Protonation of the proximal histidine ligand in heme peroxidases.

Heimdal, Jimmy; Rydberg, Patrik; Ryde, Ulf

Published in:
The Journal of Physical Chemistry Part B

DOI:
10.1021/jp710038s

2008

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

Protonation of the proximal histidine ligand in haem peroxidases

Jimmy Heimdal, Patrik Rydberg, and Ulf Ryde *

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

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

2017-04-10

Abstract

The haem peroxidases have a histidine group as the axial ligand of iron. This ligand forms a hydrogen bond to an aspartate carboxylate group by the other nitrogen atom in the side chain. The aspartate is not present in the globins and it has been suggested that it gives an imidazolate character to the histidine ligand. Quantum chemical calculations have indicated that the properties of the haem site strongly depend on the position of the proton in this hydrogen bond. Therefore, we have studied the location of this proton in all intermediates in the reaction mechanism, using a set of different quantum mechanical and combined experimental and computational methods. Quantum refinements of a crystal structure of the resting Fe^{III} state in yeast cytochrome *c* peroxidase show that the geometric differences of the two states are so small that it cannot be unambiguously decided where the proton is in the crystal structure. Vacuum calculations indicate that the position of the proton is sensitive to the surroundings and to the side chains of the porphyrin ring. Combined quantum and molecular mechanics (QM/MM) calculations indicate that the proton prefers to reside on the His ligand in all states in the reaction mechanism of the peroxidases. QM/MM free energy perturbations confirm these results, but reduce the energy difference between the two states to 12–44 kJ/mole.

Key Words: haem, proximal ligands, QM/MM, quantum refinement, proton transfer.

Introduction

Haem is one of the most important cofactors in biological systems. It is involved in a great wealth of proteins with functions ranging from electron transfer, storage and transport of small molecules like O_2, and the catalysis of a large number of reactions, e.g. in catalases, oxidases, and peroxidases.[1,2]

Haem consists of an iron ion in the centre of a porphyrin ring, which provides four nitrogen ligands in an equatorial plane. One or two axial ligands complete the octahedral coordination around the ion. These axial ligands vary extensively among the various haem proteins. For example, cytochromes have two axial ligands, typically histidine (His) or methionine, globins and haem peroxidases have a His ligand, cytochrome P450, NO synthase, and chloroperoxidase have a cysteine ligand, whereas the catalases have a tyrosine ligand. It has been shown that the axial ligand changes the chemical properties and reactivity of the haem group.[3,4,5,6,7,8]

It has also been realised that not only the axial ligand is important for the reactivity of the iron ion, but also second-sphere ligands. For example, the axial tyrosine ligand in catalases is hydrogen bonded to an arginine ligand. Likewise, the His ligand in haem peroxidases, but not in the globins, is hydrogen bonded (by the non-coordinating nitrogen atom of the imidazole side chain) to the carboxylate group of a nearby aspartate (Asp) residue. A similar metal–His–carboxylate ligand motif is found in many other proteins with different metals, e.g. in vitamin B_{12} proteins (Co), superoxide dismutases (Mn or Fe), alcohol dehydrogenase (Zn), and copper nitrite reductase. It is believed that this provides a simple way to reduce the positive charge of the metal site, but it has also been suggested to tune the properties of the metal ion, e.g. by imposing an imidazolate character on the His ligand, thereby increasing the electron density on the metal ion.[3,5,7,8]

The haem peroxidases are a group of enzymes that catalyse the one-electron oxidation of various substrates using hydrogen peroxide as the electron acceptor [9]. The resting state of the enzyme contains a high-spin (HS) Fe^{III} ion, typically with a weakly bound water ligand. This state binds a H_2O_2 molecule, which replaces the water ligand. The peroxide is believed to be deprotonated to an iron-bound hydroperoxide ion, which is then reprotonated on the distal oxygen atom. This triggers the cleavage of the O–O bond and leads, after the dissociation of a water molecule, to the formation of a highly reactive intermediate, called compound I. It consists formally of a Fe^{V}=O state, but computational investigations indicate that the electronic structure is closer to a Fe^{III} ion, with an $O^{\bullet-}$ radical and a porphyrin radical [4] (in some proteins, the latter radical is instead found on a nearby tryptophane residue). This high-valent intermediate has the potential to oxidise most substrate molecules (the reduction potential is ~0.9 V [9]), which bind at the surface of the protein, close to the edge of the haem group. By the transfer of this electron, compound I is reduced to compound II, in which the electron ends up in the porphyrin ring (or in the tryptophane residue). Compound II has almost the same reduction potential as compound I [9] and it can therefore oxidise another substrate molecule, thereby ending up in the resting Fe^{III} state (after uptake of two protons). It is unclear whether the oxoferryl group of compound II is protonated (OH^-) or not (O^{2-}).[10,11]

Computational studies have shown that the properties of the Fe–His–Asp motif strongly depend on the location of the shared proton [4,5] and this is also supported by experimental studies.[12,13,14] Although the pK_a value of the imidazole ring in His (~14) is higher than that of the carboxylate group in Asp (~4) [15], it is possible that coordination of His to the positively charged Fe ion (formally +3–5) may strongly down-shift its pK_a value to the neighbourhood of that of Asp (the pK_a value of one of the water ligands of Fe^{III} in aqueous solution is shifted from 15.6 to ~2.2 [16]). In fact, semiempirical calculations have indicated that the proton resides on Asp in compound I [17]. However, later studies with density functional theory (DFT) have indicated that the proton prefers to be the His ligand, although the two states are close in energy [18,19,20,21]. The same conclusion was reached by comparing experimental and theoretical spectra [18] and the doublet–quartet splitting energies for compound I in peroxidases [22]. A

combined quantum mechanical and molecular mechanics (QM/MM) investigation of the same state gave a much larger energy difference (~48 kJ/mole), still in favour of the state with the proton on His [23].

The position of the shared proton in peroxidases has also been studied with experimental methods: Both optical and resonance Raman spectral results have been suggested to indicate that the proximal His ligand is deprotonated [24,25,26]. On the other hand, NMR studies have indicated that in the five-coordinate HS Fe^{II} and Fe^{III} states, the His ligand is protonated and neutral [27,28], whereas in the CN^- inhibited enzyme, the proton has moved to the Asp group [29,30,31,32]. This has been taken to suggest that all six-coordinate states, including compound I, should have a deprotonated imidazolate ligand [29].

Considering this potential discrepancy between experiments and calculations, and the fact that the protonation very well may vary between the various intermediates, we in this paper study the protonation of the Fe–His–Asp motif in cytochrome *c* peroxidase for all relevant states in the reaction mechanism. Since the energy difference between the two states has been shown to be small [18] and strongly depend on the modelling of the surrounding protein [23], we test several different methods: We employ accurate DFT calculations in vacuum and a continuum solvent with a larger haem model than has been used before. Moreover, we perform QM/MM calculations, taking a full account of the surrounding protein. These calculations are then improved by QM/MM free energy perturbations along the proton-transfer path to estimate dynamic and entropic effects. Finally, we also perform a quantum refinement of a crystal structure of the resting state. The results indicate that the proton prefers to be on the His ligand in all states.

Methods

Quantum mechanical calculations

Quantum mechanical (QM) calculations were performed with density functional theory, using the Becke 1988–Perdew 1986 (BP86) functional [33,34] as implemented in the Turbomole package [35]. These calculations employed the 6-31G* basis set for all atoms [36], except iron, for which we used the DZP basis sets of Schäfer et al. [37,38]. The calculations were sped up by expansion of the Coulomb interactions in auxiliary basis sets, the resolution-of-identity approximation [39,40]. Structures were optimised until the change in energy between two iterations was below 2.6 J/mole (10^{-6} a.u.) and the maximum norm of the internal gradients was below 10^{-3} a.u.

To obtain more accurate energies, single-point calculations were performed with the three-parameter hybrid functional B3LYP method [41,42]. In these calculations, the 6-311+G(2d,2p) basis set was used for the light atoms, whereas the DZP basis set for iron was 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.

Continuum solvation effects were estimated by single-point calculations using the continuum conductor-like screening model (COSMO) [43,44]. 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 80. For the generation of the cavity, a set of atomic radii has to be defined. We used the optimised COSMO radii in Turbomole (1.3, 2.0, 1.83, and 1.72 Å for H, C, N, and O, respectively [45], and 2.0 Å for the iron).

