Title
Studies on Secondary Coordination Sphere Effects in Metal and Enzyme Catalysis

Author
Fuller, Jack Terrell

Publication Date
2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Studies on Secondary Coordination Sphere Effects in Metal and Enzyme Catalysis

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Chemistry

by

Jack Terrell Fuller III

2021

ABSTRACT OF THE DISSERTATION

Studies on Secondary Coordination Sphere Effects

in Metal and Enzyme Catalysis

by

Jack Terrell Fuller III

Doctor of Philosophy in Chemistry

University of California, Los Angeles, 2021

Professor Anastassia N. Alexandrova, Chair

In metal catalysts and enzymes, the primary coordination sphere of the metal and/or substrate determines basic reactivity. The secondary coordination sphere tunes the reactivity through weaker interactions and therefore usually presents the best handle for optimizing artificial enzymes and metal catalysts. We studied an artificial metalloenzyme system and found that the enzyme improves activity through a pi-stacking interaction in the secondary coordination sphere. The dynamics of this interaction altered reactivity among different mutants. In a natural enzyme, we found that the electric field generated solely by the primary and secondary coordination spheres was sufficient to reproduce experimental energy surfaces. We also studied a redox-active bimetallic system, finding that the redox potential could be tuned by simple electrostatics in the secondary coordination sphere if the primary coordination sphere was rigid.

The dissertation of Jack Terrell Fuller III is approved.

Justin Caram

Daniel Neuhauser

Jose Rodriguez

Anastassia N. Alexandrova, Committee Chair

University of California, Los Angeles

2021

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SCHEMES

ACKNOWLEDGMENTS

Funding was received from a George Gregory Fellowship and computing resources from the UCLA Hoffman2 Cluster and XSEDE.

Images were rendered using UCSF Chimera[1] (developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311), Chemcraft (https://www.chemcraftprog.com), Jupyter Notebook, ChemDraw, and Microsoft Office.

Chapter 4 is reproduced in part with permission from Fuller, J., III; Wilson, T. R.; Eberhart, M. E.; Alexandrova, A. N. Charge Density in Enzyme Active Site as a Descriptor of Electrostatic Preorganization. *J. Chem. Inf. Model.* **2019**, *59*, 2367−2373. Copyright 2019 American Chemical Society.

Chapter 5 is reproduced in part from Kang, K.; Fuller, J., III; Reath, A. H.; Ziller, J. W.; Alexandrova, A. N.; Yang, J. Y. Installation of Internal Electric Fields by Non-Redox Active Cations in Transition Metal Complexes. *Chem. Sci.* **2019**, *10*, 10135–10142 with permission from the Royal Society of Chemistry.

Reference:

1. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem.* **2004,** *25*, 1605–1612.

VITA

Missionary, The Church of Jesus Christ of Latter-Day Saints, 2009-2011

B.S. in Chemistry, Brigham Young University, 2014

M.S. in Chemistry, Brigham Young University, 2016

CHAPTER 1: INTRODUCTION

The basic reactivity of a metal catalyst is determined by the identity of the metal and its primary coordination sphere, or the atoms immediately bound to it. The secondary coordination sphere tunes the reactivity through weaker interactions such as hydrogen bonds, charge transfer, and steric interactions. Enzymes usually control the secondary coordination sphere more than simple organometallic complexes in solution. We studied effects originating from the secondary coordination spheres of metal catalysts, including metalloenzymes, and enzymes in general. We principally used density functional theory and molecular dynamics.

Density functional theory (DFT) is based on the Hohenberg-Kohn theorem[1] which proved that the ground-state electron density of a system uniquely determines its energy. The exact density functional is unknown, so we use approximate functionals. Kohn-Sham theory[2] decomposes the density into orbitals to yield more accurate energies in a sort of "corrected" Hartree-Fock calculation. Kohn-Sham theory was used to develop many functionals including M06, M06-2X,[3] and ωB97X-D.[4] DFT calculations are particularly efficient, having costs like those of Hartree-Fock calculations but yielding higher accuracy.[5]

In practice, orbitals and densities are approximated using basis sets of gaussian functions. These gaussian functions are contracted to form smaller, more efficient basis sets resembling hydrogen-like orbitals. Larger basis sets yield higher accuracy at increased computational cost. Often, for efficiency, geometries are optimized with smaller basis sets, and final electronic energies are calculated with larger ones. The def2-SVP basis set is small and ideal for geometry optimizations, and the def2-TZVP basis set is larger, yielding accurate energies.[6] These basis sets include effective core potentials starting with fifth-row elements; the inner-most electrons are replaced with effective potentials mimicking their effect on the more chemically relevant outer

electrons. This usually saves the computational cost of explicitly modeling them without sacrificing significant accuracy.

DFT calculations can often be performed more efficiently using the Resolution of the Identity approximation.[7] This approximation uses sums of auxiliary basis functions to replace products of primary basis functions which reduces the number of integrals that must be calculated. Additionally, a multipole approximation can be used for distant electron density interactions, further increasing efficiency.[8]

Because density functionals in general do not adequately model long-range dispersion interactions, dispersion corrections have been developed such as the D3 correction of Grimme and co-workers.[9] This correction can be added to most functionals. Functionals such as ωB97X-D[4] were developed with empirical dispersion corrections already included.

Solvation effects can be approximated implicitly with continuum models. The Conductor-like Screening Model (COSMO) wraps the explicit system in a dielectric medium mimicking the polarization of solvent molecules.[10] This is done by tessellating the solvent-accessible surface and calculating charges for each section. The SMD method computes the bulk-electrostatic contribution to the solvation in a similar manner, but it also adds terms corresponding to cavity-creation, dispersion effects, and hydrogen bonding.[11] These additional terms are parameterized based on solvent descriptors such as refractive index, surface tension, and fraction of atoms that are aromatic carbons. This makes SMD useful with less common solvents which have not been parameterized in other solvent models.

Even with implicit solvation and smaller basis sets, proteins are computationally expensive to model using DFT. Molecular force fields must be used, especially for simulations of protein dynamics. Even with molecular force fields, protein dynamics can be expensive. Discrete

Molecular Dynamics (DMD) approximates force fields with step potentials (Figure 1-1), allowing the dynamics to be run using collision events instead of timesteps.[12] This sacrifices some accuracy for much greater efficiency.

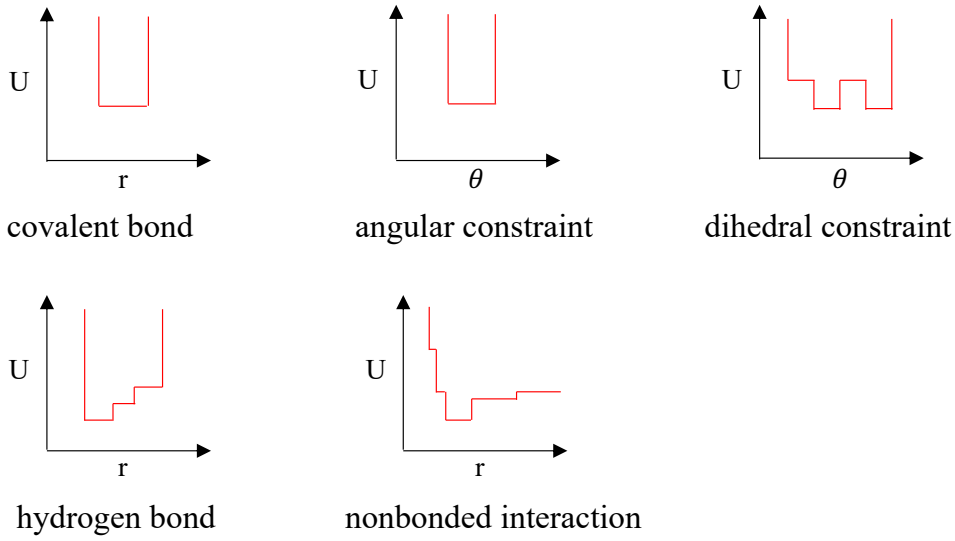


Figure 1-1. Qualitative representations of DMD step potentials.

The Alexandrova group has expanded the use of DMD to hybrid quantum mechanics/molecular mechanics simulations.[13] In the QM/DMD method, an active site is chosen, usually including the metal, substrates, and sidechains that most directly interact with them (Figure 1-2). Atoms that would connect to the rest of the protein (usually the β carbons of sidechains) are capped with hydrogens and frozen in place while an optimization is performed with an *ab initio* method (usually DFT). Then, the metal and atoms immediately bound to it are frozen while DMD is used to sample structures of the rest of the protein. This procedure is performed iteratively. This is particularly effective for simulating metalloenzymes because it allows the metal's primary coordination sphere to be treated quantum mechanically while also efficiently sampling structures of the entire protein.

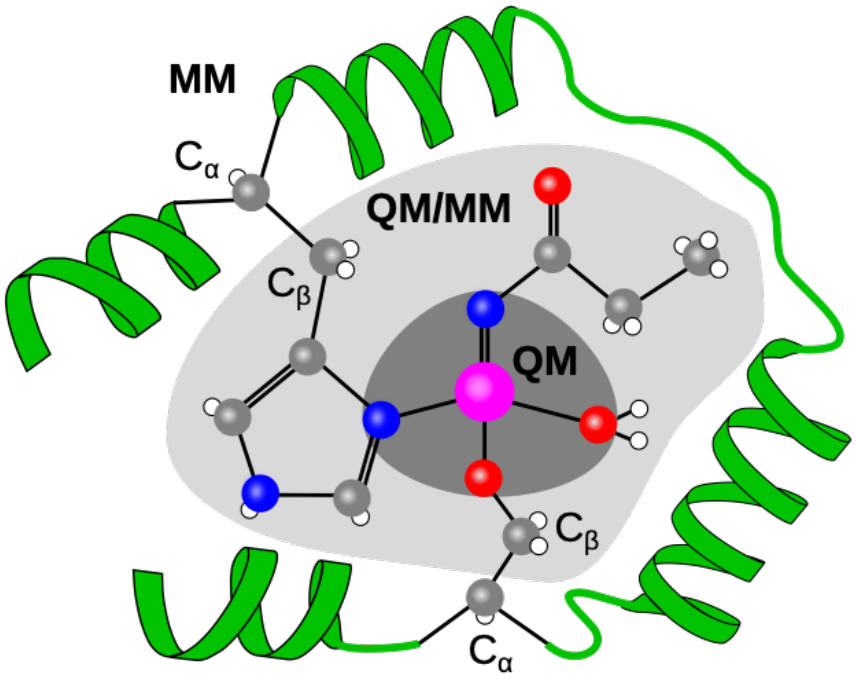


Figure 1-2. Partitioning used in the QM/DMD algorithm. The dark grey region is only treated with quantum mechanics, and the white region is only treated with DMD. The light grey region is treated with each during the corresponding phase of each iteration.

We have used these methods to study electric fields in artificial metalloenzymes, natural enzymes, and redox-active metal complexes.

References:

1. Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev. [Sect.] B* **1964**, *136*, 864–871.
2. Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev. [Sect.] A* **1965**, *140*, 1133–1138.
3. Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States,

and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215–241.

