First-Principles Transition-Metal Catalysis: Efficient and Accurate Approaches for Studying Enzymatic Systems

by

Heather J. Kulik

B.E., The Cooper Union (2004)

Submitted to the Department of Materials Science and Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Materials Science and Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2009

Author .....
Department of Materials Science and Engineering
January 7, 2009

Certified by .....
Nicola Marzari
Associate Professor
Thesis Supervisor

Accepted by .....
Christine Ortiz
Chair, Departmental Committee on Graduate Students

First-Principles Transition-Metal Catalysis: Efficient and Accurate Approaches for Studying Enzymatic Systems

by

Heather J. Kulik

Submitted to the Department of Materials Science and Engineering
on January 7, 2009, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Materials Science and Engineering

Abstract

Transition metals are ever-present as reactive centers in biological and inorganic catalytic cycles. However, the open shell character which gives $3d$ transition metals unique reactive properties also makes transition metal complexes challenging to study using traditional first principles approaches. Density functional theory is a widely popular computational approach because it recasts a many-body problem of interacting electrons into an equivalent problem of non-interacting electrons, greatly reducing computational cost. Each electron lives in the electric field of the total electron density, giving rise to a problem known as self-interaction; that is, each electron sees the total field including itself, and is therefore repelled by itself. Such an error is maximal in systems with highly localized electrons, in particular transition metals. We introduce an approach in which we augment standard density functionals with a Hubbard U term that helps to counteract the unphysical delocalization of electrons due to errors in exchange-correlation functionals. A Hubbard U approach has already been successfully applied to highly correlated systems in the solid state, but we introduce it for the first time to study the transition metal centers of molecules. This approach, we will show, is even more fitting for single-site molecules where the Hubbard U term need only counteract local effects (e.g. excessive hybridization with ligands) as opposed to multi-site systems where both short-range and long-range self-interaction problems are simultaneously present. The simplified, linear-response formulation we use in conjunction with density functional theory permits direct calculation of the Hubbard U, which is an intrinsic property of the system. We also extend this DFT+U approach by obtaining the linear-response U self-consistently

as a property of the DFT+U density, further increasing accuracy. We apply our approach to several paradigmatic systems: spin state splittings and structural properties of Fe_2 and other small molecules as well as the addition-elimination reactions of hydrogen and methane on FeO^+ to form water and methanol, respectively. We find that errors from common density functionals which are over 1.0 eV are greatly reduced to on average 0.1 eV when the DFT+U approach is implemented as compared against experiment and highly accurate but expensive quantum chemistry. We also improve structural and vibrational properties, ground state spin identification for a given configuration, and qualitative descriptions of reaction mechanism. Thanks to the minimal overhead of our DFT+U approach, we have also studied properties of systems of over one thousand electrons in size: in particular, the spin density profiles of functionalized cobalt porphyrins on a metal slab support and the reaction mechanism of the halogenating non-heme Fe(II) enzyme, SyrB2. Efficient and accurate study of transition metal chemistry paves the way for predictive and targeted design of catalysts that provide unique solutions for green chemistry and optimal harnessing of alternative energy sources.

Thesis Supervisor: Nicola Marzari
Title: Associate Professor

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Transition metal catalysis

Transition metals are ever-present as reactive centers in biological and inorganic catalytic cycles. In biological systems, mid-row $3d$ transition metals facilitate reactions ranging from methane to methanol conversion at an antiferromagnetically coupled di-metal, di-oxo center in methane monooxygenase[3, 4], to unactivated alkane halogenation by a high-energy, high-spin ferryl-oxo center in the halogenase SyrB2[5], to oxygen binding at iron porphyrins in hemoglobin[6]. Inorganic systems centered around transition metals are also designed to permit unique synthetic pathways. However, the open shell character which gives $3d$ transition metals unique reactive properties also makes transition metal complexes quite difficult to study using traditional first principles approaches. A density functional theory approach can permit the efficient study of systems of several hundred atoms in size, which is ample for an enzyme active site or inorganic catalysts. However, common implementations of density functional theory utilize mean field approximations, which are ill-suited

to studying $3d$ transition metals with highly localized, open shell orbitals[7]. It is clearly necessary to develop and pinpoint efficient and accurate approaches which are suitable for the study of transition metals which are at the heart of many critical reactions. By gaining an accurate understanding of the electronic structure at transition metal centers, we can also learn to engineer systems in order to exploit and enhance many known biological and inorganic reactions.

1.2 Thesis outline

The remainder of this thesis is outlined as follows:

- Chapter 1 reviews the framework of *ab initio* approaches employed in the thesis including both standard density functional theory and Hartree-Fock-based approaches. The discussion of and motivation for augmentation of standard exchange-correlation functionals with a Hubbard U term, employed throughout this thesis, is also included. Additionally, the building up principles of atoms and molecules from the quantum numbers of individual electrons is considered as a means of explaining the complexity of transition metal systems.

- Chapter 2 discusses the effect and role of Hubbard U approaches in describing the electronic structure of several types of diatomic molecules, which are fundamental representatives of the chemical bond. Detailed results on the iron dimer are considered which demonstrate the importance of a self-consistent U. The discussion of hydrides, oxides, and fluorides each provide unique insight into different bonding regimes: covalent bonding via $4s$-$1s$ interactions only, covalent bonding via the $3d$ manifold, and ionic bonding, respectively. The relative performance of GGA+U in describing structural, vibrational, and en-

ergetic properties of these molecules is compared against alternative exchange-correlation functionals, accurate quantum chemistry, and experiment.

- Chapter 3 considers the paradigmatic, gas phase, addition-elimination reactions of hydrogen and methane on bare FeO^+ to produce water and methanol, respectively. These reactions exhibit many of the key challenges for density functional theory in studying transition metal catalysis, which include spin surface crossings and two-state reactivity. We show how application of a GGA+U approach greatly reduces the errors exhibited by standard GGA of over 1.0 eV to around 0.1 eV, when compared against highly accurate quantum chemistry and available experiments. We also show how qualitative features important for understanding the reaction mechanism such as exothermicity, barrier heights, stationary point structure, and spin crossover are properly reproduced only with GGA+U.

- Chapter 4 discusses the applications of several transition metal-based porphyrin systems. In the first case, we look at how a U term changes the relative spin and symmetry ordering of Mn(oxo)-porphyrins with respect to other functionals. In the second case, we study the reduction of CO_2 on various oxidation states of Co-porphines to determine the gas phase reaction mechanism of CO formation, key for reduction of greenhouse gases and formation of useful byproducts. Finally, we discuss the functionalized tetrabromophenyl porphyrins (TBrPP) with cobalt and copper centers on Cu(111) metal surface, which represent some of the largest systems currently feasible for study with modern computers and GGA+U. These TBrPP-Co molecules experimentally exhibit unique Kondo temperature patterns commensurate with spin density increasing radially out from the metal center, which we reproduce only with the GGA+U approach.

- Chapter 5 focuses on the chlorination of L-Threonine by a SyrB2 active site model complex. The reaction steps were studied with GGA+U and were determined to have energetic barriers commensurate with biological conditions. Model substrates CH_4 and L-Threonine are both considered for the hydrogen abstraction and subsequent chlorination steps, and, for L-Threonine, it is found to be structurally and energetically favorable to carry out both steps in a coupled fashion. The SyrB2 halogenase shares structural similarities with hydroxylases but does not hydroxylate, and the structural reasoning for this is also presented.

1.3 First principles approaches

The theoretical approaches we will use here are generally called first principles or *ab initio* approaches. The similarity amongst all of these differing methods is that they take the many-body Schrödinger equation, which is not solvable for all but the simplest systems, and reformulate it in a way that is feasible to solve with modern computers. The approximations of these formulations vary and limit their accuracy of such approaches in differing ways. However, recent advances to modern first principles approaches have opened the door to accurate, efficient computation of the electronic structure of systems of increasing and practical size.

1.3.1 Density functional theory

The first principles method we primarily use is called density functional theory. This approach was derived initially for nearly homogenous systems in the solid state but is now used widely as well in isolated systems. We begin with a time-independent

Schrödinger equation in the adiabatic approximation. In this instance, ionic positions are treated as parameters because the nuclei are assumed to move very slowly with respect to the electrons and thus their motion is decoupled from the motion of electrons. Such an approximation only breaks down for the lightest of nuclei, namely hydrogen. Recasting the Schrödinger equation with these approximations, we get:

$$\hat{H}\Psi(r_1, ...r_N) = E\Psi \tag{1.1}$$

$$\hat{H} = \hat{T}_e + \hat{V}_{ne} + \hat{V}_{ee} = \sum_i (-\frac{1}{2}\nabla_i^2) + \sum_i v(r_i) + \frac{1}{2}\sum_{i \neq j} \frac{1}{|r_i - r_j|} \tag{1.2}$$

$$v(r) = -\sum_I \frac{Z_I}{|r - R_I|}. \tag{1.3}$$

The components of this Hamiltonian include a kinetic energy term, a Coulombic potential term for the interaction of electrons and ions, and an electron-electron interaction term. The total energy of this system may be variationally minimized to obtain a ground state wavefunction of the system. The Coulombic interaction of the ions may also be added to the expression for the energy. In practice, the many-body wavefunction which needs to be minimized is of $3N$-dimensions, where N is the number of particles. This problem is made tractable through the fundamental reformulations of density functional theory which was first expressed in works by Hohenberg and Kohn[8] and Kohn and Sham[9]. First, it was shown that a three dimensional density of the electrons may be used as a basic variable instead of the $3N$-dimensional wavefunction. For a particular ground state potential, $v(r)$, the state Ψ is fully determined, as is the charge density $\mathbf{n}(r)$. The total energy recast in these terms is then

$$F[\mathbf{n}(r)] \equiv \langle \Psi | \hat{T}_e + \hat{V_{ee}} | \Psi \rangle \tag{1.4}$$

$$E[\mathbf{n}(r)] \equiv \int v(r)\mathbf{n}(r)dr + F[\mathbf{n}(r)] \tag{1.5}$$

The $F[\mathbf{n}(r)]$ is a universal functional which, if known, would make DFT exact. However, most of the work in the field of density functional theory in the past 40 years has been instead dedicated towards defining useful approximations to $F[\mathbf{n}(r)]$ which yield accurate energies and properties. The first step toward defining $F[\mathbf{n}(r)]$ is to map the system onto a non-interacting single-electron system - typically a Slater determinant of single particle orbitals, in which case $\mathbf{n}(r) = \sum_i |\psi_i(r)|^2$. The ground state energy is then recast in terms of a kinetic energy, Hartree energy, and exchange-correlation energy.

$$E[\psi_i] = \sum_i -\frac{1}{2}\int \psi_i^*(r)\nabla^2\psi_i(r)dr + E_H[\mathbf{n}(r)] + E_{xc}[\mathbf{n}(r)] + \int v(r)\mathbf{n}(r)dr \tag{1.6}$$

The first two terms are derived from the non-interacting system while the exchange-correlation energy includes all of the many-body effects not present in the rest of the functional and thus is essentially the recasting of the aspects of $F[\mathbf{n}(r)]$ which were not known before. The simplest exchange-correlation functional which is used is the local density approximation (LDA). In LDA, the exchange-correlation functional is local and derived from a homogeneous electron gas with the density $\mathbf{n}(r)$.

$$E^{LDA}[\mathbf{n}(r)] = \int \mathbf{n}(r)\epsilon_{xc}^{hom}(\mathbf{n}(r))dr \tag{1.7}$$

An improvement upon this approach is the generalized-gradient approximation (GGA) which also takes into account the inhomogeneity of the density by incorporating the gradient into the exchange-correlation functional. The next step, inclusion of the Laplacian of the density, has also been implemented, and these functionals are called

meta-GGAs. Proponents of increasing the terms used to describe the density refer to this progression as Jacob's ladder[10], but it is not yet clear that these functionals provide a universal improvement over GGAs when we consider both cost and accuracy. All of these approaches have built in one particularly fundamental approximation that each electron lives in the field of all electrons including itself. This error, known as self-interaction, is particularly relevant in transition metal molecules, though it also affects most systems to varying degrees. One popular approach for building functionals is to mix in some percentage of Hartree-Fock exchange. This approach works to correct simultaneously some of the self-interaction problems inherent in standard semi-local functionals while also adding in some small amount of long-range Hartree-Fock exchange to the solution. The major shortcoming of these so-called hybrid functionals is that the percentage of exchange used can only be determined from a fit of the parameters to a large data set. While this approach yields fairly consistent results for organic systems, application of hybrid functionals to transition metal systems tends to yield unreliable results[11, 12]. If $F[\mathbf{n}(r)]$ were in fact known exactly, density functional theory would yield the correct solution in the adiabatic approximation to the Schrödinger equation. Unfortunately, modern functionals are still approximations, and we will investigate further the way in which they fail to accurately describe highly heterogeneous molecular systems.

1.3.2 Spin density functional theory

The effective potential of the Kohn-Sham equations contains no reference to electron spin. The only inherent constraint, even in cases of odd electrons is that:

$$\rho(\vec{r}) = \rho_\alpha(\vec{r}) + \rho_\beta(\vec{r}) \tag{1.8}$$

but there is no information about individual spin densities, as there is nothing in the potential which is spin dependent. Instead, what are referred to as spin density functionals are more commonly employed to describe the physics of open-shell systems, an approach sometimes referred to as unrestricted Kohn-Sham (UKS)[13]. In the simplest example, dissociation of H_2 molecule, a UKS approach will break the symmetry of the dissociating molecule and give rise to polarization in the system as it dissociates to doublet H atoms. Conversely, a standard Kohn-Sham will restrict the true symmetry of the molecule but give the wrong energies with respect to a UKS approach. Such shortcomings only appear at relatively long distances near dissociation and still permit accurate treatment of slightly stretched bonds, as we will later show. As a result of the description of KS orbitals in the unrestricted spin scheme, it is often considered that the expectation value, $<\hat{S}^2>$ may not correspond with a good quantum number. Instead $\hat{S}_z$ is the good quantum number in the unrestricted Kohn-Sham formalism. Additionally, the deviation of $<\hat{S}^2>$ from $<\hat{S}_z>^2$ is often referred to as spin-contamination, a problem which is known to plague Hartree-Fock approaches. In the UKS formalism, spin-contamination is not nearly as relevant as it is in HF theory because the KS wavefunction is not the true wavefunction, and the inherent correlation in the KS approach also reduces the overall amount of spin-contamination most typically observed.

Additional work was undertaken by Von Barth, Gunnarsson and others to show that spin density functional theory may be meaningfully applied to determine properties of the ground state of a well-defined spin and symmetry of a system[13, 14, 15]. That is, the energetically lowest lying state of each spatial or spin irreducible representation is in a sense a 'ground state'. Extensions to this approach to consider systems where spins are antiferromagnetically coupled can be employed, albeit somewhat problematically. Some excited states or multiplets cannot be described uniquely

by $\hat{L}^2$ and $\hat{S}^2$ and extensions to DFT are necessary for describing the energetics of these states. The simplest example of such a shortcoming is in describing the low-lying triplet and singlet states of carbon atom[14]. However, the multiplet problem in C atom is a limiting example, and the correct assumption that we can describe low-lying states with DFT is something we rely upon for the majority of this thesis but further discussion is outside of the scope of this introduction.

1.3.3 The DFT+U approach for correlated-electron systems

Following the success of LDA+U (and GGA+U) in accurately reproducing the ground state of transition metal oxide solids[7, 16], we, for the first time, apply this approach to single- or few- site molecular systems[17]. We will show in this thesis that the LDA+U approach is particularly suitable for application to single-site molecules because the U term plays a singular role in tuning the hybridization in the local bonding interactions. Traditional density functional approaches such as the local-density or generalized-gradient approximations treat transition metal systems quite poorly. These approximations use a mean field approach and assume that the charge density is delocalized and nearly homogeneous throughout the molecular framework. This assumption is particularly untrue for the $3d$ valence electrons of the transition metal centers of molecules. These $3d$-derived molecular orbitals actually closely resemble those of the isolated atomic orbitals, and therefore, energetics of such structures are often described incorrectly using standard exchange-correlation functionals.

The LDA+U or GGA+U method is an approximation which corrects LDA or GGA in order to better describe strong electronic correlations. It is based upon a Hubbard model approach for treating strongly correlated systems[18, 19]. As a

result, electronic structure approaches have incorporated Hubbard U terms in order to better treat correlated solids for quite some time[19]. A simplified, rotationally invariant LDA+U method was recently introduced by Cococcioni and de Gironcoli[7]. The major advantage of this approach, as we will show, is that the U term may be calculated as a property of the system from linear-response theory. The general form of the LDA+U functional, where LDA really means any GGA or LDA or other mean-field approximation, is as follows:

$$E_{LDA+U}[\mathbf{n}(r)] = E_{LDA}[\mathbf{n}(r)] + E_{Hub}[n^{I\sigma}_{mm}] - E_{dc}[n^{I\sigma}] \tag{1.9}$$

In this equation, $\mathbf{n}(r)$ is the electron density, while $n^{I\sigma}_{mm'}$ is a generalized atomic orbital occupation for the atom at site I which is the source of the strongly correlated electrons whose total occupation summed over all orbital angular momenta, m, is described as $n^{I\sigma} = \sum_m n^{I\sigma}_{mm}$. The $E_{LDA}[\mathbf{n}(r)]$ term represents the contribution of the energy from the LDA or GGA functional, but this functional already includes some mean-field treatment of on-site electronic interactions. It is necessary when we introduce a correction to model properly the on-site correlation, in this case $E_{Hub}[n^{I\sigma}_{mm}]$, that we also subtract out the mean field correlation in LDA using a double counting term, $E_{dc}[n^{I\sigma}]$. To recover the proper electron-electron interaction, the $E_{Hub}[n^{I\sigma}_{mm}]$ may be broken down into a Coulomb (the parameter U) and exchange interaction (the parameter J), which have been modeled formally by borrowing from atomic Hartree-Fock Slater integrals[19, 7].

By neglecting the exchange term, or essentially modeling an effective interaction, U_{eff} which is equal to $U - J$, we are able to greatly simplify previously derived expressions for the LDA+U energy functional. The double-counting term, $E_{dc}[n^{I\sigma}]$,

may be approximated in terms of the U and J parameters as:

$$E_{dc}[n^I] = \sum_I \frac{U}{2} n^I(n^I - 1) - \sum_I \frac{J}{2}[n^{I\uparrow}(n^{I\uparrow} - 1) + n^{I\downarrow}(n^{I\downarrow} - 1)] \tag{1.10}$$

With the simplification due to U_{eff} and the assumption of a renormalized spherical Hartree-Fock form, a "+U" correction may be derived[7]:

$$\begin{aligned} E_U[n^{I\sigma}_{mm'}] &= E_{Hub}[n^{I\sigma}_{mm}] - E_{dc}[n^{I\sigma}] \\ &= \frac{U}{2} \sum_I \sum_{m,\sigma} n^{I\sigma}_{mm} - \sum_{m'} n^{I\sigma}_{mm'} n^{I\sigma}_{m'm} \\ &= \frac{U}{2} \sum_{I,\sigma} Tr[\mathbf{n}^{I\sigma}(1 - \mathbf{n}^{I\sigma})] \end{aligned} \tag{1.11}$$

If we consider an occupation matrix which has been diagonalized, the above expression simplifies even further to an $E_U[n^{I\sigma}_{mm'}]$ of the form $\frac{U}{2} \sum_{I,\sigma} \sum_i \lambda_i^{I\sigma}(1 - \lambda_i^{I\sigma})$ where $\lambda_i^{I\sigma}$ is simply the occupation of an orbital type i of spin σ on site I. In general, projector operators, $P^I_{mm'}$ for localized orbital manifold m on site I may be used to build an occupation matrix. The projection operators may be defined by any number of bases, but, in this thesis, we use the atomic $3d$ orbitals, which are encoded in the pseudopotential of the atom. As we will show in Chapter 2, it is particularly key to ensure that the atomic orbitals in the pseudopotential are normalized and the best possible approximation for the appropriate charge state and electron configuration of the molecule. An alternative approach is to utilize Wannier functions as the projection manifold, but these should be kept frozen when comparing different structures[20].

It has been shown [7] that the LDA energy of an atom in contact with a reservoir of electrons will essentially behave parabolically as a function of electron occupation

with the minimum being situated at a fractional occupation number. The true energy of this model system is a statistically weighted, piecewise-linear function[21]. Therefore, the "+U" term that is added back to an LDA functional may be thought of as the correction which offsets the unphysical curvature of the original LDA energy. In order to measure the value of U from first-principles, we wish to recover the second derivative with respect to occupations of our system. In addition to the curvature in energy of the fully screened, interacting system, there exists some non-linear variation in a Kohn-Sham, non-interacting system due to rehybridization after a shift in occupation number. The effective curvature then should be the difference between the screened, interacting U and the bare, non-interacting one. This LDA+U functional therefore favors either fully occupied or empty orbitals. The tendency of the LDA+U functional to disproportionately favor integer occupations is in direct opposition to the LDA or GGA tendency to favor fractional occupations. Under most conditions, as we will show in Chapter 2, the bonds may become slightly elongated with the addition of a "+U" term, but energetics are significantly improved. The over-elongation is partially due to the fact that the projection manifold we choose is that of the isolated atomic orbitals.

In earlier approaches, super-cell constrained-density functional calculations have been applied [19, 7]. However, Cococcioni and de Gironcoli derived a linear-response approach in order to calculate the U in a much simpler way. Using Janak's theorem, we may equate the eigenvalue of state i with the derivative of the energy with respect to the occupation of the state, and this helps to recast the second-derivative problem as a first-derivative, linear-response property. Thus, the partial derivative of the energy of the state with occupation q_J with respect to q_I is -α_I, and an analogous expression is true for the bare Kohn-Sham system. The second derivative of the energy with respect to occupations, therefore, is simply the derivative of the potential

α_I with respect to occupations q_I. Such a change of variables is key because it is difficult to constrain occupations of the localized orbitals of a system. In order to further simplify the practical details of the calculation, the authors proposed a Legendre transform to the energy which made the independent variables the potential, α_I of the localized orbitals. The new energy expression could then be written as a function of α_I:

$$E[\alpha_I] = \min_{\mathbf{n}(r)} E[\mathbf{n}(r)] + \sum_I \alpha_I n_I \tag{1.12}$$

and an analogous equation may be written for the bare system. A single potential shift takes the form $\Delta V = \sum_I \alpha_I P^I$, where P^I are the projector operators, and thus affects each individual localized orbital. The linear response functions of this system with respect to a local perturbation ΔV are then:

$$\begin{aligned} \chi_{IJ} &= \frac{\partial^2 E}{\partial \alpha_I \partial \alpha_J} = \frac{\partial n_I}{\partial \alpha_J} \\ \chi^0_{IJ} &= \frac{\partial^2 E^0}{\partial \alpha^0_I \partial \alpha^0_J} = \frac{\partial n_I}{\partial \alpha^0_J} \end{aligned} \tag{1.13}$$

where I and J are neighboring sites containing localized orbitals or differing manifolds of localized orbitals on the same atomic site. Response functions are obtained by measuring the change in occupations from several small potential shifts to an already converged single point density, and the associated computational cost is negligible (see Fig. 1-1). The U associated with a single site I is then simply $U = \chi_0^{-1} - \chi^{-1}$. In this approach, special care should be taken in the inversion of the response functions. As response functions approach asymptotically small numbers, the inverses of the linear response functions will diverge. The difference of the response functions, therefore, may appear to fall within a reasonable range ($U = 0 - 10$ eV), while the value is simply an artifact of the calculation. It is key to examine the magnitude of

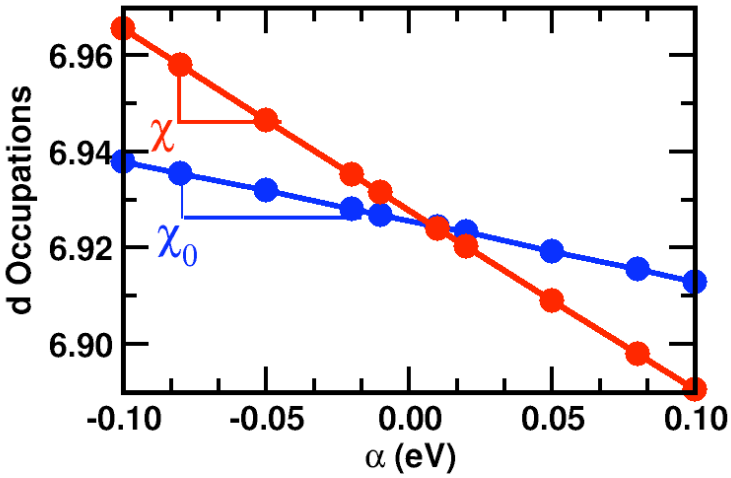


Figure 1-1: Examples of bare, χ_0, and converged, χ, linear response functions for the ^{4}TS-1 structure in the addition-elimination reaction of H_2 on FeO^+.

the response functions before inversion, as we will later discuss in greater detail in Chapter 2. The standard linear response approach of Cococcioni and de Gironcoli is straightforwardly applied to multiple manifolds of electrons, as well.

Normally, we consider only $3d$ states in transition metals to be relevant for determining a U, but the value of U is dependent strongly upon the mode of bonding. If, as is true in the case of transition metal hydrides (see Chapter 2), $4s-3d$ hybridization is strong or $4s$ orbitals dominate the molecular bonding, then the response of this manifold is also relevant. In this case, the response functions, χ_0 and χ should be expressed as matrices. These matrices are then inverted and their diagonals are subtracted to give the U_{3d} and U_{4s}. In solids, the interaction and hybridization of $3d-4s$ is not likely to be relevant, but in molecules strong mixing, particularly of $3d_{z^2}$ and $4s$ orbitals often occurs. It is the fact that these states all live in the same energetic window and strongly overlap spatially that make both manifolds relevant for treatment by LDA+U.

While the linear-response formulation has been shown to be quite useful, a number of additions to this formalism can help to increase the accuracy of the results derived from this method. In the solid state, LDA+U often opens a gap in Mott insulators which are predicted by LDA to be metals[7]. Analogously for molecules, LDA may predict the wrong ground state spin and symmetry when compared to LDA+U. In either case, the LDA ground state wavefunction differs greatly from that obtained using LDA+U, and we expect that the linear response of the LDA system does not work well to characterize the true LDA+U system. We recently proposed[17] an approach in which the true U of the system may be found in a fully self-consistent manner. To clarify our approach, we first identify in the GGA+U functional the electronic terms that have quadratic dependence on the occupations by expressing the terms as a function of a projection matrix which has been diagonalized:

$$\begin{aligned} E_{quad} &= \frac{U_{scf}}{2} \sum_I \left[\sum_i \lambda_i^I \left(\sum_j \lambda_j^I - 1 \right) \right] \\ &+ \frac{U_{in}}{2} \sum_I \sum_i \lambda_i^I (1 - \lambda_i^I). \end{aligned} \tag{1.14}$$

The first term represents the contribution already contained in the standard GGA functional, modeled here as a double-counting term, while the second term is the customary "+U" correction. Therefore, U_{scf} represents the effective on-site electron-electron interaction already present in the GGA energy functional for the GGA+U ground state when U is chosen to be U_{in}. Consistency is enforced by choosing U_{in} to be equal to U_{scf}. The U obtained from linear-response[7] (labeled here U_{out}) is also obtained by differentiating Eq. 1.14 with respect to λ_T^I:

$$U_{out} = \frac{d^2 E_{quad}}{d(\lambda_T^I)^2} = U_{scf} - \frac{U_{in}}{m}, \tag{1.15}$$

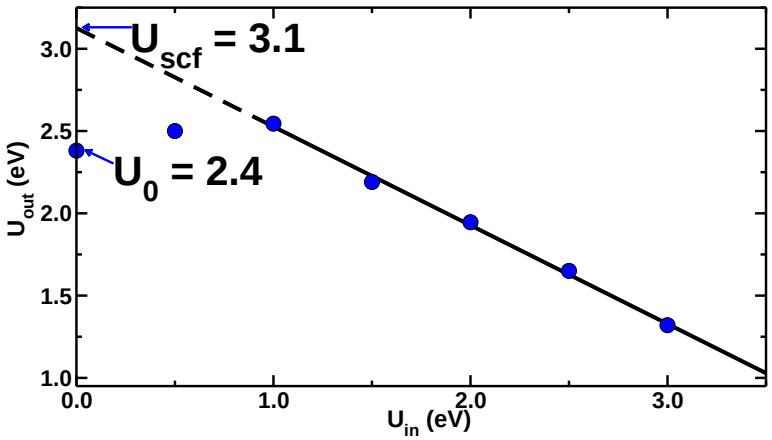


Figure 1-2: Linear-response U_{out} calculated from the GGA+U_{in} ground state of $^7\Delta_u$ Fe_2, together with the extrapolated U_{scf}. U_0 is U_{out} calculated for U_{in}=0.

where $m = 1/\sum_i (a_i^I)^2$ can be interpreted as an effective degeneracy of the orbitals whose population is changing during the perturbation (to linear order, $\delta\lambda_i^I = a_i^I \delta\lambda_T^I$ with $\sum_i a_i^l = 1$ and $\frac{d^2}{d(\lambda_T^I)^2} = \sum_{ij} a_i^I a_j^I \frac{d^2}{d\lambda_i^I d\lambda_j^I}$). Even if in principle U_{scf} depends on U_{in}, we find it to be constant over a broad interval, as apparent from Fig. 1-2: U_{out} is linear in U_{in} for the relevant range of $U_{in} \sim U_{scf}$. Thus, from few linear-response calculations for different U_{in} ground states we are able to extract the U_{scf} that should be used. This approach is particularly useful for the transition metal dimers, as illustrated in Fig. 1-2. Otherwise, the effective difference between U_0 and U_{scf} is minimized by the constraints we employ throughout this thesis. That is, for molecular systems, we constrain the total spin of the system and enforce fixed occupations without a smearing or electronic temperature in most cases. We therefore calculate the value of U for each spin and symmetry independently and average over the low-lying states. In particular, we highlight in Chapter 2 the fact that U is highly dependent upon both spin and symmetry as well as nuclear identity and configuration and must always be calculated and used as a metric to determine whether

or not an LDA+U approach is useful for improving energetics and structure. That is, if the linear response properties of a system approach zero and the subsequent U is small, there is no particular need to utilize an LDA+U approach. In general, we find that even where a U is calculated to be relatively small, LDA+U will still sensitively tune the spin splittings of molecules including the transition-metal hydrides (see Chapter 2). When we try to compare the energetics of systems with widely varying coordination numbers, as is commonly associated with reactions involving bond-breaking and dissociation, the value of U must be averaged (see Chapter 3). While averaging the value of U is a shortcoming of the approach, higher resolution may be achieved where it is deemed necessary, and even an average U is typically found to be an improvement over standard a standard LDA or GGA approach (see Chapters 4 and 5).

We will also demonstrate in Chapter 2 that LDA+U is most useful in covalently bonded metal-ligand interactions where the U term helps to reduce the tendency of LDA or GGA to overstabilize molecular orbitals which permit delocalization, those of bonding character, and destabilize those which are highly localized, such as those of non-bonding or anti-bonding character. In both the solid state and in multi-valence, multi-metal molecules, one must consider and treat both local hybridization as well as long range charge transfer and partitioning between differing metal centers (see Fig. 1-3). For these multi-site systems, a different treatment of the long range charge transfer is necessary in addition to LDA+U for short-range hybridization, but decoupling these two effects is often challenging. Constrained density functional theory is a suitable solution for the long range partitioning of charges between ions so long as the constraints can be well defined with respect to the physical system.

Shortcomings of the LDA+U approach as applied to single-site molecules are comparatively limited. We primarily identify in this work that sensitivity to pseu-

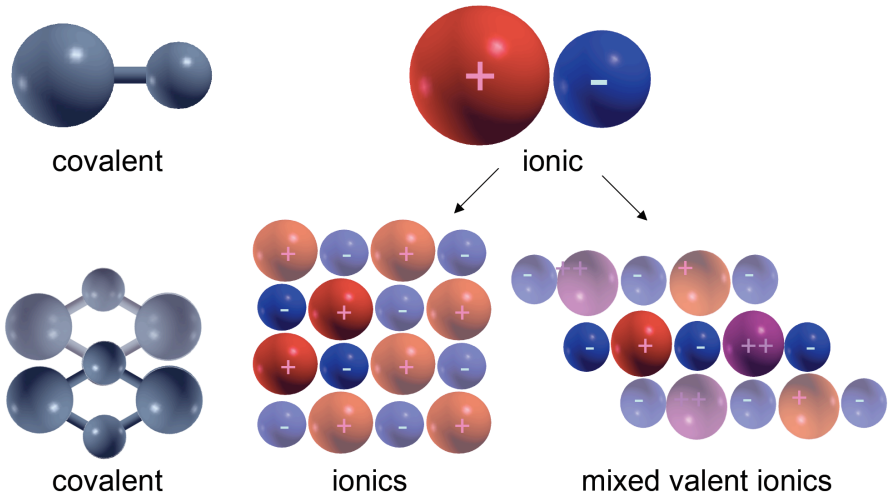


Figure 1-3: Comparison of interactions present for a molecule containing a single transition-metal center (top) and multiple metal centers present in either a molecule or a solid (bottom), where, for the ionic case, metals may have differing valency.

dopotential quality is surprisingly large, orbital manifold choice is critical where several orbitals are close in energy (e.g. $3d$ and $4s$), and that numerical errors can give rise to erroneously large values of U which are unphysical for the system in question (see Chapter 2). It has been additionally observed that, in a few cases, competition between excessive hybridization in LDA or GGA and the penalty on such hybridization in LDA+U can cause molecules to become unbound, although this is extremely rare. Overall, the use of LDA+U cannot supplant the simultaneously necessary use of chemical intuition when studying any transition-metal containing molecule. By employing LDA+U along with standard LDA, GGA, and hybrid functionals in this work, we are able to systematically improve interpretations of bonding and electronic structure as well as explain where LDA+U gives initially unsatisfactory results.

1.3.4 Post-Hartree-Fock techniques[1, 2]

Hartree-Fock (HF) calculations and corrections to these approaches have been extremely useful in the study of molecular systems. The Hartree-Fock approach generally exhibits some similarities with density functional theory when it comes to simplifying assumptions, particularly the use of the variational principle. However, the Hartree-Fock approach was derived in the limit that the wavefunction is defined by a single Slater determinant of single particle wavefunctions. Such wavefunctions are built typically from a localized basis set which themselves have limitations outside the scope of the discussion here. In Hartree-Fock theory, each electron sees the other electrons as an average field. A trial Slater determinant is used:

$$\Psi = \frac{1}{\sqrt{N!}}\mathrm{Det}\{\psi_a(1)\psi_b(2)\psi_c(3)\psi_d(4).....\psi_x(N)\} \tag{1.16}$$

and an iterative procedure to minimize the expectation value of H:

$$H = \sum_{i=1}^{N}\left(\frac{\mathbf{p_i^2}}{2m} - \frac{e^2}{4\pi\epsilon_0}\frac{Z}{r}\right) + \sum_{i<j}^{N}\frac{e^2}{4\pi\epsilon_0}\frac{1}{r_{ij}} \tag{1.17}$$

where H contains kinetic energy terms, an electrostatic Hartree (or Coulomb) term, and an exchange term. The self-consistency of the electrons is then variationally achieved with the external field. In this formulation, there is no spurious self-interaction that is present in common exchange-correlation functionals of DFT. The self-interaction which is present in the Hartree term shared also with common expressions of density-functional theory is canceled in Hartree-Fock theory exactly by the exchange term. This cancellation of errors between the Hartree and exchange terms are what motivate the use of so-called "hybrid" functionals in density func-

tional theory. The Slater determinant also introduces a degree of correlation by being anti-symmetric with respect to exchange of single particle orbitals in order to satisfy the Pauli exclusion principle. However, Hartree-Fock theory otherwise neglects correlation between electrons. The correlation energy, which is generally defined as:

$$E_{corr} = E - E_{HF} \tag{1.18}$$

is typically a small portion of the total energy, but it has been shown to be critical for accurate *ab initio* study of relative energetics and bonding. Full configuration interaction (FCI) is a post-Hartree-Fock method which accurately recovers all of the correlation energy and, in the complete basis set limit, provides an exact solution to the non-relativistic Schrödinger equation in the Born-Oppenheimer approximation. In this approach, which we will address in greater detail shortly, one constructs a wavefunction as a linear combination of the Hartree-Fock reference and all possible excitations from this reference. In fact, a CI matrix derived from the Hartree-Fock reference and all related excitations is generated, and by diagonalizing the matrix we obtain any number of roots (each corresponding to a certain electronic state of the system) which are accurate in the complete basis set limit. This approach is extremely expensive and feasible only for the smallest systems, and therefore we consider first alternative approaches which, although still very expensive, scale more favorably.

For highly-correlated electron systems, the Hartree-Fock self-consistent solution can be a poor reference. Perturbative approaches such as Møller-Plesset (MP) work in some cases as long as a high enough order of pertubative theory is used - usually MP4 is preferred. In this case, it is also necessary to ensure that there is significant energy gaps between the occupied orbitals themselves and the gap between occupied

and unoccupied orbitals. The MP2 approach is also acceptable in some cases and scales more favorably (N^4) than MP4 (N^6). If the orbitals are nearly degenerate, the perturbative approaches become unstable and yield unreliable results. For most of the systems that we will discuss, the $3d$ electrons are closely spaced in energy and Møller-Plesset is not a good candidate.

The most commonly employed single-reference approach for small systems is coupled cluster (CC). The central basis of coupled cluster theory is that the exact CI wavefunction (within the limit of the basis set) may be described as:

$$\Psi = e^{\mathbf{T}}\Psi_{HF} \tag{1.19}$$

where $\mathbf{T}$ is the cluster operator. This operator is defined as:

$$\mathbf{T} = \sum_{i=1}^{n} \mathbf{T_i} \tag{1.20}$$

where n is the total number of electrons and the operators $\mathbf{T_i}$ generate all possible determinants having i excitations from the Hartree-Fock reference. By truncating $\mathbf{T}$, the problem becomes computationally feasible. For example, in coupled cluster with doubles (CCD), we may do a Taylor expansion of the exponential and find that we simply operate on the wavefunction with the double-excitation operator:

$$\Psi_{CCD} = e^{\mathbf{T_2}}\Psi_{HF} = (1 + \mathbf{T}_2 + \frac{\mathbf{T}_2^2}{2!} + \frac{\mathbf{T}_2^3}{3!} + ...) \tag{1.21}$$

This approach includes products of excitation operators, e.g. $\mathbf{T}_2^3$ generates hextuple substitutions, which is critical for ensuring size-consistency of a method. Size consistency refers to the ability to compare the energy of a system with n particles

to a system with $n-1$ particles. Many multi-reference approaches fail to achieve size-consistency and thus yield unreliable quantities for dissociation or comparison of states with differing numbers of electrons. While coupled cluster in its CCSD(T) form (inclusion of singles and doubles with triples treated perturbatively) is considered the gold standard of single-reference quantum chemistry, it does have several shortcomings. The most obvious one is that this approach scales with N^7 and its application is thus limited to systems of the smallest size. While the perturbative triples term saves computational cost (true CCSDT scales as N^{10}), they are strongly connected to the singles excitations. If the singles excitations are large, which is often the case in highly correlated systems, the triples term can become unstable. For some cases, a unitary transformation of the Hartree-Fock orbitals to yield a zero singles amplitude can vastly improve the estimate of the triples perturbation contribution to the energy. This approach is known as Brueckner Doubles (BD(T)) and is computationally comparable for the most part to CCSD(T).

Another consideration is that the large singles amplitude can be a signature of strong multi-reference character. A measure of the singles amplitude is known as the T_1 diagnostic:

$$T_1 = \sqrt{\frac{\sum_i^{occ.} \sum_a^{vir.} (t_i^a)^2}{n}} \tag{1.22}$$

and this is occasionally used as an indication that a multi-reference approach must be used. However, large singles amplitudes as demonstrated by a large T_1 value may in many cases be more of an indicator of singles-triples instability, which may be remedied with BD(T). There also remain special cases where a bond is strongly stretched and the electrons begin to condense into their atomic states is highly multi-determinantal and a single determinant approach such as CCSD(T) will strongly overestimate the energy.

The simplest way to introduce many-body correlation is to alleviate the restriction that the solution to the problem is a single Slater determinant. If, instead, we permit a linear combination of Slater determinants to represent our state:

$$\Psi = \sum_i c_i \Psi_i = \Psi_{HF} + c_1 \Psi_1 + c_2 \Psi_2 + ... \tag{1.23}$$

then we may variationally minimize the energy of the system with respect to the coefficients, c_i. This procedure is known as multi-configuration self-consistent field (MCSCF) and is most useful for cases where a number of alternative configurations each have comparable weights and one single determinant does not dominate the solution. In order to identify the occupation of MCSCF orbitals, f_i, we sum over the Slater determinants making up the total MCSCF wavefunction, which are known as configuration state functions (CSF):

$$f_{i,MCSCF} = \sum_n^{CSFs} f_{i,n} c_n^2 \tag{1.24}$$

The occupation of a single orbital is dependent upon whether it is occupied in a CSF as well as the coefficient that CSF carries in the total wavefunction. The result is that canonical HF orbitals are often described as being partially occupied, which introduces correlation but can complicate state and symmetry assignments. The type of correlation introduced here is known as static correlation and is not very commonly found to be relevant for stabilized ground states, but is more typically applied to excited states. The MCSCF approach is the starting point for many other multi-reference approaches, but since self-consistent HF Slater determinants are often such a poor approximation for the electronic structure of transition metal containing molecules, variationally minimizing the energy through a linear combina-

tion of several Slater determinants in the MCSCF method alone may fail to improve the result significantly. By optimizing the initial wavefunction through incorporation of other reference functions, the multi-reference approaches work to introduce dynamic correlation more easily than if one started with the Hartree-Fock reference function alone.

Multi-reference methods come in forms similar to those for single-reference approaches, but they typically require more skill and experience to utilize successfully. Beginning with the MCSCF calculation, it is easy to optimize the wavefunction to a metastable minimum which creates a poor reference. Additionally, interpreting the results of multi-reference approaches takes skill and experience. The most common perturbative approach based upon an MCSCF wavefunction is CASPT2 - complete active space perturbation theory of the second kind. The active space is defined as orbitals to which electrons can be added and removed. If m electrons are selected to occupy n 'active' orbitals, the number of CSFs, N, which must be considered in the active space can grow very large:

$$N = \frac{n!(n+1)!}{(\frac{m}{2})!(\frac{m}{2}+1)!(n-\frac{m}{2})!(n-\frac{m}{2}+1)!} \tag{1.25}$$

For all but the smallest systems, a complete active space is far too large to be deemed tractable. For instance, CH_3OH has 14 valence electrons in a 12 orbital active space, giving rise to 169,884 CSFs. A slightly more challenging example, Cr_2, requires over 2 million CSFs. Instead, traditionally a restricted active space approach is used to maintain feasibility of the calculation. The most commonly employed truncation of excitations freezes core orbitals, enforces double occupation of semi-core states, freezes highly excited virtual orbitals, and permits only a limited number of excitations out of valence orbitals into a restricted unoccupied active

space. The coefficients of the CSFs from MCSCF can serve as a guide for active space selection as well. Techniques also exist which exploit chemical intuition or symmetry arguments to further reduce the active space.

Configuration interaction (CI) based approaches are among the most commonly employed multi-reference approaches. These approaches typically truncate the full CI expansion (which, if infinite, would be exact) at double excitations. In order to easily discuss CI approaches, we consider the single reference limit and we can recast the resultant wavefunction of a CI expansion as:

$$\Psi = c_0 \Psi_{HF} + \sum_{i}^{occ.} \sum_{r}^{vir.} a_i^r \Psi_i^r + \sum_{i<j}^{occ.} \sum_{r<s}^{vir.} a_{ij}^{rs} \Psi_{ij}^{rs} + ... \tag{1.26}$$

where i and j are occupied MOs and r and s are virtual orbitals in the HF reference. Thus, the first term corresponds to all single excitations from the HF reference and the second term corresponds to all double excitations. A single-reference CI matrix may be written out in order to justify the truncation at double excitations (see Fig. 1-4). From the Condon-Slater rules and Brillouin's theorem[22], matrix elements between the HF reference and triply excited and singly excited determinants are zero. The higher the excitations, the more sparse the block of the matrix becomes. Therefore, the energy of interest is dominated by double excitations, followed by singles which interact with doubles, and then triples and so on. Therefore, it is unsurprising that the truncation at CI with singles and doubles (CISD, or MRCISD) is preferred. The advantages of this approach are that the scaling is reasonably good, about N^6, and it is a variational approach which provides an upper limit to the energy. Symmetry arguments may be used to further reduce the cost of the approach. However, these approaches do have a major shortcoming, which becomes

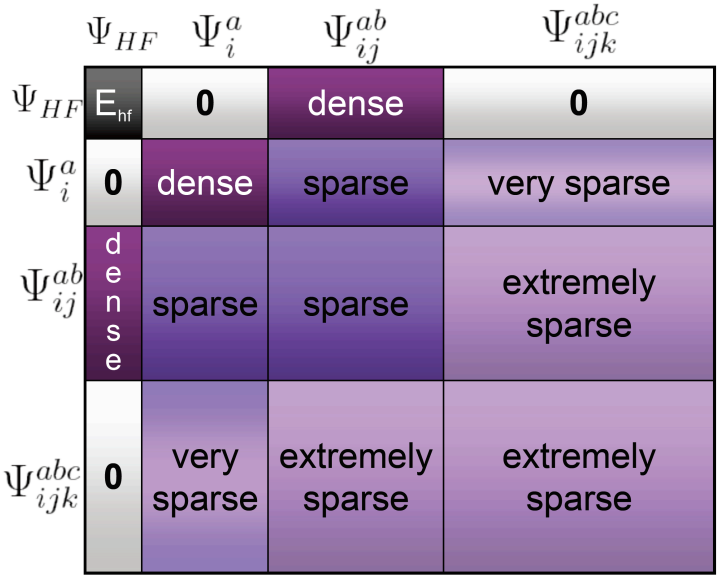


Figure 1-4: Representation of a single reference Configuration Interaction (CI) matrix in which the relative value of matrix elements is shown by varying degrees of color saturation.

apparent if we write out the CID wavefunction:

$$\Psi_{CID} = (1-c)^2\Psi_{HF} + c^2\Psi_{1\bar{1}}^{2\bar{2}} \tag{1.27}$$

where c is the coefficient determined from diagonalizing the CI matrix. If we compare the CID wavefunction of one H_2 molecule against that for two H_2 molecules separated by 25 Å, their energies are not the same. Since the two molecules are not interacting, we should be able to take the antisymmetric product of Ψ_{CID} in the above equation. However, this expression would include a term with the coefficient c^4 which corresponds to simultaneous double excitation within each molecule, a quadruply excited configuration. Approaches have been undertaken to correct for the lack of size consistency in these approaches,and the most popular of these is that

of Davidson[23], which estimates a correction for the difference in energy in a two electron system between two molecules well separated and two times the energy of an isolated molecule. This correction is written as:

$$E_Q = (1 - a_0)^2 (E_{CISD} - E_{HF}) \tag{1.28}$$

where a_0 is the coefficient of the HF determinant in the normalized, truncated CISD. Such an approach is usually referred to as CISD+Q, which is misleading. Unlike CCSD(T), the quadruples are not estimated perturbatively, and are truly justified only for the two electron systems for which the approach was derived. Much of the above derivation can be extended for the multi-reference case where a size-consistency correction is also used in the method MRCI+Q.

It is also worth noting that localized basis sets introduce variability in the results of these methods. The orbitals are often built from hydrogenic functions but also include in more sophisticated cases diffuse and high angular momentum functions. While these basis sets are relatively efficient in some cases, if a large basis set is needed to ensure accurate description of the system, calculations may become quite expensive. Results such as the total energy of a configuration can be extrapolated to the limit of infinite basis functions for only very few basis sets. In the density functional based approaches we use, the basis set consists of plane waves which can always be increased steadily and extrapolated to the infinite basis set limit. We will later discuss post-Hartree-Fock approaches for small systems as a means of benchmarking the accuracy of our density functional approaches. We find that even for closely spaced $3d$ levels and open shell systems the single-reference CCSD(T) approach is the most accurate and simplest to employ. We nevertheless compare to CASPT2 and MRCI results where the accuracy of CCSD(T) might be uncertain.

1.4 Practical considerations for DFT

Density functional calculations are completed using the QUANTUM-ESPRESSO package[24], which is a plane wave pseudopotential (PW-PP) code. Wavefunctions expressed in plane waves are useful in a solid state system or a system with a super-cell which obeys periodic boundary conditions (PBC). The plane waves included are defined simply by an energy cutoff which may be nearly continuously increased. The oscillations near the atom due to core electrons are relatively high frequency and should require many plane waves to describe. Modern solid-state codes typically employ pseudopotentials to accurately model these cores thereby reducing the number of plane waves required for the calculation. Since the core electrons do not affect bonding and are not crucial to the band structure, modeling them approximately does little to affect the results of calculations. Another cost to consider for isolated systems of a small number of systems is that to ensure isolation of each image a large amount of vacuum is included in the cell. This, combined with the number of plane waves required to describe the system, can cause calculations which use PBC to be more costly than those with localized basis sets. However, plane wave calculations can become comparably more affordable for medium to large system sizes because the number of localized basis functions begins to increase more quickly than the number of plane waves.

1.4.1 Nudged Elastic Band method for minimum energy paths

There are few approaches available computationally which can efficiently identify all critical points of a complex and steep potential energy surface (PES). The nudged elastic band (NEB) method developed by Hannes Jønsson[25, 26] is a chain of states approach for obtaining transition states and minimum energy paths (MEP) in rare

event transitions which cannot be sampled adequately by molecular dynamics. A path separating typically relaxed, stationary endpoints, denoted here as $\mathbf{R}_0$ and $\mathbf{R}_N$, is generated as an initial guess which includes several discretized points made of configurations intermediate between reactants and products, denoted as $\mathbf{R}_1$, $\mathbf{R}_2$, $\mathbf{R}_3$, and so on up to $N-1$. The total force acting on a single point in each method is the sum of a spring force along the local tangent to the path which keeps the images separated combined with the true energy minimizing force perpendicular to the path. This is written as:

$$\mathbf{F}_i = \mathbf{F}_i^s|_{\parallel} - \nabla E(\mathbf{R}_i)|_{\perp} \tag{1.29}$$

where the first term is the spring force and the second term is the true force. The spring force is a function of inter-image separation along the path:

$$\mathbf{F}_i^s|_{\parallel} = k(|\mathbf{R}_{i+1} - \mathbf{R}_i| - |\mathbf{R}_i - \mathbf{R}_{i-1}|)\hat{\tau}_i \tag{1.30}$$

where, τ_i is the local tangent to the path and k is a spring constant. The true force is also explicitly written as:

$$\nabla E(\mathbf{R}_i)|_{\perp} = \nabla E(\mathbf{R}_i) - \nabla E(\mathbf{R}_i) \cdot \hat{\tau}_i \tag{1.31}$$

Here, the force along the path is subtracted out to keep images well separated and from falling into the low energy endpoints. The initial path guess is then optimized with pseudo-molecular dynamics using the projected velocity-verlet algorithm. A shortcoming of this approach is that the images are well separated but do not usually converge to the saddle point. The result is that the energy barrier and transition state geometry must be interpolated. One modification to the original method to increase resolution at the saddle point is the inclusion of a climbing image whereby the force

applied to the image of interest, usually the highest in energy, is in the direction of increasing energy[25]. This improves resolution about the barrier where image population is most important, but under normal conditions the barrier is under-populated. As applied to the image with the highest energy, i_{max}, the climbing-image scheme applies a force to this one image:

$$\mathbf{F}_{i_{max}} = -\nabla E(\mathbf{R}_{i_{max}}) + 2\nabla E(\mathbf{R}_{i_{max}})|_{\|} \tag{1.32}$$

This is the full force of the image with the force along the path inverted, and this image does not feel the springs at all. The result is that the climbing image moves up in energy along the optimizing path and down in energy in the direction perpendicular to the band. There is no additional computational cost to the climbing image scheme (CI-NEB), and it is therefore typically implemented to get good activation energies.

