



ELSEVIER

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

Coordination Chemistry Reviews 238–239 (2003) 267–290

COORDINATION  
CHEMISTRY REVIEWS

[www.elsevier.com/locate/ccr](http://www.elsevier.com/locate/ccr)

# How iron-containing proteins control dioxygen chemistry: a detailed atomic level description via accurate quantum chemical and mixed quantum mechanics/molecular mechanics calculations

Richard A. Friesner<sup>a,\*</sup>, Mu-Hyun Baik<sup>a</sup>, Benjamin F. Gherman<sup>a</sup>, Victor Guallar<sup>a</sup>,  
Maria Wirstam<sup>a</sup>, Robert B. Murphy<sup>b</sup>, Stephen J. Lippard<sup>c,\*</sup>

<sup>a</sup> Department of Chemistry and Center for Biomolecular Simulations, Columbia University, New York, NY 10027, USA

<sup>b</sup> Schrödinger Inc., Portland, OR, USA

<sup>c</sup> Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Received 11 June 2002; accepted 27 September 2002

## Contents

Abstract	267
1. Introduction	268
2. Overview of metalloprotein active site modeling	269
2.1 Key issues in computational methodology	269
2.2 Experimental systems	270
3. Computational methodology	270
3.1 Quantum chemical methods	270
3.2 QM/MM methods	272
3.3 Analysis of the computational results	273
4. Applications to experimental systems	274
4.1 General Remarks — fundamental considerations	274
4.1.1 Reduction of dioxygen	274
4.1.2 Oxidation of methane	275
4.1.3 Redox reactions of iron complexes	276
4.2 Methane monooxygenase	277
4.2.1 Overview	277
4.2.2 Relevant experimental data	277
4.2.3 MMO active site model	278
4.2.4 Dioxygen activation	279
4.2.5 Substrate hydroxylation	282
4.2.6 Semiclassical dynamics of MMO alkane hydroxylation	284
4.3 Hemerythrin	285
4.3.1 Computational models	286
4.3.2 Reversible binding of dioxygen	286
5. Conclusions	287
Acknowledgements	288
References	288

## Abstract

Over the past several years, rapid advances in computational hardware, quantum chemical methods, and mixed quantum mechanics/molecular mechanics (QM/MM) techniques have made it possible to model accurately the interaction of ligands with

\* Corresponding authors. Tel.: +1-212-854-7607; fax: +1-212-854-7454

E-mail address: [rich@chem.columbia.edu](mailto:rich@chem.columbia.edu) (R.A. Friesner).

metal-containing proteins at an atomic level of detail. In this paper, we describe the application of our computational methodology, based on density functional (DFT) quantum chemical methods, to two diiron-containing proteins that interact with dioxygen: methane monooxygenase (MMO) and hemerythrin (Hr). Although the active sites are structurally related, the biological function differs substantially. MMO is an enzyme found in methanotrophic bacteria and hydroxylates aliphatic C–H bonds, whereas Hr is a carrier protein for dioxygen used by a number of marine invertebrates. Quantitative descriptions of the structures and energetics of key intermediates and transition states involved in the reaction with dioxygen are provided, allowing their mechanisms to be compared and contrasted in detail. An in-depth understanding of how the chemical identity of the first ligand coordination shell, structural features, electrostatic and van der Waals interactions of more distant shells control ligand binding and reactive chemistry is provided, affording a systematic analysis of how iron-containing proteins process dioxygen. Extensive contact with experiment is made in both systems, and a remarkable degree of accuracy and robustness of the calculations is obtained from both a qualitative and quantitative perspective.

© 2002 Published by Elsevier Science B.V.

*Keywords:* Iron-containing proteins; Dioxygen chemistry; Mixed quantum mechanics/molecular mechanics calculations

## 1. Introduction

Elucidation of the chemistry of metalloprotein active sites at an atomic level of detail has been a long-standing objective of bioinorganic chemistry [1,2]. The past decade has seen an exponential growth in the number of high-resolution metalloprotein crystal structures available in the Protein Data Bank (PDB), accompanied by a wealth of complementary experimental investigations ranging from spectroscopic characterization of intermediates to the synthesis of model compounds designed to mimic metalloprotein function. Despite this effort, however, the number of metalloprotein active sites for which a truly complete picture is available is vanishingly small. It has proved extremely difficult to discriminate between alternative reaction mechanisms and associated structures, or to determine precisely the relative energies of these species, by experimental means alone. Often, interpretation of experimental data in the complex condensed phase protein environment is ambiguous; in other cases, proposed intermediates have not been amenable to observation at all. Finally, even in cases where experimental measurements allow a reliable qualitative picture of the chemistry to be assembled, it is often difficult to understand why a particular structure or reactive pathway is favored, or how energetics are controlled to yield species that are precisely tuned for biological function.

In principle, accurate computational methods are capable of complementing experiments to answer these questions at an appropriate level of detail. However, formidable practical problems have had to be overcome in order to render the study of metalloproteins tractable by computational methods. Firstly, the central role of one or more metal atoms in active site chemistry, and the necessity of treating reactive events, has mandated the use of *ab initio* quantum chemical methods if the desired level of accuracy is to be achieved. While substantial progress has been made in advancing semiempirical methods and force field approaches to treat metal-containing systems and reactive events, these

approaches are not yet capable of achieving quantitative results in the absence of system-specific parameterization to experimental data. Transition metal complexes are structurally and electronically much more diverse than organic systems and thus are difficult to describe with a set of universal parameters [3,4]. Secondly, the treatment of metal-containing systems has been challenging even for *ab initio* methods. Finally, the number of atoms that can be treated via *ab initio* techniques, while increasing significantly over the past decade, is still limited. This limitation has posed a dilemma to theoreticians; use of a model containing a relatively small number of atoms is computationally efficient but can fail to capture both structural and energetic effects of the protein environment, whereas inclusion of a larger number of atoms can lead to intractable CPU times.

Over past 5 years, effective computational methods have emerged to address these problems. The widespread use of density functional theory (DFT) [5,6] deploying gradient-corrected and hybrid functionals has demonstrated that transition metal systems can be treated in a reasonably accurate, yet cost-effective fashion by *ab initio* quantum chemical methods. Advances in DFT-based algorithms, along with dramatic reductions in the cost/performance of computational hardware, allow on the order of 100–200 atoms to be modeled at a quantum chemical level on a routine basis. Progress has also been made specifically in the treatment of transition metals, including such areas as construction of an improved initial guess [7] for the DFT wavefunction and modeling of unpaired spin interactions that often give rise to complex magnetic coupling schemes (*vide infra*). Finally, the development of accurate and efficient approaches for combining quantum mechanical and molecular mechanics modeling (mixed QM/MM methods) [8–10] affords a strategy whereby the core of a metalloprotein active site can be treated by a DFT-based quantum chemical approach while the remainder of the protein is represented by a classical force field. When these techniques are appropriately combined, a methodology capable of modeling

metal-containing proteins on a routine basis can be constructed, providing accurate results for reaction chemistry at the metal center while properly including environmental effects of the protein.

In this review, we describe the approach along these lines that our research groups have pursued over the past several years. We emphasize the interplay between computation and experiment that is necessary to build an accurate picture of the stationary points, intermediates and transition states of a reaction, which characterize the essential chemistry that occurs at a metalloprotein active site. The methods are illustrated by applications to several different metalloprotein systems, all of which involve the chemistry of dioxygen. These examples demonstrate both the quantitative accuracy that is now available via computation, and the physical insight that can be obtained by complementing experimental results with calculations.

The paper is organized as follows. In [Section 2](#), we provide an overview of the central problems involved in construction of the computational methodology, and the specific experimental systems examined in the present paper. [Section 3](#) presents a brief description of the key elements of the computational methodology. [Section 4](#) begins with a general analysis of the energetics of dioxygen chemistry, and then discusses two protein systems in detail, summarizing the results we have obtained for each. Finally, in [Section 5](#), the conclusion, we summarize the results and discuss future directions.

## 2. Overview of metalloprotein active site modeling

### 2.1. Key issues in computational methodology

Metalloproteins are essential for many key life processes in biochemistry, including respiration, nitrogen fixation, and photosynthesis. While their functions are quite diverse from a chemical point of view, they share a number of common elements that must be treated in studying any of these systems, modeling of which presents challenges from both computational and theoretical perspectives. We identify these key elements as follows:

(1) Substrate binding to and product release from the metalloprotein active site. Substrate binding can involve an initial noncovalent component as well as interaction with the metal and/or atoms in the functional core of the active site. Determination of the free energy of binding to the active site, an essential precondition for carrying out subsequent chemical function, is a highly demanding task, one for which the optimal methodology is not fully worked out. The energetics of product release shares these characteristics.

(2) Determination of the structures and relative energies of key intermediates and transition states

required to carry out the specified chemistry. The basic machinery with which we approach this problem is DFT-based QM and QM/MM technology capable of handling large QM regions (150–200 atoms). Geometry optimization of proposed intermediates and transition states is by far the largest consumer of CPU time, so it is important that they be as efficient as possible. Subsequent computation of single point energies with quantum chemical basis sets large enough to achieve convergence of the energy is also necessary in a substantial number of cases. Second derivative calculations are required to provide estimates of zero point energies, which can have significant effects on activation barriers. A proper treatment of spin states is an integral aspect of modeling transition metal-containing systems, as briefly discussed below. Finally, intimate interaction between theory and experiment is necessary to characterize successfully the stationary points in complex metalloprotein systems. Experimental results provide means of constructing plausible initial guesses for the quantum chemical calculations, and also in assessing the concordance of the optimized structures with the actual physical system. In conjunction with the ability to reject proposed structures and pathways for chemical transformations based on quantum chemical energetics, a powerful approach to characterization of active site chemistry can be assembled. In turn, the theoretical results can help refine, correct and predict experimental findings.

(3) Treatment of solvation. In principle, the most rigorous approach is to include the aqueous environment of the protein explicitly; in our methodology, this task is accomplished by using a continuum solvation approach, namely solution of the Poisson-Boltzmann equation [11,12]. In practice, many enzyme active sites are physically well separated from bulk solvent and are overall electrically neutral so that long range dielectric effects are relatively unimportant. Bound waters in contrast can be quite important and must be treated explicitly, typically as part of the QM region. A reasonable approximation, adopted by many other workers, is to neutralize the ionizable residues on the surface of the protein, thus approximating the effects of dielectric screening of solvent, which dramatically reduces the long range electrostatic field to the point where it can reasonably be ignored. In our work to date, we have employed this latter approach, with the intention of investigating a more accurate treatment of solvation in the future. Thus, we do not discuss the more elaborate solvation methods further in what follows.

(4) Computation of rate constants from the potential energy surface. Assuming that the relevant QM and QM/MM calculations can be performed to generate a reasonably accurate potential energy surface, direct comparison with experimental measurements often requires evaluation of rate constants for particular

reactions. This task can be carried out by a wide range of methods, beginning with simple transition state theory and ramping up to full blown quantum mechanical simulations that include tunneling and zero point effects in a rigorous fashion. Although our work in this area is at an early stage, encouraging results are beginning to emerge; we present one example of such in some detail below.

Finally, the accuracy of the various components of the computational methodology is of critical importance to the entire modeling enterprise. Firstly, catastrophic failures of the methodology are not acceptable if the approach is to become a widely used tool that can be used routinely to complement experimental studies of new metalloprotein systems. A benchmark level of confidence in terms of the robustness of structural and energetic results needs to be established. As with experimental techniques, there will be error bars on the results that are obtained, and these in some cases may preclude drawing unambiguous conclusions about mechanistic questions. The key point, however, is that when the error bars are reduced to below a critical threshold — and we believe that such is now the case for several of the above central elements of modeling — theory and experiment can become equal partners rather than simply using theory to confirm what is already known experimentally, and to reject the theoretical results if there is any disagreement. Such a partnership already exists in the study of small molecules in the gas phase, where for many properties calculations can be more accurate and reliable than the experimental data. The extension of this vision to complex, condensed phase systems containing transition metals is an exciting prospect that appears to be within reach in the next several years.

We address this issue in detail in Section 3. We discuss the intrinsic accuracy of DFT methods for transition metal-containing systems, validation of the QM/MM methodology, and overall success of this type of modeling when applied to metalloproteins. Individual comparisons with experiment for the specific systems we have investigated are also incorporated into the sections discussing those systems. Section 3 provides a general perspective in which these and other results can be assembled to make a realistic assessment of where the methodology stands with respect to accuracy at the present time.

## 2.2. Experimental systems

We have chosen to focus our initial work primarily on iron-containing proteins that bind or activate dioxygen. Oxidation is a fundamentally important chemical process in biological systems, and the approaches that have developed evolutionarily to provide both sufficient rapidity and specificity for these reactions are among

the most sophisticated and diverse in nature. By investigating a number of such systems in detail, we are able not only to elucidate each individual set of mechanisms, but also to compare and contrast the different pathways for dioxygen chemistry enforced by the different metalloproteins. Ultimately, it is hoped that research along these lines will lead to general principles that can then be applied to other types of reactions.

The two specific metalloproteins that we discuss in this review are soluble methane monooxygenase (sMMO or MMO) and hemerythrin (Hr). MMO belongs to the family of non-heme hydroxylases and inserts an oxygen atom from dioxygen into aliphatic C–H bonds, yielding as products an alcohol and water. An external source of electrons is required to drive the catalytic cycle. Hr, in contrast, is a reversible dioxygen carrier; its function, in certain marine invertebrates, is to bind O<sub>2</sub> and then release it at required locations in the organism, much as hemoglobin is used in humans. Hr is also a non-heme diiron protein like MMO; unlike MMO, however, its active site prevents, rather than facilitates, subsequent reactions of dioxygen.

## 3. Computational methodology

### 3.1. Quantum chemical methods

The development of gradient-corrected and hybrid density functional (DFT) methods [13–15] over the past decade has qualitatively transformed our ability to treat metal-containing systems with *ab initio* quantum chemical techniques. Whereas DFT methods do not provide the capability of converging absolute and/or relative energies to an arbitrary accuracy, as is available from wavefunction-based methods, current DFT functionals deliver a combination of speed and accuracy that is simply not available with other approaches. The computational cost of a DFT calculation is comparable to that for a Hartree–Fock calculation. Both in essence require the solution of a three-dimensional field equation, and scale similarly with system size:  $\sim N^2$  for normal three-dimensional molecular structures, where  $N$  is the number of atoms in the system, and asymptotically as  $\sim N$  or  $N \ln N$  for ultra-large systems, a limit that is much easier to reach in one or two dimensions than in a normal three-dimensional enzyme active site geometry via the use of multipole technology.

