Development and Application of Computational Tools to Study Single-Electron Transfer Initiated Reactions in Solution and Enzymes

by

Kevin C. Skinner

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemistry)
In The University of Michigan
2022

Doctoral Committee:

Associate Professor Alison R. H. Narayan, Co-Chair
Associate Professor Paul M. Zimmerman, Co-Chair
Professor Nicolai Lehnert
Professor Janet L. Smith
Research Associate Professor Troy Wymore

Kevin C. Skinner

kevincs@umich.edu

ORCID iD: 0000-0001-8591-1169

Dedication

This thesis is dedicated to my family for their unwavering support throughout this journey.

Acknowledgements

I would first like to acknowledge Paul and Alison for their guidance and support throughout my tenure here at the University of Michigan. Without their guidance none of this would have been possible. I truly appreciate the ways in which they challenged me to become a better scientist, and a better person. I also want to thank Dr. Troy Wymore for his mentorship and friendship during my years at Michigan. I would also like to thank Dr. Adam Smith at The University of Akron for his mentorship during my years at Akron as an undergrad researcher. His advice was invaluable for putting my on the trajectory to grad school and the early stages of what it meant to be and think like a scientist. Finally, I would like to thank Dr. Jim Beil of Lorain County Community College for putting my on the initial path towards grad school. His advice is what put me on the path to seeking research opportunities at Akron, and also for his communication and encouragement throughout the years.

I am also thankful for both the Zimmerman and Narayan Groups. I have had countless conversations and collaborations that have enabled me to become a better scientist than I would have been without them. It has been an honor to be a part of both groups, and to have participated in all the joyous events throughout the years. Special thanks to Matt and Alan for their friendship and all the memories that we made!

I also want to thank my collaborators, for without them the science could not have been studied. Specifically, Dr. Troy Wymore, Dr. Amanda Deweyer, Dr. Cody Aldaz, Dr.

Josh Kammeraad, Dr. Tyler Doyon, Khoi Dang, Jonathan Perkins, Evan Romero, and Di Yang. Thank you for your contributions and countless discussions that helped make this science possible. I would especially like to thank Dr. Troy Wymore for instilling in me my confidence as a computational chemist and scientist, and his mentorship through these years.

I would now like to acknowledge my family. To my parents Tim and Patti, I am grateful for your support and encouragement for my journey to this thesis. To my sons Owen and Aiden, thank you for being a source of happiness and pride. I am truly happy that you are in my life and helped to bring balance. Finally, I would like to acknowledge my wife Mindea. Thank you so much for your support throughout the 13+ years we have been together. You complete me and have kept me grounded throughout my academic journey. I could not imagine getting through this without you.

Table of Contents

Lit of Tables

List of Figures

List of Abbreviations

^{1}H – proton

^{13}C – carbon-13

α-KG – α-ketoglutarate

CDFT – constrained density functional theory

CDFT/MM – constrained density functional theory/molecular mechanics

COSY – correlated spectroscopy

DE-GSM – double-ended growing string method

DFT – density functional theory

DMF – dimethylformamide

ET – electron transfer

GSM – growing string method

H-atom – hydrogen atom

HMBC – heteronuclear multiple bond correlation

HOMO – highest occupied molecular orbital

HPLC – high-performance liquid chromatography

HSQC – heteronuclear single quantum coherence

MD – molecular dynamics

MECP – minimum energy crossing point

MM – molecular mechanics

MS – mass spectrometry

NCI – non-covalent interaction

NHI – non-heme iron

NMR – nuclear magnetic resonance

OED – organic electron donor

PES – potential energy surface

PCET – proton-coupled electron transfer

QM – quantum mechanics

QM/MM – quantum mechanics/molecular mechanics

SET – single-electron transfer

SE-GSM – single-eneded growing string method

SOMO – singly occupied molecular orbital

TDAE – tetrakis(diethylamino)ethylene

TS – transition state

TST – transition state theory

TTF – tetrathiafulvene

T7H – Thymine-7-hydroxylase

Abstract

Computational modeling of (bio)chemical reactions has been around for decades using small gas-phase simulations on the quantum mechanical (QM) and large-scale simulations involving explicit solvent with combined quantum mechanical/molecular mechanical (QM/MM) techniques. Enzymes are powerful tools developed by nature to construct complex molecules that are bioactive. The reactions of enzymes can be difficult to study experimentally due to short lived intermediates, such as radicals. Computational modeling of reaction pathways of an enzyme, such as one- and two-electron mechanisms, and how the enzyme affects the reaction via active site interactions can help to guide enzyme redesign to alter the outcome or even chemistry of the enzyme. The thesis contained herein seeks to add to the computational tools available to study enzymatic and other condensed phase systems to gain insight into their mechanisms and the underlying chemical principles governing reactivity.

The first goal of this thesis was to develop a new combined constrained density functional theory/molecular mechanics (CDFT/MM) tool to study reactions involving single-electron transfer (SET) in condensed phases. Specifically, the new CDFT/MM tool was tested on two well defined organic electron donors (OEDs). Tetrakis(dimethylamino)ethylene (TDAE) and tetrathiafulvene (TTF) are OEDs that have been extensively studied by the Murphy group since the 1990s and 2000s. Reactions involving TDAE and TTF begin with a SET and result in the formation of indoles and

benzofuranyl alcohols, respectively. The CDFT/MM tool was able to describe known SET transfer processes, to identify known and unknown intermediates, and to replicate well defined chemistry.

With the CDFT/MM tool being validated, it was then applied during the study of mechanisms for the enzyme TropC. TropC is an α-ketoglutarate (αKG) dependent non-heme iron (NHI) enzyme that catalyzes a ring expansion to produce a seven-membered aromatic ring known as a tropolone, which is a scaffold observed in bioactive natural products. Initially, three reaction pathways were studied using a QM model to gain insight into potential reactions. Results from the QM model revealed that the enzyme preferred a one-electron mechanism for production of tropolones. Specifically, radical- and SET-based mechanisms were determined to be most plausible for TropC. Using the insights gained from the QM model, QM/MM modeling of TropC and the three pathways was conducted. The new CDFT/MM tool was used to study the viability of a SET within the active site of TropC. Following SET from a radical intermediate, a barrierless ring expansion was observed with a large reorganization energy. It was subsequently found that the SET pathway was potentially viable, and further study of this possibility is merited. Based on the data herein the mechanism of TropC is suggested to proceed through a one-electron radical-based ring expansion.

Chapter I

Introduction

α-Ketoglutarate-dependent (α-KG) non-heme iron (NHI) enzymes catalyze a breadth of reactions, including hydroxylation, halogenation, desaturation, epoxidation, epimerization, and skeletal rearrangements.[1-5] To achieve these transformations, this class of enzymes pairs iron with the cosubstrate α-KG to activate molecular oxygen and form a highly reactive iron(IV)-oxo species that is responsible for an initial hydrogen atom (H-atom) abstraction. The radical species generated is the divergent point from which a plethora of transformations can occur, including rebound hydroxylation, rebound halogenation, second H-atom abstraction, H-atom abstraction from tyrosine, skeletal rearrangement, and single-electron transfer (SET).[1-5] In order to understand and control subsequent transformations catalyzed by these enzymes, detailed reaction mechanisms are extremely useful. Although consensus mechanisms for many NHI transformations exist, the precise mechanism of skeletal rearrangements is not universal and can differ among NHI enzymes. In particular for these rearrangements, various mechanisms can diverge from the reactive intermediate, leading to considerable uncertainty about the true nature of the reaction.[1-5] Due to the difficulties in studying radical intermediates and processes, computations can be used to elucidate mechanisms by simulating reaction paths and provide an atomic level of detail.[6] This can be achieved by quantum mechanics (QM) or a combined quantum mechanics/molecular mechanics (QM/MM) study and has

been used to elucidate the mechanisms of radical-based skeletal rearrangements.[7-11]**1.1 Mechanism of Iron(IV)–Oxo Formation and Reactivity**

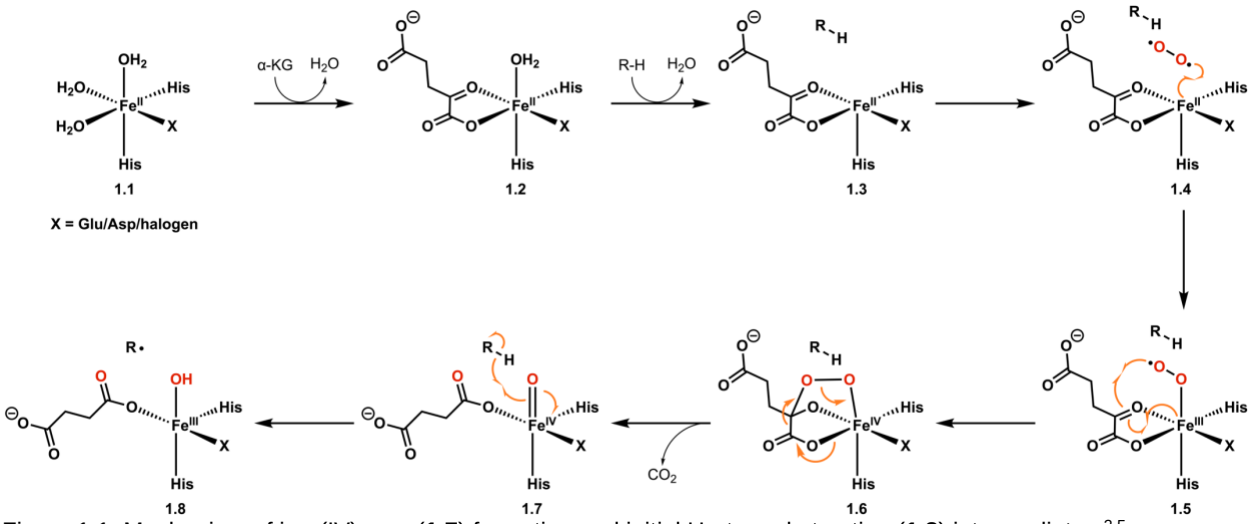

Figure 1.1. Mechanism of iron(IV)–oxo (**1.7**) formation and initial H-atom abstraction (**1.8**) intermediates.[2,5]

The active site resting state of α-KG NHI enzymes is an iron(II) center coordinated by a 2-Histidine-1-X facial triad of residues, where X is glutamate or aspartate (but a halogen for halogenation reactions), and three water molecules (Figure 1.1, **1.1**).[2,5] Displacement of two water molecules occur upon binding of α-KG in a bidentate manner (**1.2**). The final water molecule is displaced once the substrate binds (**1.3**). After the substrate is bound, molecular oxygen coordinates to the iron (**1.4**). Coupling of molecular oxygen then commences the formation of the iron(III)-superoxo species **1.5**. Radical addition of the terminal oxygen into coordinating carbonyl carbon forms the bridged iron(IV)-peroxy species **1.6**. Subsequent oxidative decarboxylation forms succinate, CO_2, and the highly reactive iron(IV)-oxo species **1.7**. From **1.7**, the aforementioned H-atom abstraction from

