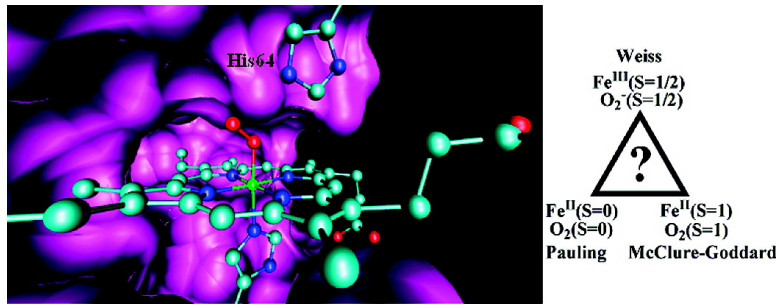
Article

# Nature of the Fe#O Bonding in Oxy-Myoglobin: Effect of the Protein

Hui Chen, Masao Ikeda-Saito, and Sason Shaik

*J. Am. Chem. Soc.*, **2008**, 130 (44), 14778-14790 • DOI: 10.1021/ja805434m • Publication Date (Web): 11 October 2008






## More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES

Published on Web 10/11/2008

# Nature of the Fe−$O_2$ Bonding in Oxy-Myoglobin: Effect of the Protein

Hui Chen,[†] Masao Ikeda-Saito,[‡] and Sason Shaik*,[†]

*Department of Organic Chemistry and the Lise Meitner-Minerva Center for Computational Quantum Chemistry, The Hebrew University of Jerusalem, Givat Ram Campus, 91904 Jerusalem, Israel, and Institute of Multidisciplinary Research for Advanced Materials, Tohoku University, Katahira, Sendai 980-8577, Japan*

Received July 14, 2008; E-mail: sason@yfaat.ch.huji.ac.il

***Abstract:*** The nature of the Fe−$O_2$ bonding in oxy-myoglobin was probed by theoretical calculations: (a) QM/MM (hybrid quantum mechanical/molecular mechanical) calculations using DFT/MM and CASSCF/MM methods and (b) gas-phase calculations using DFT (density functional theory) and CASSCF (complete active space self-consistent field) methods. Within the protein, the $O_2$ is hydrogen bonded by His64 and the complex feels the bulk polarity of the protein. Removal of the protein causes major changes in the complex. Thus, while CASSCF/MM and DFT/MM are similar in terms of state constitution, degree of $O_2$ charge, and nature of the lowest triplet state, the gas-phase CASSCF(g) species is very different. Valence bond (VB) analysis of the CASSCF/MM wave function unequivocally supports the Weiss bonding mechanism. This bonding arises by electron transfer from heme−$Fe^{II}$ to $O_2$ and the so formed species coupled then to a singlet state $Fe^{III}-O_2^-$ that possesses a dative $\sigma$(Fe−O) bond and a weakly coupled $\pi$(Fe−$O_2$) bond pair. The bonding mechanism in the gas phase is similar, but now the $\sigma$(Fe−O) bond involves higher back-donation from $O_2^-$ to $Fe^{III}$, while the constituents of $\pi$(Fe−$O_2$) bond pair have greater delocalization tails. The protein thus strengthens the $Fe^{III}-O_2^-$ character of the complex and thereby affects its bonding features and the oxygen binding affinity of Mb. The VB model is generalized, showing how the protein or the axial ligand of the oxyheme complex can determine the nature of its bonding in terms of the blend of the three bonding models: Weiss, Pauling, and McClure−Goddard.

## 1. Introduction

Myoglobin (Mb) and hemoglobin (Hb) are, respectively, dioxygen storing and transporting proteins in vertebrates.[1] Mb is present in aerobic muscle tissue, while Hb is found packed at high concentrations in red blood cells. Kendrew and Perutz solved the X-ray structures of these two hemoproteins in the 1950s in the feat that ushered a new era in X-ray crystallography of complex biomolecules. The structure of Mb, solved by Kendrew,[2] exhibits a compact and predominantly α-helical globular protein that contains a single heme unit coordinated to a histidine residue as a fifth ligand. Hb, solved by Perutz,[3] is present as an $\alpha_2\beta_2$ heterotetramer in which the individual α and $\beta$ subunits are similar to Mb both at the structural and functional level, each containing a heme unit.[1]

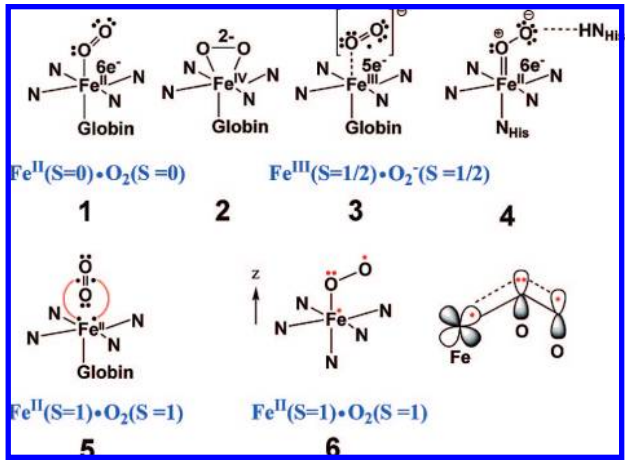
The effect of dioxygen on Hb was first discovered in the 17th century by the English physician John Mayow, who noticed that when venous blood was exposed to air or potassium nitrate (known then as saltpeter) its color changed to that of arterial blood. He ascribed this dramatic change of color to "particles" that were common to air and the saltpeter. Eventually, these particles were characterized by others and called "oxygen" by Lavoisier, an event that ushered the compositional revolution in chemistry.[4] The magnetic property of free Hb was apparently reported first by Faraday in 1845, when he subjected dried blood to a magnet. Almost 56 years later, Gamgee[5] reported that arterial blood is as diamagnetic as water. He apparently did not check venous blood nor did he note the stark difference in the magnetic properties of these two bloods. Pauling discovered this difference in an early paper with Coryell,[6] where the two authors measured the magnetic susceptibility of both deoxy-Hb and oxy-Hb and noted the profound change in the magnetic property upon dioxygen binding. Thus, from two paramagnetic species, the heme with $S = 2$ and $O_2$ with $S = 1$ in the $^3\Sigma_g^-$ ground state, one obtains a diamagnetic oxy-Hb, which meant that 24 electrons underwent pairing in the process in the four heme units! Pauling and Coryell then described the oxy-Hb using a valence bond (VB) structure, shown in **1** in Scheme 1, where $O_2$ is coordinated to $Fe^{II}$ in an end-on manner and the two moieties are in singlet states. In this model $Fe^{II}$ is $sp^3d^2$ hybridized; five of the hybrids are used to form bonds to the

[†] The Hebrew University of Jerusalem.
[‡] Tohoku University.

(1) (a) Springer, B. A.; Sligar, S. G.; Olson, J. S.; Phillips, G. N., Jr *Chem. Rev.* **1994**, *94*, 699. (b) Perutz, M. F.; Fermi, G; Luisi, B.; Shaanan, B.; Liddington, R. C. *Acc. Chem. Res.* **1987**, *20*, 309.
(2) Kendrew, J. C.; Bodo, G.; Dintzis, H. M.; Parrish, R. G.; Wyckoff, H.; Phillips, D. C. *Nature* **1958**, *181*, 662.
(3) Perutz, M. F.; Rossmann, M. G.; Cullins, A. F.; Muirhead, H.; Will, G.; North, A. C. T. *Nature* **1960**, *185*, 416.
(4) Siegfried, R. *From Elements to Atoms: A History of Chemical Composition*; American Philosophical Society: Philadelphia, 2002.
(5) Gamgee, A. *Proc. R. Soc. London* **1901**, *68*, 503.
(6) Pauling, L.; Coryell, C. D. *Proc. Natl. Acad. Sci. U.S.A.* **1936**, *22*, 210.


**10.1021/ja805434m CCC: $40.75** 

***Scheme 1.*** Models for Oxyheme




nitrogens of the porphyrin and to a histidine residue of globin, while the sixth bond is a dative bond from the filled $sp^2$ hybrid of the $O_2$ molecule in its singlet state to the remaining $sp^3d^2$ hybrid on $Fe^{II}$. Additionally, the $Fe^{II}$ retains its six electrons in three doubly filled nonbonding d orbitals, shown near the iron as $6e^-$ (the $t_{2g}$ set in an octahedral environment). This model has created two long scientific debates: one concerned the geometric structure of the complex and the other associated with the nature of the Fe−$O_2$ bonding. The latter controversy has never subsided and is alive also today, more than seven decades after Pauling suggested his model. Accordingly, our paper is focused on the nature of the Fe−$O_2$ bonding, which we derive from VB transformation of the CASSCF/MM (hybrid complete active space self-consistent field/molecular mechanical) wave function, in a manner that reveals the role of the protein. However, to understand the entire bonding controversy it is essential to first briefly describe the two aspects of the story.

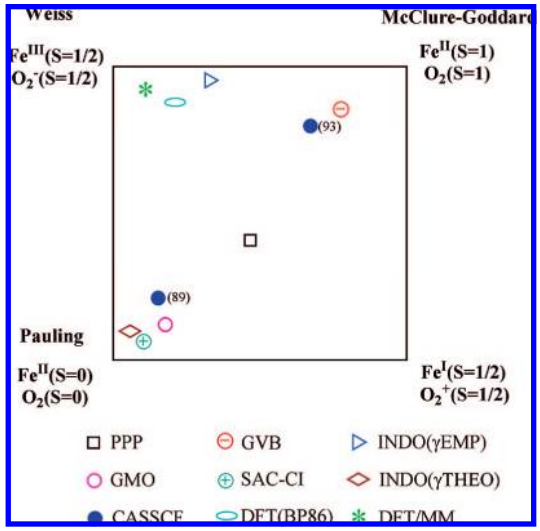
In 1956 Griffith[7] and later Gray[8] suggested the triangular structural model, **2** in Scheme 1. Even though Pauling ruled out the triangular model,[6] based on strain, nevertheless, the triangular model accounted equally well as did Pauling's model for the diamagnetism of the oxyheme complexes. This and the fact that such complexes were known at that time,[9] albeit not for $Fe^{II}$, created a strong incentive to determine the structure of Fe−$O_2$ complexes. In 1974, Collman[10] synthesized the "picket-fence" Fe−$O_2$ complexes and demonstrated that the $O_2$ molecule was bound to the iron in an end-on bent fashion, as suggested by Pauling. This was followed by a solution of the structure of oxy-Mb by Phillips and oxy-Hb by Shaanan,[11] who showed that in both cases the Fe−$O_2$ geometry involves end-on bent $O_2$ bonding, albeit to different degrees. This picture is well established to this day.[12−14] Thus, the X-ray structure determinations essentially confirmed Pauling's proposal (**1**, Scheme 1) that had been made 38−46 years before. Now, attention was drawn to the subtler question of the nature of the Fe−$O_2$ bonding.

In 1964 Weiss[15] considered Pauling's model **1** and stated "Many of the physical properties of oxyhemoglobin are at variance with such a view". On the basis of spectroscopic data and chemical behavior of synthetic Fe−$O_2$ complexes in solution, Weiss argued in favor of another model in which $Fe^{II}$ transfers a single electron to $O_2$ and the two resultant "radicals", $Fe^{III}$ and $O_2^-$, couple to a singlet state, as shown schematically in **3** in Scheme 1. In response, Pauling revised his earlier model[16] and presented the structure drawn in **4** in Scheme 1. Here, the Fe−$O_2$ bond is presented as a double bond, the proximal oxygen is taken to be neutral due to its relative electronegativity compared with $Fe^{II}$,[17] while the distal oxygen is negatively charged. Pauling reasoned that this negative charge would be stabilized by a His residue that resides nearby in the distal pocket (His64 in Mb). Whereas the oxidation state of iron in **4** is not apparent anymore and arguably it may conform to a high-valent $Fe^{IV}$, still Pauling insisted that the Weiss model is invalid and that the new model **4** represents the union of $Fe^{II}$ ($S = 0$) with $O_2$ ($S = 0$). He further postulated that the histidine ligand is protonated to prevent oxidation of $Fe^{II}$ to $Fe^{III}$.