The accuracy of DFT methods is at present best established for organic systems. For example, calibration of a series of functionals using the G2 database of small molecule atomization energies yields an average error on the order of 3 kcal mol<sup>-1</sup> [16] over a total of 148 species for the top performing functional, B3LYP [13,15,17–19]. It should be noted that total atomization

Table 1  
Theoretical and experimental bond dissociation energies (kcal mol<sup>-1</sup>)

	Calculated (B3LYP)	Experimental	Reference
Sc–H <sup>+</sup>	60.0	56.3	[27]
Sc–CH <sub>3</sub> <sup>+</sup>	60.8	55.8	[27]
Sc–O <sup>+</sup>	155.5	164.0 (±1.8)	[34]
Sc–OH <sup>+</sup>	118.8	119.0 (±2.1)	[29]
Ti–H <sup>+</sup>	56.3	53.3	[27]
Ti–CH <sub>3</sub> <sup>+</sup>	58.8	51.1	[27]
Ti–O <sup>+</sup>	147.1	158.0 (±1.6)	[34]
Ti–OH <sup>+</sup>	116.3	111.0 (±2.8)	[29]
V–H <sup>+</sup>	54.1	47.3	[27]
V–CH <sub>3</sub> <sup>+</sup>	49.9	46.0	[27]
V–O <sup>+</sup>	128.2	134.0 (±3.9)	[34]
V–OH <sup>+</sup>	96.1	101.0 (±4.4)	[29]
V–CO <sub>2</sub> <sup>+</sup>	21.0	17.3	[30]
OV–CO <sup>+</sup>	26.1	24.2	[30]
Cr–H <sup>+</sup>	36.5	31.6	[27]
Cr–CH <sub>3</sub> <sup>+</sup>	35.6	26.2	[27]
Cr–O <sup>+</sup>	72.4	86.0 (±2.8)	[34]
(CO) <sub>5</sub> Cr–CO	40.7	37.0 (±2.0)	[31]
Mn–H <sup>+</sup>	48.6	47.5	[27]
Mn–CH <sub>3</sub> <sup>+</sup>	50.2	49.1	[27]
Mn–O <sup>+</sup>	60.1	48.0 (±3.0)	[34]
Mn–OH <sup>+</sup>	76.0	79.0 (±5.7)	[32]
Fe–H <sup>+</sup>	55.7	48.0 (±1.4)	[36]
Fe–CH <sub>3</sub> <sup>+</sup>	62.3	54.6	[27]
Fe–O	91.7	92.0 (±3.0)	[36]
Fe–O <sup>+</sup>	75.0	80.0 (±1.4)	[34]
Fe–OH	72.7	77 (±4)	[36]
Fe–CO	27.1	31 (±8)	[36]
Fe–CO <sub>2</sub> <sup>+</sup>	17.0	9.5	[30]
Fe–S	69.9	76.3	[36]
Fe–F	103.2	107 (±5)	[36]
Co–H <sup>+</sup>	48.7	45.7	[27]
Co–CH <sub>3</sub> <sup>+</sup>	53.4	48.5	[27]
Co–O <sup>+</sup>	67.3	75.0 (±1.2)	[34]
Co–OH <sup>+</sup>	66.8	71.7 (±0.9)	[29]
Co–OH <sub>2</sub>	40.4	38.5	[33]
Co–CO <sub>2</sub> <sup>+</sup>	24.5	19.9	[30]
Ni–H <sup>+</sup>	45.2	38.7	[27]
Ni–CH <sub>3</sub> <sup>+</sup>	49.2	44.7	[27]
Ni–O <sup>+</sup>	54.2	63.0 (±1.6)	[34]
Ni–OH <sup>+</sup>	59.7	56.0 (±4.6)	[29]
Ni–CO <sub>2</sub> <sup>+</sup>	27.3	24.9	[30]
Cu–O <sup>+</sup>	34.2	31.0 (±2.8)	[34]
Zn–O <sup>+</sup>	35.4	38.0 (±2.8)	[34]
(CO) <sub>5</sub> Mo–CO	40.1	40.0 (±2.0)	[31]
(CO) <sub>5</sub> W–CO	44.8	46.0 (±2.0)	[31]

energies represent an exceptionally demanding test of the methodology; most chemical transformations will exhibit a greater cancellation of error, for example, isodesmic reactions. Other properties, such as ionization potentials and electron affinities, have also been extensively studied [20–26], generally with excellent results.

For transition metal-containing systems, experimental data are less readily available than for organic molecules. Nevertheless, a substantial number of comparisons have been made between theory and experiment [27–36]. Several of these are summarized in Table 1,

where we concentrate on metal–ligand bond energies. Again, prediction of bond energies is a highly demanding task, and other types of reactions are likely to exhibit smaller errors<sup>1</sup> than those listed in Table 1. It is generally observed that Becke's three-parameter hybrid exchange functional [13] gives a notable improvement, which is also a trend observed in benchmark calculations using small organic molecules. For example, the average error of M–O<sup>+</sup> bond energies [34] for the first row transition metals (Table 1) increases from 7.5 to 12.1 kcal mol<sup>-1</sup> when the B3LYP and BLYP results are compared. Our own experience confirms this trend, and the B3LYP functional is our method of choice for all DFT calculations. The accuracy obtained will typically be good enough to discriminate among proposed reaction mechanisms and structures of intermediates, while the CPU time is sufficiently modest to allow the hundreds of calculations per system required to explore properly the range of proposed alternatives (e.g. structures, protonation states, spin states) to be carried out on a routine basis.

A number of other aspects of the DFT methodology are essential in modeling of metalloprotein chemistry. Computation of the relative energetics of various spin states is important in understanding reaction pathways and in comparing the computations with experimental data. Unrestricted DFT (UDFT) methods, employing the broken symmetry methodology pioneered by Noodleman [37] for systems with multiple metals, provide a remarkably useful approach to the problem of non-standard magnetic coupling between the metal centers, as we shall illustrate below. Relativistic effective core potentials [38–40] are essential in treating heavier elements, and useful for the lower transition metal series as well. Finally, converging DFT calculations to the ground electronic state is much more challenging for metal-containing systems than it is for organic systems, where it is generally trivial to begin the calculation with an accurate initial guess, for example from extended Hückel calculations. The key to obtaining reliable results in metal-containing systems is the use of a more sophisticated initial guess methodology, which takes into account physically important features of metal–ligand interactions, such as charge transfer between the ligand and the metal.

All of the calculations discussed below have been carried out using the JAGUAR suite of ab initio electronic structure programs [41], which we have developed over

<sup>1</sup> We have recently investigated systematically the performance of B3LYP/cc-pVTZ(-f) level of theory in combination with a continuum solvation model on computing redox potentials of organometallic complexes [63] and found a theory–experiment correlation that suggests an average error for the ionization potential of organometallic complexes comparable to that found for smaller organic molecules in [16].

the past several decades. JAGUAR employs numerical methods enabling particularly efficient DFT calculations using both hybrid [13] and gradient corrected functionals [14,15], as well as a number of features specific to transition metal computations, for example an effective initial guess methodology along the lines suggested above [7]. The combination of speed and convergence available in JAGUAR has enabled us to study large QM and QM/MM models (100–200 atoms) on a routine basis, while retaining a high degree of reliability with regard to the accuracy of the results. Geometry optimizations are all carried out using the B3LYP [13,15,17,19] functional in combination with the 6-31G\*\* basis set. Iron is represented using the LACVP\*\* basis [38–40] that includes an effective core potential. Although this level of theory has proved to be sufficient for obtaining structures, we found the energies to be not always reliable. Thus, it is our standard protocol to recompute the energies in an additional single point calculation using Dunning's [42] correlation consistent triple- $\zeta$  basis cc-pVTZ(-f). For iron we utilize a modified version of LACVP\*\*, denoted as LACV3P\*\* where the exponents were decontracted to match the effective core potential with the triple- $\zeta$  basis.

### 3.2. QM/MM methods

Even with DFT methods and modern computational resources, it is impractical to treat an entire protein with a reasonably accurate QM approach. Therefore, in modeling metalloprotein active sites, there are two alternative approaches to making the computation tractable. Firstly, one can construct a purely QM model by truncating the active site core and capping the truncated regions with hydrogen atoms. This traditional approach has been used for many years with increasingly large models, as computing power and advancing QM technology has allowed treatment of larger systems. It is still a valuable one, as demonstrated in detail below. The second approach is to retain the entire protein, but to treat only a specified region around the active site core at a QM level, modeling the remainder of the protein via molecular mechanics. It is also possible to utilize semiempirical methods for the remainder of the protein, but we do not consider this methodology here. In principle, this approach allows for an optimal combination of accuracy and computational efficiency; in practice, formidable problems must be overcome with regard to producing a robust definition of the interface between the QM and MM regions. This task is relatively straightforward when there is no chemical bond between the QM and MM atoms; however, defining an accurate interface between bonded atoms has proved to be quite difficult.

Several methods for constructing the QM/MM interface between bonded atoms have been intensively

investigated by a number of research groups during the past decade. The earliest method proposed, which is still widely employed, involves the use of link atoms [43–48]. An alternative approach, which we employ, is to utilize frozen, localized molecular orbitals as the interface between a QM and MM atom [49–56]. We have described the details of this technology elsewhere, and refer the reader to those publications for a rigorous mathematical development of the theory and its implementation [54–56]. Here, we provide a brief overview of the basic ideas behind our approach, and summarize the validation tests that we have performed.

In our view, parameterization of the QM/MM interface must be approached in the same fashion as development of a new molecular mechanics force field; that is, the parameters must be optimized for the specific environment of each chemical bond for which a QM/MM interface is to be constructed. The first step in this process is to construct a library of frozen orbitals using small molecule models. For modeling metalloprotein active sites, the requirements along these lines are straightforward. We assume that the interface will be built only by making cuts in the peptidic part of the protein, with the central core containing the transition metal atom(s) and ligands treated at the QM level. Implementation of such a strategy requires treatment of QM regions on the order of 150 atoms on a routine basis, which is quite tractable using our QM software and inexpensive hardware such as a state of the art personal computer (PC). In this case, a library of frozen orbitals and associated parameters is needed for making cuts in the peptide backbone and in various side chains. For backbones we allow cuts between the  $\alpha$ -carbon and either the flanking carbonyl carbon or amide nitrogen. The side chain cuts are made between the  $\alpha$ -carbon and the  $\beta$ -carbon. We employ capped dipeptides as model systems, and have constructed and extensively tested a library of parameters to handle all cuts of the type specified above.

The parameters themselves are of the molecular mechanics form. Stretching, bending, and torsional parameters are fitted to quantum chemical data on conformational energetics of small model peptides so as to ensure that the methodology generates reasonably accurate conformational energetics. The precision available here is limited by the force field itself. In order to balance the electrostatic field from the frozen orbital, a point charge is placed along the frozen bond with the position and magnitude optimized. This type of correction is essential for reproducing properties such as deprotonation energies. Finally, van der Waals parameters for the QM region are adjusted to reproduce hydrogen bonding energies of pairs of small molecules. In general the parameterization is specific to a particular QM methodology and basis set and to the MM force field that is employed. In what follows we exclusively

utilize the OPLS-AA molecular mechanics force field for the protein [57] and have optimized parameters for a number of different QM methods. We shall discuss only the DFT implementation in the present publication, however.

Once the parameters are defined, implementation of analytical gradients allows optimization of QM/MM geometries to locate optimized structures for reactants, products, and transition states via standard geometry optimization algorithms. In addition to characterization of these stationary points, a more complete potential surface can be mapped out by the use of restrained geometry optimizations. We have not yet implemented a fully rigorous approach to the computation of analytical second derivatives for the QM/MM Hamiltonian. However, an excellent approximation can be obtained by building a truncated QM model capped with hydrogens, retaining the original QM/MM geometry, and using the usual quantum chemical methods to compute analytical second derivatives.

The accuracy of QM/MM methods can be assessed with two complementary strategies. The first is direct comparison with fully QM calculations on the same system. We have carried out extensive tests of this type, including evaluation of relative conformational energies of model dipeptides, deprotonation energies of amino acid side chains, and binding affinities of both small molecule dimers and larger systems. These data are presented in detail in Ref. [54]; a statistical summary of results in each of the categories enumerated above is given in Table 2. Some of the observed deviations are a consequence of the force field itself, while others derive from inaccuracies in the QM/MM interface. The important point of Table 2 is that the discrepancies with full QM calculations are on average smaller than those inherent in presently available DFT approaches; hence, our expectation is that the use of our QM/MM model does not represent a limiting factor in modeling of transition metal active site chemistry. These results show also that the absolute errors in both aspects of the computations are sufficiently small to allow reliable conclusions concerning reactive chemistry to be drawn in the great majority of cases. Although there will be some situations where an uncertainty of a few kcal

mol<sup>-1</sup> will preclude choosing between alternative mechanistic proposals, such is not the typical scenario we have encountered. Furthermore, when the energy difference is that small, one would expect both channels to contribute to the mechanism. A further technical point is that most of the tests discussed above involve imposition of a single QM/MM boundary region; however, we have also examined the effects of multiple cuts by considering a model for the hemerythrin active site in which a large number of QM/MM cuts are examined. This last study demonstrates that there are no systematic errors as the number of interface bonds is increased, even when these bonds are spatially quite close together, as they are in the test case.

The second strategy is to carry out QM/MM computations using the full protein, and systematically to enlarge the QM region. Several tests of this type are discussed below for hemerythrin. The results are entirely consistent with those discussed in the previous paragraphs with regard to the magnitude of the differences between the larger and smaller models; in fact, the errors are actually somewhat smaller in the full protein systems. These results obviate the concern that the error estimates discussed above might not be transferable from small models to the full protein calculations.

Our QM/MM methodology, specifically optimized for protein active sites, has been implemented in the QSITE [58] program, which has been constructed by coupling JAGUAR with the IMPACT molecular mechanics program of Levy et al. [59]. The OPLS-AA protein molecular mechanics force field developed by Jorgensen et al. [57] is used to describe the molecular mechanics part of the system. QSITE contains a graphical user interface allowing automated specification of the QM and MM regions simply by pointing and clicking on protein residues displayed visually on the screen. Our intention has been to develop the software to the point where the technology is accessible to non-experts as well as specialists in quantum chemistry calculations.

### 3.3. Analysis of the computational results

Our philosophy in understanding protein active site chemistry is to combine model quantum chemical studies, using a range of system sizes, with QM/MM modeling. The QM/MM computations provide the most accurate results for structures and energies, because they enable the effects of the protein on structures and energetics to be calculated. The interpretation of data in the absence of model studies is problematic, however. To assign specifically particular structural or energetic effects to the protein, a reference system for energies and structures in the absence of the protein is required. Although in principle many different decompositions are possible, in practice it is straightforward to specify the most useful ones from the standpoint of intuitive

Table 2

Average errors for a test set of molecules made by QM/MM (B3LYP based) and OPLS-AA in kcal mol<sup>-1</sup> compared to fully quantum mechanical benchmark calculations

Property	QM/MM	OPLS-AA
Side chain conformational energies	1.10	1.74
Alanine tetrapeptide conformational energies	1.02	1.64
Dimer binding energies	0.72	0.80
Deprotonation energies	0.52	n/a

Detailed results can be found in Ref. [54].

chemical understanding. In this section, we outline our approach as it has been deployed in the examples below.

We first consider the concept of strain energy. The central core of a metalloprotein active site may adopt a geometry in the presence of the protein scaffold that is substantially different from what it would adopt if optimized in the gas phase. In order to investigate such effects quantitatively, we have to define a computational protocol for measuring such deviations. The protocol we have employed for this purpose is as follows.

(1) Carry out a QM/MM geometry optimization of the target intermediate or transition state using the full protein.

(2) Remove a selected part of this structure from the protein, cap the dangling bonds with hydrogens, and carry out a restrained optimization in which the hydrogen atoms alone are optimized. The size of this structure is arbitrary; different selections provide different windows of physical insight. A typical choice is to retain the first coordination shell, truncating the ligands to the metal atom(s) at the  $\beta$ -carbon of the amino acid providing the ligand. To create a somewhat smaller model, truncation is made at the termination of the ligated functional group, e.g. at the C of a COOH group.

(3) Record the energy of the structure produced in (2), then carry out a full geometry optimization of this structure. Structural and energetic differences between the structures produced in (2) and (3) can then be assessed; they provide a measure of the spatial and energetic degree of strain of the intermediate in the protein, respectively.

This protocol provides a more nuanced approach to the question of protein-induced strain than a simple declaration of the central core as ‘strained’ or ‘unstrained’. In our investigations to date, we have found a quantitative range of results emerging from this procedure. Although we have not yet studied enough metalloproteins to make statistical characterization of the distribution of effects worthwhile, a sufficient accumulation of data will clearly be available in the near future and will provide bioinorganic chemists with a greatly enhanced understanding of this heretofore poorly understood phenomenon.

A second central concept in analyzing QM/MM results is relative stabilization of stationary points by protein non-bonded interactions including electrostatic and van der Waals terms as well as hydrogen bonds. These terms can readily be separated from the remainder of the energy and typically have straightforward physical interpretation. Examination of the differential contribution of these terms in structures of the reactants, transition state and products provides insight into how the protein influences energetics via these interactions.