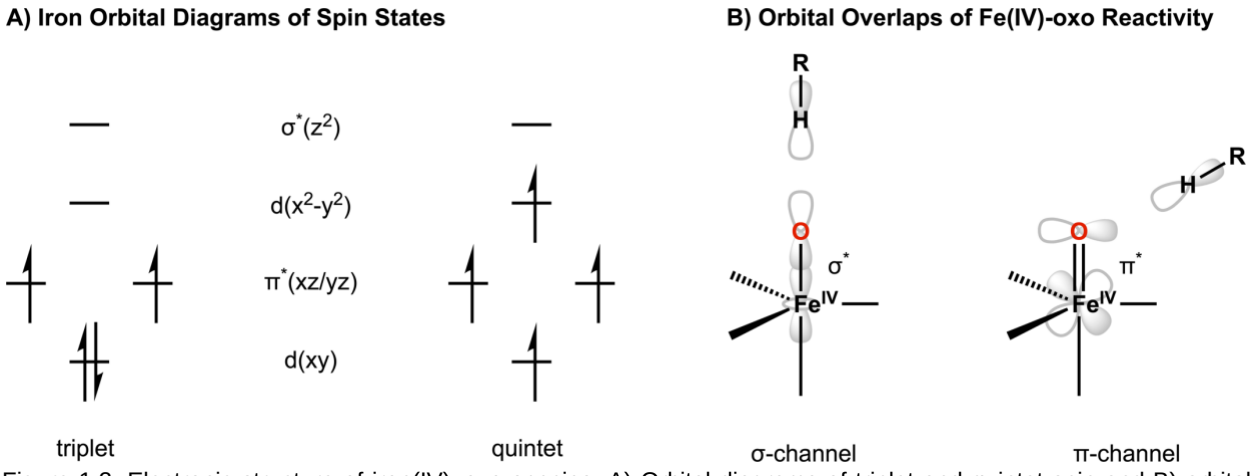

Figure 1.2. Electronic structure of iron(IV)–oxo species. A) Orbital diagrams of triplet and quintet spin and B) orbital overlaps for the σ- and π-channels.[14,15]

the substrate is performed to form a substrate radical intermediate and an iron(III)-hydroxy species (**1.8**). Iron(IV)–oxo species can have a ground triplet state outside of the enzyme, such as biomimetic complexes, but have a quintet ground state within the enzyme active site (Figure 1.2A).[12,13] The iron(IV)-oxo species will abstract an H-atom in either a σ- or π-channel, wherein the C–H–Fe angle of approach is the discriminating factor (Figure 1.2).[14,15] The angle of approach for the σ-channel is 180° and 120° for the π-channel (Figure 1.2B).[14,15] The reason for these channels is orbital overlap between the substrate and either the σ and σ* or π and π* orbitals of the iron(IV)–oxo.[14,15]

1.2 Skeletal Rearrangements of α-KG NHI Enzymes

The formation of the substrate radical and iron(III)-hydroxy species is the divergent point for α-KG NHI enzymes. From the radical substrate intermediate, skeletal rearrangements can commence in two major fashions: direct and indirect. Indirect processes occur through rebound hydroxylation[16-18] or epoxidation[19,20] followed by intramolecular nucleophilic attack or condensation.[21-24] In the biosynthetic pathway of bicyclomycin (**1.9**)

two α-KG NHI enzymes are responsible formation of the oxygen-bridged intermediate **1.10** (Figure 1.3).[19] From the piperazinedione core intermediate **1.11**, BcmE performs a hydroxylation to produce the primary alcohol intermediate **1.12**. Following further hydroxylations by the α-KG NHI enzymes BcmC and BcmG, the trihydoxylated intermediate **1.13** is produced. Next, the enzyme BcmB performs an initial desaturation to produce the exocyclic alkene intermediate **1.14** and then an epoxidation to form the intermediate **1.15**. After the epoxide is installed, the internal carbon undergoes spontaneous intramolecular attack by the primary hydroxyl installed by BcmE to form a new eight-membered, oxygen-bridged ring. Researchers propose the desaturation and epoxidation activity BcmB due to mass spectrometry (MS), 1D and 2D nuclear magnetic resonance (NMR) spectroscopy, and shunt product formation. Initially, researchers

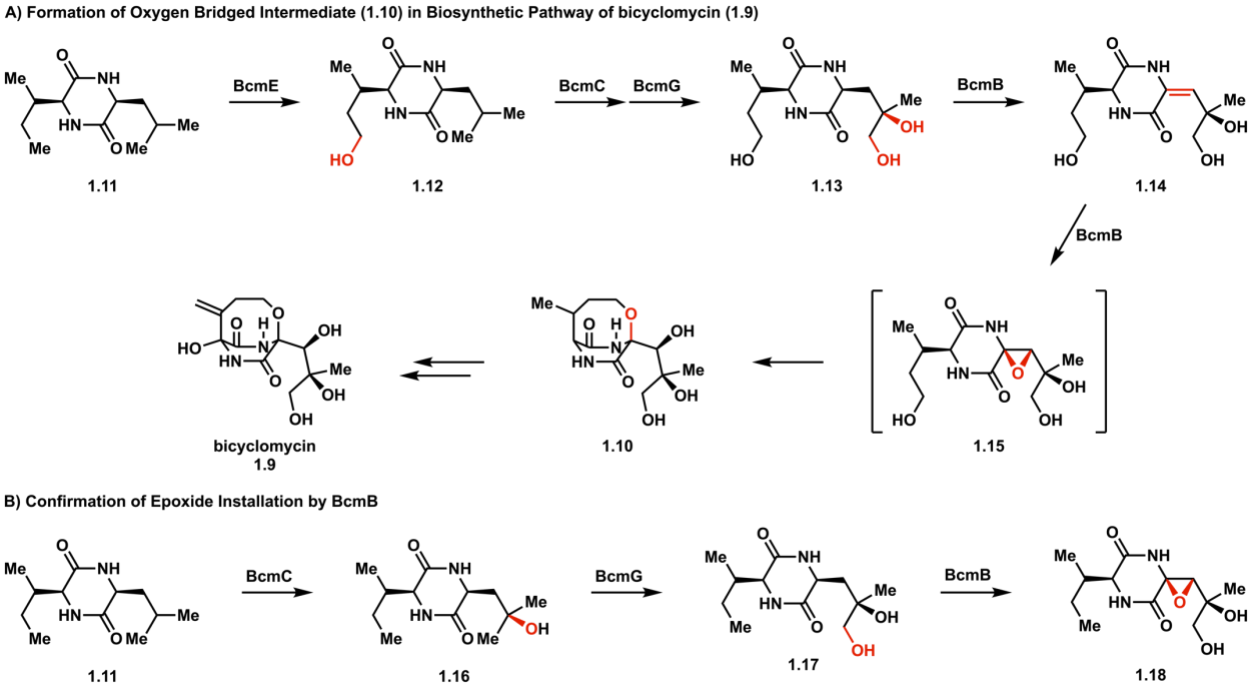

Figure 1.3. Biosynthetic pathway of bicyclomycin (**1.9**).[19] A) Biosynthesis of oxygen bridged intermediate (**1.10**). B) Confirmation of epoxide activity of BcmB. Red reflects changes to structure. Figure displays indirect ring formation via nucleophilic attack by hydroxyl group into the epoxide.

observed that both BcmC and BcmE accepted **1.11** as a substrate. Due to higher activity BmcC was chosen as the preferred enzyme. Following the first hydroxylation by BmC to form the tertiary alcohol intermediate **1.16**, BcmG produces the dehydroxylated intermediate **1.17**. BmcB then performs the epoxidation to form **1.18**, which was confirmed by MS and NMR experiments. However, the researchers were unable to react **1.18** further with any enzymes. Therefore, BcmE was declared the first enzyme to act on **1.11** and the researchers were then able to reconstitute the biosynthetic pathway of bicyclomycin (**1.9**). Throughout the correct pathway, researchers reasoned that BcmB exhibits multifunctionality (desaturation and epoxidation) due to the detection of **1.14** and **1.15** by MS.

In contrast to indirect catalysis, direct catalysis occurs through rearrangement of the radical intermediate, such as ring expansion[25,26], contraction[27,28], or ring formation[10,29,30] without participation of the iron(III)–hydroxy species.[31] These can then be terminated by a second H-atom abstraction by the iron(III)-hydroxy species to close the cycle. An example of this is the ring formation catalyzed by AndA in the biosynthetic pathway of anditomin (Figure 1.4, **1.19**).[10,32] AndA is a multifunctional α-KG NHI enzyme that is responsible for desaturation of preandiloid B (**1.20**) to form preandiloid C (**1.21**), and then the formation of the bridged ring system of andiconin (Figure 1.4, **1.22**). High-performance liquid chromatography (HPLC) assays confirmed desaturation of **1.20** by AndA, as it was detected as a minor product. The major product of the assay was **1.22**. To confirm **1.21** as an intermediate another assay in which **1.19** was exposed to AndA and showed conversion to **1.22** confirmed by MS. To understand the skeletal

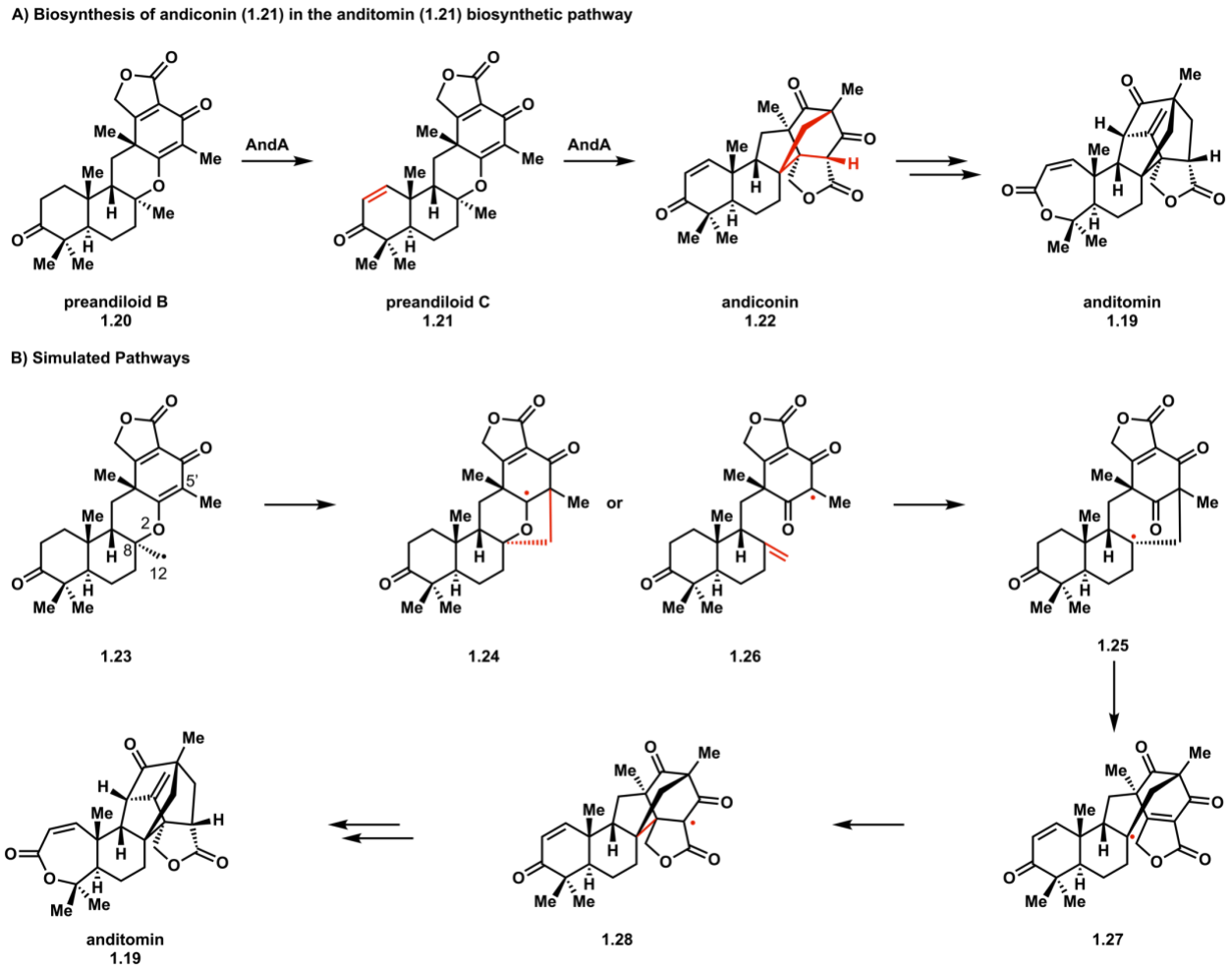

Figure 1.4. Biosynthetic pathway of anditomin (**1.19**).[10,32] A) Biosynthesis of bridged ring system by AndA. B) Structural changes from computational study. Red reflects changes to structure.

