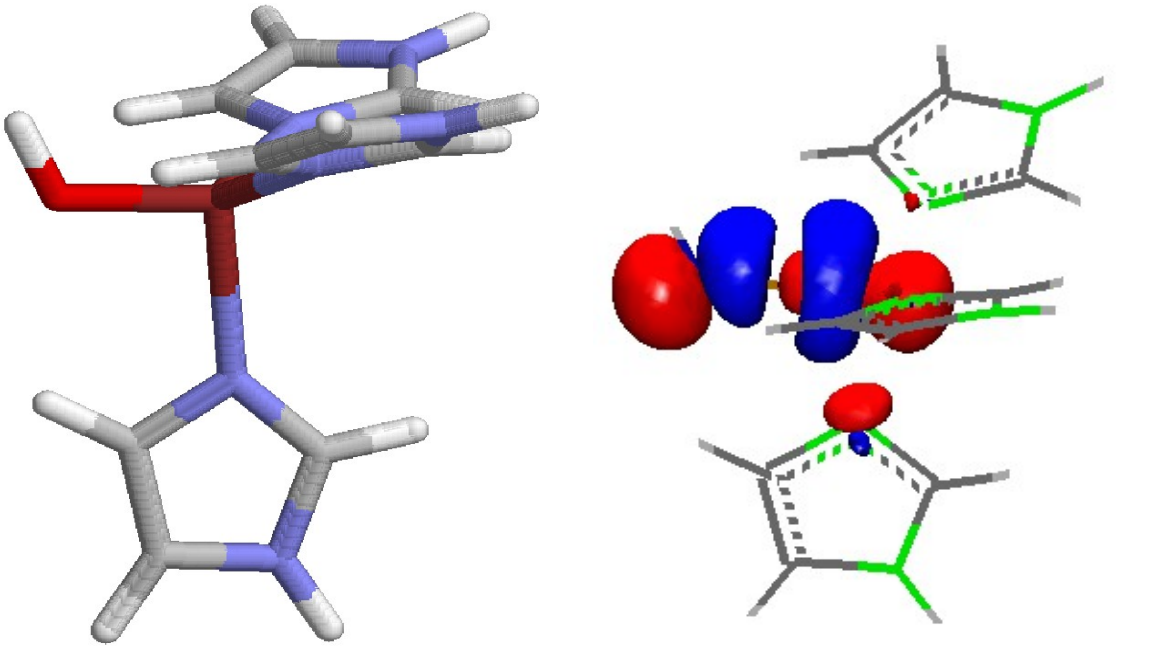
Figure 2. The optimised Tet1 structure of the $Cu(Im)_3(OH)^+$ model (a), the Trig structure of the $Cu(Im)_3(OH)^+$ (b) and $Cu(Im)_3(Imm)(H_2O)^+$ (c) models, and the Tet2 structure of the $Cu(Im)_3(OH)^+$ (d) and $Cu(Im)_3(Imm)(H_2O)^+$ (e) models. On the right-hand side, the corresponding singly-occupied orbitals of each of the complexes are shown.