QM/MM calculations

Combined QM and molecular mechanics (MM) calculations were carried out with the

ComQum program [46,47]. In this approach, the protein and solvent are split into two subsystems: The QM region (system 1) contains the most interesting atoms and is relaxed by QM methods. The MM region (system 2) consists of the remaining part of the protein and the surrounding solvent molecules and is kept fixed at the original (crystallographic) coordinates (it is relaxed in the free energy perturbations, see below). In the QM calculations, system 1 is represented by a wavefunction, whereas all the other atoms are represented by an array of partial point charges, one for each atom, taken from MM libraries. Thereby, the polarization of the quantum chemical system by the surroundings is included in a self-consistent manner. When there is a bond between systems 1 and 2 (a junction), the quantum region is truncated by hydrogen atoms, the positions of which are linearly related to the corresponding carbon atoms in the full system (the hydrogen link-atom approach) [46,48]. In order to eliminate the non-physical effect of placing point charges on atoms in the MM region bound to junction atoms (i.e. the closest neighbours of QM system), those charges are zeroed, and the resulting residual charges are smoothly distributed [46,47].

The total energy is calculated as:

$$E_{\mathrm{QM/MM}} = E_{\mathrm{QM1+ptch}} - E_{\mathrm{MM1}} + E_{\mathrm{MM12}} \quad (1),$$

where $E_{\mathrm{QM1+ptch}}$ is the QM energy of system 1 truncated by the hydrogen atoms and embedded in the set of point charges (but excluding the self-energy of the point charges). E_{MM1} is the MM energy of system 1, still truncated by hydrogen atoms, but without any electrostatic interactions. Finally, E_{MM12} is the classical energy of all atoms with normal atoms at the junctions and with the charges of the quantum system set to zero (to avoid double-counting of the electrostatic interactions). By this approach, which is similar to the one used in the Oniom method [49], errors caused by the truncation of the quantum system cancel out.

MM and MD calculations

All MM calculations were run with the sander module of the AMBER 8 software [50], using the AMBER 1999 force field [51,52]. System 1 in the QM/MM calculations was represented by charges fitted to the electrostatic potential (ESP), estimated by Turbomole with a method similar to the Merz–Kollman scheme [53], but with ~2500 points/atom. The charge on each (C) junction atom was adapted so that the total charge of the amino acid (both QM and MM atoms) was the same as the sum of QM charges of the corresponding QM fragment (disregarding a possible net charge of the amino acid outside the QM region) [47]. Thereby, we ensure that the total charge of the simulated system is an integer, but we still allow charge transfer within the QM system (the junction amino acids have non-integer total charges). Moreover, the charges on the junction atoms are changed from what is typical for a hydrogen atom to what is more typical for carbon atoms.

All bond lengths involving hydrogen atoms were constrained to their equilibrium value using the SHAKE algorithm [54]. Water solvent was described explicitly using the TIP3P model [55]. The electrostatics were treated with particle-mesh Ewald method [56,57] with a grid size of 80^3 Å, a fourth-order B-spline interpolation, a tolerance of 10^{-5}, and a real-space cut-off of 8 Å. The temperature was kept constant at 300 K using the Berendsen weak-coupling algorithm [58] with a time constant of 1 ps. The MD time step was 2 fs and the non-bonded pair list was updated every 50 ps.

QTCP

QTCP (QM/MM thermodynamic cycle perturbation) is a method to estimate free energy barriers between two reactants (A and B) at the QM/MM level [59,60]. It is based on the thermodynamic cycle in Figure 1: We estimate the free energy difference $\Delta A_{QM/MM}(\mathrm{A}\rightarrow\mathrm{B})$ by

performing three free energy perturbations, one at the MM level from A to B, ΔA_{MM}(A→B) (of course, if the difference between the two reactants is too large, this perturbation may be done in several smaller steps), and one for each of the A and B states from a description at the MM level to the QM/MM level:

$$\Delta A_{\mathrm{QM/MM}}(\mathrm{A}\rightarrow\mathrm{B}) = \Delta A_{\mathrm{MM}}(\mathrm{A}\rightarrow\mathrm{B}) - \Delta A_{\mathrm{MM}\rightarrow\mathrm{QM//MM}}(\mathrm{A}) + \Delta A_{\mathrm{MM}\rightarrow\mathrm{QM//MM}}(\mathrm{B}) \qquad (2)$$

The advantage of this approach is that the phase space needs to be sampled only at the MM level (for the A and B states), thereby avoiding the extremely time-consuming sampling at the QM/MM level (if a high-level QM theory is used). A similar approach has been used by Warshel and coworkers [61,62], but it led to severe convergence problems. We solved these by keeping the QM system fixed during the simulations [59,60]. Thereby, we also avoid the need of MM parameters for the QM system. This requires that the QM/MM calculations were run with system 2 fixed (otherwise the junction atoms will move too much).

In practice, the QTCP calculations were performed in the following way [59,60]: We started from the QM/MM calculations, but the proteins were better solvated by using an octahedral water box, extending at least 9 Å from the protein (~10 000 atoms). This system was then simulated by MD in the *NPT* ensemble (one atmosphere pressure and 300 K temperature) for 20 ps, restraining the heavy atoms by a harmonic force constant of 50 kcal/mole/Å^2 to the QM/MM structure. Thereafter, only the QM system was restrained and the system was simulated for another 50 ps in the *NPT* ensemble to allow the volume to equilibrate. After this, the QM system was moved back to the exact position in the QM/MM calculations (after translation and rotation; the QM system moves slightly during the constant-pressure simulation) and was fixed. The system was then equilibrated in the *NVT* ensemble for 200 ps and snapshots were saved every ps for an additional time of at least 400 ps.

Finally, free energies were calculated from these snapshots (strictly speaking, Helmholtz free energies are calculated, rather than the more usual Gibbs free energy, but the difference between these two quantities is small for proteins [63]). The free energy at the MM level (ΔA_{MM}(A→B) in Eqn. 2) was calculated by a standard free energy perturbation (FEP) by changing the coordinates and the charges of the QM atoms (with C junction atoms) from the A to the B state (or vice versa) and calculating the energy difference:

$$\Delta A_{MM}(A\rightarrow B) = -k_B T \ln\left\langle \mathrm{e}^{-\frac{E(R)-E(P)}{k_B T}} \right\rangle_A \qquad (3)$$

where k_B is the Boltzmann constant, T the temperature, and the angular brackets indicate an average over an MD ensemble, sampled for the A state. A test of the accuracy of this free energy is obtained by reversing the A and B states, using snapshots for the simulation of the B state instead.

Likewise, the QM/MM free energies were obtained by a FEP, in this case over the same state at two different levels of theory, viz. MM→QM/MM (owing to the cost of calculating this term, only every fifth snapshot was used in these FEPs). As has been described before [60], this was accomplished by the following equation:

$$\Delta A_{MM\rightarrow QM/MM}(X) = -k_B T \ln\left\langle \mathrm{e}^{-\frac{E'_{QM1+ptch}(X)-E_{el12}(X)}{k_B T}} \right\rangle_X \qquad (4),$$

where X is the A or B state and $E'_{\mathrm{QM1+ptch}}$ is the QM energy of the QM system embedded in a point-charge model of the surrounding protein, but excluding the energy of the point charges with themselves (the self-energy). It is essentially the same energy as $E_{\mathrm{QM1+ptch}}$ in Eqn. 1, with the small exception that in the QM/MM calculations, a few point charges near to the junction

atoms were deleted, which is not the case for $E'_{\mathrm{QM1+ptch}}$, where all point charges of the MM system were included. E_{el12} is the electrostatic interaction energy between the QM (with H junction atoms) and MM systems (excluding interactions within the two systems), calculated at the MM level (i.e. by Coulomb's law), using the same point charges for the MM system as in $E'_{\mathrm{QM1+ptch}}$ and the QM fitted point charges for the QM system (i.e. the same point charges used in the MD simulation, with the exception for the junction atoms, for which the original ESP point charges were used (in the MD simulations, the point charges on the junction atoms were adapted so that the total charge of the protein was an integer; in the calculation of E_{el12}, the sum of the point charges of the QM system is an integer, whereas that of the MM system is not an integer). Thereby, we obtain an energy that is as similar to that in the QM calculation as possible.

It should be noted that the energy in Eqn. 4 includes the total QM energy of the quantum system and therefore is large. This energy is nearly cancelled by the corresponding QM energy of the other state. Thus, the internal energy of system 1 is entirely calculated by QM. The MM energy of the QM system completely cancels out. Therefore, the MM parameters of the QM system have no influence on the results [60].

Quantum refinement

Finally, we also performed a set of quantum refinement calculations, using the software ComQum-X [64]. It can be seen as a QM/MM calculation, in which the structures are restrained towards crystallographic raw data. In ComQum-X, the MM program is replaced by the crystallographic refinement program CNS (Crystallography & NMR system) [65]. In crystallographic refinement, the coordinates, *B* factors, occupancies, etc. are 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 [66]. 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 a small, but interesting, part of the protein (system 1), in a manner completely analogous to 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

$$E_{\mathrm{ComQum\text{-}X}} = E_{\mathrm{QM1}} - E_{\mathrm{MM1}} + E_{\mathrm{MM12}} + w_{\mathrm{A}} E_{\mathrm{Xray}} \qquad (5),$$

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 [67]. 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_{MM12} and E_{Xray} forces are equal ($w_{\mathrm{A}} = 0.43$ in all calculations).

