Combination of advanced X-ray spectroscopy and quantum chemistry to determine electronic structures of metal-organic cofactors in proteins

Dissertation

zur Erlangung des Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

am Fachbereich Physik

der Freien Universität Berlin

vorgelegt von

Nils Schuth

Berlin 2018

Erstgutachter: Priv.-Doz. Dr. Michael Haumann

Zweitgutachterin: Prof. Dra. Maria A. Mroginski

Tag der Disputation: 01.06.2018

gracias a mis padres

que me han dado tanto

Table of Contents

Abstract

Proteins with metal cofactors (metalloproteins) are capable to facilitate a large variety of chemical processes and are involved in the most challenging catalytic reactions e.g. activating small and stable molecules like N_2, CO_2, H_2, H_2O, CH_4 and O_2 with superior energetic efficiency and turnover rates. Each of these processes was optimized for millennia in nature under conditions of high interest for industrial application such as atmospheric pressure, room temperature and neutral pH, while utilizing earth abundant elements. Understanding of these enzymatic reactions and metal-ligand bonding thus will be helpful in addressing challenges in synthetic catalysts. Advanced X-ray absorption and emission spectroscopy (XAS/XES) allows to determine the oxidation and spin state of metal cofactors by probing unoccupied and occupied electronic states. However, metal cofactors in reactive high-valent states are prone to radiation damage. In this thesis, XAS/XES data collection was significantly accelerated by development of time-resolved energy-sampling (TRES) detection approaches and efficiency evaluation in comparison with conventional measurement approaches. TRES facilitated monitoring of weak spectral XES features with high signal-to-noise ratio on highly diluted and X-ray sensitive materials. Quantitative correlation of XAS/XES data with quantum chemical (QC) simulations derived from density functional theory (DFT) or the complete active space self-consistent field (CASSCF) approach from molecular orbital theory allows description of key structural and electronic parameters in metalloproteins and synthetic compounds. In this thesis, QC calculation protocols were established in our laboratory and spectral simulations were benchmarked with XAS/XES data from selected metalloproteins and synthetic compounds with a special focus on heme proteins. The heme group is the dominant cofactor in proteins that involve oxygen as reactant. Hemoglobin (HB) and myoglobin (MB) are essential for transport, storage and sensing of molecular oxygen in vertebrates and their physiological function has been investigated for more than a century. However, the nature of the electronic configuration of the Fe – O_2 bonding in oxygenated HB/MB (**oxy**) was not settled so far. Mostly differing in the spin state of Fe and the O_2 ligand, apparent contradictory models have dominated the debate. We combined advanced XAS/XES with QC simulations from DFT and CASSCF to determine the electronic configuration of HB/MB in different ligation states and three porphyrin model compounds. Our finding on the spin state and electronic configuration in oxygenated heme in HB and MB provided an adequate description of the metal-ligand interaction, which merged the classical models in a holistic description of the Fe – O_2 bonding situation.

1 General introduction

Proteins are the fundamental building blocks for life on earth, transforming chemistry into biology. They consist of single (monomeric) or multiple (polymeric) peptide chains constructed from amino acids and the sequence of these organic residues is encoded in DNA. Proteins serve a multitude of purposes such as structural support in connective tissue, intra- or intercellular molecular transport or as enzymes catalyzing chemical reactions. Each of these processes and functions were optimized for millennia in nature under conditions of high interest for today's application needs. Most processes occur at ambient conditions, e.g. atmospheric pressure, room temperature and neutral pH, while utilizing abundant materials. A significant part of these reactions involve cofactors (organic, metal-organic or inorganic non-amino acid molecules) that bind to the protein matrix. Examples of protein cofactors include vitamin C (organic), vitamin B_{12} (metal-organic) and iron-sulfur clusters (inorganic). Proteins with a bound metal-organic or inorganic cofactor are denoted metalloproteins. Bound cofactors include first-row transition metals ranging from vanadium to zinc as well as molybdenum and tungsten. Metalloproteins may contribute up to 40 % of all proteins [28]. Metalloproteins are involved in the most challenging catalytic reaction processes e.g. activating small and stable molecules like N_2, CO_2, H_2, H_2O, CH_4 and O_2 with superior energetic efficiency and turnover rates. An example for the variety of metalloproteins in biological chemistry is highlighted in the bacterial nitrogen metabolic cycle, where a different metalloenzym is involved in each step of nitrogen fixation ($N_2 \rightarrow NH_3$), nitrification ($NH_3 \rightarrow NO_3^-$) and denitrification ($NO_3^- \rightarrow N_2$) and metal atoms included are vanadium, molybdenum, iron and copper [28].

Complete understanding of a metalloprotein requires knowledge of (i) function and structure of the protein and (ii) of geometric structure and electronic configuration of the cofactor for all reaction intermediates. DNA sequencing [29] and X-ray diffraction (XRD) [30] yield the protein structure, though resolving various reaction states is challenging [31]. The combination of high resolution XRD (< 2 Å) and extended X-ray absorption fine structure (EXAFS) is a powerful approach for deriving atomic-level information on cofactor structures [32], while identification of protonation state may require e.g. nuclear magnetic resonance (NMR) or neutron diffraction experiments [14, 15]. X-ray absorption and emission spectroscopy (XAS/XES) in combination with quantum chemical calculations and synthetic reference compounds has proven to be a viable method to investigate electronic configurations of metalloprotein cofactors in many systems [33-36].

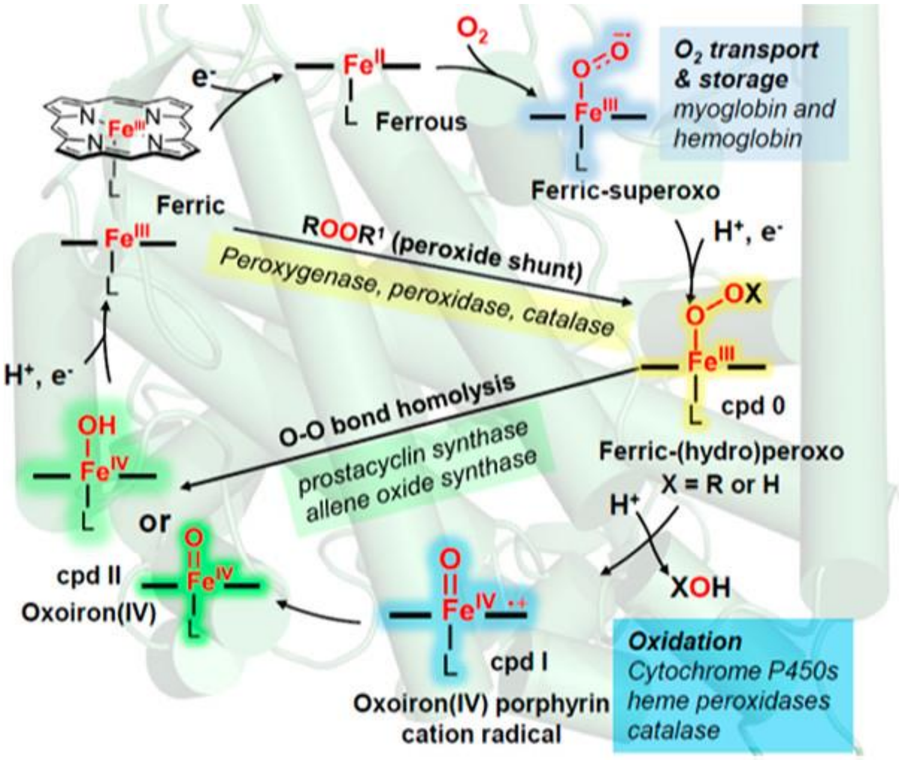
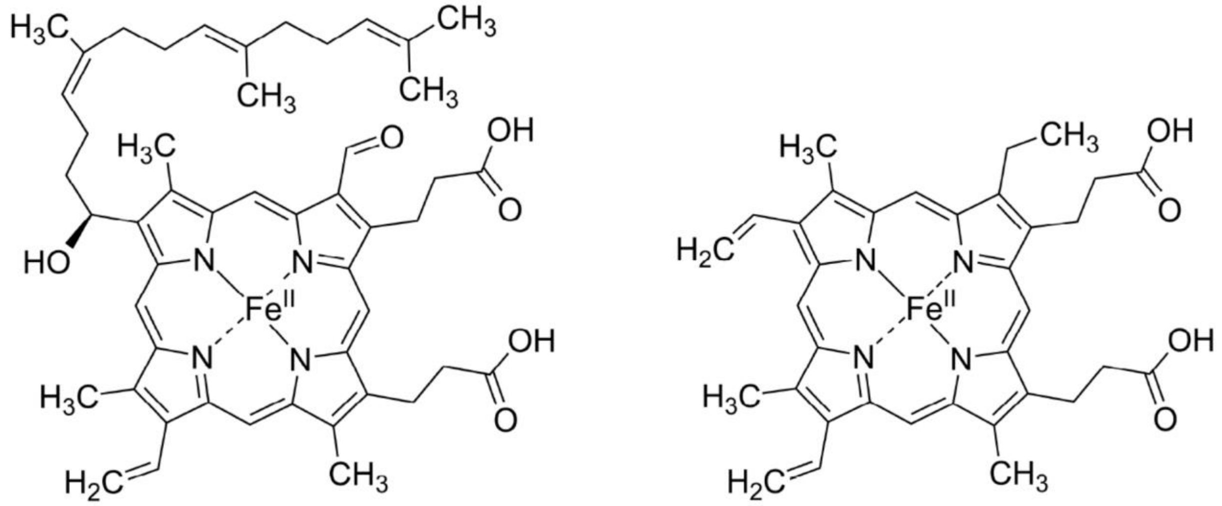


Figure 1-1: Role of heme proteins in oxygen activation processes. Heme proteins perform a large variety of biological functions by formation of metal-superoxo, -(hydro)peroxo and high-valent metal-oxo species. Reproduced from Ref. Huang, et al. [17]. Note, that the Fe(III)-superoxide bond in hemo-/myoglobin is still debated, even in this review article.

Molecular oxygen once was the most toxic waste product in the history of earth, which led to a major species extinction wave around 2.5 billion years ago [37]. The development of photosynthesis by cyanobacteria half a billion years earlier, caused a rise in O_2 concentration in the earth's atmosphere, which resulted in the extinction of most anaerobic organisms on the planet. The evolutionary pressure to survive in a dioxygen rich environment gave rise to aerobic life on earth. As a result, one of the most abundant metal-organic cofactors in nature is the heme, which consists of an iron atom incorporated into a porphyrin ring [17]. The most prominent species are heme A or heme B (see Figure 1-2) [38, 39]. Enzymes with heme cofactors play an important role in oxygen management and are capable of catalyzing reductive and oxidative chemistry involving molecular oxygen. The function of heme enzymes (see Figure 1-1) range from hydrogen peroxide degradation by *catalases*, activation of hydrogen peroxide by *heme peroxidases*, oxygenation by *heme oxygenases*, detoxification through hydroxylation by *cytochrome P450*, heme degradation (utilizing heme as substrate and cofactor) by *heme oxygenases*, to dioxygen transport, storage and sensing by *heme globins* [38, 40, 41]. Consequently, research on heme proteins constitutes a large part of the field of



Heme A **Heme B**

Figure 1-2: Heme A and Heme B, the two most common heme cofactors in nature.

bioinorganic chemistry. For example, cytochrome P450 was identified in more than 18.000 distinct proteins [38]. The number of known heme proteins has been growing ever faster due to advances in genomic analysis and hemoglobins today are found in all kingdoms of life (bacteria, plants, animals) [17, 39]. The investigation of the reaction cycles has remained a challenge. In many cases, conflicting experimental results and interpretation of theoretical calculations in terms of the oxidation and spin state of iron in the heme systems is at the basis of mechanistic controversy. This work investigates the potential of combining advanced X-ray spectroscopy with various quantum chemical theories to yield the local electronic configuration of iron in metalloproteins. For this purpose, hemoglobin and myoglobin were chosen as established benchmark proteins.

1.1 Hemoglobin and myoglobin

The function of hemoglobin (HB), as oxygen-transport protein in vertebrate blood, and myoglobin (MB), as oxygen-storage protein in vertebrate muscle tissue, has been investigated for more than a century. The first isolation of HB was reported in 1840 [42] and Faraday first observed the diamagnetic behavior of arterial blood five years later [43]. Faraday concluded that humans are diamagnetic "*for all the substances of which he is formed, including the blood, possess this property*". Both proteins can be present in their deoxygenated (**deoxy**), oxygenated (**oxy**), carbon monoxide inhibited (**carboxy**) or ferric (**met**) forms. The monomeric MB binds a single heme B cofactor, while tetrameric HB binds four (see Figure 1-3). Already early

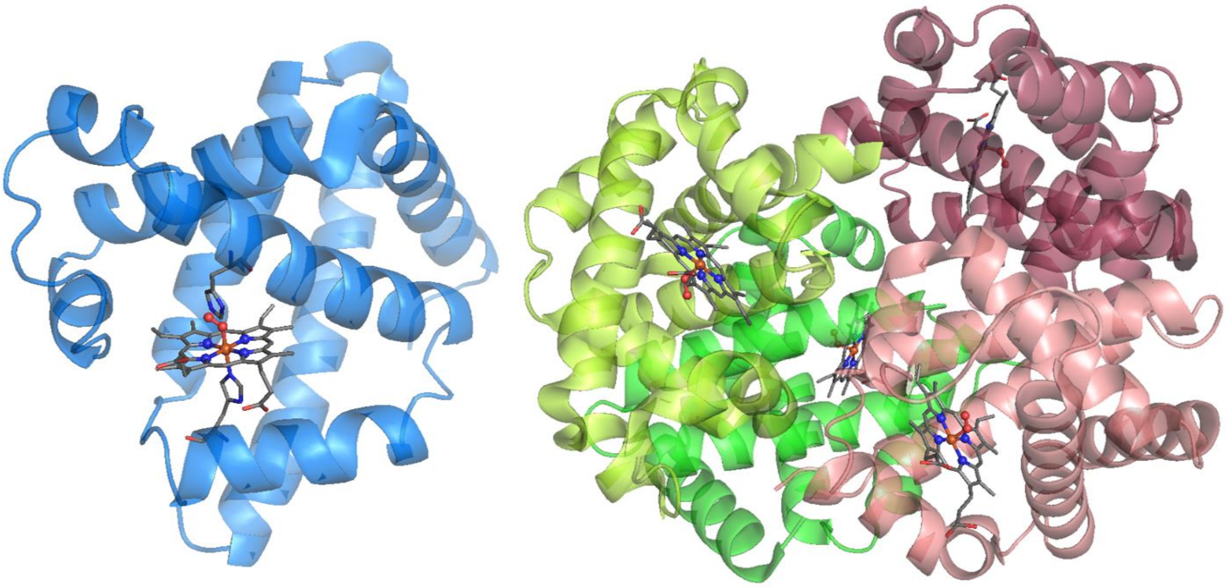


Figure 1-3: Structure from X-ray crystallography of oxygenated monomeric myoglobin (left, MB, PDB entry 1A6M) and tetrameric hemoglobin (right, HB, 2DN1) showing the heme B cofactors (ball-and-sticks) and protein subunits (cartoon). The axial and distal histidine amino acid are shown in the MB structure. The four globin subunits of HB are colored individually.

experiments in the 1920s showed that the four heme cofactors in HB bind O_2 in a cooperative manner, where every oxygen molecule could bind easier than the previous, which is essential for O_2 transport [44]. This implies a remarkable long-range interaction between the heme cofactors due to structural changes of the protein subunits. Pauling and Coryell discovered in 1936, by measuring the magnetic susceptibility of various HB species in an external magnetic field, that **deoxy** HB showed paramagnetic behavior with four unpaired electrons ($S = 4/2$), while **oxy** and **carboxy** HB were diamagnetic with no unpaired electrons ($S = 0$) [45, 46], which was later confirmed by other authors [47]. The first protein structure of HB from XRD revealed the quaternary structure shown in Figure 1-3, where **oxy** HB was in a relaxed (R) and **deoxy** HB in a tense (T) state [48], thus demonstrating the long range effect of the local change in electronic configuration. Perutz and Kendrew were awarded the Nobel Prize in 1962[1] for publishing the XRD structures of HB and MB, which also were the first protein XRD structures ever [48-50]. Higher resolution XRD structures later revealed that (i) O_2 binds in a bent end-on manner to iron (see Figure 1-4) and (ii) the iron atom is out-of-plane in **deoxy** and more in-plane in **oxy** HB [51, 52]. This connected the previously discovered long-range cooperative dioxygen binding of HB with the local change in spin state upon binding of O_2 through the

---

[1] "The Nobel Prize in Chemistry 1962". Nobelprize.org. Nobel Media AB 2014. Web. 10 Jan 2018. <http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1962/>

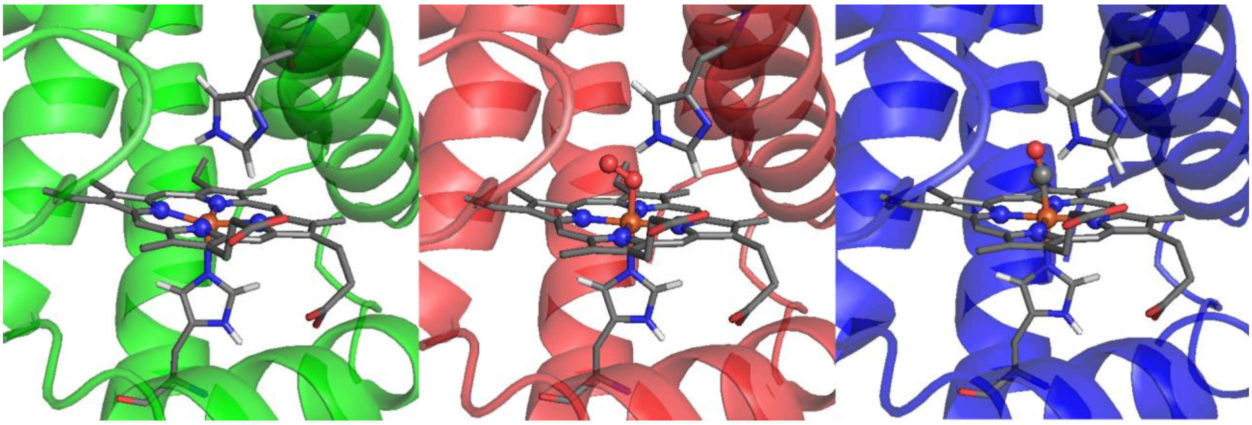


Figure 1-4: Heme cofactor in myoglobin in the ***deoxy*** *(green cartoon),* ***oxy*** *(red) and* ***carboxy*** *(blue) state from high resolution (1.15 Å) XRD structures (PDB entries* ***deoxy****: 1A6N,* ***oxy****: 1A6M,* ***carboxy****: 1A6G) [2]. Heme B cofactor, axial/distal histidine and O_2/CO ligand are shown as ball-and-sticks (iron: orange, nitrogen: blue, oxygen: red, carbon: grey). The axial and distal histidine are shown in protonation states (hydrogen: white) derived from neutron scattering and nuclear magnetic resonance experiments [14, 15]. Note the out-of-plane versus in-plane iron atom in* ***deoxy*** *versus* ***oxy/carboxy*** *and the bent end-on orientation of the O_2 ligand in* ***oxy****.*

connection of the axial histidine residue [53]. The protonation of the axial and distal histidine residues were revealed via NMR and neutron diffraction experiments [15, 47]. The geometry and protonation of the heme cofactor in HB and MB in each of the three states (shown in Figure 1-4) were therefore experimentally established, but the electronic configuration of the Fe – O_2 bond has given rise to ongoing controversy [17, 39, 54].

Magnetic susceptibility and electron paramagnetic resonance (EPR) experiments showed that a quintet Fe(II) and a triplet O_2 would form a singlet Fe – O_2 thus, six electrons pair upon O_2 binding. Yet, the two unpaired electrons on O_2 allowed for several possible configurations and the three most prominent models are shown in Figure 1-5 [55]. The first model was proposed by Pauling and Coryell in their publication in 1936 (Figure 1-5 left). They interpreted the equal magnetic susceptibility of **carboxy** and **oxy** HB by assigning the same electronic configuration represented as two resonant Lewis structures [45]. Both the O_2 and CO molecules act as innocent ligands and the dioxygen molecule bind via a dative σ bond to the iron atom [45, 55, 56]. Consequently, iron would retain a d^6 configuration and a singlet state.

A different model was suggested by Weiss in 1964 (Figure 1-5 center), who argued that (i) molecular oxygen has been found to be a non-innocent ligand to iron [57] and (ii) optical absorption spectra of **oxy** HB are more similar to hydroxo- and cyano-HB, which both have an Fe(III) low-spin configuration. Weiss proposed that Fe(II) and O_2 form an Fe(III) – O_2^-

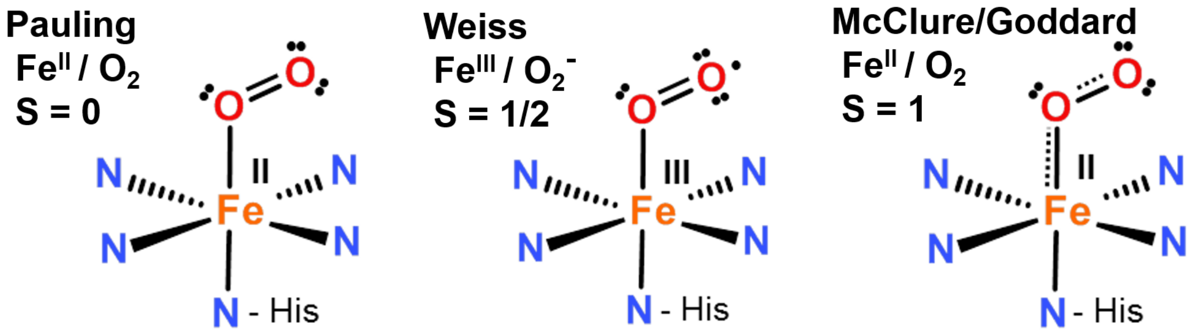


Figure 1-5: Schematic Fe – O_2 bonding in oxygenated heme proteins in the Pauling (left), Weiss (middle) and McClure/Goddard (right) models. The formal oxidation state of iron and the total spin on iron and O_2 as well as lone pairs on oxygen atoms are indicated.

complex through electron transfer from iron to the oxygen ligand. Both Fe(III) and O_2^-, in a low-spin doublet state, would form a singlet state via antiferromagnetic coupling (afc). Additionally, the optical absorption spectrum of **carboxy** HB also matches hydroxo- and cyano-HB. Therefore, Weiss argued that CO also binds as a non-innocent ligand [57].

McClure suggested in 1960 that O_2 remains in its triplet state and therefore iron converts from a high-spin/quintet to an intermediate-spin/triplet configuration (Figure 1-5 right). The four unpaired spins would then couple to a singlet state [58]. He based this model (i) on the order of ligand-field strengths, where O_2 could potentially reside on the border between high-spin and low-spin octahedral fields, and (ii) the rapid and reversible absorption, which made it unlikely that O_2 "*undergoes a great deal of electronic rearrangement*". This model did not receive much attention until Goddard and Olafson in 1975 reported ab-initio multiconfigurational calculations [59, 60], which predicted two single-occupied iron orbitals that paired with two single-occupied O_2 orbitals. Goddard and Olafson described their configuration as an 'ozone model' with a four-electrons/three-center π bonding.

Experimental investigations on the electronic configuration of **oxy** has remained inconclusive. Resonance Raman spectra showed comparable oxidation and spin state marker bands for **oxy**, **carboxy** and low-spin ferric hemoproteins, e.g. **met** [61, 62]. Mössbauer spectroscopy showed similar isomer shifts for **carboxy** and **oxy**, which were more in line with a series of Fe(III) hemoglobin complexes in low-spin and high-spin configurations [63-65]. However, the quadrupole splitting of **deoxy**, **carboxy** and **oxy** was similar to Fe(II) model complexes and suggested unpaired electrons in Fe(3d) orbitals for **oxy**. Therefore, resonance Raman and Mössbauer isomer shifts favored the Weiss model while the quadrupole splitting suggested a McClure/Goddard configuration. X-ray emission spectroscopy indicated an increase in unpaired Fe(3d) electrons from **carboxy** to **oxy**, which would argue against Pauling and in favor of Weiss or McClure/Goddard [19].

The theoretical results by Goddard and Olafson were followed by a large body of quantum chemical calculations published over the past 40 years [55, 66-77]. Hartree-Fock (HF) calculations did not predict a singlet **oxy** ground-state [55]. Thus, computational methods had to include electron correlation to describe the oxygen-bound heme spin state. Suitable calculations therefore are single-determinantal density functional theory (DFT) and multiconfigurational post-HF methods. Assigning computational results to one of the three proposed models was most often done by calculating the charge on O_2 and analyzing the orbital composition. DFT calculations were interpreted in favor of the Weiss model and post-HF methods apparently supported all three models [55]. Evaluating these differing computational results had to be made based on discriminative experimental data. However, the problem when comparing post-HF methods and DFT is that few common experimental quantities are accessible for heme complexes from both theories. These include (i) optimized geometries, (ii) EPR spectra and (iii) relative energies. Post-HF and DFT correctly predict the bent end-on Fe – O_2 coordination and a singlet ground-state, which is EPR-silent. The O_2 dissociation energy from complete active space with second-order perturbation calculations (CASPT2) is in good agreement with calorimetry measurements [77, 78]. DFT and CASPT2, on the other hand, correctly predicted the Mössbauer parameters [55, 76]. Without clear-cut evidence, which computational method is in best agreement with the experimental data, the discussion had turned towards the question if a single-determinantal method like DFT is capable of describing an apparently multiconfigurational system at all.

Conclusion

The central problem in determining which of the various models (Pauling, Weiss, McClure/Goddard) provides a suitable description of the Fe – O_2 bonding is the complex interplay between the model framework and the interpretation and correlation of experimental and theoretical results. Experimentalists often had interpreted their results in terms of structures at the edge of the parameter span with well-defined oxidation states. Theoreticians had developed complex frameworks but relation of theoretical results to the experimental benchmarks is often challenging. In the present work, this dilemma was approached by comparison of predictions from both branches of theory (post-HF and DFT) with advanced X-ray spectroscopy data in order to develop and apply a framework for reliable determination of the electronic configuration of iron in heme proteins.

1.2 Scope of this thesis

Synchrotron based X-ray absorption and emission spectroscopy (XAS/XES) facilitates investigation of geometric and electronic properties of transition metal compounds. Technical improvement of third-generation synchrotron facilities provides high photon flux, which permits XAS/XES experiments even on samples with low metal concentrations such as proteins with metal cofactors. However, metal centers in proteins are often sensitive to X-ray exposure and show rapid reduction of metal ions and ligand degradation. Furthermore, interpretation of advanced XAS/XES experiments is challenging and quantitative analysis of data requires correlation with results from quantum chemical simulations. The primary aim of this thesis was to develop and combine approaches for combined advanced X-ray spectroscopy and quantum chemistry and their implementation and application to relevant metalloproteins.

These endeavors included the following main tasks:

I. Development of time-resolved energy-sampling XAS/XES approaches to study weak X-ray absorption/emission feature of radiation sensitive materials
II. Quantitative evaluation of conventional versus energy-sampling XAS/XES methods in terms of efficiency and signal quality and implementation of software tools
III. Application of advanced X-ray and QC methods to selected metalloproteins with a focus on heme proteins
IV. Collection of complete XAS/XES data sets for hemoglobin and myoglobin (HB/MB) in wide range from cryogenic to near-ambient temperatures and data analysis
V. Determination of key structural and electronic parameters from the X-ray data
VI. Implementation in the laboratory, documentation and application of advanced quantum chemistry methods (e.g. TD-DFT, MP2 and CASSCF) for spectral simulation, quantitative correlation with XAS/XES data and evaluation of various metal systems
VII. Addressing the long-standing debate surrounding the iron-oxygen interaction in O_2 transporting hemoproteins with the aim to provide a consistent interpretation.

The following section describe the methodological progress and scientific outcome of these accomplishments.

2 Experimental methods

Synchrotron based spectroscopy was the main experimental method that was applied in this work. Metal centers in proteins and model complexes were investigated with X-ray absorption and emission spectroscopy (XAS/XES). This section will provide an introduction into XAS and XES and possible applications when investigating biological or synthetic metal complexes.

Core-hole electron or K-edge spectroscopy is element specific as energies (E) for excitation of 1s electrons scale approximately with the atomic number (Z) ~ $Z^{2.16}$ [79]. This results in large energy differences between 1s electron binding energies for first-row transition metals (Mn: 6539 eV, Fe: 7112 eV, Co 7709 eV) [80]. K-edge spectroscopy probes occupied and unoccupied states centered on the absorber atom. The transition probability (spectral intensity) and energy difference between states (spectral energy) are affected by the first-sphere ligand environment. Additionally, X-ray spectroscopy can be applied to samples in any aggregate state, at cryogenic temperatures or at room temperature. Thus, XAS and XES experiments are well suited to investigate the molecular environment and electronic structure of metal sites in proteins.

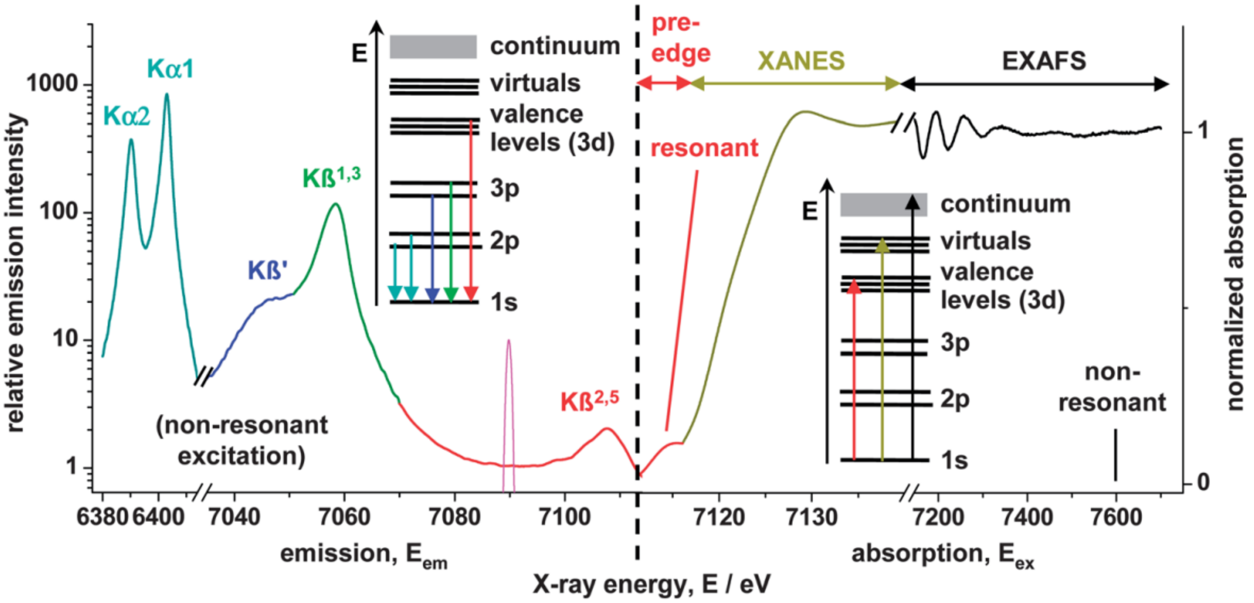


Figure 2-1: Schematic X-ray absorption (right) and emission (left; note the logarithmic y-axis) spectra of an iron complex on a common energy scale. The insets exemplify the electronic transitions in an atomic level picture, which correspond to the absorption and emission spectra. The magenta peak shows the Kβ detection energy resolution of ~ 1 eV. Note that incident energies for non-resonant excitation were set to 7800 eV in this work. Reproduced from Ref. Lambertz, et al. [26] with permission from The Royal Society of Chemistry.

An overview of the spectral energy range and nomenclature used in this work is shown in Figure 2-1. XAS corresponds to transitions into ‘unoccupied states’. Note that this represents a common generalization. A more accurate description would be a resonantly excited transition from a non-degenerate ground-state into an excited state by creation of a metal core-hole. The excited state corresponds to a multiplet of possible electronic configurations, which in some cases can be a single electronic configuration and therefore a non-degenerate state. The difference in spin-density between these two non-degenerate states takes the shape of an unoccupied molecular orbital in ligand field theory. It is important to keep this in mind when talking about absorption and emission spectra and ‘occupied’ or ‘unoccupied’ states. X-ray absorption in the pre-edge region corresponds to resonant excitations into unoccupied localized metal(d) or ligand(s,p) molecular orbitals (MOs) and thus is sensitive to oxidation and spin state as well as ligand type and coordination [23, 25, 81-83]. The X-ray absorption near edge structure (XANES) probes unoccupied MOs with delocalized metal(p) character and is very sensitive to the metal oxidation state and coordination and is therefore, e.g., used to monitor changes due to radiation damage by X-ray exposure [25, 26, 81-85]. Extended X-ray absorption fine structure (EXAFS) provides accurate metal – ligand bond lengths, ligand types and coordination numbers [25, 26, 82, 83]. XES gives insight into occupied molecular orbitals. The Kα emission (2p → 1s) is affected by metal oxidation and spin states [23, 86] but is more often used for XAS detection. A more accurate method to study oxidation and spin states is the Kβ main-line emission (Kβ’ and $K\beta_{1,3}$) corresponding to metal(3p) → 1s transitions [23, 25, 26, 81, 87]. The Kβ satellite line (Kβ’’ and $K\beta_{2,5}$) can give insight into the ligand environment surrounding the metal as well as the nature of chemical bonds [33, 83]. These different spectroscopic features and their applications are presented in more detail in the following sections.

2.1 X-ray absorption spectroscopy

XAS measurements probe unoccupied valence molecular orbitals as a function of incident beam energy (E). The absorption coefficient (μ) can be detected directly through measurement of the beam intensity before (I_0) and after (I) a sample of certain thickness (x).

$$-\mu(E)x = \ln\frac{I}{I_0} \tag{2-1}$$

XAS in transmission mode requires samples with metal concentrations > 10 mM (i.e. non-protein samples) in order to detect a sufficient change in beam intensity. An alternative is to collect absorption spectra using either (i) broad-band Kα ('total fluorescence') or (ii) narrow-band Kβ detection [88, 89]. Combining XAS and XES represents a resonant excitation into an intermediate state followed by electronic decay into a final state, so called Resonant Inelastic X-ray scattering (RIXS) [90, 91]. In case (i), all unoccupied levels in the valence shell multiplet are probed and respective core-to-valence (ctv) spectra are similar to transmission-detected spectra. In case (ii), a fraction of the multiplet is excited and the ctv spectrum represents a transect through the multiplet of unoccupied states. The latter approach allows to focus on different emission features and e.g. collection of spin-polarized absorption spectra.

XAS at the K-edge is divided into three regions (see Figure 2-1). In terms of excitation energy, the (weak) pre-edge feature, is followed by the XANES and EXAFS regions. Strict distinction between the last two regions is generally neither possible nor necessary. XANES spans the energy range where absorption is mainly associated with the absorbing atom itself. EXAFS refers to the energy region where a photoelectron is created (excitation into the continuum), which is used to perform experiments employing the EXAFS equations.

The pre-edge region

K-edge absorption spectra of first-row transition metals typically show a weak absorption feature (~ 1 – 10 % of the K-edge rise) a few eV prior to the K-edge, the so-called pre-edge. This feature was assigned to metal(1s) → metal(3d) transitions as it was visible for all first-row transition metals except Zn(II), which has a $3d^{10}$ configuration [92]. These resonant excitations of metal core-electrons into unoccupied states can be used to determine oxidation and spin states of metal complexes, which is of special interest in investigations of e.g. bond activation through active sites. The number, position and energy splitting of pre-edge features shown in Figure 2-2 can be interpreted according to a molecular orbital scheme in a local ligand field environment. The two absorption peaks of a tetrahedral high-spin Fe(II) compound (red line in A) correspond to transitions into unoccupied t_{2g}/e_g orbitals, which upshift by ~1.5 eV for the tetrahedral high-spin Fe(III) compound (blue line in A). The two octahedral complexes show the same upshift from Fe(II) to Fe(III) but absorption is too small to identify individual features. This is due to symmetry allowed p-d orbital mixing. The pre-edge peaks of the two octahedral high- and low-spin Fe(II) compounds in B correspond to transitions into unoccupied t_{2g}/e_g orbitals (high-spin/red line) or e_g orbitals (low-spin/blue line). However, the first peaks in B show an energy upshift of ~ 1 eV despite no formal change in oxidation state. Thus, the position of pre-edge features can be related to the metal electronic configuration, even without quantum chemical

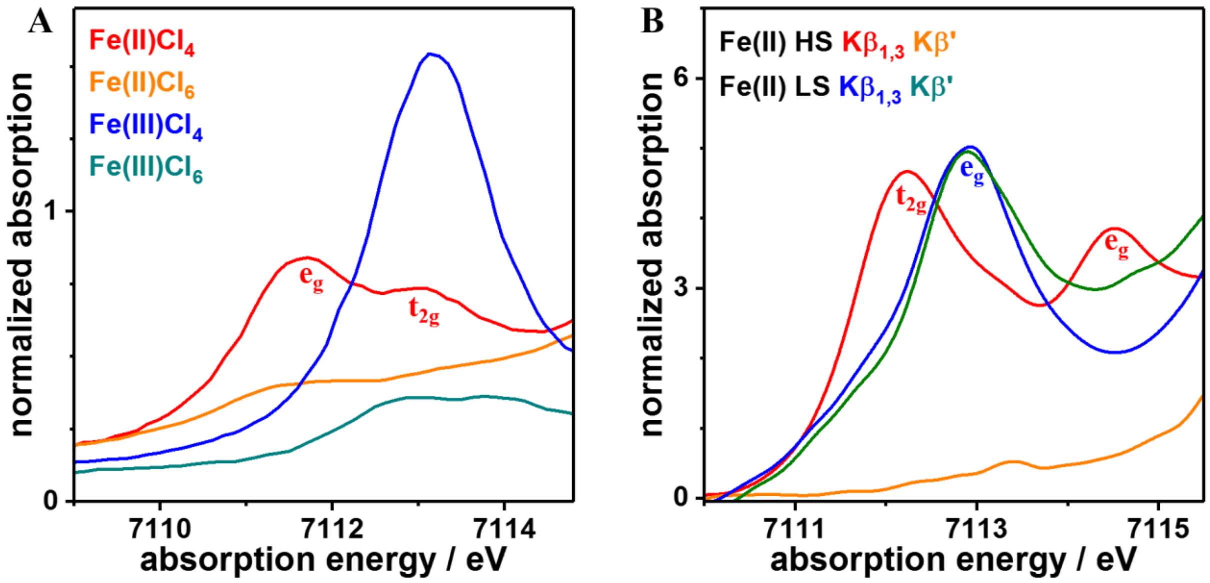


*Figure 2-2: Experimental data for pre-edge absorption of iron compounds in different coordination, oxidation and spin states. (**A**) Shown iron compounds are in a high-spin state (Fe(II): S = 4/2; Fe(III): S = 5/2) in a tetrahedral (Cl_4 ligand) or octahedral (Cl_6 ligand) coordination environment. Reprinted (adapted) with permission from Ref. DeBeer George, et al. [24]. Copyright 2008 American Chemical Society. (**B**) Narrow-band Kβ-detected spectra for $K\beta_{1,3}$ (7057 eV) or Kβ' (7045 eV) emission energies of an Fe(II) ($Fe^{II}N_6$; $Fe(phen)_2(NCS)_2$ with phen = 1,10-phenanthroline) complex. The complex is in a low-spin configuration < 115 K and high-spin configuration > 195 K. The positions of t_{2g} and e_g orbital sets for a high-spin/low-spin tetrahedral (A) and octahedral (B) geometry are indicated. Reprinted (adapted) with permission from Ref. Mebs, et al. [25]. Copyright 2015 American Chemical Society.*

calculations, but complications can arise when comparing compounds with unknown spin states.

Pre-edges can vary significantly in intensity due to coordination symmetry, with small absorption for octahedral and large absorption for tetrahedral geometries (see Figure 2-2A). The metal centered 1s → 3d excitations are formally dipole forbidden transitions ($\Delta l = \pm 1$) and quadrupole allowed ($\Delta l = 0, \pm 2$) in a geometry with inversion symmetry, e.g. a perfect octahedral environment with the absorber metal atom at the inversion center. However, breaking inversion symmetry results in metal(3d/4p) orbital mixing and the consequence are significant dipole transition probabilities [81]. Quadrupole transition oscillator strengths are about ~ 1 % the size of dipole transitions but the metal(3d/4p) overlap is orders of magnitude smaller than the overlap of metal(1s/3d) orbitals. Both transition mechanisms can accordingly be important to explain differences in intensities and the amount of orbital mixing depends on the coordination symmetry of the metal site. Consequently, tetrahedral geometries result in larger orbital overlap and thus, larger pre-edge absorption. This is the reason for the increase in pre-edge absorption from octahedral to tetrahedral complexes shown in Figure 2-2A. Distorting symmetry, e.g. through shorter bond lengths due to different ligands, also increases metal(3d/4p) orbital mixing. Thus, a metal in a distorted octahedral symmetry shows larger pre-edge intensities than in a perfect octahedron. The effect of symmetry on pre-edge intensity leads to the general rule that absorption decreases with increasing coordination number [20, 35]. Electronic excitations involving metal-to-(non-metal)ligand charge transfer (CT) are dipole allowed s → p transitions. These occur in general at a higher energy than ctv transitions without CT due to inverse distance dependency [35].

The spin state dependent features of the Kβ main-line emission (3p → 1s, see Figure 2-5A and following section on Kβ main-line emission) can be used to detect spin polarized pre-edges [87]. X-ray fluorescence emission detectors with high energy-resolution ($\Delta E < 2$ eV) can focus on $K\beta_{1,3}$ emission (Fe: ~ 7060 eV) or Kβ' emission (Fe: ~ 7045 eV). These features are split due to strong coupling between the metal(3d) and metal(3p) orbitals. The Kβ' feature intensity is proportional to resonant excitations of core electrons with spin parallel to unpaired metal(3d) electrons (spin-up) while the $K\beta_{1,3}$ peak corresponds to excitations with antiparallel spin (spin-down). Consequently, an increase in Kβ' or $K\beta_{1,3}$ intensity stems from unoccupied spin-up or spin-down molecular orbitals. Thus, comparing $K\beta_{1,3}$ and Kβ' detected pre-edge spectra shows if increasing the number of metal(3d) electrons is possible by spin-up or spin-down excitations. Two examples are shown in Figure 2-2B. The Fe(III) compound (d^5 configuration) shows two peaks for $K\beta_{1,3}$ detected spectra (red line) while the Kβ' detected pre-edge shows no emission (orange line). The net spin could therefore only be decreased, which is in line with a d^5 high-spin configuration and the ctv transitions correspond to t_{2g} and e_g orbitals. The Fe(II) sample (d^6 configuration) shows a single peak for $K\beta_{1,3}$ and Kβ' detected spectra. Thus, there is no difference in spin-up or spin-down excitation, which is consistent with a low-spin configuration with unoccupied e_g orbitals. These two examples correspond to the two extremes of spectra for first row transition metals and increase/decrease of unpaired 3d electrons gradually increases/decreases the difference between $K\beta_{1,3}$ and Kβ' detected pre-edge spectra.

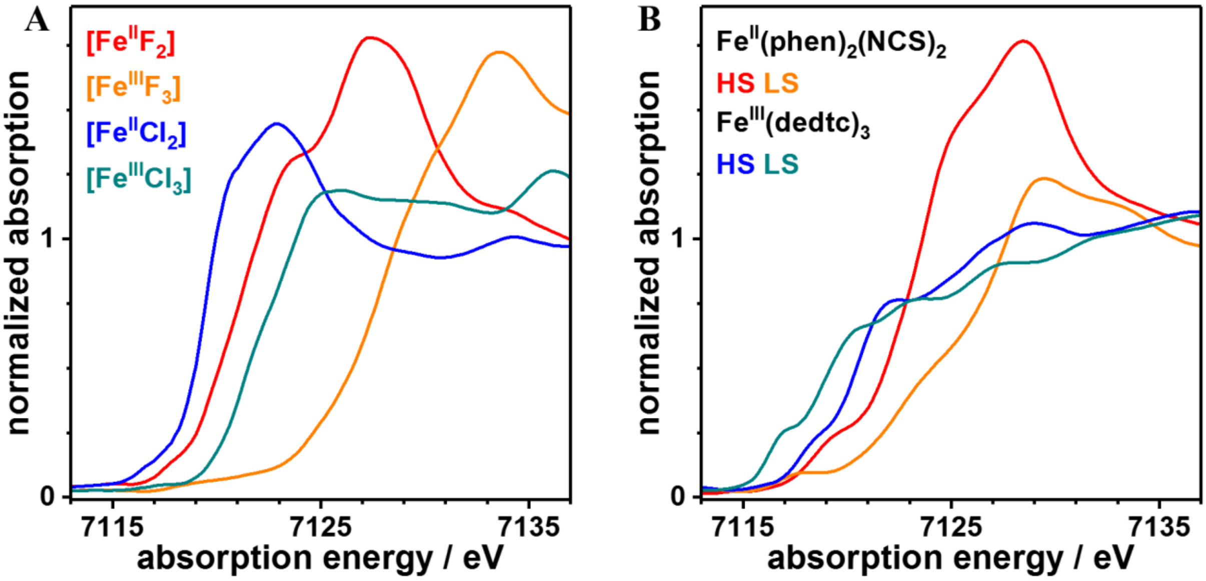


*Figure 2-3: Influence of oxidation state, ligand environment and spin state on shape and position of K-edge absorption (XANES). (**A**) XANES spectra of crystalline octahedral high-spin Fe(II) and Fe(III) complexes. The spectra were detected in transmission with a Si(220) double-crystal monochromator with a resolution of ~ 1.4 eV at temperatures < 55 K. Reprinted (adapted) with permission from Ref. Westre, et al. [20] Copyright 1997 American Chemical Society. (**B**) XANES spectra of an octahedral Fe(II)N$_6$ (Fe(phen)$_2$(NCS)$_2$ see Figure 2-3) and Fe(III)S$_6$ (Fe(dedtc)$_3$ with dedtc = N,N'-diethyldithiocarbamate) complex in high-spin (> 300 K) and low-spin (< 80 K) configuration. The average iron-ligand distances increases from low-spin to high-spin by ~ 0.2 Å. Reprinted (adapted) with permission from Ref. Mebs, et al. [25] Copyright 2015 American Chemical Society.*

The XANES region

While excitations in the pre-edge mostly stem from formally dipole forbidden 1s → 3d transitions, XANES features correspond to dipole allowed 1s → 4p transitions. The XANES region is, for transition metals, right above the fermi energy. Thus, the created photoelectron has low kinetic energy and a large mean free path. The electron can therefore interact with atoms up to the second coordination sphere. Hence, the excitations are very delocalized and mixed with ligand valence states. These excited states can involve metal-to-metal transitions, metal-to-ligand charge transfer (MLCT), ligand-to-metal charge transfer (LMCT) or a combination of these [81, 93, 94]. Quantum chemistry simulations involving only single-electron excitation can consequently be inadequate to calculate excited states in the XANES region. However, even if single-excitations would be sufficient, the number of necessary excited states would still be very large. Multiple scattering approaches (e.g. FEFF [95]) are capable to predict trends in K-edge energies but interpretation is a challenge as multiple scattering results from geometric structure. Consequently, the electronic configuration is only indirectly reflected. Nonetheless, XANES represents an excellent method for fingerprint studies e.g. by comparison to known model compounds.

The K-edge position is related to the ionization potential, which depends on the valence charge density and screening effects [94]. Changes in charge density are in general primarily attributed to change in formal oxidation state. Yet, the oxidation state is not the only factor affecting charge density. Additional influences are bond lengths, bond types (covalent versus ionic), type and number of coordinated ligands and the coordination geometry itself. Many of these factors are interdependent themselves and a clear separation is therefore difficult. The effect of the oxidation state can be seen in Figure 2-3A, were the increase in formal oxidation state causes an upshift in the edge. However, the absolute upshift varies for different metal ligands. Harder ligands (e.g. fluoride and oxygen) show a larger upshift than softer ligands (e.g. chloride or sulfur). K-edge shifts can also be affected by a change in spin state as can be seen in Figure 2-3B. The XANES spectra of the Fe(II) complex (red and orange lines) show a downshift upon change from low-spin to high-spin. On the other hand, the Fe(III) complex (blue and green lines) has an upshift for the same transition.

The EXAFS region

X-ray absorption processes discussed so far corresponded to resonant excitation of core electrons into unoccupied valence orbitals. At excitation energies exceeding any discrete excited states starts the spectral region of extended X-ray absorption fine structure (EXAFS) [96]. EXAFS is a versatile technique, which allows for element specific investigation of local geometries. It can be applied to solid, liquid or gaseous samples as well as compounds in solution. Consequently, EXAFS complements other methods such as crystallography (XRD). For EXAFS, the X-ray photon energy (E) exceeds the vacuum-energy of a core electron (E_0) and a free electron with energy $\Delta E = E - E_0$ is created. The initial state (Ψ_i) is the ground-state of the compound ($1s^2C^0$) and the final state (Ψ_f) is the product of a core hole and a free electron wave ($1s^1C^1$). This electron can be described as a free electron wave propagating spherically from the absorber with a wave vector (k) (see Figure 2-4A):

$$k(E) = \sqrt{\frac{2m_e(E - E_0)}{\hbar^2}} \tag{2-2}$$

using the electron mass (m_e) and the reduced Planck constant ($\hbar$). The final state is a superposition of the free electron wave centered on the absorbing atom and backscattered electron waves from surrounding atoms. Thus, constructive or destructive interference modulates the final state, depending on the electrons energy and the type and distance of backscattering atom (see Figure 2-4A). This results in an angle-dependent amplitude modulation of the final state wavefunction, which leads to an energy dependent modulation of the absorption coefficient. The electron path can involve one scattering atom (two-leg path) or more atoms. The transition rate from initial to final state is determined by Fermi's golden rule, normalized by the density of states (ρ), and is proportional to the absorption coefficient (μ). The perturbation causing the transition is the absorbed X-ray photon which, as a first order approximation, can be expressed as a dipole transition [97].

$$\mu \propto \left|\left\langle \Psi_f \middle| \hat{\epsilon}\, \vec{r}\, e^{i\vec{k}\vec{r}} \middle| \Psi_i \right\rangle\right|^2 \rho\left(E_f - E_i - \hbar\omega\right) \approx \left|\left\langle \Psi_f \middle| \hat{\epsilon}\, \vec{r} \middle| \Psi_i \right\rangle\right|^2 \rho\left(E_f - E_i - \hbar\omega\right) \tag{2-3}$$

The X-ray photon is defined by the electric polarization vector ($\hat{\epsilon}$), its momentum ($\hbar\vec{k}$) and energy ($\hbar\omega$).

The measured absorption coefficient is shown in Figure 2-4B. The K-edge absorption (μ) sits on a background of lower-energy photoelectron processes, e.g. L-Edge absorption, Rayleigh (elastic) and Compton (inelastic) scattering (blue line in Figure 2-4B). The absorption above the K-edge can be described as a background absorption (μ_0, green line in Figure 2-4B) modulated by a correction function (χ).

$$\mu(E) = \mu_0(E)\left(1 + \chi(E)\right) \Rightarrow \chi = \frac{\mu}{\mu_0} - 1 \tag{2-4}$$

This fractional change in absorption is the fine structure in EXAFS (see Figure 2-4C) and is usually shown in k-space and weighted with k^3 (Figure 2-4D). The former allows for a Fourier transformation (see Figure 2-4D) while the weighting compensates for inelastic losses. The peaks in the Fourier transform reflect the interatomic distances. However, the distance is about 0.3 – 0.4 Å shorter than the actual bond length as the fine structure oscillations are phase shifted (Φ). Bond lengths, coordination numbers and scattering atom types are determined by fitting of the fine structure oscillations in k-space or energy-space. The experimental oscillations are a sum of individual oscillations, each representing a scattering path (r_j) between scattering atoms. Each, so called 'shell', has an individual phase shift (Φ_j) and amplitude (A_j). The phase shift stems from the interaction of the photoelectron with the coulomb potential of the absorbing and any backscattering atom.

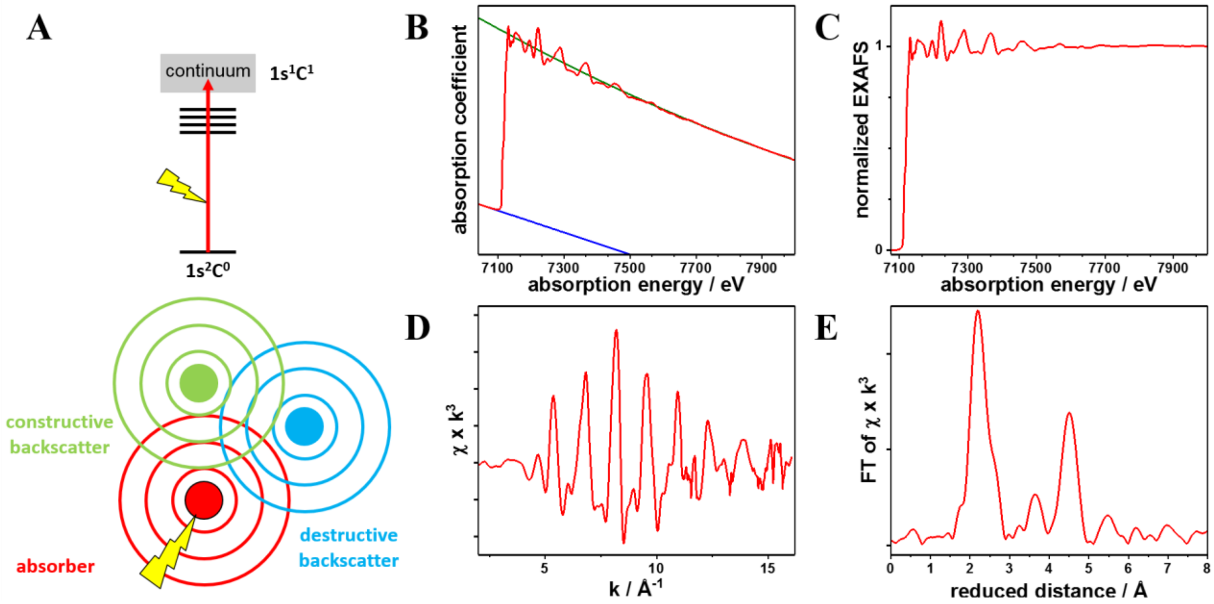


*Figure 2-4: Summary of extended X-ray absorption fine structure (EXAFS) process and data evaluation. (**A**) Processes resulting in EXAFS. The initial state is the core-electron in its ground-state, which is perturbed by the X-ray photon. This perturbation results in the final state of a free electron wave that can backscatter from nearby atoms. The backscattered waves lead to constructive or destructive interference, which affects absorption probability (see text). (**B**) X-ray absorption of an iron foil (red) with linear background (blue) and second order polynomial normalization (green). (**C**) EXAFS spectrum after background subtraction and normalization. (**D**) The extracted EXAFS oscillations (χ). The absorption energy is transformed into the wave vector of the free electron (k) and the oscillation k^3 weighted. (**E**) Fourier transform of the EXAFS oscillation.*

$$\chi(k) = S_0^2 \sum_j A_j(k, r_j)\ \sin\left(2kr_j + \phi_j(k)\right) \tag{2-5}$$

The total oscillation amplitude is scaled by an amplitude reduction factor (S_0), which is approximated as element and energy independent. Each oscillation amplitude is a function of coordination number (N_J) and type of backscattering atoms, which determines the scattering amplitude (f_j).

$$A_j(k, r_j) = N_j \frac{f_j(k)}{kr_j^2} e^{-\frac{2r_j}{\lambda(k)}} e^{-\sigma_j^2 k^2} \tag{2-6}$$

The amplitude is reduced by (i) the $1/r^2$ dependence of the electron wavefunctions intensity, (ii) coherence loss due to the electron mean free path (λ) and (iii) the Debye-Waller factor (σ_j). The latter models thermal motion and represents the mean square variation of the electron path. The phase functions, scattering amplitudes and the mean free path can either be estimated via model complexes or calculated e.g. with FEFF [95]. This rather simple theoretical approach leads to an impressive experimental accuracy when determining atomic distances (± 0.02 Å), coordination numbers (± 20 %) and ligand types (Z ≈ ± 1). However, the quality of the initial guess is important and e.g. XRD structures are helpful. The complexity of this method arises from the fact that the oscillations represent an average of all absorbers. Thus, studying complexes with multiple absorbing atoms poses a considerable challenge.

Conclusion

X-ray absorption spectroscopy (XAS) is a powerful tool to provide structural and electronic information on a local atomic level by detection of (i) core-electron transitions into unoccupied valence states (pre-edge, X-ray absorption near edge structure; XANES) or (ii) core-electron promotion into the continuum influenced by surrounding ligands (extended X-ray absorption fine structure; EXAFS). XAS was used by us to investigate metal cofactors in many metalloproteins, for example hemoproteins (see, e.g. publications I, IV, V); hydrogenases (VII) and oxidases (IX) as well as synthetic compounds (II, VI, VIII) [83, 98-105].

The pre-edge absorption can provide information on changes in coordination, geometric distortion, metal oxidation and spin state as well as the number and type of metal ligands. Pre-edge data in combination with density functional theory (DFT) or time-dependent DFT (TD-DFT) spectral simulations allows to determine and analyze key electronic parameters (see section 3.1). The weak absorption of the formally dipole-forbidden core-to-valence transitions requires suitable synchrotron facilities with high X-ray flux, high energy resolution and superior signal-to-noise ratio (for example ID26 at the ESRF, see section 2.3). XANES represents an excellent method to identify molecular features and redox levels of metal complexes, by comparison to known model complexes, and to monitor changes e.g. due to ligand activation or X-ray induced radiation damage. However, conclusions on the electronic structure are only reliable for homogeneous sample series with regard to spin state, ligand geometry and covalency. Comparison of XANES spectra detected under high and low X-ray flux conditions facilitates monitoring of radiation damage effects. EXAFS provides atomic distances at high precision (± 0.02 Å), coordination numbers (± 20 %) and ligand types (Z ≈ ± 1) of metal-specific absorption sites, which is a powerful method to determine and refine key structural parameters of metal cofactors.

All these methods were here employed to characterize a variety of metalloproteins and synthetic materials, with a focus on heme proteins.

2.2 X-ray emission spectroscopy

XES provides a way to investigate occupied molecular orbitals by analyzing the emission energy (E_{em}) dependent fluorescence after excitation. The majority of XES measurements involve non-resonant excitation, with excitation energies (E_{ex}) far above the K-edge (> 500 eV). Emission spectra of first-row transition metals after non-resonant excitation show (i) weak spectral features close to the Fermi energy (satellite line), which sit on the end of (ii) another major transition (main-line), which is two orders of magnitude stronger (see Figure 2-1). Satellite line emission corresponds to transitions from valence levels into the core-hole. These valence levels relate to the metal-ligand bonds and thus contain information about the ligand environment [26, 33, 81]. Main-line emission corresponds to dipole allowed metal core-to-core transitions. The shape is primarily influenced by metal 3p – 3d exchange and thus the metal(3d) configuration [25, 36, 81, 87]. Additional insight into the metal(3d) configuration can be gained by main-line emission after resonant excitation of a core electron into an unoccupied 3d orbital.

The Kβ satellite line

X-ray emission spectra show weak features, the Kβ satellite line (see Figure 2-1), below the Fermi energy. First, in order of increasing emission energy, occur the Kβ'' features, second the $K\beta_{2,5}$ peak and, in some cases, KLβ features, which correspond to multiple electron excitations [23, 27]. These valence-to-core (vtc) transitions are affected primarily be the metal-ligand environment. Thus, Kβ satellite line spectra can be used as a sensitive fingerprint of the ligand environment. Its features are sensitive to ligand type, chemical bonding and atomic distance and less depended on the oxidation state. The method is used in combination with reference compounds and/or simulated vtc spectra from density functional theory (DFT, see section 3.1 below) to identify ligands. Consequently, the ligand-sensitive satellite lines can complement methods, which are less sensitive to ligand types such as EXAFS.

The vtc transition intensity depends on the amount of metal(3p) and ligand(2s/p) orbital mixing [23, 33, 81, 106]. Metal centered dipole 4p → 1s or quadrupole 3d → 1s transitions in first-row transition metal compounds are of weak intensity compared to ligand-to-metal transitions. Data collection is challenging for samples with low metal concentration (< 10 mM). The weak intensity is due to the small metal-ligand orbital overlap and large transition energy.

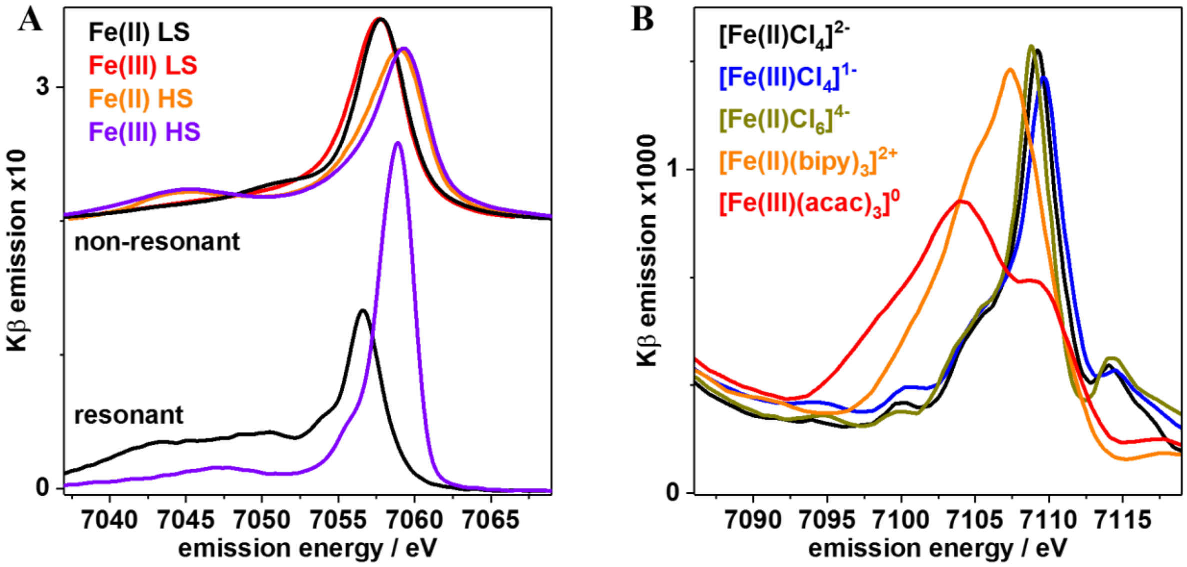


*Figure 2-5: The influence of metal spin state and ligands on main-line (A) and satellite line emission (B). (**A**) Shown are main-line emission spectra after non-resonant (E_{ex} = 7800 eV) and resonant (E_{ex} = 7113 eV) excitation of low-spin and high-spin iron compounds (Fe(II) LS: octahedral, FeN_6, $Fe(TPP)(Py)_2$; Fe(III) LS: oct, FeC_6; $[Fe(II)(CN)_6]^{-4}$; Fe(II) HS: oct, FeN_6, $Fe(phen)_2(NCS)_2$; Fe(III) HS: square pyramidal, FeN_4Cl, Fe(TPP)Cl; TPP = tetraphenylporphyrin, Py = pyridine). See text and Figure 2-6 and Figure 2-7 for further discussion. Spectra are background corrected and normalized to unity area. (**B**) Shown are satellite line emission spectra of high-spin tetrahedral (black and blue) and octahedral Fe(II/III) complexes with chloride, nitrogen (orange) or oxygen (red, acac = acetylacetonato) ligands. Total Kβ emissions are normalized to unity area. Reprinted (adapted) with permission from Ref. Lee, et al. [27]. Copyright 2010 American Chemical Society.*

The even smaller Kβ’’ feature at lower energies corresponds to L(2s) orbitals while the larger $K\beta_{2,5}$ feature at higher energies stems from, e.g., L(2p) orbitals [33, 94]. The energies of these features depends on the ligand type while the orbital overlap of metal-ligand orbitals affects transition intensities. Orbital overlap can increase due to (i) decrease in metal-ligand distance or (ii) covalent versus ionic ligand bonding [23]. The symmetry affects possible orbital mixing and consequently the number of satellite line features, where an octahedral symmetry results in fewer mixed orbitals than a tetrahedral symmetry [33]. Examples of how varying ligands affect satellite lines are shown in Figure 2-5B. The two tetrahedral iron chloride complexes show only negligible variations, despite a change in metal oxidation state. The small feature at ~7115 eV corresponds to KLβ transitions. The satellite line is also unaffected by a transition from tetra- to octahedral symmetry. However, the vtc spectra in the case of nitrogen (orange line) or oxygen

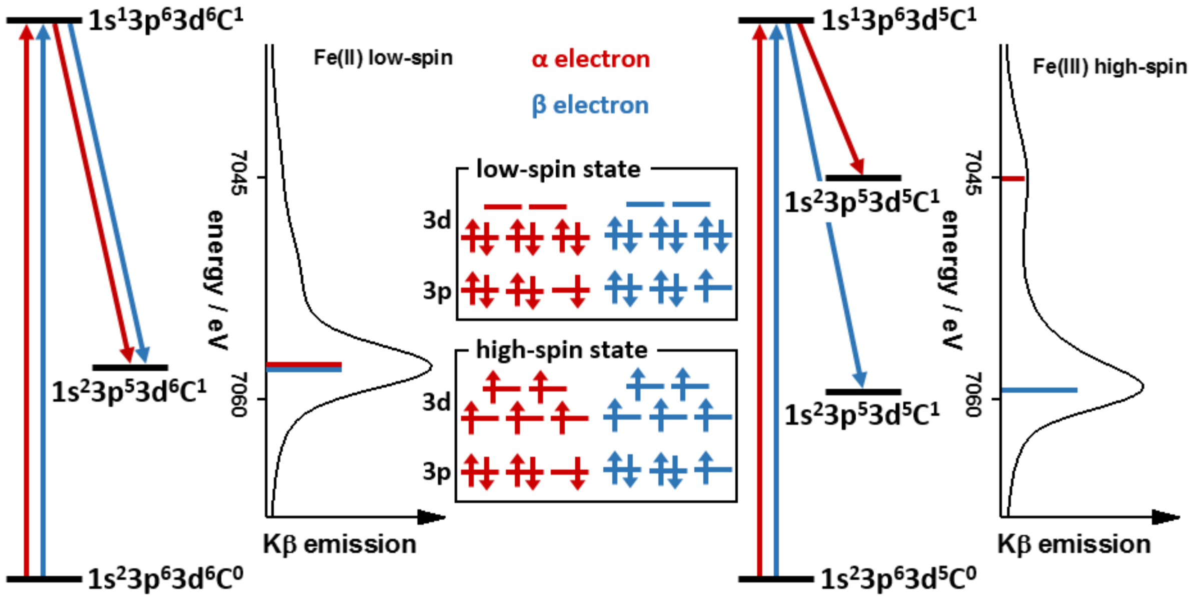


*Figure 2-6: Kβ main-line emission schematic in the case of an Fe(II) low-spin compound (left) and an Fe(III) high-spin compound (right) after non-resonant excitation. The final state electron configuration is shown in the center. (**left**) An Fe(II) compound in a low-spin ground-state absorbs a X-ray photon and emits an α (red arrow) or β (blue arrow) core electron into the continuum ($1s^2C^0 \rightarrow 1s^1C^1$). The created α/β core hole is filled by a 3p electron of equal spin, which gives rise to the Kβ$_{1,3}$ feature. (**right**) The ground and intermediate state in the case of an Fe(III) high-spin compound differ only in the unpaired $3d^5$ configuration. The final states are split by the strong exchange interaction between the net 3p and 3d electrons. The final state with an α spin 3p hole gives rise to the Kβ' feature while the β spin hole gives rise to the Kβ$_{1,3}$ feature.*

(red line) coordination show significant shifts. In conclusion, Kβ satellite lines can give insight into ligand coordination, especially when combined with calculations from e.g. DFT (see section 3.1). However, the weak intensity of satellite lines is challenging for experiments on protein samples (see section 4).

