Computational modelling of oxygenation processes in enzymes and biomimetic model complexes

de Visser, Sam P.; Quesne, Matthew G.; Martin, Bodo; Comba, Peter; Ryde, Ulf

Published in:
Chemical Communications

DOI:
10.1039/C3CC47148A

2014

Chemical Communications

Dynamic Article Links ►

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

Computational modelling of oxygenation processes in enzymes and biomimetic model complexes

Sam P. de Visser,*[a] Matthew G. Quesne,[a] Bodo Martin,[b] Peter Comba*[b] and Ulf Ryde*[c]

Received (in XXX, XXX) Xth XXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX

DOI: 10.1039/b000000x

With computational resources becoming more efficient and more powerful and at the same time cheaper, computational methods have become more and more popular for studies on biochemical and biomimetic systems. Although large efforts from the scientific community have gone into exploring the possibilities of computational methods on large biochemical systems, such studies are not without pitfalls and often cannot be routinely done but require expert execution. In this review we summarize and highlight advances in computational methodology and its application to enzymatic and biomimetic model systems. In particular, we emphasize on topical and state-of-the-art methodologies that are able to either reproduce experimental findings, e.g., spectroscopic parameters and rate constants, accurately or give predictions on short-lived intermediates and fast reaction processes in nature. Moreover, we give examples of processes where certain computational methods dramatically fail.

Introduction

Nature has a large arsenal of efficient catalysts that are able to perform regioselective as well as stereospecific substrate conversions. These chemical systems, i.e. enzymes, work at ambient temperature and pressure and have specific functions such as biodegradation, bioconversion or biosynthesis of compounds. Enzymatic reactions influence the general health and well-being of the organism and as such are important. Often the enzymes contain a transition metal active site, where the catalysis takes place. As the abundance of iron relative to other metals is large on planet Earth, many enzymes utilize iron in their active site,[1,2] but there are also numerous examples of copper-,[3] vanadium-,[4] molybdenum-,[5] and other transition metal-containing enzymes as well as non-transition metal containing enzymes. In addition, enzymes sometimes use clusters of two or more transition metal atoms as catalytic centre as is, for instance, the case in the diiron enzyme ribonucleotide reductase[6] or the multi-metal cluster in photosystem II.[7]

The mononuclear iron-containing enzymes can be split into the heme[1] and nonheme[2] categories of which we show the active site structure of two typical examples in Fig 1: namely for the heme enzyme cytochrome P450 and the nonheme iron enzyme taurine/α-ketoglutarate dioxygenase.[8,9] The crystal structure coordinates of the former were taken from the human enzyme cytochrome P450 2A13, which is a liver enzyme involved in the detoxification of endogenous compounds. The heme group is linked to the protein via a direct bond of the metal with the thiolate group of Cys_{439}, the axial ligand. It has been proposed that this axial ligand fine-tunes the electronic properties of the oxidant and gives it the functional properties to act as a monoxygenase.[10] In contrast, other heme enzymes, such as peroxidases typically have a histidine group as axial ligand,[11] while catalases use a tyrosinate group.[12] As usual, key intermediates of the catalytic cycles of enzymes are short-lived, and, consequently difficult to detect and characterize. This means that experimental work in the field often needs to be complemented by theoretical studies. Computational modelling has been instrumental in assisting experiment into establishing the fundamental properties and changes upon ligand substitution in heme enzymes and their catalytic features.[13] For instance, more than 10 years before the first characterization of the active species of P450 enzymes, theory proposed a high-valent iron(IV)-oxo heme cation radical (Compound I) that was found to react with substrates via radical mechanisms with low barriers.[14] Only a few years ago, Rittle and Green characterized Compound I using electron paramagnetic resonance (EPR), absorption and Mössbauer spectroscopy.[15]

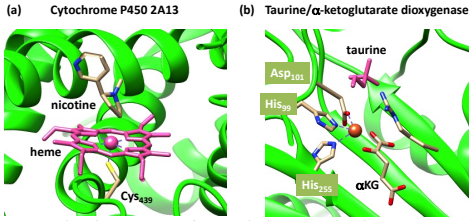


Fig. 1 Active site structures of two typical mononuclear iron enzymes: (*a*) Cytochrome P450 from 4EJG pdb file. (*b*) taurine/α-ketoglutarate dioxygenase from 1OS7 pdb file.

As a comparison, we also show in Fig 1(*b*) the crystal structure of a nonheme iron enzyme, where the metal is bound via a

characteristic facial triad represented by two histidine and an aspartate amino acid. This enzyme is involved in the biodegradation of taurine and functions as a sulphur source. Also here, computational modelling assisted in establishing differences in properties between heme and nonheme iron enzymes and the origin of the efficiency of hydrogen-atom abstraction.[16] However, axial and equatorial ligand effects in substrate oxidation by heme and nonheme iron complexes have been extensively reviewed previously and will not be covered here.[17]

It is clear from these examples that theory can assist experimental studies related to biocatalysis and biotransformations and give insight into reaction mechanisms and pathways. Unfortunately, these studies are not routinely done, but may involve caveats and pitfalls. To fully appreciate and understand the computational decisions on method and model selection, we review here computational approaches with relevance to transition metal catalysis. We will start with a brief overview of the basic methods in the field and their general accuracy and reproducibility and then give some recent examples on how theory guided and assisted experimental work.

Methods

Quantum mechanics and density functional theory

According to quantum mechanics (QM), all properties of a molecular system can be obtained from the wave function (Ψ), which is the solution of the Schrödinger equation,[18] Eq 1, and expresses the total energy of the molecule (E) as an eigenfunction of the Hamilton operator (H).

$$H\,\Psi = E\,\Psi \qquad (1)$$

The Hamilton operator includes the kinetic and potential energy contributions of all the electrons and nuclei within a molecule as described in Eq 2 and 3.[19] Here, $\hbar$ is Planck's constant divided by 2π, m_e is the electron mass, e is the elementary charge and ε_0 represents the permittivity in vacuum, r_{ij} are the distances between nuclei and electrons, whereas the distances between two electrons and two nuclei are described by either r_{ij} and R_{IJ}, respectively. The Hamiltonian in Eq 2 is split into four components that represent the kinetic energy of the electrons, the electron-nuclei attractions, the electron-electron repulsions and the nucleus-nucleus repulsions, respectively. Within the Born-Oppenheimer approximation the nucleus-nucleus interactions is constant and the electronic energy E is calculated for a rigid molecular geometry.

$$H = -\frac{\hbar^2}{2m_e}\sum_i^{N_e}\nabla_i^2 - \sum_i^{N_e}\sum_I^{N_p}\frac{Z_I e^2}{4\pi\varepsilon_0 r_{Ii}} + \frac{1}{2}\sum_{i,j}^{N_e}\frac{e^2}{4\pi\varepsilon_0 r_{ij}} + \frac{1}{2}\sum_{I,J}^{N_p}\frac{e^2}{4\pi\varepsilon_0 R_{IJ}} \qquad (2)$$

$$\nabla_i^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \qquad (3)$$

The solutions to the Schrödinger equation clearly depend on the number of electrons (N_e) and the number of nuclei (N_p) in the chemical system and will increase dramatically in complexity as the size of the system expands. Because of its complexity, the Schrödinger equation can only be solved exactly for the hydrogen atom and the $H_2^{+\bullet}$ ion, while for any larger chemical system approximations need to be used. This has led to the development of a large range of computational chemistry methods that all have different features and accuracies. For accurate results, thorough benchmarking and calibrating is required, especially for systems that contain transition metals whereby the consistency and reproducibility of methods is tested by a comparison to higher-level methods or against experimental data, as shown in several examples below.

The most basic solution to the Schrödinger equation comes from Hartree-Fock (HF) theory, in which it is assumed that each electron moves in the average field of all the other electrons and the one-electron wave functions are expanded into a set of known (hydrogen-like) orbitals; the basis set. The solution of the Schrödinger equation is not the end of the calculation as shown schematically in Fig. 2. Thus, after one has created a starting geometry and selected a method, i.e. a Hamilton operator description, and basis set, or wave function model, one can solve the Schrödinger equation and obtain the energy associated with the starting structure. Subsequently, the geometry is modified and a new energy is calculated. Thereafter, the energy is minimized with respect to all geometric coordinates in the chemical system and when the first and second derivatives of the energy with respect to the degrees of freedom reach a certain threshold value, the structure is considered converged, i.e. an optimized geometry is found. A frequency calculation, *vide supra*, will then establish whether a structure is a local minimum or a higher-order saddle point.

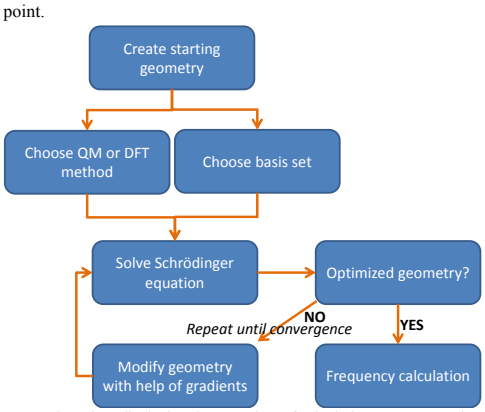


Fig. 2 Flow chart displaying the procedure of calculating a structure using either QM or DFT methods.

The HF method gives reasonable geometries but its energies are not good enough to reproduce chemical reaction mechanisms and experimental barrier heights. Improved results are generally obtained with Møller-Plesset second-order perturbation theory (MP2)[20] that gives much more accurate energies and chemical structures as compared to HF.[20] Even better results can be

obtained with coupled-cluster calculations using singles, doubles and perturbatively treated triples (CCSD(T)),[21] which is currently considered as the gold-standard method in QM.[22] Unfortunately, MP2 and especially CCSD(T) are computationally demanding and even with current resources they can only be applied to chemical systems with less than about 50 and 20 atoms, respectively.[23] For biosystems and biomimetic complexes with often well over 100 atoms computationally cheaper methods are needed.

As an alternative to the wave function methods, DFT was development by Kohn and co-workers in the 1960s.[24] They showed that there is a one-to-one correspondence between the wave function and the electron density, $\rho(r)$. The latter is a functional of the three Cartesian coordinates (r), whereas the wave function is a function of the coordinates of every electron in the chemical system; therefore, this was a major simplification. In DFT, the electronic energy (E_{el}) is calculated from a functional of the electron density, which unfortunately is not known. However, a typical expression is given by the sum of the integrals in Eq 4. The first term contains an integral over the Laplacian over the atomic orbitals and represents the kinetic energy of the electrons. The second and third terms involve the electron-nucleus attractions and the electron-electron repulsions, respectively. The final exchange-correlation term (E_{xc}) has the largest uncertainty as no exact equations are known for this component; it is normally split into individual exchange and correlation components, E_x and E_c, where the exchange energy results from ferromagnetic interactions of two electrons with the same spin (α or β) in different orbitals, while the correlation energy is determined by the electron pairing energies of two electrons in the same molecular orbital.

$$E_{el} = -\frac{1}{2}\sum_i \int \phi_i(\mathbf{r}_1)\nabla^2\phi_i(\mathbf{r}_1)d\mathbf{r}_1 + \sum_A \int \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}_1|}\rho(\mathbf{r}_1)d\mathbf{r}_1 + \frac{1}{2}\int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}d\mathbf{r}_1 d\mathbf{r}_2 + E_{xc} \quad (4)$$

There are many available exchange and correlation energy descriptions and two widely used ones are the Slater exchange (E_x^{Slater}), Eq 5,[25] and the Vosko, Wilks and Nusair correlation energy (E_c^{VWN}), Eq 6.[26] The Slater exchange equation includes a scale factor (α_{ex}) that has a value of $^2/_3$ for a uniform electron gas. The Local Density Approximation (LDA) uses $E_{xc} = E_x^{Slater} + E_c^{VWN}$, which generally gives a good improvement over HF theory, but is not good enough for most chemical systems.

$$E_x^{Slater} = -\frac{9}{4\alpha_{ex}}\left(\frac{3}{4\pi}\right)^{\frac{1}{3}}\sum_\gamma \int \left[\rho_1^\gamma(\mathbf{r}_1)\right]^{\frac{4}{3}} d\mathbf{r}_1 \quad (5)$$

$$E_c^{VWN} = \int \rho_1(\mathbf{r}_1)\varepsilon_c\left[\rho_1^\alpha(\mathbf{r}_1), \rho_1^\beta(\mathbf{r}_1)\right]d\mathbf{r}_1 \quad (6)$$

In the 1990s DFT methodology made a major leap due to the development of functionals that depend on the gradient of the electron density, e.g., E_x^{Becke},[27] and hybrid density functional methods.[28] Becke discovered that calculated energies had a systematic error and showed that the results could be improved by using a fraction of the exchange energy from a HF calculation, E_x^{HF}. He combined it with different exchange and correlation functionals using three fitted parameters (A, B, and C). An example of such a hybrid DFT method is B3LYP, given in Eq 7,[28,29] which has become immensely popular due to its low computational cost and relatively high accuracy. The three parameters A, B, and C were optimized against experimentally determined electron affinities, ionization potentials and proton affinities for a test-set of molecules; technically making it a semi-empirical method.

$$E_{xc}^{B3LYP} = AE_x^{Slater} + (1-A)E_x^{HF} + B\Delta E_x^{Becke} + E_c^{VWN} + C\Delta E_c^{LYP} \quad (7)$$

In recent years it has been found that the B3LYP method has weaknesses, especially for transition metal containing systems, and alternative DFT methods have been suggested. Unfortunately, no universal and consistent method has been developed yet, so that a range of different methods is in current use. For example, it has been shown that a modified hybrid DFT method with 15% E_x^{HF} (rather than the standard 20%), termed B3LYP*, gives better energetics for some transition metal complexes,[30] but as will be discussed below for other properties higher amounts of E_x^{HF} may be needed. In addition, both HF and standard DFT methods cannot describe dispersion interactions. Recent work of the Grimme group has shown that a simple empirical dispersion correction significantly improves calculated energies: B3LYP-D.[31]

Typically, the results of DFT calculations show only limited dependence on the functionals used. However, for transition metal complexes with close-lying spin states the ordering and spin-state energy splitting often strongly depend on the functional chosen, and particularly on the amount of E_x^{HF}. It is of paramount importance, therefore, to calibrate DFT calculations on transition metal complexes to either experimental data or to do test calculations with alternative methods.

Geometries, frequencies and free energies

The solution of the Schrödinger equation is not the end of the calculation as shown schematically in Fig. 2. Thus, after one has created a starting geometry and selected a method and basis set, the QM method gives the energy associated with the starting structure. Thereafter, the energy can be minimized with respect to the geometric coordinates of the chemical system using the first derivatives of the energy with respect to the coordinates. A frequency calculation can establish whether a structure is a local minimum or a higher-order saddle point.

