Supplementary information:

Ligand Discrimination of Myoglobin in Solution: An Iron L-Edge X-Ray Absorption Study of the Active Centre

Kathrin M. Lange,[a] Ronny Golnak,[b] Sébastien Bonhommeau[c] and Emad F. Aziz[b*]

[a] *Max-Born Institut Berlin, Max-Born-Str. 2a 12489 Berlin*

[b] *Joint Ultrafast Dynamics Lab in Solutions and at Interfaces (JULiq)*

Helmholtz-Zentrum Berlin für Materialien und Energie,

Albert-Einstein-Strasse 15, 12489 Berlin, Germany

and Freie Universität Berlin, FB Physik,

Arnimallee 14, D-14195 Berlin, Germany

[c] *Université de Bordeaux, Institut des Sciences Moléculaires, CNRS UMR 5255, F-33400 Talence, France*

E-Mail: Emad.Aziz@helmholtz-berlin.de & Emad.Aziz@fu-berlin.de

I. Sample preparation and UV-Vis Spectra

The samples have been prepared as have shown previously.[1-4] Briefly, the 2 mM aquometmyoglobin (metMb) sample was prepared by dissolving horse skeletal muscle myoglobin (Sigma Aldrich) into a 50 mM sodium phosphate buffer (pH 7). The buffer was sparged before with deoxygenated nitrogen gas for 8 hours. The 2 mM MbCN sample was obtained by adding NaCN in tenfold excess dissolved in a 0.1 M sodium hydroxide solution to a metMb solution prepared as described above. For the 2 mM ferrous samples (deoxyMb) $Na_2S_2O_4$ was used as a reducing agent, which was added in 2 fold excess to a deoxygenized metMb solution under helium atmosphere upon constant stirring. The 2 mM MbNO sample was prepared by adding $NaNO_2$ in 10 fold excess dissolved in buffer solution to a deoxyMb solution. For the 2 mM MbCO and MbO_2 samples, a deoxyMb solution was sparged with CO gas and O_2 gas, respectively.

Via UV-Vis spectroscopy successful ligation was controlled by comparing the characteristic intensities and positions of the α-, β- and Soret band to references from literature (see Fig.S1).[5-7] Further UV-Vis measurements were carried out while the sample flows in front of X-ray for ensuring proper ligation stability and monitoring possible sample damage.

Note that, we cannot exclude the presence of a small Fe(II) fraction (i.e. deoxyMb fraction) in the metMb and MbCN. However, it is a negligible fraction based on the UV/vis spectra. Especially the Q-band of the UV/Vis spectra is sensitive to the oxidation state and even small changes will alter it markedly.[8,9] See figure 2 in Yi *et al.*,[8] and figure 2 in Sugawara *et al.*[9] In order to ensure a complete oxidation of metMb and MbCN Fe-centers, several approaches can be employed. While Svistunenko *et al.* (page 596)[10] used ammonium persulphate to guarantee a complete final oxidation, Harada *et al.*,[11] used potassium ferricyanide. Nevertheless, for our purpose, we preferred to keep the original sample (without additives) to avoid potential interferences in our spectra.

