Large Scale ab Initio Quantum Chemical Calculation of the Intermediates in the Soluble Methane Monooxygenase Catalytic Cycle

Barry D. Dunietz,[†] Michael D. Beachy,[†] Yixiang Cao,[†] Douglas A. Whittington,[‡] Stephen J. Lippard,*[,‡] and Richard A. Friesner*[,†]

Contribution from the Departments of Chemistry, Columbia University, New York, New York 10027, and Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received June 21, 1999

Abstract: Ab initio DFT quantum chemical methods are applied to study intermediates in the catalytic cycle of soluble methane monooxygenase hydroxylase (MMOH), a dinuclear iron-containing enzyme that converts methane and dioxygen selectively to methanol and water. The quantum chemical models reproduce reliably the X-ray crystallographic coordinates of the active site for the oxidized diiron(III) and reduced diiron(II) states to a high degree of structural precision. The results inspired a reexamination of the X-ray structure of reduced MMOH and revealed previously unassigned electron density now attributed to a key structural water molecule. The quantum chemical calculations required construction of a model containing about 100 atoms, which preserved key hydrogen bonding patterns necessary for structural integrity. Smaller models were unstable for the reduced form of the enzyme, an observation with significant mechanistic implications. The large model was then used to investigate the catalytic intermediates H_{peroxo}, formed upon the addition of dioxygen, and Q, the active species that reacts with methane. The structures, which differ significantly from alternatives proposed in the literature, are consistent with the experimentally available information concerning the spin states, geometries, and thermodynamics of formation of these intermediates. Other models that have been proposed, particularly in the case of Q, are ruled out in our calculations by energetic considerations, which have a simple physical interpretation. A bound water molecule is critical in assembling the catalytically active species Q.

Introduction

The soluble methane monooxygenase (sMMO) system catalyzes the transformation of methane and dioxygen to methanol and water under ambient conditions in methanotrophic bacteria.[1,2] The catalytic core of the hydroxylase (MMOH) enzyme active site is built around a pair of hydroxo-bridged iron atoms, the spin state and charge of which change from one catalytic intermediate to another. Similar diiron cores occur in other dioxygen-dependent enzymes, such as the ribonucleotide reductase R2 component, toluene 3- and 4-monooxygenases, and stearoyl-ACP Δ^9 desaturase.[3]

The structure of MMOH has been determined by X-ray crystallography in two oxidation states that occur in the catalytic cycle.[4] Since that time, there has been an intense experimental and theoretical effort to supply a detailed molecular picture of the process, including the structures of key intermediates for which crystallographic data are not available and the transition state involved in the reaction of the enzyme with methane.[1,2] Despite significant progress in obtaining new and revealing experimental data, however, agreement concerning the mechanism has not yet been achieved. This situation is not surprising given the complexity of the problem. A wide variety of possible structures for the intermediates, which differ by the placement of protons, water molecules, and carboxylate groups, are plausible, given the existing crystallographic and spectroscopic data.

In principle, a solution to this dilemma can be aided by carrying out ab initio quantum chemical calculations of a sufficiently faithful representation of the active site to obtain optimized structures and energies of proposed intermediates and transition states. Although quantum chemical methods are at present too expensive to contemplate an exhaustive search of the active site phase space, it has become feasible to investigate on the order of tens or even hundreds of proposed structures, most of which one would hope could be eliminated as inconsistent with the experimental data. The key problem has been to be able to devise a model that maintains fidelity to the crystallographic data but at the same time is tractable to calculations at a sufficiently high level of theory to afford reliable energetics. The increasing capabilities of computational hardware, coupled with advances in electronic structure correlation methods and algorithms, have now made this a viable proposition.

Previous ab initio work on the sMMO problem[5,6] employed a level of theory, the B3LYP version of hybrid density functional theory (DFT),[7,8] with large basis sets, possibly satisfying the second criterion specified above. The structural models employed did not extend beyond the diiron core and associated ligands, however. Although the use of such models is essential

* To whom correspondence should be addressed.
[†] Columbia University.
[‡] Massachusetts Institute of Technology.

(1) Liu, K. E.; Lippard, S. J. *Adv. Inorg. Chem.* **1995**, *42*, 263−289.
(2) Wallar, B. J.; Lipscomb, J. D. *Chem. Rev.* **1996**, *96*, 2625−2657.
(3) Nordlund, P.; Eklund, H. *Curr. Opin. Struct. Biol.* **1995**, *5*, 758−766.
(4) Rosenzweig, A. C.; Nordlund, P.; Takahara, P. M.; Frederick, C. A.; Lippard, S. J. *Chem. Biol.* **1995**, *2*, 409−418.
(5) Siegbahn, P. E. M.; Crabtree, R. H. *J. Am. Chem. Soc.* **1997**, *119*, 3103−3113.
(6) Siegbahn, P. E. M. *Inorg. Chem.* **1999**, *38*, 2880−2889.
(7) Johnson, B. G.; Gill, P. M. W.; Pople, J. A. *J. Chem. Phys.* **1993**, *98*, 5612−5626.
(8) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372−1377.
(9) Ricca, A.; Bauschlicher, C. W. *Theor. Chim. Acta* **1995**, *92*, 123−131.

10.1021/ja9920967 CCC: $19.00
Published on Web 03/11/2000

in the initial stages of exploring active site chemistry, there are necessarily attendant uncertainties regarding the structural and energetic effects of the protein residues surrounding the catalytic core. In the present paper, we have applied the same level of theory to much larger models, which are capable of reproducing the crystallographic coordinates with as few artificial restraints as possible. Based upon work in the literature investigating the accuracy of hybrid DFT methods for transition metal containing systems, we estimate the B3LYP results for the relative energetics of the various structural intermediates to be within ~5 kcal/mol of the actual energies.[9−13] This level of precision should be sufficient to identify the correct intermediate and transition state structures, provided that the structural model is of sufficient quality. The models considered here, containing about 100 atoms, represent to our knowledge a substantial increase in size and complexity as compared to previous ab initio quantum chemical work in the literature on transition metal-containing enzymes. Furthermore, we show that such models are critical in several cases to maintain structural fidelity with the crystallographic data, which we believe is important in establishing biological relevance.

Our goal in the work described in the present article has been to obtain structures and energies of the four experimentally identified species in the catalytic cycle, the oxidized diiron(III) enzyme, the reduced diiron(II) enzyme, H_{peroxo}, and Q, consistent with currently available experimental data. We have accomplished this task and find only one plausible structure for Q, the species believed to be active in the methane reaction step, which is substantially different from any structure proposed in the literature to date. Subsequent papers will investigate in detail reaction pathways between the various proposed species using the same methods and models employed herein.

Relevant Experimental Data and Prior Theoretical Work

Following the original isolation and characterization of sMMO in the 1970s, extensive efforts have been directed at elucidating the mechanism by which bacteria oxidize methane. Much of this work has focused on the non-heme dinuclear iron site of catalysis, and the presence of two electronically coupled iron atoms has permitted investigation by numerous spectroscopic techniques. Enzymes isolated from both *Methylococcus capsulatus* (Bath) and *Methylosinus trichosporium* OB3b have been examined in detail.[14,15] The results provide the experimental framework by which to compare and judge the results of DFT calculations, and by which to generate initial structures for sampling intermediates in the catalytic cycle.

The dinuclear iron active site is located in the hydroxylase component of the sMMO system, MMOH. Mössbauer and X-ray absorption spectroscopic (XAS) investigations of purified MMOH in the resting state revealed Fe(III) coordination environments composed of oxygen and nitrogen donor atoms.[16,17] Electron paramagnetic resonance (EPR) investigations

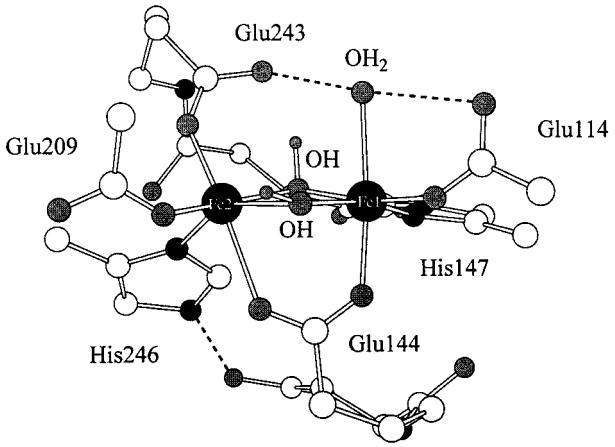

Figure 1. Oxidized MMOH crystal structure taken from the coordinates in ref 4. The bridging water molecule has been rendered as a hydroxide ion (see also ref 23). Only coordinated residues are identified.

at <10 K gave no signal, indicating a diamagnetic ground state for the two Fe(III) atoms, consistent with the Mössbauer results and indicative of antiferromagnetic exchange coupling.[16−18] Extended X-ray absorption fine structure (EXAFS) and electron nuclear double resonance (ENDOR) spectroscopic studies demonstrated that the two iron atoms were bridged by a hydroxo, rather than an oxo, ligand. This assignment is consistent with the magnitude of the coupling constant and the absence of any visible absorption band for oxidized MMOH.[16,19−21] All of these findings are consistent with the X-ray crystal structure determination of MMOH from *Methylococcus capsulatus* (Bath) and *Methylosinus trichosporium* OB3b, which revealed an Fe−Fe separation of ~3.1 Å and coordination shells that were composed solely of glutamate, histidine, hydroxide, and solvent-derived ligands (Figure 1).[4,22,23]

Analogous spectroscopic characterizations were performed on chemically reduced MMOH. Mössbauer spectroscopy and XAS revealed two divalent iron atoms.[16,17] An Fe−Fe separation was not determined, but an EPR signal with $g = 16$ was assigned to an integer spin system in which the two Fe(II) atoms were ferromagnetically coupled.[16,24,25] EXAFS and MCD experiments demonstrated that the iron atoms were most likely 5-coordinate.[16,25] The X-ray crystal structure of reduced MMOH is consistent with these findings. The loss of the bridging hydroxide ions, a carboxylate shift for Glu243, and loosely bound water molecules render the iron atoms 5-coordinate (Figure 2).[4] The resulting available coordination sites are important in providing positions for dioxygen to bind.