A QM calculation gives the electronic energy of the studied system. In order to convert these energies into values that can be compared to experiments, one needs to add vibrational, entropic and often also solvent effects. The former two can be obtained from a frequency calculation on the optimized geometry, which establishes the second derivatives of the energy with respect to the coordinates, i.e. the degrees of freedom. These derivatives are

used to calculate the vibrational energy levels of a molecule and give the zero-point energy (ZPE), see Fig 3. ZPE is the lowest vibrational energy level (ν_0) in which the system will reside at 0 K temperature. Fig 3 shows the potential energy profile along one degree of freedom (e.g., a reaction coordinate) between two local minima S1 and S2 with electronic energies E_{S1} and E_{S2}, respectively. These two local minima are separated from each other by a transition state (**TS**). Higher energy vibrational levels of the local minimum S1 are given as ν_1, ν_2 and ν_3 in Fig 3. A Boltzmann distribution at a specific temperature will determine the degree of occupation of each individual vibrational energy level.

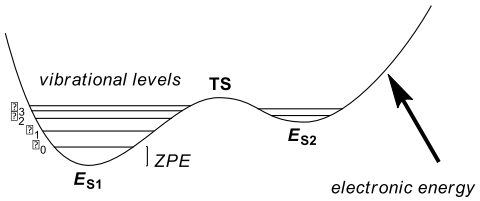


Fig. 3 Potential energy profile between two local minima S1 and S2 each with their own vibrational energy levels (ν) and ZPE.

In addition, using the information of the geometry and the vibrational energy levels the entropy can be calculated from the individual partition functions (Q) for the translational, rotational and vibrational entropy components. Generally, the entropy of a molecule is calculated from the partial derivative of the free energy (G) to the temperature (T) at constant pressure (p), Eq 8. In this equation R is the gas constant and N is the total number of particles.

$$S = -\left(\frac{\partial G}{\partial T}\right)_p = -\left(\frac{\partial}{\partial T}\left[-RT\ln\frac{Q}{N}\right]\right) \qquad (8)$$

The partition function is the product of all molecular variables and includes the translational (Q_{tr}), rotational (Q_{rot}) and vibrational (Q_{vib}) contributions of the molecule as well as those arising from internal rotations (Q_{int}) and electronic spin (Q_{es}), Eq 9. For a molecule obeying the harmonic oscillator, rigid-rotor and Ideal Gas Law conditions equations have been derived for these partition function components, Eqs 10 – 12.[32] In these equations k_B is Boltzmann's constant, h is Planck's constant, m is the molecular mass, T the temperature and σ the symmetry number. The moments of inertia A, B and C in Eq 11 are dependent on the Cartesian coordinates of the molecule with respect to its centre of mass.

$$Q = Q_{tr} \times Q_{rot} \times Q_{vib} \times Q_{\text{int}} \times Q_{es} \qquad (9)$$

$$Q_{tr} = \left(\frac{2\pi m k_B T}{h^2}\right)^{\frac{3}{2}} V \qquad (10)$$

$$Q_{rot} = \frac{8\pi^2}{\sigma}\left(\frac{2\pi k_B T}{h^2}\right)^{\frac{3}{2}}\sqrt{ABC} \qquad (11)$$

$$Q_{vib} = \prod_{k=1}^{n}\left(1 - e^{-\frac{h\nu}{k_B T}}\right)^{-1} \qquad (12)$$

Quantum chemical software packages usually include these entropy derivations in order to approximate the free energy of a molecule. The obtained free energies of activation can then be compared with experimentally determined values or rate constants. We did that on several occasions and found that free energies of activation agree to within 3 kcal mol^{-1} from experiment using B3LYP with a triple-ζ basis set and solvent corrections.[33]

To further test the method dependence on the calculated free energy of activation we did a set of test calculations as shown in Fig 4. Thus, Pestovsky and Bakac[34] reported experimental studies on the reactivity of $[Fe^{IV}(O)(H_2O)_5]^{2+}$ with dimethylsulfoxide and measured a free energy of activation of 10.5 kcal mol^{-1} at 298 K. To find out how well computations can reproduce this, we calculated the catalytic mechanism with the B3LYP, BP86, PBE1PBE, M06 and B3LYP-D DFT methods as well as with MP2. Fig 4 displays the optimized geometries of the sulfoxidation transition states and the corresponding free energies of activation. In this particular example the DFT methods are within a couple of kcal mol^{-1} from the experimental value and B3LYP and BP86 are closest. Interestingly, the computationally expensive MP2 method performs very poorly and gives a much higher barrier of 39.0 kcal mol^{-1}. Clearly, computational time in this example does not correlate with higher accuracy.

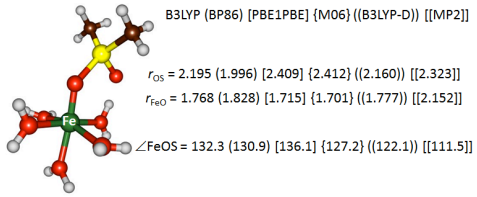


Fig. 4 Optimized geometry of the oxygen atom transfer transition state from $[Fe^{IV}(O)(H_2O)_5]^{2+}$ to dimethylsulfoxide. Bond lengths are given in angstroms, the angle in degrees and the imaginary frequency in the TS in wavenumbers. Also given are free energies of activation in kcal mol^{-1}.

Environmental (solvent) effects

Calculated energies using computational chemistry methods refer to gas-phase data. In order to compare computational results to experiment, the solvent needs to be mimicked in into the calculations. This can be done by simply adding explicit solvent molecules to the chemical model and reoptimizing the structures. However, calculations with a large number of atoms require more resources and are not always practical, especially since solvent molecules can exist in many possible conformations.[35] Therefore, a more common approach is to add solvent corrections to the energetics using an implicit solvent model.[36] This can either be done at single point level, whereby the energy is recalculated with a solvent model on the gas-phase optimized geometry, or by reoptimizing the structure using a solvent model.

Most QM software packages include implicit solvent models like the polarizable continuum model (PCM), the conductor-like screening method (COSMO) or the density solvation model (SMD).[37] In a PCM calculation, the molecule is placed in a cavity that is surrounded by the dielectric continuum, characterized by a dielectric constant, and a reaction fluid builds up at the cavity surface, described by a number of surface charges, perturbing the wave function. The cavity is normally defined by a set of atomic radii. The electrostatic part of the solvation energy is supplemented by non-polar terms, which typically include dispersion, repulsion and cavitation terms.

A common approach to investigate enzyme reactions is to cut out a small model of active site from the protein (50 – 200 atoms) and study it with QM methods, called the QM-cluster approach.[38] In order to ensure that the geometry of the model does not change too much from that in the protein, one or several atoms are fixed at the positions where the protein is truncated to form the QM model. Moreover, a continuum-solvation method is commonly employed to model the surrounding protein in an approximate manner. A dielectric constant of ~4 is typically used,[39] a quite arbitrary choice, but fortunately it has been shown that as the size of the QM system is increased to 150 – 200 atoms, the results obtained with different values of the dielectric constant typically converge.[38a,40]

Quantum mechanics/molecular mechanics (QM/MM).

Another approach to include the effect of the surrounding into QM calculations is to use the quantum mechanics/molecular mechanics (QM/MM) approach.[41] The philosophy of the QM/MM technique is to divide the chemical system, such as an enzyme, into two or more regions (shells), Fig 5: the inner core that is calculated with accurate QM or DFT methods and the outer region, which contains the remaining protein and solvent is calculated with computationally cheap methods, like molecular mechanics (MM).[42] The inner core is the region of interest and contains the basic features of the enzyme active site and the substrate, but is held in position by interactions with the rest of the protein. These chemical constraints ensure that the active site remains in a conformation that is relevant for the enzyme. Ideally, this would combine the accuracy of QM methods with the speed of MM methods.

In practice, the QM/MM approach is implemented by an energy function of the type:

$$E_{\mathrm{tot}}(r_1,r_2) = E_{\mathrm{QM}}(r_1) + E_{\mathrm{MM}}(r_2) + E_{\mathrm{QM/MM}}(r_1,r_2) \quad (13)$$

where r_1 and r_2 are the coordinates of the QM and MM systems, respectively, and E_{QM} and E_{MM} are the QM and MM energy functions. $E_{\mathrm{QM/MM}}$ is the energy of the QM/MM interface and can be implemented in many different ways, carefully avoiding double-counting of interaction terms. A simple way to implement a QM/MM method without any need to modify the QM or MM software packages is via the following equation:

$$E_{\mathrm{tot}}(r_1,r_2) = E_{\mathrm{QM}}(r_1) + E_{\mathrm{MM}}(r_1,r_2) - E_{\mathrm{MM}}(r_1) \quad (14)$$

In Eq 14 one essentially does three calculations, namely a QM calculation on the QM system, and MM calculations are performed both for the total system and for the isolated QM system.

A complication arises when the QM and MM systems are connected by covalent bonds (so called junctions).[41] Then, the QM system needs to be capped in some way to fulfil the valences of the dangling bond. This can be done either by adding extra atoms (the link-atom approach), usually hydrogen atoms, although halogen-like pseudoatoms have also been tried, or by using localised orbitals that place frozen hybrid orbitals at the boundary. With a careful calibration, both approaches can give results of a similar quality[43] and therefore, the hydrogen link-atom approach is most common, because it is simple and can be used without any modifications of the QM software. However, it should be kept in mind that the junction problem is one of the largest sources of errors for the QM/MM approach, the influence of which can be reduced by moving the junctions away from the site of interest (making the QM system larger).

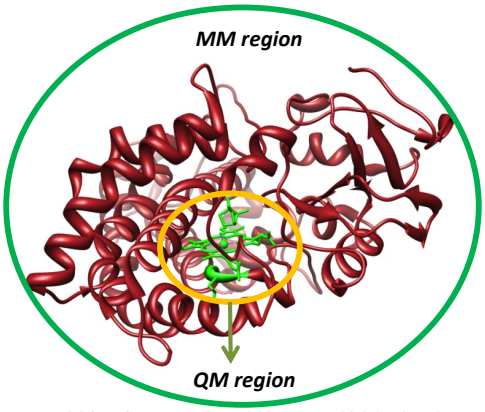


Fig. 5 Division of an enzymatic model into a QM (high-level) and MM (low-level) region.

Another important issue is how to treat the interactions between the QM and MM systems.[41] Van der Waals interactions are normally properly treated by MM, e.g. by the simple energy function in Eq 14. However, electrostatic interactions are more complicated to describe. The simplest solution is to treat these at the MM level (as in Eq 14), giving what is called mechanical embedding. It requires a MM representation of the QM system, typically a point-charge model obtained by a fit to the QM electrostatic potential. However, this ignores the polarisation of the QM system by the MM system and *vice versa*. Therefore, it is more common that a point-charge model of the MM system is included in the QM calculations and the charges of the QM system are zeroed in into the MM calculations, which is called electrostatic (or electronic) embedding. Then, the QM system is polarised by the MM system, whereas the MM system is still not polarised. It has been argued that this may lead to an overpolarisation of the QM system[44] and it has been much discussed which charges (especially around junctions) to include in the point-charge model.[41] It should also be noted that mechanical embedding reduces the link-atom problem, because the full MM calculation (and therefore the electrostatic interactions) is performed without any link atoms. With

electrostatic embedding, the link-atom problem is severe, because charges are positioned close to the artificial link-atoms. Corrections to this have been suggested,[45] but it has been shown that they often do more harm than good.[44] In fact, a carefully designed molecular-embedding method typically gives better results than electronic embedding. However, the ideal procedure would be to use polarised embedding, i.e. employing a polarisable MM method and including both point charges and polarisabilities (and possibly also higher-order multipoles) in the QM calculations. Only a few applications have been performed along those lines so far.[41,46]

QM/MM codes are available in most common QM and MM software packages (e.g. Gaussian, ADF, Amber, CHARMM, Schrödinger), as well as in some independent interfaces (e.g. ChemShell and COMQUM).[47] A promising approach is to combine QM/MM methods with experimental data (X-ray reflections, NMR, EXAFS) to obtain an ideal compromise between theory and experiments.[48]

The advantage of QM/MM methods compared to QM-cluster calculations is that the surroundings are explicitly accounted for, avoiding a bias caused by the selection of the modelled system or any need of fixing atoms at the periphery of the QM model. The disadvantage of QM/MM is that the modelled system is so large that it is hard to ensure that all structures along a reaction pathway reside in the same local minimum (a change in the hydrogen-bond pattern of a single water molecule far from the active site can change the total energy by ~7 kcal mol^{-1}, although it is probably not relevant for the reaction). Many solutions have been suggested to such problems, e.g. by keeping the surroundings fixed at the starting (crystal) structure, by running reactions back and forth until the energy stabilises, or by calculating reactions for many different structures, taken from snapshots from molecular dynamics simulations.[41] However, the preferred solution is to calculate free energies, rather than pure energies (from minimised structures).[41,45a,49] This can be done at the MM and semi-empirical QM/MM level with free-energy perturbation methods. Unfortunately, it is a challenge for high-level (including DFT-type) QM/MM methods, owing to the cost of these calculations and differences in the MM and QM potential energy surfaces.

Finally, it should be mentioned that both QM and QM/MM methods have problems with convergence of the calculations with respect to the size of the QM system. Convergence studies of QM-cluster calculations have shown that reaction energies can change by 3 – 14 kcal mol^{-1} as the QM system is enlarged from 27 – 83 to 161 – 220 atoms.[40] Other studies have shown that a difference of 14 kcal mol^{-1} can remain even for systems of ~400 atoms and that the energy may differ by 11 kcal mol^{-1} between QM systems of 300 and 1,600 atoms.[40c] The convergence is not improved by geometry optimisations.[50] However, QM/MM methods also have similar problems, due to the limited accuracy of the MM methods and the link-atom problem: Electrostatic-embedded QM/MM calculations gave a mean absolute deviation similar to QM-cluster calculations over 40 sizes of the QM system (40 – 446 atoms), of 4 and 5 kcal mol^{-1}, respectively.[44] Only a carefully designed mechanical-embedding QM/MM method could bring this difference down to 1.7 kcal mol^{-1}. Therefore, we currently recommend the use of QM/MM geometry optimisations with a quite small QM system, followed by a single-point QM calculation with a very big QM system, involving all groups within ~4.5 Å of the minimal QM system and all buried charged groups in the protein, and moving all junctions three residues away from the minimal QM system, typically 600 – 1,000 atoms.[50,51]

The QM/MM methodology starts off from experimentally determined crystal structure coordinates available from the internet.[52] Often these pdb files refer to a resting state structure of the enzyme rather than its catalytically active species. Therefore, the structure often needs to be converted to a later structure in the catalytic cycle. In addition, the pdb files typically do not contain hydrogen atoms, so that they need to be added, as well as a number of explicit solvent molecules. Generally, this is done in the set-up where the user carefully checks protonation states of residues, salt and disulphide bridges and ions. Histidine groups in proteins can exist in three possible protonation states: deprotonated, singly protonated (with two different tautomers) or doubly protonated. Therefore, all His residues need to be manually and visually investigated for likely hydrogen bonding donor and acceptor groups, as well as their solvent accessibility. Counterions can be added to the surface to balance the overall charge of the systeml. After this initial set up the system is equilibrated, and heated to room temperature with molecular dynamics (MD) simulations in a stepwise protocol.[41,53]

Results

In this section we will give examples of several computational studies that either supported experiment or had predictive value and guided experimental work in the field. In addition, we discuss bioinorganic examples where computation was calibrated against experimental rate constants, crystal structures and spectroscopic parameters. Thus, from the ground state electronic wave function it is, in principle, possible to calculate many molecular properties beside geometries, by calculating derivatives of the energy (such as e.g. the electric dipole moment as a first derivative of the external field). For properties involving electronic transitions, however, at least a reasonable approximation to the excited state electronic wave functions is needed, which often requires the application of more involved computational methods, as detailed below. These examples highlight the challenges that computational work needs to overcome. We will start, however, with a section on enzymatic reaction mechanism, and, in particular, focus on topics that are hard to study experimentally. In the second part of this section, we will discuss several biomimetic studies that were aimed at establishing a suitable method and technology to address the chemical systems.