rearrangement catalyzed by AndA, researchers turned to quantum mechancial (QM) simulations to investigate two pathways that converge. Both pathways were simulated beginning from the primary radical intermediate **1.23**. Initial C_{12}–$C_{5'}$ bond formation forms the bridged intermediate **1.24**, which would then undergo C_8–O_2 bond cleavage to produce the tertiary radical intermediate **1.25**. The second pathway would proceed in the reverse order, with C_8–O_2 bond cleavage to form the exo-alkene intermediate **1.26**, and then C_{12}–$C_{5'}$ bond formation to form **1.25**. Simulations of first pathway produced an initial high barrier for C_{12}–$C_{5'}$ bond formation, followed by a reasonable barrier for C_8–O_2 bond

cleavage. However, barriers for the reverse pathway were both reasonable. This indicated that the preferred pathway would be initial C_8–O_2 bond cleavage and then C_{12}–$C_{5'}$ bond formation. Following formation of **1.25**, a series of conformational changes were simulated to have small barriers to produce the flipped tertiary radical intermediate **1.27**. From **1.27** C_8–$C_{2'}$ bond formation would result in the formation of the tertiary radical intermediate **1.28**. Currently the mechanism for reduction of this intermediate to **1.22** is unknown, which is the case for other α-KG NHI enzymes.[32]

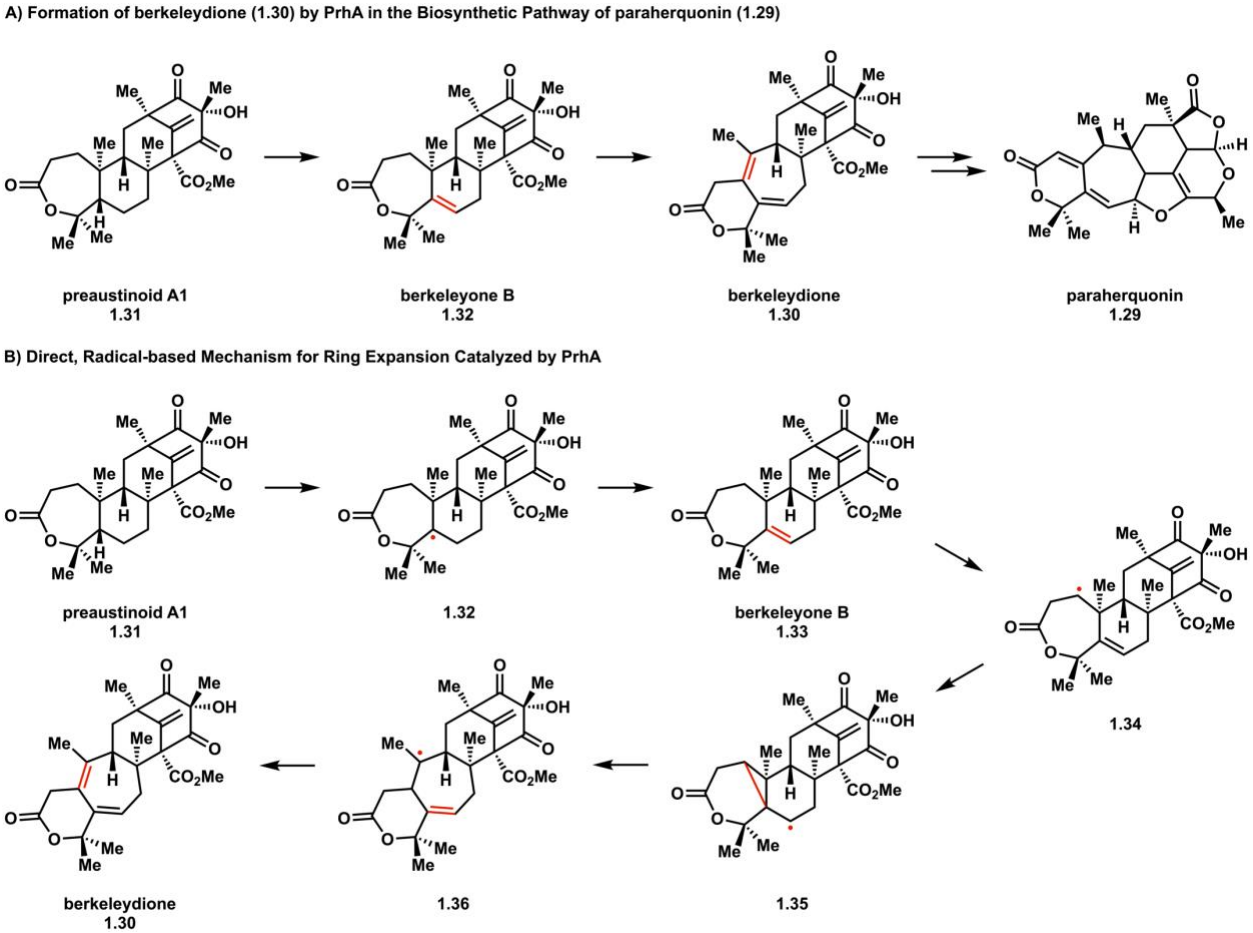

Figure 1.5. Biosynthetic pathway of paraherquonin (**1.29**).[25] A) Reactions catalyzed by PrhA. B) Radical-based mechanism of desaturation and ring expansion for formation of berekelydione (**1.30**). Red reflects changes to structure.

Ring expansion is another example of direct catalysis catalyzed by α-KG NHI enzymes. In the biosynthetic pathway of paraherquonin (**1.29**), berkeleydione **1.30** is formed by the α-KG NHI enzyme PrhA whose function was first described by Abe and coworkers in 2016 (Figure 5).[25] Beginning from the preaustinoid A1 intermediate **1.31**, PrhA first performs an H-atom abstraction to form the radical intermediate **1.32**, and then a second terminating H-atom abstraction to form berkeleyone B (**1.33**). The enzyme then catalyzes the ring expansion of **1.32** to form **1.30**. PrhA was identified as the enzyme responsible due to production of **1.30** only being detected in the presence of the enzyme. Researchers proposed formation of **1.33** first due to detection of the intermediate by HPLC and characterization by 1D and 2D NMR experiments. From **1.33**, an iron(IV)-oxo species can perform an H-atom abstraction to produce the radical intermediate **1.34**. Radical addition into the double bond results in the cyclopropyl intermediate **1.35**. Opening of the ring forms a more stable tertiary radical (**1.36**) and less sterically strained system. Termination of the radical by the iron(III)-OH species via a second H-atom abstraction results in **1.30**. Based on their experimental observations, Abe and coworkers proposed PrhA to be a multifunctional α-KG NHI enzyme that performs a desaturation and then ring expansion via direct catalysis.

Direct catalysis can also result in ring contraction, which is a reaction catalyzed by the α-KG NHI enzyme AusE (Figure 1.6).[27] In the biosynthetic pathway of austinol (**1.37**), AusE performs an initial desaturation of preaustinoid A1 (**1.31**) to produce preaustinoid A2 (**1.38**) and then ring contraction to form preaustinoid A3 (**1.39**). Researchers determined AusE as the enzyme responsible for formation of **1.39** based on HPLC in vitro assays in which **1.38** is only converted to **1.39** in the presence of AusE. Researchers

initially proposed an indirect mechanism (rebound hydroxylation, ring contraction via elimination of installed hydroxyl group) for AusE,[25] however, they more recently proposed a direct mechanism based on the discovery of AusE' in a separate gene cluster[33]. This mechanism involves an initial desaturation followed by H-atom abstraction, ring contraction, and then a terminating second H-atom abstraction. Interestingly, the sequence identity between AusE' and PrhA is 92%, which is higher than the sequence identity between AusE and PrhA (78%). AusE' and PrhA also share a common intermediate: preaustinoid A1 (**1.31**). Analysis of the crystal structures for AusE' and PrhA revealed that there are only three differences among residues in the active site and that the substrate binds in a different pose than what was initially considered.[34] Researchers proposed that the residues interact with the **1.31** in a manner that influences the conformation within the active site and subsequentially which H-atom is initially abstracted. Mutagenesis studies were performed, wherein the residues were exchanged to those in the opposite enzyme. This resulted in interconversion of the reactions performed by the variant enzymes, where variant PrhA produced preaustinoid A3 (**1.39**) and variant AusE produced berkeleyone B (**1.32**).

Although the mechanism for AusE was updated to reflect a direct mechanism based on sequence similarity to PrhA and substrate binding, experiments were not performed to confirm the feasibility of a radical-based pathway. However, QM/MM studies of the PhrA double mutant PrhA-V150L/A232S have suggested that AusE catalyzes the ring contraction of preaustinoid A2 (**1.38**) to preaustinoid A3 (**1.39**) via a direct mechanism.[11] Using a model based on the crystal structure, the researchers simulated barriers for the first (**1.40**) and second H-atom abstractions for desaturation to be feasible

