AKADEMIE VĚD
ČESKÉ REPUBLIKY

Teze disertační práce

ke získání vědeckého titulu „doktor věd“ (DSc.)

ve skupině věd **chemických**

THEORETICAL CALCULATIONS OF PHYSICO-CHEMICAL AND SPECTROSCOPIC PROPERTIES OF BIOINORGANIC SYSTEMS

Komise pro obhajoby doktorských disertací v oboru fyzikální chemie

Jméno uchazeče: Lubomír Rulíšek

Pracoviště uchazeče: Ústav organické chemie a biochemie AV ČR

Místo a datum: Praha, 15. června 2012

Summary

Among the various essential elements in biocatalysis, metalloproteins play a specific role by catalysing reactions that would not occur under physiological conditions. The presence of metal ions is thus crucial for the oxidation/reduction processes, electron transfer, spin-forbidden reactions and 'difficult reactions', such as N_2, O_2, C–H bond breaking. These processes are intimately involved in the fundamental elements of life, e.g. respiration and photosynthesis. Enormous efforts, both experimental and theoretical, have been exerted to understand the structure and function of metalloproteins (and bioinorganic systems in general). While experiments (e.g., X-ray crystallography, various spectroscopic techniques, electrochemistry) are crucial in initial phases of our understanding to a particular system, theoretical calculations complement these data by providing a unique one-to-one structure-energy mapping. As such, they play indispensable role in elucidating the reaction mechanisms of bioinorganic systems, provide an insight into the phenomena of metal-ion selectivity, and shed light on the physicochemical principles (laws) governing the behaviour of these systems.

In the presented thesis, our efforts in the area of *theoretical bioinorganic chemistry* are compiled. It includes theoretical studies of reaction mechanisms of selected metalloproteins (often containing polynuclear active sites with open-shell metal ions): trinuclear copper site found in multi-copper oxidases (MCOs), dinuclear non-heme iron Δ^9-desaturase (Δ^9D), dinuclear zinc glutamate carboxypeptidase II (GCPII), and mononuclear manganese super-oxide dismutase (MnSOD). All of them represent highly challenging systems for quantum chemical methodology either due to the complexity of the electronic structure of the active site (MCOs, MnSOD, Δ^9D) or a conformational complexity of the substrate in presumably 'electronically simpler' GCPII. It is shown that only by a tight interplay between theory and experiment (X-ray, spectroscopic, kinetic, and thermodynamic data) it is possible to formulate consensus reaction mechanisms for these complicated systems.

Also, closely related to bioinorganic chemistry is our ongoing research concerning the metal-ion selectivity with the long-term vision of *in silico* design of smaller artificial metalloenzymes (in the literature sometimes nicknamed 'artzymes'). Currently, it is at the stage of the design of shorter peptide sequences selectively binding selected metal ions. Condition *sine qua non* for successful accomplishment of this project is the ability to quantitatively calculate the complexion (free) energy changes associated with the binding of metal ions in biomolecules and theoretically predict the metal-ion selectivity.

Finally, methodological issues originating from our calculations of thermodynamic, kinetic, and spectroscopic properties of bioinorganic systems gave rise to modest contributions to quantum chemical and combined quantum and molecular mechanical (QM/MM) theory.

Introduction

In general, chemical and biochemical sciences have made enormous progress by exploring both experimental and theoretical methods to understand the fundamental aspects of (bio)chemical reactions. The same is true for a more specific (and specialised) field of bioinorganic chemistry which mostly focuses on metalloproteins as one of the essential catalytic elements in nature. In fact, it is estimated that between one-quarter and one-third of all enzymes contain metal ions in their active site.[1] The set of reactions catalysed by metalloproteins is very broad and metal ions play a crucial role in catalysing processes associated with large reaction barriers, such as activation of C-H,[2] O=O,[3] or N≡N[4] bonds, and spin-forbidden reactions.[5] They have key importance in the functionality of countless oxidoreductases, which make up approximately 25% of the known enzymes on Earth,[6] and in long-range electron transfer processes.[7]

It is, therefore, highly desirable to study coordination of metal ions in biomolecules and the function of metalloproteins at all levels of the system description. And indeed, the ground-breaking experimental studies have often been supported by theoretical work elucidating the underlying physicochemical principles. It has been a natural consequence of the recent progress in quantum chemical methodology and the increase of computer power that have led to a situation where the methods of theoretical chemistry complement and at the same time rival their experimental counterparts.[8] In this way, we have begun to understand the fundamental physicochemical features of metalloenzyme catalysis.[9]

The major advantage of using quantum chemistry in studies of biomolecules is complementarity of information that can be obtained from the calculations. The energy profiles, changes in electron densities and molecular structures during catalysis (including the transition state structures along a reaction pathway) are important system descriptors which are difficult, if not impossible, to obtain by other techniques. Thus, it is nowadays possible to conveniently model systems containing up to ~500 atoms using standard techniques, such as the popular density functional theory (DFT) methods.[10,11] This size of the system is not accessible for the high-level wave function methods; however, recent advances[12,] have considerably extended our possibilities in their usage for bioinorganic systems of a realistic size (approaching 100 atoms).

The quality and reliability of a particular theoretical study are assessed by comparing the calculated data with their experimental counterparts. In the realm of metalloproteins, the experimental methods are mostly represented by X-ray crystallography, electronic paramagnetic resonance (EPR), including the electron-nuclear double resonance technique (ENDOR), nuclear magnetic resonance (NMR) and various standard spectroscopic techniques, such as absorption spectra, circular dichroism (CD), including magnetic (MCD) and vibrational (VCD) spectra,

resonance Raman spectroscopy, Mössbauer spectroscopy and X-ray absorption techniques (EXAFS, XANES).[13] However, the task of calculating the experimental data with spectroscopic accuracy is far from being trivial.[11] Especially for systems with complicated electronic structure (coupled polynuclear metal sites, spin multiplets) new methods need to be developed, applied and tested. Once it is demonstrated that these methods are quantitatively predictive, it truly opens new horizons in the elucidation of molecular structures of reactive intermediates.

