Reductive cleavage of the O-O bond in multicopper oxidases: a QM/MM and QM study

Srnec, Martin; Ryde, Ulf; Rulisek, Lubomir

Published in:
Faraday Discussions

DOI:
10.1039/c004476h

2011

Reductive cleavage of the O–O bond in Multicopper Oxidases: QM/MM and QM Study

Martin Srnec,[a] Ulf Ryde[b] and Lubomír Rulíšek*[a]

DOI: 10.1039/b000000x [DO NOT ALTER/DELETE THIS TEXT]

The key step in the reaction mechanism of multicopper oxidases (MCOs) – cleavage of the O–O bond in O_2 – is investigated using combined quantum mechanical and molecular mechanical (QM/MM) methods. This process represents a reaction pathway from peroxy intermediate after it accepts one electron from the nearby type 1 Cu site to the experimentally observed native intermediate, which is the only fully oxidized catalytically relevant state of MCOs. Scans of the QM(DFT)/MM potential energy surface has allowed us to obtain estimates of the activation energies. Furthermore, vacuum calculations on a smaller model of the active site have allowed us to estimate the entropy contributions to the barrier height and to obtain further insight into the reaction by comparing the small cluster model with the QM/MM model that includes whole protein. Owing to the complicated electronic structure of these low-spin exchange coupled systems, multireference quantum chemical calculations at the second-order perturbation theory complete-active space (CASPT2) were attempted to benchmark the barrier heights obtained at the DFT(B3LYP) level. Our best estimate of the activation barrier is ΔG = 60–65 kJ.mol^{-1}, in a good agreement with the experimental barrier of ~55 kJ.mol^{-1} that can be inferred from the experimental rate constant, k > 350 s^{-1}. It is also shown that the reaction involves protonation of the O_2 moiety before the bond cleavage. The proton may come either from Cu-T2 ligand or from the surroundings, most likely a nearby carboxylate residue.

1 Introduction

In the last two decades, quantum chemical (QM) and combined quantum and molecular mechanical (QM/MM) calculations have proved to be a very useful tool in elucidating the reaction mechanisms of metalloproteins.[1,2,3,4,5,6,7] Theoretical modeling can, in principle, map the correspondence between conformational space and energy landscape, and characterize crucial points along the reaction coordinate, most notably the transition states (TSs).[8] However, there are two strict prerequisites to provide quantitatively correct answers and unambiguous conclusions: the methods should be accurate enough and the models should contain as little approximations as possible. It is a non-trivial task to meet both of these conditions, because it is not clear, for example, whether a large cluster model of the enzyme active site can give correct answers,[3] **xxx You may consider to add our recent article L. Hu, J. Eliasson, J. Heimdal, U. Ryde (2009) "Do quantum mechanical energies calculated for small models of protein active sites converge?", J. Phys. Chem. A, 113, 11793–11800, where we show that QM-only calculations have severe**

convergence problems. or whether one has to resort to QM/MM modeling which includes the full protein molecule.[1]

We consider the multicopper oxidases (MCOs) as an excellent example to address this question. MCOs are enzymes that couple four one-electron oxidations of a substrate with the four-electron reduction of molecular oxygen to water[9,10]

$$O_2 + 4e^- + 4H^+ \rightarrow 2H_2O \qquad (1)$$

This reaction takes place at a trinuclear copper cluster (TNC), whereas the substrate is oxidized at a type 1 copper site (Cu-T1) which is ~13 Å away from the TNC and is linked to it via bifurcated $(Cu\text{-}TNC)_2$-$(His)_2$-Cys-Cu-T1 link. The link is assumed to provide the electron transfer (ET) pathway between the two sites. The key aspects of the MCO reaction mechanism have been revealed by combining spectroscopic[11,12,13] and structural[14,15] information with QM/MM calculations,[16] mostly for the prototypical enzyme of the MCO family – laccase.[17,18] The QM/MM calculations were further supported by multireference calculations of the MCO spectroscopic properties[19] and combined extended X-ray absorption fine-structure EXAFS/QM/MM calculations.[20] This has led to the consensus reaction mechanism depicted in Figure 1.

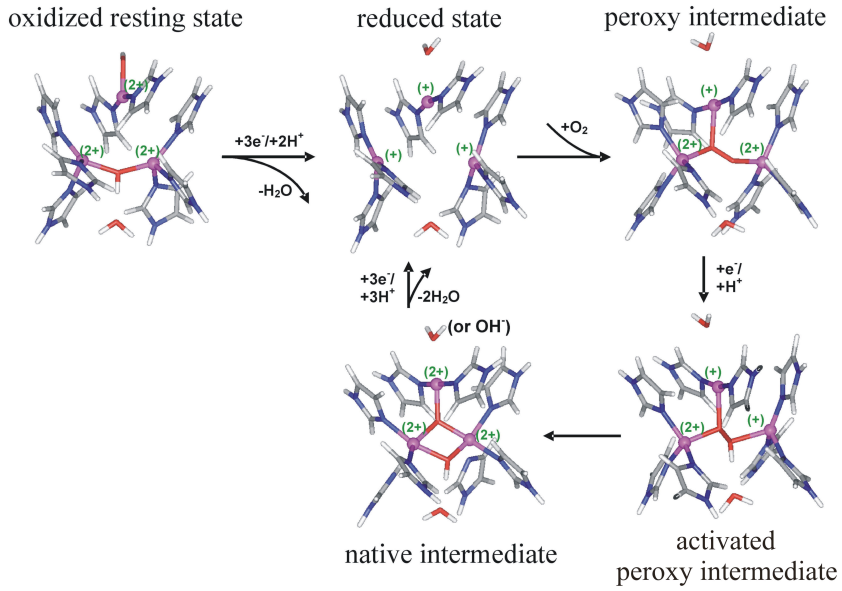

Figure 1: The consensus MCO reaction mechanism.[16]

The reaction starts with the fully reduced form of the enzyme, presumably with a water molecule weakly coordinated to the type 2 copper (Cu-T2) ion and no bridging moiety between the two type 3 copper (CuT3) ions.[16,21] The incoming dioxygen is immediately reduced, giving a peroxide-level intermediate (**PI**),[11] with a peroxide ion bound in the centre of the TNC. The reaction is completed by the uptake of one

electron from Cu-T1 site (resulting in an activated peroxy intermediate **NI'**) to yield the native intermediate (**NI**),[13] a spectroscopically characterized, catalytically relevant, fully oxidized form of the enzyme. The **NI** has one μ_3-oxo ligand bridging all three copper ions and a μ_2-hydroxo ligand bridging the two CuT3 ions.[16,19] The resting oxidized state is sugested to not be involved under catalytic turnover conditions[13] since its transformation into the fully reduced form is too slow compared to the overall reaction rate. The mechanism in the form depicted in Figure 1 was first proposed by Rulíšek *et al.*,[16] later confirmed experimentally by Yoon *et al.*,[22] and has been thoroughly reviewed recently.[21]