The third major type of term in the energy is the molecular mechanics energy of the MM region. Differences in this energy between structures indicate reorganization of the protein in response to changes in the QM region. One has to be very careful to make sure that such energy differences are not due to trapping in different local minima, a very large number of which exist on the MM potential surface, leading to comparisons of the total energy of the two structures that are not meaningful. Problems of this type can be circumvented by making sure to optimize the various states in the QM region starting from the same MM conformation. If necessary this process can be carried out iteratively. If one intermediate in a catalytic cycle exhibits a much lower MM energy, other structures can then be optimized starting from that MM conformation, and the lowest total energy selected in the end if there are a number of different MM starting points, from which to choose. A more rigorous approach to this problem involves the use of simulated annealing of the MM region, rather than minimization, after every QM geometry optimization step. We are currently testing such approaches, which appear quite promising.

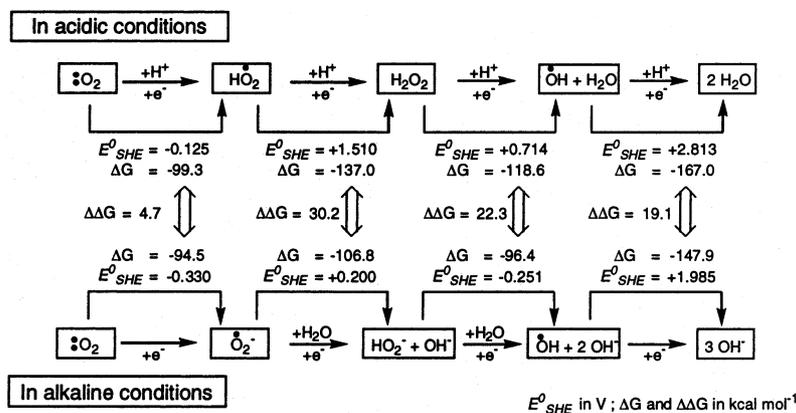
The final important energy term is the intrinsic energy of the QM region at the specified QM/MM geometry. We define this term to include the interactions of the QM charge density with the frozen orbitals, as well as the energy associated with the specialized QM/MM parameters. Where to include the interface terms is arbitrary; we do not believe that separating them out would increase physical insight. Energy differences in this term reflect both inherent energetic differences of the various chemical species comprising the intermediates and transition state under study, and electronic reorganization of the QM wavefunction in response to the external protein electric field. Separation of these two effects is possible in principle by comparing the energies of QM model systems evaluated at the QM/MM geometry, and contrasting this with the QM energy differential seen in the QM/MM energy; we have not extensively pursued this decomposition in all cases that follow.

## 4. Applications to experimental systems

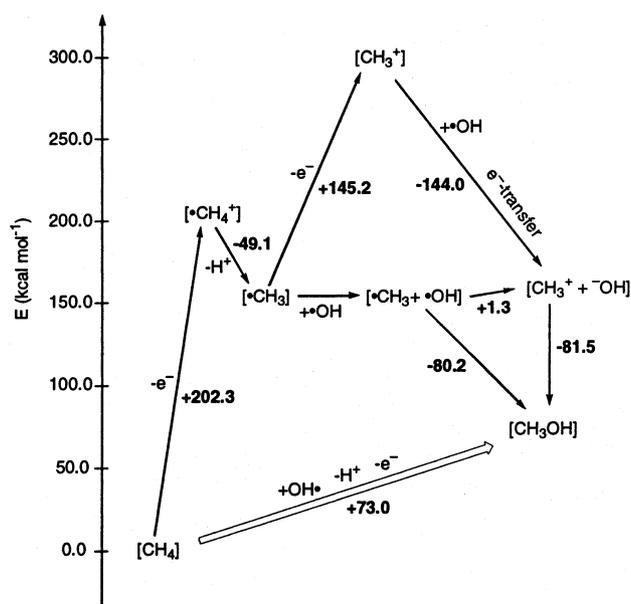
### 4.1. General remarks — fundamental considerations

#### 4.1.1. Reduction of dioxygen

A key feature in both enzymatic reactions reviewed here is the reduction of dioxygen. Whereas in MMOH the dioxygen chemistry has the purpose of generating a species that ultimately activates C–H bonds, the formation of a peroxo species is the final stage of the



Scheme 1. Latimer diagrams of the stepwise reduction of dioxygen in acidic and basic conditions [59].



Scheme 2. Non-catalytic oxidation of methane in water computed on B3LYP/cc-pVTZ(-f)++ level of theory.

dioxygen-reduction cycle in Hr. Simply writing down the Latimer diagram of dioxygen, which is the most common way of listing stepwise one-electron redox potentials [60], and the energies required to transform methane oxidatively to methanol as shown in Schemes 1 and 2, respectively, reveals important clues about what the catalyst has to accomplish. In Scheme 1, the experimental standard potentials [60] are listed and also converted to their respective free reaction energies,  $\Delta G$ , by using the simple standard relationship  $\Delta G = -FE^{\circ}_{abs}$ .  $F$  is the Faraday constant and  $E^{\circ}_{abs}$  the absolute potential, which can be easily obtained from the experimental standard potential  $E^{\circ}_{SHE}$  referenced to the standard hydrogen electrode by adding the absolute potential of the standard hydrogen electrode (SHE) of 4.43 eV [61].

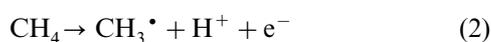
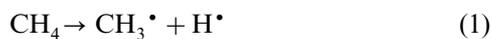
Scheme 1 shows not surprisingly that reduction is greatly favored in acidic conditions, where protons are abundant in excess and stabilize the increasing negative charge on oxygen by forming covalent and/or hydrogen bonds. Nature makes heavy use of this pH dependence of redox-reactions and utilizes the protein to control the local pH, that is, the accessibility of protons at the reactive center. The first reduction of molecular dioxygen to the superoxide species is the most difficult step, having a reduction potential of  $-0.125$  and  $-0.330$  V corresponding to free reaction energies of  $-99.3$  and  $-94.5$  kcal mol<sup>-1</sup> in acidic and alkaline conditions, respectively. The next step, formation of peroxide from superoxide, is thermodynamically more downhill, which is a clear indication of the intrinsic instability of the superoxide species towards disproportionation. Thus, superoxide intermediates in catalytic cycles are expected to have a short life time and generally to be more reactive. The neutral OH-radical formed from the peroxy species is even more unstable with respect to disproportionation. It is the strongest oxidant among all intermediates, with a redox potential of 2.813 and 1.985 V corresponding to an overall driving force of  $-167.0$  and  $-147.9$  kcal mol<sup>-1</sup> in acidic and alkaline conditions, respectively. As presented below, this powerful oxidant is required to perform the difficult task of C–H bond activation, whereas the finely tuned mechanism of reversible dioxygen binding in Hr requires careful control of proton assistance to the electron transfer between dioxygen and hydroperoxide.

#### 4.1.2. Oxidation of methane

Scheme 2 provides a calculated<sup>2</sup> thermodynamic map of methane oxidation in water in the absence of a

<sup>2</sup> All structures fully optimized on the B3LYP/cc-pVTZ(-f)++ level. Solvation energies are evaluated using the continuum solvation model as implemented in Jaguar ( $\epsilon = 80.37$ ).

catalyst. The first oxidation step, generation of the methane radical cation, is by far the most difficult and is uphill by  $202.3 \text{ kcal mol}^{-1}$ . This value corresponds to a computed standard redox-potential [62,63]  $E_{\text{calc}}^{\circ}$  of  $-4.34 \text{ V}$  referenced to the SHE, which is out of reach for most oxidation agents available at realistic conditions. Even the strongest oxidant identified above, the neutral OH radical with a reducing power of  $-167 \text{ kcal mol}^{-1}$  in acidic conditions, would not be able to oxidize methane without further assistance. Once formed, the methane radical cation is not surprisingly a strong Lewis acid and deprotonation is energetically downhill by  $49.1 \text{ kcal mol}^{-1}$ . Thus, it is in general advantageous to couple the oxidation of methane with proton transfer, which formally constitutes a hydrogen atom abstraction. Hence, it is not surprising that in MMOH the demanding first step of the C–H activation is intimately coupled to proton transfer. The transfer of a proton and an electron can be understood formally either as hydrogen abstraction (Eq. (1)) or proton-coupled electron transfer (Eq. (2)). Hydrogen abstraction implies homolytic C–H bond cleavage and thus the reaction energy would simply be  $104 \text{ kcal mol}^{-1}$ , the well-known C–H bond strength in methane.



Proton-coupled electron transfer includes formally the oxidation of  $\text{H}^{\bullet}$  to give a proton, which adds an energy term for the oxidation of a hydrogen atom. To understand the energetics of the C–H activation catalysis, it is helpful to consider the proton-coupled electron transfer. In a realistic chemical setting, both electron and proton transfer are most likely coupled and difficult to study independently. By casting the problem as a proton-coupled electron transfer, however, rather than formally formulating  $\text{H}^{\bullet}$  as the product, C–H bond activation divides into redox and acid/base components. These are well understood concepts in inorganic chemistry and thus allow for deriving a more intuitive understanding of the catalytic cycle. In Scheme 2, we show two possible pathways by which the reaction can progress after the first proton-coupled electron transfer step. The methyl radical could react with the neutral OH radical<sup>3</sup> to form methanol in a radical-recombination type mechanism, or it could undergo a second oxidation to form a methyl cation that reacts with hydroxide. Energetically, the oxidative formation of the methyl cation is uphill by  $145.2 \text{ kcal mol}^{-1}$ , which strikingly matches the calculated reduction energy for

<sup>3</sup> We assume the addition of a neutral OH radical because our calculations indicated that the first step of the catalytic cycle of MMOH involves a radicaloid oxygen that shows some resemblance with a neutral OH radical.

the OH radical of  $-144.0 \text{ kcal mol}^{-1}$  ( $E_{\text{calc}}^{\circ} = 1.81 \text{ V}$ ). The result is essentially thermoneutral formation of the charge-separated species  $\text{CH}_3^+$  and  $\text{OH}^-$  from the initial radical pair.<sup>4</sup> Thus, both pathways are equally possible thermodynamically and have been proposed in the catalytic cycle of MMOH, as will be pointed out below.

#### 4.1.3. Redox reactions of iron complexes

The magnitude of energies required to activate both dioxygen and methane without any catalyst demonstrates the need for a very versatile reaction center that can supply multiple electrons at high potentials. At the same time, the highly reactive intermediate, in particular the odd electron species, must either be separated carefully from the environment or formed in situ when the substrate is already bound to the reaction center to control the selectivity and avoid undesirable side reactions. Transition metals are attractive candidates for serving as electron reservoirs, because electrons can be added and removed from metal-d orbitals without creating highly unstable radicaloid species. In addition, the range of possible oxidation states at the metal center introduces a tunable redox agent, where the driving force of the redox reaction is not only a function of oxidation state, but can also be controlled by changing the first shell ligand coordination. In multinuclear metal systems there is an additional cooperative mechanism present, which is best evidenced in the magnetic behavior of the iron centers that can either display ferro- or antiferromagnetic coupling of the unpaired electrons to each other. An extensive and systematic assessment of these control mechanisms is currently underway in our laboratories, but is beyond the scope of this report. Only one illustrative example of such a structure–redox relationship is described in greater detail for the proton-coupled first electron transfer step of the methane activation cycle performed by MMOH below.

The iron centers in MMOH exist in oxidation states of II, III and IV at different stages of the catalytic cycle, whereas in Hr, the oxidation states of II and III are accessed. Fe(II) and Fe(III) are also the most common oxidation states of simple inorganic iron complexes in solution. Although Fe(IV) species exist in alkaline

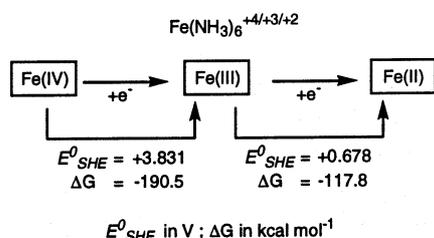
<sup>4</sup> Note that we use the redox potentials from our DFT calculations in Scheme 2, whereas the experimental values are given in Scheme 1. Thus, there is a slight mismatch of the redox potentials for the redox pair  $\text{OH}/\text{OH}^-$ . Our calculations on the B3LYP/cc-pVTZ(-f)++ level of theory gave a potential  $E_{\text{calc}}^{\circ}$  of  $1.81 \text{ V}$ . Experimentally,  $1.98 \text{ V}$  was reported in alkaline conditions [60]. This difference of  $0.17 \text{ V}$  corresponds to  $3.9 \text{ kcal mol}^{-1}$  and is only a slightly larger error than the usual margin of uncertainty in electrochemical measurements and uncritical for the purpose of establishing a detailed understanding of the reaction profile.

environments, they are usually unstable and commonly decompose in solution. Scheme 3 shows computed redox potentials for a typical iron complex in an octahedral ligand environment, the hexaammine complex  $[\text{Fe}(\text{NH}_3)_6]^{x+}$  ( $x = +2, +3, +4$ ). For reduction of the putative Fe(IV) ion, our calculation predicts a driving force of  $-190.5 \text{ kcal mol}^{-1}$ , corresponding to a redox potential  $E_{\text{SHE}}^0$  of  $+3.831 \text{ V}$ , whereas  $+0.678 \text{ V}$  is predicted for the Fe(II)/Fe(III) couple. Compared to the redox series shown above for dioxygen and methane, the difference between the first and second redox potentials is remarkably large. More importantly, the redox potential for the Fe(III)/Fe(IV)-couple is in the correct order of magnitude to promote the most difficult oxidation steps in both the dioxygen and methane oxidation cycles. Although the redox potentials of the iron complexes in their respective peptidyl framework are certainly different, these simple calculations establish approximately the region of redox potentials accessible for the different redox couples. They also make clear that the highly endothermic first step of methane activation, even with proton-coupling, which has been estimated to be  $\sim 153 \text{ kcal mol}^{-1}$  requires the oxidative power of an Fe(IV) center. Reversible binding of dioxygen in Hr to form the peroxo species, however, can be accomplished with the relatively mild redox behavior of the Fe(II)/Fe(III)-couple.

## 4.2. Methane monooxygenase

### 4.2.1. Overview

Methane monooxygenase (MMO), an enzyme that methanotrophic bacteria use to convert methane to methanol under ambient conditions [64–67], has been the subject of significant interest over the past 15 years. An understanding of the MMO catalytic process could lead to the development of an improved industrial process for converting methane to methanol, which would in turn increase the utility of methane as an energy source [68,69]. Such knowledge could also spur development of general biomimetic alkane hydroxylation and hydrocarbon oxidation catalysts. Having the versatility to oxidize a wide variety of hydrocarbons [70–75], methanotrophs could be used in bioremediation of contaminated land and water areas [73]. In



Scheme 3. Computed redox potentials of  $\text{Fe}(\text{NH}_3)_6^x$  ( $x = +4, +3, +2$ ) in aqueous solution ( $\epsilon = 80.37$ ).

addition, methanotrophs play a role in controlling global warming by limiting the amount of methane, a greenhouse gas, emitted into the atmosphere [76].

The oxidation of methane by dioxygen to form methanol and water occurs at non-heme, dinuclear iron active sites within the hydroxylase component (MMOH) of the multi-component soluble MMO (sMMO) system. The principle steps of the MMOH catalytic cycle are known and consist of four key intermediates shown schematically in Fig. 1. The resting form of MMOH is  $\text{H}_{\text{ox}}$ , an Fe(III)Fe(III) state [77–83]. Reduction of  $\text{H}_{\text{ox}}$  yields an Fe(II)Fe(II) form termed  $\text{H}_{\text{red}}$ .  $\text{H}_{\text{red}}$  reacts with dioxygen eliminating water and producing Fe(III)Fe(III)  $\text{H}_{\text{peroxo}}$  (also called compound P), which then decays to form compound Q. It is now widely accepted that Q [84–89], a high-valent Fe(IV)Fe(IV) species, is the catalytically potent intermediate that activates C–H bonds in a variety of substrates.