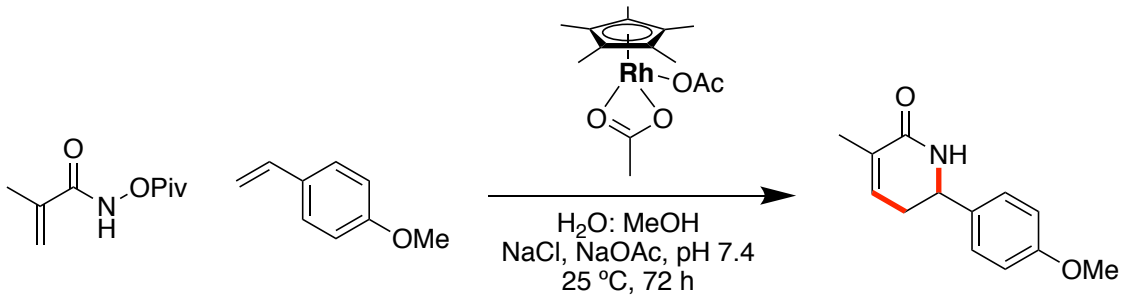
4. Chai, J.-D.; Head-Gordon, M. Long-Range Corrected Hybrid Density Functionals with Damped Atom–Atom Dispersion Corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.
5. Mardirossian, N.; Head-Gordon, M. Thirty Years of Density Functional Theory in Computational Chemistry: An Overview and Extensive Assessment of 200 Density Functionals. *Mol. Phys.* **2017**, *115*, 2315–2372.
6. Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
7. Vahtras, O.; Almlöf, J.; Feyereisen, M. W. Integral Approximations for LCAO-SCF Calculations. *Chem. Phys. Lett.* **1993**, *213*, 514–518.
8. Sierka, M.; Hogekamp, A.; Ahlrichs, R. Fast Evaluation of the Coulomb Potential for Electron Densities Using Multipole Accelerated Resolution of Identity Approximation. *J. Chem. Phys.* **2003**, *118*, 9136–9148.
9. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
10. Klamt, A.; Schüürmann, G. COSMO: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and its Gradient. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799–805.

11. Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113*, 6378–6396.
12. Ding, F.; Tsao, D.; Nie, H.; Dokholyan, N. V. Ab Initio Folding of Proteins with All-Atom Discrete Molecular Dynamics. *Structure (Oxford, U. K.)* **2008**, *16*, 1010–1018.
13. Sparta, M.; Shirvanyants, D.; Ding, F.; Dokholyan, N. V.; Alexandrova, A. N. Hybrid Dynamics Simulation Engine for Metalloproteins. *Biophys. J.* **2012**, *103*, 767–776.

CHAPTER 2: RH-SAV METALLOENZYMES

Metalloenzymes combine the reactivity of metal centers with the controllable environment of enzymes. In nature, some metalloenzymes, such as carbonic anhydrase, operate near the diffusion limit. Artificially designed metalloenzymes rarely achieve such reactivity. One exception is the recent work of Hartwig and co-workers.[1] They took a P450 enzyme, inserted iridium, and mutated a few key residues to make a catalyst for an abiological reaction (carbene insertion) with activity near that of natural enzymes. With the goal of making successes like this more frequent, we studied artificial metalloenzymes to identify design principles.

Rovis and co-workers developed an artificial rhodium enzyme using the biotin-streptavidin interaction.[2] Streptavidin binds biotin with a dissociation constant of ~10^{-14}. Therefore, covalently attaching an organometallic catalyst to biotin creates a secondary coordination sphere from streptavidin. Streptavidin is usually tetrameric, but a monomeric mutant was used in this case. The chosen organometallic catalyst was a Rh piano-stool complex catalyzing the formation of a six-membered lactam from N-pivaloylmethacrylamide and p-methoxystyrene (Scheme 2-1).



Scheme 2-1. Organometallic catalyst and reaction chosen as focus for metalloenzyme design of Rovis and co-workers.[2]

Rovis and co-workers found that the interaction with streptavidin increased the yield of the biotinylated Rh catalyst from 15 percent to 99 percent in 72 hours and made the reaction highly enantioselective.[2] They additionally performed mutations across the protein scaffold, most of

which decreased turnover number (TON) compared to wild-type (Figure 2-1). We collaborated with them to explore how the interaction with streptavidin improves activity as measured by TON, how mutations affect this interaction and therefore activity, and which mutations could improve activity. From these results, we wanted to identify general principles for designing metalloenzymes.

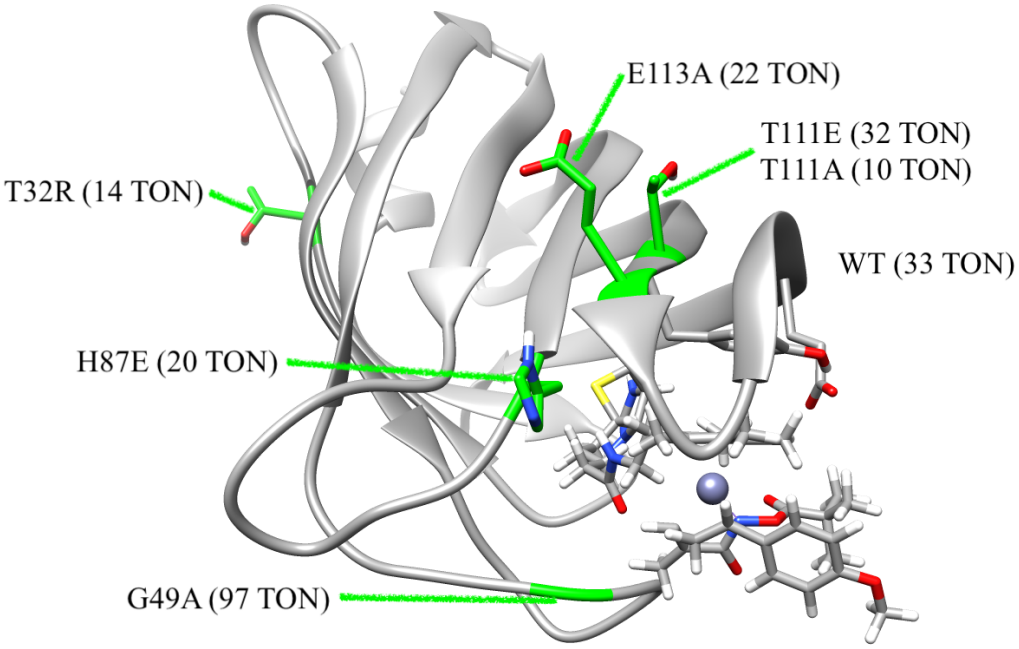


Figure 2-1. Mutations and corresponding TON for representative mutants.

We assumed streptavidin improved activity by lowering the free energy barrier for the rate-determining step of the mechanism, so we used density functional theory to calculate the mechanism of a small model of the catalyst. We used the M06 functional[3] with the def2-TZVP and def2-SVP basis sets[4] for electronic energies and geometries, respectively, and we used the multipole-accelerated resolution of the identity approximation.[5–9] We also added dispersion corrections.[10] We used the Conductor-like Screening Model[11] to implicitly model solvation with a

dielectric of 80. All quantum-mechanical calculations were performed in TURBOMOLE.[12–17] We based our exploration on the mechanism proposed by Rovis and co-workers.[2]

We found that the mechanism began with a bidentate acetate ligand becoming monodentate as the methacrylamide substrate associated (Figure 2-2). N–H activation formed a covalent metal-substrate bond and released acetic acid coupled with alkene association. This allowed C–H activation, which formed a five-membered rhodacycle and released a second equivalent of acetic acid. After association of the styrene substrate, migratory insertion increased the ring size to seven and formed the new C–C bond. Oxidative addition and reductive elimination formed the N-heterocycle, and protodemetallation led to product release.

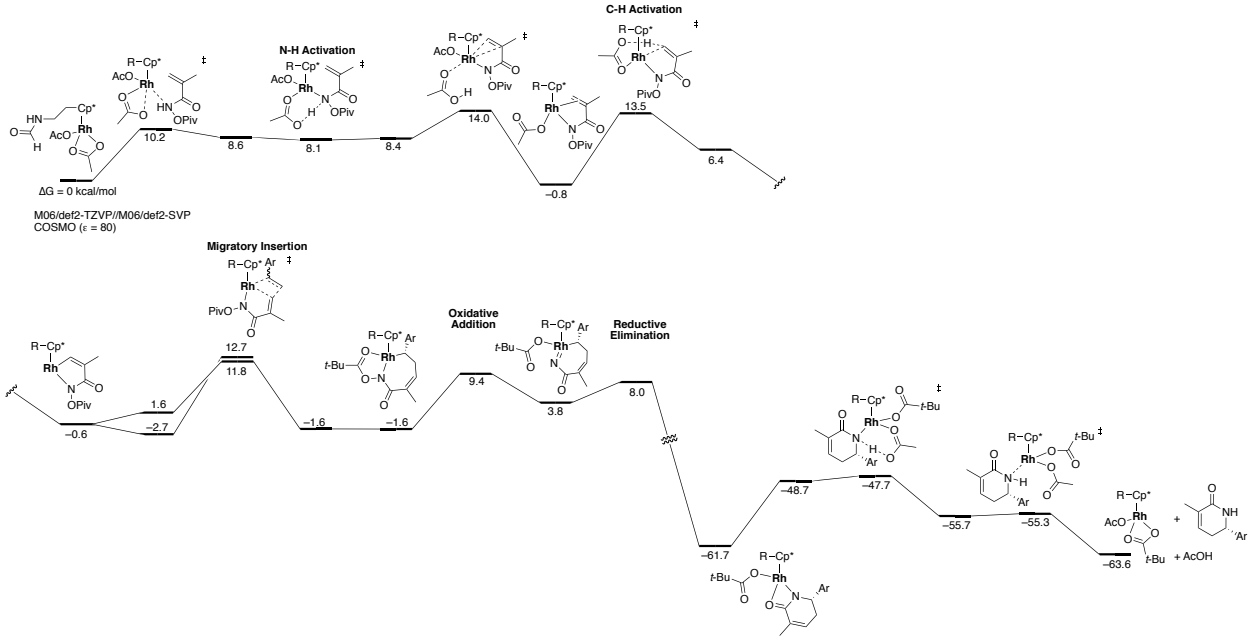


Figure 2-2. Calculated free energy surface for the small model.

The four highest free energy barriers were all near 14 kcal/mol, making the rate-determining step ambiguous. However, our collaborators performed isotope-exchange experiments which suggested that the N–H activation and C–H activation steps were reversible. This implied that either the migratory insertion or the protodemetallation step was rate-

determining. We chose to focus on the migratory insertion barrier because it was slightly larger than the protodemetallation barrier and because preliminary calculations had estimated a smaller protodemetallation barrier.

The residue Y112 appeared to be closest to the metal center, engaging in a pi-stacking interaction with the cyclopentadienyl ring of the base Rh catalyst (Figure 2-3). A hydrogen bond interaction between Y112 and E124 also seemed likely. We therefore added 4-methylphenol and acetate to our small model and recalculated the migratory insertion barrier. We found that it decreased by about 5 kcal/mol (Figure 2-4), which explained the activity difference when the biotinylated catalyst was first added to streptavidin. We hypothesized that electron donation from glutamate and tyrosine to the metal center was responsible for this effect as suggested by population analysis results.