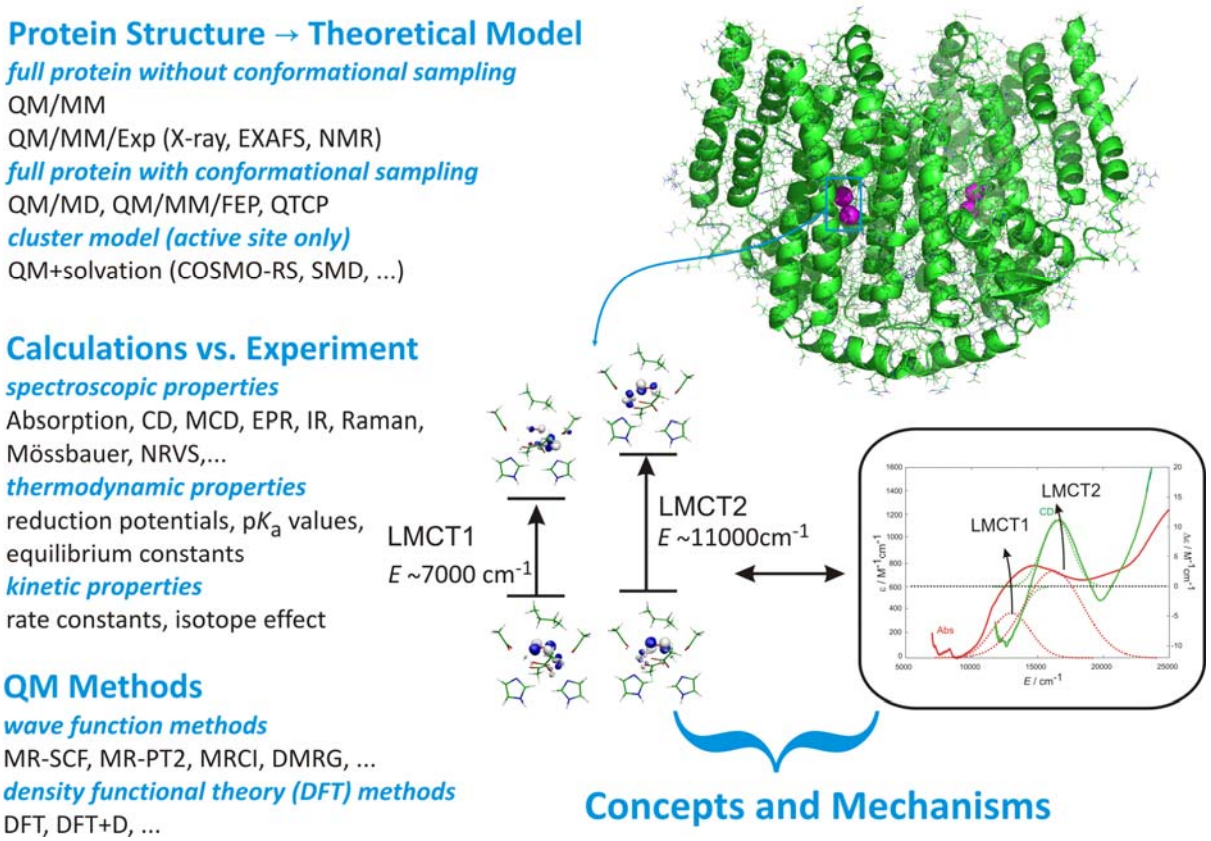

Figure 1: The arsenal of tools and methods available in theoretical bioinorganic chemistry and an outline of the general computational strategy in studying the structure, properties and mechanisms of bioinorganic systems. As an illustrative example a dinuclear non-heme diiron enzyme Δ^9 desaturase is used (in its 1,2,-μ-peroxodiferric state).[14]

Reaction Mechanisms of Metalloproteins: Correlating Theory and Experiment

Multi-Copper Oxidases. The multicopper oxidases (MCOs) form a class of enzymes playing a variety of physiological roles in organisms, having in common the presence of two copper sites that accommodate four copper ions in total. Utilising these specific features, the MCOs couple

four one-electron oxidations of a substrate with a four-electron reduction of molecular oxygen to water[15,16]:

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

This reaction takes place at a trinuclear copper cluster (TNC), whereas the substrate is oxidised at a type 1 copper site (Cu-T1), which is ~13 Å away from the TNC. The two sites are connected via a bifurcated $(Cu_{TNC}–His)_2–Cys–Cu_{T1}$ protein chain (where His–Cys–His are three consecutive residues in the protein, of which the two histidines are ligands of two copper ions in the TNC and the Cys is coordinated to the Cu-T1 ion as cysteinate, Figure 2). This structural arrangement is assumed to provide an efficient electron-transfer (ET) pathway between the two sites, transferring the four electrons that are needed for the dioxygen reduction from the Cu-T1 site to the TNC. Therefore, the oxidation states of four copper ions span the full spectrum between $(Cu^{2+})_4$ and $(Cu^{+})_4$.

There are at least three reasons making the MCOs very interesting (and tempting) for a computational bioinorganic chemist. The first is the inherent (*in vacuo*) instability of many plausible intermediates in the catalytic cycle of MCOs (i.e. the geometrical arrangements of the TNC) as a consequence of the four distinct redox states of the three copper ions (i.e. $(Cu^{2+})_3$, $(Cu^{+})(Cu^{2+})_2$, $(Cu^{2+})(Cu^{+})_2$ and $(Cu^{+})_3$) coupled with the various accessible protonation states of the copper ligands originating from water or dioxygen (e.g. oxo, hydroxo and peroxo species).[17] Therefore, the TNC site (the Cu ions with their first-sphere ligands) possesses a high positive charge (+3 or +4 according to the suggested reaction mechanism), which is partly compensated for by two carboxylate residues in the second coordination sphere of the TNC, which are conserved throughout the MCO family (Figure 2).[18] The second reason is the complicated electronic structure of the TNC. In the putative structure of the so-called native intermediate, **NI**, the TNC contains three unpaired spins at the vertices of a triangle, all coupled via an O_2^{2-} molecule in the centre. This leads to so-called spin-frustration, which means that the exchange coupling between the three pairs of Cu^{2+} ions cannot be satisfied. In this oxidation state, there are two doublet states and one quartet state close in energy (within a few hundred cm^{-1}).[19] The third reason is the unique opportunity to couple the theoretical calculations directly to experimental data and provide their theoretical interpretation.

To tackle this challenging system, our studies involved most of the "theoretical bioinorganic chemistry tools" depicted in Figure 1, specifically quantum mechanics (QM; density-functional theory and multi-reference self-consistent field – e.g. CASSCF/CASPT2) calculations,[17,20,21,22] combined QM and molecular mechanics (QM/MM) modelling, ranging from standard QM/MM optimisations[17] to the combination of QM/MM optimisation with raw EXAFS

data[23] (QM/MM-EXAFS) to address the O_2 reactivity in the TNC and QM/MM free-energy perturbations[24] (QTCP) to accurately address phenomena such as the Cu-T1 → TNC electron transfer (reorganization energies).[25] Our achievements and all the history of MCO theoretical studies are well-documented in great details in our recent comprehensive review.[26]

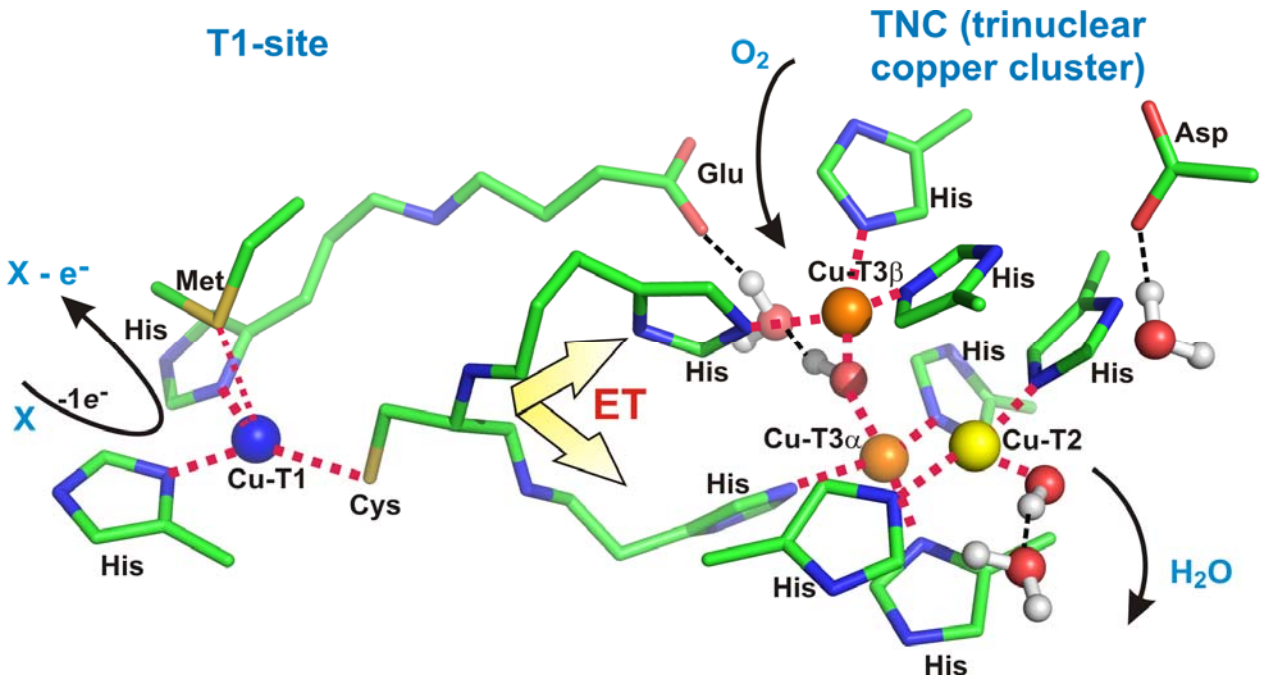

Figure 2: The general architecture of the trinuclear copper cluster site (the site of the four-electron $O_2 \rightarrow H_2O$ reduction) and of the Cu-T1 site (the site of one-electron oxidations of organic substrates or metal ions).