Following crystallographic custom, no hydrogen atoms were included in the MM region of the ComQum-X calculations, because 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.

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 [68], 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 [64,68].

The protein

All calculations in this investigation are based on the $P2_12_12_1$ crystal structure of yeast cytochrome *c* peroxidase at 1.7 Å resolution [69] (*R* factor 0.202). Coordinates, occupancies, *B* factors, and structure factors were obtained from the Protein Data Bank, access code 2cyp. The structure was collected for the resting high-spin Fe^{III} state with a weakly bound water ligand of the iron ion (Fe–O distance 2.40 Å). Alternative conformations in two residues (Thr-63 and Leu-161, two atoms in each) were ignored (only the major conformation with occupancy 0.65–0.7 was included in the calculations).

The QM system (system 1) consisted of the haem group, the water ligand, an imidazole group as a model of His-175, and an acetate group as a model of Asp-235. In most calculations, all the side chains of the haem group were included (92 atoms in total), whereas in some calculations, the side chains were replaced by a hydrogen atom (56 atoms). In calculations of other intermediate states of the haem group, the water ligand was deleted or replaced by other molecules (HO_2^-. OH^-, O^{2-}, or CN^-) and the oxidation state of the iron ion was changed.

For the QM/MM calculations, hydrogen atoms were added with the AMBER package [50]. All lysine, arginine, aspartate, and glutamate residues were charged. The protonation status of the histidine residues was determined by inspection of the local surroundings and hydrogen-bond structure. This gave protonation of $N^{\delta 1}$ for residues 52 and 175 (the distal and proximal His residues), whereas the other four His residues were protonated on both N atoms and therefore positively charged. Before running the QM/MM calculations, the positions of hydrogen atoms and solvation water molecules were optimised by a 60 ps simulated annealing with molecular dynamics (MD; from 300 to 0 K), followed by a MM minimisation.

Result and Discussion

QM calculations in vacuum and in a continuum solvent

We have studied the protonation of the His–Asp link in haem peroxidases, i.e. whether the proton prefers to be on the His ligand (HisH form) or the Asp group (AspH form). First, we studied the protonation in vacuum, using models of two different sizes, viz. with or without the side chains of the porphyrin ring. In both cases, we studied five different states in the catalytic cycle of the peroxidases, viz. the Fe^{III} resting state (without a water ligand in HS state or with the water ligand in both the HS and low-spin (LS) states), the hydroperoxy complex ($Fe^{III}HO_2^-$, in the LS state), compound I (both in the LS and intermediate-spin (IS) states), compound II (protonated or not, i.e. with a O^{2-} or a OH^- ligand, both in IS state), and also the experimentally studied $Fe^{III}CN^-$ complex (LS state).

From the results in Table 1, it can be seen that without the side chains, the proton prefers to be on the Asp ligand (negative energies) for the $Fe^{III}H_2O$, Fe^{III}, $Fe^{IV}OH$, and $Fe^{III}CN$ complexes (by 3–19 kJ/mole), whereas the HisH state is more stable for the other four complexes. These results are in agreement with earlier theoretical investigations of compound I [18,19,20,21,23] and they show that the most stable protonation state quite strongly varies with the oxidation state and distal ligand of the iron ion. However, the results are sensitive to solvation effects and in a continuum solvent with a dielectric constant of 80, the HisH state is stabilised by 4–18 kJ/mole. Consequently, the HisH state is most stable for all models, except for the LS $Fe^{III}H_2O$, Fe^{III} and $Fe^{IV}OH$ complexes, although the energy difference is always small (–4 to +10 kJ/mole). On the other hand, if the energies are recalculated with the B3LYP method and larger basis set, the AspH state is stabilised for five of the models by 16–27 kJ/mole (but destabilised for the other four by 1–10 kJ/mole). Therefore, the AspH state is predicted to be most stable for all complexes, except $Fe^{III}HO_2$, $Fe^{IV}O$, and $Fe^{III}CN$. Test calculations show that the effect comes almost entirely from the change of the density functional, rather than from

the increased basis set.

Interestingly, the results are quite different for the models with the porphyrin side chains: In this case, the HisH state is most stable (by 1–19 kJ/mole) for all models, except the five-coordinate Fe^{III} complex. In fact, for most models, only the HisH state is a local minimum. For the other states, the energy differences in Table 1 were obtained by constraining the H–O_{Asp} distance to 1.13 Å (a typical distance in the stable AspH complexes). Moreover, the results are insensitive to solvation effects or to a change in the method and basis set: The results change by less than 7 kJ/mole, with a varying sign. With B3LYP, the two states are degenerate to within 3 kJ/mole for the two $Fe^{III}H_2O$ complexes and the five-coordinate Fe^{III} complex, whereas the HisH state is more stable by 7–23 kJ/mole for the other complexes.

Even if the large model with the porphyrin side chains is more realistic, it is not entirely evident that the energies obtained with that model are more correct. Two of the side chains are propionate groups and those peripheral charged groups can cause problems in vacuum optimisations. However, as can be seen in Figure 2, the conformations of these two groups are very similar in the two protonation states. In fact, the geometry of the complexes changes little during the movement of the proton, except for a tilt of the acetate group. Therefore, the stabilisation of the HisH state by the porphyrin side chains must be entirely electrostatic in origin. As can be seen from Figure 2, the propionate groups are both on the same side of the porphyrin ring, relative to the moving proton, and it can be imagined that they will slightly attract the moving proton towards His.

Moreover, it is also known that unpaired spin may appear on propionate groups on haem, and this is normally interpreted as an artefact of vacuum calculations [23,70,71,72]. It can be seen in Table S1 (in the supplementary material) that significant spin densities are indeed found on the propionate groups for all models, except the $Fe^{III}HO_2$ and $Fe^{IV}O$ complexes at the BP/6-31G* level. However, the spin nearly or completely disappears for the calculations in a water-like continuum solvent and especially for the big-basis B3LYP calculations for all complexes, except the $Fe^{IV}OH$ and compound I models. Considering that the relative energies of the HisH and AspH states do not change with solvation or the method, this indicates that the stability of the two protonation states is not sensitive to the spin on the propionate groups. Consequently, we think that the calculations with side chains are most realistic and therefore most reliable.

Interestingly, the models without side-chains show significant spin densities on the carboxylate group of the Asp ligand in the HisH state. Again, the spin disappears for the B3LYP or continuum-solvent calculations, except for the $Fe^{IV}OH$ and compound I models. For the models without side chains, there was a dependence of the relative stabilities of the HisH and AspH states on solvation or the method, but the variation does not correlate with the spin density. A minor spin on the Asp carboxylate atoms are also found in the two calculations of compound I with porphyrin side chains (Table S1), but it disappears in the B3LYP calculations.

QM/MM calculations

The vacuum calculations did not unambiguously settle the location of the shared proton in the His–Asp link in the peroxidases and they showed that the results may be sensitive to the surroundings of the haem group. Therefore, we next run a set of QM/MM optimisations of seven models in cytochrome *c* peroxidase (the same models as in the vacuum calculations, except the LS $Fe^{III}H_2O$ model and the HS Fe^{III} model without any axial ligand, which are not involved in the peroxidase reaction cycle). All calculations involved a full porphyrin model with side chains, because these calculations are more realistic and in the QM/MM calculations, the surroundings of the propionate groups (as well as all other groups) are included by a point-charge model. On the other hand, this also means that the status of the surrounding protein affects the results. In particular, we have to decide the protonation state of the distal His residue, which forms a hydrogen bond to the axial iron ligand. For most models,

we have assumed that this residue is protonated on the $N^{\delta 1}$ atom, but not on the $N^{\varepsilon 2}$ atom. However, for the $Fe^{III}CN^-$ complex, experimental evidence indicates that that the distal His residue is protonated on both nitrogen atoms (and therefore is positively charged) [27,28]. Therefore, this protonation state is used for the CN^- complex. For compound I, we tested both protonation states.

The QM/MM optimised structures of the various models are described in Table 2. The geometries are as expected for such complexes [4]. The $Fe–N_{Por}$ bond lengths are shorter in the LS and IS states (2.01–2.02 Å) than in the HS states (~2.08 Å). The $Fe–N_{His}$ bond length is normally shortened when the proton is moved to Asp, but the effect is less than 0.04 Å and for some complexes, it even increases slightly. The Fe–O bond length depends strongly on the ligand. It is long for water (~2.32 Å), 1.77 Å for HO_2^-, 1.62–1.66 Å for compound I and for unprotonated compound II, and 1.79 Å for protonated compound II. It typically increases by ~0.01 Å when the proton is moved to Asp. The Fe–CN bond length is 1.87 Å.