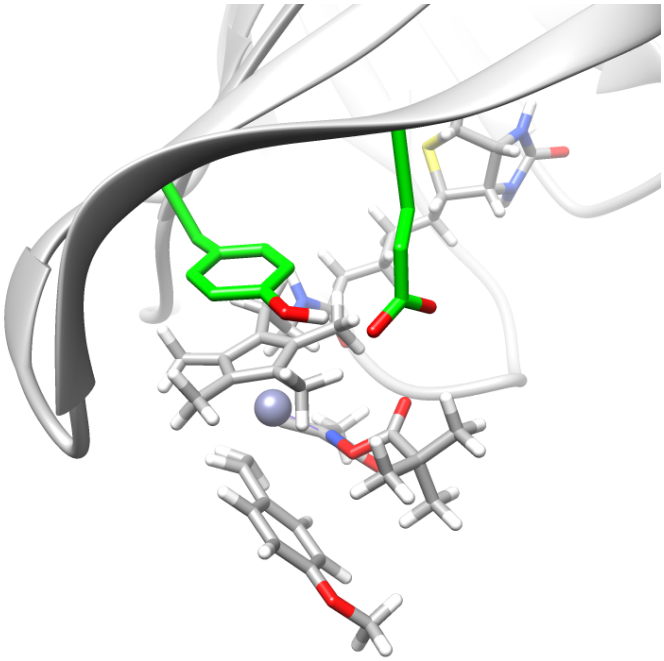


Figure 2-3. Pi-stacking interaction of Y112 and hydrogen bonding interaction of E124.

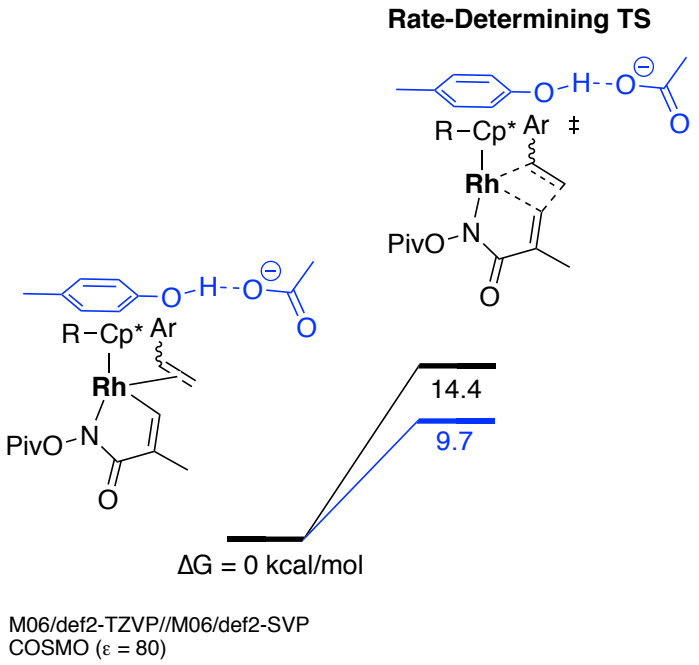


Figure 2-4. Effect of Y112 and E124 sidechains on migratory insertion barrier in small model.

However, in the small model, the sidechain analogs were unconstrained, whereas in the protein, the backbone likely enforced a less favorable interaction. We assumed that the free energy barrier would increase as these sidechains were restrained further and further from the metal center and that this increase would asymptotically approach $\Delta\Delta G^{\ddagger} = 5$ kcal/mol. This was confirmed by a scan of the barrier with increasing Rh-Y112 distance, which also showed exponential curvature (Figure 2-5). This curvature implied that the rate increased with a double exponential dependence as the Rh-Y112 distance decreased, suggesting that only structures with short distances contribute significantly to the overall rate.

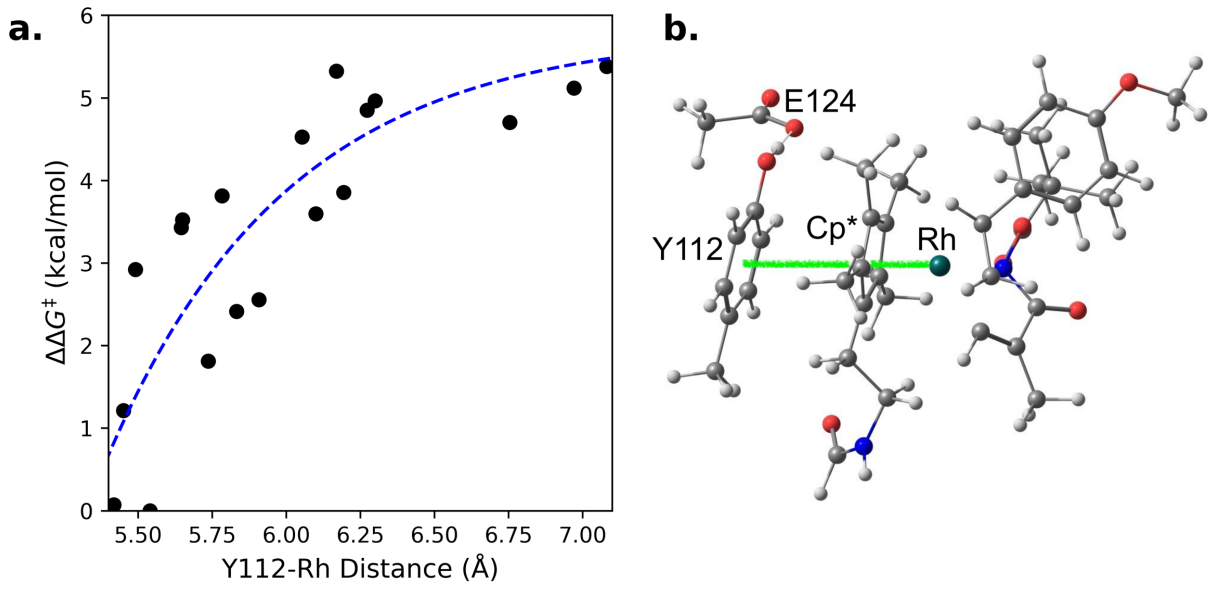


Figure 2-5. a) Scan of change in barrier with Y112-Rh distance. b) Definition of Y112-Rh distance.

To test whether this interaction could account for differences in activity in different mutants, we performed QM/DMD simulations of several representative mutants. For each mutant, we ran 5 replicates of 40 QM/DMD iterations. Each iteration ran for 10,000 time steps, which is approximately equivalent to 500 ps. Structures were saved every 10 time steps (0.5 ps). This yielded 200,000 structures analyzed per mutant. In each structure, we measured the distance between Rh and the center of the phenol ring of Y112.

We plotted probability density functions of the Rh-Y112 distances from our QM/DMD simulations and noticed that there was correlation between TON and probability at short Rh-Y112 distances (Figure 2-6). This agreed with the double exponential dependence of the rate on the distance from the small model scan. However, the most active of the selected mutants did not show correlation. This prompted us to look for additional important interactions.

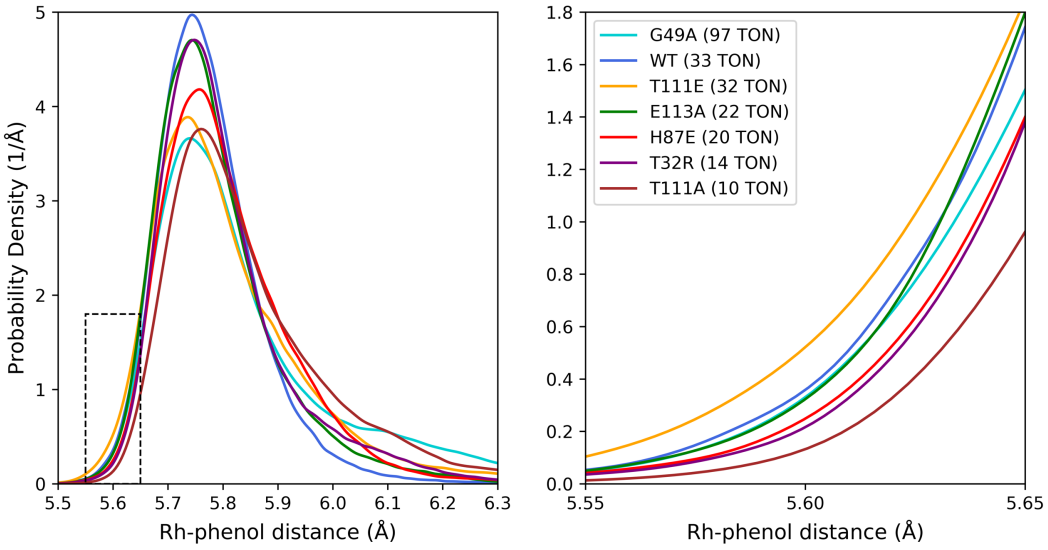


Figure 2-6. Probability density functions for the distance between Rh and the center of the phenol ring of Y112 for selected mutants.

We noticed in some of the QM/DMD simulations that the S119 residue could donate a hydrogen bond to Y112 (Figure 2-7). According to our working hypothesis, this would decrease electron donation to the Rh and decrease the favorable effect of Y112 and E124, increasing the barrier and decreasing activity. To test this, we added methanol to our small model, positioning it to donate a hydrogen bond to 4-methylphenol, and recalculated the migratory insertion barrier (Figure 2-8). The barrier increased by ~3 kcal/mol, supporting our hypothesis. This suggested that the most active mutants should have the highest probability of having a short Rh-Y112 distance while also not having a S119-Y112 hydrogen bond.

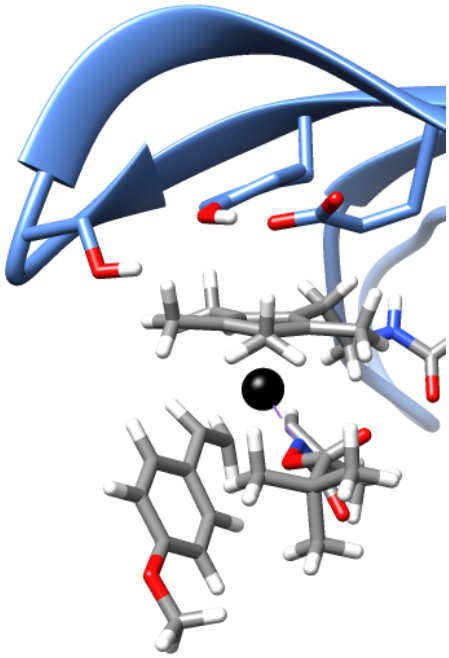


Figure 2-7. Snapshot from T111E mutant showing potential for S119-Y112 hydrogen bond.

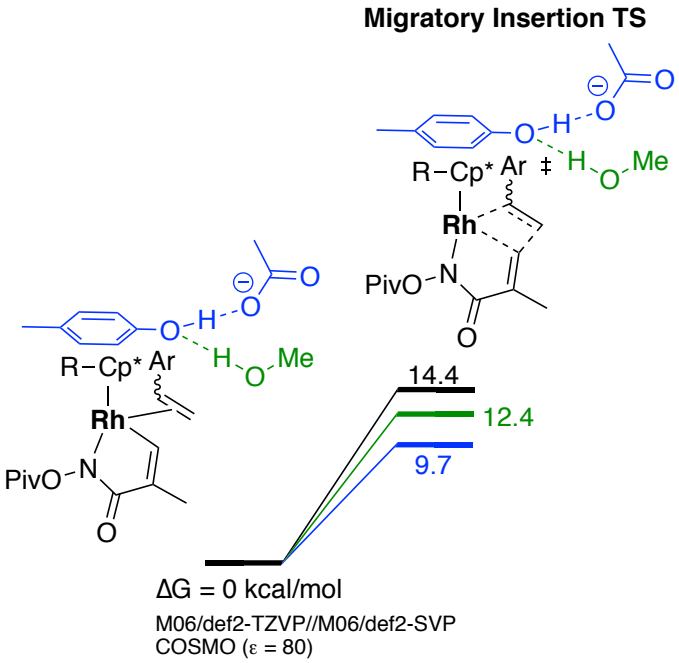


Figure 2-8. Effect of S119 sidechain (methanol) on migratory insertion barrier.