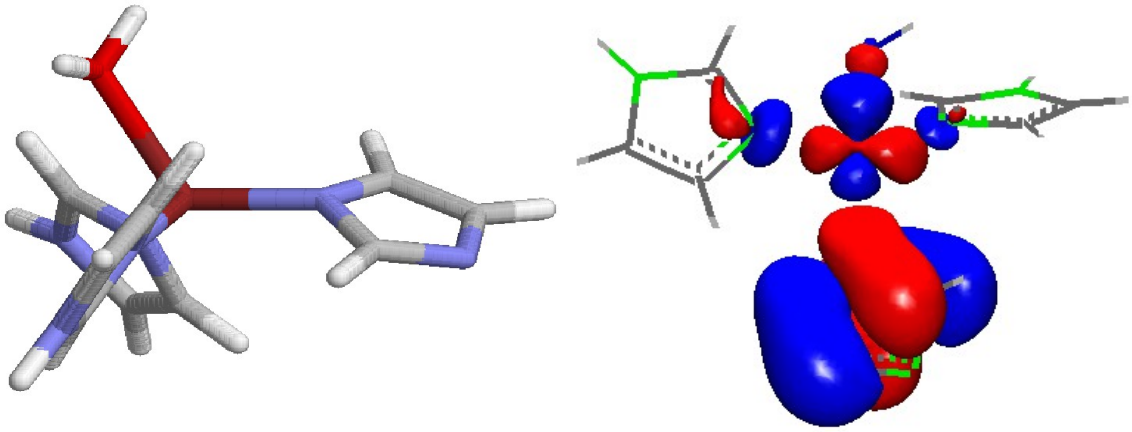
a



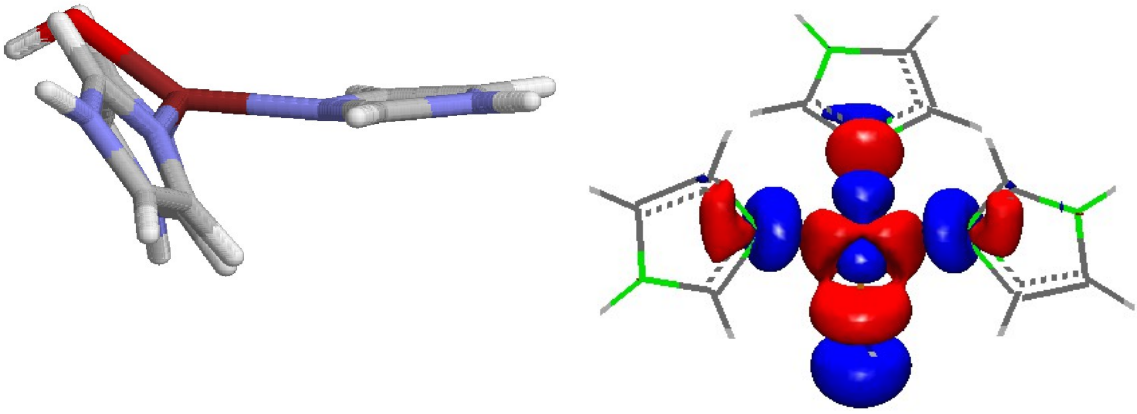
b



c

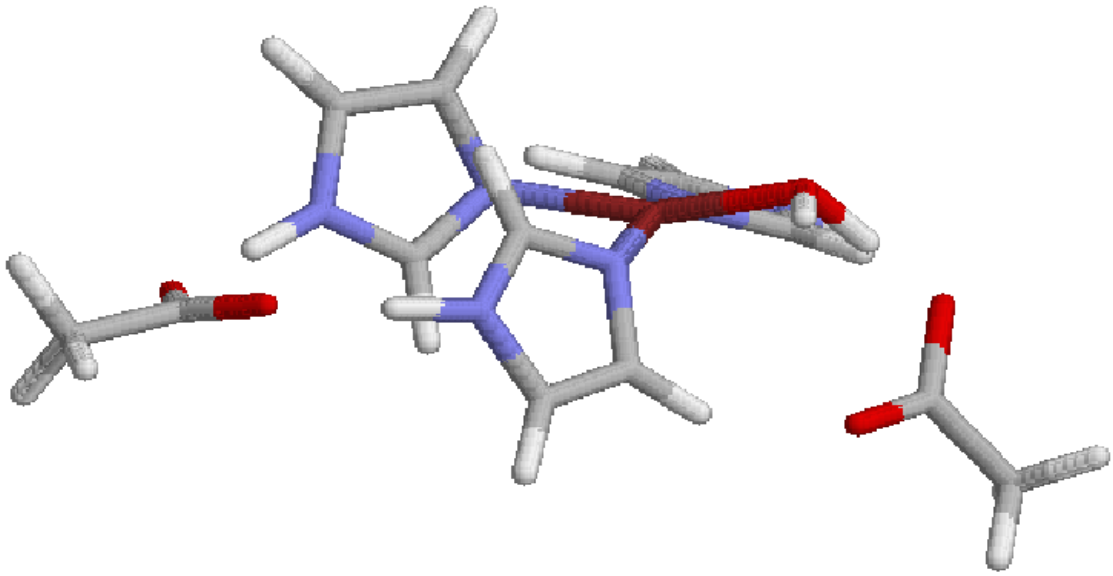


d

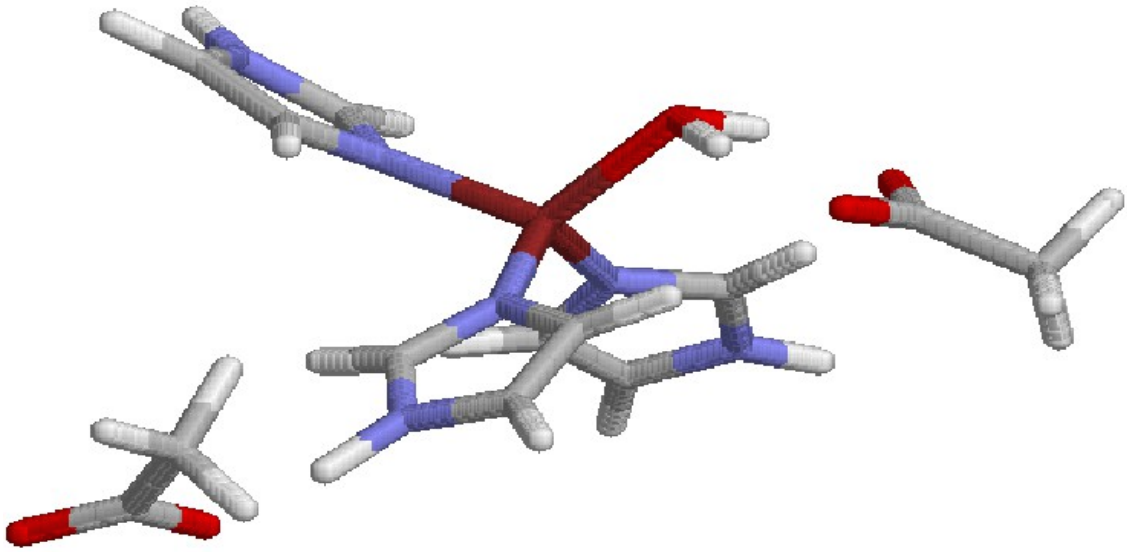


e

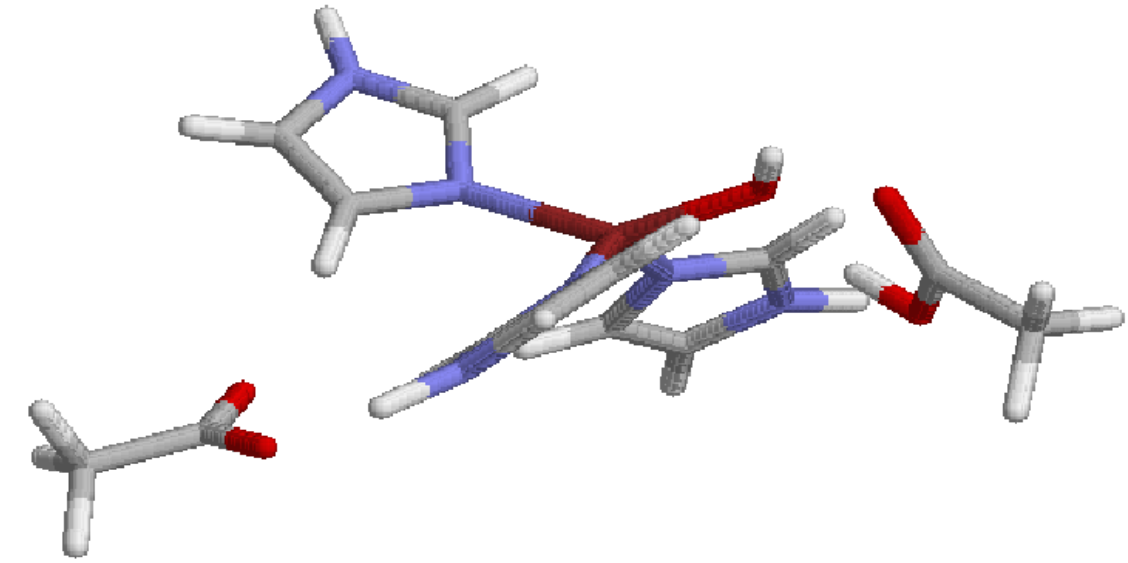
Figure 3. Vacuum structures of the various protonation states of the $Cu(Im)_3(H_2O)(Ac)_2$ model. The following structures are shown: the Tet1 structure of the Wat model (a), the Tet2 structures of the Wat (b), OH (c), Imm (d), and Both (e) models. In addition, the singly-occupied orbital of the Hyd Tet2 model (f) and the reduced Wat structure (g) are shown.