As expected, the electronic structures are the same as in previous calculations in similar complexes [4], cf. Table 2: The H_2O and CN^- models exhibit rather pure Fe^{III} states and the HO_2^- model is mainly Fe^{III}, but with 0.2 spin on the ligand and 0.1 spin in the porphyrin. Compound I has essentially three unpaired electrons: one on Fe, one on the O-ligand, and one in the porphyrin ring (but somewhat delocalised on His or Asp). Thus, it is can formally be described as $Fe^{III}O^{-\bullet}Por^{\bullet}$. Unprotonated compound II has a similar electronic structure, but it has lost the unpaired electron in the porphyrin ring. Protonated compound II, finally, is essentially $Fe^{IV}OH^-$, although there are 0.26 spin on OH^- and only ~1.6 spin on Fe. Some spin is still found on the propionate groups or on the Asp ligand in the HisH state for all complexes except $Fe^{III}HO_2$ and $Fe^{IV}O$, even with the point-charge model of the surrounding protein. However, this spin completely disappears in the B3LYP calculations. (cf. Table S1). The electronic structure does not change significantly when the proton is moved. The only larger change observed in the spin densities is that ~0.18 *e* spin on the Asp ligand moves to the His ligand in compound I. The spin on Fe changes by less than 0.05 *e* and that on the ligand by 0.01 *e* or less.

The relative energies of the HisH and AspH states are compiled in Table 3. The QM/MM energies ($E_{QM/MM}$, defined in Eqn. 1) show that in the protein, the proton prefers to be on the His residue for all models by 12–48 kJ/mole. Thus, the protein stabilises the HisH state by 4–30 kJ/mole (cf. the E_{Vac} column, which shows the vacuum results from Table 1). We can understand this effect of the protein by dividing the QM/MM energy into its components. First, we look at the $E_{QM1+ptch}$ term in Eqn. 1 (i.e. the QM energy of the quantum system, including the point-charge model). It can be seen that this energy is similar to $E_{QM/MM}$ (within 4 kJ/mole). This shows that the MM energy ($E_{MM12} - E_{MM1}$ in Eqn. 1) is insignificant.

Therefore, we can continue to compare $E_{QM1+ptch}$ with the QM energy of the quantum system without the point-charge model, E_{QM1}. Interestingly, $E_{QM1+ptch}$ is always 3–33 kJ/mole more positive than E_{QM1}, which shows that the protein stabilises the HisH state by electrostatic interactions. On the other hand, E_{QM1} is 1–9 kJ/mole more negative than the vacuum QM energy (identical to E_{QM1}, but calculated for the optimum vacuum geometry). This shows that the geometry in the protein favours the AspH state, probably as an effort to counteract the electrostatic effects. However, the direct electrostatic effect is stronger, and therefore, the net effect of the protein is to stabilise the HisH state.

Finally, we have extrapolated the QM/MM energy with a single-point B3LYP calculation with large basis set. Those results are shown in the fourth column in Table 3 (E_{Ext}) and show that B3LYP+large basis set stabilise the HisH state by another 13–21 kJ/mole, except for the two compound I models with a protonated distal His, which are hardly affected by B3LYP. These are the best QM/MM energies and they indicate that the HisH state should be 29–64 kJ/mole more stable than the AspH state, again except the two compound I models with HIP, for which the difference is 15–17 kJ/mole. The reason why these two models behave differently is that they are the only models for which the AspH state is a local minimum.

We have also calculated QM/MM energies for an intermediate along the proton transfer reaction coordinate, viz. a state with a O_{Asp}–H bond length of 1.35 Å, to get an indication about the barrier for this reaction (E_{Int} in Table 3). However, this energy is always lower than the energy of the AspH state, indicating that the energy rises monotonically throughout the reaction (except of course for the two compound I with a protonated distal His).

QTCP calculations

Considering the large electrostatic effects of the surrounding protein in the QM/MM calculations, it is motivated to refine those energies with better methods, because it is often observed that a point-charge model exaggerates electrostatic effects. Therefore, we run a set of QTCP calculations, based on the QM/MM structures.

As can be seen from Eqn. 2, the QTCP free energy consists of three terms: one MM free energy from the reactant to the product (ΔA_{MM}(A→B) in Eqn. 2) and two MM→QM/MM free energies at the reactant and product states, respectively. Moreover, the former is run in both directions, giving both a forward and a reverse free energy with opposite signs. We have also run all perturbations in two steps, via an intermediate with the H–O_{Asp} distance restrained to 1.35 Å (to improve the convergence and get a feeling of the activation barrier of the reaction). Finally, we extrapolated the QTCP energies with a B3LYP calculation with large basis set [59,60]. All these energy terms are collected in Table 4.

First, it can be seen that the MM free energies are reasonably converged, even if we only use one intermediate step in the reaction: The difference between the forward and reverse free energy is less than 2 kJ/mole (0.7 kJ/mole on average). The MM→QM/MM free energies are somewhat less converged, with a six times larger standard deviation than the MM free energies. Therefore, the statistical uncertainty in the total energies can be estimated to ~5 kJ/mole.

It can be seen that the final extrapolated energies are all positive, indicating that the HisH state is more stable than the AspH state for all models (by 12–44 kJ/mole). However, the energy differences are 9–41 kJ/mole smaller than in QM/MM, except for the two compound I complexes, for which the difference is minimal (<3 kJ/mole), and the two compound I models with a protonated distal His, for which the QTCP energy difference actually larger (by 21–27 kJ/mole) than the QM/MM energy difference. As an effect, all compound I models get a similar energy difference. This confirms that QM/MM tends to exaggerate electrostatic effects, probably owing to the fact that dynamic effects are missing and solvation is inadequate. The MM free energies, which include the stabilising effect of the protein are always positive, but only by 3–21 kJ/mole. They are typically dominated by electrostatic interactions, especially with the distal His residue (when charged), the distal arginine residue (Arg-47), a protonated His residue that forms an ionic pair with one of the propionate side chains of the porphyrin (His-180), and a water molecule rather close to the proximal His residue.

Finally, we note that the QM/MM-FE approach [60,73,74,75], which estimates the free energy from a sum of the MM free energy an a single-point QM calculation, thereby avoiding 79 QM calculations for each state, gives almost the same results as the full QTCP approach; the two methods differ by at most 14 kJ/mole, with a mean absolute difference of 4 kJ/mole.

The resting Fe^{III} state has the smallest energy difference, and this is the only state for which the energy difference is so small that it may be within the uncertainty of the method. Therefore, we must conclude that the theoretical calculations quite consistently predict that the proton will stay on the His ligand in all states during the catalytic cycle.

Quantum refinements

Finally, we have investigated whether it is possible to decide the position of the shared

proton using quantum refinement. Quantum refinement is essentially standard crystallographic refinement, in which the MM potential (which is used in all, except the most accurate structures to supplement the crystallographic raw data and give accurate bond lengths and angles) is replaced by QM calculations for a small, but interesting, part of the structure [64]. It has been shown that quantum refinement may improve crystal structures locally [76] and that it can be used to decide the protonation state of metal-bound solvent molecules in medium-resolution crystal structures [77,78].

We have employed this method on a 1.7-Å resolution crystal structure of cytochrome *c* peroxidase from yeast, which is in the resting Fe^{III} state with a weakly bound water molecule (Fe–O distance of 2.40 Å) [69]. We re-refined the structure using the crystallographic raw data for both the HisH and AspH protonation state. Then, we tried to decide which structure fits the experimental electron density best by looking at a number of different quality criteria, viz. the crystallographic *R* factors, the difference of the QM system geometry and energy, compared to the structure optimised in vacuum, and the electron density maps [77,78]. The calculations were performed on both models with and without the porphyrin side chains in the quantum system.

The results are gathered in Table 5. It can be seen that there are only small differences between the structures optimised in vacuum and in the protein, less than 0.03 Å for the Fe–N_{His} bond and 0.06 Å for the Fe–N_{Por} bonds. This is in accordance with our and others finding that metal–ligand distances calculated by density functional method agree with accurate experimental ones to within 0.07 Å [79,80]. However, for the Fe–O_{Wat} bond, the difference is larger, up to 0.15 Å. The reason for this is that the axial water molecule is so weakly bound that its position is determined more by the surroundings than by the Fe–O interaction. Similar results have been obtained with other weakly bound axial ligands [80]. The distance of the Fe ion out of the porphyrin plane is also smaller in the protein than in vacuum, by ~0.12 Å.

The distances to the shared proton is also well preserved in the refinements: The shortest bond is exactly the same in the two calculations without side chains (where both states are local minima, both in vacuum and in the protein). However, in the calculations with side chains, the preferred structure changes from the HisH state in vacuum to the AspH state in the protein and only one state is a local minimum (the other state was obtained by a constrained optimisation). The reason for this is that the two states are so close in energy (the energy difference is ~1 kJ/mole both in vacuum and in the protein).