A second technique, referred to as variable springs, also aims to improve resolution of the standard NEB approach [26]. Resolution of images around the saddle point is key, particularly if a climbing image is used. Additionally, a MEP may contain a particularly steep barrier region while other portions of the path may contain relatively shallow, low-energy regions. In order to populate the key regions closest to the barrier, therefore, stronger springs are used near the saddle point. There is implicitly no difficulty in varying spring strengths because they do not interfere at all with the convergence of the elastic band to the minimum energy path. In the simplest implementation of variable springs, a linear dependence on the energy is used such that images with low energy are connected by weaker spring constants:

$$k_i' = k_{max} - \Delta k \left(\frac{E_{max} - E_i}{E_{max} - E_{ref}} \right) \tag{1.33}$$

which applies if $E_i \geq E_{ref}$. If $E_i \leq E_{ref}$, then the multiplier inside the parentheses is simply 1. In all cases, $E_i = \max\{E_i, E_{i-1}\}$, or the higher energy of the two images connected by the spring. The value of E_{ref} is most typically chosen as the higher energy endpoint of the initial starting point. A wide range of spring constants may be used, but excessively large or small values may slow convergence.

In the practical implementations of NEB, several shortcomings exist, particularly when the approach is applied to the study of isolated molecules. That is, NEB was originally envisioned for the study of MEPs relevant to crystals rather than isolated systems. Most implementations of NEB, including that in PWscf[24], express the positions of images along the path in cartesian coordinates. There is no direct implementation to isolate out rotation and translation in the code, such as with a z-matrix formalism for treating reactions in terms of internal coordinates (bond lengths and angles, in particular). Additionally, it is crucial to fix at least one atom as well as the translation of one or more atoms in order to freeze out rotations of the complex in iterations of the path. This should reduce or eliminate spurious translation and rotation which leads to lengthening of the path and oscillation of the maximum path forces without reaching convergence. Other concerns for molecular systems include the fact that reorganization of molecules can involve several bond breaking and making steps in one NEB path. The result is that linear interpolation of images used by the code often leads to creation of strongly-repulsive images and a poor initial path, although recently implemented mass-weighted coordinates reduce the effect of some of these problems.

It is possible that several mechanisms and MEPs exist, and care must be taken to use chemical intuition to narrow down the likely few to consider. For systems with increasing complexity, such as proteins, only a single MEP can be studied in full and must be chosen with care. In this case, a metadynamics technique which has

been successfully employed in protein active sites better treats collective variables rather than translation-dependent reaction coordinates[27, 28]. The metadynamics approach propagates a fictitious dynamics in a coarse grained potential energy surface which consists of several collective variables intended to define the reaction coordinate and the approach explores the free energy profile by filling it up with gaussians. Alternatively, a method known as potential mean force permits the molecular dynamics study of a system during a bond breaking event, but this approach requires that the general form of the reaction coordinate is known[29]. Several schemes to make NEB more applicable to protein active sites have been considered, particularly in which hard degrees of freedom near the active site are treated by NEB while softer degrees of freedom of the total protein environment are excluded from the forces on the images to varying degrees[30]. Overall, for NEB calculations on systems of varying size considered in this thesis, we find a good initial path guess must always be given. This path provides realistic internuclear separations not simply obtained from interpolation of ionic coordinates of the reactants and products but also a chemically reasonable transition state. A climbing image method and variable springs should always be used. Where a GGA+U approach is employed, the paths are converged at integer increments of U, in some cases starting from the initial path of the previous value of U. Several hundred iterations are typically required to converge the paths to forces less than 0.03 eV/Å. A large cutoff is in some cases needed to accurately converge path forces. The NEB calculations are initially completed at a low cutoff and then repeated at higher cutoffs until forces on the transition state are truly eliminated as verified by a single point energy calculation.

1.4.2 Convergence of charged systems

Structural relaxations are carried out using PWscf[24] with the PBE-GGA[31] functional both with and without a Hubbard U term. PWscf, a plane wave pseudopotential code, is more typically used to complete solid state physics calculations. To complete a gas phase calculation, a large cell must be used to isolate the molecule from its periodic image. Extra care must be taken for charged systems, as is the case with the system of interest here[32]. In order to identify convergence with respect to cell size, a quartet intermediate (discussed in Chapter 3) was relaxed in cells of increasing size. The convergence of bond lengths and angles as well as projected $3d$ orbital occupations calculated are presented below in Figure 1-5.

In calculations with periodic boundary conditions, a compensating jellium background is added to cells in which the system has a net charge. When a charged system is repeated infinitely, it by necessity has infinite charge and therefore infinite energy, and the jellium is necessary to counteract this. The density of the compensating jellium background is a function of cell size and therefore affects the calculated energy of the system. Comparing charged cell energies computed at different cell sizes or charged cell systems to neutral systems is not straightforward. It has been shown that the long-range interaction of the image with the jellium decays very slowly. As found by Makov and Payne[32] for cubic cells with side length L, the energy behaves as follows:

$$E = E_0 - \frac{q^2\alpha}{2L} + \frac{2\pi qQ}{3L^3} + O(L^{-5}) \tag{1.34}$$

where E_0 is the true energy of the isolated system, q is the charge of the system, α is the Madelung constant of the cell, and Q is the quadrupole moment. A Makov-Payne correction to the total energy yields accurate total energies even when the data points

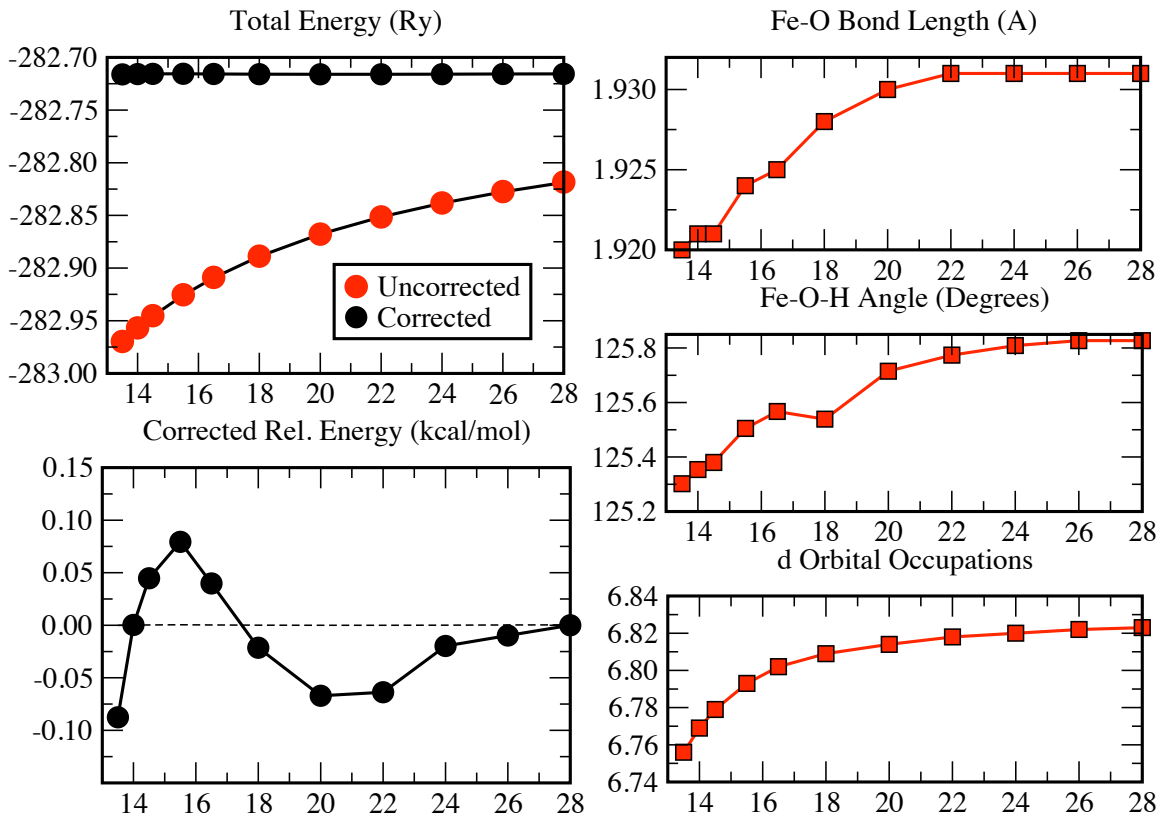


Figure 1-5: Convergence with respect to cell size of several properties of an electronic state of ^{4}Int-3$_{H_2}$ (see Chapter 3). At left, the top graph shows the difference between the corrected and uncorrected energies versus cell size, and a closer view of the corrected energies are shown at bottom. At right, the Fe-O bond length (Å), Fe-O-H angle (^{o}) and total $3d$ projected occupations are shown from top to bottom as a function of cell size.

available are for relatively small cell sizes. However, as is clear in Figure 1-5, the Makov-Payne correction is not self-consistent, and may only be used as an estimate for total energy, and not for other critical details such as structural parameters or the occupation matrix of the states. For nearly converged accuracy a 22 Bohr cell is ideal, while for preliminary study a 16.5 Bohr cell was used to increase the efficiency of calculations. As a compromise, a cell size of 20 Bohr is used for most calculations

as this nearly eliminates image effects on structure and occupations, but a Makov-Payne correction is still employed when an absolute number is needed for a charged system's energy. It is also of note that since the terms in the energy dependence are primarily cell size and charge dependent, systems containing similar ionic structures and charge in the same cell will not require the correction for comparison of relative energies.

It was necessary to be cautious in choosing an energy cutoff for the plane waves for the system of interest even though ultrasoft pseudopotentials were used which should reduce significantly the wavefunction cutoff. Multiplet splittings were conserved at 40 Ry cutoff for the wavefunction and 480 for the charge density (see Table 1.1). Using a higher dual for lower wavefunction cutoff did not yield better results. Again,

Hubbard U	30/300	35/400	40/480	50/600
0 eV	-0.124	-0.063	-0.048	-0.055
5 eV	0.407	0.405	0.396	0.395

Table 1.1: Convergence with respect to plane wave cutoff of energy splitting between a sextet and quartet Int-3 state for both GGA and GGA+U.

it was considered that for some calculations on equilibrium geometries a lower cutoff might be useful to increase the speed of the calculation (e.g. 30/360 is sufficient for resolving some features). However, it was later observed that forces on images computed in the nudged elastic band method were not converged with respect to this cutoff. The forces calculated on each self-consistent image iteration are a component of the force which iterates the path. If the forces on each image are not converged with respect to cutoff then the transition state geometry could not be sufficiently converged as well. Therefore, calculations are currently being repeated in order to be certain that forces on transition states are converged with respect to higher cutoffs.

1.5 Quantum numbers in atoms and molecules

Fundamental for the description of open shell $3d$ transition metal complexes is an understanding of the underlying combination of single particle orbitals which interact and give rise to collective electronic states, as well as the ways in which first principles approaches can describe such states. Since $3d$ valence molecular orbitals in transition metal complexes closely reflect that of the lowest-lying states of the isolated ion, it is key to understand first the electronic structure of transition metal atoms and then to later describe rules for molecular bonds formed to these atoms. The electrons of an isolated atom each have orbital and spin angular momenta which interact and couple to each other. The relative strength of different types of coupling varies depending upon the atomic number of the element. For the elements considered in this thesis, up to the first row of transition metals and lighter, the Russell-Saunders coupling scheme predominates[22].

1.5.1 Atomic term symbols from Russell-Saunders coupling

The Russell-Saunders coupling scheme is based upon the assumption that the spins of the electrons in the atom couple most strongly. That is, the total resultant spin quantum number, S, arises from adding the individual m_s for each electron. The second strongest interaction is the coupling between the orbital angular momenta of the two electrons, creating a total orbital angular momentum of the atom which is L. Finally, the next strongest interaction, spin-orbit coupling, is the coupling between spin and orbital momenta of an electron which gives rise to a total angular momentum quantum number J. In an alternative to Russell-Saunders coupling known as j-j coupling, the total angular momenta of each electron is coupled to another and L and S are no longer meaningful quantum numbers. The building up principles of

Russell-Saunders coupling are illustrated in Fig. 1-6. Term symbols for the atom are then written in the form ${}^{2S+1}L_J$, where $2S+1$ is often referred to as the multiplicity of the atom, L is the total angular momentum of the atom, and J is the coupling of L and S. The symbols for total orbital angular momentum are related to those for isolated electrons:

$$L = S(0), P(1), D(2), F(3), G(4), H(5), ... \tag{1.35}$$

with most low energy term symbols corresponding to $L = 4$ or less. In the illustration, the two electrons with $l = 2$ and $l = 3$ are vectorially summed to yield a total orbital angular momentum of 4, or G. The multiplicity based on the two parallel spins giving $S = 1$ is 3, which is also referred to as a triplet. The total angular momentum, J, is determined as the vectorial sum of L and S, which, in this case is 4. The term symbol for this two electron atom example would thus be written as ${}^3G_4$.

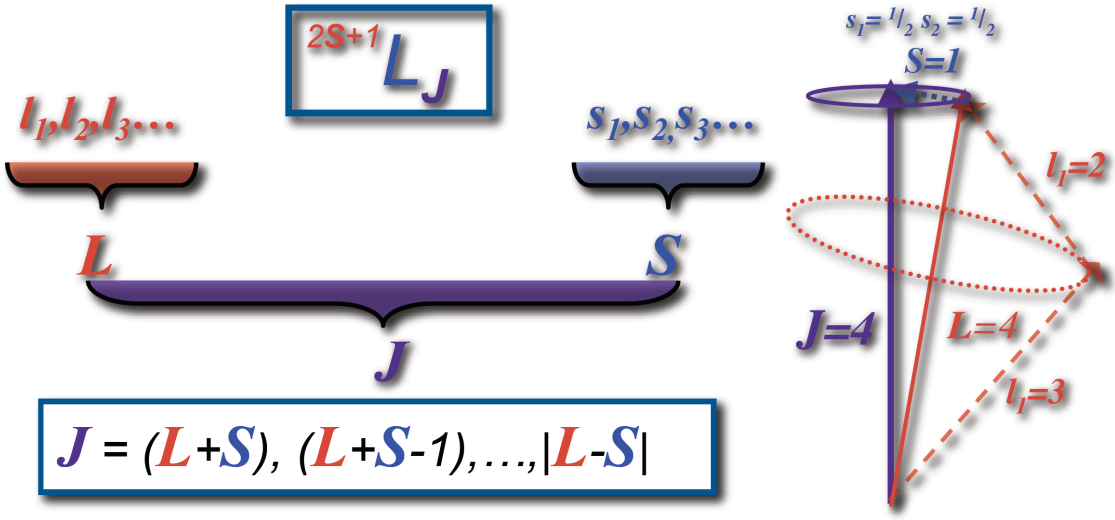


Figure 1-6: Building up principles for atomic term symbols in terms of isolated electron quantum numbers in the regime of Russell-Saunders coupling.

Building up principles for diatomic molecules have been in the literature for some time, and many mirror the common approaches for describing the total angular momentum of an atom[22, 33]. Many of these approaches assume an adiabatic limit in which there is an infinitely slow change of the internuclear distance, and the electrons respond and fully equilibrate to that distance. A diatomic molecule differs from an atom in that the electric field of an atom is spherically symmetric, while the symmetry of the electric field of a molecule is reduced to only axial symmetry about the internuclear axis. As a consequence, the constant of motion is the projection of the net orbital angular momentum, M_L, of all electrons onto the internuclear axis. In this case, the direction of motion of all electrons does not change the energy and only changes the sign of M_L, and the electric field of the internuclear axis does not split the energy of states differing only in sign. Only the different values of $|M_L|$ have different energy, and the orbital angular momentum term most often used, therefore, is:

$$\Lambda = |M_L| = 0, 1, 2...L \tag{1.36}$$

where Λ can range from zero to L. The term symbol corresponding to Λ is commonly given as a Greek letter:

$$\Lambda = \Sigma, \Pi, \Delta, \Phi... \tag{1.37}$$

which is similar to but not equivalent to the letters assigned to the total orbital angular momenta of atoms. Analogously to atoms, the net spin of the molecule is determined from a vectorial sum of the individual electrons' spins. The value of M_S for molecules is:

$$\Sigma = M_S = S, S-1, S-2, ... - S \tag{1.38}$$

usually referred to as Σ and may have up to $2S + 1$ values. The total angular

momentum of the electrons, therefore, is a sum of the spin and orbital terms or:

$$\Omega = |\Lambda + \Sigma|. \tag{1.39}$$

As long as Λ does not correspond to a $M_L = 0$ state, there are $2S + 1$ splittings of the molecular state, and this number is often referred to as a multiplicity. As is the case for atoms, this multiplicity is used as a superscript in the term symbol, and the value of $\Lambda + \Sigma$, analogous to J for atoms, is added as a subscript. The multiplicity is restricted by the total electron count further – that is, molecules with an even number of electrons have odd multiplicities (singlet, triplet, quintet, septet, and so on), while molecules with an odd number of electrons have even multiplicities (doublet, quartet, sextet, octet, and so on). In addition to the angular momenta, the symmetry of the wavefunction is also key. In particular, Σ states are referred to as either Σ^+ or Σ^-, where the $+$ refers to a wavefunction which when reflected at a plane passing through both nuclei remains unchanged, while $-$ indicates that the sign of the wavefunction is changed upon reflection. For homonuclear diatomic molecules, a center of symmetry exists as a consequence of the field remaining unaltered by reflection of the nuclei at the midpoint of the internuclear axis. Electronic eigenfunctions which change sign when reflected at the center are referred to as ungerade (u), while those that stay the same are gerade (g). The total symmetry of the molecule may be determined by multiplying the u's and g's of individual orbitals, where gerade behaves like the number +1, while ungerade is -1. Therefore, a two electron H_2 molecule occupying orbitals with $l = 0$ and $l = 1$ both in spin up has a term symbol $^3\Pi_u$.

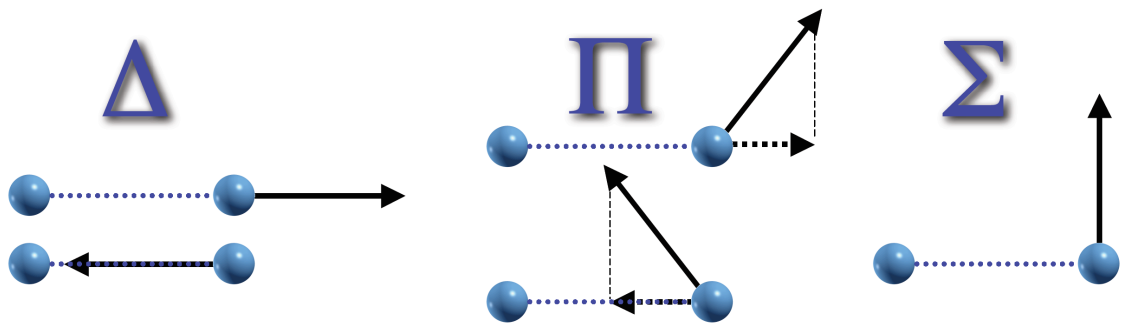


Figure 1-7: Building up principles for molecular term symbols in terms of isolated atom angular momenta, S and D, in the regime of Russell-Saunders coupling.

1.5.2 Building up principles for term symbols

A discussion of the term symbols is key to understand how the relatively high spin transition metals can form diatomic molecules which have a large number of states within a relatively small energetic splitting. For a molecule, the possible electronic states may be determined in a number of ways: the molecule may be built up by bringing together the two separate atoms and using their individual atomic states to identify molecular states, we may imagine a zero nuclear separation case in which the total electron count of the molecule is used to identify an atom with the same number of electrons as the molecule and it is then split, and finally, orbitals may be sequentially occupied at reasonable internuclear distances, and the net momenta of these orbitals may be used to determine the molecular state. The first type is often referred to as an infinite separation case, and the second is referred to as the united atom. We'll consider first and primarily the infinite separation case, as it is the most useful technique when only the experimental ground states of the isolated atoms are known. Wigner and Witmer derived from quantum mechanics rules for determining which molecular states arose from separated atoms, and these rules are

rigorously true for the Russell-Saunders coupling limit, which applies to all cases we consider[22]. The quantum number of the molecule, Λ, is equal to the sum of the orbital angular momenta of the separated atoms:

$$\Lambda = |M_{L_1} + M_{L_2}| \tag{1.40}$$

for atoms 1 and 2. The differing orientations of L_i of the atoms give rise to different values of Λ, and these correspond to different energies in the electric field of the molecule. For example, atomic states with character S ($L_1 = 0$) and D ($L_2 = 2$) may combine to yield $\Lambda = \Sigma, \Pi$, or Δ (see Fig. 1-7). As the atomic angular momenta of the two atoms increase, the number of molecular term symbols also increases dramatically, as indicated in Table 1.2.

	g + g or u + u	g + u
S+S	Σ^+	Σ^-
S+P	Σ^- Π	Σ^+
S+D	Σ^+ Π Δ	Σ^-
S+F	Σ^- Π Δ Φ	Σ^+
P+P	$\Sigma^+_{(2)}$ Σ^- $\Pi_{(2)}$ Δ	Σ^+ $\Sigma^-_{(2)}$
P+D	Σ^+ $\Sigma^-_{(2)}$ $\Pi_{(3)}$ $\Delta_{(2)}$ Φ	$\Sigma^+_{(2)}$ Σ^-
P+F	$\Sigma^+_{(2)}$ Σ^- $\Pi_{(3)}$ $\Delta_{(3)}$ $\Phi_{(2)}$ Γ	Σ^+ $\Sigma^-_{(2)}$
D+D	$\Sigma^+_{(3)}$ $\Sigma^-_{(2)}$ $\Pi_{(4)}$ $\Delta_{(3)}$ $\Phi_{(2)}$ Γ	$\Sigma^+_{(2)}$ $\Sigma^-_{(3)}$
D+F	$\Sigma^+_{(3)}$ $\Sigma^-_{(2)}$ $\Pi_{(5)}$ $\Delta_{(4)}$ $\Phi_{(3)}$ $\Gamma_{(2)}$ H	$\Sigma^+_{(3)}$ $\Sigma^-_{(2)}$

Table 1.2: Building up principles for molecular term symbols in terms of isolated atom angular momenta in the regime of Russell-Saunders coupling.

As the orbital angular momentum of each of the separated atoms increases, the possible number of molecular term symbols also increases, with a combination of D and F states giving rise to 7 different values for Λ which are split further into 20 states non-degenerate in energy. The multiplicity of molecular states may also be determined, assuming Russell-Saunders coupling. The two spins of the separated atoms, $\mathbf{S_1}$ and $\mathbf{S_2}$, add together to form a resultant $\mathbf{S}$:

$$S = S_1 + S_2, S_1 + S_2 - 1, ...|S_1 - S_2| \tag{1.41}$$

and the multiplicity of the molecule is given by $2S+1$. The molecules derived from separated singlet atoms only give rise to singlets, and a singlet combined with another multiplicity always retains the multiplicity of the second atom. For cases where both atoms are high spin, however, more complicated combinations arise. For example, a quartet and triplet give rise to doublets, quartets, and sextets. In realistic transition metal containing molecules, the transition metal often has several high spin, low-lying states which each combine with the multiplets of the other atom to yield several dissociation limits which each in themselves have several potential spins and angular momentum term symbols. The energetic splittings between each state, especially for states derived from the same dissociation limit, are quite small for these systems, and first principles approaches which introduce errors that affect different term symbols to varying degrees can yield the wrong energetic ordering of these states.

The previously described method in which a molecule's spectroscopic term symbol is determined from combinations of those of the isolated atoms is useful, particularly when no electronic structure calculations have been carried out. There is, however, the opposite possibility, that an electron density has been determined for a given molecule, spin, and interatomic distance, but its term symbol has not been deter-

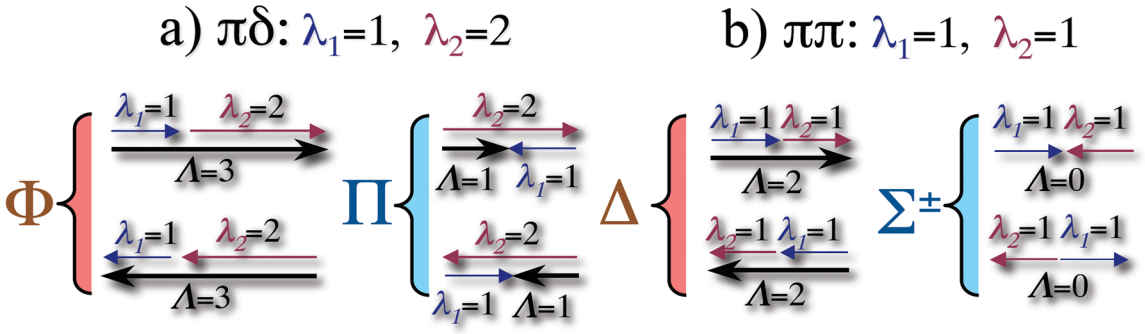


Figure 1-8: Building up principles for molecular term symbols from the two isolated electrons orbital angular momenta, λ_1 and λ_2. At left in case a, $\lambda_1 = 1$ and $\lambda_2 = 2$ give rise to Φ and Π molecular states, while, at right in case b, $\lambda_1 = \lambda_2 = 1$ give rise to Δ or $\Sigma^{\pm}$ molecular states.

mined by the electronic structure code. In this case, an alternative approach in which we assign the angular momenta of electrons individually and then combine them to determine molecular term symbols, is ideal. In a molecule, the individual electrons have only one good quantum number, λ, the orbital angular momentum for all interatomic distances. In the very short interatomic distance, the united atom's n and l become relevant, while at very long distances, the good quantum numbers are n and l of each of the separated atoms. The electron's orbital angular momentum at moderate interatomic distances takes on several values:

$$\lambda = \sigma, \pi, \delta, \phi... \tag{1.42}$$

Typically these are enumerated by their energetic ordering which is typically closely related to atomic orbital orderings, following linear combination of atomic orbital - molecular orbital (LCAO-MO) theory. The λ of orbitals may be added:

$$\Lambda = \sum \lambda_i \tag{1.43}$$

in order to determine the total molecular orbital angular momentum. In instances where the complete manifold of spin up and spin down orbitals is occupied, such as two equivalent electrons in σ orbitals, the net contribution to the molecular orbital angular momentum, Λ, is zero, giving rise to a Σ^+ state. Therefore, only valence electrons need to be considered, and amongst those electrons, only the ones in which there is partial occupation need to be added to determine Λ. In Fig. 1-8, we show the molecular term symbols which may arise from molecular occupation of $\pi\delta$ or π^2. We see that the λ_1 and λ_2 may be added with the same sign to give, in the first case, $\Lambda = 3$ or Φ, or the two may be of opposite sign giving $\Lambda = 1$ or Π. The spatial orbitals used in most codes, such as π_{xz}, correspond to linear combinations which either add or subtract spherical harmonics. By convention, those which add the spherical harmonics are assigned positive λ, while those which subtract are assigned negative λ. This permits differentiation between Π and Φ states, as we consider them in case a in Fig. 1-8. In the second case, π^2, if the electrons possess the same spatial symmetry, they can add to produce a Δ state. If the signs of the π orbitals occupied are opposed, then they cancel to produce a $\Sigma^{\pm}$ term, not to be confused with the Σ^+ term which arises from π^4 configurations. It is in fact quite common that term symbols need to be manually assigned from electron configuration even though most localized basis set electronic structure codes usually aim to determine it directly. The space group of diatomic molecules, $D_{\infty h}$, is not employed in most electronic structure codes which instead use D_{2h}. The resulting lower symmetry gives rise to degeneracy between different orbitals and makes it challenging to identify the atomic term symbols and, in turn, the molecular term symbols.

Several theories exist to explain elementary molecular bonding in order to determine which of the hundreds of possible electronic states for a given ionic configuration is most likely to be stable and observed experimentally. It is of particular note that

unstable states, ones which lack a local minimum in the potential energy curve, are still real states that can perturb other states which do possess minima. Early theories of bonding, such as that of Heitler-London, focused on exchange energy in cases like H_2 where antiparallel alignment of the electrons stabilized an electronic state, while parallel spins produced an unstable electronic state. More advanced theories typically rely on the valence molecular orbitals as having bonding or anti-bonding character which either add to or subtract from the stability of a bond. By utilizing Russell-Saunders coupling, the valence molecular orbitals may be used as a starting point for identifying the total angular momentum of a given state and its relative bond order and stability. Such an approach is useful over a relatively large number of interatomic distances, but its applicability to points near a dissociation limit is suspect. In particular, at relatively long bond lengths, avoided crossings with weakly bound or nearly dissociative states can give rise to unusual dissociation limits or the creation of minima in formerly fully dissociative states, as we will now show.

1.5.3 Interactions between electronic states

In the adiabatic limit, von Neumann and Wigner[33] showed that two electronic states of the same species cannot across each other but instead will always "avoid" each other. Depending upon the energies of two states at both their equilibrium interatomic distances as well as at the infinite limit, the effect the avoided crossing can have on the two potential energy curves varies (see Fig. 1-9). Within lower-order approximations, the two curves may be found to intersect (as shown in purple in Fig. 1-9). However, for most more accurate treatments, the two curves avoid each other, and the perturbation to the apparent energy curves is limited. This case is most often found when two bound molecular states are close in energy to each other because

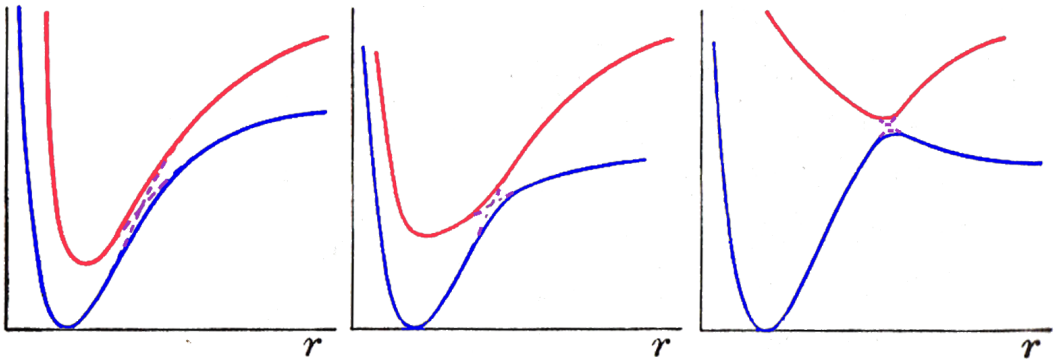


Figure 1-9: Three figures denoting avoided crossings between a lower energy state (in blue) and higher energy state (in red) as well as their crossing curves (in purple) from lower order approximations, modified from [22]. In the leftmost graph, the two curves nearly cross but only barely, while in the middle, the interaction clearly alters the dissociation energies, and in the last case, a dissociative curve obtains a minimum, while the lower energy curve exhibits a maximum in the potential.

the one with the higher energy minimum comes within the region of the dissociation energy of the other lower state. For some intermediate cases, the intersection is such that a weakly bound state crosses a strongly bound state, forming a stronger minimum in the higher energy curve and a lower dissociation energy in the lower energy curve. Lastly, in the most extreme case, the avoidance of the two curves is very strong, and a completely repulsive state crosses a strongly bound state, creating a maximum in the lower energy state and a bound minimum in the higher energy state. This potential maximum has been observed experimentally through pre-dissociation studies (e.g. in AlH). Avoided crossings in diatomic molecules are illustrative of phenomena for all bond breaking events in larger complexes. Additionally, it is key to consider a full manifold of low-lying electronic states with first principles methods in order to identify whether such avoided crossings occur and use them

to identify which values are most likely to correspond to experimental properties including dissociation energies. Overall, a purely theoretical consideration of the manifold of electronic states which could form upon the uniting of a transition metal atom with another atom has already shown that a crucial component of difficulty in studying transition metal chemistry is the sheer number of low-lying states which must be considered.

Chapter 2

Diatomic Molecules

Diatomic molecules formed by transition metals and single atom ligands are the simplest and most fundamental representatives of chemical bonding in much larger catalytic complexes. The previous chapter introduced the concept that the high spin and orbital angular momentum of most isolated transition metal atoms will inherently give rise to a large manifold of low-lying molecular states. We consider more deeply here the practical implications of this inherent challenge in studying transition metal- dimers, hydrides, oxides, and fluorides. Expansive studies of transition metal diatomic molecules have been completed with a variety of hybrid functionals and meta-GGAs in recent years[34, 35, 36, 37, 38], but this study of transition metal molecules with a GGA+U approach is the first of its kind. Of the molecules considered here, the iron dimer represents a particularly challenging case as several states of the isolated Fe atom are within 1 eV of each other, and they are all high-spin and high-angular momentum, giving rise to high spin states such as nonets (S=4) and dodecets (S=5). However, direct metal-metal bonds are not common in catalytic systems; multiple-metal centers in such systems are typically bridged with another

atom like oxygen or sulfur. In order to examine bonds we expect to be the building blocks of catalytic mechanisms, we look instead primarily at the bonds transition metals form with organic ligands. The hydrides are useful for understanding the role that $4s$ electrons play in transition metals, as the majority of bonding in these systems occurs via M$4s$-H$1s$ interactions. These systems also require the introduction of a Hubbard U term on the $4s$ manifold, termed U_{4s}, which we discuss in detail. Transition metal oxides are the most canonical representatives of key covalent bonds in reaction intermediates, as high-spin metal-oxo bonds are expected to participate in a number of key catalytic cycles which are discussed in more detail in Chapters 3 and 5. Finally, we contrast the oxide and hydride cases with the ionic transition metal fluorides. The ionic nature of the fluorides diminishes the role and necessity of a U term, since integer electron donation rather than fractional electron sharing predominates in ionic bonds. Lastly, the size and relative simplicity of diatomic molecules makes them particularly ideal to study nuances of the method. The differences in isoelectronic compounds, pseudopotential dependence of the U, and numerical stability of the method are all considered here through the study of diatomic molecules.

2.1 The iron dimer

Transition metal dimers and small clusters present a formidable challenge for both state of the art experimental and theoretical techniques. Yet these systems are of fundamental importance in our understanding as both a model for di-metal enzyme active sites[3] and as building blocks of nanoparticle catalysts which are highly relevant for chemical vapor deposition nanotube synthesis and growth[39]. For nearly thirty years the ground state of neutral Fe_2 has been identified as $^7\Delta_u$ ($3d^{14}$: $\sigma_g^2\pi_u^4\delta_g^3\pi_g^{*,2}\delta_u^2\sigma_u^{*,1}$, $4s^2$: σ_g^2) in numerous theoretical studies [40, 41, 42, 43, 44, 45, 46,

47, 48, 49] and some experimental studies[50, 51]. However, nothing about the configuration, structural, or vibrational properties of this electronic state is consistent with available experimental findings. Many later studies of the iron dimer have been based upon early theoretical studies which were plagued by both computational and technical limitations. Early density functional studies lacked accuracy in exchange-correlation functional approximations which may have over-stabilized certain electronic states[42, 44]. A trade-off between basis set size and active space size in early configuration interaction studies limited the accuracy of calculated relative energetics of the hundreds of states of various spins and symmetries considered[47, 48]. Despite the obvious need to revisit and to reconsider the discrepancy between experimental and theoretical findings of the past 30 years, recent comprehensive studies of transition metal diatomic molecules[52, 49, 37] continue to assume the $^7\Delta_u$ assignment to be the correct one for the ground state of iron dimer. There have, however, been some recent high-level quantum chemistry calculations - multi-reference configuration interaction (MRCI)[53] as well as large basis-set, single-reference coupled cluster calculations[54] which provide evidence for a more consistent alternative ground state $^9\Sigma_g^-$ for the iron dimer. All of this suggests that GGA+U, a method already successful at improving calculations on correlated-electron transition metal solids, will also provide a comprehensive assessment of the low-lying states of the iron dimer. A GGA+U approach is particularly appealing because it can also be extended to much larger systems due to its favorable scaling in comparison to much more limited post-Hartree-Fock approaches.

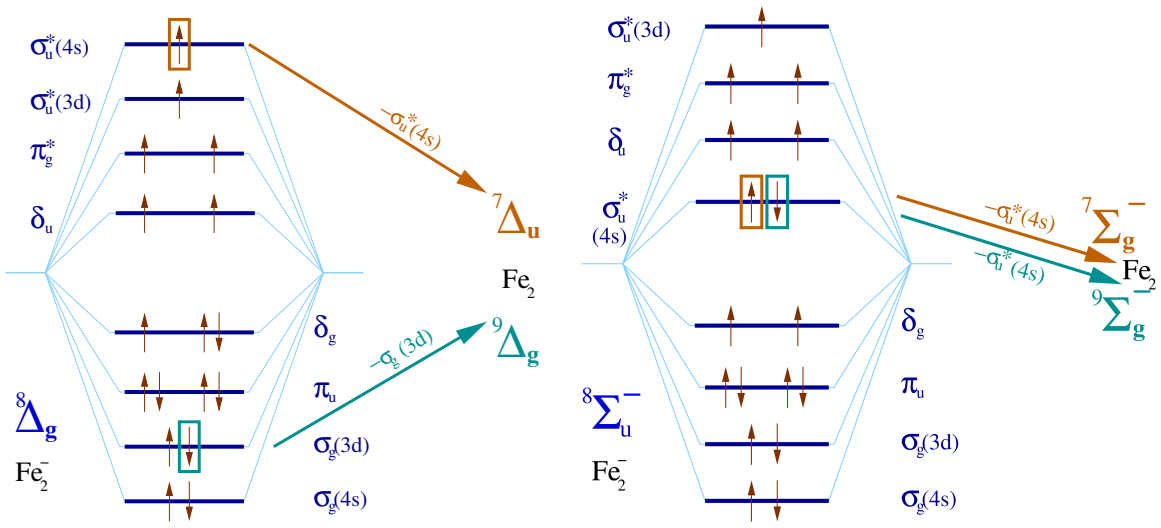
While experimental results are far from complete, they have provided significant clues which help to limit the scope of potential electronic states for the low-lying states of the iron dimer. Several spin multiplicities are good candidates for the dimer; these include anything from an antiferromagnetically coupled singlet to a

high-spin nonet. The large variety of states is due to the close spacing of the two lowest states of the neutral iron atom. The ground state of the Fe atom, ^{5}D (valence electrons: $3d^6 4s^2$), is experimentally only 0.87 eV lower in energy with respect to the excited ^{5}F state with $3d^7 4s^1$ configuration. Additionally, the lowest triplet Fe state, ^{3}F $3d^7 4s^1$, lies only 1.50 eV above the ground state. While states derived by unifying some combination of these three lowest symmetries are likely the best candidates for the lowest-lying states of the iron dimer, this still leaves a large number of possible electron configurations to consider varying from $3d^{14}4s^2$ to $3d^{12}4s^4$ configurations of any number of spin multiplicities. Stern-Gerlach experiments have shown the dimer to have a relatively large magnetic moment, thus ruling out an anti-ferromagnetic state[55], and interpretations of Mössbauer hyperfine interactions confirm the likelihood of a high spin state[56, 57]. Experimental Mössbauer results also indicated a $4s$ occupation per atom of 1.47 in the dimer by comparing isomer shift results with those for the iron atom which has a known ground state configuration of $3d^6 4s^2$[58], but it is worth noting that $4s^3$ configurations have been largely overlooked in nearly all theoretical studies. Both the bond length and frequency have been measured experimentally in rare-gas (both Ne and Ar) matrices and found to be 2.02 Å and 300 cm^{-1}, respectively[59]. This low fundamental frequency suggests a fairly weak bond order, also in contradiction with the proposed $^7\Delta_u$ state which has been suggested to have a formal bond order of three[60].

The piece of experimental evidence which provides the most information about the relative energies of electronic states is the photoelectron spectrum of Fe_2^-. This spectrum is remarkably simple, with just two peaks corresponding to transitions to Fe_2 [51]. This simplicity is also in direct opposition to the myriad of low-lying iron dimer electronic states which have been studied theoretically[47, 48]. Experimental photoelectron spectra typically access only the bound electronic states which differ

from the anion by a single electron detachment. It has been argued that strong electron correlation in systems such as the iron dimer could permit access to neutral states which possess greater differences from the anionic dimer in electronic structure. However, as this possibility is far less likely, we consider here first the theoretical results on states which are derived by a single electron loss from the anion. The experimental photoelectron spectrum of Fe_2^- displays two prominent peaks, one at 1.0 eV and the other 0.53 eV above the first peak. Structurally, the neutral states were found to have identical properties within the error of the experiment (r_e=2.02 ± 0.02 Å and ω_e=300 ± 20 cm^{-1}) and are more bound than the anion which has a bond length of 2.10 ± 0.04 Å and a vibrational frequency of 250 ± 20 cm^{-1}, indicating an increase in bond order from the anion to the neutral dimer. By studying both the low-lying states of the anion and the corresponding allowed neutral states of the neutral dimer, a recent MRCI study has identified an excitation scheme in which the lowest-lying anion state is $^8\Sigma_u^-$ and the two neutral peaks correspond to $^9\Sigma_g^-$ and $^7\Sigma_g^-$ respectively. This MRCI study also finds the previously proposed neutral iron dimer ground state, $^7\Delta_u$, to be 0.7 eV above the true ground state of the dimer. While both the alternative states $^9\Sigma_g^-$ and $^7\Sigma_g^-$ have been considered previously, less accurate methods previously found these states to be significantly excited with respect to $^7\Delta_u$[49, 48, 58].

In order to build a GGA and GGA+U description of the states present in the experimental photoelectron spectrum, we confirm the lowest lying states of the anion to be $^8\Delta_g$ ($3d^{14}$: $\sigma_g^2\pi_u^4\delta_g^3\pi_g^{*,2}\delta_u^2\sigma_u^{*,1}$, $4s^3$: $\sigma_g^2\sigma_u^*$) and $^8\Sigma_u^-$ ($3d^{13}$: $\sigma_g^2\pi_u^4\delta_g^2\pi_g^{*,2}\delta_u^2\sigma_u^{*,1}$ $4s^4$: $\sigma_g^2\sigma_u^{*,2}$). While $^8\Delta_g$ is the lowest anion state with GGA by 0.52 eV, a self-consistent GGA+U approach favors the $^8\Sigma_u^-$ state by more than 0.38 eV. The difference in state splittings between the two methods is nearly an electron volt, and the GGA+U results are in much better agreement with the previous MRCI study. For all calculations,



Scheme 1. **Scheme 2.**

Figure 2-1: Two proposed excitation schemes for the two lowest neutral states derived from known lowest-lying anion states, $^8\Delta_g$ (left, GGA) and $^8\Sigma_u^-$ (right, GGA+U) by a single electron ionization (indicated by arrows and associated orbitals).

we obtain the Hubbard U from the GGA ground state (U_0) as well as with a self-consistent approach (U_{scf}) and numerically average over all states and structures of interest. Again, We use an average U_{scf} because energies at different U are not comparable, and, thus, to obtain the best relative energetics we must average the U calculated over the relevant species.

Beginning with the GGA anion ground state, $^8\Delta_g$, the two lowest neutral dimer states which differ by a single electron loss are $^7\Delta_u$ and $^9\Delta_g$, respectively. The $^7\Delta_u$ state ($3d^{14}$: $\sigma_g^2\pi_u^4\delta_g^3\pi_g^{*,2}\delta_u^2\sigma_u^{*,1}$, $4s^2$: σ_g^2) is produced through ionization of the spin up $4s$ σ_u molecular orbital and therefore contradicts structural details from the experiments. The $^9\Delta_g$ state ($3d^{13}$: $\sigma_g^2\pi_u^4\delta_g^2\pi_g^{*,2}\delta_u^2\sigma_u^{*,1}$, $4s^3$: $\sigma_g^2\sigma_u^*$) is produced by the loss of the spin down $3d$ σ_g orbital and thus is expected to possess different structural properties from $^7\Delta_u$, also inconsistent with experiments. The electronic

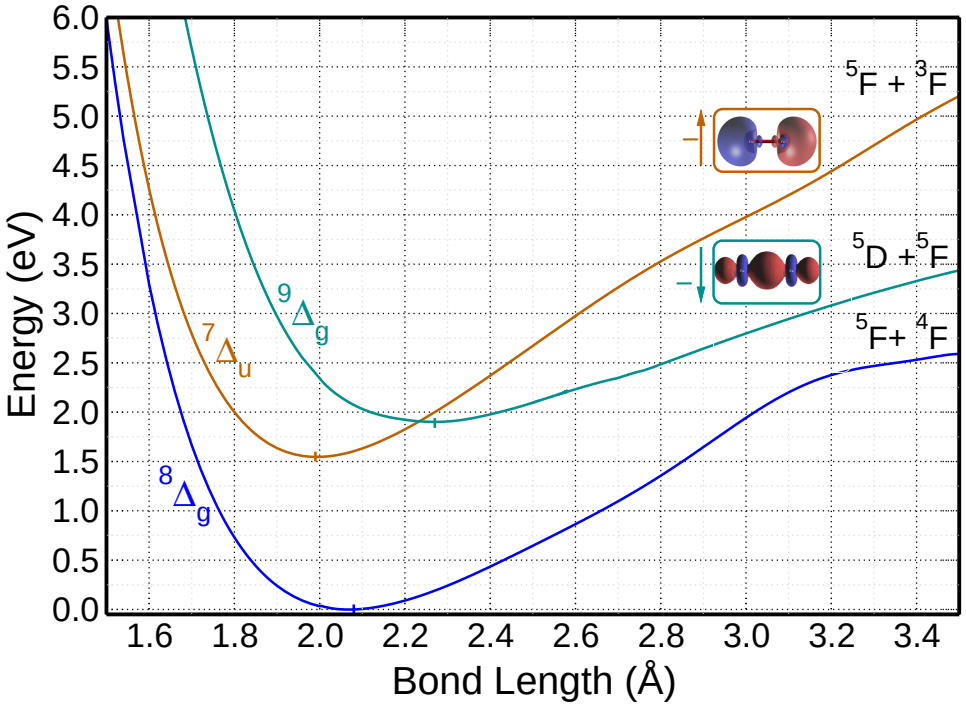


Figure 2-2: Potential energy curves for the proposed GGA ground state anion, $^8\Delta_g$ and the two related low-lying neutral states, $^9\Delta_g$ and $^7\Delta_u$.

configuration of these states is illustrated in Fig. 2-1 and the calculated potential energy curves (see Fig. 2-2) of these states confirm the bond order analysis. By using a Makov-Payne estimate[32] for the true, aperiodic total energy of our charged Fe_2^- system, we are able to estimate the energy required to ionize the $^8\Delta_g$ anion to the neutral state $^7\Delta_u$ to be 1.5 eV with the GGA approach. The GGA ionization energy fails to agree with experiments as does the splitting of the two lowest neutral states (0.40 eV). Additionally, the bond length of $^7\Delta_u$ is decreased (r_e = 1.99 Å) by 0.21 Å with respect to $^8\Delta_g$ (r_e = 2.20 Å) while the $^9\Delta_g$ state exhibits an increased bond length (r_e = 2.26 Å) with respect to the anion. Not surprisingly, the trend in fundamental frequencies for these states also fails to fit our experimental picture, with the $^7\Delta_u$ frequency of 413 cm^{-1} being too markedly increased from the anionic

State	B3LYP	GGA	$+U_0$ (2eV)	$+U_{scf}$ (3eV)	CCSD(T)	MRCI[a]
$^8\Sigma_u^-$	0.00	0.00	0.00	0.00	0.00	0.00
$^8\Delta_g$	0.14	-0.52	0.04	0.38	0.40	0.45
$^9\Sigma_g^-$	0.00	0.00	0.00	0.00	0.00	0.00
$^7\Sigma_g^-$	0.34	0.65	0.66	0.60	0.55	0.62
$^7\Delta_u$	0.18	-0.12	0.48	0.72	0.86	0.69
$^9\Delta_g$	0.36	0.28	0.36	0.41	0.38	0.45

Table 2.1: Spin splittings (in eV) for Fe_2^- and Fe_2 at several levels of theory. (a) Ref. [53].

State	GGA	GGA+U_{scf}	CCSD(T)	MRCI[a]	Expt.[b]
$^8\Sigma_u^-$	2.20, 305	2.20, 301	2.24, 276	2.23, 272	2.1, 250
$^8\Delta_g$	2.07, 360	2.08, 355	2.12, 321	2.4, –	–
$^9\Sigma_g^-$	2.11, 339	2.13, 335	2.17, 296	2.18,299	2.0,300
$^7\Sigma_g^-$	2.10, 335	2.12, 331	2.16, 304	2.17,310	2.0,300
$^7\Delta_u$	1.99, 413	2.00, 419	2.00, 404	2.25,195	–
$^9\Delta_g$	2.26, 285	2.26, 280	2.28, 220	2.35, –	–

Table 2.2: Bond lengths, r_e (Å), and harmonic frequencies, ω_e (cm^{-1}), for Fe_2^- and Fe_2, compared to experiment (here, fundamental frequencies, ω_0). (a) Ref. [53]. (b) Ref. [51].

ω_e of 360 cm^{-1}, while $^9\Delta_g$ shows a reduced frequency of 285 cm^{-1}.

The GGA+U approach provides a much more consistent prediction of the excitation scheme both with respect to experiment and the highly-accurate MRCI results already published and is a remarkable improvement over GGA. Beginning with the GGA+U ground state anion $^8\Sigma_u^-$ ($3d^{13}$: $\sigma_g^2\pi_u^4\delta_g^2\pi_g^{*,2}\delta_u^2\sigma_u^{*,1}$ $4s^4$: $\sigma_g^2\sigma_u^{*,2}$), the two lowest states which differ from ionization of a single electron are $^9\Sigma_g^-$ and $^7\Sigma_g^-$, in accordance with previous findings in the MRCI study [53]. The two states have nearly identical valence electron configuration (both are $3d^{13}4s^3$) but differ simply by the spin of the $4s$ σ_u^* orbital ionized, spin down in the case of the nonet and spin up in the case of

the septet as shown in Figure 2-1. Therefore, we may expect that both bond length and frequency of the two neutral states should be nearly identical to each other. The loss of the anti-bonding orbital suggests a decrease in bond length and increase in frequency with respect to the anion, consistent with experiment. To obtain the most accurate relative energetics of the three states, we use an average U_{scf} of 3 eV which, we note, provides improved energetics with respect to U_0 from GGA, which is only 2 eV. The potential energy curves of this system, shown in Fig. 2-3, are consistent with experiment. The ionization energy of the $^8\Sigma_u^-$ state is approximately 1.05 eV upon inclusion of the Makov-Payne correction. The splitting of the two lowest neutral states is roughly 0.6 eV, in good agreement with the experimental splitting of 0.53 eV. The trend in structural properties is also consistent with experiment. The two neutral states exhibit nearly identical structural properties (see Table 2.2). The bond length, r_e, is observed to decrease by 0.08 Å from the anion to the neutral states (from 2.20 Å to 2.12 Å), while the fundamental frequency is observed to increase by roughly 35 cm^{-1}. While the absolute values of the bond length and frequency are not identical to experiment, we note that these discrepancies are likely due to 1) experimental uncertainty and 2) the anharmonic effects which cannot be fully eliminated in the experimental measurement of ω_0 but are excluded in our theoretical measurement of the harmonic frequency.