In 1960 McClure[18] suggested an alternative Fe−$O_2$ bonding model whereby $Fe^{II}$ participates in its intermediate spin state with $S = 1$ and the two unpaired electrons on the ferrous center pair up in a Heitler−London fashion[19] with those in the $O_2$ ($S = 1$) to a total of a singlet state as we show schematically in **5**. Harcourt supported this model in a series of papers,[20] and so did Seno et al.[21] Interestingly, neither Pauling nor Weiss relate to the McClure model in their discussion. However, the model starts to be part of the debate when Goddard and Olafson suggested "the ozone model" to describe the $FeO_2$ bonding in 1975.[22] In this model, **6** in Scheme 1, the bonding involves $Fe^{II}(S = 1)\bullet O_2(S = 1)$ union as in the McClure model with a $\sigma_{Fe-O}$ bond involving overlap of the iron $d_{z^2}$ orbital with the in-plane $\pi^*$ orbital of $O_2$ and a 4-electron/3-center $\pi$ bonding like in ozone. The GVB-CI calculations of Goddard and Olafson on small model systems[22,23] reproduced the singlet ground states, the bent end-on binding of $O_2$, the presence of a low-lying triplet state, and the magnitude of the quadrupole coupling Mössbauer parameter. However, the sign of the quadrupole coupling

(7) Griffith, J. S. *Proc. R. Soc. Ser. A* **1956**, *235*, 23.
(8) Gray, H. B. *Adv. Chem. Ser.* **1971**, *100*, 365.
(9) Jones, R. D.; Summerville, D. A.; Basolo, F. *Chem. Rev.* **1979**, *79*, 139.
(10) (a) Collman, J. P.; Gagne, R. R.; Reed, C. A.; Robinson, W. T.; Rodley, G. A. *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71*, 1326. (b) Collman, J. P.; Gagne, R. R.; Reed, C. A.; Halbert, T. R.; Lang, G.; Robinson, W. T. *J. Am. Chem. Soc.* **1975**, *97*, 1427.
(11) (a) Phillips, S. E. V. *Nature* **1978**, *273*, 247. (b) Phillips, S. E. V. *J. Mol. Biol.* **1980**, *142*, 531. (c) Phillips, S. E. V.; Schoenborn, B. P. *Nature* **1981**, *292*, 81. (d) Hanson, J. C.; Schoenborn, B. P. *J. Mol. Biol.* **1981**, *153*, 117. (e) Shaanan, B. *Nature* **1982**, *296*, 683. (f) Shaanan, B. *J. Mol. Biol.* **1983**, *171*, 31.
(12) Momenteau, M.; Reed, C. A. *Chem. Rev.* **1994**, *94*, 659.
(13) Dickinson, L. C.; Chien, J. C. W. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 1235.
(14) (a) Kitagawa, T.; Ondrias, M. R.; Rousseau, D. L.; Ikeda-Saito, M.; Yonetani, T. *Nature* **1982**, *298*, 869. (b) Brucker, E. A.; Olson, J. S.; Phillips, G. N., Jr.; Dou, Y.; Ikeda-Saito, M. *J. Biol. Chem.* **1996**, *271*, 25419. (c) Dube, H.; Kasumaj, B.; Calle, C.; Saito, M.; Jeschke, G.; Diederich, F. *Angew. Chem., Int. Ed.* **2008**, *47*, 2600.
(15) Weiss, J. J. *Nature* **1964**, *202*, 83.
(16) Pauling, L. *Nature* **1964**, *203*, 182.
(17) Pauling, L. *Stanford Med. Bull.* **1948**, *6*, 215.
(18) McClure, D. S. *Radiation. Res. Suppl.* **1960**, *2*, 218.
(19) Shaik, S.; Hiberty, P. C. *A Chemist's Guide to Valence Bond Theory*; John Wiley & Sons, Inc.: Hoboken, NJ, 2008.
(20) (a) Harcourt, R. D. *Int. J. Quantum Chem.* **1971**, *5*, 479. (b) Harcourt, R. D. *Biopolymers* **1972**, *11*, 1551. (c) Harcourt, R. D. *Inorg. Nucl. Chem. Lett.* **1973**, *9*, 475. (d) Harcourt, R. D. *Chem. Phys. Lett.* **1990**, *167*, 374.
(21) (a) Seno, Y.; Otsuka, J.; Matsuoka, O.; Fuchikam, N. *J. Phys. Soc. Jpn.* **1972**, *33*, 1645. (b) Otsuka, J.; Matsuoka, O.; Fuchikam, N.; Seno, Y. *J. Phys. Soc. Jpn.* **1973**, *35*, 854.
(22) Goddard, W. A., III; Olafson, B. D. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 2335.
(23) Olafson, B. D.; Goddard, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 1315.

***Scheme 2.*** Qualitative Results of Various Theoretical Methods




constant was opposite to the one reported experimentally,[24] and this discrepancy was noted later by Karplus et al.[25] The GVB calculations further revealed that the $O_2$ moiety is virtually neutral in discord with the Weiss model. However, the calculations did not support the Pauling model either since the $Fe^{II}$ and $O_2$ moieties in bonding were not in their singlet states.

The intense Weiss−Pauling exchange in the *Nature* issue of 1964[15,16,26] and addition of the McClure−Harcourt−Goddard model has set in motion a debate over the correct bonding in iron−oxy complexes of Mb and Hb and their synthetic analogs. This debate stretches from 1936 where the first paper was published to 2007 when extensive CASSCF studies were done to probe this bonding.[27,28] The debate is ongoing, and the conclusions seem to vary frequently: What is the correct $Fe-O_2$ bonding model: is it Pauling's; Weiss's, McClure−Harcourt−Goddard's; all; some; or maybe none? This is still the question. Let us briefly describe the results of this ongoing debate.

To describe these results briefly we borrow the mnemonic, shown in Scheme 2, as the one used by Hall et al.[29] to summarize the model assignments reached by various computational techniques. Three corners of the square, in Scheme 2, represent the three-named models and an additional one involving transfer of an electron from $O_2$ to $Fe^{II}$ and singlet coupling of the two radicals. Inside the square there are various circles, squares, and triangles that correspond to the assigned model by a given computational method.[30] We added also data up to 2000, including early semiempirical methods, i.e., INDO-SCF-CI[31] and extend SCF-CI Pariser−Parr−Pople (PPP) type,[25] ab initio methods, i.e., generalized molecular orbital (GMO) approach with configuration interaction,[29a] symmetry-adapted cluster with configuration interaction (SAC/SAC-CI),[32] and complete active space self-consistent field (CASSCF) method,[33] density functional theory (DFT) method,[34] and DFT/MM approach.[35] There are no Hartree−Fock (HF) calculations since these calculations do not produce the correct spin state as the ground state.[36] Thus, quite early in the debate it was recognized that the problem requires post-HF calculation that includes a good level of electron correlation.

Generally, the assignment to a given model is made based on the charge on the $O_2$ moiety and analyses of occupancies of the MO resembling those of $O_2$ and iron. Inspection of the square shows that each model has some support from some methods, but the Pauling model is supported somewhat more than others, primarily because the charge on $O_2$ is generally quite low (less than $-0.1e$) in most of the calculations with the exception of one semiempirical INDO result that gave a significant charge of $-0.75e$[31c] and supported the Weiss model. Early post-HF ab initio calculations do not seem to agree. Thus, while an initial CASSCF calculation[33a] supported the Pauling model, a subsequent improved calculation[33b] supported the McClure−Harcourt−Goddard model. Similarly, the symmetry-adapted cluster CI (SAC-CI) calculations[32] supported the Pauling model. In contrast to these post-HF results, DFT[34,37] and DFT/MM calculations[35] unequivocally support the Weiss model, showing in addition to a significant charge on $O_2$ also an open-shell singlet ground state with two singly occupied orbitals, one on the iron and one on the $O_2$.

Recent extensive CASSCF calculations with large active spaces and good basis sets by Roos and Ryde[27] gave more consistent results. Inspection of the charge on $O_2$, $-0.20e$, led initially to a conclusion that supports the Pauling model more. However, subsequent consideration of the configurations, which contribute to the CASSCF wave function, caused the authors to modify their conclusion and suggest that the CASSCF wave function had an equal mixture of the Pauling and Weiss models. More recent CASSCF calculations[28] reinforced this conclusion. Furthermore, based on their results Ribas-Ariño and Novoa concluded that DFT methods fail to properly describe the electronic structure of oxyheme. This last conclusion was based

(24) (a) Lang, G.; Marshall, W. *J. Mol. Biol.* **1966**, *18*, 385. (b) Lang, G.; Marshall, W. *Proc. Phys. Soc.* **1966**, *87*, 3. (c) Spartallan, K.; Lang, G.; Collman, J. P.; Gagne, R. R.; Reed, C. A. *J. Chem. Phys.* **1975**, *63*, 5375.
(25) (a) Huynh, B. H.; Case, D. A.; Karplus, M. *J. Am. Chem. Soc.* **1977**, *99*, 6103. (b) Case, D. A.; Huynh, B. H.; Karplus, M. *J. Am. Chem. Soc.* **1979**, *101*, 4433.
(26) Weiss, J. J. *Nature* **1964**, *203*, 183.
(27) (a) Jensen, K. P.; Roos, B. O.; Ryde, U. *J. Inorg. Biochem.* **2005**, *99*, 45. (b) Erratum: Jensen, K. P.; Roos, B. O.; Ryde, U. *J. Inorg. Biochem.* **2005**, *99*, 978.
(28) Ribas-Ariño, J.; Novoa, J. J. *Chem. Commun.* **2007**, 3160.
(29) (a) Hall, M. B.; Newton, J. E. *Inorg. Chem.* **1984**, *23*, 4627. (b) Bytheway, I.; Hall, M. B. *Chem. Rev.* **1994**, *94*, 639. Note however that we removed the labels of the coordinates in the square.
(30) (a) Pople, J. A.; Beveridge, D. L. *Approximate Molecular Orbital Theory*; McGraw-Hill: New York, 1970. (b) Blomquist, J.; Nordén, B.; Sundbom, M. *Theor. Chim. Acta* **1973**, *28*, 313. (c) Johnson, K. H. *Adv. Quantum Chem.* **1973**, *7*, 143. (d) Hall, M. B. *Chem. Phys. Lett.* **1979**, *61*, 461. (e) Nakatsuji, H.; Hirao, K. *J. Chem. Phys.* **1978**, *68*, 2035. (f) Nakatsuji, H. *Chem. Phys. Lett.* **1978**, *59*, 362. (g) Koch, W.; Holthausen, M. C. *A Chemists Guide to Density Functional Theory*; Wiley-VCH: Weinheim, 2000.
(31) (a) Loew, G. H.; Kirchner, R. F. *J. Am. Chem. Soc.* **1975**, *97*, 7388. (b) Kirchner, R. F.; Loew, G. H. *J. Am. Chem. Soc.* **1977**, *99*, 4639. (c) Herman, Z. S.; Loew, G. H. *J. Am. Chem. Soc.* **1980**, *102*, 1815.
(32) Nakatsuji, H.; Hasegawa, J.; Ueda, H.; Hada, M. *Chem. Phys. Lett.* **1996**, *250*, 379.
(33) (a) Yamamoto, S.; Kashiwagi, H. *Chem. Phys. Lett.* **1989**, *161*, 85. (b) Yamamoto, S.; Kashiwagi, H. *Chem. Phys. Lett.* **1993**, *205*, 306.
(34) Rovira, C.; Kunc, K.; Hutter, J.; Ballone, P.; Parrinello, M. *J. Phys. Chem. A* **1997**, *101*, 8914.
(35) Rovira, C. *J. Mol. Struct. (THEOCHEM)* **2003**, *632*, 309.
(36) (a) Dedieu, A.; Rohmer, M.-M.; Veillard, B. A. *J. Am. Chem. Soc.* **1976**, *98*, 3717. (b) Nozawa, T.; Hatano, M.; Nagashima, U.; Obara, S.; Kashiwaga, H. *Bull. Chem. Soc. Jpn.* **1983**, *56*, 1721.
(37) (a) Jensen, K. P.; Ryde, U. *J. Biol. Chem.* **2004**, *279*, 14561. (b) Franzen, S. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 16754. (c) Nakashima, H.; Hasegawa, J.-Y.; Nakatsuji, H. *J. Comput. Chem.* **2006**, *27*, 426. (d) Unno, M.; Chen, H.; Kusama, S.; Shaik, S.; Ikeda-Saito, M. *J. Am. Chem. Soc.* **2007**, *129*, 13394. (e) Sigfridsson, E.; Ryde, U. *J. Biol. Inorg. Chem.* **1999**, *4*, 99. (f) Strickland, N.; Harvey, J. N. *J. Phys. Chem. B* **2007**, *111*, 841.