### 4.2.2. Relevant experimental data

The wealth of experimental data for the MMO catalytic cycle, obtained from enzymes isolated from *Methylococcus capsulatus* (Bath) [74,77–80,90–99] and *Methylosinus trichosporium* (OB3b) [71–73,100–109], can be separated into two categories—data pertaining to the activation of dioxygen from  $\text{H}_{\text{red}}$  to Q and data pertaining to the reaction of hydrocarbon substrate with Q. Such a distinction is possible because the formation and decay of intermediates in the dioxygen activation phase of the MMO cycle can be studied in the absence of substrate.

$\text{H}_{\text{red}}$  was determined by Mössbauer and EPR spectroscopy [81,82,85,110,111] to contain two weakly ferromagnetically coupled high-spin ferrous iron centers. The first consensus intermediate following  $\text{H}_{\text{red}}$  precedes  $\text{H}_{\text{peroxo}}$  formation, backed by the decay of  $\text{H}_{\text{red}}$  being faster than  $\text{H}_{\text{peroxo}}$  formation and by kinetic [88] and  $^{18}\text{O}_2$  isotope effect studies [112,113]. This intermediate lacks the characteristic diferrous EPR signal and the optical characteristics of  $\text{H}_{\text{peroxo}}$  [65].  $\text{H}_{\text{peroxo}}$ , the first spectroscopically detected intermediate after  $\text{H}_{\text{red}}$ , forms at a rate independent of dioxygen concentration [88], which further supports the existence of the earlier species. In the MMO Bath system,  $\text{H}_{\text{peroxo}}$  formation

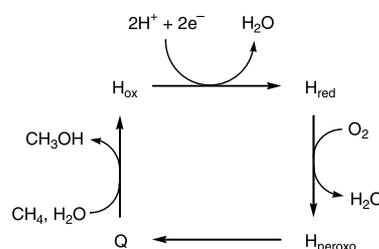


Fig. 1. Overview of the MMO catalytic cycle.

was pH independent [85] with a rate constant of  $1\text{--}2\text{ s}^{-1}$  at  $4\text{ }^{\circ}\text{C}$  [88], and experiments indicated the possibility of the sequential formation of two differing peroxo species [84,85,88]. On the other hand,  $\text{H}_{\text{peroxo}}$  formation was strongly pH dependent and displayed a solvent KIE in the MMO OB3b system [112] and had a formation rate of  $9\text{--}12\text{ s}^{-1}$  at  $4\text{ }^{\circ}\text{C}$  and pH 7 [102]. Mössbauer [84,85,100] spectroscopy indicated that  $\text{H}_{\text{peroxo}}$  contains two electronically equivalent high-spin antiferromagnetically coupled Fe(III) centers.

In the final step of dioxygen activation,  $\text{H}_{\text{peroxo}}$  decays to form Q with a rate matching that of Q formation [85,112]. Despite significant efforts, the structure of Q has to date not been determined and remains elusive, although many important characteristics are known from spectroscopic and reactivity studies. Experiments on the MMO Bath system showed  $\text{H}_{\text{peroxo}}$  decay to be pH independent, have no solvent KIE, have a  $k_{\text{obs}}$  of  $0.45\text{ s}^{-1}$  at  $4\text{ }^{\circ}\text{C}$  [85], and to result most likely from homolytic cleavage of the O–O bond in  $\text{H}_{\text{peroxo}}$ . However, results from the MMO OB3b system indicate that P decay is strongly pH dependent, has a significant solvent KIE [112], has a  $k_{\text{obs}}$  of  $2.4\text{ s}^{-1}$  at  $4\text{ }^{\circ}\text{C}$  and pH 7 [102], and support heterolytic cleavage of the O–O bond in  $\text{H}_{\text{peroxo}}$ . In either case, Mössbauer spectroscopy shows Q to have two nearly equivalent high-valent Fe(IV) atoms which are antiferromagnetically coupled [84,85,87,100].

A considerable amount of experimental data similarly exists concerning the kinetics of substrate reaction with Q and has been thoroughly reviewed recently [65]. The rate of Q decay in the presence of methane was measured to have a  $k_{\text{obs}}$  of  $2\text{--}10\text{ s}^{-1}$  at  $7.8\text{ }^{\circ}\text{C}$  [88]. In experiments with MMO Bath and radical clock substrate probes, radical-derived ring-opened products are not observed, implying a radical lifetime of less than  $\sim 150\text{ fs}$  and arguing against the feasibility of a free radical rebound mechanism [91,114]. Other experiments with chiral ethane ( $[1\text{-}^2\text{H}_1, 1\text{-}^3\text{H}_1]\text{ethane}$ ) as the substrate yielded a retention of configuration of  $\sim 70\%$  [92,103]. Such experiments rule out a completely concerted reaction since the stereochemistry is lost in  $\sim 30\%$  of the products. A radical-type mechanism is also operative a fraction of the time in radical clock experiments with 1,1-dimethylcyclopropane in the MMO OB3b system

[104]. The experimentally observed oxidation states at each iron center and magnetic coupling are summarized in Table 3.

#### 4.2.3. MMO active site model

Despite extensive experimental work mentioned above devoted to deriving a molecular level understanding of the enzymatic catalysis, many crucial details of the reaction are still unclear and thus attracted much theoretical attention [35,115–130], most notably involving density functional theory (DFT) [5,6]. Several of these studies are concerned with the exact structure and characteristics of the unknown key intermediate Q. The strategy of our computational approach is different from that of others in that we explicitly include the experimentally observed antiferromagnetic coupling of the iron centers and use a large scale model that contains approximately 100 atoms. The broken symmetry (BS) orbital approach proposed by Noodleman [37] affords a reasonable working approximation to the intrinsically multi-configurational electronic structure of an antiferromagnetically coupled dinuclear complex in the DFT framework. There is credible evidence that the thermodynamics of the catalytic cycle can be reproduced with smaller model systems and neglecting the antiferromagnetic coupling, consistent with a relatively weak magnetic interaction [35,119–130]. However, both chemical intuition and an in-depth electronic structure comparison of the ferromagnetically and antiferromagnetically coupled models indicate that key details of the catalytic cycle can only be obtained by employing antiferromagnetic coupling and a larger model [131,132]. A too small computational model may give rise to artificial attractive or repulsive forces between substrate and small ligands that represent peptidyl fragments or allow access to the metal center, which is in reality unavailable due to steric crowding from the first shell ligand coordination.

Our studies indicate that a  $\sim 60$ -atom model, which includes imidazole ligands representing histidine and acetate ligands replacing glutamate moieties, is a very reasonable one for the hydroxylation cycle, once Q has been formed, and recovers most results obtained with the 100-atom model at least semi-quantitatively. However, even this mid-sized model displays serious shortcomings for the dioxygen activation phase, where the rigidity of the extended peptidyl framework is needed to enable a consistent catalytic pathway leading to formation of Q. The exact impact of the extended ligand structure and the protein is subject of a detailed future QM/MM study. The use of ferromagnetic coupling between the metal centers is equally inadequate since it results in a serious dilemma of a spin forbidden electron transfer scenario, which is often ignored entirely in the literature. It is clear that the overall two-electron oxidation of an alkane substrate, for example, requires transfer of two electrons with opposite spin from

Table 3  
Iron atom oxidation states and couplings in the MMO catalytic cycle from  $\text{H}_{\text{red}}$  to Q

Species	Fe1 ox. state	Fe2 ox. state	Coupling
$\text{H}_{\text{red}}$	II	II	Ferromagnetic
$\text{H}_{\text{superoxo}}$	II	III	Ferromagnetic
Asymmetric $\text{H}_{\text{peroxo}}$	III	III	Antiferromagnetic
Symmetric $\text{H}_{\text{peroxo}}$	III	III	Antiferromagnetic
Q	IV	IV	Antiferromagnetic

substrate to the diiron core because both electrons originate from the same C–H  $\sigma$ -orbital. A ferromagnetically coupled diiron core is by default only able to accept one electron for each iron center with parallel spin. Thus, the assumption of a ferromagnetically coupled diiron complex does not allow for analyzing the electronic structure of the one-electron reduced intermediate of the Q-fragment in a straightforward manner to identify features that promote the second electron transfer. In fact, the one-electron intermediate of a ferromagnetically coupled model would require some type of a spin exchange mechanism to enable the second electron transfer at all. A consistent electronic picture and its analysis, however, are crucial for deriving an understanding of the reaction in terms of a structure–reactivity relationship that can ultimately be exploited to optimize biomimetic constructs in a non-peptidyl framework. A systematic study of this electronic structure–reactivity relationship is currently underway in our laboratories. In the following, we summarize the results of a study prerequisite to deriving such an in-depth understanding that is primarily concerned with generating a catalytic cycle that is both consistent with experimental results and intrinsically complete.

All calculations on the MMOH active site presented herein were carried out using a  $\sim 100$ -atom model of the MMOH active site based upon the crystallographically determined coordinates [78] for  $H_{\text{red}}$  (Fig. 2). In determining the model size, the essential goal [115,116] was to maintain key structural features seen in the crystallographic data [77,78]. The smallest initial model would be approximately 60 atoms in size and include the two iron atoms, any coordinated water molecules, and amino acid side chains coordinated to the iron atoms (i.e. Glu114, Glu144, His147, Glu209, Glu243, His247) with the backbone  $C_{\alpha}$  atoms replaced by capping hydrogen atoms. Optimization of this size model for  $H_{\text{red}}$ , however, leads to significant differences with the crystal structure [115,116]. Specifically, the Fe–Fe distance increases by 0.8 Å, Glu243 changes from having

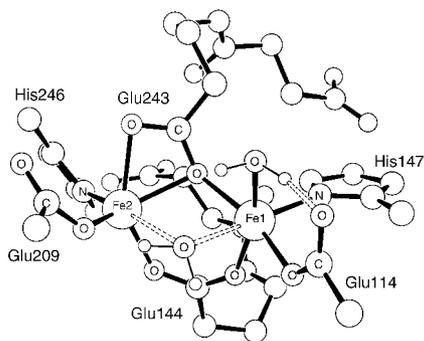


Fig. 2. An illustration of the core arrangement of the reduced form of MMOH found in the crystal structure, adapted from Ref. [77]. Residues coordinated to the iron atoms are labeled only. Dashed lines represent hydrogen bonds.

one oxygen bridging the two iron atoms (Fig. 2) to being a bridging  $\mu$ -1,2 carboxylate, and ligand orientations are generally qualitatively different. The RMSD of the optimized small model  $H_{\text{red}}$  with the MMOH crystal structure was 1.653 Å [115,116]. The model was therefore increased in size to add protein residues which constrain the diiron core and surrounding ligands. Asp143 and Asp242, which form hydrogen bonds to the histidine rings of His147 and His246 respectively, were included to stabilize the geometry [116]. Not only did this change lead to the conservation of key structural motifs, but the RMSD with the MMOH crystal structure was also reduced to 0.411 Å. The two carboxylates in the second coordination shell were protonated, so that the model would then have zero net charge, as in the MMOH active site which is buried in a low dielectric medium. Lastly, an artificial constraint was included between the  $\alpha$  carbons of Asp143 and Asp242 in order to constrain these two portions of the protein backbone from drifting apart. These backbone segments, which are not constrained by hydrogen bonds to each other, do not separate in the crystal structure due to side chain packing interactions. The distance of this constraint from the diiron core, and the very small gradient inducing drift between the two segments, assessed by carrying out unconstrained calculations, assures virtually no impact on the active site chemistry.

#### 4.2.4. Dioxygen activation

The dioxygen activation phase of the MMO catalytic cycle (Fig. 3) begins with the diferrous  $H_{\text{red}}$  form. As suggested by experimental work [85,110,111],  $H_{\text{red}}$  was modeled with high spin  $S=2$  iron atoms coupled ferromagnetically. A geometry optimization of the crystal structure (Fig. 2) maintained the water molecule *trans* to the histidine rings as being coordinated to only Fe1. The optimized model displays an Fe–Fe distance about 0.35 Å longer than found in the crystal structure [77], a difference which is in part due to an energetically flat Fe–Fe potential surface [115].  $H_{\text{red}}$  decays with loss of the water molecule *trans* to the histidines into solution and subsequent reaction with dioxygen. The loss of this water molecule creates open coordination sites on each iron atom, both iron atoms then being five-coordinate, which are available for dioxygen to occupy.

The next calculated intermediate may correspond to the  $H_{\text{peroxo}}$  precursor and is a mixed-valent Fe(II)Fe(III) superoxo intermediate, here termed  $H_{\text{superoxo}}$ . Transfer of one electron from Fe2 to dioxygen creates a superoxo  $O_2^-$  ligand bound in a bidentate fashion to Fe2. Mulliken spin and charge populations analyses (Tables 4 and 5, respectively) confirm that Fe2 is principally involved in  $H_{\text{superoxo}}$  formation. The O3–O4 bond distance shows considerable lengthening versus an  $O_2$  bond length of 1.214 Å. Bidentate binding of superoxo

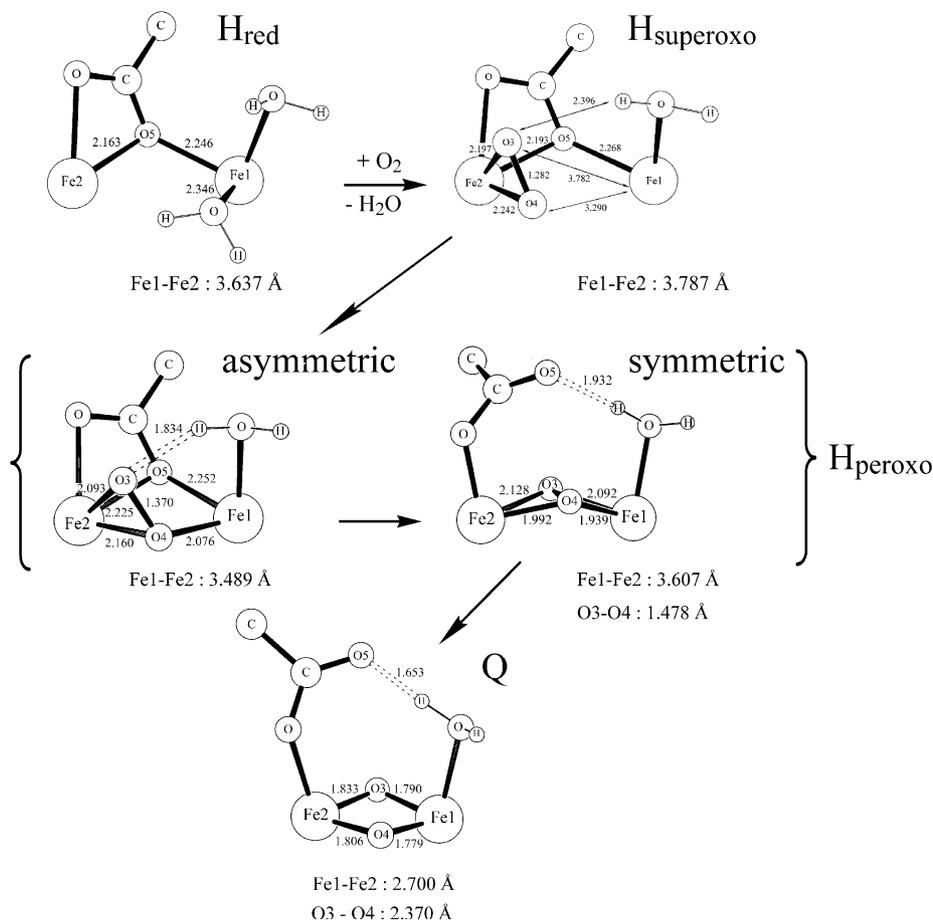


Fig. 3. Detail of the MMO catalytic cycle from  $H_{red}$  to Q. Core structures of the minimized structures are shown. Roman numerals indicate iron atom oxidation states. The numbers indicate distances (Å). Fe1–Fe2 distances are also noted. Dashed lines represent hydrogen bonds.