We therefore reanalyzed the QM/DMD simulations, simultaneously measuring the S119-Y112 hydrogen bond distance and the Rh-Y112 distance for each structure. This allowed us to plot

two-dimensional probability density functions (Figure 2-9). We observed that the most active mutant, G49A, had the highest probability of having structures with Rh-Y112 distances less than 5.65 Å and Y112-S119 hydrogen bond distances between 3.5 Å and 6 Å. This agreed with our hypothesis if we considered hydrogen bonds to be broken beyond 3.5 Å.

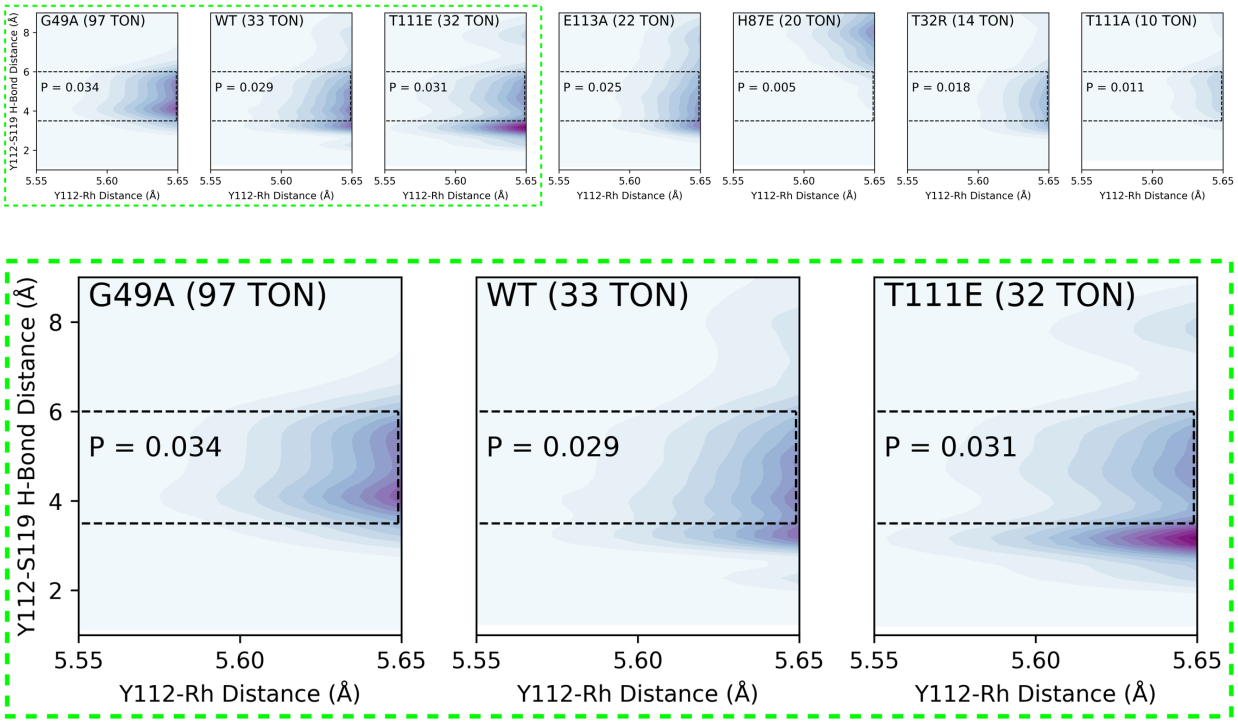


Figure 2-9. Two-dimensional probability density functions for selected mutants with plots for the three most active mutants enlarged. The boxed region is labeled with the integrated probability.

However, the upper limit on Y112-S119 distances for conformations to be included in the reactive space (6 Å) was empirical. Including probability for conformations with longer distances degraded correlation among the most active mutants but improved correlation for the mutant H87E. We postulated that water molecules, only modeled implicitly here, might replace S119 in hydrogen bonding to Y112 and similarly increase the barrier when the Y112-S119 distance was greater than 6 Å.

These results led us to hypothesize that we could improve activity by mutating S119, eliminating the hydrogen bond completely. In particular, the mutant T111E was most likely to have short Rh-Y112 distances, but these conformations usually also had short S119-Y112 distances. We therefore simulated the double mutants T111E S119X, where X was A, D, E, and G. The alanine and glycine mutants were chosen to remove the interaction, and the aspartate and glutamate mutants were chosen to replace the hydrogen bond donor with hydrogen bond acceptors. We hoped to see preservation of the high probability density at short Rh-Y112 distances.

While our simulations were underway, our collaborators synthesized T111E S119A and tested it. The TON was poor compared to most mutants, and our simulations suggested that this was because the S119A mutation decreased probability density at short Rh-Y112 distances (Figure 2-10). We therefore suggested the S119G mutation instead, which showed a probability density closer to that of the single mutant. However, the T111E S119G mutant also had low TON. Since the serine to glycine mutation would likely increase solvent exposure of Y112, this result again suggested to us that water could replace S119 in forming the detrimental hydrogen bond.

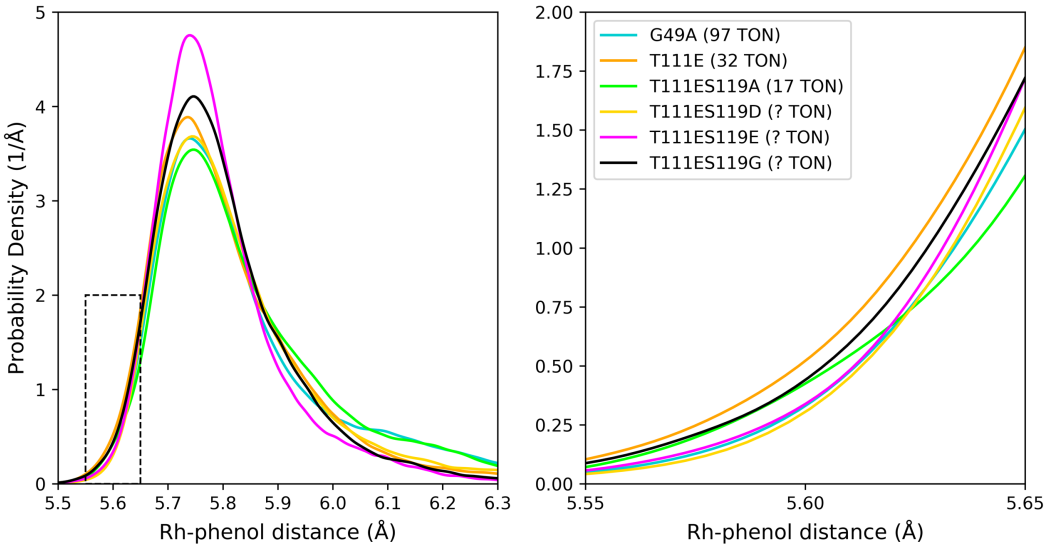


Figure 2-10. Probability density functions for T111E S119X double mutants compared to single mutant (T111E) and most active mutant (G49A).

In conclusion, we found that electron-donation resulting from a key interaction in the secondary coordination sphere improved catalysis. Mutations throughout the protein scaffold affected activity by altering this interaction. Protein dynamics were necessary to explain these effects. As a general principle of metalloenzyme design, we suggest that it is important to optimize the dynamics of the secondary coordination sphere.

References:

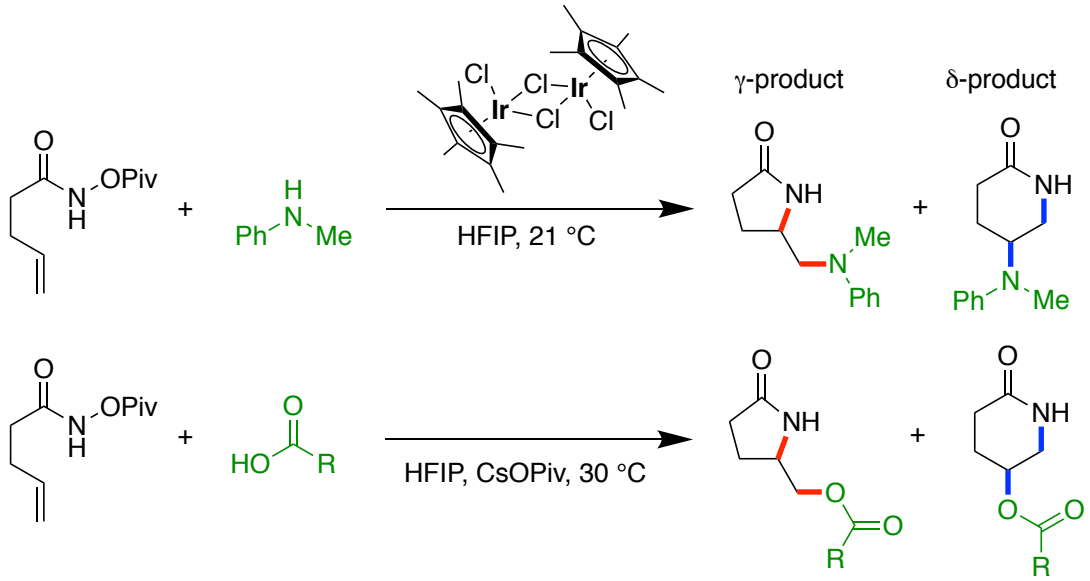
1. Dydio, P.; Key, H. M.; Nazarenko, A.; Rha, J. Y.-E.; Seyedkazemi, V.; Clark, D. S.; Hartwig, J. F. An Artificial Metalloenzyme with the Kinetics of Native Enzymes. *Science* **2016,** *354,* 102–106.

2. Hassan, I. S.; Ta, A. N.; Danneman, M. W.; Semakul, N.; Burns, M.; Basch, C. H.; Dippon, V. N.; McNaughton, B. R.; Rovis, T. Asymmetric δ-Lactam Synthesis with a Monomeric Streptavidin Artificial Metalloenzyme. *J. Am. Chem. Soc.* **2019,** *141,* 4815–4819.
3. Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
4. Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
5. Sierka, M.; Hogekamp, A.; Ahlrichs, R. Fast Evaluation of the Coulomb Potential for Electron Densities Using Multipole Accelerated Resolution of Identity Approximation. *J. Chem. Phys.* **2003**, *118*, 9136–9148.
6. Vahtras, O.; Almlöf, J.; Feyereisen, M. W. Integral Approximations for LCAO-SCF Calculations. *Chem. Phys. Lett.* **1993**, *213*, 514–518.
7. Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. Auxiliary Basis Sets to Approximate Coulomb Potentials. *Chem. Phys. Lett.* **1995**, *240*, 283–290.
8. Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. Auxiliary Basis Sets for Main Row Atoms and Transition Metals and their Use to Approximate Coulomb Potentials. *Theor. Chem. Acc.* **1997**, *97*, 119–124.
9. Ahlrichs, R. Efficient Evaluation of Three-Center Two-Electron Integrals over Gaussian Functions. *Phys. Chem. Chem. Phys.* **2004**, *6*, 5119–5121.

10. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
11. Klamt, A.; Schüürmann, G. COSMO: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and its Gradient. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799–805.
12. TURBOMOLE V6.6 2014, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from http://www.turbomole.com.
13. Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. Electronic Structure Calculations on Workstation Computers: The Program System TURBOMOLE. *Chem. Phys. Lett.* **1989**, *162*, 165–169.
14. von Arnim, M.; Ahlrichs, R. Performance of Parallel TURBOMOLE for Density Functional Calculations. *J. Comput. Chem.* **1998**, *19*, 1746–1757.
15. Häser, M.; Ahlrichs, R. Improvements on the Direct SCF Method. *J. Comput. Chem.* **1989**, *10*, 104–111.
16. Treutler, O.; Ahlrichs, R. Efficient Molecular Numerical Integration Schemes. *J. Chem. Phys.* **1995**, *102*, 346–354.
17. von Arnim, M.; Ahlrichs, R. Geometry Optimization in generalized natural internal coordinates. *J. Chem. Phys.* **1999**, *111*, 9183–9190.