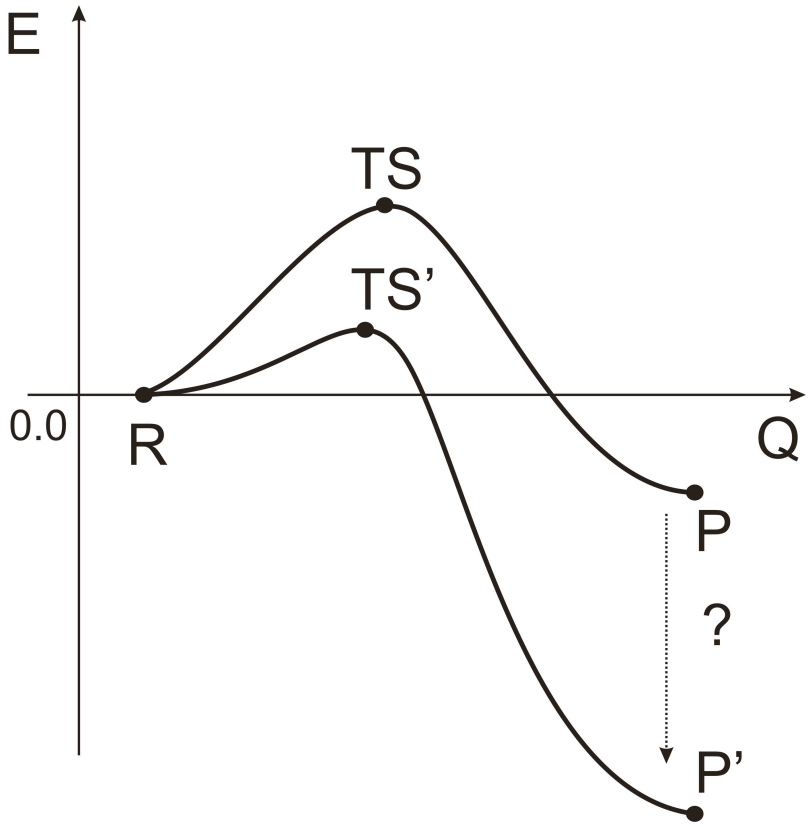

Figure 2: Overstabilization of the product state may lead to an artificial lowering of the activation barrier.[26] A similar effect was recently shown on the model reactions of boron compounds.[23]

The energetics of the O_2-cleavage reaction step (**NI'** → **NI** pathway) was recently studied in detail by Yoon and Solomon[24] using a cluster model (i.e., a truncated system representing active site with many atoms at the edge of the QM system fixed at their crystallographic positions) and an activation barrier of ~20 kJ.mol^{-1} was found, indicating a facile reaction. However, there are three issues related to the presented MCO reaction coordinate worth to address by QM/MM calculations that

consider the whole protein. First, in the cluster model, one has to fix atoms at the borderline of the QM system, an approximation that is automatically taken care of (avoided) in the QM/MM calculations. For example, the calculated reaction energy of the O_2-cleavage in MCO using cluster model (ΔE = –213 kJ.mol^{-1})[24] originates from the high exothermicity (ΔH) of the uncatalyzed $H_2 + H_2O_2 \rightarrow 2\ H_2O$ reaction (ΔH = –347 kJ.mol^{-1})[25] and one may expect this large exothermicity is even more buffered by the protein. In our opinion, this large thermodynamic driving force in the MCO cluster model may imply an artificial lowering of the activation barrier, according to the Hammond postulate (see Figure 2).[26]

The accuracy of the cluster-model approximation is even more relevant for MCOs (in comparison with other metalloproteins) since the TNC in many of its oxidation and protonation states is inherently unstable *in vacuo*.[16] Its stability is maintained in the protein by the presence of neutralizing carboxylate residues that are conserved in MCO family[16] and four His-X-His motifs (with His being ligands of copper ions in TNC) that are sewing the TNC site together. This raises the second issue, namely what size of the model system is needed in QM calculations. Only after explicitly including carboxylate moieties into the calculations were correct structure of the PI obtained in the QM calculations.[24] Finally, despite DFT methods are widely used in the calculations of bioinorganic systems and most often provide results of satisfactory accuracy, one may attempt to carefully benchmark them, especially for systems with complicated electronic structures like the exchange-coupled copper ions in oxidized TNC.

The aim of this work is to perform a careful investigation of the O_2 cleavage reaction in the MCOs — by QM/MM methods, performing a structural and energetic characterization of the activation barriers for various protonation states of the active site, complementing the *in vacuo* data obtained by Yoon and Solomon.[24] The QM/MM description quite naturally includes both electrostatic and sterical effect of the protein environment and can in principle yield correct energetics along the reaction pathway. Moreover, we also attempt to estimate entropic contributions acquired from a cluster model and to benchmark DFT methods by high-level multireference *ab initio* calculations. This can complete our understanding of the reaction mechanism of MCOs at the theoretical level that includes full protein environment.

2 Computational Details

Combined Quantum Mechanical and Molecular Mechanical Calculations

All QM/MM calculations were carried out using the ComQum program.[27,28] It is a combination Turbomole 5.7[29] for the QM part with AMBER 8[30] and the Cornell *et al.* force field[31] for the MM part. In this approach, the protein and solvent are split into three subsystems: The QM region (system 1, S1) contains the most interesting atoms and is relaxed by QM/MM forces. System 2 consists of all residues within 6 Å of any atom in system 1 and is relaxed by a full MM minimization in each step of the QM/MM geometry optimization. Finally, system 3 contains the remaining part of the protein and surrounding solvent molecules, and is kept fixed at the original coordinates. In the quantum chemical calculations, the QM system is represented by a wave function, whereas all the other atoms are represented by an array of partial point charges, one for each atom, taken from Amber libraries. The total QM/MM

energy is then calculated as:

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

where E_{QM1} is the QM energy of system 1, including a point-charge model of the surroundings (electric embedding), E_{MM123} is the MM energy of the full system, but ignoring the electrostatic interactions between system 1 and the other systems (to avoid double-counting), and E_{MM1} is the MM energy of system 1. The computational protocol is essentially identical to the one used previously in characterization of key intermediates in MCO[16]. In Ref. 16, further details of QM/MM procedure can be found.