Table 4  
Mulliken spin populations for intermediates in the MMO catalytic cycle from  $H_{red}$  to Q

Species	Fe1	Fe2	O3	O4
$H_{red}$	3.77	3.75	n/a	n/a
$H_{superoxo}$	3.70	3.89	0.78	0.79
Asymmetric $H_{peroxo}$	−4.01	3.98	−0.15	0.07
Symmetric $H_{peroxo}$	−4.02	4.00	−0.10	−0.09
Q	−3.49	3.54	−0.06	−0.03

Table 5  
Mulliken charge populations for intermediates in the MMO catalytic cycle from  $H_{red}$  to Q

Species	Fe1	Fe2	O3	O4
$H_{red}$	0.63	0.72	n/a	n/a
$H_{superoxo}$	0.74	0.81	−0.14	−0.12
Asymmetric $H_{peroxo}$	0.87	0.85	−0.32	−0.31
Symmetric $H_{peroxo}$	0.99	1.01	−0.40	−0.43
Q	0.80	0.80	−0.60	−0.53

to Fe2 provides stability to the superoxo intermediate and indicates an advantage of a side-on attack by dioxygen during superoxo formation. End-on attack by dioxygen would have yielded a situation with an unstable O3 directed outwards from the diiron core. Other stabilization resulting from side-on attack comes from the interaction of O3 with the water coordinated to Fe1. The superoxo structure was taken to be ferromagnetically coupled, with the one unpaired electron on the superoxo moiety oriented parallel to the five unpaired spins on the ferric Fe2 and the four unpaired spins on the ferrous Fe1. Formation of the superoxo intermediate is  $5.71 \text{ kcal mol}^{-1}$  endothermic versus reactants consisting of dioxygen and  $H_{red}$  without the water molecule that was trans to the histidine rings.

$H_{superoxo}$  then decays to form an asymmetric Fe(III)Fe(III)  $H_{peroxo}$  intermediate, with a structure first proposed by Dunietz, et al. [115]. Mulliken spin and charge populations (Tables 4 and 5, respectively) show that Fe1 has now become oxidized as Fe2 had been oxidized in the superoxo intermediate. O3 and O4 of the peroxo unit have been reduced as shown by their Mulliken charge populations, and each have spin populations near zero as expected for an  $O_2^{2-}$  species.

Compared with their positions in  $H_{\text{superoxo}}$ , O3 and O4 have both shifted to assume positions closer to Fe1, with O4 forming a new bond with Fe1. A key hydrogen bond is also formed between O3 and the water coordinated to Fe1. Antiferromagnetic coupling was employed in this intermediate. The asymmetric peroxy intermediate is 0.94 kcal mol<sup>-1</sup> lower in energy than  $H_{\text{superoxo}}$ .

Our calculations indicate that there is no direct redox path connecting the asymmetric  $H_{\text{peroxy}}$  with the next major reaction intermediate, Q. Instead, a thermodynamic driving force of 8.1 kcal mol<sup>-1</sup> affords formation of a symmetric isomer of  $H_{\text{peroxy}}$  (Fig. 3). This structure is highlighted by a bridging non-planar  $\mu\text{-}\eta^2\text{:}\eta^2$  (butterfly) peroxy group reminiscent of a  $H_{\text{peroxy}}$  structure put forth by Siegbahn [133]. The carboxylate group of Glu243, which had one oxygen atom in a bridging position, has undergone a shift to a position in which O5 can form a hydrogen bond to the water coordinated to Fe1. This shift in turn allows O3 to rotate in order to assume the bridging position opposite to O4 and previously occupied by O5. In the resulting structure, each iron atom (still coupled antiferromagnetically) now has an essentially equivalent coordination environment. Mulliken spin populations (Table 4) remain nearly identical in the two structures, while tighter binding of the peroxy moiety in the symmetric structures leads to slightly larger charge populations (Table 5) for the symmetric  $H_{\text{peroxy}}$ . These results point to a thermodynamic preference for homolytic O–O bond cleavage and imply a pH independent rate for  $H_{\text{peroxy}}$  decay [85]. The isomerization leads clearly to an additional weakening of the O–O bond in the peroxy moiety, as indicated by a bond elongation of 0.11 Å (Fig. 3), and prepares the subsequent bond cleavage.

The symmetric  $H_{\text{peroxy}}$  decays to form compound Q. This transformation can be viewed as an oxidative addition of the peroxy moiety to the diiron center to afford an  $\text{Fe(IV)}_2(\mu\text{-O})_2$  diamond core structure. The hydrogen bond between O5 and the water coordinated to Fe1 is preserved and strengthened in Q. Formation of Q from the symmetric  $H_{\text{peroxy}}$  is 11.04 kcal mol<sup>-1</sup> downhill. The reduction of the peroxy group is evidenced by the Mulliken charge populations (Table 5) for O3 and O4. Both the spin and charge populations show the oxidation of the iron atoms from ferric  $d^5$  to ferryl  $d^4$  configurations. In agreement with spectroscopic data [85,100], the structure for Q was treated with antiferromagnetic coupling and resulted in Fe1 and Fe2 residing in nearly identical coordination environments.

The necessity of generating the catalytically active di- $\mu$ -oxo core of Q indicates the importance of the carboxylate shift which occurs between  $H_{\text{superoxo}}$  and the symmetric  $H_{\text{peroxy}}$ . Without the Glu243 O5 atom repositioning to form the hydrogen bond, O3 would not have been able to serve as the second bridging oxygen atom in either the symmetric  $H_{\text{peroxy}}$  or Q.

Our suggestion for the structure of Q is significantly different from that previously proposed by Siegbahn [122] that was based on DFT calculations using mid-sized models and assuming a low-spin  $S = 1$  configuration of the d-electrons on each iron atom. Siegbahn also identified a  $\text{Fe(IV)}_2(\mu\text{-O})_2$  system, but suggested that Glu243 served as a bridging carboxylate (in addition to the bridging carboxylate from Glu144). Siegbahn's Q structure would be arrived at from our proposed Q geometry by displacing the remaining water molecule from Fe1. However, the structure with two bridging carboxylates is  $\sim 50$  kcal mol<sup>-1</sup> higher in energy than the currently proposed  $H_{\text{superoxo}}$ ,  $H_{\text{peroxy}}$ , and Q structures [116]. The energy difference is attributable to the considerable solvation of the water molecule in our Q geometry. The water molecule is bonded to Fe1 and hydrogen bonded to oxygen atoms having large partial negative charges in both Glu114 and Glu243. Considering the energies of the two hydrogen bonds, the penalty for removing the water molecule is estimated to be 25–30 kcal mol<sup>-1</sup> [116].

The oxidation states and the magnetic coupling between the iron centers throughout the dioxygen activation phase of the MMO catalytic cycle are presented in Table 3. The relative energies of the aforementioned intermediates are shown in Fig. 4. Formation of the symmetric  $H_{\text{peroxy}}$  and Q from  $H_{\text{red}}$  are overall exothermic by 3.32 and 14.36 kcal mol<sup>-1</sup>, respectively. The activation energy barriers in Fig. 4 reflect approximate preliminary barrier heights from early explorations of the reaction coordinate from  $H_{\text{red}}$  to Q.

Formation of the asymmetric  $H_{\text{peroxy}}$  from the superoxy intermediate is kinetically viable if the barrier between the two structures is small. Initial mappings of the reaction coordinate, rough transition state searches, and the minimal structural difference between the two intermediates indicate such to be a definite possibility. A large barrier is likely between the two peroxy structures, with conversion from the symmetric to asymmetric  $H_{\text{peroxy}}$  requiring both a carboxylate shift and a reorganization of oxygen atoms O3 and O4 originating from dioxygen. Based on data obtained so far for these two shifts, each of these two steps will contribute  $\sim 8$ – $10$  kcal mol<sup>-1</sup> to the overall activation barrier. Such a large total barrier between the two peroxy structures indicates that they should be observable, and suggests the need for additional Mössbauer spectroscopy to be done on these peroxy intermediates. The large activation energy for the peroxy interconversion reaction has also warranted the simultaneous investigation of a path directly from the superoxy intermediate to the symmetric  $H_{\text{peroxy}}$  in order to discover a potentially lower overall barrier to generating the symmetric  $H_{\text{peroxy}}$ . Preliminary barrier heights along this reaction path are significantly lower than the barrier

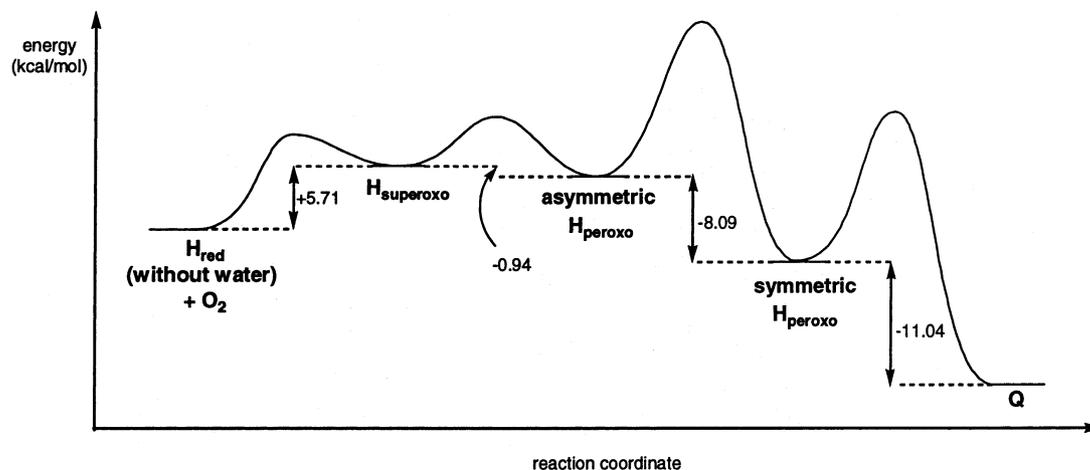


Fig. 4. Reaction coordinate for the MMO catalytic cycle from  $H_{red}$  to Q.

for Q formation from the symmetric  $H_{peroxo}$ , which in turn would explain why the symmetric compound  $H_{peroxo}$  is the first spectroscopically observed intermediate after  $H_{red}$  [65]. This finding would also agree with experimental rate constants showing the  $H_{peroxo}$  formation rate being faster than the  $H_{peroxo}$  decay rate [85,88,102].

Research is currently underway to explore superoxo structures with alternate orientations of the unpaired electron in  $O_2^-$  and with antiferromagnetic coupling of the iron atoms. The possibility of transient ferromagnetically coupled symmetric and asymmetric  $H_{peroxo}$  intermediates is also being addressed. Transition states and reaction coordinates between the various intermediates are similarly being determined so that the mechanism of dioxygen activation can be fully elucidated.

#### 4.2.5. Substrate hydroxylation

In our studies of the reaction of Q with methane [117], substrate was assumed to approach the core via a path *trans* to His147 and His246 as proposed first by Lippard et al. [78], which is also the approach most consistent with a hydrophobic binding cavity that exists in this location of the full protein structure. Various alternatives for the activation of methane by Q were explored, with the only energetically reasonable result being a straight-on approach by methane to the bridging oxo atom *trans* to the histidine ligands. Although prominent alternatives [124–130,134] were proposed in the literature, the bridging oxo group is now widely accepted as the most likely site of the initial nucleophilic attack. Our calculations predict an enthalpy of activation of 17.9 kcal mol<sup>-1</sup> for this reaction, which reduces to 14.7 kcal mol<sup>-1</sup> if differential vibrational zero-point-energy is included. Van der Waals energy for methane within the MMO substrate pocket prior to hydroxylation can be estimated to be ~6 kcal mol<sup>-1</sup> [135], with current work in our lab aiming to calculate this number definitively. The net enthalpy of activation is then ~8.5 kcal mol<sup>-1</sup>,

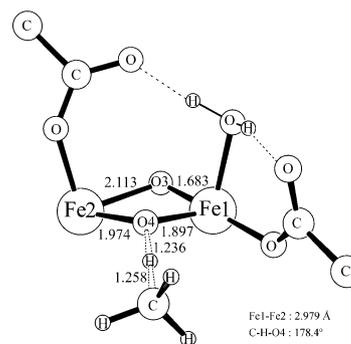


Fig. 5. Core structure of the minimized transition state resulting from hydrogen atom abstraction in the reaction of methane with Q. The numbers indicate distances (Å). The Fe1–Fe2 distance is also noted. Single dashed lines represent hydrogen bonds.

and with our calculation of  $\Delta S^\ddagger$  equalling  $-29.0$  cal mol<sup>-1</sup>K<sup>-1</sup>, the calculated free energy of activation at 25 °C is ~17 kcal mol<sup>-1</sup>, which is in good agreement with experiments [88], as are the individual computed enthalpy and entropy terms.<sup>5</sup> The structure of the rate limiting first transition state, TS-1, depicted in Fig. 5 reveals a notable structural distortion of the Fe<sub>2</sub>O<sub>2</sub> core compared to Q (Fig. 3). The Fe1–O3 bond shortens by 0.11 Å whereas Fe2–O3 is elongated by 0.28 Å. Both Fe1–O4 and Fe2–O4 bonds are elongated by 0.12 and 0.17 Å, respectively. At first inspection, the increase of  $\alpha$ -spin density at Fe2 (Table 6) accompanied by approximately the same amount of  $\alpha$ -spin loss at the methyl-carbon indicates direct electron transfer from CH<sub>4</sub> to Fe2. However, at the same time there is a notable increase of  $\beta$ -spin density at both O3 and O4, which is not consistent with a simple electron transfer mechanism from substrate to Fe2 through O4, in which

<sup>5</sup> The free energy of activation of 17.1 kcal mol<sup>-1</sup> translates into a rate constant of 1.81 s<sup>-1</sup> at 25.0 °C and can be compared with experimental rate constants of 2–10 s<sup>-1</sup> at 7.8 °C from Ref. [89], which translate to barriers of 15.1–16.0 kcal mol<sup>-1</sup>.

Table 6  
Mulliken spin populations during the reaction of Q with methane

Species	Fe1	Fe2	O3	O4	H	CH <sub>3</sub>
Q	−3.49	3.54	−0.06	−0.03	n/a	n/a
TS-1	−3.39	4.17	−0.38	−0.28	−0.02	−0.43
I-BR	−3.38	4.24	0.01	−0.26	−0.02	−0.92
I-C	−3.44	4.23	0.04	−0.28	−0.02	−0.87
TS-BR	−3.49	4.23	0.10	−0.22	0.00	−0.94
TS-C	−3.44	4.23	0.04	−0.28	−0.02	−0.82
Methanol complex	−4.17	4.20	−0.01	−0.15	0.00	0.00

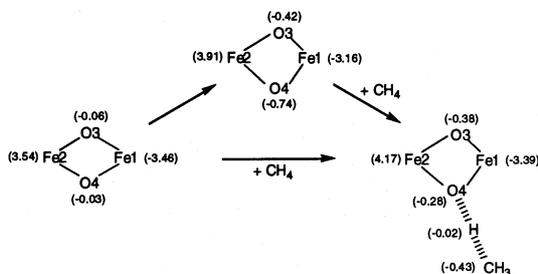


Fig. 6. Mulliken spin densities of Q, TS-1 and Q at the distorted geometry found in TS-1.

case an increase of  $\alpha$ -spin density at O4 is expected. A more intuitive picture arises when the Fe<sub>2</sub>O<sub>2</sub> core distortion and substrate attachment components of the [Q+CH<sub>4</sub>]→TS-1 step are computationally decomposed [132], by carrying out a single point calculation on the diiron moiety at the geometry found in TS-1 without the H··CH<sub>3</sub> group attached (Fig. 6). Interestingly, an  $\alpha$ -spin density of 3.91 at Fe2, which is a clear indicator for a d<sup>5</sup>-Fe(III) center, is observed even without substrate binding. Thus, the Fe<sub>2</sub>O<sub>2</sub> core distortion can be understood to assist and/or trigger an intramolecular electron-transfer mainly from O4 to Fe2, which generates a radicaloid oxygen O4 resembling a neutral O–H radical with an excess  $\beta$ -electron spin density of 0.74. In this

formal picture, the first one-electron oxidation of the substrate is consequently fueled by the intrinsic instability of a radicaloid oxygen-moiety with a formal oxidation state of –I.