CHAPTER 3: IR-SAV METALLOENZYMES

With the Rovis group, we explored an iridium metalloenzyme also built from the biotin-streptavidin system to see whether the interaction identified in the previous chapter would again be operative. The unincorporated iridium piano-stool complex catalyzed the formation of five- and six-membered lactams from unactivated alkenes (tethered to hydroxamate esters) and various nucleophiles (Scheme 3-1).[1,2] Initial experimental results showed mild improvement in turnover number by incorporation into streptavidin.[3]



Scheme 3-1. Five- and six-membered lactam formation catalyzed by iridium piano-stool complex.

We proceeded in the same way as with the Rh catalyst, using the computational methods listed in the previous chapter to calculate the mechanism with a small model and identify the rate-determining step. We examined the reaction with aniline as nucleophile as a prototypical reaction. Mechanistic steps considered included nitrene formation, cycloaddition, aziridine formation, and nucleophilic attack of a coordinated alkene. Additionally, we considered pathways leading to formation of byproduct.

Initially, we assumed a neutral working catalyst with a chloride ligand. We calculated a free-energy barrier of 13.8 kcal/mol for concerted metalation-deprotonation which would form an Ir–N bond and HCl (Figure 3-1). From this intermediate, we examined several possibilities. We found it most favorable for the reaction to proceed by forming the Ir(V)-nitrenoid intermediate proposed in previous studies[4] with a total barrier of 26.0 kcal/mol. Alternatively, nucleophilic attack of the alkene by aniline to form six- or seven-membered metallacycles had barriers of 31.2 and 33.5 kcal/mol, respectively. We ruled out this pathway because subsequent total barriers for Ir(V)-nitrenoid formation rose to 41.8 and 49.2 kcal/mol.

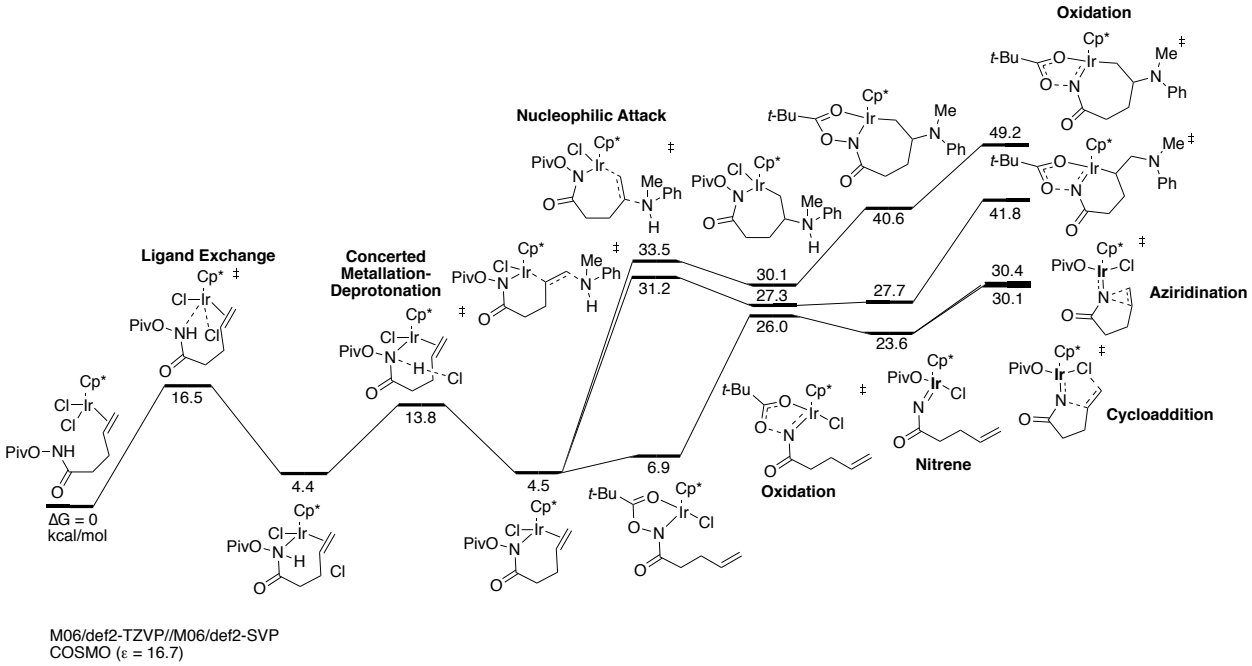
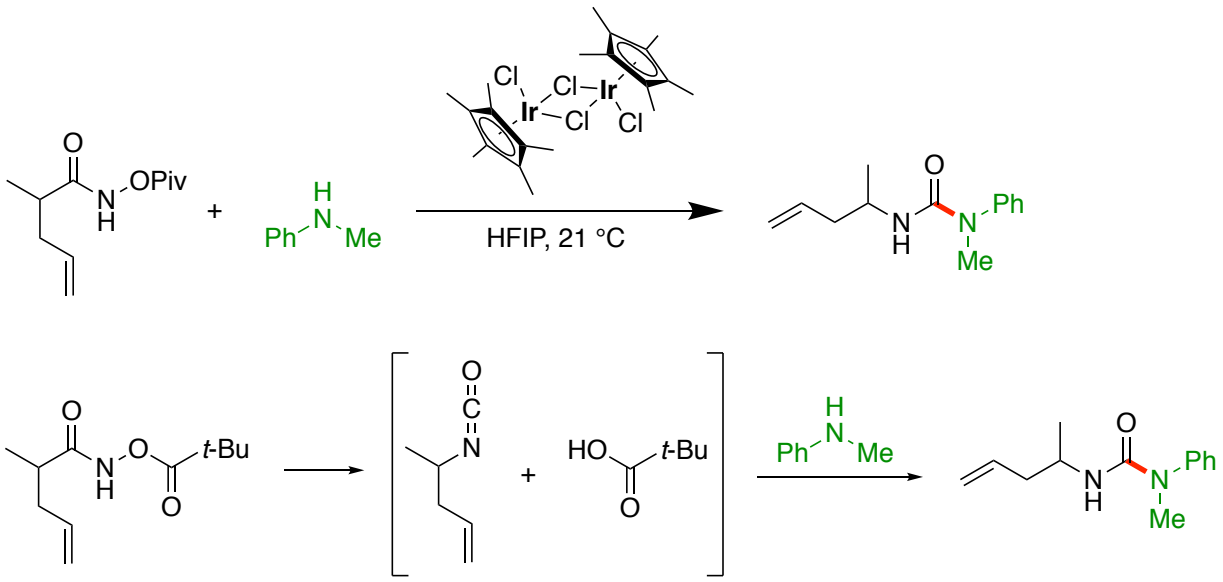


Figure 3-1. Initial mechanism with neutral catalyst.

From the nitrene, we calculated aziridination and cycloaddition pathways. We explored several cycloaddition possibilities, each of which included C–N bond formation between alkene and nitrene. The second carbon of the alkene could form a bond with iridium, pivalate, or chloride. We found C–Cl bond formation to be lowest with a barrier of 30.1 kcal/mol. For aziridination, the barrier was 30.4 kcal/mol.

During their study of α-substituted hydroxamate substrates, Conway and Rovis observed significant formation of urea byproduct, which they hypothesized resulted from Lossen rearrangement (Scheme 3-2).[1] However, they did not observe urea formation without α-substitution. We calculated a concerted Lossen rearrangement transition state at 30.2 kcal/mol and a stepwise Lossen rearrangement at 28.9 kcal/mol. This transformation was calculated to be highly exergonic, suggesting it was irreversible. As the substrate in our calculation did not have a substituent at the α-position, this did not agree with experiment. We looked for a concerted path to the aziridine, bypassing nitrene formation, but reaction path optimizations collapsed to stepwise paths. Beginning at this point in our study, we refined geometry optimizations with the larger def2-TZVP basis set, but this did not change the relative ordering of product vs byproduct transition states for this pathway (Figure 3-2).



Scheme 3-2. Urea byproduct formation by Lossen rearrangement with α-substituted substrates and proposed Lossen rearrangement pathway.

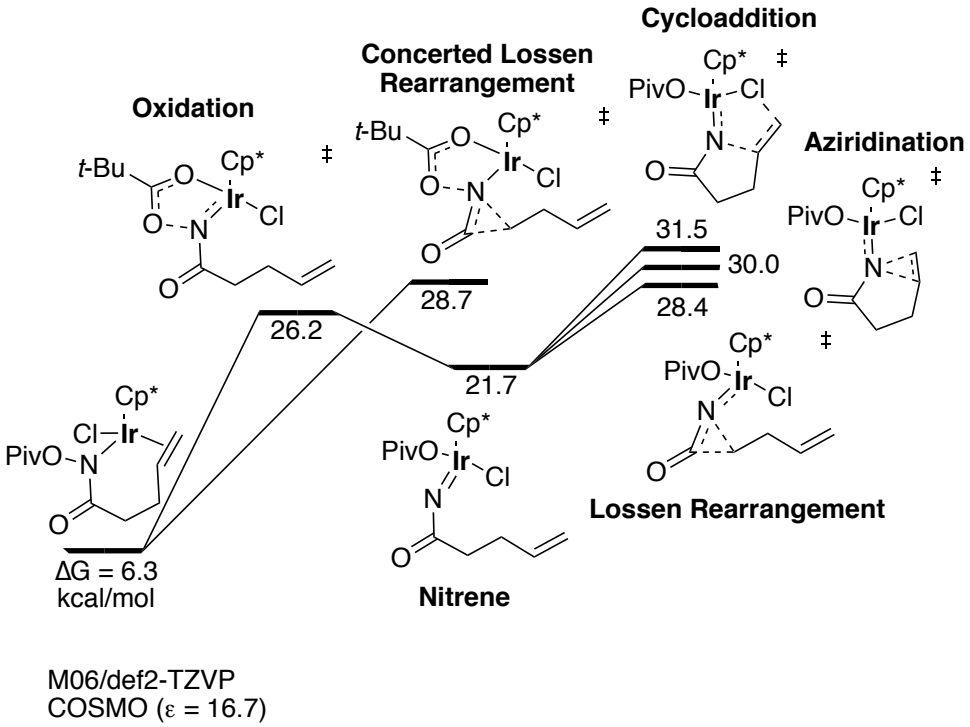


Figure 3-2. Product-formation pathways compared to Lossen-rearrangement pathways.

Our collaborators hypothesized that the chloride ligands might all dissociate, so we recomputed the initial steps with a cationic working catalyst. Using the larger def2-TZVP basis set for geometry optimization, we calculated that aziridine formation was favored with a barrier of 20.7 kcal/mol (Figure 3-3). The lowest barrier to Lossen rearrangement was for the concerted pathway at 21.4 kcal/mol, and the barrier to cycloaddition was 22.4 kcal/mol. Since this ordering of transition states agreed with experiment, we concluded that the working catalyst was cationic.