Protein Setup

All QM/MM calculations are based on the 1.4-Å structure of CueO (PDB code 1KV7).[32] This structure was selected because it had the best resolution among the published MCO structures at the beginning of this investigation. In addition, it is a monomer and lacks glycosylated surface residues, which are common in the eukaryotic proteins. Hydrogen atoms were added to the crystal structure and the total system was solvated in a sphere of water molecules with a radius of 38 Å. The positions of the hydrogen atoms and solvent water molecules were then optimized by a 90-ps simulated-annealing calculation with molecular dynamics followed by a conjugate-gradient energy minimization of their positions. We assumed the normal protonation state at pH 7 for all amino acids, except for the copper-bound Cys residue, which was assumed to be deprotonated. For the His residues, the protonation status was decided form a detailed study of the hydrogen-bond network around the residue and the solvent accessibility. Thus, His-71, 111, 113, 393, 395, 446, and 448 were assumed to be protonated on the N_δ^1 atom, His-73, 390, and 452 on the N_ε^2 atom, and the other eight His residues on both these atoms. The space created by a disordered loop, missing in the CueO crystal structure (residues 380–402; more than 22 Å from the trinuclear copper cluster) was filled by water molecules. System 1 (the QM system) consisted of the following residues: the three copper ions of TNC and their eight His ligands, O_2 or water derived ligands of the three copper ions, two carboxylate groups close to the TNC (Asp-82 and Glu-453), as well as three water molecules bridging between them with the TNC. The size of the quantum system was approximately 105 atoms (cf. Fig. 3).

Quantum chemical calculations

All quantum chemical calculations were performed at the density functional theory (DFT) level. Geometry optimizations were carried out using the Perdew–Burke–Ernzerhof (PBE) functional.[33] The DFT/PBE calculations were expedited by expanding the Coulomb integrals in an auxiliary basis set (the resolution-of-identity approximation, RI-J).[34,35] All the geometry optimizations were carried out using the 6-31G(d) basis set for all atoms,[36] except for copper, for which we used the DZP basis sets of Schäfer *et al.*[37] (referred to as the DZP basis set). More accurate energies were then estimated by single-point calculations using a larger basis set: the def2-TZVP[38] and Becke's three-parameter hybrid functional (B3LYP).[39] The structures were optimized until the change in energy between two iterations was below 0.026 $J.mol^{-1}$ (10^{-8} a.u.) and the maximum norm of the internal gradients was below 5.0 $kJ.mol^{-1}Å^{-1}$ (10^{-3} a.u). Zero-point energies, thermal corrections to the Gibbs free energy, and entropic terms were obtained from a normal-mode analysis

on small *in vacuo* models (unlike system 1 in QM/MM model, these models do not contain carboxylate groups and two additional waters) of the active site (*i.e.* ~85 atoms) using the same method and software as for the geometry optimizations. They were calculated at 298.15 K and 1 atm pressure, using an ideal-gas, rigid-rotor harmonic-oscillator approximation.[40] All the calculations were carried out for the antiferromagnetically coupled S = 1/2 potential energy surface, which is a ground state for both reactants and products. In DFT, it rather means a M_S=1/2 state, since we use the spin-flipped Kohn-Sham determinants starting from the unrestricted high-spin configurations.

The complete active space self-consistent field (CASSCF),[41] and complete active space second-order perturbation theory (CASPT2)[42] calculations were carried out using MOLCAS 7.0 program.[43] The active space comprises fifteen electrons distributed in nine orbitals as schematically shown in Figure 3. In all the state-averaged CASSCF calculations (over the two near degenerate doublet states), a level shift of 5.0 a.u. was used to improve the convergence of the multireference wave function. In the CASPT2 calculations, an imaginary level shift of 0.2 a.u. was used to eliminate intruder states.[44] The ANO-S basis set[45] with the following contractions was used: Cu [5s4p2d1f], C, O, N [3s2p], H [2s].

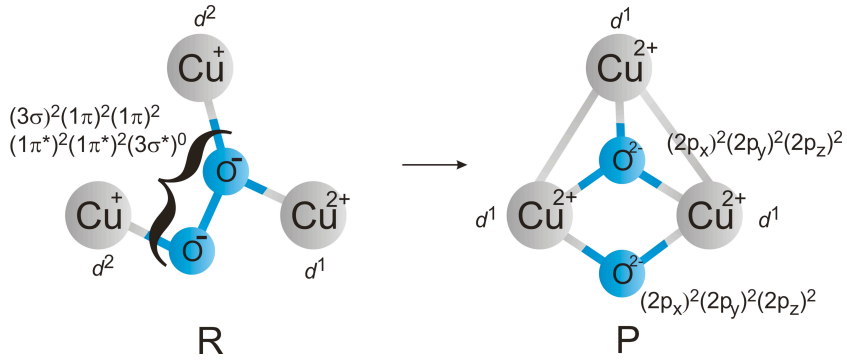

Figure 3. Schematic description of a minimum active space, including the most important orbitals involved in the catalytic O–O bond cleavage by MCOs.

Scans of QM/MM potential energy. Approximate transition states in QM/MM models

Due to the lack of the QM/MM analytical second derivatives in ComQum, two different strategies for mapping QM/MM potential energy surface and searching for transition-state structures along reaction coordinates were adopted in this study. The first (and more conventional) approach (*cf.* Figure 4A) starts from the reactant structure (**NI'**) and scans the elongation of the O–O bond. Once product structure is reached, scan in the reverse direction is carried out. These forth and back scans should converge to a potential-energy profile independent of the starting (*i.e.* reactant or product) structure. The disadvantage of such procedure is the relatively slow convergence to a stable pathway. Alternatively, another strategy in the search for the QM/MM TS was adopted (*cf.* Figure 4B), which can be described as follows:

(i) the TS of the gas-phase cluster (*in vacuo* QM model) with one imaginary frequency is found; (ii) QM/MM optimization of the whole protein is carried out with the key atoms in the quantum regions (S1) restrained to the TS obtained in QM model (*i.e.*, distances between the Cu1, Cu2, Cu3, $O1_{peroxo}$, and $O2_{peroxo}$ atoms); (iii) the QM/MM optimization is repeated but now only with one restrained distance between two oxygen atoms in peroxo species ($d_{O\text{-}O}$). Starting from this structure (denoted as $\mathbf{S_i}$), two nearby structures ($\mathbf{S_{i-1}}$, $\mathbf{S_{i+1}}$) with restrained distances $d_{O\text{-}O}(\mathbf{S_{i-1}}) = d_{O\text{-}O}(\mathbf{S_i})$-0.05 Å and $d_{O\text{-}O}(\mathbf{S_{i+1}}) = d_{O\text{-}O}(\mathbf{S_i})$+0.05 Å, respectively, are optimized. If $E(\mathbf{S_{i-1}}) < E(\mathbf{S_i}) > E(\mathbf{S_{i+1}})$, then the structure $\mathbf{S_i}$ is considered as the TS in the QM/MM model, otherwise $\mathbf{S_{i-2}}$, $\mathbf{S_{i+2}}$ structures are taken into account. This procedure is repeated until the maximum on the PES is found ($\mathbf{S_{TS}}$) which corresponds to approximate QM/MM TS. The nature of the TS can be simply verified by optimization of the $\mathbf{S_{TS-1}}$ and $\mathbf{S_{TS+1}}$ structures converging to the reactants and products, respectively.