After TS-1 is traversed, one possibility is for the system to follow a radical intermediate pathway (Fig. 7). Formation of the radical intermediate from the transition state is 11.5 kcal mol<sup>−1</sup> downhill with the methyl carbon moved back by 0.44 Å from the oxo bridge and the abstracted H atom forming a standard O–H bond at the new hydroxo bridge. The transition state for methanol formation is achieved by rotating the O–H group to make a 93° angle with the C–O axis. The activation energy of 3.9 kcal mol<sup>−1</sup> is too large to explain the chiral ethane results [92,103], with the time scale for rotation (~5–10 ps) being sufficient for the ethyl radical to racemize completely. The product, a methanol complex in which methanol is bound to the diiron core, forms as the CH<sub>3</sub> spontaneously moves toward the O atom of the bridging OH group. Fe–O bonds *trans* to the histidines remain essentially symmetric during methanol formation and in the methanol complex. Methanol then leaves the core in a barrierless fashion.

Alternatively, the system could follow a concerted pathway (Fig. 8) from the hydrogen atom transition

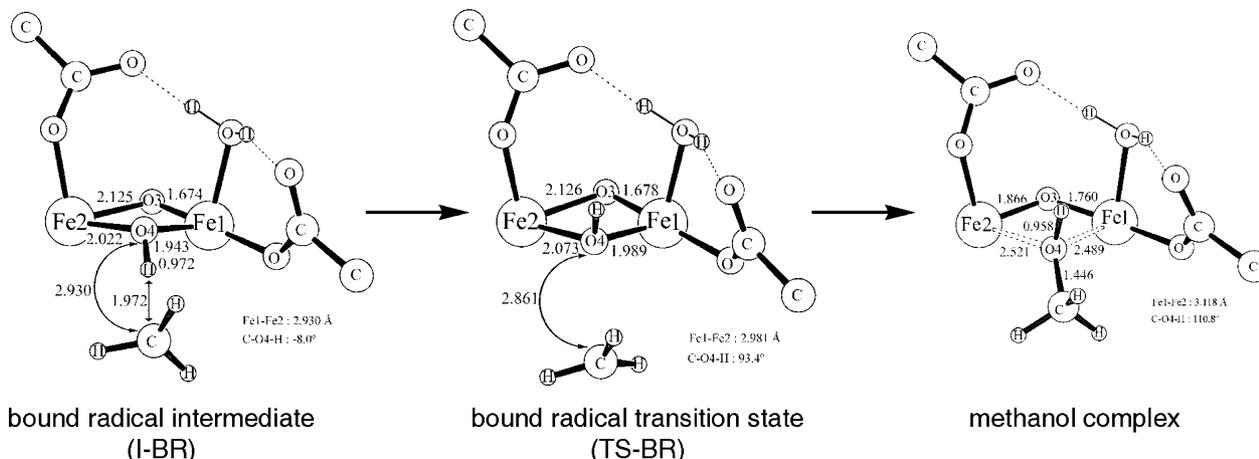


Fig. 7. Detail of the radical intermediate pathway for methanol formation. Core structures of the minimized structures are shown. Single numbers indicate distances (Å). Fe1–Fe2 distances are also noted. Single dashed lines represent hydrogen bonds.

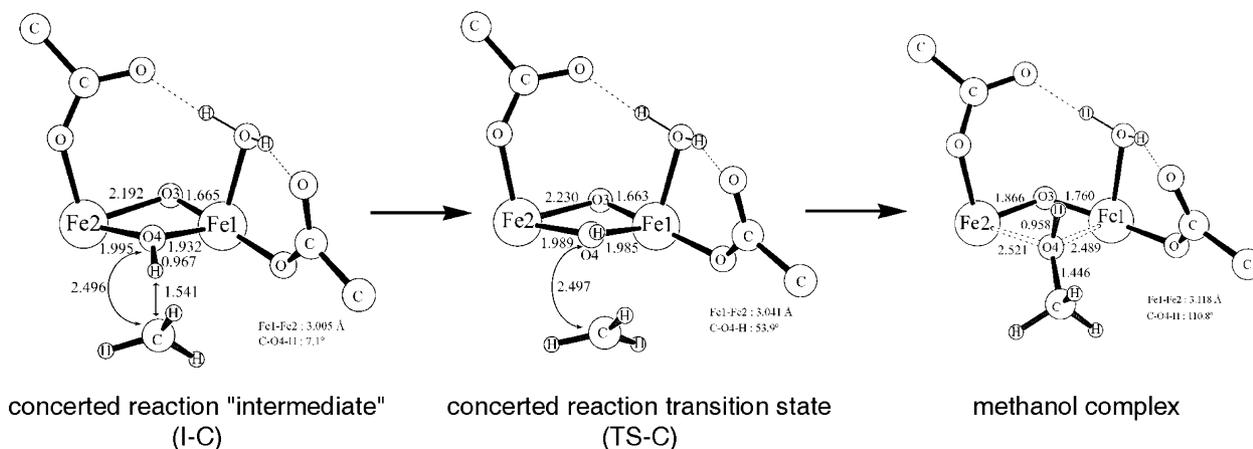


Fig. 8. Detail of the concerted reaction pathway for methanol formation. Core structures of the minimized structures are shown. Single numbers indicate distances (Å). Fe1–Fe2 distances are also noted. Single dashed lines represent hydrogen bonds.

state to product. The first step in this mechanism would be the formation of a structure with the abstracted H forming the O–H bond with minimal movement of all other atoms. This ‘intermediate’ is 3.2 kcal mol<sup>-1</sup> downhill from the transition state and does not involve recoil of the CH<sub>3</sub> group as in the radical pathway. The next step requires evolving a geometry in which the O–H group has an angle of 80–90° with the C–O axis, from which methanol formation can proceed spontaneously, as already noted. The transition state occurs when the H atom has rotated ~40° and the CH<sub>3</sub> group has rotated ~15° in the opposite direction. The barrier for this process, only 1.3 kcal mol<sup>-1</sup>, would allow substantial retention of configuration, unlike the chiral ethane experiments [92,103], and with the velocity implied by the radical clock experiments [91,104,114].

A comparison of the energetics of the two pathways for methanol formation is given in the reaction coordinate for Q reacting with methane (Fig. 9).

Mulliken spin populations for Q, the hydrogen abstraction transition state, the radical and concerted intermediates, the radical and concerted transition states, and the methanol complex are shown in Table 6. Q has two antiferromagnetically coupled high-spin ferryl Fe(IV) centers. The hydrogen abstraction transition state shows the CH<sub>3</sub> group developing radical character. The spin population on the CH<sub>3</sub> group is near unity in the radical and concerted intermediates and transition states. In the methanol complex, the spin population is zero in the methanol moiety, and the iron atoms have been reduced to two high-spin ferric Fe(III) atoms, still antiferromagnetically coupled.

These results suggest that the findings of the chiral ethane experiments [92,103] can be explained by the existence of both a radical and a concerted channel. The radical clock experiments [91,104,114] could be rationalized by the larger, heavier substrates employed therein finding the radical rebound channel less energetically and/or kinetically favorable. Dynamics calculations on the potential energy surface, described in the next section, are necessary for computation of the partitioning ratio between the two pathways.

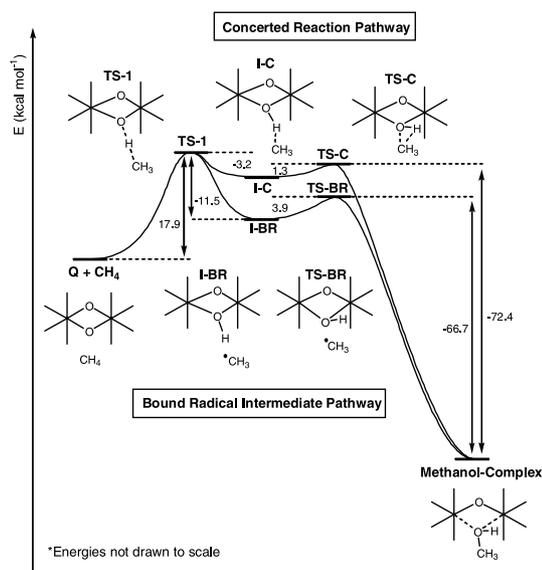


Fig. 9. Electronic reaction energy profile of methane hydroxylation promoted by Q. Both the radical-intermediate and concerted-reaction pathways are shown.

#### 4.2.6. Semiclassical dynamics of MMO alkane hydroxylation

Molecular dynamics (MD) methods that include quantum effects in the description of reaction dynamics offer a powerful approach to elucidate enzymatic reaction mechanisms at the atomic level. Recently, we [118] combined semiclassical MD with quantum chemistry calculations to provide a detailed model of the methane and ethane MMO hydroxylation reactions. Quantum effects are introduced at two levels. First, quantum initial conditions are simulated using the

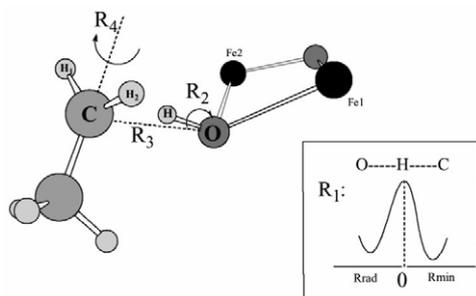


Fig. 10. Reaction coordinates for the reduced potential surface employed to model the ethane hydroxylation dynamics.

Wigner function for the Boltzmannized operator with the harmonic normal mode approximation. Secondly, tunneling corrections were improved through a WKB treatment of sudden proton transfer at classical turning points. Specific details for both approximations can be followed in detail in Ref. [118].

The ethane hydroxylation dynamics were simulated by reducing the dimensions of the electronic potential energy surface (PES) as shown in Fig. 10.  $R_1$  is the O–H/C–H distance before/after the proton coupled first electron transfer transition state,  $R_2$  is the C··H–O angle,  $R_3$  is the C··O distance, and  $R_4$  accounts for the C–C bond rotation that could exchange the H isotopes in a chiral ethyl radical and afford racemized products. The PES is obtained through optimization of all the remaining degrees of structural freedom for several representative configurations of  $R_1$ – $R_4$ . Our model of the active site of the enzyme (Fig. 2) contains a total of 105 atoms with ethane being the substrate.

The main energetic parameters in the ethane hydroxylation involve the energy barrier for proton coupled electron transfer along the  $R_1$  coordinate,  $12.5 \text{ kcal mol}^{-1}$ , and the rotational barriers along  $R_2$  and  $R_4$ ,  $3.90$  and  $4.06 \text{ kcal mol}^{-1}$ , respectively. These energy parameters are strongly coupled with  $R_3$ , the C–O distance. In order to measure the different mechanisms of the hydroxylation reaction and the amount of retention of configuration, we sampled a trajectory flux following the hydrogen atom transfer. The dynamics and trajectory analysis describes 16% of the reaction following a concerted pathway, where the system has enough energy, or can rapidly redistribute the energy, to rotate immediately after the transition state in the hydrogen atom transfer coordinate. A close examination of the trajectories following the concerted path, indicates that the system spends 20–40 fs in the transition state region until there is an energy transfer from  $R_1$  and  $R_4$  leading to a rotation in  $R_2$ . All the concerted trajectories retain the initial configuration since there is not enough time or energy to rotate along  $R_4$ . The fraction of reactants following the concerted channel increases by almost 1/3 when tunneling contributions are considered. This large quantum effect agrees with the experimentally measured

kinetic isotope effect; for example,  $k_{\text{H}}/k_{\text{D}}$  is 3.8 for hydroxylation of R–CH<sub>3</sub>CHTD [92,103].

The remaining 84% of the trajectories follow a bound radical recoil/rebound pathway, where several vibrational periods in  $R_3$  are required until the energy is effectively redistributed to  $R_2$ . This periodic motion along  $R_3$  involves mainly substrate diffusion in the active site. The radical intermediate exhibits a strongly bonded character, where the C–O distance is lower than  $3.0 \text{ \AA}$  for half of the periodic motion along  $R_3$ . The mean lifetime of the radical intermediate derived from our calculations is 300 fs. One might therefore explain the failure to observe radical-derived products. Integrating  $R_4$  along the time axis for the recoil/rebound pathway, we compute 69–84% retention of configuration, depending upon the initial energy in  $R_4$ , in good agreement with the observed experimental value of 72% for *R*-chiral ethane.

### 4.3. Hemerythrin

Hemerythrin (Hr) is a metalloprotein responsible for dioxygen transport in several marine invertebrates [136]. X-ray diffraction studies on both the deoxygenated (deoxyHr) and the oxygenated (oxyHr) forms [137] identified a dinuclear iron center at the active site. In deoxyHr, the two Fe(II) atoms [138,139] adopt five- and six-coordinate geometries having approximately trigonal bipyramidal and octahedral local symmetry, respectively, and are connected by two carboxylate groups from aspartate (Asp106) and glutamate (Glu58) moieties in addition to a water-derived hydroxo bridge [137,140–142]. The remaining coordination sites are occupied by five terminally bound histidine ligands (Fig. 11). Binding of dioxygen is accompanied by a two-electron redox process affording a Fe(III)/Fe(III) state and a hydroperoxide that is coordinated terminally to the initially 5-coordinate iron, causing a change of local structure from the trigonal bipyramidal motif to an octahedral arrangement [143]. Changes in bond distances [137] and magnetic properties [144] together with spectroscopic evidence for the presence of a hydroperoxo group in oxyHr [145] led to the proposal that the electron transfer from the metal centers to the incoming dioxygen is coupled to proton-transfer from the bridging hydroxide to the newly formed peroxide. The overall reaction for Hr is illustrated in Fig. 11. The reversibility of dioxygen binding requires that the deoxy

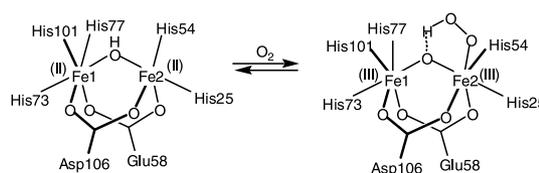


Fig. 11. Reversible O<sub>2</sub> binding to hemerythrin.

and the oxy forms of Hr in solution are relatively close in free energy. The experimental equilibrium constant of  $2.5 \times 10^5 \text{ M}^{-1}$  [136], measured under standard temperature and pressure conditions, implies that oxyHr is lower in free energy by  $7.3 \text{ kcal mol}^{-1}$ . Our initial objective has been to reproduce this free energy difference quantitatively via the use of QM and QM/MM calculations [135].

#### 4.3.1. Computational models

For the isolated QM model of Hr, the terminal histidine ligands were replaced by imidazole groups and the bridging aspartate and glutamate ligands by acetate moieties, resulting in a model cluster with a net charge of +1. For the QM/MM model, the cuts were made at side-chains of the first sphere amino acid ligands. One entire subunit of the protein was included in the model, giving a 2007-atom model for oxyHr. The active site of Hr is buried inside the protein cavity and hence isolated from bulk solvent; in this situation, the overall core moiety must be electrically neutral. This neutrality is achieved in the present system by the presence of the ionized, anionic carboxylate of Glu24, which forms a strong hydrogen bond with a core histidine. It neutralizes the formally positive charge assigned to the diiron core unit. Except for Glu24 and amino acids on the surface of the proteins forming salt bridges, all remaining ionizable residues in the MM region were neutralized. As discussed above, this neutralization in effect accounts for the screening imposed by the high dielectric solvent. Crystallographic water molecules were included in the model but were shown to have a very small effect on  $\text{O}_2$  binding.