One additional piece of experimental evidence worth considering is the bond strength, or dissociation energy of the anion and neutral iron dimers. While the bond order appears to nominally increase from the anion to the neutral states, corresponding to a shorter bond length, the experimental bond strength decreases roughly from 1.53 ± 0.17 eV for the anion to 0.78 eV ± 0.17 eV for the lowest neutral state. Upon first glance of Figure 2-3, it appears that the bond strength, roughly 1.5 eV for the anion is in agreement while the neutral state demonstrates a much higher bond

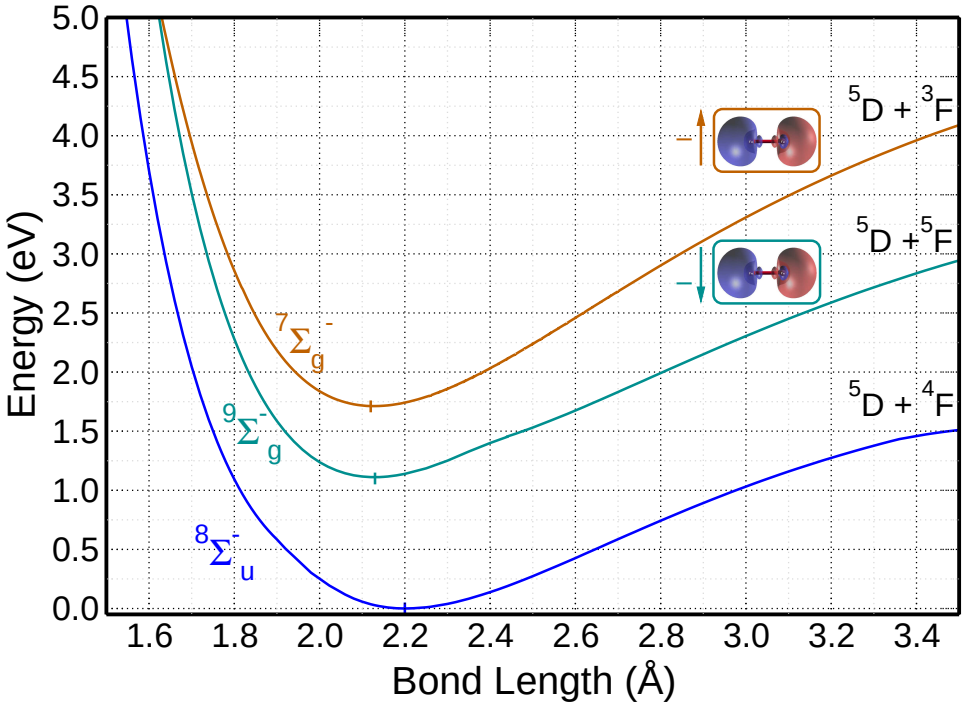


Figure 2-3: Potential energy curves for the proposed GGA+U ground state anion, ${}^8\Sigma_u^-$ and the two related low-lying neutral states, ${}^9\Sigma_g^-$ and ${}^7\Sigma_g^-$.

strength of roughly 1.9 eV. However, we note that the GGA+U anion state dissociates into a neutral ^{5}D Fe atom and a negatively charged ^{4}F Fe ion, both in their ground state. The proposed dissociation for the lowest neutral state, ${}^9\Sigma_g^-$ dissociates into a ground state neutral ^{5}D Fe and an excited state ^{5}F Fe. Once we incorporate the strong likelihood of an avoided crossing with a nonet state that has a dissociation limit of two ^{5}D neutral Fe atoms, we recover a bond strength of roughly 0.8 eV, in excellent agreement with experiment. Overall, our GGA+U results are in remarkable agreement with both experiment and high-level quantum chemistry in the form of both single-reference CCSD(T) and multi-reference MRCI (see Tables 2.1 and 2.2).

2.2 Transition metal hydrides: the role of U_{4s}

The electronic structure of transition metal hydrides helps to highlight the broader applicability of a DFT+U approach to other valence manifolds aside from $3d$ states. The dominant mode of bonding in the metal hydrides (M and H, respectively) is via M ($4s$)-H($1s$) σ-bonding interactions. The density from the $3d$ manifold plays a more indirect role in bonding through limited M ($3d_{z^2}$)-H($1s$) σ-bonding interactions as well as the significant $4s$-$3d_{z^2}$ hybridization present in σ_{4s} orbitals. All other $3d$ density is highly localized on the metal, and these states closely resemble the isolated atomic density. It follows, therefore, that the U_{3d} of all of these systems is quite low, ranging from a zero value to less than three. In turn, a U_{4s} should be considered to ensure accurate structural and energetic descriptions of the electronic states in question. We consider several prototypical cases: early- (CaH, ScH), mid- (CrH, MnH, FeH), and late- (CuH, ZnH) transition metals in order to illustrate key trends in the hydrides.

The early and late transition metal hydrides share in common low values of U_{3d} (between 0 and 1.5 eV) associated with low levels of σ-bonding from $3d$-$1s$ hybridization and nearly empty or, conversely, nearly filled $3d$ levels. Additionally, the lowest-lying electronic state for the early and late transition metal hydrides are typically well separated by over 2.0 eV and so only one spin state is investigated. On the other hand, the splitting of the lowest electronic states of the approximately half-filled $3d$ manifold cases, CrH, MnH, and FeH, can be very sensitive to the method employed[37, 61, 62, 63]. The overall trends for the value of U_{3d} show a maximum at the mid-row transition metal hydrides, between 1.5 and 3 eV, compared to the empty or filled $3d$ transition metal hydrides (see Fig. 2-4). The value of U_{3d} is most critical for the half-filled hydrides; thus, it is important to note that the linear-

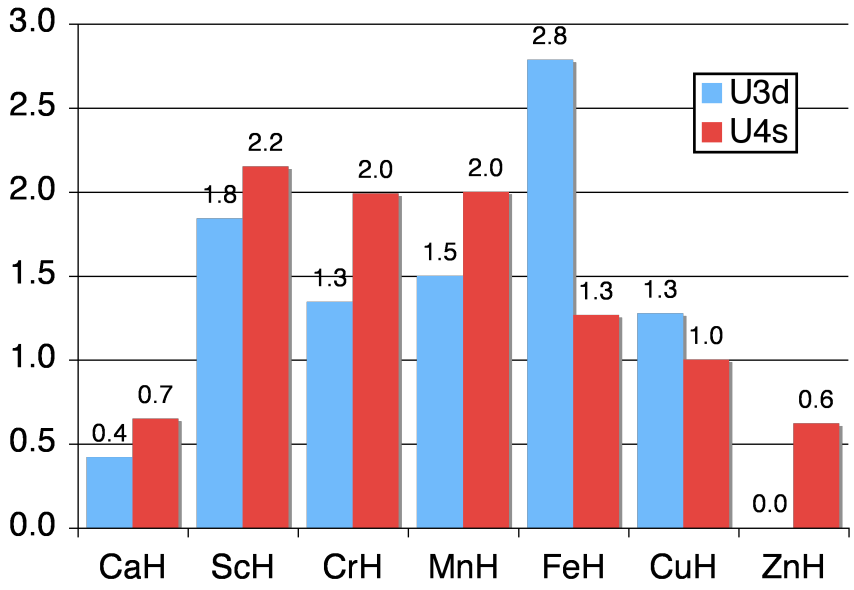


Figure 2-4: Values of Hubbard U (in eV) for several transition metal hydrides ordered with respect to their placement in the periodic table. The values of U_{3d} (blue) and U_{4s} (red) are shown to be maximum for the mid-row transition metal hydrides and decrease for nearly full shell cases, such as ZnH, or nearly empty cases, such as CaH.

response approach we use to assign values of U is key here both for selecting the appropriate projection manifolds and whether a Hubbard U term is even necessary for descriptions in bonding. The values of the U_{4s} exhibit a subtle trend similar to that of the U_{3d} values. The lowest U_{4s}, about 0.6-0.7 eV, is observed at the extrema - CaH and ZnH. Moving towards the center of the periodic table, values of U_{4s} increase from 1.0 (CuH) to over 2.0 eV (ScH, CrH, MnH), in several cases exceeding the value of U_{3d} for that system. Overall, the values of linear-response U on the $3d$ and $4s$ manifolds confirm the intuitive chemical picture that $4s$ bonding interactions are highly relevant, particularly in the mid-row cases where there is a competition between $4s$- and $3d$- derived orbital occupation.

		B3LYP	GGA	GGA+U_{3d}	GGA+$U_{3d/4s}$	CCSD(T)	Expt.[64]
CaH	r_e	1.985	1.970	1.974	1.978	2.020	2.000
$^2\Sigma^+$	ω_e	1292	1295	1287	1284	1296	1298
	D_e	2.00	1.95	1.93	1.83	1.70	1.78
ScH	r_e	1.750	1.760	1.772	1.783	1.787	1.775
$^1\Sigma^+$	ω_e	1621	1603	1591	1596	1601	1547
	D_e	3.40	2.59	2.15	2.04	2.10	2.06
ScH	r_e	1.844	1.844	1.847	1.877	1.900	
$^3\Delta$	ω_e	1460	1481	1446	1453	1467	
	D_e	2.14	2.33	2.16	2.07	2.08	
CrH	r_e	1.632	1.600	1.626	1.620	1.630	
$^4\Sigma^+$	ω_e	1737	1709	1706	1741	1727	
	D_e	1.99	1.96	1.90	1.88	1.91	
CrH	r_e	1.662	1.650	1.664	1.668	1.652	1.656
$^6\Sigma^+$	ω_e	1632	1657	1611	1586	1531	1581
	D_e	2.38	2.16	2.07	1.98	2.05	2.03
MnH	r_e	1.632	1.592	1.600	1.620	1.619	
$^5\Sigma^+$	ω_e	1737	1729	1695	1692	1711	
	D_e	1.99	1.84	1.75	1.69	1.71	
MnH	r_e	1.735	1.714	1.726	1.730	1.740	1.731
$^7\Sigma^+$	ω_e	1507	1546	1530	1527	1552	1548
	D_e	1.70	1.75	1.62	1.38	1.40	1.35
FeH	r_e	1.558	1.560	1.564	1.587	1.570	1.589
$^4\Delta$	ω_e	1734	1916	1853	1824	1829	1827
	D_e	1.81	2.13	2.00	1.78	1.65	1.70
FeH	r_e	1.682	1.670	1.681	1.701	1.685	
$^6\Delta$	ω_e	1581	1626	1565	1539	1577	
	D_e	2.31	1.81	1.73	1.51	1.58	
CuH	r_e	1.484	1.487	1.495	1.500	1.486	1.463
$^1\Sigma^+$	ω_e	1854	1973	1887	1924	1884	1941
	D_e	2.72	2.79	2.68	2.64	2.58	2.63
CuH	r_e	1.585	1.590	1.598	1.598	1.625	
$^3\Pi$	ω_e	1485	1502	1477	1477	1603	
	D_e	0.30	0.49	0.47	0.47	0.73	
ZnH	r_e	1.622	1.605	1.605	1.610	1.605	1.594
$^2\Sigma^+$	ω_e	1512	1543	1543	1503	1527	1603
	D_e	1.00	1.05	1.05	0.95	0.98	1.01

Table 2.3: Structural properties of several transition metal hydrides.

The structural, vibrational, and binding properties of a subset of all first-row, transition metal hydrides were calculated. First, these properties were calculated for the ground state spin and recorded in Table 2.3. Additionally, for cases where a second spin's lowest electronic state resided within 1-2 eV of the ground state, properties were also calculated for this second state (see Table 2.3). For each electronic state, the structural and vibrational properties were calculated at the GGA and GGA+U levels of theory and compared against available experimental values[64]. In addition to the standard approach of applying only a U_{3d}, the combined effect of the linear-response U_{3d} and U_{4s} were considered (referred to hereafter as GGA+$U_{3d/4s}$). For the nearly empty (CaH, ScH) and filled (ZnH) hydrides, the small values of linear-response U correspond to minimal changes in structural and energetic properties when comparing GGA and GGA+U approaches. The equilibrium bond lengths for these three cases are lengthened by under 0.01 Å upon addition of a U term. The harmonic frequencies of these extrema also vary by less than 50 cm^{-1} with addition of the U term. The absolute percentage change in the harmonic frequency is quite small, about 2%. Dissociation energies for the nearly empty and the nearly filled hydrides also vary minimally, usually by less than 0.10 eV. These structural structural results confirm the trend implied by the values of linear-response U: a GGA description of the properties of these systems is sufficient.

2.2.1 Septet and quintet MnH: the effect of U on energetics

For a significant number of the mid-row hydrides, the splitting of the two lowest spin states is quite small because they correspond to alternatively ferromagnetic and anti-ferromagnetic coupling of the hydrogen 1s spin with the lowest energy spin and configuration of the isolated transition metal atom. As the quintessential example,

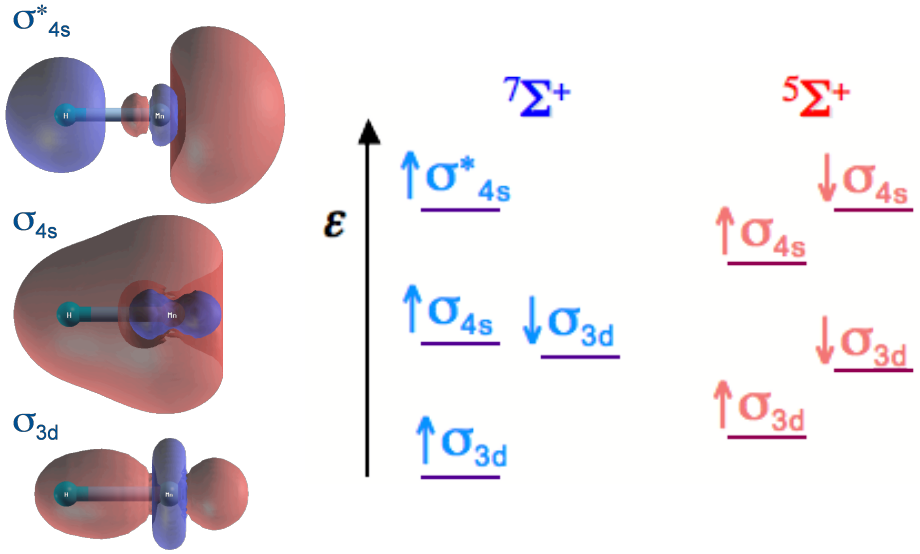


Figure 2-5: The valence orbitals which participate in bonding for the $^7\Sigma^+$ and $^5\Sigma^+$ states of MnH (left) as well as qualitative occupations an ordering of the molecular orbitals in both of these electronic states (right).

we will consider manganese hydride (MnH) in its two lowest-spin electronic states. Isolated Mn atom has a $^6S_{5/2}$ term symbol and corresponds to $3d^5 4s^2$ valence occupation. The hydrogen 1s density can couple to the Mn atom density to produce either a high spin septet state, $^7\Sigma^+$, or, alternatively, a low-spin $^5\Sigma^+$ state. Early density functional calculations have shown a preference for the low-spin $^5\Sigma^+$ state [37, 61]; however, the experimental ground state of MnH has been determined to be $^7\Sigma^+$[65, 66, 67]. The major electronic difference between the two states is that $1s$ density in the high-spin state, $^7\Sigma^+$, populates exclusively spin up $4s$-derived orbitals, including bonding (σ_{4s}) and antibonding (σ^*_{4s}). The quintet preferentially occupies exclusively bonding σ_{4s} molecular orbitals in both the up and down spin channels (see Fig. 2-5).

Initially, one would expect the $^5\Sigma^+$ state to be lower in energy, owing to the tra-

ditional picture of molecular bonding in which bonding orbitals are occupied before the associated anti-bonding orbitals. However, there are several meaningful physical explanations for why the $^7\Sigma^+$ state is in fact the ground state. The difference in bond lengths, Δr_e from the quintet to the septet is only 0.12 Å (from 1.62 to 1.73 Å for GGA+$U_{3d/4s}$). While we expect the bond order to be reduced nominally by one in going from $^5\Sigma^+$ to $^7\Sigma^+$, the σ^*_{4s} state is only weakly antibonding and possesses what appears to be intermediate between canonical non-bonding and anti-bonding character, a feature previously observed experimentally in several photoelectron spectra [68]. It follows, also, that the placement of σ^*_{4s} density on the Mn site minimizes repulsion from the other density owing to the $3d$ density almost as much as the σ_{4s} orbital does via $4s - 1s$ hybridization, but in this case it achieves it by occupying a diffuse orbital away from the molecular bond (see Fig. 2-5). Finally, the presence of H $1s$ density in the minority spin channel for $^5\Sigma^+$ induces charge transfer by Mn $3d$ states into the spin down σ_{3d} orbital where the density had been relatively weak in the case of the septet. This actually reduces the net magnetic moment of the $3d$ states further, increasing the overall energetic cost of forming the $^5\Sigma^+$ state with respect to $^7\Sigma^+$.

With GGA, the $^5\Sigma^+$ and $^7\Sigma^+$ states are nearly degenerate. However, upon inclusion of a $U_{3d/4s}$ term, we recover a splitting of 0.21 eV, in excellent agreement with both CCSD(T) and experiment [65]. The structural properties, on the other hand, change relatively little with the $^5\Sigma^+$ and $^7\Sigma^+$ states exhibiting similar bond elongation (0.01-0.03 Å) from their GGA values of 1.592 Å and 1.714 Å, respectively to 1.620 and 1.730 Å. With a slight increase in bond length, the commonly observed reduction in harmonic frequency for GGA+$U_{3d/4s}$ occurs by about 20-40 cm^{-1}, or roughly 1-2%, from the GGA values. Lastly, the most significant change, associated with the difference in relative energy at the minima, the septet and quintet

dissociation energies decrease from 1.75 eV to 1.38 eV and from 1.84 eV to 1.69 eV, respectively. Only upon inclusion of both the U_{3d} and U_{4s} is the correct splitting of the two states recovered, with a traditional GGA+U_{3d} still underestimating $\Delta E_{7\rightarrow 5}$ by about 50% at 0.13 eV. While the combined GGA+$U_{3d/4s}$ approach modulates structure in these hydrides only very weakly, the relative spin state ordering and proper ground state is only determined with a U term.

2.3 Transition metal oxides

The transition metal oxides are representative of the reactive intermediates in reactions ranging from the addition-elimination formation of methanol from methane[17, 69] to alkane halogenation in enzyme active sites[5]. The highly covalent bonds formed between transition metals and oxygen are concomitant, typically, with significant charge transfer to the oxygen, yielding a short bond. This charge transfer makes the oxygen highly reactive for hydrogen abstraction or insertion into a hydrocarbon. The charge transfer and bonding is a function of the relative contributions of $3d$ electrons into the σ and π molecular orbitals, and a U term plays a key role in tuning the strength of the molecular bond in these systems. In order to better investigate the role of GGA+U in the structural properties of these systems, we consider the lowest two spin and state symmetries of early- (CaO, TiO, VO), mid- (MnO, FeO), and late- (CoO, NiO, ZnO) transition metal oxides.

The values of the linear-response U in both traditional and self-consistent forms for all of the relevant transition metal oxides were calculated (see Table 2.3). For the essentially $3d^0$ configuration of CaO, a zero U_0 on the $3d$ manifold is obtained. In this case, we obtain a small U on the $4s^1$ states which are responsible for the bonding in this molecule. Similarly, the nearly full $3d^{10}$ ZnO molecule exhibits no

	L.S.	U_0	U_{scf}	H.S.	U_0	U_{scf}	$U_{scf,av}$
CaO	$^1\Sigma^+$	*1.25*	*1.51*	$^3\Pi$	*1.45*	*1.62*	*1.57*
TiO	$^1\Delta$	5.01	5.12	$^3\Delta$	5.72	6.03	5.58
VO	$^2\Sigma^-$	4.67	4.82	$^4\Sigma^-$	4.39	4.46	4.64
MnO	$^4\Pi$	3.32	3.53	$^6\Sigma^+$	3.39	3.40	3.47
FeO	$^7\Sigma^+$	3.14	3.15	$^5\Delta$	2.94	3.01	3.08
CoO	$^4\Delta$	3.71	3.73	$^6\Sigma^+$	3.65	3.68	3.71
NiO	$^1\Pi$	3.89	4.06	$^3\Sigma^-$	4.23	4.34	4.20
ZnO	$^1\Sigma^+$	0.00	0.00	$^3\Pi$	0.00	0.00	0.00

Table 2.4: Linear-response values of Hubbard U (in eV) for the transition metal oxides CaO, TiO, VO, MnO, FeO, CoO, NiO, and ZnO in their low spin (L.S.) and high spin (H.S.) states. The standard result, U_0, is shown, as well as the self-consistently obtained U_{scf} for each state and an average value for both states, $U_{scf,av}$. The value $U_{scf,av}$ is used to determine the energetic splittings between the two states. For CaO, the U listed is obtained on the $4s$ manifold, as the $3d$ state occupation is minimal, and this is denoted in italics in the table.

measurable U on the $3d$ states. Intermediate transition metal oxides exhibit values of U between around 3 and 5.5 with the highest values of U being for TiO at 5.6 and VO at 4.7, both of which are relatively early transition metal oxides. The values of the linear-response U are overall quite a bit higher than were determined previously for the transition metal hydride systems, confirming the idea that the strong covalent bonding lends itself to a GGA+U approach.

Structural properties were calculated for the early transition metal oxides (CaO, TiO, VO, MnO) and these results as well as available experimental values are shown in Table 2.3[64]. For these cases, the largest difference between GGA and GGA+U is immediately observed to be the dissociation energies. Using a GGA approach, the minima of the potential energy curves of early transition metal oxides is over-stabilized. The $3d$ electrons are strongly delocalized in the σ and π molecular bonds compared to the isolated atomic $3d$ states, and this makes the GGA energy at the minimum artificially lower than it should be. In turn, this error manifests itself

		B3LYP	GGA	GGA+U_{scf}	CCSD(T)	Expt.[64]
CaO	r_e	1.814	1.816	1.829	1.848	1.822
$^1\Sigma^+$	ω_e	788	770	745	772	732
	D_e	6.00	5.15	4.97	5.92	> 4.80
CaO	r_e	2.078	1.939	1.958	1.991	
$^3\Pi$	ω_e	554	591	560	563	
	D_e	3.28	3.25	2.95	5.02	
TiO	r_e	1.585	1.598	1.617	1.618	
$^1\Delta$	ω_e	1090	976	925	929	
	D_e	4.80	5.92	4.78	5.11	
TiO	r_e	1.612	1.623	1.651	1.626	1.620
$^3\Delta$	ω_e	1029	1040	961	1026	1009
	D_e	7.09	7.35	6.93	7.07	6.98
VO	r_e	1.567	1.548	1.564	1.559	
$^2\Sigma^-$	ω_e	1063	1180	1110	1132	
	D_e	5.25	5.21	4.98	5.02	
VO	r_e	1.579	1.598	1.628	1.582	1.589
$^4\Sigma^-$	ω_e	1045	1053	991	1010	1011
	D_e	6.60	6.75	6.49	6.54	6.50
MnO	r_e	1.604	1.605	1.619	1.609	
$^4\Pi$	ω_e	945	882	863	886	
	D_e	3.33	4.62	3.25	2.97	
MnO	r_e	1.633	1.628	1.654	1.651	1.646
$^6\Sigma^+$	ω_e	900	874	788	831	840
	D_e	4.73	5.46	3.91	3.83	3.88

Table 2.5: Structural properties of early (CaO, TiO, VO, and MnO) transition metal oxides at several levels of theory. Bond lengths (Å), frequencies (cm^{-1}), and dissociation energies (eV) are compared against experimental values, where available.

most typically in overestimates of dissociation energies by as much as 1-2 eV. By comparison, effects on the equilibrium bond length are much more subtle. The largest difference between GGA and GGA+U equilibrium bond lengths is around 0.03 Å, with most being around 0.01-2 Å. In most cases, the GGA number is a slight underestimate of the experimental or CCSD(T) bond length, but the GGA+U value in several cases may be an overestimate. This over-elongation is due to some inherent penalty on all bonding interactions, but this effect is systematic and thus may be understood when interpreting results. The harmonic frequencies decrease from GGA to their GGA+U values by on average 40 cm^{-1}, with the largest decreases being about twice that at 80 cm^{-1}. The experimental fundamental frequencies can only approximately be compared to the harmonic frequencies, ω_e, we measured. However, GGA in general overestimates the ω_e of most of the molecules while the value from GGA+U is a slight underestimate for the same reasons the bond is slightly over-elongated. In these early transition metal oxides, particularly CaO, TiO, and VO, the only molecular orbitals populated are bonding ones. This means that the self-interaction errors present in GGA cooperate with the true chemical hybridization present in the bonding molecular orbitals. We will later show that for oxides where there is a presence of anti-bonding or nonbonding states the GGA results will exhibit much larger errors with respect to experiment. Regardless of the nature of the early transition metal oxides, other more subtle features such as splittings are still quite sensitive to the U value, as we will later show.

The mid- to late- transition metal oxides, FeO, CoO, NiO, and ZnO were studied, and their structural properties are reported in Table 2.3. The two lowest spin and symmetry states were considered for each molecule. As was previously observed, ZnO is nearly $3d^{10}$ and has a zero linear-response U. The bond lengths, frequencies, and dissociation energies, of both the ground state $^1\Sigma^+$ ZnO and the excited triplet

		B3LYP	GGA	GGA+U_{scf}	CCSD(T)	Expt.
FeO	r_e	1.573	1.589	1.605	1.591	
$^7\Sigma^+$	ω_e	988	973	935	949	
	D_e	4.03/6.12	4.01/6.01	3.03/5.06	3.12/5.27	
FeO	r_e	1.611	1.604	1.623	1.618	1.616
$^5\Delta$	ω_e	903	948	895	901	880
	D_e	5.34	5.38	4.35	4.16	4.22
CoO	r_e	1.658	1.652	1.659	1.637	
$^6\Sigma^+$	ω_e	814	764	744	749	
	D_e	1.54	3.84	3.53	3.61	
CoO	r_e	1.591	1.613	1.634	1.643	1.629
$^4\Delta$	ω_e	1001	881	849	825	853
	D_e	4.14	5.03	4.05	3.87	3.99
NiO	r_e	1.758	1.642	1.691	1.670	
$^1\Pi$	ω_e	613	838	654	668	
	D_e	2.55	2.81	2.75	2.75	
NiO	r_e	1.626	1.641	1.653	1.649	1.627
$^3\Sigma^-$	ω_e	884	852	833	818	838
	D_e	3.73	4.85	4.08	4.01	3.92
ZnO	r_e	1.711	1.698	1.698	1.702	
$^1\Sigma^+$	ω_e	767	762	762	763	770
	D_e	3.75	3.96	3.96	3.89	
ZnO	r_e	1.884	1.868	1.868	1.871	
$^3\Pi$	ω_e	537	504	504	523	
	D_e	1.35	1.45	1.45	1.42	

Table 2.6: Structural properties of late (FeO, CoO, NiO, and ZnO) transition metal oxides at several levels of theory. Bond lengths (Å), frequencies (cm^{-1}), and dissociation energies (eV) are compared against experimental values, where available.

are in excellent agreement with CCSD(T) and available experimental values. For the other three, FeO, CoO, and NiO, structural properties are significantly improved with GGA+U_{scf}. For both triplet and quintet FeO, as found previously for sextet MnO, the GGA dissociation energies are an overestimate of CCSD(T) and experimental values by nearly 1 eV, while GGA+U values are within 0.1 eV agreement on average. This trend in improved dissociation energies is also observed for the ground states $^4\Delta$ CoO and $^3\Sigma^-$ NiO, where the GGA+U dissociation energies are within 0.1-0.2 eV of the experimental value while GGA exhibits errors as much as 1 eV. The effect of the U on the bond length is again subtle with the largest change, Δr_e, being about 0.05 Å for singlet NiO, while the average Δr_e is about 0.02 Å. The mid- to late- transition metal oxides have a $3d$ atomic density which has been distributed between both bonding and anti-bonding molecular orbitals. The GGA+U functional indirectly penalizes the hybridization present in bonding orbitals and therefore counteracts the relative destabilization present in GGA of the anti-bonding manifold. In the late transition metal oxides, this is key, and the harmonic frequencies manifest the shift in relative occupancy of bonding and anti-bonding orbitals most significantly. For instance, the harmonic frequency decreases by over 100 cm^{-1} in the case of singlet NiO when comparing GGA against GGA+U. The other states exhibit harmonic frequency decreases on the order of about 50 cm^{-1} as well. Because the relative populations of bonding and anti-bonding orbitals is so critical for these oxides, the relative energies of different states can vary significantly. We will later show that FeO^+, among others, exhibits significant dependence upon the value of U to determine splitting between differing spin states.

	L.S.	U_0	U_{scf}	H.S.	U_0	U_{scf}	$U_{scf,av}$
ScF	$^1\Sigma^+$	2.64	2.70	$^3\Sigma^+$	2.52	2.60	2.65
CrF	$^4\Sigma^+$	0.93	1.94	$^6\Sigma^+$	1.83	2.04	1.99
FeF	$^4\Delta$	2.34	2.60	$^6\Delta$	1.54	1.80	2.20
CuF	$^1\Sigma^+$	1.00	1.85	$^3\Sigma^+$	0.84	1.63	1.74

Table 2.7: Linear-response values of Hubbard U (in eV) for the transition metal fluorides ScF, CrF, FeF, and CuF in their low spin (L.S.) and high spin (H.S.) states. The standard result, U_0, is shown, as well as the self-consistently obtained U_{scf} for each state and an average value for both states, $U_{scf,av}$. The value $U_{scf,av}$ is used to determine the energetic splittings between the two states.

2.4 Transition metal fluorides: the U in partially ionic systems

Several transition metal fluorides ranging from very little $3d$ occupation (ScF) to mid-row $3d$ elements (CrF, FeF) and finally nearly full $3d$ elements (CuF) have been studied. The fluorides are unique because of the molecular bond in these cases is likely to be strongly ionic. Fundamentally, the U term modulates hybridization between the TM center's $3d$ states and the states of the ligand. In the case of a purely ionic limit where an electron is fully transferred from the metal center to the fluoride, no measurable covalent hybridization should be present. Of course, we will show that the transition metal fluorides do not exhibit a purely ionic chemical bond, and therefore, the value of U is non-zero. However, compared to isoelectronic transition metal oxides, the fluorides demonstrate considerably reduced values of linear-response U.

The linear-response values of U were calculated for the two lowest spin states of ScF, CrF, FeF, and CuF and are shown in Table 2.7. The values of U for all the fluorides considered is similar, between 1.75 and 2.75 eV on average with no

discernible trend related to the nature of the transition metal. If any distinction may be made, it should be noted that the value for the nearly fully occupied $3d$ manifold CuF exhibits a slightly lower value of U_0 and U_{scf} on average. The effects that the modest values of U have on structure of the fluorides is limited.

Structural properties were calculated for the transition metal fluorides, ScF, CrF, FeF, and CuF, and are presented along with available experimental results in Table 2.8[64]. Since the values of U_{scf} were small, it follows that the structural changes here are much smaller compared to the oxides. The differences in bond lengths between GGA and GGA+U are subtle. The elongation of the bonds in GGA+U are around 0.01-0.02 Å with the largest being about 0.03 Å. The GGA+U bond lengths do not improve upon GGA with respect to experimental values, and this suggests that for some structural properties of the fluorides a GGA approach is sufficient. Decreases that occur for GGA+U harmonic frequencies with respect to the GGA values are on average about 20-30 cm^{-1} and also do not significantly improve the result when compared against experiment. The dissociation energies, particularly for $^1\Sigma^+$ ScF and $^1\Sigma^+$ CuF, are reduced by 0.3-0.8 eV in improved agreement with experimental numbers. Overall, the ionic character of these molecules means that the issues in covalently bonded molecules such as partitioning between bonding and antibonding states are less relevant. The molecules where there is partial covalent character, such as $^1\Sigma^+$ ScF show the most significant improvement with application of a U term.

2.5 Spin splittings of diatomic molecules

While structural and vibrational properties are highly relevant for comparing to experiment and understanding the chemical bonding, the fundamental aspect of transition metal chemistry we aim to capture in any first principles approach is identifying

		B3LYP	GGA	GGA+U_{scf}	CCSD(T)	Expt.[64]
ScF	r_e	1.795	1.797	1.811	1.801	1.788
$^1\Sigma^+$	ω_e	709	745	722	736	736
	D_e	6.73	6.86	6.02	6.17	6.21
ScF	r_e	1.855	1.842	1.873	1.868	
$^3\Sigma^+$	ω_e	644	685	639	588	
	D_e	5.49	5.61	5.12	5.78	
CrF	r_e	1.783	1.774	1.781	1.826	
$^4\Sigma^+$	ω_e	647	680	670	606	
	D_e	3.95	3.71	3.61	4.84	
CrF	r_e	1.8	1.794	1.816	1.795	1.784
$^6\Sigma^+$	ω_e	636	640	604	656	664
	D_e	4.36	4.82	4.41	4.25	4.61
FeF	r_e	1.759	1.742	1.762	1.769	
$^4\Delta$	ω_e	660	656	636	657	
	D_e	4.45	4.47	4.40	4.40	
FeF	r_e	1.790	1.785	1.801	1.815	1.780
$^6\Delta$	ω_e	651	631	599	599	663
	D_e	4.30	4.41	4.35	4.37	
CuF	r_e	1.772	1.780	1.794	1.775	1.745
$^1\Sigma^+$	ω_e	600	596	578	591	623
	D_e	4.47	4.71	4.41	3.95	4.46
CuF	r_e	1.762	1.795	1.795	1.763	
$^3\Sigma^+$	ω_e	625	611	603	638	
	D_e	2.15	2.22	2.20	3.59	

Table 2.8: Structural properties of early (ScF), mid (CrF, FeF), and late (CuF) transition metal fluorides at several levels of theory. Bond lengths (Å), frequencies (cm^{-1}), and dissociation energies (eV) are compared against experimental values, where available.

spin state ordering and splittings. The self-interaction problems in GGA are manifested in an undue stabilization of delocalized, or bonding orbitals, concomitant with an unnatural penalty on the more localized nonbonding and antibonding orbitals. Since the low-lying states of a molecule typically differ in the relative contribution of these types of orbitals, we can anticipate that GGA will most likely disfavor states with higher anti-bonding character. The hybrid functionals like B3LYP in theory reduce self-interaction error, but the correction by inclusion of Hartree-Fock exchange is only approximate. The result, as we will show, is that B3LYP fails to systematically over-stabilize one set of states or under-stabilize another, and thus it is challenging to predict when B3LYP results will be good estimates or poor ones. Experimentally, these spin state splittings may be determined primarily through anion photoelectron spectroscopy, which, as we previously explained, gives access to the lowest-lying states which differ by single electron detachment from the anion ground state. In many cases, this data may be used to definitively yield the splitting between the two lowest symmetry and spin states[70, 65]. An alternative experimental approach is high-resolution infrared emission spectroscopy, which can indirectly identify excited spin states by observation of perturbations in the spectra of the ground state spin states[71].

The hydrides are among the simplest systems in which we may study spin splittings because the electronic structure of the molecule is most akin to the bare transition metal with only slight perturbation to the $4s$ and $3d_{z^2}$ levels from hybridization with H $1s$. We previously showed that for MnH, GGA finds the $^7\Sigma^+$ and $^5\Sigma^+$ states to be degenerate, while $^7\Sigma^+$ is the experimental ground state. The main difference between the two states is that an additional anti-bonding σ^* orbital is occupied in the $^7\Sigma^+$ state, and therefore the failure of GGA to stabilize this state sufficiently is unsurprising. The B3LYP results, on the other hand, are even worse with a $^5\Sigma^+$

	B3LYP	GGA	GGA+U_{3d}	GGA+$U_{3d/4s}$	CCSD(T)	Expt.
ScH $^1\Sigma^+ \rightarrow\ ^3\Delta$	0.12	-0.13	0.08	0.12	0.12	
CrH $^6\Sigma^+ \rightarrow\ ^4\Sigma^+$	1.18	1.65	1.52	1.40	1.37	1.39
MnH $^7\Sigma^+ \rightarrow\ ^5\Sigma^+$	-0.12	0.01	0.13	0.21	0.22	0.21
FeH $^4\Delta \rightarrow\ ^6\Delta$	0.50	0.57	0.45	0.26	0.27	0.24
CuH $^1\Sigma^+ \rightarrow\ ^3\Pi$	2.41	2.24	2.31	2.43	2.45	2.43

Table 2.9: Energy splitting in eV of the two lowest symmetry and spin states of several hydrides. In the cases of CrH[71, 70],MnH[65],FeH[65], and CuH[72], experimental values are available and compared against the B3LYP, GGA, GGA+U, and CCSD(T) results. Negative values indicate an incorrect ground state.

state being stabilized by over 0.12 eV with respect to the ground state $^7\Sigma^+$. Comparison to experimental results, where available, for ScH, CrH, MnH, FeH, and CuH are shown in Table 2.5. These results show that GGA is systematically improved upon, particularly in the cases of MnH and ScH where GGA fails to correctly assign the ground state spin and symmetry. In each of the other cases, GGA splittings disagree with experiment by at least 0.2 eV, and there is no clear difference between systems where the transition is high spin to low spin versus low spin to high spin transitions. GGA+U and particularly with a U on both $4s$ and $3d$ states greatly improves estimates of spin state splittings and reduces errors by about tenfold. The sole application of a U_{3d} improves upon GGA results, but in every case the results may be further improved upon by inclusion of a U_{4s}.

The B3LYP results do not fare much better than GGA, with errors in ground state assignment (MnH) as well as errors around 0.2 eV for several other cases (FeH, CrH). The hybrid functional serendipitously reproduces experimental and quantum chemistry splittings for ScH and CuH with good accuracy, but the unreliability of the method makes it unusable as a tool in predictive transition metal chemistry. Work is ongoing in assessing the accuracy of GGA+U splittings for the oxides and fluorides.

2.6 Trends from isoelectronic $^6\Sigma^+$ states

In order to isolate the effects of electron affinity, covalency or ionicity, and net molecular charge on the role that GGA+U plays, we consider an isoelectronic series of molecules which have the same spin and symmetry associated with the $^6\Sigma^+$ term symbol. We have studied five molecules including two charged species, FeO^+ and CrO^-, two highly covalent molecules, FeN and MnO, and the conversely ionic CrF. Calculations of the Hubbard U for each of these species shows that the charged, covalent species FeO^+ exhibits a larger U_{scf} of 5.5 eV compared to its closest neutral comparison, FeN at 4.38 eV. By comparison, when the increase in charge facilitates an increase in relative ionicity, as in CrO^-, the U_{scf}, in this case 2.85 eV, decreases with respect to the closest neutral comparison, MnO at 3.41 eV but is still larger than the neutral ionic CrF at 2.04 eV. Overall, these results confirm that more covalent species exhibit much higher values of U than their ionic counterparts. On the other hand, charged systems usually exhibit larger values of U_{scf} unless the difference in charge results in a relative increase in ionicity. The results are also manifested by the equilibrium bond lengths and harmonic frequencies of the species. For FeN, FeO^+, and MnO, the equilibrium bond lengths are all around 1.65 Å and the harmonic frequencies reside around 780-850 cm^{-1}. The more ionic CrO^- and CrF, on the other hand, exhibit longer bond lengths around 1.8 Å and lower harmonic frequencies around 600 cm^{-1}. The effect of a U_{scf} term is greatest on the more covalent group of $^6\Sigma^+$ states although the overall effects on bond length and frequency are modest. The greatest difference with inclusion of a U_{scf} term is the dissociation energy, which decreases by as much as 1.5 eV in the case of $^6\Sigma^+$ MnO. More clues to the differences between these isoelectronic states lie in the differences between the corresponding fractional eigenvalues for each state in the occupation matrix (see

	GGA			GGA+U_{scf}		
	$\sigma_\uparrow$	$\sigma_\downarrow$	$\pi_\downarrow$	$\sigma_\uparrow$	$\sigma_\downarrow$	$\pi_\downarrow$
FeN	0.87	0.48	0.66	0.91	0.33	0.73
FeO^+	0.94	0.41	0.55	0.98	0.24	0.64
MnO	0.72	0.25	0.24	0.81	0.20	0.17
CrF	0.61	0.12	0.07	0.57	0.15	0.09
CrO^-	0.27	0.16	0.07	0.27	0.14	0.06

Table 2.10: Occupations of up and down σ as well as spin down π orbitals from eigenvalues of the occupation matrix for each isoelectronic $^6\Sigma^+$ state.

Table 2.6).

All states have roughly integer occupations corresponding to spin up manifold $\pi^2\delta^2$, but the relative occupations of a partially occupied spin up σ and spin down σ and π^2 vary widely among the isoelectronic series. The more covalent FeN, FeO^+, and MnO exhibit large spin up σ occupations corresponding to the filling of both σ and σ^* orbitals. The U_{scf} increases the total occupation of spin up σ further, while the spin-down bonding σ occupations decrease. These changes are concomitant with the small increase in bond length and more significant decreases in dissociation energy observed for these molecules. The π occupations correspond to about one half of an electron in each orbital, which corresponds to the physically meaningful case of two bonding π molecular orbitals with half a $3d$ electron contributing to each of the orbitals. The more ionic CrF and CrO^-, on the other hand, exhibit much lower occupations in the aforementioned orbitals due to enhanced charge transfer into the ligand. The spin up σ occupation is reduced to about half, corresponding to only occupation of bonding σ orbitals, which the $3d$ density which was in the covalent σ^* orbitals, being transferred ionically instead to the F or O^{2-} ligands. The other molecular orbitals in the minority spin are similarly significantly reduced from the covalent counterparts. Also, the U does not significantly alter the values of the occupations from their GGA values.

Overall, for all of these molecules, the difference between total GGA and GGA+U occupations is minimal, but the significant difference is an enhancement of majority spin anti-bonding orbitals and a slight decrease in the occupation of minority spin bonding states. These results also demonstrate that the occupation matrix can be a valuable tool for predicting the role and anticipated magnitude of a U term prior to a linear response calculation. The ability to tie chemical intuition to values of U is key, particularly in cases we will show shortly where numerical errors can obfuscate the true U of the system.

2.7 Practical considerations in GGA+U

In addition to calculating properties of several types of transition metal containing diatomic molecules in order to learn more about the role the U plays on the chemical bond, these molecules also provide excellent test cases for understanding the practical limits of our approach.

2.7.1 Pseudopotential dependence

We have briefly mentioned that the value of the self-consistent, linear-response U is dependent upon the definition of the projection manifold. It is self-evident that the value of U should vary if a different qualitative form of projection is used such as atomic orbitals versus maximally localized Wannier functions. However, we will consider here in more detail the potential role of differences in pseudopotentials can play in the value of Hubbard U as well as the properties of a transition metal molecule. In particular, it is worth recalling that energies at different values of U are not comparable. This means that properties such as band gaps for the solid state at

different values of U are also not likely to be comparable. On the other hand, what can be compared is the structural and vibrational properties as well as other density derived properties such as spin state splittings at an average value of U appropriate for a given pseudopotential.

We consider here the role that differences in atomic projections as a consequence of pseudopotential choice can have on both the determination of the Hubbard U and its subsequent effect on physical properties. It is important to note that the value of U obtained for the same structure but with differing pseudopotentials may easily differ by as much as 2-3 eV, particularly if the pseudopotentials were generated in different oxidation states. Although this value might seem high, it reflects the changes that atomic orbitals undergo as a function of the total atomic charge. In addition, as the value of U corresponds to how much correlation is missing from the standard LDA or GGA, it follows that the poorer the match between the oxidation states of the pseudopotential and the physical system, the higher the value of the linear-response U will be. For reference, we provide here comparison of FeO^+ sextet and quartet state properties calculated using pseudopotentials generated in various oxidation states. The Fe^{2+} ($d^{5.5}s^{0.5}$), $Fe^{0.5+}$ ($d^{6.5}s^{1.0}$), and Fe (d^6s^2, used throughout the remainder of this paper) pseudopotentials had values of U_{scf} of 7.0, 4.9, and 5.5 eV, respectively. As the electronic configuration of FeO^+ states is either d^6s^1 or d^7s^0, it is reasonable that the second pseudopotential produces the lowest value of U.

In comparing LDA+U or GGA+U calculations of differing pseudopotentials, total energies and, subsequently, HOMO-LUMO or band gap values for different values of U are not likely to be comparable. However, the properties derived from each spin and symmetry's ground state density, including frequencies, bond lengths, and state splittings, may be compared. The GGA results for each of the three cases show significant differences for some structural properties and energetic splittings (see Table

	$Fe^{2+}(d^{5.5}s^{0.5})$	$Fe^{0.5+}(d^{6.5}s^{1})$	Fe $(d^{6}s^{2})$
GGA (U= 0 eV)			
$^{6}\Sigma^{+}$ r_e(Å), $\omega_e(cm^{-1})$	1.59, 902	1.63, 906	1.62, 901
$^{4}\Phi$ r_e(Å), $\omega_e(cm^{-1})$	1.52, 1033	1.57, 1040	1.56, 1038
$\Delta E_{6\rightarrow4}$ (eV)	0.76	0.82	0.84
GGA+U_{scf} (eV)	7.0	4.9	5.5
$^{6}\Sigma^{+}$ r_e(Å), $\omega_e(cm^{-1})$	1.65, 745	1.66, 751	1.66, 749
$^{4}\Phi$ r_e(Å), $\omega_e(cm^{-1})$	1.76, 606	1.75, 613	1.75, 612
$\Delta E_{6\rightarrow4}$ (eV)	0.50	0.52	0.54

Table 2.11: Comparison for several pseudopotentials of structural properties including equilibrium bond length (r_e(Å)) and harmonic frequency ($\omega_e(cm^{-1})$) and state splittings (in eV) of the $^{6}\Sigma^{+}$ and $^{4}\Phi$ states of FeO^{+}. The GGA+U results are obtained at the respective U_{scf} of each pseudopotential.

2.11). While the harmonic frequencies for $^{4}\Phi$ and $^{6}\Sigma^{+}$ of 1040 and 900 cm^{-1} are identical for all pseudopotentials within the accuracy of the fitting procedure, more significant differences are observed in the bond lengths. The most ionic pseudopotential ($d^{5.5}s^{0.5}$) yields the shortest GGA value with $^{4}\Phi$ and $^{6}\Sigma^{+}$ at 1.52 and 1.59 Å, respectively, while the other two pseudopotentials yield bond lengths within 0.01 Å of each other. The GGA spin splittings are quite close at 0.82 and 0.84 eV for the $Fe^{0.5+}$ and Fe pseudopotentials, respectively, while the more ionic Fe^{2+} is 0.76 eV. Upon augmentation of a Hubbard term using the U_{scf} appropriate to each respective pseudopotential, the bond lengths of all three pseudopotentials are brought into agreement with each other within 0.01 Å, the harmonic frequencies agree within 6-7 cm^{-1}, and the spin splittings for all three pseudopotentials agree within 0.04 eV. Overall, an even smaller spread in values is observed for GGA+U than for GGA (see Table 2.11). Even a poor pseudopotential choice will remedy itself so long as the proper value of U, U_{scf}, is used. These observations demonstrate that perfect pseudopotential selection is not mandatory, but the best pseudopotential choice is

one which clearly reflects the oxidation states and charges of the physical system of interest.

2.7.2 Numerical stability of the linear-response calculation

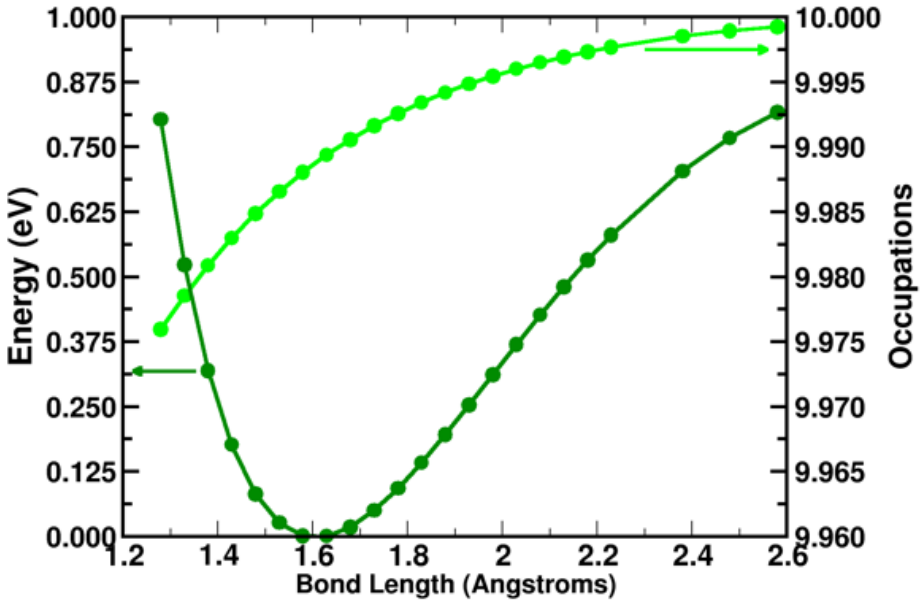


Figure 2-6: The binding energy curve (dark green) and the bond length dependent $3d$ projection occupations (light green) for $^2\Sigma^+$ ZnH.

The approach used to determine the value of the U for each electronic state suffers from numerical instabilities which should be evaluated. The value of U normally serves as a probe to the chemical nature of the molecule and should reflect what is known about the bonding interactions in the molecule. In our approach, we calculate a response function, χ of the occupations as a function of a potential shift, α, but we must invert it to obtain the U of our system. Typically when the manifold of interest contains nearly integer occupations, the response functions become small, nearly zero. In turn, when these must be inverted and the converged and bare are

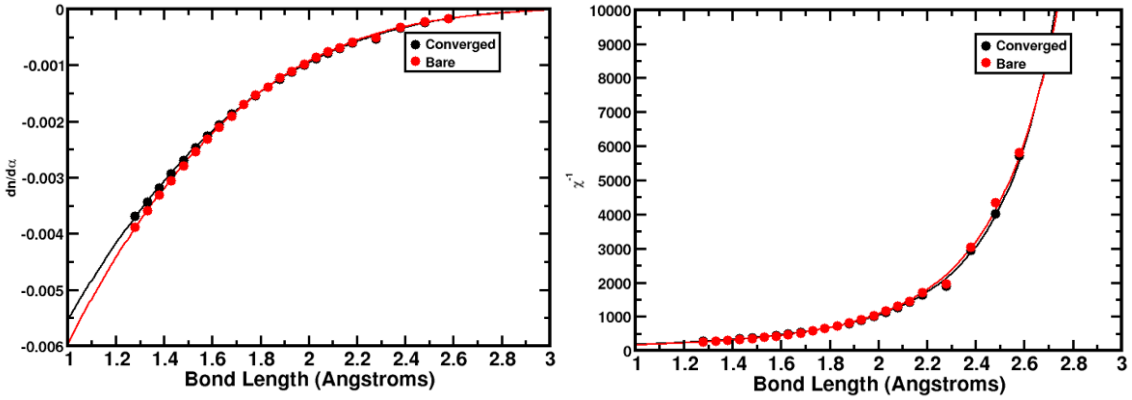


Figure 2-7: The bare (red) and converged (black) response functions, χ, for the $^2\Sigma^+$ ZnH state over a range of bond lengths (left) as well as the inverse of the response functions, χ^{-1} (right). The best fit polynomial curves are also shown as solid lines.

subtracted from each other, numerical instability arises. An example of this problem is best manifested by the ground state of ZnH, $^2\Sigma^+$, which has nearly exactly a $3d^{10}$ configuration. The occupations are their lowest at 9.97 under significant compression of the bond but rise quickly to 9.99 at the equilibrium and asymptotically approach exactly 10.00 as the molecule is dissociated (see Fig. 2-6). The relatively short bond, around 1.60 Å, is formed via interactions of Zn $4s$ atomic orbitals with H $1s$. Because the occupation matrix on the $3d$ manifold of ZnH is qualitatively as well as nearly quantitatively $3d^{10}$, the manifold should intuitively yield very little or zero response.

The bare and converged response functions of the $3d$ manifold of ZnH were measured and are shown in Fig. 2-7. Convergence criteria on total energy and diagonalization as well as higher cutoffs for the plane wave basis set can increase the accuracy of the projections calculated from the density. However, regardless of tight convergence criteria, there will always be a small amount of scatter on the data which should in particular affect the bare response function because this property is cal-

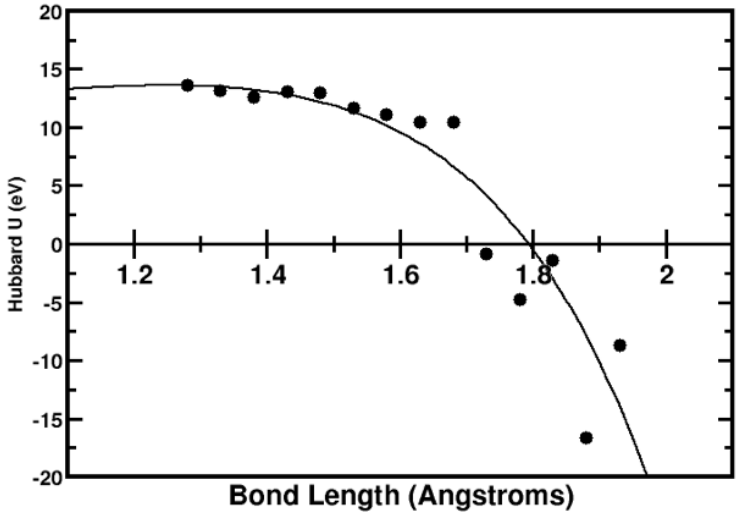


Figure 2-8: The value of U calculated from linear response at various bond lengths for ZnH from the inverse linear response functions is shown in black circles. Additionally, the difference of the two best fit lines of the linear response functions is shown as a black curve.

culated from the first self-consistent step, the initial diagonalization of the density upon response to the potential shift, α. The response functions over a variety of bond lengths are very small for both the bare and converged cases and they are also qualitatively identical. The total amount of the response is on the order of -0.002 $\frac{e^-}{eV}$ for bond lengths near equilibrium. A best fit polynomial to both the bare and converged also are nearly identical for most bond lengths. However, when the two curves and data are inverted, the reciprocal of the response functions are very large, on the order of 4-5000. The small numerical noise in the two response curves gives rise therefore to a large value of U of around 11 eV.