As sMMO proceeds through its catalytic cycle (Figure 3), two intermediates form that display absorptions in the visible

(10) Ricca, A.; Bauschlicher, C. W. *J. Phys. Chem.* **1995**, *99*, 5922−5926.
(11) Ricca, A.; Bauschlicher, C. W. *J. Phys. Chem.* **1997**, *101*, 8949−8955.
(12) Glukhovstev, M. N.; Bach, R. D.; Nagel, C. J. *J. Phys. Chem.* **1997**, *101*, 316−323.
(13) Blomberg, M. R. A.; Siegbahn, P. E. M.; Svensson, M. *J. Chem. Phys.* **1996**, *104*, 9546−9554.
(14) Colby, J.; Dalton, H. *Biochem. J.* **1978**, *171*, 461−468.
(15) Fox, B. G.; Froland, W. A.; Dege, J. E.; Lipscomb, J. D. *J. Biol. Chem.* **1989**, *264*, 10023−10033.
(16) DeWitt, J. G.; Bentsen, J. G.; Rosenzweig, A. C.; Hedman, B.; Green, J.; Pilkington, S.; Papaefthymiou, G. C.; Dalton, H.; Hodgson, K. O.; Lippard, S. J. *J. Am. Chem. Soc.* **1991**, *113*, 9219−9235.
(17) Fox, B. G.; Surerus, K. K.; Münck, E.; Lipscomb, J. D. *J. Biol. Chem.* **1988**, *263*, 10553−10556.
(18) Woodland, M. P.; Patil, D. S.; Cammack, R.; Dalton, H. *Biochim. Biophys. Acta* **1986**, *873*, 237−242.
(19) Ericson, A.; Hedman, B.; Hodgson, K. O.; Green, J.; Dalton, H.; Bentsen, J. G.; Beer, R. H.; Lippard, S. J. *J. Am. Chem. Soc.* **1988**, *110*, 2330−2332.
(20) DeRose, V. J.; Liu, K. E.; Kurtz, D. M., Jr.; Hoffman, B. M.; Lippard, S. J. *J. Am. Chem. Soc.* **1993**, *115*, 6440−6441.
(21) Thomann, H.; Bernardo, M.; McCormick, J. M.; Pulver, S.; Andersson, K. K.; Lipscomb, J. D.; Solomon, E. I. *J. Am. Chem. Soc.* **1993**, *115*, 8881−8882.
(22) Rosenzweig, A. C.; Frederick, C. A.; Lippard, S. J.; Nordlund, P. *Nature* **1993**, *366*, 537−543.
(23) Elango, N.; Radhakrishnan, R.; Froland, W. A.; Wallar, B. J.; Earhart, C. A.; Lipscomb, J. D.; Ohlendorf, D. H. *Protein Sci.* **1997**, *6*, 556−568.
(24) Hendrich, M. P.; Münck, E.; Fox, B. G.; Lipscomb, J. D. *J. Am. Chem. Soc.* **1990**, *112*, 5861−5865.
(25) Pulver, S.; Froland, W. A.; Fox, B. G.; Lipscomb, J. D.; Solomon, E. I. *J. Am. Chem. Soc.* **1993**, *115*, 12409−12422.

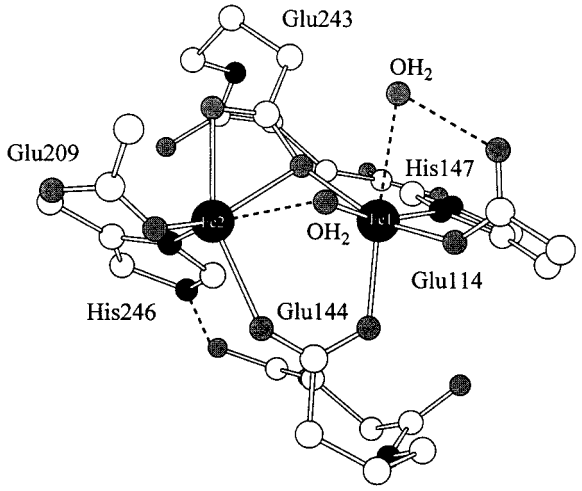

Figure 2. Reduced MMOH crystal structure, taken from the coordinates in ref 4. Only coordinated residues are identified.

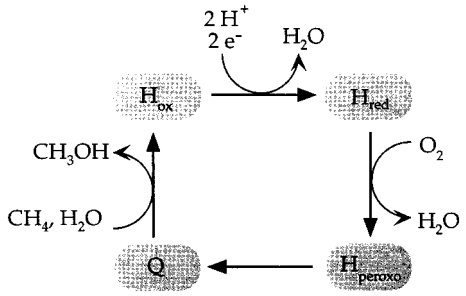

Figure 3. Simplified reaction cycle for MMOH.

range. The first of these intermediates, termed H_{peroxo}, has λ_{max} values of 420 and 725 nm.[26] This intermediate appears at a rate exceeding 1−2 s^{-1} at 4 °C and decays with a rate of 0.5 s^{-1}, as observed by stopped-flow techniques.[26,27] Concurrently with the decay of H_{peroxo}, an intermediate termed Q forms with a λ_{max} value at α420 nm.[27,28] The rate of Q decay significantly increases when methane is present, compared to its decay rate in the absence of substrate.[26,28] This result strongly suggests that Q is the intermediate that performs the oxidation chemistry. Freeze-quench trapping techniques have allowed for the spectroscopic examination of H_{peroxo} and Q by Mössbauer and EXAFS spectroscopy. A sharp, symmetric quadrupole doublet was discovered in a Mössbauer spectrum of H_{peroxo}, characteristic of two high-spin Fe(III) atoms in electronically equivalent environments.[29] The isomer shift (δ) was greater than that typically observed for carboxylate-bridged dinuclear Fe(III) centers, but a μ-η^1,η^1-peroxo diiron(III) model compound and several other dinuclear iron enzymes have since been examined that display similar parameters.[30−33] Mössbauer investigations of Q indicated two nearly equivalent Fe atoms in the Fe(IV) oxidation state.[29,34] EXAFS spectra of Q samples from *M. trichosporium* OB3b were fit by modeling a single short Fe−O bond of 1.77 Å per Fe and an Fe−Fe separation of 2.46 Å.[35] Both intermediates are diamagnetic as revealed by Mössbauer experiments at high magnetic field strength.[27,34]

Theoretical approaches to understanding the sMMO system have been previously described. In the simplest of these cases, the interaction of a methane molecule with free FeO^+ in the gas phase was studied by using DFT.[36] The preferred mechanism proceeded through a four-center transition state and involved formation of a $HO-Fe^+-CH_3$ intermediate, followed by recombination and release of methanol. Subsequent DFT calculations on a coordinatively unsaturated model of Q, in which the ligands consisted of one bridging formate, two bridging oxo anions, and three or four water molecules, continued to support the idea of a two-step concerted process involving hydrogen atom and methyl migrations.[37] An intermediate containing $Fe-CH_3$ and Fe−OH bonds was postulated, which then allowed recombination to produce methanol. The key assumption in these calculations was the assignment of the Fe atoms in Q as 5-coordinate, based upon the pre-edge areas determined in the EXAFS work with *M. trichosporium* OB3b MMOH.[35] Extended Hückel calculations have also been applied to this system, and the principal results from that analysis suggest that an unsaturated Fe center is needed to interact with a C_{3v}- or D_{2d}-distorted CH_4 molecule, forming an Fe−C bond in the reaction pathway.[5,6,38] The transiently 5-coordinate carbon atom may undergo pseudorotation in some cases, which would account for the observation of chiral ethane products with inverted stereochemistry.[39,40] Finally, DFT calculations have been performed on dinuclear iron centers that more faithfully represent the components of the active site of the enzyme.[6] From these results, it was suggested that a di(μ-oxo)diiron(IV) Q structure abstracts a hydrogen atom from methane and then rebounds to form methanol in a classical P-450-type mechanism.

Because the reliability of calculations depends on the accuracy of the starting geometry, as well as the choice of structures examined, we sought to develop a model for use in DFT calculations that would most accurately represent the effects of the protein environment. We have used that model to reproduce the oxidized and reduced structures observed in X-ray crystal structure determination of MMOH, and to generate reasonable structures and energies for the intermediates H_{peroxo} and Q. Future work will use this information to address questions of reaction pathway and hydrocarbon oxidation mechanism.

Computational Methods

All calculations were carried out with the Jaguar v3.5 suite of ab initio quantum chemistry programs. The DFT methodology used in Jaguar, including accuracy tests and performance characteristics, has

(26) Valentine, A. M.; Stahl, S. S.; Lippard, S. J. *J. Am. Chem. Soc.* **1999**, *121*, 3876−3887.
(27) Liu, K. E.; Valentine, A. M.; Wang, D.; Huynh, B. H.; Edmondson, D. E.; Salifoglou, A.; Lippard, S. J. *J. Am. Chem. Soc.* **1995**, *117*, 10174−10185.
(28) Lee, S.-K.; Nesheim, J. C.; Lipscomb, J. D. *J. Biol. Chem.* **1993**, *268*, 21569−21577.
(29) Liu, K. E.; Wang, D.; Huynh, B. H.; Edmondson, D. E.; Salifoglou, A.; Lippard, S. J. *J. Am. Chem. Soc.* **1994**, *116*, 7465−7466.
(30) Kim, K.; Lippard, S. J. *J. Am. Chem. Soc.* **1996**, *118*, 4914−4915.
(31) Bollinger, M. J., Jr.; Krebs, C.; Vicol, A.; Chen, S.; Ley, B. A.; Edmondson, D. E.; Huynh, B. H. *J. Am. Chem. Soc.* **1998**, *120*, 1094−1095.
(32) Pereira, A. S.; Small, W.; Krebs, C.; Tavares, P.; Edmondson, D. E.; Theil, E. C.; Huynh, B. H. *Biochemistry* **1998**, *37*, 9871−9876.
(33) Broadwater, J. A.; Achim, C.; Münck, E.; Fox, B. G. *Biochemistry* **1998**, *38*, 12197−12204.
(34) Lee, S.-K.; Fox, B. G.; Froland, W. A.; Lipscomb, J. D.; Münck, E. *J. Am. Chem. Soc.* **1993**, *115*, 6450−6451.
(35) Shu, L.; Nesheim, J. C.; Kauffmann, K.; Münck, E.; Lipscomb, J. D.; Que, L., Jr. *Science* **1997**, *275*, 515−517.
(36) Yoshizawa, K.; Shiota, Y.; Yamabe, T. *Chem. Eur. J.* **1997**, *3*, 1160−1169.
(37) Yoshizawa, K.; Ohta, T.; Yamabe, T. *Bull. Chem. Soc. Jpn.* **1998**, *71*, 1899−1909.
(38) Yoshizawa, K.; Ohta, T.; Yamabe, T.; Hoffman, R. *J. Am. Chem. Soc.* **1997**, *119*, 12311−12321.
(39) Priestley, N. D.; Floss, H. G.; Froland, W. A.; Lipscomb, J. D.; Williams, P. G.; Morimoto, H. *J. Am. Chem. Soc.* **1992**, *114*, 7561−7562.
(40) Valentine, A. M.; Wilkinson, B.; Liu, K. E.; Komar-Panicucci, S.; Priestly, N. D.; Williams, P. G.; Morimoto, H.; Floss, H. G.; Lippard, S. J. *J. Am. Chem. Soc.* **1997**, *119*, 1818−1827.