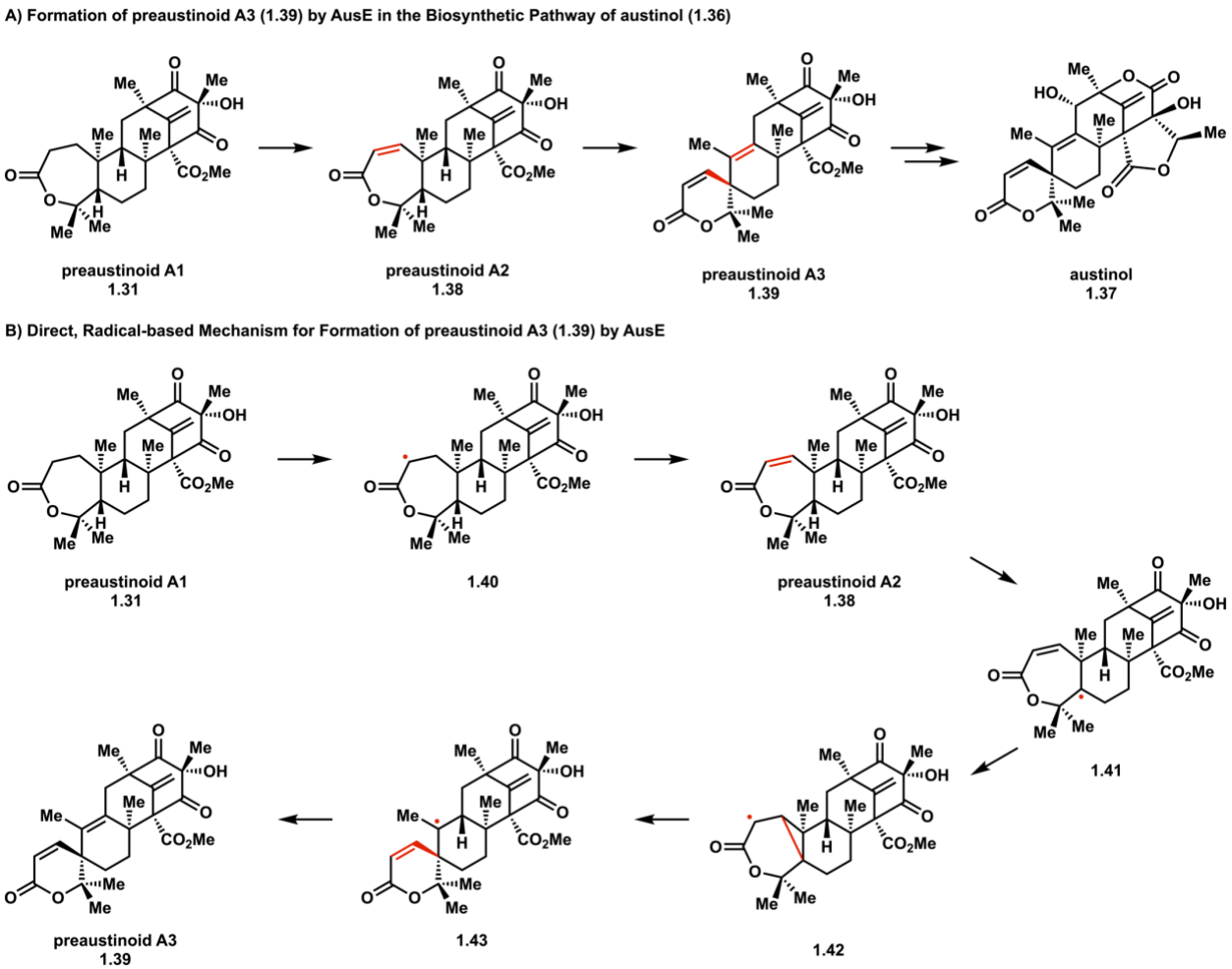

Figure 1.6. Biosynthetic pathway or preaustinoid A3 (**1.39**).[27] A) Reactions catalyzed by AusE. B) Radical-based mechanism of desaturation and ring contraction for formation of preaustinoid A3 (**1.39**). Red reflects changes to structure.

for formation of **1.38** (Figure 1.6B). Next, simulation of the ring contraction step to form **1.39** was performed, which involves: an initial H-atom abstraction (**1.41**), cyclopropane ring formation (**1.42**), ring contraction (**1.43**), and a terminating second H-atom abstraction. Barriers for these processes were all reasonable. Based on these observations, it was concluded AusE can perform ring contraction via a direct mechanism.

1.3 SET Events in α-KG NHI Enzymes

Another form of direct catalysis involves the use of SET from the substrate to the iron(III)–OH species, which have been more recently postulated based on experiment[35-40] and simulation,[8,9,41,42] but the process is somewhat elusive due to the short lifetimes of radical intermediates and computational limitations.[43] Using QM/MM simulations the enzymes ScoE[8] and Aj_EasH[9] revealed SET processes within their mechanisms. ScoE is an α-KG NHI enzyme that catalyzes isonitrile formation in (*R*)-3-isocyanobutanoate (**1.44**) from the (*R*)-3-((carboxymethyl)-amino)-butanoate precursor (Figure 1.7, **1.45**). The next step simulated was the decarboxylation step to form the isonitrile group, which involved an initial H-atom abstraction. The barrier for initial H-atom abstraction was calculated to be feasible. During simulation of the subsequent C–C breakage for decarboxylation, researchers observed an atypical proton-coupled electron transfer (PCET), in which an electron is transferred to the iron(III)–OH species at the same time as a proton is transferred from a nearby arginine. This was confirmed by analysis of the frontier molecular orbitals and natural population analysis. The spin density on the iron increased from 2.76 e^- to 3.67 e^-, and the substrate decreased from 0.95 e^- to zero. At the same time a proton from a nearby arginine was observed to have transferred to the iron(III)–OH species. This atypical PCET resulted in the restoration of the iron(II)–OH_2 and formation of the product (**1.44**).

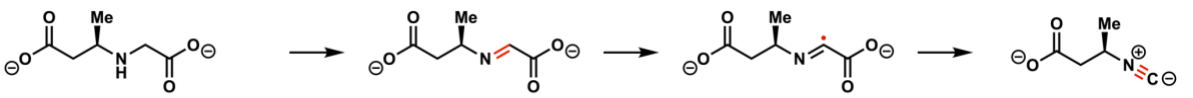

Figure 1.7. Mechanism for biosynthesis of (R)-isocyanobutanoate by ScoE.[8] Red reflects changes to structure

Simulations of the skeletal rearrangement catalyzed by the α-KG NHI enzyme Aj_EasH also revealed a SET event during biosynthesis of the iminium intermediate (**1.45**) from the enamine intermediate (**1.46**) in the cycloclavine biosynthetic pathway (Figure 1.8, **1.47**).[9] Three different mechanisms were originally proposed for Aj_EasH and included hydroxylation followed by elimination of water, hydride transfer, or radical-based skeletal rearrangement followed by heteroatom-assisted SET.[44] However, after crystal structure analysis the radical-based skeletal rearrangement pathway was favored due to the hydrophobicity of the active site.[45] With these observations in mind, Yan and Liu performed both QM and QM/MM simulations of this transformation based on the crystal structure.[9] The QM model (Model A) included the iron(IV)–oxo, 2-His-1-glu facial triad, succinate, and the substrate. Initial H-atom abstraction was simulated and produced a low barrier for the production of the tertiary radical intermediate **1.48**. Next, researchers simulated the formation of (+) and (-) forms of **1.45** via radical addition into the double bond, which would also produce the transient secondary radical intermediate **1.49**. Barriers for this process were inaccessible and the products were also overall thermodynamically uphill. However, there were large relative changes in the position of the substrate that would otherwise be avoided by restrictions of an enzyme active site.

To better understand the influence the enzyme environment has on the reaction, researchers used two different QM/MM models (Model B/C) based on different substrate

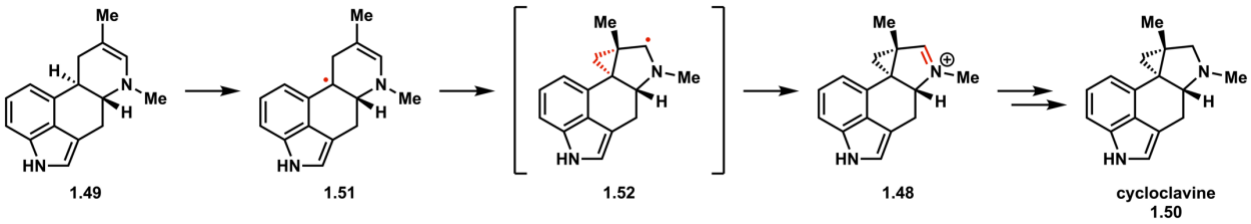

Figure 1.8. Mechanism for biosynthesis of the iminium intermediate (**1.48**) from the cycloclavine (**1.50**) biosynthetic pathway.[9] Red reflects changes to structure.

bindings. Simulations of H-atom abstraction produced reasonable barriers for both models. As was done with Model A, both the (+) and (–) forms of the iminium intermediate **1.45** were simulated for both Model B and Model C. For both models, formation of (+)-**1.45** was found to be kinetically inaccessible. Also simulated to have a high barrier was formation of (–)-**1.45** by model C. Only production of (–)-**1.45** was found to have a reasonable barrier. Both models suggest that the dominant product would be (–)-**1.45**. The barrier for (–)-**1.45** in Model B was much lower than any of the other barriers and was attributed to a heteroatom-assisted SET from the substrate to the iron. The spin decreased from 0.92 e^- to 0.02 e^- for the tertiary radical intermediate **1.47** and (–)-**1.45** respectively, and the spin on the iron(III)–OH increased from 3.00 e^- to 3.91 e^-. This was the only instance of SET observed between Models B/C and was reasoned as accessible due to the shorter distance between the substrate and iron, suggesting a distance dependence for SET.

From the results of the three models (A/B/C) researchers highlighted that the difference in models and their results are from how the substrate interacts with the surrounding environment. In Model A, the QM only model, there was a lack of active site to maintain substrate positioning, which led to results that conflicted with experimental findings. With the enzyme environment modeled by MM, the simulations produced results consistent with experimental observations. (–)-**1.45** is favored over (+)-**1.45** for both models and was rationalized due to steric clashing of the cyclopropyl ring with nearby active site residues. The barrier for (+)-**1.45** in Model B is much larger than (–)-**1.45**, but this is attributed to the shorter distance between the iron and enamine that allows for SET to occur. The three models highlighted the importance of including the enzyme

environment in simulations in order to be able to capture the influence of active site residues on substate binding, which allows for more concrete conclusions to be drawn about the mechanisms of the enzymes being simulated.

1.4 Computational Modeling of Reactions Within Enzymes

Simulations of α-KG NHI enzymes are performed in two major fashions: cluster models and solvated enzyme mode:ls.[6,46-51] Cluster models differ from the latter in that only the active site and residues identified as important are modeled using QM methods. Cluster models then partially fix atoms to simulate an enzyme environment, such as to avoid large rotational or translational movements. This type of modeling allows for potentially more residues to be treated with QM methods and provides a better representation of the active site without the full enzyme being present. In contrast, full

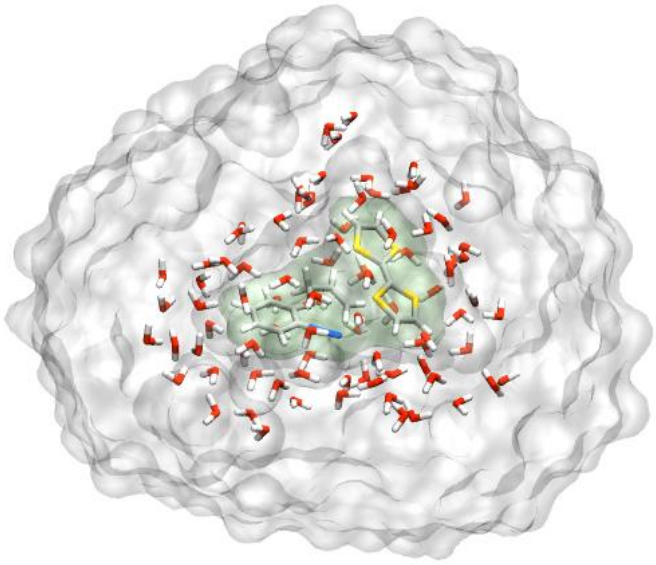

Figure 1.9. Simple scheme highlighting different regions of a solvated QM/MM model. QM region in green contains electronically interesting molecules, such as electron donors or acceptors. MM region (gray) includes all other molecules, such as solvent or counterions.

enzyme models rely on the use of a combined QM and MM approach. QM is defined for the substrate, metals if present, cofactors, and residues identified as important to catalysis, and MM for the remaining enzyme and solvent. QM/MM modeling provides a way to treat the electronically interesting region using accurate QM methods and the remaining environment using MM methods (Figure 1.9). This combination provides the accuracy of QM methods and the immense computational time savings available to MM methods.

When accounting for the energies of the QM and MM regions, there exist two methods: additive and subtractive.[6,52,53] Subtractive coupling first calculates the energy of the entire system at the MM level, then the QM energy of the QM region is added, and finally, the MM energy of the QM region is subtracted. For additive coupling, the QM region is embedded in the MM, and the regions are calculated individually and combined, and then the coupling between the two regions is calculated and added. The QM embedding is achieved in one of three ways: mechanical, electrostatic, or polarization, of which mechanical and electrostatic are most common for biological systems.[6,52,53] For mechanical embedding, the interactions between the two regions are done on the MM level and therefore mechanical embedding is most commonly implemented in the subtractive method.[53] For electrostatic embedding, the MM charges are part of the QM energy calculation. This results in the QM atoms being polarized by the MM atoms, but the MM atoms remain the same, and therefore electrostatic embedding is most commonly implemented in additive methods.[53] Building upon electrostatic embedding, polarization embedding allows the MM region to be polarized by the QM region as well.[53] Electrostatic embedding allows for the QM region to feel the effects of the MM region (solvent and

enzyme environment), therefore the work in this chapter will be done using electrostatic embedding within an additive method.