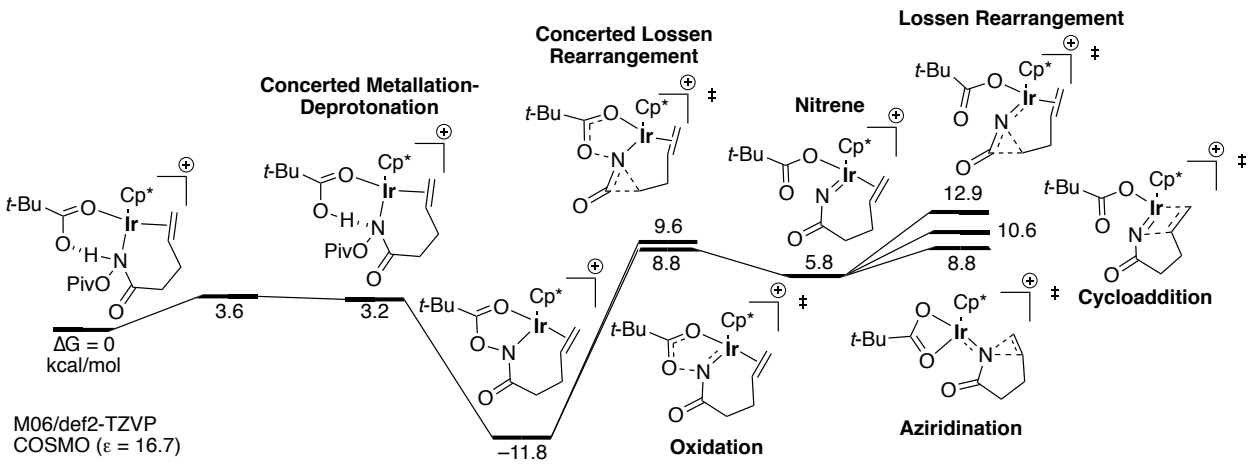


Figure 3-3. Cationic product-formation pathways and Lossen-rearrangement pathways.

Aziridination formed a bicyclo[3.1.0] aziridine which could be opened by nucleophilic attack, forming either a γ- or a δ-lactam with barriers of 18.9 kcal/mol or 19.8 kcal/mol, respectively (Figure 3-4). These calculated barriers agreed with experiment where formation of γ-lactam was favored. Proton transfer with a barrier of 11.6 kcal/mol led to product release.

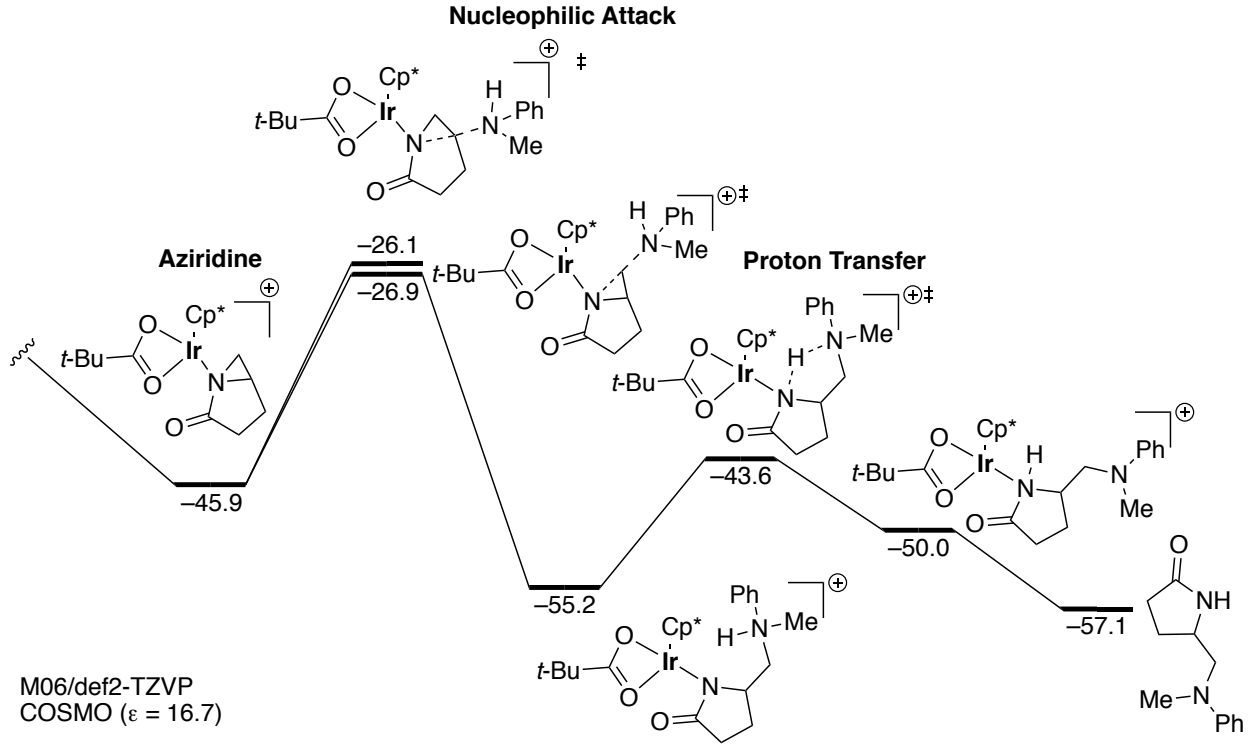


Figure 3-4. Cationic mechanism final steps.

According to the final calculated mechanism, the rate-determining step would be nitrene formation (20.6 kcal/mol) or aziridination (20.7 kcal/mol). However, the original small-catalyst experiments and these calculations were performed in hexafluoroisopropanol (dielectric of 16.7), whereas the enzyme operated in water (dielectric of 80), so we recalculated these two barriers with water as solvent. This brought the barrier for nitrene formation to 21.4 kcal/mol and the barrier for aziridination to 23.7 kcal/mol.

To explore the effect of streptavidin, we added sidechain analogs for Y112 and E124 and recomputed these barriers (Figure 3-5) as done in the previous chapter. The barriers dropped by ~2 kcal/mol to 19.4 and 21.1 kcal/mol, respectively. Addition of the S119 sidechain had almost no effect on the nitrene-formation barrier (19.5 kcal/mol) but almost completely reversed the favorable Y112 + E124 effect on the aziridination barrier (23.4 kcal/mol). This agreed with experiment that incorporation of the small catalyst into streptavidin would lead to mild improvement in activity.

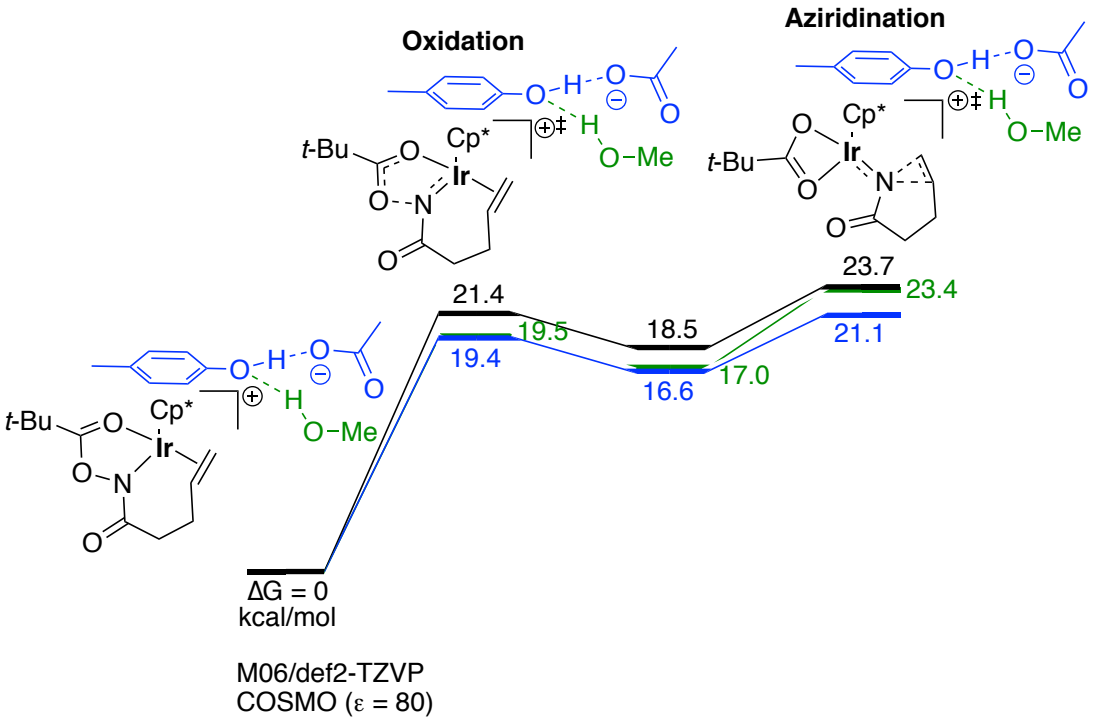


Figure 3-5. Effect of Y112 and E124 sidechains on rate-determining barriers.

We concluded that the same Y112 + E124 interaction was operative in this metalloenzyme, and it was similarly affected by S119. However, we also concluded that the original organometallic catalyst was less susceptible to enhancement from this secondary coordination sphere interaction.

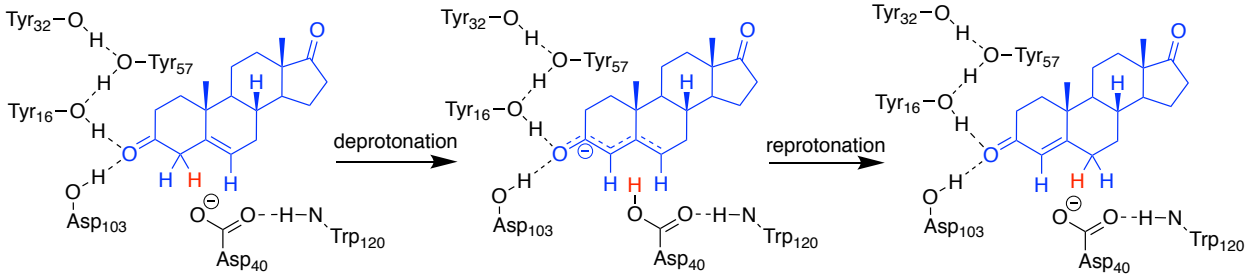
References:

1. Conway, J. H., Jr.; Rovis, T. Regiodivergent Iridium(III)-Catalyzed Diamination of Alkenyl Amides with Secondary Amines: Complementary Access to γ- or δ-Lactams. *J. Am. Chem. Soc.* **2018**, *140*, 135–138.
2. Lei, H.; Conway, J. H., Jr.; Cook, C. C.; Rovis T. Ligand Controlled Ir-Catalyzed Regiodivergent Oxyamination of Unactivated Alkenes. *J. Am. Chem. Soc.* **2019**, *141*, 11864–11869.
3. Lei, H.; Rovis, T. Unpublished results.

4. Park, Y.; Heo, J.; Baik, M.-H.; Chang, S. Why is the Ir(III)-Mediated Amido Transfer Much Faster Than the Rh(III)-Mediated Reaction? A Combined Experimental and Computational Study. *J. Am. Chem. Soc.* **2016**, *138*, 14020–14029.

CHAPTER 4: KETOSTEROID ISOMERASE

Ketosteroid isomerase (KSI) catalyzes proton shuttling with high efficiency (Scheme 4-1). Boxer and co-workers calculated a large electric field (140 MV/cm) at the carbonyl of an inhibitor bound to KSI using simulations and vibrational Stark spectroscopy.[1] Markland and co-workers showed that over 90 percent of this field came from the closest two residues: Y16 and D103.[2] To see what effect the small electric field from the secondary coordination sphere and the rest of the protein might have, we calculated the mechanism with and without external electric fields.