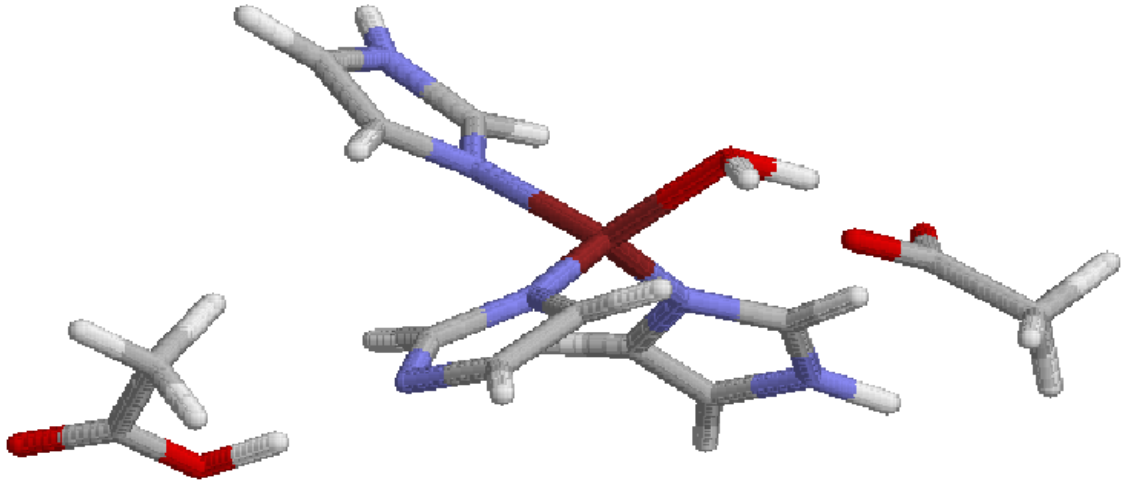
a) Wat Tet1



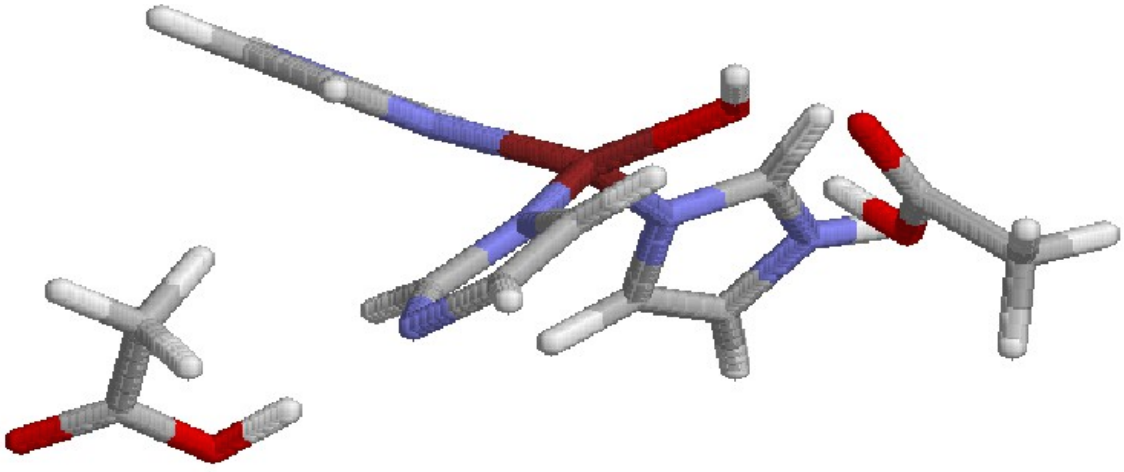
b) Wat Tet2



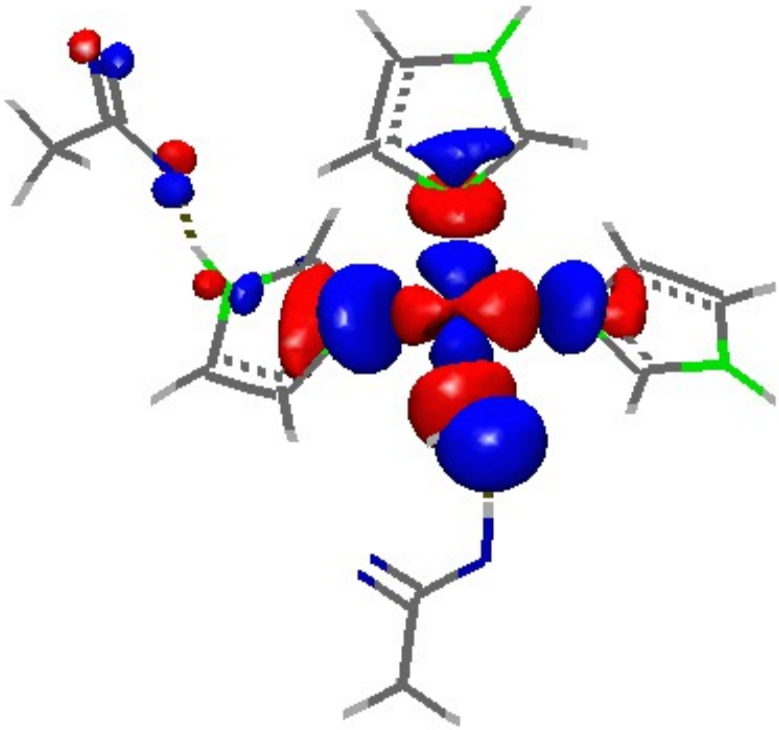
c) Hyd Tet2



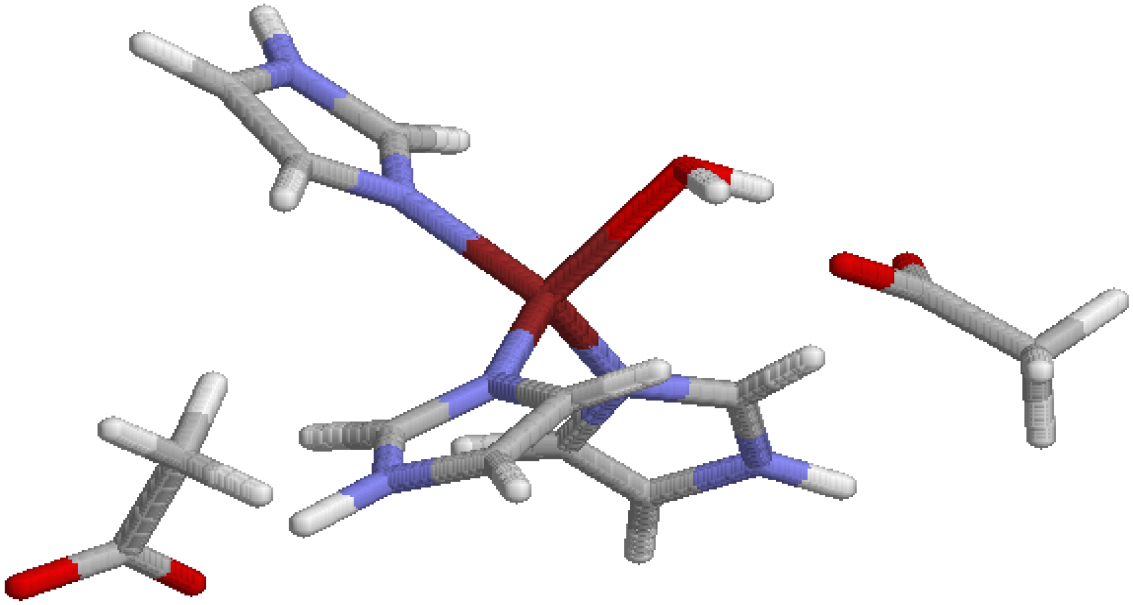
d) Imm Tet2



e) Both Tet2

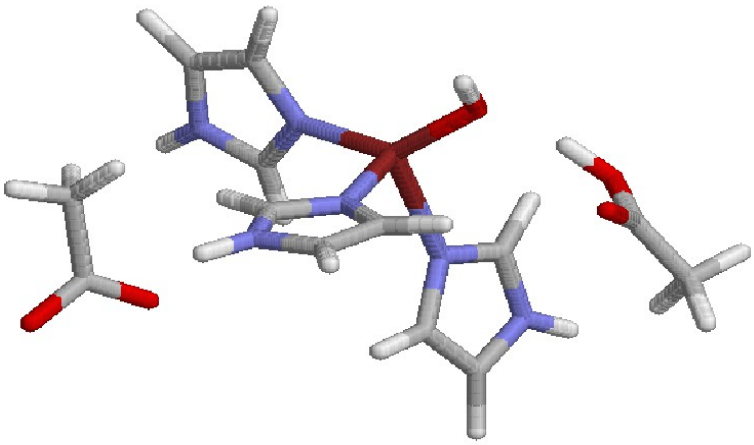


