Univerzita Karlova
Přírodovědecká fakulta

Mgr. Lubomír Rulíšek, CSc. DSc.

MODELLING PHYSICO-CHEMICAL PROPERTIES OF BIOINORGANIC SYSTEMS

MODELOVÁNÍ FYZIKÁLNĚ-CHEMICKÝCH VLASTNOSTÍ BIOANORGANICKÝCH SYSTÉMŮ

Habilitační práce

Praha, 2018

Acknowledgments

I wish to thank to all of my mentors, students, postdoctoral fellows, and collaborators; without their hard work, enthusiasm, and fresh ideas there would be much less to present in my habilitation thesis.

First of all, I thank Zdeněk Havlas who introduced me to the world of quantum chemistry and for that I am greatly indebted to him. Ulf Ryde – supervisor of my postdoctoral stay at Lund University, Sweden - has directed my career towards enzymatic reaction mechanisms, QM/MM methods and the field of metal ions in biomolecules. My long-term collaborator and mentor, Prof. Ed Solomon, Stanford University, U. S. A., taught me how to correlate quantum-chemical calculations to bioinorganic spectroscopy which hopefully prevents us from walking astray from the nature‘s design of metalloproteins, or biomolecules in general.

Martin Srnec, my first Ph.D. student, has started his independent and successful career in bioinorganic chemistry few years ago and I consider it as a great satisfaction for a supervisor. Andras Rokob, the most brilliant postdoctoral fellow ever working in my group has significantly contributed to many developments and his spirit is still alive in the group. Ondro Gutten, Daniel Bím, and Vojtěch Klusák have been very smart and hard-working Ph.D. students and it is (or has been) my privilege to work with them. Various postdoctoral fellows from more than five countries have left their footprints in the Rulíšek group at IOCB, Prague. Last but not least, Michal Straka joined the group as the senior scientist at the beginning of 2017 and greatly contributed to the diversity of the topics studied and to the depth of the scientific discussions carried out on daily basis.

The support from the IOCB and the creative and multidisciplinary environment within the institute shall be gratefully acknowledged. Financial support from the Grant Agency of the Czech Republic, and Gilead Sciences, Inc. through the joint research center have been vital for functioning of the group and for the research to be carried out.

Finally, I want to express my gratitude to my ex-wife Karolína, my partner Hedvika, my three children Filip, Lucie, and Aneta for their continuous support – I would not enjoy my professional life without happiness in my personal life.

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 cleavage (or formation). 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. While experiments (e.g., X-ray crystallography, various spectroscopic techniques, and 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. By correlating experimental and theoretical data, the reaction mechanisms of bioinorganic systems can be elucidated. This not only sheds light on the physicochemical principles (laws) governing the chemical behaviour of bioinorganic systems, but also provides an insight into the phenomena of metal-ion selectivity. Ultimately, we hope to understand the fundamental question: „Why Nature selected particular metal ion(s) to perform specific task(s)? " In this thesis, our work concerning computational treatment of transition-metal ion containing systems (most notably metalloproteins) is compiled. The work itself has been published in more than twenty research articles that are integral part of the thesis.

Specifically, our efforts in the area of *theoretical bioinorganic chemistry* are exemplified on 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), and mononuclear manganese super-oxide dismutase (MnSOD). All of them represent highly challenging systems for quantum chemical methodology, mostly due to the complexity of the electronic structure of the active site (MCOs, MnSOD, and Δ^9D). It has been 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. The similarities and differences between

Closely related to bioinorganic chemistry is our 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, our extensive efforts in computations of electrochemical properties of bioinorganic systems are summarized in the thesis.

Table of Contents

1. 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.[8] 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.[9] In this way, we have begun to understand the fundamental physicochemical features of metalloenzyme catalysis.[10]

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.[11,12] This size of the system is not accessible for the high-level wave function methods; however, recent advances[13,14,15] 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).[16] However, the task of calculating the experimental data with spectroscopic accuracy is far from being trivial.[12] 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.