Glutamate Carboxypeptidase II. Glutamate Carboxypeptidase II (GCPII) is an enzyme, which plays an important role in the regulation of levels of one of the neurotransmitters – *N*-Ac-Asp-Glu (NAAG) – in synapses. In collaboration with experimentalists, we proposed a detailed and consistent picture of the reaction mechanism of this highly interesting enzyme at the atomic level (Figure 3).[27] The best estimate of the reaction barrier was calculated to be $\Delta G^{\#} \approx 22(\pm 5)$ kcal.mol^{-1}, which is in a reasonable agreement with the experimentally observed reaction rate constant ($k_{cat} \approx 1$ s^{-1}). Besides, using QM/MM methodology, we have predicted the three-dimensional structure of GCPII homologue – GCPIII - and pinpointed the similarities and differences in the active site architecture between these two enzymes.[28] Some of these were later confirmed by X-ray crystal structure of the GCPIII.[29] Furthermore, the preliminary computational results enabled us to draw mechanistic explanations of the observed k_{cat}/K_M variations upon the point mutations in GCPII.[30] Finally, we provided a rationale for the interaction of novel substrate-based inhibitors of GCPII with enhanced lipophilicity.[31]

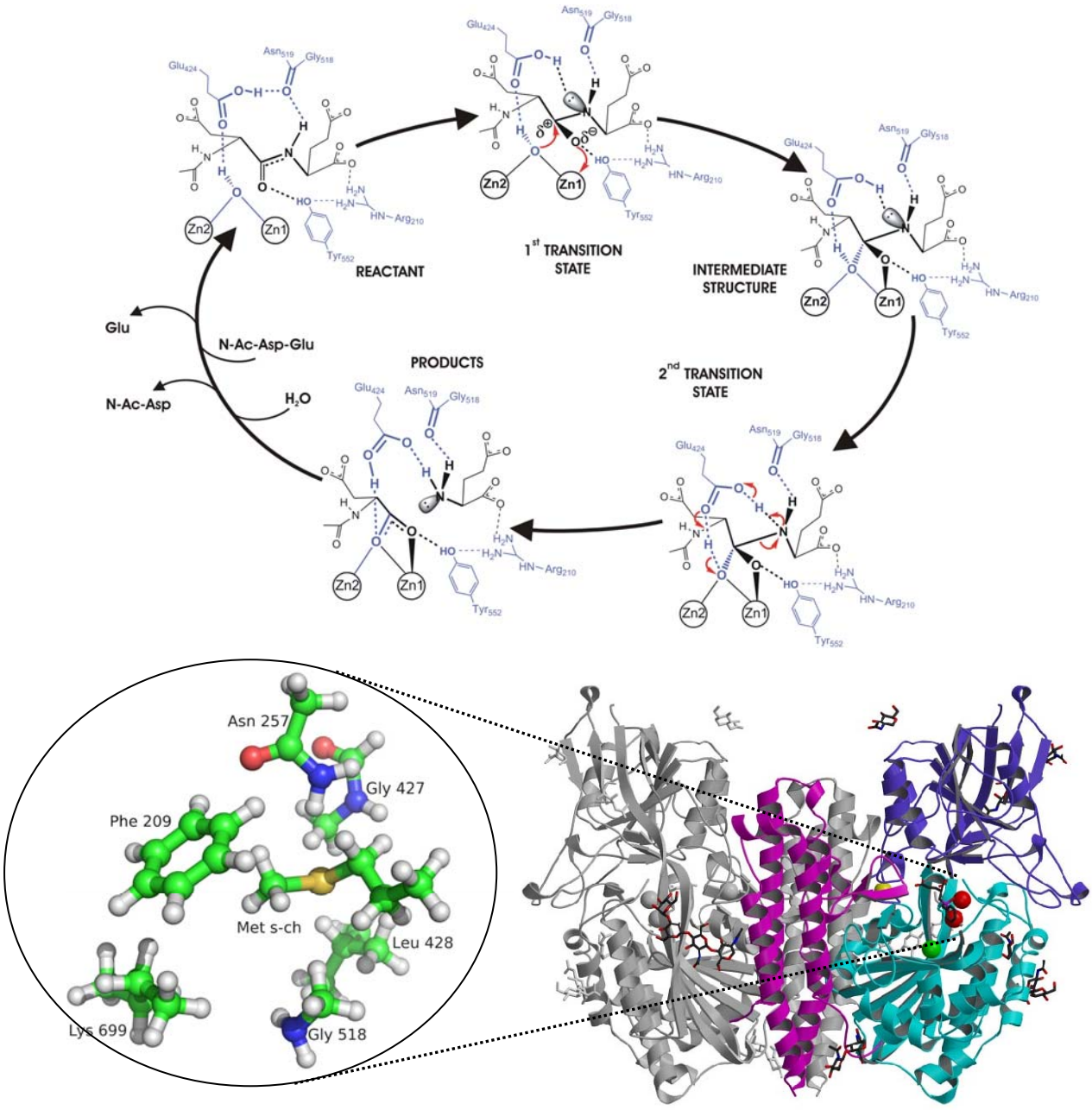

Figure 3: *(upper part)* The important steps/structures throughout the GCPII reaction pathway as obtained by QM/MM calculations (the energy profile is not depicted); *(lower part)* the qualitative insight into the GCPII inhibition and the GCPII structure illustrating the mapping of interactions in the GCPII active site, and prediction of the GCPIII structure.

Manganese Superoxide Dismutase. Superoxide dismutases (SODs) are enzymes that convert two molecules of the poisonous superoxide radical into molecular oxygen and hydrogen peroxide:

$$M^{III}SOD_{OH} + O_2^{\bullet -} + H^+ \rightarrow M^{II}SOD_{H2O} + O_2 \quad (4)$$

$$M^{II}SOD_{H2O} + O_2^{\bullet -} + H^+ \rightarrow M^{III}SOD_{OH} + H_2O_2 \quad (5).$$

According to the metal ion present as the cofactor in the active site, they can be divided into three classes: MnSODs, FeSODs, and CuZnSODs.

In manganese SODs (MnSODs), the redox active manganese ion cycles between the Mn^{2+} and Mn^{3+} oxidation states and accomplishes its enzymatic action in two half-cycles (corresponding to the oxidation and reduction of $O_2^{\bullet-}$, cf. Eqs. 4 and 5). However, the exact reaction mechanism for this disproportionation reaction was unknown. As can be seen from Eqs. 4 and 5, of an utmost importance for understanding the reaction mechanism is the protonation state of the bound water/hydroxide molecule at the MnSOD active site. Therefore, we considered it as the suitable case for the application of the combined QM/MM/X-ray study (so-called quantum refinement, first developed in the group of Prof. Ulf Ryde). Using the method for the MnSOD, we have clearly shown that the solvent ligand is OH^- in the oxidized (Mn^{3+}) state, whereas it is H_2O in the reduced (Mn^{2+}) state.[32] Moreover, we have also shown that the putative oxidised structure is to a large extent reduced during data collection, so that it contains a mixture of the Mn^{2+} and Mn^{3+} structure. The details of the study can be found in Ref. 32.