Part 1: Enzymatic models

Catalytic mechanism of cysteine dioxygenase

Cysteine dioxygenase (CDO) is a vital enzyme for human health that catalyses the metabolism of cysteine, and, thereby, detoxifies the body.[54] In contrast to most nonheme iron dioxygenases that contain a facial 2-His/1-Asp/Glu active site feature, in CDO there is a 3-His metal coordination. As most intermediates in the catalytic cycle are short-lived, little is known about its catalytic mechanism after dioxygen binding.

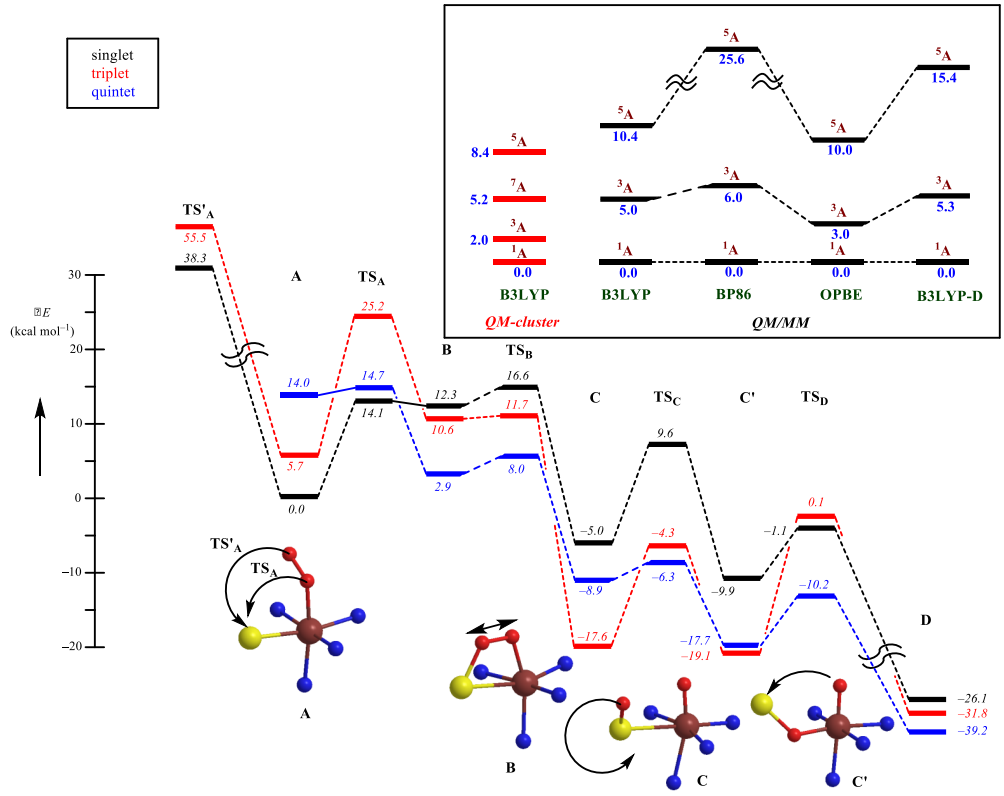


Fig. 6 Potential energy profile obtained at the QM/MM level of theory for competing mechanisms for oxygen activation by CDO enzymes on the singlet, triplet and quintet spin state surfaces. Energies were obtained with triple-ζ basis set and include ZPE corrections. The inset shows the spin state ordering of the iron(III)-superoxo complex (**A**) as calculated with several DFT methods using QM-cluster models as well as QM/MM.

Computational studies were performed using both QM-cluster model and QM/MM approaches,[53,55] and because CDO has a tight substrate-binding pocket with a large number of stabilizing hydrogen bonds and salt bridges linked to the substrate that keep the substrate in a rigid orientation, a large QM region was chosen to include this hydrogen-bonding network. The QM-cluster model and the QM region in the QM/MM calculations included iron(III)-superoxo with substrate cysteinate and the amino acid side chains of His_{86}, His_{88}, His_{140}, Arg_{60}, His_{155}, Tyr_{157} and Cys_{93}. The computational studies initially investigated the iron(III)-superoxo complex that was calculated in the lowest lying singlet, triplet, quintet and septet spin states using QM-cluster models and QM/MM. The work tested a range of DFT methods of which we show the B3LYP,[28,29] BP86,[27,56] OPBE[57] and B3LYP-D[31] results in Fig 6. All methods give an open-shell singlet spin state as the ground state with orbital occupation $\pi^*_{xy}{}^2\ \pi^*_{xz}{}^2\ \pi^*_{yz}{}^1\ \pi^*_{OO}{}^1$ that represents a superoxo radical antiferromagnetically coupled to a metal-oxo π*-type unpaired electron. With most QM/MM methods the nearest excited spin state is the triplet state, which is 3.0 – 6.0 kcal mol^{-1} higher in energy and the quintet spin state is more than 10 kcal mol^{-1} above the ground state. These studies are well reproduced with QM-cluster models[55] that give the same spin-state ordering and minor differences in relative energies. In addition, the QM-cluster calculations also located a septet spin state at 5.2 kcal mol^{-1} above the singlet spin ground state. Clearly there are a number of close-lying spin-state structures within 5 kcal mol^{-1} from the electronic ground state, which may result in a Boltzmann distribution and a mixture of low-lying geometries and spin-states at room temperature. Indeed for a biomimetic nonheme iron(III)-superoxo complex 12 different close-lying spin-state structures were calculated,[58] which highlights the complications in the description of iron(III)-superoxo complexes. Nevertheless, all computational studies on CDO agree that the iron(III)-superoxo has an open-shell singlet configuration. Addition of $O_2^{\bullet-}$ to the substrate-bound resting state of CDO enzyme led to a stable structure that was characterized spectroscopically and identified as a most likely septet spin state of a ferromagnetically coupled iron(III), $S = {}^5/_2$, coupled to an $S =$

½ radical.[59] Currently, it is not clear whether this intermediate is a catalytic cycle intermediate and why it has a significant lifetime for spectroscopic detection. It may very well be that the other spin states are more reactive and do not let the iron(III)-superoxo accumulate over time.

Subsequently, the catalytic mechanism leading to cysteine sulfinic acid products was calculated with QM-cluster and QM/MM methods. The lowest energy mechanism starts off from an open-shell singlet iron(III)-superoxo (**A**) with the formation of an S–O bond via a transition state $\mathbf{TS}_A$ to give the ring-structure **B**. Past the transition state, the system will cross-over to the triplet spin state ^{3}**B**, which is slightly lower in energy than ^{1}**B**. The ring-structure has a weakened O–O bond and splits into an iron(IV)-oxo state with a bound sulfoxide. Similar to biomimetic nonheme iron(IV)-oxo complexes with hexacoordination the triplet spin state is lower in energy than the quintet spin state.[60] The final stages of the mechanism include a sulfoxide isomerization to form **C'** followed by oxygen-atom transfer to give cysteine sulfinic acid products **D**.

An alternative mechanism was recently suggested, based on a crystal structure that appeared to give a stable persulfenate intermediate.[61] Therefore, it was also tested whether the distal oxygen atom of the iron(III)-superoxo could attack the sulphur atom of cysteinate to form a persulfenate structure via transition state $\mathbf{TS'}_A$. However, this pathway was found to be substantially higher in energy with barriers >30 kcal mol^{-1} as well as high endothermicity. Consequently, the persulfenate structure is a high-energy structure that cannot take part in the catalytic mechanism. The fact that it was observed experimentally implicates that it arises from a dead-end side reaction.

Thus, theory has predicted a most likely mechanism for CDO enzymes and predicted a multistate reactivity pattern on several close-lying spin state surfaces and a rate-determining O–S bond formation step.

Bioengineering of *S*-mandelate synthase into *R*-mandelate synthase

Further combined experimental and computational studies on nonheme iron dioxygenases were focused on *S*-*p*-hydroxymandelate synthase.[62] This is an α-ketoacid-utilizing enzyme that converts *p*-hydroxyphenylpyruvate on an iron(III)-superoxo centre into *p*-hydroxyphenylacetate and an iron(IV)-oxo species. The latter abstracts a hydrogen atom from *p*-hydroxyphenylacetate and the subsequent rebound gives *S*-mandelate products enantioselectively. QM-cluster calculations on its catalytic mechanism leading to *R*- and *S*-mandelate products were performed and the reaction was shown to proceed with a common initial hydrogen-atom abstraction leading to a radical intermediate in either the *pro*-*R* or *pro*-*S* configuration. We then calculated the transition states for radical rebound to form *R*-mandelate and *S*-mandelate products: $\mathbf{TS}_{reb,R}$ and $\mathbf{TS}_{reb,S}$. These structurally different transition states were then re-inserted into the crystal structure of the enzyme and the active site was investigated. An analysis of these structures led to suggestions of putative active-site mutations that could reverse the enantioselectivity of the enzymatic reaction. When these predictions were subsequently tested experimentally they showed a full enantioselectivity reversal from dominant *S*-mandelate products to *R*-mandelate products in the mutant. Thus, theory guided experimental work and enabled the biosynthesis of the first artificial nonheme iron enzyme whereby an enantioselectivity reversal was obtained.

Comparison of families of mononuclear Mo enzymes

Molybdenum is the most common transition metal in sea water.[63] Therefore, it is not surprising that it is used by biosystems in a range of different enzymes, such as nitrogenases and oxygen-transfer enzymes. In the nitrogenases, Mo is coordinated to an unusual dithiolene ligand, called molybdopterin (MPT). However, the nature of the other Mo-ligands varies extensively, and three families have been identified. In the dimethylsulfoxide reductase (DMSOR) family, the metal coordinates in its oxidised Mo(VI) state to two MPT ligands, one oxo group, and a protein-derived ligand (Ser, Cys, or SeCys; Fig 7). In the sulfite oxidase (SO) family, Mo(VI) coordinates to one MPT, two oxo groups, and a Cys ligand. In the xanthine oxidase (XO) family finally, Mo(VI) coordinates instead to one MPT, one oxo group, one hydroxide, and one sulfido group. It is interesting to understand why these enzymes have such different active sites. Such a question is quite hard to address with experimental methods, but with QM-cluster calculations, much insight can be gained.[64]

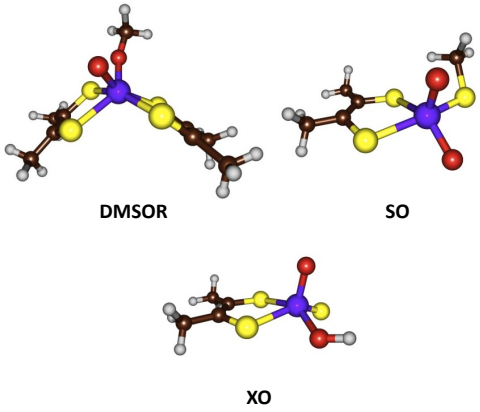


Fig. 7 Active site structures of three families of Mo-containing oxo-transfer enzymes: DMSOR, SO, and XO.

From the description above, we saw that the active sites of the three families can be described by $Mo^{VI}O(MPT)XY$, where *X* and *Y* are varying ligands, being MPT and Ser/Cys/SeCys for the DMSOR family, O^{2-} and Cys for the SO family, and S^{2-} and OH^-. Thus, the *X* ligand has a double negative charge and the *Y* ligand a single negative charge. With QM-cluster calculations, the intrinsic properties and reactivities of the three enzyme families can be studied, as well as intermediate sites with other combinations of *X* and *Y* ligands. For example, the driving force of a general oxo-transfer reaction can be studied:

$$Mo^{VI}O(MPT)XY + Z \rightarrow Mo^{IV}(MPT)XY + ZO \quad (15)$$

where Z is a typical substrate. $Z = SO_3^{2-}$ gives an exothermic reaction for all ligands, with the SO ligands in the middle. On the

other hand, Z = $S(CH_3)_2$ gives endothermic energies for all ligands, illustrating that the native reaction goes in the opposite direction (from $(CH_3)_2SO$). Moreover, the DMSOR ligands give the lowest reaction energy, probably to conserve energy. In fact, the *X* = MPT ligand reduces the reaction energy by ~25 kcal mol^{-1}. For Z = neutral xanthine, *X* = O^{2-} or S^{2-} (including the native ligands of XO) give endothermic reactions.[65] This agrees with the consensus view that the actual substrate of the enzyme is deprotonated anionic xanthine, for which all ligands give exothermic reactions. Again, the native XO ligands give energies closes to zero to conserve the energy.

Second, the reoxidation of the active site was studied, i.e.

$$Mo^{IV}(MPT)XY + H_2O \rightarrow Mo^{VI}O(MPT)XY + 2\,H^+ + 2\,e^- \quad (16)$$

In principle, it can be divided into two one-electron transfer and two proton-transfer steps, thus giving two redox potentials and two acidity constants (which depend on the order of the reactions), but the results of the net reaction are easiest to interpret. The ligand sphere of DMSOR (and all other combinations with *X* = MPT) give a positive redox potential for the net reaction in Eq 16, indicating that the reduced state is most stable, which is the correct direction for the re-reduction of the active site after the reduction of DMSOR. For the other sets of ligands, the redox potential is negative, and indicates that the oxidised state is instead most stable, again in accordance with the direction of the reactions (both SO and XO oxidise their substrates). Thus, the ligands seem to have been selected to make the re-reduction or re-oxidation of the active sites thermodynamically favourable.

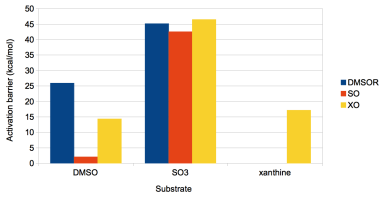


Fig. 8 Activation barriers (in kcal mol^{-1}) for substrate activation by the active sites of the three families in Fig 7. The substrates used were DMSO, SO_3^{2-} and xanthine, respectively.

Finally, the actual reaction mechanisms (and therefore the activation barriers) for the typical reactions of each of the three families were studied. It turns out that the active sites of all three families can catalyse the DMSOR reaction with reasonable barriers. In fact, the native DMSOR ligands give the highest barrier, owing to the less favourable reaction energy discussed above (Fig 8). Likewise, all three active sites can catalyse the SO reaction with similar barriers, although the native ligands give the lowest activation barrier (by 3 – 5 kcal mol^{-1}). However, for the more complicated XO reaction, only the native model was active. In particular, the S^{2-} ligand seems to be essential as a hydride acceptor.