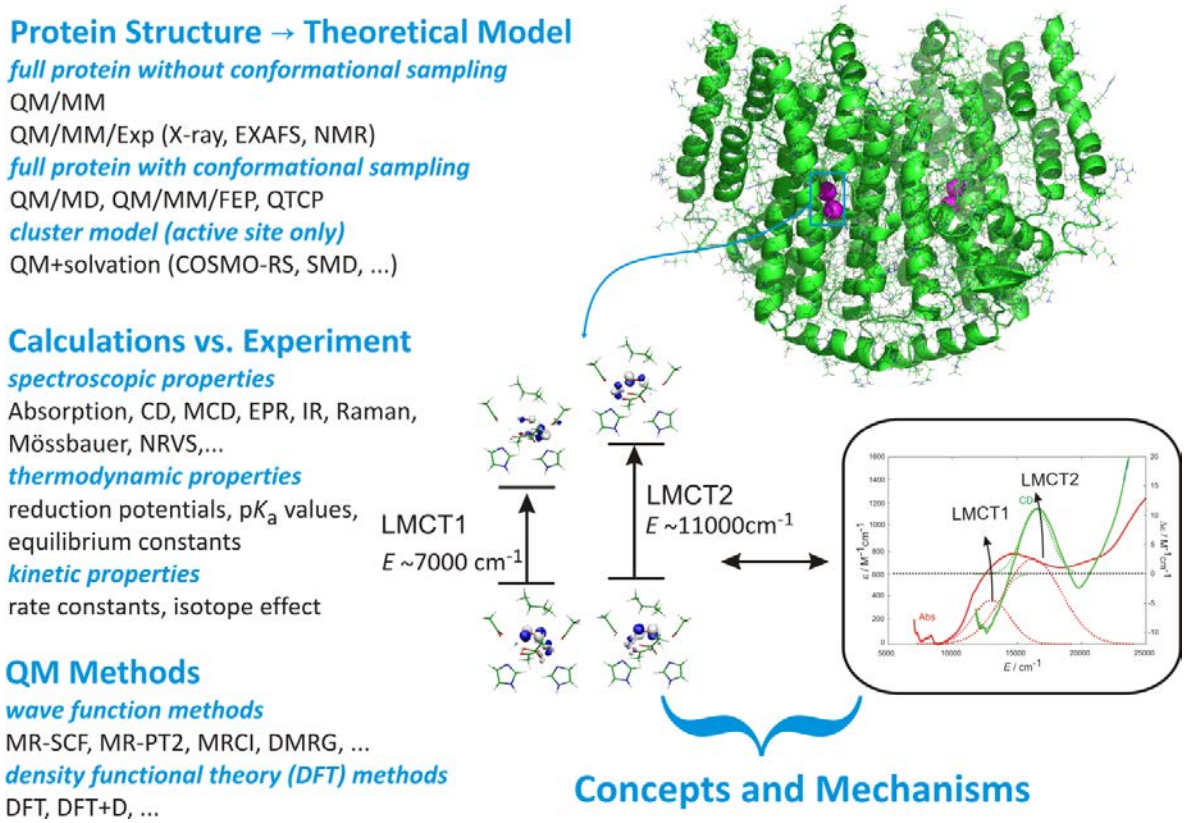

Figure 1: The arsenal of tools and methods available in theoretical bioinorganic chemistry and an outline of the general computational strategy in studying the structure, properties and mechanisms of bioinorganic systems. As an illustrative example a dinuclear non-heme diiron enzyme Δ^9 desaturase is used (in its 1,2-μ-peroxodiferric state).[17]

2. 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[18,19]:

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

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).[20] 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).[21] 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}).[22] 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,[20,23,24,25] combined QM and molecular mechanics (QM/MM) modelling, ranging from standard QM/MM optimisations[20] to the combination of QM/MM optimisation with raw EXAFS data[26] (QM/MM-EXAFS) to address the O_2 reactivity in the TNC and QM/MM free-energy perturbations[27] (QTCP) to accurately address phenomena such as the Cu-T1 $\rightarrow$ TNC electron transfer (reorganization energies).[28] Our achievements and all the history of MCO theoretical studies are well-documented in great details in our recent comprehensive review.[29]

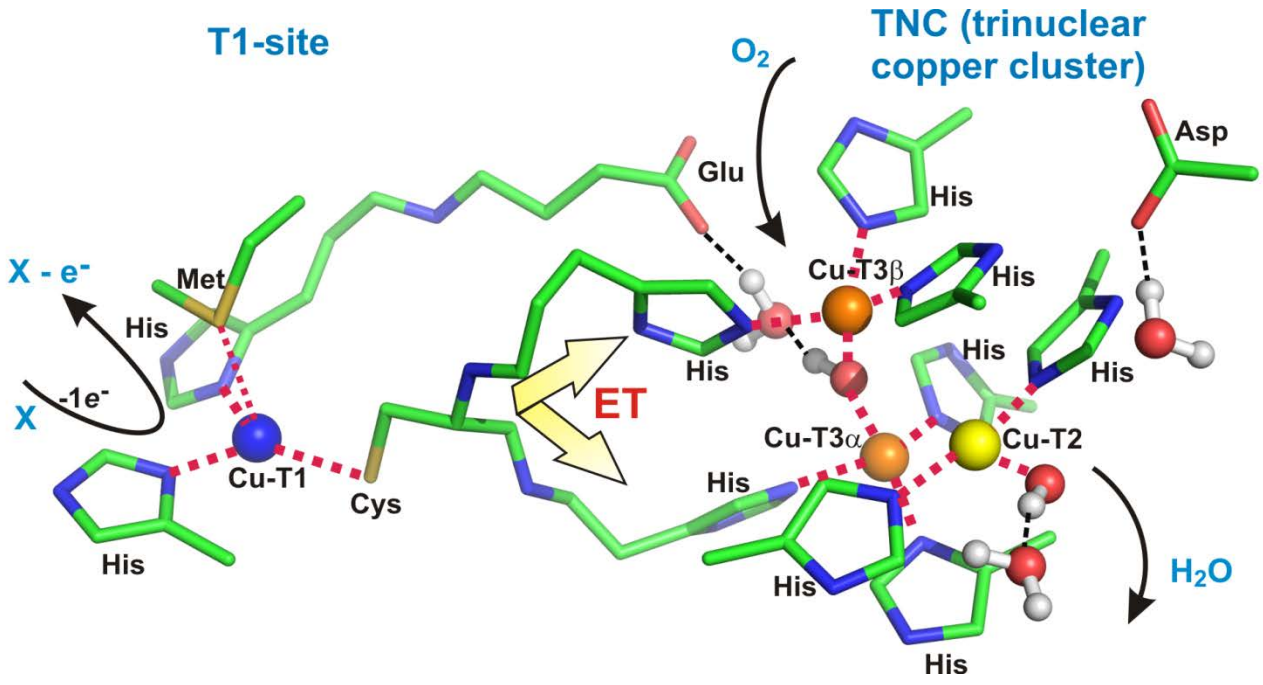

Figure 2: The general architecture of the trinuclear copper cluster site (the site of the four-electron $O_2 \rightarrow H_2O$ reduction) and of the Cu-T1 site (the site of one-electron oxidations of organic substrates or metal ions).

Our (likely) “last bow” in the field of MCOs structure/function correlations is represented by two contributions, both of them carried out in collaboration with Prof. Ulf Ryde (Univ. Lund, Sweden) and one of them also with our experimental collaborator Sergey Shleev (Univ. Malmö, Sweden). In the latter[30] we investigated an appealing possibility of uphill intramolecular electron transfer by combining electrochemical data and quantum chemical calculations. For one of the members of MCO family, bilirubin oxidase, the calculated reorganization energies, together with measured reduction potentials and known structural parameters (distance between the primary redox site at 13 Å distant mononuclear Cu-T1 site and the trinuclear Cu-T2T3 copper cluster as the final acceptor of electrons) lead to a surprising finding that at least one of the redox steps in the overall transfer of four electrons might be endergonic. This has invoked a thorough study[31] of all redox and protonation states of presumably all plausible intermediates in the full catalytic cycle of the MCOs – employing multitude of methods of molecular modelling, such as 2-QM/MM, QM/MM-FEP (free energy perturbation). The full catalytic cycle is depicted in Figure 3 and is consistent with recent efforts in the Solomon lab that combine QM calculations for structural models of active site (neglecting the rest of the protein) with the state-of-the-art spectroscopic data.