The Kβ main-line

Kβ main-line spectra are dominated by the electronic state of the metal. The core-to-core emission of first-row transition metals gives immediate insight into whether the metal(3d) electrons are in a high-spin or low-spin configuration [25, 36, 87, 94], as is shown in Figure 2-5A. The emission of high-spin complexes after non-resonant excitation shows a weak/broader peak at ~ 7045 eV, the Kβ' feature, and a strong/sharper peak at ~ 7060 eV, the Kβ$_{1,3}$ feature. The Kβ' feature decreases for the low-spin complexes and the main peak increases and downshifts by ~ 2 eV. This trend is conserved for varying site symmetry/coordination number (compare octahedral/orange versus square pyramidal/violet) and ligand types (nitrogen/black versus carbon/red). However, small changes in Kβ' intensity, after (non-)resonant excitation, can be correlated with net unpaired metal(3d) spins [23, 25]. This XES feature is of special interest for this work, as it directly addresses the question of the electronic configuration at the metal absorber. Consequently, a more detailed discussion of Kβ main-line emission after resonant or non-resonant excitation is presented.

Kβ main-line detection after non-resonant excitation

The Kβ main-line emission results from dipole-allowed radiative 3p → 1s electronic decay after previous resonant (i.e. into metal(3d) levels) or non-resonant (into the continuum) core level (1s) electron excitation (see Figure 2-6). For non-resonant excitation the sequence of states is initial state ($1s^2 3p^6 3d^n C^0$) → intermediate state ($1s^1 3p^6 3d^n C^1$) → final state ($1s^2 3p^5 3d^n C^1$). The line is split into the $K\beta_{1,3}$ and Kβ′ features due to spin-polarization of the $K\beta_{1,3}$ and Kβ′ features (3p → 1s decay) in response to strong metal(p-d) electron exchange coupling (see Figure 2-6) [107-109]. The relative line intensity and shape changes, in response to a change in the nominal number of unpaired metal d-spins, can be rationalized in a multiplet framework and are largely due to modulations of the exchange integrals between the 3p hole and the 3d electrons in the final state [36, 107, 109-112]. An increase of the Kβ′ emission intensity with respect to the $K\beta_{1,3}$ emission (for non-resonant excitation) due to coupling between an increasing number of (nominal) unpaired metal(3d) spins and the 3p hole in the final state has been shown for various first-row transition metal compounds with increasing oxidation and/or spin state [16, 23, 88, 89, 110, 111, 113-119]. However, the Kβ main-line splitting and relative intensity and energy of the Kβ′ feature is further influenced by the covalency of the metal-ligand bonding (i.e. delocalization of metal(3d) orbitals onto the ligands) [16, 36]. The latter effect has been shown to decrease the relative intensity of the Kβ′ feature significantly for example in iron compounds with sulfur instead of nitrogen ligands at Fe(III) or Fe(II) centres [16] and in compounds with increasing bonding covalency and similar Fe(III) centres [36]. Accordingly, the correlation of the number of unpaired Fe(3d) electrons to the Kβ′ intensity shows a smaller slope in 'covalent' vs. 'ionic' coordination environments [16], meaning the Kβ emission probes the effective rather than the formal Fe(d) spin count. Clear relations between the Kβ main-line spectral features and the (formal) Fe(d) spin count (such as an increase in Kβ′ intensity) are therefore only expected within relatively homogeneous series of compounds. The overall shape of the Kβ main-line in a series of compounds with comparable covalency is dominated by the number of net unpaired 3d spins and not by oxidation states. The main-line of a low-spin, intermediate-spin and high-

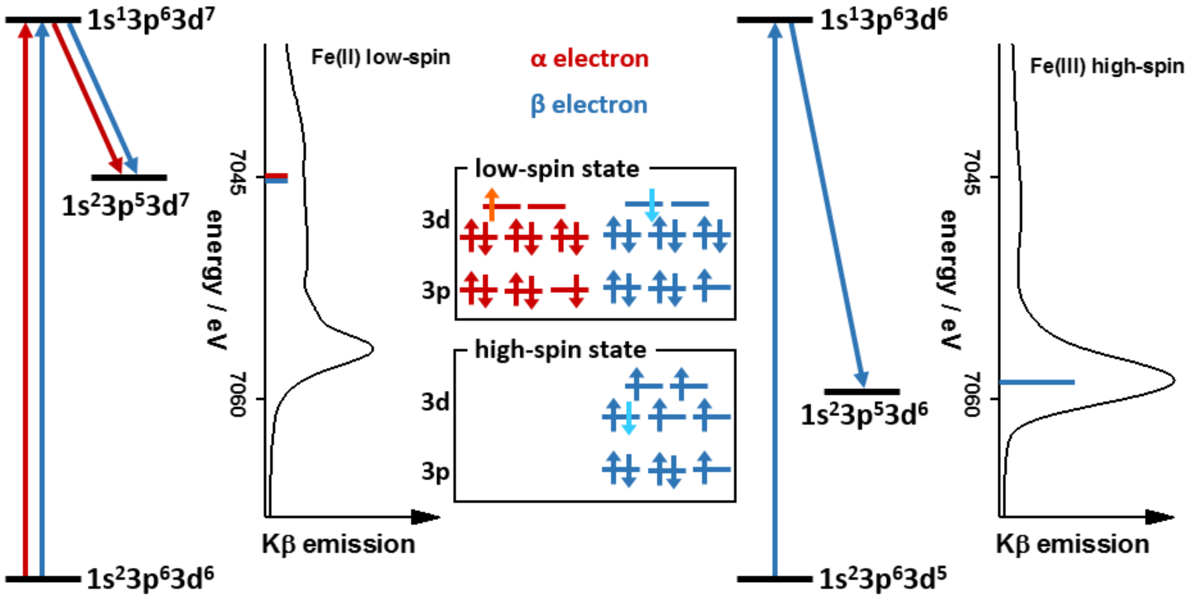


*Figure 2-7: Kβ main-line emission schematic for an Fe(II) low-spin compound (left) and an Fe(III) high-spin compound (right) after resonant excitation. The final state electron configuration is shown in the center. (**left**) An Fe(II) compound in a low-spin ground-state absorbs a X-ray photon, which excites an α (red arrow) or β (blue arrow) core electron into an unoccupied metal(3d) level ($1s^2 3d^6 \rightarrow 1s^1 3d^7$). The created α/β core hole is filled by a 3p electron of equal spin. Both final states give rise to the Kβ' feature (see spectrum) as the unpaired α/β 3d spin coincides with a 3p of opposite flavor. (**right**) No excitations of α spins are possible for an Fe(III) high-spin compound. The final state therefore is formed entirely by β spin transitions, which reduces the net count of unpaired 3d spins by one and gives rise to the $K\beta_{1,3}$ feature.*

spin Fe(II) compound are expected to differ significantly in $K\beta_{1,3}$ feature position and Kβ' feature intensity, while Fe(II) and Fe(III) high-spin compounds are not.

Kβ main-line detection after resonant excitation

For resonant 1s electron excitation into the valence levels at incident energies in the pre-edge absorption region, creation of an additional unpaired Fe(d) spin or spin pairing in the final state can occur ($1s^2 3p^6 3d^n \rightarrow 1s^1 3p^6 3d^{n+1} \rightarrow 1s^2 3p^5 3d^{n+1}$). The resulting Kβ spectrum (see Figure 2-7) corresponds to a transect through the so-called resonant inelastic X-ray scattering (RIXS) plane (i.e. intensity plotted versus emission and excitation energy) in the direction of constant excitation energy [23, 120]. The shape depends on the excitation energy due to probing of different unoccupied valence, i.e. Fe(d), levels. For an Fe(II) low-spin configuration in an octahedral ligand field an excitation of α or β electrons into the unoccupied e_g orbitals is possible, which increases the net unpaired 3d electron count by one. In both cases an Fe(3p) electron of equal spin fills the empty core hole and the 3p-3d exchange interaction is positive (see inset Figure 2-7). The final states therefore correspond to the Kβ’ feature. No α electron excitation is possible in an Fe(III) high-spin compound because all Fe(3d) orbitals are occupied. The β spin excitation reduces the net Fe(3d) count and the final state has negative 3p-3d exchange interaction because of parallel spins, which gives rise to the K$\beta_{1,3}$ feature.

Conclusion

X-ray emission spectroscopy (XES) at the Kβ emission lines provides information on metal spin states and redox levels (Kβ main-line) ligand types (Kβ satellite lines). XES here was used to investigate metal cofactors in hemoproteins (publications I, IV, V); hydrogenases (VII), oxidases (IX) and synthetic compounds (II, VI, VIII) [83, 98-105].

Kβ satellite line emission spectroscopy is a uniquely ligand-sensitive method. Emission peak positions primarily depend on ligand type and intensities on bond covalency and length. The combination of satellite line emission with reference compounds and computational valence-to-core (vtc) spectra from density functional theory (DFT, see section 3.1) gives insight into key electronic and structural parameters with regard to metal ligation.

The Kβ main-line emission shows a weak side-feature (Kβ’) with an amplitude that is (for non-resonant excitation) proportional to the number of unpaired metal(3d) electrons. Resonant excitation results in an anti-proportional correlation of the Kβ’ amplitude and the metal(3d) spins. Kβ main-line spectroscopy is therefore particularly useful to determine metal spin states. In combination with reference compounds, relative Kβ’ intensities can be correlated with formal unpaired metal(3d) counts and therefore gives insight into the local electronic configuration of metal cofactors. Calculation of Kβ main-line features on the basis of structures would be desirable and some progress has been made with *ab initio* molecular orbital (MO) theory, but is, so far, limited to high-spin complexes [36]. Therefore, only correlation studies involving ground-state electron densities from DFT or MO theory and formal unpaired metal(3d) electrons presently are feasible.

Collection of high-quality Kβ spectra, however, remains an experimental challenge for samples with low metal concentrations such as proteins and is further complicated for X-ray sensitive samples in high oxidation states. In this work, an alternative approach, time-resolved energy-sampling, was developed and employed to characterize radiation-sensitive materials (see section 4).

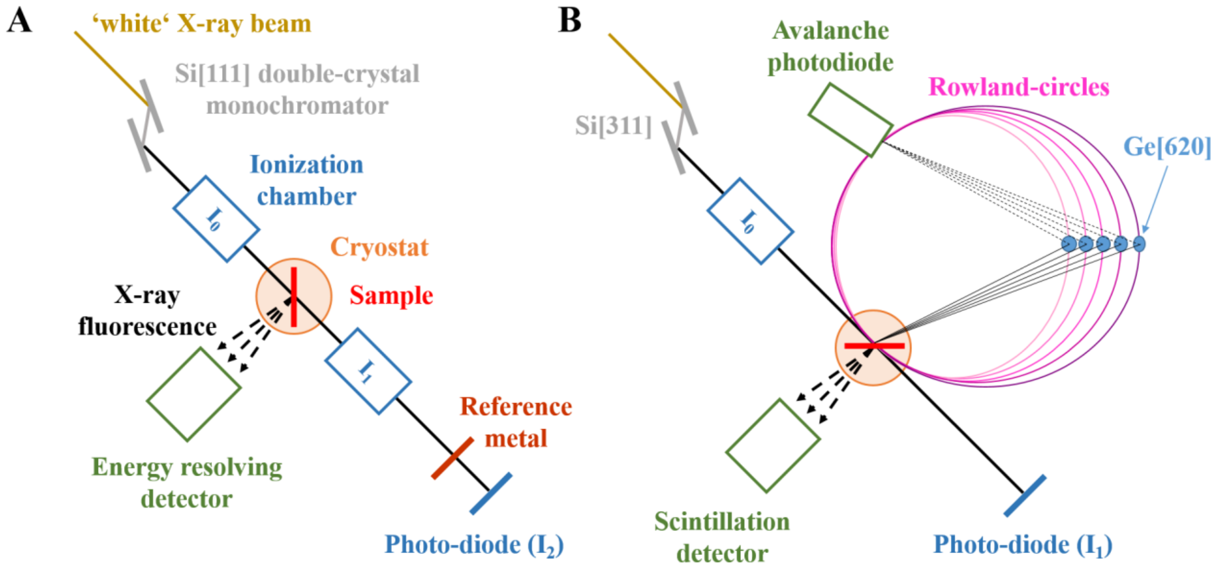


*Figure 2-8: Scheme of experimental setup at (**A**) the KMC-3 beamline at the Berlin Electron Storage Ring Society for Synchrotron Radiation (BESSY II) and (**B**) the ID26 beamline at the European Synchrotron Radiation Facility (ESRF, France). For further description, see text below.*

2.3 Materials and methods

The data in this work was collected mainly at two synchrotron sources. This section provides a brief overview of (i) the synchrotron facilities where respective experiments were conducted and (ii) the data analysis process.

Synchrotron facilities

Advanced XAS/XES experiments were carried out at the beamline ID26 at the European Synchrotron Radiation Facility (ESRF, Grenoble, France). It is equipped with an absorption/emission spectrometer that permits measurement of emission and absorption spectra with excellent contrast (high signal-to-noise ratio). The high emission energy resolution (≈ 1 eV) of the vertical-plane Rowland circle crystal spectrometer allows to analyze samples with low metal concentration (< 10 mM). However, the high photon flux at the triple-undulator beamline ID26 results in rapid X-ray radiation damage, which necessitated implementation and evaluation of time-resolved approaches for XES (see section 4).

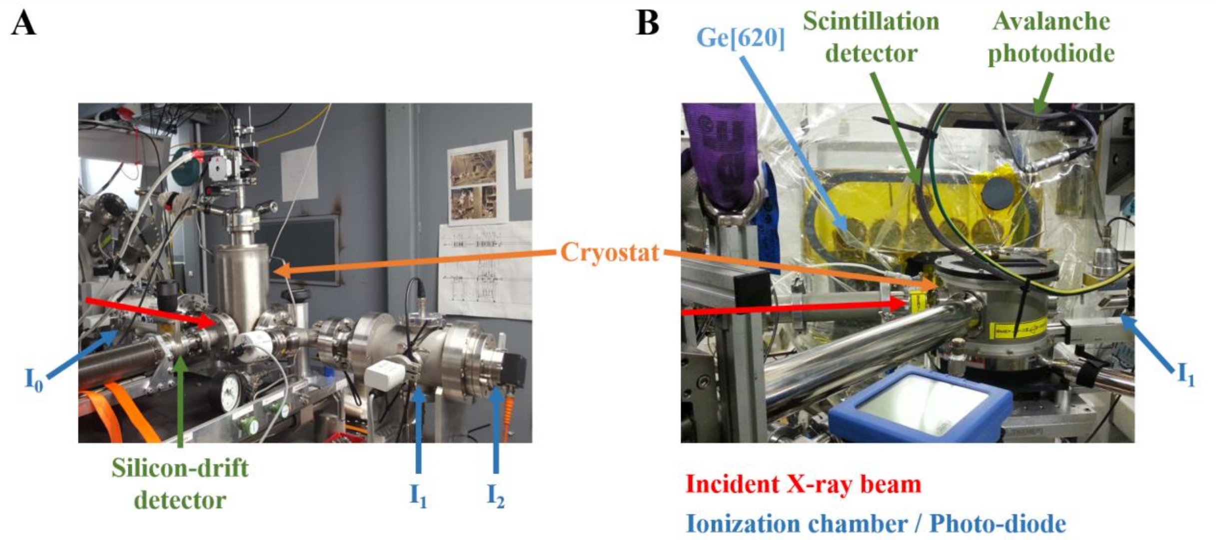


*Figure 2-9: Images of beamline setup at (**A**) the KMC-3 beamline at BESSY II and (**B**) the ID26 beamline at the ESRF. For a schematic representation, see Figure 2-8 and the text for details.*

The bending-magnet KMC-3 beamline at BESSY II applies a photon flux that is three orders of magnitude smaller compared to ID26. This is ideal for conventional XAS (XANES, EXAFS) on radiation-sensitive samples.

ID26 beamline at the ESRF

High-resolution Fe K-edge XAS and Kβ-XES spectra were collected at the triple-undulator beamline ID26 at the ESRF (Grenoble, France). The schematic setup is shown in Figure 2-8B and Figure 2-9B shows a photo of the experiment. Samples were held in a liquid helium cryostat from Cryovac (Troisdorf, Germany), at an angle of 45° relative to the incident beam. The 'white' X-ray beam (FWHM ≈ 150 eV, spectral range 2.4 – 30 keV) is created with three mechanically independent undulators. Incident beam energy was selected by a Si[311] double-crystal monochromator with ~ 0.2 eV bandwidth, providing a flux of ~ $5x10^{11}$ photons per second and mm^2 at the Fe K-edge [25, 26, 83, 121, 122]. Three water-cooled flat mirrors focus the X-ray beam and the spot size (shaped by slits) was about ~ 0.2 – 0.3 mm in vertical and ~ 0.3 – 0.4 mm in horizontal direction. The mirrors are coated (Pt, Si and Pd) to reject higher harmonics. Data were corrected for variations of incident beam intensity (electron storage ring current, undulator beam profile) by measuring photon flux with an ionization chamber before the sample (I_0). X-ray absorption was detected in transmission mode or using fluorescence detection. Absorption in transmission (see equation (2-1)) was detected via a photo-diode behind the sample (I_1, see figures above). X-ray fluorescence was detected by (i) a scintillation detector and (ii) a vertical-plane Rowland-circle spectrometer with an avalanche photodiode detector (apd). The scintillation detector has an active area of ~ 20 mm^2 and was positioned at ~ 1 m distance to the sample and perpendicular to the incident beam. The detector was shielded by a 10 µm manganese foil against scattered incident X-ray photons. The detected signal is therefore predominantly from iron Kα fluorescence (broad-band total fluorescence detection). The vertical-plane Rowland-circle emission spectrometer uses five spherically bend Ge[620] analyzer crystals at the Fe K-edge. The Rowland-circle has a diameter of 1 m and the emission energy bandwidth is ~ 1 eV. The apd is shielded by a ~ 1 mm slit against scattered incident X-ray photons. The fluorescence photons path goes through a balloon filled with helium gas in order to reduce absorption and scattering (see Figure 2-9B) [25]. The emission spectrometer was used for XAS experiments by narrow-band Kβ fluorescence detection at a constant emission energy e.g. at the maximum of the $K\beta_{1,3}$ feature. All three detectors were used simultaneously for rapid-scan (~ 1 s) XANES and pre-edge spectra collection. However, only the porphyrin reference compounds had metal concentrations high enough to detect a sufficient absorption change in transmission mode. The very short scan period is achieved by synchronous movement of the undulator gap and the Bragg angle of the monochromator. The X-ray beam was appropriately attenuated by stacked Al foils of 20 µm thickness (filter box from XIA, ~ 30 % flux attenuation at the Fe K-edge per foil). X-ray exposure time periods for unperturbed data collection were determined from X-ray fluorescence time-scans on the K-edge rise and successive XANES scans on single sample spots used to monitor X-ray induced sample modifications. Kβ emission spectra were collected using a point-by-point time-resolved energy-sampling approach (~ 1 s data acquisition per 0.35 eV step) as previously described (for details see ref. [121], section 4 and Figure 4-2). Monochromator and emission spectrometer energy axes were calibrated with a precision of ± 0.1 eV using a Gaussian fit to the elastic scattering peak of the incident X-rays for monochromator energy calibration and according calibration of the emission spectrometer energy [121]. Kβ′ amplitudes were derived from data within 7044 – 7046 eV.

KMC-3 beamline at BESSY II

Lower-resolution XANES and EXAFS spectra were collected at beamline KMC-3 at BESSY II (Helmholtz-Center Berlin, Germany). The schematic setup can be seen in Figure 2-8A and the actual setup in Figure 2-9A. The bending-magnet beamline uses a Si[111] double-crystal monochromator for energy selection (2 – 14 keV) with a resolution of ~ 2 eV. The sample was held in a liquid helium cryostat. XAS experiments in transmission mode use ionization chambers before and after the sample (I_0/I_1; see Figure). Energy calibration was done by measurement of X-ray absorption of a reference metal foil with a third ionization chamber (I_2).

A germanium detector (Camberra, USA) was positioned perpendicular to the incident beam in an evacuated tube connected to the cryostat [123].

Data processing

Most protein samples proved to be sensitive to X-ray exposure (see section 5.5). This led to a safe data acquisition time < 5 s (e.g. **oxy** HB/MB at 250 K) at the ID26 beamline. Consequently, high signal-to-noise ratios were achieved by averaging of many spectra collected on separate sample spots. The number of averaged spectra differed depending on experimental method, compound and temperature. Up to about 700 scans were collected to yield the final XAS/XES spectra. XES experiments were also conducted using a time-resolved energy-sampling (TRES) approach (see section 4), where one spectrum was reconstructed from individual measurements on several sample spots instead of one spectrum per sample spot. Data processing was done with in-house software or self-written scripts in python 2.7 or OriginPro 2017. Of future interest is the processing of TRES scans. One spectrum measured with this approach consisted of ~ 200 scan files, which had to be matched and processed. Descriptions of data processing of the different experimental data are given below.

EXAFS: Spectra were first detector dead time corrected. The total detected count rate (incoming count rate; ICR) and the spectral count rate in the region of interest (ROI) energy window (outgoing count rate; OCR) are not directly proportional. The correction is calibrated via a saturation curve according to

$$\mathrm{OCR} = \mathrm{A}\left(1 - \mathrm{e}^{-\mathrm{k}\,\mathrm{ICR}}\right) \tag{2-7}$$

Amplitude (A) and saturation rate (k) are fit parameters taken from the saturation curve. The spectra were then corrected for energy and time dependent variations in photon flux (I_0, see scheme in Figure 2-8) and afterwards energy calibrated by fitting the first inflection point of the metal reference absorption spectra, which is set to 7112 eV for iron. Spectra were averaged; a linear background was subtracted and spectra were normalized with a polynomial of second or third order. The EXAFS oscillations (χ, see equation (2-4)) were extracted by fitting the base absorption above the K-edge energy ($\mu_0(E)$) as a polynomial. The k^3-weighted EXAFS oscillation spectra were simulated with in-house software. The necessary phase functions were calculated using FEFF9 and structures from XRD for HB/MB samples or from DFT geometry optimizations for heme model compounds [95]. See Figure 2-4 for a summary of EXAFS data processing.

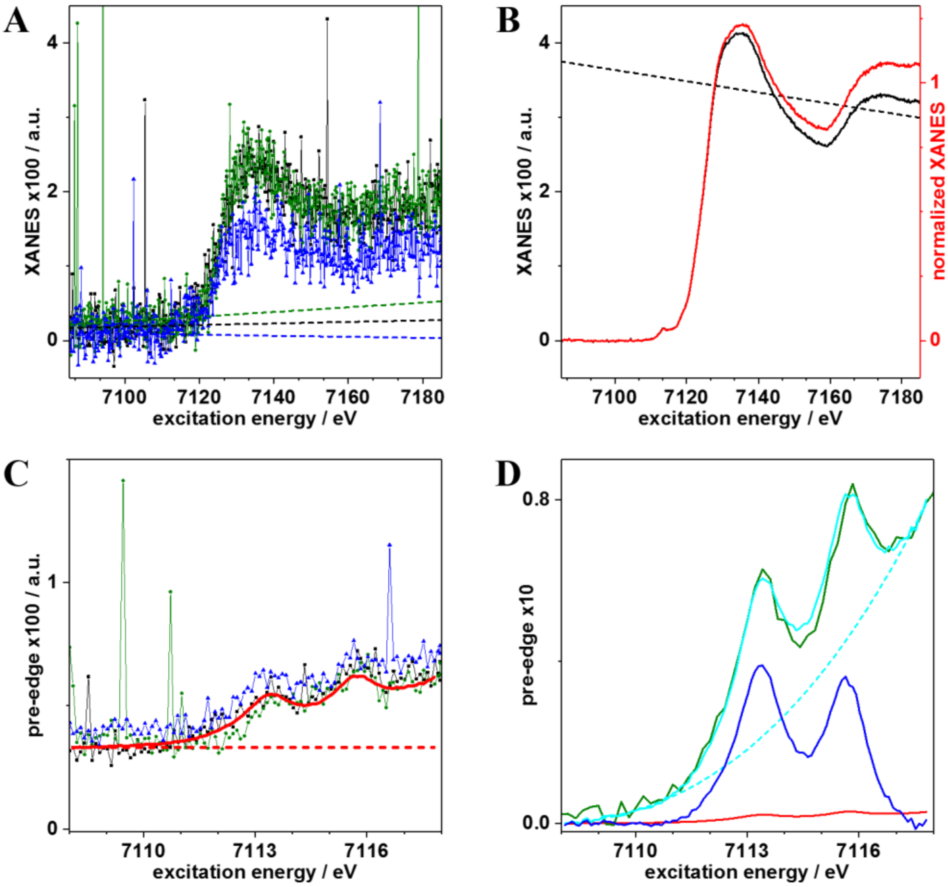


*Figure 2-10: Exemplary data processing for XANES (A, B) and pre-edge (C, D) spectra. (**A**) XANES scans (total fluorescence; corrected for beam flux), which were collected with a scintillation detector on different sample positions, show spikes due to high-energy photons from cosmic radiation. A linear background (dashed lines) is subtracted from the cleaned spectra (not shown) (**B**) XANES spectra after averaging (black line) with linear fit (black dashes) and corresponding normalized XANES spectrum (red line). (**C**) Pre-edge total fluorescence measurements (single scans, blue & green; averaged spectrum, red line) and constant background (red dashes). (**D**) The averaged spectrum (red line) is scaled (cyan line) according to the normalized XANES spectrum (olive green). A polynomial background (cyan dashes) is subtracted to yield the final pre-edge spectrum (blue line). Spectra are from **oxy** myoglobin (A/B) and **carboxy** hemoglobin (C/D) and were collected at 250 K.*

XANES: Individual spectra were first normalization to beam flux (I_0). From total fluorescence (Kα detected via scintillation detector) high-energy photon spikes were removed (see Figure 2-10A). A linear background before the K-edge rise was subtracted from spectra by fitting an energy range of ~ 2 – 4 eV. Spectra were then averaged and normalized with a linear fit (see

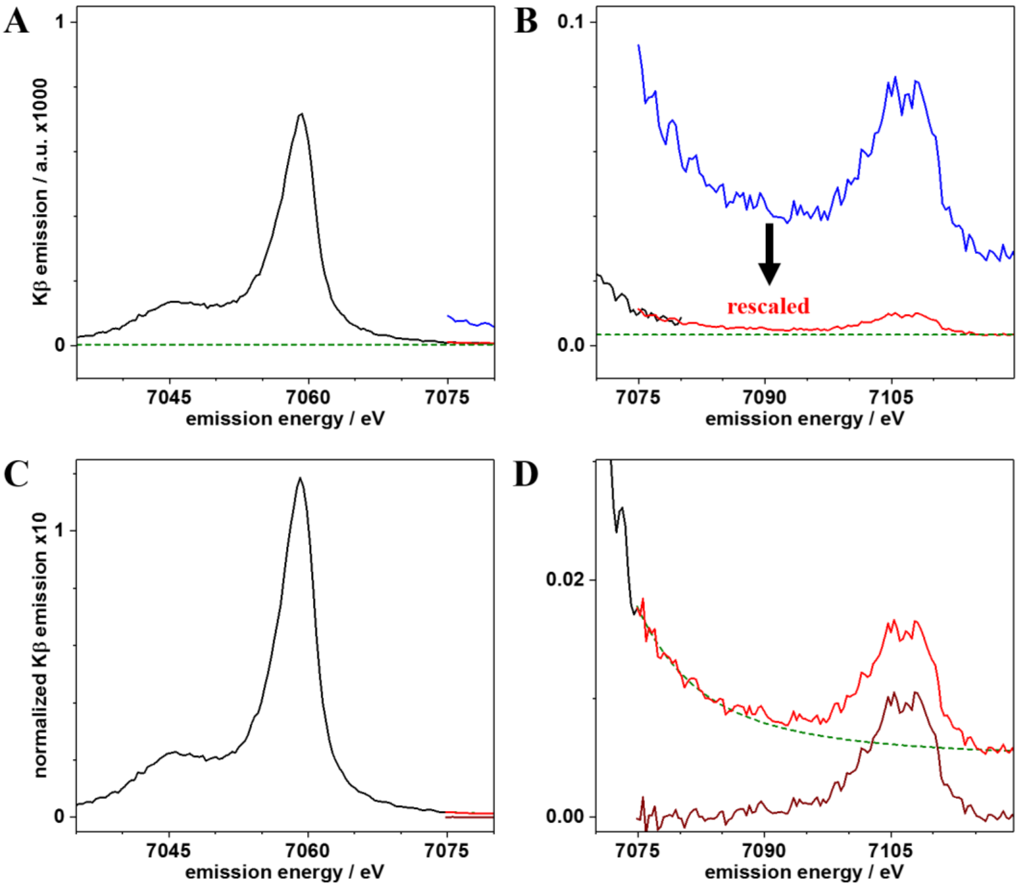


*Figure 2-11: Exemplary data processing for Kβ main-line (A/C) and satellite line (B/D) scans. (**A/B**) Shown are beam photon flux corrected and averaged main-line (black) and satellite line (blue) spectra, the rescaled satellite line (red) and the corresponding background (green dashes). (**C/D**) Kβ emission after background subtraction and normalization to unity area (black and red). The final satellite line (dark red) stems from subtracting a fitted Lorentzian (green dashes). Shown spectra are from **deoxy** myoglobin at 20 K.*

Figure 2-10B). XANES spectra were normalized so that they resemble normalized EXAFS spectra.

Pre-edges: Individual spectra were first beam flux corrected and cleaned similar to XANES spectra. Pre-edge spectra energy ranges, unlike XANES or EXAFS, do not include a background region to maximize sampling time. Spectra were therefore averaged before subtracting a constant background (see Figure 2-10C). The resulting spectra were scaled to match normalized XANES spectra (see Figure 2-10D). A polynomial background was subtracted from the resulting spectrum in order to remove the K-edge and yield the final pre-edge spectrum. Narrow-band Kβ detected spectra were processed similar to total fluorescence spectra.

Kβ main-line and satellite line: Individual spectra were first normalized to beam flux and then averaged. Kβ main-line and satellite line spectra were measured with an energy range overlap. The Kβ satellite line spectra were scaled to match the Kβ main-line tail at the spectral overlap. The Kβ spectra were interpolated, the background removed and normalized to unity area. The satellite line was corrected for the main-line tail by fitting and subtracting a Lorentzian shaped peak. Kβ main-line emission after resonant excitation of a core 1s electron into an unoccupied iron 3d orbital was processed in a similar way.

3 Computational methods

Since Schrödinger introduced wavefunctions to quantum mechanics, two things became immediately clear. First, a wavefunction that solves this equation would provide a complete description of a given system. Second, it is impossible to find an exact solution analytically for any system with more than one electron. Therefore, several methods have been developed in order to numerically solve or approximate the exact wavefunction. Among the most prominent in the field of quantum chemistry (QC) is the molecular orbital (MO) theory, which is based on the approximation of the single many-particle equation by many single-particle equations through the Hartree-Fock (HF) method. Another approach is density functional theory (DFT), which approximates the exact exchange-correlation energy by an electron density functional. As examples, MO theory and DFT have been applied in the past to describe bonding of dioxygen to iron heme complexes [55, 59, 60, 66-74, 76, 77, 124]. However, the results from both approaches converged towards different descriptions of Fe – O_2 bonding. Although X-ray spectroscopy (XAS/XES) experiments have been conducted on hemoproteins and Fe heme complexes, to our knowledge no thorough comparison of X-ray spectroscopy and quantum chemical calculations has been conducted. Consequently, a wide range of QC calculations were applied in this work and protocols implemented in our laboratory. The technical aspects of performing computational calculations had to be realized, e.g. setting up calculations and choosing or writing software to extract and visualize key information from resulting output files. These aspects are presented in section 3.3 and 3.4. On the other hand, interpretation of QC calculations requires insight into the theoretical framework. Both, DFT and multiconfigurational MO theory have been discussed extensively in the past and a lot is "textbook knowledge". However, the interpretation of complex electronic configurations and comparison of results from different branches of theory, as in our investigation of heme systems, requires a level of scholar knowledge that appears unreasonable to presuppose. The following two sections intention is to introduce, document and discuss theoretical aspects that are key for a combination of DFT and MO theory with advanced X-ray spectroscopy. Presentation of QC theory is challenging and therefore discussed in terms of well-established and topical examples (H_2, ozone) where possible. Nonetheless, theoretical QC is rather abstract and, to that end, the conclusion of each section provides a summary.

3.1 Density functional theory

Combination of various experimental methods with density functional theory (DFT) has found tremendous success in the past [125]. Increase in computation power, access to high-performance computing and development of user-friendly program packages like ORCA and Gaussian has led to a remarkable increase in publications that involve some sort of DFT calculations. While application of DFT has become more accessible, interpretation of computational results can require insight into the complex and fascinating framework of the theory. The aim of this thesis was to combine DFT-based calculations with advanced X-ray spectroscopy and this section presents an introduction into (i) the formalism of DFT and (ii) the prediction of core-to-valence (ctv) and valence-to-core (vtc) spectra from DFT.

Introduction into the DFT formalism

MO theory approximates the real wavefunction of a system in order to calculate and minimize the total energy of the system. However, a quantum mechanical wavefunction is not the only theoretical quantity that can provide the total energy of molecular systems. Any observable or non-observable quantity, which describes (i) the total number of electrons (*N*), (ii) the position of the nuclei and (iii) their atomic number (*Z*) can express the Hamiltonian and thus yield the total energy of a system.

A physical observable, which directly relates to the total number of electrons, is the electron density (ρ), which is calculated from the wavefunction of a system using the electron density operator ($\hat{\rho}(\boldsymbol{r})$).

$$\rho(\boldsymbol{r}) = \langle \Psi | \hat{\rho}(\boldsymbol{r}) | \Psi \rangle \tag{3-1}$$

$$N = \int \rho(\boldsymbol{r}) d\boldsymbol{r} \tag{3-2}$$

An example for the electron density of a molecule is shown in Figure 3-1B. The electron density (usually visualized as an isosurface plot) contains the number of electrons in the system, but the nuclear coordinates and atomic numbers are not yet included. The position of the nuclei can be derived by focusing on the core electrons, which are densely packed around each nucleus. The electron density therefore forms a local maximum at each nuclear position (see Figure 3-1C). Note that the nuclei themselves do not contribute directly to equation (3-2).

The final information to be derived from the electron density is the atomic charge at each atom position, but the absolute electron density at the exact nuclear coordinate is not sufficient.

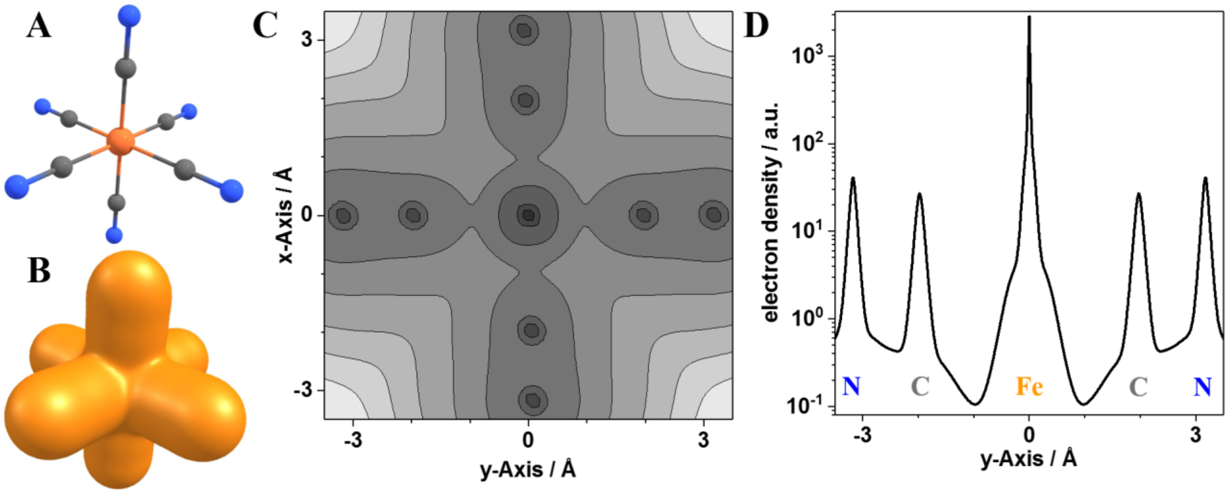


*Figure 3-1: Example of DFT calculation for ferricyanide (S = ½), which forms part of Prussian blue. Shown are (**A**) the optimized geometry, (**B**) an isosurface plot (0.1 au) of the electron density, (**C**) logarithmic electron density contour plot in the x-y plane and (**D**) the electron density (note the log scale) along the y-axis in (B) at x = z = 0. (A) Geometry was optimized using the SV(P) basis set with the BP86 functional and modeling solvation using the conductor-like screening model (COSMO) with a dielectric constant (ε) = 2. Cartesian axes run parallel to molecular axes with the iron atom at their origin. (C) Contour lines mark the change by an order of magnitude.*

Rather, the atomic charge affects the shape of the local maximum in the electron density as core electrons move closer to the nucleus with increasing atomic charge. Consequently, the electron density has a steeper rise around the nuclear position for higher Z values (see Figure 3-1D). The atomic charge is numerically derived from the tangent of the local maximum in the electron density at the position of the nucleus A from

$$\left.\frac{\partial\bar{\rho}(r_A)}{\partial r_A}\right|_{r_A=0} = -2Z_A\rho(r_A) \qquad (3\text{-}3)$$

where r_A is the spherical distance from nucleus A and $\bar{\rho}(r_A)$ is the spherically averaged electron density. The electron density thus holds all necessary information to calculate the total energy of the system. Yet, the problem of how to calculate this electron density without the prerequisite of solving the Schrödinger equations of the system remains. The solution is to realize that the wavefunction in equation (3-1) can be constructed by another (trial) electron density.

$$\rho(\boldsymbol{r}) = \langle\Psi(\rho')|\hat{\rho}(\boldsymbol{r})|\Psi(\rho')\rangle \qquad (3\text{-}4)$$

The electron density is thereby calculated from a wavefunction in an implicit form, otherwise known as a *functional* equation.

Early electron density functionals date back to the 1930s in work from Dirac [126] and Slater [127, 128] (among others) and are now known as Xα methods named after the empirical value α in the exchange energy functional. These functionals were derived by assuming a free electron gas, thus dramatically simplifying the wavefunction. This method was primarily used in solid-states physics during that time due to its relatively low demand of computation power. This made Xα methods favorable for very large systems, but too inaccurate to perform simulations on a molecular scale.

It was not until the work of Hohenberg and Kohn in 1964 that DFT made an impact on computational chemistry [129, 130]. These authors first proved that the electron density of a non-degenerate ground-state uniquely determines the external potential ($V_{ext}(r)$) in the Schrödinger equation, which is the Hohenberg-Kohn (HK) existence theorem. This exact external potential defines the exact Hamiltonian and thus the wavefunction. This makes the exact non-degenerate ground-state wavefunction a functional of the electron density. The HK existence theorem only applies to the lowest-energy electron density for each point symmetry (ground-state). Excited states with the same symmetry as the ground state cannot be calculated by DFT. A theory to determine excited state excitation-energies in order to model X-ray absorption spectroscopy is presented later in this section. The first theorem stated that the exact wavefunction would be determined by the exact electron density, but does not offer a method how to search and identify it. The HK variational theorem provides just that. It ensures that the energy of a system determined by an electron density obeys a variational principle. Thus, any trial density (ρ_{trial}) is higher in total energy than the exact electron density of the ground-state (ρ_0) which represents a global minimum in total energy of the system.

$$\left\langle \Psi[\rho_{\text{trial}}] \middle| \hat{H}[\rho_{\text{trial}}] \middle| \Psi[\rho_{\text{trial}}] \right\rangle = E_{trial} \geq E_0 = \left\langle \Psi[\rho_0] \middle| \hat{H}[\rho_0] \middle| \Psi[\rho_0] \right\rangle \tag{3-5}$$

With these two theorems at hand, one knows that a unique electron density exists and how to identify it. What is still missing is a methodology how to construct trial electron densities in order to reach the ground-state. Otherwise, DFT would still require solving the Schrödinger equation. Kohn and Sham proposed a solution to this problem the year after publishing their theorems derived above.

Kohn-Sham self-consistent field

The problem in solving the correct Hamiltonian and calculating the wavefunction is the electron-electron interaction. Kohn and Sham proposed, instead of trying to calculate the electron density by solving the equations for the real system of interacting electrons, one can construct the same electron density for a hypothetical system of non-interacting electrons [131].

The wavefunction can then, similar to HF theory, be expressed as a single Slater determinant. However, the energy of this non-interacting system has to be corrected. The missing factor is the influence of exchange and correlation to the energy. Assuming that the energy correction to the electron density of the non-interacting system is known and using the HK existence theorem, both fictitious and real system have the same electron density. As both electron densities are equal, they both have the same number of electrons, nuclei positions and charges. The Hamiltonian for this fictitious system is the sum of one-electron operators, the wavefunction is a single Slater determinant and the total energy is the sum of the one-electron eigenvalues. The energy functional can then be written as follows [4].

$$E[\rho(\boldsymbol{r})] = T_e[\rho(\boldsymbol{r})] + V_{Ne}[\rho(\boldsymbol{r})] + V_{ee}[\rho(\boldsymbol{r})] + \Delta T[\rho(\boldsymbol{r})] + \Delta V_{ee}[\rho(\boldsymbol{r})] \tag{3-6}$$

The terms on the right-hand side are the kinetic energy of the electrons (T_e), the attraction between nuclei and electrons (V_{ne}), the classical repulsion between electrons (V_{ee}), the difference between the kinetic energy of the system of interacting and non-interacting electrons (ΔT) and the correction to the classical Coulomb repulsion by quantum mechanics (ΔV_{ee}). The electron density of a wavefunction made up of non-interacting one-electron orbitals (φ_i) can be calculated with $\rho = \sum_i^N \langle \varphi_i | \varphi_i \rangle$ were the sum runs over the total number of electrons (N). Each independent particle then has an individual kinetic energy and is influenced by an individual potential. The equation is therefore transformed from single many-body operators (unsolvable) to a sum of many single-body operators (solvable). Consequently, equation (3-6) transforms into:

$$\begin{aligned} E[\rho(\boldsymbol{r})] = \sum_i^N \Bigg(\left\langle \varphi_i \middle| -\frac{1}{2}\nabla_i^2 \middle| \varphi_i \right\rangle - \left\langle \varphi_i \middle| \textstyle\sum_k^{nuclei} \dfrac{Z_k}{|\boldsymbol{r_i} - \boldsymbol{r_k}|} \middle| \varphi_i \right\rangle \\ + \frac{1}{2} \left\langle \varphi_i \middle| \int \frac{\rho(\boldsymbol{r'})}{|\boldsymbol{r_i} - \boldsymbol{r'}|} d\boldsymbol{r'} \middle| \varphi_i \right\rangle \Bigg) + E_{xc}[\rho(\boldsymbol{r})] \end{aligned} \tag{3-7}$$

Note that each electron interacts with the complete electron density instead of each individual electron and that the exchange-correlation energy correction (E_{xc}) is unique for each system and therefore depends on the actual electron density of the system. Hence the name 'exchange-correlation functional'. In order to minimize the energy through variation of the one-electron orbitals, one has to solve the eigenvalue equations $(h_i^{KS} - \varepsilon_i)\varphi_i = 0$ with the Kohn-Sham one-electron operator (h^{KS})

$$h_i^{KS} = -\frac{1}{2}\nabla_i^2 - \sum_k^{nuclei} \frac{Z_k}{|\boldsymbol{r_i} - \boldsymbol{r_k}|} + \int \frac{\rho(\boldsymbol{r'})}{|\boldsymbol{r_i} - \boldsymbol{r'}|} d\boldsymbol{r'} + \frac{\delta E_{xc}}{\delta \rho} \tag{3-8}$$

These equations drastically reduce the complexity of the problem, except that the exact exchange-correlation energy correction is unknown. This functional contains both correction factors from equation (3-6). DFT is, to this point, an exact theory, which would produce correct absolute energies if one would know the exact exchange-correlation functional. This is in contrast to HF theory, which is an approximation from the start. The correct exchange-correlation (XC) functional is almost never known and has to be approximated. Different methods have been developed to calculate the XC functional from the electron density. This has been and still is an active field of research [132, 133].

Exchange-correlation functionals

As the exchange-correlation (XC) functional depends on the electron density, the first approximation is the dependence of the XC functional on local electron density values, which is the local density approximation (LDA) or, in a spin-unrestricted calculation, local spin density approximation (LSDA). The most commonly used L(S)DA functionals are based on the Xα methods. The next step is to improve the XC functionals from LDA by inclusion of the electron density gradient. These are functionals with generalized gradient approximation (GGA). The GGA exchange functional by Becke [134] in combination with the GGA correlation functional by Perdew [135, 136], abbreviated BP or BP86, is one example often found in literature [27, 83, 90, 99]. Becke used the known exchange energies of six noble gases to optimize the correction to the LDA functional. The GGA correlation functional developed by Lee, Yang and Parr (LYP) is purely empirical as it is derived from helium atom fits [137]. As GGA functionals demonstrated the success of fitting empirical data to improve the correction to LDA the next step was to optimize the relative weight of the corrections themselves as well. This and the inclusion of HF exchange into the functional resulted in so-called ‘hybrid’ functionals like B3PW91 and B3LYP [138, 139], the latter has proven to be very successful [140-142]. The B3LYP functional reads:

$$E_{xc}^{B3LYP} = (1-\mathrm{a})\, E_x^{LSDA} + a\, E_x^{HF} + b\, \Delta E_x^{B} + (1-c)\, E_c^{LSDA} + c\, E_c^{LYP} \qquad (3\text{-}9)$$

The total exchange-correlation energy is a weighted combination of exchange and correlation energies (E_x, E_c). The exchange energy is a combination of LSDA with Becke’s gradient correction and HF theory while the correlation energy is a combination of LSDA and LYP. The three parameters are a = 0.2, b = 0.72 and c = 0.81 and were optimized using a similar expression as in the equation above, but using the gradient correction to the correlation energy by Perdew and Wang [138]. As mentioned before, DFT is exact when the real XC functional is known. However, this is not true when including approximate XC functionals because the HK variational theorem only applies to the real XC functional. Although the electron density is optimized by a variational approach, this does not mean that the resulting total energy is correct or that the XC functional, which produces the lowest energy, is automatically the closest to reality [4].

Antiferromagnetic coupling via broken-symmetry

The ability of understanding the nature of a chemical bond or reaction between two subsystems of a molecule requires knowledge about the electronic structure. A specific question of interest is if spins (or radicals) in these two subsystems interact though ferromagnetic coupling (fc) or antiferromagnetic coupling (afc). The magnetic coupling between two subsystems of a molecule A and B with well-defined spin S_A and S_B can be described via the Heisenberg interaction Hamiltonian $\widehat{H}_H$ (the circumflex indicates an operator) [5, 143].

$$\widehat{H}_H = -2J\hat{S}_A\hat{S}_B \tag{3-10}$$

In it, J is the magnetic coupling constant and $\hat{S}_i$ is the local spin-operator of a subsystem. As a result of the sign convention, J value > 0 favor parallel spins and therefore fc and J values < 0 describe afc. The total spin-operator can then be represented by the individual local spin-operators via $\hat{S} = \hat{S}_A + \hat{S}_B$. The possible expectation values for $\hat{S}^2$ are S(S+1) and S ranges from $S_A + S_B$ to $|S_A - S_B|$. The Hamiltonian therefore leads to a spin ladder with either ascending or descending total spin.

Restricted Kohn-Sham (RKS) wavefunctions often predict wrong energetic orders and differences of open-shell singlet and triplet states if afc is involved [4, 5]. This is because open-shell singlet states are intrinsically not single-determinantal. A solution to this is to use an unrestricted kohn-sham (UKS) wavefunction, which allows α and β orbitals to break spatial spin symmetry. Permitting broken-symmetry (BS) solutions improves predictions of order and energy splitting between e.g. singlet and triplet states. This comes at the cost of an unphysical spin density for a singlet state, which should be zero at all points in space. The BS wavefunction (in DFT) for a singlet state with afc between e.g. two transition-metal atoms shows spatially unpaired α MOs on one metal atom and β MOs on the other metal atom. The spin density is therefore positive on one metal and negative on the second and only zero as an integral over the entire space. The reason lies in the approach of constructing a multiconfigurational solution with a single-configurational system. When only considering afc between two electron spins on two subsystems (e.g. two transition metals; A/B), the singlet broken-symmetry wavefunction takes the form:

$$|\Psi^{BS}\rangle = |\varphi_A \bar{\varphi}_B| \tag{3-11}$$

Here any contributions besides the two magnetic orbitals (φ_i) involved in the afc are ignored. The subscript indicates the metal site and the bar indicates a spin down MO. Note that these two orbitals are not restricted to be orthogonal in space. It is possible to express this wavefunction through states obtained from full configuration interaction (Full-CI) (for more details see [5, 144]):

$$|\Psi^{BS}\rangle = c_{ion}|\,{}^1\Phi_{ion}\rangle + c_{DR}|\,{}^1\Phi_{DR}\rangle + c_T|\,{}^3\Phi\rangle \tag{3-12}$$

The first and second term (${}^1\Phi_{ion}$, ${}^1\Phi_{DR}$) describe singlet states and the third term (${}^3\Phi$) is a triplet state. These Full-CI states come with their own set of actually orthogonal (spin- and space-wise) magnetic orbitals (η_i).

$$\begin{aligned} |\,{}^1\Phi_{ion}\rangle &= 2^{-\frac{1}{2}}(|\eta_A\bar{\eta}_A| + |\eta_B\bar{\eta}_B|) \\ |\,{}^1\Phi_{DR}\rangle &= 2^{-\frac{1}{2}}(|\eta_A\bar{\eta}_B| - |\bar{\eta}_A\eta_B|) \\ |\,{}^3\Phi\rangle &= 2^{-\frac{1}{2}}(|\eta_A\bar{\eta}_B| + |\bar{\eta}_A\eta_B|) \end{aligned} \tag{3-13}$$

The first singlet state is a superposition of two configurations. Both electrons are in a single orbital either located on subsystem A or on B. This is an ionic closed-shell configuration. The second singlet state and the triplet state are both a superposition of two configurations where one electron (each with opposite spin) is found in each subsystem. The two configurations either couple to a singlet diradical (${}^1\Phi_{DR}$) or triplet diradical (${}^3\Phi$) open-shell state. The balance between the two singlet states reflects the magnitude of afc. A closed-shell configuration (no afc) would only contain the singlet ionic configuration. Increasing the weight of the singlet diradical contribution corresponds to an increasing afc. The triplet diradical state stems from the non-orthogonal magnetic orbitals from the BS-DFT wavefunction. The non-orthogonal magnetic orbitals from DFT can be transformed into the orthogonal magnetic orbitals from Full-CI, which gives rise to the coefficients (c_i) in equation (3-13). The contribution coefficients are such that $c_{ion}^2 + c_{DR}^2 + c_T^2 = 1$.

Only the singlet ionic contribution and the triplet diradical contribution are affected by optimization of the molecular orbitals [5]. Thus, any increase in the relative contribution of a diradical character has to come from increasing the triplet state weighting. The result is that, although the net spin over the entire system might be zero, the local spin density can deviate from zero for a BS singlet state. The expectation value of $\hat{S}^2$ in this case differs from zero as well. An expectation value $\langle \hat{S}^2 \rangle = 1$ can for example be interpreted as an equal mixture of singlet ($\langle \hat{S}^2 \rangle = 0$) and triplet ($\langle \hat{S}^2 \rangle = 2$) configurations. This is known as spin contamination and of special concern in HF calculations, where it represents an actual intrusion of higher spin states into the wavefunction. The interpretation though might not always be as straightforward in DFT as in HF Theory. While the wavefunction of the non-interacting system, from which the electron density is derived, may show spin contamination, the wavefunction of the actual interacting system may not. Remember that the non-interacting wavefunction was a hypothetical assumption. This means that the electron density (and thus the geometry) calculated with DFT can be correct, while the wavefunction, from which the electron density is calculated, is not.

Good agreement between experimental measurements of magnetic coupling constants and theoretical calculations have been reported when using hybrid functionals like B3LYP [143]. Comparison with Full-CI calculations revealed that although the B3LYP functional predicted the correct sign of the magnetic coupling constant, the correct magnitude was due to error cancelation. DFT calculations with B3LYP simultaneously overestimate the energy difference between high-spin and broken-symmetry states and underestimate the magnitude of separated localized spins [145]. The usually correct sign of the magnetic coupling constant results in very good agreement of DFT geometries and experimental structures. The electronic structure on the other hand has to be carefully evaluated [144].

Prediction of spectral features from XAS/XES with DFT

X-ray absorption and emission spectroscopy (XAS/XES) probes transition probabilities between the unperturbed ground-state of a system and various excited states, induced by absorption of a high-energy photon. Simulation of spectral features therefore should require calculation of high-energy excited state. However, DFT is a ground-state method and (in general) incapable to express systems in an excited state. DFT based methods have nonetheless been successful in simulating valence-to-core (vtc) and core-to-valence (ctv) spectra from XAS/XES [27, 35]. This section presents the theory of these methods and discusses possible applications.

Calculation of valence-to-core transitions

Valence-to-core (vtc) XES is a powerful tool to investigate the local ligand environment of transition metal compounds and metalloproteins. The Kβ satellite region comprises the $K\beta_{2,5}$ and Kβ'' features. The $K\beta_{2,5}$ feature stems from electronic decay of valence electrons with ligand(p) characters and the Kβ'' from valence MOs with ligand(s) character to the metal(1s) core hole [23]. In the XES process, the energy of a X-ray photon (ω_{in}) is absorbed and a photoionized state is created by emitting a core electron into the continuum. This state can decay in either a non-radiative manner, e.g. by emitting an Auger electron, or radiative manner by emitting a photon (ω_{out}). The interaction between photon and ground-state can be treated by perturbation theory and the absorption-emission process of XES described by Fermi's golden rule [81, 94]. The absorption-emission process (not including the absorbed and emitted photon), in the case of a first-row transition metal complex, can be described as a two-step process (see Figure 3-2).

$$|\mathrm{G}\rangle \rightarrow |\mathrm{I}\rangle \rightarrow |\mathrm{F}\rangle \tag{3-14}$$

$$|1\mathrm{s}^2 V^n C^0\rangle \rightarrow |1\mathrm{s}^1 V^n C^1\rangle \rightarrow |1\mathrm{s}^2 V^{n-1} C^1\rangle \tag{3-15}$$

The electronic ground-state (G) corresponds to metal core MOs ($1s^2$), ligand valence MOs (V^n) and an "empty" continuum (C^0). The ground-state is transformed into the intermediate state (I) upon absorption of an X-ray photon and emission of a core electron with energy $\Delta E_{GI} = E_I - E_G$ into the continuum. The intermediate state then transforms into the final state (F) by decay of a valence electron and emission of a photon with energy $\Delta E_{IF} = E_F - E_I$. The number of emitted photons per unit of time therefore depends on all permutations of transitions from G → I and I → F, which take Lorentzian form caused by state lifetime broadening (Γ_i), see ref. [33] for details.

$$I(\omega_{in}, \omega_{out}) \propto \sum_I \sum_F \omega_{in} \omega_{out}^3 \left|\langle G|\hat{O}|I\rangle\right|^2 \left|\langle I|\hat{O}|F\rangle\right|^2 \times \frac{\Gamma_I}{(\omega_{in} - \Delta E_{GI})^2 + \Gamma_I^2/4} \times \frac{\Gamma_F}{(\omega_{in} - \omega_{out} - \Delta E_{IF})^2 + \Gamma_F^2/4} \tag{3-16}$$

In order to calculate the spectrum at least one intermediate state configuration and all possible permutations of final states would be necessary, which is almost never feasible. Instead, the intermediate and final states are approximated from the DFT ground-state wavefunction. In DFT the ground-state is constructed from canonical molecular orbitals (CMOs), each CMO (φ_k) has a well-defined energy (ε_k) due to the Kohn-Sham equation. These individual MOs are not observable by themselves. However, approximating the vtc-XES process as a one-electron mechanism with help of DFT has proven to be valuable when investigating the ligand environment [33].

$$\hat{H}^{KS}|\varphi_k\rangle = \varepsilon_k|\varphi_k\rangle \tag{3-17}$$

The states in equation (3-15) can thus be approximated with help of the ground-state CMOs.

$$|\mathrm{G}\rangle = \left|\varphi_{1s}^{(1)}\bar{\varphi}_{1s}^{(1)} \cdots \varphi_{Vi}^{(0)}\bar{\varphi}_{Vi}^{(0)} \cdots\right\rangle \tag{3-18}$$

$$|\mathrm{I}\rangle \approx \left|\varphi_{1s}^{(0)}\bar{\varphi}_{1s}^{(1)} \cdots \varphi_{Vi}^{(1)}\bar{\varphi}_{Vi}^{(1)} \cdots\right\rangle \tag{3-19}$$

$$|\mathrm{F}\rangle \approx \left|\varphi_{1s}^{(1)}\bar{\varphi}_{1s}^{(1)} \cdots \varphi_{Vi}^{(0)}\bar{\varphi}_{Vi}^{(1)} \cdots\right\rangle \tag{3-20}$$

The subscript refers to the MO configuration, the superscript to the MO occupation number and the bar to spin-down (β) orbitals. Note that β spin excitations are omitted for clarity but included in the calculations. This significantly simplifies equation (3-16). The G → I transition is approximately constant as the promoted electron is not modeled and remaining occupied CMOs are unchanged. The I → F transitions only depend on the overlap between metal core electron orbitals and individual valence orbitals. These approximations result in a very fast computation, even for very large systems..

$$\sum_{I}\sum_{F}\left|\langle I|\hat{O}|F\rangle\right|^2 \sim \sum_{\sigma,i}\left|\langle\varphi_{1s}^{\sigma}|\hat{O}|\varphi_{Vi}^{\sigma}\rangle\right|^2 \tag{3-21}$$

The transition operator can be expanded and expressed through the electric dipole ($\hat{D}$), magnetic dipole ($\hat{\mu}$) and electric quadrupole ($\hat{Q}$) operators. The oscillator strengths of each corresponding transition is then [27, 146]:

$$\begin{aligned} f_{IF} &= \Delta E_{IF}\left|\langle I|\hat{O}|F\rangle\right|^2 \\ &\approx -\frac{2}{3}\Delta E_{IF}\left|\vec{d}_{IF}\right|^2 - \frac{2}{3}\alpha^2\Delta E_{IF}|\vec{\mu}_{IF}|^2 - \frac{2}{3}\alpha^2\Delta E_{IF}^3\sum_{a,b}|(Q_{ab})_{IF}|^2 \end{aligned} \tag{3-22}$$

where $\vec{d}_{IF}$, $\vec{\mu}_{IF}$ and $(Q_{ab})_{IF}$ are the corresponding off-diagonal matrix elements of the electric/magnetic dipole and electric quadrupole operators and (a,b) are virtual CMOs. Although contributions of all three matrixes are included in the calculations, the electric dipole transitions exceed the remaining two by ~ 2 orders of magnitude. The energy difference in this one-electron approach is reduced to the difference between CMO energies ($\Delta E_{IF} = \varepsilon_{Vi} - \varepsilon_{1s}$). These approximations make the calculations of vtc features computationally feasible but do include important assumptions. The intermediate state does not include any adjustments of the remaining occupied MOs due to the core hole. However, inclusion of a larger atom core charge when modeling the transitions did not significantly alter the results [147]. The ligand centered MOs apparently are shielded from the change in charge. The photoionization could result in a

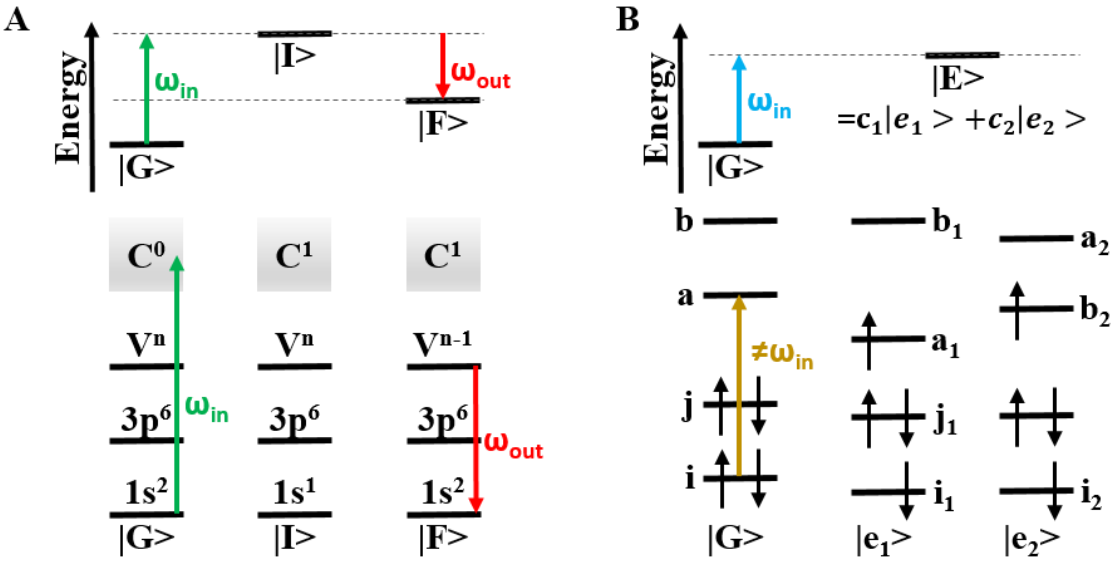


*Figure 3-2: Model of theory employed for calculation of valence-to-core (vtc) transitions from DFT (A) and core-to-valence (ctv) transitions from time-dependent DFT (TD-DFT) (B). (**A**) The vtc transitions are calculated between states constructed from the ground-state (G). The intermediate/final states (I/F) are constructed using the ground-state wavefunction with occupation number of metal(1s)/valence orbital reduced by one. The vtc transition energy is approximated as the energy difference between the two canonical MOs (CMOs). No electronic rearrangement of MOs in the intermediate or final state is included. (**B**) The energy of the excited state is adjusted for electronic relaxation in the case of ctv transitions via TD-DFT. The influence of correlation and exchange between occupied MOs (i, j) and virtual MOs (a, b) upon transition from i → a is calculated via a DFT functional. The transition energy is therefore not equal to an energy difference between CMOs. The excited state (E) includes single excitations and does not account for secondary transitions from e.g. j → b as a response to the first excitation. It does include linear combinations of different single excited states to form pure singlet or triplet states. Note that the excited state is never actually calculated; only the transition frequency and the single excited state coefficients are derived.*

multitude of different intermediate states influencing intensity and position of transitions. Multi-electron processes in the satellite line, like the KLβ feature at ~ 7115 eV in Figure 2-5, cannot be modeled with this approach. In fact, no multi-electron processes, like shake-up or shake-off events are modeled. Transition energies depend on the chosen functional and the included core potential and, to some extent, on the basis set. Theoretical spectra therefore have to be aligned with experimental spectra with an energy shift.