The U calculated from the data points as well as the difference of the two best fit polynomials is shown in Fig. 2-8. The scatter of the data from the best fit curve demonstrates some of the error bar obtained in inverting these infinitesimally small

numbers. Additionally, the curve diverges first staying flat and then crossing to negative numbers at intermediate bond lengths. Overall, the best approximation we can make on the true U of the system is that the U is zero. Great care must be made in assessing the actual size of response functions, those less than 0.01-0.1 are unlikely to yield numerically robust values of U. Additionally, the root of this problem lies in the fact that when we calculate U, it is based upon a full manifold of 10 electrons in the case of the $3d$ states. However, as only one electron exhibits a small fraction of hybridization in this case, the response function should really be renormalized in terms of the number of electrons accessible to potentially yield a response. This work is relevant also for many other systems and is ongoing.

Chapter 3

Addition-elimination reactions on FeO^+

3.1 Introduction

The gas-phase reactions of bare FeO^+ cations with hydrogen, methane, nitrogen gas, and other species have been widely studied as fundamental processes which can provide clues into the behavior of larger condensed-phase metal-oxo systems[73]. These paradigmatic reactions have been studied by numerous mass spectrometric techniques in detail[73, 74, 75, 76], but initial experimental results were surprisingly discordant with classical transition-state theories[74]. The reaction of FeO^+ with H_2 has been under particularly extensive study for some time now both theoretically[77, 78, 79, 80, 81, 82, 83] and experimentally[73, 74, 76, 84] as the most fundamental example of the two-state reactivity paradigm. Additionally, the reaction of FeO^+ with CH_4 has been studied in detail, primarily with density functional theory[85, 86, 87, 88, 89, 90, 91] and, to a lesser extent, experiment[73, 92, 93, 94]. Methane

oxidation on FeO^+ is particularly relevant because of methane's greater utility as a fuel upon conversion to methanol[95, 96], as well as the relevance to enzymatic systems such as methane monooxygenase[4, 97].

Reactions of FeO^+ with CH_4 and H_2 are both known to be exothermic by about 10 kcal/mol and 37 kcal/mol, respectively. However, these reactions are very inefficient: only about 10% of all collisions lead to products in the case of methane and 1% of all collisions lead to products for the hydrogen case. This inefficiency is particularly surprising because the same-spin nature of the ground state reactant and product species, $^6\Sigma^+$ FeO^+ and ^{6}D Fe^+, make these reactions spin-allowed[78]. There is an observed decrease in rate constant with increasing energy or temperature for both reactions, which is also in direct contradiction with the expected Arrhenius-like behavior for a spin-allowed reaction[79]. It is thus possible that the low-lying, excited quartet spin state surface also plays a role in the reaction. We refer to the two species by their spin multiplicities, defined as $2S+1$ where S is the net spin of the molecule, 5/2 for the sextet and 3/2 for the quartet. The resulting reaction inefficiency was proposed to be due to the crossing of multiple relevant spin surfaces for the reactions[79, 80], a feature now expected to be relevant in larger systems which include the enzyme horseradish peroxidase[98] and model non-heme inorganic iron alkane hydroxylation catalysts[99]. It has also been found that radical side products can form by dissociation of the reacting complex to $FeOH^+$ and radical H or CH_3, which are weakly endothermic or exothermic by roughly 3 kcal/mol, respectively. The relative exothermicity of the dissociated radical side reaction is key in interpreting and predicting branching ratios between the two types of reactions.

While theoretical considerations pointed to crossings between the sextet and quartet potential energy surfaces as the source for this reaction inefficiency [79, 80]; it is nevertheless notable that subsequent studies[76, 79, 80, 81, 82], which employed den-

sity functional techniques, have failed to agree with experimental results or highly-accurate quantum chemistry calculations. We will consider in detail here the reactions of FeO^+ with H_2 and CH_4 as a challenging test for a Hubbard U density functional approach, which we have recently introduced to improve semi-local or hybrid density functional descriptions of transition metal complexes[17]. We specifically consider the generalized-gradient approximation (GGA) in density functional theory (DFT) for its efficiency and widespread use in the community[31]; nevertheless, the Hubbard correction introduced here could be applied in conjunction with other exchange-correlation approximations. In order to gauge the success of our approach on both of these reactions, we compare DFT+U (GGA+U) results with both standard GGA results and highly-accurate quantum chemistry approaches. Post-Hartree-Fock techniques are able to recover nearly all of the correlation energy and may be used as an accurate benchmark or reference for the DFT calculations, but these approaches suffer from poor scaling which limit them to very small scales such as the reactions under consideration here (i. e. no more than 10 atoms).

3.1.1 Experimental study of the reaction of H_2 with FeO^+

Experimental study of the paradigmatic reaction of H_2 with FeO^+ using Fourier Transform ion-cyclotron-resonance (FTICR) has revealed a highly exothermic (ΔH_{rxn} = -37 kcal/mol) reaction which is seldom observed - as little as 1 in 1000 collisions leads to the formation of products[74]. Work using guided-ion beam tandem mass spectrometry identified a 0.6 eV reaction barrier along the high spin surface [75]. Therefore, for the reaction to proceed at low temperatures, the reactants must be excited to the low spin state to be activated for the reaction along the low-spin surface, which is expected to be barrierless based upon experimental measurements[73].

Shaik's model of "two-state reactivity" (TSR) predicts that two spin inversion points - or points where there is a crossing of the spin surfaces - occur at the entrance and exit channels for the reaction[100]. For the reaction, kinetic isotope effects were also determined to be negligible or nonexistent, indicating a limit on efficiency that is not activation-energy dependent. In some experimental approaches, it was possible to observe alternative side-reactions, but these are a small percentage of overall products[75]. Such side reactions include an endothermic formation of $FeOH^+$ and H radical (or, in the case of methane, a weakly exothermic formation of $FeOH^+$ and CH_3). Also, in the reverse reaction, where bare quartet Fe^+ could be formed, reactions with H_2O and CH_3OH were as much as 50 times more efficient than those with the sextet ion[75]. Overall, several mass spectrometric methods have been employed to study this reaction, and they qualitatively agree about exothermicity, a low efficiency, and the lack of barriers along the low-spin surface[73]. There do exist quantitative discrepancies in efficiency results depending on the conditions employed in the method, but this is likely due to differences in each technique such as pressure regime, thermalizing time, calibration technique, and interaction time. It is also worth noting that this addition-elimination reaction occurs primarily at low temperature, but at higher temperatures alternative reaction pathways become competitive[73].

3.1.2 Computational study of the addition-elimination reaction

Intriguing experimental results for this reaction prompted a great number of theoretical studies over the past 15 years. Calculations on FeO^+ using complete active space second order perturbation theory with Fock matrix off-diagonal elements neglected

(CASPT2D) initially confirmed the role of spin-crossovers which would explain the low experimentally observed efficiency[78]. The approximations at the CASPT2D level of theory led to the improper ordering for the quartet FeO^+ electronic states; $^4\Delta$ was identified as the ground state while $^4\Phi$ is now known to be the ground state. These results, in hindsight, also call into question the remaining results obtained with the method for this reaction, which we will address later in our own studies.

Common density functionals were later employed to calculate spin-orbit coupling (SOC) constants of geometries along the PES using an extension of the formalism of LeFebvre-Brion and Field [79]. Spin inversion probabilities may be calculated using SOC values as a crude estimate and these calculations concur with the low-experimentally found efficiency, but this approach still requires further effort to yield a quantifiable and reliable rate theory[74]. The SOC constants predicted bottlenecks to spin-inversion at both the entrance and exit channel. In particular, calculations indicated anisotropy of the SOC constants corresponding to different geometries and that SOC constants at the exit channel approach zero. In an effort to obtain quantitative results for the reaction coordinate of this system, common density functionals - BP86 and B3LYP, as well as a lesser known FT97 functional - were used with a localized basis set (Wachters on the metal and a Dunning-style TZ2P basis set on light atoms) to calculate the reactants, intermediates, products and transition states present along both spin surfaces [80]. While DFT accurately predicts a steeper sextet surface and shallow quartet surface in accordance with an experimentally observed low-efficiency reaction or occasionally fast and exothermic reaction, the quartet surface is not predicted to be nearly shallow enough. All three functionals predict a more stabilized quartet Fe^+ product than the sextet, and by using some rough estimates to correct this erroneous splitting, the researchers suggested that a spin splitting favoring the sextet at the third intermediate of about 3 kcal/mol should exist.

It has remained a challenging task to obtain computational studies which concur with the theoretical two-state reactivity model and with the experiment, and in an effort to further reconcile the theoretical picture with computational results, CCSD(T) single point energies were computed in 1999 on B3LYP geometries for this reaction[82]. Although some error is introduced by using relaxations obtained with B3LYP, we will later show that this error is relatively small, and in fact the results of this study were quite consistent with experiment and theoretical predictions. The CCSD(T) results found accurate splittings within 0.02 eV for the reactant and product complexes, a reduced barrier along the quartet surface of at most 0.32 eV, and a sextet stabilized by 0.12 eV at the third intermediate. Further effort has been made to improve upon both CCSD(T) and density functional results by optimizing the basis set to yield results which agree with experimentally known splittings[81]. Most recently, diffusion quantum Monte Carlo calculations were completed using B3LYP structures in the CASINO code. This cursory study failed to correctly account for spin-inversion at the exit channel and yielded inaccurate splittings for several intermediates. We expect major sources of error to be derived from the sheer number of low-lying states as well as the difficulty in choosing an accurate reference function for this method.

3.1.3 Methods

Calculations have been carried out primarily in two forms - plane-wave, density functional calculations, and localized basis set, post-Hartree-Fock calculations, which we used for a highly-accurate reference to our DFT results. The plane-wave density functional calculations were completed with the Quantum-ESPRESSO package[24] using the Perdew-Burke-Ernzerhof (PBE)[31] generalized gradient approximation

(GGA) in its standard form and augmented with a linear-response, Hubbard-like U term, as previously outlined[7, 17]. The linear-response calculation of the Hubbard U term is also implemented in this package. Ultrasoft pseudopotentials were employed with a plane-wave cutoff of 40 Ry for the wavefunction and 480 for the charge density for most structural relaxations to ensure accurate convergence of forces as well as spin and symmetry state splittings[1]. Structural relaxations and nudged elastic band (NEB) optimizations[26] were carried out at integer values of the Hubbard U to further assess the U dependence of structure and energetics. In order to improve the resolution of steep reaction barriers, the climbing-image and variable springs methods were employed in conjunction with the NEB calculations[25]. For diatomic molecules, the harmonic frequencies and anharmonic contributions were obtained with high-order polynomial fits to the potential energy surface.

Post-Hartree-Fock approaches were employed to provide an accurate but computationally very expensive reference to compare the density functional calculations against. The reaction with hydrogen (n_{at}=4) was studied primarily with GAUSSIAN 03[101], and the reaction with methane (n_{at}=7) was studied with MOLPRO[102]. The primary single-reference method employed was coupled cluster with singles, doubles, and perturbative triples (CCSD(T)). The T1 diagnostic was used as a tool to identify structures with potentially strong multi-reference character or instability in the triples term, and select CASPT2 and MRCI calculations were employed to verify the CCSD(T) results. All multi-reference calculations on both reactions were carried out using MOLPRO, which was available to us on a 64-bit architecture. The same Pople-style 6-311++G(3df,3pd) basis set was used for all

[1]State symmetry assignments were made based upon the absolute total angular momentum, $|\Lambda|$, of the density and the overall symmetry of the component densities, λ, which remain as "good" quantum numbers in our density functional formalism. Additionally, the preserved spin quantum number of our systems is $\hat{S}_z$.

post-HF calculations to ensure consistency.

In order to construct the CCSD(T) hydrogen on FeO^+ reaction coordinate, we estimated equilibrium CCSD(T) geometries for each intermediate. The low dimensionality of the potential energy surface permitted calculation of a fine mesh of single point energies about 0.01-0.02 Å apart and deduction of a minimum energy configuration. These geometries are also useful to determine errors in geometries in the GGA and GGA+U methods. For transition states, we used GGA+U or GGA geometries to calculate single point CCSD(T) energies since discretizing the PES at the transition states was too computationally expensive[2].

3.2 Linear-response U of the stationary points

We consider first the value of U calculated for each intermediate and transition-state for the reactions with both hydrogen and methane, which helps us identify the broad similarities of the two cases. In order to obtain a global reaction coordinate, we must use energies calculated with a single value of Hubbard U, which we choose as an average, $U_{scf,av}$, over all states. Since total energies at differing values of U are not comparable, it is important to pay attention to large deviation in U_{scf} at any stationary point from the global average. The overall $U_{scf,av}$ we obtain for the H_2 reaction is 4.93 eV, and the value for the CH_4 reaction is very similar at 5.09 eV. The U_{scf} is on average larger than the value calculated from GGA (U_0) for most geometries. We find the smallest discrepancy between the self-consistent and non-self-consistent values of U for reaction intermediates, while the differences for

[2]We note that energetic differences at the CCSD(T) level of the different TS geometries were small and did not noticeably affect barrier height estimates.

transition-states are as large as 1.5 eV, as is in the case of ^{6}TS-2$_{CH_4}$. For most single points in these two reactions, the individual U_{scf} values are within 1 eV of the $U_{scf,av}$. Based upon typical cases (e.g. Int-3 splittings in Fig. 3-9 and sextet TS-1 barriers in Table 3.3), employing GGA+U with a value of $U_{scf,av}$ which is 1 eV too small or too large for the appropriate structure likely yields energy differences in the range of 0.01-0.1 eV, and this uncertainty can, in many cases, be decreased by examining splittings with a locally averaged value of U_{scf}.

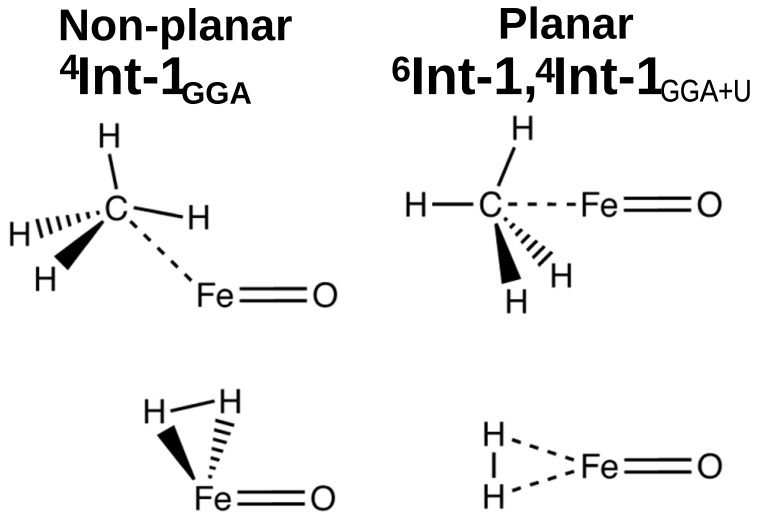


Figure 3-1: Comparison of the first intermediate along sextet and quartet spin surfaces with GGA and GGA+U for reactions of FeO^+ with CH_4 (top) and H_2 (bottom).

Trends in the values of U_{scf} are very consistent for the two reactions' quartets and sextets (see Table 3.1). While the exact numbers between the two reactions are not identical, the similarities highlight the fact that coordination and electronic configuration contribute most strongly to the calculated U value. Overall, we observe a larger U_{scf} in the range of 5.5-6.5 at the entrance channel, where the iron center is weakly 3- or 2- fold coordinated in the Int-1 structure(see Fig. 3-1). At the first

reactant barrier, U_{scf} decreases monotonically for all TS-1 and Int-2 structures, and this change is concomitant with a reduction in FeO^+ bond order and coordination.

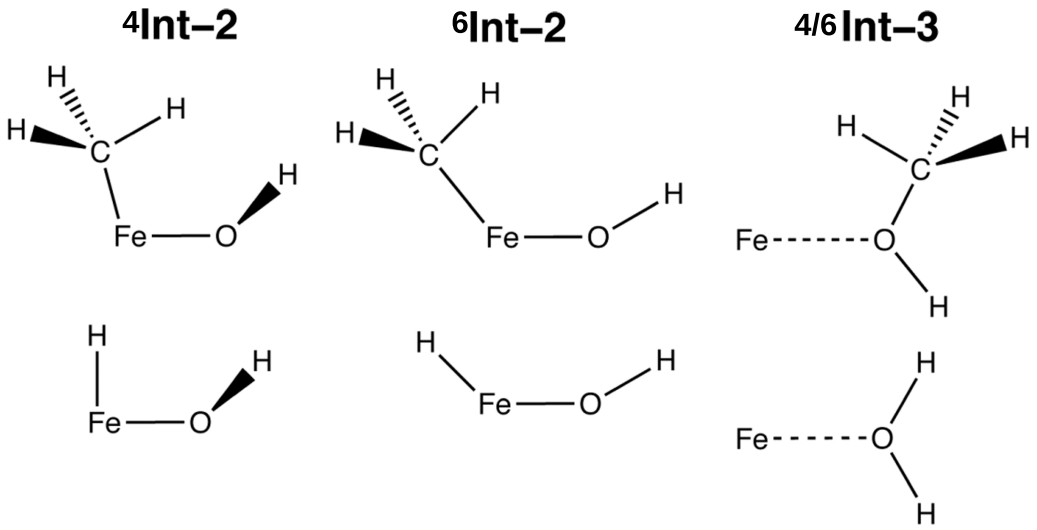


Figure 3-2: Comparison of the second and third intermediates along sextet and quartet spin surfaces with GGA and GGA+U for reactions of FeO^+ with CH_4 (top) and H_2 (bottom).

At the second reaction barrier, the quartet and sextet surfaces exhibit different linear-response properties. This behavior of U_{scf} is strongly affected by differences in the chemical coupling of the quartet and sextet Fe^+ centers with the organic ions in the two spin surfaces. The $3d^6 4s^1$-like sextet's two minority spin electrons are equally divided between the $3d$ and $4s$ manifolds, and this results in an increased prominence of the $U_{scf,4s}$ for ^{6}TS-2_{H_2} and an overall lower $U_{scf,3d}$ around 3-4 eV for both reactions. While the extension to include $4s$ states, and thus a U_{4s}, is straightforward, it plays a role exclusively in the class of weakly bound molecules where the $4s$-$3d$ hybridization dominates and $3d$ and $4s$-derived molecular states are close in energy. The sextet transition state, especially in the methane case, is indeed quite weakly bound, and we will later discuss this in greater detail. The transition state along the quartet

	Quartet					Sextet				
	H_2		CH_4				H_2		CH_4	
	U_0	U_{scf}	U_0	U_{scf}			U_0	U_{scf}	U_0	U_{scf}
FeO^+	5.84	5.50	5.84	5.50		FeO^+	5.37	5.48	5.37	5.48
Int-1	5.89	6.31	6.21	6.23		Int-1	6.46	6.41	6.07	6.17
TS-1	5.34	5.74	5.97	6.05		TS-1	4.47	4.84	4.59	4.59
Int-2	4.87	4.95	4.67	5.10		Int-2	4.18	4.17	3.93	4.37
TS-2	4.52	5.73	4.68	5.47		TS-2	3.27	3.17	2.72	4.35
Int-3	1.81	1.89	2.15	2.33		Int-3	4.79	5.00	5.32	5.47
Avg	4.71	5.02	4.92	5.11		Avg	4.76	4.84	4.67	5.07

Table 3.1: U_0 and U_{scf} for Intermediates (Int) and Transition-States (TS) along the H_2 and CH_4 reaction coordinates. Quartet surface values in left table, sextet at right.

surface, which has a $3d^7$ character, is more tightly bound and the U_{scf} of ^{4}TS-2, is close to $U_{scf,av}$. All Int-3 structures at the exit channel closely resemble a weakly bound Fe^+ ion with the product water or methanol (see Fig. 3-2). Consequently, the Fe^+ charge density is increasingly atomic-like, and the values of the Hubbard U are significantly lower at this point in the reaction surface for the quartet at around 2 eV (the sextet displays a U_{scf} considerably closer to the average).

We highlight that the U_{scf} for each stationary point of a given configuration and electronic symmetry is calculated separately. We also consider the energetic evolution of each state individually and thus take into account any changes in state energetics ordering with increasing values of Hubbard U. Since these two reactions involve unsaturated metal-centers for which the number and character of metal-ligand bonds change across various spin states, they provide a far more challenging probe of the robustness of our method compared to more typical problems which involve saturated metals or fewer changes in metal-ligand bonding. From the trends we have thus far observed in U_{scf} along the reaction coordinate, we anticipate the GGA+$U_{scf,av}$ energetic description to be optimal for the first half of the reaction,

while it should be slightly poorer for the second half of the reaction (particularly at^{6}TS-2). Even when an accurate description of the global reaction coordinate may be difficult to achieve, a locally averaged U_{scf} at a key stationary point can greatly improve the structural and energetic estimate[17].

3.3 Fe^+: a difficult test for density functionals

Shaik's argument in 1998 that the failure of DFT to predict a second spin-crossing prior to the third intermediate has led to a number of high level calculations by several groups in order to rationalize this behavior. One group proposed to optimize the basis set so that the experimental Fe^+ sextet quartet splitting was preserved[103]. However, most of the $FeO^+ + H_2$ reaction coordinate more closely depends upon FeO^+ chemistry which, while related to Fe^+ energetics, is not necessarily improved with the new basis set. For our GGA+U study, it is somewhat unnecessary to apply a U term to an isolated Fe^+ iron atom which has no ligands because there is no reservoir from which to add or remove electrons. We have measured the U of the quartet Fe^+ in which a $3d^7$ configuration represents the ground state and an excited $3d^6 4s^1$ state lies within an eV of it. In this case, the U is small but measurable at about 1 eV, and this is likely due to the close interplay of $3d$ and $4s$ manifolds.

With an isolated atom, the PWscf code by default symmetrizes the charge density with all of the symmetry operations available. This manifests itself in an occupation matrix, for instance, for the case of $3d^6 4s^1$ $^6Fe^+$ which has 0.2 electrons in each of the five spin down $3d$ orbitals. This symmetrized solution does not yield a reliable energy estimate for these calculations in the case of GGA and this solution is also strongly penalized in GGA+U. We use instead a constraint available in the Car-Parrinello code in the ESPRESSO package[104, 24]. This forces integer occupation

and thus yields a reliable energy. We estimate the sextet-quartet splitting to be about 7.31 kcal/mol, which is slightly larger than the experimental splitting (see Table 3.2). Our CCSD(T) code and basis, we also note, slightly underestimates the Fe^+ splitting. Default symmetry constraints are a concern for highly symmetric

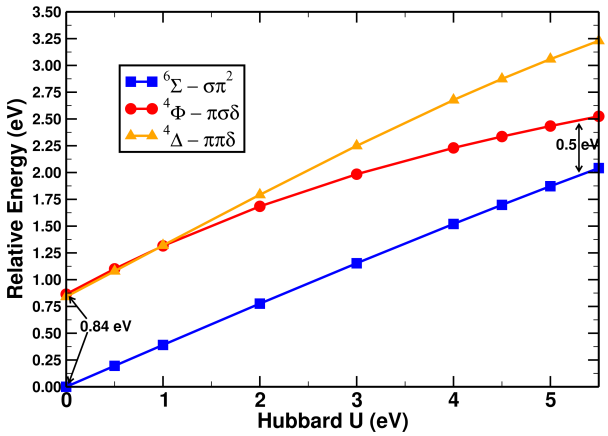
Splitting kcal/mol	CP code	PWscf $U=0$	PWscf $U=5$	g03 CCSD(T)	g03[82] CCSD(T)	Expt.
$^4Fe^+ \rightarrow ^6Fe^+$	7.31	21.44	25.37	4.14	5.25	5.35

Table 3.2: Comparison of splittings for different basis sets and methods of Fe^+.

ionic configurations. It can indeed yield erroneous energetics and structure. However, most of these cases are limited to isolated atoms and some small molecules. It may be useful in these cases to reduce the degree of symmetry enforced and verify that the energy and occupation matrix of the system remains unchanged.

3.4 The low-lying states of the FeO^+ reactant

The structure and energetics of low-lying states of FeO^+ are described with greatly improved accuracy with GGA+U over standard GGA by comparing against high level quantum chemistry methods (CCSD(T)). For both GGA and GGA+U, the FeO^+ ground state is a $d^6s^1\ ^6\Sigma^+$ with a full spin up $3d$ and $4s$ manifolds, as well as fractional occupation of spin down π and $\sigma_{d_{z^2}}$ orbitals. The total $3d$ orbital occupations are near 6.45. The molecular σ bonds exhibit strong hybridization between $4s$ and $3d_{z^2}$-derived atomic orbitals, hence the intermediate occupations of $3d$ states. Elongation of the FeO^+ bond from 1.62 Å (GGA) to 1.66 Å (GGA+U, U= 5.5 eV) is associated with enhancement of spin up $3d$ density in π^* states and subsequent reduction of the partial $3d$ occupation of spin-down π orbitals. This change in hybridization



Method	r_e	ω_e	$\omega_e x_e$
	$^6FeO^+$		
GGA	1.62	901	328
GGA+U	1.66	749	432
CCSD(T)	1.66	724	434
	$^4FeO^+$		
GGA	1.56	1038	332
GGA+U	1.75	612	172
CCSD(T)	1.70	633	188

Figure 3-3: The (a) Relative energy of FeO^+ states (b) Equilibrium bond lengths, r_e (Å), harmonic frequencies, ω_e (cm^{-1}), and anharmonicities, $\omega_e x_e$ (cm^{-1}) for the $^6\Sigma^+$ and $^4\Phi$ states of FeO^+.

also reduced the harmonic frequency by about 150 cm^{-1} in improved agreement with CCSD(T) results[17]. Overall, these results are in close correspondence to the ground state which has been experimentally identified as having a 1.643 Å bond length[84].

Several low-lying quartet FeO^+ states are close in energy, including $^4\Phi$, $^4\Delta$, and $^4\Pi$. While $^4\Phi$ and $^4\Delta$ are nearly degenerate in energy with GGA, GGA+U stabilizes $^4\Phi$ preferentially by nearly 0.75 eV over the other states, in agreement with ours and other CCSD(T) results(see Fig. 3-3)[17, 82, 105]. The equilibrium bond length of $^4\Phi$ increases significantly from 1.58 (GGA) to 1.75 Å (GGA+U) and the harmonic frequency decreases by from 1038 to 612 cm^{-1} due to an even more pronounced increase in π^* occupation in this case than in the sextet. The GGA+U bond length is slightly long compared to the 1.70 Å value for CCSD(T). The next highest state, $^4\Delta$, exhibits less severe elongation from 1.55 Å to 1.62 Å.

For both the quartet and the sextet, two competing factors exist - a physically

relevant hybridization of the Fe $3d$ electrons with O $2p$ electrons to form the bonding π orbitals as well as the spurious tendency of the $3d$ electrons to delocalize - the latter of which is corrected and greatly reduced by the Hubbard U term. It is of particular note that the GGA+U sextet-quartet splitting (0.54 eV) is in remarkable agreement with the CCSD(T) value (within three hundredths of an eV) and frequencies for the sextet and the quartet are within 25 cm^{-1} and 21 cm^{-1} of their CCSD(T) values, respectively compared to a discrepancy as large as 400 cm^{-1} for GGA[17]. Obtaining the correct spin splittings for FeO^+ low-lying states is crucial to ensure that the entire reaction coordinate is well-described.

3.5 The first reaction intermediate

Free hydrogen or methane complexes weakly with FeO^+ at the metal site to form the first intermediate (Int-1) of the reaction, which has not been isolated experimentally (see Fig. 3-1)[73]. Along the sextet surface (both with GGA and GGA+U), the electronic structure of the FeO^+ state is only weakly perturbed by the bound CH_4 or H_2, as is evidenced by the fact that the Fe-O bond lengthens by only 0.01 Å with respect to its bare FeO^+ value, or 1.63 Å for GGA and 1.67 Å for GGA+U. The preferred binding site for the reactant is at the Fe site with the reactant-Fe bond along the same direction as the Fe-O bond. The H-H bond length, 0.79 Å, is lengthened by only 0.01 Å with respect to the isolated molecule. The bond between Fe and the center of the H-H bond is quite long, 1.96 Å for GGA and 1.93 Å for GGA+U. In $^6\text{Int-1}_{CH_4}$, the methane tetrahedron flattens slightly, with the three hydrogens closest to the C-Fe bond exhibiting a dihedral of 135° instead of the idealized value of 120°. The Fe-C bond is comparably long at 1.95 Å as well. While the effect is subtle, GGA+U does in fact tune the relative energy of this intermediate.

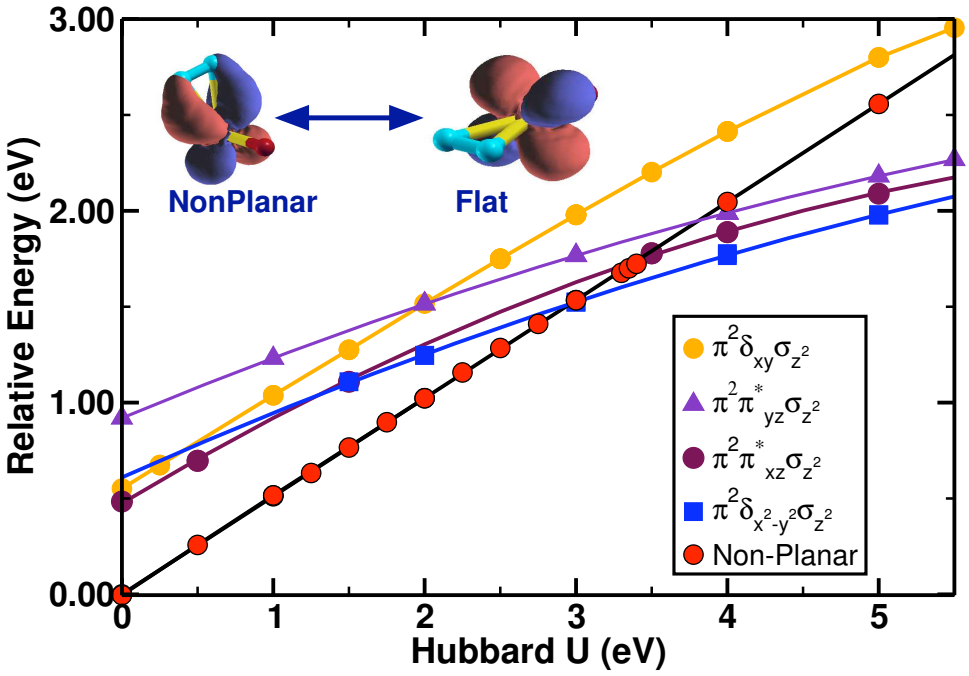


Figure 3-4: Relative energy of flat and non-planar ^{4}Int-1 states for GGA and GGA+U approaches.

In contrast to the sextet, the ^{4}Int-1 exhibits a more dramatic response structurally and energetically to the value of U, and thus demonstrates the utility of a Hubbard U approach. While the binding of hydrogen or methane to the FeO cation occurs parallel to the metal-oxo bond in the sextet case, the ligand may instead preferentially bind in a manner perpendicular to the bond. Early density functional calculations predicted that this non-planar geometry for the ^{4}Int-1_{H_2} was the preferred structure[79, 80]. However, we showed that this stabilization by GGA was not preserved in the GGA+U result for the reaction with hydrogen. Our GGA+U results reversed the preference of GGA and stabilized the planar structure for U above 3 eV and by nearly 0.5 eV at a U of 5 eV, with respect to the lowest-lying planar

geometry electronic state (see Figs. 3-1 and 3-4). In the nonplanar orientation, the angle between the midpoint of the H-H bond and FeO^+ ranged from 87^o to 91^o. In contrast to the sextet structure, the bond between H_2 and Fe is much shorter, ranging from 1.73 Å for GGA to 1.79 Å for GGA+U (3 eV). For the planar geometry, quartet Int-1 closely resembles the sextet structure with an elongated, Fe-H_2 bond. Additionally, since this structure electronically resembles a weakly perturbed FeO^+ molecule, the large number of isolated FeO^+ quartet states leads to a large number of closely spaced electronic states within about 0.5 eV of each other derived from both $^4\Phi$ and $^4\Delta$ variants (see Fig. 3-4). In fact, the flat structure favored by GGA is closely related to $^4\Delta$ with short FeO^+ and Fe-H_2 bonds at 1.57 Å and 1.94 Å respectively. The quartet Int-1 preferred instead by GGA+U is $^4\Phi$ derived with a longer, 1.79 Å FeO^+ bond.

In both planar and non-planar orientations, the primary bonding occurs via bonding overlap between H 1s density and the FeO^+ bonding σ molecular orbitals ($3d_{z^2}$+$2p_z$). In the non-planar geometry, additional binding of the complex occurs via overlap of the two hydrogens' 1s density with the δ_{xy} atomic-like orbitals of the metal center. The GGA preference of the nonplanar hydrogen binding site by nearly 0.5 eV may be viewed as an unphysical artifact of self-interaction error similar to the erroneous energetic ordering of binding sites of small organic molecules on metal surfaces[106]. Importantly, our CCSD(T) calculations confirm the GGA+U result that the planar geometry is more energetically favorable[17]. The non-planar geometry is not a stable sextet intermediate because the favorable interaction occurs via the minority spin channel, which is not present in the sextet case.

It follows that GGA shows a similar preference for the non-planar geometry for methane reaction ^{4}Int-1 complexes (see Fig. 3-1). This effect is reduced in the methane case for several reasons: the density of only one out of four hydrogens is

available for overlap with the localized orbitals of Fe, it is not possible to position the hydrogens or carbon to overlap with δ states, and, moreover, C-Fe interactions are unable to contribute to formation of a bent structure. Indeed, while a non-planar GGA structure is preferred for the ^{4}Int-1$_{CH_4}$ by about 0.11 eV, GGA+U stabilizes the planar structure for all values of U greater than 1 eV and overall by about 0.5 eV at the globally averaged U of 5 eV. MRCI results for the methane reaction agree with this state ordering, and our ^{4}Int-1 GGA+U results demonstrate the dramatic improvement in both structural and energetic description brought about by addition of a Hubbard U term.

3.6 The first reaction barrier

The first catalytic step in the reaction of FeO^+ with methane or hydrogen is the abstraction of a single hydrogen from the reactants by oxygen. The transition state associated with this step for all reactants and spin surfaces is four-centered and all bonds to the soon-to-be abstracted hydrogen are highly stretched. Along the sextet surface, GGA's have been shown to properly reproduce the characteristic experimentally steep barrier(see Table 3.3)[17, 79, 80]. This large activation barrier for both sextet reaction surfaces stems from the short and tight bond of the $^6\Sigma^+$ FeO^+ molecule which, in the GGA transition state, is stretched by 0.1 Å to 1.72 Å. The H_2 reactant bond is stretched to 1.05 Å, and a new, weak O-H interaction at 1.34 Å is formed. This transition-state structure changes very little from GGA to GGA+U, albeit with a slightly longer Fe-O bond and shorter H-H bond, and the barrier height estimates are similar for GGA (1.01 eV) and GGA+U (0.97 eV). The methane reaction has a similarly high activation barrier along the sextet surface which changes relatively little from GGA (1.35 eV) to GGA+U (1.29 eV). Structurally, the Fe-O

bond is stretched to the same extent as in the hydrogen reaction and the C-H bond is stretched by 0.3 Å from its equilibrium value of 1.10 Å. Changes in structure as a result of augmenting a "+U" term are negligible. While the forward barriers for both reactions do not change significantly with the value of U, the change in the back-reaction barriers (or the step's exothermicity) for both the hydrogen and methane case is quite large.

In contrast to the sextet surfaces, the quartet reaction barrier demonstrates considerable sensitivity to the value of the Hubbard U, as we would have predicted after the study of ^{4}Int-1. The nudged elastic band reaction path we study uses the planar ^{4}Int-1 as an endpoint, although it is a metastable minimum for GGA and low to intermediate U values. The path also passes through a non-planar GGA ^{4}Int-1-like structure before reaching the stretched transition-state which is the global minimum for GGA but disappears as a stable intermediate for intermediate values of GGA+U. We previously showed[17] that a GGA barrier for the hydrogen reaction is steep, at about 0.4 eV with respect to the non-planar Int-1, and the quartet TS-1 lies above the reactant energies, in contrast to experimental evidence(see Fig. 3-5)[73]. With GGA+U, a shallow reaction surface is recovered with an activation energy of 0.20

	CH_4		H_2	
U	ΔE_{fwd}	ΔE_{back}	ΔE_{fwd}	ΔE_{back}
0	1.35	1.88	1.01	1.60
1	1.35	1.88	1.01	1.68
2	1.35	1.89	1.00	1.77
3	1.34	1.92	0.99	1.85
4	1.33	1.96	0.98	1.94
5	**1.29**	**2.00**	**0.97**	**2.02**

Table 3.3: Forward and back-reaction barriers as a function of U for the first sextet barrier of the reactions of FeO^+ with methane and hydrogen in eV.

eV, and the transition state lies 0.16 eV below the ground state reactants. The reduction in the barrier estimate is likely derived from the weakening of the Fe-O bond in the transition state from 1.61 Å to 1.72 Å, thereby permitting increased bonding with the reactant complex.

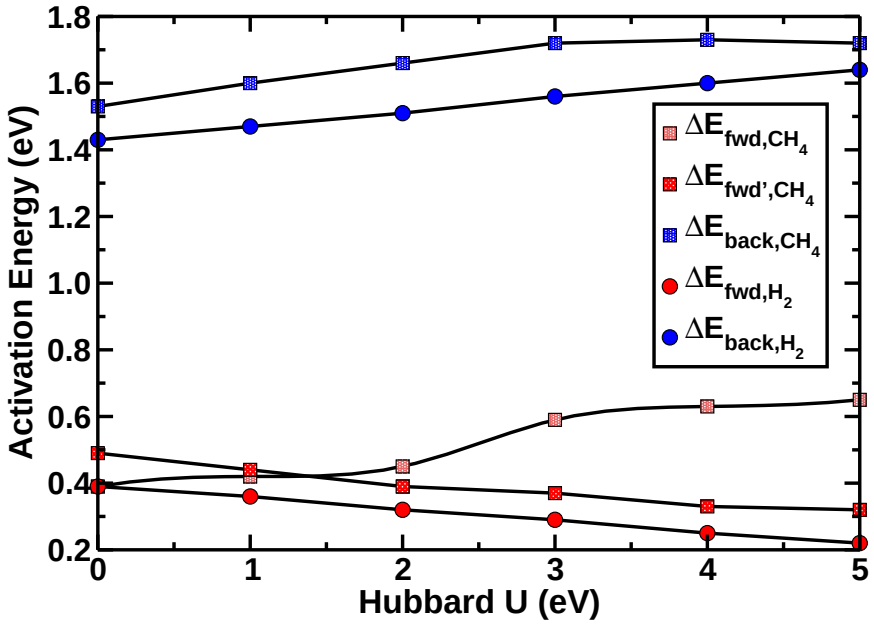


Figure 3-5: Forward and back-reaction barriers as a function of U from U= 0 (GGA) to U= 5 (the global $U_{scf,av}$) for the first quartet barrier of the reactions of FeO^+ with methane and hydrogen in eV. For the methane reaction, $\Delta E'_{fwd}$ refers to the forward reaction barrier calculated from the non-planar geometry.

As with the hydrogen case, the GGA ^{4}Int-1_{CH_4} non-planar geometry appears in the GGA reaction coordinate as a local minimum stabilized by 0.1 eV. The transition state also resembles its analogous sextet TS-1 with the methyl group sitting nearly perpendicular to the Fe-O bond and the leaving hydrogen shared equally between the carbon and oxygen. The local minimum from the non-planar structure disappears in the GGA+U reaction coordinate, and the overall reaction barrier increases to nearly 0.65 eV for U= 5 eV as a result of the reduced relative energy of the planar

^{4}Int-1 endpoint (see Fig. 3-5). If we choose instead as a reference point either the non-planar geometry or the crossing point of the quartet and sextet coordinates, the barrier estimate is reduced even further to 0.32 eV or 0.25 eV, respectively. Additionally, this GGA+U transition-state also lies below the reactant energy by 0.4 eV, compared to 0.3 eV for GGA. The quartet surface for both reactions demonstrates considerable sensitivity to the value of U applied, which must be close to the U_{scf} in order to accurately reproduce experimental and accurate theoretical estimations.

3.7 The second reaction intermediate

Upon abstraction of a hydrogen, H_{abs}, from the reactants, the hydrogen in turn binds to oxygen to form the stable Int-2, and the Fe-O bond lengthens further. This low symmetry structure possesses only a single low-lying electronic state for each spin and the energetic splitting of the sextet and quartet structures is close in energy ($\Delta E \approx \pm$ 50 meV) for both reactions. The sextet state is the GGA and GGA+U ground state for the methane case, while GGA+U preferentially stabilizes the hydrogen reaction's ^{4}Int-2 over the sextet, reversing the GGA ordering. The sextet structures have a planar configuration with the ligands at obtuse angles relative to the line of the Fe-O bond; the quartet structures instead prefer a ligand-Fe-O angle of about 90 degrees and an O-H bond directed out of the plane. For both reactions and spin surfaces, the Fe-O bond is elongated with respect to the TS-1 value. In ^{6}Int-2, the Fe-O bond is about 1.74 Å for GGA and lengthens further to about 1.77 Å for GGA+U. Total elongation for the GGA+U ^{4}Int-2 is 0.08 Å with respect to its 1.69 Å GGA value. The O-H bond is quite short, at about 0.98 Å, an equilibrium value for O-H bonds in many molecules. The Fe-H bond is comparably longer at about 1.60 Å for both GGA and GGA+U. The Fe-H bond length is similar for both the quartet, where the

character is nominally $3d^7$ and sextet, which is nominally $3d^6 4s^1$. This suggests that the Fe-H bonding is dominated by $1s - 3d_{z^2}$ interactions. The second reaction step, of which Int-2 is the starting point, is the concerted translation of the ligand (H or CH_3) at the iron site to a midway point atop the FeO bond coupled to a rotation of the already abstracted hydrogen. This geometric rearrangement culminates in the formation of the water or methanol leaving group and corresponds to a dramatic change in coordination environment for the Fe center: from two strong ligands, to one strong and one weak ligand, and finally to a single very weak ligand. These coordination changes provide a stringent test for whether GGA+U can treat each of them with equal accuracy.

3.8 The second reaction barrier

Previous results[17, 80] have shown the quartet barrier to lie below the sextet barrier, and the geometric structure plays a significant role in the relative barrier heights. In the ^{4}Int-2_{H_2} structure, a ligand-Fe-O angle of 90 degrees requires less translation to position the ligand directly over the Fe-O bond. The GGA reaction barrier for the hydrogen reaction is about 0.42 eV, but it is significantly reduced by GGA+U (U= 5 eV) to 0.13 eV (see Table 3-6). The Fe-O bond elongates further from the length in Int-2 to 1.80 Å in the GGA+U ^{4}TS-2_{H_2}. The transferring hydrogen is shared nearly equally between Fe and O, with bond lengths of 1.60 Å and 1.44 Å, respectively. The GGA+U reaction barrier for the hydrogen reaction is consistent with the shallow, barrierless quartet surface predicted by experimental evidence and our own highly accurate theoretical results[17, 74, 80].

For the methane reaction, the second quartet GGA barrier is very steep at nearly 1.4 eV, while with GGA+U the activation energy is halved to about 0.7 eV(see

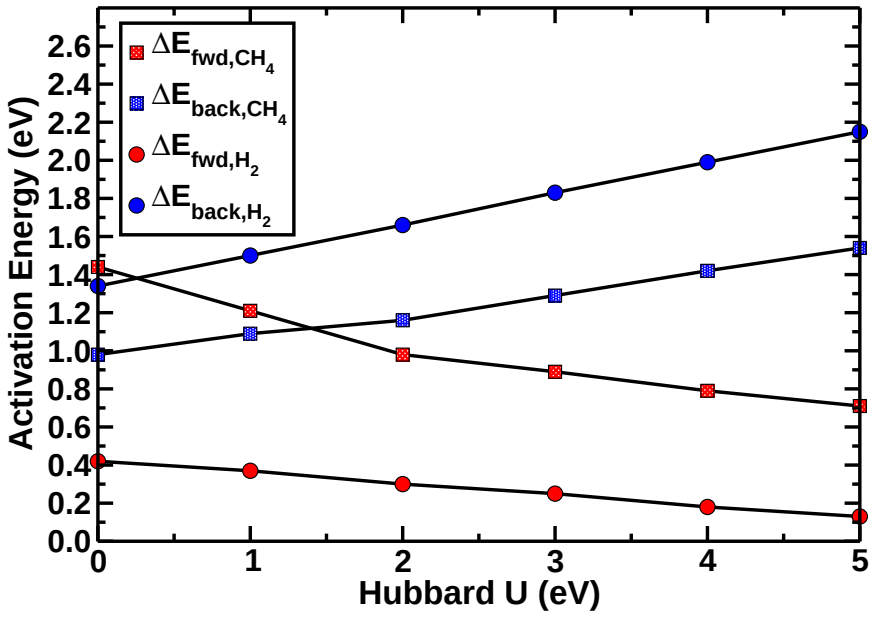


Figure 3-6: Forward and back-reaction barriers as a function of increasing U from 0 eV (GGA) to 5 eV (the global $U_{scf,av}$) for the second quartet barrier of both reactions with methane and hydrogen.

Table 3-6). The GGA ^{4}TS-2$_{CH_4}$ possesses rather long Fe-C and C-O bond lengths, suggestive of a weakly bound methyl radical and $FeOH^+$ complex. However, as U increases above 2 eV, the Fe-C bond shortens, and there is an overall increase in the CH_3-$FeOH^+$ binding for the transition state. The GGA+U results clearly favor a bound transition state, while the GGA results remain ambiguous. The relative energy to dissociated radical products is even more key in interpreting the ^{6}TS-2 results, as we will later show. While the GGA+U quartet surface barrier is larger than its hydrogen counterpart, ^{4}TS-2$_{CH_4}$ remains below the reactant energies and requires less rearrangement than the sextet surface does. Generally, at the exit channel we observe a shallower quartet surface with GGA+U than we did with GGA, and we also observe an increase in the back reaction barrier consistent with increased exothermicity of the reaction.

The second barrier along the sextet surface is highly sensitive to the value of the Hubbard U but in divergent ways for the hydrogen and methane reactions. Along the hydrogen reaction coordinate, the GGA sextet surface is relatively steep with an activation energy of 1.29 eV. The GGA+U activation barrier is markedly decreased from GGA to 0.82 eV with the $U_{scf,av}$ of 5 eV (see Table 3.4). We have previously shown that this value is an underestimate, and the decrease in activation energy with the value of U for the hydrogen reaction is due to two factors. First, the U_{scf} calculated from the $3d$ states is low at about 3.27, and second, the $4s$ electrons play an increasingly important role in bonding in this region of the sextet reaction coordinate. The GGA geometry of ^{6}TS-2, however, is not dissimilar to the ^{4}TS-2 structure with an Fe-H bond length of 1.50 Å and stretched O-H bond of 1.65 Å, which elongates further by about 0.03 Å for GGA+U. The different reaction barrier behavior of the two spin surfaces originates from both electronic structure differences (as evidenced by the differing values of linear-response U) and the differing degrees of displacement from the Int-2 geometry required to reach the transition state.

Specifically, the sextet reaction coordinate is generally derived from Fe^+ atomic states which are $3d^64s^1$ in configuration (unlike the quartet $3d^7$). The role of the $4s$ becomes most pronounced for weaker, long-distance binding interactions associated with the stretched bonds of the transition state. By extending the linear response calculation to a matrix formulation which takes into account response of the $4s$ and $3d$ simultaneously, we obtain a $U_{scf,4s}$ of 4 eV and an increased $U_{scf,3d}$ of 4 eV as well. While a matrix formulation was already present in the original derivation[7], its application to studying both $3d$ and $4s$ electrons simultaneously has not been as crucial for solid-state applications as it is here in molecules with strong $3d$-$4s$ hybridization as well as close spacing of the $3d$ and $4s$ levels. By applying a Hubbard correction on both the $4s$ and $3d$ electrons (4 eV for both, as mentioned before and

	CH_4		H_2	
U	ΔE_{fwd}	ΔE_{back}	ΔE_{fwd}	ΔE_{back}
0	1.87	1.26	1.22	2.01
1	1.68	1.39	1.19	2.23
2	1.51	1.52	1.10	2.41
3	1.36	1.69	1.00	2.64
4	1.18	1.79	0.90	2.86
5	**1.03**	**1.90**	**0.82**	**3.01**

Table 3.4: Forward and back reaction barriers as a function of U for the second sextet barrier of both reactions with methane and hydrogen.

as obtained from linear response), we obtain an improved estimate of the barrier height of 1.16 eV, as compared to the 0.82 eV value obtained from U_{3d}=5 eV. The back reaction barrier remains unchanged at 3.01 eV. Both forward and back reaction barrier estimates are in improved agreement with the CCSD(T) values of 1.1 and 3.0 eV, respectively. Alternatively, a locally averaged U_{scf} of the $3d$ states alone, 3.6 eV, provides an improved barrier estimate of about 0.96 eV. For cases where the role of the $3d$ states are reduced, a low U_{scf} can provide a useful reminder to consider also binding interactions due to $4s$ electrons which are otherwise overshadowed by the $3d$ states for complexes with short bond lengths and strong metal-ligand $3d$ hybridization.

The methane ^{6}TS-2 barrier varies strongly with Hubbard U by decreasing drastically from 1.87 eV for GGA to 1.03 eV for GGA+U (see Table 3.4). Unlike ^{6}TS-2_{H_2}, this transition state exhibits a larger $U_{scf,3d}$ of 4.35 relatively close to the $U_{scf,av}$ value of 5 eV, while the linear response U_{4s} is nearly zero, highlighting how the formalism permits detection of which manifolds are most relevant. The GGA ^{6}TS-2_{CH_4} structure, like the quartet analog, resembles a methyl radical loosely bound to an $FeOH^+$ fragment. The GGA coordination distances of 2.83 Å for Fe-C and 2.33 Å for C-O suggest that very diffuse interactions are responsible for binding. The long

Fe-C bond length in the GGA structure corresponds to weak bonding via overlap of the Fe 4s density with the C 2p which is also weakly anti-bonding with respect to C-O bonding. The loosely bound GGA structure points to significant competition between the elimination step and dissociation of the transition state to radical side products which are, in fact, both endothermic by 0.4 eV and around 0.05 eV, respectively. The similarity of the GGA TS-2^6 to a free methyl radical and $FeOH^+$ moiety is further substantiated by comparing the transition state geometries with isolated products. The C-H bond lengths and H-C-H angles are nearly identical to the isolated methyl radical (1.085 Å and 119°), and there is a very slight distortion from the plane of CH_3 with a dihedral of 170°. The nearly dissociated nature of the ^{6}TS-2_{CH_4} GGA structure is much greater than in the quartet case as a result of the spins of the isolated radical fragments - quintet $FeOH^+$ and doublet CH_3 - which couple with greater bonding interaction if the spins are opposed, as in the quartet, than if the majority spins match, as is the case in the sextet.