Thus, we can conclude that the ligands seem to have been selected to give exothermic reactions (but typically by a minimal amount of energy), both for the main reaction and the regeneration of the active site, as well as reasonable activation barriers. This illustrates how QM-cluster calculations can give insight in the evolutionary design of active sites of different enzyme families.

Reaction mechanism of DMSO reductase

DFT methods often experience problems in estimating the energy difference of chemical reactions where the transition metal changes oxidation state.[66] This is nicely illustrated by a recent study of the reaction mechanism of the enzyme DMSOR.[67] As mentioned above, the active site of the enzyme contains a Mo ion bound to two molybdopterin cofactors and a serine residue from the protein. The $(CH_3)_2SO$ substrate binds to this site with Mo in the +IV oxidation state (Fig 9). Then, dimethylsulfide dissociates, leaving a Mo(VI) ion with an oxy group. The latter dissociation is the rate-limiting step.

Different dispersion-corrected DFT methods give varying results for the activation barrier of this reaction (and also for the reaction energy), ranging from 5.5 to 33 kcal mol^{-1} (Fig 10).[67] The barrier height follows quite closely the amount of exact exchange in the functional, e.g. 5.5 – 7.9, 11, 18, 21, and 33 kcal mol^{-1} for pure functionals, TPSSH-D3, B3LYP-D3, PBE0-D3, and BHLYP-D3 with 0, 10, 20, 25, and 50% exact exchange, respectively. In order to determine which of these results is most accurate, local CCSD(T) calculations were performed with extrapolations to complete basis sets and to canonical methods. These gave a barrier of 16 kcal mol^{-1}, i.e. closest to B3LYP-D3. However, other states in the reaction were more accurately predicted by the pure functionals. Thus, *there is not a single DFT method that gives good results for all states*. Instead, it is better to use the LCCSD(T) energies and add corrections for solvation, entropy, and thermal effects from the DFT calculations. In fact, the time seems to be ripe to start to use LCCSD(T) (which currently can be used for complexes with up to 70 – 90 atoms) calculations also for metal-containing systems, provided that the multiconfigurational character is not too large.

Ulf Ryde 10/21/13 7:16 AM
Deleted: a

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

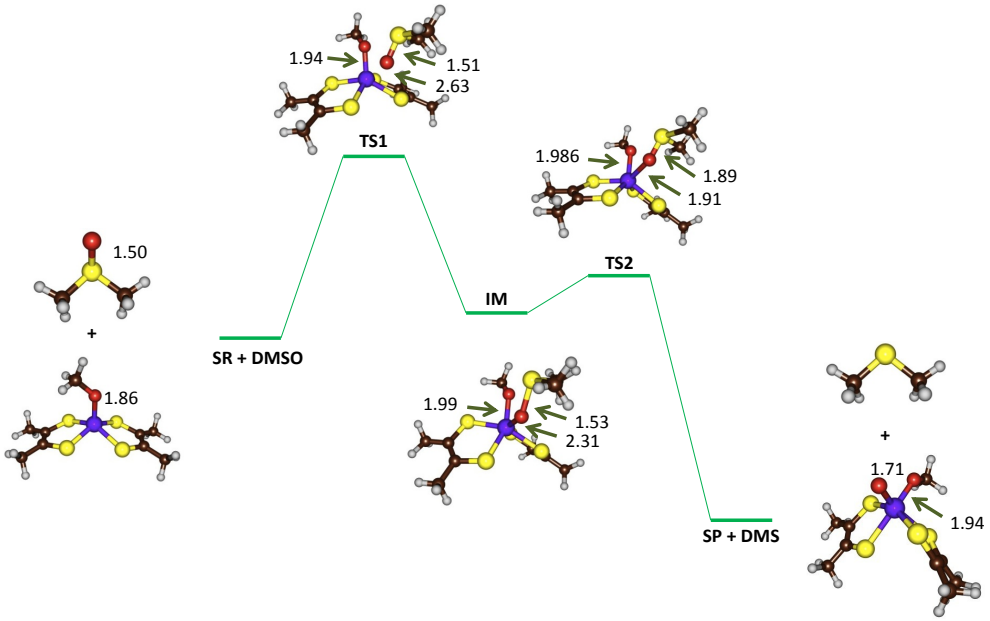


Fig. 9 Optimized geometries (B3LYP-D3/def2-TZVPP+ZORA+COSMO(ε=4) level) of critical points along the reaction mechanism of DMSOR with bond lengths in angstroms.

.

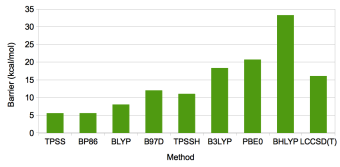


Fig. 10 Rate-limiting activation barriers (in kcal mol^{-1}) for the DMSOR reaction as calculated with various DFT methods (all including the DFT-D3 dispersion correction) and LCCSD(T).[67]

With such an approach, a barrier of 15 kcal mol^{-1} was obtained[67] in perfect agreement with the experimental estimate[68] (which of course is fortuitous). There are several previous theoretical studies of the same enzyme, which all have reported similar barriers, 8.8 – 19 kcal mol^{-1}.[69] This is unexpected, because none of them included dispersion effects, which are pronounced for this barrier, 17 kcal mol^{-1}, and they employ both pure and hybrid DFT functionals. However, looking closer at the earlier studies, it can be seen that they either used a too small basis set (that gives a too low barrier by ~17 kcal mol^{-1}) or the pure BP86 functional (that gives a ~10 kcal mol^{-1} too low barrier), in both cases compensating for the missing dispersion. Thus, the previous studies obtained the correct answer for the wrong reason.

The take-home message is that the seemingly simple Mo oxy-transfer reactions are quite complicated to treat with QM methods. Large basis sets must be used (at least triple-ζ quality), dispersion and solvation effects need to be included (also the non-polar solvation terms), and the DFT methods should ideally be calibrated against more accurate coupled-cluster or multiconfigurational methods.

Key intermediates of multicopper oxidases

An example of a case for which QM calculations have solved a problem that later has been confirmed by experiments comes

from the multicopper oxidases (MCO). This group of enzymes contains at least two copper sites: A mono-nuclear type 1 (T1) centre, with a Cu ion coordinated to at least one cysteine and two histidine residues, and a tri-nuclear centre, consisting of a type 2 copper ion coupled to a binuclear type 3 copper centre in a triangular fashion (the T23 site).[70] All three Cu ions in the T23 site coordinate to two or three histidine residues and, in the oxidised resting state two solvent-derived ligand, one terminal to the type 2 ion and one bridging between the type 3 ions.

The MCOs catalyse oxidation of various substrates at the T1 site, coupled to the reduction of O_2 to H_2O at the T23 site. By spectroscopic methods, two intermediates in the catalytic cycle have been characterized, the peroxy and native intermediates.[70] The former arises when O_2 binds to the fully reduced (Cu^{I}_3), leading to an immediate reduction of O_2 to the peroxide level. However, it was originally not known how O_2 bound to the T23 cluster – the experiments could not decide whether it binds inside the cluster, coordinating to all three Cu ions, or on its periphery, coordinating only to two of the ions.[71] In the native intermediate, all three copper ions are fully oxidised (Cu^{II}_3), the O–O bond is cleaved, and both O atoms are fully reduced. However, again it could not be decided whether it contained a central O^{2-} ion coordinated to all three Cu ions or three OH^- ions, bridging each pairs of Cu ions.[72]

At this point, Rulíšek et al[73] performed a QM/MM study of the MCO reaction. They could obtain structures of both the suggested structures of the peroxy and the native intermediates, as can be seen in Fig 11. The structures were similar to what was suggested by experiments, although the peroxide ion bound in a diagonal manner inside the T23 site, and no other structural interpretations of these states could be found even if many other starting configurations were tried. Already from Fig 11, it can be seen that the peripheral binding of both intermediates seems to be more strained than the central ones. Moreover, approximate energy considerations (the complexes do not contain the same atoms so the energies are not directly comparable, but by removing or adding atoms, an approximate comparison can be done) indicate that the central binding is favoured by 11 – 34 and 28 – 34 kcal mol^{-1}. Thus, the QM/MM calculations indicated that the central structures on the left-hand side of Fig 11 were the correct interpretation of these intermediates. The only problem was that the preferred peroxy-intermediate structure had an incorrect ground state (ferromagnetically, rather than antiferromagnetically coupled).

This problem was later solved by advanced calculations using multiconfigurational perturbation theory (CASPT2), which showed that all intermediates are antiferromagnetically coupled.[74] These calculations also gave further support to the central structures by giving better excitation energies for low-lying excited states. Additional support for these structures were also obtained by calculating EPR tensors.[75]

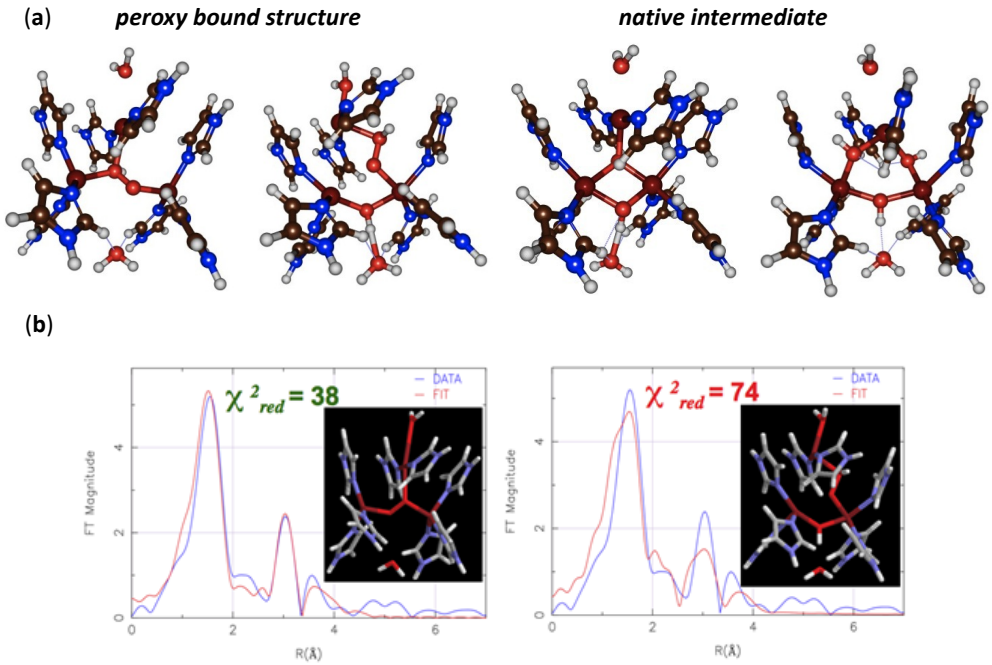


Fig. 11 (*a*) QM/MM structures for the two suggested structural interpretations of the peroxy and native intermediates with central (left) or peripheral coordination (right). (*b*) QM/MM-EXAFS structures as well as experimental (blue) and calculated (red) EXAFS spectra for the peroxy adduct.

Chemical Communications

Dynamic Article Links ►

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

EXAFS data for the Cu–Cu distances in the two intermediates were also available,[67,68] but both structures agreed equally well. However, the EXAFS data contain much more information than only Cu–Cu distances. A combined QM/MM and EXAFS refinement, gave structures that are an ideal compromise between QM and EXAFS data and showed that the central structure fits the EXAFS data appreciably better than that with peripheral binding of HO_2^- (Fig 11*b*).[76] Subsequently, spectroscopic and inorganic modelling studies have confirmed that the central binding is the correct mode of both the peroxy and native intermediates of the MCOs.[70,77] Altogether, these studies show the predictive power of the QM calculations and the strength of combining various spectroscopic and theoretical investigations.

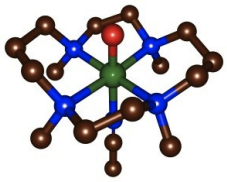
Part 2: Biomimetic systems

Predictive studies of spectroscopic properties

Calculating spectroscopic properties of molecules is a challenging task and often requires computationally expensive methods, such as CCSD(T), in combination with a large basis set. As these methodologies are not always accessible for large chemical systems of well over 50 atoms, such as biomimetic model complexes, users often resort to DFT modelling. A typical strategy applies conventional single-reference (hybrid) DFT with a polarized triple-ξ quality basis set for the calculation of EPR *g* and *A* values, as well as d-d transitions and spin distributions. It should be noted, however, that at this level of theory one would not expect to obtain highly accurate results and even qualitatively they can be incorrect, e.g., when there is the unphysical preference for higher spin states or unrealistic energies for spin transitions.[78] Unfortunately, rigorous electronic property calculations with correlated wave function treatments such as CCSD(T)[21] in the single-reference domain or CASSCF[79] (and subsequent perturbations, such as CASPT2[80] or NEVPT2[81]) in the multi-reference domain, require large (better than triple-ξ quality) basis sets and are, despite recent developments,[78,82] often prohibitively expensive in terms of computer time. For this reason, a good strategy is to try to correlate the DFT-derived properties to experimental data where at all possible. This approach allows explaining and to some extent predicting, spectroscopic properties, but only for a particular problem, where the applied methods have been previously validated.

Prediction of the d-d electronic transitions of a nonheme iron(IV)-oxo complex

The prediction of the electronic d-d transitions of [*trans*-Fe(O)(NCCH$_3$)TMC]$^{2+}$, TMC = 1,4,8,11-tetramethyl-1,4,8,11-tetraazacyclotetradecane, was studied with high-level *ab initio* (SORCI[83]) methods, Time-Dependent DFT (TD-DFT[84]) and a Ligand Field method based on DFT, LF-DFT.[85] The use of a LF-DFT method is possible in this particular case because of the localized character of the 3d electrons, which gives a mainly ionic interaction between the metal and its ligands: the ligands donate electrons into the partly filled 3d and virtual 4s and 4p valence-shell of orbitals on the metal. The bonding molecular orbitals (MOs) are, therefore, ligand-centred, whereas the antibonding orbitals are metal-centred. The resulting d^n configuration on the central metal then leads to the electronic d-d transitions for emission and absorption.[85a] The concept of the LF-DFT method[86] is based on the validity of constructing a sub-Hamiltonian for the d-d block on the metal taking into account that the d-d transitions are energetically well separated from the charge transfer transitions.[87]



r_{FeO} = 1.65 (1.64)
$r_{FeN,average}$ = 2.14 (2.13) ± 0.02
$r_{FeN,axial}$ = 2.11 (2.12)

S = 1

Fig. 12 Geometry-optimized structure of [*trans*-Fe(O)(NCCH$_3$)TMC]$^{2+}$ from a BP86 (B3LYP) optimization,[85a,88] for the electronic ground state S = 1.