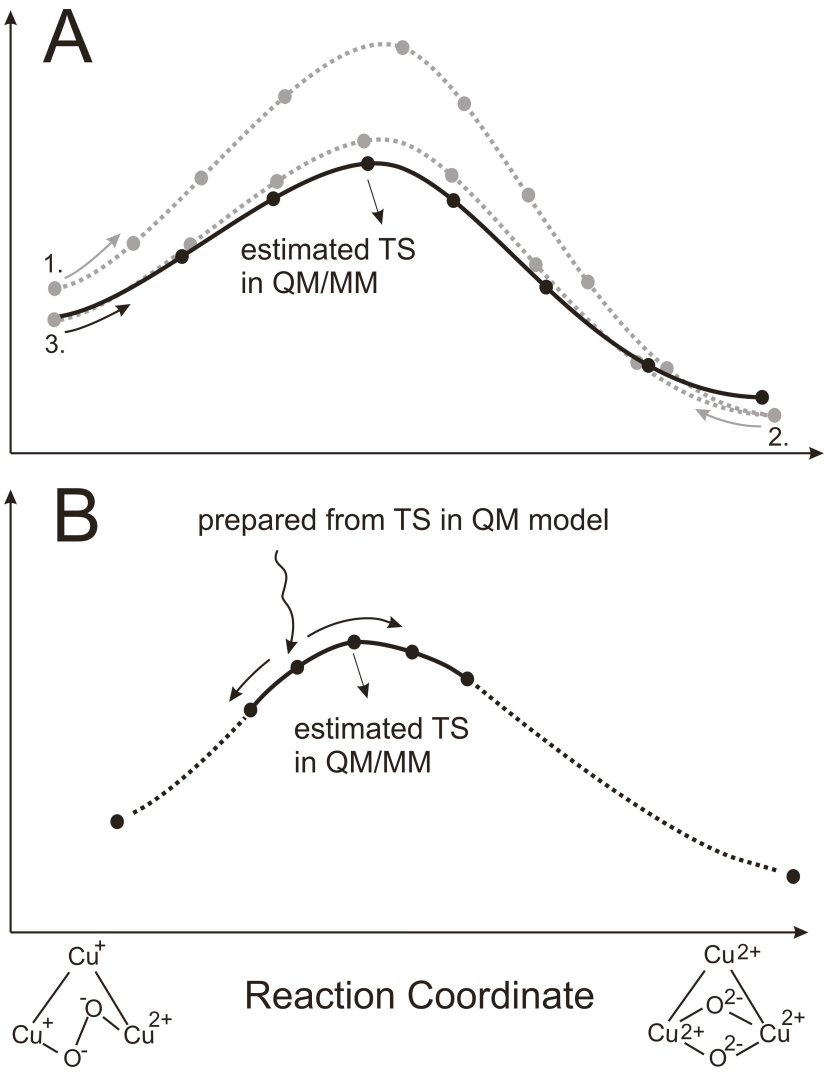

Figure 4: Two different strategies to estimate the transition state (TS) in QM/MM model. The computational details are described in the text.

3 Results and Discussion

QM/MM reaction coordinates of O_2 cleavage in MCOs.

Six alternatives for the reaction pathways corresponding to the cleavage of the O–O bond in MCOs were studied. These included two protonation states of the peroxide moiety (O_2^{2-} and O_2H^-), and three possibilities concerning the Cu-T2 ligand (H_2O, OH^-, no ligand). Two representative pathways (with H_2O as Cu-T2 ligand) are depicted in Figure 5. We postulated[16] the existence of transient species – activated peroxy intermediate (**NI'**) – structural analogue of the **PI** with the Cu-T1 site oxidized and one more electron available in TNC (denoted **PI + e** by Yoon and Solomon[24]). It is, therefore, the last chemically distinct species prior to the cleavage

of O–O bond. The product of the reaction is the experimentally characterized **NI** in the μ_3-oxo binding mode of O^{2-} which was shown to be energetically most favorable than three μ_2-hydroxo bridges on all three Cu-Cu sides.[16]

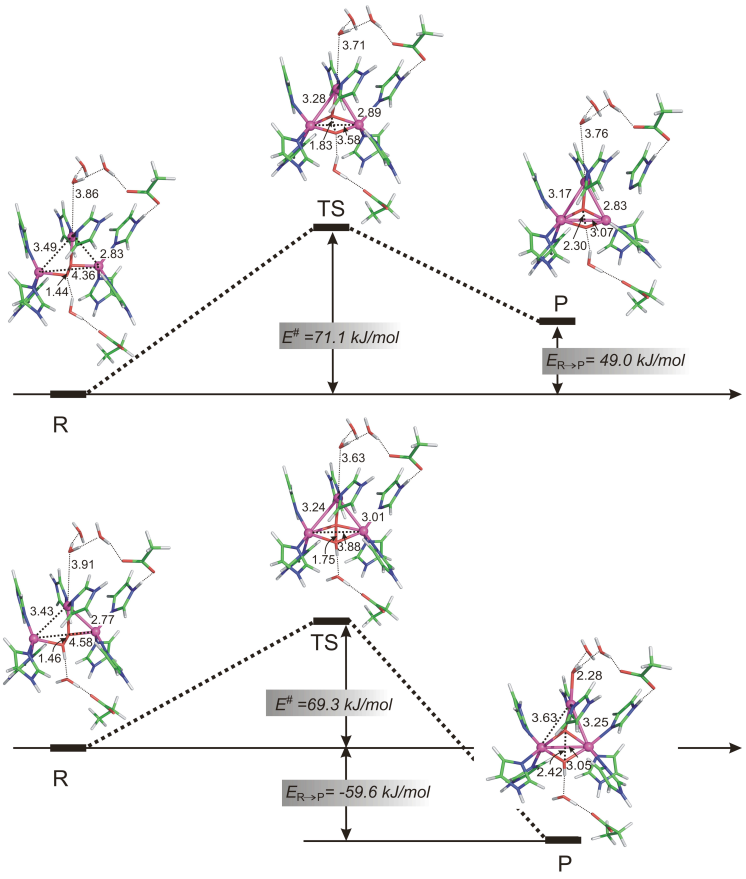

Figure 5: The QM/MM optimized structures for two of the six studied **NI'** → **NI** reaction pathways (O_2 cleavage in MCOs), starting from the two states with the Cu-T2 ligand H_2O and with either O_2^{2-} or HO_2^- in the centre of the TNC cluster. All distances are in Å.