f) Hyd Tet2; electronic structure

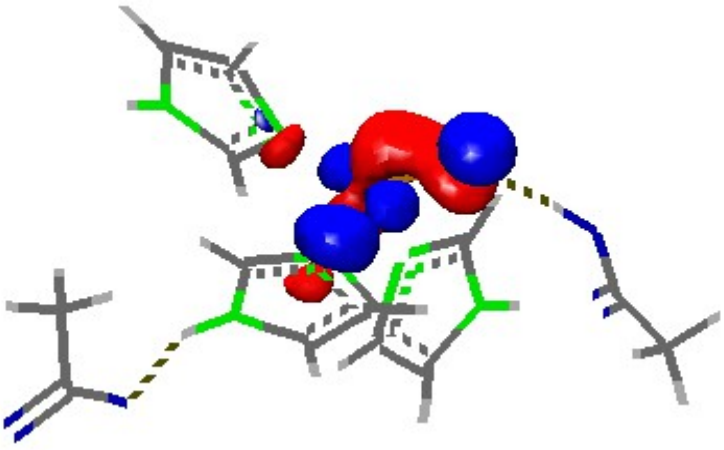


g) Wat Red

Figure 4. The QM/MM structure of the Hyd state of the $Cu(Im)_3(H_2O)(Ac)_2$ model, optimised inside nitrite reductase from *A. cycloclastes* [14].