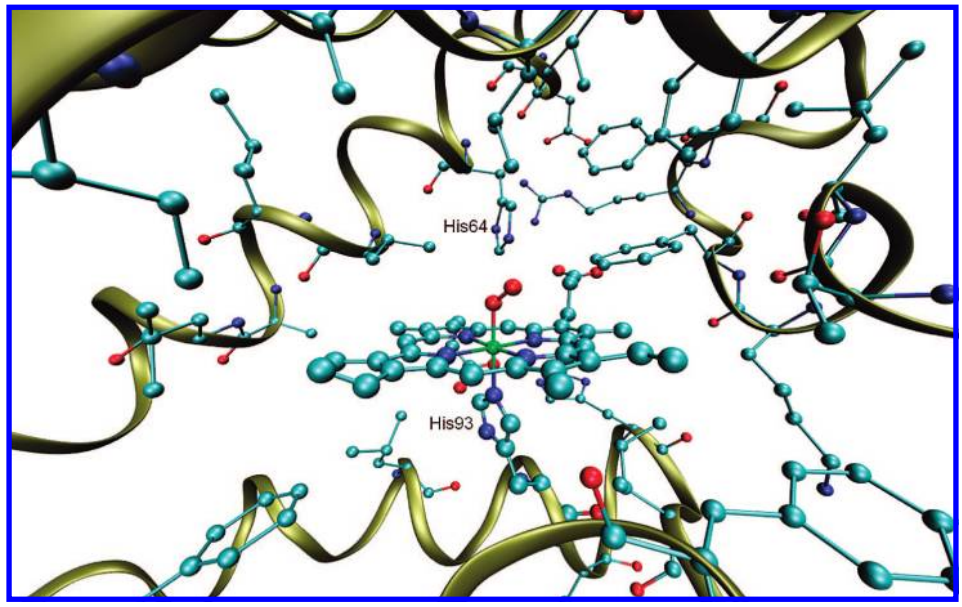



***Figure 1.*** Active site of oxy-Mb (PDB code 2Z6S).[37d]

on the facts that the CASSCF calculations revealed a small charge transfer to $O_2$ (−0.20e) and a multiconfigurational wave function. In contrast, DFT methods yield generally larger charge transfer to $O_2$ and by construction are obviously not multiconfigurational. Furthermore, it turns out that the triplet state in the CASSCF calculation[28] has significantly more multiconfigurational character than the singlet ground state, while the Weiss model requires both states to have the same electronic structure, save the spin coupling, as indeed obtained by the DFT calculations. Thus, it seems that DFT is missing out seriously on this problem. However, as we shall see latter, this is not the case: DFT/MM in fact gives a reasonably good Fe−$O_2$ bonding description.

Thus, despite many theoretical studies, 'what is the actual Fe−$O_2$ bonding in oxy-Hb and oxy-Mb?' remains a question.[29b,38] Specifically, except for the DFT/MM calculations by Rovira,[35] all of the above studies used rather simplified models and performed gas-phase calculations. Does the protein not have any effect? In fact, a mere inspection of the situation in oxy-Mb illustrated in Figure 1[37d] shows that the O−O bond eclipses one of the Fe−N bonds in the porphyrin plane, while in all gas-phase calculations the O−O bond bisects the N−Fe−N angle in the porphyrin plane. This feature is common, in fact, to oxy-Mb and oxy-Hb.[11,37d,39] Thus, the protein affects the conformation of oxy-Mb, but it exerts even more important effects. The first one is the distal histidine residue His64 (E7), which was found to form a hydrogen bond (H bond) with the distal O atom of $FeO_2$ in Mb (see Figure 1) and Hb.[11c,40] Indeed, key residue His64 has been widely explored and is thought to play an important role in the relative binding of Mb to ligands such as $O_2$, CO, etc.[1,37e,41,42] To date, only two papers addressed this factor explicitly in calculations but both were mainly focused on energetic evaluation rather than on electronic structure exploration.[35,37d] The second effect of the protein is the bulk polarization by the protein environment of Mb and Hb on the electronic structure of oxyheme active center.

Clearly, therefore, the nature of the Fe−$O_2$ bonding cannot be discussed without consideration of the protein environment. Furthermore, since all the bonding models in the corners of the square in Scheme 2 (see also Scheme 1) are essentially VB-type models, we must find a direct way of extracting this information from the wave function rather than relying on indirect probes like the charge on $O_2$ and/or resemblance of delocalized MOs to MOs of the $O_2$ and heme fragments. We therefore decided to undertake such a study with the following strategy. (a) In the first step, we study oxy-Mb using DFT/MM calculations with five different functions, three pure (GGA) functionals and two hybrid functions. (b) Subsequently, we perform in-protein CASSCF/MM calculations as well as gas-phase CASSCF(g) calculations to elucidate the effect of the protein. (c) Finally, we shall project the CASSCF wave function to a valence bond wave function in a manner that will enable us to make a clear assignment of the correct bonding description. On the basis of these calculations we shall also address the question whether the DFT method is appropriate to treat the oxyheme system.

## 2. Computational Details

**2.1. QM(DFT)/MM Methodology and Software.** The QM-(DFT)/MM calculations were done using ChemShell[43] interfaced

(38) Kaupp, M.; Rovira, C.; Parrinello, M. *J. Phys. Chem. B* **2000**, *104*, 5200.
(39) Vojtechovsky, J.; Chu, K.; Berendzen, J.; Sweet, R. M.; Schlichting, I. *Biophys. J.* **1999**, *77*, 2153.
(40) Lukin, J. A.; Simplaceanu, V.; Zou, M.; Ho, N. T.; Ho, C. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 10354.
(41) Berg, J. M.; Tymoczko, J. L.; Stryer, L. *Biochemistry*, 6th ed.; Freeman: New York, 2007.
(42) (a) Olson, J. S.; Mathews, A. J.; Rohlfs, R. J.; Springer, B. A.; Egeberg, K. D.; Sligar, S. G.; Tame, J.; Renaud, J. P.; Nagai, K. *Nature* **1988**, *336*, 265. (b) Braunstein, D.; Ansari, A.; Berendzen, J.; Cowen, B. R.; Egeberg, K. D.; Frauenfelder, H.; Hon, M. K.; Ormos, P.; Sauke, T. B.; Scfoll, R.; Schulte, A.; Sligar, S. G.; Springer, B. A.; Steinbach, P. J.; Young, R. D. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 8497. (c) Springer, B. A.; Egeberg, K. D.; Sligar, S. G.; Rohlfs, R. J.; Mathews, A. J.; Olson, J. S. *J. Biol. Chem.* **1989**, *264*, 3057. (d) Rohlfs, R. J.; Mathews, A. J.; Carver, T. E.; Olson, J. S.; Springer, B. A.; Egeberg, K. D.; Sligar, S. G. *J. Biol. Chem.* **1990**, *265*, 3168. (e) Olson, J. S.; Phillips, G. N., Jr *J. Biol. Inorg. Chem.* **1997**, *2*, 544. (f) Lim, M. H.; Jackson, T. A.; Anfinrud, P. A. *J. Biol. Inorg. Chem.* **1997**, *2*, 531.

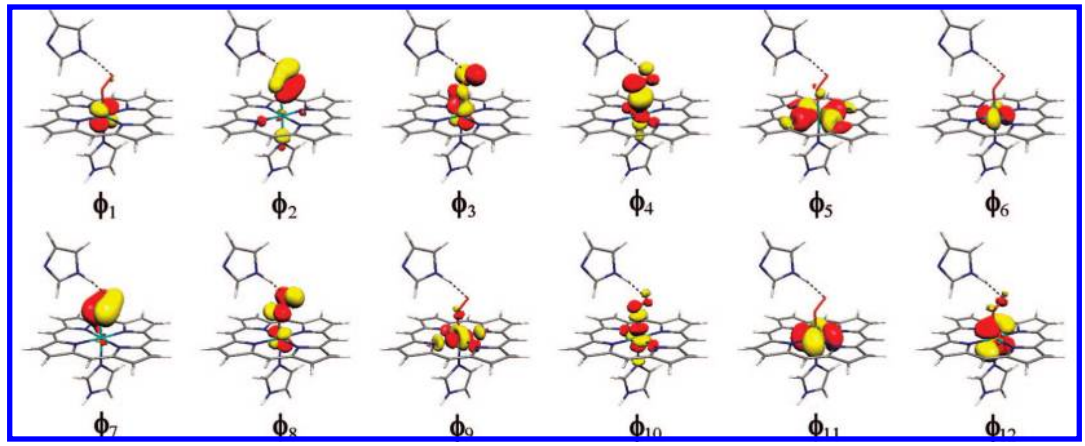



***Figure 2.*** CASSCF active orbitals of oxy-Mb in model I.

to Turbomole[44] and DL_POLY,[45] which together handle the QM(DFT) and MM calculations. A series of functionals, from pure GGA ones to hybrid ones including BP86,[46] BLYP, B3LYP,[47] PBE, and PBE0,[48] was used for the QM(DFT) part, and the CHARMM22 force field[49] was used for the MM part throughout the work. The active region of the protein, which is described by DFT, interacts with the rest of the protein (MM part) by electrostatic and Lennard−Jones interactions. The electronic embedding scheme[50] was used to account for the polarization effect of the QM part induced by the protein environment. The dangling bond at the QM/MM boundary was saturated by a hydrogen-link atom and treated in the framework of the charge-shift method.[50]

DFT/MM geometry optimizations were performed using a basis set called B1, which is comprised of the Wachters all-electron basis set[51] augmented with diffuse d and polarization f functions (8s7p4d1f), on iron, a double-$\zeta$ basis set augmented with polarization and diffuse functions, 6-31+G(d),[52] on the first coordination sphere, and 6-31G[52] on the rest of the atoms. The energy was corrected by QM(DFT)/MM single-point calculations with a larger basis set B2, which is identical to B1 with respect to iron, while the rest of the atoms are described by 6-31++G(d,p) basis set.[52] To test the effect of the environment on the electronic structure we also calculated a gas-phase model with DFT (so-called "model II"; see next subsection).

**Charge Analyses.** The NBO charge population analysis was employed in this work.[53] The orbital plot was done using the graphic program MOLEKEL.[54]

**Spin Contamination.** Due to the spin contamination of the symmetry-broken UDFT results for the lowest open-shell singlet state, we applied Yamaguchi's spin-projected correction[55] for our calculated energy of the symmetry-broken singlet state as shown below

$$E_S = (E_C - aE_{S+1})/(1-a), a = [\langle S^2\rangle_C - s(s+1)]/2(s+1) \quad (1)$$

$E_C$ is the spin-contaminated energy for the singlet state, $E_{S+1}$ is the energy for the triplet state (it has only small spin contamination), and $\langle S^2\rangle_C$ is the calculated spin expectation value of the spin-contaminated singlet state.

**2.2. CASSCF/MM and CASSCF(g) Methodology and Software.** The optimized geometry and point charges generated from the DFT/MM calculation were used as input for CASSCF/MM calculations, which were performed by the Gaussian 03 program.[56] We tried several kinds of active spaces for the CASSCF before we determined the final choice of the 14 electrons distributed in 12 orbitals. The active orbitals consists of iron $3d_{xy}$, $3d_{yz}$, $3d_{xz}$, $3d_{x^2-y^2}$, $3d_{z^2}$, $4d_{xy}$, and $4d_{xz}$, four O−O $\pi$, $\pi^*$ (perpendicular to and in the Fe−O−O plane), that is, five iron 3d orbitals, two 4d orbitals, four $\pi$ and $\pi^*$ of $O_2$, and one $b_{1g}$ type (using $D_{4h}$ symmetry porphine terminology) combination of the ligand orbital of heme nitrogen atoms which form a Fe−N (Porp) $\sigma$ bond. The two iron 4d orbitals corresponding to the occupied $3d_{xy}$ and $3d_{xz}$ orbitals were also included in the active space to account for the double-shell effect (the radial correlation in the d shell).[57] These CASSCF canonical orbitals are shown in Figure 2, and their main constituents are listed in Table 1.