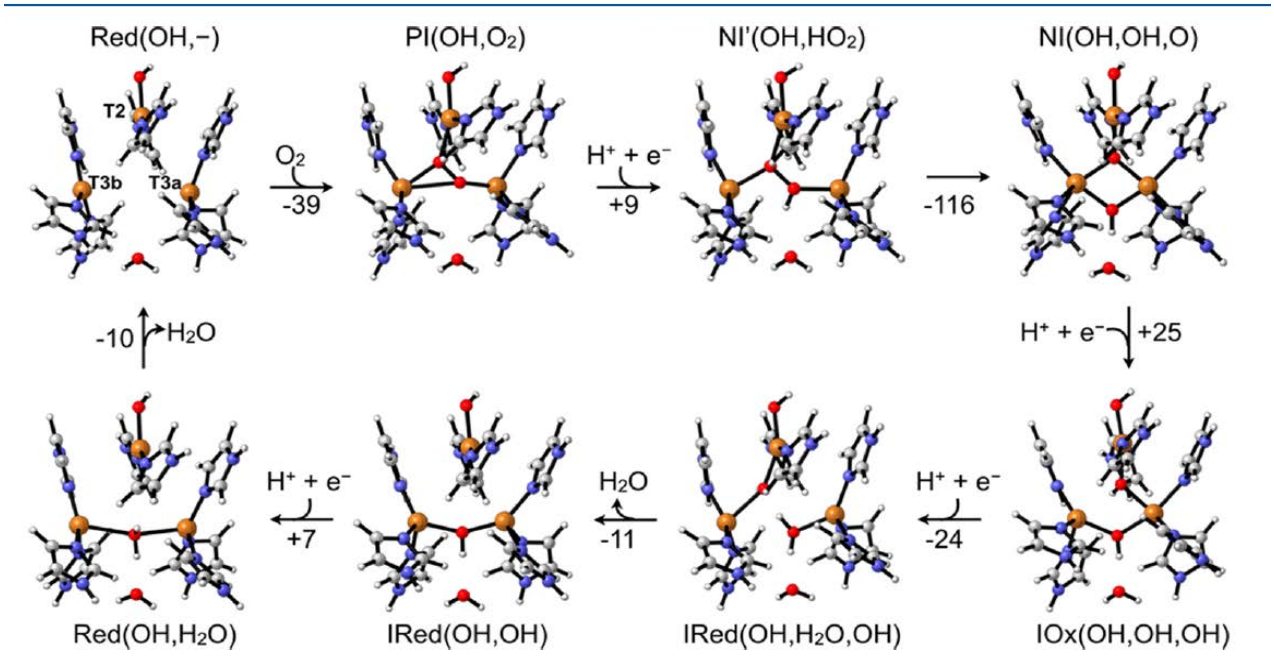

Figure 3: Suggested catalytic cycle of the MCOs. Calculated reaction free energies (in kJ.mol^{-1}) are given for each step, assuming a pH of 7, the fact that the electrons come from the Cu_{T1} site and that the Cu_{T1} site is reduced in the steps not involving electron transfer.

Δ^9 Desaturase: Prominent Non-Heme Diiron Enzyme. 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.[32]

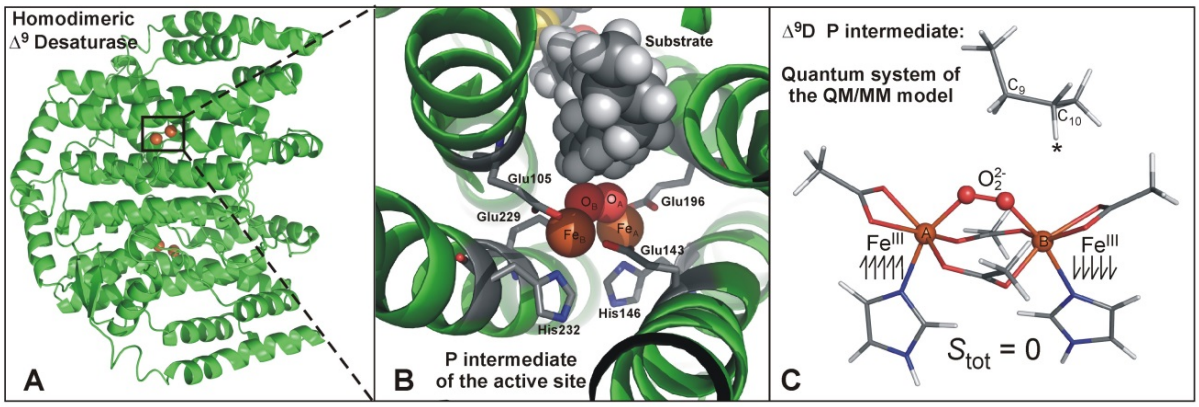

Figure 4. **A.** Homodimer of Δ^9 Desaturase with two $NHFe_2$ active sites (orange spheres). **B.** One of the two active sites in the **P** intermediate (with 1,2-μ binding mode for O_2) in the presence of the substrate. The ligating residues are displayed as sticks (the structure is taken from the QM/MM model). **C.** The quantum region of the QM/MM model of the Δ^9D **P** intermediate. As shown experimentally, the **P** intermediate has the $S = 0$ ground state arising from two antiferromagnetically-coupled high-spin ($S_{Fe} = 5/2$) Fe^{III} centers that are bridged by two Glu residues and a peroxide ligand. The star indicates an aliphatic H atom that, during the catalysis, has to be activated to trigger the insertion of the double bond between C_{10} and C_9 sites.