Bond lengths for the more tightly bound GGA+U (U= 5 eV) transition state are shortened to 2.49 Å for Fe-C and 2.26 Å for C-O. The GGA+U structural differences are largely due to an increased population of three-center bonds involving the Fe 3d_{xz}, O 2p, and C 2p orbitals in lieu of weaker interactions between Fe 4s and C 2p. In order to demonstrate in greater detail how the Hubbard U term tunes the TS-2^6 binding strength, we measure the changes in Fe-O bond length (Δr_{FeO}) as well as the difference between the TS-2^6 CH_3 dihedral and the planar value in the isolated methyl radical ($\Delta\angle_{\mathrm{CH_3}}$). With increasing values of U up to 2 eV, we at first see a decrease in the Fe-O bond and slight decrease in the dihedral (see Table 3.5). However, for U= 3 eV, the difference in the Fe-O bond length increases as charge transfer out of the Fe-O bond and into the CH_3 moiety increases concomitant with a dihedral change from an increased polarization of the Fe-O bond. These changes

U	$\Delta E_{\mathrm{Rct-TS2^6}}$	$\Delta E_{\mathrm{FeOH^+-TS2^6}}$	Δr_{FeO}	$\Delta \angle_{\mathrm{CH_3}}$
0	0.02	0.36	3.16%	5.75%
1	-0.22	0.48	2.88%	5.69%
2	-0.46	0.56	2.61%	5.24%
3	-0.67	0.61	4.83%	9.10%
4	-0.92	0.68	4.95%	9.64%
5	**-1.07**	**0.78**	**5.13%**	**10.02%**

Table 3.5: ^{6}TS-$2_{\mathrm{CH_4}}$ structural and energetic changes with U. The ΔE are in eV and the structures are percent difference from the bond length or dihedral for the isolated moiety at the same value of U as in the transition-state.

in electronic structure are evidenced by a 5% increase in Fe-O bond length and 10% difference of the dihedral in the U= 5 eV structure. The transition state is also stabilized by over 0.6 eV when compared to the dissociated products and a now weakly exothermic side product. An accurate estimate of the relative height of this barrier both to reactants and to the radical side products is a particularly sensitive test for any *abinitio* method; GGA+U is in very good agreement with experimental findings and the MRCI reference calculation, while GGA fails to provide a physically reasonable description of the reaction.

3.9 The third reaction intermediate

The exit channel Int-3 resembles an isolated Fe^+ ion weakly ligated by the leaving oxygen of either the water or methanol (see Fig. 3-2). The weak metal-ligand interaction produces many closely-spaced low-lying electronic states differing only by their minority spin orbital occupations, which presents a challenge to resolve ordering of states separated by a few hundredths of an electron-volt. Based upon the $\Delta E_{6\rightarrow 4}$=0.23 eV splitting of isolated Fe^+, the sextet Int-3 should be most stable, but common density functionals erroneously predict the quartet instead to be most

stable[17, 79, 80]. Since the metal-ligand binding is equally weak in both reactions, the Int-3 structures of both reactions should be studied together. For both the case of the quartet and the sextet, the spin up electron configuration is $3d^5$, and so these electrons play no role in the state ordering. In the sextet case, a single $3d$ electron resides in the spin-down manifold, and the π and δ orbitals retain their degeneracy. For all ^{6}Int-3 structures, occupation of a δ orbital which minimizes overlap with any of the water molecules is preferred (see Fig. 3-7). The occupation of a π orbital yields an electronic state roughly 0.1 eV higher in energy. Population of the σ orbital is prohibitive, as it corresponds to the maximal overlap with the remainder of the electron density, and it resides roughly 0.5 eV above the lowest state. None of these relative symmetry orderings change significantly with increasing values of U, although the geometry does change. The two lowest GGA sextet Int-3_{H_2} states have a long Fe-O bond length of about 2.06 Å which lengthens to 2.14 Å for GGA+U. For the methane case, the GGA equilibrium Fe-O bond length is slightly shorter, about 2.03 Å, but it lengthens the same amount to about 2.11 Å with GGA+U.

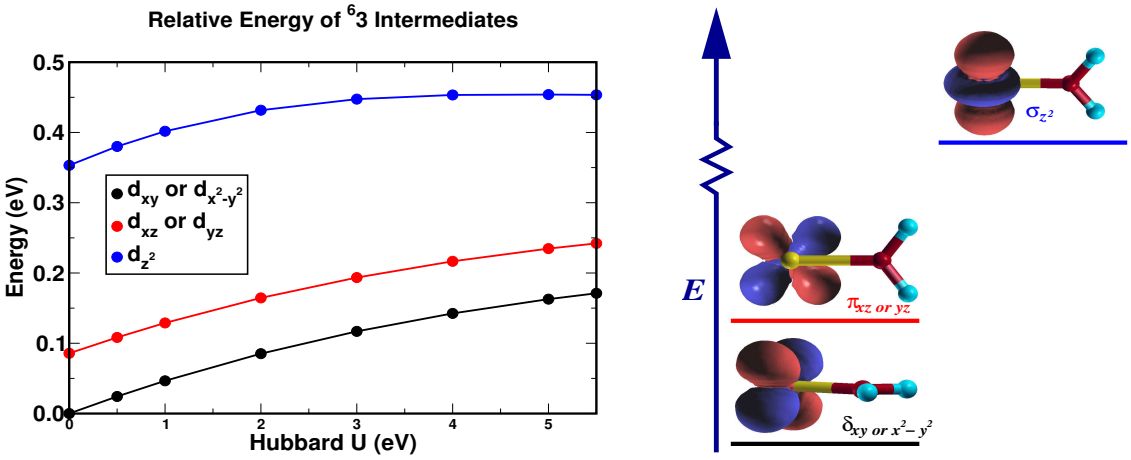


Figure 3-7: (a) Relative Energy in eV of Int-3 sextet states (b) Qualitative Energy Diagram of Int-3 sextet states.

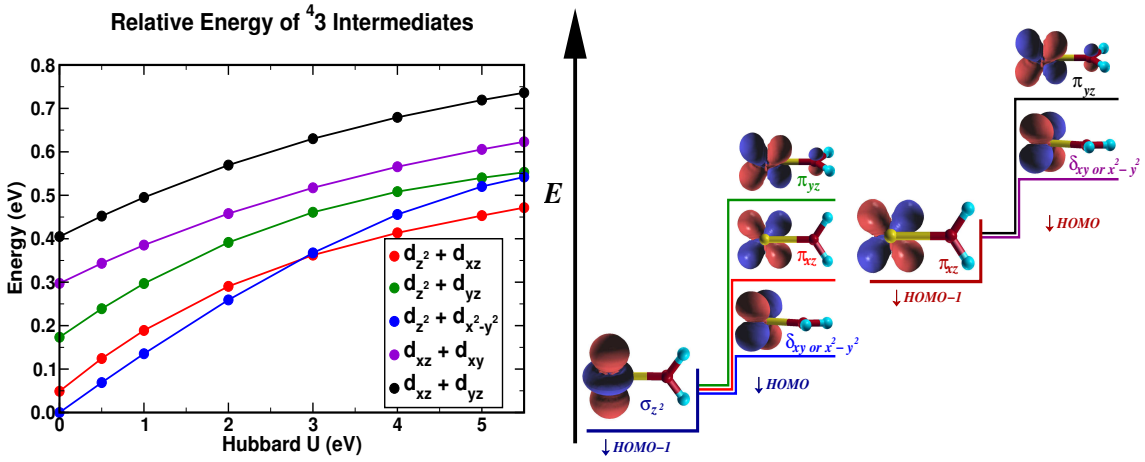


Figure 3-8: (a) Relative Energy in eV of Int-3 quartet states (b) Qualitative Energy Diagram of Int-3 quartet states.

For the quartet Int-3 structures, GGA Fe-O bonds are about 1.93 Å for the lowest Int-3_{H_2} states and lengthen to about 2.03 Å for U= 5 eV. The methanol complex is similar with a GGA bond length of 1.90 Å extending to 1.99 Å for GGA+U. With quartet states, a greater number of permutations of the occupied orbitals is possible, and the degeneracy present in the sextet Int-3 π and δ orbitals is now broken (see Fig. 3-8). Population of the $\pi_{d_{xz}}$ orbital yields states lower in energy than those occupying $\pi_{d_{yz}}$ because $\pi_{d_{yz}}$ forms a weakly repulsive anti-bonding orbital with stray $2p_z$ density from the oxygen. The two lowest states favor instead σ orbital occupation in combination with either a $\delta_{d_{x^2-y^2}}$ or $\pi_{d_{xz}}$ orbital. The symmetries which are derived from alternative permutations of the two $3d$ electrons reside roughly 0.1 to 0.4 eV above these two states. For GGA, the lowest state occupies the δ orbital and is favorable by about 0.04 eV, but at a U of around 3 eV, the two lowest states reverse their ordering and, instead, the π occupation is preferred by about 0.09 eV (at U = 5 eV). However, regardless of the U applied, the occupation of a $3d_{z^2}$-like orbital is

strongly favored.

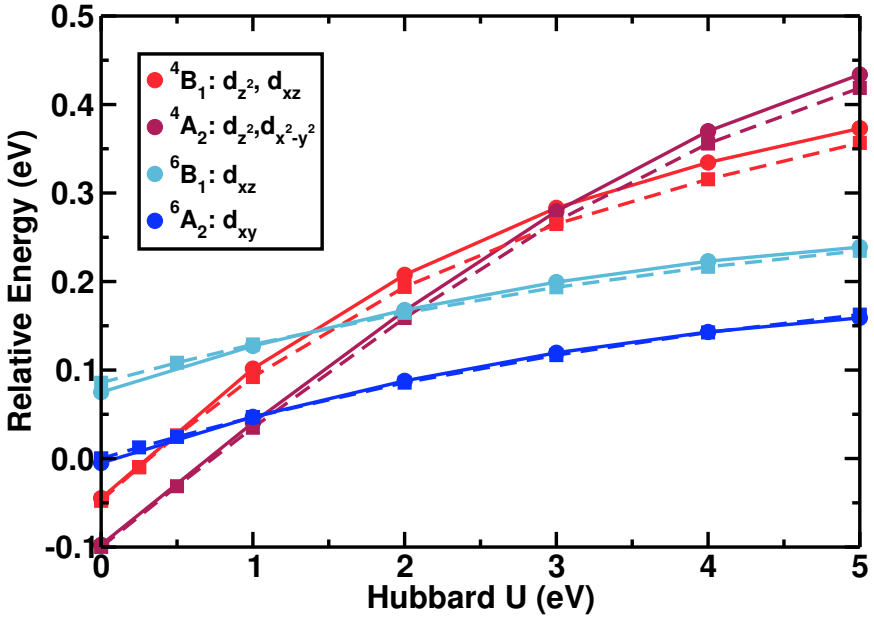


Figure 3-9: Comparison of the energy of the two lowest lying sextet Int-3 states (blue) and quartet Int-3 states (red) for both methanol (solid lines) and water (dashed) ligands.

While GGA erroneously predicts the Int-3 ground state to be a quartet, with increasing U the two lowest sextet states become lower in energy than the lowest quartets. This is an improvement upon GGA, which found the quartet to be lower in energy by 0.1 eV and, thus, does not predict spin inversion at the exit channel. At a U of 5 eV, we find the sextet-quartet ordering reversed and a state splitting of about 0.21 eV. This value may be an overestimate compared to our own and other CCSD(T)[82] calculations, which place the sextet to quartet splitting closer to 0.12 eV. Because the splitting is very sensitive to the application of the Hubbard U, it is advantageous to obtain the splitting at the true U of this system. Each state has an individual U_{scf} which should be calculated separately and then averaged

locally over all ^{4}Int-3 and ^{6}Int-3 states, rather than using the globally averaged U_{scf} which is used to construct a global reaction coordinate. By employing an average U_{scf} of 3.5 eV from the two lowest electronic states, the GGA+U splitting is found to be 0.12 eV, in improved agreement with CCSD(T) results. The lowest sextet and quartet GGA+U electronic states also exhibit consistent symmetry with those obtained from CCSD(T). Interestingly, as we show in Figure 3-9, the state splittings and their dependence on Hubbard U is nearly independent of ligand identity, thus showing that the GGA errors and GGA+U corrections are both systematic in nature. This analysis suggests that Int-3_{CH_4} sextet-quartet splittings at a locally averaged U, 0.13 eV, should be an improved estimate over the 0.22 eV value at the global $U_{scf,av}$.

3.10 Discussion of the reaction coordinates

We have shown that for the addition-elimination reaction of hydrogen on FeO^+, the GGA+U reaction coordinate provides a substantial improvement over GGA when compared against experiment or highly-accurate CCSD(T) reference calculations. The GGA reaction coordinate (see Figs. 3-10 and 3-11) is in direct contradiction to available experimental evidence and theory because it 1) underestimates the exothermicity by over 1 eV, 2) the quartet surface is too steep to rationalize a "shallow, barrierless" reaction[74, 80], 3) spin-splittings at intermediates show large errors, and 4) there is no spin-crossover at the exit channel[3]. The majority of these errors in the GGA reaction coordinate are localized to the second reaction barrier (see Fig. 3-11), with the exception of the errors in the quartet Int-1 structures and barriers in the first barrier surface (see Fig. 3-10). While the first GGA sextet TS barrier largely

[3]The reaction coordinates of different methods have been aligned at the $^6\text{Int} - 1$. The curves associated with GGA and GGA+U reaction coordinates are splines of points from the minimum energy path as determined by NEB calculations.

agrees with CCSD(T) results, the second GGA barrier exhibits much greater errors for the exothermicity and state ordering at Int-3. In contrast, the GGA+U (U= 5 eV) reaction coordinate (see Figs. 3-10 and 3-11) correctly estimates and predicts the reaction exothermicity, the shallow but excited quartet surface, barrier heights, and spin crossover at the exit channel (see Figure 3-11). Systematic improvements on GGA+$U_{scf,av}$ results on the largest discrepancies (e.g. at the hydrogen ^{6}TS-2 barrier) were achieved by including explicitly the 4s manifold, highly relevant for isolated transition metal complexes, or by locally averaging the value of U_{scf}. Additionally, the portions of the reaction coordinate which are properly estimated by GGA in the first TS barrier (see Fig. 3-10), namely the sextet barrier steepness and structures, are preserved in GGA+U, while the errors of the GGA approach in the second TS barrier (see Fig. 3-11) are greatly improved upon.

In order to further measure the quantitative accuracy of the GGA+U hydrogen on FeO^+ reaction coordinate, we estimated equilibrium CCSD(T) geometries for each intermediate. The low dimensionality of the potential energy surface permitted calculation of a fine mesh of single point energies about 0.01-0.02 Å apart and deduction of a minimum energy configuration. These geometries are also useful to determine errors in geometries in the GGA and GGA+U methods. For transition states, we used GGA+U or GGA geometries to calculate single point CCSD(T) energies since discretizing the PES at the transition states was too computationally expensive[4].

To further measure the quantitative accuracy of the GGA+U reaction coordinate, we average the error of the GGA+U and GGA splittings for each intermediate with respect to the CCSD(T) splittings for that intermediate. We also compare here to B3LYP results which we obtained but have not discussed in detail because they

[4]We note that energetic differences at the CCSD(T) level of the different TS geometries were small and did not noticeably affect barrier height estimates.

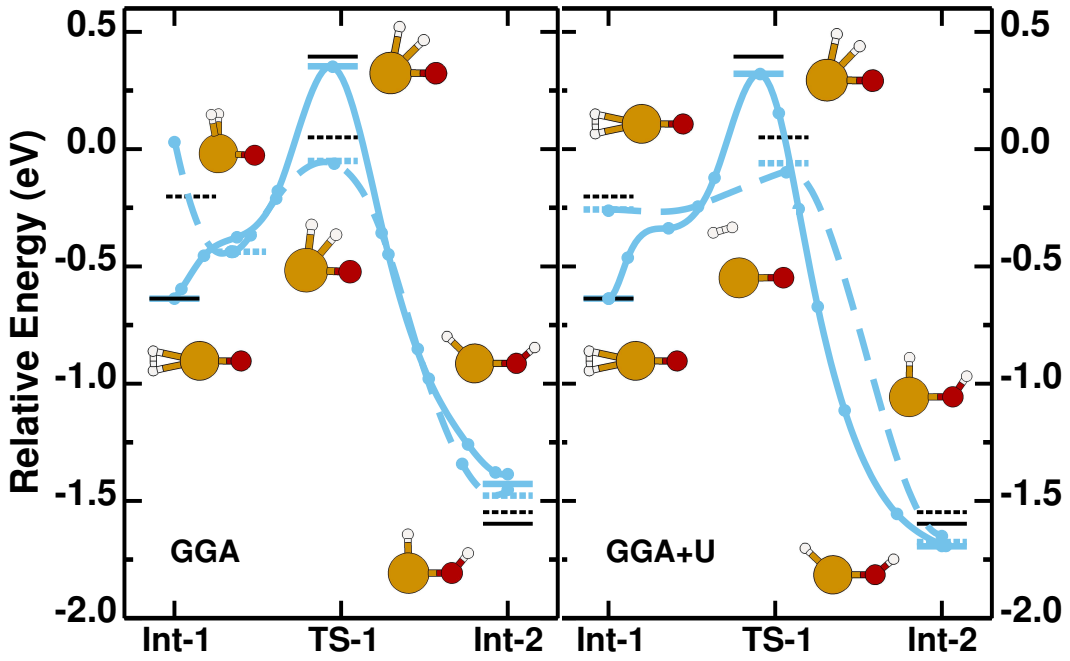


Figure 3-10: The TS-1 NEB minimum energy paths (blue) for GGA (left) and GGA+U (right) are compared against CCSD(T) (black) reference energies for both quartet (dashed) and sextet (solid) surfaces of the reaction of FeO^+ with H_2. Here, U is the globally averaged $U_{scf,av}$ of 5 eV. The results of the different methods were aligned only once at ^{6}Int-1.

exhibited the same inaccuracies which were previously discussed in the literature[79, 80]. The average error for the five spin splittings along the reaction coordinate is greatly reduced to 0.04 eV with GGA+U from GGA which exhibits an average error of 0.20 eV. The commonly employed B3LYP functional performs poorly with an average error of 0.30 eV in splittings. Not only does GGA+U improve errors by five-fold over GGA, but using the CCSD(T) geometries we found from interpolating a potential energy scan, we also observe that GGA+U on average produces improved geometries. Mean errors for GGA+U geometries are reduced from 4.3 pm (for GGA) to 2.2 pm, and are slightly worse than B3LYP geometries with an average error of

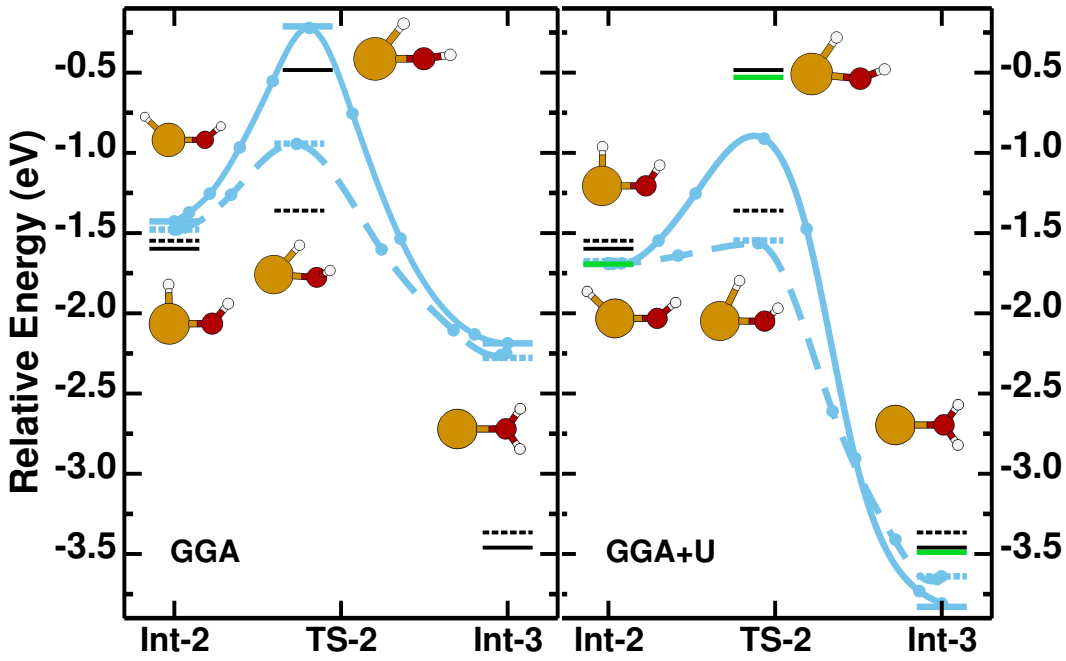


Figure 3-11: The TS-2 NEB minimum energy paths (blue) for GGA (left) and GGA+U (right) (both in blue) are compared against CCSD(T) (black) reference energies for both quartet (dashed) and sextet (solid) surfaces of the reaction of FeO^+ with H_2. The results of different methods were aligned only once at ^{6}Int-1. The GGA+U results were determined at the globally averaged $U_{scf,av}$ of 5 eV, excluding the points calculated at locally averaged U_{3d}= 4 eV and U_{4s}= 4eV (shown in green).

1.3 pm. Although the B3LYP functional shows a slight improvement in geometries, the seven-fold improvement in spin-splittings GGA+U provides is much more key. The improvement of the B3LYP geometries over GGA and GGA+U could also be attributed to differences between the bonds hydrogen forms in all-electron, localized basis set methods and plane-wave pseudopotential methods.

The GGA description of FeO^+ on methane demonstrates even greater discordance with theory and known experimental results than the hydrogen reaction. In fact, using GGA we predict both the reactions to create methanol and to form radical side products to be endothermic by 0.2 and 0.4 eV, respectively (see Fig. 3-12).

	$TS-1^4$		$TS-1^6$		$TS-2^4$		$TS-2^6$	
Method	CH_4	H_2	CH_4	H_2	CH_4	H_2	CH_4	H_2
GGA	-1.14	-1.04	-0.54	-0.59	0.46	-0.92	0.61	-0.79
GGA+U	-1.07	-1.42	-0.71	-1.05	-0.84	-2.02	-0.87	-2.19
Post-HF	-0.99	-1.35	-0.75	-0.96	-0.87	-1.82	-0.80	-1.87

Table 3.6: Net energy required for reaction steps (in eV) for GGA, GGA+U (U= 5 eV), and Post-HF methods, where the Post-HF method for the H_2 reaction is CCSD(T) and the CH_4 reaction is MRCI.

Additionally, the sextet TS-2 barrier resides slightly above the reactant energies. The remaining GGA discrepancies are comparable to those for the reaction with hydrogen and include spin-splittings at intermediates, barrier estimates, and the absence of a spin-crossover at the exit channel. When we instead compare our GGA+U results to highly accurate MRCI calculations, we observe quantitative agreement between the two methods (see Fig. 3-13). The GGA+U quartet surface is shallower but not barrierless, in agreement with MRCI results. We should also note that experimental results point to a possible kinetic isotope effect for the methane reaction, though none is observed for the hydrogen reaction[73], suggesting that classical barrier estimates may need also to be considered. We see an improved estimate of exothermicity of both main and side reactions and now also preserve spin crossover at the exit channel. The description of the quartet Int-1 geometry is also improved. A comparison of the accuracy of GGA+U geometries is beyond the scope of this study due to the higher dimensionality of the PES over that for the four atom reaction.

We can further examine the accuracy of GGA+U in treating the energetics of these reactions by decomposing the exothermicity for all spins and reaction steps, as summarized in Table 3.6. While GGA produces errors as large as 1.4 eV for the methane reaction and 1.1 eV for the hydrogen reaction, GGA+U is in very good agreement with accurate quantum chemistry. The largest errors, such as the 0.32

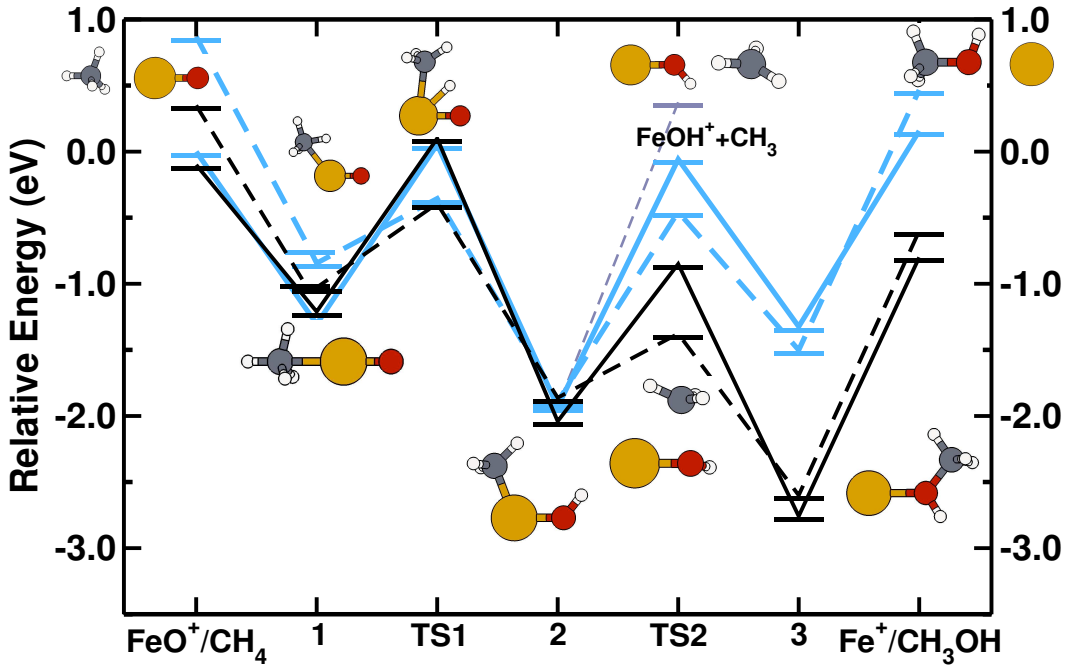


Figure 3-12: Comparison of GGA (blue) and MRCI (black) reaction coordinates for quartet (dashed) and sextet (solid) surfaces of the reaction of FeO^+ with CH_4.

eV underestimate at the sextet TS-2_{H_2} barrier can be systematically improved by extensions that either include a U_{4s} or a locally averaged U_{scf}.

Using the PBE-GGA approximation, we find that activation barriers, stationary-point spin splittings, and reaction energies exhibit errors as large a 1.4 eV for both the methane and the hydrogen systems discussed. By augmenting our GGA exchange-correlation functional with a "+U" term, we reduced these errors by an order of magnitude to, on average, 0.1 eV with respect to the best available quantum chemical methods (such as CCSD(T) and MRCI) as well as experimental values. Importantly, the Hubbard U term is not used as a fitting parameter but it is a true linear-response property of the transition metal complex which may in principle augment any exchange-correlation density functional. The practical limitations to this

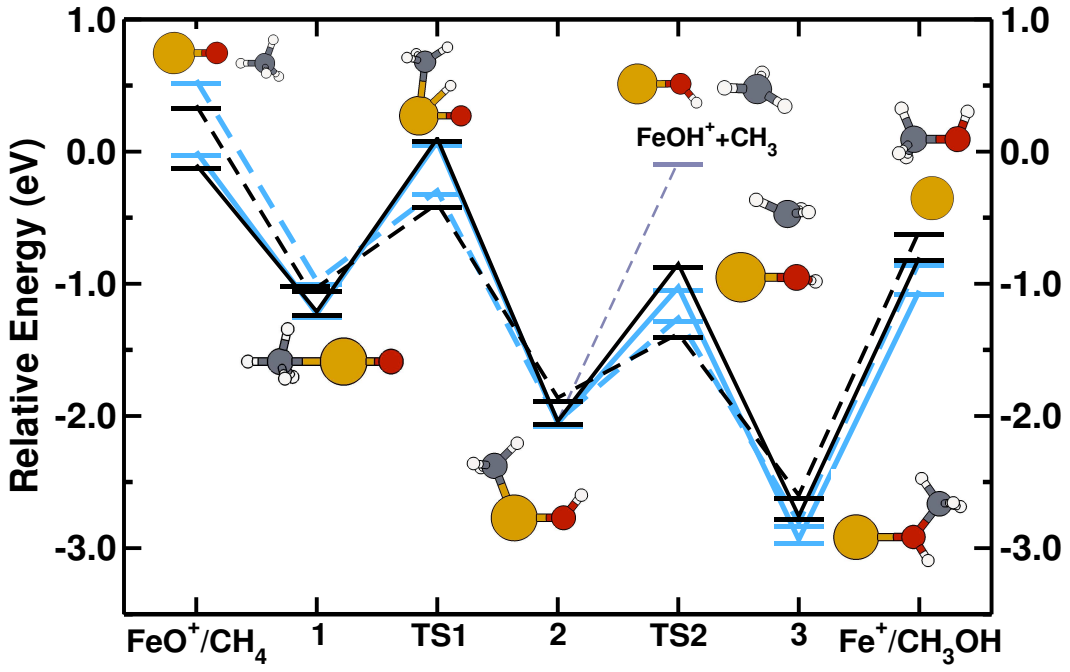


Figure 3-13: Comparison of GGA+U (U= 5 eV) (blue) and MRCI (black) reaction coordinates for quartet (dashed) and sextet (solid) surfaces of the reaction of FeO^+ with CH_4.

approach stem from the complexity of the systems which we wish to study: evolution of the coordination environment along a global reaction coordinate may result in local deviations from a globally averaged Hubbard U or bring to the forefront the role of 4s contributions. Overall, the DFT+U results (here, GGA+U) have been shown to provide for systematic improvement over GGA for all systems considered thus far by treating, for the first time within a DFT framework, the energetics of differing spin surfaces and electronic states with the same level of accuracy. The inexpensive linear-response U calculation also acts as a probe for the relative utility of the DFT+U approach for a given transition metal complex; that is, if the Hubbard U calculated is small or nearly zero, a standard description may be sufficient. We believe that this work paves the way for studying a myriad of large-scale systems

which contain transition metals with both accuracy and efficiency.

3.11 Further considerations: accuracy of quantum chemical approaches

While CCSD(T) results on the hydrogen addition-elimination reaction appear to agree well with experiment and our GGA+U results, we also employed a multi-reference approach to reconcile previous inconsistencies with older high-level, multi-reference calculations. Early CASPT2D calculations indicated some discrepancies with both experiment and our own CCSD(T) results particularly at the Int-2 splitting. The CASPT2D method zeroes all off-diagonal elements of the Fock matrix, and is a method which was used historically to reduce computational cost of the CASPT2 method[77]. To verify that our CCSD(T) method produced results consistent with a thorough CASPT2 result, we completed CASPT2 calculations with our 6-311++G(3df,3pd) basis set using MOLPRO[102]. We chose as a reference the energy of $^6\Sigma^+$ FeO^+ and used CCSD(T) geometries for the CASPT2 calculations. We found that while previous results identified an 11 kcal/mol separation for the Int-2 quartet and sextet, our CASPT2 results are much more consistent with CCSD(T)

	Sym.	Largest Coeff.	Percent of Wavefunction
$^6FeO^+$	$^6\Sigma^+$	0.92	87%
2^4	A''	0.90	84%
2^6	A'	0.98	97%
3^6	A_2	0.99	99%

Table 3.7: Electronic state symmetries obtained in CASPT2 which are in agreement with previously published results as well as our own. We also provide the coefficient of the dominant determinant as well as the contribution that this state makes to the overall wavefunction.

	CCSD(T)[a]	CCSD(T)[b]	CASPT2D[c]	CASPT2[d]
2^4	-34	-30	-14	-28
2^6	-35	-31	-25	-31
3^6	-76	-70	-67	-69

Table 3.8: Comparison of relative energies in kcal/mol of intermediates along the FeO^+/H_2 reaction surface referenced to the total energy of $^6\Sigma^+$ FeO^+ and H_2. (a) Our own CCSD(T) numbers in this work with CCSD(T) geometry and 6-311++G(3df,3pd) basis set. (b) Ref. [82] used modified Ahlrichs TZVP basis for Fe and Pople 6-311++G(2df,2p) for H and O with B3LYP geometries. (c) Ref. [78] used ANO [8s7p6d4f2g] basis for Fe, [3s2p1d] and [5s4p3d2f] for H and O, BP86 geometries. (d) Our own CASPT2 with 6-311++G(3df,3pd) basis set at CCSD(T) geometries.

(see Table 3.8). All three of our approaches predict a 1 to 3 kcal/mol splitting. Also, the relative energy of the ^{6}Int-3 is placed similarly in all three methods at about -69 kcal/mol to -76 kcal/mol. We therefore rationalized the discrepancies in previous studies and also demonstrated that the dominant configuration for all three states considered made up in most cases over 90% of the total wavefunction (See Table 3.7). This means that a single reference approach such as CCSD(T) is likely sufficient for this system in any case.

Chapter 4

Metalloporphyrins

Metal-oxo complexes of the mid-to-late transition metals play key roles as catalysts in biochemical redox reactions of both natural and synthetic origin. Porphyrins are a particularly well-known organic scaffold upon which inorganic transition metal catalysts are built. The porphyrin is a heterocyclic macrocycle which consists of four pyrrole groups linked together via carbon bridges which preserve Hückel's rule for aromaticity (that is, 4n+2 π electrons are delocalized over the molecule)[107]. The porphyrin is often functionalized to lend unique properties to the molecule, and it is referred to simply as a porphine in its unfunctionalized form. A metal may bind readily at the center of the porphyrin with four planar bonds to neighboring nitrogens stemming from the pyrrole groups. This structure leaves the metal open to bonding in the two remaining positions known as the axial sites. The metal oxidation state within the porphyrin is most typically nominally +2 or +3 as a result of hybridization with the planar ligands[108]. This oxidation state preference makes the middle transition metals most amenable to reactive chemistry in terms of donating bonding electrons, particularly manganese, iron, and cobalt.

One of the most ubiquitous instances of iron porphyrin is in hemoglobin, a major component of blood in humans which selectively binds and releases O_2 at the axial site depending upon the local concentration of oxygen gas[109]. High-valent oxoiron porphyrins are also involved in the substrate oxygenation reactions of the well-studied cytochrome P450 and in oxidation reactions of other enzymes such as chloroperoxidase [6]. Subsequently, numerous substitutions of the porphyrin scaffold to include other transition metals which would be best suited to carry out widely varying chemistry have been completed. Porphyrins remain of great interest in study in a number of cases because of their flat structure. The porphyrin rigidly grasps the transition metal atom but can also easily be held to a protein or surface via bonding from one side of the flat molecule while the other side remains open to reactions or binding. Additionally, the system is well characterized, and its synthesis is straightforward as is its spectroscopic characterization[107, 110]. I will describe in detail three different porphyrin systems which I have studied that each are of interest for different reasons. The first, Mn(oxo)-porphyrins, are of interest to understand the discrepancies in describing porphyrin systems with different functionals. In the second case, we will use unfunctionalized cobalt porphyrins in order to understand the mechanism of CO_2 reduction to CO on negatively charged porphyrins. Finally, I will describe a newly observed phenomenon of delocalized spin on bromophenyl functionalized cobalt porphyrins when adsorbed on a Cu(111) surface which has great promise for the field of spintronics. Ultimately, the porphyrin remains a great tool for scientists - experimentalists and theorists alike - to use in order to create reaction pathways which catalyze novel functionalizations.

One practical consideration for many of these porphyrin systems is that in order to further reduce computational cost and to also more easily assign electronic states and orbital symmetries, we enforce the C_{2v} symmetry - having xz and yz mirror

planes - of the molecule in our calculations. We note, however, that for porphyrins and the other molecules we have studied thus far, imposing symmetrization of the charge density associated with the imposed ionic symmetry does not significantly alter the final self-consistent solution. Such an approach does break down in the limit of a single transition metal atom which has an infinite number of symmetry operations but for which density functional theory necessarily finds the energy of an orbital of d_{z^2} type to be inequivalent to the other four $3d$ orbitals. The total energy of the symmetrized solution in most molecular systems, however, is essentially identical to a solution with no imposed symmetry as long as no unphysical constraint has been made such as forcing an atom to remain in the plane of the rest of the molecule.

4.1 Spin states of Mn(oxo)-porphyrins

Manganese-oxo complexes are thought to be relevant in particular in the photosynthetic oxidation of water to produce oxygen. Biomimetic systems such as Mn(oxo)-salen catalysts are also known to be efficient in hydroxylation and epoxidation of many organic substrates including aliphatic hydrocarbons[111]. It is therefore of great interest to obtain a consistent and complete understanding of the relationship between electronic structure and reactivity in these systems. The short lifetime of intermediates and the close proximity of numerous electronic states in such systems often limits the scope of experiments[112]. The numerous states and spins of each system also present a considerable challenge to theoretical approaches which are still limited to lower than desired accuracy. We consider here Mn(V)-oxo porphyrins which have been studied and well characterized experimentally using spectroscopic techniques such as UV-Vis and ^{1}H NMR[113]. These complexes experimentally contain porphyrin rings that are functionalized to enhance stability. The Mn metal is

proposed to have a high-valent state which is diamagnetic and corresponds to the double occupation of a $3d_{xy}$ atomic-like orbital. For the purpose of computational efficiency, we consider an unfunctionalized Mn(oxo)-porphyrin which is six-fold coordinated with a water ligand bound to the manganese from below for a total of 41 atoms and a net charge of +1. The Mn pseudopotential employed was generated in a neutral oxidation state, which could make it challenging to reproduce experimental results.

GGA and hybrid functionals are known to disagree with respect to state-ordering on similar complexes, and so we employ the GGA+U functional in an effort to provide a consistent picture of state ordering because for a system of this size accurate post-Hartree-Fock quantum chemical calculations are not currently feasible. For this complex, results of our collaborators initially suggested that GGAs such as the BPW91 functional stabilize a singlet state for this complex by more than 0.4 eV with respect to the next lowest state, a triplet. Hybrid functionals, on the other hand, find all electronic states to be closely spaced, with a different triplet and a quintet being the lowest states. These two lowest states differ only by the spin-coupling of the relevant $3d$ electrons and, therefore, it is physically reasonable that they are closely spaced in energy. The singlet is the least stable electronic state for hybrid functional calculations. We completed GGA and GGA+U total energy calculations and structural relaxations on this system in a 34 Bohr cell with a 30 Rydberg cutoff for the wavefunction and a 300 Rydberg cutoff for the charge density.

Our GGA calculations confirm previous BPW91 results. The singlet state is stabilized and a triplet lies 0.4 eV above. Another triplet and quintet are about 0.8 eV above the ground state, and an open-shell singlet lies even higher in energy. We calculated the self-consistent U of each of these states and found the average to be 5.3 eV, slightly lower than the U_0 of 5.38 eV. The range of U calculated for

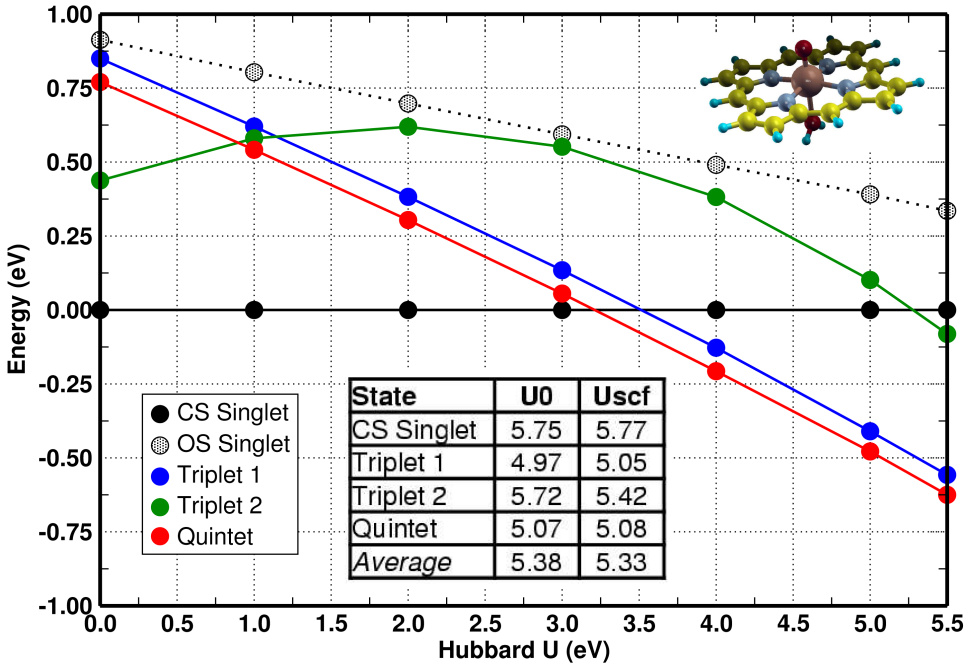


Figure 4-1: Relative Energies of quintet, triplet, and singlet Mn(oxo)-porphyrin states.

the different states was relatively narrow with the largest U being for the singlet, about 5.77 eV, and the smallest U was about 5.05. With GGA+U, the relative state ordering is reversed to agree much more closely with previous B3LYP results. The quintet is the lowest-lying state with the analogous triplet being approximately degenerate in energy to it. For a U of 5.3 eV, both of these states lie about 0.55 eV below the closed-shell singlet which was the GGA ground state. Also, the triplet favored by GGA is stabilized in GGA+U by about 0.05 eV with respect to the closed shell singlet, but it is still highly excited. The open-shell singlet is highly excited in both GGA and GGA+U and is therefore not likely to be relevant. The next step for understanding relative energetics is to study whether one can better mimic a Mn(V)=O porphyrin as it would exist in solution by replacing the water ligand with

an OH group. An -OH ligand should be more electron withdrawing, thus stabilizing the low $3d$ occupation Mn center observed experimentally. Of course, comparison of results using pseudopotentials generated in various oxidation states could be key (as we showed in Chapter 3) to properly describe an Mn(V) center. However, the electron-withdrawing peripheral groups attached to the ring system, which are absent from our current model, could be important in stabilizing the singlet that is observed experimentally.

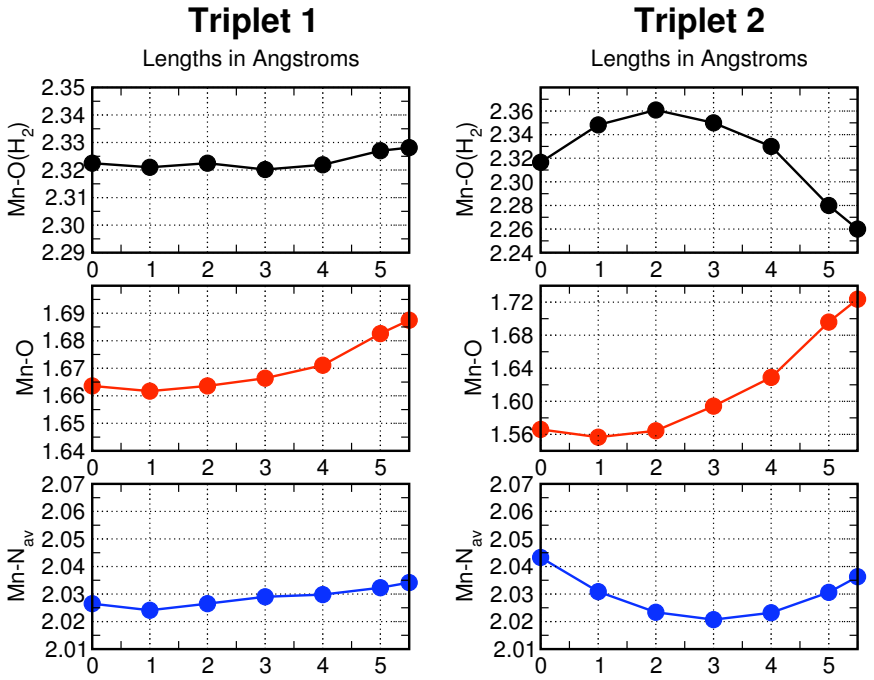


Figure 4-2: Variation in critical bond lengths with U for the two triplet Mn-porphyrin states.

GGA+U affects structural properties of the Mn(oxo)-porphyrins as much as it does electronic properties. While in the case of Fe(oxo)-porphyrins, previously studied in our group, GGA+U appeared to lead to unphysical bond elongation, no such behavior was observed in this case[114]. The most directly relevant orbitals in the

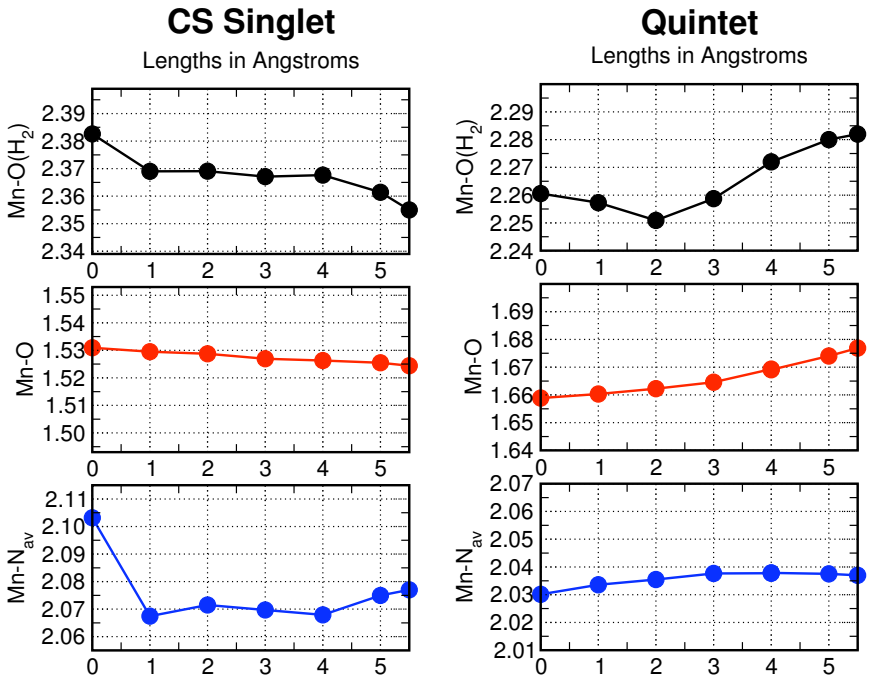


Figure 4-3: Variation in critical bond lengths with U for quintet and singlet Mn-porphyrin states.

manifold of states to which the U is applied includes δ orbitals which lie in the plane of the porphyrin ring and π^* orbitals which are perpendicular to the ring and lie along the Mn-oxo bond. Typically in density functional calculations, π^* orbitals are under-populated in favor of other orbitals which are more delocalized as a consequence of self-interaction errors. In both triplets, we observe an Mn-oxo bond elongation with increasing U that corresponds to an increase in population of the π^* orbitals. In the case of the GGA-stabilized triplet the Mn-oxo elongation is quite large, increasing from 1.56 Å to nearly 1.72 Å (see Fig. 4-2). The GGA+U stabilized triplet demonstrated less bond elongation, increasing only 0.03 Å from 1.66 Å at GGA to 1.69 Å at GGA+U (5.5 eV). As would follow from the singlet electronic structure, there is no occupation of π^* orbitals and the Mn-Oxo bond actually shortens with increasing U

(see Fig. 4-3). The quintet demonstrates comparable bond elongation to its analogous triplet state. We note that in monitoring the Mn-water bond that it was looser for both GGA and GGA+U than previously reported in localized-basis set studies. We anticipate this to be due to the difference in plane-wave and localized-basis set calculations. In particular, localized basis sets with a limited set of diffuse functions can force the electrons to be overly localized. The average Mn-N distance was similar for most states at about 2.03 Angstroms and did not increase much with increasing U. It is more relevant that it out-of-plane displacements (not pictured) were reduced in both the GGA triplet and the singlet with increasing U as well as the other two states (where it was less pronounced, initially, in any case). In comparing corroles to porphyrins, Shaik et. al [6] previously found that the out-of-plane displacement of the metal was more pronounced in corrole systems and also seemed to play a role in stabilizing high-valent low-spin manganese electronic states. Useful future study remains for this system in observing results for other other substituents and coordinating ligands because metal-oxo porphyrins are fundamental oxidative catalysts which are ubiquitous in enzymatic complexes. Ultimately, this system provides a useful test for understanding the root of discrepancies between GGA functionals and hybrid functionals and how such discrepancies may be remedied by application of a Hubbard U term.

4.2 Reduction of CO_2 at cobalt porphines

Mitigation of the amount of CO_2 in the atmosphere is highly relevant for environmental preservation, with potential schemes for converting CO_2 into useful products being of the greatest interest. Electrochemical reduction of CO_2 to CO for the first step in the formation of synthesis gas at cobalt porphine centers has been observed

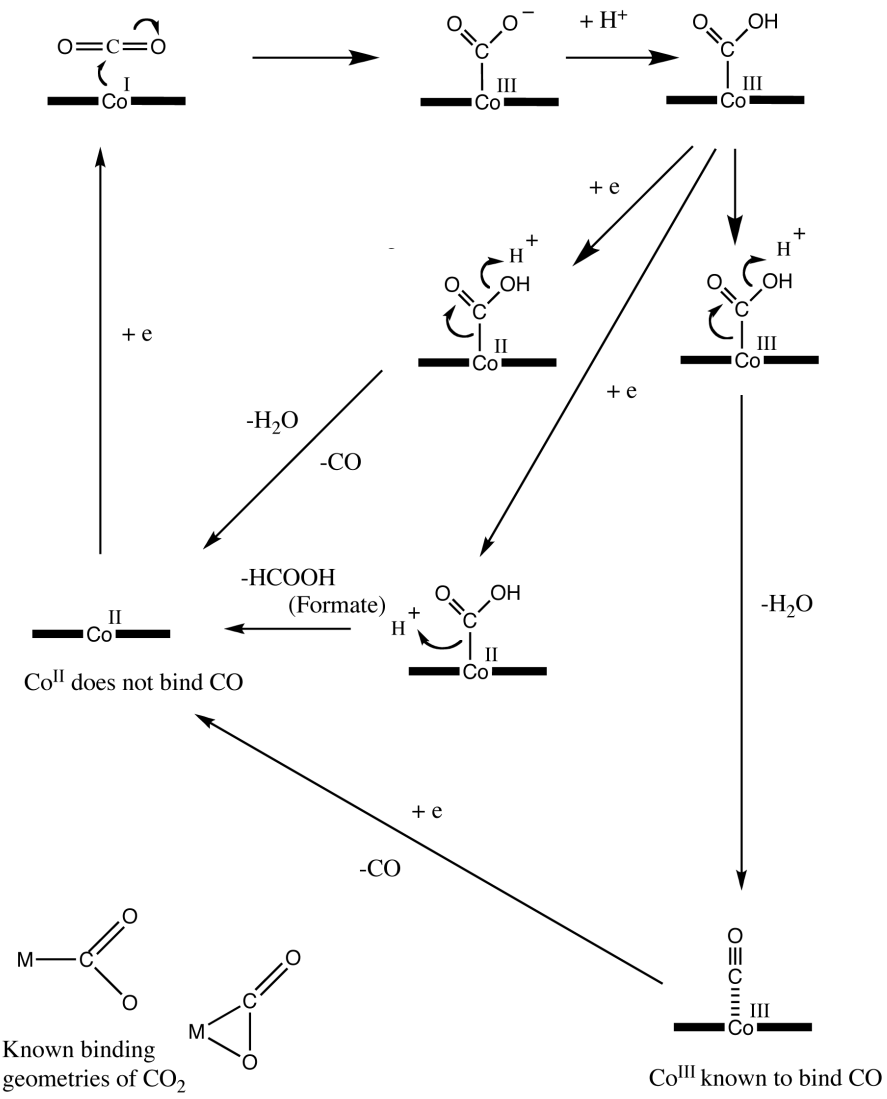


Figure 4-4: Proposed steps and oxidation states for reduction of CO_2 to CO on Co porphine based on previously completed electrochemical experiments on this and other systems.

experimentally[115]. However, the precise mechanism of this reaction is not yet known, and we will thus investigate the relevant oxidation states of cobalt likely to play key roles in the reaction. Based upon similar experiments, there are several potential pathways for the reduction of CO_2 to CO, as shown in Fig. 4-4. Experimentally, the reaction is carried out under aprotic solvent - either in acetonitrile or in high pH solutions - in order to maximize solubility of CO_2 gas[115]. The first step of the reaction is believed to be the binding of CO_2 molecule to cobalt porphyrin in a highly reduced oxidation state (Co(0) or Co(I)). Charge transfer to the CO_2 moiety from Co is then believed to occur resulting in a bound carboxylic acid group. Once the charge transfer has occurred, the CO_2^- is now accessible to attack by water or OH^- to result in a leaving oxygen. This leaves behind a bound CO which may be released under a number of conditions including change in oxidation state to one in which CO is less well bound (e.g. Co(II)). Thus, the first step to identify the most likely reaction mechanisms is to determine the relative binding energies of the different moieties, a most challenging task even for modern DFT approaches.

charge/spin	N/A	CO	CO_2	COOH	Avg.	Max
Co(0) -2/2	5.0	6.3	6.9	–	6.1	6.9
Co(I) -1/1	6.9	7.1	7.4	–	7.1	7.4
Co(II) 0/2	4.8	6.0	5.2	–	5.3	6.0
Co(III) +1/3	4.3	5.1	4.5	5.7	4.6	5.7
Co(III) +1/1	5.8	7.3	5.4	6.1	6.2	7.3
Avg.	5.6	6.3	5.6	5.9	5.9	

Table 4.1: Values of U_{scf} determined for equilibrium structures of various charges and spins of CoP with different bound moieties. The averages for each bound molecule and each oxidation state are also shown.