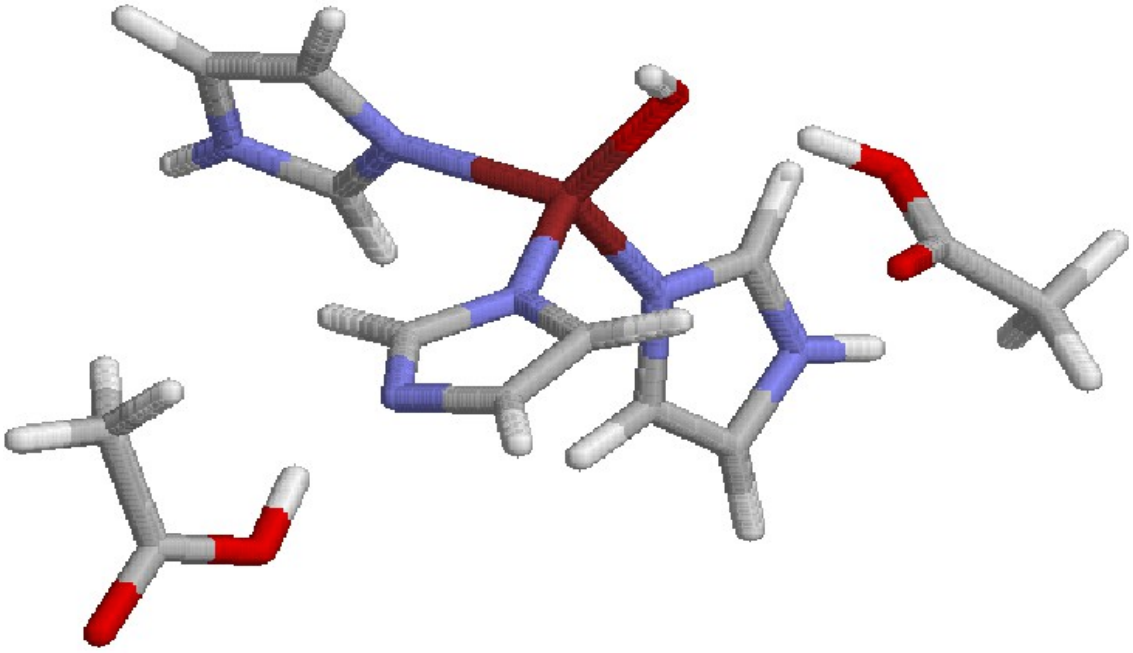


a

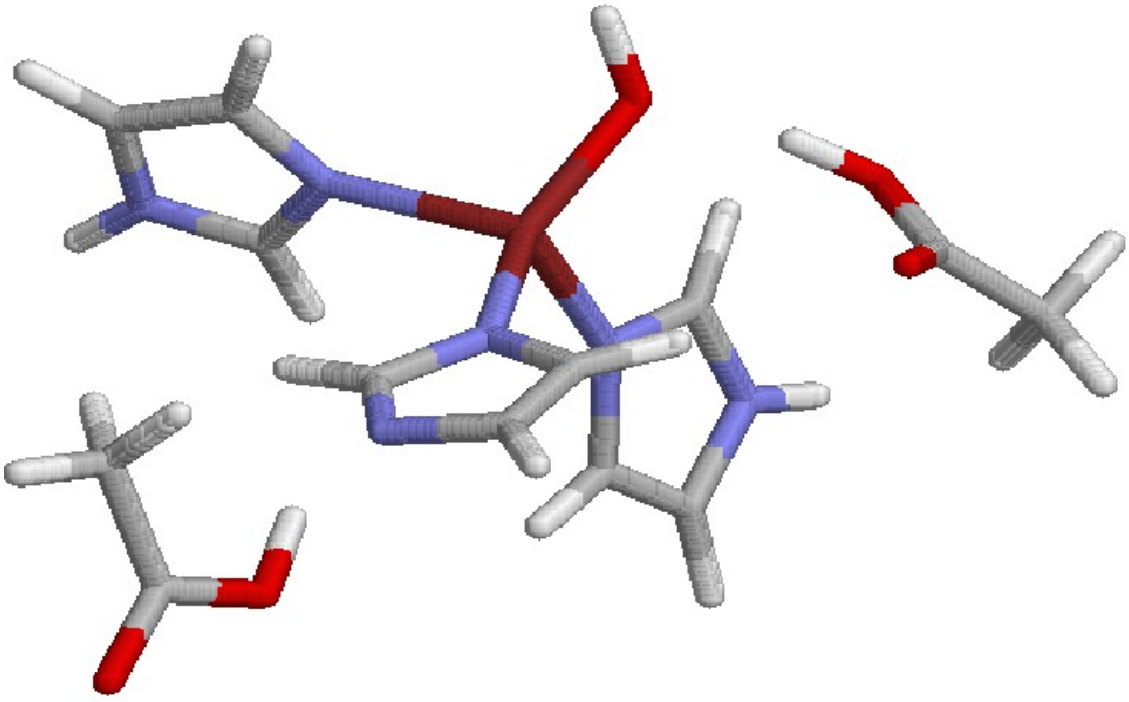


b

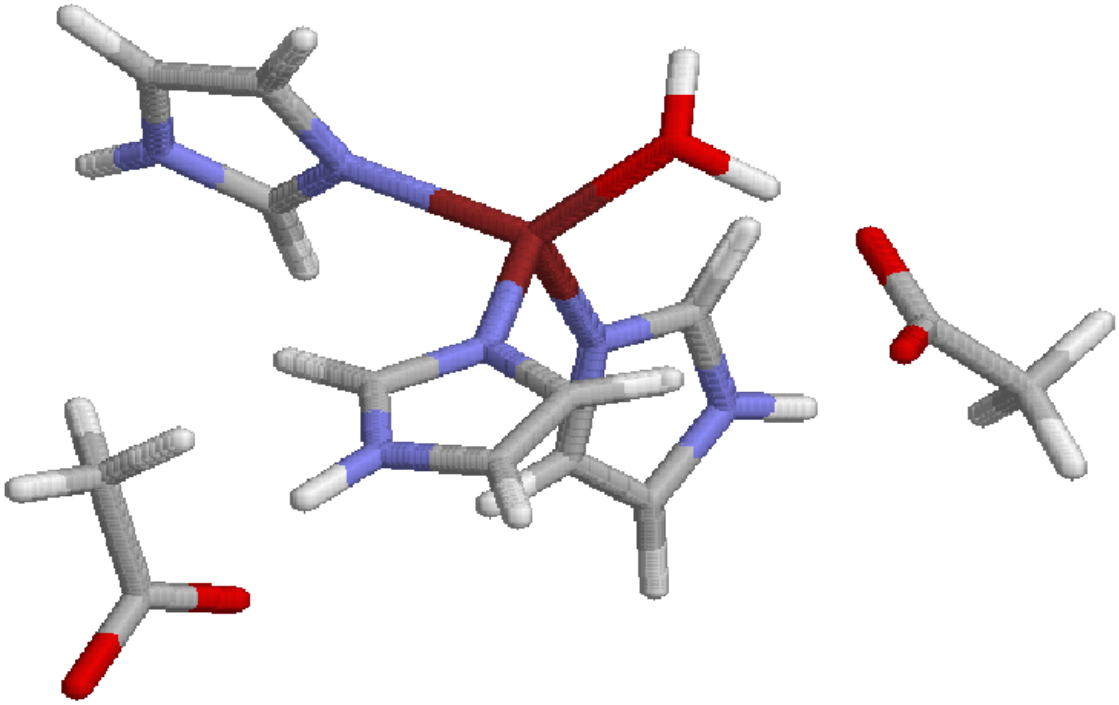
Figure 5. Structures optimised with quantum refinement of the crystal structure of nitrite reductase from *A. cycloclastes* [14]: Both, Hconf2 (a), Both, Hconf3 (b); Wat, Hconf1 (c), Hyd Hconf1 (d), Imm Hconf1 (e), Both Hconf1 (f). In addition, the open-shell orbital of the Both Hconf1 state is also shown (g).