To explore the possible effect exerted by the protein environment on the oxyheme at the CASSCF level we performed both CASSCF/MM calculations, referred to as "model I", as well as a gas-phase species, "model II", which does not include the point charges generated in the QM/MM calculation nor does it include a His64 moiety that is H bonded to the $O_2$ moiety. Thus, this latter

(43) ChemShell 2.05b4; Sherwood, P.; et al. *J. Mol. Struct. (THEOCHEM)* **2003**, *632*, 1.
(44) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165.
(45) Smith, W.; Forester, T. *J. Mol. Graph.* **1996**, *14*, 136.
(46) (a) Becke, A. D. *J. Chem. Phys.* **1986**, *84*, 4524. (b) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822.
(47) (a) Becke, A. D. *Phys. Rev. A* **1988**, *36*, 3098. (b) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785. (c) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648. (d) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.
(48) (a) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865. (b) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1997**, *78*, 1396. (c) Perdew, J. P.; Ernzerhof, M.; Burke, K. *J. Chem. Phys.* **1996**, *105*, 9982.
(49) MacKerell, A. D., Jr; et al. *J. Phys. Chem. B* **1998**, *102*, 3586.
(50) Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996**, *100*, 10580.
(51) (a) Wachters, A. J. H. *J. Chem. Phys.* **1970**, *52*, 1033. (b) Hay, P. J. *J. Chem. Phys.* **1977**, *66*, 4377. (c) Bauschlicher, C. W., Jr.; Langhoff, S. R.; Partridge, H.; Barnes, L. A. *J. Chem. Phys.* **1989**, *91*, 2399.
(52) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257.
(53) (a) Glendening, E. D.; Reed, A. E.; Carpenter, J. E.; Weinhold, F. *NBO*, Version 3.1. (b) Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735.
(54) Flükiger, P.; Lüthi, H. P.; Portmann, S.; Weber, J. *MOLEKEL 4.3*; Swiss Center for Scientific Computing: Manno, Switzerland, 2000, 2002.
(55) Yamaguchi, K.; Jensen, F.; Dorigo, A.; Houk, K. N. *Chem. Phys. Lett.* **1988**, *149*, 537.
(56) Frisch, M. J.; et al. *Gaussian 03*, Revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.
(57) Roos, B. O.; Andersson, K.; Fülscher, M. P.; Malmqvist, P. A.; Serrano-Andrés, L.; Pierloot, K.; Merchán, M. *Adv. Chem. Phys.* **1996**, *93*, 219.

***Table 1.*** Characters of the CASSCF Active Orbitals

| MO | type[a] |
|---|---|
| $\phi_1$ | Fe $3d_{xz}$ |
| $\phi_2$ | $O_2$ $\pi$ |
| $\phi_3$[b] | Fe $3d_{yz}$ + $O_2$ $\pi_\perp$* |
| $\phi_4$ | Fe $3d_{z^2}$ + $O_2$ $\pi$* |
| $\phi_5$ | (Por) N$\sigma$ + Fe $3d_{x^2-y^2}$ |
| $\phi_6$ | Fe $3d_{xy}$ |
| $\phi_7$[b] | $O_2$ $\pi_\perp$ |
| $\phi_8$[b] | Fe $3d_{yz}$ − $O_2$ $\pi_\perp$* |
| $\phi_9$ | Fe $3d_{x^2-y^2}$ − (Por) N$\sigma$ |
| $\phi_{10}$ | Fe $3d_{z^2}$ − $O_2$ $\pi$* |
| $\phi_{11}$ | Fe $4d_{xy}$ |
| $\phi_{12}$ | Fe $4d_{xz}$ |

[a] The approximately $z$ axis points to the Fe−O direction, and the $x$ and $y$ axes point to two nearly orthogonal Fe−$N_{porp}$ bond directions. [b] $O_2$ $\pi_\perp$ and $\pi_\perp$* represent O−O $\pi$ bonding and antibonding orbitals perpendicular to the Fe−O−O plane; $O_2$ $\pi$ and $\pi$* represent O−O $\pi$ bonding and antibonding orbitals in the Fe−O−O plane.

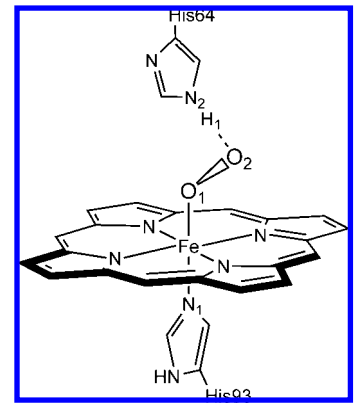
calculation corresponds to the gas-phase situation and will also be referred to as CASSCF(g). Notably, the CASSCF(g) calculations were done on the DFT/MM-optimized geometry of the oxyheme system. No symmetry constraint was used in our calculation.

Ab initio multiconfigurational calculations require extended basis sets to obtain reliable results. Accordingly, the CASSCF calculations used the following basis set: for iron we employed a cc-pvdz basis set (6s5p3d1f) for first-row transition metal developed by Peterson et al.;[58] for the five nitrogen atoms and two oxygen atoms in the immediate coordination sphere of iron we used the Dunning's cc-pvdz[59] basis set; for the rest of the atoms the 6-31G[52] basis set was used. The total number of basis functions for model I is 442.

**2.3. Setup of the System.** To prepare suitable initial structures for the QM/MM calculations we started from our recently determined X-ray structure of the oxy sperm whale Mb species (PDB code 2Z6S).[37d] We built a complete model of solvated protein by adding missing hydrogen atoms and a 16 Å thick water solvent layer. The entire system consisted of 12 344 atoms, including 9315 atoms in the solvent part. The system was further prepared according to standard procedures (see Supporting Information). We performed the QM/MM calculations on the coordinates obtained after simple force field energy minimizations and short MD on the inner solvent layer (in order to remove close contacts).

**2.4. Mössbauer Spectroscopic Calculations.** Since Mössbauer spectra of oxy-Mb are available and the sign of the quadrupole coupling constant is negative and can save a probe of the correctness of the calculations, a full Mössbauer calculation was done for the various DFT/MM calculations. The Mössbauer parameters were evaluated with the program ORCA[60] using single-point B3LYP and BP86 calculation at the DFT/MM-optimized geometries. The isomer shift was evaluated from the electron density at the iron nucleus.[61] Iron was described by the triply polarized core properties basis set CP(PPP),[61] and the other atoms were described by the SV(P) basis set[62] with the inner s functions left uncontracted. For the iron atom, an enhanced integration grid was used, and the overall integration accuracy was increased to 7. MM point charges were included in these calculations to probe the effect of the protein environment.

***Scheme 3.*** Schematic View of the QM Region Used in the QM/MM Calculations




**2.5. Charged Residues and Total Charge of the System.** The total charge of the so-generated system was −1, corresponding to the following protonation states of various residues: aspartic acid (Asp) and glutamic acid (Glu) residues were assumed to be ionized, i.e., negatively charged (Asp20, Asp27, Asp44, Asp60, Asp122, Asp126, Asp141, Glu4, Glu6, Glu18, Glu38, Glu41, Glu52, Glu54, Glu59, Glu83, Glu85, Glu105, Glu109, Glu136, Glu148), and arginine (Arg) and lysine (Lys) residues were used as positively charged (Arg31, Arg45, Arg118, Arg139, Lys16, Lys34, Lys42, Lys47, Lys50, Lys56, Lys62, Lys63, Lys77, Lys78, Lys79, Lys87, Lys96, Lys102, Lys133, Lys140, Lys145, and Lys147). On the basis of p$K_a$ estimation using the PROPKA program[63] plus careful visual checking of the environment of each histidine residue, all histidine residues were singly protonated; hence, each was kept neutral. There are two negative charges on the heme moiety of the propionate side chain of the protoporphyrin.

**2.6. QM Region and Optimized QM/MM Region.** Our chosen QM part was comprised of an iron−oxy porphyrin complex (without the side chains of heme) with its proximal ligand His93 (modeled as imidazole) and distal ligand His64 (modeled as imidazole), which forms a H bond with the $O_2$ coordinated to the iron. The QM region is illustrated in Scheme 3.

In addition to the QM region, 27 additional residues and 5 crystal waters around the heme were allowed to relax in the QM/MM geometry optimization. As a result, the total QM/MM region included Ile28, Leu29, Leu32, Phe33, Lys42, Thr39, Leu40, Phe43, Arg45, Phe46, Asp60, Leu61 His64, Thr67, Val68, Leu69, Ala71, Leu72, Leu89, Ser92, His93 (for technical reasons, this residue includes oxyheme), His97, Ala94, Ile99, Tyr103, Leu104, Ile107, Ile111, Phe138, Wat442, Wat454, Wat456, Wat509, Wat552.

## 3. Results

To explore the electronic structure of oxy-Mb, we presented our results in two separate sections. Section 3.1 summarizes the results of the DFT/MM calculations, while section 3.2 involves the CASSCF/MM results. We focus on the main results here, and more technical details were collected in the Supporting Information deposited with this paper.

**3.1. DFT/MM. 3.1.1. Electronic Structures and Optimized Geometries.** In agreement with previous UDFT calculations on the oxyheme system,[34,37a,d] our UDFT/MM calculations indicate that the ground state of oxy-Mb is an open-shell singlet state. Scheme 4 depicts the electronic structure of the lowest singlet and triplet states of oxy-Mb in terms of orbital occupancy diagrams. One can see that in the symmetry-broken UDFT/MM calculation the lowest singlet and triplet states possess similar

(58) (a) Balabanov, N. B.; Peterson, K. A. *J. Chem. Phys.* **2005**, *123*, 064107. (b) Balabanov, N. B.; Peterson, K. A. *J. Chem. Phys.* **2006**, *125*, 074110.
(59) (a) Dunning, T. H., Jr *J. Chem. Phys.* **1989**, *90*, 1007. (b) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796. (c) Woon, D. E.; Dunning, T. H., Jr *J. Chem. Phys.* **1993**, *98*, 1358.
(60) Neese, F. *ORCA*, Version 2.4, Revision 10; Max-Planck-Institut für Bioanorganische Chemie: Mülheim a.d. Ruhr, Germany, 2004.
(61) Neese, F. *Inorg. Chim. Acta* **2002**, *337*, 181.
(62) Schäfer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1992**, *97*, 2571.
(63) Li, H.; Robertson, A. D.; Jensen, J. H. *Proteins* **2005**, *61*, 704.

***Scheme 4.*** Electronic Configurations of the Lowest Singlet and Triplet States in DFT/MM Calculations of Oxy-Mb

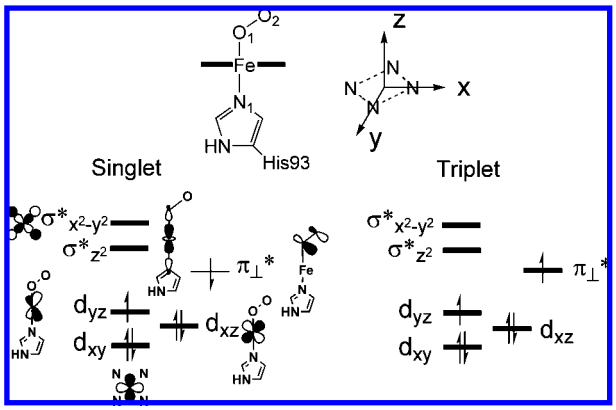



***Table 2.*** Computed Relative Energies (kcal/mol) of Oxy-Mb at the DFT/MM B2 Levels after Spin-Projected Energy Correction for the Lowest Singlet State

| | DFT/MM[a] | | | | |
|---|---|---|---|---|---|
| state | B3LYP | PBE0 | BP86 | BLYP | PBE |
| singlet | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| triplet[a] | 7.36 | 7.80 | 3.95 | 3.41 | 3.99 |
| quintet-1[a,b] | 24.75 | 23.09 | —[d] | —[d] | —[d] |
| quintet-2[a,c] | 24.86 | 22.35 | 28.39 | 27.72 | 28.16 |

[a] Using the structure optimized for the singlet ground state at the corresponding DFT/MM B1 level. [b] With configuration $d_{xy}^2 d_{yz}^1 \sigma^*_{z^2}{}^1 d_{xz}^1 \pi^*_\perp{}^1$. [c] With configuration $d_{yz}^2 d_{xy}^1 \sigma^*_{x^2-y^2}{}^1 d_{xz}^1 \pi^*_\perp{}^1$. [d] Could not obtain the corresponding configuration.