been described elsewhere.[41] For the calculations presented below, we estimate that Jaguar runs ~5 to10 times faster than the corresponding calculations would in Gaussian 92. In addition, for many of the calculations we have employed a parallel version of Jaguar, running on four nodes of an Origin 2000 parallel supercomputer at NCSA. The enhanced throughput obtained by a combination of fast algorithms and parallelization was essential in obtaining results in a reasonable time frame.

The B3LYP functional is used for all DFT calculations presented.[7,8] The superiority of performance of this functional as compared to first generation GGA functionals such as BLYP[7,42] for both structural and energetic results has been well documented in the literature, for example, in computing the heats of atomization of the extended G2 database.[43] For transition metal-containing systems, the evidence is more anectodal, but still quite convincing;[9−13] we expect the error in the B3LYP results to be less than 5 kcal/mol for relative energies. This level of error does not yield completely quantitative thermodynamics, but is sufficient to rule out a large number of incorrect proposed structures of intermediates.

With regard to basis set usage, we follow for the most part the approach developed in previous work.[5,6] Initially, explorations of various chemical structures were carried out with the 6-31G basis set (double-ζ, or DZ, level). Structures with competitive energies were then further optimized with polarization functions on the heteroatoms near the dinuclear iron core (i.e. with the 6-31G* basis on those atoms). For the iron atoms, the Los Alamos LACVP (and LACVP**) basis, with an effective core potential for the irons, was used for geometry optimizations.[44−46]

These basis sets, while useful for geometry optimization, are inadequate for accurate evaluation of relative energies of qualitatively different structures. For this purpose, basis sets of at least triple-ζ quality are required. After some experimentation, we determined that the cc-pVTZ (-f) basis set of Dunning was suitable for the organic atoms at the core of the system;[47] the effects of f functions were specifically tested and found to be very small. For the iron atoms, we use a basis set that is a modified version of the LACVP basis set in which exponents were decontracted to form a triple-ζ quality basis compatible with the Fe effective core potential (J. Perry, personal communication). This basis set, denoted LACV3P, was optimized to yield accurate relative energies for various charge states of the iron atom, a key issue in the present work, and has an excellent performance. The larger basis set was used to calculate the energies of the different species, and the energies reported here were obtained at this level. The additional correlation energy involved with the larger basis set was required to achieve reliable energy comparisons between the different states; use of a DZP basis, for example**,** leads to qualitatively incorrect orderings of the various intermediates.

Considerable effort is often required to converge the wave function to the ground electronic state for transition metal-containing systems, a problem exacerbated in the present case by the presence of two coupled iron atoms and the size of the system. In this paper, we use a new automated initial guess methodology,[48] which greatly facilitates such convergence. Without this tool, it would have been extremely difficult to investigate the ~100 different structures studied in the course of this work, not all of which are discussed below, in a relatively short time frame. The initial guess method was in general capable of robustly converging HF and DFT wave functions without human intervention, as long as the energy shift parameter (necessary in DFT due to the small HOMO-LUMO gap) was set to a reasonable value.

The systems here contain numerous open shell orbitals on the high spin iron atoms in which each atom contains up to 5 parallel unpaired spins. To handle this aspect of the chemistry, we use an unrestricted DFT (UDFT) methodology recently implemented in Jaguar. As in unrestricted Hartree−Fock calculations, some error is introduced due to spin contamination, but this error appears to be relatively modest in its effects on thermodynamic properties, as is evidenced by the excellent performance of the B3LYP method for open shell systems in the G2 database. It should be noted, however, that for systems of the present type, with a large number of unpaired spins residing on transition metal atoms, few examples have been quantitatively investigated and compared to experiment. This shortcoming could lead to a magnitude of error difference from that expected from the G2 results, although there is no particular evidence that such is actually the case.

A final important aspect of the system is that in three of the four intermediates (oxidized enzyme, $H_{peroxo,}$ and Q) the unpaired spins on the iron atoms are antiferromagnetically (AF) coupled, leading to a system comprising a manifold of different spin-coupled states. The procedure by which the singlet state is projected out is described below. In the reduced enzyme, the spins are ferromagnetically (F) coupled. The AF state can be represented, approximately, in a UDFT formalism, but obtaining a converged wave function of this type is nontrivial using our standard initial guess algorithm. Consequently, we employ a localized projection approach similar in spirit to that implemented by others.[49] First, a ferromagnetic state is converged; the open shell DFT density orbitals in this case are delocalized over the two iron atoms. A Boys localization scheme is then used to produce a localized representation of these orbitals. The localized orbitals can then be employed to build a good initial guess for the singlet state, simply by assigning paired spins to the corresponding localized orbitals on each iron atom.

This approach, which for a symmetric dimer would involve "broken symmetry," is identical with that used in work on transitional metal dimers.[49−51] The AF state in such systems is a linear combination of spin eigenfunction determinants, which in a conventional UDFT type picture implies a high degree of spin contamination. One can proceed a step further, however, and use the F and AF calculations to model the spin manifold by means of a Heisenberg spin operator formalism. In the limit where the spatial overlap of the open shell orbitals on the two transition metal centers is small, and the iron atoms can be treated as approximately equivalent (a reasonable although not rigorous approximation), one can model the spin manifold as described in ref 49 and extract a J value for the Heisenberg Hamiltonian from the splitting between the F (high spin) and AF (broken symmetry) states. The Heisenberg model can then be used to compute the energy of the singlet spin state in the manifold, which for AF coupling will be lower than the energy of the broken symmetry state. We carry out this treatment below for the various MMO intermediate structures that we have generated and report (when relevant) J values and corrected singlet state energies.

The foregoing protocol contains uncontrolled approximations, and a rigorous treatment of the spin manifold, for example via a multiconfigurational self-consistent field methodology, would be more accurate if carried out at a sufficiently high level of theory. Such a treatment is difficult to accomplish at present with models of this size and complexity. Given the magnitude of the energies involved in the spin couplings, our view is that the approximations made here constitute a reasonable first step toward understanding the effects of spin state on the relevant potential energy surfaces. Future work will undertake a more detailed investigation of these issues.

Structural Models

(a) Reduced and Oxidized Forms of MMOH. We begin with a simple objective: to reproduce the core X-ray crystal structure for the oxidized and reduced forms of the enzyme to a reasonable level of fidelity. The smallest model for which such a comparison is meaningful is one in which the ligands to the iron are represented by the actual amino acid side chains,

(41) Friesner, R. A.; Murphy, R. B.; Beachy, M. D.; Ringnalda, M. N.; Pollard, W. T.; Dunietz, B. D.; Cao, Y. X. *J. Phys. Chem. A* **1999**, *103*, 1913−1928.
(42) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098−3100.
(43) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **1997**, *106*, 1063−1079.
(44) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 270−283.
(45) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 299−310.
(46) Wadt, W. R.; Hay, P. J. *J. Chem. Phys.* **1985**, *82*, 284−298.
(47) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007−1023.
(48) Vacek, G.; Perry, J. K.; Langlois, J.-M. *Chem. Phys. Lett.* **1999**, *310*, 189−194.
(49) Zhao, X. G.; Richardson, W. H.; Chen, J. L.; Li, J.; Noodleman, L.; Tsai, H. L.; Hendrickson, D. N. *Inorg. Chem.* **1997**, *36*, 1198−1217.
(50) Noodleman, L. *J. Chem. Phys.* **1981**, *74*, 5737−5743.
(51) Noodleman, L.; Baerends, E. J. *J. Am. Chem. Soc.* **1984**, *106*, 2316−2327.

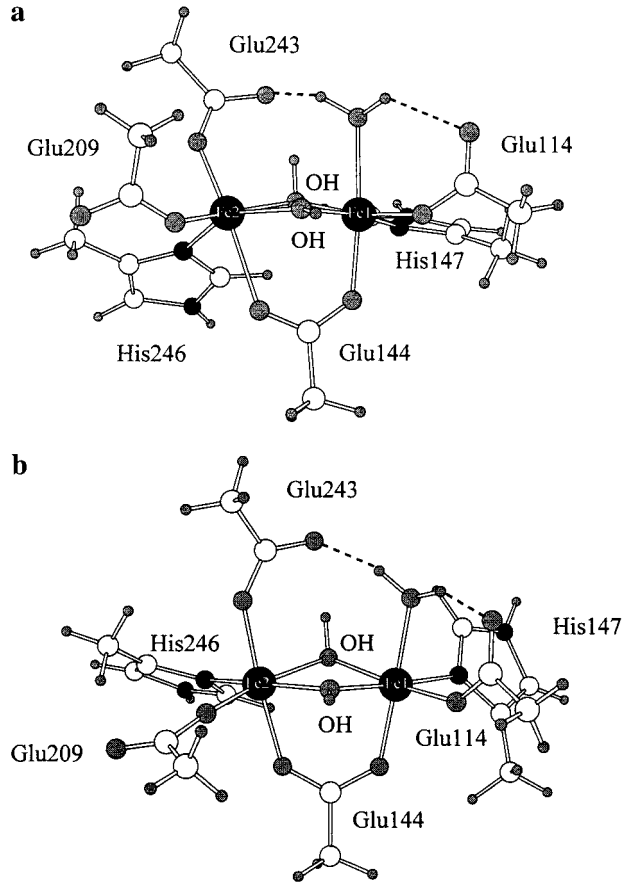

Figure 4. (a) Oxidized small model starting structure. (b) Optimized oxidized small model.