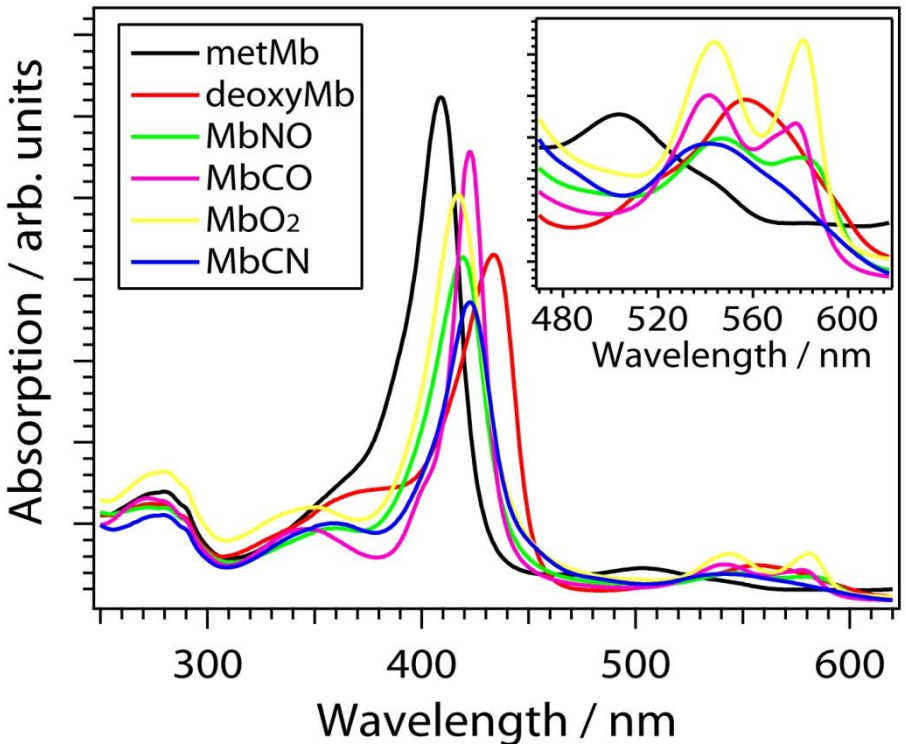

Figure S1: UV-Vis spectra of myoglobin upon ligation to different ligands.

II. Multiplet calculations and charge transfer quantification

In order to extract quantitative information from $L_{2,3}$-edges absorption spectra, experimental XA spectra were modelled using the so-called charge transfer Multiplet theory (CTMT).[12,13] This approach takes into account all the electronic Coulomb interactions as well as the spin–orbit coupling on any electronic open shell and treats the geometrical environment of the absorbing atom through crystal field potential. The spectrum is calculated as the sum of all possible localized transitions in atoms for an electron excited from the 2p level to a 3d

level. No attention is paid to the electric dipole allowed 2p - 4s transitions that are experimentally and theoretically found to be negligible compared to the 2p-3d main channel.

In the simplest formulation, a pure $3d^n$ configuration is attributed to the 3d transition ions in the ground state and transitions between $2p^6 3d^n$ ground state and $2p^5 3d^{n+1}$ final excited state are calculated. The inter-electronic repulsions are introduced through Slater–Condon integrals, F^2_{dd} and F^4_{dd} for the initial state and F^2_{dd}, F^4_{dd}, F^2_{pd}, G^1_{pd} and G^3_{pd} for the final state. The Slater–Condon integrals are calculated through an atomic Hartree–Fock model and are scaled down by a reduction factor κ which reflects the electronic delocalisation and the covalence of the chemical bonds. The atomic spin–orbit coupling parameters, ζ_{3d} and ζ_{2p}, are calculated considering the monoelectronic potential around the ion. The octahedral surrounding of the metal ion is represented by an octahedral (O_h) ligand field potential whose strength is parameterized by 10Dq. In this case, 3d metal orbitals are split into t_{2g} (d_{xz}, d_{xy}, d_{yz}) and e_g (d_z^2,$d_{x^2-y^2}$) sets. For a tetragonal (D_{4h}) ligand field, in case of Jahn–Teller distortions for instance, two other parameters must be added, namely Ds and Dt, to account for the deformation and the 3d metal orbitals further decompose into b_{2g} (d_{xy}), e_g (d_{xz},d_{yz}), a_{1g} (d_z^2) and b_{1g} ($d_{x^2-y^2}$) hence. A sketch comparing O_h and D_{4h} symmetries was provided in previous work on hemoglobin.[14] Note that the local symmetry of the iron sites in all haem species (either five-coordinated pyramidal or six-coordinated planar) is similar to the D_{4h} symmetry. We can underscore here that considering a local symmetry lower than D4h would have left our calculations unaltered provided the degeneracy e_g (d_{xz},d_{yz}) orbitals is not lifted.[15]

Covalent mixing of the metal valence d-orbitals with the ligand valence p-orbitals is simulated using a charge-transfer model, which, in the case of ligand-to-metal charge transfer (LMCT), adds a $d^{n+1}L^+$ configuration above the d^n ground state, where $3d^{n+1}$ stands for an extra 3d electron coming from the ligands and L^+ for the corresponding hole on a ligand orbital. The $d^{n+1}L^+$ configuration is set at an energy Δ_{LM} above the d^n configuration and these two states are coupled by configuration interaction (CI), represented by the mixing term $T_i = \langle d^n|h|^{dn+1}L\rangle$, where h is the molecular Hamiltonian the T_i is proportional to metal-ligand overlap for each of the i symmetry blocks. It is represented in octahedral symmetry by the two charge transfer integrals $T(t_{2g})$ and $T(e_g)$ and in tetragonal symmetry by the four integrals $T(b_{1g})$, $T(a_{1g})$, $T(b_{2g})$ and $T(e_g)$.For a donor ligand system, the ground and LMCT states are written as linear combinations of $|3d^n\rangle$ and $|3d^{n+1}L^+\rangle$ wave functions, while the L-edge excited state is a linear combination of $|2p^5 3d^{n+1}\rangle$ and $|2p^5 3d^{n+2}L\rangle$ wave functions. The coefficients of these linear combinations are functions of T and Δ_{LM} for the ground state and T' and Δ_{LM}' for the excited state, where $\Delta_{LM}' = \Delta_{LM} + U\text{-}Q$ (also called the charge transfer