Prior to applying QM/MM to hemerythrin, the accuracy of the QM/MM interface for the three types of amino acids (His, Glu, Asp) present in the first coordination sphere was tested in benchmark calculations on small model systems. In the first set of tests, the amino acid for which the test was carried out was included in full, including models of their nearest residual neighbors, while the remaining amino acid residues were replaced by ligands of minimum size ( $\text{NH}_3$  for His, and  $\text{COO}^-$  for Glu and Asp). Using such a model, both full QM calculations and mixed QM/MM single-point energy calculations were performed and the  $\text{O}_2$  binding energy and spin populations obtained with the two methods were compared. Whereas the QM and the QM/MM dioxygen binding differed by only  $1 \text{ kcal mol}^{-1}$  for the His and Glu amino acids, a somewhat larger deviation of  $2 \text{ kcal mol}^{-1}$  was obtained for the small Asp residue. The spin populations on iron, both for the oxy and the deoxy forms, are also very similar in the two methods, within 0.02 units for all three amino acids investigated. In these test calculations, only one cut for each amino acid was included, while there are on average 3.5 amino acids binding to each of

the two iron atoms in the active site of Hr. In a second set of tests, the accuracy for models that include multiple cuts was investigated by comparing QM and QM/MM results for a mononuclear iron center coordinating up to three full histidine groups. For these models, the  $\text{Fe(II)} \rightarrow \text{Fe(III)}$  oxidation by hydrogen abstraction from coordinating water molecules was investigated. It was found that the presence of three cuts resulted in a deviation of  $0.8 \text{ kcal mol}^{-1}$  relative to an extrapolated QM model and the presence of two cuts resulted in a deviation of  $0.2 \text{ kcal mol}^{-1}$ . The errors are not systematic and a larger error ( $2.0 \text{ kcal mol}^{-1}$ ) was actually obtained for the model containing one cut. Due to the computational cost of using full QM models, the above tests were made for single-point energies at the structure found in  $2.0 \text{ \AA}$ -resolution X-ray diffraction studies. This limitation means that the electronic structure of the QM region is not in a minimum energy configuration, which complicates the description of the interaction between the two regions. The final test was made for a realistic Hr model for which one entire subunit of the protein was included and full geometry optimizations were performed. In this test, one of the seven side-chain cuts was replaced by a backbone cut, located further away from the metal center. Replacing a side-chain with a backbone cut and reoptimizing the structures had an effect of only  $0.04 \text{ kcal mol}^{-1}$  on the  $\text{O}_2$  binding energy, and a maximum effect of 0.09 on the iron spin population. Taken together, the tests indicate that the deviation between a full QM and a mixed QM/MM treatment is within  $2 \text{ kcal mol}^{-1}$  for  $\text{O}_2$  binding in hemerythrin. There is no indication that a QM/MM model containing several cuts is less accurate than a model containing only one cut. For the only test where a realistic protein model was used and a full geometry optimization was performed, the deviation between a model with a cut in the backbone and that with a cut in the side-chain is less than  $0.1 \text{ kcal mol}^{-1}$  for  $\text{O}_2$  binding.

#### 4.3.2. Reversible binding of dioxygen

Purely quantum mechanical calculations of a small model of the active site including the ligands of the first coordination shell around the diiron core give an electronic dioxygen binding enthalpy of  $-5.3 \text{ kcal mol}^{-1}$ . The reaction, however, is entropically unfavorable mostly owing to loss of entropy associated with dioxygen. Consequently, the addition of translational, rotational and vibrational entropy corrections shifts the equilibrium in favor of the deoxy form by  $13.8 \text{ kcal mol}^{-1}$ . Adding these two terms together,  $\text{O}_2$  is unbound by  $8.5 \text{ kcal mol}^{-1}$  of free energy, disagreeing by  $15 \text{ kcal mol}^{-1}$  with the experimental estimate of  $-7.3 \text{ kcal mol}^{-1}$  for the reaction free energy [136]. Naturally, the question arises whether or not consideration of the protein environment will improve the theory–experiment correlation. In principle, there are three main

factors that will constitute the effect of the protein. Firstly, partial charges on the peptide residues will interact electrostatically with the diiron unit and also cause a potentially notable redistribution of the electron density at the diiron core. The second, straightforward effect gives rise to what is widely recognized as ‘strain energy’ and is due to steric constraints imposed by the structure of the protein on the ligand moieties directly connected to the metal center affording an otherwise non-equilibrium ligand arrangement. Lastly, other non-covalent interactions, most importantly driven by van der Waals forces, are expected to afford a substantial difference when the small quantum mechanical model is compared to the full peptidyl system. Quantifying these effects in a consistent manner is one of the most attractive features of the QM/MM approach.

In the case of Hr, the electrostatic component is anticipated to give preference to oxyHr, where a higher degree of charge separation is expected due to charge transfer from the iron centers to the dioxygen fragments. Explicit evaluation of Coulombic interactions between the peptidyl domain (MM-region) and the electron density of the QM-region revealed a differential effect of  $3.2 \text{ kcal mol}^{-1}$  in favor of oxyHr, where the most dominant interaction is between the carboxylate fragment of Glu24 and the diiron core, which becomes oxidized in the course of the reaction. Intuitively, a substantial strain effect would be expected given that the first shell coordination sphere, especially around Fe2, reorganizes significantly. The strain energy can be computed explicitly by carrying out a single point calculation on a small QM model using a structure directly derived from the QM/MM equilibrium geometry and comparing it to the fully relaxed structure of the purely QM model calculations. For this system, the electronic structure of the iron-dimer including the first coordination ligands is highly flexible and the strain energy is estimated to be less than  $1 \text{ kcal mol}^{-1}$  although the orientation of the imidazole moieties are notably different in the QM and QM/MM models [135]. The largest effect of the peptidyl environment arises from van der Waals interactions between the bound dioxygen and the hydrophobic pocket around the diiron core, which amounts to a  $6.3 \text{ kcal mol}^{-1}$  preference towards oxyHr. Lastly, a correction term needs to be introduced for the solvation of dioxygen in water. Using Henry’s law the free energy of solvation of dioxygen at room temperature can be estimated to be  $+3.9 \text{ kcal mol}^{-1}$ . With these effects taken into account, the overall free energy of the dioxygen binding is computed to be  $-5.2 \text{ kcal mol}^{-1}$ , which is in remarkably good agreement with the estimate of  $-7.3 \text{ kcal mol}^{-1}$  obtained from the experimentally measured equilibrium constant [146].

The kinetics of reversible dioxygen binding, i.e. the exact mechanism of dioxygen binding to the diiron core,

is highly complicated owing to the unpaired electrons on the iron centers that couple antiferromagnetically to each other. Thus, the excess electrons on each of the  $d^6$ -Fe(II) centers that transfer to the incoming dioxygen, presumably in stepwise fashion, have opposite spins. However, triplet dioxygen can only accept two electrons that have the same spin, which inevitably generates a situation that can be best described as a ‘spin-forbidden’ redox process. Brunold and Solomon [147] proposed to overcome this intrinsic problem by initial attack of dioxygen at the coordinatively unsaturated iron center that is accompanied by transfer of the first electron. Then a superexchange mechanism assisted by strong coupling due to delocalization of the redox active orbitals between the two iron centers via the oxo-bridge was proposed to flip the spin on the superoxo moiety and trigger the second electron transfer from the remote iron site to form the peroxo ligand. Currently, we are exploring the kinetics of this reaction using both the large QM model and the QM/MM model. The latter is particularly challenging since artificial electronic effects due to the QM-MM boundaries, in our approach the frozen orbital that connects the QM and MM domains, have to be carefully separated from physically meaningful electronic distortions. Note that this is an intrinsic problem of any QM/MM approach and one of the pressing open questions for the QM/MM technology.

Preliminary explorations of the reaction energy profile point to a slightly different mechanism from what has been suggested by Brunold and Solomon and involves initial attack of dioxygen at the bridging hydroxo-group accompanied by an outer sphere electron transfer from the coordinatively saturated iron site giving a superoxo moiety, followed by double-exchange with the resulting mixed-valence diiron complex to enable the second electron transfer. Work is in progress to obtain a more quantitative picture for the kinetics of dioxygen binding in the large scale QM and QM/MM framework.

## 5. Conclusions

The results described above reflect the considerable progress that we, and other groups, have made over the past several years in modeling metalloprotein active sites. For both experimental systems considered, qualitative agreement with experimental data has been quite good, and quantitative agreement has been demonstrated in several cases, such as the computation of the free energy differential between the oxy and deoxy forms of Hr. Although improvements in methodology are still an important direction to pursue, it is apparent that the current computational tools — particularly the use of QM/MM methods in conjunction with QM methods — can provide useful information for a range of

metalloproteins. Widespread adoption of these methods, and deployment across a much broader range of enzymes, would yield a qualitative advance in our characterization and understanding of how metalloproteins function, and the role of the protein in controlling structures and energetics of the key intermediates and transition states required for biological function. Continuing reductions in computational cost/performance ratio will systematically lower the resources required to carry out such studies; even now, calculations on QM/MM models containing 100–150 atoms can routinely be performed using a two-processor PC, costing a few thousand dollars.

With regard to control of dioxygen chemistry by metalloenzymes, some significant conclusions can be drawn already from the studies that we have carried out to date. In the case of Hr, oxygenation chemistry does not occur because the O–O bond is not cleaved and no Fe(IV) species is involved. In this case, the key role of the protein is to stabilize the oxy form versus the deoxy form, tuned to the point where dioxygen can appropriately be bound and then released in the relevant environments. The principal features enabling this stabilization are the hydrogen bonding of the peroxy species, electrostatic stabilization through Glu24 of the electron deficient imidazole in the oxy form, and van der Waals interactions between the protein and dioxygen/peroxy moiety.

The catalytic reaction pathways of both the dioxygen and methane activation cycles in MMO are significantly influenced by hydrogen bonding structures formed by the first shell ligands around the diiron centers. The hydrogen bonding patterns are altered by two mechanisms: (1) carboxylate shifts exhibited by Glu243; and (2) addition of dioxygen and transformation of the position of the oxygen atoms as the various intermediates are produced. The rate determining step in the hydroxylation reaction involving the oxidatively potent Fe(IV) centers manifests a classical proton-coupled electron transfer transition state in which the hydrogen is roughly halfway between the carbon to which it was originally attached.

There is still much to be learned about both systems, and a great deal of work to do before theory and experiment can truly be said to have been brought into concordance at the level of detail required for a truly reliable atomic level picture. One important area is computation of spectroscopic observables, such as optical spectra and Mössbauer spectra [148], for a given theoretical structure, thus allowing a more convincing identification of experimental species detected by these techniques. A second is characterizing kinetic processes at the same level of detail as intermediates, and directly computing rate constants, kinetic isotope effects, and other experimentally observable dynamical quantities; some progress along these lines has been made for

MMO, but even for that system there is a lot more to accomplish. Ultimately, one would like to make predictions using the theoretical models that can be tested experimentally, for example in site directed mutagenesis experiments. The ability to make predictions that can be confirmed by experimental data will be the best signal that the equal partnership between theory and experiment envisioned above is within reach.

## Acknowledgements

This work was partially supported by grants from the NIH to R.A.F. (GM40526) and S.J.L. (GM32134), and by the Environmental Molecular Sciences Institute at Columbia University, supported by the NSF and DOE. Computational support from the NSF NPACI program, and from Pacific Northwest National Laboratories, is also acknowledged.

## References

- [1] A. Messerschmidt, R. Huber, T. Poulos, K. Wieghardt, *Handbook of Metalloproteins*, Wiley, Chichester, 2001.
- [2] S.J. Lippard, J. Berg, *Principles of Bioinorganic Chemistry*, University Science Books, Mill Valley, CA, 1994.
- [3] T.R. Cundari, *J. Chem. Soc. Dalton Trans.* (1998) 2771.
- [4] P. Comba, M. Zimmer, *J. Chem. Educ.* 73 (1996) 108.
- [5] R.G. Parr, W. Yang, *Density Functional Theory of Atoms and Molecules*, Oxford University Press, New York, 1989.
- [6] T. Ziegler, *Chem. Rev.* 91 (1991) 651.
- [7] G. Vacek, J.K. Perry, J.M. Langlois, *Chem. Phys. Lett.* 310 (1999) 189.
- [8] R.E. Christoffersen, D.W. Genson, G.M. Maggiora, *J. Chem. Phys.* 54 (1971) 239.
- [9] R.E. Christoffersen, G.M. Maggiora, *Chem. Phys. Lett.* 3 (1969) 419.
- [10] A. Warshel, M. Levitt, *J. Mol. Biol.* 103 (1976) 227.
- [11] B. Marten, K. Kim, C. Cortis, R.A. Friesner, R.B. Murphy, M.N. Ringnalda, D. Sitkoff, B. Honig, *J. Phys. Chem.* 100 (1996) 11775.
- [12] D.J. Tannor, B. Marten, R.B. Murphy, R.A. Friesner, D. Sitkoff, A. Nicholls, M.N. Ringnalda, W.A.I. Goddard, B. Honig, *J. Am. Chem. Soc.* 116 (1994) 11875.
- [13] A.D. Becke, *J. Chem. Phys.* 98 (1993) 5648.
- [14] J.P. Perdew, Y. Wang, *Phys. Rev. B* 45 (1992) 13244.
- [15] C.T. Lee, W.T. Yang, R.G. Parr, *Phys. Rev. B* 37 (1988) 785.
- [16] L.A. Curtiss, K. Raghavachari, P.C. Redfern, J.A. Pople, *J. Chem. Phys.* 106 (1997) 1063.
- [17] J.C. Slater, *Quantum Theory of Molecules and Solids: the Self-Consistent Field for Molecules and Solids*, vol. 4, McGraw-Hill, New York, 1974.
- [18] A.D. Becke, *Phys. Rev. A* 38 (1988) 3098.
- [19] S.H. Vosko, L. Wilk, M. Nusair, *Can. J. Phys.* 58 (1980) 1200.
- [20] L.A. Curtiss, P.C. Redfern, K. Raghavachari, J.A. Pople, *J. Chem. Phys.* 109 (1998) 42.
- [21] G.S. Tschumper, H.F. Schaefer, *J. Chem. Phys.* 107 (1997) 2529.
- [22] F. Deproft, P. Geerlings, *J. Chem. Phys.* 106 (1997) 3270.
- [23] B.G. Johnson, P.M.W. Gill, J.A. Pople, *J. Chem. Phys.* 98 (1993) 5612.