Interestingly, there are larger differences between the quantum refined structures and the original crystal structure: Both the Fe–N_{His} and Fe–N_{Por} distances are shorter in the original crystal structure, by 0.09–0.15 Å and 0.03–0.09 Å, respectively. In fact, the Fe–N_{Por} distances in the original crystal structure are shorter than what is expected for a HS Fe^{III} ion. This indicates that the original refinement used a poor force field for the Fe ion. Alternatively, it might indicate that the iron ion is partly in the LS state, but there is no indication that the quantum-refined structures have too long Fe–N_{Por} distances from the electron density maps (Figure 3). Therefore, we expect that the detailed distances in re-refined structures are more accurate (the average error in distances in crystal structures of this resolution is ~0.3 Å [81,82]). This is also supported by the fact that the quantum-refined structures have a lower *R* factor than the original crystal structure.

The results in Table 5 show that there are only minor differences in the geometries of the two protonation states, besides the position of the shared proton. The largest difference is observed for the Fe–N_{His} distance in the calculations without side chains (0.04 Å) and this difference is reduced to 0.01 Å in the calculation with side chains (note that the vacuum structures behave in a similar way).

The crystallographic *R* factor indicates that the HisH structure fits the experimental structure factors slightly better. However, the difference is very small, 0.18669 compared to ~0.18674. This illustrates that the *R* factor is a global quantity that is insensitive to small local differences in the structure.

Therefore, we have also calculated the residue (real-space) *R* factor for the four residues in the QM system (His-175, Asp-235, the haem group, and the water ligand). It typically shows a much larger variation. Unfortunately, for the present cases, it does not show any difference between the HisH and AspH states in any of the calculations (irrespectively if it was based on a normal or omit electron-density map). This indicates that the actual differences between the two states are very small. The extensive difference in the residue *R* factor between the calculations with or without side chains is caused entirely by the water ligand, which seems to be better described in the calculations with side chains.

Next, we looked at the differences between the structures optimised in vacuum and in the protein for the two protonation states. From Table 5 (column Δr_{QM}), it can be seen he average difference in the Fe–ligand distances is 0.002–0.007 Å larger for the AspH state than for the HisH state (somewhat more if the floppy Fe–O distance is included).

A more general estimate is offered by the strain energy (ΔE_{QM1}), i.e. the energy difference of the QM system optimised in vacuum and in the protein. As can be seen from the last column in Table 5, it is always lower for the AspH state (by 1–9 kJ/mole). This indicates that this state actually fits the experimental electron density better than the HisH state. Of course, this energy includes the intrinsic (QM) difference between the two protonation states, which is small with side chains, but 10 kJ/mole without the side chains. If we instead compare with the best vacuum structure of each protonation state, the AspH structure is still lowest with side chains (by 2 kJ/mole), whereas the other state is 1 kJ/mole lower without side chains.

Finally, we have also studied the electron-density maps calculated from the experimental structure factors and from the quantum refined structures. Figure 3 shows a superposition of the structures of the two protonation states (both with and without side chains). It can be seen that there are only minor differences in the position of the atoms, primarily for the His and Asp groups. The figure also shows the difference ($f_o - f_c$) electron density maps at the 3σ level for the two protonation states (i.e. the significant differences between the experimental data and the quantum refined structures). It can be seen that for most features, the blue and green, as well as the red and yellow volumes are identical. However, there are two positive features, for which the blue volumes are significantly larger than the green ones, one close to the coordinating atom of the His ligand and one close to the C^{β} atom of Asp-235 (most pronounced for the calculations without side chains, but discernible also in the quantum refinement with side chains in the QM system). This indicates that the AspH model actually fits the experimental electron density slightly better than the HisH model. However, this difference is not reflected in the residue *R* factor, indicating that it may depend on details of the map calculation and the density level considered.

In conclusion, the quantum refinements give only small differences between the two protonation states and the various quality criteria give differing results, especially when calculations both with and without side chains are considered. This indicates that the two protonation states are too similar to be discerned by quantum refinement of this structure.

Conclusions

In this paper, we have presented a thorough theoretical investigation of the protonation status of the Fe–His–Asp motif in all reaction intermediates for cytochrome *c* peroxidase. We have shown that although we move a proton only 0.3–0.6 Å along a hydrogen bond, the position of this shared proton is strongly sensitive to details of the calculations, including effects from the surrounding protein, which may change the relative energies by up to 40 kJ/mole.

Therefore, this at first seemed to be a proper project for the quantum refinement approach, which has successfully been used to decide the protonation status of metal-bound solvent molecules [77,78]. Unfortunately, it turned out that the geometric difference between the two protonation states was too small to obtain any reliable trends in the refinements (cf. Table

5). Therefore, we did not pursue such an investigation for the other states in the peroxidase reaction mechanism.

Instead, we tried to deduce the protonation by comparing the energies of the two states. In vacuum, the AspH state is often the most stable state, especially if the porphyrin side chains are not included. However, with the side chains or in a continuum solvent, the HisH state is most stable, except perhaps for the resting Fe^{III} state.

In the protein, the HisH state is even more stabilised and in the QM/MM calculations, this state is 15–64 kJ/mole more stable than the AspH state for all models studied. If dynamic and entropic effects are included by the QTCP approach, the energy difference is reduced somewhat, but all models are still most stable in the HisH state (by 12–44 kJ/mole). These QM/MM free energy perturbation results are close to the best that can be obtained with today's computer resources and software.

Thus, our computational results indicate that the proton shared in the Fe–His–Asp motif resides on the His ligand in all states in the peroxidase mechanism. This is in agreement with previous theoretical studies on compound I, based on DFT calculations [18,19,20,21,23]. However, the results show that the relative energies of the two protonation states in the various intermediates vary by 32 kJ/mole. Thus, it is normally not enough to study the proton transfer for a single intermediate and extrapolate the results for all other intermediates. Moreover, our results show that the protein may change energies obtained with isolated models in vacuum by over 60 kJ/mole and even QM/MM results may change by over 40 kJ/mole if dynamic and entropic effects are considered. This has strong bearings on the theoretical study of other metalloproteins.

Our results are also in agreement with experimental NMR studies for the resting Fe^{III} state [27,28]. However, NMR studies of the $Fe^{III}CN^-$ state have indicated that the proton has moved to Asp [29,30,31,32], although the effect is mostly discussed in gradual terms (imidazolate character) and it is sensitive to mutations on both the proximal and distal side of the haem ring [31]. This was the reason why we included this state in our investigation, but we see no evidence for such a movement: Our QTCP predictions indicate that the HisH state is 32 kJ/mole more stable for this model and none of the calculations indicate any increased tendency towards the AspH state for the CN^- complex. Further investigations are needed to settle this discrepancy, although mechanistically, the CN^- complex is of minor interest.

Acknowledgements

This investigation has been supported by funding from the research school in pharmaceutical science (FLÄK), the Swedish Research Council, and by computer resources of LUNARC at Lund University.