energy) with U being the 3d-3d electron repulsion (so-called Hubbard energy) and Q the 2p-3d repulsion (core hole energy). In case of metal-to-ligand charge transfer (MLCT), the ground state configurations are composed of $3d^n$ and $3d^{n-1}L^-$ configurations where L^- represents an extra electron originating from the metal ion. Δ_{ML} (Δ'_{ML}) determine the average energy separation between $3d^n$ and $3d^{n-1}L^-$ configurations in the ground (final) state. It fulfils the relation $\Delta'_{ML} = \Delta_{ML} + Q - U$. In the frame of our CTMT calculations, κ=80% (standard value), ζ_{3d} =0.6 eV, ζ_{2p} = 8.2 eV, 10Dq=1.3 eV, Ds=Dt=0, Δ_{LM}=0, Δ_{ML}=-1 eV and U-Q=-1 eV. The other parameters which allow us to obtain the best agreement between experimental and calculated XA spectra for deoxyMb and metMb are provided in table S1.

The purpose of the calculation lies in the computation of intensities of all possible transitions. In order to compare them to the experimental $L_{2,3}$-edges XAS spectra, the theoretical transitions must be convoluted by a broadening function which takes into account the intrinsic core-hole lifetime and the instrumental resolution. In the following simulations, the full width at half maximum (FWHM) of the lifetime Lorentzian broadening function is 0.2 eV over the L_3-edge and 0.4 eV over the L_2-edge. The FWHM of the instrumental Gaussian broadening function is fixed at 0.2 eV. All simulations are carried out considering a temperature of 300 K.

Table S1: Ligand-to-metal (LM) and metal-to-ligand (ML) charge transfer integrals used to reproduce experimental XA spectra for deoxyMb and metMb in D_{4h} symmetry. Ground state configurations are also indicated.

	deoxyMb	metMb	metMb with increased MLCT
Ground state configuration	94.3 % d^5L^- + 5.6 % d^6 + 0.1 % d^7L^+	6.7 % d^4L^- + 65.0 % d^5 + 28.3 % d^6L^+	15.0 % d^4L^- + 63.5 % d^5 + 22.5 % d^6L^+
$T(b_1)_{LM}$	1.0	2.4	2.4
$T(a_1)_{LM}$	1.0	0	0
$T(b_2)_{LM}$	0	1.2	1.2
$T(e)_{LM}$	0	1.7	1.2
$T(b_1)_{ML}$	1.0	0.7	1.0
$T(a_1)_{ML}$	0.4	0.7	1.0
$T(b_2)_{ML}$	0	0	0
$T(e)_{ML}$	0	0	0

Figure S2 presents calculated XA spectra for deoxyMb, metMb and a metMb sample with increased MLCT. Despite the use of very different ground state configurations, XA spectra associated with metMb and deoxyMb proteins look very similar after normalization relative to the L_2-edge known to be little affected by fluorescence quenching. The main discrepancy between these spectra is observed at the high-energy side of the L_2-edge (~721.7 eV), where the calculated intensity is overestimated for deoxyMb and underestimated for metMb compared to their corresponding experimental value (Fig.3 of the body text). Upon increasing σ back-donation (MLCT) in metMb, we observe an intensity decrease of the main spectral feature centred at 707.4 eV, while the intensity of the pre-edge peak at ~705.6 eV rises slightly. This simulation shows that the decrease in intensity observed for MbCN (expected to be in the +III oxidation state like metMb) can be partly reproduced by adding MLCT. In the present case, the amount of d^4L^- configuration more than doubled, since it goes from 6.7% in metMb to 15.0% (see Table S1). In MbCN, we can expect an even much higher increase but the occurrence of fluorescence quenching (see body text), which is not taken into account in our calculations, prevent us from evaluating precisely the amount of MLCT.