In the case of [*trans*-Fe(O)(NCCH$_3$)TMC]$^{2+}$, the structure was first carefully optimized with BP86[85a] and then studied with LF-DFT, ΔSCF and time-dependent DFT (TD-DFT). The ΔSCF method is similar to a normal DFT calculation, but one or more valence electrons are placed into higher-lying Kohn-Sham orbitals. As a comparison, the optimized geometries of [*trans*-Fe(O)(NCCH$_3$)TMC]$^{2+}$ at the BP86 and B3LYP levels of theory are shown in Fig 12 and the MO diagram is given in Fig 13.[85a,88]

As can be seen from Table 1, both LF-DFT and TD-DFT produce acceptable results, whereas SORCI gives an incorrect ground state. For the ground state to $^3A_1$ transition, the TD-DFT value is closer to experiment than the LF-DFT value, but on the other hand the TD-DFT calculation fails to predict the ground state to $^3A_2$ transition.[89] As the size of the complex is already challenging for a SORCI treatment, not all transitions could be obtained; on the other hand, some transitions, such as the $^1A_1$ and ^1A spin-flip transitions, cannot be explained by the LF-DFT method.

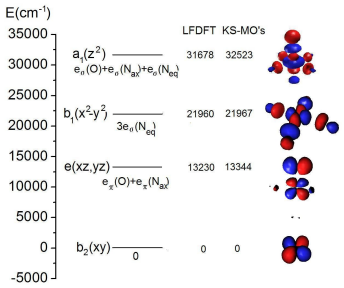


Fig. 13 MO energy diagram for [*trans*-Fe(O)(NCCH$_3$)TMC]$^{2+}$ calculated with LF-DFT. The energies of the LF-DFT orbitals are compared with those resulting from the average-of-configuration KS calculation. Orbital contours (pertaining to values of the density of 0.05 au) are plotted, but the individual atoms are omitted for clarity.

Table 1 Theoretical and experimental vertical excitation energies (cm^{-1}) within D_{4h} symmetry of the $S = 1$ structure of [*trans*-Fe(O)(NCCH$_3$)TMC]$^{2+}$.

Electronic state	LF-DFT	SORCI	ΔSCF	TD-DFT (SAOP)	assignment	Exp (polarization)
$^3A_2$	0	3226	0	0	$xy^2\ xz^1\ yz^1$ [a]	-
$^5A_1$	911	0	6488	8146	$xy \rightarrow x^2-y^2$	-
$^1B_2$	10036	-		6170	spin-flip [b]	-
$^5B_1$	10629	-		13759	$xy \rightarrow z^2$	-
$^1B_1$	10979	-		6314	spin-flip [b]	-
^3E	11968	-	10486	9472	$xy \rightarrow xz,yz$	~10500(xy)
^3E	15680	-	12668	13726	$xz,yz \rightarrow z^2,x^2-y^2$ [d]	13000(xy)
$^1A_1$	16924	10484		11589	spin-flip [c]	-
$^3A_1$	19347	-	12570	18372	$xy \rightarrow x^2-y^2$	17000(z)
^5E	19359	-		-	$xy(xz,yz) \rightarrow z^2,x^2-y^2$ [e]	-
^1E	20360	9436		13070	$xy \rightarrow xz,yz,x^2-y^2$ [d]	-
$^3B_2$	22012	-		[f]	$xy \rightarrow x^2-y^2$	-
^3E	22184	-	17848	18739	$xz,yz \rightarrow z^2,x^2-y^2$ [d]	-
$^3B_1$	22599	-		19662	$xy \rightarrow x^2-y^2$	-
$^3A_1$	23799	-		-	$xy \rightarrow x^2-y^2$	-
$^3A_2$	26525	-	23948	-	$xy \rightarrow z^2,x^2-y^2$ [d]	25000(z)

[a] ground state configuration. [b] spin-change within the ground state configuration. [c] spin-change within the ground state configuration with significant contributions from excited state configurations. [d] substantial multiconfigurational character. [e] two-electron transition. [f] not found in the energy range below the ligand-to-metal charge-transfer excitations. [g] calculated using the reported $S = 1$ DFT geometry optimized structure[89] showing negligible deviation from a D_{4h} symmetry.

Prediction of the d-d electronic transitions of a binuclear Cu^{II} bispidine complexes

Copper(I) complexes of tetra- or pentadentate mono- or dinucleating bispidine ligands (bispidine = methyl-2,4-bis(2-pyridinyl)-3,7-diazabicyclo-[3.3.1]-nonane-9-diol-1,5-dicarboxylate) are known to take part in oxygen activation reactions. For instance, in their reaction with catechols to quinones they have been shown to form unusually stable end-on [{(bispidine)Cu}$_2$(O$_2$)]$^{2+}$ complexes.[90] see Fig 14. For this binuclear complex, the metal centres are either antiferromagnetically or ferromagnetically coupled into an open-shell singlet or triplet spin state, respectively. In the present case, the two spin state surfaces for the oxidation of catechol to quinone by [{(bispidine)Cu}$_2$(O$_2$)]$^{2+}$ are close to being degenerate,[90] especially on the reactant side of the reaction mechanism. Particularly important for the description of the free energies along the reaction mechanism are both solvent and dispersion effects due to small energy differences between the available spin states, whereby the omission of one of these effects can easily change the spin state ordering. Also note that the treatment of solvent effects may vary depending on the solvent model and program package used (for instance, the COSMO[91] vs. PCM[47b,92] models). Similar situations occur in the description of the dispersion treatment through either the Grimme-D3 model[93] or the use of long-range dispersion functionals.[94] For the treatment of dispersion, no single "best" method has emerged to date; such calculations should therefore be conducted with at least two separate program packages or set-ups/approaches, and the relative energetics should match to have some confidence in the results.[90]

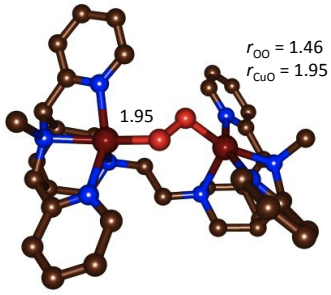


Fig. 14 Optimized geometry of the end-on [{(bispidine)Cu}$_2$(O$_2$)]$^{2+}$ complex.[90]

The d-d spectrum for [{(bispidine)Cu}$_2$(O$_2$)]$^{2+}$ was determined experimentally and subsequent TD-DFT computations led to the spectrum shown in Fig 15. As expected the spectra match only qualitatively, and among the three mentioned methods for the prediction of d-d transitions, i.e. the LF-DFT, SORCI and TD-DFT methods, probably the latter is the least accurate. TD-DFT also has fundamental problems.[95] In the case of the binuclear complex above, however, the semi-quantitative picture still allows to assign the experimentally observed transitions to the computed ones.

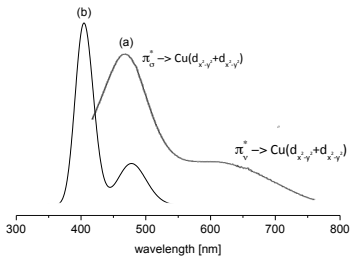


Fig. 15 Experimental (*a*) and TD-DFT calculated (*b*) absorption spectrum of end-on $[\{(bispidine)Cu\}_2(O_2)]^{2+}$. The transitions reflect $\pi^*_\sigma \rightarrow Cu(d_{x2-y2}+d_{x2-y2})$ [(*a*) 486 nm and (*b*) 405 nm] and $\pi^*_v \rightarrow Cu(d_{x2-y2}+d_{x2-y2})$ [(*a*) 650 nm and (*b*) 486 nm].[90]

Prediction of exchange coupling constants

Oligonuclear transition metal complexes with (multiple) unpaired electrons located on each metal centre can exist in a number of possible electronic states with the unpaired electron spins either ferro- or antiferromagnetically coupled. The ferro- or antiferromagnetic coupling is quantified through the exchange coupling constant *J*, Eq 17, where a negative *J* value corresponds to an antiferromagnetically coupled ground state and a positive one to a ferromagnetically coupled ground state. The magnetic properties of binuclear transition metal complexes is described via the Heisenberg-Dirac-van Vleck spin Hamiltonian (H_{HDvV}), Eq 17.

$$H_{HDvV} = 2\, J_{12} S_1 S_2 \qquad (17)$$

The broken-symmetry approach, Eq 18, can be used for a DFT-based calculation of the exchange coupling constants.[96] In this equation, E_{HS} and E_{BS} represent the energies of the high-spin (HS) and low-spin broken-symmetry (BS) states, whereas $\langle S^2\rangle_{HS}$ and $\langle S^2\rangle_{BS}$ are their spin-expectation values, respectively. The broken symmetry state refers to a symmetry lowering due to a spin-flip of an otherwise identical electronic configuration.

$$J_{12} = \frac{E_{BS} - E_{HS}}{\langle S^2\rangle_{HS} - \langle S^2\rangle_{BS}} \qquad (18)$$

These coupling constants can be calculated with most standard computational chemistry software packages including Jaguar,[47a] Gaussian,[47b] and ORCA.[97] The latter is especially simple to use as it directly outputs the broken symmetry-derived coupling constants on request.

Accurate predictions of *J* values depend strongly on the method and basis set used, as follows from calculations on the bisphenolate-bridged dicopper(II) complex, shown in Table 2. Thus, the hybrid DFT methods B3LYP,[28,29] B3P86,[28,57] and B3PW91[28,98] reproduce the experimental value within reasonable accuracy, which contrasts the results obtained from pure DFT methods that due to their poor description of the broken-symmetry state, do not give accurate results. The basis set should be of triple-ζ quality on the metal and its first coordination sphere, but can be much smaller for the atoms not involved in the electronic coupling, which should speed up the calculation.[99]

Table 2 Exchange coupling constant *J* for bisphenolate-dicopper(II) calculated using Eq 18 by DFT with different functionals and basis sets.[a,b]

Method	J_{G09} [c]	J_{Orca}	J_{Jaguar}
B3LYP/TZV	–229	–231	–231
B3P86/TZV	–238	–227	–241
B3PW91/TZV	–228	–230	–227
BLYP/TZV	–838	–838	–854
BP86/TZV	–861	–834	–880
BPW91/TZV	–831	–832	–848
PBE/TZV	–841	–841	–854
SVWN/TZV	–1156	–1178	–1181
B3LYP/3-21G	–103	–99	–114
B3LYP/TZVP	–215	–214	–231
B3LYP/6-31G*/TZVP [d]	–237	–216	–239

[a] in cm^{-1} [b] $J_{exp} = -298\ cm^{-1}$. [c] initial guess obtained with Jaguar.[47a] [d] Basis set: TZVP for Cu^{II} and the donor atoms, 6-31G* for the remaining atoms.

For the prediction of coupling constants, it is important to do an initial geometry optimization of the structures with DFT, rather than a single-point calculation on an available X-ray structure. This procedure avoids possible errors with the DFT method – as, strictly speaking, properties derived from the Kohn-Sham (KS) wave function are only valid at stationary points on the PES. On the other hand, correlating properties derived from DFT-relaxed geometries to experimentally determined ones also may have problems as the environment (e.g. solvent) used in the DFT calculation may be imperfectly modelled. Nonetheless one can show that *J* coupling constants for a number of oligonuclear transition metal complexes containing Cu, Mn, Fe, V, Ni, and Cr can be obtained with similar accuracy from DFT-optimized structures.[99]

Prediction of EPR *g* and *A* tensors of Cu(II) hexapeptides

The prediction of EPR *g* and *A* tensors with DFT methods is not straightforward, as we will exemplify with results on mononuclear copper(II) with 18-membered azacrown-6-macrocycle.[100] Fig 16 shows the B3LYP/6-31G* optimized structure for a cyclic copper(II) pseudo-hexapeptide overlaid with its corresponding X-ray structure. As can be seen from the overlays in Fig 16, even this modest method and basis set predicts the structure of the complex reasonably well. On the other hand, the accurate prediction of *g* and *A* tensors is much more complicated and the corresponding spin Hamiltonian (H_s) has the form as given in Eq 19.[101] and is essential for accurate predictions. It depends on the external magnetic field (*B*), the electron and nucleus spin-operators for the total spin of the system (*S* and *I*), the *g*-tensor, the hyperfine coupling constants (*A*) and copper (^{63,65}Cu) and nitrogen (^{14,15}N) isotopes.

$$H_s = \sum_{i=x,y,z} \left\{ \mathbf{B}_i \bullet g_i \bullet \mathbf{S}_i + \mathbf{S}_i \bullet A_i\left(^{63.65}\mathrm{Cu}\right) \bullet \mathbf{I}_i\left(^{63.65}\mathrm{Cu}\right) - g_n \beta_n \mathbf{B}_i \bullet \mathbf{I}_i \right\} + \sum_{j=1}^{3,4} \left\{ \mathbf{S}_i \bullet \mathbf{A}_i\left(^{14.15}\mathrm{N}\right) \bullet \mathbf{I}_i\left(^{14.15}\mathrm{N}\right) - g_n \beta_n \mathbf{B}_i \bullet \mathbf{I}_i \right\} \qquad (19)$$

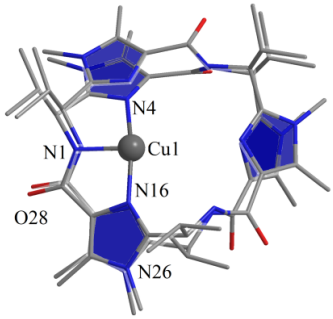


Fig. 16 Overlay of the X-ray structure of the copper(II) complex of hexapeptide H3L1 with the DFT optimized geometry (Gaussian 03, B3LYP/6-31G*); hydrogen atoms and co-ligands omitted for clarity.

DFT calculations using a number of different functionals and basis sets have been tested for three complexes with the program packages ORCA 2.6[97] and Mag-RESpect.[102] The results are given in graphical form in Fig 17 for the *g*-tensor (part *a*) and the *A*-tensor (part *b*) components. Usually, the *x* and *y* components of the *g*-tensor are obtained accurately with a hybrid DFT method with double to triple-ξ quality basis set, but difficulties in the correct description of the g_z component arises in this case. The calculated g_z value depends partly on the chosen basis set and, for instance, improves in quality upon changing the basis set from a split-valence basis set (SVP) to a much larger Wachters-type basis set[103] for copper and the first ligand sphere. In order to obtain values that are close to experiment, though, the amount of exact exchange in the DFT method needs to be modified (often arbitrarily, to a value of 38 or 40%, or alternatively through the use of a functional such as BHLYP which contains 50% exact exchange). Even with these modifications, all components of the *g*-tensor cannot be accurately reproduced with "standard" DFT methods. The MAG-Respect program gives values very close to experiment, employing the spin Hamiltonian as given in Eq 19, but unfortunately is computationally very demanding.[100] A similar situation occurs for calculations on the *A*-tensor, but as can be expected for the hyperfine coupling, accurate predictions critically depend on a very good basis set (even for the A_x and A_y components). Also here, improvement of the results is obtained by changing the total amount of exact exchange in the DFT method, whereby values beyond the 20% E_{HF} in B3LYP is needed, although in this case the A_z values are overestimated by a factor of 2. Only with spin-orbit coupling (SOC) treatments one can expect to obtain accurate predictions for the *A*-tensor.