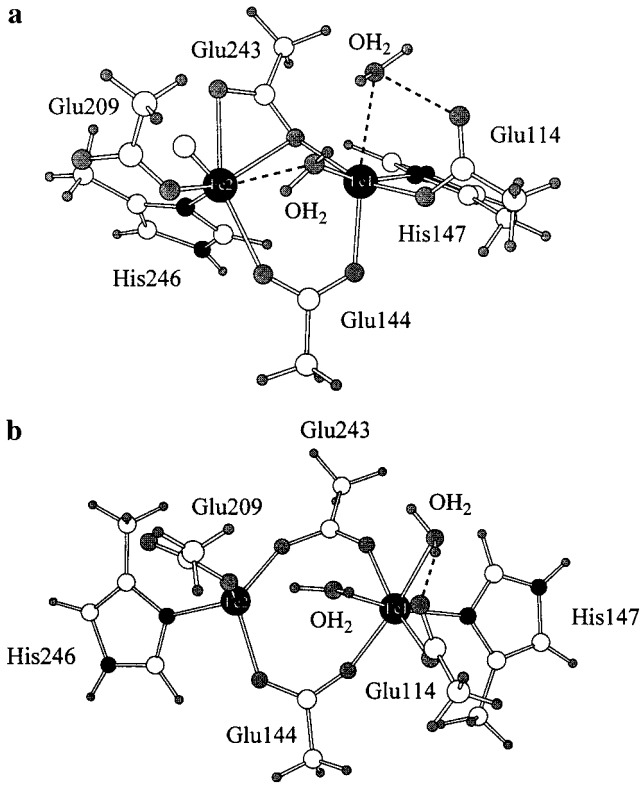

Figure 5. (a) Reduced small model starting structure. (b) Optimized reduced small model.

Table 1. RMSD (Å) of the Oxidized and Reduced Models from the MMOH Crystal Structure

model	species	all atoms	no hydrogens
small	reduced	1.653	1.455
large	reduced	0.411	0.320
small	oxidized	1.243	0.894
large	oxidized	0.634	0.501

but where the backbone C_α carbons are removed and replaced with capping hydrogen atoms. Figures 4a and 5a present such structures, taken from the X-ray data, for the oxidized and reduced forms of the enzyme. These "small" models contain about 60 atoms.

We then subject each of these structures to geometry optimization at the B3LYP/DZ level. The optimized structures are shown in Figures 4b and 5b. The oxidized structure displays some movement upon optimization. In particular, the histidine bound to Fe1 significantly changes its torsion angle around the Fe−N bond. Nonetheless, the optimized structure is fundamentally similar to that in the crystal structure, with a RMSD from the latter of 1.243 Å. The deviations observed are primarily due to rotations of the histidine imidazole moieties.

The reduced enzyme structure behaves very differently, however. The optimized structure in Figure 5b differs in several major respects from the crystal structure. First of all, the Fe−Fe distance has increased from 3.3 to 4.1 Å, clearly beyond the limits of experimental error. Second, the unusually positioned carboxylate of Glu243 in Figure 5a, which had one oxygen atom in a bridging position between the two iron atoms, has been converted to a μ-1,2-carboxylate, a very different structural motif. Third, the details of the ligand orientations are qualitatively different between the two structures. The RMSD of this model from the crystal structure is 1.653 Å (Table 1). We therefore conclude that the structural model presented in Figure 5 is inadequate for studying MMOH chemistry.

The discrepancy between DFT and X-ray crystallographic results for the reduced enzyme structure must either occur because the quantum chemical methodology has grossly failed, which seems unlikely, or because additional protein residues constrain the diiron core and surrounding ligands in a manner not built into the model. Examination of the X-ray structure reveals that there is in fact a network of hydrogen bonds that might serve to stabilize the observed geometry. In particular, by adding the residues Asp143 and Asp242 that hydrogen bond to the histidine rings, we can further tie together the coordination environment around the dinuclear iron center in a manner that better approximates the biological system. After visualization of the structure and extensive computational experiments, we have constructed a model that has the requisite structural properties. This model, containing about 100 atoms total, is shown in Figures 6a and 7a for oxidized and reduced MMOH, respectively. The overall model has no net charge because the active site is buried in a low dielectric protein medium, which requires the two carboxylates in the second coordination shell be protonated. Without significant solvent access, net ionization would be energetically unfavorable. Finally, we include coordinated waters and/or hydroxide ions revealed by the X-ray crystallographic analysis. These solvent-derived ligands will turn out to be essential components in the chemistry that follows.

Optimization of the structures leads to reasonable agreement with the crystallographic results. In studying some proposed forms of the intermediates, however, we observed a tendency in some cases for the two portions of the protein backbone in the structure to drift apart. These two components are the only ones not specifically constrained to each other by hydrogen bonds. To represent faithfully the structure as observed in the

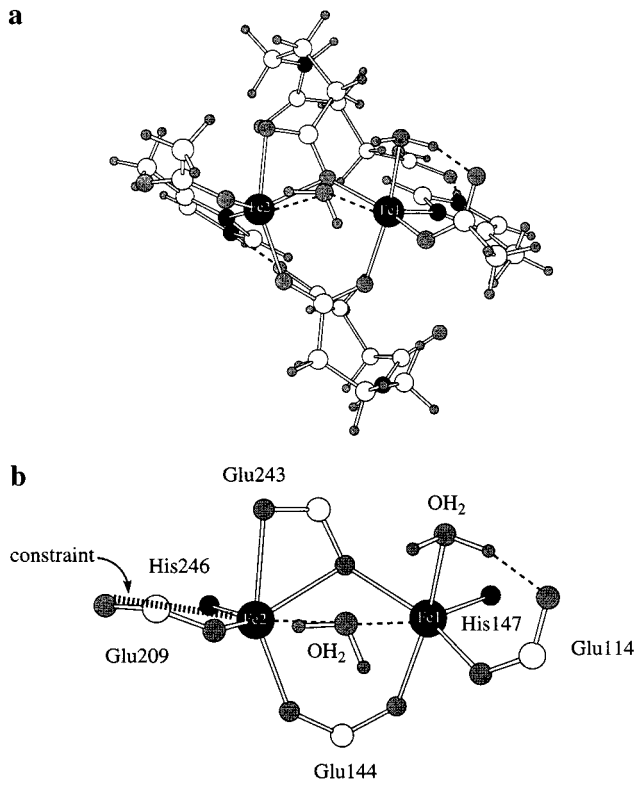

Figure 6. (a) Minimized reduced large model. (b) Core of the minimized reduced large model depicting only the atoms for which the added polarization functions of the DZP basis set were used. Bond distances are given in Table 7.

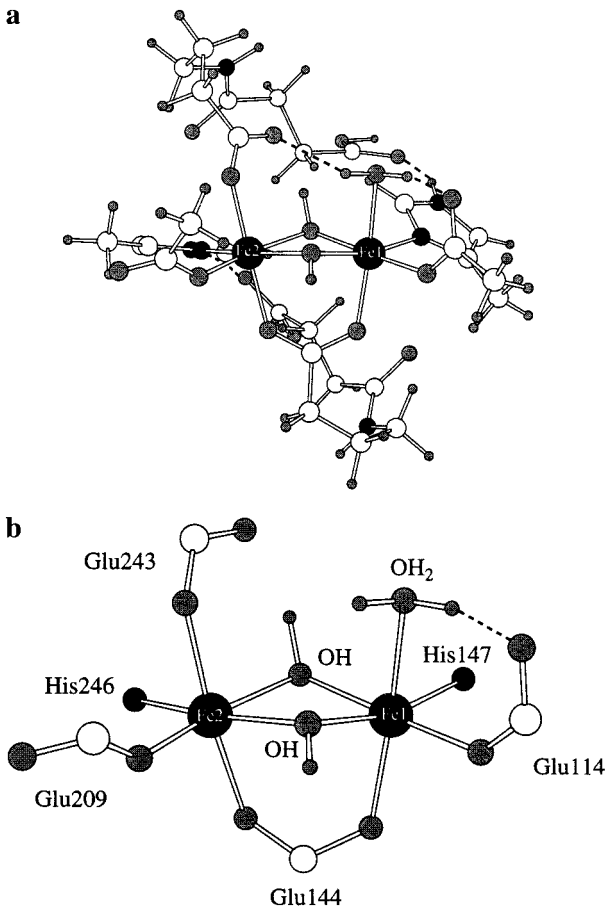

Figure 7. (a) Minimized oxidized large model. (b) Core of the minimized oxidized large model depicting only the atoms for which the added polarization functions of the DZP basis set were used. Bond distances are given in Table 7.

protein, we therefore introduced a single artificial constraint between the β carbon atoms of Asp143 and Asp242, depicted in Figure 8 for the oxidized enzyme. We do not believe that this constraint will incorrectly bias the active site chemistry because it is distant from the diiron core and only a relatively small force is exerted in all cases. On the contrary, we suggest that this constraint more or less properly represents the constraining effects of the protein 4-helix bundle on the atoms in the model.[22] In the MMOH structure, the drift that was observed in certain DFT calculations is prevented by amino acid side chain packing interactions found in the bundle. Henceforth, all results shown incorporate this constraint.

In carrying out optimizations of the reduced enzyme, one feature of the structure deviated substantially from the experimental results and could not be rationalized on the basis of imprecision in locating heavy atom coordinates in the analysis of the electron density. The Glu209 carboxylate ligand bonded to Fe2 in Figure 6 has an oxygen atom that appears to make no hydrogen bonds to any other group in the enzyme; it is "dangling" out into empty space. In the calculations, this oxygen atom shifts to become liganded to the 5-coordinate Fe atom, lowering the energy of the system by 15−20 kcal/mol. This positioning is clearly inconsistent with the X-ray data and in addition would adversely affect the subsequent dioxygen addition step. The question then arose as to why the enzyme fails to adopt the conformation that we compute, given that its energy is so much lower and there is no apparent barrier to the motion required to build the structure.