Calculating core-to-valence transitions with time-dependent DFT

Correlating Kβ satellite line emission spectra with predictions from DFT calculations allows to investigate ligand-sensitive occupied valence-bonds. Calculation of core-to-valence (ctv) from XAS would give insight into metal(3d/4p) mixing of unoccupied valence levels and therefore allow analysis of metal oxidation/spin state and first sphere ligation (atom types and symmetry) [148]. It is possible to use the approach from vtc to simulate ctv spectra. However, unoccupied MOs do not contribute to the total energy of a wavefunction. The orbital shapes and energies are therefore not optimized but exited states with additional valence electrons are expected to result in significant reconfiguration in the valence shell. The transition itself is highly localized which is at odds with the spatially delocalized canonical molecular orbitals (CMOs). This argues against approximating ctv spectra as static one-electron transitions similar to vtc spectra. It would be necessary to calculate excited states, but DFT is limited to time-independent systems i.e. ground-states. Calculating excited states from post-HF methods (e.g. CASSCF) is difficult. However, this only computes transition energies and not transition intensities. Instead, it is possible to calculate the linear response of the DFT ground-state to a time-dependent electric field perturbation [24, 148, 149]. The time-dependent response of a ground-state wavefunction to this perturbation is calculated via time-dependent DFT (TD-DFT).

The foundation that an exact time-dependent electron density ($\rho(r,t)$) determines a unique time-dependent external potential ($V(r,t)$) (up to a spatially constant, time-dependent function), which determines a time-dependent wavefunction ($\Psi(r,t)$) (up to a time-dependent phase factor) was laid by the Runge-Gross theorem [150]. The time-dependent wavefunction therefore is a functional of the time-dependent electron density:

$$\Psi(r,t) = \Psi[\rho(r,t)](t)e^{-i\alpha(t)} \tag{3-23}$$

The time-dependent electron density ($\rho(r,t)$) of a real interacting system can be represented by the time-dependent electron density of a non-interacting reference system ($\rho_{ni}(r,t)$), similar to the approach from Kohn and Sham in DFT. This non-interacting system can then be represented by a single Slater determinant, which is constructed from time-dependent molecular orbitals ($\varphi_i(r,t)$).

$$\rho(r,t) = \rho_{ni}(r,t) = \sum_{i}^{N} |\varphi_i(r,t)|^2 \tag{3-24}$$

These molecular orbitals (MOs) are the solution of the time-dependent Schrödinger equation, which can be rewritten as the time-dependent Kohn-Sham (TD-KS) operator.

$$i\frac{\partial}{\partial t}\varphi_i(r,t) = \left(-\frac{1}{2}\nabla_i^2 - \sum_k^{nuclei}\frac{Z_k}{|r_i - r_k|} + \int\frac{\rho(r')}{|r_i - r'|}dr' + \frac{\delta A_{xc}[\rho]}{\delta\rho(r,t)} + V(r,t)\right)\varphi_i(r,t) \tag{3-25}$$

The individual parts are analogous to the time-independent KS operator. The kinetic energy is represented by the 1st term on the right hand side, the 2nd represents the electron-nuclei coulomb interaction and the 3rd the coulomb interaction between the electron and the charge distribution of all other electrons. The 4th term (the 'exchange-correlation' part of the so-called action integral, see [149]) contains, analogue to time-independent KS, all exchange and correlation effects and the last term represents any external potential. The time-dependent exchange-correlation (XC) term is unknown and is approximated by a standard DFT ground-state XC functional. This is done on the assumption of an adiabatic local density approximation (ALDA). The response of MOs, in ALDA, is slow compared to the transition from the ground-state to the excited state. The transition is therefore sudden, from an electronically relaxed ground-state into an electronically relaxed excited state. This permits to first calculate the excited state and afterwards the transition instead of requiring to simulate the transition though time. Instead of direct calculation of the excited states, TD-DFT calculates resonant transition frequencies (ω), which cause an infinite response in the TD-KS equations. In this way, the transition frequencies between the ground-state and excited states are found without need for explicit calculation of the excited states. The transition frequencies are calculated by solving the following nonstandard eigenvalue type equation [24, 149].

$$\begin{pmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{B} & \boldsymbol{A}\end{pmatrix}\begin{pmatrix}\boldsymbol{X}\\ \boldsymbol{Y}\end{pmatrix} = \omega\begin{pmatrix}\boldsymbol{1} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{-1}\end{pmatrix}\begin{pmatrix}\boldsymbol{X}\\ \boldsymbol{Y}\end{pmatrix} \tag{3-26}$$

A and **B** are themselves matrices and **X** and **Y** are transition vectors. The matrix elements of **A** and **B** for a transition from an occupied orbital (i) into a virtual orbital (a), for the case of a hybrid DFT functional, depends on the fraction of Hartree-Fock exchange in the functional (C_{HF}) and interactions with all other occupied (j) and unoccupied (b) orbitals.

$$A_{ia,jb} = (\varepsilon_a - \varepsilon_i)\delta_{ij}\delta_{ab} + (ia|jb) - C_{HF}(ij|ab) + (1 - C_{HF})(ia|\hat{f}_{xc}|jb) \tag{3-27}$$

$$B_{ia,jb} = (ia|bj) - C_{HF}(ib|aj) + (1 - C_{HF})(ia|\hat{f}_{xc}|bj) \tag{3-28}$$

The 1st term in (3-27) corresponds to the energy difference between origin orbital (φ_i) and target orbital (φ_a) and only appears on the diagonal of **A**. The remaining terms are the response of the exchange-correlation interaction to the excitation, which models the electronic reconfiguration of the electron density after excitation.

$$(ia|jb) = \int\int \varphi_i^*(r)\varphi_a(r)\frac{1}{|r-r'|}\varphi_b^*(r')\varphi_j(r') \tag{3-29}$$

$$(ia|\hat{f}_{xc}|bj) = \int\int \varphi_i^*(r)\varphi_a(r)\frac{\delta^2 E_{xc}}{\delta\rho(r)\delta\rho(r')}\varphi_b^*(r')\varphi_j(r') \tag{3-30}$$

The effective two-electron exchange-correlation integral in the 2nd and 4th term in equation (3-27) account for coupling between an excitation (i → a) with any other excitation (j → b) and the 3rd term adjusts for both excitations. The **B** matrix on the other hand includes coupling between an excitation (i → a) and a 'de-excitation' (b → j). The **B** matrix is therefore only of interest if the wavefunction at $t = 0$ already contains double excitations. This should not be the case for the ground-state and **B** is therefore neglected, which is the Tamm-Dancoff approximation (TDA). The remaining problem to be solved then is the diagonalization of **A**.

$$\boldsymbol{AX} = \omega\boldsymbol{X} \Leftrightarrow (\boldsymbol{A}-\omega)\boldsymbol{X} = 0 \tag{3-31}$$

This equation closely resembles the equation for a configuration interaction with only single excitations. The resulting excited state is therefore a linear combination of single excitation wavefunctions and can thus form pure singlet or triplet states (see Figure 3-2B). Solving this equation gives the transition amplitudes (or coefficients) (X_{ia}) and transition frequencies of the ctv excitation, which can then be used to calculate dipole and quadrupole transition moments. Inclusion of quadrupole allowed transitions is necessary as, depending on coordination and spin state, both dipole and quadrupole 1s → (3d/4p) transitions can be (partially) allowed. The inclusion of ALDA and TDA does not include any double excitations, which have a negligible contribution as the large energy difference between metal(1s) MO and virtual MOs should dominate any (several orders of magnitude smaller) energy difference between metal(2s) MO or valence MO and virtual MOs [24]. The absolute transition energies are strongly affected by the applied functional due to the core potential [24]. The relative energies vary less as the behavior of the core potential cancels out and energy shifts have been found to be appropriate [24, 151]. Special attention has to be given to charge transfer states, as TD-DFT does not correctly model the long-range 1/R interaction. Analyzation of TD-DFT results poses its own challenges. While the origin orbital is well-defined as the core metal MO, the target orbital in the excited state can be somewhat elusive. The excited states themselves are not directly calculated but constructed from linear combinations of single excited ground-state MOs. If the excited state corresponds to one single-excited state or is largely dominated by one excitation

($X_{ia} > 0.9$) it is possible to evaluate the target MO from the virtual MO from the ground-state it was constructed from. This is not possible when the excited state is constructed by several single-excited ground-state MOs. For these cases the difference density is considered, which equates to the difference between the spin density of the excited state and the ground-state. The resulting difference density equals the spin density of the excited state target MO. The ground-state origin MO is strongly localized on the absorbing atom and therefore almost not visible. The nature of the target MO can then be assigned by visual inspection.

Conclusion

Density functional theory (DFT) has been the "golden standard" in bioinorganic chemistry. DFT is capable to yield accurate structural parameters via geometric optimization in an efficient manner even for large metal complexes. It is able to reproduce correct energetic spin state ordering even for transition metal compounds that include antiferromagnetic coupling (afc) and DFT has been proven to be a highly versatile tool in calculating experimental parameters from EPR spectroscopy, Mössbauer spectroscopy, IR spectroscopy and X-ray absorption and emission spectroscopy (XAS/XES).

XAS/XES transitions include large energy differences between excited states and the unperturbed ground-state. DFT-based methods are nonetheless able to calculate core-to-valence (ctv) and valence-to-core (vtc) transitions with good reliability. The vtc transitions in the Kβ satellite line emission are approximated as one-electron processes, utilizing the ground-state DFT wavefunction, without any electronic relaxation as a response to the perturbation. The ctv transitions in the pre-edge absorption correspond to calculated excited states from time dependent DFT (TD-DFT) and therefore do account for electronic relaxation.

Meaningful interpretation of key electronic parameters from DFT requires significant insight into the theoretical framework. The electron density in DFT can express the Hamiltonian of a system and, thus, be used to calculate the total energy. It is constructed from a single-determinantal model-system wavefunction (Kohn-Sham self-consistent field). The energy of the DFT wavefunction is approximately corrected for missing exchange and correlation interactions through an exchange-correlation (XC) functional. For transition metals, description of afc is facilitated by the broken-symmetry formalism, which remedies the restrictions of spatial symmetry within the DFT wavefunction. This approach usually results in the correct order of spin states but unphysical spin densities (e.g. for open-shell singlet states), which have been found to depend on the chosen DFT functional [152, 153]. However, analysis of the electronic configuration from DFT calculations is often based on the model-system wavefunction, which is not necessarily a suitable and correct description of the wavefunction of the real system. Therefore, conclusions on the electronic configuration based on DFT wavefunctions have to be evaluated with great care.

Here, (TD-)DFT was used for geometric optimization and spectral simulation of ctv and vtc X-ray spectra from metal cofactors and synthetic metal compounds, thereby facilitating quantitative benchmarking of calculated features, which leads to a deep understanding of the spectroscopic features and to reliable interpretation of electronic structures.

3.2 Complete active space self-consistent field

In density functional theory (DFT), the shape of an electron density is optimized by energy minimization. This electron density is constructed from a wavefunction of a fictitious non-interacting system. This approach with high computational efficiency has the drawback that the electron density does not include the same amount of information as a wavefunction, as it is not an eigenfunction of every quantum mechanical operator (especially spin). Analysis of DFT calculations often focuses on the DFT wavefunction but this becomes less reliable the more electron-electron interactions are present in the real system [154]. The Hartree-Fock (HF) method optimizes the energy of an approximated wavefunction itself and post-HF approaches include electron-electron correlation interaction with a multiconfigurational description of the wavefunction. This section is an introduction into the complete active space self-consistent field (CASSCF) approach. The theoretical framework is presented first and illustrated with the example of hydrogen molecule dissociation. The effective unpaired electron density is introduced as to analyze diradical character in CAS wavefunction. To conclude, CASSCF and EUED are used to describe the electronic configuration of ozone, which provided a theoretical framework for our investigation of heme proteins.

Introduction into multiconfigurational MO theory

The electrons in the Hartree-Fock method each interact with the static charge density distribution of all other electrons. Consequently, the wavefunction of a single many-electron system is approximated by the linear combination of many one-electron systems. The HF wavefunction can then be expressed as a single Slater determinant. The MOs are optimized using the energy from the Fock operator (similar to the KS operator), which returns the energy of each individual MO. In this way, a wavefunction itself is optimized which is an eigenfunction of quantum mechanical operators. It includes exchange interaction but does not include correlation (DFT does). Therefore, the difference between total energy and HF energy, in the HF limit of an infinite basis set, equals the correlation energy. Electron correlation can be included through optimizing the energy of the many-electron system using the Hamilton operator on the HF wavefunction instead of the Fock operator. The energy from the Hamilton operator is the lowest possible energy for a single-configurational representation. This represents the upper limit of the total energy of the real system, as the approximation of non-interacting electrons could theoretically be true. The expectation value of the Hamiltonian can be reduced further by constructing the wavefunction from more than one electronic configuration, e.g. degenerate states. This expanded wavefunction should be closer to the real solution according to the variational principle. There is no upper limit for adding more configurations other than the maximum number of permutations of states, thus advancing from a single-configurational self-consistent field (SCF) to a multiconfigurational SCF (MCSCF). The MCSCF wavefunction (Ψ_{MCSCF}) is a weighted linear combination of configuration states (ψ_i) with differing orbital occupation but equal total spin (S).

$$|\Psi_{MCSCF}^{S}\rangle = C_0|\psi_0^S\rangle + C_1|\psi_1^S\rangle + \cdots = \sum_k C_k|\psi_k^S\rangle \tag{3-32}$$

$$|\psi_k^S\rangle = \sum_l c_l|\phi_l\rangle \tag{3-33}$$

Each term in equation (3-32) represents a configuration with an individual factor (C_i) for weighting and normalization. The leading term is the HF wavefunction, which in most cases is the dominant configuration. The configuration states are constructed from a shared set of molecular orbitals (MOs) with MO coefficients (c_l) (see equation (3-33)). Thus, a minimization of the Hamiltonian expectation value needs to optimize the weight (C_k) of each state as well as the MO coefficients (c_l) and the energy gradient needs to vanish for each weight and coefficient. An accurate description of the electronic configuration would hence need all possible permutations of occupied/virtual MOs and optimize the entire MO space spanned by the basis set. This full configuration interaction (Full-CI) approach very quickly exceeds today's computational feasibility. Instead, not all MOs have to be included when trying to find configurations from which to construct the MCSCF wavefunction. Typically, core MOs from transition metals or MOs that correspond to atoms outside the second coordination sphere can be excluded. The same logic can be applied to the virtual orbitals. The wavefunction can therefore be segmented into three separate sections, the internal, external and active space. The internal space is constructed as a fully occupied restricted closed-shell system while the external space is the unoccupied closed-shell part. The active space corresponds to the linear weighted combination of state configurations. These were individually selected in early applications of this theory [155], which is only feasible for small molecules. The active space is instead constructed from all possible permutations of a selected number of electrons (n) in a selected number of orbitals (m) with equal total spin. This is called a complete active space, CAS(n,m), calculation. The orbitals in the active space can be occupied by zero, one or two electrons in each configuration. The final optimized orbitals do not have a unique energetic eigenvalue, which could be discussed in a similar fashion as canonical molecular orbitals (CMOs) in DFT

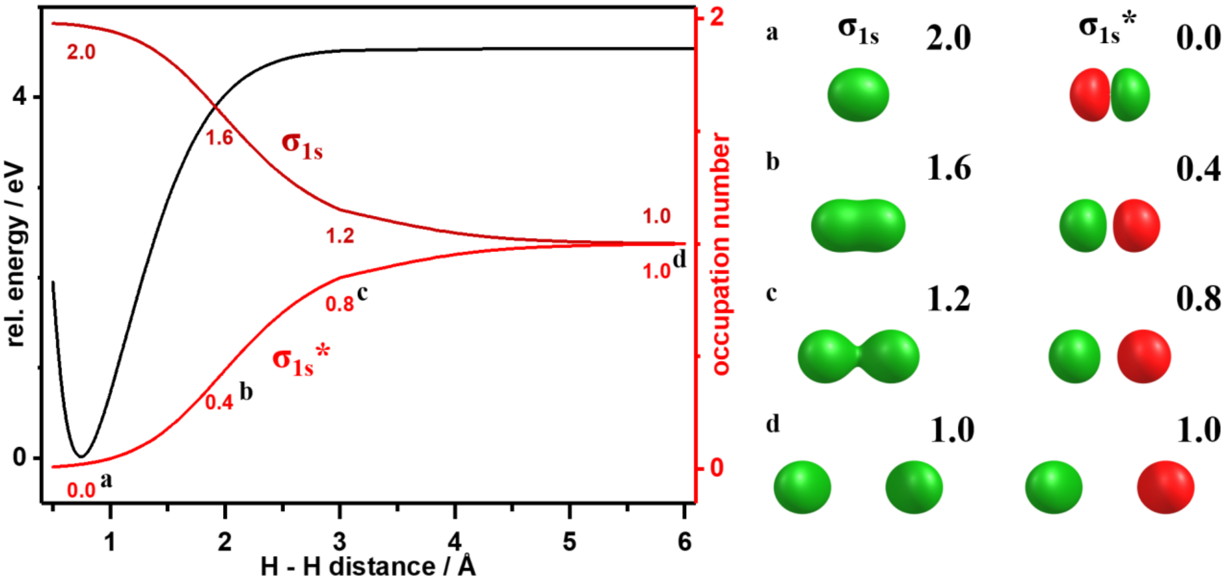


Figure 3-3: Potential energy surface (PES) scan for two hydrogen atoms from CAS(2,6) calculations with natural molecular orbitals (NMOs) in the active space. Shown are the relative total energy (black, left axis) and the occupation numbers (right axis) of the bonding σ_{1s} (dark red) and antibonding σ_{1s} (red) MO. Selected NMOs along this PES are shown on the right, with their corresponding distance and occupation numbers highlighted in the PES plot. Note that NMOs cannot break spatial symmetry. The gas-phase calculation was performed with the TZVP basis set and a large grid (GRID5) in ORCA 3.0.3. Isosurfaces are drawn at ± 0.04 au.*

or HF theory. Instead, the discussion turns to the occupation number (*ON*) of an orbital (*l*) in the active space.

$$ON[l] = \sum_k (on)_l C_k^2 \tag{3-34}$$

The occupation number of an orbital is the weighted sum of the occupation in all configurations and corresponds to the probability to finding an electron, independent of its spin, in this orbital. The sum of occupation numbers of all orbitals in the active space equals the number of electrons in the active space. Consequently, a single-configurational HF or CASSCF wavefunction results in integer occupation numbers. The occupation numbers are non-integer if more than one configuration is present in the resulting wavefunction. Thus, occupation numbers of orbitals in the active space of a CASSCF calculation can cover the entire range from more occupied MOs, with occupation numbers close to 1 or 2, to more unoccupied MOs (ON ≈ 0). One way to guide interpretation of a MCSCF wavefunction is to look for correlating MOs, which have a combined occupation number equal to 1 or 2. Two correlating orbitals (a correlation pair) describe static electron correlation, which is the principal reason to perform CASSCF calculations. The inclusion of electron pair correlation describes the separation of electrons in an orbital and is a necessary quantity when describing molecule dissociation. An example for the bond dissociation of a hydrogen molecule is shown in Figure 3-3 and corresponding CAS state configurations in Table 3-1. In the hydrogen molecule at equilibrium distance (~ 0.75 Å), both electrons are most probably found in the bonding σ_{1s} orbital (ON[σ_{1s}] = 2) and the wavefunction is almost single-configurational (σ_{1s} state contribution ~ 100 %). Increase of the H – H distance increases the probability to find one electron on each atom (ON[σ_{1s}*] > 0, σ_{1s}* state contribution > 0), while the probability to find electrons between each atom at very large distances vanishes (ON[σ] = ON[σ*] = 1, contributions of σ_{1s} and σ_{1s}* state = 50 %) [155]. Increasing electron separation therefore yields higher occupation of antibonding orbitals. Accordingly, the CAS wavefunction becomes two-configurational with increasing H_2 bond length. The correlation between bonding and antibonding orbital hence reflects the separation of the electron pair. Note that the single-configurational wavefunction represents a closed-shell singlet state while the two-configuration wavefunction corresponds to an open-shell singlet state. If relative energies were reversed one would speak of antiferromagnetic coupling (afc). The bond dissociation energy of dihydrogen estimated by the CAS(2,6) calculation in Figure 3-3 is 4.54 eV and comes very close to the experimental literature value of 4.52 eV [156], especially compared to the estimate of 7.53 eV when using B3LYP.

Success of a multiconfigurational SCF (MCSCF) calculation such as CASSCF highly depends on the selected active space. The first aspect is that each included "occupied" MO does have at

Table 3-1: Effect of H_2 dissociation on weight of state configurations in CAS wavefunction

	CAS state contribution [%]			
H – H distance / Å	0.75	1.95	3	5
σ_{1s} state	98	79	58	50
σ_{1s}* state	2	21	42	50

Show is the relative weight of CAS state configurations contributing to the multiconfigurational wavefunction from CAS(2,6) calculations at varying H_2 bond distances. The CAS wavefunction is essentially single-configurational at shorter bond lengths and gradually becomes two-configurational at larger distances. While the first configuration (σ_{1s} state) consists of a double-occupied bonding σ and empty antibonding σ MOs, the second (σ_{1s}* state) has two electron in the σ* and none in the σ MO. Occupation numbers, relative energies and MOs are shown in Figure 3-3.*

least another “unoccupied” MO to form a correlation pair. Otherwise, a completely occupied or empty MO forms part of the active space, which severely hinders convergence. If a good selection of MOs is included, all occupation numbers range from at least 0.02 to 1.98. Another aspect are the influence of initial guess orbitals. The SCF calculation must find the optimal set of configuration weights and MO coefficients, which makes the energy gradient very unstable. An initial guess without optimal orbital shape is consequently likely not to converge or converge to a local minimum. The last part is the factorial growth of the number of state configurations, which have to be considered for CAS(n,m) calculation. This leads to an upper limit regarding feasible computations, although this limit is constantly increasing with growing computational power.

Effective unpaired electron density

The distribution of unpaired electrons in high-spin molecules can be assigned by comparing the spin-up and spin-down molecular orbitals (MOs) of an spin-unrestricted wavefunction. The number of unpaired spins corresponds to the difference in occupied spin-up and spin-down MOs and the localization of unpaired occupied MOs can be addressed by unrestricted Hartree-Fock corresponding orbitals (UCOs, see section 3.4). However, the total spin of a singlet wavefunction is zero at all points in space and integrated over the entire wavefunction. This makes it difficult to interpret the local electronic configuration of a singlet state with antiferromagnetic coupling (afc). In density functional theory (DFT), analysis of the unsymmetrical distribution of spin-up and spin-down density yields the amount and localization of diradical character in a DFT wavefunction with broken spatial symmetry. In a complete active space self-consistent field (CASSCF) calculation, multiconfigurational singlet states are composed of a superposition of several state configurations each with S = 0 and, thus, no unsymmetrical spin distribution regardless of afc. This section presents the effective unpaired electron density (EUED), which yields the amount and localization of diradical character in a multiconfigurational wavefunction.

When Löwdin presented his work about density matrices, he noted about the one-particle reduced density matrix ($\rho^{(1)}$) that it can provide a sufficient criterion to prove if a wavefunction can be expressed as a single slater determinant [3]. This would be the case if the eigenvalues (the occupation number) of its eigenorbitals (the natural molecular orbitals, NMOs) are either unity or zero. Takatsuka et al. used this realization in reverse and defined a density function, using the one-particle density matrix, which characterizes the spatial separation of spin-up and spin-down electrons [157].

$$D(\boldsymbol{r},\boldsymbol{r}') = 2\rho(\boldsymbol{r},\boldsymbol{r}') - \int \rho(\boldsymbol{r},\boldsymbol{r}'')\rho(\boldsymbol{r}'',\boldsymbol{r}')d\boldsymbol{r}'' \tag{3-35}$$

This function calculates the distribution of 'odd' electrons and the trace of the matrix yields the total number of effective unpaired electrons (N_D) [158, 159].

$$N_D = tr\ D(r,r') \tag{3-36}$$

The distribution can be simplified significantly if the wavefunction is expressed by natural molecular orbitals (NMOs, Φ_i) as they diagonalize the one-particle density matrix.

$$\rho(r,r') = \sum_i n_i \phi_i(r)\phi_i^*(r') \tag{3-37}$$

Inserting equation (3-37) into equation (3-35) yields:

$$\begin{aligned} D(r,r') &= 2\sum_i n_i \phi_i(r)\phi_i^*(r') - \int \sum_i n_i \phi_i(r)\phi_i^*(r'') \sum_j n_j \phi_j(r'')\phi_j^*(r')\, d\boldsymbol{r}'' \\ &= \sum_i n_i(2-n_i)\phi_i(r)\phi_i^*(r') \end{aligned} \tag{3-38}$$

All terms with natural orbitals (ϕ_i, ϕ_j) with $i \neq j$ vanish since the NMOs are orthonormal. Therefore, when expressing a wavefunction with NMOs, all off-diagonal elements in equation (3-35) vanish and the trace in equation (3-36) can be derived from the NMO occupation numbers (n_i).

$$N_D = \sum_i n_i(2-n_i) \tag{3-39}$$

The number of effective unpaired electrons, or 'odd' electrons, (N_D) is ≥ 0. It can only be zero if the occupation number for each NMO is zero or two. A wavefunction can therefore be expressed by a single Slater determinant if the total number of 'odd' electrons is zero and all natural orbitals are doubly occupied or empty. If a singlet wavefunction has a positive number of 'odd' electrons, it cannot be expressed by a single Slater determinant (see ozone example in Figure 3-4 and Figure 3-5). Thus, at least two natural orbitals must have fractional occupation to a certain extent.

The EUED function provides a theoretical indication to the spatial separation of electron pairs in a given molecular system. It originates either from non-zero spin density (in non-singlet states) or from correlation effects [160]. Because it uses natural molecular orbitals, it is an intrinsic part of a wavefunction and does not depend on any localization criteria. The EUED is equal to the spin density ($\rho_u(r) = \rho_\alpha(r) - \rho_\beta(r)$) in the case of a restricted open-shell wavefunction

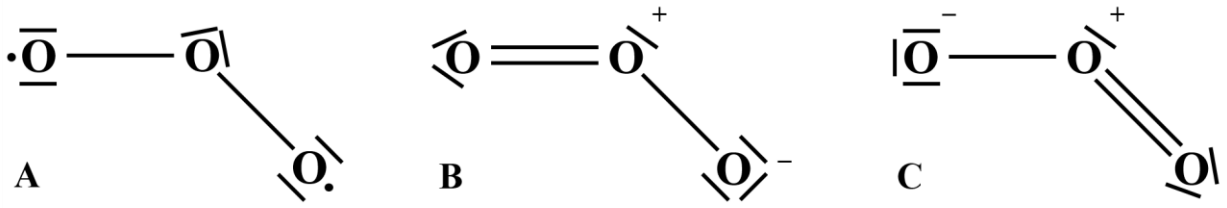


Figure 3-4: The bonding structures and electronic configurations of ozone. Note that only (B) and (C) represent resonant Lewis structures.

[159]. It can be assigned to different atoms in a molecule by using e.g. a Mulliken, a Löwdin or a hard-shell weighting. The number of 'odd' electrons is only comparable for similar systems and has been used to characterize chemical reactions or bonding situations [160-162]. The theoretical boundary for the total number of 'odd' electrons (N_D) for a system with N electrons is $0 \leq N_D \leq 2N$. The upper boundary is mathematically possible but apparently physically implausible. It would correspond to a correlated wavefunction without unoccupied orbitals and therefore a highly 'diluted' state [163].

The ozone molecule as a multiconfigurational example

Ozone is a good example for a multiconfigurational wavefunction [155]. It has a triangular structure with a O – O bond distance of 1.27 Å and a O – O – O bond angle of 117° [18]. Its bond distance lies in between the double bond of O_2 (1.21 Å) and the single bond of H_2O_2 (1.47 Å). This symmetry and the singlet ground-state allows for several possible configurations (see Figure 3-4), which are comprised of one diradical (**A**) and two ionic configurations (**B**, **C**). It is not obvious which of these competing configurations is the dominant one or even if one exists. This is a perfect example for a multiconfigurational complete active space self-consistent field (CASSCF) calculation and especially topical for our work on heme proteins due to the proposal of Goddard and Olafson to describe the bonding situation in dioxygen bound hemoglobin in terms of an ozone configuration [59, 60]. This section therefore presents CASSCF calculations for ozone.

The result of a CASSCF calculation for ozone is shown in Figure 3-5, which is a geometry optimization of a CAS(8,7) wavefunction for the singlet ground-state. The active space includes bonding and antibonding σ orbitals (**a** – **d**) as well as a non-bonding π orbital describing lone pairs on the two terminal oxygen (**f**) and bonding and antibonding π /π* orbitals parallel to the plane spanned by the three oxygen atoms (**e**, **g**). The three π orbitals are expected to be the most affected by electron correlation due to the three resonance structures. The configuration of the

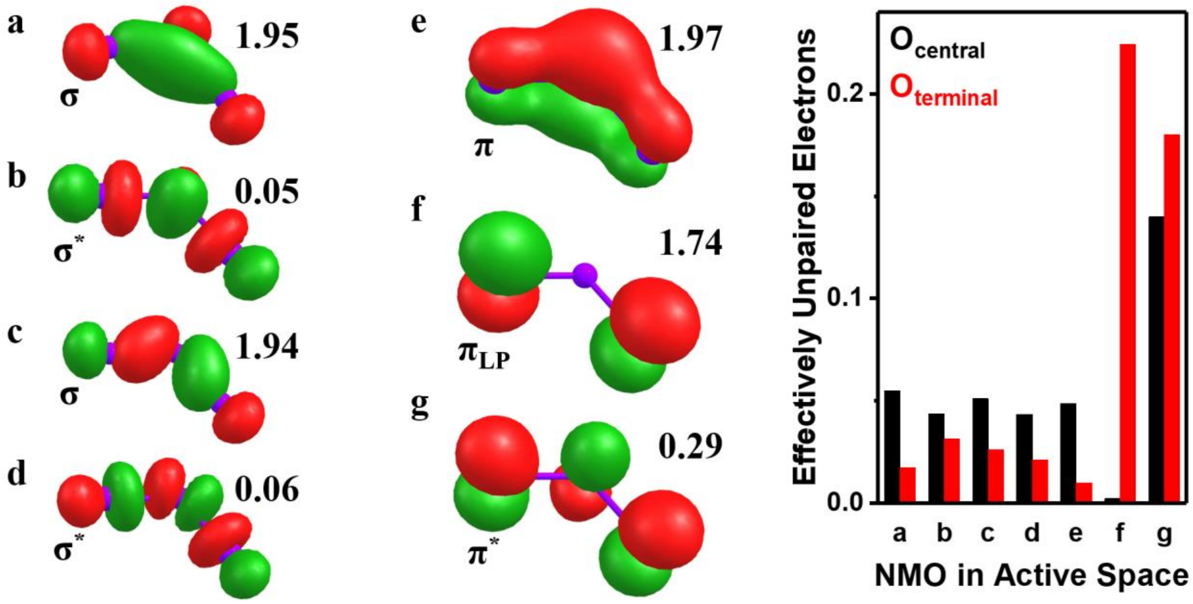


Figure 3-5: Example for a CASSCF calculation and the effective unpaired electron density (EUED). Shown is the active space (as natural molecular orbitals (NMOs)) of a CAS(8,7) calculation of a singlet ozone. Oxygen atoms and bonds are shown in purple for clarity. The configuration of the three largest contributing state configurations are shown in Table 3-2. Occupation numbers are shown for each NMO, which are drawn at an isovalue of ± 0.1 au. The geometry of the ozone molecule was optimized using the CAS(8,7) energies. The result is an atom distance of 1.29 Å (lit: 1.27 Å) and a bond angle of 115° (lit: 117°). The EUED shown on the right are calculated according to equation (3-39) and assigned with a Löwdin weighting. Each terminal oxygen has an EUED of 0.51 and the central oxygen an EUED of 0.39. The total number of 'odd' electrons in the active space is 1.41.

three largest CAS state configurations contributions to the wavefunction are shown in Table 3-2. All bonding or non-bonding MOs are double-occupied in the first state. The second largest configuration corresponds to a double excitation from the non-bonding π MO into the antibonding π* MO. This way, both π and π* orbitals, which span all three oxygen atoms, are double-occupied. The ionic bonding structure is represented by the first state while an increase in diradical character results from an increase in the weight of the second configuration. The remaining CAS state with weights larger than 0.4% result from two single or one double excitation from σ MOs **a**/**c** into σ* MOs **b**/**d** and/or (non-)bonding π MO **e**/**f** into π* MO **g** (see Figure 3-5). The resulting wavefunction is therefore clearly multiconfigurational with two mayor states (~90%) and a large amount of minor states which, as a whole, make up a significant contribution (~ 10 %). This multiconfigurational self-consistent field (MCSCF) can be interpreted as a two-configurational SCF (TCSCF). This would be difficult to represent by a single-determinantal method, which favors the largest contributing (ionic) configuration.

Analyzing the wavefunction with help of occupation numbers (ON) of natural molecular orbitals (NMOs) yields correlation partner orbitals. The ON sum of partner orbitals has to be close or equal to a multiple of two and thus, the total occupation number of corresponding electron pairs and each correlation group represents the charge separation due to static electron correlation.

Each bonding σ orbital (**a**, **c**) forms a correlation pair with one antibonding σ* orbital (**b**, **d**). The static electron pair correlation (and thereby charge separation) for the NMO pair **c**/**d** is larger compared to **a**/**b**, presumably because electrons in **c** orient along the O – O bonds while electrons in **a** cover the space spanned by the O – O – O angle. The remaining four electrons are distributed over the three remaining molecular orbitals. This is the '*three-center four-electron pi bond*' Goddard and Olafson referred to in their work about oxygen binding in hemoglobin [59, 60], when expressing the ozone wavefunction as NMOs. Be aware that the correlation between bonding π MO (**e**) and the non-bonding π MO (**f**) and antibonding π* MO (**g**) arises from the 10 % contribution from the large number of low-weight CSFs, which would be excluded when treating the wavefunction as a TCSCF.

Table 3-2: Configuration and weight of CASSCF states for singlet ozone.

CAS state weight	CAS state configuration						
	a	b	c	d	e	f	g
0.81	2	0	2	0	2	2	0
0.11	2	0	2	0	2	0	2
0.01	2	0	1	1	1	2	1
Total occupation numbers	1.95	0.05	1.94	0.06	1.97	1.74	0.29

*Shown are the three largest contributions to the wavefunction of the CAS(8,7) calculation and the total occupation numbers of resulting natural molecular orbitals (NMOs), which are shown in Figure 3-5. The NMOs are sorted in terms of correlation partners, which are two pairs of bonding and antibonding σ MOs (**a**/**b** and **c**/**d**) and a trio (**e**, **f**, **g**) with a non-bonding π orbital describing lone pairs on the two terminal oxygen atoms and a bonding and antibonding π/π* orbitals reaching over all three oxygen atoms.*

Further insight into the electronic configuration can be gained by analyzing the effective unpaired electron density (EUED) described in the previous section. The total EUED of ozone is 1.41 and mainly arises from the four-electron π bond (1.02), the majority of which come from the lone-pair MO (**f**, EUED = 0.45) and the π* MO (**e**, 0.50). The two σ bond correlation pairs contribute with 0.18 (**a**, **b**) and 0.21 (**c**, **d**) to the diradical character, which makes up about 30 % of the total EUED and shows the importance of including the large amount of individually small contributions into the active space.

Distributing the odd electrons according to a Löwdin population analysis leads to an EUED of 0.51 on each terminal oxygen and 0.39 on the central oxygen. This is completely in line with the resonance Lewis structure of ozone, which involves two electronic configurations in which the central oxygen is positively charged while either one of the two terminal oxygen is negatively charged (see Figure 3-4B and C). The EUED therefore successfully depicts the resonance hybrid and can bridge the complex bonding situation in a MCSCF wavefunction and formal chemical interpretations. Note, that the number of 'odd' electrons cannot be negative.

The population analysis for different geometry optimizations with HF, DFT and CASSCF are shown in Table 3-3. It shows a significantly larger charge separation for a selection of single-determinantal methods compared to the multiconfigurational CASSCF calculation. The amount of charge separation, and therefore ionic character, decreases from HF to HF-DFT hybrid to DFT to CASSCF. This results in smaller bonding angles (less repulsion between the two

Table 3-3: Comparison of atomic charge distributions from Hartree-Fock (HF), DFT with hybrid (B3LYP) or GGA (BP86) functional and CASSCF calculations for singlet ozone.

Values from	atomic charge on $O_{central}$	atomic charge on $O_{terminal}$	$O_{central}$ – $O_{terminal}$ distance /Å	O – O – O angle
HF	0.32	-0.16	1.20	119°
DFT(B3LYP)	0.25	-0.12	1.26	118°
DFT(BP86)	0.23	-0.12	1.28	118°
CAS(8,7)	0.19	-0.10	1.29	115°
Experiment [18]	-	-	1.27	117°

Tight geometry optimizations were performed using the TZVP basis set and atomic charges are from a Löwdin population analysis. HF and DFT calculations resulted in pure spin states (no net spin density). For further discussion of the CASSCF calculation, see Table 3-2, Figure 3-5 and text.

terminal oxygen) and larger $O_{central} – O_{terminal}$ distances (less attraction between central and terminal oxygen). However, both pure and hybrid DFT functionals result in geometries closer to experiments than HF and CASSCF, despite the supposedly more accurate expression of the ozone wavefunction by CASSCF. Note, that the DFT wavefunction does not need to be 'correct' for a correct electron density. The results from CASSCF can be improved by including dynamic electron correlation, which is included in DFT, through perturbation theory e.g. CASPT2 or NEVPT2 [77, 164]. However, performing geometry optimizations with e.g. NEVPT2 is difficult as there are no analytic gradients available (yet).

Conclusion

Molecular orbital (MO) theory is an *ab initio* approximation to the quantum mechanical wavefunction. Complete active space self-consistent field (CASSCF) calculations introduce static electron correlation to MO theory, which well describes effects such as charge separation within a pair of (anti-)bonding orbitals and therefore antiferromagnetic coupling (afc). Such correlation is represented by a weighted superposition of different state configurations within an active space, thereby resulting in a multiconfigurational wavefunction. The superposition yields fractional occupation numbers of orbitals in the active space, which is an advanced method to characterize interatomic bonding.

Singlet states in MO theory yield no net spin density in a wavefunction regardless of electronic configuration such as ionic bonding or afc. However, the diradical character in a CASSCF wavefunction can be represented by the effective unpaired electron density (EUED). Stemming from fractional occupation numbers in natural molecular orbitals, the EUED by Takatsuka, et al. [157] represents unpaired ("odd") electrons that are due to static electron correlation.

Here, the CASSCF method was implemented in the laboratory and is outlined and applied to well-established examples (H_2, ozone), defining an interpretational framework for the later on treated heme systems. As an example, it is shown that the total charge separation through dissociation of dihydrogen yields unity occupation numbers in a bonding σ and antibonding σ* orbital, which is the result of two state configurations with equal weight. The diradical character resulting from afc is represented similarly. In the ozone example, the active space shows static correlation between four electrons distributed in a bonding $\pi(O_3)$, antibonding $\pi^*(O_3)$ and non-bonding π orbital that describes lone pairs on the terminal oxygen atoms. This situation can be considered as a four-electron three-orbital electronic configuration.

CASSCF with a large active space as necessary to describe complex electronic configurations in transition metal complexes is challenging in terms of computational time and performance but also on a conceptual level. CASSCF is not a “black box“ and consequently requires careful preparation and selection of active space orbitals (see section 3.3). Therefore, preceding optimization of geometrical parameters with density functional theory and subsequent calculation of a multiconfigurational CASSCF wavefunction is recommended.

In this work, the combination of CASSCF with advanced X-ray spectroscopy was the key to resolve the long-standing controversy of the nature of Fe – O_2 bonding in heme proteins.

3.3 Computational procedures in detail

This section provides a detailed description of how quantum chemistry calculations in this work were performed. The choice for level of theory applied at each individual step is laid out and explained.

Starting geometries for calculations on heme species

A large number of crystallographic structures for hemoglobin (HB) and myoglobin (MB) proteins at very good resolution below 1.5 Å are available in the RCSB protein data bank (PDB). HB and MB samples turned out to be very sensitive to X-ray exposure (see section 5.5) with rapid changes in spin state and binding or unbinding of ligands. It is therefore reasonable to expect that the XRD structures might be influenced by radiation damage as X-ray crystallography is conducted at a photon flux at least ten times larger than in this work. Different starting geometries were therefore compared in DFT simulations. All Heme B sites present in a PDB structure were extracted together with the axial and distal imidazole compounds of the respective histidine residues as can be seen in Figure 3-6. The side chains of the porphyrin ring were removed and replaced by hydrogen atoms. The histidine was protonated as a N3-H tautomer [14, 15]. A crystal structure from the cambridge structural database (Refcode: NIWLUA) was used as an initial guess for the Fe(II) low-spin (LS) model compound ($Fe(TPP)(Py)_2$, **1**) [22]. No entries are presently available for the Fe(II) high-spin (HS) ($Fe(TPP)(THF)_2$, **2**) and Fe(III) HS (Fe(TPP)Cl, **3**). Compound **1** was therefore used as the basis for compounds **2** and **3** (see Figure 3-6) and altered with the GaussView 5 editor. The coordinate system was chosen with the x-y plane parallel to the porphyrin with each axis pointing along an Fe – N_{porph} bond, which set the z axis parallel the axial ligands.

DFT geometry optimizations

Starting geometries were optimized via spin-unrestricted DFT calculations using the ORCA 3.0.3 [165] and the Gaussian09 [166] program packages. Geometry optimization of HB/MB XRD structures were performed with ORCA while the porphyrin model compounds were optimized using Gaussian. Optimizations with Gaussian were done with the B3LYP hybrid functional [138] and Ahlrichs triple-ζ-valence with polarization (TZVP) [167] basis set without any solvation model. Optimizations with ORCA were conducted using approximations for the coulomb integral (J, resolution of identity (RI) approximation, auxiliary basis set TZVP/J) and the HF exchange integral (K, chain of spheres (COSX) approximation) and, if not stated otherwise, simulating solvation in water via the conductor like screening model (COSMO)

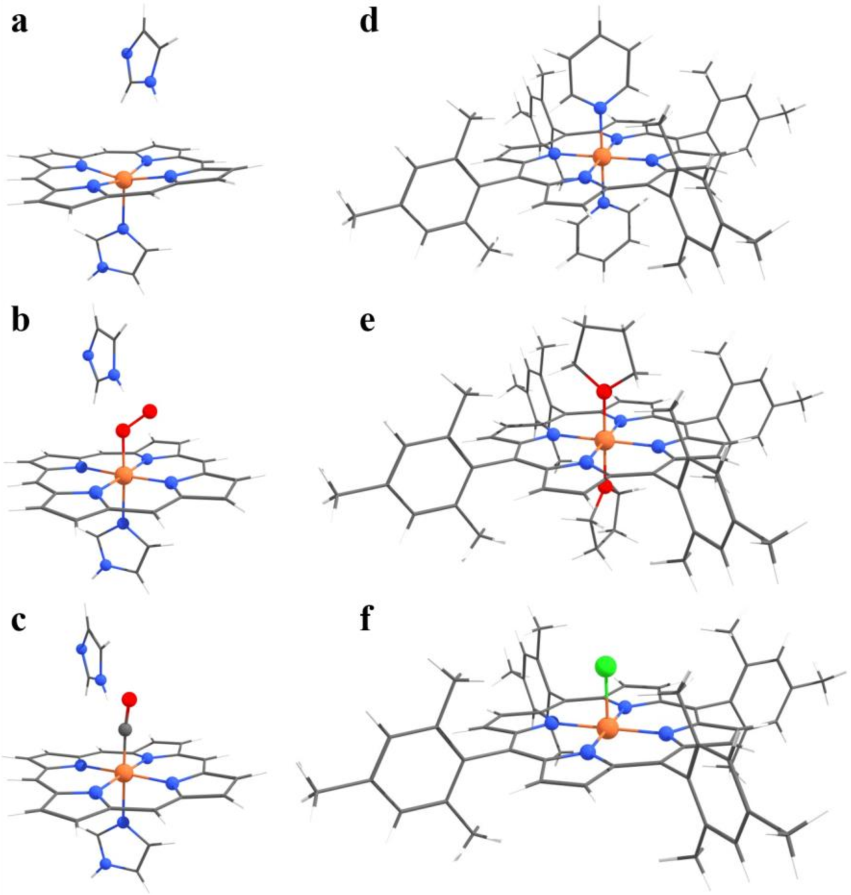


*Figure 3-6: Exemplary starting geometries for performing quantum chemical simulations. (**a**-**c**) Geometries shown are from crystallographic structures of human hemoglobin with 1.25 Å resolution. The individual structures are from hemoglobin in its deoxy form (**a**) (pdb code: 2DN2/ chain identifier: B / residue sequence number: 147), oxy form (**b**) (2DN1/ A/ 142) and carbonmonoxy form (**c**) (2DN3/ A/ 142) [7]. The Fe(II) low-spin porphyrin model complex (Fe(TPP)(Py)$_2$) (**d**) is from the Cambridge structural database (CSD) (Refcode: NIWLUA) [22] while the Fe(II) high-spin (HS) (Fe(TPP)(THF)$_2$) (**e**) and Fe(III) HS (Fe(TPP)Cl) (**f**) were constructed from (d) with GaussView 5 (TPP = tetraphenylporphyrin, Py = pyridine, THF = tetrahydrofurane). Oxygen atoms are shown in red, iron in orange, nitrogen in blue, carbon in grey, chloride in green and hydrogen atoms are indicated as white sticks.*

[168-170]. Carbon atom coordinates of the porphyrin ring as well as carbon and nitrogen atom positions of the distal histidine were constrained for heme optimizations. All carbon atom coordinates were constrained for the compound 2. Calculations for the **oxy** state were done with and without broken-symmetry formalism. Tight convergence tolerances and default DFT and HF exchange grids were used. Potential energy surface (PES) scans were done with equal settings but the initial guess geometry was pre-optimized using the the BP86 functional [134].

Calculation of ctv and vtc spectra with DFT and TD-DFT

Ctv spectra from TD-DFT and vtc spectra from DFT were calculated (see section 3.1) with the improved Ahlrichs TZVP (def2-TZVP) [171] basis set, the RI and COSX approximation on a dens DFT integration grid (ORCA Grid4, 4x more grid points than default gird). For TD-DFT, the minimum number of excited states to calculate can be estimated from ligand field theory. In the case of heme in its carboxy form (Fe(II) LS in an octahedral geometry) two transitions into the unoccupied Fe d_{z2} and d_{x2-y2} orbitals and two transitions into the unoccupied CO π^* orbitals are expected. Accounting for the spin-unrestricted DFT calculation (independent excitation of α and β spins) results in 8 ctv transitions. The number of excited states calculated by TD-DFT was increased until all these transitions were found (usually maximum of 20).

Preparing the CASSCF calculations

CASSCF calculations were performed without RI and COSX approximations for the SCF step because the introduced 'noise' in the energy can interfere with the convergence[2]. However, the necessary active spaces for complexes investigated in this study were comparatively large. Iron alone contributes with a CAS(5,5) for Fe(III) (d^5) or CAS(6,5) for Fe(II) (d^6) active space and additional relevant electrons/orbitals accounting for ligands. The integral-transformation step prior to the multireference configuration interaction (MR-CI) calculation was therefore done with RI approximation and an auxiliary basis set with simultaneously J and K fitting. Reason for this was a smoother error in energy. In order to prepare for the intended CASSCF calculations first the geometries from prior PES scans or geometry optimizations were optimized again using DFT (B3LYP, def2-TZVP, def2-TZVP/JK) on a dense grid for DFT (GRID4) and HF exchange (GRIDX4) with RI and COSX options turned on.

The resulting structures and wavefunctions were then used for a single point MP2 calculation (Møller-Plesset second-order perturbation theory) to generate MP2 natural molecular orbitals ($MP2_{NMOs}$). The $MP2_{NMOs}$ are advantageous because of (i) an optimized orbital shape of occupied and virtual MOs (in contrast to DFT and HF) and (ii) the $MP2_{NMOs}$ are ordered according to their occupation number (in descending order from 2 – 0) and therefore weighted by their importance for a multiconfigurational wavefunction. The $MP2_{NMOs}$ were the starting point in generating an initial guess for the CASSCF calculation.

Performing CASSCF calculations

CASSCF is not a black box and successful convergence and/or meaningful results can only be achieved with substantial knowledge about the system in question. The most important step

---

[2] ORCA 3.0.3 manual and CEC Summer School 2016

when carrying out CASSCF calculations is the selection of initial guess MOs for the active space. Automated initial guesses will only be successful for very small systems (see e.g. H_2 in Figure 3-3 and ozone in Figure 3-5) and $MP2_{NMOs}$ are a good starting point. Any $MP2_{NMO}$ with an occupation number between 0.02 and 1.98 should be considered to be included in the active space as this represents a large amount of static correlation. However, this would mean that, in the case of **oxy**, a CASSCF calculation with 150 MOs and 184 electrons should be performed. As this exceeds available computational resources, another selection criterion was necessary. K-edge XAS and XES is focused on iron and therefore any MO included in the active space should be from iron, its first coordination sphere or ligands of special interest like the terminal oxygen in the **oxy** and **carboxy** state. The initial guess MOs were therefore localized molecular orbitals (LMOs) created with the utility program orca_loc provided by ORCA 3.0.3, which were inspected and selected for each CAS calculation. The range of $MP2_{NMOs}$ to be localized can be freely set and this permits to localize occupied and virtual $MP2_{NMOs}$ independent from each other. $MP2_{NMOs}$ are already ordered by occupation number and therefore weighted by the degree of electron correlation. The criterion for the orbital window was increased gradually, starting from the highest occupied MO, until the necessary number of one or two center bond-like orbitals of interest were produced. Figure 3-7 shows an example for such a procedure. The smallest orbital range (A) with $MP2_{NMO}$ occupation numbers in the MP2 wavefunction between 1.9 and 0.1 already includes two MOs representing π and π* Fe – O bonds formed from an $Fe(d_{xz/yz})$ and an $O_2(\pi^*)$ MO. The unoccupied $Fe(d_{z2})$ centered $MP2_{LMO}$ with a small $2p_z$ contribution from the $O_{central}$ atom misses an occupied counterpart to interact with. Remember, the active space is only stable if every included MO correlate with at least one other MO. The remaining $MP2_{LMO}$ (not shown) within this localization range are located on carbon atoms of the porphyrin. Increasing the orbital window in (B) reveals the $O_2(\sigma_p)$ and $O_2(\sigma^*_p)$ $MP2_{LMOs}$ with small contributions from the porphyrin nitrogen atoms in the $O_2(\sigma_p)$ and $Fe(d_{z2})$ $MP2_{LMO}$. These small contributions can cause so-called ‘intruder’ states, which drift into the active space and hinder convergence. The $Fe(d_{z2})$ $MP2_{LMO}$ is still missing a correlation partner. The next step therefore is to further increase the orbital window (shown in Figure 3-7C), which removes some or all porphyrin contributions. This range also contains a $MP2_{LMO}$ resembling $O_2(\pi_z)$, which is the missing correlation partner for $Fe(d_{z2})$. The π Fe – O bond splits into two $MP2_{LMOs}$, which does not result in complications as long as both $MP2_{LMOs}$ are included in the active space. This selection of four double-occupied and three empty $MP2_{LMOs}$ forms a good initial guess for

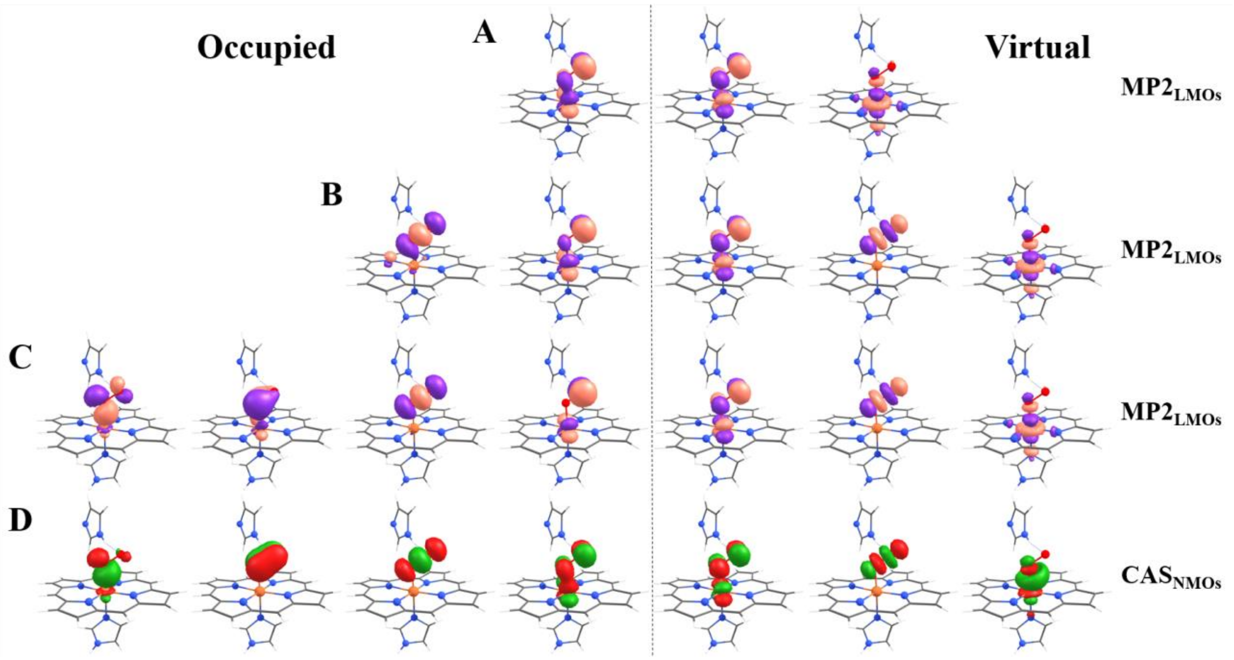


*Figure 3-7: Influence of varying orbital localization ranges on shape and number of localized molecular orbitals (LMOs) from MP2 calculations ($MP2_{LMOs}$). Only $MP2_{LMOs}$ are shown that are localized on the iron atom and/or the central and/or terminal oxygen atom (pink and purple). (**A - C**) All shown $MP2_{LMOs}$ were generated from natural molecular orbitals (NMOs) of the same previous MP2 calculation ($MP2_{NMOs}$, not shown). $MP2_{NMOs}$ are first localized in an orbital window ranging from occupation numbers (ONs) ≥ 1 to occupation numbers smaller than an upper threshold (ON_{up}), the occupied orbital range. The resulting, partially localized wavefunction is then localized between $ON_{low} < ON < 1$, the virtual range. Used thresholds (ON_{up}, ON_{low}) are (1.90, 0.1) for (A), (1.94, 0.05) for (B) and (1.97, 0.03) for (C). (**D**) Active space NMOs (red and green) from a CAS(8,7) calculation utilizing the $MP2_{LMOs}$ from (C) as initial guess ($Fe\text{-}O_{central}$=1.875 Å, see text for more information on calculations). The isosurfaces are drawn at ± 0.05 au.*

a CAS calculation. Each orbital has a counterpart to interact with in form of corresponding π, π*, σ and σ* MOs. All are localized on the three atoms of interest with almost no contributions from the porphyrin nor the axial or distal histidine. The number of CSFs in the complete active space is reasonable with 490 possible permutations for the singlet state in question. The generated wavefunction file was used as an initial guess and the selected $MP2_{LMOs}$ were rotated into the active space. The wavefunction of a converged CASSCF calculation was used as an initial guess for a subsequent calculation with a larger active space with more orbitals of interest. The former active space was used without alteration and the former internal and external space were localized independently (CAS_{LMOs}). This approach of multistep CASSCF calculations proved to be more successful than including all orbitals of interest at once.

Conclusion

Procedures for execution and interpretation of density functional theory (DFT) and complete active space self-consistent field (CASSCF) based calculations were designed, implemented and extensively applied during this work. In this respect, I am grateful to the lectures in the summer school "Physical Methods in Molecular and Heterogeneous Catalysis" of the Max-Plank-Institute for Chemical Energy Conversion (MPI CEC) who encouraged me in attempting the CASSCF approach. Combination of various commercial and freeware software tools (ORCA, Gaussian, GaussView and Chemcraft) with self-written scripts (Python, Fortran, Origin) was achieved, which was necessary to extract and visualize electronic structure properties from large output files of quantum chemical calculations. The results of the computational work are presented in detail in the following section.

3.4 A short note on orbital representation

The range of methods used in this work includes a large amount of underlying theory, which cannot be discussed completely or adequately during this thesis. There is nonetheless one topic, which has often caused confusion during discussions and therefore might need to be explored briefly. When presenting results from quantum chemical computations the discussion quickly centers on molecular orbitals (MOs). While the correct wavefunction is unique, the MOs which form it are not, as they are not a rigorous part of quantum mechanics. Although the wavefunction is constructed by MOs, any normalized linear combination of these MOs will not change the energy of the wavefunction nor the expectation value of any quantum mechanical operator. This allows an infinite number of MO representations, which density functional theory (DFT) takes advantage of. Any conclusion on the basis of MOs therefore needs to choose a representation suitable for the discussed topic. The MO representations used in this work and their possible application are thus briefly introduced in this section.

Canonical molecular orbitals

A converged wavefunction is constructed from a linear combination of MOs (φ_k) which are eigenfunctions of the Fock operator ($\hat{H}^F$) (in the case of Hartree-Fock (HF) Theory) or the Kohn-Sham operator ($\hat{H}^{KS}$) (in the case of DFT) and their eigenvalues are the MO energy (ε_k).

$$\hat{H}^{F/KS}|\varphi_k\rangle = \varepsilon_k|\varphi_k\rangle \qquad (3\text{-}40)$$

These canonical molecular orbitals (CMOs) characterise the first MO representation anyone probably encounters when performing a HF or DFT calculation for the first time [4, 6]. The CMO energies are of great interest for the interpretation and simulation of spectroscopic data. They can form the first approximation when discussing/calculating transition energies as in, but not limited to, XES transitions from DFT. One misconception when discussing CMOs is the concept of electron delocalization. CMOs are often spread out along the entire molecule (see Figure 3-8A), which does not imply that an actual electron would be delocalized along the entire molecule. The CMO isosurface represent spatial regions with equal probability to find an electron with the CMOs energy. However, two electrons of equal energy do not necessarily stem from the same atom/bond. As an example, the spin-down (β) CMO in Figure 3-8A has isosurface localized on the O_2 ligand, the distal histidine and opposing parts of the porphyrin. This does not reflect an electron that uses a combination of delocalization and quantum tunnelling but rather 2p MOs of equal energy. A CMO represents the converged linear combination of various basis functions to one energy, which does not invoke any hard restrictions from e.g. ligand field theory. Once the wavefunction has converged though, its

CMOs can be transformed according to different criteria. These transformed MOs can be more helpful in understanding various key properties of the wavefunction, depending on the transformation criteria.

Natural molecular orbitals

Natural Molecular Orbitals (NMOs) (Φ_k) of a wavefunction ($|\Psi\rangle = |\Phi_1\Phi_2\cdots\Phi_k\rangle$) are eigenfunctions of the one-particle reduced density matrix ($\rho^{(1)}$) and their eigenvalue is the occupation number (n_k) of the NMO [3]. The distinction between particle and electron is of importance because, as discussed in this section, a single MO, depending on the choice of its representation, not automatically corresponds to an electron.

$$\hat{\rho}^{(1)}|\phi_k\rangle = n_k|\phi_k\rangle \tag{3-41}$$

This is the simplest way to expresses the wavefunction on the basis of an independent-particle model [172]. These NMOs are an intrinsic part of the wavefunction and need no further criterion, hence the term ‘natural’. They are basis set independent, which is an advantage when comparing results. The eigenvalues (occupation numbers, n_k) represent the average number of particles in each of the NMOs. They range from $0 \leq n_k \leq 2$ for restricted and $0 \leq n_k \leq 1$ for unrestricted wavefunctions. The wavefunction can be expressed as a single Slater determinant if all occupation numbers take integer values in a multiconfigurational calculation. These NMOs cannot be assigned a single molecular energy but do represent a single particle. This is a large step closer towards chemical interpretation. The bottom NMO shown in Figure 3-8B is an example were the Fe – O_2 π bond is intrinsically assigned a single-particle character in the wavefunction.

One problem arises from the fact that NMOs are symmetry adapted just like the CMOs [173]. Take two oxygen atoms separated by a large enough distance, large enough to call them independent. Their NMOs e.g. 2p orbitals still have contributions on both atoms in a single orbital (see H_2 dissociation in Figure 3-3) while any chemist would draw them separately. An example for symmetry adaptation can be seen in the upper NMO in Figure 3-8. Here the four Fe – N σ bonds are combined into one symmetric Fe – N_{Porph} bond. This limits the possibility to assign a single NMO conceptually with a single electron or bond.

Localized molecular orbitals

The CMOs and NMOs presented above express different quantities of a wavefunction e.g. MO energy and MO occupation. Both representations typically feature MOs which spread over more than three atoms e.g. due to spatial symmetry. This renders them difficult for discussion

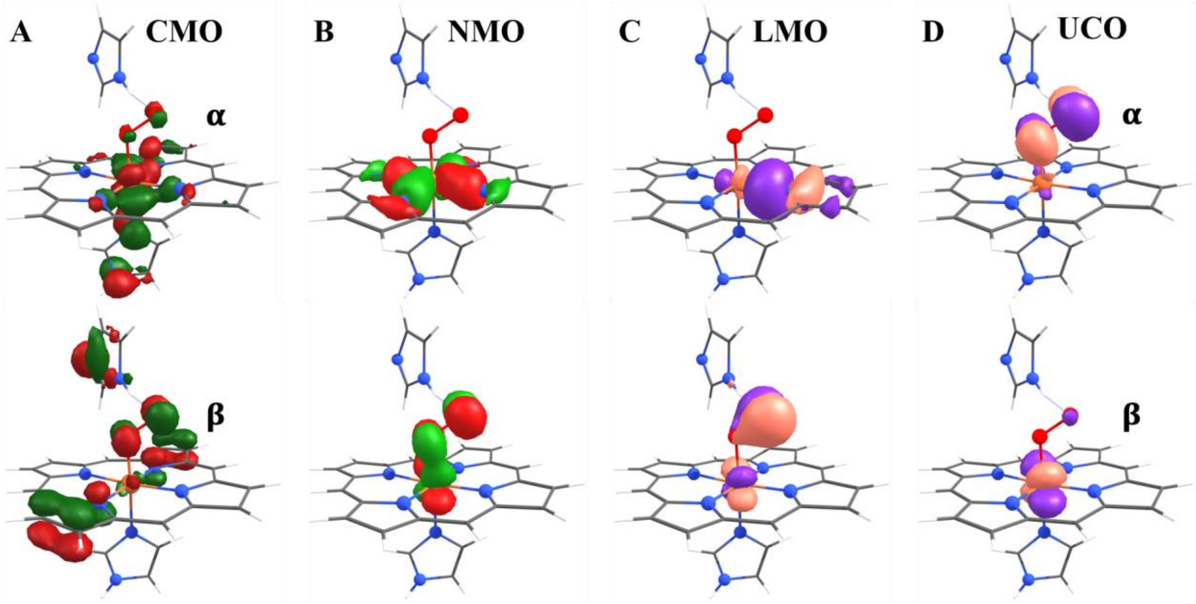


*Figure 3-8: Examples for different molecular orbital representations. Shown are (**A**) canonical molecular orbitals (CMOs), (**B**) natural MOs (NMOs), (**C**) localized MOs (LMOs) and (**D**) UHF corresponding orbitals (UCOs) [3-6]. Color code refers to the intrinsic (red and green) or external (pink and purple) restriction-imposing nature of a representation. CMOs are drawn with darker colors then NMOs. Shown pairs of CMOs, NMOs and LMOs stem from different MOs, whereas the UCOs are from the same MO. The overlap integral between α and β part of the UCO MO ≈ 0.3, indicating a broken-symmetry wavefunction (see text). The shown CMOs and UCOs have opposite spin. Isosurfaces are drawn at ± 0.05 au. All MOs are generated from the same geometry optimized DFT wavefunction. The Fe – $O_{central}$ distance was constrained to 1.875 Å, other details see text.*

of chemical concepts, which often involve three atoms or less. This is where localized molecular orbitals (LMOs) become necessary.