Spin distribution and stability of a Cu(II) bispidine complex

For the well-characterized bispidine copper(II) complex, $[Cu(L^1)(Cl)]^+$, L^1 = methyl-2,4-bis(2-pyridinyl)-3,7-diazabicyclo-[3.3.1]-nonane-9-diol-1,5-dicarboxylate, Fig 18, an attempt was made to find a DFT methodology to predict structures that correctly predict the energetically favoured isomer, EPR *g*- and *A*-tensors and d-d transitions. As expected, there is not one single density functional method that performs best when all experimentally accessible observables are taken as a reference. Instead, for each property to be reproduced, a functional/method/basis set combination has to be carefully chosen and benchmarked. As we have demonstrated above, several DFT methods are suitable for the calculation of d-d transitions and EPR parameters, but in this section we will focus on the spin density distribution. We calculated spin densities for the complex given in Fig 18 and its structural isomer (where Cl is bonded perpendicular to the pyridine plane). The spin density on copper can vary strongly is dependent on the DFT method chosen,[103,104] and in our example it ranges from 0.57 using pure DFT functional to 0.93 for HF. As a matter of fact, the spin density on copper is linearly dependent on the amount of exact exchange admixture to the (hybrid) functional. The estimate for the experimental value[105] for the spin densities is reproduced quite accurately with the SORCI method, but for the DFT method the amount of exact exchange which would best fit the reference values would be 61%). A similar linear trend is observed for the spin density on the ligands, although for the chlorine ligand a value of 38% exact exchange would give the most accurate reproduction of the reference values.

To summarize, there does not appear to be a straightforward method to calculate electronic properties such as d-d transitions, EPR parameters or spin distributions from the existing DFT functionals without either using more specialized methods such as LF-DFT, Mag-ReSpect or SORCI for EPR values and d-d transitions, or to carefully tune the functional to reproduce a given property. The size of many model systems prohibits the use of a more computationally demanding method, and one has to resort to trial and error for picking the right combination of DFT method and basis set to compensate effects such as too covalent metal-to-ligand interactions in pure DFT methods. Still, the choice of an established functional can often be advantageous, because the weaknesses and strengths of a well-known functional such as B3LYP or BP86 are well understood, especially for geometry optimizations, which often give the right trends for a given property, if not the experimentally observed values. If accurate predictions of experimentally observables are needed, though, no single DFT method will give satisfactory results without careful tuning to experimental and/or high level *ab initio* reference data.

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

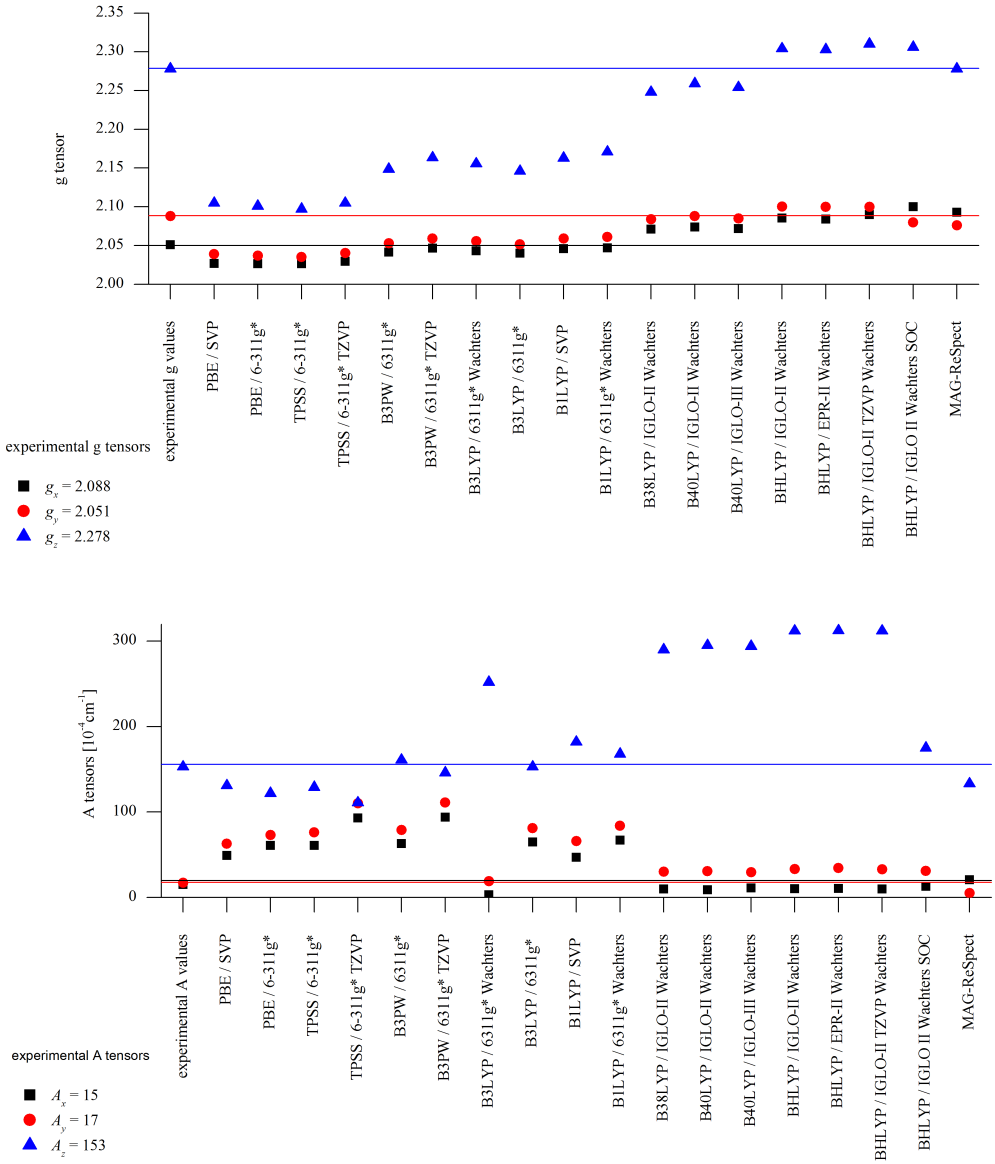


Fig. 17 Calculated EPR parameters of the mononuclear copper(II) complex of H3L1 [H2L1Cu^{II}(HOMe)]$^+$ as obtained with a range of DFT methods. (*a*) *g*-tensors [in 10^{-4} cm^{-1}]. (*b*) *A*-tensors [in 10^{-4} cm^{-1}]. Most accurate calculated parameters are depicted in bold.

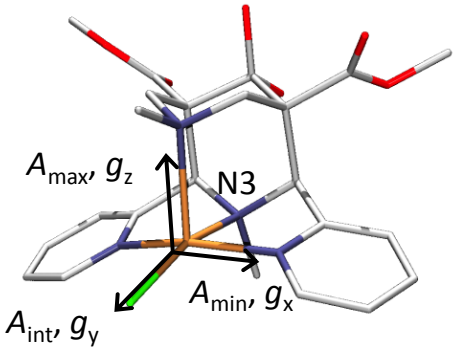
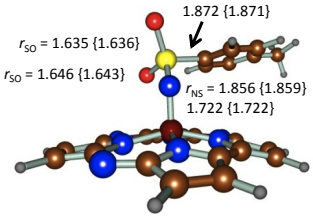


Fig. 18 X-ray structures of $[Cu(L^1)(Cl)]^+$ with Cl *trans* to N3.

Prediction of Mössbauer parameters

In a combined experimental and computational effort, the electronic properties and reactivity of a high-valent iron(IV)-imido complex was investigated.[106] Multiply bonded iron-nitrido and imido complexes have relevance to biological systems, including the enzyme nitrogenase that catalytically reduces nitrogen molecules to ammonia.[107] Our studies focused on iron(IV)-*N*-tosyl-corrolazine, $[Fe^{IV}(Cz^{+\bullet})NTs]$, which was investigated by several spectroscopic methods, namely UV-Vis, Mössbauer and EPR spectroscopic methods, alongside theory. The optimized geometries of the lowest lying doublet and quartet spin states calculated with UB3LYP and a double-ζ basis set are given in Fig 19. In analogy to cytochrome P450 Compound I,[13,14] $[Fe^{IV}(Cz^{+\bullet})NTs]$ has close-lying doublet and quartet spin states with three unpaired electrons: two in π^*_{FeN} orbitals (π^*_{xz}, π^*_{yz}) and a third in a corrolazine π^* orbital (a"). These three unpaired electrons are either antiferromagnetically or ferromagnetically coupled to the overall doublet or quartet spin states, respectively. We find a small energy splitting between the doublet and quartet spin states in favour of the low-spin state. This is in agreement with EPR studies that characterized it as an $S = ½$ spin state.

In addition, we calculated the Mössbauer and EPR parameters of $^2[Fe^{IV}(Cz^{+\bullet})NTs]$ in Orca,[97] which we report at the bottom of Fig 19 alongside the experimentally determined values. Spectroscopic parameters were calculated at the B3LYP[28,29] level of theory on the optimized geometries reported in Fig 19 in combination with the CP(PPP) basis set on Fe coupled to a TZVP basis set on the rest of the atoms.[107] The quadrupole splitting (ΔE_Q) was calculated from the electric field gradients V_i ($i = x, y,$ and z), the elementary charge of a proton/electron (e), the nuclear quadrupole moment $Q(^{57}Fe)$ of 0.16 barn, and the asymmetry parameter of the nuclear quadrupole tensor (η) using Eqs 20 and 21.



$\Delta E+ZPE+E_{solv}$ = 0.0 {5.7} kcal mol^{-1}

S =½:	δ	ΔE_Q	A_x	A_y	A_z
Exptl:	–0.05	2.95	26	26	3.5
Theory:	–0.12	+2.92	–4.8	24.0	30.9

$^2[Fe^{IV}(Cz^{+\bullet})NTs]$ $\{^4[Fe^{IV}(Cz^{+\bullet})NTs]\}$

Fig. 19 DFT optimized geometries of $^{2,4}[Fe^{IV}(Cz^{+\bullet})NTs]$ with bond lengths in angstroms. Also given are calculated Mössbauer and EPR spectroscopic parameters (in mm s^{-1}) of the S = ½ state as compared to experiment..

$$\Delta E_Q = \frac{1}{2} eQ \bullet V_z \bullet \sqrt{1 + \frac{\eta^2}{3}} \qquad (20)$$

$$\eta = \left(V_x - V_y\right)/V_z \qquad (21)$$

The isomer shift δ was calculated from the spin density at the iron nucleus ρ_0(Fe) using fit-parameters as implemented in Orca,[97] whereas the magnetic hyperfine parameters A_i ($i = x, y,$ and z), were obtained using the scalar relativistic zero-order regular approximation (ZORA) at the B3LYP level of theory. These calculated variables were then used to fit the experimental Mössbauer spectrum, taking into account that the Euler angles had to be rotated due to differences in the principal axis between the experimental and computational studies. Nevertheless, as follows from Fig 19, the computational methods reproduced the experimentally determined Mössbauer and EPR parameters excellently. Calculated values of δ and ΔE_Q are within 0.10 mm s^{-1} from experiment, although it should be noted that this type of agreement cannot be expected for all transition metal complexes at this level of theory. Therefore, precaution should be used to predict experimental Mössbauer and EPR spectroscopic parameters from DFT calculations.

Conclusions

Since the formulation of the Schrödinger equation and the development of computational quantum chemistry almost a century ago, theory has come a long way. Major advances in the field have been made, in particular in the area of methods development. Especially, computational resources have become more efficient and powerful that enables computational studies of relatively large (bio)chemical systems. Some of these techniques are highly accurate and are able to reproduce experimentally

measured variables well. However, for challenging chemical systems such as biomimetic transition metal complexes as well as biochemical systems, some major challenges remain. Currently, there is not a single universal method that is applicable to transition metal chemistry. Therefore, considerable testing and benchmarking alongside experiment is still needed. However, it appears that theory has established a strong foothold in biochemistry and biomimetic chemistry and more and more computational studies are done alongside experiment. As such theory can make predictions and guide experiment in the field and address reaction mechanisms and spectroscopic parameters of short-lived intermediates.

Acknowledgements

This investigation has been supported by grants from the Swedish research council (UR, project 2010-5025). It has also been supported by computational resources provided by the National Service of Computational Chemistry Software (NSCCS) to SdV and the Swedish National Infrastructure for Computing (SNIC) at Lunarc at Lund University (UR). MGQ thanks the BBSRC for a studentship. The EU-COST Network is thanked for funding for a short-term-scientific-mission to Heidelberg of SdV. PC and BM are grateful for financial support from the German Science Foundation (DFG) and the University of Heidelberg.

Notes and references

[a] *Manchester Institute of Biotechnology and School of Chemical Engineering and Analytical Science, the University of Manchester, 131 Princess Street, Manchester M1 7DN, United Kingdom. Fax: +44 1613065201; Tel: +44 1613064882; E-mail: sam.devisser@manchester.ac.uk*

[b] *Anorganisch-Chemisches Institut, Universität Heidelberg, Im Neuenheimer Feld 270, 69120 Heidelberg, Germany. Fax: +49 6221 546617; Tel: +49 6221 548453; E-mail: Peter.Comba@aci.uni-heidelberg.de*

[c] *Department of Theoretical Chemistry, Lund University, Chemical Centre, P. O. Box 124, SE-221 00 Lund, Sweden. Fax: +46 46 2228648; Tel: +46 462224502; E-mail: Ulf.Ryde@teokem.lu.se*

† Electronic Supplementary Information (ESI) available: Figures with calculated EPR *g* and *A* values using a range of DFT methods. See DOI: 10.1039/b000000x/

1 a) M. Sono, M. P. Roach, E. D. Coulter and J. H. Dawson, *Chem. Rev*., 1996, **96**, 2841; b) J. T. Groves, *Proc. Natl. Acad. Sci. USA*, 2003, **100**, 3569; c) P. R. Ortiz de Montellano, Ed. *Cytochrome P450: Structure, Mechanism and Biochemistry*. 3rd ed.; Kluwer Academic/Plenum Publishers, New York, 2004; d) B. Meunier, S. P. de Visser and S. Shaik, *Chem. Rev*., 2004, **104**, 3947; e) A. W. Munro, H. M. Girvan and K. J. McLean, *Nat. Prod. Rep*., 2007, **24**, 585; f) K. M. Kadish, K. M. Smith and R, Guilard, Eds. *Handbook of Porphyrin Science*. World Scientific Publishing Co., New Jersey, 2010; g) P. R. Ortiz de Montellano, *Chem. Rev*., 2010, **104**, 932; h) S. P. de Visser and D. Kumar, Eds. *Iron-containing enzymes: Versatile catalysts of hydroxylation reaction in nature*. RSC Publishing, Cambridge (UK), 2011.

2 a) E. I. Solomon, T. C. Brunold, M. I. Davis, J. N. Kemsley, S.-K. Lee, N. Lehnert, F. Neese, A. J. Skulan, Y.-S. Yang and J. Zhou, *Chem. Rev*., 2000, **100**, 235; b) T. D. H. Bugg, *Curr. Opin. Chem. Biol*., 2001, **5**, 550; c) M. J. Ryle and R. P. Hausinger, *Curr. Opin. Chem. Biol*., 2002, **6**, 193; d) M. Costas, M. P. Mehn, M. P. Jensen and L. Que Jr, *Chem. Rev*., 2004, **104**, 939; e) S. V. Kryatov, E. V. Rybak-Akimova and S. Schindler, *Chem. Rev*., 2005, **105**, 2175; f) M. M. Abu-Omar, A. Loaiza and N. Hontzeas, *Chem. Rev*., 2005, **105**, 2227; g) A. R. McDonald and L. Que Jr., *Coord. Chem. Rev*., 2013, **257**, 414; h) D. Buongiorno and G. D. Straganz, *Coord. Chem. Rev*., 2013, **257**, 541; i) S. Al-Attar and S. de Vries, *Coord. Chem. Rev*., 2013, **257**, 64.