References

1 Fraústo da Silvia, J.J.R; Williams, R.J.P. The biological chemistry of the elements; Clarendon Press: Oxford, 1994.
2 Kaim, W.; Schwederski, B. Bioinorganic chemistry: inorganic elements in the chemistry of life.; John Wiley & Sons: Chichester, 1994.
3 Poulos, T.L. J. Biol. Inorg. Chem. 11996, 1, 356-359.
4 Rydberg, P.; Sigfridsson, E.; Ryde, U. J. Biol. Inorg. Chem. 2004, 9, 203-223.
5 Jensen, K.P.;Ryde, U. Mol. Phys. 2003, 101, 2003-2018.
6 Jensen, K.P.;Heimdal, J.;Rydberg, P.; Ryde, U.; in: Morokum K.; Musaev, J. (Eds.). Computational modeling of principles and mechanisms of transition metal based homogeneous catalytic processes; John Wiley & Sons: Chichester, in press 2007.
7 Peisach, J. Ann. N. Y. Acad. Sci, 1975, 244, 187,
8 Morrison, M.;Schonbaum, R. Annu. Rev. Biochem. 1976, 45, 881.
9 Gajhede, M.;in: Messerschmidt, A.; Huber, R.; Poulos, T.;Wieghardt, K.; (Eds.), Handbook of Metalloproteins; John Wiley & Sons: Chichester, 2001; pp. 195-209.
10 Hersleth, H.-P.; Ryde, U.;Rydberg, P.; Görbitz, C.H.;Andersson, K.K. J. Inorg. Biochem. 2006, 100, 460-476.
11 Behan, R.K.; Green, M.T. J. Inorg. Biochem. 2006, 100, 448-459.
12 Stanford, M.A.; Swartz, J.C.; Phillips, T.E.; Hoffman, B.M. J. Am Chem. Soc. 102 1980, 102, 4492-4499.
13 Nicholls, P. Biochim. Biophys. Acta 1962, 60, 217-226.
14 Morrison, M.;Schonbaum, R.G. Annu. Rev. Biochem. 1976, 45, 861-888.
15 Ullman, G.M.; Knapp, E.-W. Eur. Biophys. J. 1999, 28, 533-551.
16 Baes, C.F.; Mesmer, R.E. The hydrolysis of cations; Wiley: New York, 1976
17 Du, P.; Loew, G.H. Biophys. J. 1995, 68, 69-80.
18 D.L. Harris, Loew, G.H. J. Porphyr. Phthalocya. 2001, 5, 334-344.
19 Wirstam, M.; Blomberg, M.R.A.; Siegbahn, P.E.M. J. Am. Chem. Soc. 1999, 121 10178-10185.
20 Menyhárd, D.K.;Náray-Szabó, G. J. Phys. Chem. B 1999, 103, 227-233.
21 de Visser, S.P.; Shaik, S.; Sharma, P.K.;Kumar, D.; Thiel, W. J. Am. Chem. Soc. 2003, 125, 15779-15788.
22 Green, R. J. J. Am. Chem. Soc. 2000, 122, 9495-9499.
23 Derat, E.; Cohen, S.; Shaik, S.; Altun, A.; Thiel, W. J. Am. Chem. Soc. 2005, 127 13611-13621.
24 Mincey, T.; Traylor, T.G. J. Am. Chem. Soc. 1979, 101, 765-766.
25 Stein, P.; Mitchell, M.; Spiro, T.G. J. Am. Chem. Soc. 1982, 102, 7795-7797.
26 Desbois, A.; Mazza, G.; Stetzkowski, F.; Lutz, M. Biochim. Biophys. Acta. 1984, 785, 161-176.
27 La Mar, G.N.; de Ropp, J.S. Biochem. Biophys. Res. Commun. 1979, 90, 36-41.
28 La Mar, G.N.; de Ropp, J.S. J. Am. Chem. Soc. 1982, 104, 5203-5206.
29 de Ropp, J.S.; Thanabal, V.; La Mar, G.N. J. Am. Chem. Soc. 1985, 107, 8268-8270.
30 Thanabal, V.; de Ropp, J.S.; La Mar, G.N. J. Am. Chem. Soc. 1988, 110, 3027-3035.
31 Satterlee, J.D.; Erman, J.E.; Mauro, J.M.; Kraut, J. Biochemistry 1990, 29, 8797-8804.
32 Banci, L.; Bertini, I.; Turano, P.; Ferrer, J.C.; Mauk, A.G. Inorg. Chem. 1991, 30, 4510-4516.
33 Becke, A.D. Phys. Rev. A 1988, 38, 3098-3100.
34 Perdew, J.P. Phys. Rev. B 1986, 33, 8822-8824.
35 Treutler, O.; Ahlrichs, R. J. Chem. Phys. 1995, 102, 346-354.
36 Hehre, W.J.; Radom, L.; Schleyer, P.v.R.; Pople, J.A. Ab initio molecular orbital theory; Wiley-Interscience: New York, 1986.
37 Schäfer, A.; Horn, H.; Ahlrichs, R. J. Chem. Phys. 1992, 97, 2571-2577.
38 Schäfer, A.; Huber, C.; Ahlrichs, R. J. Chem. Phys. 1994, 100, 5829-5835.

39 Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. Chem. Phys. Lett. 1995, 240, 283-290.
40 Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. Theor. Chem. Acc. 1997, 97, 119-124.
41 Becke, A.D. J. Chem. Phys. 1993, 98, 1372-1377.
42 Hertwig, R.H.; Koch, W. Chem. Phys. Lett. 1997, 268, 345-351.
43 Klamt, A.; Schüürmann, J. J. Chem. Soc. Perkin Trans. 1993, 2, 799-805.
44 Schäfer, A.; Klamt, A.; Sattel, D.; Lohrenz, J.C.W.; Eckert, F. Phys. Chem. Chem. Phys. 2000, 2, 2187-2193.
45 Klamt, A.; Jonas, V.; Bürger, T.; Lohrenz, J.C.W. J. Phys. Chem. A 1998, 102, 5074-5085.
46 Ryde, U. J. Comput.-Aided Mol. Design 1996, 10, 153-164.
47 Ryde, U.; Olsson, M.H.M. Int. J. Quantum Chem. 2001, 81, 335-347.
48 Reuter, N.I.; Dejaegere, A.;Maigret, B.; Karplus, M. J. Phys. Chem. A 2000, 104, 1720-1735.
49 Svensson, M.; Humbel, S.; Froese, R.D.J.; Matsubara, T.; Sieber, S.; Morokuma, K. J. Phys. Chem. 1996, 100, 19357-19363.
50 Case, D. A.; Darden, T. A.; Cheatham, III, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Wang, B.; Pearlman, D. A.; Crowley, M.; Brozell, S.; Tsui, V.; Gohlke, H.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Schafmeister, C.; Caldwell, J. W.; Ross, W. S.;Kollman, P. A. *AMBER 8*, 2004, University of California, San Francisco.
51 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.
52 Wang, J.; Cieplak, P.; Kollman, P.A. J. Comput. Chem. 2000, 21, 1049-1074.
53 Besler, B.H.; Merz, K.M.; Kollman, P.A. J. Comput. Chem. 1990, 11, 431
54 Ryckaert, J.P.; Ciccotti, G.; Berendsen, H.J.C. J. Comput. Phys. 1977, 23, 327-341.
55 Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.; Impey, R.W.; Klein, M.L. J. Chem. Phys. 1983, 79, 926-935.
56 Darden, T.; York, D.; Pedersen, L. J. Chem. Phys. 1993, 98, 10089-10092.
57 Essmann U.; Perera, L.; Berkowitz, M.L.; Darden, T.; Lee, H.; Pedersen, L.G. J. Chem. Phys. 1995, 103, 8577-8593.
58 Berendsen, H.J.C.; Postma, J.P.M.; van Gunsteren, W.F.; DiNola, A.; Haak, J.R. J. Chem. Phys. 1984, 81, 3684-3690.
59 Rod, T.H.; Ryde, U. Phys. Rev. Lett. 2005, 94, 138302.
60 Rod, T.H.; Ryde, U. J. Chem. Theory Comput. 2005, 1, 1240-1251.
61 Muller, R.P.; Warshel, A. J. Phys. Chem. 1995, 99, 17516-17524.
62 Olsson, M.H.M.; Hong, G.; Warshel, A. J. Am. Chem. Soc. 2003, 125, 5025-5039.
63 Chan, H.S. Proteins 2000, 40, 543-571.
64 Ryde, U.; Olsen, L.; Nilsson, K. J. Comput. Chem. 2002, 23, 1058-1070.
65 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.
66 Kleywegt, G.J.; Jones, T.A. Methods Enzymol. 1997, 277, 208-230.
67 Pannu, N.S.; Read, R.J. Acta Cryst. A 1996, 52, 659-668.
68 Engh, R.A.; Huber, R. Acta Cryst. A 1991,47, 392-400.
69 Finzel, B.C.; Poulos, T.L.; Kraut, J. J. Biol. Chem. 1984, 259, 13027-13036.
70 Schöneboom, J.C.; Cohen, S.; Lin, H.; Shaik. S.; Thiel, W. J. Am. Chem. Soc. 2004, 126, 4017-4034.
71 Altun, A.; Guallar, V.; Friesner, R.A.; Shaik, S.; Thiel, W. J. Am. Chem. Soc. 2006, 128,

3924-3925.
72 Altun, A.; Shaik, S.; Thiel, W. J. Comput. Chem. 2006, 27, 1324-1337.
73 Chandrasekhar, J; Jorgensen, W. L. J. Am. Chem. Soc. 1985, 107, 2974-2975.
74 Zhang, Y.; Liu, H.; Yang, W. J. Chem. Phys. 2000, 112, 3483-3492.
75 Ishida, T.; Kato, S. J. Am. Chem. Soc. 2003, 125, 12035-12048.
76 Ryde, U.;Nilsson, K. J. Am. Chem. Soc. 2003,125, 14232-14233.
77 Nilsson, K.; Ryde, U. J. Inorg. Biochem. 2004, 98, 1539-1546.
78 Rulíšek, L.; Ryde, U. J. Phys. Chem. B 2006, 110, 11511-11518.
79 Neese, F. J. Biol. Inorg. Chem. 2006, 11, 702-711.
80 Ryde, U. Dalton Trans. 2007 607-625.
81 Fields, B.A.; Bartsch, H.H.; Bartunik, H.D.; Cordes F.; Guss, J.M.; Freeman, H.C. Acta Cryst. D 1994, 50, 709-730.
82 Cruickshank, D.W.J. Acta Cryst. D 1999, 55, 583-601.

Table 1. Energy differences (kJ/mole) between models where the proton resides on Asp or on His. A negative sign indicates that the AspH structure is most stable. The energies were either calculated at the BP86/6-31G* (method 1) in vacuum or in a continuum solvent with a dielectric constant of 80, or at the B3LYP/6-311+G(2d,2p) (method 2) level.