A vast range of different localization schemes has been developed. These can be separated into *ab initio* and *a priori* methods. The latter rely on extrinsic definitions of bonds, lone pairs and core orbitals and are therefore, to a certain extent, pre-determined e.g. natural bond order (NBO) analysis [173-175], which probably comes closest to a Lewis structure. In a NBO analysis core orbitals and lone pairs are localized first, followed by bonding and antibonding orbitals and, as a final step, Rydberg-like orbitals. *Ab initio* methods rely on an intrinsic constrain to optimize the final orbitals. Examples are minimal spatial extension for each MO [176] or maximum electron repulsion inside each MO [177], which result in very compact orbitals.

The method by Pipek and Mezey used in this work (provided by ORCA) minimizes the number of atoms where a MO is localized by means of minimizing the atomic overlap integrals between all MOs [178]. This criterion produces σ and π bonds and maximally distinct LMOs. The upper LMO in Figure 3-8D shows one of the four Fe – N_{Porph} σ bonds generated by this approach which were spread over all N_{Porph} atoms when expressed as NMOs. The lower LMO highlights the drawback of minimizing atoms per LMO. The three-atom NMO is separated into two diatomic NMOs after localization (only one shown), both involving an Fe($d_{xz/yz}$) and one O(p_y) (also compare Figure 3-7A and C). LMOs nonetheless provide an excellent means to produce MOs that reflect a more Lewis-like representation. They have been used in this work to produce chemically reasonable, efficient and stable initial guess MOs for CASSCF calculations. LMOs were the most reliable orbital representation, compared to CMOs, NMOs and other representations not mentioned in this work, for successful CASSCF convergence.

Unrestricted Hartree-Fock corresponding orbitals

A well-established way to investigate molecules involving transition metals with antiferromagnetic coupling (afc) are DFT calculations with broken-symmetry [153]. The resulting spin-unrestricted wavefunction contains semi-localized magnetic orbitals, which correspond to the interacting single-occupied molecular orbitals [5]. The magnetic orbitals are orthogonal in their spin but have, in general, little overlap in space. These orbitals are thus, in general, hard to identify if large molecules are involved and even more so when comparing CMOs. Unrestricted Hartree-Fock corresponding orbitals (UCOs) provide a way to extract these non-orthogonal magnetic orbitals [5].

The UCOs are optimized for a maximum overlap integral ($A_{\alpha\beta}$) between each spin-up (α) and spin-down(β) orbital. The generated UCOs fall into three different categories. In the first category are UCOs that have an overlap of unity and correspond to the closed-shells of a spin-unrestricted wavefunction. In the second are orbitals, which have no overlap at all because of α MOs without a β MO as counterpart. These represent the single-occupied MOs (SOMOs) of states with multiplicity > 1 where the number of α MOs exceeds the number of β MOs. The remaining MOs are spin up MOs with a spin down MO counterpart but with an overlap integral between zero and one. These are the interacting SOMOs of the broken-symmetry wavefunction. An example for this can be seen in Figure 3-8D. Shown are α and β UCOs of a heme **oxy** state with a broken-symmetry wavefunction. These are the only two UCOs with an overlap integral unequal unity ($A_{\alpha\beta} \sim 0.3$). This representation reflects in a clear manner that the broken-symmetry in the DFT wavefunction occurs between the O_2(π*) orbital and an Fe(d) orbital. The overlap integral of two interacting SOMOs can be used to quantify the diradical character (D_{BS}) of the broken-symmetry wavefunction.

$$D_{BS} = 1 - A_{\alpha\beta}^{2} \tag{3-42}$$

The diradical character approaches unity for a small overlap and therefore large spatial separation. This is indicative of a strong coupling between the two spatially separated centers. A large overlap (small separation) indicates a smaller coupling due to a smaller spatial separation. The overlap integral between these UCOs can be used to quantify the diradical or ionic nature of a broken-symmetry wavefunction. This procedure also extracts SOMOs of non-singlet states.

Conclusion

Quantum mechanical wavefunctions can be constructed from a set of molecular orbitals (MOs). These orbitals do not represent an observable and also are not unique as any normalized linear combination of a given set of MOs results in the same wavefunction. Consequently, different MO representations facilitate visualization of selected parameters of a wavefunction. The four MO representations used in this thesis were introduced here in detail.

Canonical molecular orbitals (CMOs) diagonalize the Fock/Kohn-Sham operator and have well-defined energies, which makes them a suitable first approximation to calculate transition energies and analyse $K\beta_{2,5}$ valence-to-core emission spectra. However, CMOs are often spread out over several atoms, which does not necessarily imply electron delocalization. ***Natural molecular orbitals*** (NMOs) represent the intrinsic definition of a spin particle in a wavefunction. While NMOs do allow characterization of interatomic bonding, they should not be interpreted as representing an electron. NMOs were used to prepare, perform and analyse CASSCF calculations. The NMO occupation number is the basis for effective unpaired electron density (EUED) calculation. ***Localized molecular orbitals*** (LMOs) result in maximally distinct MOs, most closely corresponding to a Lewis representation. Unlike CMOs and NMOs, which diagonalize intrinsic operators, LMOs can be localized with a variety of criteria to produce e.g. an accurate representation of π and σ bonds. As a result, LMOs allow interpretation of electron delocalization, which was used here to initially select MOs for CASSCF calculations. ***Unrestricted Hartree-Fock corresponding orbitals*** (UCOs) yields MOs, which show with maximal spatial overlap. This separates the open-shell valence MOs contributing in a broken-symmetry or restricted open-shell wavefunction and allows analysis of diradical character from MOs involved in antiferromagnetic coupling. A proper MO representation is required to visualize and interpret the key properties of a wavefunction (energy, spin, diradical character) for reliable interpretation of quantum chemical results.

4 A time-resolved energy-sampling approach for Kβ XES data collection

Investigation of the electronic structure of metals with varying ligand environments via XES has been a viable tool in the past. The metal spin states can be accessed by the Kβ$_{1,3}$ emission [23, 36] and insight into the interplay of ligand and metal centered orbitals can be gained from Kβ$_{2,5}$ emission [23]. However, because the Kβ$_{1,3}$ feature is one and the Kβ$_{2,5}$ feature two orders of magnitude smaller than the dominant Kα emission, for protein samples, with typical metal concentrations well below 10 mM, measurement of Kβ emission spectra represent an experimental challenge. Relatively low signal count rates of 10^2-10^4 s^{-1} for the Kβ$_{1,3}$ and 1-10^2 s^{-1} for the Kβ$_{2,5}$ emission are often observed (see Figure 4-1A). It is therefore important to decrease the scattering background as much as possible, which, in most cases, rules out the option of dispersive emission spectra collection. An alternative is an energy step-scan approach with multi-crystal-analyzer X-ray emission spectrometers. This approach provides the necessary low background count rate due to the high energy resolution, large solid angle detection and large distance between detector and sample. This comes at the cost of considerably longer times for recording of consecutive emission energies. Emission data measured later in the scan might therefore be influenced by X-ray exposure. This method also has a larger dead time due to involvement of many (crystal) motor movements of the spectrometer. XES is further complicated if samples are sensitive to radiation damage (see Figure 4-1B), i.e. photoreduction of metal centers in higher oxidation states [83]. This effect limits the maximal “safe” X-ray exposure period for each sample position, which necessitates repeated data collection and decreases the data acquisition efficiency due to accumulating dead time. This section presents an alternative approach, which combines time-resolved XES data collection with the superior signal-to-noise (S/N) ratio of a multi-analyzer crystal spectrometer and high acquisition efficiency at short X-ray exposure periods. The general procedure for conducting such experiments is outlined and the efficiencies of the conventional energy step-scan (CES) and the time-resolved energy-sampling (TRES) approaches are compared.

Time-resolved energy-sampling XES

XES on very dilute samples requires a high energy resolution and a low scattering background, which are provided by X-ray emission spectrometers with spherically bent analyzer crystals (i.e. focusing of the monochromatic emission in two dimensions on a small (1 mm^2) slit on the detector. Such spectrometers use an energy step-scan approach to collect emission spectra over

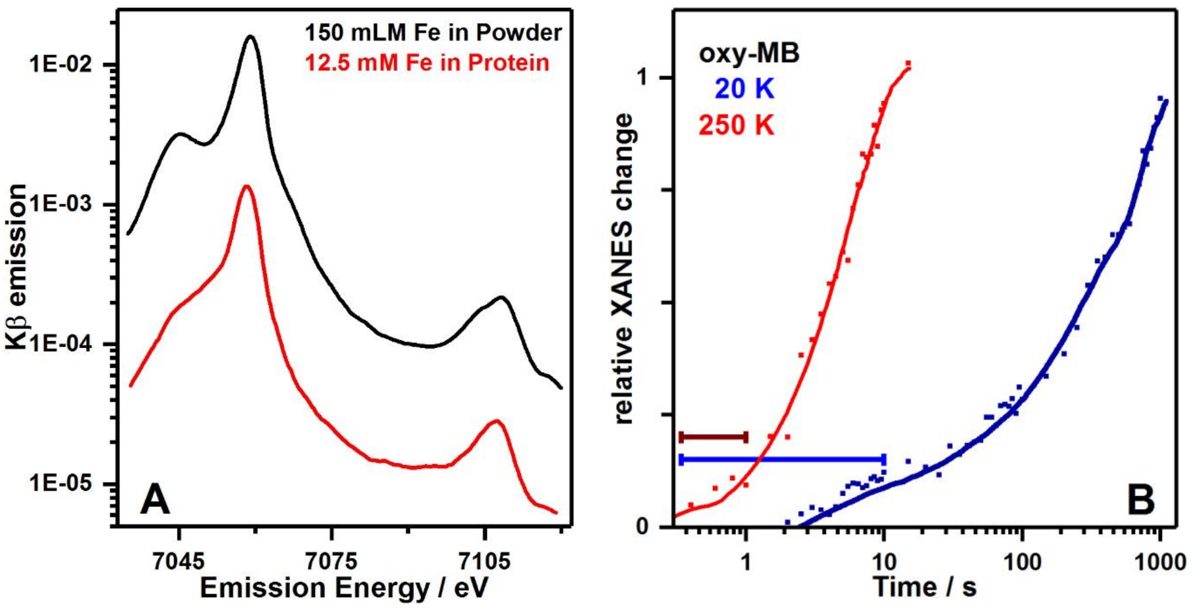


*Figure 4-1: Influence of metal concentration and sensitivity to X-ray exposure on Kβ XES (**A**) Kβ emission spectra from high-spin Fe(III) porphyrin reference compound (powder sample) and oxygenated hemoglobin protein in frozen solution. A more than one order of magnitude decreased emission intensity of the protein necessitates a 10-fold longer data acquisition period to obtain a similar signal-to-noise ratio as for the porphyrin. (**B**) Acceleration of radiation damage, caused by X-ray exposure, on oxy-HB due to a change in temperature, which is monitored by fluorescence changes at 7123 eV. The by two orders of magnitude faster sample damage at 250 K leads to a smaller "safe" X-ray exposure period (colored bars), which results in 100 times more energy scans, using the step-scan approach, to yield the same S/N at both temperatures. Lines represent moving average to guide the eye.*

a wide energy range. In the TRES approach, emission data are collected at a single energy on separate sample spots over a longer acquisition time, compared to conventional step-scan XES, and the complete emission spectrum is constructed from these data points. Due to sample thickness inhomogeneity, a second round of data acquisition on each spot at a constant reference emission energy is required. The advantage of the time-resolved energy-sampling approach is that at each emission energy and sample spot, the total "safe" X-ray exposure time can be used, which results in a significant increase in data acquisition efficiency. The individual experimental steps for TRES are as follows.

I – Determination of the "safe" X-ray exposure time

Radiation-damage occur in two steps, the 'primary' scattering event and 'secondary' diffusion phase [179]. In the primary step, a high-energy photon is absorbed or inelastically scattered. This induces a cascade of secondary electrons that in turn create free radicals, which move

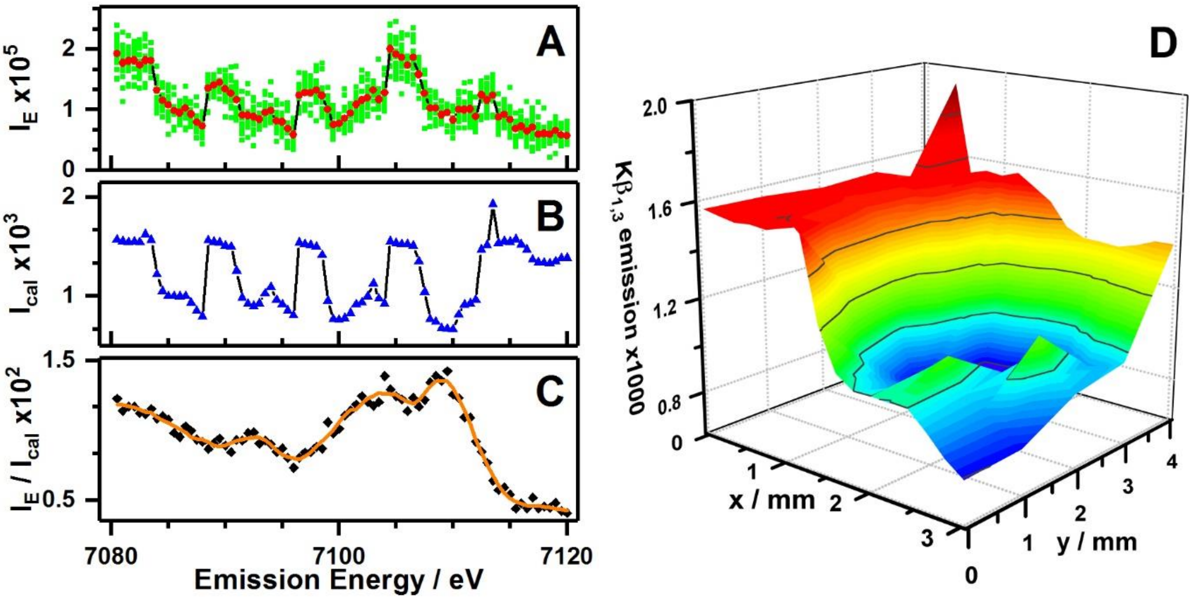


*Figure 4-2: Example of $K\beta$ XES measured with a time-resolved energy-sampling approach. Data were collected on a R2lox1 protein binding a di-iron cofactor from Geobacillus kaustophilus (4 mM). (**A**) X-ray fluorescence time traces (green squares) collected for 15 s at different emission energies and different sample positions (averaged data = red circles). (**B**) Fluorescence data used for amplitude calibration (blue triangles) measured for 10 s at 7059 eV on the same spot, shown for clarity at corresponding emission energies from (A). (**C**) $K\beta_{2,5}$ spectrum (black diamonds) (smoothed data = orange line). (**D**) Emission intensity variation from (B) across the sample area corresponding to local changes in metal concentration.*

along the protein backbone and break weak bonds. This process is expected to be relatively temperature independent as it occurs on timescales smaller than atomic motions. In the secondary step free radicals diffuse through the solvent and react with other parts of the protein. This step is proportional to solvent diffusion and therefore strongly temperature dependent above the glass transition temperature. As the rate of radiation damage on a sample is almost unpredictable, but usually faster for highly oxidized metal sites, it has to be quantified for each sample and temperature.

The shape and position of the metal K-edge in XAS is sensitive to structural (coordination) and electronic changes (spin, oxidation state) at metal cofactors in metalloproteins [25, 180, 181]. Monitoring of K-edge changes therefore provides a reliable way to determine the maximal X-ray exposure period for which radiation damage is absent. Detecting of radiation damage is achieved by measuring an intensity time course for example of total fluorescence (Kα emission) at fixed incident energy close to the center of the K-edge rise. Two examples can be seen in Figure 4-1B. The fluorescence intensity rises as the K-edge shifts towards lower energy. This is probably due to changes in the ligand environment and oxidation and spin state changes at the iron sites. The time-scans further highlight the effect of temperature on these processes. Especially when measuring above the protein glass-transition temperature (~200K), more rapid radiation damage is often observed. The maximal “safe” time period per sample position in these examples was shorter than 10 s at 20 K but shorter than 1 s at 250 K.

II – Sample area required for TRES

The amount of sample (i.e. sample area) that is required for TRES data collection depends on the beam size and “safe” acquisition period. Sample inhomogeneity and limitations in sample position accuracy also have to be considered. The beam size used in this work ranged from 0.15 – 0.35 mm^2 with larger areas resulting in more reliable intensity calibration. A complete Kβ main-line or satellite line emission spectrum can usually be collected on a sample (1 mm thick) with ≥ 25 mm^2 area (i.e. at least 25µl sample at ~1 mM metal concentration)

III - Fluorescence data collection at each emission energy on a different sample position

The number of energy points (i.e. sample positions), the total sampling period per sample position and the time resolution (e.g. 100 × 0.1 s) need to be set to the “safe” X-ray exposure time period. Longer sampling periods may provide insight into the nature of the photodamage. Figure 4-2A shows a TRES experiment to measure the $K\beta_{2,5}$ emission of a protein sample. The S/N ratio is reasonable for Kβ data acquisition but considerable inhomogeneity per sample spot in sample thickness are observed. This makes the next step of calibration necessary.

IV – Normalization of fluorescence data

A second round of sampling is carried out at the previous sample positions but using a constant reference energy, for example at the $K\beta_{1,3}$ peak maximum (see mid Figure 4-2A). The reference scan provides insight into the sample “thickness” (i.e. relative metal content) at each spot (if the scattering background is constant or negligible). Figure 4-2B shows that even for considerable sample inhomogeneity, $K\beta_{2,5}$ spectra with reasonable S/N are obtained. The statistics of the reference data should be significantly better than the spectral data, so that no noise is added in the normalization step. Sampling times in the range of 25% - 100% of the spectral sampling time were used here because the $K\beta_{1,3}$ emission was 50-100 times larger than the measured $K\beta_{2,5}$ features. Normalization at the spectral data is achieved by dividing the spectral data by the reference data (see bottom Figure 4-2A). For the shown example a $K\beta_{2,5}$ spectrum with good S/N ratio from fluorescence count rates as low as 20 s^{-1} was measured with time-resolved energy-sampling in less than half of the time which would be required to obtain

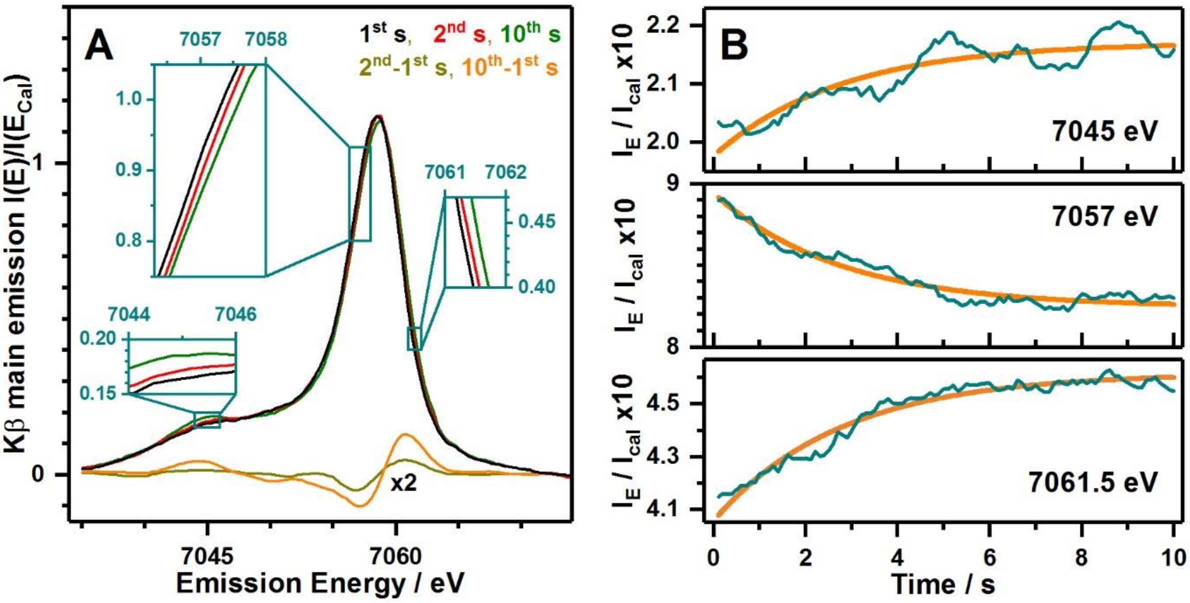


*Figure 4-3: Selecting time ranges unaffected from X-ray exposure after completing measurements and opportunity for kinetic analysis. (**A**) Kβ main-line emission spectra from **oxy** myoglobin samples (~ 10 mM in solution) at 250 K. The fluorescence at each emission energy was sampled for 10 s ($\Delta t = 0.1$ s). Shown spectra are mean time traces in the ranges from [0; 1] (black),]1; 2] (red) and]9; 10] s (green). The spectra were background corrected and difference spectra scaled for comparison. The Kβ' feature is increasing and the $K\beta_{1,3}$ peak upshifting with time as highlighted by the insets and difference spectra. (**B**) Time traces of fluorescence emission were averaged for an energy range of ± 1 eV at 7045, 7057 and 7061.5 eV. The three traces were each fit with a single exponential function with global time constant ($\tau = (2.8 \pm 0.5)$ s). Shown are smoothed time traces (green) and single-exponential fits (orange).*

a comparable S/N using the step-scan approach. We found that the normalization procedure is able to correct for concentration variations of up to 50% in the sample for samples with metal concentrations exceeding 5 mM. Samples with a smaller concentration are more easily affected by variations in concentration along the sample and repeated energy-sampling scans, i.e. on a different sample, are required to improve statistics.

The here presented TRES scan approach permits to compare spectral time traces and therefor monitor changes due to e.g. radiation damage, which allows to extract the radiation free spectrum. An example for is shown in Figure 4-3. Here the TRES scan with 10 s sampling time is split into spectra at various times (see Figure 4-3A) and difference spectra yield time dependent changes. The shift of the $K\beta_{1,3}$ and the rise of the Kβ' feature indicate a spin change of the iron site. Time traces from three emission energies, which correspond to the largest changes in difference spectra (see Figure 4-3B), where fitted with single-exponential functions and a global time constant $\tau = (2.8 \pm 0.5)$ s. Therefore, ~ 10 % of sample has undergone a change in spin state after 0.3 s and the final spectrum can be constructed accordingly.

In the next section, the efficiencies of the step-scan and TRES approach are compared to provide a quantitative rational decide which method should be applied to increase efficiency during limited synchrotron beamtime.

Evaluating the efficiencies of step-scan versus energy-sampling approaches

The efficiency (μ) of measuring a single XES spectrum is defined as the ratio of data sampling time (*t*) and the total time required for acquisition of the spectrum (*T*). For the energy step-scan approach, the total measuring time is the sum of the sampling time and the dead time (t_d) i.e. for motor movement. For a TRES scan, the spectral sampling time was considered to be equal to the time of normalization data acquisition. Be aware that the normalization data acquisition usually was almost two-fold shorter than the spectral data acquisition.

$$\mu_{CES} = \frac{t}{t + t_d} \tag{4-1}$$

$$\mu_{TRES} = \frac{t}{2\,t + t_d} \tag{4-2}$$

The efficiencies of CES and TRES measurements are exemplified by data collected at beamline ID26 at ESRF (Grenoble, France) in Figure 4-4. It is immediately clear that a CES scan is always more efficient when considering only one measurement to collect a single XES spectrum. This is due to two reasons; (i) the efficiency for TRES cannot rise above 0.5 as normalization data acquisition adds to the total measuring period and (ii) the dead time is larger (for a single measurement) because of the sample-stage motor movements for sample position changing. However, a different picture is obtained when comparing the two methods in relation to the sampling period that is allowed at each sample position (see Figure 4-4B). Now the energy-sampling approach is more efficient for a sampling period per sample spot smaller than 100 s or less, thus the energy-sampling approach becomes highly favorable for short data acquisition on radiation sensitive samples. In order to decide which method to use, an estimate of the acquisition time to achieve spectra of equal S/N ratio. The measuring time (t_m) necessary to obtain a fluorescence emission signal with the targeted S/N ratio needs to be determined. The S/N ratio of a spectrum with total counts $C = C_s + C_{bg}$ (C_s = counts from sample signal, C_{bg} = counts from scattering background):

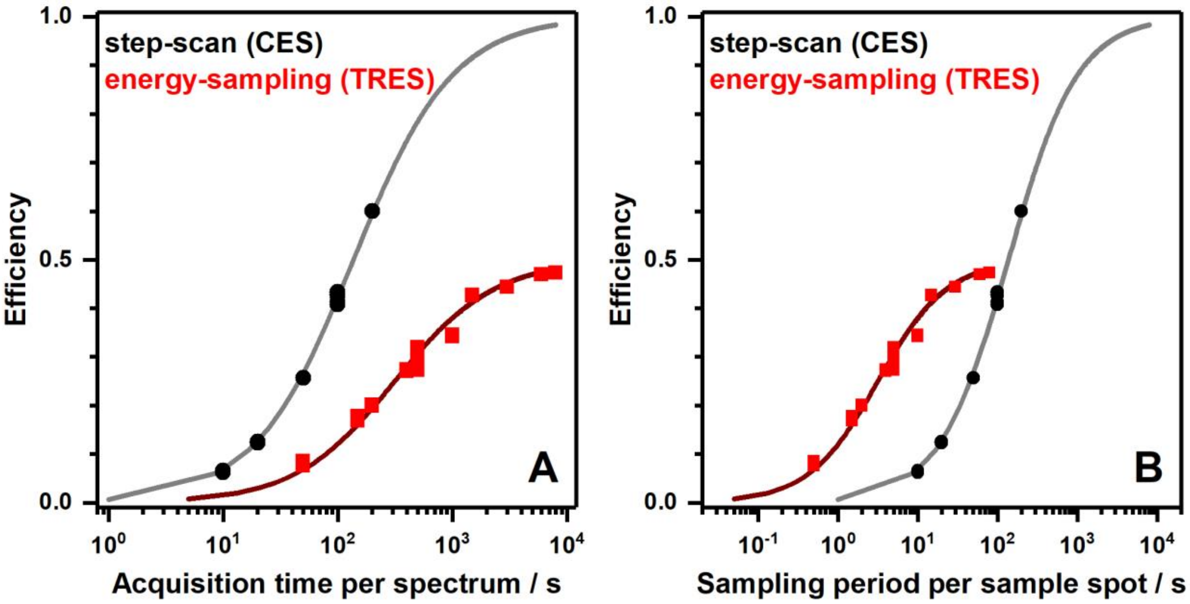


*Figure 4-4: Data acquisition efficiency in XES. (**A**) Efficiency as a function of total spectrum measurement time. (**B**) Efficiency as a function of sampling period per sample position. The efficiency is defined as the total spectral sampling period divided by the time of data acquisition (spectral + reference data). Lines show fits of dead time (t_d). Step-scan is fitted with $f(t)=t/(t+t_d)$ and energy sampling with $g(t)=t/(2t+t_d)$. Fitted dead times are 137 s (red) and 630 s (black). The efficiency here does not take the S/N ratio into account and is therefore not indicative of the spectral quality.*

$$S/N = C_s/\sqrt{C_s + C_{bg}} \tag{4-3}$$

Expressing absolute counts with sampling time (t_m) and count rates k_s/k_{bg} (i.e. $c_s = t_m\, k_s$) leads to:

$$S/N = \frac{t_m\, k_s}{\sqrt{t_m(k_s + k_{bg})}} \Longrightarrow t_m = S/N^{\,2}\ \frac{(k_s + k_{bg})}{k_s^2} \tag{4-4}$$

Typical values for K$\beta_{2,5}$ emission measurements at a 3rd generation source (ESRF) and the undulator beamline ID26 (~10^{12} photons s^{-1} at Fe K-edge) for a sample with 10 mM iron were $k_s = 30\ s^{-1}$ and $k_{bg} = 5\ s^{-1}$ which, for a S/N ratio of 10, resulted in a measuring period per emission energy of $t_m \approx 4$ s. This value when multiplied with the number of points per spectrum (n) yields the total measuring time. If the “safe” sampling period per sample position (t_S) is shorter than the time it takes to complete a measurement at one position, two or more sample spots are required to collect the data at a single emission energy. The number of repetitions (x) for the

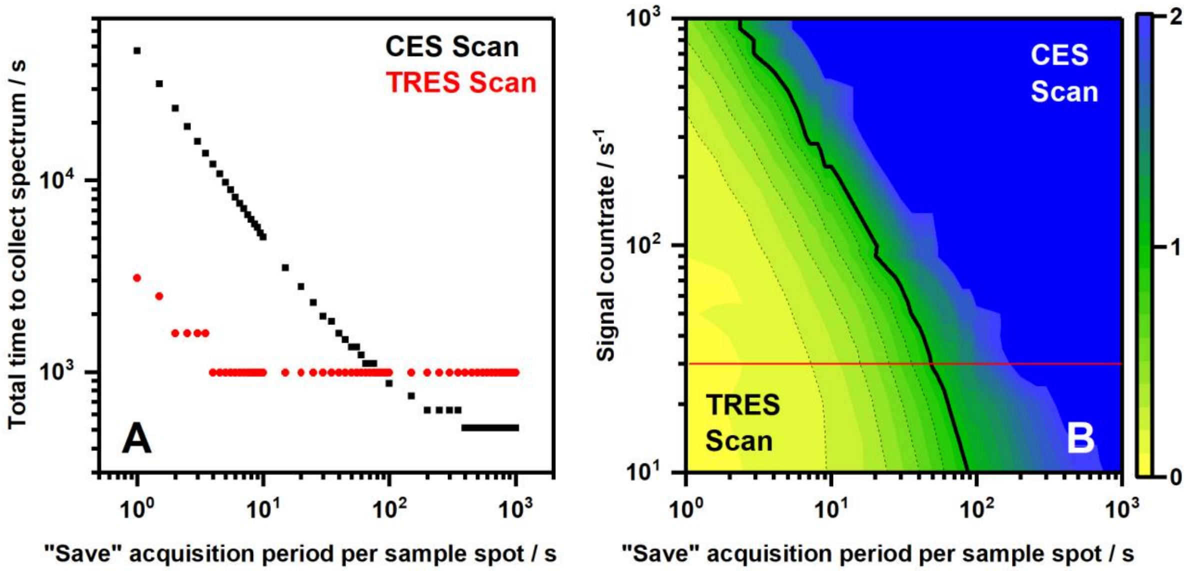


*Figure 4-5: Comparison of conventional energy step scan (CES) and time-resolved energy-sampling scan (TRES) approaches. (**A**) Total time for acquisition of a spectrum with a S/N ratio of 10 for a signal count rate of 30 s^{-1} and a background count rate of 5 s^{-1}. Values were calculated using equations (4-7) and (4-8) (100 points per spectrum and dead times of 120 s [CES scan] and 600 s [TRES scan]). The stepwise behavior is due to the ceiling and floor function for the number of repetitions (see equations (4-7) and (4-8)). (**B**) Ratio of total data acquisition time of CES scan versus TRES scan measurements. The dead times, background and S/N ratio were as in (A). Unity ratio is indicated as a black line and the red line indicates a signal count rate as used in (A).*

energy-sampling approach depends on the ratio of t_m to t_S. Step-scan repetitions (y) on the other hand also depend on the number of points per spectrum (z).

$$x = \left\lceil \frac{t_m}{t_S} \right\rceil \tag{4-5}$$

$$y = \left\lceil z\, \frac{t_m}{t_S} \right\rceil \tag{4-6}$$

Assuming that the sample surface can adopt 200 non-overlapping beam positions, the sample positioning time (t_p) is about 300 s, which, together with the individual dead times (t_d), leads to the following total measuring time (T):

$$T_{CES} = z\, t_m + y\, t_d^{CES} + \left\lfloor \frac{y-1}{200} \right\rfloor t_p \tag{4-7}$$

$$T_{TRES} = 2\, z\, t_m + x\, t_d^{TRES} + \left\lfloor \frac{y-1}{2} \right\rfloor t_p \tag{4-8}$$

Again, the TRES calibration measurement is of equal length as the spectrum measurement in equation (4-8). A comparison of T values for the two approaches is shown in Figure 4-5A. CES scans are more efficient for samples that are insensitive to X-ray exposure ("safe" X-ray exposure > 100 s). The TRES scanning time becomes significantly shorter (up to two orders of magnitude) for shorter "safe" X-ray exposure periods (< 100 s).

Conclusion

Two approaches for collection of weak X-ray emission spectra were quantitatively compared. Software was developed that allows to estimate by use of key experimental parameters, which method should be applied for given X-ray beamline conditions. In order to decide whether to use the conventional energy step-scan (CES) or the time-resolved energy-sampling (TRES) approach, only a small set of setup-specific parameters needs to be considered. These include the scattering background magnitude, motor movement times and resulting dead times for step-scan and energy-sampling experiments. The sample-specific parameters, (the signal count rate and "safe" X-ray exposure period) have to be determined (i.e. from XANES data). Graphical representations of the "setup" and "sample" specific parameters were designed for easy access. The calculation procedures were implemented in a program that facilitates to decide, which data collection method is superior. The software is available to interested users. The methods were extensively tested at the high-flux ID26 beamline. In this case, $K\beta_{2,5}$ emission spectra of samples with very low metal concentrations should be measured using the conventional energy step-scans if the "safe" X-ray exposure period exceeds 100 s. Such long "safe" periods, however, to our experience are not accessible even for "radiation hard" dry powder material samples at cryogenic temperatures. If the "safe" X-ray exposure period decreases by one order of magnitude or more, even higher concentrated samples (~ 10 – 50 mM metal) should be measured using the time-resolved energy-sampling scan. Monitoring weak X-ray spectral features with time-resolved energy-sampling is a superior approach with high efficiency for samples with low concentration and/or high sensitivity to X-ray exposure, to collect high-quality data.

5 Effective intermediate-spin iron in O_2-transporting heme proteins

This chapter has been published in part in the following publications.

N. Schuth, S. Mebs, D. Huwald, P. Wrzolek, M. Schwalbe, A. Hemschemeier, M. Haumann, *Effective intermediate-spin iron in O2-transporting heme proteins.* Proc Natl Acad Sci U S A, 2017. **114**(32) https://doi.org/10.1073/pnas.1706527114 [105]

N. Schuth, M. Haumann. *X-ray spectroscopy reveals iron spin state in oxygenated hemoproteins.* ESRF Spotlight on Science [102] and ESRF Highlights 2017 [103]

5.1 Introduction

Proteins binding an iron-porphyrin (heme) cofactor are nature´s versatile dioxygen (O_2) managers, facilitating O_2 sensing, transport, activation chemistry, and catalysis [39]. Scrutiny over decades has been devoted to the classical O_2-transporting myoglobin (MB) and hemoglobin (HB) proteins of vertebrates [38, 39, 56, 182, 183]. Cooperative vs. non-cooperative O_2 binding to heme iron in tetrameric HB or monomeric MB is textbook knowledge (Figure 5-1). Renewed interest in the topic stems from ongoing discovery of globins in plants and microorganisms [184-186]. A controversy remains on the nature of the iron-oxygen bonding in HB/MB, dating back to pioneering work of Pauling, Monod, and Perutz starting in the 1930´s [56, 187].

Three physiologically relevant heme species are distinguished in HB/MB, for which crystal structures at ~ 1 Å resolution are available [2, 7]: the ligand-free (**deoxy**), carbon-monoxide inhibited (**carboxy**), and O_2-bound (**oxy**) forms, besides of the ferric (met) species from in vitro autoxidation (Figure 5-1). Magnetic susceptibility experiments [45, 47] yielded paramagnetic behavior in **deoxy** and **met**, which correspond to high-spin (HS) iron in **deoxy** (d^6 Fe(II); spin (S) multiplicity, M = 2S + 1 = 5) and **met** (d^5 Fe(III); M = 6). **Carboxy** showed diamagnetic behavior due to low-spin (LS) Fe(II) configuration (M = 1) [188]. For **oxy**, a diamagnetic (M = 1) cofactor was observed as well [45, 56, 188]. However, the O_2 triplet ground-state provides several options for spin-flipping or charge transfer (Figure 5-1): Pauling suggested a neutral singlet O_2 bound to a formal LS Fe(II) [45, 189], McClure suggested a triplet O_2 with antiferromagnetic coupling (afc) to the intermediate-spin (IS) Fe(II) [58], an IS Fe(II) with d-level population inversion in an ozone-like model was also suggested by Goddard [59], and Weiss suggested an Fe(III) afc-coupled to a superoxo ($O_2^{\bullet-}$) ligand [57]. The possible resonance

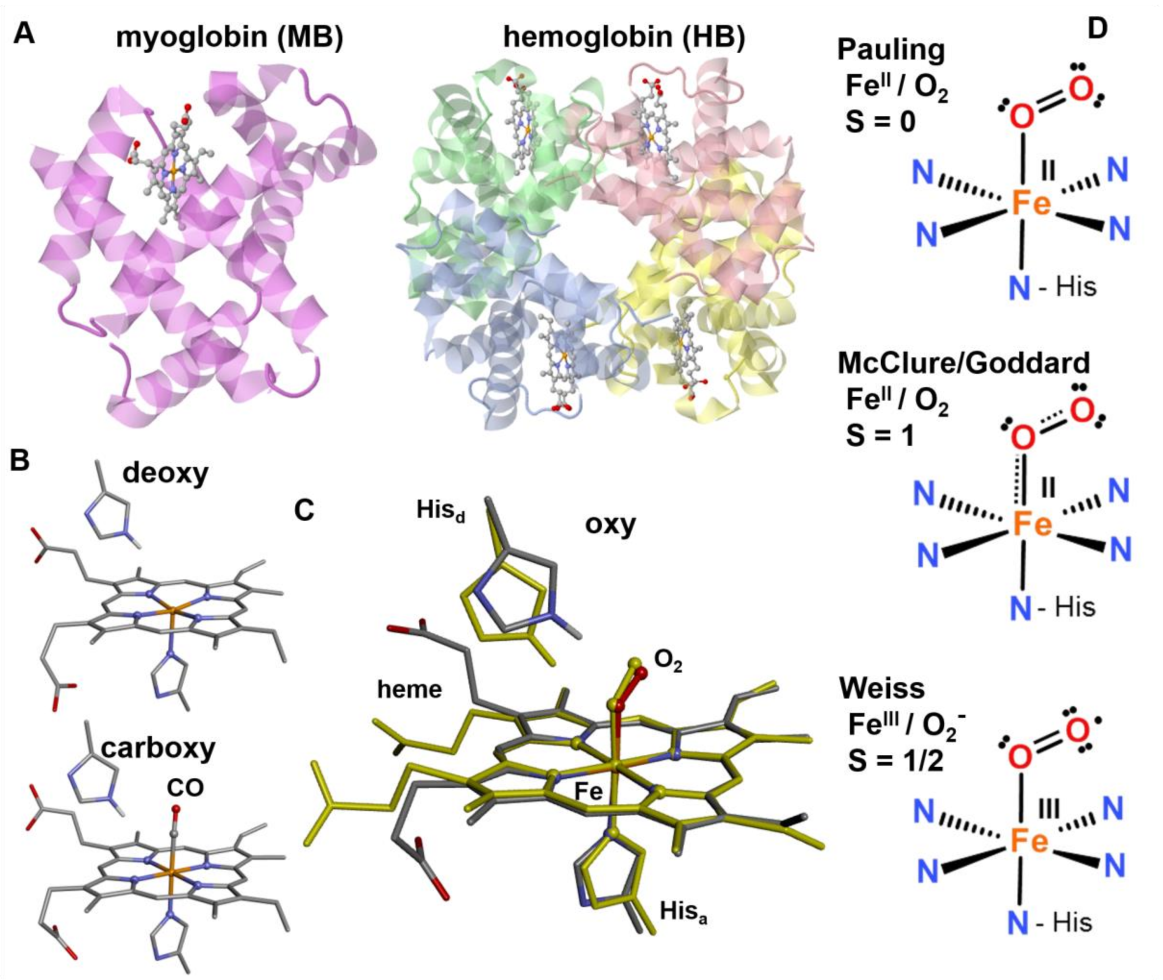


*Figure 5-1: Heme protein structures and O_2 binding. (**A**) Crystal structures of human MB/HB (left/right, PDB entries 3RGK/2DN2). (**B**) Structures of **deoxy**/**carboxy** hemes in MB (PDB entries 3QM6/3QM7, 0.91/0.96 Å resolution). (**C**) Overlay of **oxy** hemes in MB/HB (balls-and-sticks/yellow-sticks, PDB entries 1A6M/2DN1, 1.00/1.25 Å resolution). (**D**) Schematic Fe – O_2 bonding in **oxy** in the Pauling (top), McClure/Goddard (middle), and Weiss (bottom) models ($His_{p/d}$, proximal/distal histidine; atom color code: grey, C; blue, N; red, O; orange, Fe).*

character of these Lewis structures was emphasized already in early theoretical work [190]. DFT, providing access to experimental observables [54, 191], and CASSCF quantum chemical calculations facilitated determination of many structural and spectroscopic features of the O_2 bonding [39, 190]. A McClure-like **oxy** [71], as well as Pauling/Weiss descriptions have emerged, to some extent depending on the theoretical framework used to interpret multi-configurational wavefunctions [55, 71, 76]. Modification of porphyrin groups and heme/ligand geometries by the protein matrix and hydrogen-bonding may bias the Fe – O_2 interaction towards the diverse function of heme cofactors [39, 192-194].

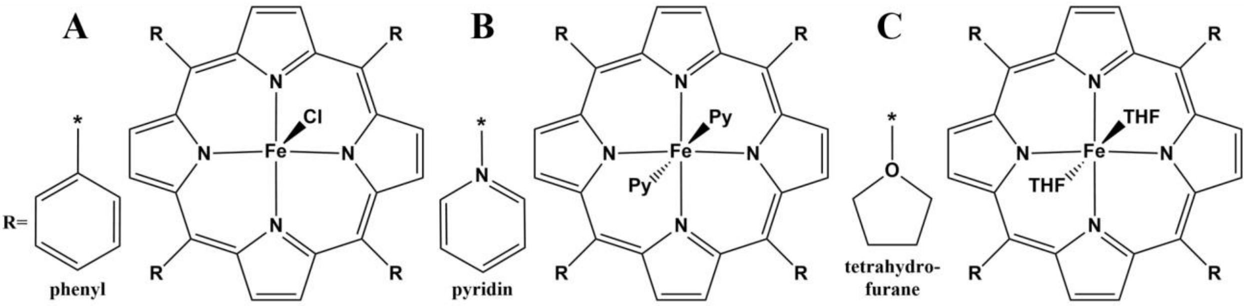


*Figure 5-2: Chemical structures of (**A**) Fe(III)(TPP)Cl (high-spin), (**B**) Fe(II)(TPP)(Py)$_2$ (low-spin) and (**C**) Fe(II)(TPP)(THF)$_2$ (high-spin); TPP = tetraphenylporphyrin, Py = pyridine, THF = tetrahydrofurane.*

Core level spectroscopy is powerful for electronic structure analysis of transition metal complexes [23, 195]. X-ray absorption spectroscopy (XAS) probes unoccupied valence levels by resonant 1s electron excitation in the pre-K-edge region (core-to-valence transitions, ctv), ligation symmetries and oxidation/spin states in the XANES, and bond lengths in the EXAFS, whereas X-ray emission spectroscopy (XES) of the Kβ lines probes occupied levels via electronic decay from metal-3p (main-lines) or -valence (satellite lines; valence-to-core transitions, vtc) levels to the core hole [196]. The ctv/vtc spectra can be calculated by DFT methods for benchmarking [197, 198]. 3p – 3d spin-coupling splits the Kβ main-line emission into the Kβ′ and $K\beta_{1,3}$ features and the Kβ′ intensity is a measure of metal spin state and bonding covalency [25, 36, 199]. XAS has been used to address structural metrics, solution vs. crystalline materials, and redox states in HB/MB [200-204]. In this respect, ctv data were interpreted in favor of the Weiss model of O_2 binding [1, 201]. Systematic Kβ XES on hemoproteins has not been done yet and is challenging on protein solutions due to radiation-induced modifications [205]. Additionally, temperature effects may contribute, as for **met** HB/MB partial spin crossover (i.e. high-spin population) was found [206]. We used advanced X-ray spectroscopy to study heme species in MB and HB proteins in solution. According to our results, structural and electronic features were similar for both proteins and constant at 20 – 260 K. Static and kinetic XAS/XES and (TD-)DFT/CASSCF analyses revealed high-spin Fe(II) in **deoxy**, low-spin Fe(II) in **carboxy**, and intermediate-spin iron in a delocalized Fe – O_2 bonding situation in **oxy**, which converges towards a unified view of the metal – ligand interaction.

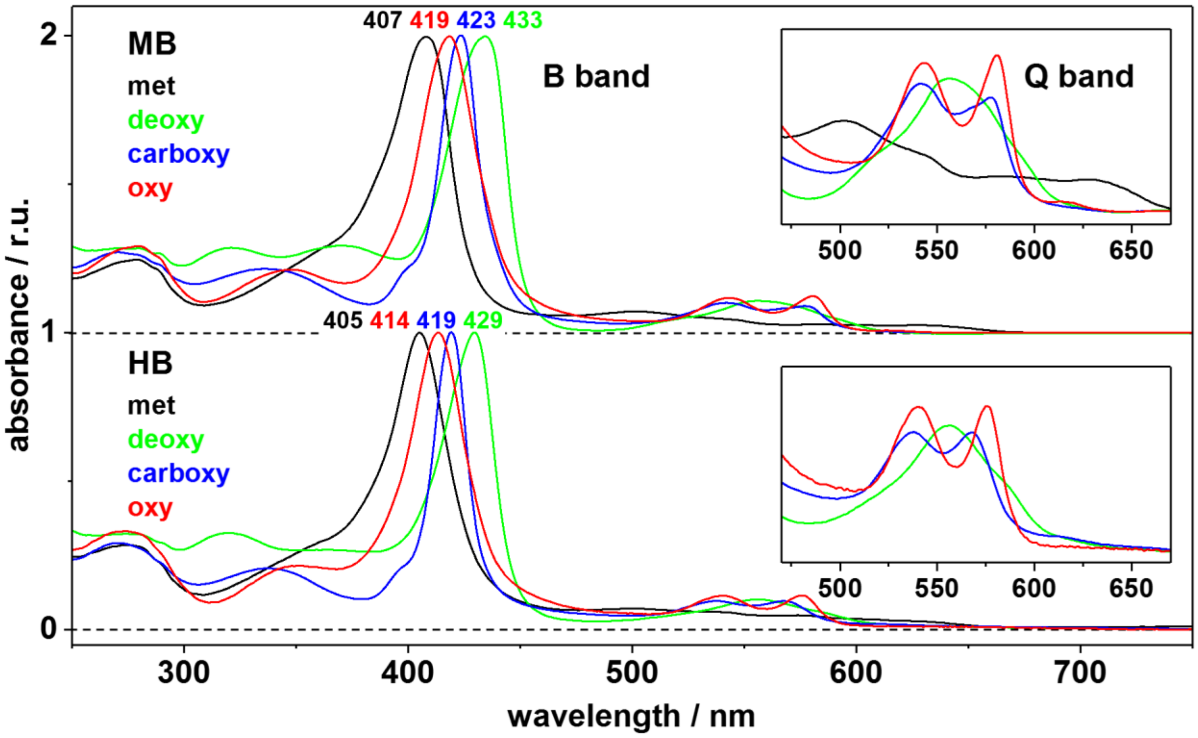


*Figure 5-3: Optical absorption spectra of MB and HB. Spectra were normalized at the Soret band amplitudes at the indicated wavelengths. Insets: magnification of spectra in the Q band region. Spectra correspond to aliquots (~12 µM) of HB/MB protein preparations for XAS/XES in buffer solutions under N_2 (**met/deoxy**) or CO/O_2 (**carboxy/oxy**) atmospheres.*

5.2 Preparation of heme protein and model complex samples

Heme protein samples were prepared and characterized with the group of A. Hemschemeier (D. Huwald) during several research stays at the Ruhr-University Bochum. Bovine hemoglobin (HB) and myoglobin (MB) proteins (met) were purchased from Sigma-Aldrich, dissolved in buffer A (20 mM MES, 10 mM NaCl, pH 7) at 12 mM heme, and reduced with sodium-dithionite (60 mM) to yield **deoxy**. MB proteins stem from bovine heart muscle tissue. MB/HB **deoxy** solutions (350 µL) were incubated under pure carbon monoxide or oxygen gas for 10 min to derive **carboxy** or **oxy** [207]. Sample aliquots (50 µl) were loaded into Kapton-covered acrylic-glass holders and frozen in liquid nitrogen. Sample handling was done in a glove box under 100 % dry N_2 atmosphere. Optical absorption spectra of HB/MB (12 µM in buffer A under the respective atmosphere in a sealed 1 cm cuvette) were collected on a Shimadzu UV-2450 spectrometer and are shown in Figure 5-3. The absorption spectra show two principal spectral regions, which are a fingerprint for porphyrin complexes. The high intensity absorption around ~400 nm is the B- or Soret band (named after its discoverer), which gives heme proteins the red color. The low intensity absorption at ~500 – 600 nm is the Q band, which correspond to transitions between porphyrin-centered states. These states have to be described by multiconfigurational methods with perturbation theory and large active spaces, e.g. CASPT2 or NEVPT2, for correct transition energies, which indicates large static and dynamic electron correlation on the porphyrin ring [164]. Energy splitting and relative intensity between Soret and Q-band depends on the metal atom and axial ligands and changes are in good agreement with previous publications [1, 123, 208].

Porphyrin reference compounds were chosen (i) to cover the entire spin range possible with Fe(II) and Fe(III) compounds, from low-spin Fe(II) (S = 0) to high-spin Fe(III) (S = 5/2) and (ii) to maintain a comparable ligand environment with the porphyrin ring and five or six coordinated iron in an octahedral or square pyramidal geometry. Porphyrin compounds were synthesized in the workgroup of M. Schwalbe by P. Wrzolek (HU-Berlin). Porphyrin compounds ($Fe(TPP)(Py)_2$, **1**; $Fe(TPP)(THF)_2$, **2**; Fe(TPP)Cl, **3**; see Figure 5-2) were synthesized as described [1, 209] and grinded with boron nitride (w/w 1:10) to yield homogenous powder samples for XAS/XES.

Conclusion

Heme protein samples were prepared during several visits to the Ruhr-Universität Bochum with kind support by our collaborators (D. Huwald, group of A. Hemschemeier).Samples were characterized with UV/VIS absorption spectroscopy and shifts in the Soret and Q bands indicated near-quantitative formation of **deoxy**, **oxy** and **carboxy** states in hemoglobin and myoglobin samples. Porphyrin model complexes were synthesized by P. Wrzolek in the group of M. Schwalbe (Humboldt-Universität zu Berlin). These reference compounds cover a wide spin state range (S = 0 – 5/2) while maintaining a comparable coordination geometry. Quantitative and comparative analysis of the heme proteins and the porphyrin model complexes is presented in the next chapter.

5.3 Heme species quantification

MB and HB proteins in solution were prepared with ligand-free (**deoxy**), carbon monoxide inhibited (**carboxy**), and oxygenated (**oxy**) hemes. Similar optical absorption spectra of MB and HB (see Figure 5-3) with typical Soret band shifts (MB/HB: **met**, 407/405 nm; **deoxy**, 433/429 nm; **carboxy**, 423/419 nm; **oxy**, 419/414 nm) and differences in the charge-transfer region (500-600 nm) agreed with previous reports [201, 210]. In this section, analysis from of XANES and EXAFS spectra are presented to evaluate the sample quality by comparison to previous results e.g. from XRD. Additionally, the absence of radiation damage (high X-ray dose) and temperature induced changes in electronic configuration (spin crossover) through experiments below and above the Glass transition (~ 200 K).

Species sensitive fingerprint from XANES

XANES spectra at the iron K-edge of three sets of MB/HB samples and three sets of porphyrin model compounds were collected at 20 K and 260 K at two monochromator resolutions and using broad-band Kα or narrow-band Kβ emission detection (see Figure 5-4 and Figure 5-5). The XANES spectral shape represent a sensitive fingerprint, which can be used to verify the absence of X-ray exposure or temperature induced changes. K-edge shapes of HB and MB states did not change with increasing temperatures, for sufficiently short X-ray exposure (~ 1 s, see section 5.5). The K-edge energies at half-rise for a given protein and state at 20 K and 260 K were similar within ± 0.1 eV (**deoxy**, 7121.0 eV; **carboxy**, 7122.3 eV; **oxy**, 7123.9 eV; see Table 5-1 and insets in Figure 5-4), in agreement with previous reports [1, 200, 201]. This indicates (i) the absence of temperature induced changes in electronic configuration due to spin crossover and (ii) the absence of radiation damage (X-ray dose) by reduced exposure time and beam attenuation. Experiments to determine the sufficiently short save X-ray exposure time are presented in section 5.5. The XANES of HB/MB in the **oxy** state shows no significant difference in shape, whereas the K-edges for hemoglobin samples in the **deoxy** and **carboxy** state share a sharper edge profile than respective myoglobin samples. These changes reflect minor differences in heme geometry. Thus, XAS indicates similar ground-states for both proteins in all three states. The most notable difference between all three states in hemoglobin and myoglobin is the shift in edge position. The XANES for HB/MB upshifts from **deoxy** to **carboxy** by ~ 1.3 eV and from **deoxy** to **oxy** by ~ 2.9 eV. The **deoxy → oxy** upshift has previously been interpreted as a change in oxidation state in favor of the Weiss model [1]. However, the **deoxy → carboxy** upshift does not stem from a change in oxidation state but

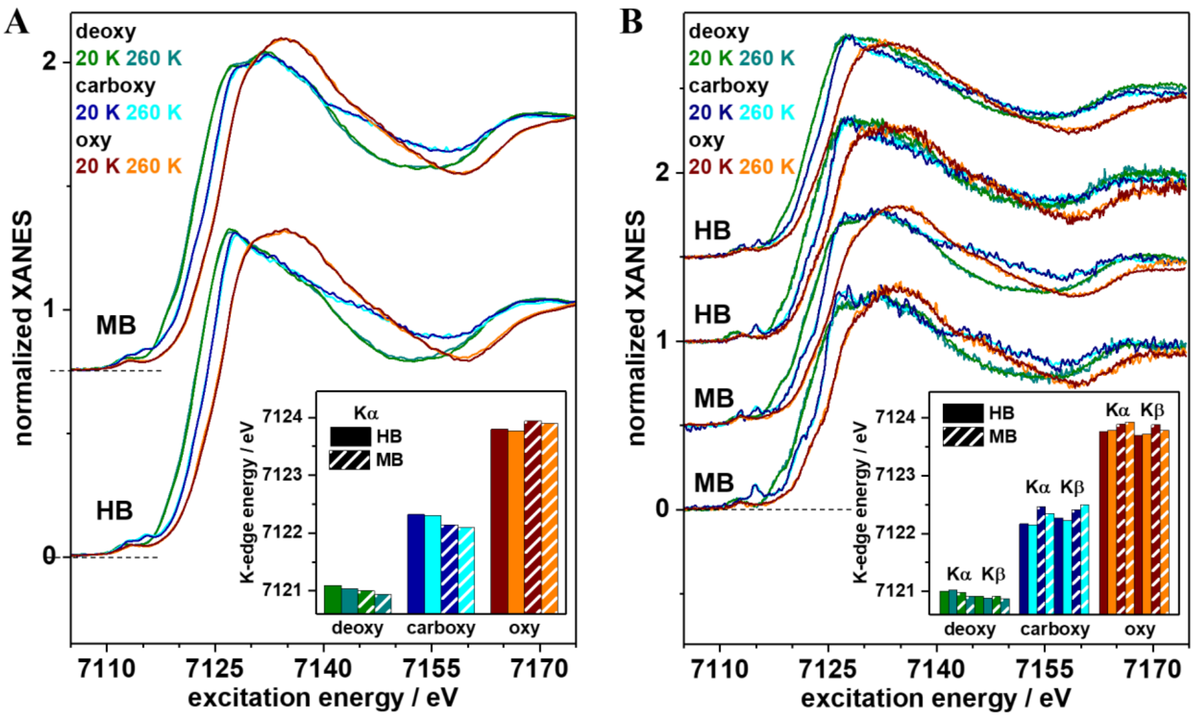


*Figure 5-4: Normalized lower- and higher-resolution XANES spectra of MB and HB at 20 K and 260 K. Spectra of the three heme species at 20/260 K (vertically shifted) were collected using a Si[111] (**A**, at beamline KMC3/Berlin) and Si[311] (**B**, at beamline ID26/Grenoble) double-crystal monochromator and broad-band (~150 eV, "total-fluorescence") Kα detection or narrow-band (~1 eV) Kβ emission detection (for energies at the maximum of the respective* $K\beta_{1,3}$ *emission line, see Figure 5-8). The insets show respective K-edge energies at 50 % level (see Table 5-1). Note within noise limitations similar K-edge energies for both detection methods, which are also similar to the data in the inset on the left. The lower signal-to-noise ratio for spectra collected at the ID26 (right) is due to the shorter save time for data acquisition, the smaller beam size and the smaller detector angle area. Spectra on the left correspond to EXAFS experiments (see Figure 5-6).*

corresponds to a change in coordination symmetry, ligand type and iron spin state. Carbon and nitrogen are softer ligands compared to oxygen, which is also visible in the broader XANES shape for **oxy**-HB/MB, **met**-HB and porphyrin **2** compared to **deoxy**, **carboxy** and porphyrin **1**. The K-edge shifts, due to a change in oxidation state, spin state or ligand type, increases for harder ligands (see Figure 2-3). The K-edge for **met** hemoglobin (see Figure 5-5), which is a six-coordinated Fe(III) low-spin complex with a water molecule as sixth ligand, even has an additional upshift compared to **oxy** of ~ 0.8 eV. For the model complexes, K-edge energies of porphyrin **1** and **2** are equal to **carboxy** because of XANES shape variation by axial ligands.

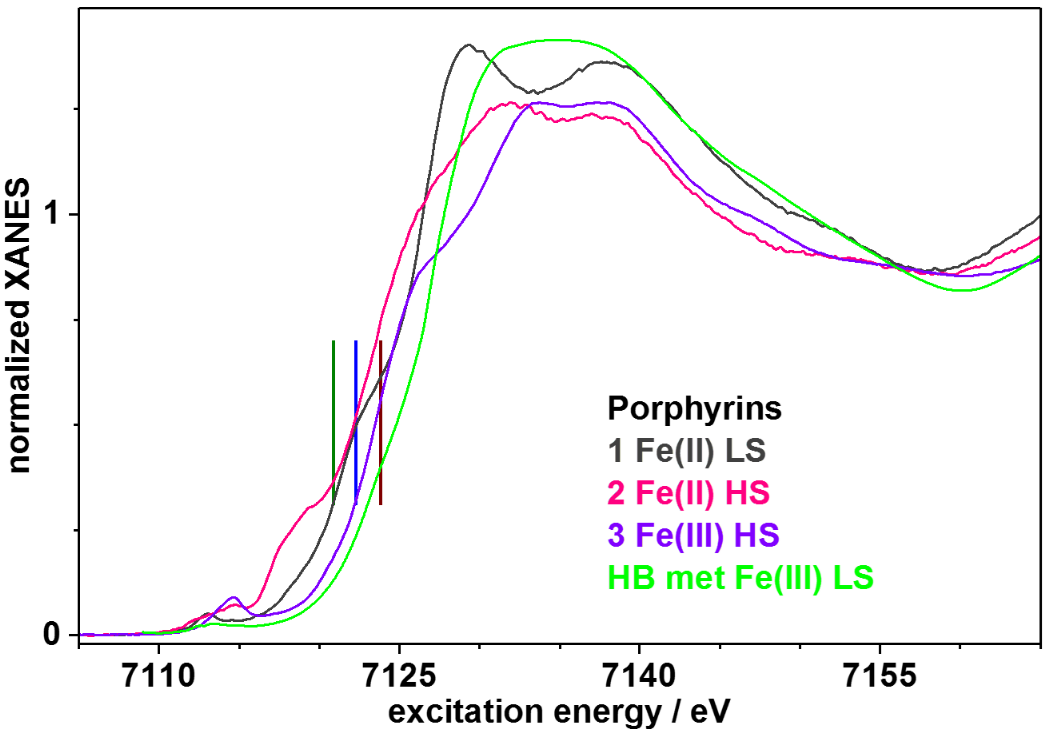


Figure 5-5: Normalized high-resolution XANES spectra (Si[311] double-crystal monochromator, Kα-fluorescence detection) of porphyrin compounds ***1****,* ***2****, and* ***3*** *and of* ***met*** *HB (spectrum reproduced from ref. [1]). K-edge energies from HB/MB states are indicated as color coded vertical lines (olive:* ***deoxy****, blue:* ***carboxy****, wine red:* ***oxy****; see Table 5-1).*

Table 5-1: Fe K-edge energies of heme species and porphyrin complexes.

	K-edge energy / eV			
	MB	HB	Porphyrins	
deoxy	7120.9(1)	7121.0(1)	**1**	7122.1(1)
carboxy	7122.3(1) [7122.4(4)]	7122.2(1) [7122.3(3)]	**2**	7122.3(1)
oxy	7123.9(1) [7123.9(4)]	7123.8(1) [7123.9(6)]	**3**	7123.5(1)
met	[7124.7]	-		

K-edge energies (at 50 % level) refer to mean lower-/higher-resolution XANES data in Figure 5-4 and Figure 5-5 or (in brackets) were estimated from spectra reported in literature [1, 8-13], which were shifted so that respective energies for deoxy aligned with our data (the range of energy variations is given in parenthesis).

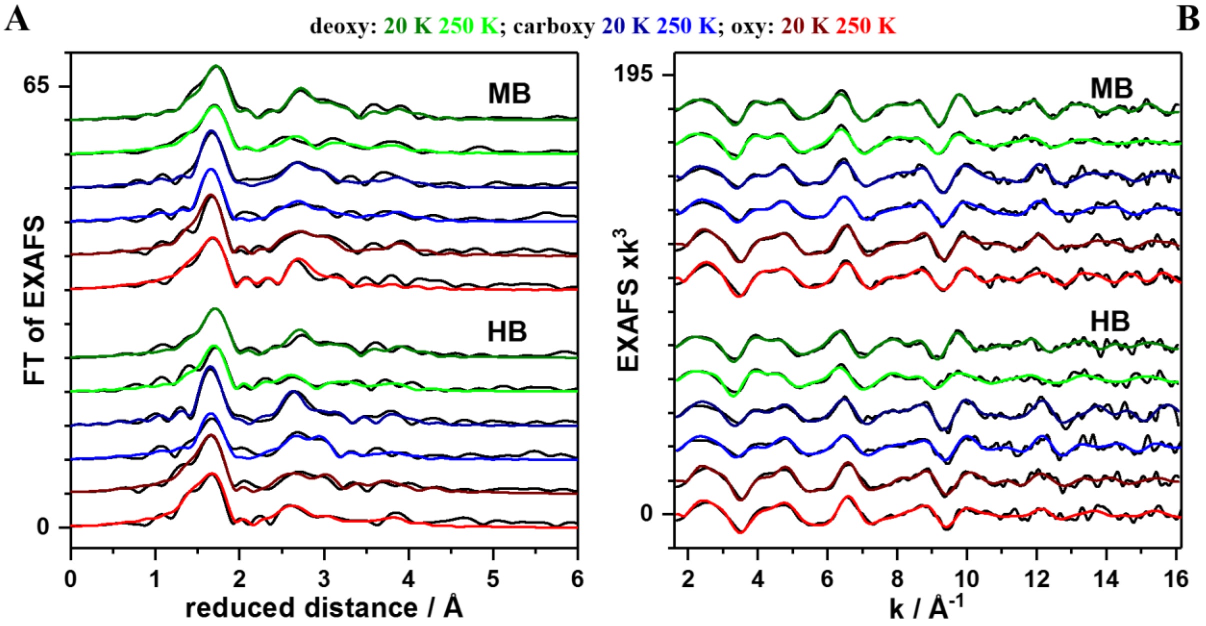


*Figure 5-6: EXAFS spectra of MB and HB at 20 K and 260 K. (**A**) Fourier-transforms calculated for k-values of 1.6 – 16.2 Å^{-1} with cos windows extending over 10 % at each k-range end. (**B**) Corresponding EXAFS oscillations from experimental data (black lines) and simulations (colored lines) with parameters shown in Figure 5-7 and Table 5-2.*

The models by Pauling, McClure/Goddard and Weiss all include changes in site symmetry and iron spin state, whereas only Weiss includes a change in oxidation state. All three models therefore could explain the upshift. Consequently, the shift in K-edge for **deoxy**/**oxy** is not indicative of a change in iron oxidation state (to favor the Weiss model). Experimental features that are more closely related to the metal spin state are presented in later sections.