The initial interest in the project was directed towards identifying the relative binding energies of carbon monoxide and carbon dioxide by these molecules. Ulti-

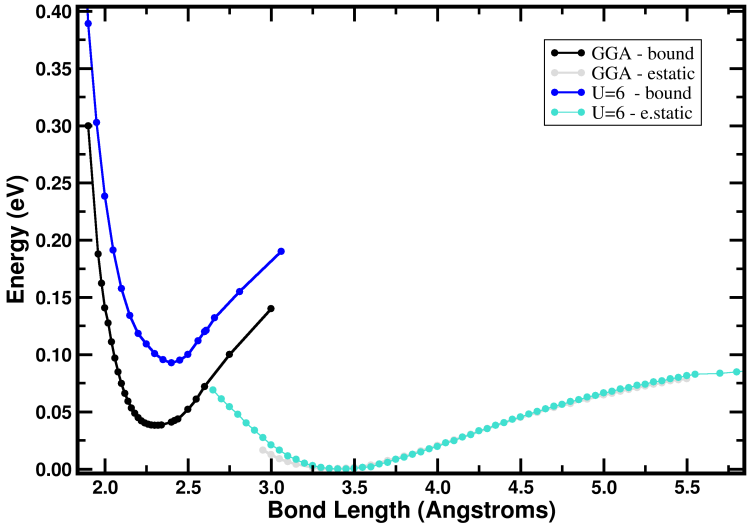


Figure 4-5: Binding energy curves in two regimes for Co(I)P-CO_2. There is a short bond length regime for both GGA and GGA+U as well as a shallow minimum at longer bond lengths due to weak electrostatic interactions.

mately, a reaction coordinate permitting the reduction of carbon dioxide to carbon monoxide will also be investigated, but this depends strongly on the starting complex. It is also well known that binding energies and spin state splittings are very sensitive to poor approximations made in the exchange-correlation functional. We have used GGA+U to study the Co(0)-Co(III) porphines both uncoordinated as well as with CO, CO_2 and COOH bound. The linear-response values of U were calculated for all species and determined to be highest for the CO-bound species (average of 6.3 eV versus around 5.5 eV for the other species). The overall average of all calculated species, 6 eV, was used for the majority of GGA+U binding energy calculations, but the individual values of linear-response U are shown in Table 4.2. It is of note that the values of U do not vary widely between the bare porphine and the system with a bound molecule. Therefore, inclusion of an interatomic distance-dependent

U formulation would not significantly affect the resulting binding energy curves.

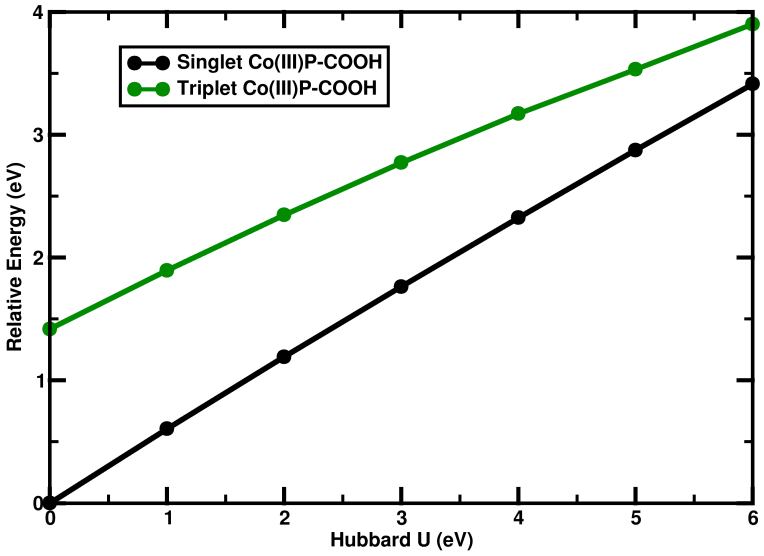


Figure 4-6: Relative energy of singlet and triplet Co(III)P-COOH states versus value of Hubbard U.

Experimentally, it is believed that neutral CO_2 binding to reduced cobalt porphyrin is the first reaction step. Initial investigation showed that oxidized Co(II)P and Co(III)P failed to bind CO_2 at all, while the reduced Co(0)P also failed to bind the molecule and was determined to not be a very likely oxidation state in the experimental system. Instead, Co(I)P-CO_2 was investigated more thoroughly, and the resulting doublet binding curves for GGA and GGA+U are shown in Fig. 4-5. At short distances, a relatively strong binding regime is observed for both GGA and GGA+U at about 2.30 and 2.33 Å respectively. However, at longer bond lengths, a lower energy state emerges with a much longer 3.5 Å equilibrium bond length, indicative of an electrostatic interaction. While such interactions are at the very limit of what semi-local density functionals are likely to properly describe, it is still possi-

ble to consider the implications of this result. The presence of two binding regimes suggests either an avoided crossing produces a secondary long-range minimum or that the shorter bond length, metastable minimum is the result of charge transfer. In order to investigate further the latter option, it is therefore of use to investigate the binding of a carboxylic acid group to the porphine.

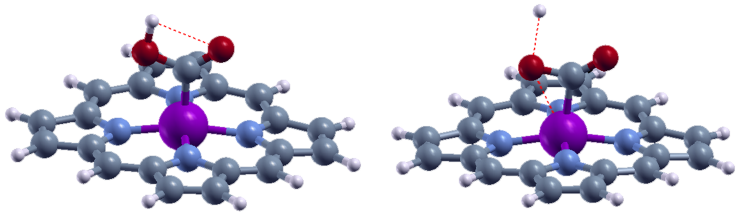


Figure 4-7: Relaxed structure of Co(III)P-COOH GGA+U molecule shows a strongly bound carboxylic acid motif. Stretching of the O-H bond to dissociation leaves behind a CO_2 moiety partially stabilized by a secondary Fe-O bond in addition to the direct Fe-C bond.

The binding of COOH to CoP yields most stably the Co(III) oxidation state on CoP, regardless of the initial starting charge of the molecule. That is, if excess charge is applied, it is observed to go to the porphine ring rather than the Co center. The singlet state of Co(III)P-COOH is well bound and stabilized relative to the triplet state in GGA by 1.5 eV but in GGA+U by only 0.5 eV (see Fig. 4-6). The singlet prefers double occupation of all π molecular states followed by δ_{xy} occupation. The triplet state on the other hand prefers single occupation of π_{xz} and $\delta_{x^2-y^2}$. The Co-C bond in singlet Co(III)P-COOH is 1.88 Å which is considerably shorter than the metastably bound Co(I)P-CO_2 complex. One way in which we may investigate the differences in CO_2H and CO_2 binding is by pulling the proton off of the carboxylic acid moiety. We find the CO_2^x structure remains bound, with the weakly hydrogen bonded proton to the CO_2^x moiety and a slightly elongated Co-C bond to 1.99 Å(see

Fig. 4-7). It also appears that the structure is partially stabilized by a secondary side-on bond between one of the oxygen atoms and cobalt. The binding of CO_2 requires further investigation but is highly suggestive that reduction of CO_2 to a carboxylic acid is key in the binding steps, at least for gas phase reactions.

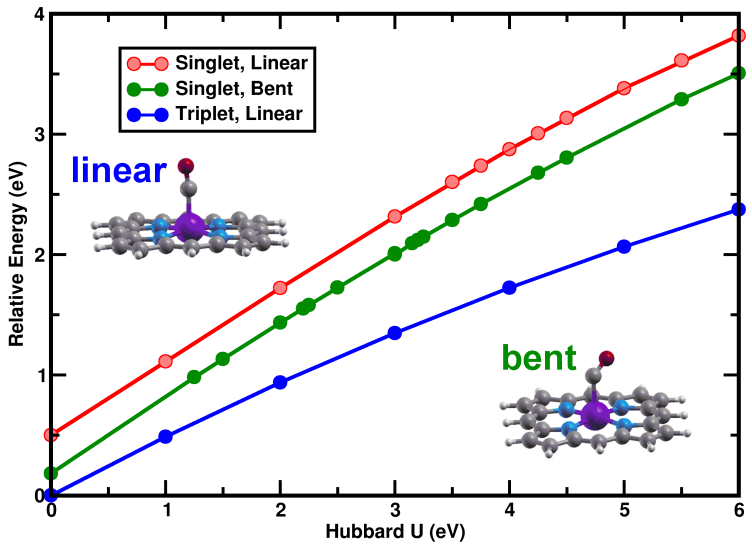


Figure 4-8: Relative energy of singlet and triplet Co(I)P-CO complexes versus value of Hubbard U. The triplet strongly prefers a linear Fe-CO geometry while the singlet bent Fe-CO structure is lower in energy than the linear one.

state	GGA D_e	GGA r_e	U=6 D_e	U=6 r_e
linear singlet	0.35 eV	1.92 Å	0.25 eV	1.93 Å
bent singlet	0.15 eV	2.16 Å	0.08 eV	2.15 Å
triplet	0.65 eV	1.81 Å	0.12 eV	1.88 Å

Table 4.2: Binding energies and structures of GGA and GGA+U Co(I)P-CO molecules.

In addition to CO_2 binding, the relative strength of CO binding is crucial in understanding whether or not CO poisoning is likely to occur on the CoP catalysts.

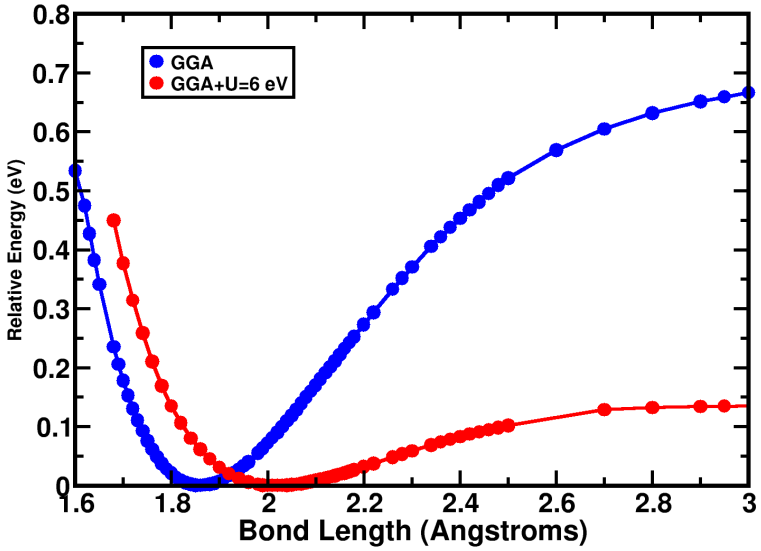


Figure 4-9: Binding energy curves of doublet Co(II)P-CO for GGA (blue) and GGA+U, U= 6 eV (red). The large GGA binding energy is reduced greatly with GGA+U.

The geometry and strength of binding of CO to the catalytic centers also provides a key test case for our theoretical approaches. We will investigate first the binding of CO to Co(I)P which is not believed to be highly relevant experimentally but is of interest to see whether or not the active reduced catalyst might be poisoned by the presence of CO. The triplet state is stabilized with respect to the singlet state by around 0.2 eV for GGA and to over 1 eV for GGA+U at a value of 6 eV (see Fig. 4-8). The GGA triplet is tightly bound with a D_e of 0.65 eV, but this is drastically reduced to a much more reasonable value of 0.12 eV with GGA+U (see Table 4.2). The singlet exhibits two configurations manifested structurally by a linear Fe-C-O bond or a bent bond. The triplet exhibits the shortest bond length of 1.81 Å in GGA and 1.88 Å in GGA+U, but the singlets exhibit higher dissociation energies. The bent singlet is stabilized in energy by about 0.25 eV, and its dissociation energies

in GGA and GGA+U, 0.15 eV and 0.08 eV, are both quite low. This structure also corresponds to a longer bond at 2.15 Å for both GGA and GGA+U. The linear singlet is relatively tightly bound, 0.35 eV and 0.25 eV, respectively, for GGA and GGA+U with a Co-C bond shortened to 1.93 Å. These results do initially suggest that CO could become irreversibly bound if the reaction mechanism led to population of the linear Co(I)P-CO singlet state.

Trends observed for Co(I)P-CO binding are also apparent for Co(II)P-CO and Co(III)P-CO. The doublet Co(II) porphine is experimentally observed not to bind CO, but the GGA structure has a short 1.8 Å bond length and a high dissociation energy of about 0.75 eV (see Fig. 4-9). GGA+U at $U_{scf,av}$=6 eV instead yields a much lower gas phase D_e of about 0.12 eV and a longer 2.05 Å bond length. This value could be even further modulated by the presence of solvent. Lastly, Co(III)P-

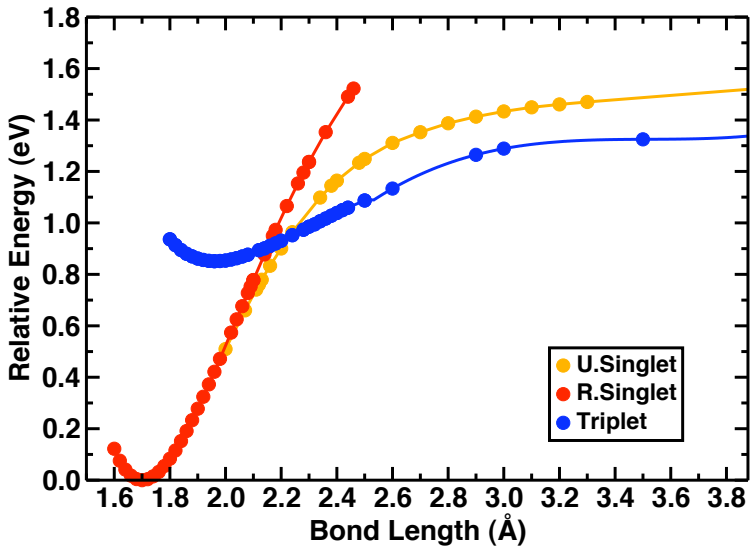


Figure 4-10: Binding curves and relative energies for GGA of Co(III)P-CO restricted and open shell singlets (red and orange) as well as triplets (blue).

CO structure singlets and triplets are initially tightly bound for GGA. The singlet

Co(III)P-CO has the largest dissociation energy of about 1.5 eV and the shortest equilibrium bond length of about 1.7 Å. The triplet structure is metastable by nearly 1 eV (see Fig. 4-10) but has a lower dissociation energy at about 0.5 eV and an equilibrium bond length of 1.95 Å. The Co(III)P-CO structures are unlikely

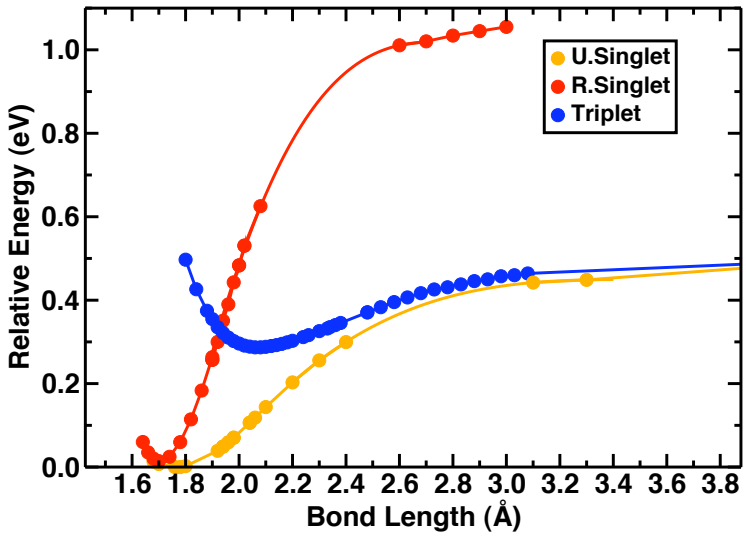


Figure 4-11: Binding curves and relative energies for GGA+U of Co(III)P-CO restricted and open shell singlets (red and orange) as well as triplets (blue).

to be long lived as reduction to Co(II)P by the electrode is followed by subsequent release of CO. However, GGA+U binding curves (see Fig. 4-11) show a markedly decreased binding energy for both singlets and triplets. The singlet is still stabilized with respect to the triplet but now only by 0.25 eV for $U = 5$ eV. At a U of 6 eV, the singlet dissociation energy is reduced to about 0.29 eV and the triplet state is very weakly bound at about 0.06 eV. The equilibrium bond lengths of the GGA+U singlets and triplets are also elongated at about 1.8 Å and 2.05 Å, respectively. Overall, GGA results suggest CO poisoning to be quite likely, while GGA+U yields binding energies much more consistent with known experimental parameters.

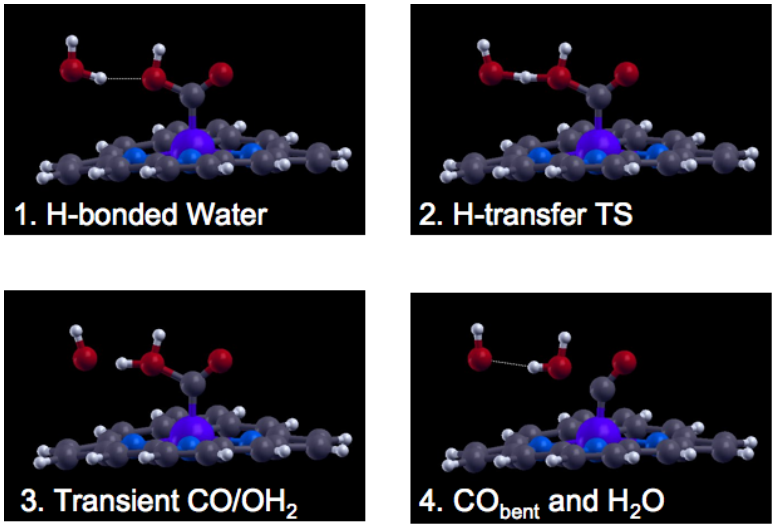


Figure 4-12: Structures relevant in the formation of CO, OH^-, and H_2O from attack of -COOH on H_2O in the following order: 1) hydrogen bonding of the water to COOH, 2) transfer of an H atom to the oxygen 3) short-lived bound water 4) spontaneous dissociation of water to leave behind CO bound to the Co(III)P.

Reaction mechanisms for the reduction of CO_2 to CO are the ultimate goal now that the relative binding energies and oxidation states are known. Experimentally, aprotic solvents are preferred, but in the gas phase, we exploit protonation in stabilizing CoP-CO_2^- for the ease of the calculation. In order to investigate the reduction of CO_2, we have considered the donation of a proton from a water molecule in order to facilitate reduction. The rate limiting step for this potential mechanism is the donation of a proton to Co(III)P-COOH. Upon protonation, we observe the dissociation of water from Co(III)P-CO(OH_2) to be barrierless and spontaneous with a hydrogen bond between water and OH^- being formed (see Fig. 4-12). The remaining Fe-CO structure is observed in a bent structure which was shown to have a low dissociation energy.

Investigation for alternative mechanisms which do not rely as heavily on the

presence of protonated Co(III)P-COOH and water is ongoing. Additionally, presence of solvent in the form of both implicit, model dielectrics as well as explicit molecules could significantly alter descriptions of binding and reaction mechanisms. However, a plausible mechanism has emerged whereupon charge transfer to CO_2 stabilizes binding to the initially reduced Co(I)P to form Co(III)P-COOH followed by addition-elimination attack on the carboxylic acid moiety to release water and leave behind a bent Co(III)P-CO structure which is released again upon further reduction. The role of GGA+U has been shown to be crucial in these systems as it universally brings the binding energy of various key moieties to the porphine into much greater agreement with experiments.

4.3 Cobalt tetrabromophenyl-porphyrins on copper surfaces

The study of molecules on metal surfaces has long been known to be a particular challenge for DFT and other theoretical methods[106]. For DFT, the nature, type and strength of adsorption for a molecule on a surface is very sensitive to exchange-correlation errors known as self-interaction. As the molecule contacts a surface, its density spreads out and hybridizes with the more delocalized states of the surface. In turn, self-interaction errors result in over-hybridization and improper bonding geometries. Therefore, Cobalt tetra-bromophenyl porphyrins (TBrPP-Co) on Copper surfaces (in this case, Cu(111)) are a particularly compelling application for GGA+U because the Hubbard U term must work to counteract not only self-interaction errors in the description of the isolated molecule but also the way in which the adsorption and charge transfer on Cu(111) is described.

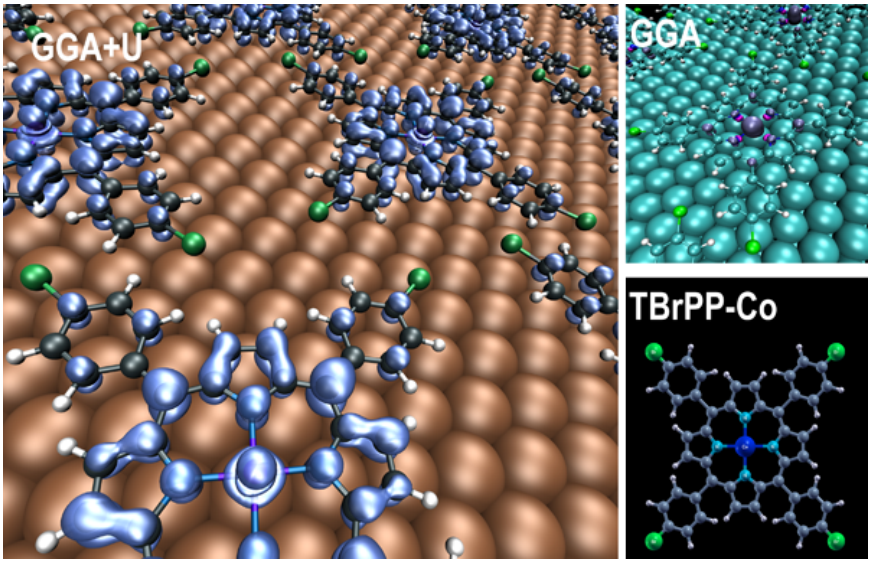


Figure 4-13: Total spin density of TBrPP-Co and Cu(111) complex for GGA+U (left) and GGA (top, right). A color-coded structure of the isolated molecule is also provided for comparison (bottom, right).

Spin-electron interactions involving molecules which are magnetic or have a net spin are of great interest for their potential in the field of spintronics[116, 117]. On metallic surfaces, many body interactions between a magnetic impurity and the free electron host can generate a Kondo resonance[118]. The stronger the spin-electron coupling, the higher the Kondo temperature, which has been particularly observed in molecules[118, 119, 120]. Experimentalists have previously observed unique Kondo temperature patterns for the self-assembled monolayers of these TBrPP-Co molecules[119, 121]. In order to closely follow the experimental setup, we employed a monoclinic unit cell, a=c=17.05 Å b=18.73 Å and α= 113.7°. This geometry was chosen to best mimic the experimentally observed alignment of the self-assembled monolayer of porphyrins at 7^o off the [110] direction on the Cu(111) surface, a roughly 16 Å separation between the centers of nearest neighbor molecules[119].

The initial geometry was generated using the constraint that the cobalt center is likely to bind directly to a Cu atom. During structural relaxations, however, the molecule was placed at a Cu site and allowed to relax both vertically and horizontally on the surface. The two layers of Cu(111) surface in the unit cell were not allowed to relax from the calculated bulk lattice parameter. The bulk fcc Cu unit cell lattice parameter was determined using a 8x8x8 k-point mesh to be 3.63 Å, which is within 0.5% of the experimental 3.61 Å value. The total unit cell contains 100 Cu atoms and 77 atoms from the TBrPP-Co molecule for a total of 1357 electrons (1100 derived from Cu and 257 from the molecule). However, several test calculations were also carried out using only one layer of Cu. Electronic structure calculations were carried out on the isolated molecule, and, in this case, the bromophenyl groups were not permitted to rotate out of plane from the rest of the molecule. In the gas phase, these functional groups will tend to rotate to reduce the total energy of the system, but for proper comparison to the complexed molecule on the surface, the lowest total energy flat structure is required. The cutoffs employed for these calculations were 25 Ry for the wavefunction and 250 Ry for the density. For cases where a Hubbard U was employed, we applied it to the cobalt $3d$ states only, and we did not apply it to the Cu surface. Preliminary results which did employ a U on the Cu surface suggested that this added step was unnecessary.

Isolated TBrPP-Co molecules exhibit a doublet ground state with the spin located exclusively at the cobalt center. The apparent oxidation state of the cobalt in the porphyrin molecule, an overall neutrally charged system, is Co(II) $3d^7$, which is consistent with similar porphyrin systems[122]. Two low-lying configurations exist for the isolated molecule, separated by about 0.25 eV for GGA and decreasing only slightly to about 0.21 eV for a GGA+U (U_{scf}= 5.9 eV). Both states have greatly reduced, minimal $3d_{x^2-y^2}$ density, and the lower energy state favors a single spin

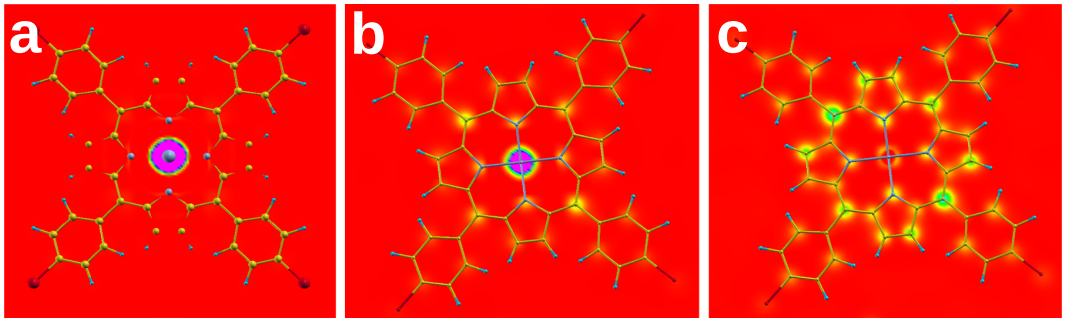


Figure 4-14: (Left,a) Spin density color plane of isolated TBrPP-Co molecule for both GGA and GGA+U as compared against (middle, b) GGA spin density for TBrPP-Co on Cu(111) and (right, c) GGA+U spin density for TBrPP-Co on Cu(111).

up occupation of the $3d_{z^2}$ derived orbital while the higher energy state has a singly occupied $3d_{xz}$ based orbital. The localization of the spin density to the cobalt center present in the molecule is apparent in both GGA and GGA+U (for all values of U) calculations of the two lowest electronic states of the TBrPP-Co molecule.

4.3.1 U dependence of TBrPP-Co spin density

Upon complexation of the molecule with Cu(111) surface, the role of the Hubbard U term becomes much more critical. The value of U_{scf} for the molecule-surface complex is 5.4 eV, relatively close to that of the isolated molecule. As may be seen in the color-plane cut of the density for the molecule-surface complex in Fig. 4-14, there is enhancement of the spin density in the ring of the molecule, particularly at the lobes connecting the main porphyrin to the bromophenyl functional groups. However, the majority of the spin in the GGA system resides still at the cobalt center. By applying U=5.4 eV, we are able to recover a delocalized electron density which, as we will later show, underlies the apparent unique Kondo temperature behavior observed experimentally in this system. The differences between GGA and GGA+U become

even more apparent by comparing the 3D spin-density for the two calculations at the same value for the isosurface (see Fig. 4-13).

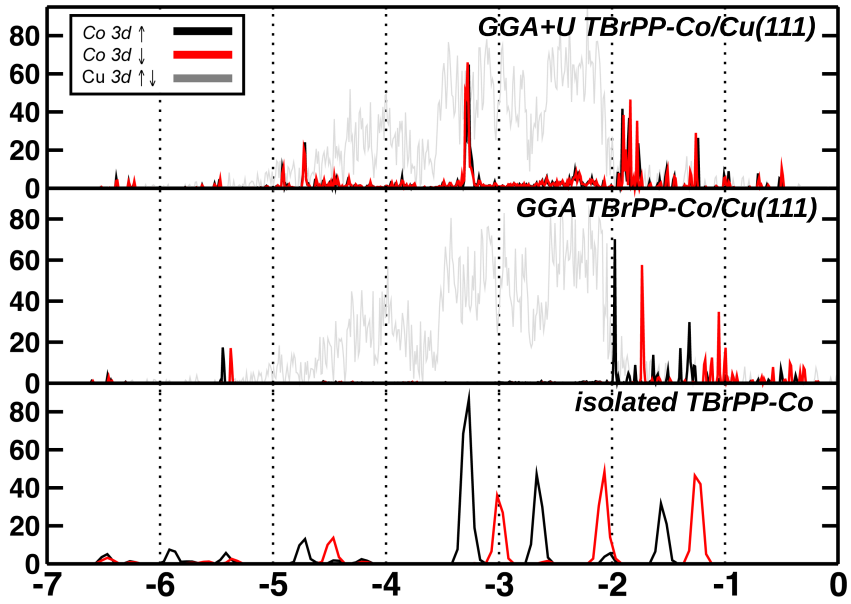


Figure 4-15: Comparison of three PDOS results for TBrPP-Co porphyrins. The three cases are isolated molecule (bottom), molecule on two layers of Cu(111) using GGA (middle) or GGA+U (top). All three PDOS have been aligned with respect to the $3d$ states of cobalt and a slightly larger Gaussian broadening has been used on the molecular states than in the case of the complex. Where appropriate, Cu $3d$ states from the slab are shown in gray, the spin up $3d$ states of Co are shown in black, and the spin down are shown in red.

Additionally, a projected density of states (PDOS) for all three systems, isolated and complexed (with and without a U term), helps provide additional clues into the changes in electronic structure when a U is applied (see Fig. 4-15). Since there is no qualitative difference in the electronic structure of the isolated molecule upon addition of a "+U" term, we consider only the PDOS of one isolated GGA molecule as a reference for the $3d$ states against the GGA and GGA+U complex results.

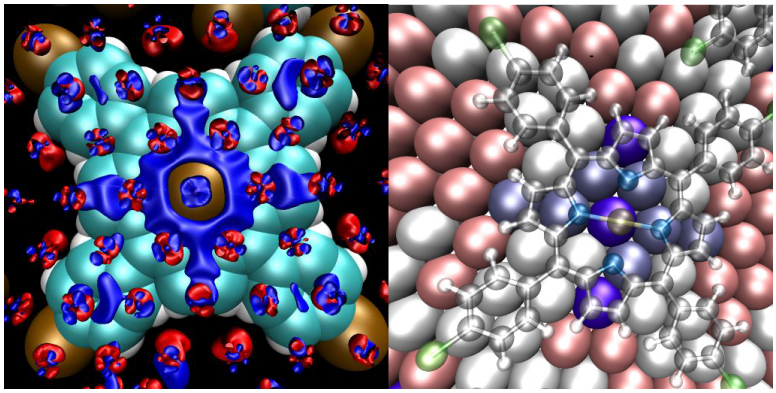


Figure 4-16: (Left) Density difference of Cu slab when molecule is added to surface. The greatest density difference is a depletion of density in a cross surrounding the cobalt center. (Right) Color coded atom-by-atom charge differences showing which atoms have donated charge (blue) or slightly enhanced charge (red) as aligned with the stick figure of the molecule on top.

For both GGA and GGA+U, the presence of the Cu 3*d* band greatly shifts the molecular Co 3*d* states. The majority of the 3*d* states in the isolated molecule span across 6 eV, with the majority fitting in a 2 eV window near the HOMO. For the GGA complexed system there is almost no Co 3*d* density within the 3 eV window of the Cu 3*d* band, rather Co 3*d* states are shifted to exclusively lie above the Cu 3*d* band. The GGA system also exhibits different eigenvalues for spin up and spin down state, which is common when the occupation of a single orbital of the same spatial symmetry is energetically favorable for one spin but double occupation is prohibitive (i.e. in localized 3*d* systems with a given spin). On the other hand, the GGA+U 3*d* eigenvalues overlap almost perfectly and yield an integrated spin of about four electrons in both spin channels (vs. 4 up and 3 down in the GGA system). Another interesting feature in the GGA+U PDOS is an apparent shift downward of select Co 3*d* states. In addition to the states which reside above the Cu 3*d* band

fitting in a smaller energy window of about 0.75 eV versus 1.1 eV, there are also several states well within the Cu 3d window both at the lower energy end and in the middle. This additional component of 3d-3d overlap, likely due to δ interactions, could be contributing to differences in charge transfer and spin density observed between the two cases. For both GGA and GGA+U, alignment with Cu 3d metal states greatly changes the behavior and occupation preference of the Co 3d orbitals but not necessarily in the same way.

Charge transfer occurs from the Cu(111) surface to the molecule and drives the reorganization of the states on the cobalt center and the porphyrin ring. The extent of charge transfer from the metal to TBrPP-Co was determined by calculating the difference in the Löwdin charges of atoms between the metal-porphyrin complex and the sum of the isolated molecule and slab values[123]. Of the approximately

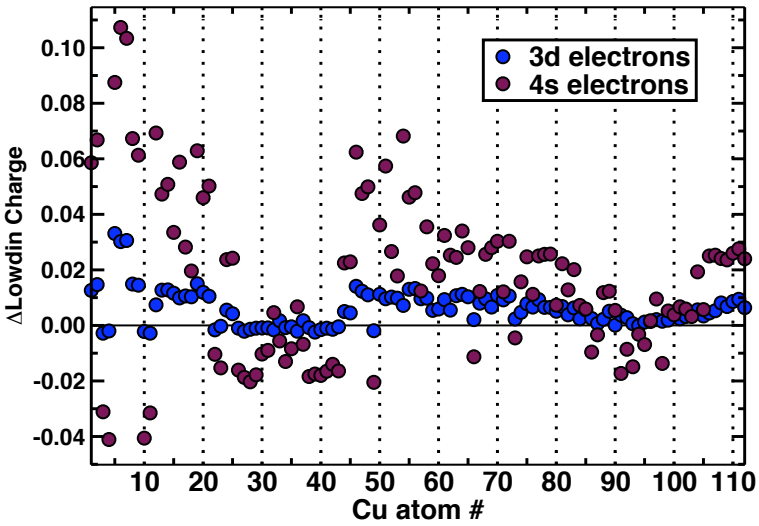


Figure 4-17: Löwdin charge differences for 4s and 3d components of individual Cu atoms in the Cu(111) slab ordered by atom number. The largest differences occur on the 4s electrons of select first layer Cu atoms, while the 3d charges are less significantly changed.

2.75 electrons transferred from the copper metal to the molecule in the GGA+U system, about 0.5 electrons were transferred to the cobalt center, and 2.25 electrons were transferred to the ring. The four bromides in the molecule additionally lost a small amount of charge, about 0.25 electrons total, which appear to have been transferred to the porphyrin ring as well. The largest regions of reduced charge lie in a cross corresponding to the cobalt and pyrrole groups of the porphyrin ring (see Fig.4-16). The charge density is depleted on the upper copper layer primarily from the $4s$ states (see Fig. 4-17). The density differences in the second layer are already within mathematical error of our approximations, and this justifies our use, by computational necessity, of only two layers of the metal (as shown in Fig. 4-17).

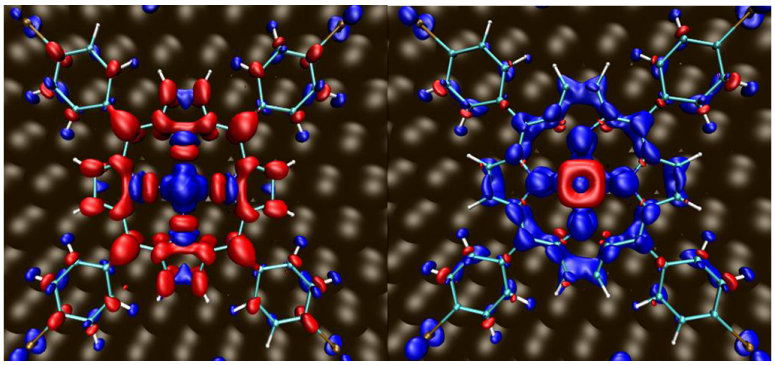


Figure 4-18: TBrPP-Co molecular spin density differences for spin up density (left) and spin down density (right). The greatest enhancements (red) are in the ring for the spin-up density and in the Co-center for the spin-down density, while there is an apparent depletion (blue) in some density close to the surface for the spin-down, reduction in spin up cobalt center density, as well as overall depletion of density at the four bromides.

In addition to identifying the sites of greatest depletion of charge on the Cu(111) surface, we may further examine density differences between the gas phase and complexed TBrPP-Co molecule (see Fig. 4-18). For the Co site, charge density is do-

nated from Cu to the minority spin $3d_{z^2}$ orbital, which is unoccupied in the gas phase molecule. For the majority spin Co $3d$, there is an apparent reduction in the density (see Fig. 4-19). Instead, the spin up density increases greatly on the outer ring, while it decreases significantly in the spin-down channel, particularly at the underside of the molecule closest to the Cu surface. There is also a net decrease in both spin up and down density for the four Br atoms. The precise reason for reorganization at Co is considered in more detail later, but we also would like to note that some charge transfer occurs to the GGA molecule when complexed with Cu(111). While the charge transfer occurs, it is reduced to about 0.3 electrons, down 40% from the value for the GGA+U system. The charge transfer to the porphyrin ring is relatively comparable for GGA and GGA+U systems as well, roughly 2 electrons, but as we previously observed, GGA charge transfer occurs to both spin up and spin down density without as much enhancement of the spin on the ring.

	Co $3d$ $n_\uparrow$					Co $3d$ $n_\downarrow$						
	σ	π	$\delta_{x^2-y^2}$	δ_{xy}	$n_\uparrow$	σ	π	$\delta_{x^2-y^2}$	δ_{xy}	$n_\downarrow$	$n_{\uparrow-\downarrow}$	Mag
Iso.												
GGA	0.99	0.97	0.60	0.99	4.52	0.09	0.93	0.53	0.99	3.48	1.04	1.0,1.4
$+U$=5.9	0.99	0.99	0.60	1.00	4.56	0.04	0.98	0.48	0.99	3.47	1.08	1.0,1.3
[Iso]$^{-2}$												
GGA	0.96	0.97	0.49	0.99	4.38	0.95	0.55	0.42	0.99	3.46	0.91	1.0,1.5
$+U$=5.9	0.99	0.99	0.64	1.00	4.60	0.98	0.53	0.50	0.99	3.52	1.08	1.0,1.5
[IsoCu]$^{-2}$												
GGA	0.81	0.98	0.55	0.99	4.31	0.73	0.90	0.53	0.99	4.06	0.26	2.0,2.4
$+U$=5.9	0.79	0.95	0.52	0.99	4.19	0.79	0.95	0.52	0.99	4.20	-0.01	2.0,2.0
On slab												
GGA	0.85	0.94	0.44	0.99	4.17	0.68	0.85	0.42	0.99	3.78	0.39	1.0,1.2
$+U$=5.4	0.77	0.97	0.33	0.99	4.02	0.79	0.96	0.33	0.99	4.04	-0.02	1.0,1.1

Table 4.3: Occupations of Co $3d$ projection matrix for several configurations of TBrPP-Co molecule including isolated (Iso.) molecule which is neutral, negatively charged, and negatively charged with an axial Cu atom ligand, as well as the molecule on Cu(111) slab. The difference between spin up and spin down occupations on Co $3d$ is also displayed as well as the total system magnetization followed by the absolute magnetization, where the latter is the absolute value of integrated spin density.

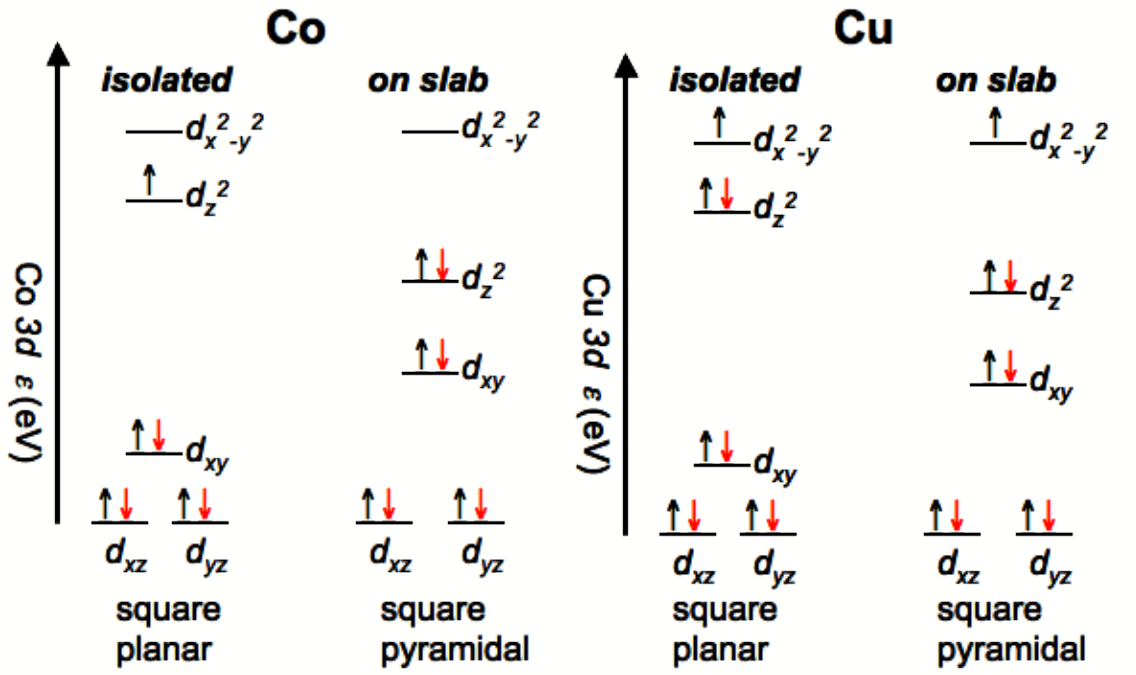


Figure 4-19: Qualitative molecular eigenvalue ordering for TBrPP-Co (left) and TBrPP-Cu (right) isolated compared to ordering on a slab of Cu(111). The addition of axial ligand from the slab to TBrPP-Co reflects the crystal field ordering shift from square planar to square pyramidal and from Co(II) to Co(I). For TBrPP-Cu, there is no change in oxidation state preference when the molecule is adsorbed on Cu(111).

We thus may revisit the isolated molecule as a simpler test system to try to understand under what conditions the spin density of the molecule may reorganize. As shown in Table 4.3, addition of charge to initially neutral isolated porphyrins does not result in significant increase of minority spin density at the cobalt center. Rather, the overall charge at the exterior of the ring increases, populating orbitals of both spins. Instead of enforcing charge transfer by increasing the total number of electrons in the system, we are able to mimic in the gas phase the signatures of the spin reorganization apparent upon complexation with the metal surface by adding a

single Cu atom as an axial ligand to our system. By adding only a single Cu atom to the system, we still see a shift in the oxidation state of the cobalt center's $3d$ electrons from a $3d^7$ Co(II) doublet to a $3d^8$ Co(I) singlet (see Fig. 4-19). This shift is associated with an increase to double occupation of a $3d_{z^2}$-like orbital, which was previously observed only in the GGA+U molecule-metal complex. The reduction in cost for doubly occupying the $3d_{z^2}$ orbital is most easily explained in the framework of a crystal field theory problem. The addition of the fifth ligand in the axial position causes the highest occupied level to come down in energy, which, accompanied by charge transfer from the Cu atom, facilitates the transition from Co(II) to Co(I). One would anticipate $3d_{z^2}$ population to be the most favorable unoccupied orbital to populate because it would have the greatest overlap with the density from Cu atoms, unlike the also unoccupied $3d_{x^2-y^2}$ orbital which is perpendicular to the Co-Cu bond.

4.3.2 Contrast of TBrPP-Cu on Cu(111) with TBrPP-Co

It is possible that this unique charge-transfer induced Kondo behavior is highly specific to the metal center used in the molecule. Indeed, one could also consider the nature of the metal surface, but based on the effects described, it is already known that the metal acts as a reservoir of electrons and that it is the local chemistry of the metal center which truly modulates the behavior of the system. The cobalt center of the TBrPP molecule was replaced with a copper atom and the structure of the molecule was relaxed again. The isolated molecule has a roughly $3d^9$ doublet character. Upon complexation to the Cu(111) surface, there is no reorganization of the spin density from the Cu center to the ring. This may be understood by much of the same crystal field theory argument as was previously described (see Fig. 4-19).

In addition to the reduced likelihood of charge transfer between Cu on the surface

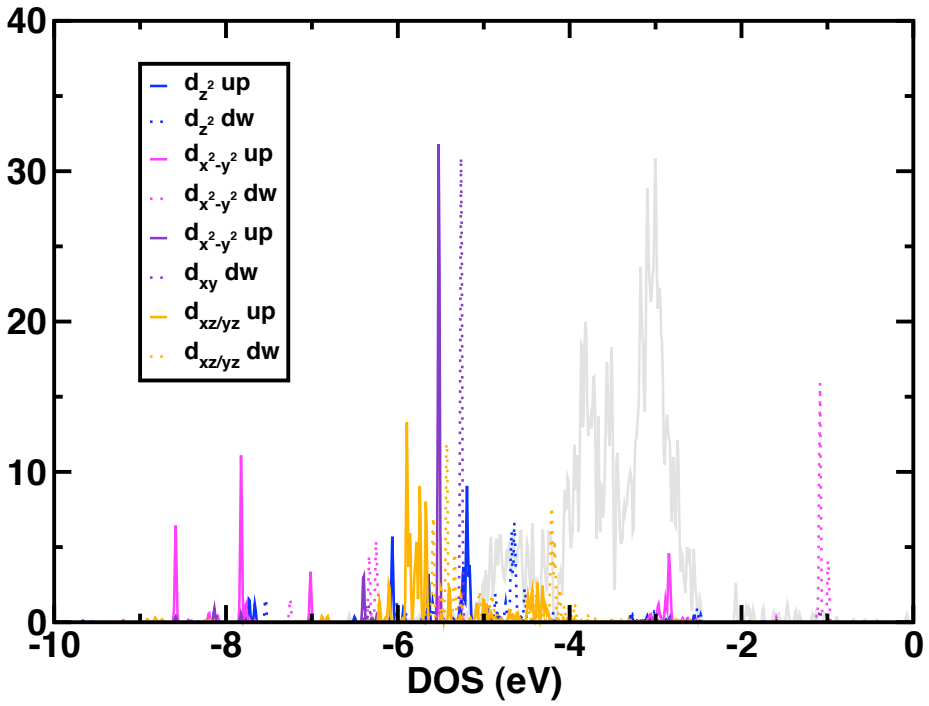


Figure 4-20: Projected density of states (PDOS) for TBrPP-Cu porphyrins on one layer of Cu(111) surface. States from the Cu(111) surface are shown in gray. The Cu center $3d$ states are each individually described with their spin up component (solid) and spin down component (dashed) shown separately. The $d_{x^2-y^2}$ spin-down orbital is high in energy and unoccupied while the remainder of $3d$ states are primarily at the low-energy end of the Cu (111) $3d$ band.

and Cu in the molecule with respect to Co, the Cu center in the molecule is already nominally a Cu(II) $3d^9$ doublet. Charge transfer into the highest occupied orbital, $3d_{x^2-y^2}$ would be highly prohibitive, and this orbital is orthogonal to the direction of the Cu_{metal}-$\text{Cu}_{molecule}$ bond (see Fig. 4-20). As a result, the Cu(II) character is preserved from the isolated molecule even upon complexation with Cu(111) surface. This confirms also experiments previously completed by our collaborators, Saw-Wai Hla and coworkers[119]. While Co is unlikely to be the only metal to exhibit unique

spin signatures in a porphyrin upon complexation with a metal surface, a recipe emerges based upon simple arguments for oxidation states and orbital occupations. By knowing the oxidation state and $3d$ electron count of the metal in the molecule to be complexed, it should be possible to roughly predict whether or not spin reorganization would occur upon complexation.

4.3.3 Kondo resonance in TBrPP-Co

Ultimately, the theoretical investigation of TBrPP-Co on Cu(111) was motivated by experimental observations of Kondo resonance from dI/dV tunneling spectroscopy measurements. The Kondo temperature is extracted from the dI/dV data via a prominent feature state near zero bias (see Fig. 4-21). At the center of a TBrPP-Co

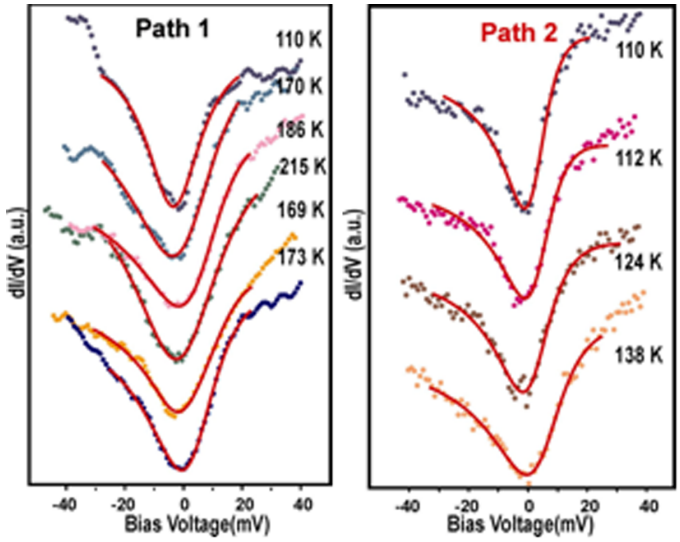


Figure 4-21: Sample sequences of Kondo resonances recorded along the path towards the bromophenyl group (path 1) and towards the pyrrole group (path 2). The curves fit to the data yield the listed Kondo temperatures at right.

molecule inside a SAM on Cu(111) the Kondo temperature is large, but experimen-

talists also observe Kondo resonance throughout the entire porphyrin unit of the molecule[119, 124]. The variation of the Kondo temperature across the molecule is manifested in the tunneling spectroscopy data measured at discrete points along two paths: one diagonal and one across (along a Co-N bond) the molecule. The dI/dV spectra measured at various locations over the porphyrin unit reveal higher Kondo temperatures moving away from the center of the molecule. The measured data are fitted using Ujsaghy's formula and the resulting Kondo temperatures are plotted as a function of the distance from the center of the molecule[125]. The Kondo temperature increases from 106 K to 212 K for path 1 as the tip is moved away from the molecule center and decreases to 170 K over the bromophenyl unit. For path 2, it increases from 111 K to 151 K (see Fig. 4-21). Remarkably, the highest Kondo temperature is found when the STM tip is located near a protruding lobe of the molecule at the linker position between the main porphyrin ring and the bromophenyl group. By using an expression for the strength of the electron interaction in terms of the

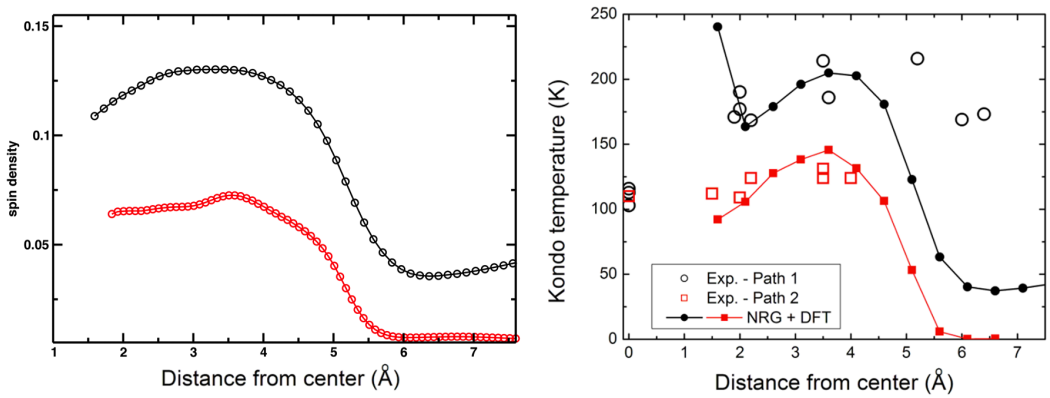


Figure 4-22: (Left) Spin density averaged in 2 Å window widthwise and 0.5 Å window smoothing lengthwise along direction towards bromophenyl (1, black) and pyrrole (2, red). (Right) Experimental Kondo temperature and calculated theoretical Kondo temperature based on average spin density for directions 1 and 2.

density, our own theoretical results can be used to predict Kondo temperature behavior along the molecule. For such calculations, we averaged the spin density along a direction into a 2 Å window widthwise. Additionally, we smoothed the density along the direction away from the molecule by windowing the data along this direction. The spatial variation of the spin densities along the paths 1 and 2 obtained from DFT may be incorporated into numerical renormalization group (NRG) calculations in order to extract Kondo behavior along the molecule. Since Kondo resonance is a many body effect not directly observable from DFT, the NRG extension is a necessary step. The connection between DFT and experiment is obtained as follows: the Kondo temperature T_K is related to $u(x) = 2\rho J(x)$, where J is a coupling term and ρ is a density. Assuming the spin density, n_{DFT}, is related to ρ, the distant dependent Kondo temperature reveals a good fit of the density once renormalized to a constant. Kondo temperature increase is therefore a direct result of the net spin density in the molecular orbital. In very good agreement with experiments, we see a greater density along the diagonal direction which corresponds to the direction toward the bromophenyl groups compared to the direction toward the pyrrole group of the ring (see Fig. 4-22). Additionally, from Löwdin charges, we are able to identify the atoms in the molecule with the greatest local spin polarization and these are identified to be the C linker atoms to the bromophenyl groups in the molecule.