Model	Energy difference (kJ/mole)					
Side chains	with side chains			without side chains		
Method	1	1	2	1	1	2
Dielectric constant	1	80	1	1	80	1
$Fe^{III}H_2O$ LS	1.1[a]	0.9	-3.3	-21.3[b]	-1.9	-24.2
$Fe^{III}H_2O$ HS	1.5[a]	2.8	-1.7	-10.0	0.4	-26.0
Fe^{III} HS	-2.6[b]	1.3	-2.8	-12.6	-3.7	-30.9
$Fe^{III}HO_2^-$ LS	15.5[a]	11.1	20.1	1.3[a]	5.9	3.9
$Fe^{V}O^{2-}$ LS	8.5[a]	11.0	13.5	1.5	5.6	-14.3
$Fe^{V}O^{2-}$ IS	10.2[a]	9.5	13.9	6.2	9.7	-17.6
$Fe^{IV}O^{2-}$ IS	18.5[a]	13.4	23.2	3.0[a]	8.6	6.0
$Fe^{IV}OH^-$ IS	4.9[a]	4.0	7.2	-8.8	-2.2	-35.4
$Fe^{III}CN^-$ LS	16.8[a]	9.7	17.9	-3.1[b]	6.6	7.3

[a] The AspH state was obtained by constraining the O–H distance to 1.13 Å
[b] The HisH state was obtained by constraining the N–H distance to 1.10 Å

Table 2. Geometries and spin densities in the optimised QM/MM structures. Int is an intermediate state with the H–O_{Asp} bond constrained to 1.35 Å. HIP means that the distal His ligand was assumed to be protonated on both nitrogen atoms. Constrained bond lengths are shown in bold face.

Model	State	Distances (Å)				Spin density (*e*)				
		Fe–N_{Por}	Fe–N_{His}	Fe–O	H–O	Fe	Porphyrin	Ligand	His	Asp
$Fe^{III}H_2O$ HS	HisH	2.07	2.05	2.31	1.57	4.14	0.55	0.05	0.16	0.04
	Int	2.08	2.04	2.32	**1.35**	4.15	0.55	0.05	0.19	0.00
	AspH	2.08	2.01	2.34	**1.13**	4.14	0.51	0.04	0.24	0.00
$Fe^{III}O_2H^-$ LS	HisH	2.02	2.01	1.77	1.65	0.88	-0.08	0.22	-0.01	0.00
	Int	2.02	2.02	1.77	**1.35**	0.89	-0.10	0.22	-0.02	0.00
	AspH	2.02	2.01	1.78	**1.13**	0.90	-0.12	0.22	-0.02	0.00
$Fe^{V}O$ LS	HisH	2.02	2.11	1.62	1.68	0.99	-0.48	0.66	-0.04	-0.18
	Int	2.02	2.08	1.62	**1.35**	0.96	-0.49	0.66	-0.13	-0.01
	AspH	2.02	2.07	1.62	**1.13**	0.95	-0.40	0.67	-0.26	0.00
$Fe^{V}O$ IS	HisH	2.02	2.11	1.65	1.70	1.42	0.46	0.94	-0.03	0.20
	Int	2.02	2.12	1.65	**1.35**	1.38	0.52	0.93	0.14	0.02
	AspH	2.02	2.12	1.65	**1.13**	1.37	0.39	0.93	0.31	0.00
$Fe^{V}O$ HIP LS	HisH	2.02	2.11	1.63	1.60	1.14	-0.55	0.71	-0.05	-0.25
	Int	2.02	2.13	1.64	**1.35**	1.13	-0.55	0.71	-0.22	-0.11
	AspH	2.02	2.11	1.64	1.08	1.12	-0.40	0.71	-0.45	0.00
$Fe^{V}O$ HIP IS	HisH	2.02	2.12	1.66	1.72	1.42	0.42	0.85	-0.02	0.29
	Int	2.02	2.13	1.66	**1.35**	1.42	0.47	0.85	0.20	0.12
	AspH	2.02	2.13	1.66	1.08	1.41	0.38	0.85	0.43	0.00
$Fe^{IV}O$ IS	HisH	2.02	2.12	1.65	1.69	1.17	-0.02	0.91	-0.02	0.00
	Int	2.02	2.12	1.65	**1.35**	1.17	-0.02	0.91	-0.02	0.00
	AspH	2.02	2.11	1.66	**1.13**	1.17	-0.02	0.91	-0.02	0.00
$Fe^{IV}OH^-$ IS	HisH	2.01	2.05	1.78	1.64	1.56	0.05	0.26	-0.01	0.09
	Int	2.01	2.04	1.79	**1.35**	1.62	0.08	0.26	0.02	0.00
	AspH	2.01	2.02	1.79	**1.13**	1.62	0.04	0.26	0.06	0.00
$Fe^{III}CN^-$ LS	HisH	2.01	1.98	1.87	1.60	0.97	0.01	-0.01	0.03	0.00
	Int	2.01	1.98	1.87	**1.35**	0.97	-0.01	-0.01	0.05	0.00
	AspH	2.01	1.97	1.88	**1.13**	0.95	-0.02	-0.01	0.08	0.00

Table 3. The relative energies (kJ/mole) of the HisH and AspH states in the QM/MM calculations of cytochrome *c* peroxidase. Int is the extrapolated energy of an intermediate state with the O_{Asp}–H distance fixed to 1.35 Å. $E_{QM/MM}$ is the total QM/MM energy, defined in Eqn. 1, $E_{QM1+ptch}$ and E_{QM1} are the QM energy of the QM system, with or without a point-charge model of the surrounding protein, respectively. E_{Ext} is the QM/MM energy, extrapolated by the B3LYP method and a larger basis set and E_{Int} is the corresponding extrapolated QM/MM energy of an intermediate with a H–O_{Asp} distance of 1.35 Å. E_{Vac} and E_{Wat} are the corresponding vacuum and continuum water results for models with porphyrin side chains from Table 1. A positive energy indicates that the HisH state is most stable.

Model	$E_{QM/MM}$	$E_{QM1+ptch}$	E_{QM1}	$\boldsymbol{E_{Ext}}$	E_{Int}	E_{Vac}	E_{Wat}
$Fe^{III}H_2O$ HS	15.5	16.9	-5.8	**28.9**	12.6	1.5	2.8
$Fe^{III}O_2H^-$ LS	43.8	47.4	14.7	**58.8**	24.9	15.5	11.1
$Fe^{V}O$ LS	20.0	17.8	3.7	**40.9**	12.6	8.5	11.0
$Fe^{V}O$ HIP LS	12.0	9.1	6.4	**14.9**	16.7	8.5	11.0
$Fe^{V}O$ IS	23.7	24.8	6.8	**41.3**	20.3	10.2	9.5
$Fe^{V}O$ HIP IS	16.4	16.7	6.7	**16.6**	18.9	10.2	9.5
$Fe^{IV}O$ IS	48.0	50.6	17.9	**63.8**	29.0	18.5	13.4
$Fe^{IV}OH^-$ IS	19.2	20.2	-1.0	**32.1**	16.8	4.9	4.0
$Fe^{III}CN^-$ LS	25.8	28.3	15.5	**40.9**	17.5	16.8	9.7

Table 4. QTCP free energies (kJ/mole) for the proton transfer from the HisH form to the AspH form, via an intermediate (Int) with H–O_{Asp} constrained to 1.35 Å. The total free energy is a sum of a MM free energy, run both in the forward and reverse directions and a MM→QM/MM free energy (QM), which is given relative that of the HisH state. In the second step, the total energy also includes the MM energy of the previous step. The extrapolated energy is the total energy plus the energy correction from a B3LYP calculation with the large basis set. A positive energy indicates that the HisH state is most stable.

Model	State	QM	Forward	Reverse	Mean	Total	Extrapolated
$Fe^{III}H_2O$ HS	Int	-3.3	4.9	-4.1	4.5	1.2	8.5
	AspH[a]	-13.6	7.0	-8.3	7.6	-1.4	12.0
$Fe^{III}HO_2^-$ LS	Int	1.5	5.0	-5.4	5.2	6.6	15.8
	AspH[a]	9.9	5.0	-7.0	6.0	21.1	36.1
$Fe^{V}O$ LS	Int	10.7	3.7	-3.8	3.8	14.5	14.8
	AspH[a]	10.2	2.7	-3.9	3.3	17.2	38.2
$Fe^{V}O$ IS	Int	24.1	5.2	-5.4	5.3	29.4	34.3
	AspH[a]	19.9	0.8	-0.9	0.9	26.0	43.7
$Fe^{V}O$ HIP LS	Int	3.9	3.9	-3.9	3.9	7.7	5.5
	AspH[a]	19.4	9.4	-8.8	9.1	32.5	35.4
$Fe^{V}O$ HIP IS	Int	9.3	6.1	-6.7	6.4	15.7	17.3
	AspH[a]	22.6	15.3	-14.2	14.8	43.8	44.0
$Fe^{IV}O$ IS	Int	6.4	4.3	-3.9	4.1	10.5	20.5
	AspH[a]	-0.6	3.7	-4.3	4.0	7.5	23.3
$Fe^{IV}OH^-$ IS	Int	6.4	2.1	-1.4	1.7	8.1	15.7
	AspH[a]	-0.2	6.4	-6.7	6.6	8.1	21.0
$Fe^{III}CN^-$ LS	Int	5.7	1.5	-0.9	1.2	6.9	15.4
	AspH[a]	13.5	2.3	-1.4	1.9	16.5	31.6

[a] This state was obtained by constraining the H–O_{Asp} distance to 1.13 Å.