***Table 3.*** Spin Distribution of the Lowest Singlet and Triplet Oxy-Mb at the DFT/MM B2 Levels[a]

| functional | spin state | Fe | $O_1$ | $O_2$ | His93 | His64 | Porp |
|---|---|---|---|---|---|---|---|
| B3LYP | singlet | 0.98 | −0.40 | −0.58 | 0.01 | 0.00 | −0.01 |
| | triplet | 1.07 | 0.38 | 0.59 | 0.00 | 0.00 | −0.04 |
| PBE0 | singlet | 0.99 | −0.40 | −0.59 | 0.01 | 0.00 | −0.02 |
| | triplet | 1.15 | 0.39 | 0.56 | −0.01 | 0.00 | −0.08 |
| BP86 | singlet | 0.67 | −0.28 | −0.42 | 0.02 | 0.00 | 0.01 |
| | triplet | 1.21 | 0.38 | 0.42 | 0.01 | 0.00 | −0.01 |
| BLYP | singlet | 0.71 | −0.30 | −0.44 | 0.01 | 0.00 | 0.02 |
| | triplet | 1.19 | 0.38 | 0.41 | 0.01 | 0.00 | 0.01 |
| PBE | singlet | 0.68 | −0.28 | −0.42 | 0.01 | 0.00 | 0.01 |
| | triplet | 1.21 | 0.38 | 0.41 | 0.01 | 0.00 | −0.01 |

[a] Using the structure optimized for the singlet ground state at the corresponding DFT/MM B1 level.

electronic configurations. The only difference is that the unpaired electrons in iron $3d_{yz}$ and $\pi_\perp^*$ of $O_2$ are ferromagnetically coupled in the triplet state but antiferromagnetically coupled in the singlet state.

Table 2 collects the DFT/MM B2 calculated vertical energy gaps of the lowest lying triplet and quintet states, while Table 3 displays the corresponding spin populations. As seen in Table 2 the lowest triplet state resides a few kcal/mol higher than the ground singlet state. After spin projection for correction of the spin-contaminated singlet state (see Tables S2−S5, Supporting Information), the energy gap gets systematically larger in the hybrid functional case (7−8 kcal/mol) than that in the GGA functional case (3−4 kcal/mol). We noted that the spin projection of the symmetry-broken UDFT calculation is still being debated.[64] Without a convincing experimental result for the existence of a low-lying triplet state in Mb,[65] it is still not possible to judge whether the result after spin projection is better or worse than that before it. However, as will be shown in section 3.2.1, the singlet−triplet gap after spin projection is consistent with the result at the CASSCF/MM level while the uncorrected value is not. For the quintet state our DFT/MM calculation produced two kinds of configurations with the hybrid functionals. The difference between these quintet states is that in quintet-1 two unpaired electrons reside in $d_{yz}$ and $\sigma^*_{z^2}$ orbitals, whereas in the quintet-2 two unpaired electrons reside in $d_{xy}$ and $\sigma^*_{x^2-y^2}$ orbitals. The vertical energy gaps of the quintet states vis-à-vis the ground singlet state are much larger than 20 kcal/mol. In Table 3 the spin population provides detailed evidence about the above electronic configuration assignment (For the natural orbitals see Figure S1 in the Supporting Information).

***Table 4.*** DFT/MM-Optimized Structures (Bond Lengths [Angstroms], Bond Angles [deg]) of Oxyheme in the Singlet Ground State of Oxy-Mb Using Different Functionals for the QM Part

| | DFT/MM | | | | | |
|---|---|---|---|---|---|---|
| variable[a] | B3LYP | PBE0 | BP86 | BLYP | PBE | crystal structure |
| $Fe-O_1$ | 1.868 | 1.856 | 1.799 | 1.819 | 1.796 | 1.83[c]/1.81[d] |
| Fe−N (Por)[b] | 2.038 | 2.025 | 2.035 | 2.051 | 2.035 | 1.98[c]/2.01[d] |
| $Fe-N_1$ (His93) | 2.032 | 2.009 | 2.033 | 2.061 | 2.033 | 2.08[c]/2.06[d] |
| $O_1-O_2$ | 1.310 | 1.292 | 1.322 | 1.338 | 1.318 | 1.25[c]/1.24[d] |
| $N_2-O_1$ (His64) | 2.998 | 2.975 | 3.022 | 3.021 | 3.020 | 2.93[c]/3.02[d] |
| $N_2-O_2$ (His64) | 2.788 | 2.761 | 2.754 | 2.785 | 2.752 | 2.76[c]/2.67[d] |
| $H_1-O_2$ | 1.772 | 1.744 | 1.722 | 1.757 | 1.720 | −/− |
| $\angle FeO_1O_2$ | 120.75 | 120.34 | 122.12 | 122.61 | 122.19 | 124.2[c]/122.5[d] |

[a] See Scheme 3 for atom labels. [b] Average value of four Fe−N bonds between Fe and porphyrin. [c] Reference 37d. [d] Reference 39.

Table 4 displays the optimized geometries for the ground singlet state at the DFT/MM B1 level alongside data for two available X-ray structures. It can be seen that generally the agreement with the experimental structures is good and there are neither large differences nor apparent bias among the employed functionals. Thus, the geometry optimized at the B3LYP/MM B1 level will be used in the subsequent parts of this paper.

The calculated NBO charge population indicates (see Table S8, Supporting Information) that there is apparent charge transfer to $O_2$, i.e., −0.51e at the B3LYP/MM level and −0.50e at the PBE0/MM level (both B2 data). Using GGA functionals the magnitude of the transferred charge is a little bit smaller (−0.41e to −0.44e at the B2 level depending on functional used) but still significant.

**3.1.2. Mössbauer spectroscopic parameters.** The computed Mössbauer spectroscopic parameters of oxy-Mb at the B3LYP/MM and BP86/MM levels are shown in Table 5 alongside experimental data.[66] The quadrupole splitting $\Delta E_Q$, chemical isomer shift $\delta$, and asymmetric parameter $\eta$ agree reasonably well with the experimental values. The magnitude of the calculated $\Delta E_Q$ is a little bit larger than the previous DFT (BP86) value for a gas-phase model calculation[27] (−2.52 mm/s) and

(64) (a) Ruiz, E.; Alvarez, S.; Cano, J.; Polo, V. *J. Chem. Phys.* **2005**, *123*, 164110. (b) Adamo, C.; Barone, V.; Bencini, A.; Broer, R.; Filatov, M.; Harrison, N. M.; Illas, F.; Malrieu, J. P.; Moreira, I. de P. R. *J. Chem. Phys.* **2006**, *124*, 107101. (c) Ruiz, E.; Alvarez, S.; Cano, J.; Polo, V. *J. Chem. Phys.* **2006**, *124*, 107102.

(65) (a) Savicki, J. P.; Lang, G.; Ikeda-Saito, M. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 5417. (b) Cerdonio, M.; Congiu-Castellano, A.; Mogno, F.; Pispisa, B.; Romani, G. L. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 398. (c) Cerdonio, M.; Congiu-Castellano, A.; Calabrese, L.; Morante, S.; Pispisa, B.; Vitale, S. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 4916.

(66) Boso, B.; Debrunner, P. G.; Wagner, G. C.; Inubushi, T. *Biochim. Biophys. Acta* **1984**, *791*, 244.

***Table 5.*** Calculated Mössbauer Spectroscopic Parameters for Oxy-Mb at the DFT/MM Level

| | $\Delta E_Q$ (mm/s) | $\delta$ (mm/s) | $\eta$ |
|---|---|---|---|
| oxy-Mb[a] | −2.758 | 0.331 | 0.056 |
| oxy-Mb[b] | −2.671 | 0.356 | 0.049 |
| experimental values[c] | −2.28 | 0.27 | 0.01 |

[a] Calculated at the B3LYP/MM level using the optimized geometry at the B3LYP/MM B1 level. [b] Calculated at the BP86/MM level using the optimized geometry at the B3LYP/MM B1 level. [c] Reference 66 from sperm whale oxy-Mb at 4.2 K.

***Table 6.*** Computed Relative Energies (kcal/mol) of the Singlet and Triplet (T-1 and T-2) States of Oxy-Mb at the CASSCF/MM and CASSCF(g) Levels

| state | CASSCF/MM[a] | CASSCF(g)[a] |
|---|---|---|
| singlet | 0.00 | 0.00 |
| T-1[b,c] | 8.78 | 12.23 |
| T-2[b,d] | 15.35 | 7.25 |

[a] Using the structure optimized for the singlet ground state at the B3LYP/MM B1 level, active space (14,12) used. [b] Triplet states, obtained from state-averaged calculation with equal weight for the two triplet states. [c] The dominant configuration in T-1 is $d_{xy}^2 d_{xz}^2 d_{yz}^1 \pi^*_\perp{}^1$. [d] T-2 has a substantial multiconfigurational character.

also near the data of a similar conformation of $O_2$ and proximal His93 from previous DFT calculations for a variety of structures.[67] The calculated $\eta$ agrees with the experimental value in Mb better than the previous DFT (BP86) gas-phase result (0.24).[27]

**3.2. CASSCF/MM and CASSCF(g) Calculations.** As mentioned in section 2, to explore the effect exerted by the protein environment on the electronic structure of oxy-Mb, we probed two models with the CASSCF method. The first model (model I) is same as that in the DFT/MM calculation, thus leading to CASSCF/MM calculation. While the second one (model II), on which CASSCF(g) calculations are based, is generated from model I by neglecting the polarization of the protein environment (point charge of the MM part) and omitting the important distal His64 residue, i.e., similar with the usual model systems adopted in many previous calculations, but keeping the DFT/MM-optimized geometry for oxyheme. By CASSCF/MM and CASSCF(g) calculations we tried to uncover the nature of the $Fe-O_2$ bond in oxy-Mb. Since the above shown B3LYP/MM-optimized geometry agrees well with the experimental structure, we used this structure in the CASSCF/MM and CASSCF(g) calculations presented below.

**3.2.1. Calculated Relative Energies and Properties.** Table 6 collects the calculated relative energies of the singlet and two lowest triplet states for oxy-Mb at the CASSCF/MM and CASSCF(g) levels. It is seen that irrespective of the environment the ground state is a singlet state. However, the triplet state behaves in a manner that depends on the environment: Thus, with CASSCF/MM (model I) the lowest triplet state is T-1, which has a dominant configuration $d_{xy}^2 d_{xz}^2 d_{yz}^1 \pi^*_\perp{}^1$, lying 8.78 kcal/mol higher than the ground singlet state; a second state, T-2, with a significant multiconfiguration character lies higher in energy, ca. 15.35 kcal/mol. By contrast, in the CASSCF(g) calculation the state ordering is inverted; the lowest triplet state is T-2 while T-1 lies higher. As just mentioned, T-1 is the triplet state with a single configuration akin to Scheme 4 ($d_{xy}^2 d_{xz}^2 d_{yz}^1 \pi^*_\perp{}^1$) while

(67) Godbout, N.; Sanders, L. K.; Salzmann, R.; Havlin, R. H.; Wojdelski, M.; Oldfield, E. *J. Am. Chem. Soc.* **1999**, *121*, 3829.

***Table 7.*** NBO Charge Population of the Low Singlet and Triplet States in Oxy-Mb at the CASSCF/MM CASSCF(g) Levels[a]

| | spin state | Fe | $O_1+O_2$ | His93 | His64 | Porp |
|---|---|---|---|---|---|---|
| CASSCF/MM | singlet | 1.37 | −0.44 | 0.18 | −0.04 | −1.06 |
| | T-1 | 1.42 | −0.54 | 0.19 | −0.04 | −1.02 |
| | T-2 | 1.54 | −0.46 | 0.17 | −0.04 | −1.21 |
| CASSCF(g) | singlet | 1.28 | −0.22 | 0.13 | | −1.20 |
| | T-1 | 1.42 | −0.16 | 0.11 | | −1.37 |
| | T-2 | 1.49 | −0.06 | 0.10 | | −1.53 |

[a] Using the structure optimized for the singlet ground state at the B3LYP/MM B1 level. The CASSCF used an active space of (14,12).