While we can directly observe the zero temperature density and spin-density of our system, direct comparisons to experimental parameters should also be made to confirm that we are reasonably reproducing the experimental system. Approximate imaging of the highest occupied molecular orbital (HOMO) and lowest-occupied molecular orbital (LUMO) is obtained experimentally through dI/dV tunneling spectroscopy maps. Such spectroscopic images are more coarse-grained than what we are likely to see theoretically, but they suggest that the HOMO consists of density which

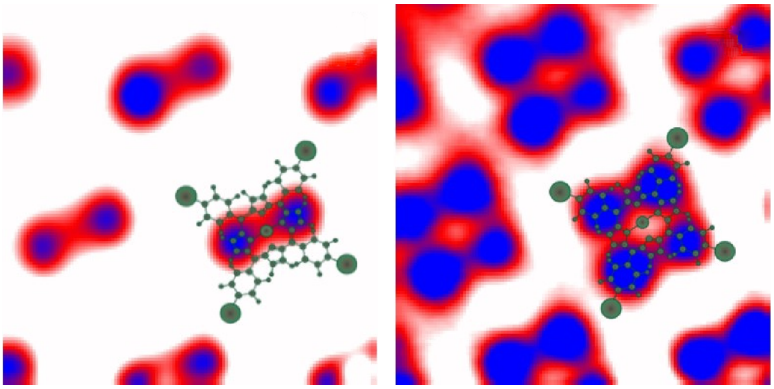


Figure 4-23: (Left) The highest occupied molecular orbital (HOMO) as imaged experimentally by dI/dV STM spectroscopy as compared to (right) the lowest-unoccupied molecular orbital (LUMO).

largely resides in the middle of the molecule across the two pyrrole groups and the Co center (see Fig. 4-23). The LUMO, on the other hand, exhibits no spin density at the Co center and the density is instead at four lobes associated with the bromophenyl groups of the molecule. Our own theoretical isosurfaces roughly agree with this experimental picture (see Fig. 4-24). The shape of the LUMO also helps to explain the strong spin density observed in the ground state calculations of the complex. There is majority spin occupation of the lobes and regions of the ring which is in turn unoccupied in the minority spin (i.e. it corresponds to the LUMO).

Overall, the study of TBrPP-Co on Cu(111) represents the leading edge of what even the best theory and computation can currently access. Self-interaction challenges as a result of both the isolated molecule and the surface have been successfully counteracted using the GGA+U approach. As a result, we have reproduced the proper spin density profile of a large 177 atom, 1357 electron system. Using an extension to our ground state density functional approach, we have also been able to explain experimentally observed novel spin density profiles and Kondo temperature

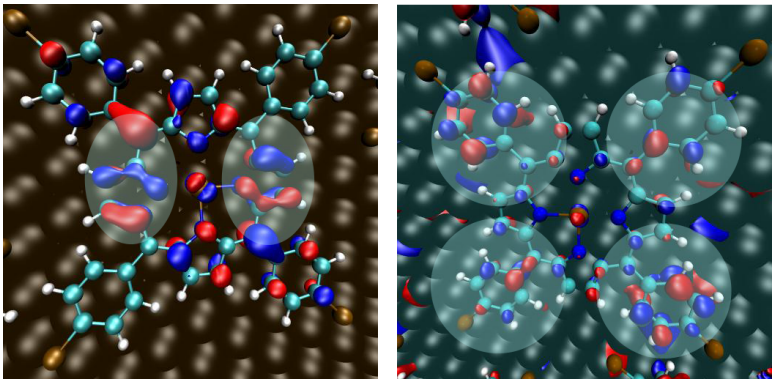


Figure 4-24: (Left) Highest occupied molecular orbital (HOMO) and (right) lowest unoccupied molecular orbital (LUMO) for the TBrPP-Co molecule on Cu(111). The regions experimentally identified as associated with the HOMO and LUMO in experimental dI/dV imaging have been highlighted - rough qualitative agreement of the locations of highest density are observed.

behavior. This new behavior observed theoretically and experimentally in molecules on surfaces both suggests that modulation of the oxidation states of these molecules will permit modulation of the spatial distribution of the spin and provides a recipe for such behavior based also upon the the transition metal's identity.

Chapter 5

SyrB2: A non-heme iron, α-ketoglutarate-dependent halogenase

Many man-made and over 4500 natural halogenated molecules have been characterized, with chlorine being the predominant modifier[126]. Natural organic halogenation compounds with known therapeutic capabilities include vancomycin and chlortetracycline[127]. Halogenation is also a common means of modulating a drug's biological activity - nearly 20% of pharmaceutical drugs are halogenated[128, 129]. It is observed that synthesized therapeutics exhibit 4-10 fold increase in efficacy upon chlorination of a single site, but chlorination of unactivated alkanes can be quite challenging[129, 130]. There are several types of chlorinating enzymes reliant upon metal centers: an iron Heme, a vanadium center, or a two-histidine (His) non-heme Fe^{2+} center are three common examples[131, 132]. The first two types require hydrogen peroxide as a co-substrate to the alkane, while the non-heme Fe^{2+} center

utilizes dioxygen instead of hydrogen peroxide. The hydrogen peroxide based enzymes, termed haloperoxidases, are believed to form a hypohalite (-OX, where X is the halogen) intermediate, while the oxygen-dependent halogenases are believed to oxidize ground state X^- ions to high energy radicals or cations which then easily form C-X bonds[133]. Of all types of chlorinating enzymes, the two-His dependent Fe^{2+} center enzymes are particularly attractive because they are known to form C-Cl bonds at typically unreactive, aliphatic carbons[134]. Mechanisms of this latter class of halogenases may be inferred from numerous hydroxylases[135, 136], but the differences which cause strong preference of halogenation to the point of exclusion of hydroxylation is not well understood. While structures of halogenating active sites have recently been determined[5], this field has relatively few first-principles studies, despite the fact that electronic structure is key to understanding the particular nuances of the halogenase activity[5, 137, 138].

The particular non-heme halogenase we will consider occurs naturally in a *Pseudomonas syringae* biosynthetic pathway and is part of a larger set of proteins which create an antifungal peptide called Syringomycin E[134]. The ninth residue of the Syringomycin E lipo-nonapeptidolactone is a chlorinated threonine (4-Cl-L-Thr). A flexible prosthetic arm on the protein SyrB1 is activated and loads threonine onto one end via thiolation. The 66 kDa SyrB1 protein has two domains - one domain is responsible for tethering the amino acid by thiolation (the T domain) to the flexible phosphopantetheine arm while the other is associated with ATP adenylation (the A domain)[134]. The L-Thr-S-SyrB1 complex then delivers the substrate into the center of the halogenase, which is named SyrB2. SyrB1 was studied for incorporation of various amino acids, and the formation of L-Thr-S-SyrB1 was observed to be at least 30 times faster and with 60 times higher selectivity than the incorporation of any other amino acid, with the next most likely being L-serine. It is known that

binding of the threonine and SyrB1 arm likely excludes the water from the active site and permits the incorporation of reagents into the SyrB2 active site[5]. Since the enzyme only chlorinates the substrate when loaded on the long SyrB1 arm, computational models should in the future accommodate the as yet not well understood changes the arm may induce in the surrounding protein environment. The SyrB2 protein reacts to produce 4-Cl-Thr-S-SyrB1, which is later used in the final steps by the SyrE protein to complete the formation of Syringomycin E. SyrB2 has been shown to be capable of bromination in the absence of chlorine, although preference for chlorine over bromine is observed by a factor of about 180; no halogenation with fluorine or iodine nor any hydroxylation has been observed[5, 139]. This detail is of particular interest because of the significant homology between this halogenase and oxygenases. Mutation of the alanine which permits chlorine binding in SyrB2 to an aspartic acid that is present in hydroxylases fails to induce hydroxylation, while binding of chlorine is not observed in hydroxylases in which aspartic acid has been replaced by alanine[5, 138]. It is therefore of particular interest to identify what aspect of the electronic structure of this enzyme precludes hydroxylation.

Non-heme halogenases are known to be homologous to a family of hydroxylases, which are comparatively well studied[136]. Several hydroxylases are heme-dependent monooxygenases, but a class of lesser studied non-heme iron hydroxylases, most frequently found in microbial systems, shares the most homology with the halogenases[140, 141]. The non-heme iron hydroxylases contain a motif of two histidines and a carboxylate, known as the "facial triad"[142], and their reactivity is dependent upon decarboxylation of a metal ligating co-substrate such as α-ketoglutarate (α-KG; see Figure 5-1). The non-heme hydroxylases and halogenases share the majority of the characteristics of the active site motif, but few, if any, enzymes are biologically capable of both hydroxylation and halogenation. Most hy-

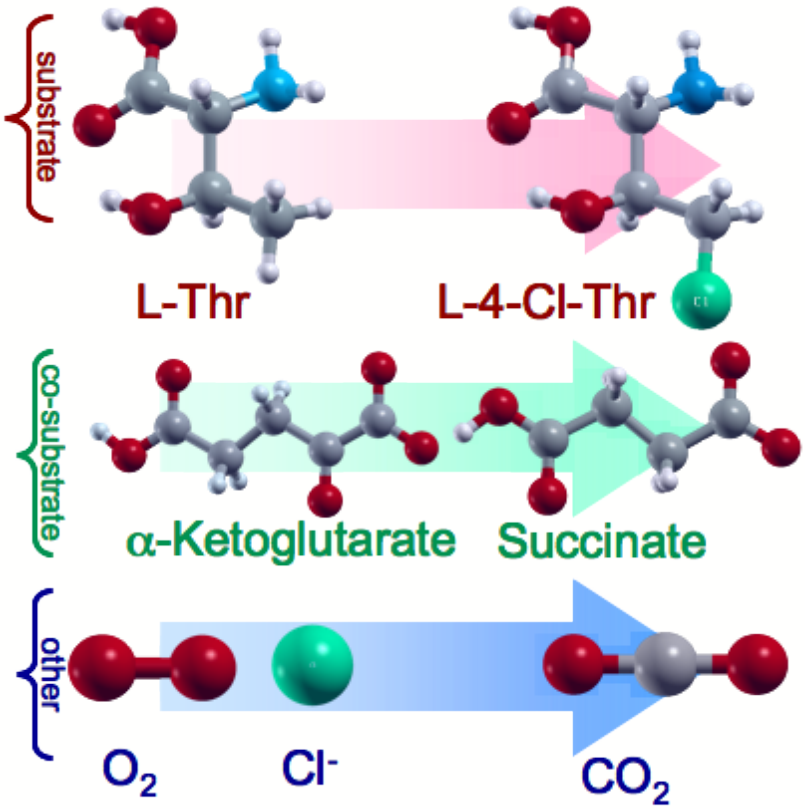


Figure 5-1: Conversion of substrate, co-substrate, and reagents (from top to bottom) in the SyrB2 halogenation pathway from reactant form (left) to products (right).

droxylases act on free amino acid or polypeptide chains, but very recent biochemical work on the as yet uncrystallized and previously overlooked hydroxylase, SyrP, which hydroxylates L-Asp to L-3-OH-Asp only when the substrate is loaded on a phosphopantetheine prosthetic group similar to the one requisite in SyrB2[143]. Since both SyrP and SyrB2 functionalize tethered amino acids and are from the same biosynthetic pathway, comparison of the crystal structures of the two enzymes and subsequent first principles studies could reveal what subtle differences dictate preferred functionalization mechanisms. One particularly well-characterized hydroxylase is the

taurine/α-KG dependent dixoygenase, known as TauD, which has been studied by site-directed mutagenesis techniques[144]. These experiments showed that the His99 that is coplanar with αKG and Fe(II) is a prerequisite for an active enzyme, while the distal His255 can be substituted with as much as 81% activity maintained. Substitution of Asp101 is only possible if it is to an alternative carboxylic acid such as a D101E variant. While mutation of TauD's aspartic acid to alanine did not reproduce halogenase activity, addition of formate was able to stimulate some enzyme activity by potentially forming an intermediate which mimics the role that Asp101 plays[144]. Other experiments on TauD have demonstrated a large kinetic isotope effect (KIE) in stopped-flow absorption and Mössbauer experiments which further suggests that high-spin Fe(IV)=O intermediate decay is strongly coupled to C-H bond cleavage[145]. The measured KIE of around 37 suggests that quantum-mechanical tunneling is likely to be key in the C-H bond breaking step[144].

Recent combined biophysical and biochemical study has led to the elucidation of X-ray crystal structure of SyrB2 without the substrate bound[5]. The published crystal structure of SyrB2 with bound Fe(II) and α-KG was obtained at a resolution of 1.60 Å. The protein is comprised of 2465 atoms including 310 protein residues in addition to co-substrates and ligands[5]. This catalytic cycle is known to be Fe(II), α-KG, Cl^- and O_2 dependent, and the reaction also produces succinate, CO_2, and OH^- ion for charge balance (see Fig. 5-1)s. While much of the active site motif, two histidines and bidentate α-KG, is conserved with respect to the hydroxylases, SyrB2 is the first structure of a mononuclear iron protein that does not coordinate the Fe(II) with a carboxylate but rather is instead coordinated by a Cl^- ion. Of the remaining two free sites of the hexacoordinated iron, a chlorine occupies the fifth site, opposite one of the coordinated oxygens from α-KG while the remaining axial site in the unactivated form is shown experimentally to be occupied by a water molecule.

This water molecule is believed to leave once the substrate binds, and at this time SyrB2 has not yet been co-crystallized with its SyrB1-S-Thr substrate. During the catalytic cycle, it is believed that a high-valent Fe(IV)=O center is generated via O_2 attack and decarboxylation of α-KG as follows:

$$Fe^{II}{-}O_2 + \alpha KG \rightarrow Fe^{IV}{=}O + \text{succinate} + CO_2. \tag{5.1}$$

The Fe(IV)=O intermediate can then abstract a hydrogen from the substrate, leaving behind a radical:

$$Fe^{IV}{=}O + R{-}CH_3 \rightarrow Fe^{III}{-}OH + R{-}CH_2^{\bullet}. \tag{5.2}$$

The methyl radical of the substrate then abstracts a neutral Cl, potentially in a radical rebound mechanism which mirrors hydroxylase activity:

$$[Fe^{III}{-}OH]{-}Cl + R{-}CH_2 \rightarrow Fe^{III}{-}OH + R{-}CH_2Cl. \tag{5.3}$$

In hydroxylases, the OH radical in the axial position would instead be delivered to the substrate rather than Cl, since an aspartic acid is located in the hydroxylases where Cl is bound in halogenases[128]. The absence of a CO_2^- containing residue in this part of the active site leaves a place for a halide to bind to the Fe^{2+}. The halide has also been shown to improve the binding of Fe^{2+} center to the enzyme by nearly three orders of magnitude[5]. One of the reasons that this enzyme does not hydroxylate the substrate is likely due to the lower reduction potential of the chloride radical compared to hydroxide. The chloride reduction potential is E_m = +1.36 V *vs.* standard hydrogen electrode (SHE) while OH reduction is E_m = +2.02 V *vs.* SHE. Therefore, chlorination of the alkyl radical is likely very thermodynam-

ically favorable[133]. Using a first principles approach, we hope to comprehensively study the steps of the chlorination reaction and to understand conclusively why no hydroxylation is observed.

Several key aspects of the protein environment help to stabilize the hexacoordinate iron center and its ligands in the halogenase. The α-KG, which helps to bind the iron, is stabilized by several hydrogen bonding interactions at either end of the molecule. Near the Fe-coordinated end, an arginine (Arg254) stabilizes both the carboxylate of α-KG as well as water molecules present in the active site. At the opposite end of α-KG, three hydrogen bonds stabilize the bare oxo end - two from another arginine (Arg248) as well as with a tryptophan (Trp145) which stacks atop α-KG. In addition to stabilization by hydrogen bonding, it is believed that the phenylalanine (Phe104) and tryptophan (Trp145) π-stacking also helps to position the α-KG molecule relative to the iron. At the other end of the coordination site, the chloride ion sits in a hydrophobic pocket composed primarily of alanine and phenylalanine residues. The chloride ion is coordinated to at least two water molecules which are stabilized by hydrogen bonding with other residues in the active site. This water coordination of the chloride may be critical to ensuring the elongated Fe-Cl bond of about 2.44 Å in comparison to an average hexacoordinate ferrous iron Fe-Cl bond length found in the literature of about 2.30 Å[146, 147]. Our first principles approach should ultimately take into account any effect of hydrogen bonding interactions of peripheral residues, and we would like to understand the role of Fe(II)-Cl bond elongation in chlorination of the methyl radical.

A recent kinetic study of a homologous protein, CytC3, a halogenase that is in this case derived from *Streptomyces*, works towards further elucidating what happens after a high-energy ferryl-oxo intermediate is formed[148]. The CytC3 enzyme behaves similarly to SyrB2 but instead chlorinates L-2-aminobutyric acid. This study

included Mössbauer experiments on active intermediates in the reaction. At least two Fe(IV) intermediates were observed, both with an apparent quintet spin. These two intermediates were observed in roughly a ratio of 4:5. Kinetic details also show that the ratio of these two intermediates is unchanged with time, thus indicating they are in a rapid equilibrium. Within 50s of this reaction, the Fe(IV) intermediates have decayed significantly, and instead the Mössbauer results reveal a signature typical of high-spin Fe(II) complexes[149]. This is likely a result of progression to a later part of the reaction cycle including possibly a complex with products or regeneration of the reactant complex. The same interconverting high spin Fe(IV) intermediates were observed in a brominating reaction by CytC3, albeit with a shift to 1:4 preference for one of the Fe(IV) intermediates[150]. Additionally, EXAFS results in the brominating reactants shows preference for a long Fe(II)-Br bond of 2.53 Å, while the reactive Fe(IV) intermediate appears to exhibit a Fe-Br bond of about 2.43 Å[150].

5.1 Selecting SyrB2 model size and calculation details

Studying SyrB2 from a first principles approach is critical for understanding the electronic structure of transition states and intermediates which help to energetically facilitate unactivated hydrocarbon halogenation under mild, biological conditions. Unfortunately, construction of a realistic active site for first principles calculations is not trivial. The full length of α-KG, as well as the residues which help to stabilize it, extends several angstroms away from the Fe center and the believed location of the substrate upon binding. This adds considerable computational cost to the problem in itself. In addition, we must consider how to construct the hydrophobic pocket in

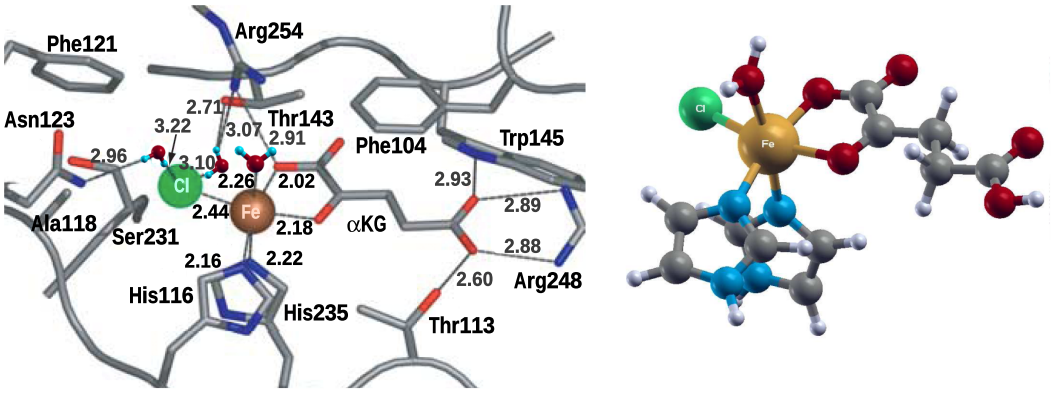


Figure 5-2: a) Cartoon of the SyrB2 active site with experimental distances modified from [5] and b) minimal 38 atom model of SyrB2 with only direct Fe(II) ligands as used in calculations.

which Cl is bound. One possibility is to treat this just with vacuum in our supercell. While vacuum is not necessarily the most accurate environment for Cl, it may be useful for preliminary calculations. An improved scenario would involve simulation and model of the protein environment with a modified dielectric constant. Even the most minimal structure of the active site, which includes 38 atoms, could elucidate the electronic structure of intermediates in the catalytic cycle.

This starting complex would include the five membered rings of the His residues, with the carbon backbone bond terminated and replaced by a hydrogen, at two of the coordination sites. The α-KG would be included in its entirety to accurately model the third and fourth coordination sites with the bare carboxylate present in the protein environment being hydrogenated here. Lastly, we would include a Cl^- ion and a water molecule at the fifth and sixth coordination sites, respectively. Sufficient vacuum would be included to eliminate long range image effects. While most of the calculations presented in this thesis utilize this minimal model, some caution should be taken in attempting to draw conclusions about the full enzyme or likely

candidates for active model complexes in solution. Recent work on model inorganic complexes has shown that dioxygen binding and activation at a metal site can be extremely sensitive to residues available for second shell hydrogen bonding[151]. Model inorganic compounds which share structural similarities to SyrB2 have been shown to be structurally sound, but the same steric bulk that is required to stabilize the Fe(II) center also prevents the key oxidative decarboxylation step[152].

Preliminary study of this small model has shown that the water molecule in the sixth site is very weakly bound, and the neighboring protein environment is likely to be critical to ensuring its placement in the solved structure. However, the low-binding energy of the H_2O is not terribly relevant considering that it leaves once the substrate binds. We know that hydrogen bonding in the form of free water molecules (as in the protein structure) or in neighboring residues is likely to play a role in the elongated Fe(II)-Cl bond. However, a small model complex may still prove sufficient to elucidate catalytic steps. Early calculations on this system showed that the Fe-O distances of both α-KG coordination sites are shortened by nearly 0.10 Å, due to the lack of the stabilizing protein environment, especially at the uncoordinated end of the molecule. The histidine site distances are largely unchanged, although the fully relaxed angle the His residues form with each other and the iron site does not always reflect that observed in the crystal structure. The Fe-Cl bond is shortened to around 2.30 Å in the calculations, though it is speculated that the reactivity of this enzyme is significantly enhanced by Fe-Cl bond elongation of 2.44 Å observed in the crystal structure (see Fig. 5-2). We do find that using our GGA+U methodology we achieve an improved estimate of 2.41 Å for the Fe-Cl bond and the Fe-O distances are also in agreement with the experiment. We have not included in the simulation any free water molecules to provide hydrogen bonding interactions with the Cl^-, however. As a result of using the minimal model, we also observe displacement of

the histidine rings with respect to the experimental structure to a varying degree for different intermediates.

Density functional calculations on larger, more realistic models of the active site are also possible. Inclusion of the residues which stabilize the placement of α-KG with their carbon backbones hydrogen terminated as well as the residues which stabilize the placement of waters which stabilize the Cl ion yields an active site with roughly 162 atoms. This model still requires fixing some of the residue locations since they lack the carbon backbone of the protein to enforce their location. Larger spheres may be cut out around the active site. Generally, the limit of accurate density functional calculations would be at around 300-450 atoms. A larger active site model (see Figure 5-3) of about 450 atoms permits inclusion of some secondary residues and carbon backbone and corresponds to a sphere of 12 Å radius centered around Fe. Larger models are useful for providing accurate estimates of the role of peripheral residues. Where a full carbon backbone and residue may not be included, we always cap the relevant atoms with hydrogens to ensure that the electronic properties and charge of the system exhibit a similarity to the full protein environment. While such larger model calculations are ongoing, they have not been included in this thesis. We can expect based upon previous results that residues in the active site are likely to help stabilize metal and co-substrate binding and weakly modulate metal-ligand interactions, particularly in the decarboxylation step[153].

Calculations were carried out on the minimal model system which had as few as 34 atoms and as many as 51 atoms when including the threonine substrate. Plane-wave calculations were carried out using the PBE-GGA exchange-correlation functional and ultrasoft pseudopotentials with a wavefunction and charge density cutoff of 30 and 300 Ry, respectively. The functional was augmented with a Hubbard U term calculated from linear-response and averaged over several intermediates to be 4.5

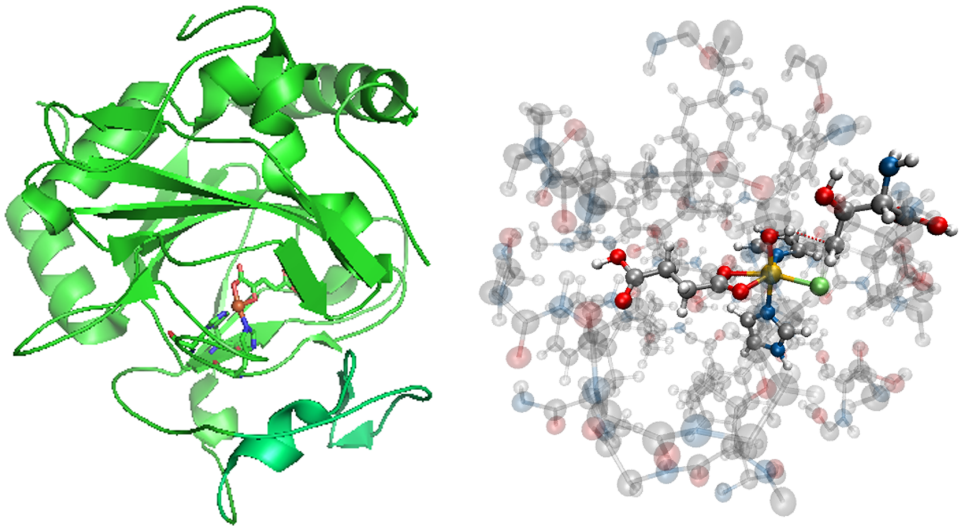


Figure 5-3: Complete SyrB2 protein shown as a ribbon structure with primary ligands to iron shown as ball and stick model (left) and smaller 450 atom model centered around α-KG and iron (right). The 450 atom model (semi-transparent) represents the upper end of calculations which are feasible with DFT, and within it is shown the minimal 51 atom model, which includes only primary ligands and the substrate (opaque).

eV. The intermediates were placed in a large cubic cell roughly 18 Å on each side to ensure there were no periodic image effects, and neutrality of the complete model system was always maintained. In most cases, energetic barriers were determined with the nudged-elastic band (NEB) approach in the formulation which includes a climbing image and variable springs for better resolution around the transition-state. Forces on the images on the path for the NEB calculations were minimized to less than 0.05 eV/Å before convergence was ascertained. Additionally, in some cases with no barrier, binding curves or dissociations were determined via a constrained relaxation in which the interatomic distance between two species was fixed and all other degrees of freedom were permitted to relax. Analysis of the oxidation states

of relevant species was obtained from the occupation matrix of the molecular states projected onto the localized $3d$ atomic states of iron.

5.2 Fe(IV)=O formation and hydrogen abstraction

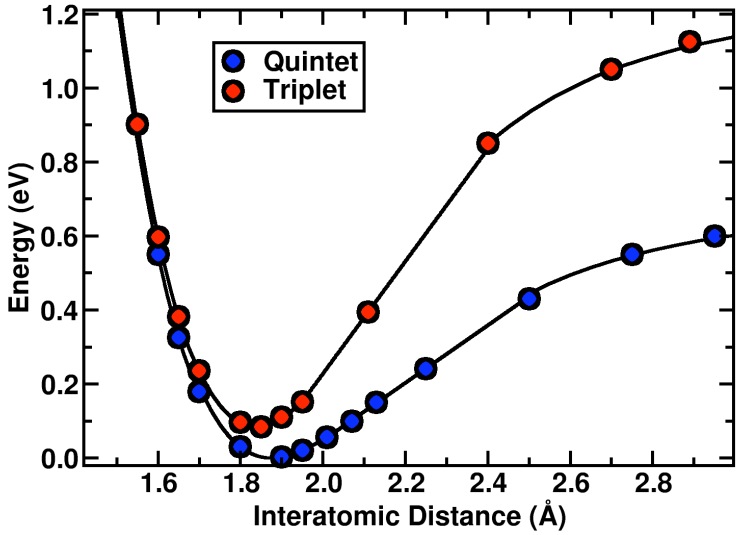


Figure 5-4: Binding curves of dioxygen in the axial site of the SyrB2 model compound for the lowest state (quintet, blue) and second lowest state (triplet, red). Dissociation energies of the complexes are moderate (around 1.3 eV) and the iron-O_2 bond length is relatively short (1.9 Å).

The first step for halogenation is the activation of the enzyme by formation of a high energy iron-oxo intermediate. The iron-oxo intermediate is formed through the attack of an oxygen from O_2 (bound to Fe in the axial position) on α-ketoglutarate (α-KG). Binding of O_2 is favorable, with the lowest electronic state, a quintet, corresponding to a relatively short, 1.85 Å equilibrium bond length between the iron

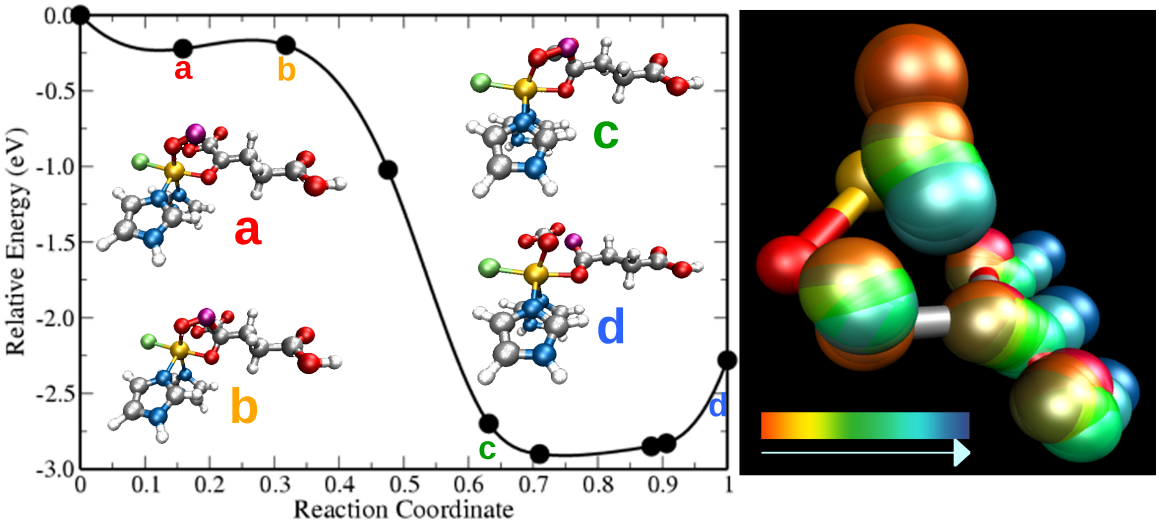


Figure 5-5: Decarboxylation of α-ketoglutarate to form succinate occurs via binding of one oxygen atom from O_2 to a carbon to induce sp^3 hybridization (left, see structures along the reaction pathway labeled a-d). The reaction is nearly barrierless and exothermic with a slight energetic barrier associated with complete dissociation of CO_2. The total reorganization due to additive attack of the carbon by oxygen and subsequent decarboxylation is shown (right) from red to blue with initial movement of the carbon out of plane and subsequent return to the plane upon decarboxylation being the most notable feature.

center and O_2. The quintet exhibits significantly bigger bond length and lower dissociation energy around 0.6 eV, as compared to the shorter bond of 1.8 Å and higher dissociation energy for the triplet, which is around 1.3 eV. This difference could be a source of autoxidation in the enzymatic system; that is, population of the triplet manifold could inactivate catalyst. The oxygen molecule bond length is elongated by over 0.1 Å with respect to the value of isolated gas phase oxygen, indicating a decrease in bond order, further facilitating for bond cleavage in the decarboxylation step. Upon addition of the oxygen to α-KG, decarboxylation occurs to leave behind the Fe(IV)=O intermediate. Calculations of this decarboxylation step show that once the additive oxygen from O_2 is in the vicinity of the α-KG carbon backbone,

the electronic structure at the carbon begins to evolve. The sp^2 hybridization of the conjugated carbon transforms to sp^3 hybridization, as evidenced by the planar structure around this carbon doming to yield a tetrahedral structure. This step is essentially barrierless and is followed by the exothermic release of CO_2 from the sp^3 activated carbon. Concomitantly, the O-O bond of the axial dioxygen breaks, and the newly formed succinate molecule returns to the plane of the other two iron ligands. The newly added oxygen now in succinate may bond with Fe and produce a bidentate succinate. Our electronic structure calculations weakly prefer a monodentate structure, as evidenced by one short Fe-O bond and one long one. Overall, the calculations predict the decarboxylation step to be barrierless and exothermic by nearly 3.0 eV(see Fig. 5-5). This estimate is reduced slightly to about 2.5 eV when including the endothermic release of CO_2 molecule to the gas phase. It was considered that the CO_2 might bind to or interact with either the iron directly or other active site ligands. However, it was found that the CO_2 would not preferentially bind at an empty iron ligand binding site. We now believe this endothermicity is an artifact of self-interaction from release of CO_2, and it is not likely that the CO_2 would remain bound to the succinate or the metal center.

The newly formed iron-oxo intermediate upon decarboxylation has been identified spectroscopically [148]. Experimentally, the electronic structure of this intermediate has been identified as a high spin, highly oxidized Fe(IV) which forms a double bond with the oxygen atom. Calculations confirm the experimentally observed electronic structure; that is, the oxidation state is nominally Fe(IV). Four spin up electrons doubly occupy the δ and π states, and there is significant charge transfer from the iron atom into the oxygen atom in the spin up σ orbital. The Fe-O bond length is extremely short at about 1.65 Å (see Fig. 5-6), which is comparable to the very short bond length of the isolated gas-phase molecule, $^6\Sigma^+$ FeO^+. Additionally, the

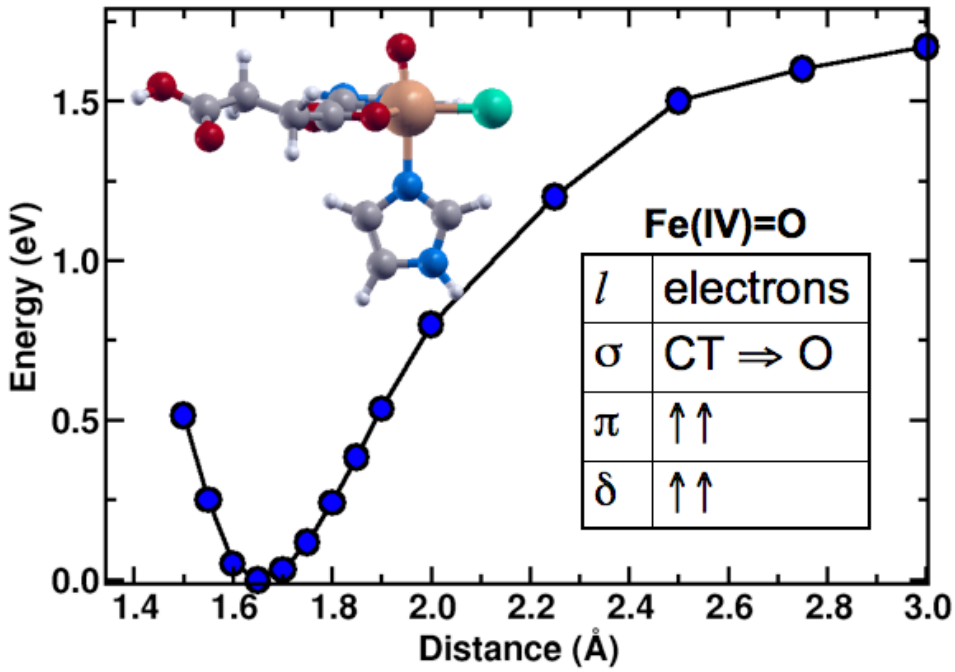


Figure 5-6: Structural and electronic properties of the axial Fe(IV)=O intermediate with binding energies of the oxygen atom exceeding 1.6 eV and a short 1.65 Å bond in the quintet ground state. Occupations from projections onto the localized 3*d* manifold strongly suggest an Fe(IV) oxidation state.

bond length is reduced by nearly 0.8 Å in the oxo-intermediate with respect to a weakly bound water molecule in the same site. The significant charge transfer from iron into the oxygen atom, roughly 2 electrons, makes this particular species ideal for abstraction from relatively unreactive substrates.

The abstraction of hydrogen from the substrate is a key step in the overall halogenation reaction. However, this hydrogen abstraction step is actually quite common in a large number of α-KG dependent enzymes, several of which are involved in hydroxylation or other reactions[128]. The first step for our calculations is to ensure that theoretical treatments of the hydrogen abstraction step are accurate. While rad-

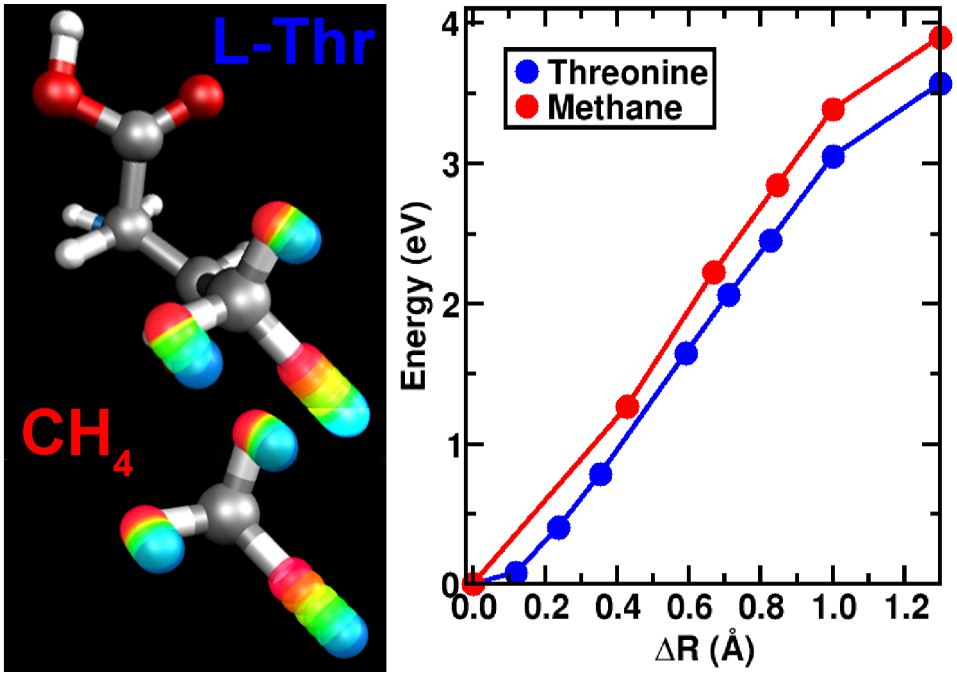


Figure 5-7: Comparison of hydrogen abstractions from the isolated gas phase methane and L-threonine molecules. The constrained hydrogen abstraction yields the same relaxation of the remaining hydrogens on the radical formed for both species (left). The energetic cost of the bond dissociation is estimated to be greater than 4.0 eV for methane and greater than 3.6 eV for L-threonine (right).

ical formation is sometimes perceived to be a unique challenge for density functional approaches, our calculations show that the energetics of these reactions may be calculated accurately. Gas phase calculations on isolated methane and L-threonine were carried out through constrained relaxations in which a hydrogen was gradually abstracted from the relevant methyl group while the rest of the molecule was permitted to relax freely. Both methane and L-threonine exhibit a flattening of the remaining atoms bound to the methyl carbon as the proton is abstracted. For methane, a CH_3 radical remains which is described by the three HCH angles each being 120^o (see Fig. 5-7). In L-threonine, the relaxation of the remaining bound hydrogens mirrors

that of the methane with the exception that a hydrogen in methane is replaced by the carbon backbone of the molecule. The energetics of this bond stretching and dissociation were also calculated (see Fig. 5-7). The bond dissociation energy of a C-H bond in methane is known to be experimentally about 4.45 eV. Our calculations based upon a 1.25 Å stretching from the equilibrium C-H distance give an energetic cost of about 4.0 eV. Corrections to consider the fully isolated fragments revises this value up further to about 4.2 eV. Overall, the bond dissociation energy calculated theoretically for methane is in very good agreement with the experimentally known value. It is also known experimentally that small decreases on the order of 0.25-0.5 eV are observed when comparing hydrogen abstraction from a methyl group as compared to larger, bulkier hydrocarbons which form more stable radicals. Our L-threonine calculations qualitatively follow this trend. At the same 1.25 Å stretching of the C-H bond in L-threonine, the energetic cost for bond dissociation is about 0.4 eV lower. This decreases slightly in the limit of completely isolated fragments to about 0.3 eV. Overall, these results show that we may be confident in determining barrier heights of hydrogen abstraction from methane and L-threonine using the SyrB2 model complex.

The formation of methyl radical via hydrogen abstraction was studied prior to studying hydrogen abstraction from L-threonine because of the greatly reduced cost in terms of fewer atoms and degrees of freedom for transition-state path-finding approaches. The abstraction of a single hydrogen from methane using the Fe(IV)=O intermediate of the SyrB2 model complex was studied with the nudged elastic band (NEB) method. The energetic cost of hydrogen abstraction by the model complex from methane is greatly reduced from the isolated gas phase value. While methane hydrogen abstraction was theoretically found to require about 4.0 eV in the gas phase, the NEB transition-state barrier height of abstraction with SyrB2 is less than 0.4

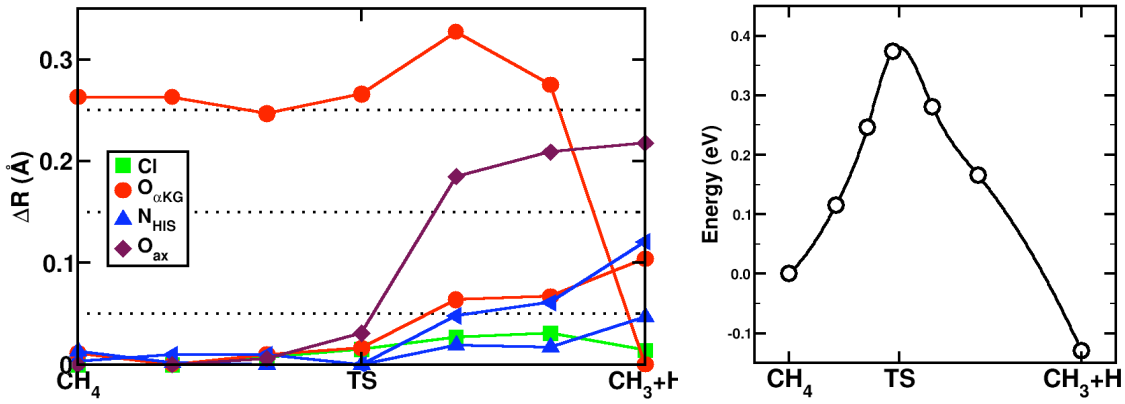


Figure 5-8: Evolution of key structural measurements as a function of hydrogen abstraction progression from methane (left) shows that after abstraction a shorter, bidentate succinate geometry is preferred while the other bond lengths become longer. The energetic barrier of hydrogen abstraction from methane is about 0.37 eV (right) and is exothermic.

eV (see Fig. 5-8). While this barrier is still high for ambient temperature enzymatic systems, considerations including barrier height modulation by second shell ligands in the enzyme as well as proton tunneling could make this reaction quite feasible. Additionally, this reaction step is actually weakly exothermic by just slightly less than 0.15 eV.

The use of the NEB approach also permits us to track structural properties along the minimum-energy path. The critical bond lengths for the hydrogen abstraction are plotted in Fig. 5-8 in terms of their displacement from the shortest values, which in most cases, corresponds to the value in the relaxed Fe(IV)=O intermediate. The majority of the bonds lengthen significantly upon the beginning of the formation of an O-H bond between the model complex and the abstracted hydrogen (see Fig. 5-9). This feature is likely due to the increasing electron density at the Fe center concomitant with a transition from an Fe(IV)-like oxidation state to Fe(III). As a result, less electron density is participating in net bonding interactions, leading to

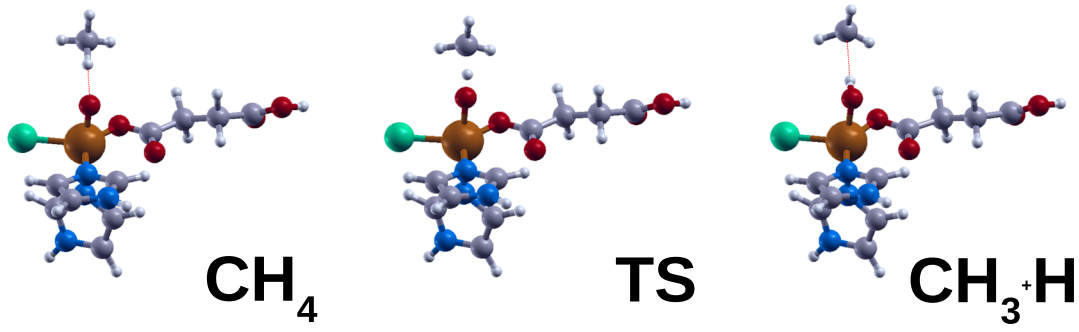


Figure 5-9: Structures in the reaction pathway of hydrogen abstraction from methane to form methyl radical. The first step is formation of a hydrogen bond to Fe(IV)=O_{ax} (left), the transition state shares the hydrogen equally between the methyl group and oxygen (center), and lastly the abstracted hydrogen in the form of hydroxyl is weakly hydrogen bonded to the methyl radical (right).

longer bond lengths between iron and its ligands. The elongation of the Fe-N(His) and Fe-Cl bonds are relatively subtle at around 0.1 Å on average. A more dramatic switching occurs at the transition state for the axial oxygen - from a bond length of about 1.65 Å for the isolated model complex to about 1.85 Å. Along the path, the extension of the Fe-O bond is nearly an instantaneous switch with none of the path images living at an intermediate distance between the short and long Fe-O bond regimes. In the reactant complex, Fe(IV)=O hydrogen bonds to a single hydrogen from methane. The transition state shares equally the abstracted hydrogen between model complex and methyl radical and the products. An elongated Fe-OH_{abs} bond is a signature of the product complex, which weakly hydrogen bonds to the methyl radical (see Fig. 5-9). Another interesting structural change observed was a decrease of the bond length of iron with with one of the equatorial oxygens derived from the succinate molecule. In the model complex oxo-intermediate, one short Fe-O bond exists from succinate, while the other is only weakly interacting. Upon binding of

the proton to the axial oxygen, the initially shorter succinate Fe-O bond lengthens slightly but the other formerly long Fe-O bond drastically shortens. This structural change corresponds to preferential bidentate binding of the succinate once the O-H_{abs} interaction is fully formed.

5.3 Chlorination and catalyst regeneration

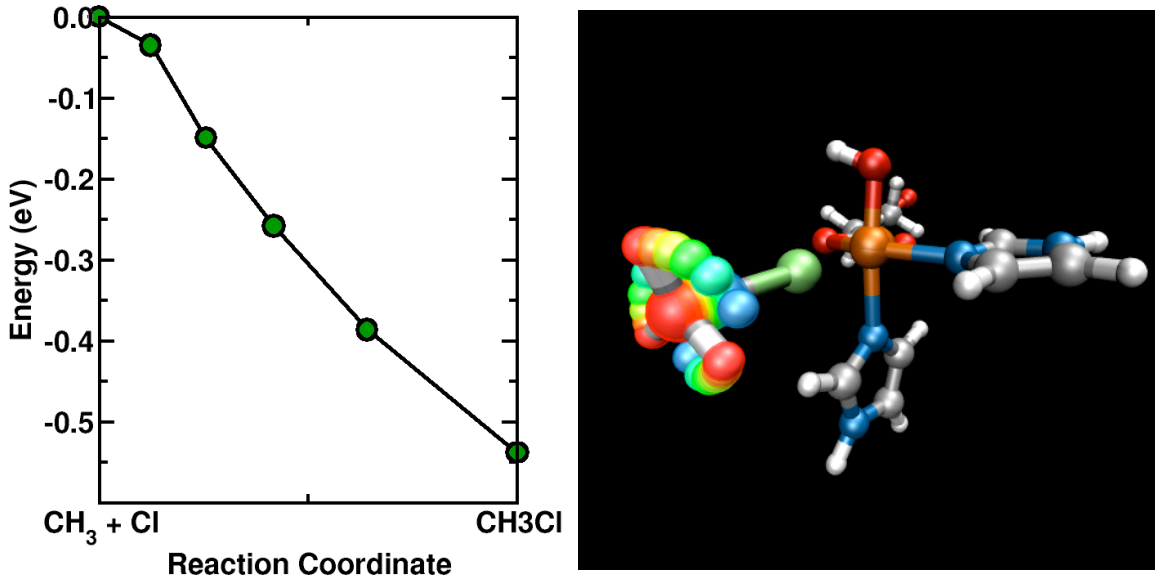


Figure 5-10: Chlorination of methyl radical by SyrB2 model compound is barrierless and exothermic by 0.6 eV (left) and the structure corresponds to doming of the remaining three hydrogens upon binding of chlorine (right).

While hydrogen abstraction is a common step for enzymes in several biosynthetic pathways, the step which follows in SyrB2, chlorination, is unique particularly to this and only a handful of other enzymes. Following the formation of methyl radical, we investigated the chlorination of the radical to determine the energetic cost of this step. These calculations were carried out with a constrained relaxation in which the

methyl-Cl distance was reduced but the other degrees of freedom were relaxed freely. We observed the chlorination of the free methyl radical to be completely barrierless (see Fig. 5-10). The methyl radical chlorination is also exothermic by nearly 0.6 eV. Formation of the C-Cl bond is associated with a doming of the C-H bonds to return to a tetrahedral environment around the methyl carbon. Previously, it was noted that mutated hydroxylation enzymes which replace the aspartic acid with an alanine do not halogenate at any experimentally observable level. These results on the model complex suggest that the lack of halogenation is not due to an energetic barrier but rather a function of how well the enzyme is able to bind the chlorine. Calculations on methane have demonstrated that hydrogen abstraction barriers are greatly reduced using the Fe(IV)=O intermediate present in the SyrB2 catalyst and that the chlorination should occur spontaneously in the presence of a methyl radical.