Conclusive computational evidence concerning the MnSOD reaction cycle was obtained using the QM/MM approach (i.e., modelling the catalyzed reaction in the protein environment).[33] These QM(DFT)/MM results were further complemented by CASSCF/CASPT2/MM single-point energy calculations for the most plausible models to account properly for the multireference character of the various spin multiplets. The results indicate that the oxidation of $O_2^{\bullet-}$ to O_2 most likely occurs by an associative mechanism following a two-state (quartet–octet) reaction profile. The barrier height is estimated to be less than 25 kJ.mol^{-1}. On the other hand, the conversion of $O_2^{\bullet-}$ to H_2O_2 is likely to take place by a second-sphere mechanism, *i.e.* without direct coordination of the superoxide radical to the manganese centre. The reaction pathway involves the conical intersection of two quintet states (Figure 4), giving rise to an activation barrier of ~60 kJ.mol^{-1}. The activation barriers along the proposed reaction pathways are in very good agreement with the experimentally observed reaction rates of SODs ($k_{cat} \approx 10^4$–10^5 s^{-1}).

Finally, it can be mentioned that in the subsequent study[34] we used the quantum and molecular mechanical thermodynamic cycle perturbation method[24] (QTCP) of Prof. Ulf Ryde to address two key issues related to the function of SODs in general: reduction potential of the metal ion in the active site (Mn, Fe) which is fine-tuned (by the protein) to the specific value, different from its value in solution and the pK_a values of the metal-bound water-derived species. Though the calculated reduction potentials were quite dependent on many details of the computational procedure and system setup, the final calculated values of 0.3 V were in a very good agreement

with the experiment.[34] Thus, we have shown that it is in principle possible (though extremely difficult) to address these complex phenomena computationally.

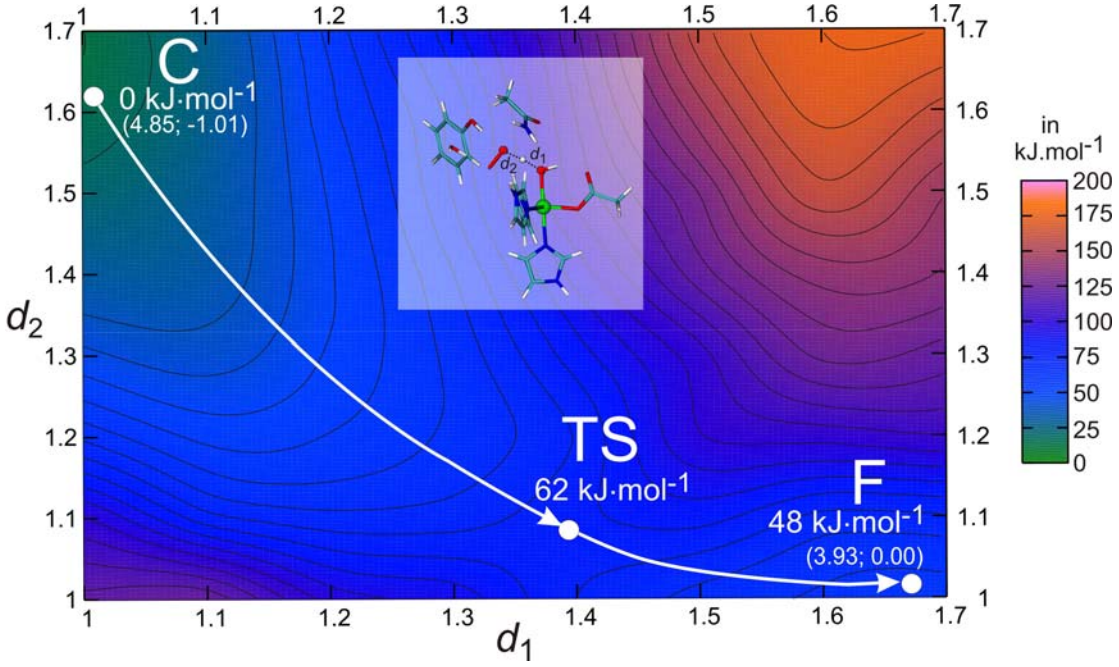

Figure 4. The two-dimensional QM(B3LYP/6-31G*)/MM potential-energy surface of the proton-coupled electron transfer from Mn–OH_2 to O_2 in the SS2 mechanism. The interatomic distances are in Å. The numbers (x; y) in the graph depict the spin densities at the manganese and O_2 centres.

Δ^9 Desaturase. In the recent study[14] we published our first computational data (complemented by a small amount of experimental data from Solomon's laboratory –MCD spectra) on the most challenging system studied insofar in our laboratory - Δ^9 desaturase (Δ^9D). Δ^9D is a non-heme diiron enzyme and therefore, the constitution of its active site by itself represents a formidable task for contemporary computational methods. The ultimate goal is to decipher the reaction mechanisms for this type of enzymes and one may expect that along the reaction pathway many reactive intermediates containing iron-oxo units are to be encountered, as well as radicals ensuing from the homolytic cleavage of the C-H bonds, and many other 'non-trivial' species. While the application of DFT methodology might be probably the only practical solution to the problem, its accuracy has not yet been fully grasped. It is difficult, if not impossible, to benchmark DFT methods even on smaller systems (as is normally the standard procedure), as the CASPT2 (RASPT2) or MRCI calculations would necessitate more active orbitals to be included in the calculations than is within the reach of current computational power. Still, the correlation of

several calculated (DFT) spectroscopic parameters with their experimental counterparts (absorption, CD, vibrational, and Mössbauer) for multiple structural models of the experimentally defined peroxodiferric intermediate (**P**) enabled us to lock in the molecular structures present in the initial steps of the catalytic cycle. We suggested that protonation of the peroxide moiety, possibly preceded by water binding in the Fe_A coordination sphere, could be responsible for the conversion of the **P** intermediate in Δ^9D into a form capable of hydrogen abstraction.

Theoretical Calculations of Metal Ion Complexation in Peptides and Proteins

Complexation (Free) Energies of Metal Ions in Biomolecules and Metal Ion Selectivity. In general, the level of metal-ion selectivity enhancement is dependent on the ability to predict and control the specificity of the site (i.e. an optimal combination of metal-binding residues and the preferred coordination number).[35] The predictions are frequently based on the observed frequency of the occupation of the coordination sites in metalloproteins[36,37] or use semiquantitative concepts, like the HSAB theory of Parr and Pearson.[38] Only recently has the design become assisted by advances and accumulated experience in quantum (theoretical) bioinorganic chemistry. Using high-level quantum chemical methods coupled with efficient algorithms for solving Poisson-Boltzmann equation to account for the effect of environment, many important properties of the first- and second-shell ligand binding have been calculated,[39,40] e.g. the effect of the carboxylate-binding mode on metal binding/selectivity and function in proteins.[41]