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.[12] 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 (Figure 4)

The next step in understanding the reaction mechanism of Δ^9D is computation of the activation energy for the (presumably) rate-determining step – abstraction of the first hydrogen from the alkane chain of the substrate. This can be achieved either homolytically or heterolytically (as hydride anion). It has been soon realized that most of the standard and routinely used DFT functionals fail to describe the complex electronic structure of the "enzyme+O_2+susbstrate" system. The span of computed activation energies by DFT is enormous (10-200 kJ.mol^{-1}, *c.f.* Figure 5A) which precludes any quantitative assessment and understanding of the Δ^9D reactivity. To this aim, density-matrix renormalization group (DMRG) complete active space second order perturbation theory (DMRG-CASPT2) calculations were carried out for the model active site (of approx. size depicted in Figure 4C) in various states of the catalytic cycle.[33]

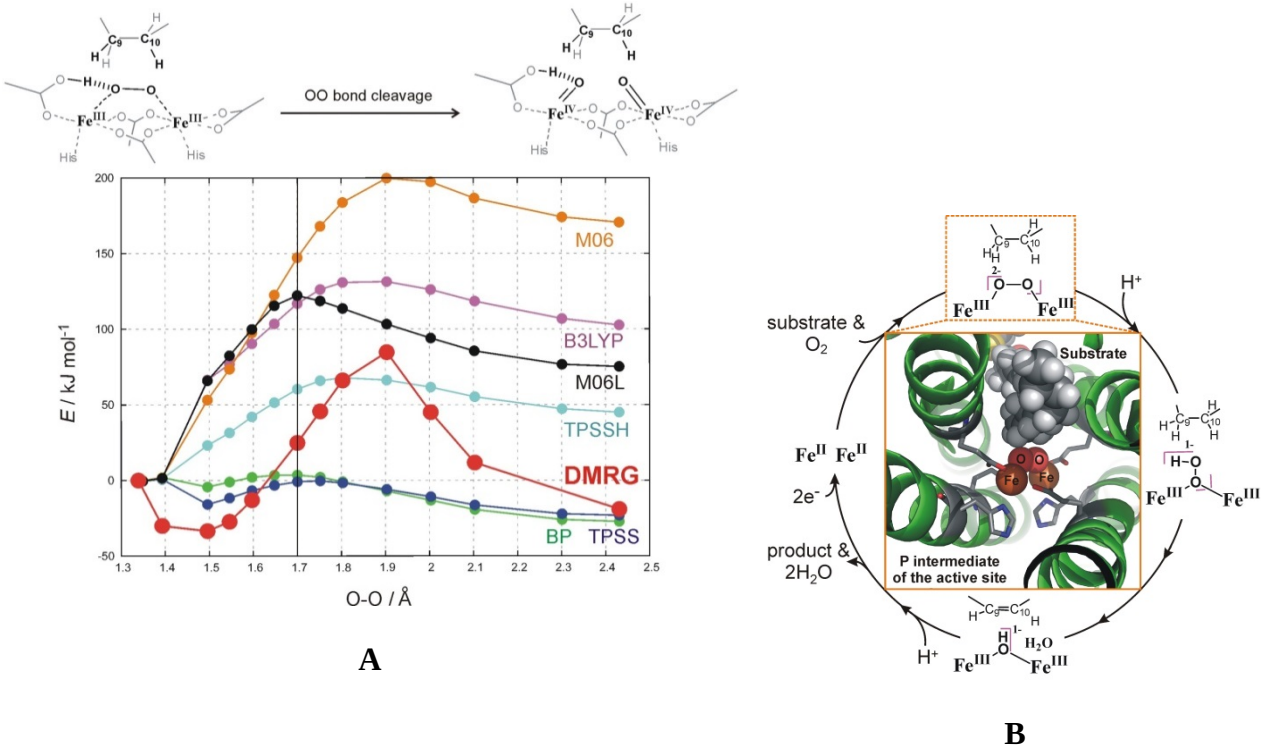

Figure 5: A. Performance of various DFT functionals for one of the plausible steps in the mechanism, O-O bond cleavage **B.** The Δ^9D reaction mechanism as suggested by the QM/MM and DMRG-CASPT2 calculations..

The computed DMRG-CASPT2 barriers were in agreement with the experimental rate constants and to our knowledge, it has been the first application of the DMRG-CASPT2 for the bioinorganic reactivity. It did not, however, solve all the issues which precludes full understanding of the Δ^9D reactivity. Most notably, the question why (and when) the enzyme desaturates and not hydroxylates has not yet been satisfactorily answered. It is a subject of intense and ongoing computational efforts.

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

$$M^{III}SOD_{OH} + O_2^{\bullet -} + H^+ \rightarrow M^{II}SOD_{H_2O} + O_2 \quad (2)$$

$$M^{II}SOD_{H_2O} + O_2^{\bullet -} + H^+ \rightarrow M^{III}SOD_{OH} + H_2O_2 \quad (3).$$

According to the metal ion present as the cofactor in the active site, they can be divided into three classes: MnSODs, FeSODs, and CuZnSODs.[34]

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. 2 and 3). However, the exact reaction mechanism for this disproportionation reaction was unknown. As can be seen from Eqs. 2 and 3, 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.[35] 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. 35.

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).[36] 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 6), 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[37] we used the quantum and molecular mechanical thermodynamic cycle perturbation method[27] (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 experiment.[37] Thus, we have shown that it is in principle possible (though extremely difficult) to address these complex phenomena computationally.