Metal-ligand bond lengths from EXAFS

EXAFS spectra of HB/MB samples were collected at 20 K and 260 K up to a k-range of 16 Å^{-1} and are shown in Figure 5-6. Precise iron-ligand bond lengths could be determined, which were compared to structures from X-ray crystallography (see Table 5-2). Atomic distances were invariant with temperature; in agreement with the XANES spectra (see Figure 5-7). The amplitudes for the same heme species in the Fourier-transform decreased with increasing temperature, which stems from an increase in Debye-Waller parameter due to an increase in thermal motion. Coordination numbers close to unity were found for the Fe – N_{His}, Fe – CO and Fe – O_2 bonds and coordination numbers close to four for Fe – N_{Porph} in MB/HB. This further supports the results from optical absorption with near quantitative (> 90%) presence of **deoxy**, **carboxy** and **oxy** in the investigated samples and absence of temperature induced

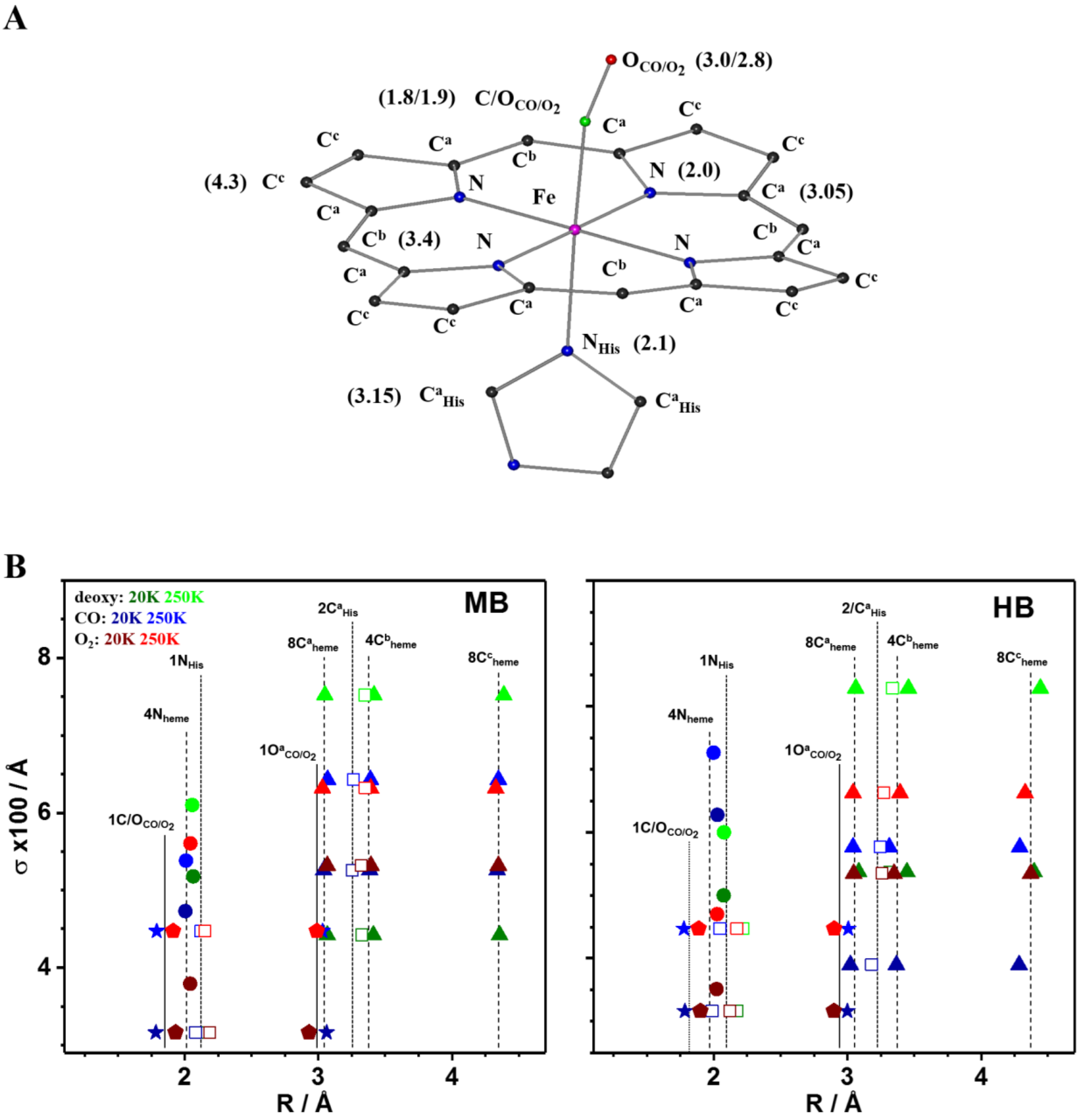


*Figure 5-7: Interatomic distances from EXAFS analysis. (**A**) Schematic heme structure with marked atoms in the first, second, and third coordination spheres of the iron center in deoxy, carboxy, and oxy. (**B**) EXAFS fit results corresponding to spectra at 20/260 K in Figure 5-4A and Figure 5-6 (R, interatomic distance; σ, Debye-Waller parameter; numbers in the annotations of the vertical lines represent coordination numbers, N, as used in the fits). Vertical lines were drawn at respective mean distances to emphasize similarities for MB/HB at 20/260 K. Interatomic distances from EXAFS and crystal structures are compared in Table 5-2.*

geometrical variations. EXAFS analysis showed similar distances between the iron atom and its first ligands for hemo- and myoglobin in all three states. The Fe – $O_{central}$ (1.90 Å), Fe – $C_{carboxy}$ (1.79 Å), Fe – $N_{porphyrin}$ (**deoxy**, 2.07 Å; **carboxy**, 2.01 Å; **oxy**, 2.04 Å), and Fe – N_{His} (**deoxy**, 2.18 Å; **carboxy**, 2.07 Å; **oxy**, 2.16 Å) distances for MB/HB in all three states and at 20/260 K deviated by < 0.02 Å. These distances in the first and second iron coordination spheres mostly agreed within < 0.05 Å with crystal structures. Notably, the distance from axial ligands in **oxy** showed a significantly larger spread than any other iron first sphere distance in the available XRD structures. Most likely, these variations are due to X-ray induced ligation changes in XRD (radiation damage), which are discussed in a section 5.5. These results verified near-quantitative (> 90 %) presence of hemo- and myoglobin in the **deoxy**, **carboxy**, and **oxy**

Table 5-2: Interatomic distances in heme species from XAS and XRD.

	interatomic distance / Å									
	Fe-C/O(=O)		Fe(-C/O)=O		Fe-$N_{porphyrin}$		Fe-N_{His}		Fe-$C_{porphyrin/His}$	
	XAS	XRD	XAS	XRD	XAS	XRD	XAS	XRD	XAS	XRD
MB										
deoxy	-	-	-	-	2.06(2)	2.04(2)	2.18(2)	2.20(2)	3.20	3.19
carboxy	1.79(1)	1.82(2)	3.03(2)	2.94(4)	2.00(1)	2.01(3)	2.09(2)	2.07(1)	3.18	3.15
oxy	1.91(2)	1.83(2)	2.97(3)	2.64(10)	2.04(1)	1.99(3)	2.16(2)	2.12(10)	3.19	3.15
HB										
deoxy	-	-	-	-	2.07(2)	2.07(2)	2.17(2)	2.21(2)	3.21	3.23
carboxy	1.78(1)	1.72(3)	3.01(2)	2.87(5)	2.01(1)	2.01(1)	2.04(2)	2.11(3)	3.15	3.17
oxy	1.89(2)	1.91(19)	2.91(3)	2.98(16)	2.03(1)	2.03(3)	2.15(2)	2.12(11)	3.17	3.19

[a]*XAS data refer to mean values from EXAFS at 20/260 K in Figure 5-7 (full range of deviation from mean value in parenthesis). XRD (X-ray crystallography) values correspond to two (**deoxy**), four to five (**carboxy**), or four to five (**oxy**) structures for MB/HB in the protein data base (PDB) at better than 2.5 Å resolution. The PDB entries are as follows and XRD data collection temperatures given in brackets if they are not reported (nr) or unequal to 100 K. MB/deoxy: 2QM6, 3RGK [277 K], 1MBD [nr]; MB/carboxy: 2QM7, 3E55, 1A6G; MB/oxy: 1A6M, 1LTW [295 K], 5HAV, 1MBO [nr], 2Z6S; HB/deoxy: 2DN2 [300 K]; HB/carboxy: 2DN3, 4ESA, 1J40 [35 K], 4HRR, 1IRD; HB/oxy: 2DN1, 3A0G [291 K], 1HBI [nr], 2R80, 1THB [nr]. Standard deviations are given in parenthesis.*

states, showed that the iron site structures are similar in MB and HB, and revealed the absence of temperature-induced cofactor changes.

Conclusion

Complete sets of XANES and EXAFS spectra of two different heme proteins (HB, MB) in three ligation states were collected at 20 K and 260 K and analyzed. The spectra revealed the absence of temperature induced variations related to changes in geometry or electronic configuration (e.g. temperature induced spin crossover). Accordingly, the determined parameters are likely relevant also under physiological conditions. Bond lengths from EXAFS analysis were in good agreement with crystal structures and the Fe – O_2 bond length was refined to 1.9 Å. XANES spectra at low X-ray flux/low energy resolution and high X-ray flux/high energy resolution showed no significant differences and were in good agreement with previous studies. Radiation damage, which is a problem in experiments at high X-ray flux (e.g. crystallography), was clearly absent in our XAS data. The XANES/EXAFS results verified (i) near-quantitative ligand binding in HB/MB samples, (ii) physiologically relevant electronic configurations at cryogenic temperatures and (iii) absence of radiation damage at cryogenic and near-ambient temperatures. However, the differences in the XANES/EXAFS spectra of **deoxy** and **oxy** are not sufficient to discriminate between the Pauling, Weiss and McClure/Goddard models of Fe – O_2 bonding. Therefore, extended time-resolved high-resolution XAS and XES studies were performed as described in the next section.

5.4 Spin states from Kβ XES

Variations in Kα and Kβ XES correspond to changes in metal spin and oxidation state and the Kβ main-line has been used to investigate metal spin states in the past [25, 87, 211]. Specifically the relative $K\beta_{1,3}$ amplitudes and the Kβ′ feature at ~7045 eV reflect the local electronic configuration of the absorbing metal. This section shows the Kβ main-line emission experiments that were aimed to yield local iron spin state configuration.

Non-resonantly excited Kβ main-line emission of HB and MB with **deoxy**, **carboxy**, and **oxy** hemes and the three porphyrin model complexes (**1**, LS Fe(II); **2**, HS Fe(II); **3**, HS Fe(III); see Figure 5-2) were collected at 20 K and 260 K. The spectra are shown in Figure 5-8. For X-ray sensitive samples (see next section) a time-resolved energy-sampling approach (see section 4) [121] instead of a conventional energy step-scan was applied. Both proteins in all states and the porphyrin model complexes showed temperature-independent Kβ main-line spectra (Figure 5-8). Kβ’ and $K\beta_{1,3}$ emission was measured for different temperatures ranging from 20 K – 260 K for **oxy** HB/MB. The spin state sensitive Kβ’/$K\beta_{1,3}$ ratio is shown in Figure 5-9. It was also found to be temperature invariant. The shapes of the main-line emission for HB and MB samples in equal states were the same within noise limits and $K\beta_{1,3}$ peak energies for MB/HB

Table 5-3: Fe $K\beta_{1,3}$ line energies of heme species.

heme samples		$K\beta_{1,3}$ energy / eV	porphyrin model complexes	$K\beta_{1,3}$ energy / eV
carboxy	HB	7058.0(1) [7057.9(2)]	**1** Fe(II) LS	7058.0(1)
	MB	7057.9(1)		
oxy	HB	7058.4(1) [7058.3(2)]	**2** Fe(II) HS	7058.5(1)
	MB	7058.3(1)		
deoxy	HB	7058.6(1)	**3** Fe(III) HS	7058.7(1)
	MB	7058.6(1)		

*$K\beta_{1,3}$ line energies correspond to the first-moment position ($M = \Sigma[I(E)*E] / \Sigma [I(E)]$) of non-resonant spectra in Figure 5-8 (mean of 20/260 K data, energy range in parenthesis). The first moment positions are calculated for emission energies (E) within 7055 – 7062 eV. Energies in brackets are estimated from Kβ data in ref. [21] (energies were shifted to match **deoxy** value and energy differences scaled to match data in Figure 5-8, estimated error in parenthesis) [23].*

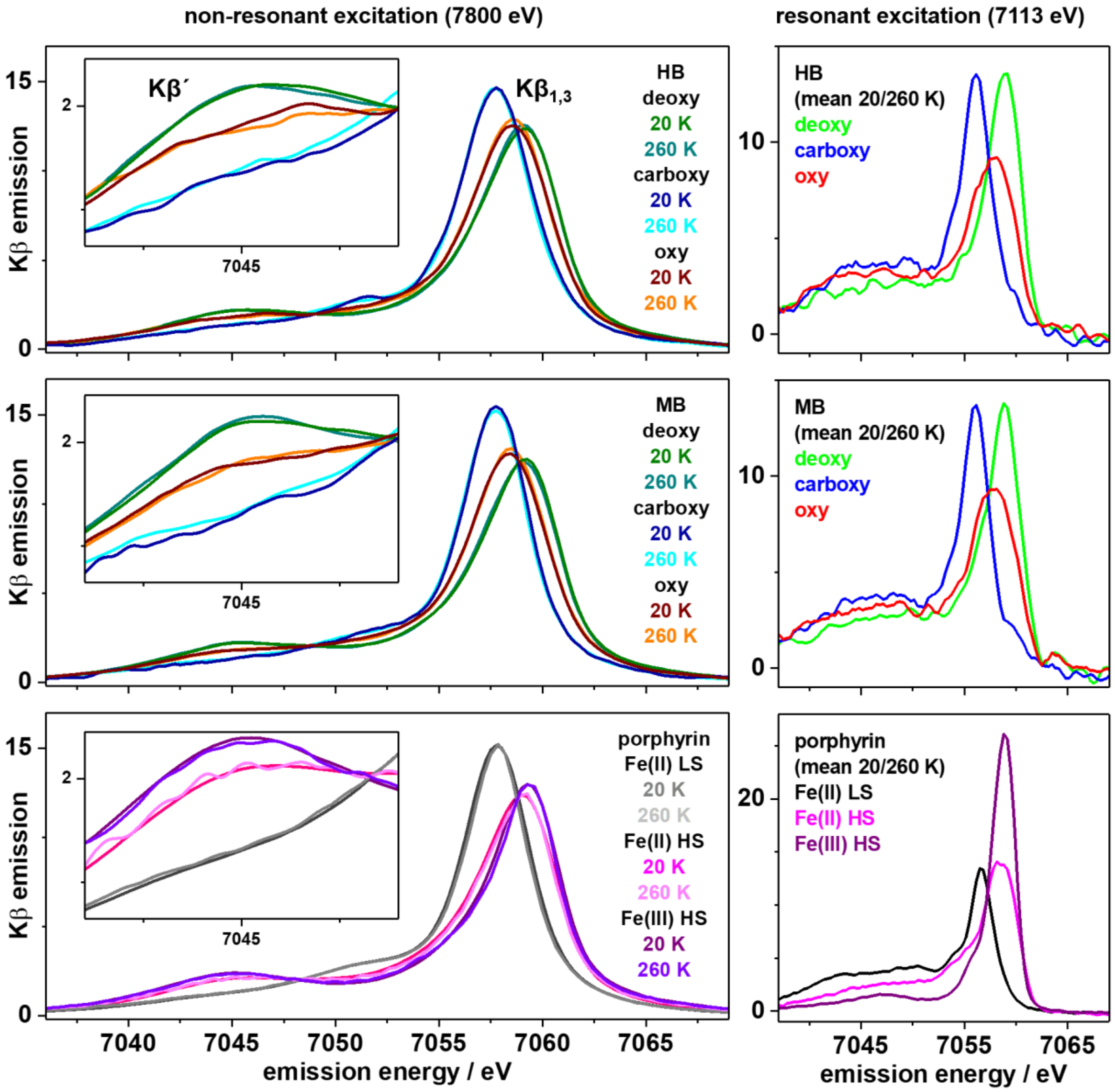


Figure 5-8: Kβ emission spectra of MB, HB, and porphyrin references. Left: non-resonantly excited, normalized Kβ main-line spectra at 20 K or 260 K and magnified Kβ′ feature in the inset (top, MB solution; middle, HB solution; bottom, porphyrin compound powders). Right: resonantly excited, normalized Kβ main-line spectra (mean of data at 20 K and 260 K).

agreed within ± 0.1 eV (see Table 5-3). Hemoproteins and porphyrin compounds in a low-spin (**carboxy**, **1**) or a high-spin (**deoxy**, **2**, **3**) configuration showed equal first-moment energies. In contrast, the $K\beta_{1,3}$ peak energy for **oxy** HB/MB is closer to first-moment positions of high-spin ($\Delta E \approx 0.2$ eV) than low-spin ($\Delta E \approx 0.4$ eV) compounds. This trend persists when compared to average first-moment positions from iron compounds in different oxidation and spin states (LS: 7057.7(3), HS: 7058.6(2); maximum spread in parenthesis) shown in Figure 5-10A. $K\beta_{1,3}$ first-moment positions are influenced by the ligand environment [23].

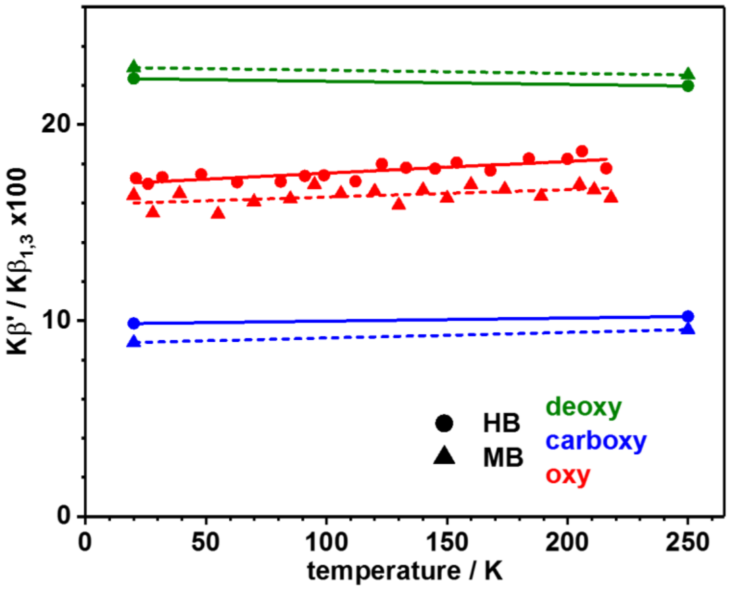


Figure 5-9: Temperature-independence of Kβ line ratios in MB and HB. For the three heme species, intensity ratios at 20 K and 260 K correspond to main-line emission amplitudes at 7045 eV (Kβ′) and the Kβ$_{1,3}$ *maximum of non-resonantly excited spectra in Figure 5-8 without background subtraction. For oxy, the further data points were derived from absolute emission intensities at 7045 eV (Kβ′) and 7057 eV* (Kβ$_{1,3}$)*, which were measured (0.5 s acquisition) on separate sample spots during stepwise lowering of the temperature from 260 K to 20 K and amplitude normalized to the mean 20/260 K intensity ratios. Lines show linear fits to the data points (dashed lines, MB; solid lines, HB). Slight apparent line slopes result from minor variations in background due to sample thickness variations or temperature and are not related to iron spin state changes within the 20 – 260 K temperature range.*

The variation in intensity of the Kβ’ feature intensity reflects the local electronic configuration (oxidation/spin state) of the iron atom (see section 2.2). The Kβ′ amplitudes for MB/HB, the porphyrins, and for non-porphyrin compounds were therefore determined and plotted versus the formal number of unpaired Fe(3d) spins, which is shown in Figure 5-11. The ~ two-fold increase of the Kβ′ amplitude for an Fe(3d) spin count increase from 0 in low-spin Fe(II) to 5 in high-spin Fe(III) was well described by a linear relation. The Kβ′ amplitudes of **deoxy** MB/HB agreed well with an Fe(3d) count of 4 as in the porphyrin compound **2**, in agreement with a high-spin Fe(II) configuration. The Kβ′ amplitudes of **carboxy** MB/HB were centered close to the 0 unpaired Fe(3d) spins expected from a low-spin Fe(II) configuration as in porphyrin compound **1**. For **oxy** MB/HB, the Kβ′ amplitudes varied only within ~ 5 % and the mean value was in best agreement with an Fe(3d) spin count of 2. Although the analysis shows a very good correlation between Kβ′ amplitudes and Fe(3d) counts, the absolute Kβ′ amplitudes

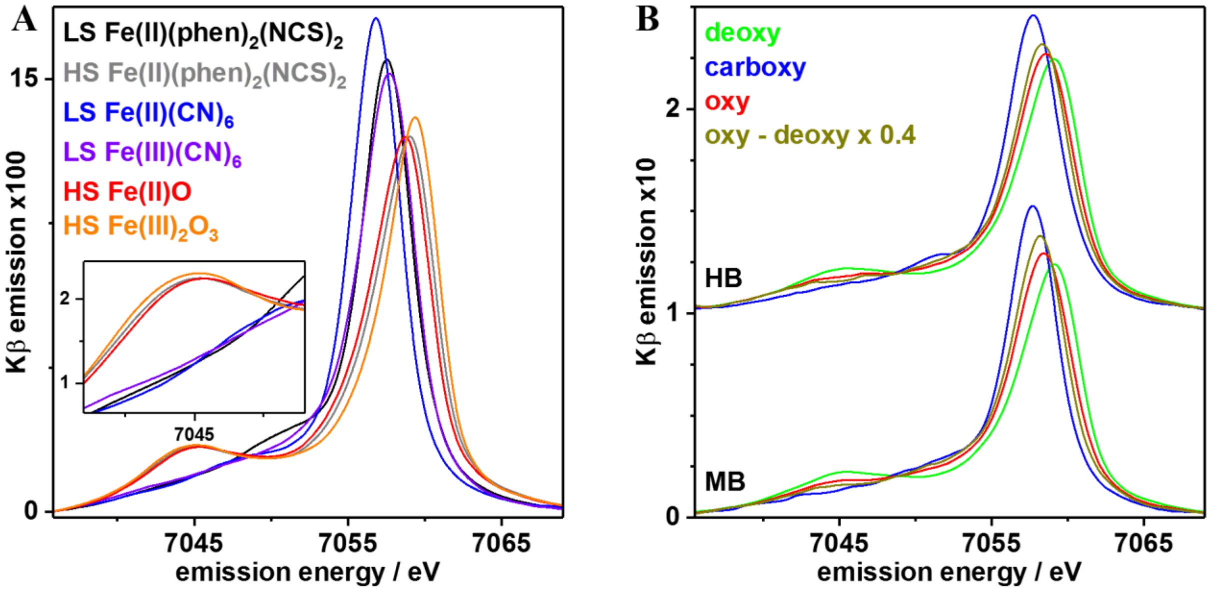


*Figure 5-10: (**A**) Kβ spectra of non-porphyrin compounds. Normalized spectra of microcrystalline (powder) materials of the indicated oxide or cyanide complexes and of the molecular compound were normalized and measured at 20 K, except of the spectrum of HS Fe(II)(phen)$_2$(NCS)$_2$ (phen = 1,10-phenanthroline) showing abrupt and complete spin crossover (LS → HS transition) at ~175 K [16], which was measured at 260 K. The Inset shows the Kβ' feature. (**B**) Purity of Kβ spectra of oxy MB/HB. Mean spectra at 20/260K for the three heme species are compared to (re-normalized) spectra of oxy after subtraction of 40 % of the deoxy spectra. This procedure resulted in distorted line shapes (i.e. in the Kβ′ region) and a respective large deoxy contribution was excluded based on the UV/vis and XAS data.*

changes are rather small. Consequently, the presence of small populations with higher or lower unpaired Fe(3d) count in the heme samples could significantly affect the results. Small variations may stem from incomplete reduction of met heme to **deoxy** heme or presence of **deoxy** heme by incomplete **oxy** heme formation during sample preparation. However, the conversion of the Kβ′ amplitude of **oxy** HB/MB to the smaller Kβ′ amplitude for 1 spin (as for low-spin Fe(III)) required subtraction of, i.e., a ~ 40 % contribution of **deoxy** from the **oxy** spectra (see Figure 5-10B). Such a subtraction resulted in distorted line shapes and was unlikely regarding near-quantitative **oxy** formation deduced from the optical absorption and XAS data (see previous sections). Formation of **deoxy** through X-ray exposure was excluded by temperature independent spectra from **carboxy** MB/HB. The **carboxy** samples showed the same X-ray sensitivity at 20 K as **oxy** samples but radiation damage at 260 K was absent even after prolonged X-ray exposure (see next section). A low-spin Fe(III) contribution from met heme would even lead to underestimation of the Kβ′ amplitude of **oxy**. An in-depth fit analysis of the Kβ spectra revealed detailed relations between peak energies/intensities and converged towards the same spin assignments as above (Figure 9-1– 3 in the appendix). These results strongly suggested that the high-spin Fe(II) in **deoxy** MB/HB is converted to an apparent intermediate-spin (IS) iron species in **oxy** upon O_2-binding to the metal.

A single earlier Kβ study on dried HB/MB at room temperature is available [19]. Respective Kβ' amplitudes are shown in Figure 5-11 (black stars/squares). Intensities from hemin, **deoxy** HB and **carboxy** HB are in good agreement (after alignment) with data from Figure 5-8 for Fe(3d) counts of 5, 4 and 0 and the Fe(3d) count of ~ 1 for Fe(III) in met HB/MB agreed with dominance of the low-spin species [206]. The Kβ′ amplitude of **oxy** HB suggested an Fe(3d) count of 2, hence yielding the same result as our data. The IS Fe(II) population in **oxy** therefore seems not to depend on the protein hydration level and persists at ambient temperatures.

The varying iron spin configurations result in different emission spectra after excitation of a core electron into an unoccupied Fe(3d) orbital (see section 2.2). Thus, resonantly excited Kβ spectra were collected at an energy in the K-pre-edge region (E_{exc} = 7113 eV; right in Figure 5-8). The spectra for MB/HB samples and the porphyrin compounds revealed pronounced shape differences. Despite significant variation, low- and high-spin compounds still yield similar first-moment positions/$K\beta_{1,3}$ peak energies (LS: 7056.9(1) eV, HS: 7058.4(2) eV; maximum spread in parenthesis) while **oxy** HB/MB (7058.0(1) eV) remains closer to the HS compounds. The Kβ′ feature increased for decreasing Fe(3d) counts and a plot of the Kβ′ amplitudes is shown as an inset in Figure 5-11. The number of unpaired Fe(3d) electrons in this plot is rationalized assuming a final state after photon emission with (i) no antiparallel unpaired 3d spins and (ii) no spin-flip of 3d spins during excitation and/or emission. Consequently, a low-spin d^6 configuration yields a low-spin d^7 configuration upon exciting an electron and **carboxy** and porphyrin compound **1** gain one unpaired Fe(3d) electron. The high-spin d^6/d^5 configurations (**deoxy**, compounds 2/3) are converted into high-spin d^5/d^4 states with a by one unit diminished unpaired Fe(3d) number. A low-spin d^5 state of **oxy** (as suggested by Weiss) would than be transformed into LS or IS d^6 configurations and, on average, yield spectra corresponding to one unpaired Fe(3d) spin (LS: Fe(3d) = 0, IS: Fe(3d) = 2). Accordingly, an IS d^6 state (McClure and Goddard) results in a final state with LS or HS d^7 configuration, which corresponds to an average of 2 unpaired Fe(3d) spins (LS: 1, HS: 3). Alignment of Kβ' intensity from **deoxy** and **carboxy** HB/MB and porphyrin compounds in the above manner showed a linear relation (Figure 5-11). For **oxy** HB/MB, the Kβ′ amplitudes were close to 2 unpaired Fe(3d) spins, which indicated a higher spin state than in **carboxy** and further supported an apparent intermediate-spin iron atom in **oxy**.

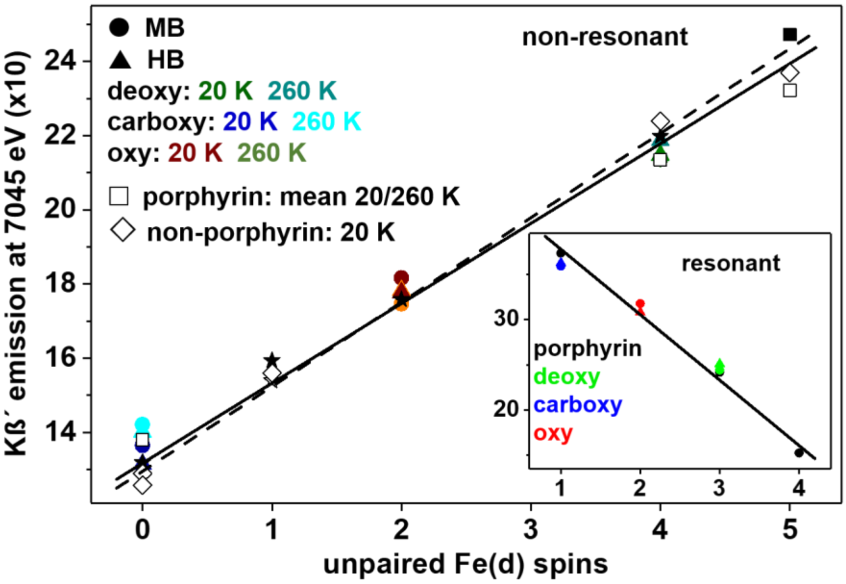


Figure 5-11: Correlation of Kβ´ intensity and formal unpaired Fe(3d) spin count. Data points refer to mean amplitudes (at 7045 ± 0.5 eV) of Kβ spectra in Figure 5-8 (left) and Figure 5-10A (non-porphyrin compounds). Kβ data for dehydrated ***deoxy****,* ***carboxy****, and* ***oxy*** *HB and* ***met*** *(at Fe(3d) = 1) MB/HB (★) and solid hemin (■) at room temperature derived from ref. [19] were scaled to our data at Fe(3d) = 0. Fe(3d) spins correspond to: 0, LS Fe(II); 1, LS Fe(III); 4, HS Fe(II), and 5, HS Fe(III). Kβ´ intensities for oxy MB/HB were placed at Fe(3d) = 2 for a formal IS Fe(II) (solid or dashed lines, linear fits to porphyrin/non-porphyrin or porphyrin data). Inset: data from Kβ spectra in Figure 5-8 (right, for Fe(3d) count calculation see the text; line, linear fit to porphyrin data).*

Conclusion

For the first time, quantitative Kβ emission spectra for non-resonant and resonant excitation were collected for two different heme proteins in three states each and for several synthetic porphyrin model complexes in a wide temperature range. Spectra were not affected by temperature variation in agreement with the XAS results. The detected changes in relative Kβ’ amplitudes straightforwardly corresponded to varying metal spin and oxidation states. Kβ’ amplitudes were directly proportional to the formal unpaired Fe(3d) spin count. This revealed the number of 0 unpaired electrons in low-spin Fe(II) samples (**carboxy**, **1**), 4/5 in high-spin Fe(II)/Fe(III) samples (**deoxy**, **2**, **3**), and, most important, 2 effective unpaired electrons for **oxy**. These results clearly indicate that the electronic configuration of the iron center in oxygenated heme proteins (**oxy**) corresponds to an effective intermediate-spin Fe(II). This result differs from the proposed LS Fe(II) or Fe(III) configurations as suggested by Pauling and Weiss. A quantitative interpretation of our unexpected finding of 2 effective unpaired spins in **oxy** was attempted and achieved by the advanced quantum chemical calculations described further below.

5.5 X-ray induced ligation changes

XAS and XES experiments at synchrotron beamlines with high photon flux have to account for radiation damage from X-ray exposure, especially for compounds in higher oxidation states. The radiation sensitivity of a sample is largely unpredictable and save exposure times can vary from a few seconds to minutes. XANES is particularly sensitive to metal and ligand centered variations and is well suited to monitor radiation damage. This section presents XAS/XES spectra from radiation damage experiments.

X-ray fluorescence time traces were recorded at a fixed excitation energy in the K-edge rise to monitor radiation-induced changes at the iron centers (see Figure 5-13 and inset in Figure 5-12A). Fe(II) porphyrin compounds **1** and **2** showed no changes in time-scans of ~ 5 min at 20 K and 260 K (not shown) and consequently, no shifts in K-edge position. The high-spin Fe(III) compound **3** showed a rise in X-ray fluorescence time traces at 260 K, which was absent at 20 K. The time-scan data was in good agreement with a single-exponential fit (τ ~ 150 s). The increase in fluorescence reflects an energy-downshift of the K-edge (Figure 5-12A). The edge shape was not affected in addition to the shift. Kβ main-line spectra showed a downshift of $K\beta_{1,3}$ peak position and a decrease in Kβ’ intensity. These results, together with the mono-phasic behavior, revealed a metal-centered reaction and indicate an Fe(III) to Fe(II) reduction without a change in spin state, coordination symmetry or ligation.

In contrast, X-ray fluorescence time traces of HB/MB samples revealed remarkable differences in response to X-ray exposure at 20 K or 260 K. Both proteins showed similar time-scan traces for the same state and temperature. **Deoxy** showed no change in absorption at cryogenic temperature during ~ 3.5 min of X-ray exposure while **oxy** and **carboxy** revealed a bi-phasic absorption increase. The latter traces were well described by bi-exponential fits ($\tau_{\mathbf{oxy}}$ ~ 5 / 75 s; $\tau_{\mathbf{carboxy}}$ ~ 10 / 150 s). Finally, the fluorescence of **oxy** and **carboxy** converged to the intensity of **deoxy** after long X-ray exposure. At 260 K **carboxy** absorption remained constant for the ~ 3.5 min of exposure while the absorption of **deoxy** decreased with a single phase (τ ~ 25 s) down to the intensity of **carboxy**. Time traces from **oxy** showed an initial increase (τ ~ 8 s), surpassing the **carboxy** level, and a subsequent decrease (τ ~ 18 s) to the **carboxy** amplitude. Note, that the decrease of **oxy** and **deoxy** had similar kinetics. The Kβ main-line from **deoxy** at 260 K after prolonged X-ray exposure of (~ 2.5 min) revealed a transition from a high- to a low-spin configuration, as shown in Figure 5-14A, which coincided with an upshift and change in spectral shape of the K-edge. In addition, the pre-edge increased in intensity and gained a second prominent feature. The latter corresponds to a ligand-centered ctv transition due to the

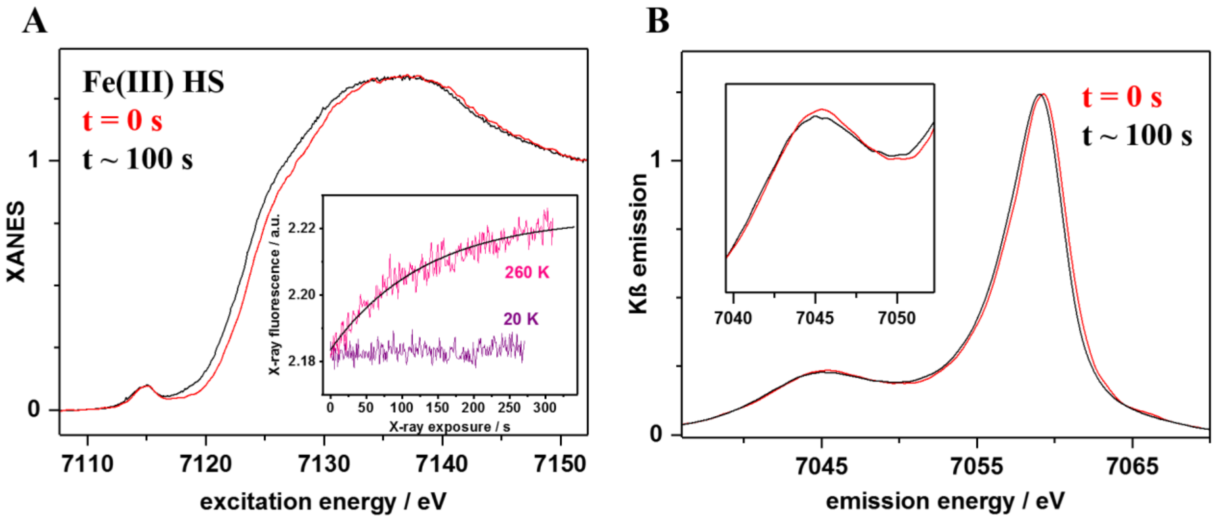


*Figure 5-12: X-ray induced change in oxidation state for porphyrin compound **3** ($[FeTPPCl]^0$, see Figure 5-2A). (**A**) Higher-resolution total-fluorescence XANES spectra at ~ 0 s and ~ 100 s of X-ray exposure. Inset: X-ray fluorescence (at 7124 eV, arbitrary units) time-scan data at 20 K and 260 K. The data was shifted to equal values at t = 0 s for clarity. The solid line shows a single-exponential fit. (**B**) Non-resonant Kβ main-line emission spectra at ~ 0 s and ~ 100 s of X-ray exposure. The inset shows the enlarged Kβ' feature. Note the decrease in Kβ′ intensity at ~100 s.*

low-spin state. These spectral changes were absent at 20 K. For **carboxy**, Kβ main-line, pre-edge and XANES measurement essentially showed the reverse behavior. Spectra at 20 K before and after X-ray exposure showed a transition from low- to high-spin, a decrease in pre-edge intensity (especially of the higher-energy feature) and a downshift in K-edge, all of which was absent at 260 K (see Figure 5-14C). Consequently, **deoxy** shows absence of photoreduction of the ferrous iron at 20 K and binds a mobile CO species at 260 K. The CO ligand is detached by X-ray exposure at 20 K where carbon oxide species are immobile and re-binding is impaired. However, similarly fast carbon oxide binding and detachment explained the invariant data for **carboxy** at 260 K. Mobile CO species may be created for example by X-ray induced decarboxylation of amino acid side chains or CO photolysis [212, 213]. For **oxy**, the bi-phasic behavior at 20 K corresponds to formation of **deoxy** by unbinding of O_2, as suggested by an increase in Kβ' amplitude, a downshift in K-edge and an increase in pre-edge, which overall resembles the observations for **carboxy** at 20 K (see Figure 5-14B).Notably, this transformation is followed by a reaction similar to **deoxy** at 260 K (decreasing Kβ' feature, two peaks in pre-edge, upshift in XANES), which suggests subsequent formation of **carboxy** from **deoxy**. Note that the initial Kβ' increase and XANES downshift further support near-quantitative

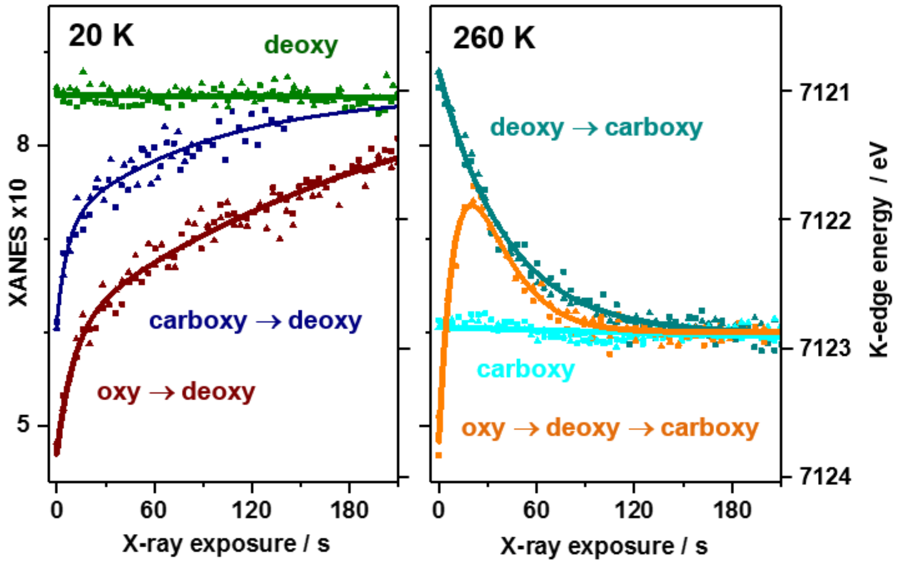


*Figure 5-13: Heme state conversion during X-ray exposure. Fluorescence changes monitored at 7123 eV in the XANES (see Figure 5-4) for indicated starting species in MB/HB (triangles/squares) at 20/260 K (left/right), approximate K-edge amplitudes/energies shown on left/right y-axes. Time traces correspond to ctv, XANES, and Kβ main-line changes (Figure 5-14) due to indicated apparent transitions between heme species. Solid lines: bi-exponential (**carboxy**/**oxy**, 20 K; **oxy**, 260 K), mono-exponential (**deoxy**, 260 K), or linear (**deoxy**, 20 K; **carboxy**, 260 K) fits.*

oxygenated HB/MB, as initially present **deoxy** would result in the opposite behavior. The bi-phasic kinetics of **carboxy** and **oxy** at 20 K suggest similar two-step process involving ligand-centered reduction. This argues against an initial superoxide bound at iron in **oxy**, but is consistent with similar bi-phasic behavior of **oxy** and **carboxy** in Resonance Raman and Mössbauer spectroscopy [61-63, 65]. The temperature dependent X-ray sensitivity of **carboxy** and **oxy** could presumably be an explanation for the variations in bond-distances of Fe – O/N$_{His}$ bonds in **oxy** from XRD in Table 5-2. The crystallographic data were obtained at about 100 K, which could cause O_2 release and CO species binding during diffraction data collection of **oxy**, while CO species in **carboxy** would presumably rebind to iron at this temperature.

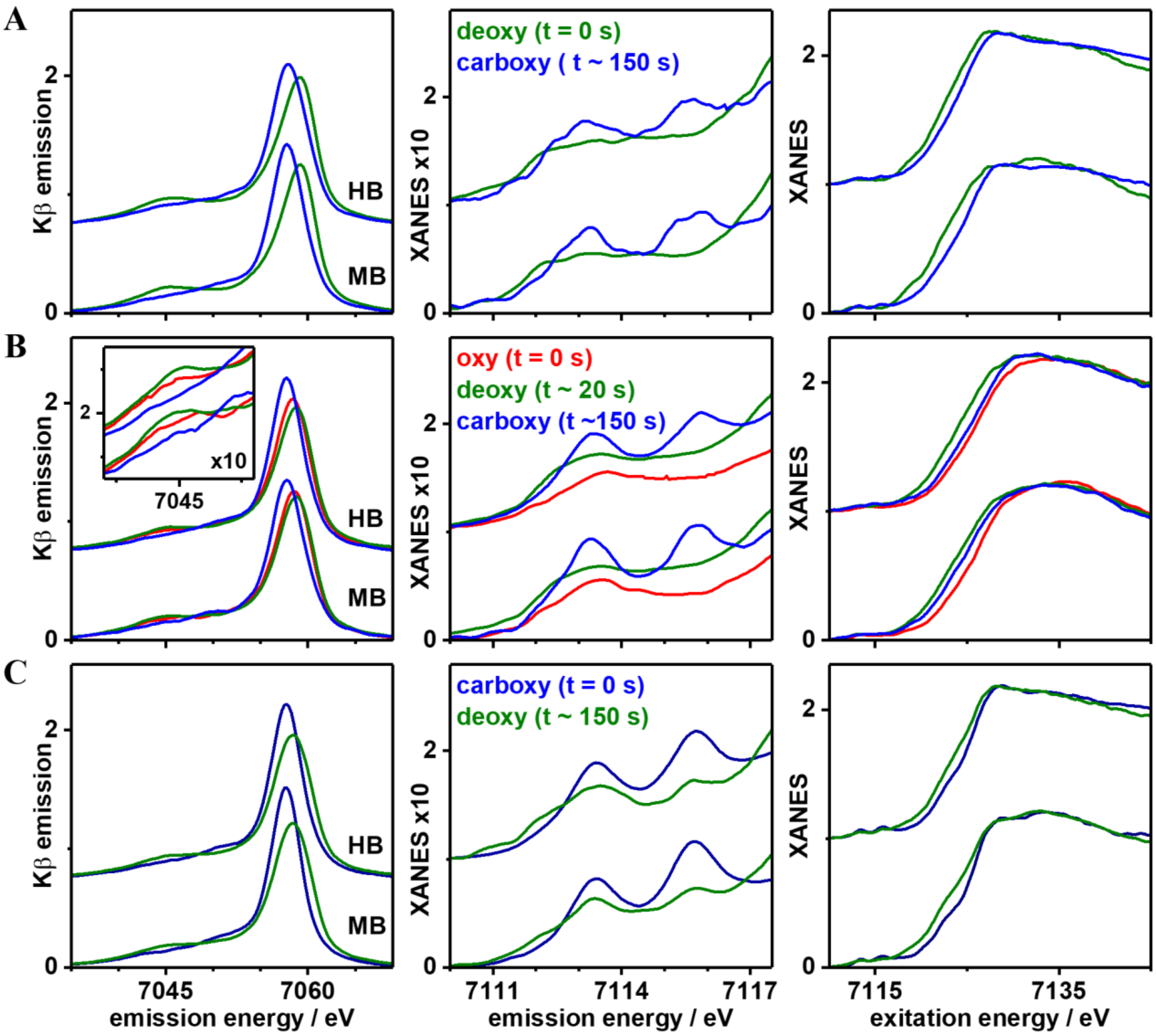


Figure 5-14: X-ray induced species modifications in ***deoxy*** *(****A****),* ***oxy*** *(****B****) and* ***carboxy*** *(****C****) HB/MB samples. Spectra (left, Kβ main-lines; middle, pre-edge absorption; right, Kα-detected XANES) were obtained at 260 K (A and B panels) or 20 K (C panels) for MB/HB containing the indicated initial heme species (at t = 0 corresponding to ~ 1 s of X-ray exposure) or after ~20/150 s of irradiation, i.e. close to the completion of X-ray induced apparent conversion of* ***deoxy*** *→* ***carboxy*** *(top), of* ***oxy*** *→* ***deoxy*** *→* ***carboxy*** *(middle), or of* ***carboxy*** *→* ***deoxy*** *(bottom) (Figure 5-13). Note, e.g., the increase of the Kβ′ feature followed by a decrease (B, left) when starting from* ***oxy****, which corroborates the virtual absence of a* ***deoxy*** *fraction in the* ***oxy*** *samples. The data indicate a reversible spin state transition LS Fe(II) → HS Fe(II) for* ***carboxy*** *→* ***deoxy*** *upon ligand release/rebinding and a IS Fe(II)→HS Fe(II)→LS Fe(II) conversion when starting from* ***oxy****.*

Conclusion

XANES time traces at 20 K or 260 K indicate pronounced changes of the heme species in HB/MB due to prolonged X-ray exposure, which is further supported by variations in Kβ main-line emission and pre-edge absorption spectra. At 20 K, X-ray exposure of **oxy** and **carboxy** samples induces rapid reduction and detachment of the axial ligand resulting in formation of **deoxy**. **Deoxy** showed no X-ray induced changes at cryogenic temperatures. At 260 K, **carboxy** was unaffected by X-ray exposure, **deoxy** species transformed into **carboxy**, but **oxy** first formed **deoxy** and subsequently **carboxy**. These radiation-induced coordination changes at the iron center presumably explain in particular the Fe – O_2 bond length heterogeneity in crystal structures. In our experiments, rapid X-ray induced modification of heme cofactors was avoided by use of the time-resolved energy-sampling approach for data collection. Combination of advanced X-ray spectroscopy methods with crystallographic structures facilitates refinement of metal site structures for reliable electronic configuration determination.

5.6 Electronic structure from vtc XES and ctv XAS

Pre-edges from XAS correspond to resonant excitations into unoccupied states and pre-edge spectra carry information about coordination symmetry, metal oxidation state and spin state [20]. Detecting X-ray fluorescence after resonant excitation with a high energy resolution yields spin-polarized pre-edge spectra because of energy splitting in Kβ main-line emission after resonant excitation (see section 2.2) [20, 25, 81]. Non-resonant Kβ satellite lines from XES are sensitive to the character of surrounding ligands [33]. This section analyses experimental valence-to-core (vtc) and core-to-valence (ctv) spectra from XES/XAS with regard to the electronic configuration of heme samples and porphyrin model complexes.

High-resolution pre-edge absorption (ctv) and non-resonantly excited Kβ satellite emission (vtc) spectra were collected on MB/HB and the porphyrin model complexes. The satellite lines are shown in Figure 5-15 and the pre-edges in Figure 5-16. Spectral features were almost identical at 20 K and 260 K, which justified averaging to obtain mean ctv/vtc spectra of the porphyrins and the heme species. The small variations in averaged spectra of HB and MB reflected only subtle differences between both proteins and were therefore also averaged to obtain spectra shown in Figure 5-17.

High-resolution Kβ satellite line emission spectra

The satellite lines showed ~ 1 eV energy differences and intensity variations of the $K\beta_{2,5}$ feature. The three high-spin complexes (**deoxy**, porphyrin **2**, **3**) showed a more pronounced double feature than porphyrin model complex **1**, **carboxy** and **oxy**. The small energy-shifts from varying axial ligands indicated that the vtc transitions were dominated by the porphyrin ligand (compare Figure 2-5B). Consequently, the first moment positions of the satellite lines in Figure 5-17A show a weak correlation with the average nucleus charge (including the four nitrogen atoms on porphyrin) surrounding the iron metal ($R^2 \approx 0.8$) in the order **carboxy** < porphyrin **1** < **deoxy** < porphyrin **2** < **oxy** < porphyrin **3**. The Kβ satellite emission is compared with vtc spectra from DFT in section 5.6.

Narrow- and broad-band detected pre-edge spectra

The spectra from ctv XAS reflected a more intricate dependence on the varying ligands. The ctv spectra of the heme species and porphyrin complexes showed 2 – 3 peaks varying by ~ 1 eV in energy and ~ 2-fold in intensity. Variations in pre-edge intensities of HB/MB samples and porphyrin compounds corresponded to differences in the coordination symmetry and bond

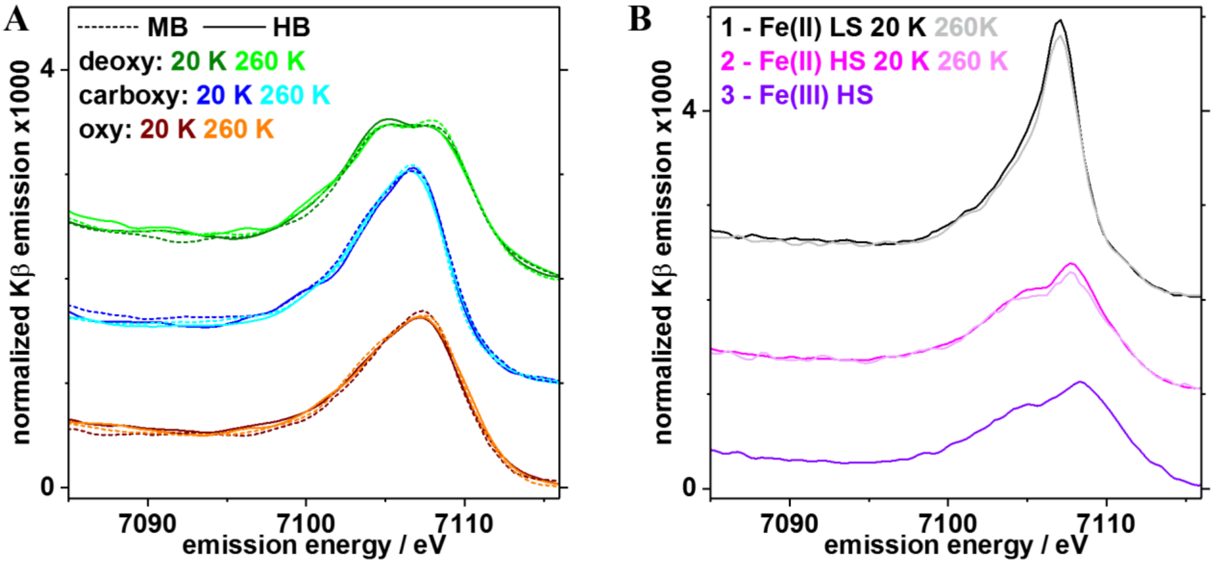


*Figure 5-15: Kβ satellite emission spectra at 20 K and 260 K (vertically shifted) of HB/MB (**A**) and porphyrin compounds (**B**) approximately normalized to the respective Kβ main-line spectra (Figure 5-8), see chapter 2.3 for details.*

lengths as confirmed by EXAFS and DFT (see following section). For the model complexes, five-coordinated compound **3** shows a larger absorption than the two six-coordinated compounds. Additionally, the asymmetric ligand environment (see Figure 5-2) of the high-spin Fe(II) model complex yields a larger absorption than the symmetric ligation of low-spin Fe(II). Areas of the ctv peaks for heme species decreased in the order **deoxy** > **carboxy** > **oxy** (excluding the second peak from **carboxy**), which reflects (i) the increase in number of coordinated ligands and (ii) the shorter Fe – CO compared to the Fe – O_2 bond length (see Table 5-2), and therefore larger distortion. Comparing pre-edge features from broad-band Kα, narrow-band $K\beta_{1,3}$ and Kβ' emission detection further supports spin state assignments from Kβ main-line emission. For compounds **1** – **3**, ctv spectra show only subtle differences when comparing Kα and $K\beta_{1,3}$ detected spectra. However, the Kβ' detected absorption of the high-spin complexes revealed no pre-edge features while low-spin Fe(II) maintained a single feature. This reflects impaired Fe(1s) spin-up electron excitation into the all occupied spin-up Fe(3d) valence levels, whereas for the low-spin porphyrin, $K\beta_{1,3}$- and Kβ′-detected spectra were similar due to similar spin-up/down Fe(1s) excitation into the unoccupied levels (see section 2.2, Figure 2-6 and Figure 2-7). The energy difference between the two features from compound 2 implied an energy splitting between d_{x2-y2} (porphyrin ligand) and d_{z2} (axial THF ligands) and thus a distorted geometry. For **carboxy**, the relative amplitude of the first and second feature varies from ~ 1:1 (Kα), ~ 1:2 ($K\beta_{1,3}$) to ~ 2:1 (Kβ'), which suggested a metal to metal transition in the

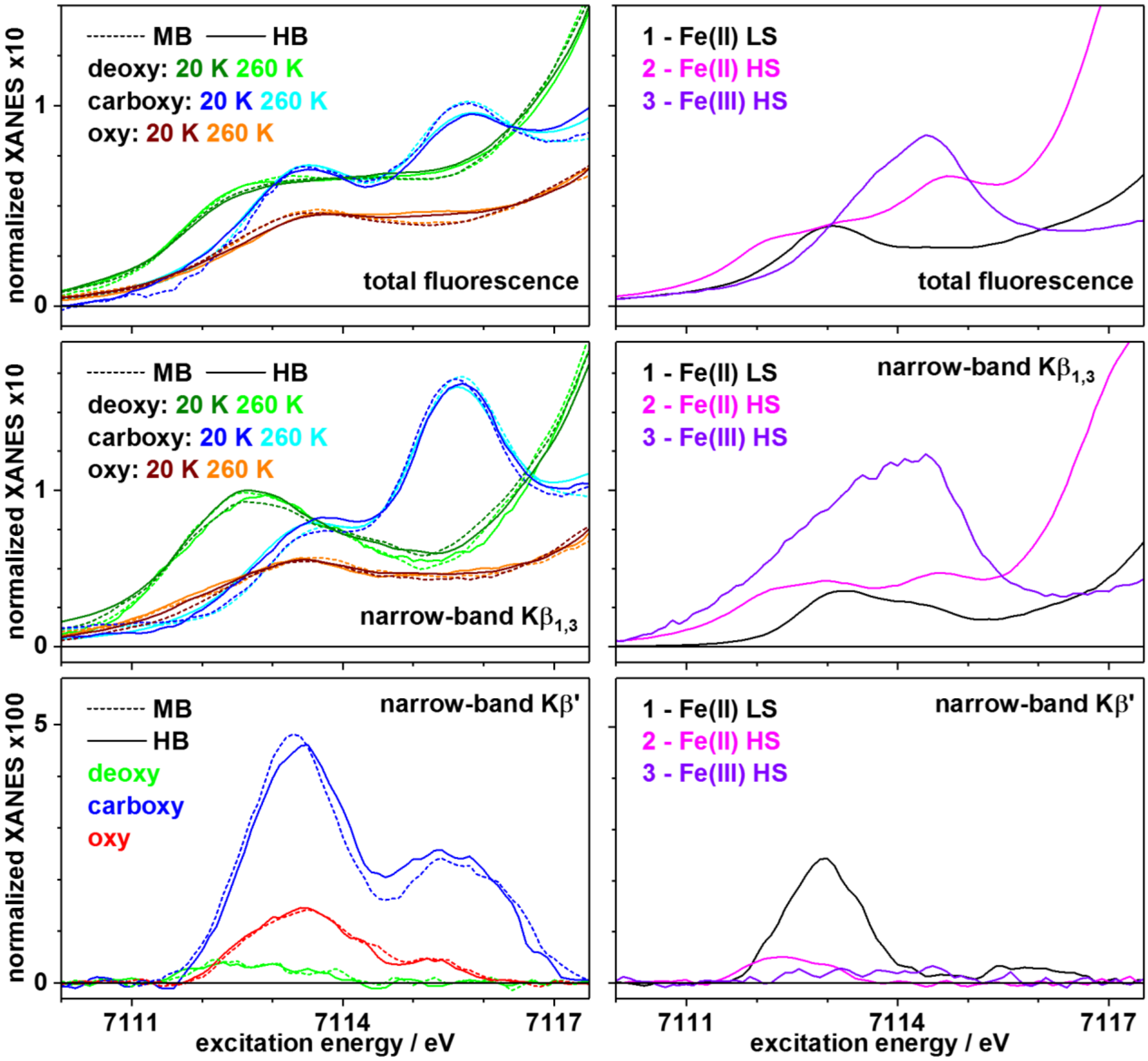


*Figure 5-16: Pre-edge absorption (ctv) spectra of MB/HB collected at 20 K and 260 K (**left**) and porphyrin compounds (**right**) collected at 20 K. Pre-edge absorption spectra at 20/260 K were measured (~1 s scan duration) at conditions similar to the XANES data in Figure 5-4B, but within a more narrow energy range (7109-7119 eV), and scaled to the respective XANES spectra. Spectra were broad-band Kα (**top**), narrow-band* $K\beta_{1,3}$ *(**mid**) or Kβ' (**bottom**) detected. The detector energy for* $K\beta_{1,3}$ *ctv spectra was set to the respective peak maximum (**deoxy**: 7059 eV, **carboxy**: 7057.5 eV, **oxy**: 7058 eV, 1 – 3: 7057.5 eV; see Figure 5-8) and to 7045 eV for Kβ' ctv spectra. The Kβ' detected ctv spectra were obtained after subtraction of smooth polynomial background curves due to the main K-edge rise.*

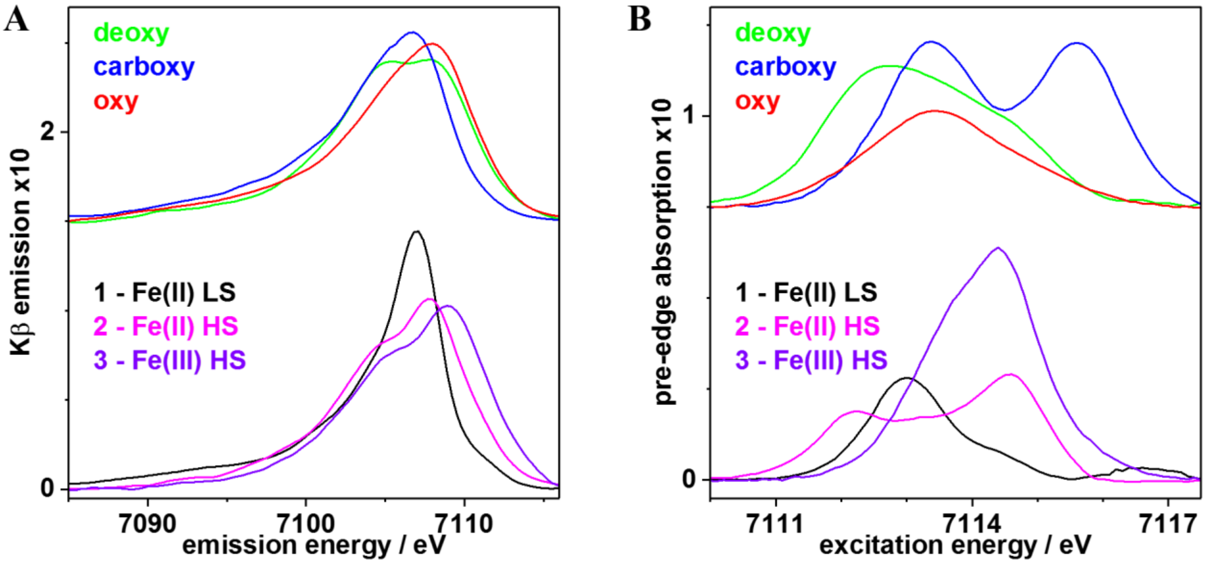


*Figure 5-17: Background subtracted Kβ satellite emission (A) and pre-edge absorption (B) spectra. Spectra correspond to averages of 20 K and 260 K of respective compounds and heme states correspond to mean HB/MB spectra. (**A**) The vtc spectra were obtained after subtraction of a Lorentzian fit and area normalization of Kβ satellite emission spectra from Figure 5-15. (**B**) The ctv spectra correspond to background corrected (smooth polynomial) total fluorescence detected pre-edges from Figure 5-16.*

first peak (larger (3p/3d) exchange interaction) and a metal to ligand excitation in the second peak (smaller interaction). Similar ctv amplitudes in Kα- and $K\beta_{1,3}$- versus vanishing absorption in Kβ'-detected spectra for **deoxy** was in agreement with high-spin Fe(II) while, for **carboxy**, the persisting first peak in all three detection modes agreed with the low-spin Fe(II) configuration. The smaller amplitudes vs. the Kα-detected spectra indicated a higher or lower spin state in **oxy** vs. **carboxy** or **deoxy**. Additionally, the first peak of **carboxy** showed an upshift from **deoxy**, which was consistent with the upshift of compound **1** and **2** (see Figure 5-17B) and literature (see Figure 2-2). Consequently, a similar behavior would predict pre-edge features from a low-spin Fe(III) complex at a similar or higher energy than compound **3**. Thus, the equal first peak position of **oxy** and **carboxy** suggested a similar oxidation state, further supporting two unpaired spins in **oxy**.

Conclusion

A complete set of high-resolution broad-band Kα and narrow-band $K\beta_{1,3}$/Kβ' detected pre-edge spectra and high-resolution satellite line spectra was collected for two heme proteins in three states and three synthetic porphyrin model compounds in a broad temperature range. Variations in $K\beta_{1,3}$/Kβ’ detected ctv spectra clearly demonstrate the high-spin (**deoxy**, **2**, **3**) and low-spin (**carboxy**, **1**) configurations and a CO-ligand centered pre-edge absorption peak in **carboxy**. Relative absorption intensities and energy variations are in line with coordination numbers and symmetry distortions from EXAFS. Satellite lines showed more pronounced double features for HS compounds than spectra from LS compounds and the emission is dominated by the porphyrin moiety. Both, ctv and vtc spectra for **oxy** indicate a non-low-spin configuration of the Fe – O_2 site. Reliable interpretation of this finding was achieved by computational considerations in the following section.

5.7 Electronic configuration from DFT

In recent years significant progress has been made in correlation of K-edge XAS and XES spectra with predictions from DFT calculations [25, 33, 35, 83]. In this section, results from DFT and TD-DFT calculations of heme states in proteins and model compounds are presented, which, in combination, yield geometric structures, as well as ctv and vtc spectra, and are used to evaluate the results of the synchrotron based experiments in the previous section.