We have contributed to the development of novel and modified metal-ion binding sites through several computational studies, in which the affinities of all the metal-binding amino acid side chains towards selected transition metal ions (Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, Cd^{2+}, and Hg^{2+}) were evaluated,[42,43] the cooperative effect accompanying the simultaneous binding of more than one AA side chain quantitatively characterised,[44] and finally the '*M^{2+}-specific*' combinations of amino acid side chains proposed for each of these metal ions.[45] The predicted *M^{2+}-specific* combinations can be considered, on the basis of the DFT/B3LYP calculations of the complexation energies of the metal ions in $[M(Y_1)\ldots(Y_n)]^{c+}$ complexes (with Y_i being an amino acid side chain), to be the best candidates for the highly selective metal-binding site in proteins or peptides. The specificity can be further increased by taking into account the preferred coordination geometry for each metal ion, based on an analysis of the Protein Data Bank (PDB) and Cambridge Structural Database (CSD) structures.[37] Consequently, the *Co^{2+}-specific* and *Ni^{2+}-specific* combinations are defined as $[Co(Y_1)\ldots(Y_6)]^{c+}$ (or $[Ni(Y_1)\ldots(Y_6)]^{c+}$) complexes (where Y_is are 'optimum' amino acid side chains) in an octahedral coordination geometry. For the Cd^{2+} and Zn^{2+} ions, the optimal

$[Zn(Y_1)\ldots(Y_4)]^{c+}$ or $[Cd(Y_1)\ldots(Y_4)]^{c+}$ sites are tetrahedral, for the Cu^{2+} square-planar, and for the Hg^{2+} (i.e., $[Hg(Y_1)(Y_2)]^{c+}$) linear. These combinations were later used as templates in our design of novel metal-binding peptide sequences (*vide infra*).

In the above systematic studies[42-45] describing our extensive efforts to address the issues of metal-ion selectivity at the end of 1990s and beginning of 2000s, we demonstrated that some basic trends in the selectivity of metal-ion binding in proteins can be already captured by the interaction (complexation) energies calculated *in vacuo*. Nevertheless, the ultimate goal is the quantitative assessment of the experimental stability constants (which is a measure of the change in the free energy of the system associated with the transfer of the metal ion from the hydrated reference state to the particular complex). The standard thermodynamic cycle used for the calculation of the overall ΔG consists of the free energy (or free enthalpy) differences in the gas-phase complexation of the ions and ligands and the difference in the solvation free energies of the complexed and free ligands. These processes are usually associated with large (free) energy changes of several hundreds of kcal.mol^{-1} (the gas-phase association of the ion…neutral or ion…ion species and their solvation/desolvation energies) that almost cancel each other out to yield the final ΔG values of several kcal.mol^{-1}. This fact yields quantitatively accurate calculations of the stability constant as extremely difficult and challenging tasks for computational chemistry.

To this aim, we have recently attempted to critically evaluate the performance of the *ab initio* and DFT electronic structure methods available and recent solvation models in calculations of the energetics associated with metal ion complexation[46]. On the example of five model complexes ($[M^{II}(CH_3S)(H_2O)]^+$, $[M^{II}(H_2O)_2(H_2S)(NH_3)]^{2+}$, $[M^{II}(CH_3S)(NH_3)(H_2O)(CH_3COO)]$, $[M^{II}(H_2O)_3(SH)(CH_3COO)(Im)]$, $[M^{II}(H_2S)(H_2O)(CH_3COO)(PhOH)(Im)]^+$ in typical coordination geometries) and four metal ions (Fe^{2+}, Cu^{2+}, Zn^{2+}, and Cd^{2+}; representing open- and closed-shell and the first- and second-row transition metal elements), we provided reference values for the gas-phase complexation energies, as presumably obtained using the CCSD(T)/aug-cc-pVTZ method, and compared it with cheaper methods, such as DFT and RI-MP2, that can be used for large-scale calculations. We also discussed two possible definitions of interaction energies underlying the theoretically predicted metal-ion selectivity and the effect of geometry optimization on these values. Finally, popular solvation models, such as COSMO-RS and SMD, were used to demonstrate whether quantum chemical calculations can provide the overall free enthalpy (ΔG) changes in the range of the expected experimental values for the model complexes or match the experimental stability constants in the case of three complexes for which the experimental data exist. The computational data also highlighted several intricacies in the theoretical predictions of the experimental stability constants: the covalent character of some metal-ligand bonds (e.g.

Cu(II)-thiolate) causing larger errors in the gas-phase complexation energies, inaccuracies in the treatment of solvation of the charged species and difficulties in the definition of the reference state for Jahn-Teller unstable systems (e.g., $[Cu(H_2O)_6]^{2+}$). While the agreement between the experimental (as derived from the stability constants) and calculated values is often within 5 kcal.mol^{-1}, in more complicated cases, it may exceed 15 kcal.mol^{-1}. Therefore, extreme caution must be exercised in assessing the subtle issues of metal ion selectivity quantitatively.

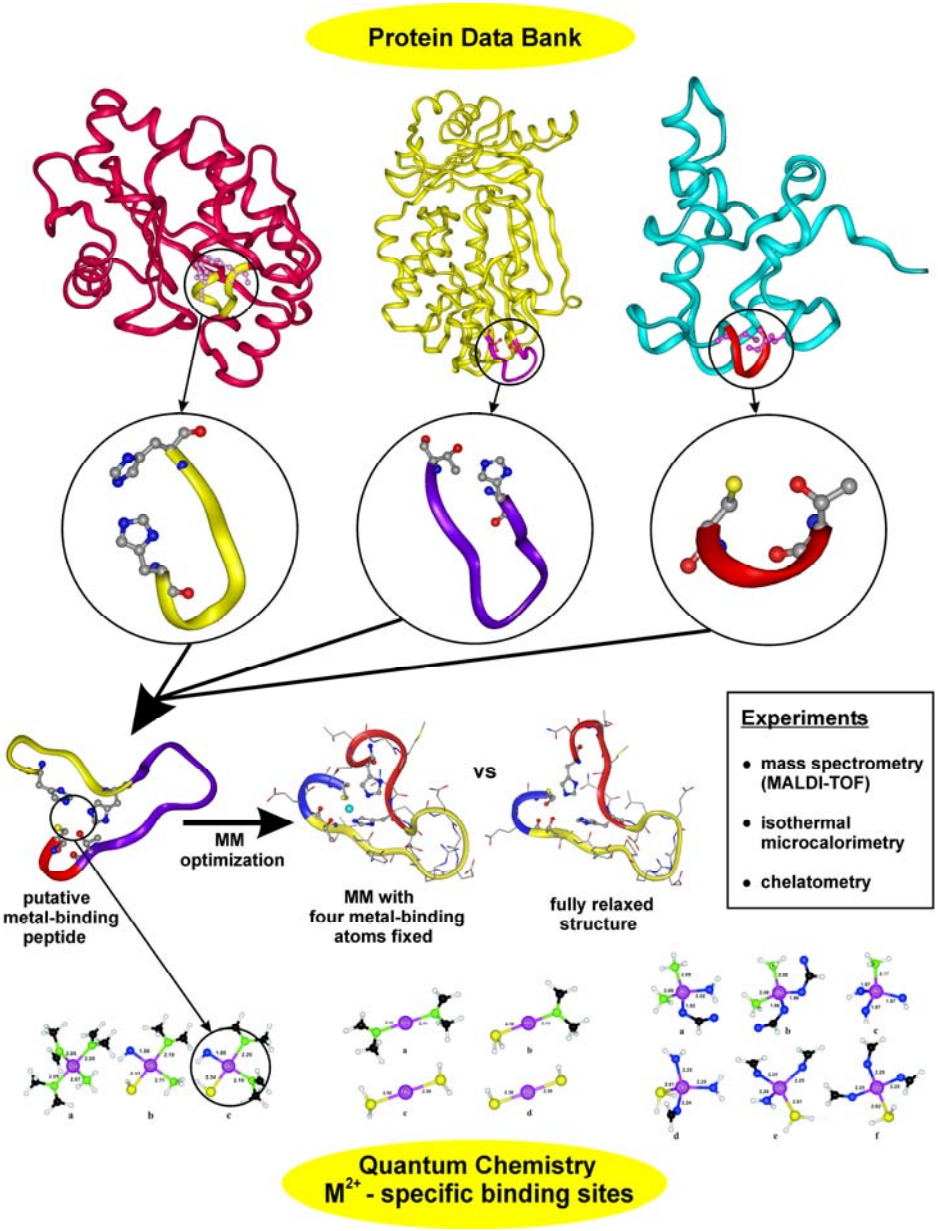

Figure 5: The overview of the novel computational strategy for the design of specific metal-binding peptide sequences.