1.5 Density Functional Theory

For QM simulations of reaction pathways and molecular properties the most common practice has been to use density functional theory (DFT), which utilizes electron densities ($\boldsymbol{\rho}$) to obtain the molecular properties[54-56] DFT operates within the Born-Oppenheimer approximation, in which the nuclei are viewed as fixed compared to the much lighter electrons. The energy calculated ($E[\rho]$) within DFT can be described by eq 1, where it is the sum of kinetic ($E_T[\rho]$), nuclear-electron attraction ($E_{ne}[\rho]$), electron-electron repulsion ($E_{ee}[\rho]$), and electron-electron interaction energies ($E_{XC}[\rho]$). The only unknown term within eq 1 is $E_{XC}[\rho]$, which describes the exchange and correlation energies, and this must be treated approximately using a density functional. DFT functionals use different approximations, such as the local (spin) density approximation or generalized gradient approximation, to calculate the exchange and correlation energies. The exchange and correlation energies can be calculated using a combination of Hartree-Fock and DFT, and results in the formation of hybrid functionals. Hybrid functionals BP86[57,58] and B3LYP[59,60] were used throughout this thesis.

$$E[\rho] = E_T[\rho] + E_{ne}[\rho] + E_{ee}[\rho] + E_{XC}[\rho] \qquad \text{eq 1}$$

However, DFT simulations are restricted to a single electronic state, whereas simulations involving SET involves multiple electronic states. To treat SET-initiated

reactions, the technique called constrained DFT (CDFT) is potentially useful.[61-63] CDFT restricts the electron density distribution to ground or SET states, by applying adding an additional weighted functional to eq 1. This functional is minimized to satisfy user defined constraints. There are a range of constraints[64-71] that can be placed on atoms(s) within a system. However, for the systems studied within this thesis, only charge, spin, or a combination of the two are needed to describe the SET processes.

1.6 Reaction Pathway Modeling

Computational and theoretical tools based on DFT can give deep, atomistic insight into reaction pathways. DFT is commonly used to study mechanisms of molecular systems via reaction pathway modeling and achieved via exploration of a potential energy surface (PES) is commenced until the relevant structures are located for a single elementary step.[72] There exist three main methods for optimizing reaction pathways on a PES: intrinsic reaction coordinate (IRC), single-ended growing string method (SE-GSM), and double-ended growing string method (DE-GSM). The IRC method begins at the transition state (TS) structure and optimizes toward the reactant and product structures.[73] However, this method requires that one obtains the correct TS structure prior to the IRC calculation, which is far from trivial. The SE-GSM and DE-GSM methods do not rely on the use of a guess TS structure, but instead uses reactant and/or product geometries. SE-GSM will start from reactant geometry and grow along a predetermined reaction coordinate until the TS is located, and then continue until a stable intermediate is identified. DE-GSM operates by using both reactant and product structures and grows towards the TS from both directions, which is known as a chain-of-states method.[74,75] However, the DE-GSM

method is limited by needing to have identified a proper product structure. To this end the SE-GSM can be used to explore the many different reaction coordinates. The DE-GSM method will be used in both Chapters II and III, where SE-GSM will also be used for Chapter III. Chapter IV utilizes a mass-weighted difference of distance method for the reaction coordinate. [76,77]

1.7 Thesis Overview

Chapters II-IV contained within this dissertation cover the development and application of a new combined constrained density functional theory and molecular mechanics (CDFT/MM) tool to study single-electron transfer (SET) reactions within solution and enzymes. Chapter V will be the conclusion and future outlook of the dissertation and CDFT/MM tool.

1.7.1 Chapter II Overview

Chapter II will cover the development and first applications of a new CDFT/MM method designed to investigate reactions involving SET. The SET-initiated reactions between the organic electron donors (OED) tetrakis(diethylamino)ethylene (TDAE) and tetrathiafulvene (TTF) and their respective aryldiazonium salts provided model reactions to apply the new CDFT/MM tool. The CDFT/MM tool is applied in the study of the two different OED systems in different explicit solvent environments (Figure 1.10). SET occurs between the OEDs and aryldiazonium substrates, and results in the loss of N_2 and cyclization of the substrates. The reactions are modeled using QM/MM models in which the OEDs and substrates are in the QM region, and the solvents and counterions are in

the MM region. The mechanisms of each system were refined, and the impact of solvent was investigated. It was found in both systems that inclusion of the solvent is important to accurately model the systems.

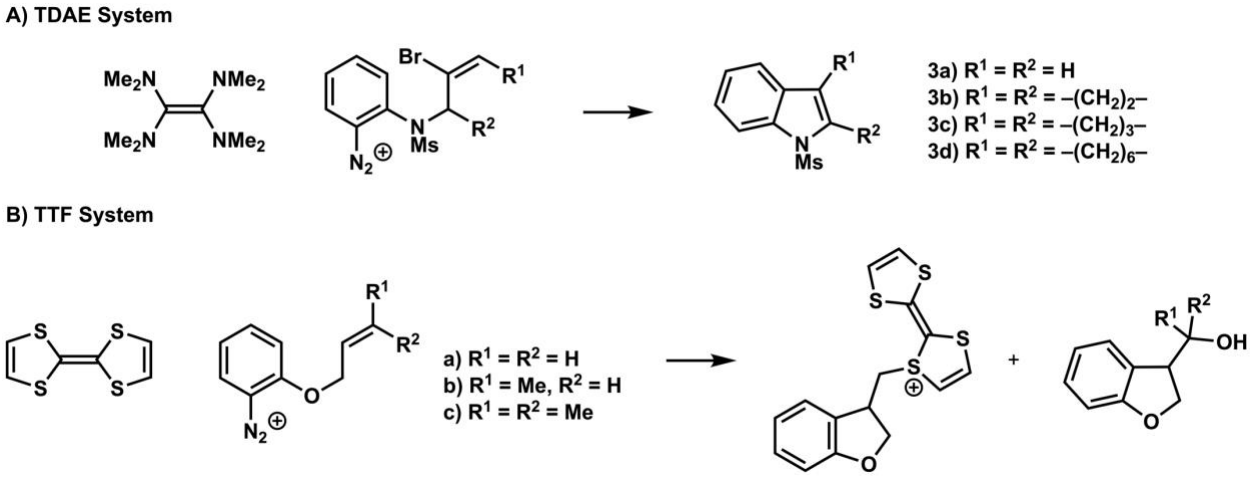

Figure 1.10. OED systems studied in Chapter II. Both systems under SET-initiated loss of N_2 and cyclization.

1.7.2 Chapter III Overview

Chapter III covers the initial modeling of the unprecedented ring expansion catalyzed by the α-ketoglutarate-dependent non-heme iron enzyme TropC.[78] TropC catalyzes the ring expansion of a nonaromatic six-membered dicarbonyl substrate to the aromatic seven-membered tropolone product stipitaldehyde within the biosynthetic pathway of stipitatic acid (Figure 1.11). Based on a partial crystal structure, a complete 3D enzyme model in a solvent explicit environment was constructed. Small active site mimics including the substrate, iron, and coordinating ligands were used to investigate a series of proposed one- and two-electron mechanisms. Originally, a two-electron mechanism of hydroxylation and semipinacol rearrangement mechanism was proposed towards the formation of the troplone product. However, one-electron pathways involving

a radical based or SET-based rearrangement were also envisioned. Based on experiments and computations, the QM models indicate a one-electron based mechanism over the originally proposed two-electron mechanism.

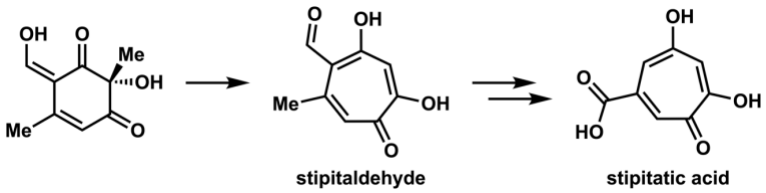

Figure 1.11. Ring expansion catalyzed by TropC.

1.7.3 Chapter IV Overview

Chapter IV will expand the modeling of TropC to the QM/MM scale using the model constructed in Chapter III. The reaction pathways initially studied in Chapter III are discussed in Chapter IV as well. This will include the use of the new CDFT/MM tool to study a SET between the substrate and the iron center. Importantly, the effects of the solvent and enzyme active site on the ring expansion catalyzed by TropC are also analyzed by way of non-covalent interactions (hydrogen bonds, van der Waal interactions). It was found that both one- and two-electron pathways are feasible, which is consistent with experimental observations. However, the one-electron pathway is favored to go through a radical-based ring expansion with the possibility of a SET-based ring expansion.

1.8 References

[1] Gao, S.-S.; Naowarojna, N.; Cheng, R.; Liu, X.; Liu, P. Recent Examples of α-Ketoglutarate-dependent Mononuclear Non-haem Iron Enzymes in Natural Product Biosynthesis. *Nat. Prod. Rep.* **2018**, *35*, 792–837.

[2] Kal, S.; Que, L. Dioxygen Activation by Nonheme Iron Enzymes With the 2-His-1-Carboxylate Facial Triad That Generate High-Valent Oxoiron Oxidants. *J. Biol. Inorg. Chem.* **2017**, *22*, 339–365.

[3] Tang, M.-C.; Zou, Y.; Watanabe, K.; Walsh, C. T.; Tang, Y. Oxidative Cyclization in Natural Product Biosynthesis. *Chem. Rev.* **2017**, *117*, 5226–5333.

[4] Wu, L.-F.; Meng, S.; Tang, G.-L. Ferrous Iron and α-Ketoglutarate-Dependent Dioxygenases in the Biosynthesis of Microbial Natural Products. *Biochim. Biophys. Acta* **2016**, *1864*, 453–470.

[5] Martinez, S.; Hausinger, R. P. Catalytic Mechanisms of Fe(II)- and 2-Oxoglutarate-Dependent Oxygenases. *J. Biol. Chem.* **2015**, *290*, 20702–20711.

[6] Van Der Kamp, M. W.; Mullholland, A. J.; Combined Quantum Mechanics/Molecular Mechanics (QM/MM) Methods in Computational Enzymology. *Biochemistry* **2013**, *52*, 2708–2728.

[7] Li, H.; Zhu, W.; Liu, Y. Mechanism of the Uncoupled Carbocylization and Epimerization Catalyzed by Two Non-heme Iron/α-Ketoglutarate Dependent Enzymes. *J. Chem. Inf. Model.* **2019**, *59*, 5086–5098.

[8] Li, H.; Liy, Y. Mechanistic Investigation of Isonitrile Formation Catalyzed by The Nonheme Iron/α-KG-Dependent Decarboxylase (Scoe). *ACS Catal.* **2020**, *10*, 2942–2957.

[9] Yan, L. Lui, Y. Insights into The Mechanism and Enantioselectivity in the Biosynthesis of Ergot Alkaloid Cycloclavine Catalyzed by Aj_EasH From *Aspergillus Japonicus*. *Inorg*. *Chem*. **2019**, *58*, 13771–13781.

[10] Nakashima, Y.; Mitsuhashi, T.; Matsuda, Y.; Senda, M.; Sato, H.; Yamazaki, M.; Uchiyama, M.; Senda, T.; Abe, I. Structural and Computational Bases for The Dramatic Skeletal Rearrangement in Anditomin Biosynthesis. *J. Am. Chem. Soc.* **2018**, *140*, 9743–9750.