Table 5. Results of the quantum refinements of cytochrome *c* peroxidase in the resting HS $Fe^{III}H_2O$ state with two different protonation states and with or without side chains on the porphyrin ring. Distances in bold face were constrained during the optimisation. Fe oop is the average distance of the Fe ion out of the porphyrin plane defined by the four nitrogen atoms. The corresponding results for QM optimisations in vacuum are also included. Δr_{QM} is the average difference in the Fe–ligand distances and ΔE_{QM1} is the energy difference of the QM system, optimised in vacuum or in the crystal structure.

Side chain?	State	Distance to Fe			Fe oop	N–H	O–H	*R* factor	residue *R*	Δr_{QM}	ΔE_{QM1}
		N_{His}	N_{Por}	O_{Wat}							
Yes	AspH	2.09	2.06–2.10	2.39	0.20	1.41	1.14	0.18673	0.104	0.029	118.5
		2.08	2.06–2.16	**2.39**	0.32	1.40	**1.14**	Vacuum			
	HisH	2.10	2.06–2.10	2.38	0.19	**1.16**	1.39	0.18669	0.104	0.028	119.4
		2.10	2.05–2.16	**2.38**	0.30	1.16	1.40	Vacuum			
No	AspH	2.04	2.04–2.07	2.31	0.21	1.60	1.04	0.18675	0.190	0.036	87.7
		2.00	2.07–2.12	2.56	0.35	1.64	1.04	Vacuum			
	HisH	2.08	2.04–2.07	2.31	0.20	1.12	1.47	0.18669	0.190	0.029	96.7
		2.07	2.07–2.12	2.50	0.31	1.12	1.48	Vacuum			
Crystal		1.95	1.97–2.04	2.40	0.16			0.18732	0.106		

Figure 1. The thermodynamic cycle forming the basis of the QTCP method.

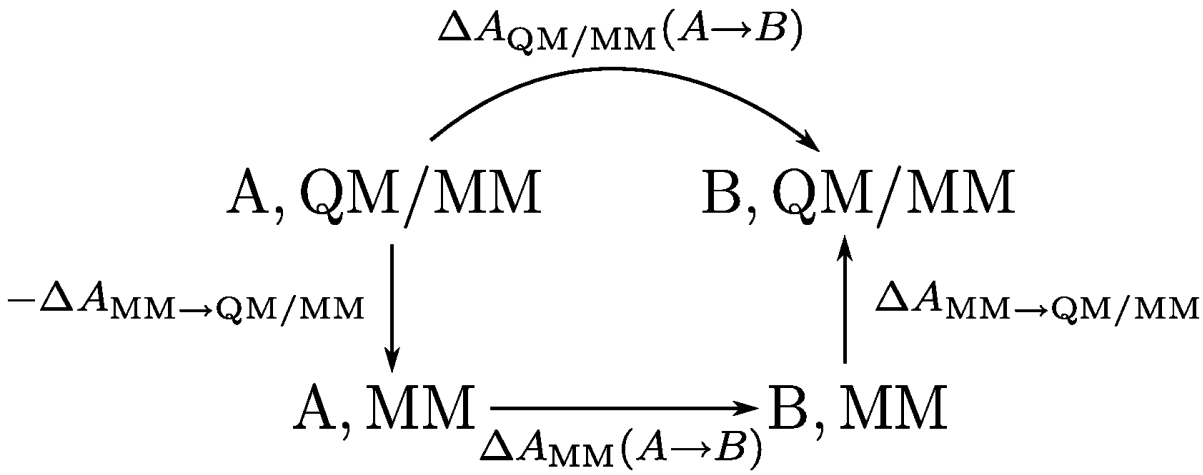

Figure 2. Overlayed structures of the HisH (green) and AspH states of the IS compound I model with porphyrin side chains, optimised in vacuum.

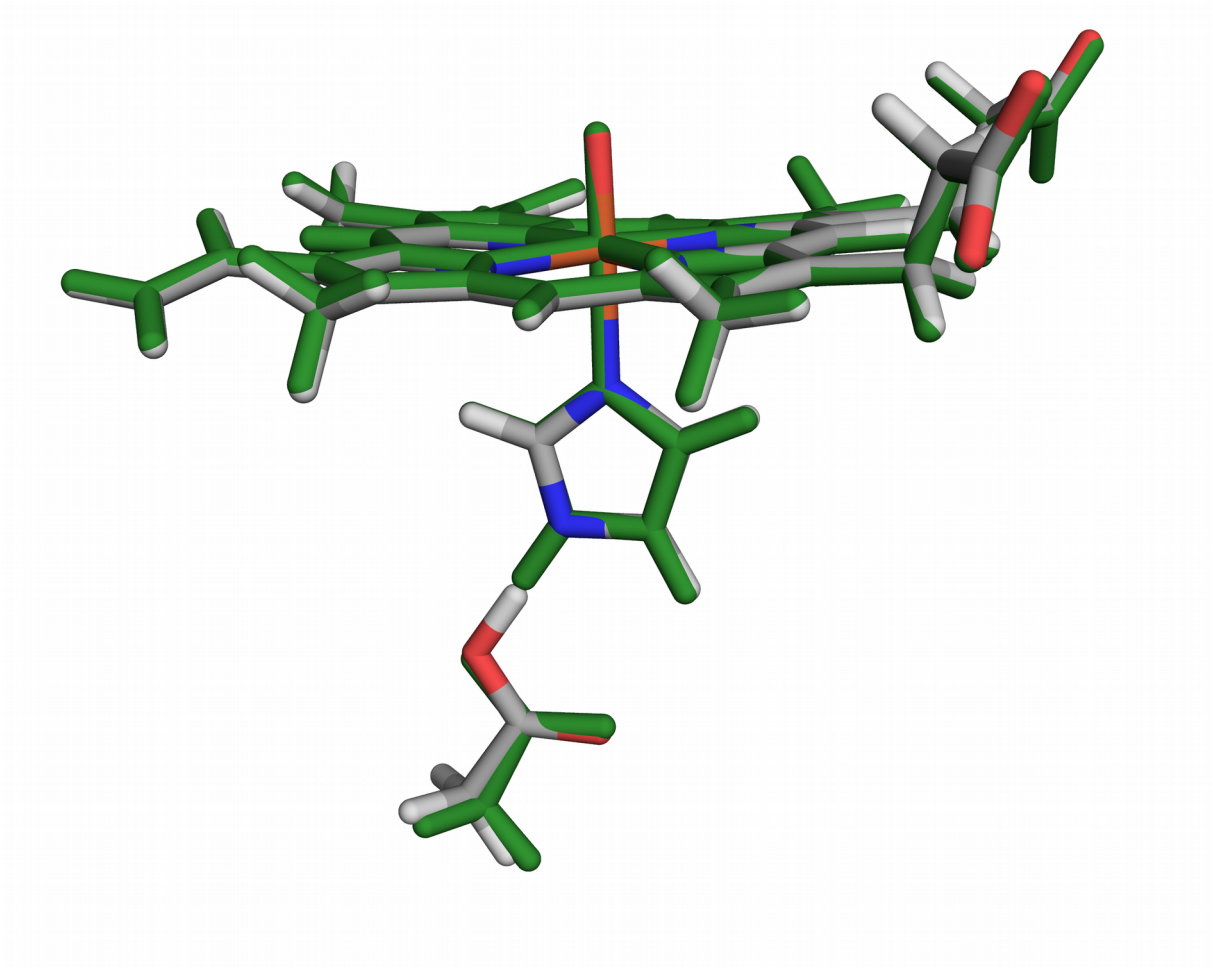

Figure 3. The quantum refined structure of the AspH (green) and HisH (atomic colours) states of cytochrome *c* peroxidase in the resting HS $Fe^{III}H_2O$ state. In addition, the difference ($f_o - f_c$) electron density maps at the 3σ level is shown (blue and red for the HisH state and green and yellow for the AspH state; blue and green for positive densities, red and yellow for negative densities). The top picture is without the porphyrin side chains in the quantum system, the bottom one with them.