Scheme 4-1. Isomerization reaction catalyzed by KSI.

We used the active site shown in Scheme 4-1 as our small model. This included the sidechains of Y32, Y57, Y16, D103, D40, and W120. D40 is the catalytic base, and the other five residues are part of the hydrogen bond network containing the substrate and catalytic base. The positions of the β-carbons were taken from structure 1OH0[3] of the PDB and frozen in place, along with capping hydrogens replacing the α-carbons.

We used the M06-2X functional[4] with the def2-TZVP and def2-SVP basis sets[5] for electronic energies and geometries, respectively, and we used the multipole-accelerated resolution of the identity approximation.[6–10] To model the dielectric effect of residues adjacent to the active site, we used the Conductor-like Screening Model[11] with a dielectric constant of 4. All quantum-mechanical calculations were performed in TURBOMOLE.[12–17] We applied uniform

external electric fields with magnitudes of 10 MV/cm, on the order of the field predicted to result from residues other than Y16 and D103, along the carbonyl bond in both directions.

The enthalpy surface has been experimentally determined for the D38E mutant.[18] Although this mutant was ~200 times less active than wild type, it was chosen for experimental analysis because it eliminated the effect of diffusion on the kinetics. We compared our results to its kinetics since the activity difference would only correspond to an energy difference of ~3 kcal/mol. The calculated electronic energy surface of our small model with no applied field closely matched this experimental enthalpy surface (Figure 4-1).

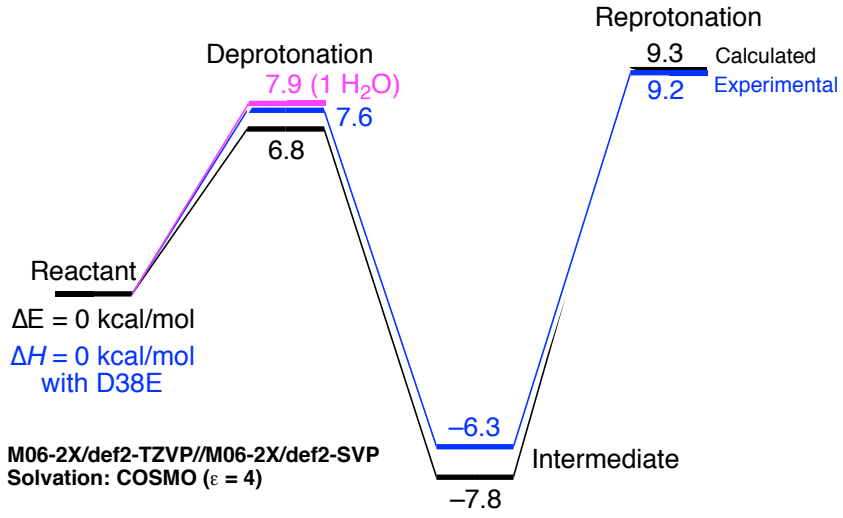


Figure 4-1. Calculated and experimental energy surfaces for KSI mechanism.

A previous study suggested that a water molecule was present in the active site during the deprotonation step and left the active site in the intermediate state.[19] We therefore also calculated the deprotonation barrier with an explicit water molecule and found a small shift in the barrier. This shift slightly improved agreement with experiment, so we continued to calculate the deprotonation barrier with and without an explicit water molecule.

Upon application of external electric fields of differing signs along the carbonyl, the energy surface shifted as predicted when considering reaction difference dipoles[1] (Figure 4-2).

The dipole of the carbonyl is greatest in the intermediate state, so electric fields of differing signs should lower one barrier while raising the other. This implies the field is optimal for catalysis when free energy barriers for each step are equal. Indeed, the free energy barriers of each step for the D38E mutant were experimentally determined to be equal.[18]

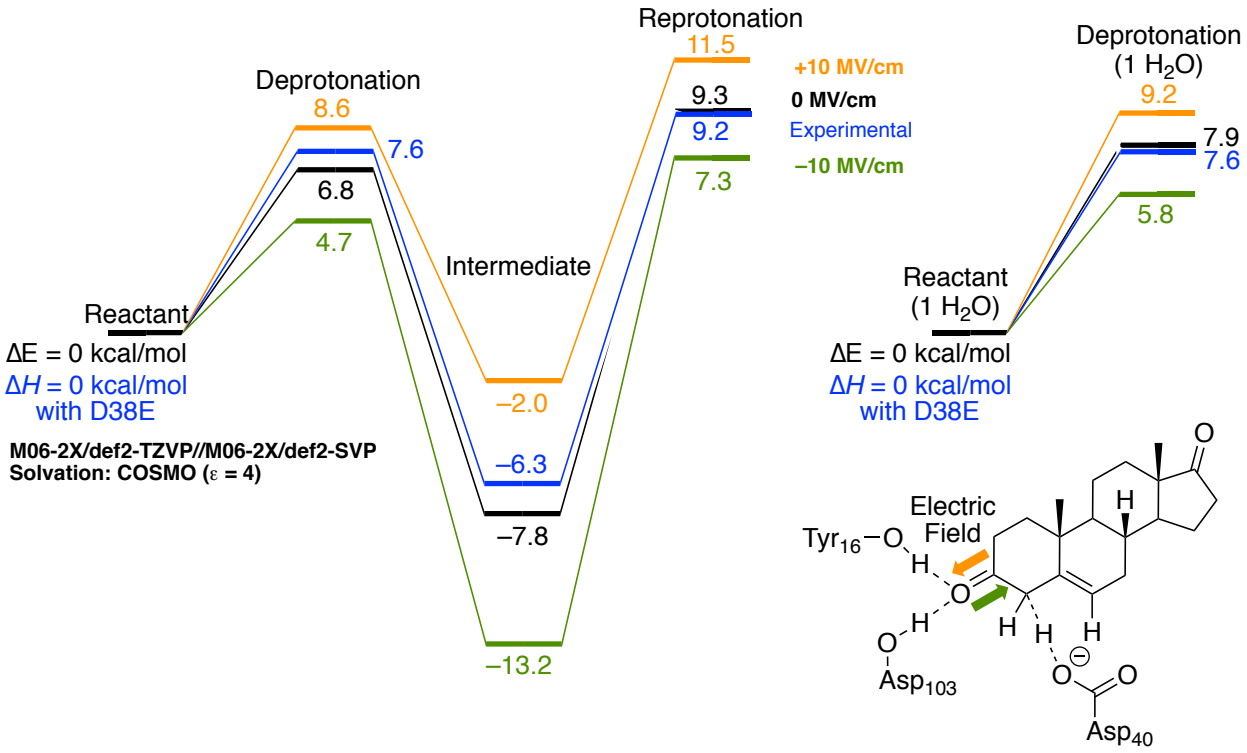


Figure 4-2. Calculated energy surfaces with electric fields applied along the carbonyl.

We therefore concluded that the optimal external field should correspond to the calculated surface most closely matching the experimental surface. Since the calculated surface with no external field already closely matched experiment, we concluded that the immediate active site was already well-positioned by the protein scaffold to create the optimal catalytic field. This suggests the main purpose of the scaffold, at least in KSI, is to maintain the positioning of the hydrogen bond network of the active site.

References:

1. Fried, S. D.; Bagchi, S.; Boxer, S. G. Extreme Electric Fields Power Catalysis in the Active Site of Ketosteroid Isomerase. *Science* **2014**, *346*, 1510–1514.
2. Wang, L.; Fried, S. D.; Markland, T. E. Proton Network Flexibility Enables Robustness and Large Electric Fields in the Ketosteroid Isomerase Active Site. *J. Phys. Chem. B* **2017**, *121*, 9807–9815.
3. Kim, S. W.; Cha, S.-S.; Cho, H.-S.; Kim, J.-S.; Ha, N.-C.; Cho, M.-J.; Joo, S.; Kim, K. K.; Kwan, Y. C.; Oh, B.-H. High-Resolution Crystal Structures of Δ^5-3-Ketosteroid Isomerase with and without a Reaction Intermediate Analogue. *Biochemistry* **1997**, *36*, 14030–14036.
4. Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
5. Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
6. Sierka, M.; Hogekamp, A.; Ahlrichs, R. Fast Evaluation of the Coulomb Potential for Electron Densities Using Multipole Accelerated Resolution of Identity Approximation. *J. Chem. Phys.* **2003**, *118*, 9136–9148.
7. Vahtras, O.; Almlöf, J.; Feyereisen, M. W. Integral Approximations for LCAO-SCF Calculations. *Chem. Phys. Lett.* **1993**, *213*, 514–518.

8. Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. Auxiliary Basis Sets to Approximate Coulomb Potentials. *Chem. Phys. Lett.* **1995**, *240*, 283–290.
9. Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. Auxiliary Basis Sets for Main Row Atoms and Transition Metals and their Use to Approximate Coulomb Potentials. *Theor. Chem. Acc.* **1997**, *97*, 119–124.
10. Ahlrichs, R. Efficient Evaluation of Three-Center Two-Electron Integrals over Gaussian Functions. *Phys. Chem. Chem. Phys.* **2004**, *6*, 5119–5121.
11. Klamt, A.; Schüürmann, G. COSMO: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and its Gradient. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799–805.
12. TURBOMOLE V6.6 2014, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from http://www.turbomole.com.
13. Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. Electronic Structure Calculations on Workstation Computers: The Program System TURBOMOLE. *Chem. Phys. Lett.* **1989**, *162*, 165–169.
14. von Arnim, M.; Ahlrichs, R. Performance of Parallel TURBOMOLE for Density Functional Calculations. *J. Comput. Chem.* **1998**, *19*, 1746–1757.
15. Häser, M.; Ahlrichs, R. Improvements on the Direct SCF Method. *J. Comput. Chem.* **1989**, *10*, 104–111.
16. Treutler, O.; Ahlrichs, R. Efficient Molecular Numerical Integration Schemes. *J. Chem. Phys.* **1995**, *102*, 346–354.

17. von Arnim, M.; Ahlrichs, R. Geometry Optimization in generalized natural internal coordinates. *J. Chem. Phys.* **1999**, *111*, 9183–9190.
18. Houck, W. J.; Pollack, R. M. Temperature Effects on the Catalytic Activity of the D38E Mutant of 3-Oxo-Δ^5-Steroid Isomerase: Favorable Enthalpies and Entropies of Activation Relative to the Nonenzymatic Reaction Catalyzed by Acetate Ion. *J. Am. Chem. Soc.* **2004**, *126*, 16416–16425.
19. Feierberg, I.; Åqvist, J. The Catalytic Power of Ketosteroid Isomerase Investigated by Computer Simulation. *Biochemistry* **2002**, *41*, 15728–15735.

CHAPTER 5: REDOX-ACTIVE BIMETALLIC COMPLEXES

The Yang group synthesized bimetallic complexes containing both redox-active transition metals and redox-inactive metals (Figure 5-1). They used Ni and Fe as transition metals and Na, K, and Ba as secondary metals. They observed that the redox potential of the transition metal shifted in response to the presence and charge of the secondary metal compared to the same transition metal in the monometallic complex shown in Figure 5-1. We collaborated with the Yang group to determine how this secondary coordination sphere effect occurred.