[11] Bai, J.; Yan, L.; Liu, Y. Catalytic Mechanism of the PrhA (V150L/A232S) Double Mutant Involved in the Fungal Meroterpenoid Biosynthetic Pathway: A QM/MM Study. *Phys. Chem. Chem. Phys.* **2019**, *25658*–25668.

[12] Phung, Q. M.; Martín-Fernández, C.; Harvey, J. N.; Feldt, M. Ab Initio Calculations for Spin-Gaps of Non-heme Iron Complexes. *J. Chem. Theory. Comput.* **2019**, *15*, 4297–4304.

[13] Geng, C.L; Ye, S.; Neese, F. Analysis of Reaction Channels for Alkane Hydroxylation by Nonheme Iron(IV)-Oxo Complexes. *Angew. Chem. Int. Ed.* **2010**, *49*, 5717–5720.

[14] Lundberg, M.; Borowski, T. Oxoferryl Species in Mononuclear Non-heme Iron Enzymes: Biosynthesis, Properties and Reactivity From a Theoretical Perspective. *Coord. Chem. Rev.* **2013**, *257*, 277–289.

[15] Ye, S.; Geng, C.-Y.; Shaik, S.; Neese, F. Electronic Structure Analysis of Multistate Reactivity in Transition Metal Catalyzed Reactions: The Case of C–H Bond Activation by Non-heme Iron(IV)–Oxo Cores. *Phys. Chem. Chem. Phys.* **2013**, *15*, 8017–8030.

[16] Matsuda, Y.; Wakimoto, T.; Mori, T.; Awakawa, T.; Abe, I. Complete Biosynthetic Pathway of Anditomin: Nature's Sophisticated Synthetic Route to a Complex Fungal Meroterpenoid. *J. Am. Chem. Soc.* **2014**, *136*, 15326–15336.

[17] Havemann, J.; Vogel, D.; Loll, B.; Keller. U. Cyclolization Of D-Lysergic Acid Alkaloid Peptides. *Chem. Biol.* ***2014***, *21*, 146–155.

[18] Dürr, C.; Schnell, H.-J.; Luzhetskyy, A.; Murillo, R.; Weber, M.; Welzel, K.; Vente, A.; Bechthold, A. Biosynthesis of the Terpene Phenalinolactone In *Stretomyces* Sp. Tü6071: Analysis of the Gene Cluster and Generation Of Derivatives. *Chem. Biol.* **2006**, *13*, 365–377.

[19] Meng, S.; Han, W.; Zhao, J.; Jian, X.-H.; Pan, H.-X.; Tang, G.-L. A Six-Oxidase Cascade for Tandem C–H Bond Activation Revealed by Reconstitution of Bicyclomycin Biosynthesis. *Angew. Chem. Int. Ed.* **2018**, *57*, 719–723.

[20] Ishikawa, N.; Tanaka, H.; Koyama, F.; Noguchi, H.; Wang, C. C. C.; Hotta, K.; Watanabe, K. Non-heme Dioxygenase Catalyzes Atypical Oxidations of 6,7-Bicyclic Systems to Form the 6,6-Quinolone Core of Viridicatin-type Fungal Alkaloids. *Angew. Chem. Int. Ed.* **2014**, *53*, 12880–12884.

[21] Lukat, P.; Katsuyama, Y.; Wenzel, S.; Binz, T.; König, C.; Blankenfeldt, W.; Brönstrup, M.; Müller, R. Biosynthesis of Methyl-Proline Containing Griselimycins, Natural Products With Anti-tuberculosis Activity. *Chem. Sci.* **2017**, *8*, 7521–7527.

[22] Li, L.; Tang, M.-C.; Tang, S.; Gao, S.; Soliman, S.; Hang, L.; Xu, E.; Ye, T.; Watanabe, K.; Tang, Yi. Genome Mining and Assembly-line Biosynthesis of the UCS1025A Pyrrolizidinone Family of Fungal Alkaloids. *J. Am. Chem. Soc.* **2018**, *140*, 2067–2071.

[23] Chen, L.; Yue, Q.; Li, Y.; Niu, X.; Xiang, M.; Wang, W.; Bills, G. F.; Liu, X.; An, Z. Engineering of *Glarea Lozoyensis* for Exclusive Production of the Pneumocandin B_0 Precursor of the Antifungal Drug Caspofungin Acetate. *Appl. Environ. Microbiol.* **2015**, *81*, 1550–1558.

[24] Jang, W.; Cacho, R. A.; Chiou, G.; Garg, N. K.; Tang, Y.; Walsh, C. T. EcdGHK Are Three Tailoring Iron Oxygenases for Amino Acid Building Blocks of the Echinocandin Scaffold. *J. Am. Chem. Soc.* **2013**, *135*, 4457–4466.

[25] Matsuda, Y.; Iwabucji, T.; Fujimoto, T.; Awakawa, T.; Nakashima, Y.; Mori, T.; Zhang, H.; Hayashi, F.; Abe, I. Discovery of Key Dioxygenases That Diverged in Paraerquonin and Acetoxydehydroaustin Pathways in *Penicillium Brasilianum*. *J. Am. Chem. Soc.* **2016**, *138*. 12671–12677.

[26] Tarhonskaya, H.; Szöllössi, A.; Leung, I. K. H.; Bush, J. T.; Henry, L.; Chowdhury, R.; Iqbal, A.; Claridge, T. D. W.; Schofield, C. J.; Flashman, E. Studies on Deacetoxycephalosporin C Synthase Support a Consensus Mechanism for 2-Oxoglutarate Dependent Oxygenases. *Biocemistry* **2014**, *53*, 2483–2493.

[27] Matsuda, Y.; Awakawa, T.; Wakimoto, T.; Abe, I. Spiro-Ring Formation is Catalyzed by a Multifunctional Dioxygenase in Asutinol Biosynthesis. *J. Am. Chem. Soc.* **2013**, *135*, 10962–10965.

[28] Jakubczyk, D.; Caputi, L.; Hatsch, A.; Nielsen, C. A. F.; Diefenbacher, M.; Klein, J.; Molt, A.; Schröder, H.; Cheng, J. Z.; Naesby, M.; O'Connor, S. E. Discovery and Reconstitution of the Cycloclavine Biosynthetic Pathway–Enzymeatic Formation of a Cyclopropyl Group. *Angew. Chem. Int. Ed.* **2015**, *54*, 5117–5121.

[29] Chekan, J. R.; Mckinnie, S. M. K.; Moore, M. L.; Poplawski, S. G.; Michael, T. P.; Moore, B. S. Scalable Biosynthesis of the Seaweed Neurochemical, Kainic Acid. *Angew. Chem. Int. Ed.* **2019**, *58*, 8454–8457.

[30] Brunson, J. K.; Mckinnie, S. M. K.; Chekan, J. R.; Mccrow, J. P.; Miles, Z. D.; Bertrand, E. M.; Bielinski, V. A.; Luhavaya, H.; Oborník, M.; Smith, G. J.; Hutchins, D. A.; Allen, A. E.; Moore, B. S. Biosynthesis of the Neurotoxin Domoic Acid in a Bloom-Forming Diatom. *Science*, **2018**, *361*, 1356–1358.

[31] Nakamura, H.; Matsuda, Y.; Abe, I. Unique Chemistry of Non-heme Iron Enzymes in Fungal Biosynthetic Pathways. *Nat. Prod. Rep.* **2018**, *35*. 633–645.

[32] Matsuda, Y.; Wakimoto, T.; Mori, T.; Awakawa, T.; Abe, I. Complete Biosynthetic Pathway of Anditomin: Nature's Sophisticated Synthetic Route to a Complex Fungal Meroterpenoid. *J. Am. Chem. Soc.* **2014**, *136*, 15326–15336.

[33] Horn, F.; Linde, J.; Mattern, D. J.; Walther, G>; Guthke, R.; Brakhage, A. A.; Valiante, C. Draft Genome Sequence of the Funcgus *Penicillium Brasilianum* MG11. *Genome Announc.* **2015**, *3*, E00724–00715.

[34] Nakashima, Y.; Mori, T.; Nakamura, H.; Awakawa, T.; Hoshino, S.; Senda, M.; Senda, T.; Abe, I. Structure Function and Engineering Of Multifunctional Non-heme Iron Dependent Oxygenases in Fungal Meroterpenoid Biosynthesis. *Nat. Commun.* **2018**, *9*, 104.

[35] Dunham, N. P.; Chang, W.-.; Mitchell, A. J.; Martinine, R. J.; Zhang, B.; Bergman, J. A.; Rajakovich, L. J.; Wang, B.; Silakov, A.; Krebs, C.; Boal, A. K.; Bollinger, J. M. Jr. Two Distinct Mechanisms for C–C Desaturation by Iron(II)- and 2-(Oxo)glutarate-Dependent Oxygenases: Importance of α-Heteroatom Assistance. *J. Am. Chem. Soc.* **2018**, *140*, 7116–7126.

[36] Dunham, N. P.; Mitchell, A. J.; Del Río Pantoja, J. M.; Krebs, C.; Bollinger, J. M. Jr.; Boal, A. K. α-Amine Desaturation of D-Arginine by the Iron(II)- and 2-(Oxo)glutarate-Dependent L-Arginine 3-Hydroxylase, VioC. *Biochemistry* **2018**, *57*, 6479–6488.

[37] Chang, W.-C.; Yang, Z.-J.; Tu, Y.-.; Chien, T.-C. Reaction Mechanism of a Nonheme Iron Catalyzed Oxidative Cyclization via C–C Bond Formation. *Org. Lett.* **2019**, *21*, 228–232.

[38] Huang, J. -L.; Tang, Y.; Yu, Chng, -P.; Sanyl, D. Jia, X.; Guo, Y.; Chang, W. -C. Mechanistic Investigation of Oxidative Decarboxylation Catalyzed by Two Iron(II)- and 2-Oxoglutarate-Dependent Enzymes. *Biochemistry*, **2018**, *57*, 1838–1841.

[39] Chen, T.-Y.; Xue, S.; Tsai, W.-C.; Chien, T.-C.; Guo, Y.; Chang, W.-C. Deciphiring Pyyrolidine and Olefin Formation Mechanism in Kainic Acid Biosynthesis. *ACS Catal.* **2021**, *11*, 278–282.

[40] Shimo, S.; Ushimaru, R.; Engelbrecht, A.; Harada, M.; Miyamoto, K.; Kulik, A.; Uchiyama, M.; Kaysser, L.; Abe, I. Stereodivergent Nitrocyclopropane Formation during Biosynthesis of Belactosins and Hormaomycins. *J. Am. Chem. Soc.* **2021**, *143*, 18413–18418.

[41] Chen, T.-Y.; Zheng, Z.; Zhang, X.; Chen, J.; Cha, L.; Tang, Y.; Guo, Y.; Zhou, J.; Wang, B.; Liu, H.-w.; Chang, W.-C. Deciphering the Reaction Pathway of Mononuclear Iron Enzyme-Catalyzed N≡C Triple Bond Formation in Isocyanide Lipopeptide and Polyketide Biosynthesis. *ACS Catal.* **2022**, *12*, 2270–2279.

[42] Ali, H. S.; Ghafoor, S.; de Visser, S. P. Density Functional Theory Study into the Rection Mechanism of Isonitrile Biosynthesis by the Nonheme Iron Enzyme ScoE. *Top. Catal.* **2022,** *65*, 528–543.