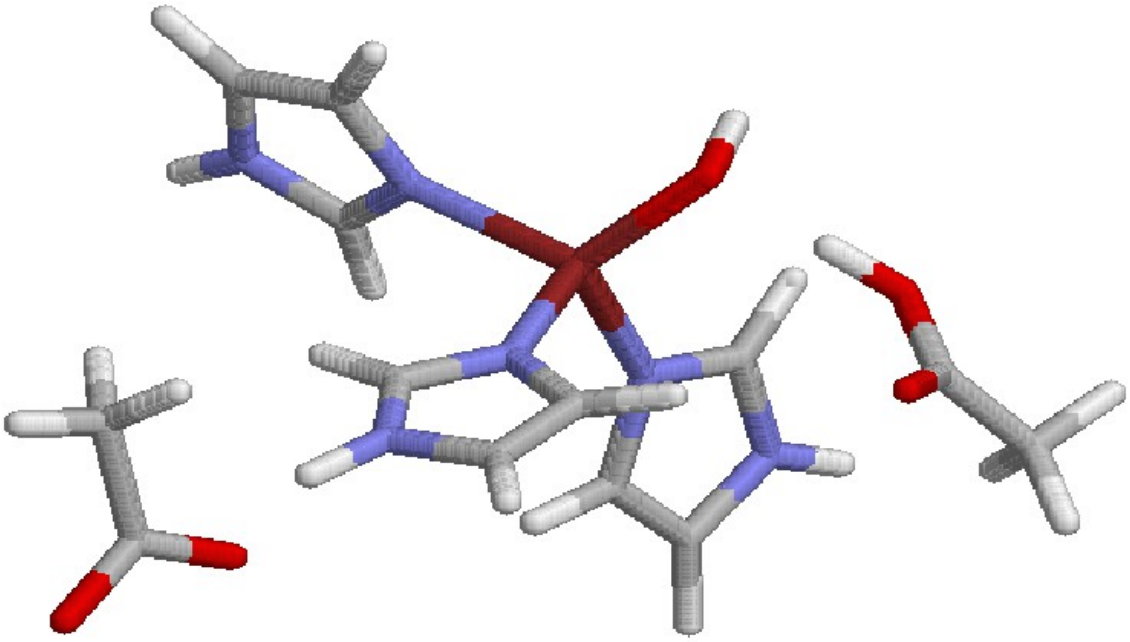
a) Both, Hconf2



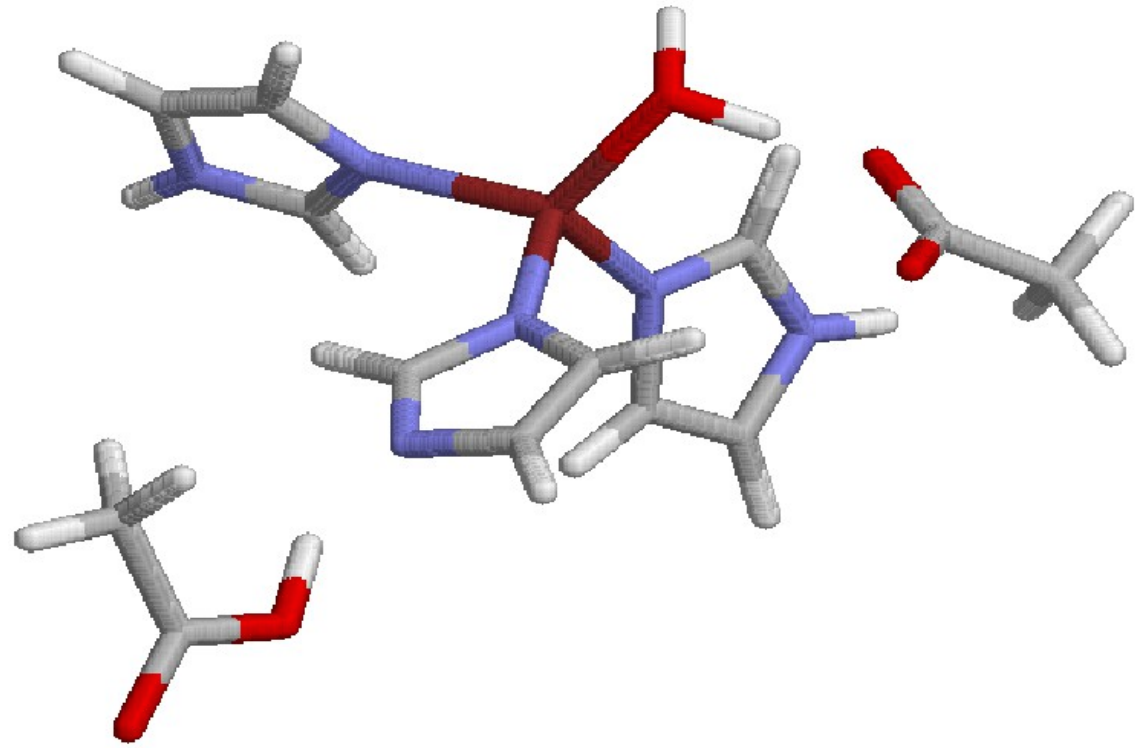
b) Both, Hconf3



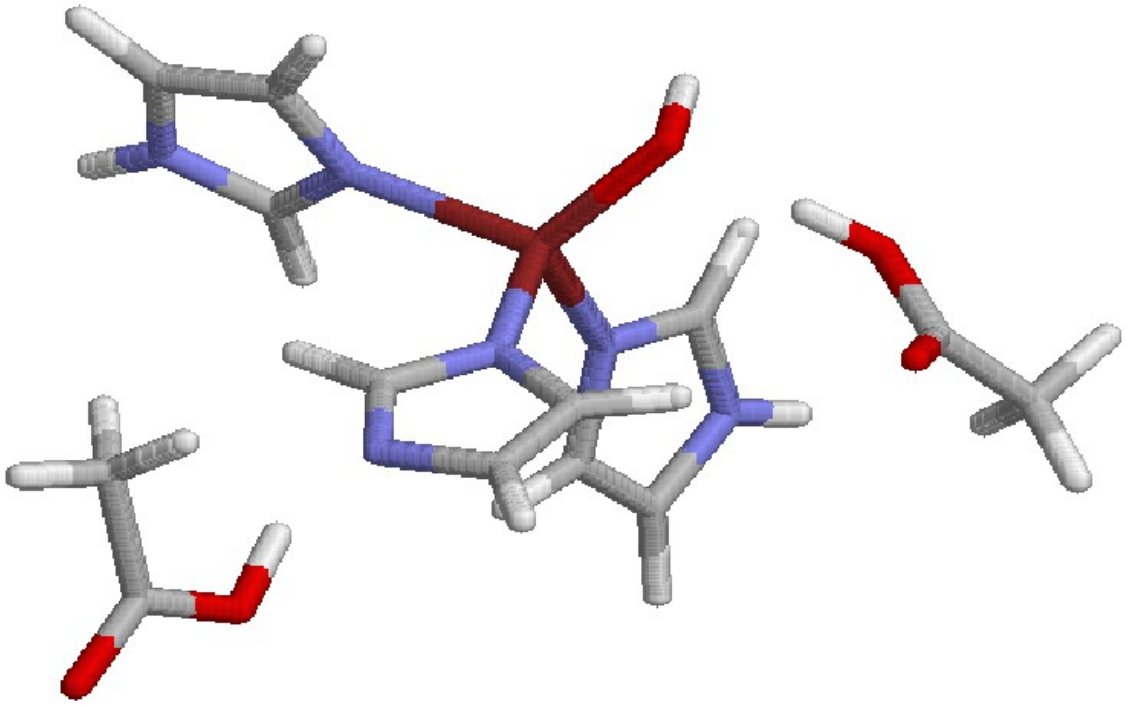
c) Wat, Hconf1



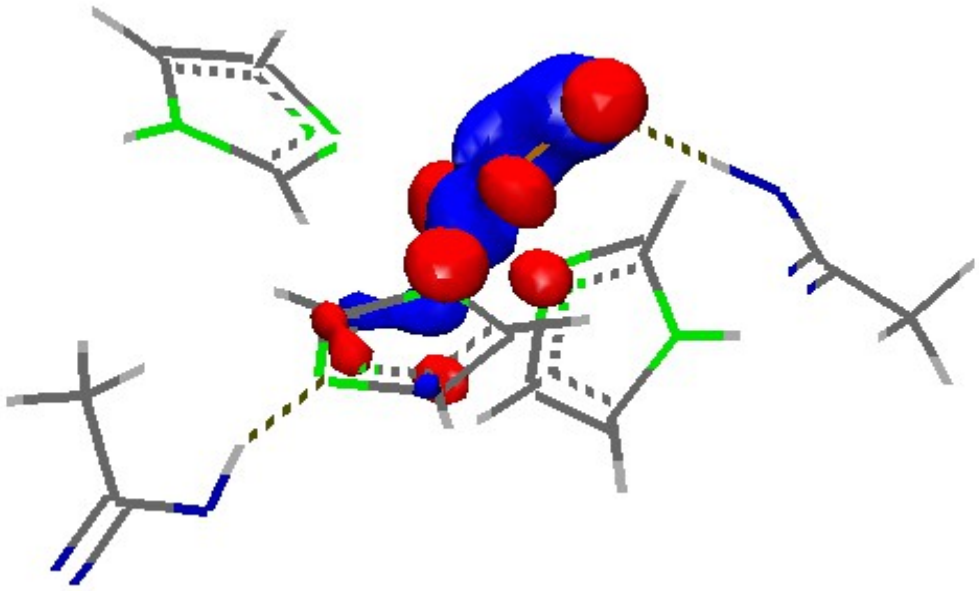
d) Hyd, Hconf1



e) Imm, Hconf1



f) Both, Hconf1



g) Electronic structure of Both, Hconf1

Figure 6. Relation between the angle of these versus the maximum N–Cu–N angle of 17 nitrite reductase structures in the Protein Data Bank. Two types of structures are evident, which are not directly related to the source organism (*Alcaligenes faecalis, A. xylosoxidans*, or *Achromobacter cycloclastes*). Structures from the following PDB files