- [24] P.M.W. Gill, B.G. Johnson, J.A. Pople, M.J. Frisch, *Int. J. Quantum Chem. Quantum Chem. Symp.* 26 (1992) 319.
- [25] B.G. Johnson, P.M.W. Gill, J.A. Pople, *J. Chem. Phys.* 97 (1992) 7846.
- [26] P.M.W. Gill, B.G. Johnson, J.A. Pople, M.J. Frisch, *Chem. Phys. Lett.* 197 (1992) 499.
- [27] M.R.A. Blomberg, P.E.M. Siegbahn, M. Svensson, *J. Chem. Phys.* 104 (1996) 9546.
- [28] J.L.C. Thomas, C.W. Bauschlicher, M.B. Hall, *J. Phys. Chem. A* 101 (1997) 8530.
- [29] A. Ricca, C.W. Bauschlicher, *J. Phys. Chem. A* 101 (1997) 8949.
- [30] M. Sodupe, V. Branchadell, M. Rosi, C.W. Bauschlicher, *J. Phys. Chem. A* 101 (1997) 7854.
- [31] M. Diedenhofen, T. Wagener, G. Frenking, in: T.R. Cundari (Ed.), *Computational Organometallic Chemistry*, Marcel Dekker, New York, 2001.
- [32] P.E.M. Siegbahn, M.R.A. Blomberg, *Annu. Rev. Phys. Chem.* 50 (1999) 221.
- [33] Y. Nakao, K. Hirao, T. Taketsugu, *J. Chem. Phys.* 114 (2001) 5216.
- [34] Y. Nakao, K. Hirao, T. Taketsugu, *J. Chem. Phys.* 114 (2001) 7935.
- [35] P.E.M. Siegbahn, M.R.A. Blomberg, *Chem. Rev.* 100 (2000) 421.
- [36] M.N. Glukhovtsev, R.D. Bach, C.J. Nagel, *J. Phys. Chem. A* 101 (1997) 316.
- [37] L. Noodleman, *J. Chem. Phys.* 74 (1981) 5737.
- [38] P.J. Hay, W.R. Wadt, *J. Chem. Phys.* 82 (1985) 270.
- [39] W.R. Wadt, P.J. Hay, *J. Chem. Phys.* 82 (1985) 284.
- [40] P.J. Hay, W.R. Wadt, *J. Chem. Phys.* 82 (1985) 299.
- [41] JAGUAR 4.1, Schrödinger Inc., Portland, OR, 2000.
- [42] T.H. Dunning, *J. Chem. Phys.* 90 (1989) 1007.
- [43] F. Maseras, K. Morokuma, *J. Comput. Chem.* 16 (1995) 1170.
- [44] V. Luzhkov, A. Warshel, *J. Comput. Chem.* 13 (1992) 199.
- [45] P.A. Bash, M.J. Field, M. Karplus, *J. Am. Chem. Soc.* 109 (1987) 8092.
- [46] U.C. Singh, P.A. Kollman, *J. Comput. Chem.* 7 (1986) 718.
- [47] T.K. Woo, L. Cavallo, T. Ziegler, *Theor. Chem. Acc.* 100 (1998) 307.
- [48] T.K. Woo, P.M. Margl, P.E. Blöchl, T. Ziegler, *J. Phys. Chem. B* 101 (1997) 7877.
- [49] J.L. Gao, P. Amara, C. Alhambra, M.J. Field, *J. Phys. Chem. A* 102 (1998) 4714.
- [50] N. Ferre, X. Assfeld, J.L. Rivail, *J. Comput. Chem.* 23 (2002) 610.
- [51] Y.K. Zhang, T.S. Lee, W.T. Yang, *J. Chem. Phys.* 110 (1999) 46.
- [52] G. Monard, M. Loos, V. Thery, K. Baka, J.L. Rivail, *Int. J. Quantum Chem.* 58 (1996) 153.
- [53] V. Thery, D. Rinaldi, J.L. Rivail, B. Maigret, G.G. Ferenczy, *J. Comput. Chem.* 15 (1994) 269.
- [54] R.B. Murphy, D.M. Philipp, R.A. Friesner, *J. Comput. Chem.* 21 (2000) 1442.
- [55] R.B. Murphy, D.M. Philipp, R.A. Friesner, *Chem. Phys. Lett.* 321 (2000) 113.
- [56] D.M. Philipp, R.A. Friesner, *J. Comput. Chem.* 20 (1999) 1468.
- [57] W.L. Jorgensen, D.S. Maxwell, J. Tiradorives, *J. Am. Chem. Soc.* 118 (1996) 11225.
- [58] QSITE, Schrödinger Inc., Portland, OR, 2001.
- [59] IMPACT 7.2, Schrödinger Inc., Portland, OR, 2001.
- [60] A.J. Bard, R. Parsons, J. Jordan, *Standard Potentials in Aqueous Solution*, Marcel Dekker, New York, 1985.
- [61] H. Reiss, A. Heller, *J. Phys. Chem.* 89 (1985) 4207.
- [62] M.-H. Baik, J.S. Silverman, I.V. Yang, P.A. Ropp, V.A. Szalai, W.T. Yang, H.H. Thorp, *J. Phys. Chem. B* 105 (2001) 6437.
- [63] M.-H. Baik, R.A. Friesner, *J. Phys. Chem. A* 106 (2002) 7407.
- [64] K.E. Liu, S.J. Lippard, *Adv. Inorg. Chem.* 42 (1995) 263.
- [65] M. Merckx, D.A. Kopp, M.H. Sazinsky, J.L. Blazyk, J. Muller, S.J. Lippard, *Angew. Chem. Int. Ed. Engl.* 40 (2001) 2782.
- [66] A.L. Feig, S.J. Lippard, *Chem. Rev.* 94 (1994) 759.
- [67] B.J. Wallar, J.D. Lipscomb, *Chem. Rev.* 96 (1996) 2625.
- [68] R.A. Periana, D.J. Taube, E.R. Evitt, D.G. Loffler, P.R. Wentrcck, G. Voss, T. Masuda, *Science* 259 (1993) 340.
- [69] A.E. Shilov, G.B. Shul'pin, *Chem. Rev.* 97 (1997) 2879.
- [70] J. Colby, D.I. Stirling, H. Dalton, *Biochem. J.* 165 (1977) 395.
- [71] M.J. Rataj, J.E. Kauth, M.I. Donnelly, *J. Biol. Chem.* 266 (1991) 18684.
- [72] K.K. Andersson, W.A. Froland, S.K. Lee, J.D. Lipscomb, *New J. Chem.* 15 (1991) 411.
- [73] B.G. Fox, J.G. Borneman, L.P. Wackett, J.D. Lipscomb, *Biochemistry* 29 (1990) 6419.
- [74] J. Green, H. Dalton, *J. Biol. Chem.* 264 (1989) 17698.
- [75] H. Dalton, *Adv. Appl. Microbiol.* 26 (1980) 71.
- [76] I.J. Higgins, D.J. Best, R.C. Hammond, *Nature* 286 (1980) 561.
- [77] A.C. Rosenzweig, P. Nordlund, P.M. Takahara, C.A. Frederick, S.J. Lippard, *Chem. Biol.* 2 (1995) 409.
- [78] A.C. Rosenzweig, C.A. Frederick, S.J. Lippard, P. Nordlund, *Nature* 366 (1993) 537.
- [79] A.C. Rosenzweig, S.J. Lippard, *Acc. Chem. Res.* 27 (1994) 229.
- [80] D.A. Whittington, S.J. Lippard, *J. Am. Chem. Soc.* 123 (2001) 827.
- [81] J.G. Dewitt, J.G. Bentsen, A.C. Rosenzweig, B. Hedman, J. Green, S. Pilkington, G.C. Papaefthymiou, H. Dalton, K.O. Hodgson, S.J. Lippard, *J. Am. Chem. Soc.* 113 (1991) 9219.
- [82] B.G. Fox, K.K. Surerus, E. Münck, J.D. Lipscomb, *J. Biol. Chem.* 263 (1988) 10553.
- [83] N. Elango, R. Radhakrishnan, W.A. Froland, B.J. Wallar, C.A. Earhart, J.D. Lipscomb, D.H. Ohlendorf, *Protein Sci.* 6 (1997) 556.
- [84] K.E. Liu, D.L. Wang, B.H. Huynh, D.E. Edmondson, A. Salifoglou, S.J. Lippard, *J. Am. Chem. Soc.* 116 (1994) 7465.
- [85] K.E. Liu, A.M. Valentine, D.L. Wang, B.H. Huynh, D.E. Edmondson, A. Salifoglou, S.J. Lippard, *J. Am. Chem. Soc.* 117 (1995) 10174.
- [86] S.K. Lee, J.C. Nesheim, J.D. Lipscomb, *J. Biol. Chem.* 268 (1993) 21569.
- [87] L.J. Shu, J.C. Nesheim, K. Kauffmann, E. Münck, J.D. Lipscomb, L. Que, *Science* 275 (1997) 515.
- [88] A.M. Valentine, S.S. Stahl, S.J. Lippard, *J. Am. Chem. Soc.* 121 (1999) 3876.
- [89] J.D. Lipscomb, L. Que, *J. Biol. Inorg. Chem.* 3 (1998) 331.
- [90] D.A. Whittington, M.H. Sazinsky, S.J. Lippard, *J. Am. Chem. Soc.* 123 (2001) 1794.
- [91] K.E. Liu, C.C. Johnson, M. Newcomb, S.J. Lippard, *J. Am. Chem. Soc.* 115 (1993) 939.
- [92] A.M. Valentine, B. Wilkinson, K.E. Liu, S. Komar-Panicucci, N.D. Priestley, P.G. Williams, H. Morimoto, H.G. Floss, S.J. Lippard, *J. Am. Chem. Soc.* 119 (1997) 1818.
- [93] D.A. Whittington, A.M. Valentine, S.J. Lippard, *J. Biol. Inorg. Chem.* 3 (1998) 307.
- [94] M.P. Woodland, D.S. Patil, R. Cammack, H. Dalton, *Biochim. Biophys. Acta* 873 (1986) 237.
- [95] J. Lund, M.P. Woodland, H. Dalton, *Eur. J. Biochem.* 147 (1985) 297.
- [96] M.P. Woodland, H. Dalton, *J. Biol. Chem.* 259 (1984) 53.
- [97] D.A. Kopp, G.T. Gassner, J.L. Blazyk, S.J. Lippard, *Biochemistry* 40 (2001) 14932.
- [98] N. Deighton, I.D. Podmore, M.C.R. Symons, P.C. Wilkins, H. Dalton, *J. Chem. Soc. Chem. Commun.* (1991) 1086.
- [99] S.K. Smoukov, D.A. Kopp, A.M. Valentine, R. Davydov, S.J. Lippard, B.M. Hoffman, *J. Am. Chem. Soc.* 124 (2002) 2657.
- [100] S.K. Lee, B.G. Fox, W.A. Froland, J.D. Lipscomb, E. Münck, *J. Am. Chem. Soc.* 115 (1993) 6450.

- [101] B.G. Fox, W.A. Froland, J.E. Dege, J.D. Lipscomb, *J. Biol. Chem.* 264 (1989) 10023.
- [102] B.J. Brazeau, J.D. Lipscomb, *Biochemistry* 39 (2000) 13503.
- [103] N.D. Priestley, H.G. Floss, W.A. Froland, J.D. Lipscomb, P.G. Williams, H. Morimoto, *J. Am. Chem. Soc.* 114 (1992) 7561.
- [104] F. Ruzicka, D.S. Huang, M.I. Donnelly, P.A. Frey, *Biochemistry* 29 (1990) 1696.
- [105] J.C. Nesheim, J.D. Lipscomb, *Biochemistry* 35 (1996) 10240.
- [106] Y. Jin, J.D. Lipscomb, *Biochemistry* 38 (1999) 6178.
- [107] S.L. Chang, B.J. Wallar, J.D. Lipscomb, K.H. Mayo, *Biochemistry* 40 (2001) 9539.
- [108] Y. Liu, J.C. Nesheim, K.E. Paulsen, M.T. Stankovich, J.D. Lipscomb, *Biochemistry* 36 (1997) 5223.
- [109] K.E. Paulsen, Y. Liu, B.G. Fox, J.D. Lipscomb, E. Münck, M.T. Stankovich, *Biochemistry* 33 (1994) 713.
- [110] B.G. Fox, M.P. Hendrich, K.K. Surerus, K.K. Andersson, W.A. Froland, J.D. Lipscomb, E. Münck, *J. Am. Chem. Soc.* 115 (1993) 3688.
- [111] M.P. Hendrich, E. Münck, B.G. Fox, J.D. Lipscomb, *J. Am. Chem. Soc.* 112 (1990) 5861.
- [112] S.K. Lee, J.D. Lipscomb, *Biochemistry* 38 (1999) 4423.
- [113] S.S. Stahl, W.A. Francisco, M. Merckx, J.P. Klinman, S.J. Lippard, *J. Biol. Chem.* 276 (2001) 4549.
- [114] S.Y. Choi, P.E. Eaton, D.A. Kopp, S.J. Lippard, M. Newcomb, R.N. Shen, *J. Am. Chem. Soc.* 121 (1999) 12198.
- [115] R.A. Friesner, B.D. Dunietz, *Acc. Chem. Res.* 34 (2001) 351.
- [116] B.D. Dunietz, M.D. Beachy, Y.X. Cao, D.A. Whittington, S.J. Lippard, R.A. Friesner, *J. Am. Chem. Soc.* 122 (2000) 2828.
- [117] B.F. Gherman, B.D. Dunietz, D.A. Whittington, S.J. Lippard, R.A. Friesner, *J. Am. Chem. Soc.* 123 (2001) 3836.
- [118] V. Guallar, B.F. Gherman, W.H. Miller, J. Lippard Stephen, A. Friesner Richard, *J. Am. Chem. Soc.* 124 (2002) 3377.
- [119] M. Torrent, D.G. Musaev, K. Morokuma, *J. Phys. Chem. B* 105 (2001) 322.
- [120] H. Basch, K. Mogi, D.G. Musaev, K. Morokuma, *J. Am. Chem. Soc.* 121 (1999) 7249.
- [121] P.E.M. Siegbahn, *Chem. Phys. Lett.* 351 (2002) 311.
- [122] P.E.M. Siegbahn, *Inorg. Chem.* 38 (1999) 2880.
- [123] P.E.M. Siegbahn, R.H. Crabtree, P. Nordlund, *J. Biol. Inorg. Chem.* 3 (1998) 314.
- [124] K. Yoshizawa, A. Suzuki, Y. Shiota, T. Yamabe, *Bull. Chem. Soc. Jpn.* 73 (2000) 815.
- [125] K. Yoshizawa, *J. Biol. Inorg. Chem.* 3 (1998) 318.
- [126] K. Yoshizawa, Y. Shiota, T. Yamabe, *J. Am. Chem. Soc.* 120 (1998) 564.
- [127] K. Yoshizawa, T. Ohta, T. Yamabe, R. Hoffmann, *J. Am. Chem. Soc.* 119 (1997) 12311.
- [128] K. Yoshizawa, Y. Yokomichi, Y. Shiota, T. Ohta, T. Yamabe, *Chem. Lett.* (1997) 587.
- [129] K. Yoshizawa, T. Yamabe, R. Hoffmann, *New J. Chem.* 21 (1997) 151.
- [130] K. Yoshizawa, R. Hoffmann, *Inorg. Chem.* 35 (1996) 2409.
- [131] M.-H. Baik, D. Lee, R.A. Friesner, S.J. Lippard, *Isr. J. Chem.* 41 (2001) 173.
- [132] M.-H. Baik, B.F. Gherman, R.F. Friesner, S.J. Lippard, *J. Am. Chem. Soc.*, in press.
- [133] P.E.M. Siegbahn, *J. Biol. Inorg. Chem.* 6 (2001) 27.
- [134] K. Yoshizawa, *J. Organomet. Chem.* 635 (2001) 100.
- [135] M. Wirstam, S.J. Lippard, R.A. Friesner, *J. Am. Chem. Soc.*, submitted for publication.
- [136] R.E. Stenkamp, *Chem. Rev.* 94 (1994) 715.
- [137] M.A. Holmes, I. Letrong, S. Turley, L.C. Sieker, R.E. Stenkamp, *J. Mol. Biol.* 218 (1991) 583.
- [138] J.L. York, A.J. Bearden, *Biochemistry* 9 (1970) 4549.
- [139] K. Garbett, D.W. Darnall, I.M. Klotz, R.J. Williams, *Arch. Biochem. Biophys.* 135 (1969) 419.
- [140] T.C. Brunold, E.I. Solomon, *J. Am. Chem. Soc.* 121 (1999) 8277.
- [141] K. Zhang, E.A. Stern, F. Ellis, J. Sandersloehr, A.K. Shiemke, *Biochemistry* 27 (1988) 7470.
- [142] A.K. Shiemke, T.M. Loehr, J. Sandersloehr, *J. Am. Chem. Soc.* 106 (1984) 4951.
- [143] R.E. Stenkamp, L.C. Sieker, L.H. Jensen, J.D. Mccallum, J. Sandersloehr, *Proc. Natl. Acad. Sci. USA* 82 (1985) 713.
- [144] J.W. Dawson, R.H. Wang, J.M. Schredder, G.R. Rossman, H.B. Gray, H.E. Hoening, *Biochemistry* 11 (1972) 461.
- [145] A.K. Shiemke, T.M. Loehr, J. Sandersloehr, *J. Am. Chem. Soc.* 108 (1986) 2437.
- [146] C.R. Lloyd, E.M. Eyring, W.R. Ellis, *J. Am. Chem. Soc.* 117 (1995) 11993.
- [147] T.C. Brunold, E.I. Solomon, *J. Am. Chem. Soc.* 121 (1999) 8288.
- [148] T. Lovell, W.-G. Han, T. Liu, L. Noodleman, *J. Am. Chem. Soc.* 124 (2002) 2890.