The calculated activation barriers and reaction energies are listed in Tables 1 and 2, respectively. These include the total QM/MM energies and the non-electrostatic (steric) MM contribution to the total QM/MM energy, E_{MM} term ($E_{MM} = E_{MM123} - E_{MM1}$). The third term in Eqn. 1, the quantum mechanical energy with the point charges included in the one-electron Hamiltonian (the QM energy, including the electrostatic stabilization of quantum system by the surrounding protein), E_{QM1}, is not included since it can be trivially calculated as the difference of the $E_{QM/MM}$ and E_{MM} terms. For comparison, the single-point energies at the optimized QM/MM geometries of S1 for the same reaction step *in vacuo* are also given ($E_{QM/vac}$).

Table 1: Calculated QM/MM activation barriers ($\Delta E^{\neq}_{QM/MM}$) for the O–O bond cleavage in CueO together with the individual contributions: $\Delta E_{QM/vac}{}^{\neq}$, the *in vacuo* QM energy change of system 1; $\Delta E_{MM}{}^{\neq}$, the difference in MM energy changes of systems 1 and 3; $\Delta G^{\neq}_{QM/MM}$, the estimate of the activation free energy (entropic and thermal enthalpic contributions are taken from Table 3). The geometries were optimized at the QM(RI-PBE/DZP)/MM level whereas the single-point energies were recalculated at the QM(B3LYP/def2-TZVP)/MM level. For each system two rows of numbers are introduced, first row corresponds to results obtained from scan B, the second row to results obtained from scan A (both procedures are depicted in the Figure 4 and described in Section 2.4.). The total energies of the reactants (**NI'**) are set to zero. All values are in kJ.mol^{-1}.

Cu-T2 ligand	Central ligand	Charge of system 1	$\Delta E^{\neq}_{QM/vac}$	$\Delta E^{\neq}_{MM}$	$\Delta E^{\neq}_{QM/MM}$	$\Delta G^{\neq}_{QM/MM}$
H_2O	O_2H^-	+1	88.7	-3.2	69.3	66.8
			89.8	-2.5	69.5	67.0
OH^-	O_2H^-	0	–	–	–	–
			38.9	6.2	59.0	–
–	O_2H^-	+1	81.7	5.0	80.6	79.4
			82.1	-0.9	75.1	73.7
H_2O	O_2^{2-}	0	64.1	0.8	71.1	72.4
			53.4	1.0	70.0	71.3
OH^-	O_2^{2-}	-1	84.6	-13.3	68.1	67.0
			61.6	5.4	64.6	63.5
–	O_2^{2-}	0	66.8	-6.4	73.9	78.4
			67.5	1.6	73.0	77.5

It can be seen that all activation barriers are quite close to each other with $\Delta E^{\neq}$ = 59–81 kJ.mol^{-1} ($\Delta G^{\neq}$ = 63–79 kJ.mol^{-1}). These values are also close to the value of ~55 kJ.mol^{-1} that can be deduced from rate of k_{cat} > 350 s^{-1} (ref. 12) using Eyring equation. **Xxx Strictly, <59 kJ.** The Cu-T2 ligand in the **PI** is normally assumed to be an H_2O and the central peroxide in the O_2^{2-} protonation state.[11,16] Our preliminary calculations indicate that reducing **PI** to **NI'** would significantly shift the pK_a value of the central O_2^{2-} ligand from 0 to 11, and therefore, we expect that reduction of T23 site is immediately followed by protonation of O_2^{2-}. However, it is not fully clear whether the proton is taken from the Cu-T2 ligand or from the surrounding protein, e.g. from two carboxylate moieties that are part of our QM/MM system (*cf.* Figure 5). **Xxx Our estimated pKa of T2-water is 13 with PB and -6 with GB for NI'-HO2.**

The fact that protonation of O_2^{2-} may precede the O–O bond cleavage can be supported by comparison of two alternative pathways: **NI'**{OH^-,-,O_2H^-} → **NI**{OH^-,OH^-,O^{2-}:C} and **NI'**{H_2O,-,O_2^{2-}} → **NI**{H_2O,O^{2-},O^{2-}:C}. [Here and below, we introduce notation used in Ref. 16 where in **X**{A,B,L:P} A denotes Cu-T2 ligand; B, if present, the Cu-T3/CuT3' μ_2-bridging moiety; L the ligand in the centre of $(Cu)_3$ triangle; and P specifies its position (e.g., P = C denotes μ_3- coordination)]. Although the calculated energy difference between corresponding **NI'** states is very similar (within ±4 kJ.mol^{-1} for the QM/MM energies of the two species), the QM/MM activation barrier associated with the **NI'**{OH^-,-,O_2H^-} state is estimated to be $\Delta E^{\neq}$ ≈ 59 kJ.mol^{-1} compared to $\Delta E^{\neq}$ ≈ 71–72 kJ.mol^{-1} obtained for the **NI'**{H_2O,-,O_2^{2-}} state. **Xxx Why is scan B and DG missing for the former in Table 1?** Note that other alternatives included in the Table 1 also have higher barriers than **NI'**{OH^-,-,O_2H^-}. Moreover, on the product side, **NI**{OH^-,OH^-,O^{2-}:C} is 122 kJ.mol^{-1} more stable than **NI**{H_2O,O^{2-},O^{2-}:C}, in agreement with the assumed protonation state of the observed NI.[13,16] **xxx Again, our calculations give indifferent results for the deprotonation of T2-water in NI(H2O;OH;O), but**

they agree with you that NI(Oh,OH,O) is more stable than NI(H2O,O,O) As can be seen in Table 2 the calculated thermodynamic driving force for the **NI'**{HO^-,-,O_2H^-} → **NI**{OH^-,OH^-,O^{2-}:C} reaction is $\Delta G_{QM/MM}$ = 88 kJ.mol^{-1} ($\Delta E_{QM/MM}$ = 101 kJ.mol^{-1}) which lowers the exothermicity of the reaction by ~120 kJ.mol^{-1} compared to the QM cluster model.[24] This fact demonstrates the role of the enzyme in buffering the excess reaction energy of this highly exothermic reaction.