These results motivated us to reexamine the original electron density map of the reduced enzyme structure.[4] In this process, we discovered electron density consistent with a previously unassigned water molecule in a position to form hydrogen bonds to the "dangling" carboxylate oxygen and to a part of the protein backbone. Unfortunately, this discovery was made toward the end of the present study. Incorporation of the water molecule and its hydrogen-bonding pattern in all of the calculations would have necessitated many months and a large expenditure of computer time. Consequently, for the present paper, we mimic the effect of this part of the hydrogen-bonding network by an additional geometric constraint, shown in Figure 6b. We intend to study the sMMO active site using QM/MM techniques soon, and the "new" water molecule will be included in this effort. For the present, the constraint is a reasonable alternative, again because the force on the constrained atom at its initial position is relatively small. Figures 6 and 7 present optimized structures of the reduced and oxidized forms of the enzyme. The final optimizations of these structures were carried out with the mixed DZ/DZP basis set using UDFT methods, with the oxidized enzyme in the high spin (AF) state and the reduced enzyme in the high spin (F) state. The mixed basis set is defined by extending the DZ basis and treating the atoms depicted explicitly in Figures 6b and 7b with the additional polarization functions of the DZP basis set. As can be seen from the figures, the optimized structures are in good agreement with the X-ray structures in most details, with acceptable energies and RMS deviations from those structures. A summary of the RMSD of both the smaller model and the final larger model of the reduced and oxidized models from the crystal structure is given in Table 1, where the impressive improvement due to the use of the more appropriate larger model, especially for the reduced structure, is demonstrated. The charge and spin multiplicity of the reduced and oxidized models as well as the other species in the

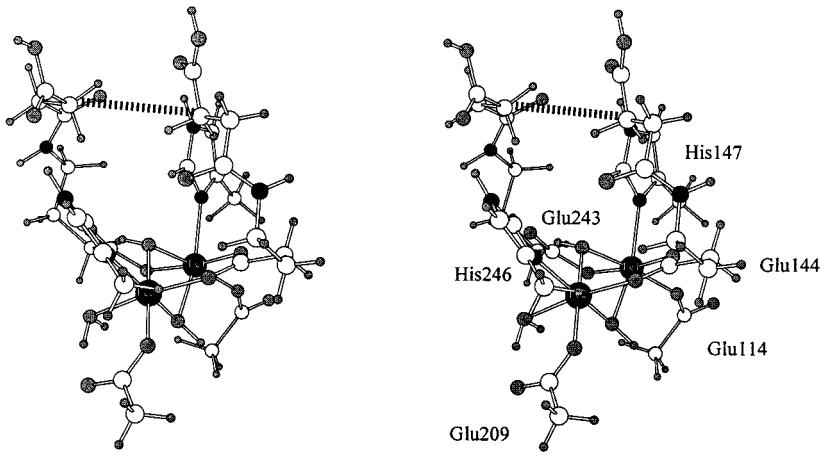

Figure 8. Stereoview of the large oxidized model depicting the geometry constraint that was applied (dashed line).

Table 2. Spin Multiplicity and Charge of the Iron Atoms in the Different States and the Fe−Fe Distance in the Optimized Structure

species	spin multiplicity/ coupling			formal oxidation state		distance (Å)
	Fe1	Fe2	Fe1−Fe2	Fe1	Fe2	Fe−Fe
reduced	5	5	9	2	2	3.5
oxidized	6	6	AF	3	3	3.1
peroxo	6	6	AF	3	3	3.4
superoxo	5	6	AF	2	3	3.4
Q-best	5	5	AF	4	4	2.67
Q-alternate	5	5	AF	4	4	2.90

thermodynamic cycle were determined according to the available experimental data and are summarized in Table 2.

The oxidized enzyme exhibits excellent quantitative agreement with the X-ray structure, including the Fe−Fe distance. This state is the lowest energy state of the catalytic cycle and, as shown above, is reasonably stable even without the protein hydrogen bonding network to provide constraints. The agreement of calculated and experimental structures is therefore not surprising. We have investigated the thermodynamic stability of alternative protonation states of the bridging oxygen and terminal water atoms and find the di(μ-hydroxo) aqua version shown in the figures to be the most favorable.

Agreement for the reduced enzyme is less quantitatively accurate; for example, the Fe−Fe distance is about 0.2 Å too large in the calculated structure. This feature is not particularly disturbing since the Fe−Fe potential surface is rather flat energetically, our description of the protein environment (particularly for the reduced enzyme) is not completely rigorous, and some errors in the bond length will arise from inaccuracies of the DFT functional. To provide a quantitative indication as to how flat the Fe−Fe potential surface is, we have performed a restrained geometry optimization of the reduced enzyme, fixing the Fe−Fe distance at the experimental value of 3.3 Å and following with large basis set single point calculations. The resulting energy difference with the lowest energy structure is only 4.3 kcal/mol. In addition, the precision of the X-ray data allows for ±0.1 Å in distance. The key features of the structure, however, such as the altered positioning of the carboxylate of Glu243, are faithfully reproduced, in contrast to the smaller model discussed above. The movement of this carboxylate through the intermediates in the catalytic cycle, together with the role of the bound water molecule, is central to the understanding of how the intermediate species are sequentially produced. Within the framework described in this paper, we cannot investigate the mechanism of this motion from the oxidized to reduced enzyme until the role of protein B, which possibly involves conformational changes in MMOH or displacement of solvent in the active site upon binding, is fully delineated.[52−54] As will be seen below, however, the Glu243 carboxylate continues to shift position as one moves through the catalytic cycle and these shifts are necessary to form the relevant intermediates.

The inability of the smaller model to reproduce accurately the structure of the reduced enzyme suggests that, in the protein, the dinuclear iron(II) species has substantial strain energy, which is imposed via the hydrogen bonding network of the protein around the active site as was discussed previously. This higher energy may have significant implications for catalysis. Reduction of the oxidized enzyme requires input of energy, some of which may be stored in the strained geometry imposed by the protein scaffolding and then used in part to drive the catalytic reaction in subsequent steps.

In summary, the results of this section provide us with confidence that our basic structural model is adequate to the task at hand, and has led to one important chemical finding, namely, the water molecule in the reduced enzyme that keeps Glu209 monodentate. We now turn to the determination of the structures and energetics of the H_{peroxo} and Q intermediates, a much more difficult task since there are no single-crystal X-ray data to guide the calculations.

(b) Model for the H_{peroxo} Intermediate. The formation of H_{peroxo} is initiated by addition of dioxygen to the system. From the experimental data, the thermodynamics of formation of H_{peroxo} are clearly downhill, so we have the requirement that reduced enzyme plus dioxygen be higher in energy than H_{peroxo} plus a water molecule. The iron atoms are both assigned as high spin Fe(III), which are AF coupled.[27] Table 2 lists the spin and charge states, Table 3 the single point energies calculated using the final basis set as described above, and Table 4 the energy comparisons of the different species involved in the cycle.

The results reported in Table 4 represent enthalpy, rather than free energy, differences between various intermediate states, with one exception. We have roughly approximated the translational/rotational entropy of moving a water molecule from a coordination position in the active site to bulk solution to obtain the energy difference between the reduced enzyme and the H_{peroxo} intermediate (see below). Note that we have not tried to make a similar approximation for dioxygen because it is not clear what the "reference" state of that molecule would be prior

(52) Green, J.; Dalton, H. *J. Biol. Chem.* **1985**, *260*, 15795−15801.
(53) Gassner, G. T.; Lippard, S. J. *Biochemistry* **1999**, *38*, 12768−12785.
(54) Coufal, D. E.; Tavares, P.; Pereira, A. S.; Hyunh, B. H.; Lippard, S. J. *Biochemistry* **1999**, *38*, 4504−4513.

Table 3. Energetics of the Different Species in the Thermodynamic Cycle[a]

species	energy (hartrees)
O_2	−150.3752
H_2O	−76.4585
CH_4	−40.5376
CH_3OH	−115.7696
oxidized	−2949.7978
reduced	−2874.5474
peroxo	−2948.4517
superoxo	−2948.4477
Q-best	−2948.4582
Q-alternate	−2871.9131
Q-alternate + H_2O	−2948.3716
reduced + O_2	−3024.9226
peroxo + H_2O	−3024.9101
Q-best + H_2O + CH_4	−3065.4542
oxidized + CH_3OH	−3065.567

[a] Energies are at the B3LYP/combined basis set as described in the text.

Table 4. Summary of Energy Comparisons

reactants	products	enthalpy diff (kcal/mol)
reduced + O_2	peroxo + H_2O(aq)	−2.2
Q-alternate + H_2O	Q-best	−54
peroxo	Q-best	−4.1
Q-best + H_2O + CH_4	oxidized + CH_3OH	−71

to its binding to the core. If it were highly confined in a protein cavity, there will be little or no translational/rotational entropy associated with that state. The relative vibrational entropies of the various states could be estimated by computing harmonic frequencies at the relevant minima. When this procedure was carried out for smaller model systems the differences in zero point energy were quite small,[5] and it is likely that the same would be true here. Our judgment is that these zero point differences would not change the qualitative conclusions to be drawn from Table 4, particularly given the already existing limitations of the calculations due to error bars in the enthalpies arising from uncertainties in the accuracy of the DFT calculations and of the structural model. This point should be rigorously checked, however. To do so for a system of this size requires the ability to compute analytical second derivatives with effective core potentials, which we have only recently implemented in Jaguar. We intend to carry out such calculations and to report them in a subsequent publication.

To add dioxygen, a binding site at the dinuclear iron center is required. This site can most easily be produced by removing the less strongly bound water depicted in the center of Figure 6b describing the reduced model. We have done so, added dioxygen, and optimized various initial guesses for the structure. We have also studied models assuming that both iron atoms are Fe(III) and for models in which one iron is Fe(III) and one is Fe(II). The former implies that the dioxygen is converted to a peroxo species, whereas the latter implies that it is at the superoxo level. Participation of a short-lived (τ < 1 ms) superoxo intermediate cannot be ruled out by the available experimental data.