Structures and spectra of heme species from different DFT functionals

The results of density functional theory (DFT) calculations often depends critically on the choice of functional. In particular, predictions of quantities like energetic order of spin states, reaction path energies or geometries may vary with the functional [4, 141, 153]. Consequently, heme geometries were optimized with four DFT functionals and resulting structures compared to bond lengths from EXAFS in section 5.3. Vtc spectra were calculated with DFT and ctv spectra with TD-DFT (see section 3.1) and compared to experimental spectra from the previous section. In order to cover a wide range of DFT functionals, four representative candidates were selected. The ‘pure’ GGA DFT functional by Becke and Perdew (BP86), the meta-GGA DFT functional by Tao, Perdew, Staroverov and Scuseria (TPSS) as well as its hybrid counterpart with 10% HF exchange (TPSSh) and the popular hybrid functional B3LYP with 20% HF exchange [134, 136-138, 214, 215]. A single heme site for each state was used as an initial guess for a constraint geometry optimization with B3LYP (**oxy**: pdb code: 2DN1/ chain identifier: A / residue sequence number: 142, **deoxy**: 2DN2/B/147, **carboxy**: 2DN3/A/142). The resulting geometry (see Figure 5-18a – c) was the initial guess for subsequent optimizations using the remaining three functionals with equal constraints. The geometries were afterwards optimized without constraints on the distorted porphyrin.

Comparing EXAFS parameters with structures from DFT

The optimized geometries, in the case of B3LYP, are shown in Figure 5-18 and parameters from DFT in Table 5-4 (**carboxy**/**deoxy** heme) and Table 5-5 (**oxy** heme). In tendency, bond lengths increased in the order BP86 < TPSS < TPSSh < B3LYP. The two non-hybrid DFT functionals converge to comparable bond lengths while the two hybrid functionals yield larger differences. Although absolute bond lengths change with functionals, some trends are conserved. The average iron distance to its first sphere ligands decreases from **deoxy** < **oxy** < **carboxy**. This reflects (i) the change in coordination from square pyramidal (**deoxy**) to octahedral (**carboxy**/**oxy**) and (ii) the stronger effect of CO versus O_2 on the trans axial ligand.

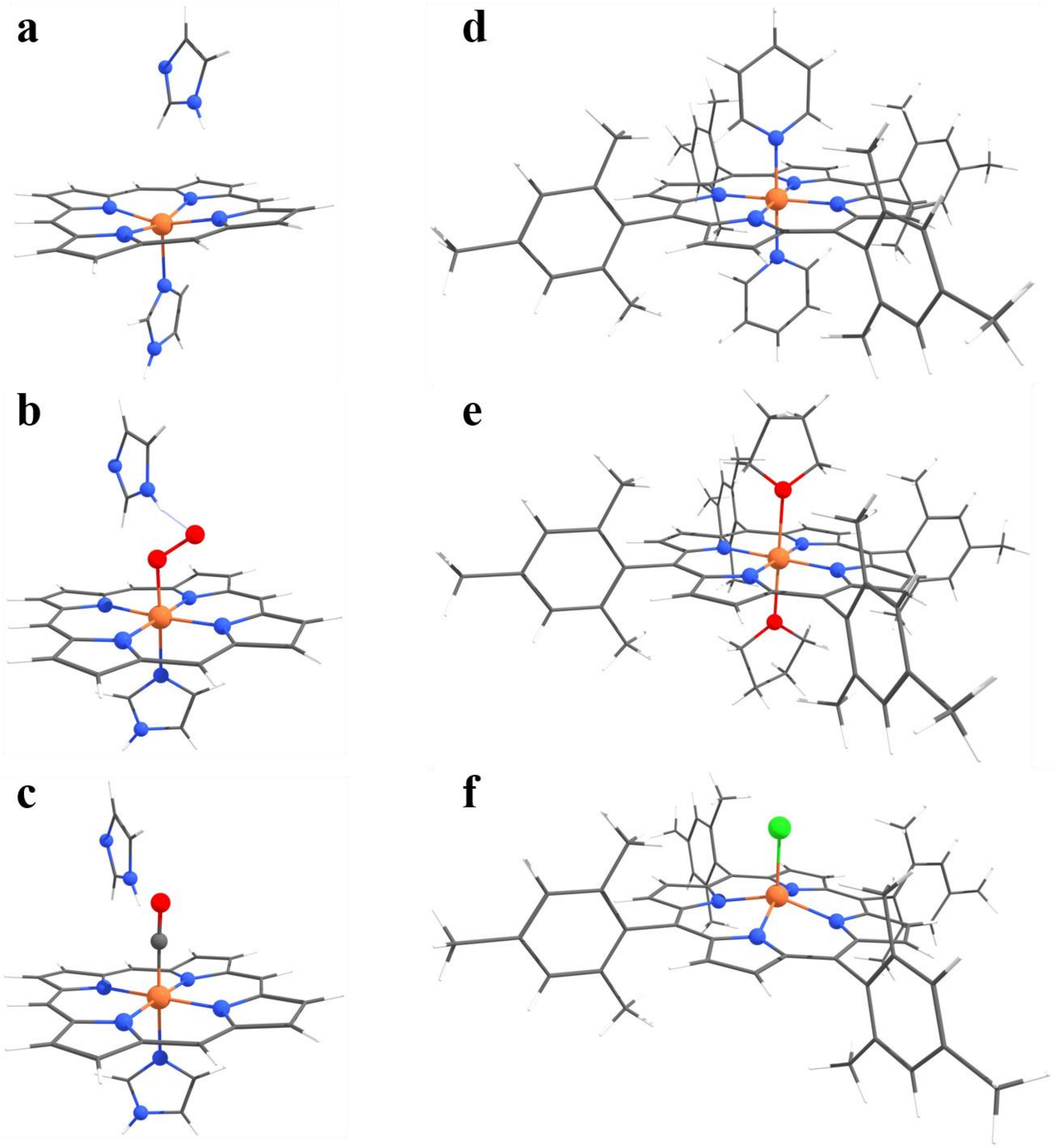


Figure 5-18: Result of geometry optimizations with DFT for ***deoxy*** *(a),* ***oxy*** *(b),* ***carboxy*** *(c), Fe(TPP)(Py)$_2$ (**1**, Fe(II) low-spin, d), Fe(TPP)(THF)$_2$ (**2**, Fe(II) high-spin, e) and Fe(TPP)Cl (**3**, Fe(III) high-spin, c). The calculations were done with the B3LYP functional and def2-TZVP(/J) basis set (see chapter 3.3). Carbon atom positions on the porphyrin ring were constrained for a – c, and f. See Figure 5-19, Figure 5-21, Table 5-4, Table 5-5 and Table 5-7 for metric parameters and calculated spectra.*

The shorter Fe – C bond length versus Fe – O_c stems from significant π-backbonding in **carboxy**. All functionals show a similar change in Fe – O_c/C_{CO} bond length from **oxy** to **carboxy** of ~ 0.1 Å.

The DFT functionals result in different Mulliken charges on iron for all three states, which increase in the same order as bond lengths. However, they all show a conservative trend with charges decreasing along **deoxy** > **oxy** > **carboxy**. Each functional results in comparable decrease in iron charge of ~ 0.4 upon binding of CO or O_2, which decrease again from BP86 > TPSS > TPSSh > B3LYP. The Mulliken charge on iron and distance to the axial histidine are inverse proportional ($R^2 > 0.92$) to the amount of diradical character in the wavefunction (see section 3.4). The diradical character decreases with increasing/changing HF exchange. It originates from a broken spatial symmetry in one Fe d_{xz} orbital (see Figure 5-22). The UCO has more contribution on porphyrin carbon atoms for the spin up part while the spin down part is larger on porphyrin nitrogen atoms. This is in part due to the distortion of the constrained porphyrin ring as extracted from the XRD structure but the same trend, though smaller in magnitude, is observed when constraints on atom coordinates are released. The remaining differences presumably stem from variations in modeling exchange interaction. The Fe – N_{His} bond length from EXAFS are in better agreement with predictions from TPSSh/B3LYP then BP86/TPSS. The good performance of B3LYP with regard to carbon monoxide compared to other methods/functionals was reported in the past [216]. Indeed, B3LYP is in best agreement with EXAFS with regard to the Fe – C_{CO} distance while all functionals predict similar Fe – O_{CO} distances. The difference between charge on CO and Fe correlate directly ($R^2 = 0.98$) with the Fe – C_{CO} distance.

In the **oxy** state, all functionals predict a broken-symmetry solution involving one electron pair, even though the broken-symmetry was not explicitly searched for. Thus, the broken-symmetry is a stable solution independent of functional. The origin of the spatial asymmetry can best be visualized with the UHF corresponding orbitals (UCOs), which can be seen in Figure 5-22. Each UCO shows a single spin up orbital on iron and a single spin down orbital on the oxygen molecule. The overlap integral between both spin parts of the UCOs, and thus the diradical character and charge transfer, is smallest for BP86 and largest for B3LYP. Therefore, all calculations predict an antiferromagnetic coupling (afc) between iron and O_2 of varying magnitude. This and the Mulliken spin density close or equal to one supports the Fe(III) low-spin configuration proposed by Weiss. The increase in afc stems from an increase in spin contamination of the DFT wavefunctions (see Table 5-6), which corresponds to an increasing admixture of a triplet state to the wavefunction. Both TPSSh and B3LYP show the largest spin contamination with approximately 50% triplet and 50% singlet configuration. Remember, that the total spin of zero is only conserved over the entire system (see section 3.1). All parameters regarding the interaction between iron and oxygen (bond length with Mulliken charge and

Table 5-4: Bond lengths, Mulliken spins, Mulliken charges, and diradical character from geometry optimizations of ***deoxy*** *and* ***carboxy*** *heme with different DFT functionals and bond lengths from EXAFS simulations.*

deoxy

functional	bond length / Å		Mulliken Spin on Fe	Mulliken Charge on Fe	Diradical character
	Fe-$N_{porphyrin}$	Fe-N_{His}			
BP86	2.09(2)	2.13	3.92	0.20	0.23
TPSS	2.09(2)	2.13	3.92	0.32	0.19
TPSSh	2.10(2)	2.16	3.94	0.40	0.14
B3LYP	2.11(2)	2.19	3.87	0.44	0.10
HB XAS	2.07(2)	2.17(2)	-	-	-

carboxy

functional	bond length / Å				Mulliken Charge	
	Fe-C_{CO}	Fe-O_{CO}	Fe-$N_{porphyrin}$	Fe-N_{His}	Fe	CO
BP86	1.73	2.90	2.01(2)	2.07	-0.43	-0.08
TPSS	1.75	2.91	2.02(2)	2.05	-0.35	-0.06
TPSSh	1.76	2.91	2.02(2)	2.07	-0.25	-0.02
B3LYP	1.79	2.93	2.03(2)	2.08	-0.13	0.01
HB XAS	1.78(1)	3.01(2)	2.01(1)	2.04(2)	-	-

Values in parenthesis represent maximum bond length ranges within a given model. The converged structure from geometry optimizations with B3LYP was used as an initial guess for the other optimizations. The diradical character in ***deoxy*** *originates from one UCO (see Figure 5-22). The wavefunctions for* ***carboxy*** *converged to a pure spin state (no net Mulliken spin and/or diradical character) All calculations were done with equal constraints and basis set (def2-TZVP(/J)). See Table 5-2 for myoglobin EXAFS parameters.*

Mulliken spin with diradical character) show strong correlation ($R^2 > 0.98$). The reason for the different behavior of these functionals is likely the varying HF exchange. B3LYP is in best agreement regarding the Fe – O_c bond length but the Fe – O_t bond distance may be underestimated by ~ 0.1 Å by all functionals, which points to a less tilted Fe – O – O angle or larger O – O bond length in the heme proteins. The spread of the Fe – N_{His} distance for each state is < 0.04 Å and the bond length is in good agreement with EXAFS for **deoxy** and **carboxy**. However, all functionals may underestimate the Fe – N_{His} in **oxy** by ~ 0.1 Å. This is consistent with QM/MM studies on a myoglobin XRD structure, that modeled parts of the protein environment and employed different DFT functionals [55]. These findings hint at difficulties of DFT to correctly model the electronic configuration of **oxy**. Nonetheless, the bond lengths from hybrid functionals are in good agreement with EXAFS with best performance of B3LYP.

Comparing core-to-valence (ctv) transitions from XAS and TD-DFT

Ctv transition energies and intensities from TD-DFT strongly depend on the degree of spin contamination and HF exchange [24, 148, 153] and TD-DFT often underestimates transition energies for excited states involving charge transfer [153]. Consequently, the differences in diradical character for the **oxy** heme state in particular should give rise to a different behavior

Table 5-5: Bond lengths, Mulliken spins, Mulliken charges, and diradical character (eq. (3-42)) from geometry optimizations of ***oxy*** *heme with different DFT functionals and EXAFS simulations.*

oxy									
functional	bond length / Å				Mulliken Spin		Mulliken Charge		Diradical character
	Fe-O_c	Fe-O_t	Fe-$N_{porphyrin}$	Fe-N_{His}	Fe	O_2	Fe	O_2	
BP86	1.81	2.73	2.02(3)	2.04	0.70	-0.68	-0.40	-0.40	0.61
TPSS	1.84	2.74	2.02(3)	2.03	0.91	-0.86	-0.28	-0.42	0.72
TPSSh	1.85	2.73	2.03(3)	2.03	1.07	-0.97	-0.18	-0.43	0.86
B3LYP	1.89	2.78	2.03(3)	2.03	1.11	-1.02	-0.08	-0.48	0.92
HB XAS	1.89(2)	2.91(3)	2.03(1)	2.15(2)	-	-	-	-	-

Values in parenthesis represent maximum bond length ranges within a given model. The converged structure from geometry optimizations with B3LYP was used as an initial guess for the other optimizations. The broken-symmetry keyword was not used in the input file. The diradical character originates from one UCO (see Figure 5-22). All calculations were done with equal constraints and basis set (def2-TZVP(/J)). See Table 5-2 for myoglobin EXAFS parameters.

of the DFT functionals. Figure 5-19 shows the experimental and TD-DFT ctv spectra and ctv difference spectra as well as difference densities from TD-DFT. Even though all functionals showed similar results in geometric structure, their predictions regarding ctv transitions show large differences in intensity and transition energies. In order to compare the theoretical transition energies, which depend on the DFT functional, a comparable energetic shift is necessary. The theoretical ctv transitions for **deoxy** heme, when aligned by peak maximum after broadening, only differ in intensity between functionals. The excitation energy of the transition into the iron $3d_{z2}$ orbital (difference density 5 in Figure 5-19C) is therefore used to align the broadened spectra of all functionals. Difference densities between excited state and ground-state reveal that each functional results in very similar orbital shapes for target valence orbitals (see Figure 5-19C). **Deoxy** excited states only consist of Fe(3d) orbitals. Core electrons are excited, in the case of **carboxy**, into iron 3d and CO π^* orbitals. The broken-symmetry of the wavefunction for **oxy** results in excitations into the iron 3d and O_2 π^* orbitals visible in the UCOs. All functionals show the same order of absolute ctv transition intensity as the background corrected XAS pre-edges with **oxy** < **deoxy** < **carboxy**. However, the in-/decrease in ctv transition areas of **oxy**/**carboxy** (shown as areas relative to **deoxy** ctv/pre-edge areas in Figure 5-20A) shows significant variations between functionals and states. The decrease in pre-edge/ctv area between **oxy** and **deoxy** is in good agreement with experiment. The experimental pre-edge area of **oxy** heme, taking into account the uncertainty in background subtraction, decreases to 60% relative to the **deoxy** heme pre-edge area, while the relative ctv areas range from ~ 40% (TPSS/B3LYP) to ~ 60% (BP86). The spread lies within the error of the experimental pre-edges. The absolute TD-DFT ctv area for **deoxy** heme is anti-proportional to the average Fe – N distance ($R^2 = 0.96$). For **oxy** heme, the ctv integral correlates with the Fe – O_c distance ($R^2 = 0.96$), the diradical character ($R^2 = 0.94$), the Mulliken charge ($R^2 = 0.97$) and spin ($R^2 = 0.94$) on iron. Consequently, the TD-DFT predictions for **deoxy** depend on the entire ligand field while the **oxy** ctv transitions primarily depend on the oxygen ligand. The ctv areas

*Table 5-6 Spin contamination in DFT wavefunction of the **oxy** state from different functionals.*

	BP86	TPSS	TPSSh	B3LYP
Spin contamination	0.62	0.75	0.92	0.96

The spin contamination originates from a broken spatial symmetry in the α and β part of one UCO (see column 'oxy' Figure 5-22)

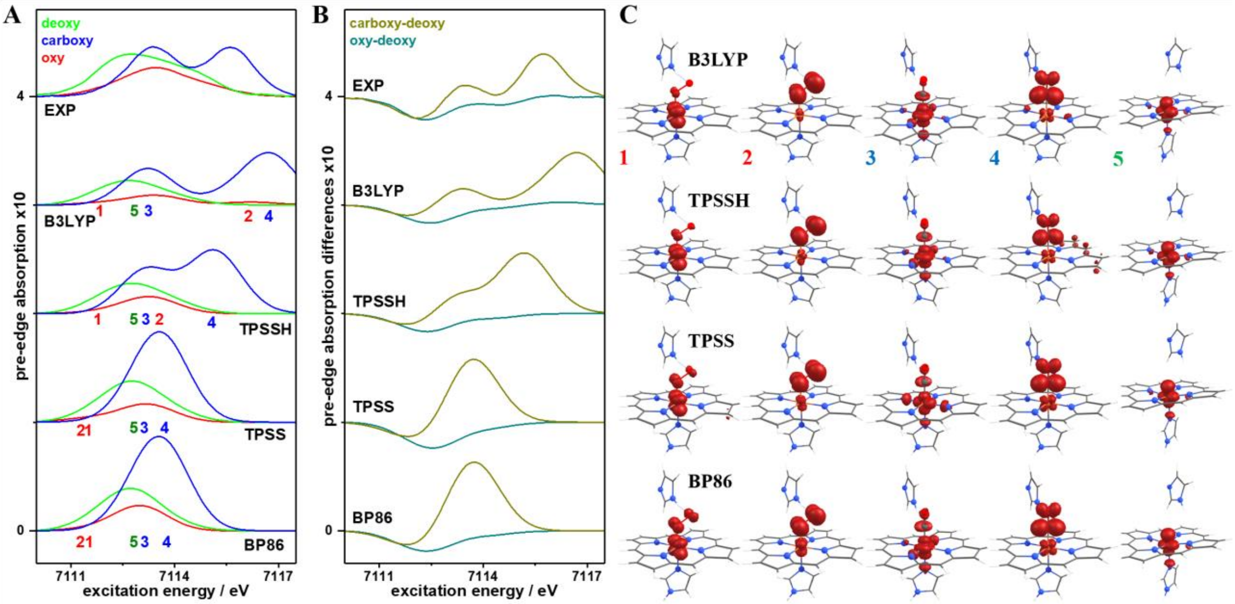


*Figure 5-19: Effect of DFT exchange-correlation functional on ctv transitions from TD-DFT. (**A**) Comparison of background corrected (smooth polynomial) experimental pre-edge spectra with ctv transitions from TD-DFT. Experimental spectra are averages of HB and MB samples at 20 K and 260 K. Selected TD-DFT transition energies are indicated and corresponding orbitals shown in (C) (red numbers: **oxy**, blue: **carboxy**, green: **deoxy**). Oscillator strengths from TD-DFT are broadened with Gaussians (σ = 0.75 eV) and scaled. TD-DFT transition energies are shifted such that the deoxy TD-DFT spectra align with the experiment and the deoxy Fe(d_{z2}) orbital (C 5) has the same energy for all functionals. The energy shifts are 181.4 eV (BP86), 169.4 eV (TPSS), 154.9 eV (TPSSh) and 151.3 eV (B3LYP). The spectra are vertically shifted for clarity. (**B**) Difference between **carboxy** and **deoxy** (dark yellow) and between **oxy** and **deoxy** (dark cyan) spectra from (A). (**C**) Difference densities between DFT ground-state and TD-DFT excited state indicating the unoccupied Fe(3d) or ligand valence MO that contains the resonantly excited Fe(1s) electron. The transition energies are indicated in (B) and isosurfaces are drawn at ± 0.005 au.*

for **carboxy** are overestimated by all functionals. The experimental pre-edge intensity increases from **deoxy** to **carboxy** by ~130%, while TD-DFT predictions range from TPSS (~170%) < BP86 (~200%) < TPSSh (~250%) < B3LYP (~270%). This is in line with previous studies indicating a tendency of TD-DFT to overestimate ctv oscillator strengths for metal-to-ligand charge transfer (MLCT) transitions involving strong π backbonding [35]. Accordingly, the major **carboxy** ctv transition intensity corresponds to transitions into the unoccupied CO π^* orbitals (see MO 4 in Figure 5-19C).

Ctv peak-maximum energies for **deoxy** are almost invariant ($\Delta E < 0.1$ eV) after energy alignment (MO 5 in Figure 5-19C). The relative ctv transition energies of **carboxy** and **oxy** are

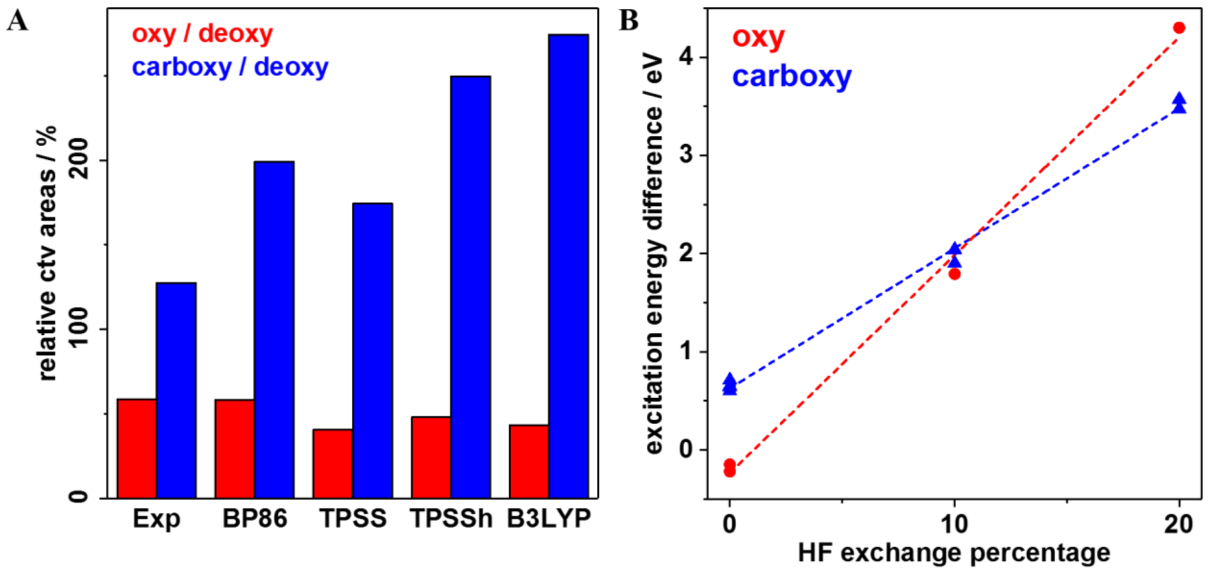


*Figure 5-20: Comparison of relative pre-edge/ctv areas (A) and energy shifts (B) from TD-DFT with different functionals. (**A**) Shown are oxy/carboxy pre-edge/ctv areas relative to the deoxy pre-edge/ctv. Intensities correspond to integrals of spectra or sum of ctv oscillator strength in Figure 5-19. (**B**) Dependence of excitation energy differences between core-to-valence transitions with and without charge-transfer (CT) on percentage of Hartree-Fock exchange in DFT functionals (BP86/TPSS = 0%; TPSSh = 10%, B3LYP = 20%). The excitation energy difference for carboxy is calculated between transitions into the Fe(3d$_{z2}$) (difference density 3 in Figure 5-19C) and the two CO(π*) orbitals (4 in Figure 5-19C, only one shown). Differences for oxy correspond to the Fe(3d$_{xz/yz}$)and O$_2$(π*) orbitals (1 and 2 in Figure 5-19C). Note that the O$_2$(π*) is lower in energy then Fe(3d$_{xz/yz}$) for BP86 and TPSS. The dashed lines are linear regressions, which for both oxy and carboxy yield R^2 = 0.996.*

invariant for metal-to-metal transitions (MO 1, 3, 5) or show systematic differences between DFT functionals for metal-to-ligand transitions (MO 2, 4). For **oxy**, the unoccupied iron 3d$_{xy}$ orbital (difference density 1 in Figure 5-19C) does not shift while the transition into the antibonding oxygen π* orbital upshifts for TPSSh and B3LYP. The **carboxy** ctv features, which show a similar behavior as **oxy**, are dominated by two transitions, the iron 3d$_{z2}$ and the antibonding CO π* orbitals (difference density 3 and 4). The 3d$_{z2}$ orbital does not change in energy while the π* orbital upshifts for TPSSh and B3LYP relative to BP86 and TPSS. Consequently, only the TD-DFT ctv spectra from the two hybrid functionals predict an energy splitting between metal-to-metal and metal-to-ligand transitions, which yields the double feature in the **carboxy** spectra in agreement with the experimental pre-edge of HB/MB.

The excited states involving MLCT from iron to the O_2 or CO ligand only shift for the two hybrid functionals. Absolute shifts cannot be compared as each functional results in different absolute energies. However, comparing energy differences between excited states of a given DFT functional is possible and shown in Figure 5-20B. The MLCT state shifts show a correlation with the fraction of HF exchange. The energy splitting of the resulting double feature, as can also be seen in the **carboxy** spectra, is under- (BP86, TPSS/h) or overestimated (B3LYP). Thus, the 'correct' fraction of HF exchange could be between 10 % and 15 %. Both TPSSh and B3LYP predict the same number of excitations. These transitions, after energy alignment, are either virtually identical in energy (for non-CT states) or show the above-mentioned shift (for MLCT states). The overestimated intensities in **carboxy** however do not improve with increasing HF exchange. The transition nature is therefore expected to be unchanged, the energy split to be tunable and the intensity to be overestimated. Pre-edges involving MLCT could be an interesting benchmark for future optimization of TD-DFT predictions.

Comparing valence-to-core (vtc) spectra from XES and DFT

The Kβ satellite line is a further experimental X-ray feature which can be simulated by DFT. It is sensitive to the ligand environment of the metal (see section 2.2). Figure 5-21 shows a comparison between experimental and theoretical spectra. All functionals predict the correct order of maximum intensities from **carboxy** > **oxy** > **deoxy**. This might be a mere coincidence as none of the simulated **deoxy** spectra shows the double feature visible in the experiment. Consequently, energy calibration was done aligning the **carboxy** peak maxima from DFT with experiment, which is indicated in Figure 5-21B. Comparing selected origin CMOs of vtc transitions shows no energy shifts comparable to those observed for ctv transitions from TD-DFT. Small differences in intensity and energy are visible for each of the transitions in Figure 5-21A, which stem presumably from the differences in geometry of each DFT functional. These CMOs also highlight the similarities in the different wavefunctions, which show very similar isosurfaces. The high-intensity vtc transitions with high iron contribution (red star, green square) show a large mixing with porphyrin orbitals, which dominate the satellite spectrum (see also Figure 5-24). All **oxy** wavefunctions have a double-occupied $O_2(\pi^*_z)$ orbital, which also highlights the interpretation of a Weiss-like configuration. The variations between functionals are more subtle and require a closer look at the difference spectra.

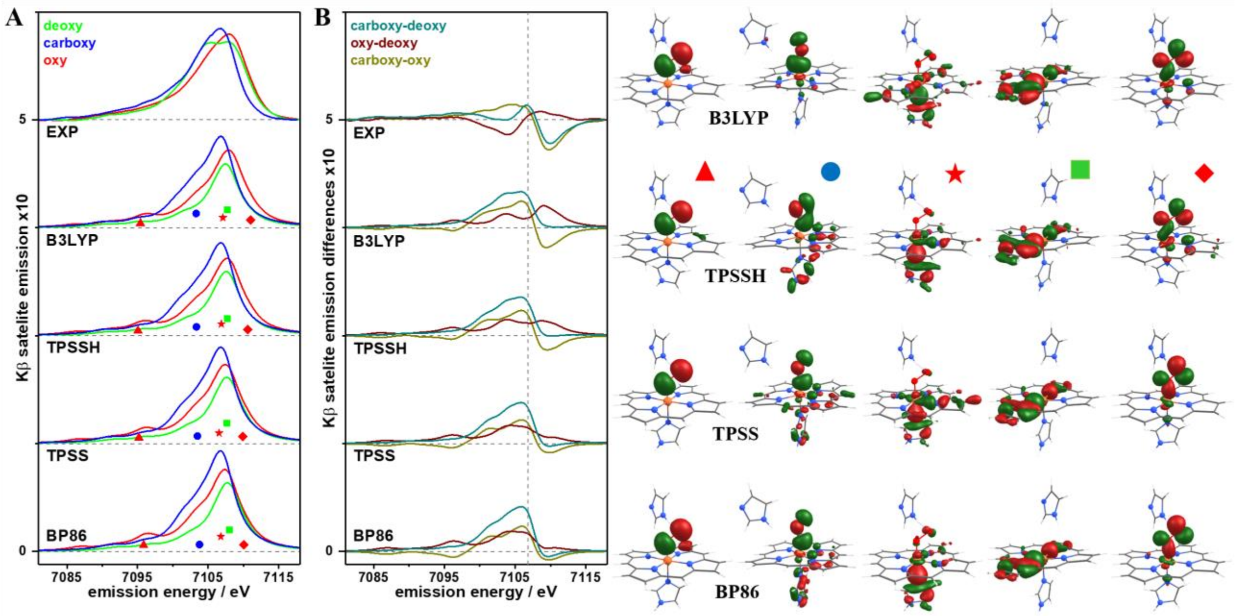


*Figure 5-21: Effect of DFT exchange-correlation functional on vtc transitions from DFT. (**A**) Comparison of background corrected (Lorentz function) and normalized (unity area) experimental Kβ satellite line spectra with vtc transitions from DFT. Experimental spectra are averages of HB and MB samples at 20 K and 260 K. Selected transition energies and intensities from DFT are indicated (color coded) and corresponding canonical MOs shown in (C). Oscillator strengths from DFT are scaled down to 2% and broadened with a Lorentzian shape (σ = 4 eV). DFT transition energies are shifted such that the carboxy vtc spectra maximum aligns with the experimental satellite line. The energy shifts are 181.4 eV (BP86), 169.7 eV (TPSS), 153.65 eV (TPSSh) and 148 eV (B3LYP). Note that theoretical ctv and vtc spectra have the same energy shift for BP86. Dashed horizontal lines indicate vertical shift. (**B**) Difference between carboxy and deoxy (dark cyan), oxy and deoxy (wine red) and carboxy and oxy (dark yellow) spectra from (A). The dashed vertical line indicates the carboxy peak maximum (7106.8 eV) used for energy alignment. (**C**) Selected canonical MOs with significant ligand contribution and relative high intensity. The transition energies and intensities are indicated in (B) and isosurfaces drawn at ±0.05 au.*

Focusing first on the region with emission energy lower than the **carboxy** maximum, one sees the effect of the missing double feature in the **deoxy** spectra. This leads to a different order of spectral intensity in this region, which is visible in the absolute and difference spectra. The relative intensity of **oxy** and **carboxy** features are in better agreement. The experimental **carboxy** spectrum is shifted to lower energies compared to the **oxy** spectrum and every functional predicts a shift in the same direction. However, the magnitude of the shift varies, which even reverses the order of maximum peak positions. B3LYP and TPSSh show a ranking

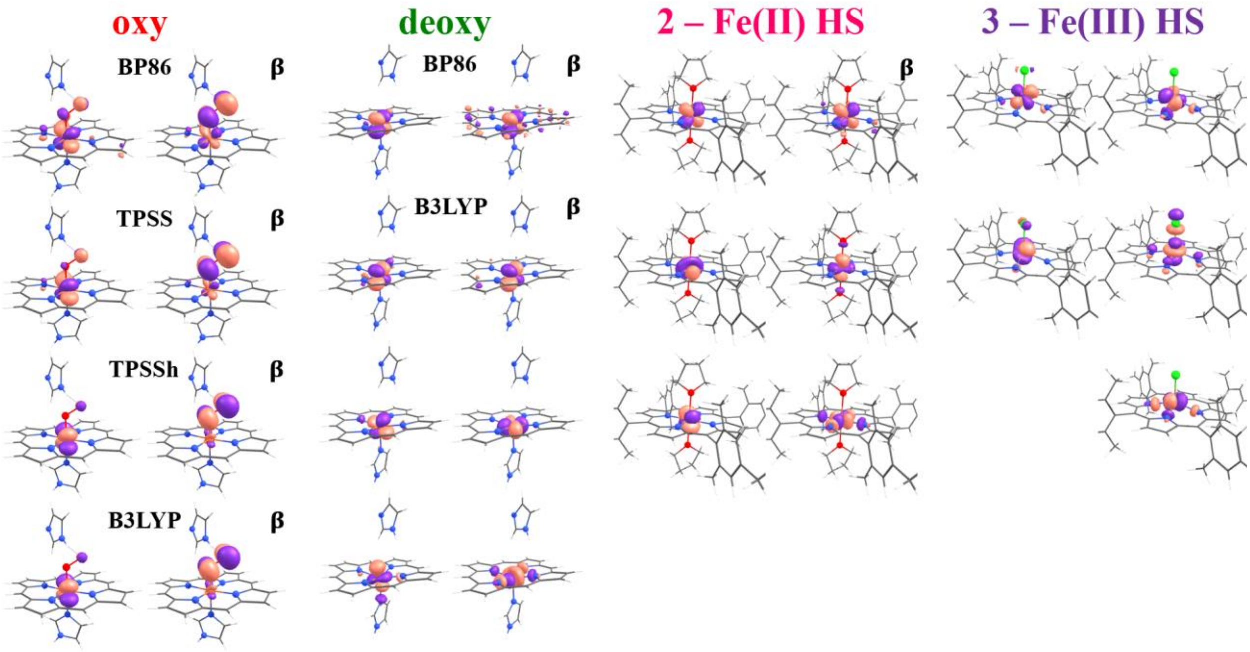


*Figure 5-22: UHF Corresponding Orbitals (UCO) with diradical character from DFT geometry optimizations. Different DFT functionals used in the case of oxy or deoxy heme are indicated. All remaining states shown were optimized using the B3LYP functional. Shown are UCOs with an overlap integral between α and β spin MOs unequal to unity. Corresponding β orbitals are omitted if the overlap integral is zero. There are no UCOs with diradical character (overlap integral unequal unity) in the case of **carboxy** and porphyrin **1** (Fe(II) LS). Isosurfaces are drawn at ± 0.05 au. See Table 5-5 – 7 and Table 5-7 for corresponding DFT parameters.*

of vtc peak-maximum energy of **carboxy** < **deoxy** < **oxy**, TPSS and BP86 interchange **deoxy** and **oxy**. The relative shift for **oxy** probably stems from the differences in diradical character between DFT functionals and corresponding charge separation. The resulting gradual increase in oxidation state would shift the entire spectrum towards higher energies. These small shifts are visible in e.g. the transitions in **oxy** from the N_{His}(2p) orbital (red star) or the O_2(π^*_z) orbital (red diamond). The best prediction in peak shift achieves B3LYP, although differences between functionals remain small. In conclusion, the consistency between satellite lines from experiment and DFT for **oxy** and **carboxy** indicates that one-electron transition approach from DFT appears to have difficulty to model vtc transitions for the high-spin iron porphyrin complex. However, the comparison between **oxy** and **carboxy** (slightly) favors again the hybrid over pure functionals and B3LYP over TPSSh.

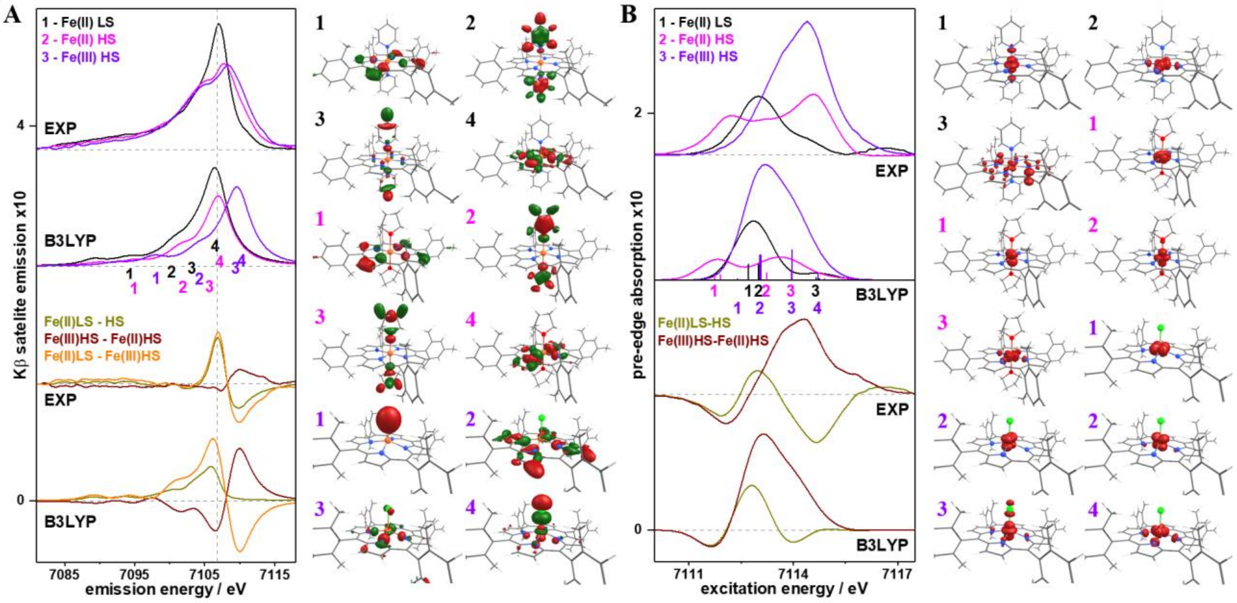


*Figure 5-23: Comparison of Kβ satellite lines (A) and pre-edges (B) from porphyrins **1** – **3**. Absolute experimental/theoretical spectra are shown on top and corresponding difference spectra below. (**A**) Mean experimental spectra (20/260 K) were background corrected (Lorentzian) and normalized to unity area. Theoretical vtc transitions from DFT are shifted (148 eV), broadened (σ = 4 eV, Lorentzian) and scaled. The vertical dashed line shows the* ***carboxy*** *peak maximum energy from Figure 5-17. (**B**) Background corrected experimental pre-edges (EXP) and theoretical (B3LYP) ctv spectra from TD-DFT. Ctv spectra were shifted (151.3 eV), broadened (σ = 0.5 eV, Gaussians) and scaled. Selected transition energies and intensities from (TD-)DFT are indicated. The indices are color coded and vertically shifted for clarity. Transitions that are very close in energy are annotated by one index but still stem from two excited states (see difference density 1 for Fe(II) HS). The corresponding canonical MOs for vtc and density differences for ctv transitions are shown to the right (isosurfaces drawn at ± 0.05 au (vtc) and ± 0.005 au (ctv)). Optimized geometries are shown in Figure 5-18. Dashed horizontal lines indicate vertical shifts of spectra.*

Ctv and vtc spectra for porphyrin model complexes

The previous section evaluated the performance of various functionals with regard to the HB/MB species and hybrid functionals showed better agreement with experiment. This section compares predictions from (TD-)DFT with the B3LYP functional of the three investigated porphyrin model complexes.

The geometries of the porphyrin model complexes were optimized with the B3LYP functional (see section 3.3 for details); first sphere iron bond lengths and Mulliken spin densities (MSDs) are shown in Table 5-7. MSDs of **carboxy**/porphyrin **1** and **deoxy**/**2** are in good agreement and close to the values from XES (Figure 5-11 in section 5.4), supporting the experimental results. The MSD on iron for porphyrin **3** is closer to four unpaired spins. The unpaired spin up UCOs (see Figure 5-22) for **deoxy** and porphyrin **2** correspond to four single-occupied MOs (SOMOs), which are predominantly centered on iron with small porphyrin contribution to the Fe(d_{x2-y2}) MO. Five SOMOs are present in the DFT wavefunction of porphyrin **3**, which are almost completely centered on iron with larger contributions from the axial Cl^{-} ligand (d_{z2}) and the porphyrin ring (d_{z2} and d_{x2-y2}), presumably due to the larger distortion of the square pyramidal coordination by the chloride versus the histidine ligand, apparently yields the MSD on iron of ~ 4.

Ctv spectra from TD-DFT and vtc spectra from DFT are shown in Figure 5-23. TD-DFT predicts two features for all three compounds and energy differences between features are in excellent agreement with XAS pre-edges. Ctv transitions stem from orbitals with large iron contributions, which resemble the single-occupied UCOs in Figure 5-22, except of the second feature for **1**, which shows contributions from the porphyrin ligand. The single feature for **1** is in line with the d^6 low-spin configuration. The large energy difference between the two peaks for **2** corresponds to a d^6 high-spin configuration with large crystal field splitting between t_{2g} and e_g levels because of the axial distortion of the octahedral symmetry. For porphyrin **3**, the left or right shoulder of the single feature correspond to transitions into the t_{2g} or e_g set of the d^5 high-spin configuration in a distorted square pyramidal symmetry. The shoulders from

Table 5-7: Bond lengths/angles, Mulliken spin densities and atomic charges in porphyrin model complexes from DFT (B3LYP/def2-TZVP)

Porphyrin compounds	bond length [Å]		Mulliken spin		Mulliken charges	
	Fe - N_{porph}	Fe - N, O, Cl	Fe	axial ligands[a]	Fe	axial ligands[a]
1 LS Fe(II)	2.01(2)	2.01(1)	0	0	-0.02	0.67
2 HS Fe(II)	2.08(2)	2.32(1)	3.84	0.05	0.52	0.07
3 HS Fe(III)	1.99(3)	2.29(1)	4.16	0.23	0.43	-0.50

*Values in parenthesis represent bond length ranges within a given model. Porphyrin compounds ((see Figure 5-18)): **1**, $Fe(TPP)(Py)_2$; **2**, $Fe(TPP)(THF)_2$; **3**, Fe(TPP)Cl; TPP = tetraphenylporphyrin, Py = pyridine, THF = tetrahydrofurane. [a]Axial ligands include $(Py)_2$ (**1**), $(THF)_2$ (**2**) and Cl (**3**). See Figure 5-22 for UCOs and Figure 5-23 for ctv/vtc spectra.*

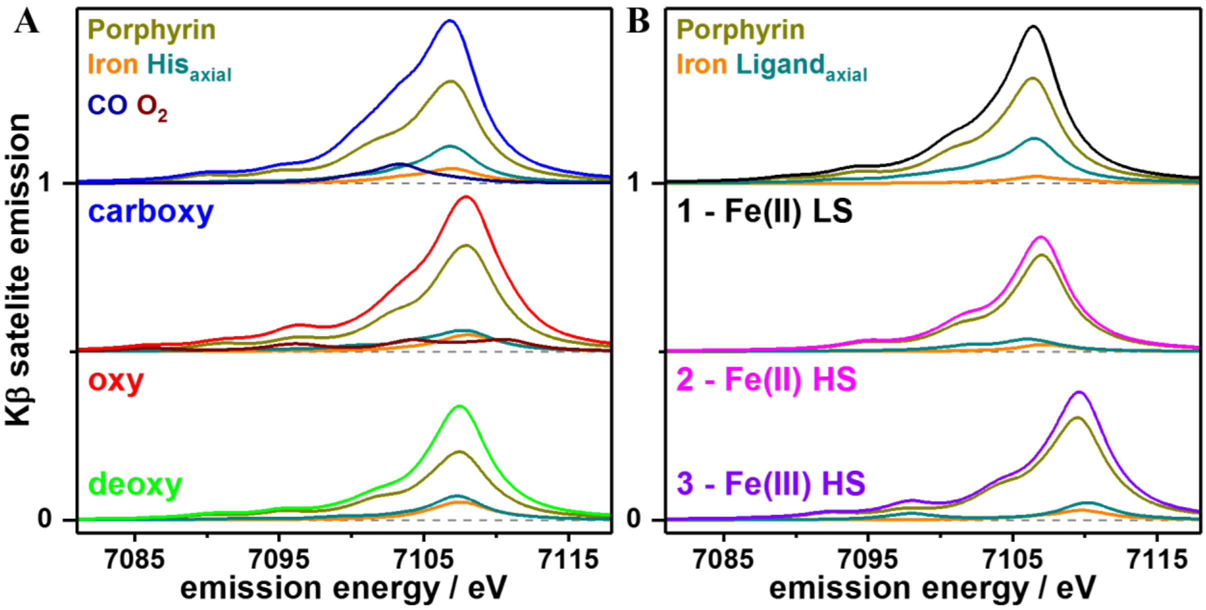


*Figure 5-24: Kβ satellite lines from DFT with single-MO transition approximation for heme proteins (A) and porphyrin model complexes (B). The contributions of different metal/ligand specific atomic orbitals to the vtc transitions are plotted individually. The relative atomic contributions from occupied valence orbitals correspond to the reduced orbital population per CMO with Mulliken weighting, which is extracted from the output files with in-house software. (**A**) The distal histidine shows no contribution in vtc spectra and is omitted in the plot. (**B**) The 'Porphyrin' assignment includes the phenyl species of the model compounds (see Figure 5-18). Axial ligands are the pyridine (**1**) and tetrahydrofuran ligands (**2**) and the chloride atom (**3**). Spectra are shifted by 148 eV and broadened with Lorentzians (σ = 4 eV). Geometries are shown in Figure 5-18. Spectra are vertically shifted (grey dashed lines).*

theoretical ctv of **3** show an opposite intensity-ratio compared to the experimental pre-edge, which could result from underestimation of the Cl^- contribution (respective orbital 3 in Figure 5-23B). Consequently, smaller metal-ligand orbital-mixing would reduce dipole oscillator strengths. Relative ctv intensities and shifts between porphyrins from TD-DFT agree well with experimental and only small deviations are revealed by the difference spectra. These results are in line with the interpretation of experimental pre-edges in section 5.6.

The vtc transitions shown in Figure 5-23A do not reproduce all experimental satellite line features but are generally consistent. Namely, (i) the vtc intensity decrease from low-spin to high-spin and (ii) the upshift in Kβ$_{2,5}$ feature energies in the order **1** → **2** → **3**. The energy upshift from porphyrin **1** to **2** is comparable (first-moment position difference: exp ~ 0.5 eV, DFT ~ 0.4 eV), yet the difference between **2** and **3** is significantly larger for DFT (~ 0.8 eV, ~

1.9 eV). The K$\beta_{2,5}$ double feature of the high-spin compounds **2** and **3** is not reproduced by DFT, which is also the case for **deoxy**.

The orbital overlap and energy difference between core and valence CMOs from DFT yield the vtc spectra shown in the previous figures (see section 3.1). The contribution of different metal/ligand species to the valence CMOs can be calculated from atomic orbital (AO) contributions using the Mulliken scheme. Contributions to the vtc spectra from heme and porphyrin compounds are shown in Figure 5-24. DFT vtc transitions originate predominantly from porphyrin. Axial ligands are responsible for subtle differences such as the rise of the low energy shoulder in **carboxy** by the CO ligand and the broader spectrum for **oxy** by two O_2 centered transitions. Emission from iron centered MOs yields almost no intensity. The results are in line with a previous report of various high-spin and low-spin mono-iron compounds, which found vtc spectra from DFT and experiment to agree well [27]. However, the two high-spin compounds in this report with pyrrole or pyridine ligands ($[Fe(III)(tpfp)Cl]^0$ with tpfp = tetra(pentafluoro)phenylporphyrin; $[Fe(II)(bipy)_3]^{2+}$ with bipy = 2,2'-bipyriine) showed similar double features as **2**, **3** and **deoxy** that were not predicted by DFT. The calculations were done using the BP86 functional, the CP(PPP) basis set [217] for iron and TZVP for the remaining atoms. Thus, the double feature presumably is independent of coordination number (5/6), symmetry (symmetric/distorted-octahedron, square-pyramidal) and axial ligand (pyrrole, chloride, oxygen and pyridine). A variety of DFT functionals (pure, meta and hybrid) and different basis sets for iron (TZVP, CP(PPP)) were incapable to model this double feature. It remains unclear whether this is due to (i) an underestimated orbital overlap and therefore underestimated oscillator strength, (ii) a two-electron process, which the one-step model cannot simulate or (iii) significant electronic relaxation present in the absorption/emission process, which is not included in the DFT vtc model.

Variation of the Fe – O_2 bond length in oxy

The analysis of Kβ' emission in section 5.4 demonstrated zero unpaired Fe(3d) electrons in porphyrin compound **1** and **carboxy**, four in **2** and **deoxy** and five in porphyrin compound **3** (see Figure 5-11). The main-line emission further showed two unpaired Fe(3d) electrons for **oxy**. These results are in good agreement with Mulliken spin densities on the iron atom from DFT with TPSSh and B3LYP functionals for all cases except for **oxy** (see Table 5-4, Table 5-5 and Table 5-7). The broken-symmetry wavefunction for **oxy** showed large diradical character between an Fe(3d) and an antibonding dioxygen π* MO oriented perpendicular to the Fe – O – O plane ($O_2(\pi^*_\perp)$), which results in ~ 1 unpaired spin on iron. BP86 predicted an even smaller diradical character with a Mulliken spin density of 0.7 on iron. This inconsistency between DFT and XES might be due to DFT converging to a particular electronic configuration out of several energetically similar solutions with different configurations. To that end, difficulties of calculating correct spin state order have been encountered in the past for several iron porphyrin complexes, where small variations in ligand bond lengths could alter the order of states [218]. Presumably, similar small variations in the Fe – O_2 bond length might result in different electronic configurations within a singlet spin state. For this reason, this section investigates the effect of (i) the Fe – O_2 distance and (ii) the initial-guess wavefunction on the electronic configuration of the converged DFT wavefunction.

The **oxy** state was optimized with constraint Fe – $O_{central}$ bond lengths between 1.7 – 2.2 Å using an automatic, a triplet or a quintet initial-guess wavefunction and the results are shown in Figure 5-25. Varying the Fe – O bond length and initial-guess wavefunction yields two singlet states. The potential energy surface (PES) with higher total energy (black line) corresponds to a Pauling-like configuration without broken-symmetry and results from the automatic initial-guess. While triplet or quintet SCF wavefunctions as initial-guesses converge to the same lower energy broken-symmetry state (red line), which has afc between iron and oxygen involving two electrons and consequently represents a Weiss-like solution. The difference between the energy minima of the Pauling and Weiss state (~ 0.8 eV) is in the order of a hydrogen bond (< 1.5 eV) [219] and the PES minimum for Weiss (1.9 Å) is in better agreement with EXAFS than Pauling (1.75 Å). Notably, the Pauling PES minimum without constraint Fe – O distance results in a Weiss-like state with a bond length of ~ 1.9 Å. Ctv spectra from TD-DFT predict a single peak for the Pauling state from excitations into Fe(3d$_{z2}$), Fe(3d$_{x2-y2}$) and an antibonding MO formed from an Fe(3d$_{xz/yz}$) MO and an $O_2(\pi^*_\perp)$ MO (see Figure 5-25B/C). The three ctv features for the Weiss state are in better agreement with experimental pre-edges (see Figure 5-17), where the single Fe(3d) – $O_2(\pi^*_\perp)$ transition splits into a lowest-energy iron-centered and a highest-energy O_2 centered transition. No canonical intermediate-spin configuration with two unpaired spins on the iron atom and the oxygen molecule was produced.

The Mulliken spin density (MSD) on the iron atom and oxygen molecule for the Weiss-like DFT wavefunction increases with increasing Fe – O distance (Figure 5-25A bottom), which corresponds to an increase in charge separation between Fe and O_2. While the MSD at the PES energy minimum (Fe: 1.1, O_2: 1.0) does not agree with two unpaired Fe(3d) electrons from XES, it still suggests that more than one Fe(3d) MO is contributing to the spin density. Additionally, the MSD decrease/increase towards smaller/larger distances (note the saddle

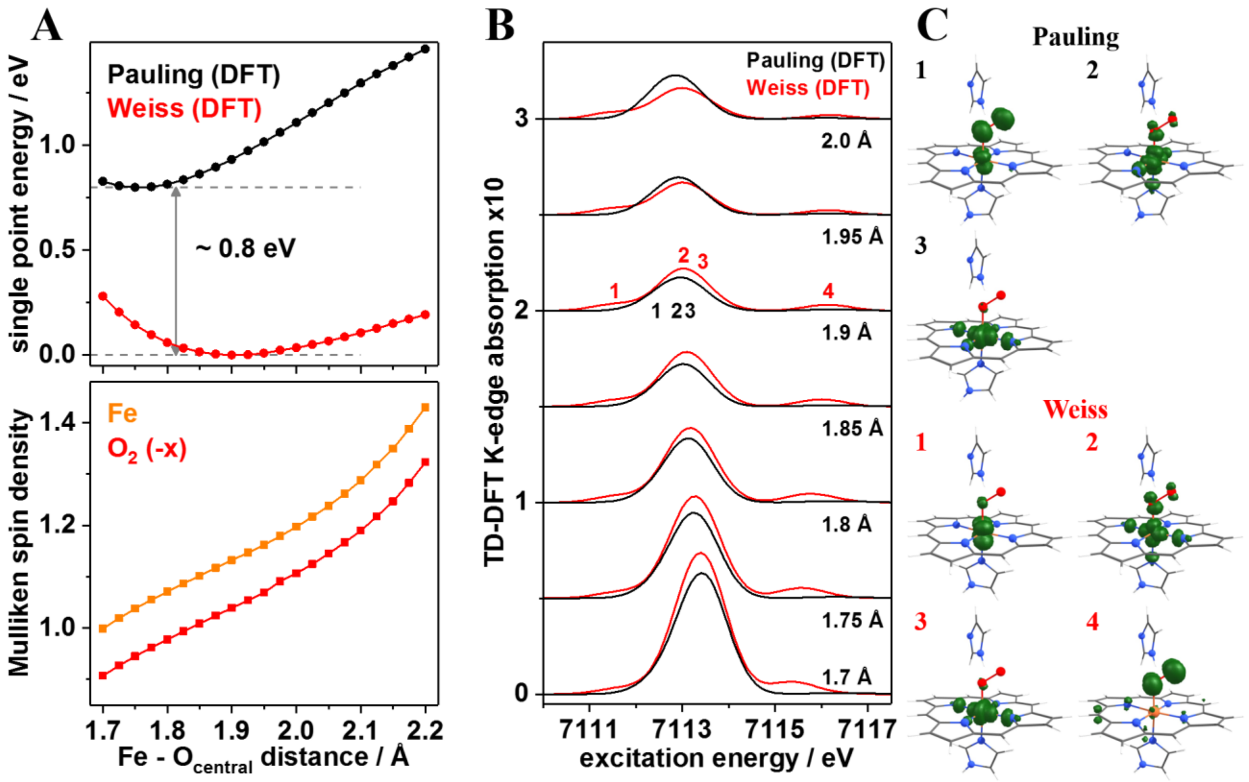


Figure 5-25: DFT energies and electron densities for ***oxy*** *from potential energy surface scans of the Fe – $O_{central}$ bond length. Data refer to geometry-optimized structures from DFT (B3LYP, TZVP) with restraint Fe – $O_{central}$ bond length. (**A**) Total energy (top) and spin density from Mulliken AO distribution (bottom) versus Fe – O_2 bond length for two singlet DFT wavefunction configurations, one with diradical character (with afc; Weiss) and one without (no afc, Pauling). Both wavefunctions used unrestricted Kohn-Sham MOs. Total energies are shifted so that the minimum for Weiss (1.9 Å) is zero. The energy difference between the two minima (Pauling: 1.75 Å) is indicated. Spin density on O_2 is inverted. UHF corresponding orbitals (UCOs) of the Weiss wavefunction are shown in Figure 5-26. Note, that the spin density for the Pauling wavefunction is zero. (**B**) Ctv transitions from TD-DFT for geometries in (A). Intensities are broadened (Gaussian shape, σ = 0.5 eV) and transition energies shifted (151 eV). Target MOs of indicated transitions are shown on the right. (**C**) Electron density differences between DFT ground-state and TD-DFT excited states (green), which correspond to target ctv MOs. Isosurfaces are drawn at ± 0.005 au.*

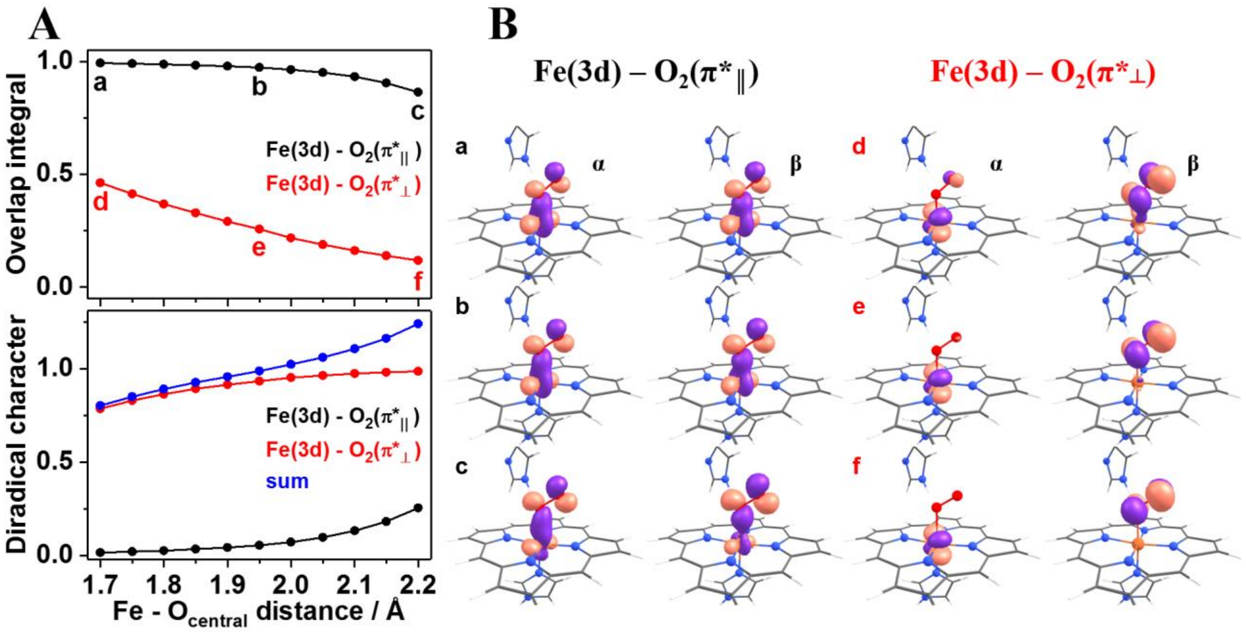


Figure 5-26: Influence of Fe – $O_{central}$ distance on broken-symmetry ***oxy*** *wavefunction. Energy and spin density are shown in Figure 5-25. (**A**) Overlap integral (top) and diradical character (bottom) of two UCOs corresponding to interaction of bonding Fe(3d) MOs with antibonding $O_2(\pi^*)$ MOs perpendicular/parallel (⊥/‖) to the Fe – O – O plane. These UCOs yield the largest broken spatial symmetry in the wavefunction. UCOs at indicated points are shown on the right. (**B**) UCOs with small (a – c) and large (d – f) diradical character, which corresponds to small/large spatial asymmetry between spin up (α) and spin down (β) MOs. Isosurfaces are drawn at ± 0.06 au. See section 3.4 for more details on UCOs.*

point around 1.9 Å) indicates at least two MOs with broken symmetry. The two UCOs, which contribute the most to the diradical character are shown in Figure 5-26. The largest spatial separation between spin up and spin down MOs stems from the previously mentioned $FeO_2(\pi^*_\perp)$ UCO. The second UCO with smaller diradical character corresponds to an Fe(3d) MO bonding with an antibonding dioxygen π* MO parallel to the Fe – O – O plane ($FeO_2(\pi^*_{||})$). The diradical character from $FeO_2(\pi^*_\perp)$ converges to unity towards larger bond lengths while it starts from zero and increases for the $FeO_2(\pi^*_{||})$ MO. MSD and the sum of diradical character of these two UCOs correlate perfectly ($R^2 = 0.997$). Extrapolation from this behavior suggests a second spatially split UCO for large bond lengths, which results in the electronic configuration suggested by McClure. Thus, the DFT wavefunction does include broken-symmetry involving two Fe(3d) and two O_2 π* MOs but not to the extent that is necessary to explain the Kβ’ feature from section 5.4.

Conclusion

Geometry optimizations of heme species, with four DFT functionals, and porphyrin model complexes, with B3LYP, were conducted and extensive systematic studies on the influence of the theory level on the cofactor structure as well as on core-to-valence (ctv) transitions from TD-DFT and valence-to-core (vtc) spectra from DFT were performed. Additionally, the potential energy surface as a function of the Fe – O_2 bond distance (1.7 – 2.2 Å, B3LYP) in **oxy** was calculated. For the heme systems, the interatomic distances (e.g. Fe – C/$O_{central}$) from B3LYP and TPSSh (hybrid functionals) were in better agreement with the experimental data than results from BP86 (generalized gradient approximation (GGA) functional) and TPSS (meta-GGA functional). Mulliken spin densities on the iron atom from non-oxy species reproduced formal unpaired Fe(3d) spin assignments from XES.

Ctv transitions. Prediction of pre-edge absorption features by TD-DFT was strongly affected by the functional and specifically the **carboxy** double peaks appeared to be a good indicator. The two hybrid functionals (TPSSh, B3LYP) showed better agreement with the experimental pre-edges for **carboxy** than the two functionals without HF exchange (BP86, TPSS). The energy separation of the two separated pre-edge absorption peaks of **carboxy** was best reproduced using TPSSh, but underestimated and therefore predicted as a peak with shoulder feature. For B3LYP, the energy separation is overestimated but, as a consequence, two clearly separated absorption peaks, in agreement with pre-edge absorption, were calculated. The ctv transitions for the porphyrin model complexes from B3LYP also agreed with experiment. Interestingly, the second pre-edge peak in **carboxy** stems from a metal-to-ligand transition involving charge transfer, as supported by our spin-polarized detected pre-edge spectra, but shifts significantly with varying Hartree-Fock exchange in the DFT functional. These findings highlight the importance of comparing different functionals when simulating pre-edges with TD-DFT.

Vtc transitions. The vtc transitions from DFT were rather independent of the functional and mostly consistent with the experimental data. The vtc double feature of **deoxy**, however, was not well described. Energy shifts in the vtc between **carboxy** and **oxy** were consistent in the DFT data. The double feature was also absent in calculated vtc spectra of the high-spin porphyrin complexes **2** and **3**, but energy shifts between heme species and porphyrin model complexes were consistent. Hybrid DFT functionals therefore seem to be recomendable to calculate bond-distances and ctv X-ray spectra and B3LYP appeared to be the superior choice, at least for heme systems. Double features in Kβ satellite line emission spectra for complexes with high-spin iron bound to porphyrin (our work) or pyridine (literature data) were not predicted by any DFT approach, presumably due to model limitations.

Potential energy surface. Irrespective of bond distance, **oxy** converged to either of two singlet states with a closed-shell configuration (Pauling) or an open-shell configuration with broken-symmetry between Fe and O_2 in one molecular orbital (Weiss). This resulted from varying the initial guess for self-consistent field convergence. The Weiss-like DFT solution (Fe(III) – $O_2^{\bullet}$) yields a significantly lower energy of the system and is in considerably better agreement with bond-distances from EXAFS and pre-edges from XAS. However, local spin analysis of the DFT wavefunction of **oxy** in the Weiss-like state yields a Mulliken spin density of 1 on iron, which is not in agreement with our Kβ XES data. As a result, bond-distances and core-to-valence transitions from B3LYP showed good agreement between experiment and theory for **oxy** and non-oxy species. Analysis of the DFT wavefunctions seems to be, in the case of **oxy**, insufficient to describe the local electronic configuration on iron. The likely reason presumably lies in the limitation of a single-determinantal wavefunction. To that end, the next section presents multiconfigurational quantum chemical calculations for **oxy** and non-oxy species in order to overcome these limitations.

5.8 Electronic configuration from CASSCF

DFT calculations in the previous section were in good agreement with bond lengths from EXAFS, XAS pre-edge transitions and, to some extent, XES satellite lines. Spin densities on iron for the non-oxy species showed good correlation with relative Kβ’ intensities. However, local spin analysis of the **oxy** DFT wavefunction was not in agreement with XES main-line experiments. Notably, the DFT wavefunction for **oxy** is the only state with large diradical character from broken symmetry in the wavefunction, which involves the Fe – O_2 bond. Possibly because the electronic configuration of this bond is of multiconfigurational nature, which has been argued in the past [55, 76, 77]. While broken-symmetry is an indicator against a single-configurational description, it is by no means proof and the discussion about the electronic configuration far from settled. Multiconfigurational quantum chemical calculations in recent years have primarily focused on the post-HF complete active space self-consistent field (CASSCF) method, occasionally in combination with second-order perturbation theory (CASPT2). Some of these approaches have reproduced Mössbauer, UV/VIS or calorimetric parameters (see section 1). In the present section, results from CASSCF calculations with large active spaces are presented and corresponding effective unpaired electron densities are compared to main-line XES from section 5.4. This method is first applied to **deoxy**, **carboxy** and the three porphyrin model complexes, to verify the applicability of the method, and then **oxy**. It should be noted, that no method is yet available to calculate K-edge ctv/vtc spectra from a CASSCF wavefunction.