were used: 1as7, 1ndr, 1nia, 1nib, 1npj, 1npn, 1oe1, 1oe2, 2afn (three subunits each); 1et5, 1et7, 1et8, 1gs6, 1gs8, 1hau, 1nic, 1nie, and 2nrd (only one subunit each).

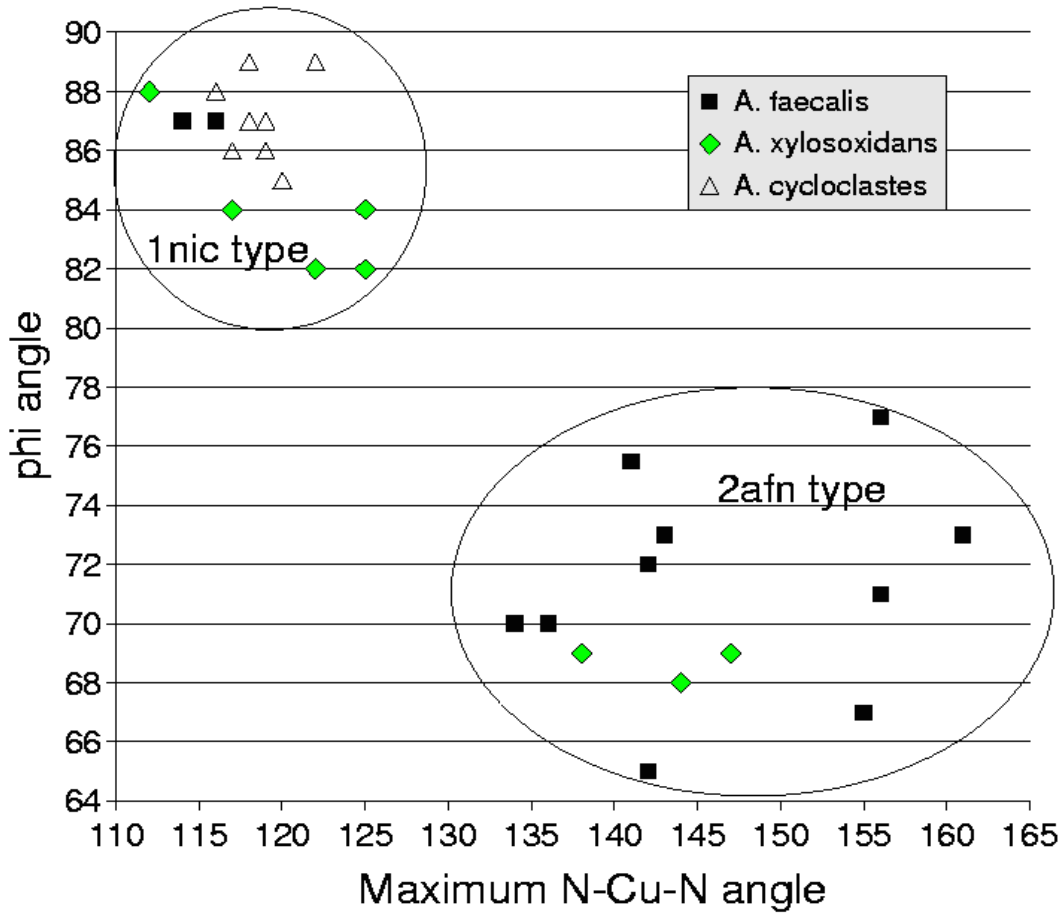


Figure 7. An overlay of the crystal structures of nitrite reductase from *A. cycloclastes* (solvent molecule, i.e. the red cross, to the right) and *A. faecalis* (solvent to the left). In addition, the electron-density $f_o - f_c$ difference map of the latter structure is show at 3.0 levels, calculated from the $_A$ weighted maps (green and blue). Note the positive residual density (green) near the alternative position of the solvent molecule.