Figures 9 and 10 present the lowest energy peroxo and superoxo species that we have identified and allow a close look at the structural differences of the cores of the two structures. Both are (correctly) lower in energy than the starting materials, assuming that one takes into account the fact that the water molecule removed from H_{peroxo} will go into solution and including the translational/rotational entropy gain as well as the

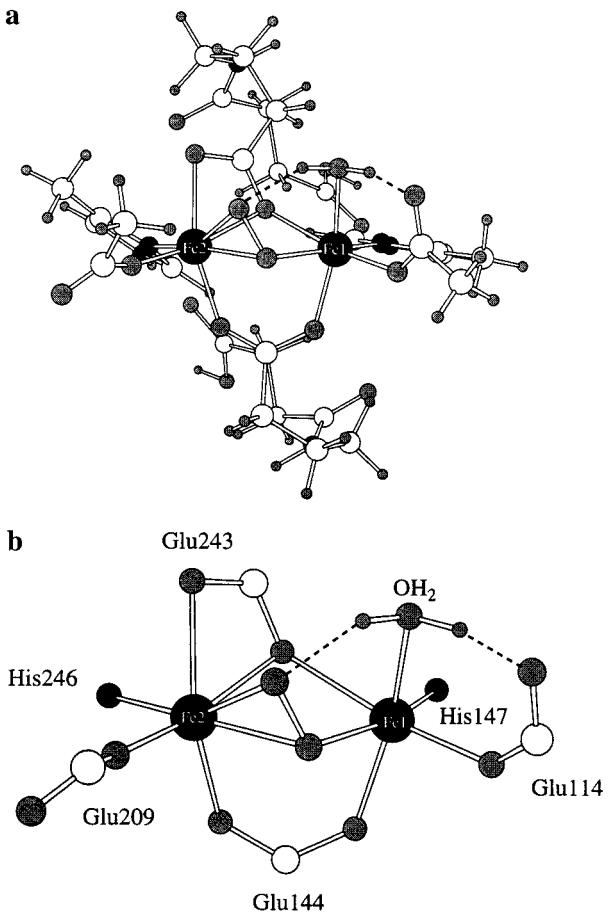

Figure 9. (a) Minimized H_{peroxo} structure. (b) Core of the minimized H_{peroxo} structure. Bond distances are given in Table 7.

standard gas phase to water free energy of transfer of this process. Large basis set single point calculations were carried out to obtain this result, data for which are summarized in Table 3. The energies of the superoxo and peroxo structures are comparable to each other, suggesting that there may be a nearly degenerate manifold of states with differing charge distribution between iron and oxygen. Distinguishing between the energetics of these two structures is beyond the capabilities of both the DFT methods and the structural model. A larger representation of the protein and a more accurate electronic structure calculational approach would be needed. We view the proposed intermediates in the figures as providing a satisfactory understanding of H_{peroxo} at present, however, allowing energetic comparisons with Q (see below) and suggesting experimental testing of the proposed geometric features.

(c) Model for Intermediate Q. We tested two basic types of Q structures. In one, the shifting carboxylate Glu243 displaces the remaining water molecule from Fe1 to form a structure with two dibridging carboxylates. A second structure was tried in which the bridging oxygen of this carboxylate instead moves out and several opportunistic hydrogen bonds are formed, leaving the water in place. Our first conclusion is that displacement of the water molecule is not energetically feasible. Figure 11 presents an optimized structure of a typical Q model of this type, which closely resembles those postulated recently.[6,35,55] The energy of this structure, including the displaced water, is 50 kcal/mol higher than that of our two low energy H_{peroxo}/$H_{superoxo}$ structures, which is far outside the range of error expected for the DFT calculations. Even if one assumes that

(55) Lee, D.; Lippard, S. J. *J. Am. Chem. Soc.* **1998**, *120*, 12153−12154.

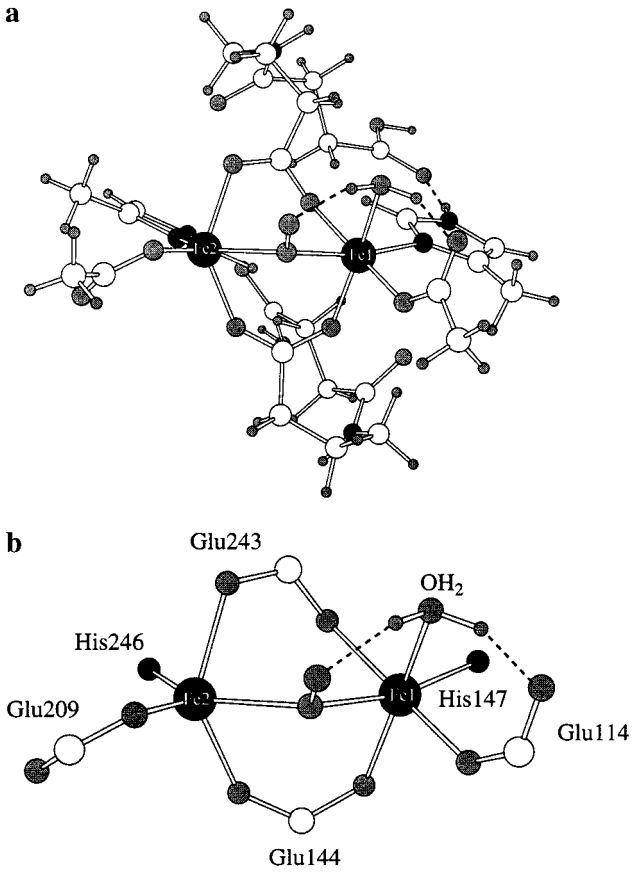

Figure 10. (a) Minimized $H_{superoxo}$ structure. (b) Core of the minimized $H_{superoxo}$ structure. Bond distances are given in Table 7.

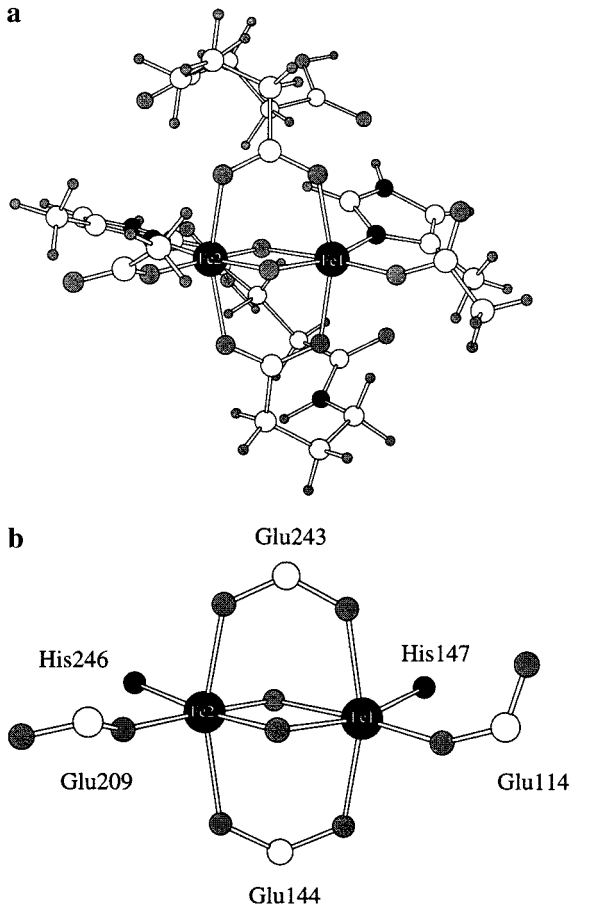

Figure 11. (a) Minimized alternative Q structure. (b) Core of the minimized alternative Q structure. Bond distances are given in Table 7.

the displaced water returns to the bulk, capturing solvation free energy of about 10 kcal/mol, the energy gap remains extremely large.

The reason for this energy gap is that the water molecule is extraordinarily well solvated in its location in H_{peroxo}. In addition to the bond to iron, it forms two hydrogen bonds with oxygen atoms possessing large partial negative charges, each worth 10−15 kcal/mol (in contrast to hydrogen bonds to typical neutral polar acceptors, which typically have values no larger than ∼5 kcal/mol). There is no other physical location for this water in which a comparable amount of energy can be recovered. We estimate the penalty for removing the water to be 25−30 kcal/mol, based on calculations in which the water was removed from H_{peroxo}.

Figure 12 presents our suggested Q structure in which one donor oxygen atom of the bridging Glu243 in H_{peroxo} has been displaced by one oxygen atom supplied originally by dioxygen. The resulting dangling Glu243 oxygen atom forms a strong hydrogen bond to a nicely positioned backbone NH group in the second coordination shell. A second interesting feature of this structure is that one of the hydrogen atoms on the water is now hydrogen bonded to the coordinating oxygen on the shifting carboxylate. This hydrogen bond is absolutely necessary if one is to achieve a short nonbonded Fe−Fe distance in the context of a model in which the water is retained. The reason is that, as we have found through numerous calculations, two eclipsed oxygen atoms in the core cannot approach each other closer than ∼3.2 Å without a bridging proton, presumably due to electrostatic repulsion. If the irons are to move to within ∼2.5 Å of one another, the oxygen atoms have to be closer than 3.0 Å, which can only occur when mediated by a hydrogen bond.

A consequence of the need for the bound water molecule is that each of the iron atoms in our Q structure is 6-coordinate. XAS experiments on samples of intermediate Q from *M. trichosporium* OB3b contained pre-edge areas consistent with an iron atom coordination number no greater than 5.[35] There is intrinsic error in this type of analysis, however, which is exacerbated by the need to subtract out the pre-edge features of other diiron species that are present in the freeze-quenched EXAFS samples. A kinetically competent form of intermediate Q may require an open coordination site, but the calculations support our 6-coordinate iron model as the lowest energy structure. We are addressing the possible requirement for an open coordination site in our ongoing work on the C−H bond activation pathway.