3 See, e.g., a) L. Quintanar, L. Rivillas-Acevedo, R. Grande-Aztatzi, C. Z. Gómez-Castro, T. Arcos-López and A. Vela, *Coord. Chem. Rev*., 2013, **257**, 42; b) S. Fukuzumi and K. D. Karlin, *Coord. Chem. Rev*., 2013, **257**, 187; c) L. Rulíšek and U. Ryde, *Coord. Chem. Rev*., 2013, **257**, 445.

4 See, e.g., a) V. Conte, A. Coletti, B. Floris, G. Licini and C. Zonta, *Coord. Chem. Rev*., 2011, **255**, 2165; b) G. Licini, V. Conte, A. Coletti, M. Mba and C. Zonta, *Coord. Chem. Rev*., 2011, **255**, 2345.

5 See, e.g., a) P. J. Gonzalez, M. G. Rivas, C. S. Mota, C. D. Brondino, I. Moura and J. J. G. Moura, *Coord. Chem. Rev*., 2013, **257**, 315; b) H. Broda, S. Hinrichsen and F. Tuczek, *Coord. Chem. Rev*., 2013, **257**, 587.

6 See, e.g., a) A. B. Tomter, G. Zoppellaro, N. H. Andersen, H.-P. Hersleth, M. Hammerstad, Å. K. Røhr, G. K. Sandvik, K. R. Strand, G. E. Nilsson, C. B. Bell III, A.-L. Barra, E. Blasco, L. Le Pape, E. I. Solomon and K. K. Andersson, *Coord. Chem. Rev*., 2013, **257**, 3; b) P. Nordlund and P. Reichard, *Annu. Rev. Biochem*., 2006, **75**, 681; c) J. Herrick and B. Sclavi, *Mol. Microbiol*., 2007, **63**, 22.

7 See, e.g., a) N. Nelson and C. F. Yokum, *Annu. Rev. Biochem*., 2006, **57**, 521; b) C. W. Cady, R. H. Crabtree and G. W. Brudvig, *Coord. Chem. Rev*., 2008, **252**, 444; c) D. J. Vinyard, G. M. Ananyev and G. C. Dismukes, *Annu. Rev. Biochem*., 2013, **82**, 577; d) P. E. M. Siegbahn, *Chem. Eur. J*., 2008, **14**, 8290.

8 N. M. DeVore and E. E. Scott, *J. Biol. Chem*., 2012, **287**, 26576.

9 J. R. O'Brien, D. J. Schuller, V. S. Yang, B. D. Dillard and W. N. Lanzilotta, *Biochemistry*, 2003, **42**, 5547.

10 J. H. Dawson, R. H. Holm, J. R. Trudell, G. Barth, R. E. Linder, E. Bunnenberg, C. Djerassi, S. C. Tang, *J. Am. Chem. Soc*., 1976, **98**, 3707.

11 a) P. R. Ortiz de Montellano, *Annu. Rev. Pharmacol. Toxicol*., 1992, **32**, 89; b) M. Hofrichter, *Enzyme Microbial Technol*., 2002, **30**, 454; c) E. L. Raven, *Nat. Prod. Rep*., 2003, **20**, 367; d) T. L. Poulos, *Arch. Biochem. Biophys*., 2010, **500**, 3; e) F. Hollmann, I. W. C. E. Arends, K. Buehler, A. Schallmey and B. Bühler, *Green Chemistry*, 2011, **13**, 226.

12 a) P. Nicholls, I. Fita and P. C. Loewen, *Adv. Inorg. Chem*., 2000, **51**, 51; b) N. G. Veitch and A. T. Smith, *Adv. Inorg. Chem*., 2000, **51**, 107; c) H.-P. Hersleth, U. Ryde, P. Rydberg, C. H. Görbitz and K. K. Andersson, *J. Inorg. Biochem*., 2006, **100**, 460.

13 a) F. Ogliaro, S. P. de Visser and S. Shaik, *J. Inorg. Biochem*., 2002, **91**, 554; b) P. Rydberg, E. Sigfridsson and U. Ryde, *J. Biol. Inorg. Chem*., 2004, **9**, 203; c) D. Kumar, R. Latifi, S. Kumar, E. V. Rybak-Akimova, M. A. Sainna and S. P. de Visser, *Inorg. Chem*., 2013, **52**, 7968.

14 a) M. T. Green, *J. Am. Chem. Soc*., 1999, **121**, 7939; b) S. P. de Visser, S. Shaik, P. K. Sharma, D. Kumar and W. Thiel, *J. Am. Chem. Soc*., 2003, **125**, 15779; c) K. Yoshizawa, Y. Shiota and T. Yamabe, *J. Am. Chem. Soc*., 1999, **121**, 147.

15 J. Rittle and M. T. Green, *Science*, 2010, **330**, 933.

16 a) S. P. de Visser, *Angew. Chem. Int. Ed*., 2006, **45**, 1790 ; b) S. P. de Visser, *J. Am. Chem. Soc*., 2006, **128**, 9813; c) A. Decker, M. D. Clay and E. I. Solomon, *J. Inorg. Biochem*., 2006, **100**, 697.

17 a) S. P. de Visser, J.-U. Rohde, Y.-M. Lee, J. Cho and W. Nam, *Coord. Chem. Rev*., 2013, **257**, 381; b) K. P. Jensen, P. Rydberg, J. Heimdal and U. Ryde *A comparison of the tetrapyrrole cofactors in nature and their tuning by axial ligands.* In *Computational modeling for homogeneous and enzymatic catalysis*, K. Morokuma and J. Musaev, Eds., Wiley-VCH, Weinheim, 2008, pp. 27–56.

18 E. Schrödinger, *Phys. Rev*., 1926, **28**, 1049.

19 a) P. W. Atkins and R. S. Friedman, *Molecular Quantum Mechanics*, 3rd Ed., Oxford University Press, 1997; b) F. Jensen, *Introduction to Computational Chemistry*, 2nd Ed., John Wiley & Sons, Ltd., Chichester, 2007; c) S. P. de Visser, "*Introduction to quantum behaviour – a primer*." In *Quantum Tunnelling in Enzyme-Catalysed Reactions*, R. K. Alleman and N. S. Scrutton (Eds.); RSC Publishing, Cambridge, 2009, Chapter 2, p 18.

20 C. Møller and M. S. Plesset, *Phys. Rev*., 1934, **46**, 0618.

21 J. A. Pople, M. Head-Gordon and K. Raghavachari, *J. Chem. Phys*., 1987, **87**, 5968.

22 S. R. Langhoff, Ed. *Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy;* Kluwer, Dordrecht, 1995.

23 J. W. W. McDouall, *Computational Quantum Chemistry*, RSC Publishing, Cambridge, 2013.

24 a) P. Hohenberg and W. Kohn, *Phys. Rev. B*, 1964, **136**, 864; b) W. Kohn and L. J. Sham, *Phys. Rev. A*, 1965, **140**, 1133.

25 J. C. Slater, *The self-consistent field for molecular and solids.* In: *Quantum Theory of Molecular and Solids*, McGraw-Hill, New York, 1974, Vol. 4.

26 S. H. Vosko, L. Wilk and M. Nusair, *Can. J. Phys*., 1980, **58**, 1200.

27 A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098.

28 A. D. Becke, *J. Chem. Phys*., 1993, **98**, 5648.

29 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785.

30 M. Reiher, O. Salomon and B. A. Hess, *Theor. Chem. Acc.*, 2001, **107**, 48.

31 T. Schwabe and S. Grimme, *Phys. Chem. Chem. Phys*., 2007, **9**, 3397.

32 J. O. Hirschfelder, C. F. Curtiss and R. B. Bird, *Molecular Theory of Gases and Liquids*, Chapman & Hall, London, 1954.

33 a) D. Kumar, S. P. de Visser and S. Shaik, *Chem. Eur. J*., 2005, **11**, 2825; b) S. P. de Visser, K. Oh, A.-R. Han and W. Nam, *Inorg. Chem*., 2007, **46**, 4632; c) A. K. Vardhaman, C. V. Sastri, D. Kumar and S. P. de Visser, *Chem. Commun*., 2011, **47**, 11044.

34 O. Pestovsky and A. Bakac, *Inorg. Chem*., 2006, **45**, 814.

35 J. H. Jensen, *Molecular Modeling Basics*. CRC Press, Boca Raton, 2010.

36 J. Tomasi, B. Mennucci and R. Cammi, *Chem. Rev*., 2005, **105**, 2999.

37 a) C. M. Breneman and K. B. Wiberg, *J. Comput. Chem*., 1990, **11**, 361; b) M. Cossi, N. Rega, G. Scalmani and V. Barone, *J. Comput. Chem*., 2003, **24**, 669; c) A. Klamt and J. Schüürmann, *J. Chem. Soc. Perkin Trans 2*, 1993, **5**, 799; d) A. V. Marenich, C. J. Cramer, and D. G. Truhlar, *J. Phys. Chem. B*, 2009, **113**, 6378.

38 a) M. J. Ramos and P. A. Fernades, *Acc. Chem. Res*., 2008, **41**, 689; b) P. E. M. Siegbahn and F. Himo, *WIREs Comput. Mol. Sci*., 2011, **1**, 323.

39 a) T. Borowski, A. Bassan and P. E. M. Siegbahn, *Biochemistry*, 2004, **43**, 12331; b) S. P. de Visser, *Chem. Eur. J*., 2006, **12**, 8168.

40 a) R. Sevastik and F. Himo, *Bioorg. Chem*., 2007, **35**, 444; b) P. Georgieva and F. Himo, *J. Comp. Chem*., 2010, **31**, 1707; c) L. Hu, J. Eliasson, J. Heimdal and U. Ryde, *J. Phys. Chem. A*, 2009, **113**, 11793.

41 a) A. Warshel and M. Levitt, *J. Mol. Biol*., 1976, **103**, 227; b) H. M. Senn and W. Thiel, *Angew. Chem. Int. Ed*., 2009, **48**, 1198; c) R. Lonsdale, K. E. Ranaghan and A. J. Mulholland, *Chem. Comm*., 2010, **46**, 2354.

42 H. M. Senn and W. Thiel, *Top. Curr. Chem*., 2007, **268**, 173.

43 N. I. Reuter, A. Dejaegere, B. Maigret and M. Karplus, *J. Phys. Chem*. 2000, **104**, 1720.

44 L. Hu, P. Söderhjelm and U. Ryde, *J. Chem. Theory Comput*., 2011, **7**, 761.

45 a) T. H. Rod and U. Ryde, *J. Chem. Theory Comput*., 2005, **1**, 1240; b) T. Vreven, K. S. Byun, I. Komáromi, S. Dapprich, J. A. Montgomery, K. Morokuma and M. J. Frisch, *J. Chem. Theory Comput*., 2006, **2**, 815.

46 a) P. Söderhjelm, C. Husberg, A. Strambi, M. Olivucci and U. Ryde, *J. Chem. Theory Comput*., 2009, **5**, 649; b) C. Steinmann, D. G. Fedorov and J. H. Jensen, *J. Phys. Chem. A* 2010, **114**, 8705.

47 a) *Jaguar* 7.6, Schrödinger LLC, New York NY, 2007; b) M. J. Frisch et al, *Gaussian*-09, Revision A.02, Wallingford, CT, 2009; c) A. K. Rappé, C. J. Casewit, K. S. Colwell, W. A. Goddard III and W. M. Skiff, *J. Am. Chem. Soc*., 1992, **114**, 10024; d) W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J. Am. Chem. Soc*., 1995, **117**, 5179; e) A. D. MacKerell, D. Bashford, J. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B*, 1998, **102**, 3586; f) G. te Velde, F. M. Bickelhaupt, S. J. A. van Gisbergen, C. Fonseca Guerra, E.-J. Baerends, J. G. Snijders and T. Ziegler, *J. Comp. Chem*., 2001, **22**, 931; g) M. Svensson, S. Humbel, R. D. J. Froese, T. Matsubara, S. Sieber and K. Morokuma, *J. Phys. Chem*., 1996, **100**, 19357; h) U. Ryde, *J. Comput.-Aided Mol. Design*, 1996, **10**, 153; i) P. Sherwood et al, *J. Mol. Struct*., 2003, **632**, 1.

48 U. Ryde, *Dalton Trans*., 2007, 607.

49 H. Hu and W. Yang, *Annu. Rev. Phys. Chem*., 2008, **59**, 573.

50 S. Sumner, P. Söderhjelm and U. Ryde, *J. Chem. Theory Comput*., 2013, **9**, 4205.

51 a) L. Hu, P. Söderhjelm and U. Ryde, *J. Chem. Theory Comput*., 2013, **9**, 640; b) R. Z. Liao and W. Thiel, *J. Comp. Chem*. 2013, doi 10.1002/jcc.23403.

52 H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, *Nucleic Acids Res*., 2000, **28**, 235.

53 D. Kumar, W. Thiel and S. P. de Visser, *J. Am. Chem. Soc*., 2011, **133**, 386.

54 a) M. H. Stipanuk, *Annu. Rev. Nutr*., 2004, **24**, 539; b) G. D. Straganz and B. Nidetzky, *ChemBioChem*., 2006, **7**, 1536; c) C. A. Joseph and M. J. Maroney, *Chem. Commun*., 2007, 3338.

55 a) S. Aluri and S. P. de Visser, *J. Am. Chem. Soc*., 2007, **129**, 14846; b) S. P. de Visser and G. D. Straganz, *J. Phys. Chem. A*, 2009, **113**, 1835.

56 J. P. Perdew, *Phys. Rev. B*, 1986, **33**, 8822.

57 a) N. C. Handy and A. J. Cohen, *Mol. Phys*., 2001, **99**, 403; b) W.-M. Hoe, A. Cohen and N. C. Handy, *Chem. Phys. Lett*., 2001, **341**, 319; c) J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett*., 1996, **77**, 3865.

58 K.-B. Cho, H. Chen, D. Janardanan, S. P. de Visser, S. Shaik and W. Nam, *Chem. Commun*., 2012, **48**, 2189.

59 J. A. Crawford, W. Li and B. S. Pierce, *Biochemistry*, 2011, **50**, 10241.

60 a) H. Hirao, D. Kumar, L. Que Jr and S. Shaik, *J. Am. Chem. Soc*., 2006, **128**, 8590; b) S. P. de Visser, *J. Am. Chem. Soc*., 2006, **128**, 15809; c) R. Latifi, M. A. Sainna, E. V. Rybak-Akimova and S. P. de Visser, *Chem. Eur. J*., 2013, **19**, 4058.

61 C. R. Simmons, K. Krishnamoorthy, S. L. Granett, D. J. Schuller, J. E. Dominy Jr., T. P. Begley, M. H. Stipanukand P. A. Karplus, *Biochemistry*, 2008, **47**, 11390.