In conclusion, the QM/MM calculations suggest that the most plausible reaction pathway is the **NI'**{OH^-,-,O_2H^-} → **NI**{OH^-,OH^-,O^{2-}:C}. This reasoning is similar to the discussed high- and low-pK_a pathways reported by Yoon and Solomon in their QM study[24] of the O_2 cleavage. However, the QM/MM calculations predict a more important role of the protonation of the peroxide than the QM calculations do (the $\Delta\Delta E_{QM}^{\neq}$ between the pathways with and without a proton was only 0.5 kcal.mol^{-1}).[24]

In Table 2 we collected the calculated reaction energies for the six studied pathways. In contrast to the activation barriers (Table 1), differences between reaction energies are much larger, ranging from ΔE = –107 kJ.mol^{-1} for **NI'**{OH^-,-,O_2^{2-}} → **NI**{OH^-,O^{2-},O^{2-}:C} pathway to ΔE = +49 kJ.mol^{-1} for the **NI'**{H_2O,-,O_2^{2-}} → **NI**{H_2O,O^{2-},O^{2-}:C} pathway. An analysis of effects that cause these significant differences in reaction energies is not straightforward. However, some differences can be inferred from geometrical parameters of the active site depicted in Figure 5. While geometries of the reactants (**NI'**) and the corresponding TSs are similar, product geometries (**NI**) vary more significantly (e.g., the H_2O···Cu-T2 distance is 2.32 Å in **NI**{H_2O,OH^-,O^{2-}:C} *vs.* 3.76 Å in **NI**{H_2O,O^{2-},O^{2-}:C}). **Xxx This difference in the Cu-O bond cannot explain the 156 kJ/mol difference (it can be easily tested by shortening the Cu-O distance).** One might hypothesize that proper sampling of the conformational space by molecular dynamics or Monte Carlo statistics could partially reduce such large differences in reaction energies but it is necessary to add in defense of the static QM/MM approach that many different initial QM/MM structures tend to converge to similar conformations. If there are larger conformational differences (deduced from $\Delta(E_{MM123}–E_{MM1})$ term) they are usually compensated by other contributions (see Table S1). **Xxx I think we should use the QM/MM-PBSA and QTCP methods on your structures to obtain more stable estimates, including dynamic effects outside the QM system. If you send us your structures, we can do that, or I may instruct Martin how to do the calculations.**

Table 2: Calculated QM/MM reaction energies ($\Delta E_{QM/MM}$) for the O–O bond cleavage in MCOs together with the individual contributions: $\Delta E_{QM/vac}$, the *in vacuo* QM energy change of system 1; ΔE_{MM}, the difference in MM energy changes of systems 1 and 3; $\Delta G_{QM/MM}$, the estimate of the activation free energy (entropic and thermal enthalpic contributions are taken from Table 3). The geometries were optimized at the QM(RI-PBE/DZP)/MM level whereas the single-point energies were recalculated at the QM(B3LYP/def2-TZVP)/MM level. The total energies of the reactants (**NI'**) are set to zero. All values are in kJ.mol^{-1}.

Cu-T2 ligand	Central ligand	Charge of S1	$\Delta E_{QM/vac}$	ΔE_{MM}	$\Delta E_{QM/MM}$	$\Delta G_{QM/MM}$
H_2O	O_2H^-	1	-93.9	-9.8	-58.8	-43.4
OH^-	O_2H^-	0	-135.1	-1.9	-101.4	-88.1
–	O_2H^-	1	-42.1	23.8	-26.4	-8.0
H_2O	O_2^{2-}	0	40.4	-0.6	49.0	51.9
OH^-	O_2^{2-}	-1	-178.2	0.1	-107.2	-78.9
–	O_2^{2-}	0	24.5	3.1	36.6	52.4

Cluster model: Entropy contributions

In this study we address the reactivity of different states of the **NI'** intermediate and attempt to elucidate the most favoured pathway. Since the differences in the TS barriers between the pathways are relatively small one has to consider also the entropic and thermal enthalpic (further denoted ΔH_{therm})) contributions that need to be added to electronic energies to obtain estimated Gibbs free energies. The entropic term is very sensitive to low-frequency modes in the vibrational partition function and therefore to proper description of geometry of the structure (i.e., local minima and first-order saddle points on the potential-energy surfaces must be exactly defined within the approximation), which is problematic in case of our ComQum QM/MM scheme. Therefore, we used small *in vacuo* QM models of the active site (see methodological section) to estimate entropy and thermal enthalpy changes ($\Delta H - T\Delta S$) in the protein. We believe such a simplification is reasonable since large conformational changes are not assumed to occur in the protein except for the active site. This assumption can be supported by the fact that the steric MM energy changes are mostly within 10 kJ.mol^{-1}.

In Table 3, the $\Delta H_{\text{therm}}^{\neq} - T\Delta S^{\neq}$ and $\Delta H_{\text{therm}} - T\Delta S$ terms contributing to the activation barrier and reaction energies are shown. In both cases, their values are similar for all six alternative pathways. The effect of $\Delta H_{\text{therm}}^{\neq} - T\Delta S^{\neq}$ on the barriers is relatively small (from –2.5 kJ.mol^{-1} to 4.5 kJ.mol^{-1}) **xxx Delete the decimal. Why is the reaction in read omitted?** implying that the essential chemistry of O_2 cleavage in MCOs is reasonably well predicted already by the values of $\Delta E^{\neq}$. Comparing the entropic and enthalpic contributions (the most dominant vibrational terms are shown in Table S2), it appears that both of them are of similar magnitude (entropy term is slightly greater) but of the opposite sign. In spite of our extensive efforts, we did not succeed in optimizing the TS of the *in vacuo* QM model corresponding to the **NI'**{OH^-,-,O_2H^-} → **NI**{OH^-,OH^-,O^{2-}:C} pathway. However, rather small scattering of the calculated $\Delta H_{\text{therm}}^{\neq} - T\Delta S^{\neq}$ values obtained for the other systems (pathways) gives us an ample amount of confidence that activation entropic and enthalpic changes in this favoured pathway will be similar and we estimate that $\Delta G_{\text{QM/MM}}^{\neq} \approx$ 60–65 kJ.mol^{-1}.