Molecular Design of Novel Metal-Binding Peptide Sequences. Concomitantly with developing the accurate computational strategies for addressing metal-ion selectivity, the major landmark in a long term project of developing a novel strategy for designing peptides with specific metal-ion (M^{2+}) chelation sites was accomplished in 2008. Ion-specific combinations of amino acid side chains that provide functional groups at the vertices of the desired coordination polyhedron are computationally predicted and linked into a single polypeptide chain (using fragments from the protein structures deposited in the Protein Data Bank).

The application of this procedure for the most *M^{2+}-specific* combinations of amino acid side chains[45] resulted in several peptide sequences (with length of 4-20 amino acids) with the potential for specific binding of six studied metal ions (Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, Cd^{2+}, and Hg^{2+}).[47] The gas-phase association constants of the studied metal ions to these *de novo* designed peptides were experimentally determined by MALDI ionization mass spectrometry using 3,4,5-trihydroxyacetophenone as a matrix, whereas the thermodynamics of the metal ion coordination in condensed phase was measured by isothermal titration calorimetry (ITC) and chelatometry methods. The data show that some of the computationally predicted peptides are potential *M^{2+}-specific* metal ion chelators. The whole strategy for the novel *in silico* design of metal-specific peptides is depicted in Figure 5 and more details about our efforts to tackle the problem of metal-ion selectivity and the design of metal-binding sequences can be also found in Refs. 37, 42-47.

Methodological Issues in Bioinorganic Chemistry

Spin-Orbit Coupling Effects in Octahedral Ru(II/III) and Os(II/III) Complexes. In general, relativistic effects are known to play an important role in bioinorganic chemistry. However, they are not often taken into considerations explicitly and this may even hide some qualitative principles governing the behaviour of metal ions. Thus, in one of our studies the notable discrepancy between the calculated and experimental reduction potentials for some octahedral complexes of osmium(II) was observed, while at the same time a satisfactory (to within 0.05-0.1 V) agreement for their ruthenium(II) counterparts was obtained.[48] It led us to a systematic investigation of its origin. To this end, reduction potentials of several $Ru^{2+/3+}$ and $Os^{2+/3+}$ octahedral complexes - $[M(H_2O)_6]^{2+/3+}$, $[MCl_6]^{4-/3-}$, $[M(NH_3)_6]^{2+/3+}$, $[M(en)_3]^{2+/3+}$ $[M(bipy)_3]^{2+/3+}$, and $[M(CN)_6]^{4-/3-}$ - were calculated using the CASSCF/CASPT2/CASSI and MRCI methods including spin-orbit coupling (SOC) by means of first-order quasi-degenerate perturbation theory (QDPT).[49] It was shown that the effect of SOC accounts for a systematic shift of approximately -70 mV in the reduction potentials of the studied ruthenium (II/III) complexes and an

approximately -300 mV shift for the osmium(II/III) complexes. SOC splits the sixfold degenerate $^2T_{2g}$ ground electronic state (in ideal octahedral symmetry) of the M^{3+} ions into the $E_{(5/2)g}$ Kramers' doublet and $G_{(3/2)g}$ quartet, which were calculated to split by 1354-1573 cm^{-1} in the Ru^{3+} and by 4155-5061 cm^{-1} in the Os^{3+} complexes. It was demonstrated that this splitting represents the main contribution to the stabilization of the M^{3+} ground state with respect to the closed shell $^1A_{1g}$ ground state in M^{2+} systems.

These results led us to the conclusion that, especially for $Os^{2+/3+}$ complexes, an inclusion of SOC is necessary to avoid systematic errors of ~300 mV in the calculated reduction potentials and only then the accurate values can be obtained (once the solvation energy is calculated with an acceptable accuracy, which is also a non-trivial task). The effect is schematically depicted in Figure 6 and more details can be found in Ref. 49.

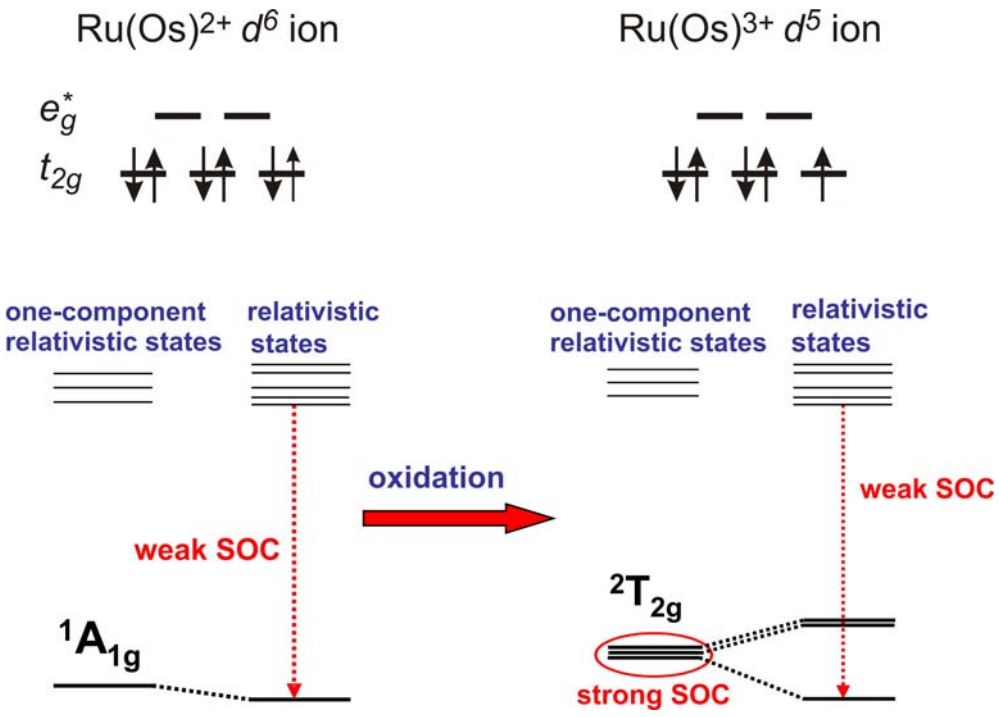

Figure 6. Schematic representation of the SOC effects on the values of the reduction potentials of $Ru^{2+/3+}$ and $Os^{2+/3+}$ octahedral complexes.

Curvature Correction for Microiterative Optimizations with QM/MM Electronic Embedding. Another contribution to the method development has been represented by the improvement of the QM/MM minimization procedure. We suggested a computationally cheap second-order correction that employs an estimated Hessian from the Davidon-Fletcher-Powell

method to tackle the problems caused by the too small curvature of the QM/MM potential energy surface. Test calculations on four metalloenzymatic systems (~100 QM atoms, ~2000 relaxed MM atoms, ~20000 atoms in total) show that our approach efficiently restores the convergence in cases where gradient correction leads to oscillations[50] and thus improves the robustness of the QM/MM minimization.