[43] Pan, J.; Wenger, E. S.; Matthews, M. L.; Pollock, C. J.; Bhardwaj, M.; Kim, A. K.; Allen, B. D.; Grossman, R. B.; Krebs, C.; Bollinger, J. M. Jr. Evidence of Modulation of Oxygen Rebound Rate in Control of Outcome by Iron(II)- and 2-Oxoglutarate-Dependent Oxygenases. *J. Am. Chem. Soc.* **2019**, *141*, 15153–15165.

[44] Jakubczyk, D.; Caputi, L.; Hatsch, A.; Nielsen, C. A. F.; Diefenbacher, M.; Klein, J.; Molt, A.; Schröder, H.; Cheng, J. Z.; Maesby, M.; O'Connor, S. E. Discovery and Reconstitution of the Cycloclavine Biosynthetic Pathway–Enzymatic Formation of a Cyclopropyl Group. *Angew. Chem. Int. Ed.* **2015**, *54*, 5117–5121.

[45] Jakubczyk, D.; Caputi, L.; Stevenson, C. E. M.; Lawson, D. M.; O'Connor, S. E. Structural Characterization of EasH (*Aspergillus Japonicus*) – An Oxidase Involved in Cycloclavine Biosynthesis. *Chem. Commun.* **2016**, *5*, 14306–14309.

[46] De Visser, S. P. Mechanistic Insight On The Activity and Substrate Selectivity of Nonheme Iron Dioxygenases. *Chem. Rec.* **2018**, *18*, 1501–1516.

[47] Ahmadi, S.; Herrera, L. B.; Chehelamirani, M.; Hostaš, J.; Jalife, S.; Salahub, D. R. Multiscale Modeling of Enzymes: QM-Cluster, QM/MM, and QM/MM/MD: A Tutorial Review. *Int. J. Quantum Chem.* **2018**, *118*, E25558.

[48] Himo, F. Recent Trends in Quantum Chemical Modeling of Enzymatic Reactions. *J. Am. Chem. Soc.* **2017**, *139*, 6780–6786.

[49] Vedin, N. P.; Lundberg, M. Protein Effects in Non-heme Iron Enzyme Catalysis: Insights From Multiscale Models. *J. Biol. Inorg. Chem.* **2016**, *21*, 645–657.

[50] Quesne, D. G.; Borowski, T.; De Visser, S. P. Quantum Mechanics/Molecular Mechanics Modeling of Enzymatic Processes: Caveats and Breakthroughs. *Chem. Eur. J.* **2016**, *22*, 2562–2581.

[51] Blomberg, M. R. A.; Borowski, T.; Himo, F.; Liao, R.-Z.; Siegbahn, P. E. M. Quantum Chemical Studies of Mechanisms for Metalloenzymes. *Chem. Rev.* **2014**, *114*, 3601–3658.

[52] Sousa, S. F.; Ribeiro, A. J. M.; Neves, R. P. P.; Brás, N. F.; Cerqueira, N. M. F. S. A.; Fernandes, P. A.; Ramos, M. J. Application of Quantum Mechanics/Molecular Mechanics Methods In The Study Of Enzymatic Reaction Mechanisms. *Wires Comput. Mol. Sci.* **2017**, *7*, E1281.

[53] Groenhof, G. Introduction to QM/MM Simulations. *Methods Mol. Biol.* **2013**, *924*, 43–66.

[54] Becke, A. D. Perspective: Fifty Years of Density-Functional Theory in Chemical Physics. *J. Chem. Phys.* **2014**, *140*, 18A301.

[55] 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.

[56] Cohen, A. J.; Mori-Sánchez, P.; Yang, W. Challenges for Density Functional Theory. *Chem. Rev.* **2012**, *112*, 289–320.

[57] Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100.

[58] Perdew, J. P.; Density-Functional Approximation for the Correlation Energy of the Inhomogeneous Electron Gas. *Phys. Rev. B* **1986**, *33*, 8822–8824.

[59] Becke, A. D. Density-functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

[60] Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-energy Formula Into a Functional of the Electron Density. *Phys. Rev. B* **1998**, *37*, 785–789.

[61] Kaduk, B.; Kowalczyk, T.; Van Voorhis, T. Constrained Density Functional Theory. *Chem. Rev.* **2012**, *112*, 321–370.

[62] Dederichs, P. H.; Blügel, S.; Zeller, R.; Akai, H. Ground States of Constrained Systems: Application to Cerium Impurities. *Phys. Rev. Lett.* **1984**, *53*, 2512–2515.

[63] Wu, Q.; Van Voorhis, T. Direct Optimization Method to Study Constrained Systems Within Density-Functional Theory. *Phys. Rev. A* **2005**, *72*, 024502.

[64] Ku, C.; Sit, P. H.-L. Oxidation-State Constrained Density Functional Theory for The Study Of Electron-Transfer Reactions. *J. Chem. Theory Comput.* **2019**, *15*, 4781–4789.

[65] Ramos, P; Pavanello, M. Constrained Subsystem Density Functional Theory. *Phys. Chem. Chem. Phys.* **2016**, *18*, 21172–21178.

[66] Plaisance, C. P.; Van Santen, R. A.; Reuter, K. Constrained-Orbital Density Functional Theory. Computational Method and Applications to Surface Chemical Processes. *J. Chem. Theory Comput.* **2017**, *13*, 3561–3574.

[67] Ziegler, T.; Krykunov, M. On The Calculation of Charge Transfer Transitions With Standard Density Functionals Using Constrained Variational Density Functional Theory. *J. Chem. Phys.* **2020**, *133*, 074104.

[68] Evangelista, F. A. Shushkov, P.; Tully, J. C. Orthogonality Constrained Density Functional Theory for Electronic Excited States. *J. Phys. Chem. A* **2013**, *117* 7378–7392.

[69] Oberhofer, H.; Blumberger, J. Electronic Coupling Matrix Elements From Charge Constrained Density Functional Theory Calculations Using a Plane Wave Basis Set. *J. Chem. Phys.* **2010**, *133*, 244105.

[70] Guo, Z.; Liu, Y.; Zhang, X. Constrained Lattice Density Functional Theory and Its Applications On Vapor–Liquid Nucleations. *Sci. Bull.* **2015,** *60*, 320–327.

[71] Carmona-Espínfola, J.; Núñez-Rojas, E.; García-Melgarejo, V.; Gázquez, J. L.; Alejandre, J. Constrained Dipole Moment Density Functional Theory for Charge Distributions in Force Fields for the Study Of Molecular Fluids. *J. Chem. Phys.* **2020**, *152*, 124116.

[72] Dewyer, A. L.; Argüelles Alonso J.; Zimmerman, P. M. Methods for Exploring Reaction Space in Molecular Systems. *Wires Comput. Mol. Sci.* **2017**, *8*, E1354.

[73] Maeda, S.; Harabuvhi, Y.; Ono, Y.; Taketsugu, T.; Morokuma, K. Intrinsic Reaction Coordinate: Calculation, Bifurcation, and Automated Search. *Int. J. Quantum Chem.* **2015**, *115*, 258–269.

[74] Sheppard, D.; Terrell, R.; Henkelman, G. Optimization Methods for Finding Minimum Energy Paths. *J. Chem. Phys.* **2008**, *128*, 134106.

[75] Fukui, K. The Path of Chemical Reactions – The IRC Approach. *Acc. Chem. Res.* **1981**, *14*, 363–368.

[76] Schlitter, J.; Constraint Methods for Determining Pathways and Free Energy of Activated Processes. *Eut. Phys. J. Special Topics* **2011**, *200*, 91–105.

[77] Schlitter, J.; Swegat, W.; Mülders, T. Distance-Type Reaction Coordinates for Modelling Activated Processes. *J. Mol. Model* **2001**, *7*, 171–177.

[78] Davison, J.; Fahad, A.; Cai, M.; Song, Z.; Yehia, S. Y.; Lazarus, C. M.; Bailey, A. M.; Simpson, T. J.; Cox, R. J. Genetic, Molecular, and Biochemical Basis of Fungal Tropolone Biosynthesis. *PNAS* **2012**, *109*, 7642–7647.

Chapter II

Simulating Electron Transfer Reactions in Solution: Radical-Polar Crossover

This chapter based on a manuscript entitled, "Simulating Electron Transfer Reactions in Solution Radical-Polar Crossover," Skinner, K. C.; Kammeraad, J. A.; et. al., currently under revision for future publication.

Abstract

Single-electron-transfer (SET) promotes a wide variety of interesting chemical transformations, but modeling of SET requires a careful treatment of electronic and solvent effects to give meaningful insight. Therefore, a combined constrained density functional theory and molecular mechanics (CDFT/MM) tool is introduced herein specifically for SET-initiated reactions. Mechanisms for two radical-polar crossover reactions involving the organic electron donors tetrakis(dimethylamino)ethylene (TDAE) and tetrathiafulvalene (TTF) were studied with the new tool. An unexpected tertiary radical intermediate within the TDAE system was identified, relationships between kinetics and substitution in the TTF system are explained, and the impact of the solvent environment of both the TDAE and TTF systems are examined. Overall, the new method is shown to usefully inform mechanistic analysis of SET-initiated reactions, and therefore is envisioned to be useful for studying reactions in the condensed phase.

2.1 Introduction

Reactions initiated through single electron transfer (SET) have been used in organic synthesis for nearly two centuries. More recently, SET processes have been leveraged to achieve novel photocatalytic[1-8] and transition-metal-free coupling reactions[9] involving hypervalent iodide reagents[10] and neutral organic electron donors (OEDs).[11-13] For all of these reactions, SET initiates a subsequent radical addition, scission, atom/group transfer, and/or coupling event.[14] Although these overarching mechanisms are broadly employed, the specific mechanistic details of SET and ensuing transformations can be challenging to study due to the short-lived nature of radical intermediates. Understanding the kinetics of SET is vital for distinguishing between mechanistic possibilities, leading to improvements in chemical and biological synthetic methods.[1,15]

To aid in the study of SET-initiated reactions, computational and theoretical tools based on density functional theory (DFT) can give deep, atomistic insights. Reaction pathway analysis can reveal the kinetics and thermodynamics of transformations involving radical intermediates, and Marcus theory can be used to describe the rate of electron transfer (k_{et}) for SET processes.[16] Quantum mechanical (QM) approaches to studying SET processes include deductive and constructive diabatization[17] and the Nelsen Four-Point method.[18,19,] In particular, the Nelsen Four-Point method has been widely used due to its simplicity and straightforwardness, and operates by obtaining the reorganization energies of the donor and acceptor compounds independently. Although this has been shown to be successful,[26] if it is desired to include the donor and acceptor in the same simulation, then an extension of DFT is required. Whereas DFT simulations typically simulate a single electronic state, SET necessarily involves more than one

electronic state. To treat SET-initiated reactions, constrained DFT (CDFT) is potentially useful[20,21,22] as the electron density distribution can be used to achieve a desired charge or spin state. CDFT has seen substantial developments extend constraints, such as oxidation state[23], dipole moment[24], orbital density[25], and others.[26-30] CDFT also permits evaluation of the electronic coupling[31-35], which is needed in Marcus theory estimates of the SET rate constant. While CDFT primarily has been used for this purpose, it has not yet been widely used[36-38] as a tool for SET reaction pathway studies. Instead, CDFT has been useful in a range of applications, such as charge transfer[39-43], long-range electron transfer[44,45], magnetic exchange couplings[46-48], excited states[49,50], and spectroscopy[51-53].