The binding characteristics of iron with the hydroxyl axial ligand is highly relevant in the mechanism of the SyrB2 enzyme. Firstly, the catalyst is known to halogenate and not hydroxylate even after significant mutations which should cause SyrB2 to mimic hydroxylation enzymes. An understanding of the relative binding energies of the hydroxyl ligand axially compared to the equatorial iron-chlorine interaction should give some clues to why this enzyme prefers halogenation. Hydroxyl binding strength is also relevant because the regeneration of the SyrB2 catalyst to its resting state requires replacement of the hydroxyl with an easily displaced water molecule. As shown in Fig. 5-11, the binding energy is quite large for the hydroxyl radical in both the lowest sextet and second-lowest, quartet state. The bond length of the quartet is around 1.8 Å and slightly shorter than the already mentioned sextet ground state. Both states exhibit binding energies over 2.0 eV and this result suggests that the bare hydroxyl cannot be easily removed. Any potential removal of the hydroxyl ligand may occur via either charge transfer from the iron to produce

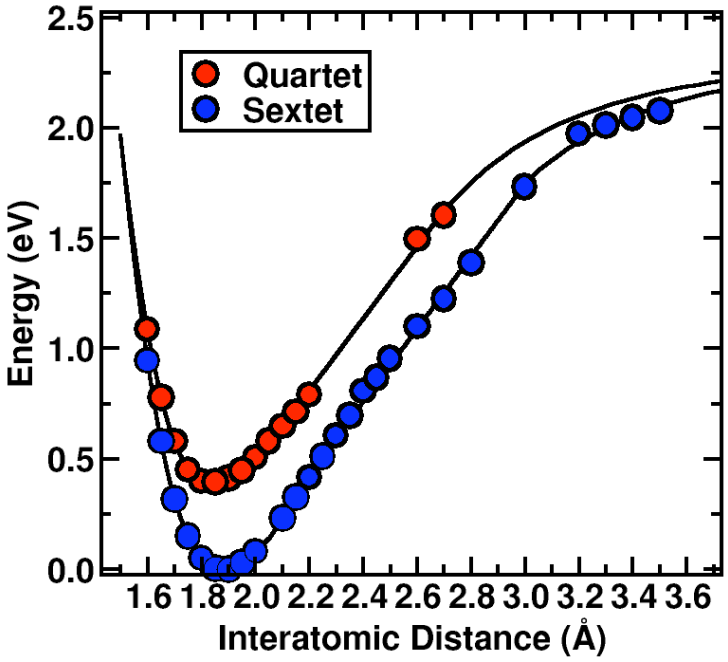


Figure 5-11: Binding energy curves for the hydroxyl axial ligand on SyrB2 model complex for both quartet (red) and sextet (blue) states. The binding energy for both states is nearly 2 eV and both exhibit relatively short, 1.8 Å bond lengths. The difference between the two states is the spin coupling of the iron and the OH radical.

a stable, dissociated OH^- or, more likely, a protonation event occurs which leaves behind a more loosely bound water molecule.

The relative binding energy of hydroxyl and chlorine ligands can help give initial clues to why the SyrB2 enzyme and model complex might prefer chlorination over hydroxylation. While not nearly as complete as a transition state path search for the two mechanisms, the relative binding energy implies the comparative cost of bond breaking in the radical attack reaction. Preliminary results indicate that the energetic cost of breaking the Fe-Cl bond is around 0.6 eV, which is much less than the nearly 2.0 eV binding energy for the axial Fe-OH bond (see Fig. 5-12). The bond length theoretically observed in most calculations is around 2.30 Å, but, as we

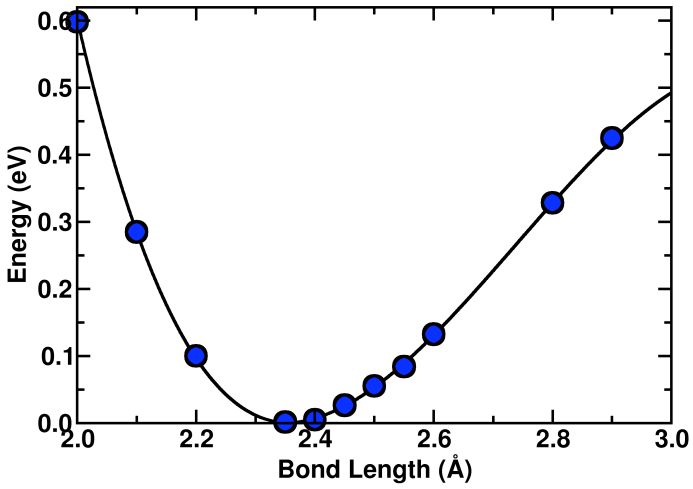


Figure 5-12: Binding curves of chlorine in an equatorial site of the SyrB2 model compound for the lowest state (sextet, blue) only. Dissociation energies of the complex is relatively low at about 0.5 eV and the iron-chlorine bond length is 2.30 Å. At the experimentally observed Fe-Cl bond length in the crystal structure of 2.44 Å, the binding energy is reduced further by about 0.1 eV.

previously noted, the Fe-Cl bond is elongated in the crystal structure of SyrB2 to be about 2.44 Å. This elongation could be due to the presence of interactions with waters or hydrophobic residues in the active site. At the experimental bond length, the binding energy is decreased by roughly 0.1 eV for the Fe-Cl bond with respect to the theoretical minimum. However, we note, the relative shallowness of the Fe-Cl binding curve suggests that Fe-Cl bond elongation is not critical for halogenation to occur. Overall these results suggest strongly that after the radical abstraction step, chlorination of any substrate should be barrierless because the energetic gain of forming a C-Cl bond outweighs the relatively small cost for breaking the Fe-Cl bond. Further, the relative binding energies of -OH and -Cl to the iron center should

modulate the relative rates of hydroxylation and chlorination[154]. In general, an enzyme of this family which is able to abstract a hydrogen to form a radical should also be able to release a chlorine atom to the radical if the binding energy to the metal center is comparable to that in SyrB2. Therefore, the most probable reason why more enzymes don't halogenate would be an inability to bind chlorine to the metal center in a location proximal to the substrate.

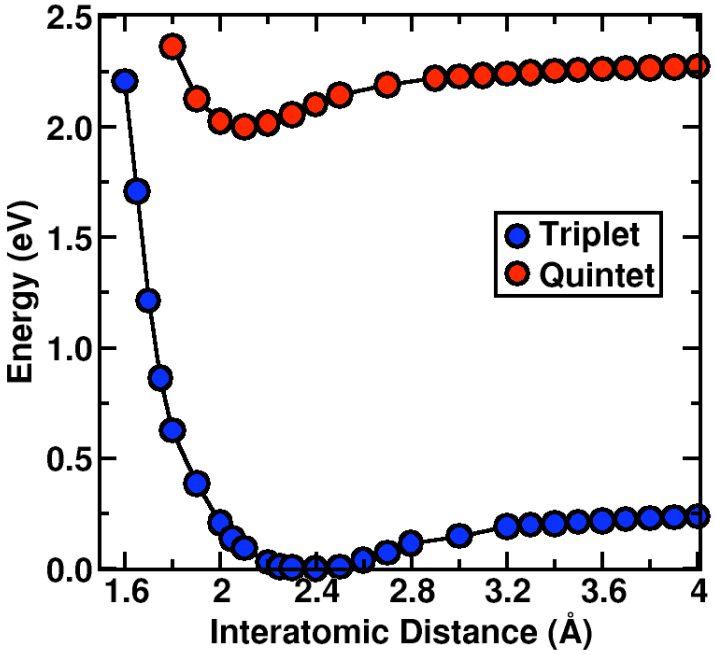


Figure 5-13: Binding curves of water in the axial site of the SyrB2 model compound for the lowest state (quintet, blue) and second lowest state (triplet, red). Dissociation energies of the complexes are low (under 0.25 eV) and the iron-water bond length is long (2.4 Å).

Of all axial ligands considered, the water binding strength is the lowest. Of course, this is unsurprising as we previously observed long 2.0-2.2 Å Fe-OH_2^+ bonds in the gas phase addition-elimination reaction of hydrogen on FeO^+. In fact, the Fe-OH_2

bond length of about 2.4 Å for the quintet and triplet SyrB2 model complexes is nearly 0.6 Å longer than that for the hydroxyl bond and over 0.85 Å longer than the strong Fe(IV)=O bond (see Fig. 5-13). Similarly, dissociation energies are reduced from as much as 2.5 eV for the oxo and hydroxyl ligands to about 0.25 eV, a ten-fold reduction in the energetic cost of dissociation. These results are suggestive that a free source of protons would protonate the hydroxyl ligand and permit release or weak binding of the water molecule. Unlike charge transfer for release of the hydroxyl ligand, the protonation step would not increase the oxidation state of the Fe center, and it is therefore a more plausible mechanism. The most likely sources of protons are the free water or hydronium ions present as a result of the enzyme being partially open to the environment. A survey of the main residues in the active site shows that there are no acidic amino acids within close proximity to the hydroxyl ligand. In fact, the only alternative to proton donation from a water would be from the carboxylic end of the succinate ligand which is about 8.0 Å away.

Calculations were carried out on the protonation step of the hydroxyl ligand to permit regeneration of water in the axial binding site. Protonation via a hydronium ion was observed to be barrierless and exothermic. Once the proton is donated from hydronium to form water, a hydrogen bond persists between the two water molecules and breaking this costs a small amount of energy, about 0.06 eV. Overall, the donation of the proton to the hydroxyl is exothermic by nearly 0.6 eV. Structural intermediates show a decreasing of the O-H bond lengths in the hydronium concomitant with the donation of the third proton. The expected elongation of the Fe-OH_2 bond is also observed (see Fig. 5-14). Protonation from a neutral water molecule is also likely to be favorable but less exothermic than for the case of the hydronium ion.

While several calculations have been carried out on methane model systems, the

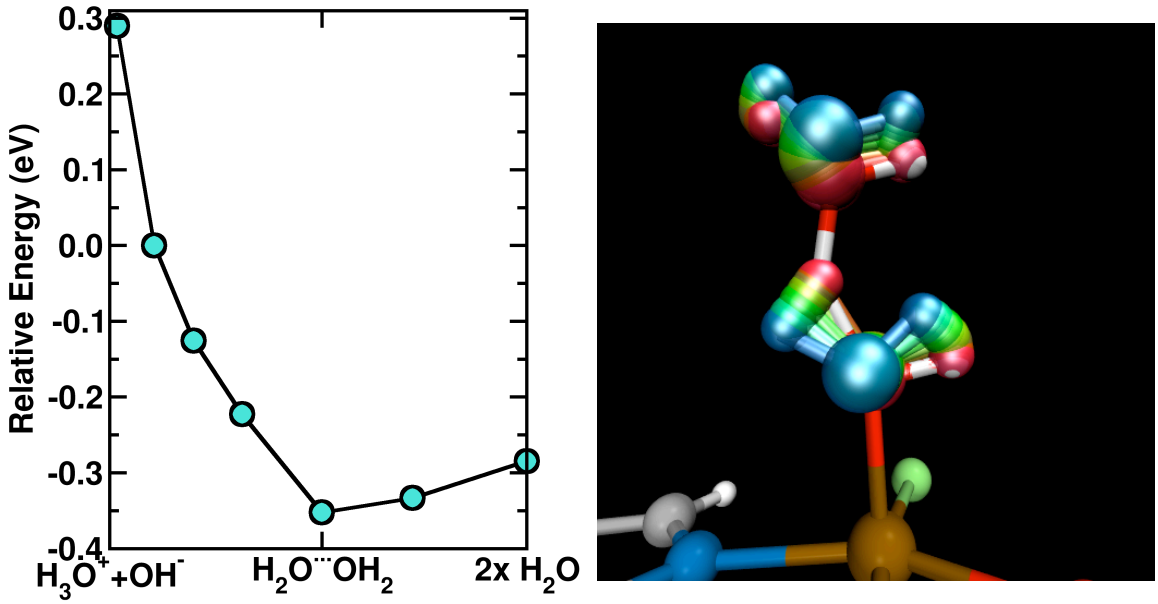


Figure 5-14: Energetics (left) and trajectory (right) of hydroxyl protonation step via proton transfer from a hydronium ion. The proton transfer is exothermic and barrierless, and the two waters preferentially hydrogen bond to each other.

ultimate goal is to understand how the SyrB2 model complex can abstract a hydrogen and subsequently chlorinate the natural L-threonine substrate. In the enzymatic system, the L-threonine is loaded on a protein arm called SyrB1 which forms a thiol-linkage with the L-threonine in place of part of the carboxylic acid group. SyrB2 does not catalyze the halogenation of unloaded substrate, and the arm most likely opens up the enzyme and deliver the substrate to the center of the active site. The substrate is most likely delivered directly over the Fe-Cl bond because the hydrophobic Cl pocket represents the area of greatest disorder in the protein, while strong stacking interactions exist along the other side between α-ketoglutarate and the nearby amino acids. In the model complex, the L-threonine was therefore positioned over the Fe-Cl bond in preparation for hydrogen abstraction. The abstraction of hydrogen from L-threonine was calculated to have a greatly reduced barrier of slightly under 0.3 eV

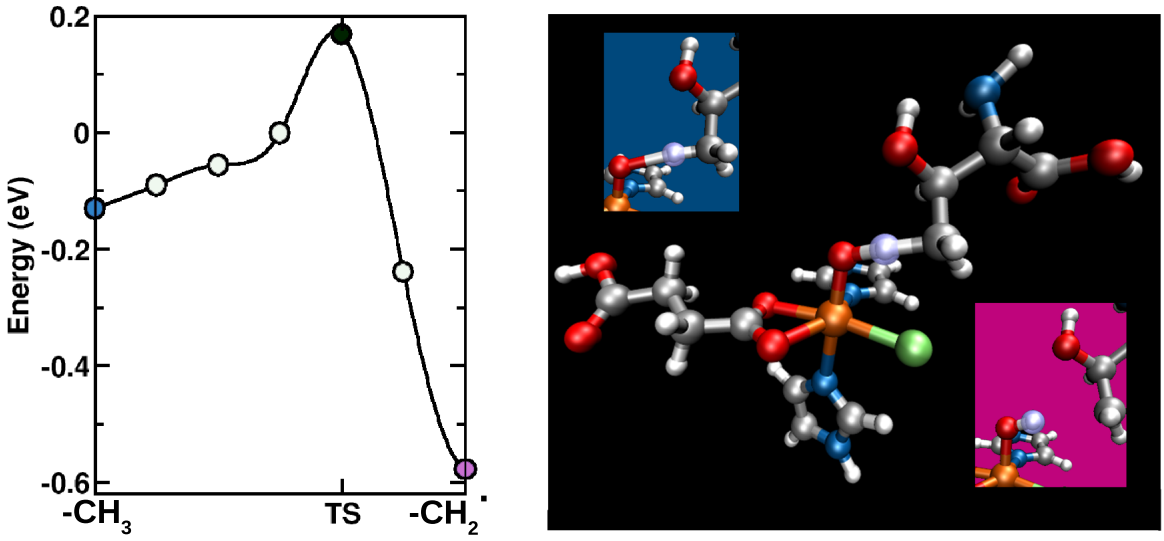


Figure 5-15: The reaction pathway of hydrogen abstraction from the methyl group of L-threonine has an activation barrier of around 0.3 eV (left) and is exothermic by 0.5 eV. The reactants, transition state, and products are shown (right) in the small inset in blue, the larger black figure, and the small pink inset, respectively.

with respect to the gas phase estimate of 3.8 eV (see Fig. 5-15). The transition-state of the hydrogen transfer is the point at which the proton is shared equally between the L-threonine and the iron-oxo species (see Fig. 5-15). Upon abstraction of the hydrogen from L-threonine, the newly formed axial OH ligand appears to stabilize slightly the L-threonine radical via hydrogen bonding to the alcohol group. The other structural trends in the SyrB2 model complex reflect those observed previously for the hydrogen abstraction on methane including of course the flattening of the ligands around the relevant carbon from tetrahedral to a planar geometry.

Structural arguments based upon the most likely location of substrate loading being directly over the Fe-Cl bond as well as the hydrogen bonding interactions between the axial hydroxyl and the -OH group on the L-threonine radical leads to the suggestion of a potentially concerted or, more likely, strongly coupled mechanism

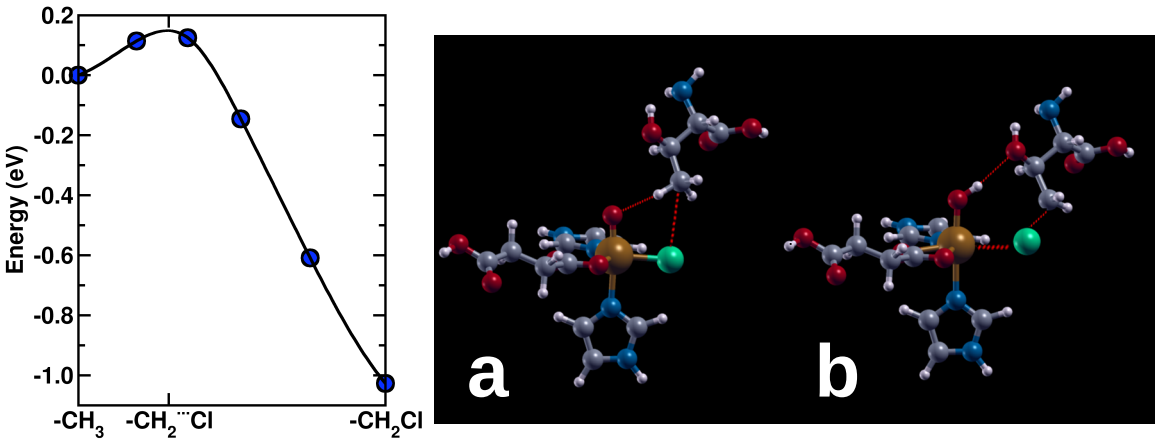


Figure 5-16: Energetics (left) and structure (right) of coupled hydrogen abstraction and halogenation mechanism. The barrier of the coupled reaction is reduced to 0.14 eV or half of what it is for the single abstraction event. At right, structure a represents the bonding interactions prior to atom abstractions, and structure b shows the chlorine atom being shared equally between the methyl carbon and iron while the hydroxyl hydrogen bonds to L-threonine.

for the hydrogen abstraction and chlorination. Since a small barrier exists on the hydrogen abstraction reaction pathway and the chlorination step was previously observed to be exothermic and barrierless, it follows that some barrier height reduction is likely when the two steps are coupled. Therefore, we carried out nudged elastic band calculations in which the chlorination and hydrogen abstraction occurred at the same time. These results show (see Fig. 5-16) that the partial stabilization by interactions between the methyl carbon and the chlorine appear to slightly reduce the barrier height estimates from the path determined only for hydrogen abstraction. The coupled reaction exothermicity is consequently quite large at about 1.2 eV. A key factor which promotes this coupled reaction is that the abstracted hydrogen forms a hydrogen bond with the L-Thr radical, further stabilizing the molecule to accept the chlorine atom. We do not find a four-centered transition state in which both

the hydrogen abstraction and chlorination steps are both occurring, and therefore it cannot be said that these steps are strictly concerted. Overall, the chlorination from this structural arrangement is amenable to constraints which would be imposed by the substrate being loaded on the SyrB1 arm. Some translation of the substrate is preferred to induce binding with chlorine, but it is only in the lateral direction.

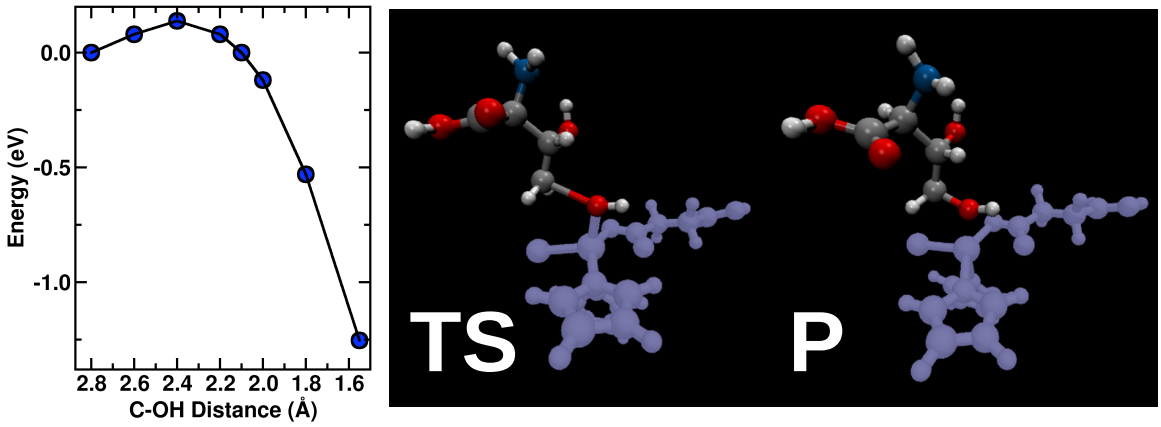


Figure 5-17: Energetics (left) and structure (right) of the hydroxylation step of the L-threonine radical. Unlike chlorination, hydroxylation exhibits a small barrier of about 0.18 eV as a result of strain and elongation of the OH bond during attack and elimination via the radical (see TS at right).

The hydrogen bond which forms between L-threonine radical and the model complex further suggests that hydroxylation as the next step would be prohibitive. A rebound attack on the hydroxyl ligand by the radical would first require rotation of the O-H bond and breaking of the hydrogen bond to the substrate. Additionally, the carbon and oxygen would have to move closer together in order for the chemical bond to form which would require either translation of the L-threonine molecule further into the active site, potentially a prohibitive step depending upon the binding interactions between the SyrB1 arm and the SyrB2 protein, or, alternatively, it would require straining the Fe-OH out of the perpendicular axial position and towards the

substrate. Preliminary calculations have revealed that this step is weakly endothermic (see Fig. 5-17). The greatest cost in undergoing hydroxylation appears to be the significant elongation of the Fe-OH bond from the equilibrium value of 1.85 Å to over 2.0 Å prior to donation to form 4-OH-L-Thr. Previously we observed a relatively tight binding curve for the Fe-OH moiety, and this has carried through to the reaction barrier estimates for hydroxylation. Although the hydroxylation reaction has a small barrier, structural considerations including the flexibility of the substrate when loaded on the SyrB1 arm are much more likely to be critical than the hydroxylation barrier height. In other enzymes in the same family as SyrB2, differences in the binding interaction of the loading arm and substrate with the active site as well as an Asp residue in place of the chlorine are likely to preferentially catalyze hydroxylation.

The halogenation reaction L-Thr to 4-Cl-L-Thr by a model complex of SyrB2 has been studied in detail. Using only the ligands directly bound to iron in the crystal structure's active site, we were able to study and understand many of the key catalytic steps which have been postulated to be relevant in the full biological system. We observed barrierless decarboxylation of α-ketoglutarate to form succinate which generates a high-energy Fe(IV)=O intermediate that has been previously characterized spectroscopically. This reactive intermediate greatly reduces the energetic cost of hydrogen abstraction to form radicals from the uncatalyzed values for both L-threonine and methane, respectively at around 3.8 and 4.0 eV to under 0.30 and 0.38 eV with abstraction completed via the model catalyst. The chlorination step was shown to be barrierless and exothermic for methane by over 0.6 eV. However, we discovered that the formation of 4-Cl-L-Thr is even more energetically favorable when the abstraction and chlorination are coupled and occur in rapid succession. The interactions of the methyl carbon with a leaving chlorine stabilize the radical

during the removal of a hydrogen via the iron-oxo ligand. Upon abstraction of the hydrogen, the chlorination step is enhanced by the presence of a hydrogen bond between the -OH ligand of the catalyst and an -OH group on L-Thr. Upon completion of the catalytic cycle, regeneration of the catalyst and removal of the tightly bound hydroxyl moiety was shown to be possible via exothermic donation of a proton from a hydronium ion to form the weakly bound water observed in the resting state of the full enzyme. Undoubtedly, second shell ligands in this case help to modulate the binding of substrates and co-substrates as well as the relative barrier heights of steps. However, using only a model complex of between 34 and 51 atoms, we were able to accurately describe a plausible mechanism for halogenation which has implications for the full enzyme. In an ongoing effort to further understand differences between hydroxylation and chlorination, we are studying the minimal model complex in which Cl is replaced by an Asp residue. Additionally, work which incorporates selective residues in the second shell up to 450 atoms is ongoing.

Chapter 6

Concluding Remarks

We have shown that a Hubbard U approach may be used fruitfully to augment any standard exchange-correlation functional used in the density functional framework. This approach, already known to be successful in the solid state physics community for improving estimates of redox potentials in batteries and correctly reproducing insulating behavior, has only for the first time in this thesis been applied to molecules. In molecules, we have shown that the Hubbard U approach helps to fix problems created by the non-cancellation of the electronic density in exchange-correlation functionals, commonly known as self-interaction. We showed that the self-interaction problems are highly influential on the method's ability to correctly assign the ground state of transition metal containing molecules because of the strong localization of the $3d$ manifold. The Hubbard U term we apply, typically to GGA but which may also be applied in conjunction with other commonly employed functionals, helps to combat the over-hybridization of $3d$ electrons, preference for bonding states over antibonding or nonbonding states, and the proclivity for fractional occupation of orbitals. While in the solid-state, the two factors of both short-range hybridiza-

tion due to direct ligands as well as long-range charge transfer-based self-interaction errors are present, in the single-site molecules addressed in this thesis, the Hubbard U term has proven most successful as it only is needed to address local hybridization.

We have demonstrated the success of our approach starting from the smallest, but not necessarily simplest systems - the high symmetry diatomic molecules. In discussing the iron dimer, we showed that errors in common density functionals had for many years clashed with experimental results and prevented a clear explanation of such experiments or identification of the ground state. We also showed how different bonding regimes identified by transition metal diatomic molecules formed with hydrogen, oxygen, and fluorine each inspired a slightly different perspective of how the Hubbard U approach may be employed. That is, for the hydrides, we found a U_{4s} to be highly relevant and necessary to describe the strong role of the $4s$ electrons in bonding, while in the fluorides we found the overall role of the U to be diminished as hybridization problems are much more key in strongly covalent systems, such as the oxides. We demonstrated that the GGA+U approach, in particular, can systematically determine the ground state and spin state splittings far better than standard GGA or other popular, but unreliable functionals such as the hybrid functional B3LYP.

Moving on from the most fundamental two atom case, we considered the small but paradigmatic reactions of hydrogen (H_2) and methane (CH_4) on bare FeO^+ molecule to form water (H_2O) and methanol (CH_3OH), respectively. These addition-elimination reactions had not yet been properly described by standard density functional approaches, and energetic errors in barriers and intermediate splittings as large as 1.4 eV for standard GGA were reduced to around 0.1 eV for GGA+U, with respect to highly accurate but expensive quantum chemistry reference. We showed that despite a limitation of our approach - the need for an average U in order to

compare a complete reaction coordinate - GGA+U still produced excellent results on the systems considered. We also demonstrated how using a locally averaged U could greatly improve results even further. In the case of the methane reaction, we were also able to show that transition-states became more bound with GGA+U where they were barely bound with respect to dissociated products for GGA, demonstrating that GGA+U tunes the system to more closely reflect the correct chemical one rather than simply shifting energetics.

Following our successes with small model systems compared to experiment and very accurate theory, we moved to the larger metalloporphyrin systems where we could only compare to experimental details and other density functional approaches. In these cases, we were in particular able to show that cobalt porphines relevant for reduction of CO_2 exhibited much improved binding energies consistent with experiment when GGA+U was used. We also considered the tetrabromophenyl cobalt porphyrins on Cu(111) surface. We were able to show a transformation of oxidation states in these molecules upon formation of a self-assembled monolayer on the Cu(111) surface from Co(II) to Co(I). This change concomitant with charge transfer to the molecule resulted in an increased spin density throughout the macrocycle, and we were only able to observe this feature with the use of GGA+U. By employing a model approach based upon the GGA+U density, we were able to predict Kondo temperatures from the values of averaged spin density. While Kondo resonance is a many-body phenomenon, the proper modeling of the ground state density in this case permitted the direct comparison against anomalous experimental Kondo temperatures which were closely tied to increased spin on the macrocycle.

Finally, we considered a key enzyme, the halogenase SyrB2, and how model systems with GGA+U could be used to elucidate missing details from the halogenating mechanism of this system. We were able to show in particular that the reaction steps

of this enzyme were nearly barrierless, and activation energies for steps such as the hydrogen abstraction were reduced by nearly 4 eV with respect to the uncatalyzed value. In addition, we have proposed that the position of the threonine substrate in this system predisposes the enzyme to rapid chlorination following hydrogen abstraction rather than the slower rebound hydroxylation step. Such work has already provided new details into the reaction of SyrB2 but further work on larger systems could provide even more information about what role second shell ligands play.

Future work on GGA+U approaches exists in both theoretical advances and applications. We wish to implement earlier work (not shown here) on allowing for variation in U along a defined reaction coordinate by integrating a variable force. Additionally, only a fraction of the electrons in the $3d$ manifold participate in linear response. It is necessary to understand whether a renormalized approach including only the relevant portion of the manifold could yield more reliable results. Finally, applications of our approach are endless, and extension to include larger approaches such as linear scaling DFT or QM/MM could permit access to study of a full enzyme with accurate GGA+U treatment at the core. Many enzymes are capable of completing unique chemistry still as yet not understood which are only now with sufficient computing power being studied. Furthermore, we have truly provided for the first time a completely predictive approach which is accurate and efficient. Importantly, the linear-response formalism provides a key to understanding how much the approach itself is even needed. A directed approach which combines screening of transition metal complexes and predictive descriptions of unique new enzymes is truly necessary for identifying new reactions and alternative routes for producing products and harnessing energy in a new, more environmentally friendly way.

Bibliography

[1] Jensen, F. *Introduction to Computational Chemistry*. Wiley (1999).

[2] Cramer, C.J. *Essentials of Computational Chemistry*. Wiley (2003).

[3] Yoshizawa, K. and Yumura, T. *Journal of Inorganic Biochemistry*, **96**(1), 257–257 (2003).

[4] Yoshizawa, K. *Journal of Inorganic Biochemistry*, **78**(1), 23–34 (2000).

[5] Blasiak, L.C., Vaillancourt, F.H., Walsh, C.T. and Drennan, C.L. *Nature*, **440**(7082), 368–371 (2006).

[6] Visser, S.P.D., Ogliaro, F., Gross, Z. and Shaik, S. *Chemistry-A European Journal*, **7**(22), 4954 (2001).

[7] Cococcioni, M. and De Gironcoli, S. *Physical Review B*, **71**(3), 035105 (2005).

[8] Hohenberg, P. and Kohn, W. *Physical Review*, **136**(3B), B864 (1964).

[9] Kohn, W. and Sham, L.J. *Physical Review*, **140**(4A), A1133 (1965).

[10] Tao, J., Perdew, J., Staroverov, V. and Scuseria, G. *Physical Review Letters*, **91**(14) (2003).

[11] Riley, K.E. and Merz, K.M. *Journal of Physical Chemistry A*, **111**(27), 6044–6053 (2007).

[12] Jacobsen, H. and Cavallo, L. *Organometallics*, **25**(1), 177–183 (2006).

[13] Gunnarsson, O. and Lundqvist, B.I. *Physical Review B*, **13**(10), 4274–4298 (1976).

[14] Vonbarth, U. *Physical Review A*, **20**(4), 1693–1703 (1979).

[15] Ziegler, T., Rauk, A. and Baerends, E.J. *Theoretica Chimica Acta*, **43**(3), 261–271 (1977).

[16] Zhou, F., Cococcioni, M., Marianetti, C.A., Morgan, D. and Ceder, G. *Physical Review B*, **70**(23), 235121 (2004).

[17] Kulik, H.J., Cococcioni, M., Scherlis, D.A. and Marzari, N. *Physical Review Letters*, **97**(10), 103001 (2006).

[18] Anisimov, V.I., Zaanen, J. and Andersen, O.K. *Physical Review B*, **44**(3), 943–954 (1991).

[19] Liechtenstein, A.I., Anisimov, V.I. and Zaanen, J. *Physical Review B*, **52**(8), R5467–R5470 (1995).

[20] Fabris, S., de Gironcoli, S., Baroni, S., Vicario, G. and Balducci, G. *Physical Review B*, **71**(4) (2005).

[21] Perdew, J.P. and Levy, M. *Physical Review Letters*, **51**(20), 1884–1887 (1983).

[22] Herzberg, G. *Molecular Spectra and Molecular Structure*. Krieger Publishing Company (1989).

[23] Davidson, E.R. and Silver, D.W. *Chemical Physics Letters*, **52**(3), 403–406 (1977).

[24] QUANTUM-ESPRESSO is a community project for high-quality quantum-simulation software, based on density-functional theory, and coordinated by Paolo Giannozzi. See http://www.quantum-espresso.org and http://www.pwscf.org.

[25] Henkelman, G., Uberuaga, B.P. and Jonsson, H. *Journal of Chemical Physics*, **113**(22), 9901 (2000).

[26] Henkelman, G. and Jonsson, H. *Journal of Chemical Physics*, **113**(22), 9978 (2000).

[27] Iannuzzi, M., Laio, A. and Parrinello, M. *Physical Review Letters*, **90**(23), 238302 (2003).

[28] Ensing, B., Laio, A., Parrinello, M. and Klein, M. *Journal of Physical Chemistry B*, **109**(14), 6676 (2004).

[29] Frenkel, D. and Ladd, A.J.C. *Journal of Chemical Physics*, **81**(7), 3188 (1984).

[30] Xie, L., Liu, H. and Yang, W. *Journal of Chemical Physics*, **120**(17), 8039–8052 (2004).

[31] Perdew, J.P., Burke, K. and Ernzerhof, M. *Physical Review Letters*, **77**(18), 3865–3868 (1996).

[32] Makov, G. and Payne, M.C. *Physical Review B*, **51**(7), 4014 (1995).

[33] Brandsden, B.H. and Joachain, C.J. *Physics of Atoms and Molecules*. Prentice Hall (1982).

[34] Gutsev, G. and Bauschlicher, C. *Journal of Physical Chemistry A*, **107**(23), 4755–4767 (2003).

[35] Gutsev, G., Andrews, L. and Bauschlicher, C. *Theoretical Chemistry Accounts*, **109**(6), 298–308 (2003).

[36] Schultz, N., Zhao, Y. and Truhlar, D. *Journal of Physical Chemistry A*, **109**(49), 11127–11143 (2005).

[37] Furche, F. and Perdew, J.P. *Journal of Chemical Physics*, **124**(4), 044103 (2006).

[38] Zhao, Y. and Truhlar, D.G. *Journal of Chemical Physics*, **124**(22) (2006).

[39] Cassell, A.M., Raymakers, J.A., Kong, J. and Dai, H. *Journal of Physical Chemistry B*, **103**(31), 6484 (1999).

[40] Ballone, P. and Jones, R. *Chemical Physics Letters*, **233**(5-6), 632 (1995).

[41] Castro, M. and Salahub, D.R. *Physical Review B*, **49**(17), 11842 (1994).

[42] Chen, J., Wang, C., Jackson, K.A. and Pederson, M.R. *Physical Review B*, **44**(12), 6558 (1991).

[43] Dhar, S. and Kestner, N. *Physical Review A*, **38**(3), 1111 (1988).

[44] Harris, J. and Jones, R. *Journal of Chemical Physics*, **70**(02), 830 (1979).

[45] Nagarathna, H., Montano, P.A. and Naik, V.M. *Journal of the American Chemical Society*, **105**(10), 2938 (1983).

[46] Noro, T., Ballard, C., Palmer, M.H. and Tatewaki, H. *Journal of Chemical Physics*, **100**(1), 452 (1994).

[47] Shim, I. and Gingerich, K.A. *Journal of Chemical Physics*, **77**(5), 2490 (1982).

[48] Tomonari, M. and Tatewaki, H. *Journal of Chemical Physics*, **88**(3), 1828 (1988).

[49] Yanagisawa, S., Tsuneda, T. and Hirao, K. *Journal of Chemical Physics*, **112**(2), 545 (2000).

[50] Rohlfing, E., Cox, D., Kaldor, A. and Johnson, K. *Journal of Chemical Physics*, **81**(9), 3846 (1984).

[51] Leopold, D.G. and Lineberger, W. *Journal of Chemical Physics*, **85**(1), 51 (1986).

[52] Barden, C.J., Rienstra-Kiracofe, J.C. and Schaefer III, H.F. *Journal of Chemical Physics*, **113**(2), 690 (2000).

[53] Hubner, O. and Sauer, J. *Chemical Physics Letters*, **358**(5-6), 442 (2002).

[54] Irigoras, A., Michelini, M., Sicilia, E., Russo, N., Mercero, J. and Ugalde, J. *Chemical Physics Letters*, **376**(3-4), 310 (2003).

[55] Cox, D., Trevor, D., Whetten, R., Rohlfing, E. and Kaldor, A. *Physical Review B*, **32**(11), 7290 (1985).

[56] Montano, P., Barrett, P., and Shanfield, Z. *Journal of Chemical Physics*, **64**(7), 2896 (1976).

[57] Guenzburger, D. and Saitovitch, E.M.B. *Physical Review B*, **24**(5), 2368 (1981).

[58] McNab, T., Micklitz, H. and Barrett, P. *Physical Review B*, **4**(11), 3787 (1971).

[59] Purdum, H., Montano, P., Shenoy, G. and Morrison, T. *Physical Review B*, **25**(7), 4412 (1982).

[60] Leopold, D., Almlof, J., Lineberger, W. and Taylor, P. *Journal of Chemical Physics*, **88**(6), 3780 (1988).

[61] Das, G. *Journal of Chemical Physics*, **74**(10), 5766–5774 (1981).

[62] Sodupe, M., Lluch, J.M., Oliva, A., Illas, F. and Rubio, J. *Journal of Chemical Physics*, **92**(4), 2478–2480 (1990).

[63] Tanaka, K., Sekiya, M. and Yoshimine, M. *Journal of Chemical Physics*, **115**(10), 4558–4564 (2001).

[64] Harrison, J. *Chemical Reviews*, **100**(2), 679–716 (2000).

[65] Stevens, A.E., Feigerle, C.S. and Lineberger, W.C. *Journal of Chemical Physics*, **78**(9), 5420–5431 (1983).

[66] Blint, R.J., Kunz, A.B. and Guse, M.P. *Chemical Physics Letters*, **36**(2), 191–194 (1975).

[67] Balfour, W.J. *Journal of Chemical Physics*, **88**(8), 5242–5243 (1988).

[68] Green, S., Alex, S., Fleischer, N., Millam, E., Marcy, T. and Leopold, D. *Journal of Chemical Physics*, **114**(6), 2653–2668 (2001).

[69] Kulik, H.J. and Marzari, N. *Journal of Chemical Physics*, **129**(13) (2008).

[70] Miller, A.E.S., Feigerle, C.S. and Lineberger, W.C. *Journal of Chemical Physics*, **87**(3), 1549–1556 (1987).

[71] Ram, R.S., Jarman, C.N. and Bernath, P.F. *Journal of Molecular Spectroscopy*, **161**(2), 445–454 (1993).

[72] Calvi, R.M.D., Andrews, D.H. and Lineberger, W.C. *Chemical Physics Letters*, **442**(1-3), 12–16 (2007).

[73] Schroder, D., Schwarz, H., Clemmer, D.E., Chen, Y.M., Armentrout, P.B., Baranov, V.I. and Bohme, D.K. *International Journal of Mass Spectrometry and Ion Processes*, **161**(1-3), 175–191 (1997).

[74] Schroder, D., Fiedler, A., Ryan, M.F. and Schwarz, H. *Journal of Physical Chemistry*, **98**(1), 68–70 (1994).

[75] Clemmer, D.E., Chen, Y.M., Khan, F.A. and Armentrout, P.B. *Journal of Physical Chemistry*, **98**(26), 6522–6529 (1994).

[76] Mercero, J.M., Matxain, J.M., Lopez, X., York, D.M., Largo, A., Eriksson, L.A. and Ugalde, J.M. *International Journal of Mass Spectrometry*, **240**(1), 37–99 (2005).

[77] Andersson, K., Malmqvist, P.A., Roos, B.O., Sadlej, A.J. and Wolinski, K. *Journal of Physical Chemistry*, **94**(14), 5483–5488 (1990).

[78] Fiedler, A., Schroder, D., Shaik, S. and Schwarz, H. *Journal of the American Chemical Society*, **116**(23), 10734–10741 (1994).

[79] Danovich, D. and Shaik, S. *Journal of the American Chemical Society*, **119**(7), 1773–1786 (1997).

[80] Filatov, M. and Shaik, S. *Journal of Physical Chemistry A*, **102**(21), 3835–3846 (1998).

[81] Chiodo, S., Kondakova, O., Michelini, M.D., Russo, N., Sicilia, E., Irigoras, A. and Ugalde, J.M. *Journal of Physical Chemistry A*, **108**(6), 1069–1081 (2004).

[82] Irigoras, A., Fowler, J.E. and Ugalde, J.M. *Journal of the American Chemical Society*, **121**(37), 8549–8558 (1999).

[83] Rosi, M. and Bauschlicher, C.W. *Journal of Chemical Physics*, **90**(12), 7264–7272 (1989).

[84] Aguirre, F., Husband, J., Thompson, C.J., Stringer, K.L. and Metz, R.B. *Journal of Chemical Physics*, **119**(19), 10194–10201 (2003).

[85] Yoshizawa, K., Shiota, Y. and Yamabe, T. *Chemistry-A European Journal*, **3**(7), 1160–1169 (1997).

[86] Yoshizawa, K., Shiota, Y. and Yamabe, T. *Organometallics*, **17**(13), 2825–2831 (1998).

[87] Bronstrup, M., Schroder, D. and Schwarz, H. *Chemistry-A European Journal*, **5**(4), 1176–1185 (1999).

[88] Yoshizawa, K., Shiota, Y. and Yamabe, T. *Journal of Chemical Physics*, **111**(2), 538–545 (1999).

[89] Yoshizawa, K., Suzuki, A. and Yamabe, T. *Journal of the American Chemical Society*, **121**(22), 5266–5273 (1999).

[90] Shiota, Y. and Yoshizawa, K. *Journal of the American Chemical Society*, **122**(49), 12317–12326 (2000).

[91] Shiota, Y. and Yoshizawa, K. *Journal of Chemical Physics*, **118**(13), 5872–5879 (2003).

[92] Aguirre, F., Husband, J., Thompson, C.J., Stringer, K.L. and Metz, R.B. *Journal of Chemical Physics*, **116**(10), 4071–4078 (2002).

[93] Schroder, D. and Schwarz, H. *Angewandte Chemie-International Edition in English*, **29**(12), 1433–1434 (1990).

[94] Schwarz, H. *International Journal of Mass Spectrometry*, **237**(1), 75–105 (2004).

[95] Bohme, D.K. and Schwarz, H. *Angewandte Chemie-International Edition*, **44**(16), 2336–2354 (2005).

[96] Schwarz, H. and Schroder, D. *Pure and Applied Chemistry*, **72**(12), 2319–2332 (2000).

[97] Rosenzweig, A.C., Frederick, C.A., Lippard, S.J. and Nordlund, P. *Nature*, **366**(6455), 537–543 (1993).

[98] Derat, E., Shaik, S., Rovira, C., Vidossich, P. and Alfonso-Prito, M. *Journal of the American Chemical Society*, **129**(20), 6346 (2007).

[99] Hirao, H., Kumar, D., Que, L. and Shaik, S. *Journal of the American Chemical Society*, **128**(26), 8590–8606 (2006).

[100] Schroder, D., Shaik, S. and Schwarz, H. *Accounts of Chemical Research*, **33**(3), 139–145 (2000).

[101] M. J. Frisch et al. (2004). See http://www.gaussian.com.

[102] H. J. Werner et al. (2006). See http://www.molpro.net.

[103] Chiodo, S., Russo, N. and Sicilia, E. *Journal of Computational Chemistry*, **26**(2), 175–183 (2005).

[104] Sit, P.H.L., Cococcioni, M. and Marzari, N. *Journal of Electroanalytical Chemistry* (2007).

[105] Matxain, J.M., Mercero, J.M., Irigoras, A. and Ugalde, J.M. *Molecular Physics*, **102**(23-24), 2635–2637 (2004).

[106] Dabo, I., Wieckowski, A. and Marzari, N. *Journal of the American Chemical Society*, **129**, 11045–11052 (2007).

[107] Goff, H., Lamar, G.N. and Reed, C.A. *Journal of the American Chemical Society*, **99**, 3641–6 (1977).

[108] Peng, S.M. and Ibers, J.A. *Journal of the American Chemical Society*, **98**(25), 8032–8036 (1976).

[109] Perutz, M.F. *Annual Reviews of Biochemistry*, **48**, 327–386 (1979).

[110] Edwards, W.D., Weiner, B. and Zerner, M.C. *Journal of the American Chemical Society*, **108**(9), 2196–2204 (1986).

[111] Angelis, F.D., Jin, N., Car, R. and Groves, J.T. *Inorganic Chemistry*, **45**(10), 4268 (2006).

[112] Groves, J.T., Lee, J. and Marla, S.S. *Journal of the American Chemical Society*, **119**(27), 6269 (1997).

[113] Jin, N. and Groves, J.T. *Journal of the American Chemical Society*, **121**(12), 2923 (1999).

[114] Scherlis, D.A., Cococcioni, M., Sit, P. and Marzari, N. *Journal of Physical Chemistry B*, **111**(25), 7384–7391 (2007).

[115] Behar, D., Dhanasekaran, T., Neta, P., Hosten, C.M., Ejeh, D., Hambright, P. and Fujita, E. *Journal of Physical Chemistry A*, **102**, 2870–2877 (1998).

[116] Park, J., Pasupathy, A.N., Goldsmith, J.I., Chang, C., Yaish, Y., Petta, J.R., Rinkoswki, M., Sethna, J.P., Abruna, H.D., l. McEuen, P. and Ralph, D.C. *Nature*, **417**, 722–724 (2002).

[117] Zhao, A.D., Li, Q.X., Chen, L., Xiang, H.J., Wang, W.H., Pan, S., Wang, B., Xiao, X.D., l. Yang, J., Hou, J.G. and Zhu, Q.S. *Science*, **309**, 1542–1544 (2005).

[118] Madhavan, V., Chen, W., Jamneala, T., Crommie, M.F. and Wingreen, N.S. *Science*, **280**, 567–569 (1998).

[119] Iancu, V., Deshpande, A. and Hla, S.W. *Physical Review Letters*, **97**, 266603 (2006).

[120] Henzl, J. and Morgenstern, K. *Physical Review Letters*, **98**, 266601 (2007).

[121] Iancu, V., Deshpande, A. and Hla, S.W. *Nano Letters*, **6**, 820–823 (20006).

[122] Walker, F.A. *Journal of the American Chemical Society*, **95**(4), 1150–1153 (1973).

[123] Lowdin, P.O. *Advances in Quantum Chemistry*, **5**(185) (1970).

[124] Perera, U.G.E., Kulik, H.J., Iancu, V., Dias da Silva, L.G.G.V., Ulloa, S.E., Marzari, N. and Hla, S.W. *Submitted to Nature Physics* (2008).

[125] Ujsaghy, O., Kroha, J., Szunyogh, L. and Zawadowski, A. *Physical Review Letters*, **85**, 2557–2560 (2000).

[126] Gribble, G.W. *Journal of Chemical Education*, **81**(10), 1441–1449 (2004).

[127] Higgins, L.J., Yan, F., Liu, P.H., Liu, H.W. and Drennan, C.L. *Nature*, **437**(7060), 838–844 (2005).

[128] Vaillancourt, F.H., Yeh, E., Vosburg, D.A., Garneau-Tsodikova, S. and Walsh, C.T. *Chemical Reviews*, **106**(8), 3364–3378 (2006).

[129] Neumann, C.S., Fujimori, D.G. and Walsh, C.T. *Chemistry & Biology*, **15**(2), 99–109 (2008).

[130] Grgurina, I., Barca, A., Cervigni, S., Gallo, M., Scaloni, A. and Pucci, P. *Experientia*, **50**(2), 130–133 (1994).

[131] Chen, X. and Van Pee, K.H. *Acta Biochimica Et Biophysica Sinica*, **40**(3), 183–193 (2008).

[132] Fujimori, D.G. and Walsh, C.T. *Current Opinion in Chemical Biology*, **11**(5), 553–560 (2007).

[133] Anderson, J.L.R. and Chapman, S.K. *Molecular Biosystems*, **2**(8), 350–357 (2006).

[134] Vaillancourt, F.H., Yin, J. and Walsh, C.T. *Proceedings of the National Academy of Sciences of the United States of America*, **102**(29), 10111–10116 (2005).

[135] Hausinger, R.P. *Critical Reviews in Biochemistry and Molecular Biology*, **39**(1), 21–68 (2004).

[136] Purpero, V. and Moran, G.R. *Journal of Biological Inorganic Chemistry*, **12**(5), 587–601 (2007).

[137] Flashman, E. and Schofield, C.J. *Nature Chemical Biology*, **3**(2), 86–87 (2007).

[138] Hewitson, K.S., Holmes, S.L., Ehrismann, D., Hardy, A.P., Chowdhury, R., Schofield, C.J. and McDonough, M.A. *Journal of Biological Chemistry*, **283**(38), 25971–25978 (2008).

[139] Vaillancourt, F.H., Vosburg, D.A. and Walsh, C.T. *ChemBioChem*, **7**(5), 748–752 (2006).

[140] Vaillancourt, F.H., Yeh, E., Vosburg, D.A., O'Connor, S.E. and Walsh, C.T. *Nature*, **436**(7054), 1191–1194 (2005).

[141] Sinnecker, S., Svensen, N., Barr, E.W., Ye, S., Bollinger, Jr., J.M., Neese, F. and Krebs, C. *Journal of the American Chemical Society*, **129**(19), 6168–6179 (2007).

[142] Hegg, E.L. and Que, L. *European Journal of Biochemistry*, **250**(3), 625–629 (1997).

[143] Singh, G.M., Fortin, P.D., Koglin, A. and Walsh, C.T. *Biochemistry*, **47**(43), 11310–11320 (2008).

[144] Grzyska, P.K., Muller, T.A., Campbell, M.G. and Hausinger, R.P. *Journal of Inorganic Biochemistry*, **101**(5), 797–808 (2007).

[145] Price, J.C., Barr, E.W., Glass, T.E., Krebs, C. and Bollinger, J.M. *Journal of the American Chemical Society*, **125**(43), 13008–13009 (2003).

[146] Goldsmith, C., Jonas, R., Cole, A. and Stack, T. *Inorganic Chemistry*, **41**(18), 4642–4652 (2002).

[147] Allen, F. *Acta Crystallographica Section B-Structural Science*, **58**(Part 3 Sp. Iss. 1), 380–388 (2002).

[148] Galonic, D.P., Barr, E.W., Walsh, C.T., Bollinger, Jr., J.M. and Krebs, C. *Nature Chemical Biology*, **3**(2), 113–116 (2007).

[149] Krebs, C., Fujimori, D.G., Walsh, C.T. and Bollinger, Jr., J.M. *Accounts of Chemical Research*, **40**(7), 484–492 (2007).

[150] Fujimori, D.G., Barr, E.W., Matthews, M.L., Koch, G.M., Yonce, J.R., Walsh, C.T., Bollinger, Jr., J.M., Krebs, C. and Riggs-Gelasco, P.J. *Journal of the American Chemical Society*, **129**(44), 13408 (2007).

[151] Lucas, R.L., Zart, M.K., Murkerjee, J., Sorrell, T.N., Powell, D.R. and Borovik, A.S. *Journal of the American Chemical Society*, **128**(48), 15476–15489 (2006).

[152] Friese, S.J., Kucera, B.E., Young, Jr., V.G., Que, Jr., L. and Tolman, W.B. *Inorganic Chemistry*, **47**(4), 1324–1331 (2008).

[153] Neidig, M.L., Brown, C.D., Kavana, M., Choroba, O.W., Spencer, J.B., Moran, G.R. and Solomon, E.I. *Journal of Inorganic Biochemistry*, **100**(12), 2108–2116 (2006).

[154] Neidig, M.L., Brown, C.D., Light, K.M., Fujimori, D.G., Nolan, E.M., Price, J.C., Barr, E.W., Bollinger, Jr., J.M., Krebs, C., Walsh, C.T. and Solomon, E.I. *Journal of the American Chemical Society*, **129**(46), 14224–14231 (2007).