T-2 has a substantial multiconfigurational character (see orbital occupation number in Figures S3 and S5, Supporting Information), precisely the same as in the recent gas-phase CASSCF(g) calculation using a model system similar to model II.[28] Thus, our calculation indicates the CASSCF/MM and CASSCF(g) calculations produce a substantially different lowest triplet state. Interestingly, B3LYP/MM gave the same T-1 state as CASSCF/MM and virtually the same excitation energy gap. As such, while CASSCF/MM and B3LYP/MM are in close agreement, the CASSCF(g) results are substantially different, thereby reflecting the impact of the H bond between the distal oxygen and His64 and the bulk polarity of the protein pocket.

The calculated NBO charge distributions are collected in Table 7. One can see that the calculated charge population for the $O_2$ moiety from CASSCF(g) is quite small, i.e., −0.22e for the ground singlet state. This is consistent with the recent gas-phase CASSCF calculations[27] (−0.20e) using a similar model system as model II for CASSCF(g). However, the CASSCF/MM calculations yield a substantially higher $O_2$ charge (−0.44e): twice as large as that coming from the CASSCF(g) calculations. On the other hand, this CASSCF/MM $O_2$ charge value is very close to the B3LYP/MM result (−0.51e) for model I. The calculated charge population of the $O_2$ moiety is more negative for T-1 than that for T-2. This explains qualitatively why in CASSCF/MM, where the H-bond interaction between His64 and $O_2$ moiety is present, T-1 is lower in energy than T-2.

To assess the impact of the protein environment on DFT, we performed gas-phase DFT single-point calculation on the model system of model II without His64 and point charge of the protein (oxyheme still hold the corresponding DFT/MM optimized geometry), hence DFT(g) calculations. As expected the $O_2$ NBO charge in the singlet ground state, ca. −0.41e, at the B3LYP/B2 level, is somewhat smaller than in DFT/MM. Furthermore, in the gas phase the GGA functionals (BP86, BLYP, and PBE) gave a completely different triplet state compared with GGA/MM; this triplet possesses two singly occupied orbitals composed of two perpendicular $\pi^*$-type orbitals of the $O_2$ moiety ($\pi_\perp^*$ and $\pi^*$) in combination with two iron $3d_{yz}$ and $3d_{xz}$ orbitals, respectively, thus conserving the triplet $O_2$ character. If the system has $C_s$ symmetry as in many previous gas-phase DFT calculations, this triplet state will be of different symmetry to the triplet state coming from GGA/MM calculation. Indeed, the same findings were reported in previous gas-phase DFT calculation using the GGA (BP86) functional on an oxyheme model system similar with our model II.[37a] With the hybrid functionals (B3LYP and PBE0) the lowest triplet state is the same as shown our DFT/MM calculation; however, the energy of this triplet state is somewhat higher than in DFT/MM (see Table S6, Supporting Information). All these variations highlight the importance of taking into account the effects of His64

***Table 8.*** Four Most Important Configuration State Functions (CSF) of the Lowest Singlet and Triplet States of Oxy-Mb from CASSCF (14,12) Calculation at the CASSCF/MM Level

| configuration | singlet | | triplet (T-1) | |
|---|---|---|---|---|
| CASSCF/MM | CSF[a] | weight (%) | CSF[a] | weight (%) |
| 1 | 222222200000 | 65.9 | 22u2222u0000 | 89.4 |
| 2 | 220222220000 | 23.2 | 22u0222u0200 | 1.4 |
| 3 | 22du222u0d00 | 0.7 | 02u2222u2000 | 0.5 |
| 4 | 222022200200 | 0.6 | u2222u2ud200 | 0.5 |
| CASSCF(g) | CSF[a] | weight (%) | CSF[a,b] | weight (%) |
| 1 | 222222200000 | 65.0 | 22u2222u0000 | 78.9 |
| 2 | 220222220000 | 16.2 | 22u0222u0200 | 5.5 |
| 3 | 22du222u0d00 | 2.5 | 222u22200u00 | 1.4 |
| 4 | 222022200200 | 1.6 | 2d2222uu0u00 | 0.8 |

[a] The CSFs are eigenfunctions of the spin operators and labeled in terms of occupations per orbital: The orbitals appear in the ordering of Figure 2 and correspond to $\phi_1-\phi_{12}$ from left to right. The numbers and letters in each digit correspond to the number of electrons and spin direction (e.g., "2" means a doubly occupied orbital, "u" means a singly occupied orbital with α spin, "d" means a singly occupied orbital with $\beta$ spin, and "0" means an unoccupied orbital). [b] T-1 is the second excited triplet state in CASSCF(g).

and the protein polarization as done here in the DFT/MM calculation. Thus, explicit consideration of His64 and the protein polarization could substantially change the electronic configuration of the lowest triplet state within the DFT approach as well as within the CASSCF method.

**3.2.2. CASSCF/MM and CASSCF(g) Wave Functions for the Lowest Singlet and Triplet State.** To better understand the nature of the bonding between the iron and the $O_2$ moiety it is necessary to analyze the CASSCF wave functions. Table 8 summarizes the four most important configurations with the highest weight for the CASSCF/MM calculations. It is seen that the singlet ground state is described by two most significant configurations with a total weight of about 90%; the next important configurations have a weight of less than 1%. For the triplet state T-1 the CASSCF/MM wave function has one most important configuration with a weight of more than 89%. Thus, now it is apparent that in CASSCF/MM the ground singlet state is a two-configuration SCF (TCSCF) wave function with a tiny contribution from other configurations, which may be considered as small perturbations on the TCSCF wave function. On the other hand, the CASSCF/MM for the lowest triplet state is a single configuration wave function with small perturbations from other configurations. The CASSCF(g) calculations give very similar results for the ground state: here the weight of the two major configurations is somewhat smaller (81%) than in the CASSCF/MM case, and the second most important configuration 2 is less important than in the CASSCF/MM case. More so, the lowest triplet state of the CASSCF is not T-1 but rather the multiconfigurational state T-2. In CASSCF(g) calculation T-1 (now being the second triplet) is also dominated by one configuration but with less weight (79%) than that in the CASSCF/MM case (89%). The impact of the protein is then apparent on all aspects of the CASSCF calculations, energy, $O_2$ charge, configurational content, and triplet-state identity.

Let us inspect the nature of the most important configurations of the CASSCF wave function in Table 8. For the lowest singlet state, configurations 1 and 2 in Table 8 are both closed shell configurations in the sense of the totally doubly occupied MO. The only difference between configurations 1 and 2 is that in the former one orbital $\phi_3$ is doubly occupied and orbital $\phi_8$ is vacant, whereas in the latter the two electrons in orbital $\phi_3$ are excited to orbital $\phi_8$. On inspection of Figure 2 one can find that $\phi_3$ and $\phi_8$ are the bonding and antibonding combinations between iron $3d_{yz}$ and $\pi_\perp$* of $O_2$ that are both perpendicular to the FeOO plane (there is some $\pi_\perp$ contribution that polarizes the $O_2$ component orbitals). The significant weight of configuration 2 in which the "antibonding" orbital $\phi_8$ is doubly occupied indicates that the Fe−$O_2$ bond cannot be described by a normal double bond with one doubly occupied $\pi$-type bonding orbital. Indeed, the calculated occupation numbers of orbitals $\phi_3$ and $\phi_8$ for the singlet ground state at the CASSCF/MM level are 1.46 and 0.58, respectively, being far from 2 or 0. This is another indication of the importance of these two orbitals for the static electron correlation in the CASSCF wave function of oxy-Mb. The fact that only two different orbitals are involved in the two main configurations of CASSCF/MM wave function also indicates that except for these two orbitals the other parts of the electronic structure of oxy-Mb could be qualitatively well described by the MO SCF procedure. We will discuss this issue again later in this work. For T-1, as shown in Table 8, its dominant configuration 1 only consists of singly occupied orbitals $\phi_3$ and $\phi_8$, which generate the identical configuration as the one obtained from DFT/MM calculation (see also Figure S5 in the Supporting Information for the orbital occupation number of T-1).

***Table 9.*** Comparison of the B3LYP/MM, CASSCF/MM, and CASSCF(g) Calculated Relative Energies of the States of Oxy-Myoglobin (kcal/mol) and $O_2$ NBO Charge Populations (in Parentheses)[a]

| | relative energy ($O_2$ charge) | | |
|---|---|---|---|
| state | CASSCF/MM[b] | CASSCF(g)[b] | B3LYP/MM[c] |
| singlet | 0.00 (−0.44e) | 0.00 (−0.22e) | 0.00 (−0.51e) |
| T-1[e] | 8.78[d] (−0.54e) | 12.23[d] (−0.16e) | 7.36 (−0.55e) |
| T-2[f] | 15.35[d] (−0.46e) | 7.25[d] (−0.06e) | —[g] |

[a] Using the structure optimized for the singlet ground state at the B3LYP/MM B1 level. [b] Active space (14,12) used. [c] At the B3LYP/MM B2 level. [d] State-averaged calculation with equal weight for two triplet states. [e] With configuration $d_{xy}^2d_{xz}^2d_{yz}^1\pi^*_\perp{}^1$ as only (DFT) or dominant (CASSCF) configuration. [f] Having substantial multiconfigurational character. [g] This configuration failed to converge and could not be obtained.

## 4. Discussion

In order to focus the discussion we assembled in Table 9 the key properties of the singlet and triplet states for the CASSCF/MM, CASSCF(g), and B3LYP/MM methods. From Table 9 it is apparent that CASSCF/MM and B3LYP/MM calculations are more consistent with each other, while CASSCF(g), which lacks the effect of the H bonding with His64 and the polarity of the protein, is quite different.

In summary, the protein environment effect is mainly composed of the distal His64 residue (E7) that forms a H bond with the $O_2$ moiety and electrostatic polarization of the protein backbone. The impact of neglecting the above two aspects in the calculation is apparent in Tables 7−9, and it is common for both CASSCF and DFT levels. Removal of the protein environment causes the following changes: (i) It reduces the amount of negative charge on $O_2$ by a factor of 2 in CASSCF, (ii) it reduces the total weight of the two major configurations in the CASSCF wave function, and (iii) it causes switching in the identity of the lowest triplet state both in CASSCF as well as in GGA functionals. Let us then turn to inspect the Fe−$O_2$ bonding in the protein as well as in the gas phase.

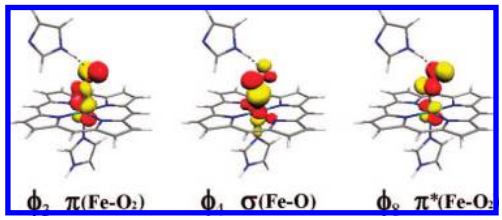



***Figure 3.*** Three key orbitals describing the Fe−$O_2$ bonding in the CASSCF calculations.

**4.1. Nature of the Fe−$O_2$ Bond in Oxy-Mb.** We recall from Scheme 1 that all the alternative models of bonding are derived from valence bond (VB) concepts and that, in fact, there is still no consensus regarding the nature of the Fe−$O_2$ bond. Since we now have a CASSCF/MM wave function that takes into account the protein influence we can take advantage of this facility and derive a VB model that is tuned to the protein effect. We recall that the CASSCF/MM wave function is dominated by two configurations and is, hence, a two-configuration SCF (TCSCF) wave function (about 90% by weight), where all other configurations have tiny contributions and can be viewed as perturbations on the TCSCF wave function. Figure 3 shows the three orbitals that are responsible for the Fe−$O_2$ bonding in the TCSCF wave function.