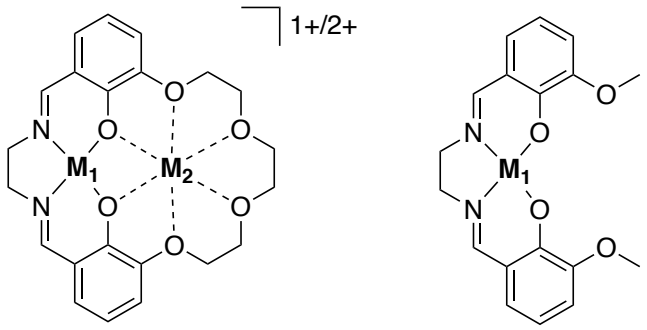


Figure 5-1. Basic structure of bimetallic and monometallic complexes. M_1 is a redox-active transition metal and M_2 is a redox-inactive metal.

We used the Gaussian 09 software[1] with the ωB97X-D functional[2] to calculate spectra and orbital energies. We used the def2-SVP basis set for geometry optimizations and the def2-TZVP basis set for final electronic energies and spectra.[3] We modeled solvent implicitly using the SMD method.[4]

Spectra calculated using time-dependent density functional theory[5] (TDDFT) showed good agreement with experimental spectra for the Ni(II) complexes listed in Table 5-1. These spectra showed minimal changes in absorption wavelengths as the charge of the secondary cation increased from 0 (nonexistent) to 1 (Na^+) to 2 (Ba^{2+}).

Table 5-1. Comparison of TDDFT and experimental absorption spectra for Ni(II) complexes in dimethylformamide.

	Ni(II)		Ni(II)-Na^+		Ni(II)-Ba^{2+}	
Assignment	**λ_{calc} (nm)**	**λ_{exp} (nm)**	**λ_{calc} (nm)**	**λ_{exp} (nm)**	**λ_{calc} (nm)**	**λ_{exp} (nm)**
d-d	563	547	565	552	607	555
MLCT	390	415	389	406	386	404
π-π*	345	349	333	345	327	344

In contrast, the orbitals shifted to lower energies as the charge of the secondary metal increased (Figure 5-2). We concluded that the effect of the secondary cation was mostly electrostatic, changing the redox potential by changing the orbital energies without altering the ligand field structure.

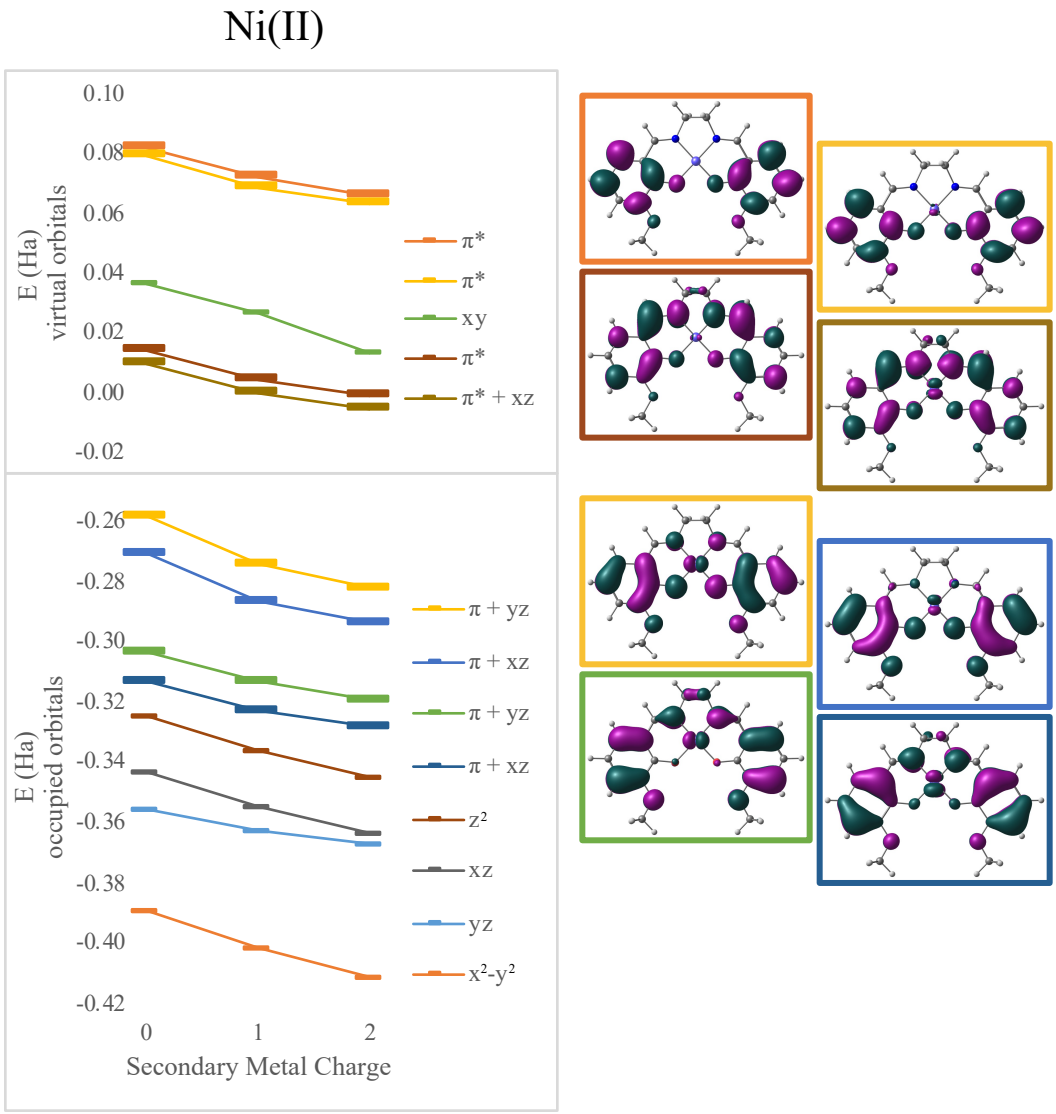


Figure 5-2. Orbital energies for Ni(II) complexes. Orbitals involved in MLCT are bolded on the left and plotted on the right (monometallic complex).

We additionally calculated the electronic structures of Fe(II) and Fe(III) monometallic and bimetallic complexes.[6] The crystal structures of these complexes revealed square pyramidal geometries with the extra ligand being solvent (acetonitrile) in the case of Fe(II) and chloride in the case of Fe(III). Importantly, the identity of the ligand changed the effect of the secondary cation

on the Fe(III/II) redox potential; the chloride complexes showed much smaller redox potential shifts than the corresponding acetonitrile complexes.

Our calculated structures showed geometric distortion of the ligand field when both secondary cation and chloride were present. Electrostatic attraction pulled the chloride and the secondary metal closer together. This changed the d-orbital composition of the highest occupied molecular orbital (HOMO) as shown in Figure 5-3. The movement of chloride and restructuring of the orbitals likely altered the secondary cation effect compared to the acetonitrile complexes. Nevertheless, orbital energies still generally decreased as the charge of the secondary cation increased.

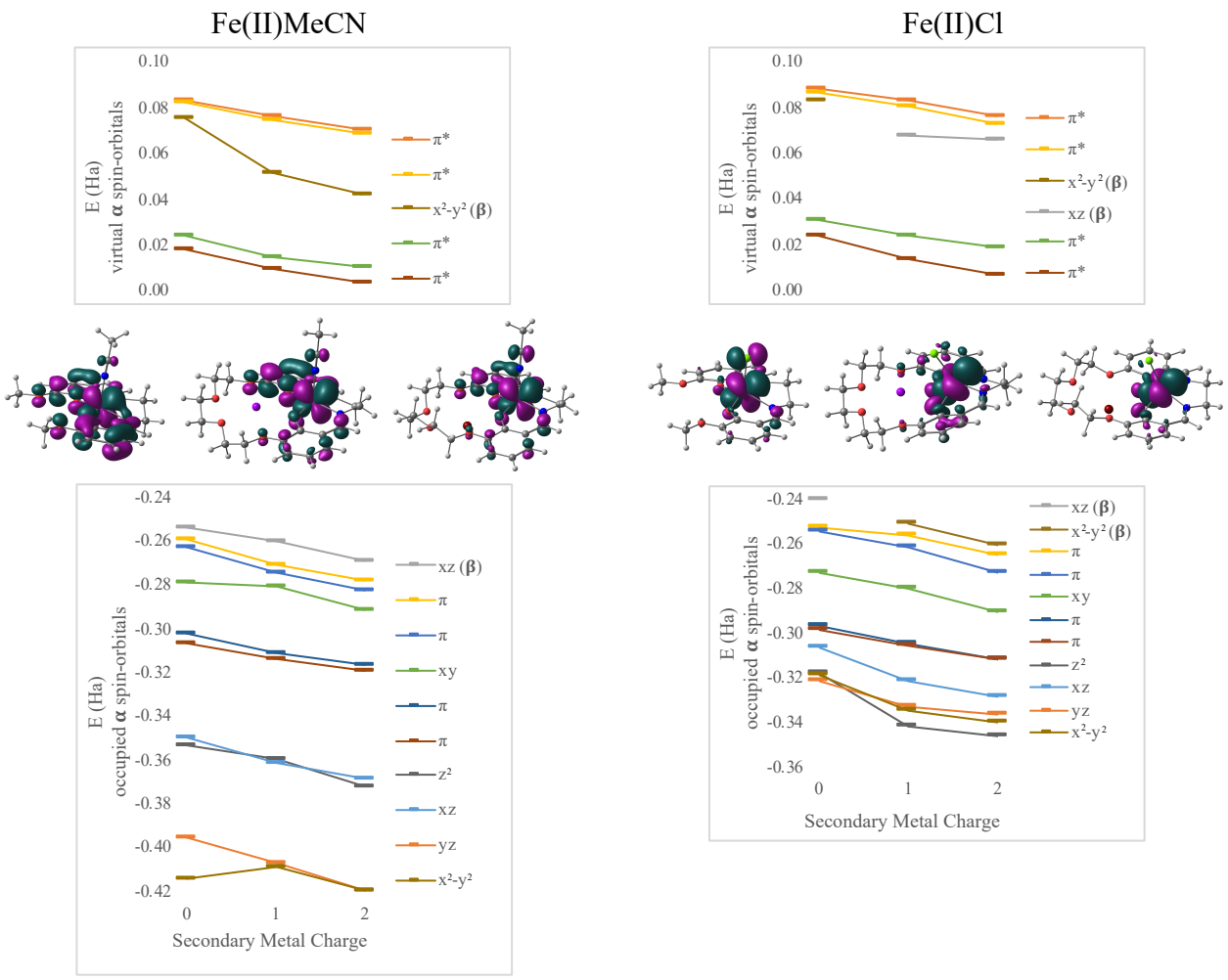


Figure 5-3. Orbital energies, geometries, and HOMOs for Fe(II) acetonitrile and chloride complexes.

In conclusion, we determined that cations in the secondary coordination sphere of transition metals can exert a purely electrostatic effect, decreasing orbital energies proportionally to the charge of the secondary cation. However, we also saw that if the ligand field was not rigid, these cations could distort it in addition to any electrostatic effects.

References

1. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, *Revision D.01*, Gaussian, Inc., Wallingford CT, 2013.
2. Chai, J.-D.; Head-Gordon, M. Long-Range Corrected Hybrid Density Functionals with Damped Atom–Atom Dispersion Corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.
3. Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
4. Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113*, 6378–6396.

5. Dreuw, A.; Head-Gordon, M. Single-Reference ab initio Methods for the Calculation of Excited States of Large Molecules. *Chem. Rev.* **2005**, *105*, 4009–4037.
6. Chantarojsiri, T.; Ziller, J. W.; Yang, J. Y. Incorporation of Redox-Inactive Cations Promotes Iron Catalyzed Aerobic C–H Oxidation at Mild Potentials. *Chem. Sci.* **2018**, *9*, 2567–2574.