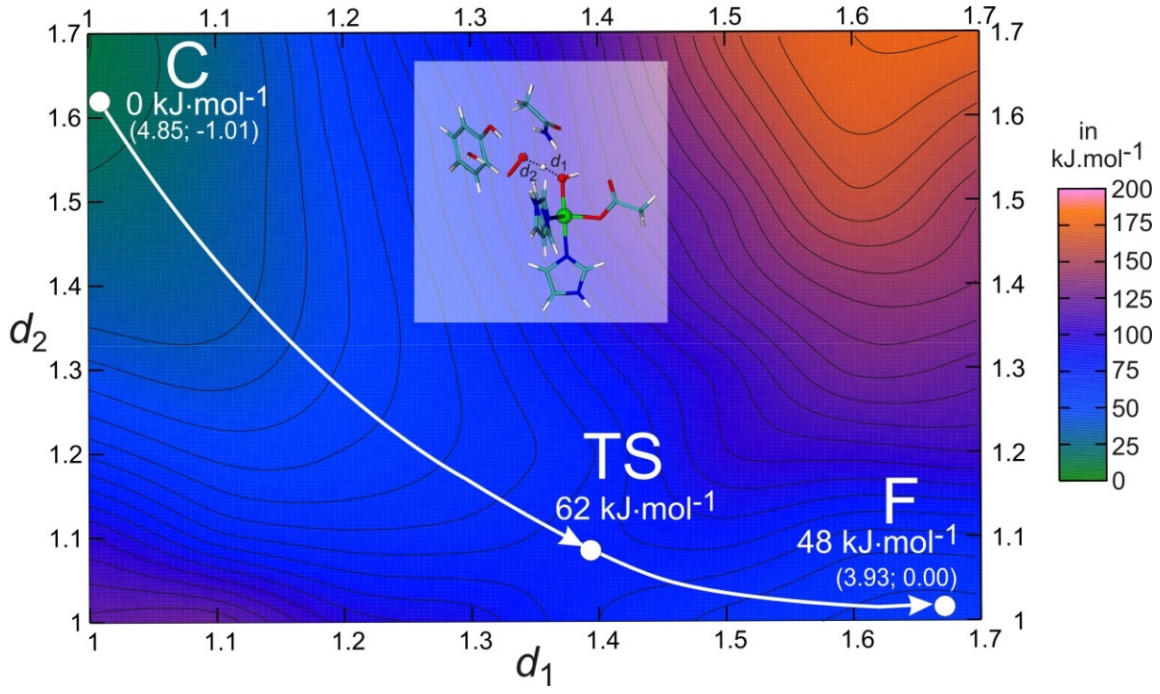

Figure 6. The two-dimensional QM(B3LYP/6-31G*)/MM potential-energy surface of the proton-coupled electron transfer from Mn–OH_2 to O_2 in the second-sphere mechanism. The interatomic distances are in Å. The numbers (x; y) in the graph depict the spin densities at the manganese and O_2 centres.

3. Theoretical Calculations of Metal Ion Complexation in Peptides and Proteins

Predicting Stability Constants (β) from the First Principles. Our long-lasting efforts to understand the metal uptake by biomolecules were to some extent accomplished by a study attempting to calculate stability constants from first principles.[38] Stability constant (β) is the most important experimental quantity describing thermodynamics of metal ion binding with various (in)organic ligands, or biomolecules. In principle, it can be calculated as the free energy change associated with the metal ion complexation, i.e. its uptake from the solution under standard conditions. Since this process is associated with interactions of charged species, large values of interaction and solvation energies are in general involved. Using the standard thermodynamic cycle (*in vacuo* complexation, and solvation/desolvation of the reference state and of the resulting complexes), one usually subtracts values of several hundreds of kcal.mol^{-1} to obtain final results in the order of units or tens kcal.mol^{-1}. We used density functional theory (DFT) and Møller-Plesset second order perturbation theory (MP2) calculations together with the conductor-like screening model for realistic solvation (COSMO-RS method) to calculate the stability constants of selected complexes – $[M(NH_3)_4]^{2+}$, $[M(NH_3)_4(H_2O)_2]^{2+}$, $[M(Imi)(H_2O)_5]^{2+}$, $[M(H_2O)_3(His)]^{+}$, $[M(H_2O)_4(Cys)]$, $[M(H_2O)_3(Cys)]$, $[M(CH_3COO)(H_2O)_3]^{+}$, $[M(CH_3COO)(H_2O)_5]^{+}$, $[M(SCH_2COO)_2]^{2-}$ – with eight divalent metal ions (Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, Cd^{2+}, and Hg^{2+}). Using the currently available computational protocols we show that it is possible to achieve *relative* accuracy of 2-4 kcal.mol^{-1} (1-3 orders of magnitude in β). However, since most of the computed values are affected by metal-dependent and ligand-dependent systematic shifts, the accuracy of the 'absolute' (uncorrected) values is generally lower. For metal-dependent systematic shifts we propose specific values to be used for a given metal ion and the current protocol. At the same time, we argue that ligand-dependent shifts (which cannot be easily removed) do not influence the metal-ion selectivity of the particular site and therefore, the selectivity can be computed to within 2 kcal.mol^{-1} average accuracy. In near future, we want to fully utilize the protocol for calculating the binding of above ions with all possible combinations of metal-binding side chains of amino-acids.

Contribution of Higher Coordination Spheres to Metal Ion Selectivity. Having accomplished the quantitative evaluation of the selectivity of the metal-binding site employing the first-shell model, a question arose to what extent is the selectivity of the particular metal-binding site in a protein determined by the first-shell coordination sphere. In an attempt to answer this question, six model peptides complexed with the set of divalent metal ions (Mn^{2+}, Fe^{2+}, Co^{2+},

Ni^{2+}, Cu^{2+}, Zn^{2+}, Cd^{2+}, and Hg^{2+}) and their various first-shell representations were studied.[39] It has been shown that the metal-induced selectivity of a system is a complex function of multiple factors, that define measure of its localization to a binding site, *i.e.* the possibility of capturing it using a 1st shell model of the system. It is essential that this model includes all of the ligands bound to the metal-ion. This requirement alone usually ensures that the binding is described reasonably well by the model and that the metal-ion is properly "shielded" from parts of a system further away, bringing the average absolute error to values no higher than few units kcal.mol^{-1}. However, some binding geometries were problematic due to insufficient shielding, e.g. linear or square planar geometry. Achieving more satisfactory accuracy requires consideration of the nature of ligands. Softer ligands, represented in the realm of metalloproteins by cysteine and methionine residues, require bulkier representation, e.g. the whole side-chain of a residue. Hard ligands, like carboxyl groups, alcohols, amines, can be described at a similar level of accuracy more economically by the functional group plus single methyl group. Imidazole is sufficient to represent histidine side-chain, *N*-methylmethylamide to represent backbone amide group, and methylamide to represent glutamine and asparagine ligands. This level of description can be expected to result in average absolute deviations below 1 kcal.mol^{-1}, which translates to maximum relative error (among studied metal ions) below 3 kcal.mol^{-1}. Non-covalent partners of ligands usually influence retention of selectivity in a negative way. The presence of charged groups can be influential, especially in case of insufficient screening or their proximity to metal ion due to small 1st shell ligand, e.g. water molecule. In case of hydrogen bonding the effect is much less significant, influencing average absolute deviation by tenths of kcal.mol^{-1}. The strength of the hydrogen bond as well as softness of the acceptor can increase this value. In summary, we have provided rigorous computational proof that most of the metal-ion selectivity is governed by the structure and composition of the 1st shell and pointed out several caveats to this general statement.