Non-oxy heme species and porphyrin model complexes

CASSCF calculations for **deoxy** and **carboxy** were carried out on DFT optimized structures (B3LYP) from the previous section. For the three porphyrin models, phenyl ligands were replaced by hydrogen atoms, and hydrogen atom positions optimized with the BP86 functional. This resulted in a considerably smaller atom numbers and computational demands without affecting the iron environment. Natural molecular orbitals (NMOs) from MP2 single point calculations were localized and selected as described in section 3.3. Active spaces were chosen to include orbitals that correspond to the iron atom or axial ligand(s). Resulting active spaces are shown in Figure 5-27 and corresponding configurations of CASSCF states and occupation numbers (ON) are shown in Table 5-8. Stable SCF convergence without porphyrin intruder states was possible for all calculations despite typically > 150 occupied or virtual $\mathrm{MP2_{NMOs}}$ with large static correlation effects located on the porphyrin ring.

CASSCF for **deoxy** has the smallest active space (15 possible configurations) and integer occupation numbers indicate that the nitrogen atom of the axial imidazole shows no correlation effects with the Fe(3d) orbitals in the active space. Accordingly, the CAS wavefunction consists of a single configuration with two double- (1 and 2 in Figure 5-27) and four single-occupied orbitals (3 – 6). The orbitals 1 and 2 correspond to a N(p) and an Fe(d_{xz}) orbital and 3 – 6 to remaining Fe(3d) orbitals (3: d_{yz}, 4: d_{xy}, 5: d_{z2}, 6: d_{x2-y2}), which is the same orbital configuration as from DFT (see Figure 5-22).

*Table 5-8: Summary for CASSCF calculations for **deoxy**, **carboxy** and porphyrin **1** – **3**.*

complex	heme		porphyrin		
	deoxy	**carboxy**	**1** Fe(II) LS	**2** Fe(II) HS	**3** Fe(III) HS
	MO occupancy (state contribution to wavefunction in %)				
CASSCF states	221111 (100)	22222200 (95.6) 22221111 (1.9) 22222020 (1.2) 22220202 (1.2)	222222222 000000000 (99.7)	2222211110000 (95.4) 2221111110011 (0.6) 2112211111100 (0.6)	22211111 (100)
MO no.	spin density/occupation number				
1	2.000	2.000	2.000	2.000	2.000
2	2.000	2.000	2.000	1.977	2.000
3	1.000	2.000	2.000	1.977	2.000
4	1.000	1.955	2.000	1.975	1.000
5	1.000	1.955	2.000	1.975	1.000
6	1.000	0.045	2.000	1.000	1.000
7		0.045	2.000	1.000	1.000
8			2.000	1.000	1.000
9			2.000	1.000	
10			0.000	0.025	
11			0.000	0.025	
12			0.000	0.024	
13			0.000	0.024	

CAS state configurations, with relative contributions in brackets, refer to double-, single- or unoccupied natural molecular orbitals (NMOs) in CAS active space. NMOs in the active spaces are shown in Figure 5-27. Spin densities/occupation numbers correspond to the weighted average of all CAS state configurations in a CASSCF wavefunction (see section 3.2) and represent probabilities to find an electron in a given NMO.

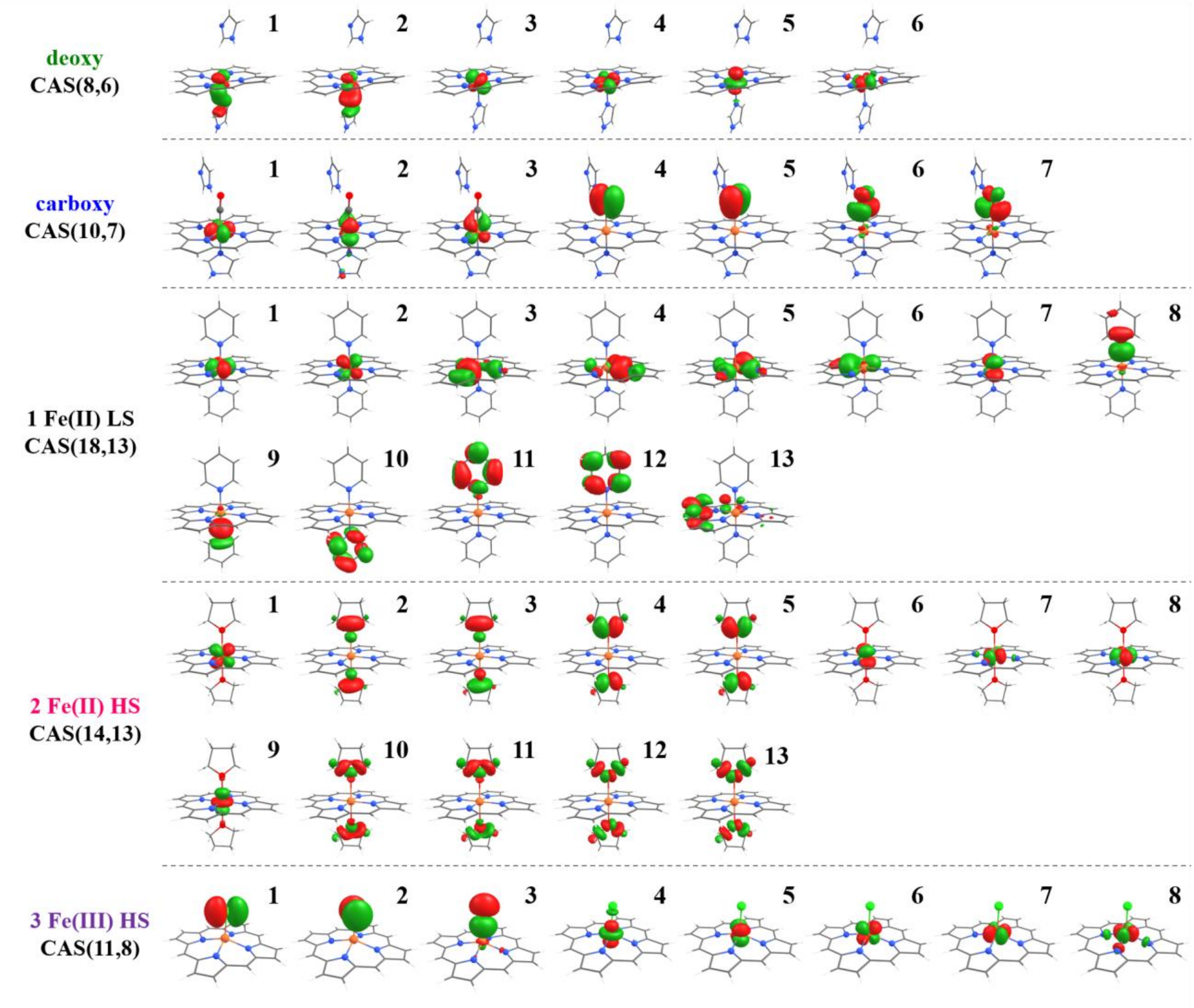


Figure 5-27: Natural molecular orbitals (NMOs) in active space from CASSCF calculations. NMOs correspond to CASSCF calculations on geometry-optimized structures (see Figure 5-18 and previous section) with the following parameters; compound/(electrons, MOs): ***deoxy****/(8,6),* ***carboxy****/(10,7), porphyrin* ***1****/(18,13), porphyrin* ***2****/(14,13), porphyrin* ***3****/(11,8). NMO isosurfaces are drawn at ± 0.05 au. NMOs are ordered by their occupation number/spin density. See Table 5-8 for CASSCF parameters. The phenyl ligands from the porphyrin compounds were replaced by hydrogen atoms for CASSCF calculations. The CAS(11,8) calculation for porphyrin* ***3*** *used a CAS(5,5) calculation (NMOs: 4 – 8) as input.*

High-spin porphyrin model compound **3** shows a similar behavior as **deoxy** with a slightly larger active space (≈ 50 configurations), which includes three double-occupied Cl(3p) orbitals and five single-occupied Fe(3d) orbitals.

For low-spin porphyrin model compound **1**, while the active space is large with > 50 k possible configurations, the result is an almost single-configurational wavefunction with double-occupied iron d_{xy}, d_{xz} and d_{yz} orbitals (the t_{2g} set in an octahedral ligand field).

In **carboxy**, three double-occupied Fe(3d) orbitals in CAS wavefunction resemble the configuration in **1**, though with a smaller active space (≈ 150 configurations). However, orbitals centered on the CO ligand do show some multiconfigurational character with the dominant configuration (> 95 %) in the wavefunction composed of double-occupied iron t_{2g}, CO(π_x) and CO(π_y) orbitals (NMO 1 – 5 in Figure 5-27). Three additional state configurations are present. These additional configurations represent the first departure from single-configurational MO theory. The Hamiltonian of a system in CASSCF not only describes the dominant configuration but also additional configurations, which are expressed as single/double excited configurations with an occupation decrease by one/two in an occupied orbital (in the dominant configuration) and corresponding increase in an unoccupied orbital, for more details see section 3.2. In **carboxy**, the second configuration contains two single excitations and the third/fourth configurations contain a double excitation from bonding CO($\pi_{x/y}$) into antibonding CO($\pi^*_{x/y}$) orbitals (4 → 7, 5 → 6). The four CAS state configurations compose almost the entire **carboxy** wavefunction (~ 99.9 %). While natural MOs centered on iron have integer occupation numbers, CO natural MOs yield fractional occupations with two electrons shared by a pair of

Table 5-9: Mulliken atomic charges/spin densities and effective unpaired electron densities (EUED) from CASSCF wavefunctions of heme and porphyrin model compounds.

Species	Mulliken atomic charges		Mulliken spin densities		Hard shell EUED[b]	
	Fe	Axial ligands[a]	Fe	Axial ligands[a]	Fe	Axial ligands[a]
deoxy	0.93	-	3.92	-	4	-
carboxy	0.67	-0.07	0	0	0	0.35
oxy	0.47	-0.55	0	0	1.99	0.32
1 Fe(II) LS	0.68	0.42	0	0	0	0
2 Fe(II) HS	0.98	0.03	3.86	0.03	4	0.37
3 Fe(III) HS	0.94	-0.62	4.77	0.05	5	0

[a] *Axial ligands include CO (**carboxy**), O_2 (**oxy**), N_{Py} (**1**), O_{THF} (**2**), or Cl (**3**) ligands (see Figure 5-18). **Oxy** refers to a CAS(14,11) calculation with an Fe – O_2 bond length of 1.90 Å. See Table 5-8/Figure 5-27 (non-oxy) and Table 5-10/Figure 5-28 (**oxy**) for CAS wavefunction configurations and natural molecular orbitals (NMOs).* [b] *EUED from occupation number in NMO is assigned to atom with largest atomic orbital contribution (hard shell weight). Individual assignment for **oxy** is shown Figure 5-28.*

related (anti-)bonding π orbitals. Therefore, bonding/antibonding CO(π) orbitals form correlation pairs due to static electron correlation, with no effect on the spin density on iron.

This motive of static electron correlation in molecules involving oxygen atoms also occurs in porphyrin model compound **2**. Occupation numbers in the CAS wavefunction of **2** show a double-occupied Fe($d_{xz/yz}$) orbital (orbital 1) and four unpaired spins in the remaining four Fe(3d) MOs (6 – 9), which is in agreement with DFT. As in **deoxy** and **3**, integer occupation numbers show no correlation effects between Fe(d) and other orbitals in the active space. The axial THF ligands however show static correlation between (anti-)bonding σ orbitals (orbitals 2 – 5, 10 – 13). This, similar to **carboxy**, results from the superposition of a mayor and several minor state configurations. The dominant state configuration (> 95 %) includes double-occupied Fe($d_{xz/yz}$)/O^{THF}(σ) orbitals (MOs 1 – 5), single-occupied Fe(3d) orbitals (MOs 6 – 9) and empty O^{THF}(σ*). The second and third configuration correspond to two single excitations from O^{THF}(σ) → O^{THF}(σ*) orbitals (2/3 → 10/11 and 4/5 → 12/13). The remaining ~ 3 % stem from several configurations with small weights (< 0.25 %), presumably from double excitations or four single excitations from σ → σ* orbitals.

Mulliken population analysis (see Table 5-9) of CAS wavefunctions yields comparable atomic charges on iron for the same low-spin (**carboxy**, **1**) and high-spin (**deoxy**, **3**) compounds and in the same order (**carboxy**/**1**, **deoxy**/**3**, **2**) as DFT (see Table 5-4, Table 5-5, Table 5-7). Mulliken spin densities from CAS versus DFT for **deoxy**, **2** and **3** are closer to results from XES, with **3** improving significantly. This is presumably due to smaller orbital mixing from Fe(d_{z2}) and Cl($3p_z$) orbitals in **3** (see Figure 5-22).

CASSCF calculations can result in multiconfigurational wavefunctions, which yield non-integer natural MO occupation numbers. These fractional spin densities allow for the analysis of e.g. open-shell versus closed-shell singlet configurations. One example is the effective unpaired electron density (EUED) by Takatsuka (see equation (3-39) in section 3.2). The EUED of an orbital can be assigned to the atom with the largest atomic orbital (AO) contribution (hard shell weight). Consequently, the EUED of the single-occupied NMOs in porphyrin **3** are entirely assigned to iron. The CAS wavefunctions of **1** and **carboxy** yield an EUED of zero, in line with a closed-shell iron. The high-spin species result in four unpaired electrons on the iron atom for **deoxy** (MO 3 – 6) and porphyrin **2** (MO 6 – 9) and five for porphyrin **3** (MO 4 – 8). Note that while the total EUED can be non-zero (**carboxy**, **2**), the hard shell assignment separates electron correlations on ligands that do not include iron.

Electronic configuration in oxy

CASSCF calculations combined with the Takatsukas effective unpaired electron density (EUED) showed perfect agreement with results from Kβ main-line XES in the previous section. The present section analyses the electronic configuration in the CAS wavefunction of the **oxy** heme system. For this purpose, first the influence of varying Fe – O_2 bond length on CASSCF calculations is presented and evaluated with regard to the experimental results from previous sections. Second, the emerging electronic configuration in **oxy** from CASSCF is compared to the proposed models from Pauling, McClure/Goddard and Weiss.

Effect of the Fe – O_2 bond length on the CAS wavefunction

Several combinations of input orbitals, active space configurations and other computational parameters were varied and tested versus a range of **oxy** structures with varying Fe – O_2 bond lengths. The **oxy** structures stem from the Weiss-like DFT state in the potential energy surface (PES) scan in section Figure 5-25 and Fe – O_2 distances were chosen to cover the energy minima of the Pauling- and Weiss-like DFT states. Two active spaces were found to converge for all Fe – O_2 bond lengths, a calculation with 8 electrons distributed in 7 orbitals (CAS(8,7)) and a CAS(14,11) calculation, which used the wavefunction from CAS(8,7) as initial guess. A complete summary of the resulting active spaces and parameters for all bond lengths is shown in the appendix in Figure 9-4/Table 9-1 (CAS(8,7)) and Figure 9-5/Table 9-2 (CAS(14,11)). The two step process that produces the converged active spaces for both CAS calculations can be seen in Figure 5-28.

For all Fe – O_2 bond lengths, both CAS(8,7) and CAS(14,11) wavefunctions are dominated (~ 90 – 95 %) by two CAS state configurations, as was observed in earlier CASSCF calculations [55, 76, 220]. The remaining ~ 5 – 10 % of the CAS wavefunction result from configurations with small contributions (< 2 %). Although differences between both CAS wavefunctions are subtle, they reveal intriguing relations to results in previous sections.

The relative total energy, shown in Figure 5-29A, for the CAS(8,7) calculations has a similar dependence on the Fe – O_2 distance as the broken-symmetry Weiss configuration from DFT with an energy minimum at ~ 1.875/1.9 Å. Increasing the active space with the CAS(14,11) calculations shifts the energy minimum towards shorter bond lengths (< 1.85 Å) and closer to the energy minimum for the closed-shell Pauling configuration. While the total energy in DFT increased by ~ 0.8 eV from the Weiss-like to Pauling-like configuration, arguing against the closed-shell wavefunction, it decreases by ~ 0.4 eV from CAS(8,7) to CAS(14,11) in favor of the larger active space. The shift in energy minimum towards smaller bond-lengths is the result

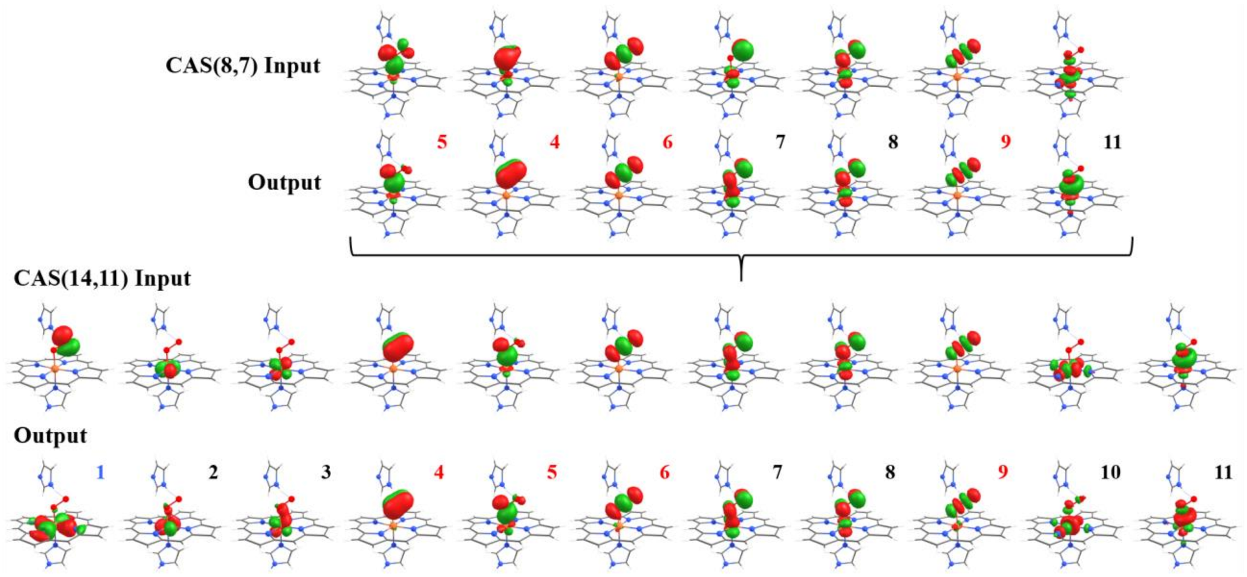


Figure 5-28: Development of the MO configurations in a two-step CASSCF procedure (Fe – O_2 bond length: 1.90 Å). Input localized molecular orbitals (LMOs) were generated as descript in section 3.3. Isosurfaces are drawn at ± 0.05 au. Geometric structure was optimized with B3LYP (see section 5.7). The CAS output orbitals are ordered in ascending occupation number (ON) and numbered according to the ON in the CAS(14,11) wavefunction (Note MO 4, 5 and 11 in CAS(8,7) output), while input orbitals are grouped with corresponding output orbitals. Color of orbital numbers reflect atom-assignment of the EUED (black: Fe, red: O_2, blue: porphyrin). For CASSCF parameters see Table 5-10. Orbital 1 corresponds to an intruder state that has larger correlation with the active space than the $O_{terminal}(p_z)$ input orbital.

of a change in weight of CAS state configurations after increasing the size of the active space (see Figure 5-29B). Both wavefunction feature a set of state configurations similar enough to be comparable. To that end, Table 5-10A lists the three dominant CAS state configurations and corresponding active spaces are shown in Figure 5-28 and Figure 5-30. The largest configuration ($\mathbf{C_A}$) in essence corresponds to double-occupied bonding and empty antibonding orbitals, which reflects a closed-shell configuration similar to **carboxy** and porphyrin **1**. The second configuration ($\mathbf{C_B}$) shows a double excitation from a π-bonding orbital between iron and dioxygen (π(FeO_2), MO 7) into the π-antibonding counterpart (π*(FeO_2), MO 8). If the CAS wavefunction would be entirely made up of $\mathbf{C_A}$ a presumably Pauling-like state would result. $\mathbf{C_B}$ however represents a charge separation and consequently diradical character between iron and dioxygen. Equal weight between $\mathbf{C_A}$ and $\mathbf{C_B}$ would correspond to a Weiss-like state (remember the example for H_2 dissociation in Figure 3-3). Thus, when the ratio of $\mathbf{C_B}/\mathbf{C_A}$ increases, the CAS wavefunction becomes more Weiss-like (or less Pauling-like) with increasing Fe – O_2 distance. While the weight of $\mathbf{C_A}$ increases ($\Delta\mathbf{C_A} \approx 1$ %) from CAS(8,7) →

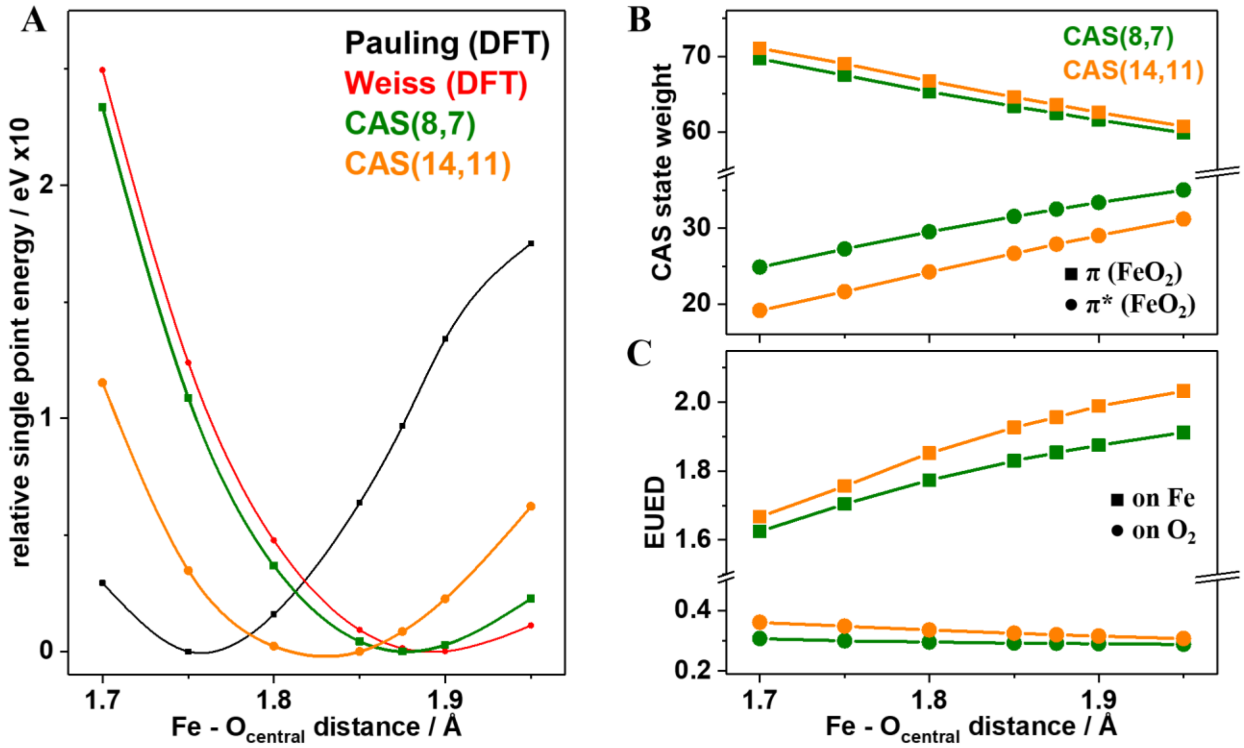


Figure 5-29: Effect of the Fe – O_2 bond length and the active space parameters on the ***oxy*** *CASSCF wavefunction. (****A****) Relative total energy (points) from DFT with (red, Weiss) and without (black, Pauling) broken-symmetry and from CASSCF. CASSCF Calculations were done on the same geometry as Weiss(DFT). The CAS(14,11) calculation used the CAS(8,7) wavefunction as initial guess and orbitals in active spaces are shown in Figure 5-28. Colored lines guide the eye. See section 5.7 for details on DFT calculation. (****B****) Weight of the two most dominant state configurations in the CAS(8,7) and CAS(14,11) wavefunctions. The first configuration (■) consists of double-occupied orbitals 4/1 – 7 (CAS(8,7)/CAS(14,11)) and empty orbitals 8 – 11 (see Figure 5-28), while the second configuration (●) has orbital 8 double-occupied instead of 7. Orbital 7 corresponds to a π-bonding orbital between iron and dioxygen (π(FeO_2)) while orbital 8 corresponds to the π-antibonding counterpart (π*(FeO_2)). (****C****) Effective unpaired electron density (EUED) on iron (■) and dioxygen (●) assigned via hard shell weighting (see colored MO numbers in Figure 5-28).*

CAS(14,11), the weight of **C$_B$** decreases (Δ**C$_B$** ≈ -5 %) for all bond lengths. This change in relative contribution therefore represents a shift towards a more closed-shell configuration, which results in an energy minimum shift towards smaller bond lengths. The third configuration (**C$_C$**) differs from **C$_A$** by two single excitations, (i) O_2(π) → π*(FeO_2) (MOs 4, 7) and (ii) O_2(σ) → O_2(σ*) (MOs 6, 9). While the second single excitation represents static electron correlation between a bonding and antibonding orbital, the first single excitation transfers an electron from an O_2 centered MO into an Fe centered MO. Consequently, the sum of occupation numbers (ON) in MO 7 and 8 is > 2. To that end, ONs show that correlation pairs/trios are preserved in both active spaces. Namely, the sums of ONs in (i) MO 5 and 11 and (ii) MO 6 and 9 yields two-electron two-orbital configurations while MOs 4, 7 and 8 correspond to four-electron three-orbital interactions (see next section). The four additional MOs in the CAS(14,11) calculations form another group as a correlation quartet. Additionally, the ONs in MO 5/11 de-/increase, which indicates slightly larger correlation between these two MOs, presumably because CASSCF compensated for the orbitals with less correlation not included in the active space of the CAS(8,7) calculations.

Both CAS wavefunctions form a singlet state and thus have no spin density. Fractional ON in the natural MOs however still allow for local spin analysis in term of Takatsukas effective unpaired electron density (EUED). The increase in **C_B/C_A** ratio results in an increase in total EUED with increasing bond length, which is almost entirely assigned to Fe or O_2 by a hard shell weighting (see Figure 5-29C). For CAS(8,7), the EUED on iron of ~ 1.9 at 1.9 Å, which is closest to the PES energy minimum from DFT and the bond length from EXAFS, is in good agreement with Kβ main-line XES. The EUED on iron stems almost entirely (~ 98 %) from MOs 7 and 8. However, the EUED on iron for the CAS(14,11) calculation of ~ 2 is in perfect agreement (see Table 5-9), which results from a decrease in **C_B/C_A** ratio and correlation in the additional, iron centered MOs. Consequently, the EUED on iron from MOs 7 and 8 decreases down to ~ 88 % and the remaining ~ 12 % stems from EUED in MOs 2, 3, 10 and 11.

Electronic configuration in oxy from CASSCF

The singlet state in **oxy** and the nature of the bond between iron and dioxygen has sparked a debate spanning more than eight decades [39, 45, 55, 57, 58, 60] (see section 1). The proposed models include (i) a neutral singlet O_2 molecule bound to a singlet Fe(II) (low-spin, d^6) by Pauling and Coryell, (ii) a neutral triplet O_2 molecule with antiferromagnetic coupling (afc) to a triplet Fe(II) (intermediate-spin, d^6) by McClure and Goddard and (iii) a charged doublet O_2^- superoxide with afc to a doublet Fe(III) (high-spin d^5) by Weiss (see Figure 5-1D). The following section discusses results derived from the **oxy** CAS(14,11) calculation for an Fe – O_2 bond length of 1.90 Å in the context of the proposed models from Pauling/Coryell, McClure/Goddard and Weiss.

Two state configurations form the majority (~ 90 %, see Table 5-10A) of the CAS(14,11) wavefunction. These two configurations (**C_A**, **C_B**), depending on their relative weighting, result in the Pauling (100 % **C_A**, 0 % **C_B**) or Weiss model (50 % **C_A**, 50 % **C_B**) as they describe the transition from a closed-shell singlet state to open-shells that couple to a singlet state (afc, see

H_2 dissociation example in Figure 3-3). However, the present ratio **C_B/C_A** of ≈ 0.5 favors neither model. Additionally, the remaining 10 % of the wavefunction stem from configurations that do not correspond to Pauling- or Weiss-like states. Though **C_C**, the third configuration, consists of two single excitations and thus contains four single-occupied orbitals, therefore it does neither describe a configuration similar to the McClure/Goddard model. The four unpaired electrons are in three O_2 centered MOs (MOs 4, 6 and 9 in Figure 5-30) and one iron centered MO (MO

Table 5-10: Parameters of CAS(8,7) and CAS(14,11) calculations at 1.9 Å Fe – O_2 bond length.

A	state configuration											state weight [%]	
MO no.	1	2	3	4	5	6	7	8	9	10	11	CAS(8,7)	CAS(14,11)
C_A	2	2	2	2	2	2	2	0	0	0	0	61.6	62.6
C_B	2	2	2	2	2	2	0	2	0	0	0	33.4	29.0
C_C	2	2	2	1	2	1	2	1	1	0	0	1.3	1.2

B	occupation number (ON) / spin density (SD)										
MO no.	1	2	3	4	5	6	7	8	9	10	11
CAS(8,7)				1.97	1.98	1.95	1.30	0.72	0.05		0.02
CAS(14,11)	1.99	1.98	1.97	1.97	1.97	1.95	1.37	0.67	0.05	0.05	0.03

C	effective unpaired electron density (EUED)										
CAS(8,7)				0.05	0.05	0.10	0.91	0.92	0.10		0.04
CAS(14,11)	0.02	0.04	0.05	0.05	0.06	0.10	0.86	0.89	0.10	0.09	0.06

D	Assigned EUED			
	total	Fe	O_2	porphyrin
CAS(8,7)	2.166	1.875	0.291	0
CAS(14,11)	2.326	1.989	0.315	0.022

*(**A**) Configurations and weights of the three dominant CAS states in CAS(8,7) and CAS(14,11) calculations. Orbitals are shown in Figure 5-28. State configuration refers to double-, single- or non-occupied natural molecular orbitals (NMOs) in active space. Note that the active space for CAS(8,7) only contains NMOs 4 – 9 and 11. (**B**) Occupation numbers (ON) in CAS NMOs, which is the weighted average of all state configurations in a CAS wavefunction. (**C**) Effective unpaired electron density (EUED) in CAS NMOs resulting from fractional occupation numbers (EUED = ON*[2-ON], see section 3.2) (**D**) Total EUED assigned to iron, dioxygen or porphyrin. EUED is assigned to the atom/molecule with largest atomic orbital contribution in a CAS orbital (see colored orbital indices in Figure 5-28).*

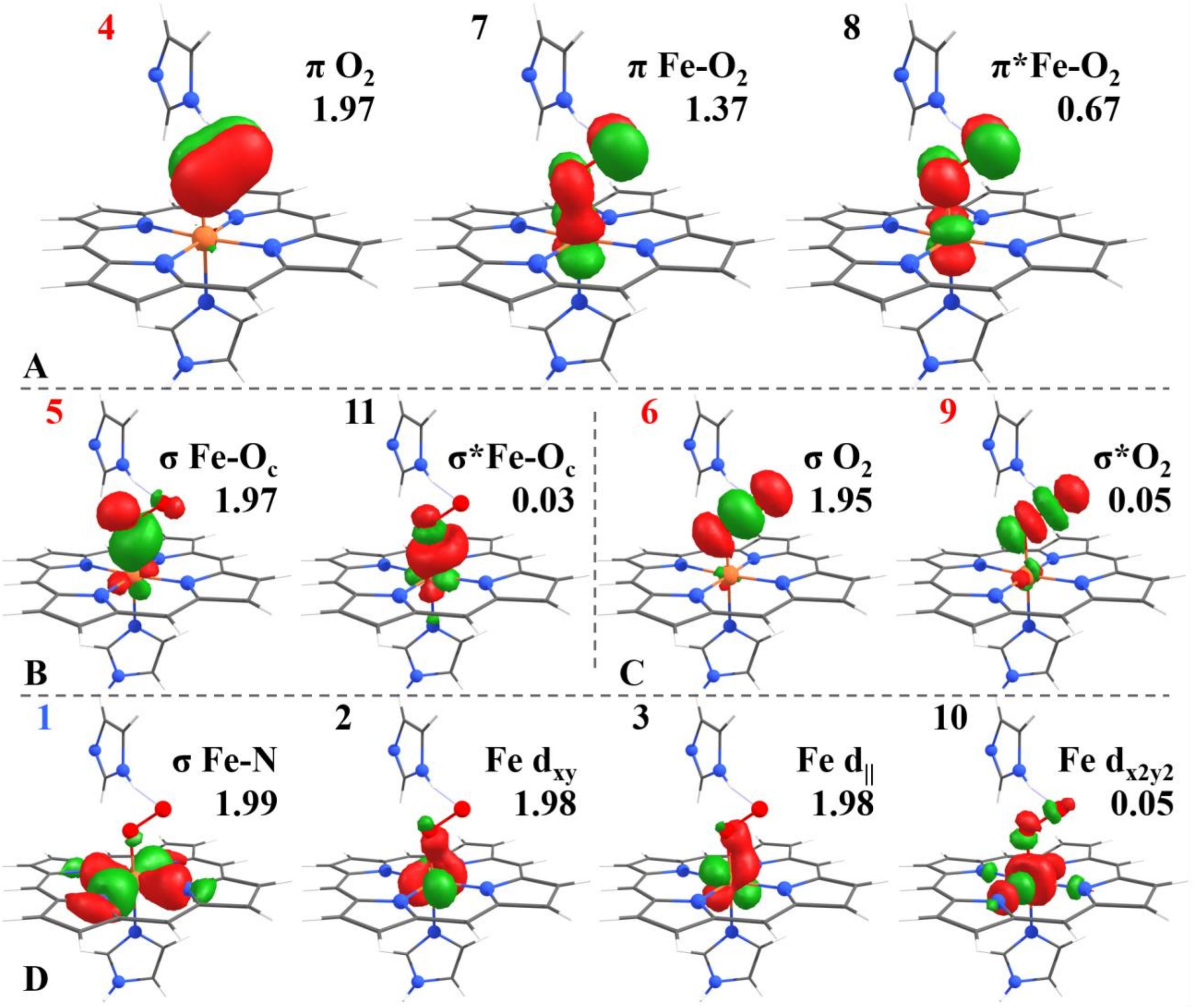


Figure 5-30: Iron-oxygen bonding from CAS(14,11) simulation of ***oxy****. Indicated are MO indices, occupation numbers and assignments to chemical bonding (O_c = Fe-bound atom, Fe $d_{||}$ = superposition of d_{xz} and d_{yz} MOs parallel to Fe – O_2 plane). Axis are according to the octahedral geometry of iron with the z-axis along N_{His} – Fe – O_c and x/y-axis along N_{porph} – Fe – N_{porph}. Lines separate orbitals into correlation groups* ***A - D****. CAS wavefunction parameters are shown in Table 5-10 and relative orbital composition in Table 5-11. Isosurfaces are drawn at ± 0.05 au. σ-donation and delocalized π/π* backbonding stabilizes the O_2 ligand. For further analysis see text.*

8). In order to achieve an intermediate-spin state as suggested by McClure or Goddard, the second configuration would need a second double excitation from another bonding into an antibonding MO e.g. MO 5 → 11. Consequently, the wavefunction cannot be assigned to a singular model. Rather, the electronic configuration from CASSCF combines elements from all three models. Accordingly, the discussion has to move from individual CAS state configurations towards a more holistic interpretation of the wavefunction. Correlation groups formed by NMOs in the active space (see Figure 5-30) help to characterize chemical bonding. The largest correlation group **D** (see Figure 5-30) consists of (i) a σ bond between iron and the porphyrin nitrogen atoms (σ(Fe-N), MO 1), (ii) an essentially double-occupied Fe(d_{xy}) orbital (MO 2), (iii) a double-occupied Fe(3d) orbital oriented parallel to the Fe – O – O plane (Fe($d_{||}$), MO 3) and (iv) an essentially empty Fe($d_{x2\text{-}y2}$) orbital. The assignments are supported by the orbital composition on iron in Table 5-11, e.g. the σ(Fe-N) MO shows a minor contribution from Fe($d_{x2\text{-}y2}$) and no contribution from O_2. Fe($d_{||}$) is a linear combination of Fe(d_{xz}) and Fe(d_{yz}) MOs. Orbital shape and atom contributions show that Fe(d_{xy}), Fe($d_{||}$) and Fe($d_{x2\text{-}y2}$) are centered predominantly on iron. The correlation pair in **C** is a double-occupied σ(O_2) and an empty σ*(O_2) orbital with negligible contribution from iron. **C** and **D** therefore have negligible influence on the Fe – O_2 bond.

The correlation duo in **B** has almost double-occupied/empty σ(Fe-O_c)/σ*(Fe-O_c) MOs, which represents a polarized σ(Fe-O_c) bond with dominant O_c(p) and weak Fe(d_{z2}) contributions, suggesting strong bond polarization (i.e. an ionic dative bond with electron density transfer from O_2 to the Fe(d_{z2}) orbital) [54, 55]. This is the bond suggested by Pauling and in contrast to a covalent σ bond suggested by Goddard. The group **A** further complicates the situation. In this group, π(Fe-O_2) and π*(Fe-O_2) orbitals (7,8) with dominant Fe(d) (perpendicular to Fe($d_{||}$)

*Table 5-11: AO contribution to NMOs in active space from CAS(14,11) calculations in **oxy***

1.9 Å		Löwdin orbital composition [%]										
MO No.		1	2	3	4	5	6	7	8	9	10	11
Fe	3d	x^2-y^2	xy	xz/yz	-	z^2	-	xy/yz	xy/yz	-	x^2-y^2	z^2
	d_{sum}	28	93	91	3	22	2	53	46	4	77	44
O_c	2p	-	-	-	x/y	z	x/y/z	x/y	x/y	x/y/z	-	z
	p_{sum}	1	3	4	58	64	47	12	23	37	6	36
O_t	2p	-	-	-	x/y	-	x/y/z	x/y	x/y	x/y/z	-	-
	p_{sum}	0	0	1	35	3	47	31	29	39	2	0

The Löwdin orbital from iron, central and terminal oxygen to NMOs shown in Figure 5-30. Shown is the sum of 3d (Fe) and 2p ($O_{c/t}$) AO contributions. Additionally, the main orbital contribution is listed according to the octahedral ligand field that surrounds the iron ion, with the z-axis along N_{His} – Fe – O_c and x/y-axis along N_{porph} – Fe – N_{porph}. Contributions below 10 % are omitted. Note that the orientation of orbitals 3, 4, 6, 7, 8 and 9 is not along a singular axis, which yields multiple mayor contributions.

in **D**) and weaker O_2(p) contributions are partially occupied (~ 1.4 and 0.7), which indicates a diradical formation with afc as suggested by Weiss. This is also supported by a Mulliken charge of about -0.5 on O_2 (see Table 5-9). However, the static correlation with an almost double-occupied π(O_2) orbital corresponds to the four-electron in three-orbitals ozone configuration suggested by Goddard (compare ozone example in Figure 3-5). Additionally, The effective unpaired electron density (EUED) from the π(Fe-O_2) and π*(Fe-O_2) orbitals of ~0.9 represents ~90 % of the total EUED on iron. Static correlation between iron centered orbitals in **D** and the dative σ(Fe-O_c) bond in **C** contribute the remaining ~10 %, which results in two effectively unpaired spins on iron (as observed in Kβ main-line XES), which, combined with the double-occupied Fe(3d) orbitals, results in an effectively intermediate-spin Fe(II) as suggested by McClure. The total EUED on O_2 is ~0.3 and the total EUED of iron and dioxygen is ~2.3. The latter value is close to the EUED of ~2.4 for a charged ozone [157].

Conclusion

Complete active space self-consistent field (CASSCF) calculations with varying active space size were carried out on porphyrin reference complexes and non-oxy heme species. The spin densities from natural molecular orbitals in the active space and the effective unpaired electron density (EUED) on iron and ligands were determined from the CAS wavefunctions. Three double-occupied Fe(3d) MOs for the porphyrin **1** and **carboxy** (low-spin Fe(II)) and four or five single-occupied Fe(3d) MOs for the **deoxy**, **2** or **3**, (high-spin Fe(II)/(III)) with integer EUEDs on iron (0, 4, or 5) fully agree with the formal oxidation and spin states and with our results from Kβ XES. This shows that CASSCF with a sufficiently large active space is well suited to describe the electronic configuration of the heme cofactor.

CASSCF on **oxy** structures with varying Fe – O_2 bond length were performed. The aim was to (i) identify a stable active space and (ii) evaluate the systematic effects of geometry variation. Two stable active space configurations emerged, CAS(8,7) (8 electrons in 7 orbitals) and CAS(14,11). Both CAS wavefunctions were dominated by two CAS state configurations ($\mathbf{C_A}$, $\mathbf{C_B}$), which differ by a double-occupied/unoccupied ($\mathbf{C_A}$) or unoccupied/double-occupied ($\mathbf{C_B}$) bonding/antibonding π(FeO_2)/π*(FeO_2) orbital. These two configurations, depending on their relative weight, could represent Fe – O_2 bonding according to Pauling or Weiss. Indeed, relative energies for **oxy** heme from CAS(8,7) and the Weiss-like state from DFT were similar. However, increasing the active space in CAS(14,11) results in a decrease of diradical character in the Fe – O_2 site, which shifts the energy minimum towards shorter bond lengths. The EUED on iron, for CAS(14,11) at an Fe – O_2 distance in agreement with EXAFS and DFT, results in effectively two unpaired spins. This is in perfect agreement with our experimental results from Kβ XES. This reveals that the **oxy** heme is neither described by the Pauling or Weiss models but the electronic configuration is much more complex.

A thorough analysis of the CAS wavefunction for **oxy** was performed. It indicates (i) a dative σ(Fe-O_c) bond as in the Pauling model, (ii) partial superoxide formation with afc between Fe and O_2 similar to the Weiss model, (iii) an ozone-like four-electron three-orbital configuration as suggested by Goddard combined with (iv) an effectively intermediate-spin Fe(II) in line with McClure. The global picture of the Fe – O_2 bonding in MB/HB therefore converges towards a three-center interaction where σ-donation and delocalized π/π* backbonding stabilize the O_2 ligand, which combines features of each of the classical models. Our analysis resolves the apparent discrepancies between the classical models of the **oxy** heme though quantitative correlation of results from spectroscopy and quantum chemistry, which benchmarks the multiconfigurational description of the Fe – O_2 bonding in the oxygenated heme cofactor.

5.9 Discussion

We used Fe Kβ-XES and K-edge XAS and quantum chemistry calculations to study molecular and electronic structures of MB and HB heme species in solution. Fast spectral changes for all species even at cryogenic temperatures and moderate X-ray flux mandated rapid-scan XAS and energy-sampling XES approaches for unperturbed data collection. Apparent bond lengths/angle variations particularly in crystal structures of **oxy** heme likely relate to such radiation-induced modifications [205]. The X-ray induced heme conversion is interesting in itself and we interpret our data as sequential two-electron reduction (at 20 K) of bound CO or O_2 followed by product detachment [221] to form apparent **deoxy** heme and replacement by mobile carbon oxide species (at 260 K) possibly derived from amino acid decarboxylation [222]. The X-ray spectral changes support our redox/spin assignments for the heme species.

Similar metrical, electronic, and kinetic parameters indicated similar metal site configurations in MB and HB, which ruled out a significant influence of amino acid environment or hydrogen-bonding variations on the iron centers and ligand bonding. Heme side-chain orientation differences in MB/HB were reflected in the XANES by slight virtual level, i.e. Fe(4p), energy changes [223] barely affecting the iron valence levels. All relevant heme features in MB/HB were invariant within 20 – 260 K, excluding population of higher spin states even far above the protein glass-transition point (~ 200 K) [224]. This finding and, e.g., gradual spin crossover in iron-porphyrins [225-227] suggest that also at room temperature the spin state in the three heme species is the same and therefore physiologically relevant.

Kβ main line emission spectra of MB/HB and porphyrin model compounds revealed a direct correlation of the formal unpaired Fe(3d) spin count to the relative Kβ' intensity, in agreement with results for other iron species [25, 36], establishing further that the iron-centered radiative 3p → 1s decay is a reliable probe of the valence level configuration. The steep correlation locates the overall iron bonding in the MB/HB hemes and the porphyrins closer to the non-covalent than covalent limit [25, 36, 90, 91]. Agreement of our and earlier Kβ data of MB/HB [19] supports the spin state changes. For the heme proteins and porphyrins, our Kβ main-line and ctv/vtc data and (TD-)DFT/CASSCF revealed ferrous high-spin (**deoxy**, porphyrin **2**) or low-spin (**carboxy**, porphyrin **1**) and ferric high-spin (porphyrin **3**) species [1, 39], emphasizing the method´s sensitivity for valence/spin state discrimination.

For **oxy**, the Kβ main-line spectra indicated two unpaired Fe(3d) spins on iron, which was compatible with the ctv/vtc data, by definition implying an intermediate-spin metal center, as suggested by McClure/Goddard [58, 59]. No significantly decreased ligand character of Fe(d) levels in **oxy** vs. **deoxy**/**carboxy** in the calculations rendered lowered iron-ligand bonding covalency as an explanation for the Kβ data of **oxy** unlikely. The Pauling model with a low-spin Fe(II) and an innocent O_2 ligand [45] was essentially excluded (as well as higher spin states, $M > 1$) by our data, strengthening previous conclusions [1, 201]. DFT converged to a Weiss-like configuration with significant diradical character between Fe and O_2, which resulted in a ferric iron and a superoxide ligand [57] and yielded ctv/vtc spectra more close to our and earlier data [1, 201]. While the prediction from DFT, of diradical character present in the electron density of the Fe – O_2 bond, is supported by experiment, it appears to be the result of error cancelation and therefore overestimated in the DFT wavefunction [55, 76]. CASSCF on **oxy** yielded a multiconfigurational wavefunction dominated by two CAS state configurations as found earlier [55, 71, 76]. It comprised in effect two unpaired spins on iron, in perfect agreement with our Kβ main-line data, thereby benchmarking the quantum chemical prediction and providing a quantitative link between theory and spectroscopy.

The emerging bonding situation includes (i) a dative σ(Fe-O_c) bond as in the Pauling model, (ii) partial superoxide formation with antiferromagnetic coupling between Fe and O_2 similar to the Weiss model, (iii) an ozone-like four-electron three-orbital configuration as suggested by Goddard combined with (iv) an effectively intermediate-spin Fe(II) in line with McClure. It combines the classical models into a more realistic, unifying view of O_2-bonding in the heme

Table 5-12: Effective unpaired electron density from CASSCF.

	species	redox & spin state	EUED	
			Fe	ligand[a]
Heme	**deoxy**	Fe(II) HS	4.00 (4)	-
	carboxy	Fe(II) LS	0.00 (0)	0.35
	oxy	Fe(II) IS[b]	1.99 (2)	0.32
Porphyri	**1**	Fe(II) LS	0.00 (0)	0.00
	2	Fe(II) HS	4.00 (4)	0.37
	3	Fe(III) HS	5.00 (5)	0.00

[a] *Axial CO (**carboxy**), O_2 (**oxy**), N_{Py} (**1**), O_{THF} (**2**), or Cl (**3**) ligands (see Figure 5-27 and Figure 5-30);* [b] *simplified annotation (see text); formal unpaired Fe(d) spins in parenthesis.*

proteins, featuring limited radical character and delocalized backbonding stabilization of the ligand as prerequisites in reversible O_2 transport.

6 General discussion

In this thesis, synchrotron-based advanced X-ray absorption and emission spectroscopy (XAS/XES) was combined with quantum chemistry calculations using density functional theory (DFT), time-dependent DFT (TD-DFT) and the complete active space self-consistent field method (CASSCF) from molecular orbital (MO) theory to determine geometric and electronic parameters of metal sites in proteins and synthetic compounds. This combination of experimental and theoretical methods is of high value for future investigations of complex metal-ligand bonding scenarios as it provides quantitative relations between experimentally accessible parameters (oxidation/spin state), molecular and electronic structures and metal-ligand bonding properties. This section provides a summary of the most important findings in relation to ongoing work in our laboratory and possible future research.

I Procedures for time-resolved energy-sampling (TRES) X-ray spectroscopy, both in absorption and emission mode, were developed and evaluated. In general, the TRES approach resulted in an about ten-fold increase in experimental efficiency, compared to conventional energy step-scan procedures, facilitating XES with high signal-to-noise ratio on highly diluted and X-ray sensitive materials. General expressions for description of experimental efficiency were derived, which depends only on a small set of synchrotron-setup and sample specific parameters. They were implemented in a software tool that facilitates making a well-founded decision whether a conventional (step-scan) or the time-resolved energy-sampling approach should be conducted. TRES is a viable approach to increase the overall experimental efficiency for maximal output during scarce beamtime at synchrotron facilities.

II Advanced X-ray spectroscopy methods were applied to various metal centers in proteins and synthetic compounds (see our publications). A special focus in this thesis is the investigation of heme proteins (hemoglobin, HB and myoglobin, MB). Our TRES method made it possible to collect complete XAS/XES data sets for the radiation-sensitive HB and MB species in a temperature range from 20 K to 260 K. Thus, TRES allows to bridge the gap between physiologically more relevant temperatures and experimentally favorable cryogenic conditions for radiation damage sensitive protein samples.

III Advanced quantum chemistry methods (e.g. TD-DFT, MP2, CASSCF) were implemented in our laboratory and the procedures and involved software were documented for future application. One focus was our development of tools and approaches for analysis and understanding of the results of these computational methods. Software tools were developed that facilitate rapid readout and graphical representation of the contents of large-scale output files from quantum chemistry software packages. Several approaches (natural bond order, topological electron density, molecular orbital representations) were quantitatively compared and guidelines for application of different MO representations were verified.

IV The combination of experimental XAS/XES data with theoretical DFT/CASSCF calculations allowed determination of key structural and electronic parameters in metal centers of proteins and synthetic compounds. The combination of XES and CASSCF, to our knowledge, has not been previously applied in literature and our results open up the field for investigations on biological and synthetic systems.

V The application of advanced X-ray spectroscopy and quantum chemical methods to heme proteins resulted in quantitative correlation between experimental and theoretical parameters and benchmarking of the computational results. Our finding on the spin state and electronic configuration in oxygenated heme in HB and MB provided an adequate description of the metal-ligand interaction, which merged the classical models in a holistic description of the Fe – O_2 bonding situation.

High-resolution XAS and XES in combination with advanced quantum chemical simulations thereby provided a versatile toolset to determine geometric and electronic structures in active sites in biological and synthetic catalysts.

The combination of effective unpaired electron density analysis from CASSCF with Kβ main-line emission (publications I, IV, V) was remarkably successful [102, 103, 105]. As the example of oxygenated HB/MB illustrates, especially for compounds where application of more rigid chemical frameworks is difficult, the multiconfigurational quantum chemical approach provides a cleaner picture of the electronic configuration. CASSCF calculations, however, remain a considerable challenge, in particularly for compounds that contain more than one transition metal ion. Thus, one should first turn to Mulliken spin densities from DFT and attempt multiconfigurational MO theory only if necessary. The CASSCF approach is of particular interest when transition metal with non-innocent ligands are involved in an EPR-silent state. Heme proteins are involved in a wide variety of catalytic processes of small-molecule chemistry. Nitrosylheme complexes are surrounded by a similar discussion as was ongoing on oxygenated HB/MB [17, 39, 54, 152]. Elucidation of the electronic configuration of iron after binding of a non-innocent nitrogen oxide may be an interesting future topic for application of advanced X-ray and quantum chemistry methods.

Theoretical satellite line spectra from DFT were able to reproduced experimental trends for a variety of metal cofactor and synthetic compounds featuring mono- or binuclear metal centers with manganese, iron and nickel ions (see publications). Yet, DFT did not predict all satellite line features, of the high-spin heme/porphyrin complexes in this work and earlier publications. This could indicate an edge-case, where the static one-electron process implied by the theoretical model does not apply. However, implementation of more dynamic simulations, as involved in the TD-DFT for pre-edge simulation, would likely require a significant enlargement of computation resources. Good agreement was obtained between pre-edge spectra and TD-DFT calculations for predominantly single-configurational heme proteins and porphyrin model compounds. Additionally, theoretical ctv transitions from open-shell singlet wavefunctions with broken-symmetry versus closed-shell singlet wavefunctions calculated using hybrid DFT functionals facilitated identification of antiferromagnetic coupling in heme proteins and dinickel complexes (II in our publications). These promising results motivated us to address the ongoing discussion (publication III) surrounding the identification of the coordination site and substrate water molecules in the oxygen-evolving Mn_4Ca cluster in photosystem II. Accordingly, we combined pre-edge absorption features of photosystem II samples in four stable catalytic states, monitored by time-resolved energy-sampling, with calculated core-to-valence spectra from TD-DFT. This ongoing work will presumably result in new insights into the complex catalytic metal cofactor of photosystem II.

With ongoing development of improved quantum chemistry methods and computation facilities, as well as inauguration of new 3rd and 4th generation synchrotron sources worldwide combined X-ray spectroscopy and quantum chemistry methods may soon become a standard in bioinorganic chemistry research.

7 List of publications

I N. Schuth, S. Mebs, D. Huwald, P. Wrzolek, M. Schwalbe, A. Hemschemeier, M. Haumann
Effective intermediate spin iron in O2-transporting heme proteins.
Proc. Natl. Acad. Sci. USA 114, 8556-8561 (2017)

II N. Schuth, H. Gehring, B. Horn, P. Holze, R. Kositzki, P. Schrapers, C. Limberg, M. Haumann
Biomimetic mono- and dinuclear Ni(I) and Ni(II) complexes studied by X-ray absorption and emission spectroscopy and quantum chemical calculations.
J. Phys. Conf. Ser. 712, 1-012134 (2016)

III N. Schuth, Z. Liang, M. Schonborn, A. Kussicke, R. Assuncao, I. Zaharieva, Y. Zilliges, H. Dau
Inhibitory and Non-Inhibitory NH3 Binding at the Water-Oxidizing Manganese Complex of Photosystem II Suggests Possible Sites and a Rearrangement Mode of Substrate Water Molecules.
Biochemistry 56, 6240-6256 (2017)

IV N. Schuth, M. Haumann
X-ray spectroscopy reveals iron spin state in oxygenated hemoproteins.
ESRF Spotlight, http://www.esrf.eu/home/news/spotlight/content-news/spotlight/spotlight301.html (2017)

V N. Schuth, M. Haumann
X-ray spectroscopy reveals iron spin state in oxygenated hemoproteins.
ESRF Highlights 2017, (2018)

VI A. Petuker, S. Mebs, N. Schuth, P. Gerschel, M. L. Reback, B. Mallick, M. van Gastel, M. Haumann, U.-P. Apfel
Spontaneous Si-C bond cleavage in (triphos-Si)-nickel complexes.
Dalton Trans. 46, 907-917 (2017)

VII R. Kositzki, S. Mebs, N. Schuth, N. Leidel, L. Schwartz, M. Karnahl, F. Wittkamp, D. Daunke, A. Grohmann, U.-P. Apfel, F. Gloaguen, S. Ott, M. Haumann
Electronic and molecular structure relations in diiron compounds mimicking the [FeFe]-hydrogenase active site studied by X-ray spectroscopy and quantum chemistry.
Dalton Trans. 46, 12544-12557 (2017)

VIII P. Schrapers, S. Mebs, Y. Ilina, D. S. Warner, C. Wörmann, N. Schuth, R. Kositzki, H. Dau, C. Limberg, H. Dobbek, M. Haumann
The binuclear nickel center in the A-cluster of acetyl-CoA synthase (ACS) and two biomimetic dinickel complexes studied by X-ray absorption and emission spectroscopy.
J. Phys. Conf. Ser. 712, 1-012029 (2016)

IX R. Kositzki, S., Mebs, J. Marx, J. J. Griese, N. Schuth, M. Högbom, V. Schünemann, M. Haumann
Protonation state of MnFe and FeFe cofactors in a ligand binding oxidase revealed by X-ray absorption, emission, and vibrational spectroscopy and QM/MM calculations.
Inorg. Chem. 19, 9869-9885 (2016)

8 References

1. Wilson, S.A., et al., *X-ray absorption spectroscopic investigation of the electronic structure differences in solution and crystalline oxyhemoglobin.* Proc Natl Acad Sci U S A, 2013. **110**(41): p. 16333-8.
2. Vojtechovsky, J., et al., *Crystal structures of myoglobin-ligand complexes at near-atomic resolution.* Biophys J, 1999. **77**(4): p. 2153-74.
3. Lowdin, P.O., *Quantum Theory of Many-Particle Systems .1. Physical Interpretations by Means of Density Matrices, Natural Spin-Orbitals, and Convergence Problems in the Method of Configurational Interaction.* Physical Review, 1955. **97**(6): p. 1474-1489.
4. Cramer, C.J., *Essentials of Computational Chemistry*. 2nd ed. 2004: John Wiley & Sons Ltd.
5. Neese, F., *Definition of corresponding orbitals and the diradical character in broken symmetry DFT calculations on spin coupled systems.* Journal of Physics and Chemistry of Solids, 2004. **65**(4): p. 781-785.
6. Autschbach, J., *Orbitals: Some Fiction and Some Facts.* Journal of Chemical Education, 2012. **89**(8): p. 1032-1040.
7. Park, S.Y., et al., *1.25 angstrom resolution crystal structures of human haemoglobin in the oxy, deoxy and carbonmonoxy forms.* Journal of Molecular Biology, 2006. **360**(3): p. 690-701.
8. Pin, S., et al., *X-Ray-Absorption Spectroscopy of Hemoglobin.* Hemoglobins, Pt C, 1994. **232**: p. 266-292.
9. Lima, F.A., et al., *Probing the electronic and geometric structure of ferric and ferrous myoglobins in physiological solutions by Fe K-edge absorption spectroscopy.* Physical Chemistry Chemical Physics, 2014. **16**(4): p. 1617-1631.
10. Eisenberger, P., et al., *Structure-Function Relations in Hemoglobin as Determined by X-Ray Absorption Spectroscopy.* Proceedings of the National Academy of Sciences of the United States of America, 1976. **73**(2): p. 491-495.
11. Bianconi, A., et al., *Local Fe Site Structure in the Tense-to-Relaxed Transition in Carp Deoxyhemoglobin - a-Xanes (X-Ray Absorption near Edge Structure) Study.* Proceedings of the National Academy of Sciences of the United States of America, 1986. **83**(20): p. 7736-7740.
12. Durham, P., et al., *X-Ray Absorption near Edge Structure (Xanes) for Co, Cn and Deoxyhemoglobin - Geometrical Information.* Embo Journal, 1983. **2**(9): p. 1441-1443.
13. Zhang, K., et al., *Active site conformation in myoglobin as determined by X-ray absorption spectroscopy.* Proteins: Structure, Function, and Genetics, 1991. **10**(4): p. 279-586.
14. Lukin, J.A., et al., *NMR reveals hydrogen bonds between oxygen and distal histidines in oxyhemoglobin.* Proc Natl Acad Sci U S A, 2000. **97**(19): p. 10354-8.
15. Phillips, S.E.V. and B.P. Schoenborn, *Neutron-Diffraction Reveals Oxygen-Histidine Hydrogen-Bond in Oxymyoglobin.* Nature, 1981. **292**(5818): p. 81-82.
16. Mebs, M., et al., *Abrupt versus gradual spin-crossover in Fe(II)(phen)2(NCS)2 and Fe(III)(dedtc)3 compared by X-ray absorption and emission spectroscopy and quantum-chemical calculations.* Inorg Chem, 2015. **54**: p. 11606-11624.
17. Huang, X. and J.T. Groves, *Oxygen Activation and Radical Transformations in Heme Proteins and Metalloporphyrins.* Chemical Reviews, 2017.
18. Trambarulo, R., et al., *The Molecular Structure, Dipole Moment, and g Factor of Ozone from Its Microwave Spectrum.* The Journal of Chemical Physics, 1953. **21**(5): p. 851-855.

19. Koster, A.S., *Electronic State of Iron in Hemoglobin, Myoglobin, and Derivatives, as Inferred from X-Ray Fluorescence Spectra.* The Journal of Chemical Physics, 1972. **56**(6): p. 3161-3164.
20. Westre, T.E., et al., *A Multiplet Analysis of Fe K-Edge 1s → 3d Pre-Edge Features of Iron Complexes.* Journal of the American Chemical Society, 1997. **119**(27): p. 6297-6314.
21. Koster, A.S., *Electronic State of Iron in Hemoglobin, Myoglobin and Derivatives, as Inferred from X-Ray-Fluorescence Spectra.* Journal of Chemical Physics, 1972. **56**(6): p. 3161-&.
22. Safo, M.K., et al., *Models of the Cytochromes. Axial Ligand Orientation and Complex Stability in Iron(II) Porphyrinates: The Case of the Noninteracting dπ Orbitals.* Journal of the American Chemical Society, 1997. **119**(40): p. 9438-9448.
23. Glatzel, P. and U. Bergmann, *High resolution 1s core hole X-ray spectroscopy in 3d transition metal complexes - electronic and structural information.* Coordination Chemistry Reviews, 2005. **249**(1-2): p. 65-95.
24. DeBeer George, S., T. Petrenko, and F. Neese, *Prediction of iron K-edge absorption spectra using time-dependent density functional theory.* J Phys Chem A, 2008. **112**(50): p. 12936-43.
25. Mebs, S., et al., *Abrupt versus Gradual Spin-Crossover in Fe-ll(phen)(2)(NCS)(2) and Fe(III)dedtc)(3) Compared by X-ray Absorption and Emission Spectroscopy and Quantum-Chemical Calculations.* Inorganic Chemistry, 2015. **54**(24): p. 11606-11624.
26. Lambertz, C., et al., *Electronic and molecular structures of the active-site H-cluster in [FeFe]-hydrogenase determined by site-selective X-ray spectroscopy and quantum chemical calculations.* Chemical Science, 2014. **5**(3): p. 1187-1203.
27. Lee, N., et al., *Probing valence orbital composition with iron Kbeta X-ray emission spectroscopy.* J Am Chem Soc, 2010. **132**(28): p. 9715-27.
28. Thomson, A.J. and H.B. Gray, *Bio-inorganic chemistry.* Current Opinion in Chemical Biology, 1998. **2**(2): p. 155-158.
29. Pettersson, E., J. Lundeberg, and A. Ahmadian, *Generations of sequencing technologies.* Genomics, 2009. **93**(2): p. 105-111.
30. Umena, Y., et al., *Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å.* Nature, 2011. **473**: p. 55.
31. Young, I.D., et al., *Structure of photosystem II and substrate binding at room temperature.* Nature, 2016. **540**: p. 453.
32. Dau, H., I. Zaharieva, and M. Haumann, *Recent developments in research on water oxidation by photosystem II.* Curr Opin Chem Biol, 2012. **16**(1-2): p. 3-10.
33. Gallo, E. and P. Glatzel, *Valence to Core X-ray Emission Spectroscopy.* Advanced Materials, 2014. **26**(46): p. 7730-7746.
34. Neese, F., *Prediction of molecular properties and molecular spectroscopy with density functional theory: From fundamental theory to exchange-coupling.* Coordination Chemistry Reviews, 2009. **253**(5-6): p. 526-563.
35. Roemelt, M., et al., *Manganese K-Edge X-Ray Absorption Spectroscopy as a Probe of the Metal-Ligand Interactions in Coordination Compounds.* Inorganic Chemistry, 2012. **51**(1): p. 680-687.
36. Pollock, C.J., et al., *K beta Mainline X-ray Emission Spectroscopy as an Experimental Probe of Metal-Ligand Covalency.* Journal of the American Chemical Society, 2014. **136**(26): p. 9453-9463.
37. Planavsky, N.J., et al., *Evidence for oxygenic photosynthesis half a billion years before the Great Oxidation Event.* Nature Geoscience, 2014. **7**: p. 283.
38. Poulos, T.L., *Heme enzyme structure and function.* Chem Rev, 2014. **114**(7): p. 3919-62.