Conclusions

This thesis consists of our contributions to the field of *theoretical bioinorganic chemistry*. Its author witnessed the development of the field over the past one and half decade. Indeed, starting (in mid 1990s) with the first attempts to model the catalytic action of metalloproteins (using small active site models of ~30-50 atoms, modelled in the gas-phase environment) we matured to a 'thoughtful and educated' usage of today's sophisticated methods and approaches, such as QM/MM schemes, advanced solvation models, QM/MM-FEP-like sampling, various modern DFT functionals, multi-reference wave function methods and many others (*cf.*, Fig. 1). These enabled us to take into account the full system in the context of the condensed-phase environment and thus slowly converge to a situation where theoretical calculations rival the experiments and sometimes even provide guidance to the experimental collaborators. It has not been, though, the rivalry, but rather a friendship and complementarity between the two that led to the overall progress in bioinorganic chemistry. On the theoretical side, this implies that the modern computational approaches tend to change from the 'interpretative domain' to a 'predictive domain'. This transfer or rather extension has to be, though, fuelled by a never-ending conceptual and methodological development. We may foresee that in near future, further progress in computational methodology will result in the situation when theoretical bioinorganic chemistry will be an integral part of the description and characterization of any system studied by bioinorganic chemists.

References

(1) Waldron, K. J.; Robinson, N. J. *Nat. Rev. Microbiol.* **2009**, *7*, 25.

(2) Balcells, D.; Clot, E.; Eisenstein, O. *Chem. Rev.* **2010**, *110*, 749.

(3) Solomon, E. I.; Ginsbach, J. W.; Heppner, D. E.; Kieber-Emmons, M. T.; Kjaergaard, C. H.; Smeets, P. J.; Tian, L.; Woertink, J. S. *Faraday Discuss.* **2011**, *148*, 11.

(4) Seefeldt, L. C.; Hoffman, B. M.; Dean, D. R. *Annu. Rev. Biochem.* **2009**, *78*, 701.

(5) Harvey, J. N. *Phys. Chem. Chem. Phys.* **2007**, *9*, 331.

(6) Dutton, P. L.; Moser, C. C. *Faraday Discuss.* **2011**, *148*, 443.

(7) Page, C. C.; Moser, C. C.; Chen, X.; Dutton, P. L. *Nature* **1999**, *402*, 47.

(8) Senn, H. M.; Thiel, W. *Angew. Chem.-Int. Edit.* **2009**, 48, 1198.

(9) Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H. B.; Olsson, M. H. M. *Chem. Rev.* **2006**, *106*, 3210.

(10) Kohn W. *Rev. Modern Phys.* **1999**, *71*, 1253.

(11) Rokob, T. A.; Srnec, M.; Rulíšek, L. *Dalton Trans.* **2012**, *41*, 5754.

(12) Neese, F.; Liakos, D. G.; Ye, S. *J. Biol. Inorg. Chem.* **2011**, *16*, 821.

(13) Lippard, S. J.; Berg, J. M. in *Principles of Bioinorganic Chemistry*, University Science Books (1994).

(14) Srnec, M.; Rokob, T. A.; Schwartz, J. K.; Kwak, Y.; Rulíšek, L.; Solomon, E. I. *Inorg. Chem.* **2012**, *51,* 2806.

(15) Messerschmidt, A. In *Multicopper oxidases*; Messerschmidt, A., Ed.; World Scientific: Singapore; River Edge, NJ, (1997); p. 23.

(16) Solomon, E. I.; Sundaram, U. M.; Machonkin, T. E. *Chem. Rev.* **1996**, *96*, 2563.

(17) Rulíšek, L.; Solomon, E. I.; Ryde, U. *Inorg. Chem.* **2005**, *44*, 5612.

(18) Augustine, A. J.; Quintanar, L.; Stoj, C. S.; Kosman, D. J.; Solomon, E. I. *J. Am. Chem. Soc.* **2007**, *129*, 13118.

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

(20) Chalupský, J.; Neese, F; Solomon, E. I.; Ryde, U.; Rulíšek, L. *Inorg. Chem.* **2006**, *45*, 11051.

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

(22) Srnec, M.; Ryde, U.; Rulíšek, L. *Faraday Discuss.* **2011**, *148*, 41.

(23) Ryde, U.; Hsiao, Y.-W.; Rulíšek, L., Solomon, E. I. *J. Am. Chem. Soc.* **2007**, *129*, 726.

(24) Rod, T. H.; Ryde, U. *Phys. Rev. Lett.* **2005**, *94*, 138302.

(25) Hu, L.-H.; Farrokhnia, M.; Heimdal, J.; Shleev, S.; Rulíšek, L.; Ryde, U. *J. Phys. Chem. B* **2011**, *115*, 13111.

(26) Rulíšek, L.; Ryde, U. *Coord. Chem. Rev.* **2012**, *256*, yy-zz. In press DOI: 10.1016/j.ccr.2012.04.019.

(27) Klusák, V.; Bařinka, C.; Plechanovová, A.; Mlčochová, P.; Konvalinka, J.; Rulíšek, L.; Lubkowski, J. *Biochemistry* **2009**, *48*, 4126.

(28) Hlouchová, K; Bařinka, C.; Klusák, V.; Šácha, P; Mlčochová, P.; Majer, P.; Rulíšek, L.; Konvalinka, J. *J. Neurochem.* **2007**, *101*, 682.

(29) Hlouchová, K.; Bařinka, C.; Konvalinka, J.; Lubkowski, J. *FEBS Journal* **2009**, *276*, 4448.
(30) Mlčochová, P.; Plechanovová, A.; Bařinka, C.; Mahadevan, D.; Saldanha, J. W.; Rulíšek, L.; Konvalinka, J. *FEBS Journal* **2007**, *274,* 4731.
(31) Plechanovová, A.; Byun, Y; Alquicer, G.; Škultétyová, Ľ; Mlčochová, P.; Němcová, A.; Kim, H.-J.; Navrátil, M.; Mease, R.; Lubkowski, J.; Pomper, M.; Konvalinka, J.; Rulíšek, L.; Bařinka, C. *J. Med. Chem.* **2011**, *54*, 7535.
(32) Rulíšek, L.; Ryde, U. *J. Phys. Chem. B* **2006**, *110*, 11511.
(33) Srnec, M.; Aquilante, F.; Ryde, U.; Rulíšek, L. *J. Phys. Chem. B* **2009**, *113,* 6074.
(34) Heimdal, J.; Kaukonen, M.; Srnec, M.; Rulíšek, L.; Ryde, U. *ChemPhysChem* **2011**, *12*, 3337.
(35) Lee, K. H.; Matzapetakis, M.; Mitra, S.; Marsh, E. N. G.; Pecoraro, V. L. *J. Am. Chem. Soc.* **2004**, *126*, 9178.
(36) Glusker, J. P. *Adv. Protein Chem.* **1991**, *42*, 1.
(37) Rulíšek, L.; Vondrášek, J. *J. Inorg. Biochem.* **1998**, *71*, 115.
(38) Parr, R. G.; Pearson, R. G. *J. Am. Chem. Soc.* **1983**, *105*, 7512.
(39) Dudev, T.; Lim, C. *J. Am. Chem. Soc.* **2002**, *124*, 6759.
(40) Dudev, T.; Lim, C. *J. Phys. Chem. B* **2001**, *105*, 10709.
(41) Dudev, T.; Lim, C. *Acc. Chem. Res.* **2007**, *40*, 85.
(42) Rulíšek, L.; Havlas, Z. *J. Am. Chem. Soc.* **2000**, *122*, 10428.
(43) Rulíšek, L.; Havlas, Z. *J. Chem. Phys.* **2000**, *112*, 149.
(44) Rulíšek, L.; Havlas, Z. *J. Phys. Chem. A* **2002**, *106*, 3855.
(45) Rulíšek, L.; Havlas, Z. *J. Phys. Chem. B* **2003**, *107*, 2376.
(46) Gutten, O.; Beššeová, I.; Rulíšek, L. *J. Phys. Chem. A* **2011**, *115*, 11394.
(47) Kožíšek, M.; Svatoš, A.; Buděšínský, M.; Muck, A.; Bauer, M. C.; Kotrba, P.; Ruml, T.; Havlas, Z.; Linse, S.; Rulíšek, L. *Chem. Eur. J.* **2008**, *14*, 7836.
(48) Vrábel, M.; Hocek, M.; Havran, L.; Fojta, M.; Votruba, I.; Klepetářová, B.; Pohl, R.; Rulíšek, L.; Zendlová, L.; Hobza, P.; Shih, I.; Mabery, E.; Mackman, R. *Eur. J. Inorg. Chem.* **2007**, 1752.
(49) Srnec, M.; Chalupský, J.; Fojta, M.; Zendlová, L.; Havran, L.; Hocek, M.; Kývala, M.; Rulíšek, L. *J. Am. Chem. Soc.* **2008**, *130*, 10947.
(50) Rokob, T. A.; Rulíšek, L. *J. Comput. Chem.* **2012**, *33,* 1197.