62 S. M. Pratter, C. Konstantinovics, C. L. M. DiGiuro, E. Leitner, D. Kumar, S. P. de Visser, G. Grogan and G. D. Straganz, *Angew. Chem. Int. Ed*., 2013, **52**, 9677.

63 R. H. Holm, E. I. Solomon, A. Majumdar and A. Tenderholt, *Coord. Chem. Rev*., 2011, **255**, 99.

64 J.-L. Li and U. Ryde, *J. Biol. Inorg. Chem*., submitted.

65 S. Metz and W. Thiel, *J. Am. Chem. Soc*., 2009, **131**, 14885.

66 J. Heimdal, M. Kaukonen, M. Srnec, L. Rulíšek and U. Ryde, *Chem. Phys. Chem*., 2011, **12**, 3337.

67 J.-L. Li, R. A. Mata and U. Ryde, *J. Chem. Theory Comput*., 2013, **9**, 1799.

68 N. Cobb, T. Conrads and R. Hille, *J. Biol. Chem*., 2005, **280**, 11007.

69 a) C. E. Webster and M. B. Hall, *J. Am. Chem. Soc*., 2001, **123**, 5820; b) A. Thapper, R. J. Deeth and E. Nordlander, *Inorg. Chem*., 2002, **41**, 6695; c) J. P. McNamara, I. H. Hillier, T. S. Bhachu and C. D. Garner, *Dalton Trans*., 2005, 3572; d) M. Hofmann, *J. Mol. Struct*., 2006, **773**, 59; e) M. Hofmann, *Inorg. Chem*., 2008, **47**, 5546; f) E. Hernandez-Marin and T. Ziegler, *Can. J. Chem*., 2010, **88**, 683; g) A. L. Tenderholt, J. J. Wang, R. K. Szilagyi, R. H. Holm, K. O. Hodgson, B. Hedman and E. I. Solomon, *J. Am. Chem. Soc*., 2010, **132**, 8359.

70 a) E. I. Solomon, A. J. Augustine and J. Yoon, *Dalton Trans*., 2008, 3921; b) E. I. Solomon, J. W. Ginsbach, D. E. Heppner, M. T. Kieber-Emmons, C. H. Kjaergaard, P. J. Smeets, L. Tian and J. S Woertink, *Faraday Discuss*., 2011, **148**, 11.

71 W. Shin, U. M. Sundaram, J. L. Cole, H. H. Zhang, B. Hedman, K. O. Hodgson and E. I. Solomon, *J. Am. Chem. Soc*., 1996, **118**, 3202.

72 S.-K. Lee, S. D. George, W. E. Antholine, B. Hedman, K. O. Hodgson and E. I. Solomon, *J. Am. Chem. Soc*., 2002, **124**, 6180.

73 L. Rulíšek, E. I. Solomon and U. Ryde, *Inorg. Chem*., 2005, **44**, 5612.

74 J. Chalupský, F. Neese, E. I. Solomon, U. Ryde and L. Rulíšek, *Inorg. Chem*., 2006, **45**, 11051.

75 S. Vancoillie, J. Chalupský, U. Ryde, E. I. Solomon, K. Pierloot, F. Neese and L. Rulíšek, *J. Phys. Chem. B*, 2010, **114**. 7692.

76 U. Ryde, Y.-W. Hsiao, L. Rulíšek and E. I. Solomon, *J. Am. Chem. Soc*., 2007, **129**, 726.

77 J. Yoon, L. M. Mirica, T. D. P. Stack and E. I. Solomon, *J. Am. Chem. Soc*., 2005, **127**, 13680.

78 a) F. Neese, D. G. Liakos and S. Ye, *J. Biol. Inorg. Chem*., 2011, **16**, 821; b) M. Güell, J. M. Luis, M. Solà and M. Swart, *J. Phys. Chem. A*, 2008, **112**, 6384; c) M. Swart, M. Güell and M. Solà, *Phys. Chem. Chem. Phys*., 2011, **13**, 10449.

79 H. B. Schlegel and M. A. Robb, *Chem. Phys. Lett*., 1982, **93**, 43.

80 a) B. O. Roos, K. Andersson, M. P. Fulscher, L. Serrano-Andres, K. Pierloot, M. Merchan and V. Molina, *J. Mol. Struct. (Theochem)*, 1996, **338**, 257; b) B. O. Roos, *Acc. Chem. Res*., 1999, **32**, 137.

81 C. Angeli, R. Cimiraglia and J. P. Malrieu, *J. Chem. Phys*., 2002, **117**, 9138.

82 a) C. Hampel and H. J. Werner, *J. Chem. Phys*., 1996, **104**, 6286; b) F. Neese, A Ansen and D. G. Liakos, *J. Chem. Phys*., 2009, **131**, 15.

83 F. Neese, *J. Chem. Phys*., 2003, **119**, 9428.

84 a) E. K. U. Gross and W. Kohn, *Adv. Quantum Chem*., 1990, **21**, 255; b) M. E. Casida, *Time Dependent Density Functional Response Theory for Molecules.* In: *Recent advances in density functional methods*, D. P. Chong (Ed), World Scientific, Singapore, 1995, Vol. 1, p. 155; c) E. K. U. Gross, J. F. Dobson and M. Petersilka, *Density functional theory of time-dependent phenomena.* In: *Density Functional Theory, Springer Series: Topics in Current Chemistry*, R. F. Nalewajski (Ed), Springer, Berlin, Germany, 1996; d) S. J. A. Van Gisbergen, J. G. Snijders and E.-J. Baerends, *J. Chem. Phys*., 1995, **103**, 9347; e) R. Bauernschmitt and R. Ahlrichs, *Chem. Phys. Lett*., 1996, **256**, 454; f) A. Rosa, G. Ricciardi, O. Gritsenko and E.-J. Baerends, *Struct. Bond*., 2004, **112**, 49; g) F. Wang and T. Ziegler, *Mol. Phys*., 2004, **102**, 2585; h) A. Dreuw and M. Head-Gordon, *Chem. Rev*., 2005, **105**, 4009; i) F. Wang, T. Ziegler, E. van Lenthe, S. J. A. Van Gisbergen and E.-J. Baerends, *J. Chem. Phys*., 2005, **122**, 204103; j) M. Seth and T. Ziegler, *J. Chem. Phys*., 2005, **23**, 144105.

85 a) M. Atanasov, P. Comba, C. A. Daul and F. Neese, *The Ligand Field Paradigm and New Insights into the Electronic Properties of Transition Metal Complexes Based on Modern Electronic Structure Calculations.* In *Models, Mysteries and Magic of Molecules*, J. C. A. Boeyens and J. F. Ogilviee (Eds.), Springer, Dordrecht, the Netherlands, 2008; b) M. Atanasov, C. A. Daul and C. Rauzy, *Chem. Phys. Lett*., 2003, **367**, 737; c) M. Atanasov, C. Daul and C. Rauzy, *Struct. and Bonding*, 2004, **106**, 97; d) C. Daul, C. Rauzy, M. Zbiri, P. Bättig, R. Bruyndonckx, E.-J. Baerends and M. Atanasov, *Chem. Phys. Lett*., 2004, **399**, 433; e) M. Atanasov, C. Rauzy, P. Bättig and C. Daul, *Int. J. Quantum Chem*., 2005, **102**, 119.

86 a) M. Atanasov, P. Comba and C. A. Daul, *Inorg. Chem*., 2008, **47**, 2449; b) M. Atanasov, P. Comba, G. R. Hanson, S. Hausberg, S. Helmle, and H. Wadepohl, *Inorg. Chem*., 2011, **50**, 6890; c) M. Atanasov, P. Comba and S. Helmle, *Inorg. Chem*., 2012, **51**, 9357.

87 M. Atanasov, P. Comba, S. Helmle, D. Müller and F. Neese, *Inorg. Chem*., 2012, **51**, 12324.

88 S. P. de Visser, R. Latifi, L. Tahsini and W. Nam, *Chem. Asian J*., 2011, **6**, 493.

89 A. Decker, J.-U. Rohde, L. Que Jr and E. I. Solomon, *J. Am. Chem. Soc*., 2004, **126**, 5378.

90 P. Comba, B. Martin, A. Muruganantham and J. Straub, *Inorg. Chem*., 2012, **51**, 9214.

91 A. Klamt and G. Schüürmann, *J. Chem. Soc. Perkin Trans.* 2, 1993, 799.

92 a) S. Miertus, E. Scrocco and J. Tomasi, *Chem. Phys*., 1981, **55**, 117; b) J. L. Pascual-Ahuir, E. Silla and I. Tunon, *J. Comp. Chem*., 1994, **15**, 117.

93 a) S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys*., 2010, **132**, 154104; b) S. Grimme, S. Ehrlich and L. Goerigk, *J. Comp. Chem*., 2011, **32**, 1456.

94 J.-D. Chai and M. Head-Gordon, *J. Chem. Phys*., 2008, **128**, 084106.

95 J. Schirmer and A. Dreuw, *Phys. Rev. A*, 2007, 122513.

96 a) A. Aizman and D. A. Case, *J. Am. Chem .Soc*., 1982, **104**, 3269; b) K. Yamaguchi, Y. Takahara and T. Fueno, *Applied Quantum Chemistry*, V. H. Smith (Ed.), Reidel Inc., Dordrecht, 1986; c) R. Sessoli and D. Gatteschi, *Angew. Chem. Int. Ed*., 2003, **42**, 268; d) J. R. Long, In *Molecular Cluster Magnets in Chemistry of Nanostructured Materials*, P. Yang (Ed.), World Scientific, Hong Kong, 2003, p 291; e) I. Rudra, Q. Wu, T. van Voorhis, *J. Chem. Phys*., 2006, **124**, 024103; f) M. Atanasov, P. Comba, S. Hausberg and B. Martin, *Coord. Chem. Rev*., 2009, **253**, 2306.

97 F. Neese, *ORCA – an ab initio, DFT and semiempirical SCF-MO package*, Bonn, Germany, Version 2.7-00, 2009.

98 J. P. Perdew, K. Burke and Y. Wang, *Phys. Rev. B*, 1996, **54**, 16533.

99 P. Comba, S. Hausberg and B. Martin, *J. Phys. Chem A*, 2009, **113**, 6751.

100 N. Dovail, PhD Dissertation, University of Heidelberg (Germany), 2010.

101 F. Neese, *J. Chem. Phys*., 2001, **115**, 11080.

102 V. G. Malkin, O. L. Malkina, R. Reviakine, A. F. Arbouznikov, M. Kaupp, B. Schimmelpfennig, I. Malkin, T. Helgaker and H. Ruud, MAG-ReSpect 1.2.

103 A. Wachters, *J. Chem. Phys*., 1969, **52**, 1033.

104 M. Güell, J. L. Luis, L. Rodríguez-Santiago, M. Sodupe and M. Solà, *J. Phys. Chem. A*, 2009, **113**, 1308.

105 M. Atanasov, P. Comba, B. Martin, V. Müller, G. Rajaraman, H. Rohwer and S. Wunderlich, *J. Comput. Chem*., 2006, **27**, 1263.

106 P. Leeladee, G. N. L. Jameson, M. A. Siegler, D. Kumar, S. P. de Visser and D. P. Goldberg, *Inorg. Chem*., 2013, **52**, 4668.

107 a) Y. Hu and M. W. Ribbe, *Acc. Chem. Res*., 2010, **43**, 475; b) B. M. Hoffman, D. Lukoyanov, D. R. Dean and L. C. Seefeldt, *Acc. Chem. Res*., 2013, **46**, 587.

108 F. Neese, *Inorg. Chim. Acta*, 2002, **337**, 181.

Author biographies

Sam de Visser graduated with a BSc in Analytical Chemistry from the College of Zeeland (the Netherlands) in 1990. He then moved to the University of Amsterdam (the Netherlands), where he obtained an MSc in Physical Chemistry (1993) and a PhD in Organic Mass Spectrometry (1997) with Prof Nico Nibbering. During his PhD studies he got in touch with computational modelling and moved into this field and specialized further during two postdoctoral positions at King's College London and the Hebrew University of Jerusalem. Since 2004 he has been a lecturer in Biophysics at the University of Manchester (United Kingdom) and is based at the Manchester Institute of Biotechnology. His research interests are in the field of biomimetic and metalloenzyme reaction mechanisms using QM-cluster as well as QM/MM methodologies. He has experience with heme enzymes (cytochromes P450, peroxidases, catalases) as well as nonheme iron dioxygenases. In addition, his group works on a range of transition metal containing biomimetic models in combination with experimental studies, where the calculations focus on reaction kinetics and spectroscopic characterization.

Sam P. de Visser

Matthew Quesne completed his undergraduate studies from the University of Lancaster (United Kingdom) with specialization in Biochemistry. He then moved to the University of Manchester, where he started his PhD studies under supervision of Dr Sam de Visser in September 2009. He has experience with computational modelling using QM-cluster and QM/MM techniques. His research interests include the understanding of biological reaction mechanisms of nonheme iron hydroxylases and halogenases. In addition, he has worked on synthetic model complexes of these nonheme iron systems.

Matthew Quesne

Bodo Martin obtained a diploma in Chemistry from the Friedrich-Alexander University in Erlangen (Germany) in 1999, and a PhD in Organic and Computational Chemistry from the same University in 2004. In 2002 he joined the workgroup of Peter Comba at the University of Heidelberg (Germany). He is interested in modelling transition metal complexes using *ab initio* and DFT methodologies as well as classical calculations of geometries, reaction mechanisms, and spectroscopic parameters. He is also involved in method development of novel classical force fields and property models for inorganic compounds.

Bodo Martin

Peter Comba obtained a diploma in Chemistry and Chemical Education from the ETH Zürich (Switzerland) and a PhD from the Université de Neuchâtel (Switzerland). After a Research Fellowship at the Australian National University, a postdoc at the University of Lausanne and a junior group leader position at the University of Basel (Switzerland), he joined the University of Heidelberg (Germany) in 1992. He is interested in fundamental properties of transition metal coordination chemistry, involving ligand design and synthesis, preparative coordination chemistry, spectroscopy as well as theoretical and computational inorganic chemistry, with projects in bioinorganic and medicinal inorganic chemistry, molecular magnetism and molecular catalysis.

Peter Comba

Ulf Ryde received his PhD in Biochemistry from Lund University (Sweden) under supervision of Prof G. Pettersson in 1991. He then moved into the field of Theoretical Chemistry at the same university, where he became a docent in 1996 and full professor 2004. During 2001 – 2007 he had a senior research position from the Swedish research council. He studies the structure and function of proteins, in particular, metalloproteins, such as blue copper proteins, heme enzymes, vitamin B_{12} enzymes, hydrogenases, and multi-copper oxidases. He has developed QM/MM methods for an accurate treatment of environmental effects, e.g. using accurate MM force fields with multipole expansions and anisotropic polarisation, and combinations of QM/MM with free-energy methods or experimental approaches, such as X-ray crystallography, NMR, and EXAFS. He also studies and develops methods to calculate ligand-binding affinities, especially combinations of MM and QM methods.

Ulf Ryde