The $\Delta H_{\text{therm}} - T\Delta S$ contributions to the reaction energies range from 13 kJ.mol^{-1} to 23 kJ.mol^{-1}. The results of the analysis displayed in Table S2 show that $-T\Delta S$ (i.e., $-T\Delta S_{\text{vib}}$) is the dominant term. It implies that the **NI** structures are entropically disfavoured compared to the corresponding **NI'** intermediates (by approximately 15 kJ.mol^{-1}).[46] **Xxx Here I stopped.**

It is also interesting to compare the energetics of the reaction obtained for QM clusters and QM/MM models of enzyme. The barriers are significantly lower in the majority of QM models with respect to QM/MM calculations (e.g., 33 kJ.mol^{-1} *vs* 73–79 kJ.mol^{-1} for **NI'**{-, -,O_2H^-}→ **NI**{-,OH^-, O^{2-}:C} pathway). As for reaction energies, the general trend is in accordance with the previous discussion: the NI products are significantly more stable in reference to NI' reactants using the gas-phase QM models compared to the corresponding QM/MM models. It is in agreement with above mentioned Hammond's postulate and our previous observations:[23] the overstabilization of the product state leads to an artificial lowering of the activation barrier (as depicted in Figure 2). Finally, it can be also mentioned that the solvation effect represented by implicit solvent model (COSMO)[47] with permittivity 4 or 20 is rather small (*cf.* Table S3). Therefore, we

may conclude that explicit inclusion of the whole protein and solvent water molecules by means of QM/MM scheme leads to significantly improved energies.

Table 3. Calculated gas-phase activation energies ($\Delta E_{QM}^{\neq}$) and reaction energies (ΔE_{QM}), along with estimation of thermal enthalpic and entropic contributions ($\Delta H_{therm}^{(\neq)} - T\Delta S^{(\neq)}$) to reaction or activation free energies ($\Delta G^{(\neq)}_{QM}$) are displayed. The results were obtained at the B3LYP/def2-TZVP//RI-PBE/DZP level whereas the thermochemical analysis ($\Delta H_{therm} - T\Delta S$ term) was carried out at the RI-PBE/DZP level. The total energies of the reactants (NI') are set to zero. All values are in kJ.mol^{-1}.

Cu-T2 ligand	Central ligand	Charge of QM cluster	$\Delta E^{\neq}_{QM}$	$\Delta H^{\neq} - T\Delta S^{\neq}$	$\Delta G^{\neq}_{QM}$	ΔE_{QM}	$\Delta H - T\Delta S$	ΔG_{QM}
H_2O	O_2H^-	3	50.4	-2.5	47.9	-113.6	15.4	-98.3
OH^-	O_2H^-	2	-	-	-	-141.7	13.3	-128.4
–	O_2H^-	3	33.4	-1.2	32.2	-118.5	18.4	-100.1
H_2O	O_2^{2-}	2	39.7	1.3	41.0	-41.7	12.9	-28.8
OH^-	O_2^{2-}	1	79.4	5.6	85.0	-95.9	22.8	-73.1
–	O_2^{2-}	2	39.8	4.5	44.3	-42.7	15.8	-26.9

The multireference CASSCF/CASPT2 calculations

The accuracy of CASSCF/CASPT2 method applied to reactions with large change in molecular electronic structure, *i.e.* processes in which bonds are broken or/and formed, strongly depends on a selected active space. In an optimal case, a full valence active space should be sufficient for accurate prediction of energetics of reaction. Unfortunately, such an active space is prohibitively too large to be computationally manageable for even a minimalistic model of bioinorganic complexes (e.g., an active site of metalloenzymes). In this study we attempted to estimate activation barriers and reaction energies of gas-phase models of the studied **NI'** → **NI** reaction at the CASSCF/CASPT2 level. As mentioned in the Computational Details section we considered the distribution of fifteen electrons in nine MOs: six in O_2 moiety (i.e., σ, 2π, 2π*, σ*) along with $3d_{Cu}$ orbitals as an adequate (minimal) active space. We may present, at the moment, the first value obtained for the **NI'**{H_2O,-,O_2^{2-}} → **NI**{H_2O,O^{2-},O^{2-}:C} pathway though obtained using smaller (13-in-9) active space. The CASSCF/CASPT2 activation barrier has been calculate to be $\Delta E_{QM}^{\neq}$ = 45.4 kJ.mol^{-1} which is in very good agreement with the value of 39.7 kJ.mol^{-1} reported in Table 3 for this pathway. A full set of CASSCF(15-in-9)/CASPT2 values is, however, necessary, to make conclusive observations.

4 Conclusions

In the current study, we investigated in detail the reductive cleavage of the O-O bond by MCOs using QM/MM method and compared the calculated data with their counterparts obtained using cluster model (QM). We have shown that the most favorable pathway involves protonation of the peroxide moiety within the TNC by the proton most likely originating from the Cu-T2 water molecule and its subsequent cleavage. The calculated activation barriers are in a good agreement with the experimental barrier and it is shown that only by including the whole protein (via QM/MM approach), we were able to buffer the large exothermicity of the reaction. We have also demonstrated that the entropic and thermal enthalpy corrections are

not playing crucial role in distinguishing between the studied pathways though they correct the calculated barriers by 5-20 kJ.mol^{-1} on an absolute scale. Finally, an attempt is presented to carry out the CASSCF/CASPT2 calculations for the studied reaction in order to validate the (in principle single-determinantal) DFT method used and the preliminary results show a reasonable agreement between the two methods.

Acknowledgment

We gratefully acknowledge the financial support from the Ministry of Education, Youth and Sports of the Czech Republic (Research Projects Z40550506 and LC512) and from the Swedish Research Council (UR).

References

[a] *Gilead Sciences & IOCB Research Center, Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nám. 2, 166 10 Praha 6, Czech Republic. Fax: +420 220 578; Tel:+420 220 183 263; E-mail: rulisek@uochb.cas.cz*
[b] *Department of Theoretical Chemistry, Lund University, Chemical Center, P.O. Box 124, S–221 00 Lund, Sweden. E-mail: Ulf.Ryde@teokem.lu.se*

† Electronic Supplementary Information (ESI) available: the protein coordinates and the point charges on all the atoms in the MM region (in the PDB format) and the equilibrium geometries of the quantum region for all of the studied structures. See DOI: 10.1039/b000000x/