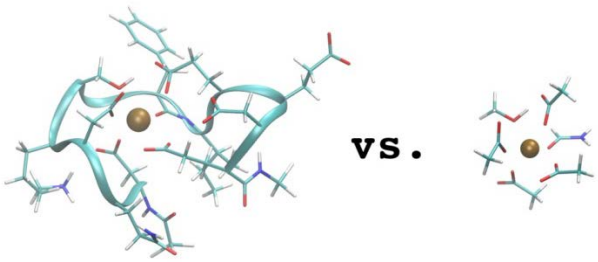

Figure 7: Factors determining metal-ion selectivity in peptidic sites were elucidated using the recently benchmarked DFT(BP86-D3//COSMO-RS) computational protocol.

4. Computational Electrochemistry and Beyond

Accurate Calculations of Reduction Potentials Employing (VTHAA-)DFT/COSMO-RS Methodology. Encouraged by the excellent agreement between the computed and experimental reduction potentials for the series of 21 substituted ferrocene/ferrocenium couples (MAD of 0.03 V for the span of 0.7 V in the measured E^0's) employing DFT(PBE-D3)/def2-TZVP//COSMO-RS protocol,[40] the notoriously difficult group 8 (Fe, Ru and Os) octahedral complexes were revisited.[41] It has been shown that COSMO-RS takes care of the second solvation sphere effects for the computed reduction potentials which were shown to be, using standard PCM models, up to 1V! This has been the case for the $[MX_6]^{2+}/[MX_6]^{3+}$ oxidations. Still, the results for even more highly (formally) charged $[MX_6]^{4-}/[MX_6]^{3-}$ redox pairs (X = CN^-, Cl^-, …) were not satisfactory.

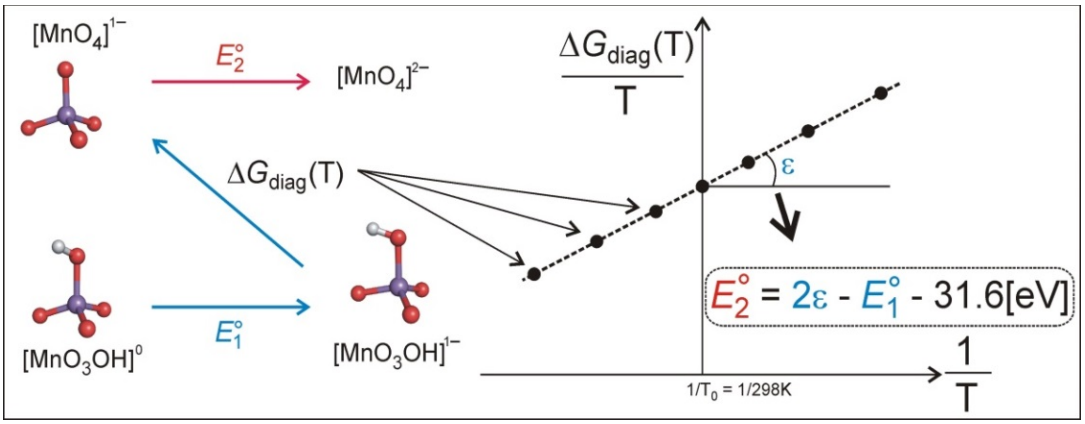

Figure 8. Computational protocol for the prediction of the reduction potential of a negatively (or positively) charged complex (E_2°) by means of the reduction potential of its protonated (or deprotonated) counterpart (E_1°). As an example, the reduction-potential calculation for $[MnO_4]^-$ is schematically depicted. Note that the constant of –31.6 eV corresponds to $2G_{solv}(H^+) - 2E^\circ_{abs}(SHE)$

To this aim, Srnec and Bím[42] devised an elegant protocol denoted variable temperature H-atom addition/abstraction (VTHAA) method. It has been shown that VTHAA method (protocol) yield accurate reduction potentials of the highly charged species by calculating the

data for their less charged cognates (Figure 8). Very recently, the protocol has been also successfully applied to a series of 47 electrochemically characterized [Fe(X)$_n$(O)] complexes which were considered by many as one of the most challenging systems for the computations of redox potentials.[43]

5. 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.[44] 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[45,46] 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+}). 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.[44] 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 9 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 Ref. 44.