List of Publications Comprising the Thesis (corresponding author(s) denoted by *):

1. Rulíšek, L.*; Vondrášek, J.: Coordination Geometries of Selected Transition Metal Ions (Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, Cd^{2+}, Hg^{2+}) in Metalloproteins. *J. Inorg. Biochem.* **1998**, *71*, 115-127.
2. Rulíšek, L.*; Havlas, Z.: Ab Initio Calculations of Monosubstituted (CH_3OH, CH_3SH, NH_3) Hydrated Ions of Zn^{2+} and Ni^{2+}. *J. Phys. Chem. A* **1999**, *103*, 1634-1639.
3. Rulíšek, L.*; Havlas, Z.: Ab Initio Calculations of $[CoY_{6-n}X_n]^{2+}$ Complexes. *J. Chem. Phys.* **2000**, *112*, 149-157.
4. Rulíšek, L.*; Havlas, Z.: Theoretical Studies of Metal Ion Selectivity. 1. DFT Calculations of Interaction Energies of Amino Acid Side Chains with Selected Transition Metal Ions (Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, Cd^{2+}, Hg^{2+}). *J. Am. Chem. Soc.* **2000**, *122*, 10428-10439.
5. Rulíšek, L.*; Havlas, Z.: Theoretical Studies of Metal Ion Selectivity. 2. DFT Calculations of the Affinity of Metal-Binding Sites in Metalloproteins for Selected Transition Metal Ions (Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, Cd^{2+}, Hg^{2+}). *J. Phys. Chem. A* **2002**, *106*, 3855-3866.
6. Rulíšek, L.*; Havlas, Z.: Theoretical Studies of Metal Ion Selectivity. 3. A Theoretical Design of the Most Specific Combinations of Functional Groups Representing Amino Acid Side Chains for the Selected Metal Ions (Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, Cd^{2+}, and Hg^{2+}). *J. Phys. Chem. B* **2003**, *107*, 2376-2385.
7. Rulíšek, L.; Solomon, E. I.; Ryde, U.*: A QM/MM Study of the O_2 Reductive Cleavage in the Catalytic Cycle of Multicopper Oxidases. *Inorg. Chem.* **2005**, *44*, 5612-5628.
8. Rulíšek, L.*; Ryde, U.: Structure of Reduced and Oxidised Manganese Superoxide Dismutase - A Combined Computational and Experimental Approach. *J. Phys. Chem. B* **2006**, *110*, 11511-11518.
9. Chalupský, J.; Neese, F; Solomon, E. I.; Ryde, U.; Rulíšek, L.*: Multireference *Ab Initio* Calculations on Reaction Intermediates of the Multicopper Oxidases. *Inorg. Chem.* **2006**, *45*, 11051-11059.
10. Srnec, M.; Chalupský, J.; Fojta, M.; Zendlová, L.; Havran, L.; Hocek, M.; Kývala, M.*; Rulíšek, L.*: Effect of Spin-Orbit Coupling on Reduction Potentials of Octahedral Ruthenium (II/III) and Osmium (II/III) Complexes. *J. Am. Chem. Soc.* **2008**, *130*, 10947-10954.
11. Kožíšek, M.; Svatoš, A.; Buděšínský, M.; Muck, A.; Bauer, M. C.; Kotrba, P.; Ruml, T.; Havlas, Z.; Linse, S.; Rulíšek, L.*: Molecular Design of Specific Metal-Binding Peptide Sequences from Protein Fragments. Theory and Experiment. *Chem. Eur. J.* **2008**, *14*, 7836-7846.
12. Srnec, M.; Aquilante, F.; Ryde, U.; Rulíšek, L.*: Reaction Mechanism of Manganese Superoxide Dismutase Studied by Combined Quantum and Molecular Mechanical Calculations and Multiconfigurational Methods. *J. Phys. Chem. B* **2009**, *113*, 6074-6086.
13. Klusák, V.; Bařinka, C.; Plechanovová, A.; Mlčochová, P.; Konvalinka, J.; Rulíšek, L.*; Lubkowski, J.*: Reaction Mechanism of Glutamate Carboxypeptidase II Revealed by Mutagenesis, X-ray Crystallography and Computational Methods. *Biochemistry* **2009**, *48*, 4126-4138.
14. Vancoillie, S.; Chalupský, J.; Ryde, U.; Solomon, E. I.; Pierloot, K., Neese, F.*, Rulíšek, L.*: Multireference *Ab Initio* Calculations of g tensors for Trinuclear Copper Clusters in Multicopper Oxidases. *J. Phys. Chem. B* **2010**, *114*, 7692-7702.
15. Srnec, M.; Ryde, U.; Rulíšek, L.*: Reductive Cleavage of the O–O Bond in Multicopper Oxidases: QM/MM and QM Study. *Faraday Discuss.* **2011**, *148*, 41-53.

16. Gutten, O.; Beššeová, I.; Rulíšek, L.*: Interaction of Metal Ions with Biomolecular Ligands: How Accurate Are Calculated Free Energies Associated with Metal Ion Complexation? *J. Phys. Chem. A* **2011**, *115*, 11394-11402.
17. Rokob, T. A.;* Rulíšek, L.: Curvature Correction for Microiterative Optimizations with QM/MM Electronic Embedding. *J. Comput. Chem.* **2012**, *33,* 1197–1206.
18. Srnec, M.; Rokob, T. A.; Schwartz, J. K.; Kwak, Y.; Rulíšek, L.*; Solomon, E. I.*: Structural and Spectroscopic Properties of the Peroxodiferric Intermediate of *Ricinus Communis* Soluble Δ^9 Desaturase. *Inorg. Chem.* **2012**, *51,* 2806-2820.
19. Rokob, T. A.; Srnec, M.; Rulíšek, L.*: Theoretical Calculations of Physico-Chemical and Spectroscopic Properties of Bioinorganic Systems: Current Limits and Perspectives. *Dalton Trans.* **2012**, *41*, 5754-5768. *Invited Perspective.*
20. Rulíšek, L.*; Ryde, U.: Theoretical Studies of Active-Site Structure, Spectroscopic Properties, and the Mechanism of Multicopper Oxidases. *Coord. Chem. Rev.* **2012**, *256*, In press DOI: 10.1016/j.ccr.2012.04.019. *Invited Review.*