From a quantitative point of view, the Q model proposed in Figure 12 is highly satisfactory. The energy is now below that of H_{peroxo}, as it should be. It is in fact the only putative Q structure meeting this criterion that we have found, independent of Fe−Fe distance. The Fe−Fe distance is 2.67 Å. This value is not quite as low as that extracted from the EXAFS data, but as in the case of the reduced enzyme Fe−Fe distance, one has to take into account both experimental error and the fact that the Fe−Fe surface is energetically very flat. To illustrate just how flat this surface is for the present case, we have optimized a version of the Q model in which the Fe−Fe distance is constrained to 2.5 Å, but is otherwise similar to that shown in Figure 12. The total energy of this structure, evaluated with our large basis set defined above, is only ∼3.5 kcal/mol higher than that at the minimized structure. It is thus quite reasonable that the influence of the full protein and/or minor errors in the quantum chemistry could shift the equilibrium position within

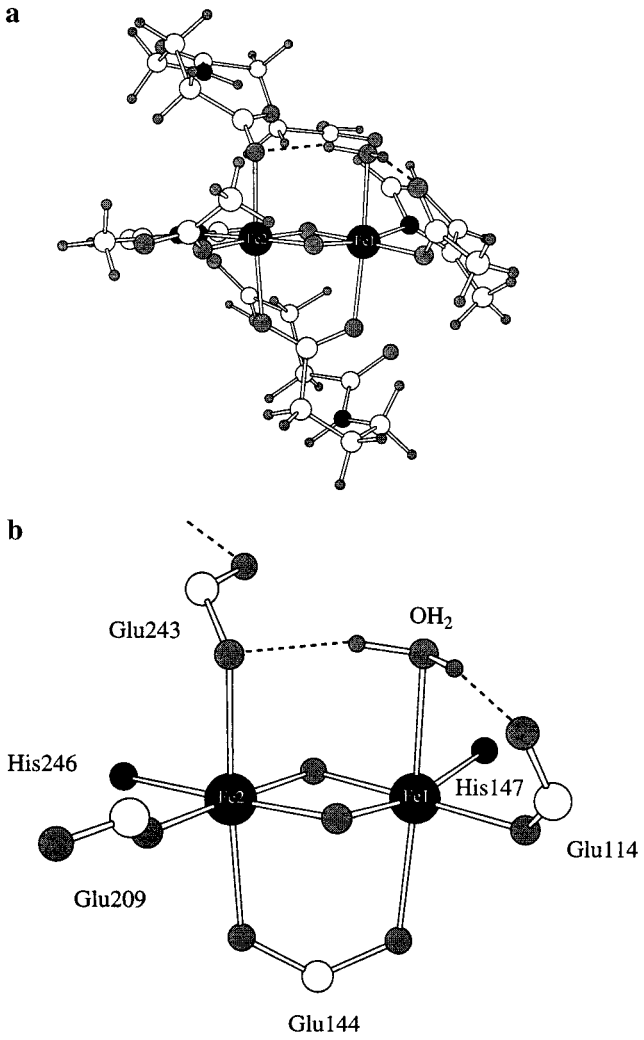

Figure 12. (a) Minimized Q structure. (b) Core of the minimized Q structure. Bond distances are given in Table 7.

Table 5. Fe−Fe Distances (Å) in H_{peroxo} and Q

spin coupling	P Fe−Fe	Q Fe−Fe
F	3.37	2.88
AF	3.36	2.67

0.1−0.2 Å from the calculated minimum. Further exploration of this issue will require additional calculations with more accurate correlation methods and a better representation of the protein environment, and probably repetition of the experiments in ref 35 to validate the degree of precision available with regard to the Fe−Fe distance.

A few significant points should be noted with regard to the sensitivity of the Q structure to basis set and spin state. First, the AF state is lower in energy than the F state as is observed experimentally. This result is encouraging with regard to the ability of the UDFT method to handle the energetics of the spin states. Second, the equilibrium distance of the AF state is about 0.2 Å shorter than that of the F state, emphasizing again that proper treatment of the spin state is required to achieve structural congruence with experiment. These results are summarized in Tables 5 and 6. Third, a DZ basis set also yields incorrect results for the Fe−Fe distance (results not shown); it is necessary to include polarization functions in the core region to achieve the results given above. These three points agree with the conclusions of ref 6, where similar results were obtained in the context of a smaller model with, as discussed above, rather different

Table 6. The H_{peroxo} and Q Energies (au) at Different Levels

spin coupling	basis set	Q (au)	P (au)	Q−P (kcal/mol)
F	lacvp	−2947.131	−2947.172	25.8
F	lacvp**	−2947.940	−2947.977	23.2
F	final basis	−2948.406	−2948.405	−0.6
AF	lacvp**	−2947.309	−2947.316	4.4
AF	final basis	−2948.458	−2948.452	−3.8

Table 7. Bond Distances (Å) of the Minimized Core Structures

	H_{red}	H_{ox}	H_{peroxo}	$H_{superoxo}$	Q (alt.)	Q
Fe1−N His-147	2.190	2.178	2.190	2.204	2.220	2.057
Fe1−O Glu114	2.020	1.989	1.981	1.988	1.864	1.877
Fe1−O H_2O	2.210	2.166	2.106	2.118		2.217
Fe1−O Glu144	2.122	2.043	2.027	1.993	2.103	2.161
Fe2−N His246	2.181	2.189	2.197	2.208	2.124	2.126
Fe2−O Glu209	1.953	1.880	1.850	1.908	1.839	1.810
Fe2−O Glu243	2.256	2.038	2.298	2.154	2.137	2.145
Fe2−O Glu144	2.056	2.212	2.109	2.080	2.100	2.120
Fe1−O bridge 1[a]	2.100	1.943	2.253		2.053	1.760
Fe1−O bridge 2[b]	2.811	2.034	2.064	2.121	2.015	1.767
Fe2−O bridge 1[a]	2.285	2.068	2.231		1.694	1.811
Fe2−O bridge 2[b]	2.754	2.060	2.157	2.333	1.953	1.807
Fe2−O peroxo			2.091			

[a] Bridge 1 is located between the two His residues. The bridging species is one oxygen atom of Glu243 in H_{red}, H_{peroxo}, and $H_{superoxo}$, a hydroxide moiety in H_{ox}, and an oxo atom in Q. [b] Bridge 2 is located between Glu209 and Glu114. The bridging species is a water molecule in H_{red}, a hydroxide moiety in H_{ox}, an oxygen atom of the peroxo in H_{peroxo}, an oxygen atom of a superoxo in $H_{superoxo}$, and an oxo atom in Q.

chemistry. We also investigated the question of whether geometry optimization with the large basis set would move the Fe−Fe distance even closer than the results reported above. This test requires an expensive calculation (DFT geometry optimization with more than 1000 basis functions, many of them highly contracted) but is quite feasible with Jaguar. The results are that the geometry remains quite close to the distance reported above, however, confirming that a DZP basis set is generally quite adequate for geometry optimization unless extremely precise results are required.

Another series of questions raised by the structure in Figure 12 is the energetics of the locations of protons, such as those on the water molecule. A general feature of the system is that it is "proton poor" in that there are numerous highly charged oxygen species that would welcome the addition of a proton. One could imagine placing these protons elsewhere, e.g., the water proton hydrogen bonded to the carboxylate could move over to that residue and hydrogen bond to the resulting hydroxide. We have tested this and numerous related possibilities and the structure we present has the lowest energy by a significant margin.

A final thermodynamic check on the model is obtained by calculating the energy of the following reaction:

$$Q + H_2O + CH_4 \rightarrow \text{oxidized} + CH_3OH$$

By means of this reaction, Q generates methanol and returns the enzyme to the oxidized state. We have carried out this calculation and find the energetics to be favorable by 70 kcal/mol.

(d) Estimation of Exchange Coupling Parameters *J*. As discussed above, our modeling of the "antiferromagnetic" state is based on the use of a broken symmetry (BS) UDFT wave function. Table 8 presents splitting energies calculated for the five species discussed above: oxidized enzyme, reduced enzyme, peroxo intermediate, superoxo intermediate, and Q intermediate. Assuming that the two iron atoms in the core are

Table 8. Splitting Energy of the Intermediates Spin States

species	spin state energy (au)		energy splitting (kcal/mol)
	antiferromagnetic	high spin	
oxidized	−2949.7978	−2949.7976	0.10
peroxo	−2948.4517	−2948.4528	−0.69
superoxo	−2948.4477	−2948.4484	−0.44
Q-best	−2948.4582	−2948.4478	6.55
reduced	−2874.5393	−2974.5474	−5.10

equivalent, we can use the formulas of ref 49 described below to extract the singlet state energy. The "broken symmetry" description is a linear combination of pure spin states. The unrestricted determinant is constructed such that the spins on one part of the molecule are aligned "up" and the spins on the other part of the system are aligned in the opposite direction. For the case where the spins on both parts of the molecule are parallel, however, a single determinant is used to describe this pure "high spin" (HS) state. A Heisenberg Hamiltonian is used

$$H_{\text{spin}} = (-2J)\mathbf{S_A}\mathbf{S_B}$$

where S_A and S_B are the monomer spin quantum numbers, and $\mathbf{S_A}\mathbf{S_B}$ is the operator dot product. The expectation values for the high spin states (HS: $\mathbf{M_A} = +\mathbf{S_A}$, $\mathbf{M_B} = +\mathbf{S_B}$) and the broken symmetry spin states (BS: $\mathbf{M_A} = +\mathbf{S_A}$, $\mathbf{M_B} = -\mathbf{S_B}$) are given by eqs 1 and 2.

$$\langle \mathbf{S_A}\mathbf{S_B} \rangle_{\text{HS}} = +S_A S_B \quad (1)$$

$$\langle \mathbf{S_A}\mathbf{S_B} \rangle_{\text{BS}} = -S_A S_B \quad (2)$$

The energy difference between the HS state and the BS state can be expressed as eq 3,

$$E_{\text{HS}} - E_{\text{BS}} = (-4J)S_A S_B \quad (3)$$

and the J value needed to evaluate the energies of the different spin coupling states can be extracted from this expression. Note also that eqs 4 and 5 give the same information.

$$E_{\text{HS}} - E_{\text{BS}} = (-J){S_{\text{MAX}}}^2 \quad (4)$$

$$E_{\text{HS}} - E_{S=0} = (-J)S_{\text{MAX}}\cdot(S_{\text{MAX}} + 1) \quad (5)$$

Finally, if the antiferromagnetic state is the ground state, one can use the extracted J value to estimate the additional stabilization that would be obtained for a pure singlet state as opposed to the admixture of spin states contained in the broken symmetry wave function. In the present case, it makes sense to compute this value only for intermediate Q (vide infra). Using the energy splitting given in Table 8 and the spin states detailed in Table 2, the J value for Q can be calculated (using eq 6) to

$$J = -(E_{\text{HS}} - E_{\text{BS}})/(4S_A S_B) \quad (6)$$

be −0.41 kcal/mol or $-143\ \text{cm}^{-1}$. This value is quite consistent with an estimate for $J < -30\ \text{cm}^{-1}$ for intermediate Q of MMOH from Mössbauer spetroscopy.[34] The stabilization energy of the singlet compared to the manifold BS state is $-4J$, or 1.64 kcal/mol.