One of these orbitals, $\phi_4$, is a $\sigma$(Fe−O) orbital made from an overlap of the 2p orbital of the proximal oxygen with the $3d_{z^2}$ orbital on iron, and it is doubly occupied in the two configurations of the TCSCF wave function. The other two, $\phi_3$ and $\phi_8$, account for the $\pi$ bonding and are alternately doubly occupied/vacant in the two configurations; configuration 1 has a doubly occupied $\phi_3$ and a vacant $\phi_8$, while configuration 2 has a doubly occupied $\phi_8$ and a vacant $\phi_3$. Furthermore, the two orbitals $\phi_3$ and $\phi_8$ constitute a bonding−antibonding pair. Thus, since all other orbitals, including $\phi_4$, are doubly occupied, the TCSCF wave function can be expressed as follows

$$\Psi = c_1|\cdots\phi_3\bar{\phi}_3| - c_2|\cdots\phi_8\bar{\phi}_8| \qquad (2)$$

where $\phi_3$ and $\phi_8$ are the corresponding $\pi$(Fe−$O_2$) and $\pi^*$(Fe−$O_2$) orbitals in Figure 3 and the three dots represent the other closed shell part. As such, much like the classical textbook problem of the $H_2$ bonding, after CI the TCSCF wave function of eq 2 can be transformed (see the Supporting Information for details) into a GVB-type wave function[68] with two new orbitals that form a singlet pair, i.e., a bond pair, as follows

$$\Psi = \frac{1}{2\sqrt{1+S^2}}[|\cdots\psi_1\bar{\psi}_2| - |\cdots\bar{\psi}_1\psi_2|] \qquad (3)$$

Here, $\psi_1$ and $\psi_2$ are the transformed orbitals that are singly occupied in eq 3 and spin paired in a Heitler−London fashion to a bond pair. The term $S$ in the normalization constant is the overlap between the so transformed orbitals, i.e., $S = \langle\psi_1|\psi_2\rangle$.

The transformed $\psi_1$ and $\psi_2$ GVB pair of orbitals extracted from the CASSCF/MM wave function is shown in Figure 4a, and their calculated overlap $S$ is rather small, about 0.25. Inspection of these orbitals shows that they are almost pure fragment orbitals, each having a very little delocalization tail on the other fragment. Thus, one of them is almost a pure $3d_{yz}$ iron orbital, while the other is an almost pure $\pi_\perp^*$ orbital of $O_2$ (for a more quantitative analysis see the Supporting Information). As such, the bond pair is described by two electrons, one

(68) See chapter 3 of ref 19.

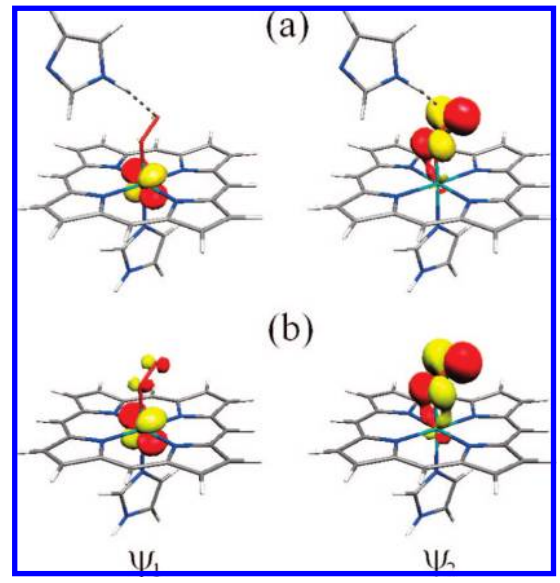



***Figure 4.*** $\psi_1$ and $\psi_2$ GVB-type orbitals transformed from CASSCF orbitals $\phi_3$ and $\phi_8$ (see Figure 3) for the (a) CASSCF/MM wave function and (b) CASSCF(g) wave function.

in $3d_{yz}$ and the other in $\pi_\perp^*$, that are spin paired and account for the $\pi$-type Fe−$O_2$ bonding. Clearly, this $\pi$ bonding arises by a single electron transfer from the heme-$Fe^{II}$ to $O_2$ with subsequent Heitler−London-type[19] pairing of the two odd electrons on the $Fe^{III}$ ($S = 1/2$) and $O_2^-$ ($S = 1/2$) fragments. This is precisely the Weiss bonding mechanism, and we may conclude that the $\pi$ bonding described by the CASSCF/MM wave function to a good approximation mirrors the Weiss mechanism. Thus, Fe−$O_2$ has a $\sigma$ bond in $\phi_4$ and a weakly coupled $\pi$ bond pair in $3d_{yz}$ and $\pi_\perp^*$. The weakness of this $\pi$ bond is indicated by the fact that the triplet state, T-1, in which the electrons in $3d_{yz}$ and $\pi_\perp^*$ share the same spin is only 8.8 kcal/mol higher than the singlet.

The orbitals in Figure 4b are the transformed orbitals of the TCSCF wave function of CASSCF(g). The resulting orbitals are similar to the ones in the protein environment (Figure 4a). However, the delocalization tails are more significant. Thus, the orbital denoted as $\psi_2$ orbital contains a significant amount of $3d_{yz}$ into $\pi_\perp^*$. Similarly, the orbital labeled $\psi_1$ involves $3d_{yz}$ with a significant amount of $\pi_\perp$ and a small amount of $\pi_\perp^*$. The significant delocalization tails in the gas-phase GVB orbitals (Figure 4b) cause, as expected, a larger overlap $S = 0.33$ compared with the corresponding GVB orbitals from CASSCF/MM wave function (Figure 4a) for which $S = 0.25$. The larger delocalization tails suggest the gas-phase oxy-Mb species has a somewhat stronger $\pi$-type Fe−$O_2$ bonding. This is confirmed by the larger energy gap between the singlet state and T-1 in the gas phase (15.4 kcal/mol) as well as by the smaller occupation number of MO $\phi_8$ that has an antibonding character (see Figures S2 and S4, Supporting Information). Still, though, the mechanism of bonding is a Weiss type, initiated by electron transfer and coupling of the $Fe^{III}$ ($S = 1/2$) and $O_2^-$ ($S = 1/2$). However, the significant delocalization tails of the orbitals tell us that in the coupling of the two fragments other modes of electron distribution make significant contributions to the final Fe−$O_2$ bonding.

**4.2. Why Is the Charge on $O_2$ Only −0.5?** While the above bonding feature better fits the Weiss mechanism, one might still wonder: what is the reason that the calculations yield an $O_2$ charge close to −0.5 and not −1.0 as expected from a

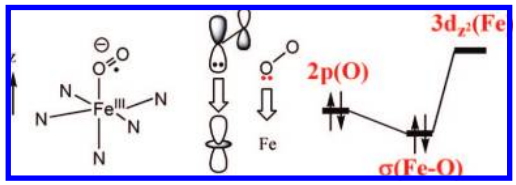



***Figure 5.*** Schematic representation of the $\sigma$(Fe−O) bond, which arises by donation from the $2p_z$ AO of the doubly filled $\pi^*(O_2^-)$ orbital to the vacant $3d_{z^2}$(Fe) orbital. On the right-hand side, this is shown using a perturbative orbital mixing diagram.

straightforward consideration of the Weiss model? To answer this question it is necessary to analyze the $\sigma$ bond between iron and the $O_2$ moiety, which is represented as the doubly occupied orbital $\phi_4$ in Figure 3 (actually its occupation number is 1.94 in the ground state in CASSCF/MM; see Figure S4 in the Supporting Information). In the Goddard ozone model (**6** in Scheme 1) it was proposed that the singly occupied iron $3d_{z^2}$ orbital forms a covalent $\sigma$ bond with the singly occupied $p_z$ orbital of $O_2$ in FeOO plane. However, our CASSCF/MM calculation shows that the $\sigma$ bond, represented by $\phi_4$, is more like a dative bond than a covalent bond. It is formed by mixing of the filled $\pi^*(O_2^-)$ orbital of the $O_2^-$ species and the vacant $3d_{z^2}$ orbital on the heme-$Fe^{III}$ moiety. As shown schematically in Figure 5, the mixed orbital should have a dominant $O_2^-$ character, but at the same time the mixing will result in some back charge transfer from $O_2^-$ into the empty $3d_{z^2}$ orbital of $Fe^{III}$, thereby diminishing the $O_2$ charge from −1.0 to closer to −0.5. Indeed, in orbital $\phi_4$ the calculated weight of the atomic orbital contributions from the $O_2$ moiety is almost three times as large as the contribution from the iron $3d_{z^2}$ orbital (see the Supporting Information). As such, the $\sigma$(Fe−O) bond in the protein environment is a highly polarized dative bond with a small covalent character.

Interestingly, in line with the back charge transfer in the mixing model in Figure 5, the calculated atomic orbital population (see Table S16, Supporting Information) shows that the $3d_{z^2}$ orbital in the CASSCF/MM calculations has an occupation number of 0.59. This can be compared with the significantly larger corresponding value in the CASSCF(g) calculation (0.72) and similarly larger value (0.68) from previous gas-phase CASSCF calculation.[27] Thus, the back charge transfer due to the orbital mixing in Figure 5 is much greater in the gas phase, and this is the reason why the CASSCF(g) calculations show an $O_2$ charge of only −0.2e. Clearly therefore, the protein environment increases the polar character of the $\sigma$(Fe−O) bond and converts it from a covalent bond in the gas-phase conditions to a dative bond. Altogether, the protein environment affects both $\sigma$(Fe−O) and $\pi$(Fe−$O_2$) bonding of Mb. We shall further address the difference between these bonding features and make predictions regarding the spectrum of Fe−$O_2$ bonding features expected for different enzymes. However, before doing so let us digress for a moment in order to show that DFT/MM describes the same Fe−$O_2$ bonding features as CASSCF/MM.

**4.3. DFT/MM and CASSCF/MM Method in Oxy-Mb.** From our DFT/MM and CASSCF/MM results (Tables 2, 6, 7, and 9) it is now apparent that after the effect of the protein environment is taken into account these two methods can give a consistent description for most of the features of oxy-Mb in the lowest singlet and triplet states. For example, as summarized in Table 9 the CASSCF/MM calculated charge population of the $O_2$ moiety is near the DFT/MM value; the calculated gap between the lowest singlet and triplet states is virtually identical for the

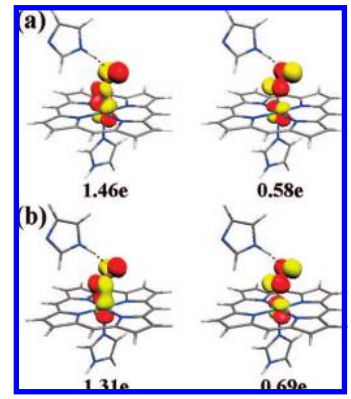



***Figure 6.*** Electron occupation numbers in the (a) canonical orbitals of CASSCF/MM and (b) natural KS orbitals of B3LYP/MM.

two methods. Furthermore, the triplet state T-1 is the same in both calculations (see Scheme 4 and Table 6). Similarly, in Figure 4 we show that the orbital transformation converts the CASSCF/MM wave function into an open-shell singlet with two singly occupied orbitals that are spin paired, much like the UDFT/MM description in Scheme 4. We can further show the striking similarity of the two methods by inspecting the canonical/natural orbitals and their occupation numbers, as done in Figure 6. It is seen that the resulting natural Kohn−Sham (KS) orbitals have very similar occupation numbers as those of the canonical orbitals of the CASSCF/MM treatment. In fact, the occupation numbers in the natural KS orbitals already signal the fact that the electronic structure is a TCSCF type, which in turn is equivalent to an open-shell singlet state.

One shortcoming of the symmetry-broken UDFT/MM approach for the ground singlet state is the spin-contamination problem due to the single determinant representation in DFT, and this feature can be corrected by spin projection that stabilizes slightly the singlet state relative to the triplet state. For the lowest triplet state, where spin contamination is negligible, the agreement of the electronic structure from CASSCF/MM and DFT/MM is straightforward, as shown from singly occupied orbitals in both methods. Thus, both CASSCF/MM and DFT/MM support the Weiss mechanism of Fe−$O_2$ bonding in which heme-$Fe^{II}$ transfers a single electron to $O_2$ and the heme-$Fe^{III}$−$O_2^-$ species are coupled into a singlet state and generate a low-lying triplet state.