39. Kepp, K.P., *Heme: From quantum spin crossover to oxygen manager of life.* Coordination Chemistry Reviews, 2017. **344**: p. 363-374.
40. Wilks, A. and G. Heinzl, *Heme oxygenation and the widening paradigm of heme degradation.* Archives of Biochemistry and Biophysics, 2014. **544**: p. 87-95.
41. Meunier, B., S.P. de Visser, and S. Shaik, *Mechanism of Oxidation Reactions Catalyzed by Cytochrome P450 Enzymes.* Chemical Reviews, 2004. **104**(9): p. 3947-3980.
42. Hünefeld, F.L., *Der Chemismus in der Thierischen Organisation*. 1840: Brockhaus, Leipzig.
43. Faraday, M., *On new magnetic actions, and on the magnetic condition of all matter*, in *EXPERIMENTAL RESEARCHES IN ELECTRICITY*. 1845, Taylor and Francis: London.
44. Ferry, R.M. and A.A. Green, *STUDIES IN THE CHEMISTRY OF HEMOGLOBIN: III. THE EQUILIBRIUM BETWEEN OXYGEN AND HEMOGLOBIN AND ITS RELATION TO CHANGING HYDROGEN ION ACTIVITY.* Journal of Biological Chemistry, 1929. **81**(1): p. 175-203.
45. Pauling, L. and C.D. Coryell, *The Magnetic Properties and Structure of Hemoglobin, Oxyhemoglobin and Carbonmonoxyhemoglobin.* Proc Natl Acad Sci U S A, 1936. **22**(4): p. 210-6.
46. Pauling, L. and C.D. Coryell, *The Magnetic Properties and Structure of the Hemochromogens and Related Substances.* Proc Natl Acad Sci U S A, 1936. **22**(3): p. 159-63.
47. Savicki, J.P., G. Lang, and M. Ikeda-Saito, *Magnetic susceptibility of oxy- and carbonmonoxyhemoglobins.* Proc Natl Acad Sci U S A, 1984. **81**(17): p. 5417-9.
48. Perutz, M.F., et al., *Structure of Haemoglobin - 3-Dimensional Fourier Synthesis at 5.5-a Resolution, Obtained by X-Ray Analysis.* Nature, 1960. **185**(4711): p. 416-422.
49. Kendrew, J.C., et al., *3-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis.* Nature, 1958. **181**(4610): p. 662-666.
50. Kendrew, J.C., et al., *Structure of Myoglobin - 3-Dimensional Fourier Synthesis at 2 a Resolution.* Nature, 1960. **185**(4711): p. 422-427.
51. Shaanan, B., *Structure of human oxyhaemoglobin at 2.1 A resolution.* J Mol Biol, 1983. **171**(1): p. 31-59.
52. Fermi, G., et al., *The crystal structure of human deoxyhaemoglobin at 1.74 A resolution.* J Mol Biol, 1984. **175**(2): p. 159-74.
53. Pauling, L., *The Oxygen Equilibrium of Hemoglobin and Its Structural Interpretation.* Proc Natl Acad Sci U S A, 1935. **21**(4): p. 186-91.
54. Shaik, S. and H. Chen, *Lessons on O-2 and NO bonding to heme from ab initio multireference/multiconfiguration and DFT calculations.* Journal of Biological Inorganic Chemistry, 2011. **16**(6): p. 841-855.
55. Chen, H., M. Ikeda-Saito, and S. Shaik, *Nature of the Fe-O2 bonding in oxy-myoglobin: effect of the protein.* J Am Chem Soc, 2008. **130**(44): p. 14778-90.
56. Bren, K.L., R. Eisenberg, and H.B. Gray, *Discovery of the magnetic behavior of hemoglobin: A beginning of bioinorganic chemistry.* Proc Natl Acad Sci U S A, 2015. **112**(43): p. 13123-7.
57. Weiss, J.J., *NATURE OF IRON-OXYGEN BOND IN OXYHAEMOGLOBIN.* Nature, 1964. **202**(492S): p. 83-&.
58. McClure, D.S., *Electronic Structure of Transition-Metal Complex Ions.* Radiation Research Supplement, 1960. **2**: p. 218-242.
59. Goddard, W.A. and B.D. Olafson, *Ozone Model for Bonding of an O2 to Heme in Oxyhemoglobin.* Proceedings of the National Academy of Sciences of the United States of America, 1975. **72**(6): p. 2335-2339.

60. Olafson, B.D. and W.A. Goddard, *Molecular Description of Dioxygen Bonding in Hemoglobin.* Proceedings of the National Academy of Sciences of the United States of America, 1977. **74**(4): p. 1315-1319.
61. Yamamoto, T., et al., *Valence and Spin State of Iron in Oxyhemoglobin as Inferred from Resonance Raman Spectroscopy.* Journal of Biological Chemistry, 1973. **248**(14): p. 5211-5213.
62. Spiro, T.G. and T.C. Strekas, *Resonance Raman-Spectra of Heme Proteins - Effects of Oxidation and Spin State.* Journal of the American Chemical Society, 1974. **96**(2): p. 338-345.
63. Lang, G. and W. Marshall, *Mossbauer effect in some haemoglobin compounds.* J Mol Biol, 1966. **18**(3): p. 385-404.
64. Weissbluth, M. and J.E. Maling, *Interpretation of quadrupole splittings and isomer shifts in hemoglobin.* J Chem Phys, 1967. **47**(10): p. 4166-72.
65. Spartalian, K., G. Lang, and T. Yonetani, *Low temperature photodissociation studies of ferrous hemoglobin and myoglobin complexes by Mossbauer spectroscopy.* Biochim Biophys Acta, 1976. **428**(2): p. 281-90.
66. Loew, G.H. and R.F. Kirchner, *Electronic-Structure and Electric-Field Gradients in Oxyhemoglobin and Oxycytochrome P-450 Model Compounds.* Journal of the American Chemical Society, 1975. **97**(25): p. 7388-7390.
67. Kirchner, R.F. and G.H. Loew, *Semiempirical Calculations of Model Oxyheme - Variation of Calculated Electromagnetic Properties with Electronic Configuration and Oxygen Geometry.* Journal of the American Chemical Society, 1977. **99**(14): p. 4639-4647.
68. Herman, Z.S. and G.H. Loew, *Theoretical Investigation of the Magnetic and Ground-State Properties of Model Oxyhemoglobin Complexes.* Journal of the American Chemical Society, 1980. **102**(6): p. 1815-1821.
69. Newton, J.E. and M.B. Hall, *Generalized Molecular-Orbital Calculations on Transition-Metal Dioxygen Complexes - Models for Iron and Cobalt Porphyrins.* Inorganic Chemistry, 1984. **23**(26): p. 4627-4632.
70. Yamamoto, S. and H. Kashiwagi, *Casscf Study on the Fe-O2 Bond in a Dioxygen Heme Complex.* Chemical Physics Letters, 1989. **161**(1): p. 85-89.
71. Yamamoto, S. and H. Kashiwagi, *Casscf Calculation on Dioxygen Heme Complex with Extended Basis Set.* Chemical Physics Letters, 1993. **205**(2-3): p. 306-312.
72. Bytheway, I. and M.B. Hall, *Theoretical Calculations of Metal-Dioxygen Complexes.* Chemical Reviews, 1994. **94**(3): p. 639-658.
73. Nakatsuji, H., et al., *Ground and excited states of oxyheme: SAC/SAC-CI study.* Chemical Physics Letters, 1996. **250**(3-4): p. 379-386.
74. Rovira, C., et al., *Equilibrium geometries and electronic structure of iron-porphyrin complexes: A density functional study.* Journal of Physical Chemistry A, 1997. **101**(47): p. 8914-8925.
75. Rovira, C., *Role of the His64 residue on the properties of the Fe-CO and Fe-O-2 bonds in myoglobin. A CHARMM/DFT study.* Journal of Molecular Structure-Theochem, 2003. **632**: p. 309-321.
76. Jensen, K.P., B.O. Roos, and U. Ryde, *O-2-binding to heme: electronic structure and spectrum of oxyheme, studied by multiconfigurational methods.* Journal of Inorganic Biochemistry, 2005. **99**(1): p. 45-54.
77. Ribas-Arino, J. and J.J. Novoa, *The mechanism for the reversible oxygen addition to heme. A theoretical CASPT2 study.* Chemical Communications, 2007(30): p. 3160-3162.
78. Johnson, C.R., et al., *Oxygen Binding Constants and Stepwise Enthalpies for Human and Bovine Hemoglobin at Ph 7.6.* Biochemistry, 1992. **31**(41): p. 10074-10082.

79. Deslattes, R.D., et al., *X-ray transition energies: new approach to a comprehensive evaluation.* Reviews of Modern Physics, 2003. **75**(1): p. 35-99.

80. Thompson, A.C., *X-Ray Data Booklet*, A.C. Thompson, Editor. 2009, Lawrence Berkeley National Laboratory: http://xdb.lbl.gov/.

81. de Groot, F., *High resolution X-ray emission and X-ray absorption spectroscopy.* Chemical Reviews, 2001. **101**(6): p. 1779-1808.

82. Schrapers, P., et al., *Sulfido and Cysteine Ligation Changes at the Molybdenum Cofactor during Substrate Conversion by Formate Dehydrogenase (FDH) from Rhodobacter capsulatus.* Inorganic Chemistry, 2015. **54**(7): p. 3260-3271.

83. Kositzki, R., et al., *Protonation State of MnFe and FeFe Cofactors in a Ligand-Binding Oxidase Revealed by X-ray Absorption, Emission, and Vibrational Spectroscopy and QM/MM Calculations.* Inorg Chem, 2016. **55**(19): p. 9869-9885.

84. Schrapers, P., et al., *Axial Ligation and Redox Changes at the Cobalt Ion in Cobalamin Bound to Corrinoid Iron-Sulfur Protein (CoFeSP) or in Solution Characterized by XAS and DFT.* PLoS One, 2016. **11**(7): p. e0158681.

85. de Groot, F.M.F., *X-ray absorption and dichroism of transition metals and their compounds.* Journal of Electron Spectroscopy and Related Phenomena, 1994. **67**(4): p. 529-622.

86. Kawai, J., et al., *Charge-transfer effect on the linewidth of Fe K alpha x-ray fluorescence spectra.* Phys Rev B Condens Matter, 1994. **50**(16): p. 11347-11354.

87. Peng, G., et al., *High-resolution manganese x-ray fluorescence spectroscopy. Oxidation-state and spin-state sensitivity.* Journal of the American Chemical Society, 1994. **116**(7): p. 2914-2920.

88. Lambertz, C., et al., *Electronic and molecular structures of the [2Fe] and [4Fe4S] units of the active-site H-cluster in [FeFe]-hydrogenase determined by spin- and site-selective XAE and DFT.* Chem Sci, 2014. **5**: p. 1187-1203.

89. Leidel, N., et al., *Site-selective X-ray spectroscopy on an asymmetric model complex of the [FeFe] hydrogenase active site.* Inorg Chem, 2012. **51**(8): p. 4546-4559.

90. Lundberg, M., et al., *Metal-ligand covalency of iron complexes from high-resolution resonant inelastic X-ray scattering.* J Am Chem Soc, 2013. **135**(45): p. 17121-34.

91. Kroll, T., et al., *Resonant inelastic X-ray scattering on ferrous and ferric bis-imidazole porphyrin and cytochrome c: nature and role of the axial methionine-Fe bond.* J Am Chem Soc, 2014. **136**(52): p. 18087-99.

92. Shulman, G.R., et al., *Observations and interpretation of x-ray absorption edges in iron compounds and proteins.* Proceedings of the National Academy of Sciences, 1976. **73**(5): p. 1384-1388.

93. DeBeer, S., et al., *X-ray Absorption Edge and EXAFS Studies of the Blue Copper Site in Stellacyanin: Effects of Axial Amide Coordination.* The Journal of Physical Chemistry B, 2000. **104**(46): p. 10814-10819.

94. Glatzel, P. and A. Juhin, *X-ray Absorption and Emission Spectroscopy*, in *Local Structural Characterisation*. 2013, John Wiley & Sons, Ltd. p. 89-171.

95. Rehr, J.J., et al., *Parameter-free calculations of X-ray spectra with FEFF9.* Phys Chem Chem Phys, 2010. **12**(21): p. 5503-13.

96. Teo, B.K., *EXAFS: Baisc Principles and Data Analysis*. 1986: Springer-Verlag.

97. Bunker, G., *A Practical Guide to X-ray Absorption Fine Structure Spectroscopy*. 2010: Cambridge University Press, New York.

98. Kositzki, R., et al., *Electronic and molecular structure relations in diiron compounds mimicking the [FeFe]-hydrogenase active site studied by X-ray spectroscopy and quantum chemistry.* Dalton Transactions, 2017. **46**(37): p. 12544-12557.

99. Petuker, A., et al., *Spontaneous Si-C bond cleavage in (TriphosSi)-nickel complexes.* Dalton Trans, 2017. **46**(3): p. 907-917.

100. Schrapers, P., et al., *The binuclear nickel center in the A-cluster of acetyl-CoA synthase (ACS) and two biomimetic dinickel complexes studied by X-ray absorption and emission spectroscopy*. Journal of Physics: Conference Series, 2016. **712**(1): p. 012029.
101. Schuth, N., et al., *Biomimetic mono- and dinuclear Ni(I) and Ni(II) complexes studied by X-ray absorption and emission spectroscopy and quantum chemical calculations.* Journal of Physics: Conference Series, 2016. **712**(1): p. 012134.
102. Schuth, N. and M. Haumann. *X-ray spectroscopy reveals iron spin state in oxygenated hemoproteins*. 2017; Available from: http://www.esrf.eu/home/news/spotlight/content-news/spotlight/spotlight301.html.
103. Schuth, N. and M. Haumann, *X-ray spectroscopy reveals iron spin state in oxygenated hemoproteins*, in *ESRF Highlights 2017*. 2018, ESRF.
104. Schuth, N., et al., *Inhibitory and Non-Inhibitory NH3 Binding at the Water-Oxidizing Manganese Complex of Photosystem II Suggests Possible Sites and a Rearrangement Mode of Substrate Water Molecules.* Biochemistry, 2017. **56**(47): p. 6240-6256.
105. Schuth, N., et al., *Effective intermediate-spin iron in O2-transporting heme proteins.* Proc Natl Acad Sci U S A, 2017. **114**(32): p. 8556-8561.
106. Glatzel, P., et al., *Reflections on hard X-ray photon-in/photon-out spectroscopy for electronic structure studies.* Journal of Electron Spectroscopy and Related Phenomena, 2013. **188**: p. 17-25.
107. Vankó, G., et al., *Probing the 3d spin momentum with X-ray emission spectroscopy: the case of molecular-spin transitions.* J. Phys. Chem. B, 2006. **110**: p. 11647-11653.
108. de Groot, F., *High-Resolution X-ray Emission and X-ray Absorption Spectroscopy.* Chemical Reviews, 2001. **101**(6): p. 1779-1808.
109. Wang, X., F. De Groot, and S.P. Cramer, *Spin-polarized x-ray emission of 3d transition metal ions - a comparison via Kalpha and Kbeta detection.* Phys. Rev. B, 1997. **56**: p. 4553-4564.
110. Tsutsumi, K., H. Nakamori, and K. Ichikawa, *X-Ray Mn Kbeta Emission-Spectra of Manganese Oxides and Manganates*. Physical Review B, 1976. **13**(2): p. 929-933.
111. Peng, G., et al., *High-Resolution Manganese X-Ray-Fluorescence Spectroscopy - Oxidation-State and Spin-State Sensitivity*. Journal of the American Chemical Society, 1994. **116**(7): p. 2914-2920.
112. Glatzel, P., et al., *Site-selective EXAFS in mixed-valence compounds using high-resolution fluorescence detection: A study of iron in Prussian Blue.* Inorganic Chemistry, 2002. **41**(12): p. 3121-3127.
113. Glatzel, P., et al., *Influence of the core hole on K beta emission following photoionization or orbital electron capture: A comparison using MnO and (Fe2O3)-Fe-55*. Physical Review B, 2001. **64**(4).
114. Hamalainen, K., et al., *Spin-Dependent X-Ray Absorption of Mno and Mnf2*. Physical Review B, 1992. **46**(21): p. 14274-14277.
115. Zaharieva, I., et al., *Towards a comprehensive X-ray approach for studying the photosynthetic manganese complex—XANES, K-alpha/K-beta/K-beta-satellite emission lines, RIXS, and comparative computational approaches for selected model complexes.* Journal of Physics: Conference Series, 2009. **190**: p. 012142, 1-6.
116. Sakurai, K., et al., *Trace chemical characterization using monochromatic X-ray undulator radiation.* Analytical Chemistry, 2000. **72**(11): p. 2613-2617.
117. Degroot, F.M.F., et al., *Charge-Transfer Multiplet Calculations of the K-Beta X-Ray-Emission Spectra of Divalent Nickel Compounds*. Journal of Physics-Condensed Matter, 1994. **6**(34): p. 6875-6884.
118. Leidel, N., et al., *Bridging-hydride influence on the electronic structure of an [FeFe] hydrogenase active-site model complex revealed by XAES-DFT.* Dalton Trans, 2013. **42**(21): p. 7539-54.

119. Leidel, N., et al., *Electronic structure of an [FeFe] hydrogenase model complex in solution revealed by X-ray absorption spectroscopy using narrow-band emission detection.* J Am Chem Soc, 2012. **134**(34): p. 14142-57.
120. Bergmann, U. and P. Glatzel, *X-ray emission spectroscopy.* Photosynthesis Research, 2009. **102**(2-3): p. 255-266.
121. Zaharieva, I., et al., *Room-Temperature Energy-Sampling Kbeta X-ray Emission Spectroscopy of the Mn4Ca Complex of Photosynthesis Reveals Three Manganese-Centered Oxidation Steps and Suggests a Coordination Change Prior to O2 Formation.* Biochemistry, 2016. **55**(30): p. 4197-211.
122. Yao, S., et al., *Biomimetic [2Fe-2S] clusters with extensively delocalized mixed-valence iron centers.* Angew Chem Int Ed Engl, 2015. **54**(42): p. 12506-10.
123. Huwald, D., et al., *Characterization of unusual truncated hemoglobins of Chlamydomonas reinhardtii suggests specialized functions.* Planta, 2015. **242**(1): p. 167-85.
124. Dedieu, A., et al., *Oxygen Binding to Iron Porphyrins - Ab-Initio Calculation.* Journal of the American Chemical Society, 1976. **98**(12): p. 3717-3818.
125. Mardirossian, N. and M. Head-Gordon, *Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals.* Molecular Physics, 2017. **115**(19): p. 2315-2372.
126. Dirac, P.A.M., *Note on Exchange Phenomena in the Thomas Atom.* Mathematical Proceedings of the Cambridge Philosophical Society, 1930. **26**(03): p. 376-385.
127. Slater, J.C., *Wave Functions in a Periodic Potential.* Physical Review, 1937. **51**(10): p. 846-851.
128. Slater, J.C., *A Simplification of the Hartree-Fock Method.* Physical Review, 1951. **81**(3): p. 385-390.
129. Hohenberg, P. and W. Kohn, *Inhomogeneous Electron Gas.* Physical Review B, 1964. **136**(3b): p. B864-+.
130. Salahub, D.R., et al., *Chapter 38 - Applied density functional theory and the deMon codes 1964–2004 A2 - Dykstra, Clifford E*, in *Theory and Applications of Computational Chemistry*, G. Frenking, K.S. Kim, and G.E. Scuseria, Editors. 2005, Elsevier: Amsterdam. p. 1079-1097.
131. Kohn, W. and L.J. Sham, *Self-Consistent Equations Including Exchange and Correlation Effects.* Physical Review, 1965. **140**(4a): p. 1133-&.
132. Grimme, S., *Semiempirical hybrid density functional with perturbative second-order correlation.* The Journal of Chemical Physics, 2006. **124**(3): p. 034108.
133. Maier, T.M., et al., *Validation of local hybrid functionals for TDDFT calculations of electronic excitation energies.* The Journal of Chemical Physics, 2016. **144**(7): p. 074106.
134. Becke, A.D., *Density-functional exchange-energy approximation with correct asymptotic behavior.* Phys Rev A Gen Phys, 1988. **38**(6): p. 3098-3100.
135. Perdew, J.P., *Erratum: Density-functional approximation for the correlation energy of the inhomogeneous electron gas.* Phys Rev B Condens Matter, 1986. **34**(10): p. 7406.
136. Perdew, J.P., *Density-functional approximation for the correlation energy of the inhomogeneous electron gas.* Phys Rev B Condens Matter, 1986. **33**(12): p. 8822-8824.
137. Lee, C., W. Yang, and R.G. Parr, *Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density.* Physical Review B, 1988. **37**(2): p. 785-789.
138. Stephens, P.J., et al., *Ab-Initio Calculation of Vibrational Absorption and Circular-Dichroism Spectra Using Density-Functional Force-Fields.* Journal of Physical Chemistry, 1994. **98**(45): p. 11623-11627.

139. Becke, A.D., *A New Mixing of Hartree-Fock and Local Density-Functional Theories.* Journal of Chemical Physics, 1993. **98**(2): p. 1372-1377.

140. Santra, B., A. Michaelides, and M. Scheffler, *On the accuracy of density-functional theory exchange-correlation functionals for H bonds in small water clusters: Benchmarks approaching the complete basis set limit.* The Journal of Chemical Physics, 2007. **127**(18): p. 184104.

141. Goerigk, L. and S. Grimme, *A thorough benchmark of density functional methods for general main group thermochemistry, kinetics, and noncovalent interactions.* Physical Chemistry Chemical Physics, 2011. **13**(14): p. 6670-6688.

142. Laurent, A.D. and D. Jacquemin, *TD-DFT benchmarks: A review.* International Journal of Quantum Chemistry, 2013. **113**(17): p. 2019-2039.

143. Ruiz, E., et al., *Toward the Prediction of Magnetic Coupling in Molecular Systems: Hydroxo- and Alkoxo-Bridged Cu(II) Binuclear Complexes.* Journal of the American Chemical Society, 1997. **119**(6): p. 1297-1303.

144. Noodleman, L., *Valence bond description of antiferromagnetic coupling in transition metal dimers.* The Journal of Chemical Physics, 1981. **74**(10): p. 5737-5743.

145. Caballol, R., et al., *Remarks on the proper use of the broken symmetry approach to magnetic coupling.* Journal of Physical Chemistry A, 1997. **101**(42): p. 7860-7866.

146. Hilborn, R.C., *Einstein Coefficients, Cross-Sections, F Values, Dipole-Moments, and All That.* American Journal of Physics, 1982. **50**(11): p. 982-986.

147. Smolentsev, G., et al., *X-ray Emission Spectroscopy To Study Ligand Valence Orbitals in Mn Coordination Complexes.* Journal of the American Chemical Society, 2009. **131**(36): p. 13161-13167.

148. Ray, K., et al., *Description of the ground-state covalencies of the bis(dithiolato) transition-metal complexes from X-ray absorption spectroscopy and time-dependent density-functional calculations.* Chemistry-a European Journal, 2007. **13**(10): p. 2783-2797.

149. Dreuw, A. and M. Head-Gordon, *Single-reference ab initio methods for the calculation of excited states of large molecules.* Chemical Reviews, 2005. **105**(11): p. 4009-4037.

150. Runge, E. and E.K.U. Gross, *Density-Functional Theory for Time-Dependent Systems.* Physical Review Letters, 1984. **52**(12): p. 997-1000.

151. Chandrasekaran, P., et al., *Prediction of high-valent iron K-edge absorption spectra by time-dependent density functional theory.* Dalton Trans, 2011. **40**(42): p. 11070-9.

152. Radon, M., E. Broclawik, and K. Pierloot, *Electronic Structure of Selected {FeNO}(7) Complexes in Heme and Non-Heme Architectures: A Density Functional and Multireference ab Initio Study.* Journal of Physical Chemistry B, 2010. **114**(3): p. 1518-1528.

153. Neese, F., *A critical evaluation of DFT, including time-dependent DFT, applied to bioinorganic chemistry.* J Biol Inorg Chem, 2006. **11**(6): p. 702-11.

154. Larsson, S., *Applications of CASSCF.* International Journal of Quantum Chemistry, 2011. **111**(13): p. 3424-3430.

155. Roos, B.O., *Chapter 25 - Multiconfigurational quantum chemistry A2 - Dykstra, Clifford E,* in *Theory and Applications of Computational Chemistry*, G. Frenking, K.S. Kim, and G.E. Scuseria, Editors. 2005, Elsevier: Amsterdam. p. 725-764.

156. Blanksby, S.J. and G.B. Ellison, *Bond Dissociation Energies of Organic Molecules.* Accounts of Chemical Research, 2003. **36**(4): p. 255-263.

157. Takatsuka, K., T. Fueno, and K. Yamaguchi, *Distribution of Odd Electrons in Ground-State Molecules.* Theoretica Chimica Acta, 1978. **48**(3): p. 175-183.

158. Bochicchio, R.C., *On spin density and hole distribution relations: valence and free valence.* Journal of Molecular Structure: THEOCHEM, 1998. **429**: p. 229-236.

159. Staroverov, V.N. and E.R. Davidson, *Distribution of effectively unpaired electrons.* Chemical Physics Letters, 2000. **330**(1-2): p. 161-168.
160. Lobayan, R.M. and R.C. Bochicchio, *Pairing and unpairing electron densities in organic systems: Two-electron three center through space and through bonds interactions.* Journal of Chemical Physics, 2014. **140**(17).
161. Staroverov, V.N. and E.R. Davidson, *Diradical Character of the Cope Rearrangement Transition State.* Journal of the American Chemical Society, 2000. **122**(1): p. 186-187.
162. Lobayan, R.M., et al., *Topology of the Effectively Paired and Unpaired Electron Densities for Complex Bonding Patterns: The Three-Center Two-Electron Bonding Case.* Journal of Chemical Theory and Computation, 2009. **5**(8): p. 2030-2043.
163. Head-Gordon, M., *Reply to comment on 'characterizing unpaired electrons from the one-particle density matrix'.* Chemical Physics Letters, 2003. **380**(3–4): p. 488-489.
164. Angeli, C., M. Pastore, and R. Cimiraglia, *New perspectives in multireference perturbation theory: the n-electron valence state approach.* Theoretical Chemistry Accounts, 2007. **117**(5): p. 743-754.
165. Neese, F., *The ORCA program system.* Wiley Interdisciplinary Reviews-Computational Molecular Science, 2012. **2**(1): p. 73-78.
166. Frisch, M.J., et al., *Gaussian 09*. 2009: Gaussian Inc. Wallingford CT 2009.
167. Schafer, A., H. Horn, and R. Ahlrichs, *Fully Optimized Contracted Gaussian-Basis Sets for Atoms Li to Kr.* Journal of Chemical Physics, 1992. **97**(4): p. 2571-2577.
168. Sinnecker, S., et al., *Calculation of solvent shifts on electronic g-tensors with the conductor-like screening model (COSMO) and its self-consistent generalization to real solvents (Direct COSMO-RS).* Journal of Physical Chemistry A, 2006. **110**(6): p. 2235-2245.
169. Neese, F., *An improvement of the resolution of the identity approximation for the formation of the Coulomb matrix.* J Comput Chem, 2003. **24**(14): p. 1740-7.
170. Izsak, R. and F. Neese, *An overlap fitted chain of spheres exchange method.* Journal of Chemical Physics, 2011. **135**(14).
171. Weigend, F. and R. Ahlrichs, *Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy.* Physical Chemistry Chemical Physics, 2005. **7**(18): p. 3297-3305.
172. Lowdin, P.O., *Quantum Theory of Many-Particle Systems .2. Study of the Ordinary Hartree-Fock Approximation.* Physical Review, 1955. **97**(6): p. 1490-1508.
173. Weinhold, F. and C.R. Landis, *NATURAL BOND ORBITALS AND EXTENSIONS OF LOCALIZED BONDING CONCEPTS.* Chemistry Education Research and Practice, 2001. **2**(2): p. 91-104.
174. Magnasco, V. and A. Perico, *Uniform Localization of Atomic and Molecular Orbitals .I.* Journal of Chemical Physics, 1967. **47**(3): p. 971-&.
175. Reed, A.E., L.A. Curtiss, and F. Weinhold, *Intermolecular Interactions from a Natural Bond Orbital, Donor-Acceptor Viewpoint.* Chemical Reviews, 1988. **88**(6): p. 899-926.
176. Boys, S.F., *Construction of Some Molecular Orbitals to Be Approximately Invariant for Changes from One Molecule to Another.* Reviews of Modern Physics, 1960. **32**(2): p. 296-299.
177. Edmiston, C. and K. Ruedenberg, *Localized Atomic and Molecular Orbitals.* Reviews of Modern Physics, 1963. **35**(3): p. 457-&.
178. Pipek, J. and P.G. Mezey, *A Fast Intrinsic Localization Procedure Applicable for Abinitio and Semiempirical Linear Combination of Atomic Orbital Wave-Functions.* Journal of Chemical Physics, 1989. **90**(9): p. 4916-4926.
179. Warkentin, M. and R.E. Thorne, *Glass transition in thaumatin crystals revealed through temperature-dependent radiation-sensitivity measurements.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 10): p. 1092-100.

180. Kositzki, R., et al., *y Protonation State of MnFe and FeFe Cofactors in a Ligand-Binding Oxidase Revealed by X-ray Absorption, Emission, and Vibrational Spectroscopy and QM/MM Calculations.* Inorganic Chemistry, 2016. **55**(19): p. 9869-9885.
181. Schrapers, P., et al., *'Sulfido and Cysteine Ligation Changes at the Molybdenum Cofactor during Substrate Conversion by Formate Dehydrogenase (FDH) from Rhodobacter capsulatus.* Inorganic Chemistry, 2015. **54**(7): p. 3260-3271.
182. Reeder, B.J., et al., *The radical and redox chemistry of myoglobin and hemoglobin: from in vitro studies to human pathology.* Antioxid Redox Signal, 2004. **6**(6): p. 954-66.
183. Bewley, K.D., et al., *Multi-heme proteins: nature's electronic multi-purpose tool.* Biochim Biophys Acta, 2013. **1827**(8-9): p. 938-48.
184. Weber, R.E. and S.N. Vinogradov, *Nonvertebrate hemoglobins: functions and molecular adaptations.* Physiol Rev, 2001. **81**(2): p. 569-628.
185. Frey, A.D. and P.T. Kallio, *Bacterial hemoglobins and flavohemoglobins: versatile proteins and their impact on microbiology and biotechnology.* FEMS Microbiol Rev, 2003. **27**(4): p. 525-45.
186. Hoy, J.A. and M.S. Hargrove, *The structure and function of plant hemoglobins.* Plant Physiol Biochem, 2008. **46**(3): p. 371-9.
187. Brunori, M., *Hemoglobin is an honorary enzyme.* Trends in Biochemical Sciences, 1999. **24**(4): p. 158-161.
188. Weissbluth, M., *Hemoglobin: cooperativity and electronic properties*. 1975, Berlin: Springer.
189. Pauling, L., *Magnetic properties and structure of oxyhemoglobin.* Proc Natl Acad Sci U S A, 1977. **74**(7): p. 2612-3.
190. Harcourt, R.D., *Transition metal complexes with CO, N2, NO and O2 ligands*, in *Bonding in Electron-Rich Molecules*, R.D. Harcourt, Editor. 2016, Springer: New York. p. 231-246.
191. Kepp, K.P., *Consistent descriptions of metal-ligand bonds and spin-crossover in inorganic chemistry.* Coordination Chemistry Reviews, 2013. **257**(1): p. 196-209.
192. Yonetani, T. and K. Kanaori, *How does hemoglobin generate such diverse functionality of physiological relevance?* Biochim Biophys Acta, 2013. **1834**(9): p. 1873-84.
193. Shikama, K., *Nature of the FeO2 bonding in myoglobin and hemoglobin: A new molecular paradigm.* Prog Biophys Mol Biol, 2006. **91**(1-2): p. 83-162.
194. Perutz, M.F., *Myoglobin and haemoglobin: role of distal residues in reactions with haem ligands.* Trends Biochem Sci, 1989. **14**(2): p. 42-4.
195. De Groot, F. and A. Kotani, *Core Level Spectroscopy of Solids*. 2008, Boca Raton, FL, USA: Taylor & Francis CRC press.
196. van Bokhoven, J.A. and C. Lamberti, *X-Ray Absorption and X-Ray Emission Spectroscopy: Theory and Applications*. 2016, New York: Wiley.
197. Pollock, C.J. and S. DeBeer, *Insights into the Geometric and Electronic Structure of Transition Metal Centers from Valence-to-Core X-ray Emission Spectroscopy.* Accounts of Chemical Research, 2015. **48**(11): p. 2967-2975.
198. Lehnert, N., S.D. George, and E.I. Solomon, *Recent advances in bioinorganic spectroscopy.* Current Opinion in Chemical Biology, 2001. **5**(2): p. 176-187.
199. deGroot, F.M.F., *Site-Selective XAFS: a new tool for catalysis research.* Topics in Catalysis: Science Publishers, 2000: p. 179-186.
200. Pin, S., et al., *X-ray absorption spectroscopy of hemoglobin.* Methods Enzymol, 1994. **232**: p. 266-92.

201. Lima, F.A., et al., *Probing the electronic and geometric structure of ferric and ferrous myoglobins in physiological solutions by Fe K-edge absorption spectroscopy.* Phys Chem Chem Phys, 2014. **16**(4): p. 1617-31.
202. Eisenberger, P., et al., *Structure-function relations in hemoglobin as determined by x-ray absorption spectroscopy.* Proc Natl Acad Sci U S A, 1976. **73**(2): p. 491-5.
203. Durham, P., et al., *X-ray absorption near edge structure (XANES) for CO, CN and deoxyhaemoglobin: geometrical information.* EMBO J, 1983. **2**(9): p. 1441-3.
204. Zhang, K., et al., *Active site conformation in myoglobin as determined by X-ray absorption spectroscopy.* Proteins, 1991. **10**(4): p. 279-86.
205. Hersleth, H.P. and K.K. Andersson, *How different oxidation states of crystalline myoglobin are influenced by X-rays.* Biochim Biophys Acta, 2011. **1814**(6): p. 785-96.
206. Feis, A., et al., *Spin state and axial ligand bonding in the hydroxide complexes of metmyoglobin, methemoglobin, and horseradish peroxidase at room and low temperatures.* Biochemistry, 1994. **33**(15): p. 4577-83.
207. Di Iorio, E.E., *Preparation of derivatives of ferrous and ferric hemoglobin.* Methods Enzymol, 1981. **76**: p. 57-72.
208. Gouterman, M., *Study of the Effects of Substitution on the Absorption Spectra of Porphin.* Journal of Chemical Physics, 1959. **30**(5): p. 1139-1161.
209. Scheidt, W.R., S.R. Osvath, and Y.J. Lee, *Crystal and Molecular-Structure of Bis(Imidazole)(Meso-Tetraphenylporphinato)Iron(Iii) Chloride - a Classic Molecule Revisited.* Journal of the American Chemical Society, 1987. **109**(7): p. 1958-1963.
210. Ratanasopa, K., et al., *Dissection of the radical reactions linked to fetal hemoglobin reveals enhanced pseudoperoxidase activity.* Front Physiol, 2015. **6**: p. 39.
211. Vanko, G., et al., *Probing the 3d spin momentum with X-ray emission spectroscopy: The case of molecular-spin transitions.* Journal of Physical Chemistry B, 2006. **110**(24): p. 11647-11653.
212. Weik, M., et al., *Specific chemical and structural damage to proteins produced by synchrotron radiation.* Proceedings of the National Academy of Sciences of the United States of America, 2000. **97**(2): p. 623-628.
213. Liu, B., et al., *Structural changes that occur upon photolysis of the Fe(II)(a3)-CO complex in the cytochrome ba(3)-oxidase of Thermus thermophilus: A combined X-ray crystallographic and infrared spectral study demonstrates CO binding to Cu-B.* Biochimica Et Biophysica Acta-Bioenergetics, 2012. **1817**(4): p. 658-665.
214. Staroverov, V.N., et al., *Comparative assessment of a new nonempirical density functional: Molecules and hydrogen-bonded complexes.* Journal of Chemical Physics, 2003. **119**(23): p. 12129-12137.
215. Tao, J.M., et al., *Climbing the density functional ladder: Nonempirical meta-generalized gradient approximation designed for molecules and solids.* Physical Review Letters, 2003. **91**(14).
216. Cohen, A.J. and Y. Tantirungrotechai, *Molecular electric properties: an assessment of recently developed functionals.* Chemical Physics Letters, 1999. **299**(5): p. 465-472.
217. Neese, F., *Prediction and interpretation of the 57Fe isomer shift in Mössbauer spectra by density functional theory.* Inorganica Chimica Acta, 2002. **337**(Supplement C): p. 181-192.
218. Groenhof, A.R., et al., *Electronic ground states of iron porphyrin and of the first species in the catalytic reaction cycle of cytochrome P450s.* Journal of Physical Chemistry A, 2005. **109**(15): p. 3411-3417.
219. Steiner, T., *The Hydrogen Bond in the Solid State.* Angewandte Chemie International Edition, 2002. **41**(1): p. 48-76.

220. Jensen, K.P., B.O. Roos, and U. Ryde, *O-2-binding to heme: electronic structure and spectrum of oxyheme, studied by multiconfigurational methods (vol 99, pg 45, 2004).* Journal of Inorganic Biochemistry, 2005. **99**(4): p. 978-978.
221. Della Longa, S. and A. Arcovito, *X-Ray-Induced Lysis of the Fe-O Bond in Carbonmonoxy-Myoglobin.* Inorganic Chemistry, 2010. **49**(21): p. 9958-9961.
222. Xu, G.Z. and M.R. Chance, *Radiolytic modification of acidic amino acid residues in peptides: Probes for examining protein-protein interactions.* Analytical Chemistry, 2004. **76**(5): p. 1213-1221.
223. Brown, W.E., J.W. Sutcliffe, and P.D. Pulsinelli, *Multiple Internal Reflectance Infrared-Spectra of Variably Hydrated Hemoglobin and Myoglobin Films - Effects of Globin Hydration on Ligand Conformer Dynamics and Reactivity at the Heme.* Biochemistry, 1983. **22**(12): p. 2914-2923.
224. Demmel, F., et al., *Vibrational frequency shifts as a probe of hydrogen bonds: thermal expansion and glass transition of myoglobin in mixed solvents.* European Biophysics Journal with Biophysics Letters, 1997. **26**(4): p. 327-335.
225. Piruzian, L.A., et al., *The Dependence of Magnetic-Susceptibility of Oxyhemoglobin and Carboxyhemoglobin in Erythrocytes on Temperature.* Izvestiya Akademii Nauk Sssr Seriya Biologicheskaya, 1984(6): p. 894-900.
226. Kent, T.A., et al., *High Magnetic-Field Mossbauer Studies of Deoxymyoglobin, Deoxyhemoglobin, and Synthetic Analogs.* Biochimica Et Biophysica Acta, 1979. **580**(2): p. 245-258.
227. Ikeue, T., et al., *Saddle-shaped six-coordinate iron(III) porphyrin complexes showing a novel spin crossover between S=1/2 and S=3/2 spin states.* Angewandte Chemie-International Edition, 2001. **40**(14): p. 2617-2620.

9 Appendix

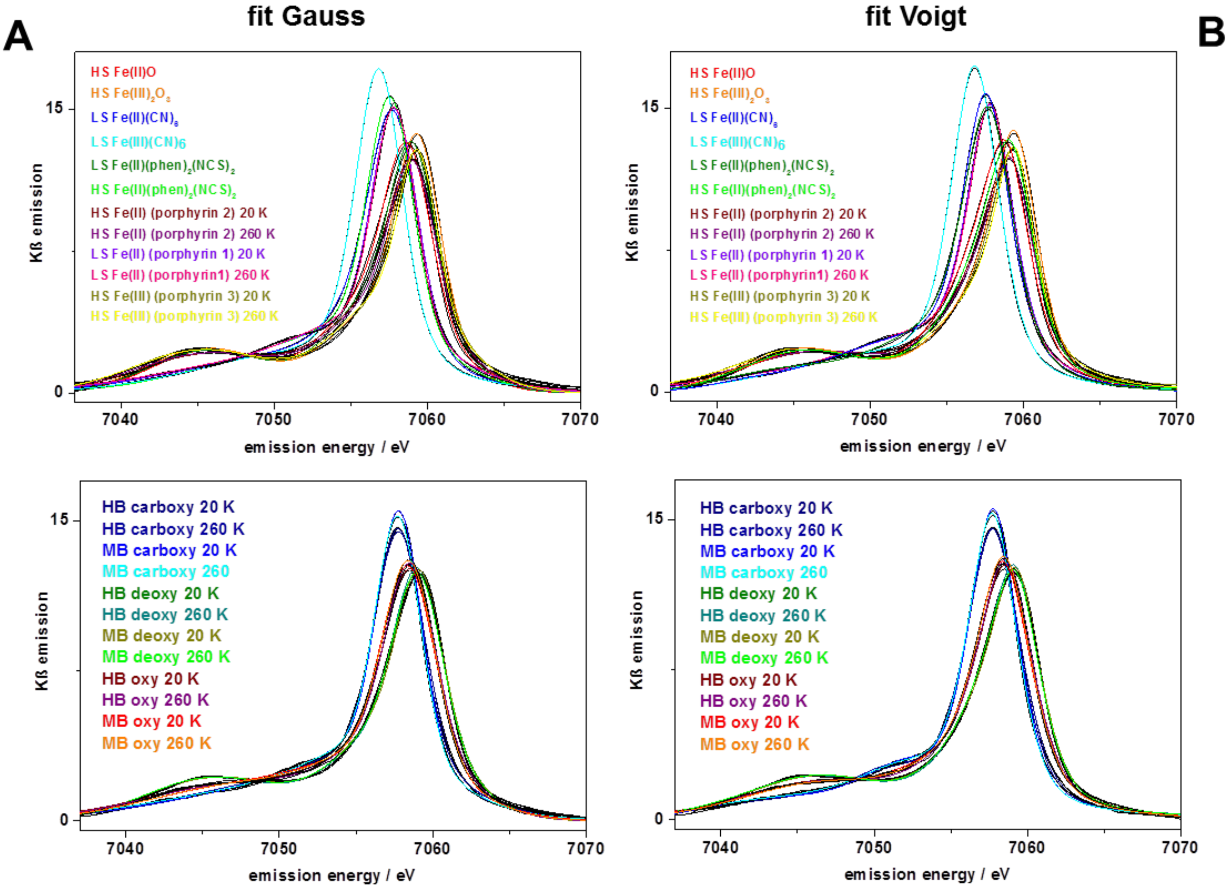


Figure 9-1: Fit analysis of Kβ main-line spectra. Top panels: spectra of porphyrin and non-porphyrin compounds. Bottom panels: spectra of MB and HB. (A) Experimental spectra (black lines) and fits with three Gaussian functions with variable band widths (colored lines). (B) Experimental spectra (black lines) and fits with three Voigt functions with variable band widths (50 % Gaussian and Lorentzian characters) and a linear baseline.

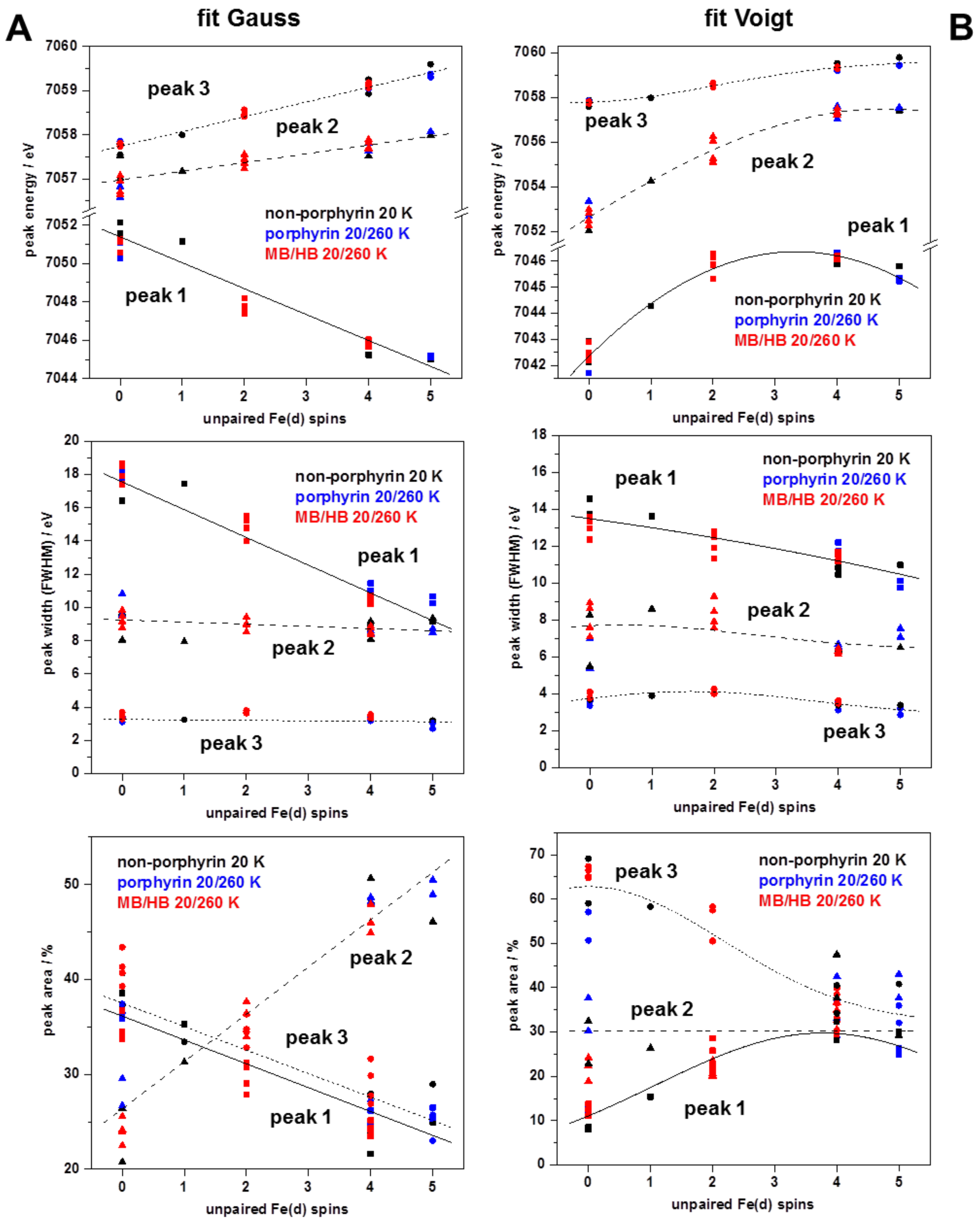


*Figure 9-2: Fit results of Kβ spectra. Data correspond to fit curves in Figure 9-1. Left: data from fits with Gaussian functions. Right: data from fits with Voigt functions. Top panels, line peak energies; middle panels, peak widths; bottom panels, integral peak areas. Lines show linear fits (left panels) or Gaussian fits (right panels) to the data for guiding the eye. Heme species in MB/HB were placed at Fe(d) spin counts of 4 (**deoxy**), 0 (**carboxy**), or 2 (**oxy**).*

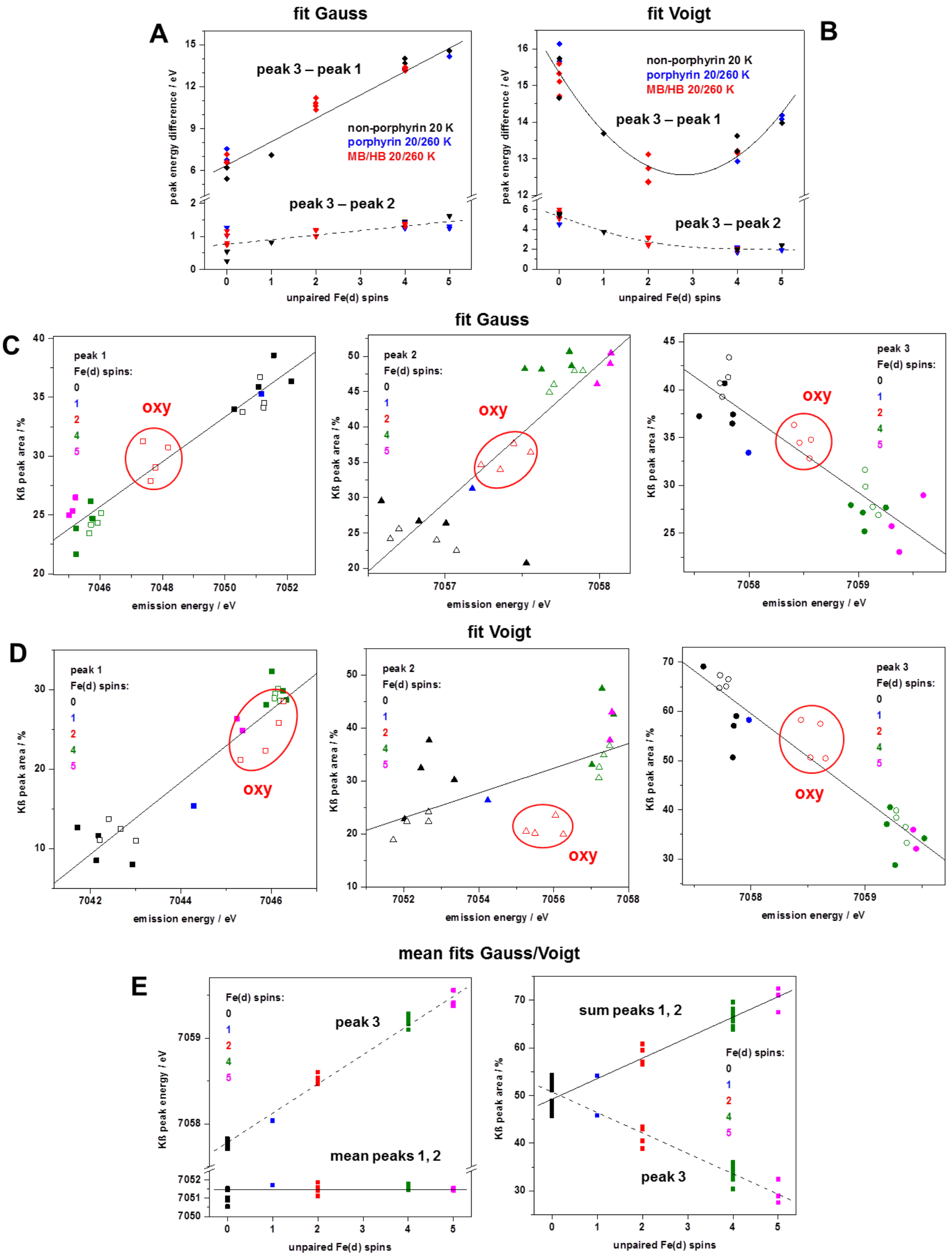


Figure 9-3: Correlations from Kβ fit analysis. Data correspond to Figure 9-1 and Figure 9-2. (A and B) Kβ line energy differences (lines show linear, left, or Gaussian, right, fits). (C and D) Kβ peak areas vs. center energies (lines show linear fits) for the indicated (formal) number of unpaired Fe(d) spins in the porphyrin/non-porphyrin compounds and HB/MB; data for ***oxy*** *(red circles) was placed at a Fe(d) spin count of 2. (E) Combined Kβ peak energies (mean (left) or sum (right) of peaks 1/2 and peak 3) plotted vs. unpaired Fe(d) spin counts (lines show linear fits). The combined Gauss/Voigt data reproduce relative energies/intensities of* $K\beta'/K\beta_{1,3}$ *features as obtained from visual inspection of Kβ spectra or Kβ′ amplitude determination (Figure 5-10 & Figure 5-11).*

Table 9-1: Parameters of CAS(8,7) calculations with varying Fe – O_2 bond length.

oxy	configuration		Fe-O_2 bond length [Å]					
MO no.	4 5 6 7 8 9 11	1.700	1.750	1.800	1.850	1.875	1.900	1.950
		contribution to CAS wavefunction [%]						
CASSCF states	2 2 2 2 0 0 0	69.7	67.5	65.3	63.4	62.4	61.6	59.9
	2 2 2 0 2 0 0	24.9	27.3	29.5	31.6	32.5	33.4	35.1
	1 2 1 2 1 1 0	1.3	1.3	1.3	1.3	1.3	1.3	1.3
	2 2 0 2 0 2 0	0.8	0.8	0.8	0.8	0.7	0.7	0.7
	1 2 1 1 2 1 0	0.6	0.6	0.7	0.7	0.7	0.7	0.7
	2 0 2 2 0 0 2	0.4	0.4	0.4	0.4	0.4	0.4	0.4
	2 2 0 0 2 2 0	0.3	0.3	0.4	0.4	0.4	0.4	0.4
		spin density/occupation number						
MO no.	4	1.972	1.973	1.973	1.974	1.974	1.974	1.975
	5	1.973	1.975	1.976	1.976	1.977	1.977	1.977
	6	1.950	1.950	1.951	1.951	1.951	1.951	1.951
	7	1.476	1.428	1.382	1.341	1.322	1.304	1.269
	8	0.551	0.599	0.643	0.684	0.703	0.721	0.755
	9	0.051	0.051	0.051	0.051	0.051	0.050	0.050
	11	0.027	0.025	0.024	0.023	0.023	0.023	0.022
		effective unpaired electron density (EUED)						
atom	Fe	1.625	1.705	1.773	1.829	1.854	1.875	1.912
	$O_{central}$	0.307	0.300	0.296	0.293	0.292	0.291	0.289
	$O_{terminal}$	0	0	0	0	0	0	0
	N_{Porph}	0	0	0	0	0	0	0

Orbitals in CASSCF active space are shown in Figure 9-4. State configuration refers to double-, single- or non-occupied MOs in active space. EUED is assigned to the atom with largest AO contribution in a CAS orbital (see colored MO numbers in Figure 5-28).

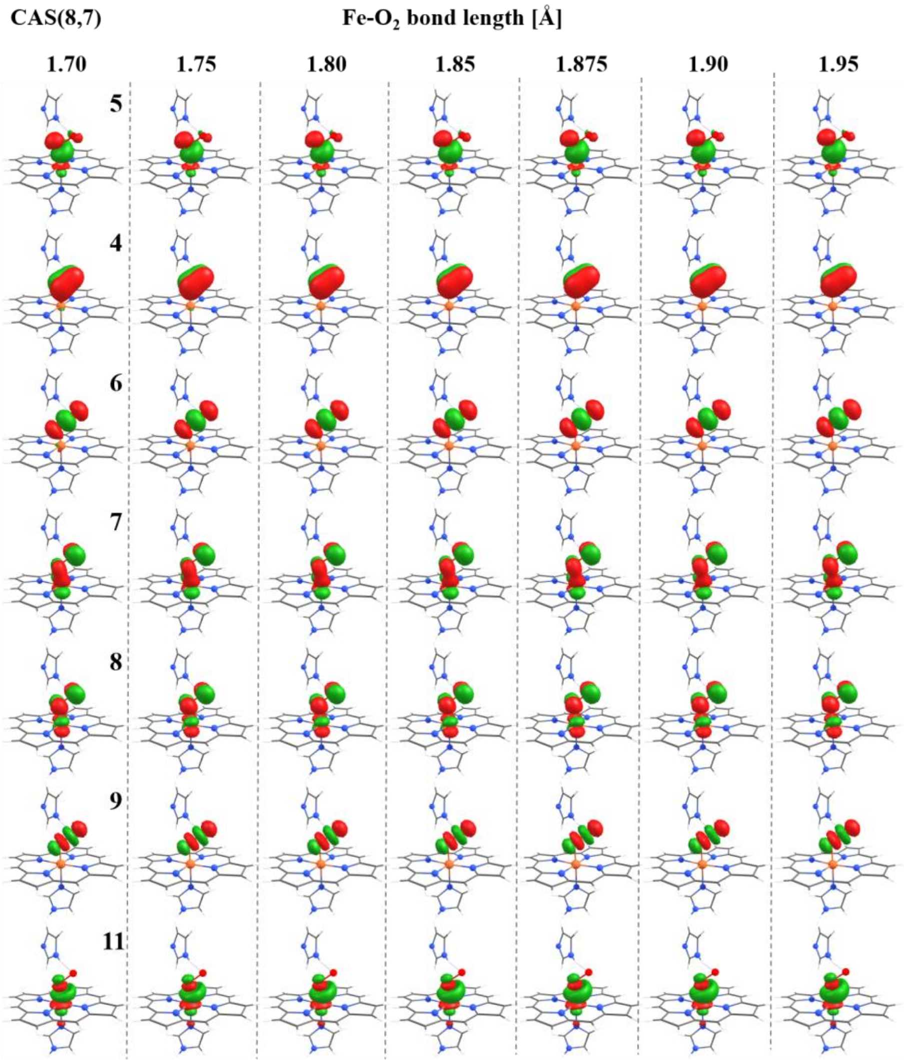


Figure 9-4: NMOs correspond to CAS(8,7) calculations for oxy after geometry-optimization (B3LYP/def2-TZVP) of structures with an increasing (and fixed) Fe – O_2 bond length. Orbitals are ordered by occupation number and named according to order of corresponding orbitals in CAS(14,11) calculations (see Figure 9-5). For corresponding CASSCF parameters see Table 9-1. Isosurfaces are drawn at ± 0.05 au.

Table 9-2: Parameters of CAS(14,11) calculations with varying Fe – O_2 bond length.

oxy	configuration		Fe-O_2 bond length [Å]					
MO no.	1 2 3 4 5 6 7 8 9 10 11	1.700	1.750	1.800	1.850	1.875	1.900	1.950
		contribution to CAS wavefunction [%]						
CASSCF states	2 2 2 2 2 2 2 0 0 0 0	71.0	69.0	66.7	64.6	63.6	62.6	60.7
	2 2 2 2 2 2 0 2 0 0 0	19.1	21.6	24.3	26.7	27.9	29.0	31.2
	2 2 2 1 2 1 2 1 1 0 0	1.0	1.2	1.0	1.1	1.2	1.2	1.2
	2 2 2 2 2 0 2 0 2 0 0	0.5	0.6	0.8	0.6	0.7	0.7	0.7
	2 2 2 1 2 1 1 2 1 0 0	0.4	0.5	0.5	0.5	0.6	0.6	0.7
	2 1 1 2 2 2 2 1 0 1 0	<0.3	0.4	0.4	0.4	0.5	0.5	0.5
		spin density/occupation number						
MO no.	1	1.984	1.985	1.987	1.988	1.989	1.989	1.990
	2	1.983	1.983	1.980	1.980	1.980	1.980	1.981
	3	1.975	1.975	1.975	1.975	1.975	1.975	1.975
	4	1.967	1.968	1.970	1.971	1.972	1.972	1.973
	5	1.958	1.960	1.964	1.966	1.968	1.969	1.971
	6	1.946	1.947	1.948	1.949	1.949	1.950	1.950
	7	1.564	1.516	1.465	1.417	1.394	1.371	1.329
	8	0.478	0.525	0.576	0.623	0.646	0.668	0.710
	9	0.056	0.054	0.053	0.053	0.052	0.052	0.051
	10	0.046	0.044	0.049	0.048	0.046	0.047	0.043
	11	0.043	0.041	0.032	0.030	0.030	0.028	0.028
		effective unpaired electron density (EUED)						
atom	Fe	1.667	1.756	1.852	1.927	1.957	1.989	2.032
	$O_{central}$	0.361	0.348	0.335	0.325	0.319	0.315	0.307
	$O_{terminal}$	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N_{Porph}	0.031	0.030	0.025	0.023	0.023	0.022	0.020

Orbitals in CASSCF active space are shown in Figure 9-5. State configuration refers to double-, single- or non-occupied MOs in active space. EUED is assigned to the atom with largest AO contribution in a CAS orbital (see colored MO numbers in Figure 5-28).

Figure 9-5: NMOs correspond to CAS(14,11) calculations for oxy after geometry-optimization (B3LYP/def2-TZVP) of structures with an increasing (and fixed) Fe – O$_2$ bond length. NMOs are ordered/named by occupation number For corresponding CASSCF parameters see Table 9-2. Isosurfaces are drawn at ± 0.05 au.

Table 9-3: Löwdin orbital composition from NMOs in the active space of CAS(14,11) for ***oxy****.*

oxy	atom																
	Fe							$O_{central}$					$O_{terminal}$				
MO No.	s	p	d					s	p			d	s	p			d
			z^2	x^2-y^2	xz	yz	xy		z	x	y			z	x	y	
1	0	0	0	26	0	0	1	0	1	0	0	0	0	0	0	0	0
2	1	0	4	0	1	1	86	0	2	0	1	0	0	0	0	0	0
3	0	0	1	0	50	36	5	1	1	1	2	0	0	0	1	0	0
4	0	0	0	0	2	1	0	0	2	21	35	1	0	0	12	23	1
5	1	3	9	0	4	5	3	4	44	9	11	0	0	0	2	1	1
6	0	0	1	0	1	0	0	0	17	22	8	1	0	16	22	9	1
7	0	0	0	0	23	29	0	0	0	4	8	1	0	0	11	21	0
8	0	0	0	0	19	26	0	0	0	8	15	1	0	0	10	18	1
9	0	0	1	3	0	0	0	6	7	21	9	5	4	13	18	7	5
10	0	0	12	65	0	1	0	0	6	0	0	0	0	0	1	0	0
11	1	6	28	14	0	1	1	4	31	2	4	1	0	0	0	0	0

Data refer to CAS(14,11) calculations (1.90 Å Fe – O_2 bond length) in Figure 5-30 and Table 5-10. $O_{central}$ is the oxygen atom directly bound to Fe, $O_{terminal}$ is the second-sphere atom of O_2. The z-axis points along the Fe – N_{His} bond and x/y-axis along Fe – N_{Porph} bonds.

Zusammenfassung

Proteine mit metallischen Kofaktoren (Metalloproteine) vollführen eine große Bandbreite an chemischen Prozessen und sind involviert in anspruchsvolle katalytische Reaktionen, z. B. die Aktivierung von kleinen und stabilen Molekülen wie N_2, H_2, H_2O, CH_4 und O_2, mit hoher energetischer Effizienz und Umsetzungsrate. Jeder dieser Prozesse wurde über Jahrmillionen von der Natur optimiert unter Konditionen, die von hohem Interesse für industrielle Anwendung sind, wie Atmosphärendruck, Raumtemperatur und neutralem pH, unter Einsatz von weitverbreiteten Elementen. Das Verständnis dieser enzymatischen Reaktionen und der Metall-Liganden Bindungen wird daher helfen, Herausforderungen in synthetischer Katalyse zu begegnen. Fortschrittliche Röntgenabsorptions- und Röntgenemissionsspektroskopie (RAS/RES) ermöglicht die Bestimmung von Oxidations- und Spinzustand von metallischen Kofaktoren durch Detektieren von unbesetzten und besetzten elektronischen Zuständen. Jedoch sind metallische Kofaktoren in hoch reaktiven hochvalenten Zuständen anfällig für Strahlenschäden. In dieser Arbeit wurde die Messung von RAS/RES Daten durch Entwicklung einer Messmethode mit zeitaufgelösten Fluoreszenzmessungen (ZAFM) und Evaluierung der Messeffizienz im Vergleich zu konventionellen Ansätzen signifikant beschleunigt. ZAFM ermöglichte die Detektion von schwachen spektralen Merkmalen mit hohem Signal-Rausch Verhältnis von dünn konzentrierten und hoch strahlungssensitiven Proben. Quantitative Korrelation von RAS/RES Daten mit quantenchemischen (QC) Kalkulationen aus der Dichtefunktionaltheorie (DFT) oder der *complete active space self-consistent field* (CASSCF) Methode als Teil der Molekülorbitaltheorie (MO-Theorie) ermöglicht die Beschreibung von wichtigen strukturellen und elektronischen Parametern in Metalloproteinen und synthetischen Verbindungen. QC Kalkulationen wurden während dieser Arbeit in unserem Labor etabliert und die Simulation von spektralen Merkmalen verifiziert mit RAS/RES Daten von ausgewählten Metalloproteinen und synthetischen Verbindungen. Ein Fokus wurde insbesondere auf Proteine der Hämgruppe gelegt, welche der dominante Kofaktor bei Sauerstoffreaktionen ist. Hämoglobin (HB) und Myoglobin (MB) sind essenziell für Transport, Aufbewahrung und Detektion von molekularem Sauerstoff in Wirbeltieren und deren physiologische Funktion seit mehr als einem Jahrhundert untersucht. Nichtsdestotrotz wurde die Natur der elektronischen Konfiguration in mit Sauerstoff angereichertem HB/MB (**oxy**) noch nicht hinreichend geklärt. Die sich widersprechenden, debattierten Modelle unterscheiden sich hauptsächlich im Spinzustand des Fe Atoms und des O_2 Liganden. Wir haben fortschrittliche RAS/RES mit QC Simulationen kombiniert um die elektronische Konfiguration von HB/MB in verschiedenen Präparationen und drei Porphyrin Modelkomplexen, zu untersuchen. Unsere Ergebnisse bezüglich des Spinzustandes und der elektronischen Konfiguration in **oxy** HB und MB lieferte eine adäquate Beschreibung der Metall-Liganden Interaktion und vereinte die klassischen Modelle in ein holistisches Bild der Fe – O_2 Bindung.

Selbstständigkeitserklärung

Ich erkläre gegenüber der Freien Universität Berlin, dass ich die vorliegende Doktorarbeit selbstständig und ohne Benutzung anderer als der angegebenen Quellen und Hilfsmittel angefertigt habe. Die vorliegende Arbeit ist frei von Plagiaten. Alle Ausführungen, die wörtlich oder inhaltlich aus anderen Schriften entnommen sind, habe ich als solche kenntlich gemacht. Diese Arbeit wurde in gleicher oder ähnlicher Form noch bei keiner anderen Universität als Prüfungsleistung eingereicht und ist auch noch nicht veröffentlicht.

Berlin den:

Nils Schuth