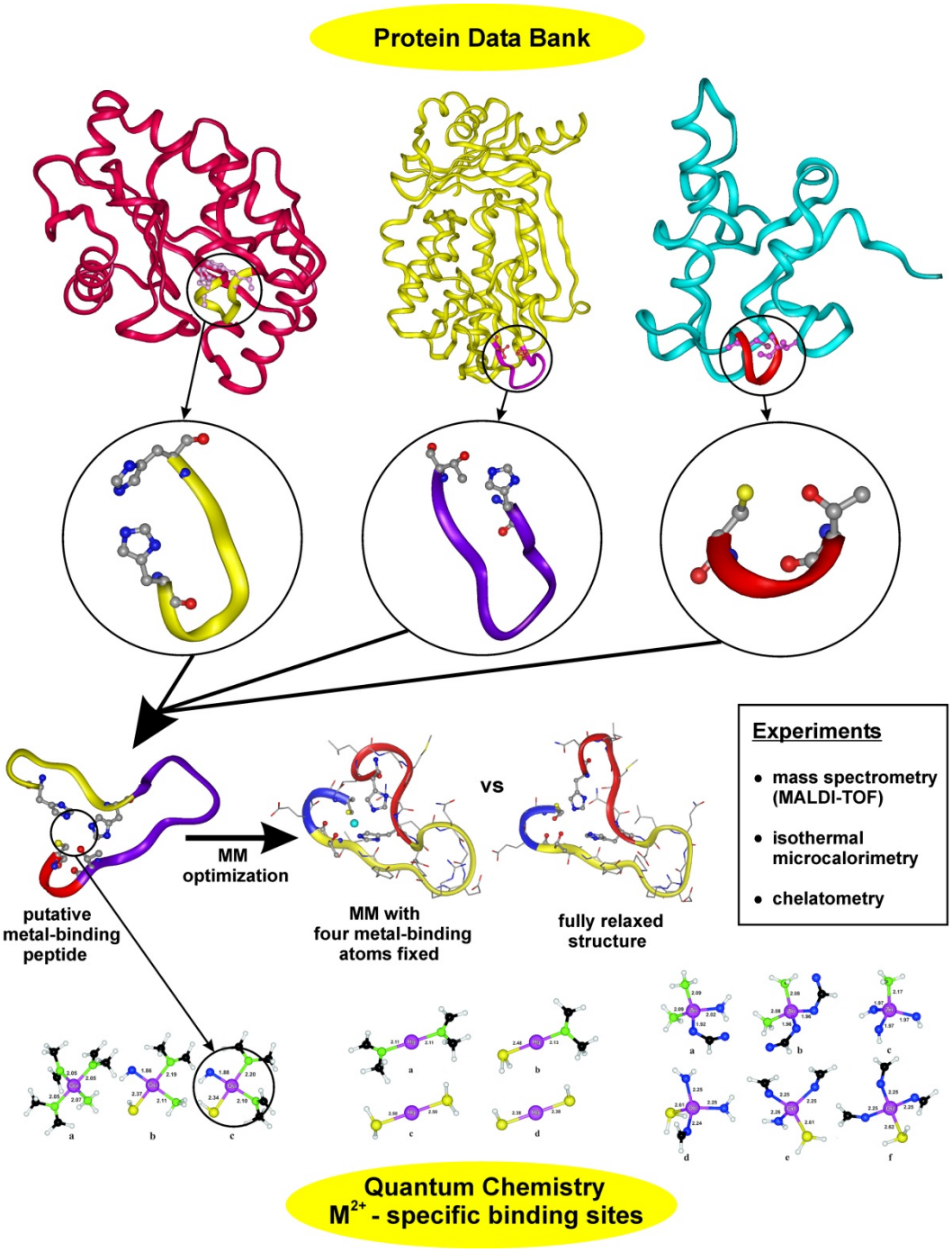

Figure 9: The overview of the novel computational strategy for the design of specific metal-binding peptide sequences.

6. 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.[47] 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).[48] 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 six fold 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 10 and more details can be found in Ref. 48.

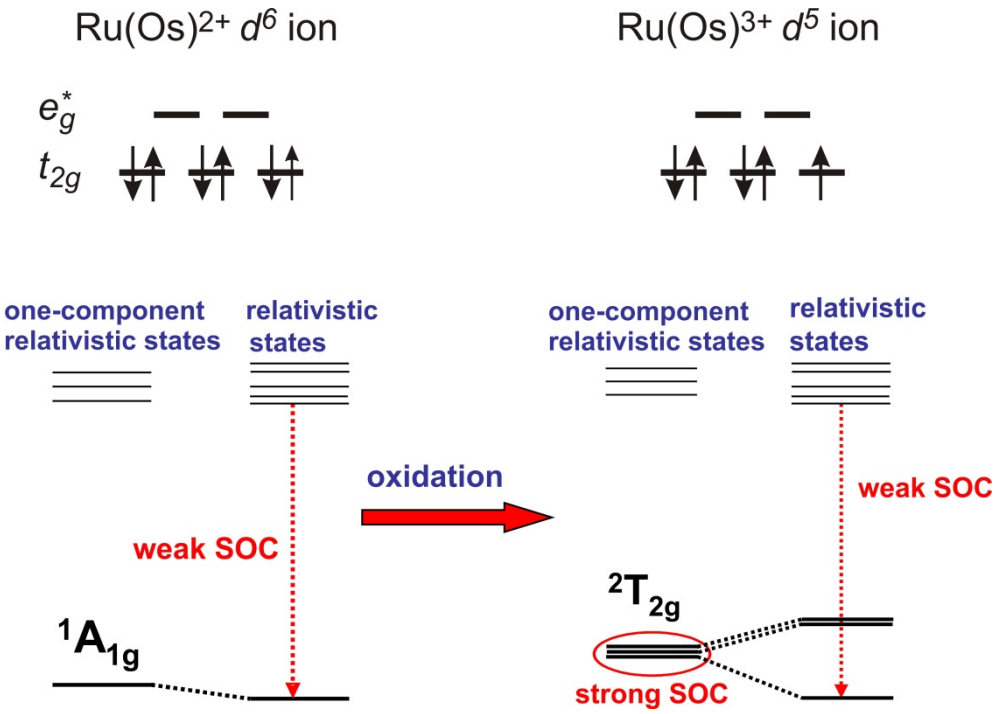

Figure 10. Schematic representation of the SOC effects on the values of the reduction potentials of $Ru^{2+/3+}$ and $Os^{2+/3+}$ octahedral complexes.

7. Conclusions

The thesis compiles some of our contributions to the field of *theoretical bioinorganic chemistry*. Its author witnessed the development of the field over the past two decades. 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.*, Figure 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.

Current computational efforts in the Rulíšek group are best characterized as "striving for quantitative accuracy in (not only) theoretical bioinorganic chemistry". Latest developments in quantum chemistry (DFT-D3 functionals, multi-reference methods, efficient coupled cluster implementations, such as DLPNO-CCSD(T),[14] solvation methods,…) provide computational chemists with tools that enable to tackle the problems of increasing complexity with near-quantitative accuracy (~10 $kJ.mol^{-1}$ in many real-world applications). If and only if we are able to quantitatively reproduce the experimental data, we may conceive more general concepts, decomposition schemes, qualitative design rules, etc.

Still, it is to be seen in the next decade whether all of these might be applied in a way to not only interpret and complement experimental data, but to drive the chemical research further by guiding the experiments, designing new materials, ligands, and catalysts *ab initio*.

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) Solomon, E. I.; Heppner, D. E.; Johnston, E. M.; Ginsbach, J. W.; Cirera, J.; Qayyum, M.; Kieber-Emmons, M. T.; Kjaergaard, C. H.; Hadt, R. G.; Tian, L. *Chem. Rev.* **2014**, *114*, 3659−3853.

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

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

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

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

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

(14) Liakos, D. G.; Neese, F. *J. Chem. Theory Comput.* **2015**, *11*, 4054-4063.

(15) Yanai, T.; Kurashige, Y.; Mizukami, W.; Chalupský, J.; Tran Nguyen Lan; Saitow, M. *Int. J. Quantum Chem.* **2015**, *115*, 283–299.