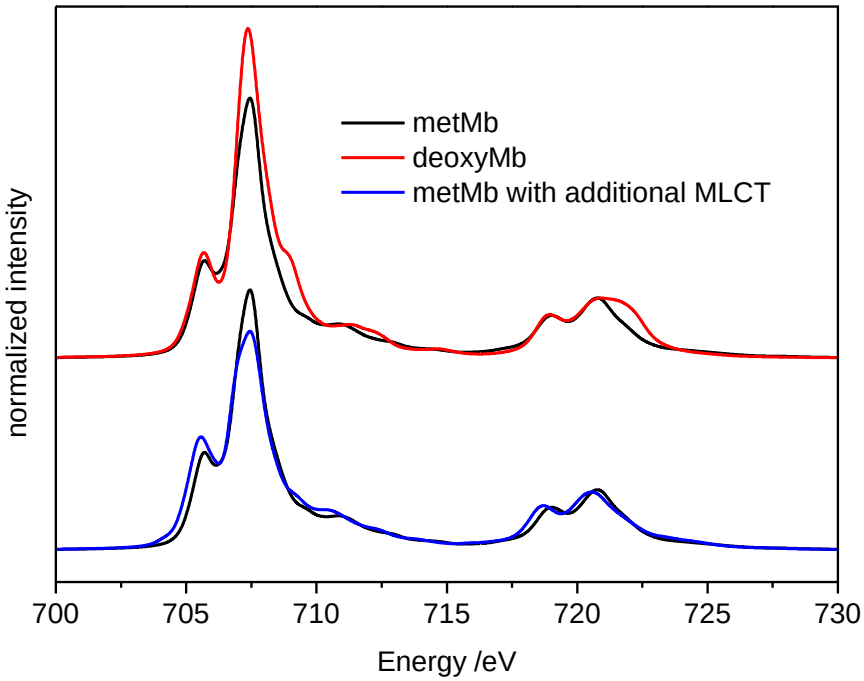

Figure S2: Calculated Fe L-edge X-ray absorption spectra associated with metMB and deoxyMb proteins. A spectrum showing the effect of an increased MLCT in metMb is also plotted.

(1) Kim, S.; Jin, G.; Lim, M. *J Phys Chem B* **2004**, *108*, 20366.
(2) Della Longa, S.; Pin, S.; Cortes, R.; Soldatov, A. V.; Alpert, B. *Biophys J* **1998**, *75*, 3154.
(3) Wang, H. X.; Peng, G.; Miller, L. M.; Scheuring, E. M.; George, S. J.; Chance, M. R.; Cramer, S. P. *J Am Chem Soc* **1997**, *119*, 4921.

(4) Alves Lima, F., Thèse École polytechnique fédérale de Lausanne EPFL, 2011.

(5) Bowen, W. J. *J Biol Chem* **1949**, *179*, 235.

(6) Antonini, E.; Brunori, M. *Hemoglobin and myoglobin in their reactions with ligands*; North-Holland Pub. Co.: Amsterdam, 1971.

(7) Kelly, J. B.; Baskin, S. I. *U.S. Army Medical Research Institute of Chemical Defense* **2006**.

(8) Yi, J.; Orville, A. M.; Skinner, J. M.; Skinner, M. J.; Richter-Addo, G. B. *Biochemistry-Us* **2010**, *49*, 5969.

(9) Sugawara, Y.; Matsuoka, A.; Kaino, A.; Shikama, K. *Biophys J* **1995**, *69*, 583.

(10) Svistunenko, D. A.; Sharpe, M. A.; Nicholls, P.; Blenkinsop, C.; Davies, N. A.; Dunne, J.; Wilson, M. T.; Cooper, C. E. *Biochem J* **2000**, *351*, 595.

(11) Harada, Y.; Taguchi, M.; Miyajima, Y.; Tokushima, T.; Horikawa, Y.; Chainani, A.; Shiro, Y.; Senba, Y.; Ohashi, H.; Fukuyama, H.; Shin, S. *J Phys Soc Jpn* **2009**, *78*.

(12) Aziz, E. F.; Eisebitt, S.; de Groot, F.; Chiou, J. W.; Dong, C. G.; Guo, J. H.; Eberhardt, W. *J Phys Chem B* **2007**, *111*, 4440.

(13) Regier, T. Z.; Achkar, A. J.; Peak, D.; Tse, J. S.; Hawthorn, D. G. *Nature Chemistry* **2012**, *advance online publication*.

(14) Aziz, E. F.; Ottosson, N.; Bonhommeau, S.; Bergmann, N.; Eberhardt, W.; Chergui, M. *Phys Rev Lett* **2009**, *102*.

(15) Hocking, R. K.; Wasinger, E. C.; Yan, Y.-L.; deGroot, F. M. F.; Walker, F. A.; Hodgson, K. O.; Hedman, B.; Solomon, E. I. *J Am Chem Soc* **2007**, *129*, 113.