**4.4. Unified VB Description of Fe−$O_2$ Bonding in Oxy-Mb.** Having discussed the singly occupied orbitals obtained from CASSCF/MM and CASSCF(g) calculations in Figure 4, we can now attempt to derive a unified VB picture of bonding in a form that will allow us to comprehend the manner whereby the bonding changes with the environment, the axial ligand of the heme, etc. The delocalization tails in the singly occupied GVB orbitals tell us in fact that the Weiss-type bonding is sustained by other VB contributions. Thus, while the configuration with the singly occupied fragment orbitals describes the covalent part of the $\pi$(Fe−$O_2$) bond, shifting one electron between these orbitals would correspond to the ionic contributions to this bond.[19] As such, transferring one electron from the singly occupied $\pi_\perp$* orbital into the singly occupied $3d_{yz}$ orbital would generate the Pauling structure, while transferring the electron from $d_{yz}$ to $\pi_\perp$* would generate the $Fe^{IV}O_2^{2-}$ structure. These secondary structures are hidden in the delocalization tails of the GVB orbitals.[68] Similarly, transferring a single electron from the $\pi$* orbital to the $3d_{z^2}$ orbital would generate the McClure−Goddard configuration. These secondary contributions can be

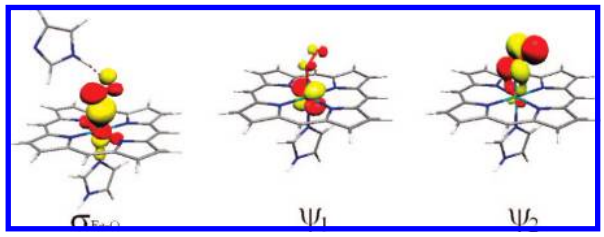


***Figure 7.*** Three orbitals responsible for Fe−$O_2$ bonding.

revealed by further expanding the wave function written in terms of the GVB-type orbitals into VB structures based on pure fragment orbitals.[19] To provide a visual aid for the equations we show in Figure 7 the three orbitals that are responsible for the Fe−$O_2$ bonding using the gas-phase $\psi_1$ and $\psi_2$ orbitals which are more general forms of these orbitals with slightly stronger mixing that depends on conditions.

The corresponding wave function that accounts for the bonding is

$$\Psi \propto |\cdots \sigma_{FeO}\bar{\sigma}_{FeO}\psi_1\bar{\psi}_2| - |\cdots \sigma_{FeO}\bar{\sigma}_{FeO}\bar{\psi}_1\psi_2| \quad (4)$$

where the $\sigma$ bond orbital and orbitals $\psi_1$ and $\psi_2$ can be approximately expressed as follows in terms of the pure fragment orbitals

$$\sigma_{FeO} = \pi^* + \delta d_{z^2}, \psi_1 = d_{yz} - \lambda\pi_\perp, \psi_2 = \pi^*_\perp + \lambda d_{yz}, \quad S_1 = \langle d_{yz}|\pi^*_\perp\rangle, S_2 = \langle d_{z^2}|\pi^*\rangle \quad (5)$$

The terms $S_1$ and $S_2$ are the fragment orbital overlaps.

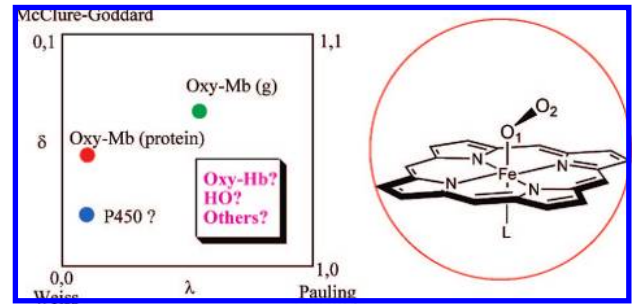
Using the expressions of the bonding orbitals in terms of fragment orbitals, the wave function in eq 4 can be expanded into VB structures based on pure fragment orbitals; the details of the full derivation are in the Supporting Information. The expansion leads to the expression in eq 6, which shows the leading terms when $\lambda$ and $\delta$ are small

$$\Psi = \sqrt{(1+S_1^2)/2}\Phi_{Weiss} + \lambda\Phi_{Pauling} + \delta\sqrt{(1+S_1^2)(1+S_2^2)}\Phi_{McClure-Goddard} \quad (6)$$

The other terms, not included in eq 6, correspond to a ferryl configuration $Fe^{IV}O_2^{2-}$, and to $Fe^{I}O_2^{+}$ and $Fe^{0}O_2^{2+}$ states, which have smaller contributions. Importantly, we can see that the wave function of the Fe−$O_2$ bond, based on pure fragment orbitals, involves a mixture of the three bonding models of Scheme 2 with coefficients that depend on the mixing parameters $\lambda$ and $\delta$ and hence on the ability of the protein environment to polarize the orbitals and decrease both $\lambda$ and $\delta$. In the case of oxy-Mb, with CASSCF/MM obviously $\lambda$ and $\delta$ are small ($\lambda$ and $\delta$ are about 0.18 and 0.55, $S_1$ and $S_2$ are about 0.04 and 0.14), making the Weiss configuration dominant with a significant contribution from the McClure−Goddard model and with the Pauling model being the least significant (the ratios of the coefficients of the Weiss/Pauling/McClure−Goddard configurations are 0.71/0.18/0.56).

In fact, even though the bonding mechanism remains a Weiss type, the Fe−$O_2$ bonding depends on the VB mixing of all the configurations. Thus, as we argued above, the McClure−Goddard configuration has the covalent component of the $\sigma$(Fe−O) bond while the Pauling configuration is the "ionic" component of the $\pi$(Fe−$O_2$) bond, which is described primarily by the covalent Weiss-type configuration. The relative contributions ($\lambda$ and $\delta$ in eq 6) of these configurations will be modulated by the

***Scheme 5.*** Schematic Representation of the Fe−$O_2$ Bonding in Oxyheme Complexes as a Function of the Orbital Mixing Parameters in Eq 6




environment and the axial ligand (L) of the iron. Scheme 5 shows qualitatively the expected behavior for different situations. Thus, considering the case of oxy-Mb, one can see that the situation in the protein, where the orbitals are polarized having very small mixing coefficients, the Fe−$O_2$ bonding is closest to the Weiss configuration. However, in the gas phase where the orbital mixing coefficients are larger the bonding blends more of the Pauling and McClure−Goddard configurations. The conclusion that in the gas phase the $\sigma$ and $\pi$ bonds of oxy-Mb are less polar than in the protein suggests that an important factor in the binding of $O_2$ by Mb is the electrostatic interaction between the $Fe^{III}$ and $O_2^-$ moieties that are stabilized by the H bonding to His64 and the bulk polarity.

Furthermore, Scheme 5 enables one to consider the effect of the axial ligand L on the Fe−$O_2$ bonding in a simple manner. Indeed, oxyheme species can either be made or actually exist in the native catalytic cycle of many heme enzymes. For example, oxy-cytochrome P450 is an important intermediate en route to formation of the ultimate active species, compound I.[69] Other examples where oxyhemes exist in the native cycle include the nitric oxide synthase (NOS)[70] and heme oxygenase (HO).[71] From the above analysis of the bonding in oxy-Mb we can expect that the "push" effect of the cysteine-derived thiolate ligand as the proximal ligand in P450[69a,e] will further stabilize the Weiss configuration and make it more dominant than in oxy-Mb. This may be true also for oxy-NOS and oxy-CPO (chloroperoxidase).[72] In general, however, the model raises more questions for the future. Thus, for example, in HO the axial ligand is identical to Mb but the hydrogen-bonding environment of the $O_2$ ligand is entirely different, residing in a very polar environment near a large water cluster. How will this change the Fe−$O_2$ bonding? How will this bonding change in other enzymes? Model systems? These are all questions for the future. Our above analysis could provide a general framework to study and analyze Fe−$O_2$ bonding in other oxyheme species.

(69) (a) Dawson, J. H.; Sono, M. *Chem. Rev.* **1987**, *87*, 1255. (b) Sono, M.; Roach, M. P.; Coulter, E. D.; Dawson, J. H. *Chem. Rev.* **1996**, *96*, 2841. (c) Denisov, I. G.; Makris, T. M.; Sligar, S. G.; Schlichting, I. *Chem. Rev.* **2004**, *105*, 2253. (d) Meunier, B.; de Visser, S. P.; Shaik, S. *Chem. Rev.* **2004**, *104*, 3947. (e) Shaik, S.; Kumar, D.; de Visser, S. P.; Altun, A.; Thiel, W. *Chem. Rev.* **2005**, *105*, 2279. (f) Shaik, S.; Hirao, H.; Kumar, D. *Nat. Prod. Rep.* **2007**, *24*, 533.
(70) (a) Groves, J. T.; Wang, C. C. Y. *Curr. Opin. Chem. Biol.* **2000**, *4*, 687. (b) Rosen, G. M.; Tsai, P.; Pou, S. *Chem. Rev.* **2002**, *102*, 1191.
(71) (a) Colas, C.; Ortiz de Montellano, P. R. *Chem. Rev.* **2003**, *103*, 2305. (b) Unno, M.; Matsui, T.; Ikeda-Saito, M. *Nat. Prod. Rep.* **2007**, *24*, 553. (c) Ragsdale, S. W. *Chem. Rev.* **2006**, *106*, 3317.
(72) Kühnel, K.; Derat, E.; Terner, J.; Shaik, S.; Schlichting, I. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 99.

## 5. Conclusion

We asked at the outset what is the nature of the $Fe-O_2$ bonding in oxy-Mb? To answer the question we performed two types of calculations: (a) In-protein QM/MM type of calculations using DFT/MM and CASSCF/MM methods and (b) gas-phase calculations using DFT and CASSCF methods, hence DFT(g) and CASSCF(g). Within the protein, the dioxygen ligand of oxy-Mb is H bonded by the distal residue His64 and the entire heme feels the bulk polarity of the protein pocket, while in the gas phase the species is devoid of these interactions. The DFT/MM-optimized geometries for the ground singlet state with several widely used functionals were found to agree reasonably well with the experimental X-ray structure. The calculated Mössbauer parameters are also in good accord with the experimental values. It was found that the protein exerts a major impact on most of the features of the oxy-Mb complex. Removal of the protein environment causes the following changes: (i) It reduces the amount of negative charge on $O_2$ by a factor of 2 in CASSCF, (ii) it reduces the total weight of the two major configurations in the CASSCF wave function, and (iii) it causes switching in the identity of the lowest triplet state both in CASSCF as well as in GGA functionals.

The $Fe-O_2$ bonding analysis was done by transforming the CASSCF wave functions into a generalized valence bond (GVB) wave function. The resulting GVB wave function shows that the bonding is composed of a $\sigma(Fe-O)$ bond and a weak $\pi(Fe-O_2)$ bond composed of singly occupied and spin-paired electrons in the $d_{yz}(Fe)$ and $\pi_\perp^*(O_2)$ fragment orbitals; these fragment orbitals are almost pure with very small delocalization tails. As such, the bonding arises by a single electron transfer from heme-$Fe^{II}$ to $O_2$ and the so formed heme-$Fe^{III}$ and $O_2^-$ fragments coupled then to a singlet state $Fe^{III}-O_2^-$ that possesses a dative $\sigma(Fe-O)$ bond and a weakly coupled $\pi(Fe-O_2)$ bond pair. Thus, the CASSCF/MM wave function shows a strong support of the Weiss bonding mechanism. The bonding mechanism in the gas phase is not different with the exception that the $\sigma(Fe-O)$ bond is less polarized (higher back-donation from $O_2^-$ to $Fe^{III}$), while the constituents of $\pi(Fe-O_2)$ bond pair have greater delocalization tails. The protein thus strengthens the $Fe^{III}-O_2^-$ character of the complex, thereby modulating the $O_2$ affinity by the polarity and H-bonding capability of the protein.

Expansion of the GVB wave function to VB structures that are based on pure VB structures provides a unified bonding model (Scheme 5), which shows how the protein or the axial ligand of the oxyheme complex can determine the nature of its bonding in terms of the blend of the three bonding configurations (Weiss, Pauling, and McClure−Goddard) with Pauling's being the least important. It shows that the oxyheme complex behaves as an electronic chameleon, changing its bonding according to its protein host and axial ligand.

**Acknowledgment.** The paper is dedicated to David Avnir for his 60th birthday. The research is supported by the Israel Science Foundation (ISF) Grant (to S.S.) and Grants-in-Aid 18370052 and 17GS0419 from JSPS and MEXT, Japan (to M.I.-S.). H.C. thanks the Golda Meir Fellowship Fund.

**Supporting Information Available:** Computational procedures, calculated energies, spin and charge population, detailed derivation in the analysis of the CASSCF wave function, *x, y, z* coordinates of the DFT/MM-optimized geometries of oxy-Mb, and complete refs 43, 49, and 56. This material is available free of charge via the Internet at http://pubs.acs.org.

JA805434M