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

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

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

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

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

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

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

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

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

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

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

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

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

(29) Rulíšek, L.; Ryde, U. *Coord. Chem. Rev.* **2013**, *257*, 445-458.

(30) Shleev, S.; Andoralov, V.; Falk, M.; Reimann, C. T.; Ruzgas, T.; Srnec, M.; Ryde, U.; Rulíšek, L. *Electroanalysis* **2012**, *24*, 1524-1540.

(31) Li, J.–L.; Farrokhnia, M.; Rulíšek, L.; Ryde, U. *J. Phys. Chem. B* **2015**, *119*, 8268–8284.

(32) Rokob, T. A.; Chalupský, J.; Bím, D.; Andrikopoulos, P. C.; Srnec, M.; Rulíšek, L. *J. Biol. Inorg. Chem.* **2016**, *21*, 619-644.

(33) Chalupský, J.; Rokob, T. A.; Kurashige, Y.; Yanai, T.; Solomon, E. I.; Rulíšek, L.; Srnec, M. *J. Am. Chem. Soc.* **2014**, *136*, 15977-15991.

(34) Sheng, Y.; Abreu, I. A.; Cabelli, D. E.; Maroney, M. J.; Miller, A.-F.; Teixeira, M.; Valentine, J. S. *Chem. Rev.* **2014**, *114*, 3854-3918.

(35) Rulíšek, L.; Ryde, U. *J. Phys. Chem. B* **2006**, *110*, 11511.

(36) Srnec, M.; Aquilante, F.; Ryde, U.; Rulíšek, L. *J. Phys. Chem. B* **2009**, *113*, 6074.

(37) Heimdal, J.; Kaukonen, M.; Srnec, M.; Rulíšek, L.; Ryde, U. *ChemPhysChem* **2011**, *12*, 3337.

(38) Gutten, O.; Rulíšek, L. *Inorg. Chem.* **2013**, *52*, 10347-10355.

(39) Gutten, O.; Rulíšek, L. *Phys. Chem. Chem. Phys.* **2015**, *17*, 14393-14404.

(40) Khobragade, D. A.; Mahamulkar, S. G.; Pospíšil, L.; Císařová, I.; Rulíšek, L.; Jahn, U. *Chem. Eur. J.* **2012**, *18*, 12267-12277.

(41) Rulíšek, L. *J. Phys. Chem. C* **2013**, *117*, 16871-16877.

(42) Bím, D.; Rulíšek, L.; Srnec, M. *J. Phys. Chem. Lett.* **2016**, *7*, 7-13.

(43) Bím, D.; Rulíšek, L.; Srnec, M.: Computational Electrochemistry as a Reliable Probe of Experimentally Elusive Mononuclear Non-Heme Iron Species. *J. Phys. Chem. C* **2018**, *accepted*.

(44) 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.

(45) Rulíšek, L.; Havlas, Z. *J. Am. Chem. Soc.* **2000**, *122*, 10428-10439.

(46) Rulíšek, L.; Havlas, Z. *J. Phys. Chem. B* **2003**, *107*, 2376-2385.

(47) 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.

(48) Srnec, M.; Chalupský, J.; Fojta, M.; Zendlová, L.; Havran, L.; Hocek, M.; Kývala, M.; Rulíšek, L. *J. Am. Chem. Soc.* **2008**, *130*, 10947.

List of Publications Comprising the Habilitation Thesis (*…corresponding author(s)):

1. 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.
2. 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.
3. 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.
4. 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.
5. Ryde, U.*; Hsiao, Y.-W.; Rulíšek, L., Solomon, E. I.: Identification of the Peroxy Adduct in Multi-Copper Oxidases by a Combination of Computational Chemistry and EXAFS Measurements. *J. Am. Chem. Soc.* **2007**, *129*, 726-727.
6. 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.
7. 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.
8. 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.
9. 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.
10. 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.
11. 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. (Correction: *J. Phys. Chem. A* **2012**, *116*, 8407-8407).

12. 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.
13. Rulíšek, L.*; Ryde, U.*: Theoretical Studies of the Active-Site Structure, Spectroscopic, and Thermodynamic Properties, and Reaction Mechanism of Multicopper Oxidases. *Coord. Chem. Rev.* **2013**, *257*, 445-458. *Invited Review.*
14. Rulíšek, L.*: On the Accuracy of Calculated Reduction Potentials of Selected Group 8 (Fe, Ru and Os) Octahedral Complexes. *J. Phys. Chem. C* **2013**, *117*, 16871-16877.
15. Gutten, O.; Rulíšek, L.*: Predicting the Stability Constants of Metal-Ion Complexes from First Principles. *Inorg. Chem.* **2013**, *52*, 10347-10355.
16. Chalupský, J.; Rokob, T. A.; Kurashige, Y.; Yanai, T.; Solomon, E. I.*; Rulíšek, L.*; Srnec, M.*: Reactivity of the Binuclear Non-Heme Iron Active Site of Δ^9 Desaturase Studied by Large-Scale Multireference *Ab Initio* Calculations. *J. Am. Chem. Soc.* **2014**, *136*, 15977-15991.
17. Gutten, O.; Rulíšek, L.*: How simple is too simple? Computational perspective on importance of second-shell environment for metal-ion selectivity. *Phys. Chem. Chem. Phys.* **2015**, *17*, 14393-14404.
18. Li, J.–L.; Farrokhnia, M.; Rulíšek, L.; Ryde, U.*: Catalytic Cycle of Multicopper Oxidases Studied by Combined Quantum- and Molecular-Mechanical Free-Energy Perturbation Methods. *J. Phys. Chem. B* **2015**, *119*, 8268–8284.
19. Bím, D.*; Rulíšek, L.; Srnec, M.*: Accurate Prediction of One-Electron Reduction Potentials in Aqueous Solution by Variable-Temperature H-Atom Addition/Abstraction Methodology. *J. Phys. Chem. Lett.* **2016**, *7*, 7-13.
20. Rokob, T. A.; Chalupský, J.; Bím, D.; Andrikopoulos, P. C.; Srnec, M.*; Rulíšek, L.*: Mono- and binuclear non-heme iron chemistry from a theoretical perspective. *J. Biol. Inorg. Chem.* **2016**, *21*, 619-644. *Invited Perspective.*