1 H. M. Senn and W. Thiel, *Angew. Chem.-Int. Edit.*, 2009, **48**, 1198-1229.
2 S. C. L. Kamerlin, M. Haranczyk and A. Warshel, *J. Phys. Chem. B*, 2009, **113**, 1253-1272.
3 P. E. M. Siegbahn and F. Himo, *J. Biol. Inorg. Chem.*, 2009, **14**, 643-651.
4 H. Hu and W. T. Yang, *Annu. Rev. Phys. Chem.*, 2008, **59**, 573-601.
5 U. Ryde, *Curr. Opin. Chem. Biol.*, 2003, **7**, 136-142.
6 B. Kirchner, F. Wennmohs, S. F. Ye and F. Neese, *Curr. Opin. Chem. Biol.* 2007, **11**, 134-141.
7 F. Neese, *Coord. Chem. Rev.*, 2009, **253**, 526-563.
8 A. Warshel, *Computer Modeling of Chemical Reactions in Enzymes and Solutions.* John Wiley&Sons, Inc., New York, 1997.
9 A. Messerschmidt, In *Multicopper oxidases*; Messerschmidt, A., Ed.; World Scientific: Singapore; River Edge, NJ, 1997; pp. 23-80.
10 E. I. Solomon, U. M. Sundaram and T. E. Machonkin, *Chem. Rev.*, 1996, **96**, 2563-2605.
11 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**, 3202-3215.
12 A. E. Palmer, S.-K. Lee and E. I. Solomon, *J. Am. Chem. Soc.*, 2001, **123**, 6591-6599.
13 S.-K. Lee, S. D. George, W. E. Antholine, B. Hedman, K. O. Hodgson and E. I. Solomon, *J. Am. Chem. Soc.*, 2002, **124**, 6180-6193.
14 K. Piontek, M. Antorini and T. Choinowski, *J. Biol. Chem.*, 2002, **277**, 37663-37669.
15 S. A. Roberts, G. F. Wildner, G. Grass, A. Weichsel, A. Ambrus, C. Rensing and W. R. Montfort, *J. Biol. Chem.*, 2003, **278**, 31958-31963.
16 L. Rulíšek, E. I. Solomon and U. Ryde, *Inorg. Chem.*, 2005, **44**, 5612-5628.
17 F. Xu, *Biochemistry*, 1996, **35**, 7608-7614.
18 G. J. Davies and V. Ducros, In *Handbook of Metalloproteins*; Messerschmidt, A., Huber, R., Wieghardt, K., Poulos, T, Eds.; Wiley, New York, 2001, pp. 1359-1368.
19 J. Chalupský, F. Neese, E. I. Solomon, U. Ryde and L. Rulíšek, *Inorg. Chem.*, 2006, **45**, 11051-11059.
20 U. Ryde, Y.-W. Hsiao, L. Rulíšek and E. I. Solomon, *J. Am. Chem. Soc.*, 2007, **129**, 726-727.
21 E. I. Solomon, A. J. Augustine and J. Yoon, *Dalton Trans.*, 2008, 3921-3932.
22 J. Yoon, L. M. Mirica, T. D. P. Stack and E. I. Solomon, *J. Am. Chem. Soc.*, 2005, **127**, 13680-13693.
23 M. Lepšík, M. Srnec, D. Hnyk, B. Grüner, J. Plešek, Z. Havlas and L. Rulíšek, *Collect. Czech Chem. Commun.*, 2009, **74**, 1-27.

24 J. Yoon and E. I. Solomon, *J. Am. Chem. Soc.*, 2007, **129**, 13127-13136.
25 P. W. Atkins, *Physical Chemistry*, Oxford University Press, Oxford, U. K., 1978.
26 G. S. Hammond, *J. Am. Chem. Soc.*, 1955, **77**, 334-338.
27 U. Ryde, *J. Comput.-Aided Mol. Design*, 1996, **10,** 153-164.
28 U. Ryde and M. H. M. Olsson, *Int. J. Quantum Chem.*, 2001, **81**, 335-347.
29 O. Treutler and R. Ahlrichs, *J. Chem. Phys.*, 1995, **102**, 346-354.
30 D. A. Case, T. A. Darden, T. E. Cheatham III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, K. M. Merz, B. Wang, D. A. Pearlman, M. Crowley, S. Brozell, V. Tsui, H. Gohlke, J. Mongan, V. Hornak, G. Cui, P. Beroza, C. Schafmeister, J. W. Caldwell, W. S. Ross and P. A. Kolman, AMBER 8, University of California, San Francisco, 2004.
31 W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J. Am. Chem. Soc.*, , 1995, **117**, 5179-5197.
32 S. A. Roberts, A. Weichsel, G. Grass, K. Thakali, J. T. Hazzard, G. Tollin, C. Rensing and W. R. Montfort, *Proc. Natl. Acad. Sci. U.S.A.*, 2002, **99**, 2766-2771.
33 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865-3868.
34 K. Eichkorn, O. Treutler, H. Öhm, M. Häser and R. Ahlrichs, *Chem. Phys. Lett.*, 1995, **240**, 283-290.
35 K. Eichkorn, F. Weigen, O. Treutler and R. Ahlrichs, *Theor. Chim. Acta*, 1997, **97**, 119-124.
36 W. J. Hehre, L. Radom, P. v. R. Schleyer and J. A. Pople, Ab initio molecular orbital theory; Wiley-Interscience: New York, 1986.
37 A. Schäfer, C. Huber and R. Ahlrichs, *J. Chem. Phys.*, 1994, **100**, 5829-5835.
38 F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3297-3305.
39 A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098-3100; C. T. Lee, W. T. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785-789; A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648-5652.
40 F. Jensen, Introduction to Computational Chemistry; John Wiley & Sons, 1999.
41 B. O. Roos and P. R. Taylor, *Chem. Phys.*, 1980, **48**, 157-173.
42 K. Andersson, P.-Å. Malmqvist and B. O. Roos, *J. Chem. Phys.*, 1992, **96**, 1218-1226.
43 G. Karlström, R. Lindh, P.-Å. Malmqvist, B. O. Roos, U. Ryde, V. Veryazov, P.-O. Widmark, M. Cossi, B. Schimmelphennig, P. Neogrady and L. Seijo, *Comput. Mater. Sci.*, 2003, **28**, 222-239.
44 B. O. Roos and K. Andersson, *Chem. Phys. Lett.*, 1995, **245**, 215-243; N. Forsberg and P.-Å. Malmqvist, *Chem. Phys. Lett.*, 1997, **274**, 196-204.
45 K. Pierloot, B. Dumez, P.-O. Widmark and B. O. Roos, *Theor. Chim. Acta*, 1995, **90**, 87-114.
46
47 A. Klamt and G. Schuurmann, *J. Chem. Soc., Perkin Trans. 2* 1993, 799-805; A. Schäfer, A. Klamt, D. Sattel, J. C. W. Lohrenz and F. Eckert, *Phys. Chem. Chem. Phys.*, 2000, **2**, 2187-1293.