The results can be briefly summarized as follows. For the oxidized enzyme, and the peroxo and superoxo intermediates, the calculated splitting between the high spin and broken symmetry states, and hence the magnitude of J, is very small. In view of the uncertainties in the proposed structures, the DFT methodology, and the broken symmetry approach, our view is that the calculations are consistent with the observation of either a F or AF ground state, as long as the experimental magnitude of J is small. Such is the case for the oxidized enzyme, which is observed experimentally to be an AF singlet with a small exchange coupling, and for the peroxo intermediate.

The calculated splittings for the reduced enzyme and Q are larger and hence can be taken to represent more reasonably predictions of the experimental spin states. In both of these cases, the calculations agree with the experimental observations: the reduced enzyme is seen to be a high spin F species whereas the Q intermediate is observed as an AF singlet. The additional stabilization energy computed for Q aids in reducing the energy of this state further below that of H_{peroxo}, a desirable outcome from the standpoint of comparison with experiment. This energy stabilization, calculated above, can be added to the enthalpy energy of the reaction producing Q species. This energy is reported in Table 4 and hence can be corrected to be greater than 5 kcal/mol.

Discussion

In this work we have obtained accurate geometric representations and relative energies of the catalytic center of soluble methane monooxygenase in its resting diiron(III) and reduced diiron(II) states. To accomplish this task, we expanded the size of the system treated by a significant factor beyond what has heretofore been employed in modeling enzymatic catalysis via correlated quantum chemical calculations. Smaller models were inadequate to reproduce the constraints imposed by the hydrogen-bonding network around the active site. The ability of the calculations to alert us to the presence of a structural constraint in the protein, identified in the difference electron density maps as a previously unassigned water molecule hydrogen bonded to the dangling oxygen atom of Glu209, lends confidence to the validity of the methodology. The complexity of the reduced form of the enzyme, while easily understandable in retrospect, required explicit calculations to make manifest.

The structures of two spectroscopic intermediates, H_{peroxo} and Q, in the catalytic cycle have also been interrogated and plausible geometries for both have been derived. Neither structure had been recognized previously, for both required key hydrogen bonding interactions with a water molecule bound to Fe1. The relative roles of the superoxo and peroxo states of the dioxygen molecule following reaction with the reduced enzyme and the detailed structural features stabilizing each state need further investigation but are of obvious importance. This hydrogen bonding and the mobile carboxylate of Glu243 control the Fe−Fe distance in a novel manner that is likely to be shared in related metalloenzyme active sites. More generally, the hydrogen bonding network around the active site is cleverly used to stabilize the various intermediates in a multiplicity of ways, and hydrogen bonding shifts define much of what happens in moving from one intermediate to another. In view of these findings, it is not surprising that aspartate residues hydrogen bonding to the histidine rings are conserved in most bacterial systems containing dinuclear iron centers that activate dioxygen.[3]

Modeling of these phenomena, as well as the overall nature of the structural constraining forces of the protein, mandates the use of a large structural model. Indeed, the present model is not even fully adequate for the reduced enzyme, as has been noted above. Furthermore, the central nature of hydrogen bonding networks means that establishing the location of protons and coordinated waters is also essential in building an accurate and energetically relevant model. This detail is one of the most challenging aspects of the research direction suggested by this paper and consumed by far the bulk of the computation time

that we have expended on this project to date. One can hope that, as more systems are studied at this level of detail, a database of results will accumulate and the intuition gained will allow fewer alternatives to be investigated in the future. Most likely, however, the solution is simply the application of sufficient computing power which, fortunately, is continuing to grow each year as predicted by Moore's law.

Since X-ray structural information is not available about the intermediates, we may inquire whether our geometric predictions (Tables 2, 5, and 7) are consistent with available spectroscopic and other experimental information. The Fe−Fe distance and some other features of Q agree well with available data, as already indicated. The unsymmetrical nature of the H_{peroxo} structure (Figure 9) is perhaps less consistent with the sharp single quadrupole doublet measured in the Mössbauer spectrum for this species,[29] but a calculation of the quadrupole splitting based on the proposed geometry would be required to make such a determination. The existence of the $H_{superoxo}$ intermediate, which is isoenergetic with H_{peroxo}, is interesting in view of work indicating the presence of a kinetically detectable intermediate in the catalytic cycle between the reduced enzyme and the peroxo state.[26,56] A recent study has suggested that protons are important in catalysis.[57] Although it is true that protons are needed in the overall reaction to form water, the observed pH dependence of H_{peroxo} and Q formation may simply reflect the need for a water molecule, rather than a hydroxide ion, to be coordinated to Fe1 to stabilize these essential catalytic intermediates. The mediation of short, nonbonded O···O contacts by protons supplied by this water molecule has important energetic consequences, as shown above. The structures presented here do not require consumption of one proton for each of the H_{peroxo} and Q formation steps, as proposed for MMOH from *M. trichosporium* OB3b.[57] We are currently investigating whether the observed rate changes upon substitution of D_2O for H_2O might be explained by hydrogen bonding effects within the context of our model. Finally, recent experimental results with the D84E mutant of *E. coli* ribonucleotide reductase R2, in which the residue analogous to Glu114 in MMOH is mutated from aspartate to glutamate, provided stabilization of the peroxo intermediate.[31,58] This added stability may be related to an increased ability to form a hydrogen bond network from the peroxo moiety to an iron-bound water and to Glu84.

The story presented here is incomplete, because it does not address the question of pathways and barriers to transformation between the various catalytic intermediates. With the exception of that between the oxidized and reduced species as discussed above, such an investigation is tractable with the present model and is currently ongoing. The possible mechanisms of methane addition have been extensively discussed in the literature, and the structural models we have elucidated do not necessarily rule out any of them. The occurrence of the water molecule in the Q structure presents additional possibilities, not previously considered, which we are incorporating in our continuing investigations. For example, the bridging oxo atoms are bound unsymmetrically between the two iron atoms with each forming a shorter bond to Fe1, which contains a terminally bound water molecule (Table 7 and Figure 12b). If a (μ-oxo)−Fe bond were required to open during catalysis, this asymmetry may suggest that the terminal oxo bond would form at Fe1.

Although we have emphasized in this discussion the benefits of using a large structural model, it is important to recognize that there may well be aspects of MMO chemistry that are described quite well by the smaller models employed in previous work.[5,6] We do not yet know how the details of the structural models of the various intermediates will impact on the energetics of the catalytic mechanism. For example, it has been proposed that the conversion of methane to methanol takes place via abstraction of a hydrogen atom from methane by one of the bridging oxo atoms, followed by rapid extraction of the resulting −OH group from the core by the methyl radical produced in the first step. There are experimental arguments against this mechanism that we shall not discuss here.[59,60] The main point we wish to make is that the barrier to H atom abstraction obtained with the model depicted in Figure 12 may (or may not) be reasonably well approximated by a much simpler model of the diiron core. Further computational experiments, currently in progress, are needed to resolve this issue.

In summary, we have developed structural and energetic models of the intermediates in the sMMO catalytic cycle, which are in good agreement with available experimental data. A significant improvement in the present methodology that we are actively pursuing is the incorporation of the entire protein in the calculations via mixed quantum mechanical/molecular mechanics (QM/MM) methods. With the use of a reliable QM/MM approach incorporating a QM region of sufficient size, investigations such as those presented here should exhibit increased reliability and become possible on a routine basis. This accomplishment will in turn allow the elucidation of metalloenzyme catalytic mechanisms at a much greater level of precision and with much less effort than has been possible in the past.

Acknowledgment. This work is supported by grants GM40526 (to R.A.F.) and GM32134 (to S.J.L.) from the National Institutes of Health and by an NSF/DOE funded EMSI grant (CHE-98-10367) to R.A.F. Computational resources were provided by the NPACI program under a grant to R.A.F. and by the Research Resource facility at Columbia supported by the NCRR division of NIH (P41 RR06892). D.A.W. is an NIH Biotechnology predoctoral trainee. We are grateful to Prof. P. E. M. Siegbahn for kindly providing a preprint of ref 6.

Supporting Information Available: Tables of atomic coordinates for the structures depicted in Figures 6a, 7a, 9a, 10a, 11a, and 12a (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA9920967

(56) Liu, Y.; Nesheim, J. C.; Lee, S.-K.; Lipscomb, J. D. *J. Biol. Chem.* **1995**, *270*, 24662−24665.
(57) Lee, S.-K.; Lipscomb, J. D. *Biochemistry* **1999**, *38*, 4423−4432.
(58) Moënne-Loccoz, P.; Baldwin, J.; Ley, B. A.; Loehr, T. M.; Martin, J.; Bollinger, J. *Biochemistry* **1998**, *37*, 14659−14663.
(59) Whittington, D. A.; Valentine, A. M.; Lippard, S. J. *J. Biol. Inorg. Chem.* **1998**, *3*, 307−313.
(60) Stahl, S. S.; Lippard, S. J. *Dioxygen and alkane activation by iron-containing enzymes*; Ferreira, G. C., Moura, J. J. G., Franco, R., Eds.; Wiley-VCH Verlag: Weinheim, Germany, 1999; pp 303−321.