For an accurate description of an SET process, reorganization of the solvent[20] following ET needs to be considered. In particular, solvent molecules prefer different orientations before and after ET to accommodate the relocation of charge, so solute-solvent interactions are essential to SET modeling.[20,54,55] To address this CDFT has been combined with molecular mechanics (MM) models to study SET of organic semiconductors[56], metals[57], molecular dyads[61,58], peptides[59,60], and proteins[61-63] in solution. The same considerations apply to SET-initiated reactions, where molecular transformations after ET will be strongly affected by the presence and motion of solvent molecules. This consideration motivates the CDFT/MM models described herein, which will be shown to have good utility in treating this interesting class of reactions.

Treating solvent effects surrounding a region where reactions or ET are taking place is optimally modeled using quantum mechanics/molecular mechanics (QM/MM) techniques.[64-67] QM/MM modeling provides a way to treat the electronically interesting region using accurate QM methods and the remaining environment using MM methods.

This combination provides the accuracy of QM methods and the immense computational time savings available to MM methods. To study SET-initiated reactions, the QM technique of choice herein is CDFT. Figure 2.1 depicts the difference in treatment of electronic states between two organic molecules within a solvent environment. When the QM region (surfaces) are treated with DFT, the positive charge (Figure 2.1A, blue surface) is located on the molecule with the diazonium group. However, using CDFT can emulate SET and the positive charge instead lies on the OED (Figure 2.1B, grey surface). While CDFT/MM has been used to study ET processes, studies have not investigated any reactions that follow the ET step. Since CDFT/MM should be able to capture the full picture of SET-initiated reactions with a realistic solvent model, this modeling strategy will be applicable to a variety of synthetically relevant transformations.

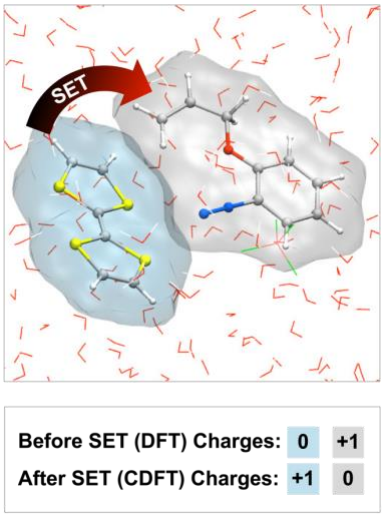

Figure 2.1. Electronic states of donor (TTF, blue) and acceptor (aryldiazonium, gray) fragments in an explicit solvent environment. DFT gives the initial electronic state, whereas CDFT can access a second electronic state with a distinct charge distribution. QM region is represented by ball-and-stick and surfaces. Lines represent the MM region.

This chapter looks into two interesting reactions via CDFT/MM, involving OEDs tetrakis(dimethylamino)ethylene (TDAE, **2.1**) and tetrathiafulvalene (TTF, **2.10**) and their

respective aryldiazonium salts (**2.2a**, **2.11a**). The mechanisms for reaction are elucidated by identifying and characterizing previously unmentioned intermediates. In addition, the impact of the solvent environment following SET is also investigated for both OED reactions. Overall, this work's detailed, atomistic mechanisms reveal how the solvent environment influences the reactions pathways, providing insights for better SET acceptor and/or donor design.

2.2 Computational Details

All 3D images were generated using the Visual Molecular Dynamics software program.[68]

2.2.1 Model Preparation

Models for the TDAE and TTF systems were prepared in CHARMM.[69] Parameters for the OEDs and aryldiazonium salts were obtained from CGenFF[70,71] 2.3.0 when possible and from unrestricted HF[72] 6-31G*[73-78] optimizations otherwise. 40 Å solvent cubes were used for each system. Solvent boxes used with periodic boundary conditions and were obtained from the CHARMM-GUI[79] (water) and Packmol[80] (dimethylformamide, DMF). DMF solvent box was subjected to an initial 5000 step MM minimization using the steepest descent algorithm and the CHARMM36[81] force field. The particle mesh Ewald method was used to describe long-range electrostatic effects. Next a two-step molecular dynamics (MD) simulation was performed with the CHARMM/OpenMM[82] interface in version c42b1 using periodic boundary conditions, particle mesh Ewald for long range electrostatic interactions, a Langevin thermostat (friction coefficient = 5), and a 2 fs timestep. The first step of the MD simulation was heating from 10 K to 298 K for 5 ps, and

the second was equilibration of the models at 298 K for 20 ps. The OEDs and aryldiazonium salts were then solvated using the solvent boxes and all solvent molecules within 2.8 Å of heavy atoms were removed. A 5000 step MM minimization was performed to using the steepest descent algorithm and the CHARMM36 force field. A force constraint of 100.0 kcal/mol $Å^2$ at a distance of 3.5 Å was applied between a sulfur atom of TTF and the primary carbon of the aryldiazonium salt to maintain proper orientation.

2.2.2 MD Simulations

Two-step MD protocols were performed with the CHARMM/OpenMM utilizing the same parameters as the DMF solvent box equilibration. The first step was heating from 10 K to 298 K for 5 ps, and the second was equilibration of the models at 298 K for 50 ps.

2.2.3 Free Energy Surfaces

Starting from the equilibrated MD models, free energy surfaces[15,17] for the TDAE system were constructed based on a total of 1000 ps of additional sampling (ten 100 ps simulations), where a frame every 1 ps was saved. For each frame, DFT/MM and CDFT/MM energy calculations were performed reflecting ground and SET conditions, respectively, using the unrestricted B3LYP[83,84] density functional and 6-31G* basis set totaling 4000 energy calculations (1000 for each state on each surface). Taking the difference in energy between the ground and SET states results in the vertical energy gap (ΔE, eq 1). Once the vertical energy gap is calculated, the probability distribution for electron transfer $P_{ET}(\Delta E)$ can be calculated using eq 2. Within the linear response approximation, the probability distribution in the form of eq 2 is assumed to be Gaussian

and will therefore produce parabolic free energy profiles consistent with Marcus theory.[16] The free energy of electron transfer ($G_{ET}(\Delta E)$) can then be calculated using eq 3, which results in quadratic Marcus curves.

$$\Delta E = E_{SET} - E_{ground} \qquad \text{eq 1}$$

$$P_{ET}(\Delta E) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(\Delta E-\mu)^2}{2\sigma^2}} \qquad \text{eq 2}$$

$$G_{ET}(\Delta E) = -k_B T \ln\left(P_x(\Delta E)\right) \qquad \text{eq 3}$$

Following construction of the curves the lowest energy point was found to be on the ground state surface with a value of 1.8 kcal/mol. To have the lowest energy value be set to 0.0 kcal/mol, 1.8 kcal/mol was subtracted from all values of $G_{ET}(\Delta E)$. The minima of the two curves can be used to calculate reorganization energy (λ), ΔG, and the effective barrier for SET transfer (TS_{SET}).

$$\lambda = (\Delta E_{SET} - \Delta E_{Ground})/2 \qquad \text{eqn 4}$$

$$\Delta G = (\Delta E_{SET} + \Delta E_{Ground})/2 \qquad \text{eqn 5}$$

$$TS_{SET} = \frac{(\lambda+\Delta G)^2}{4\lambda} \qquad \text{eqn 6}$$

Once ΔG was calculated, the values of for the SET were adjusted again adjusted.

2.2.4 QM/MM Details

The last frame of the equilibrated model was pruned to include all molecules within 18 Å of the QM region for the TDAE system, and 12 Å for the TTF system. Atoms past 10 Å

from the QM region of the TDAE system, and 7 Å for the TTF system were frozen. The QM region was selected to be the OED and corresponding aryldiazonium substrate based on bond-forming and breaking events. The MM region contains the relevant solvent and a tetrafluoroborate counterion. All simulations and calculations were performed using a development version of pDynamo[85], which is a Python program designed to perform QM and QM/MM simulations. pDynamo is linked to external software packages, such as ORCA[86], to perform QM energies and gradients. Electrostatic embedding of the MM region is used to enable the QM region to be polarized by the MM region. Also, within electrostatic embedding, the MM region is affected by the QM charges by way of Coulomb interactions. To provide support for CDFT, pDynamo was interfaced with QChem. All geometry optimizations were performed using pDynamo, where the QM energies and gradients were obtained with QChem 5.0[87] using the unrestricted B3LYP 6-31G*. All reaction path simulations were performed using the growing string method implemented in pDynamo.

2.2.5 pDynamo/QChem Interface

To interface QChem into pDynamo, cclib[88] was used to obtain energies, gradients, and charges from the QChem outputs. Both pDynamo and QChem calculate the nuclear repulsion energy, so the energy computed from QChem was removed during energy evaluations. CDFT constraints are automatically applied to user defined atom ranges.

2.3 SET-initiated Indole Synthesis

TDAE (**1**) is an OED[11-13] that initiates radical-polar crossover of aryldiazonium salts **2.2a-d** to form indoles **2.3a-d**[90] (Figure 2.2). Whereas TDAE is one of many OED reagents, it is expected to show the general principles and features of SET-initiated reactions, making it an interesting case for CDFT/MM modeling. In the first step of the proposed mechanism[90], SET occurs from **2.1** to **2.2a** to induce loss of N_2 and form the aryl radical intermediate **2.4**. Next, 5-*exo-trig* cyclization generates primary radical **2.6**. After cyclization, dissociation of bromine leads to indoline intermediate **2.7**, which can then tautomerize to form product **2.3a**. Although these proposed steps all appear plausible, limited information is available about the mechanistic details of the reaction. Toward a more thorough mechanistic understanding, the entire reaction pathway was examined

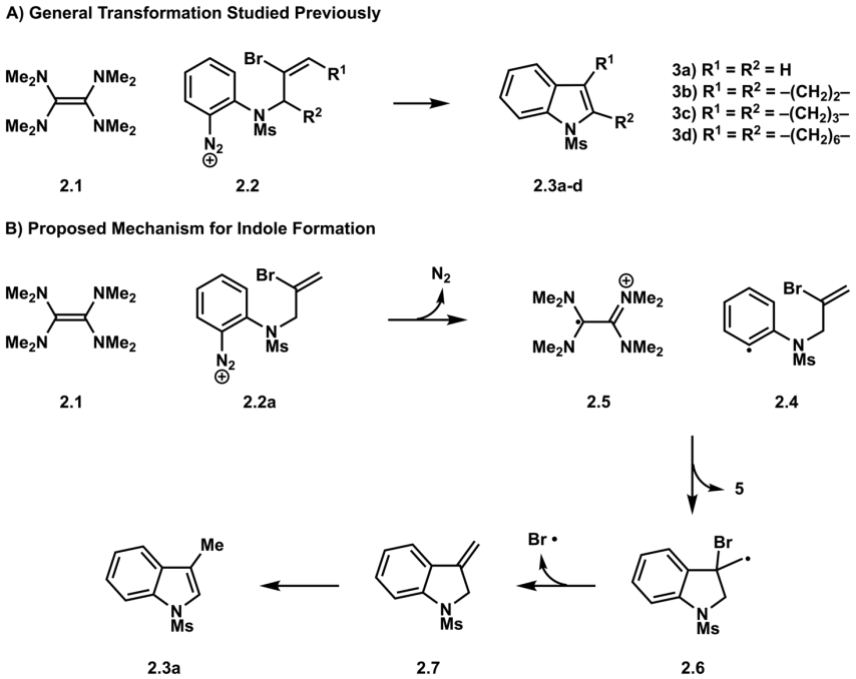

Figure 2.2. Overview of TDAE (**1**) system. A) Transformation studied by Murphy and coworkers[89] involving **1** and the aryldiazonium salts **2.2a-d**. B) Previously proposed mechanism for SET-initiated formation of indole **2.3a**.

using CDFT/MM. Two aspects of anticipated interest are the synchronicity of the SET-initiated loss of N_2 and the fate of the bromine atom.

Results from CDFT/MM for the mechanism of indole synthesis are illustrated in Figure 2.3. CDFT/MM simulates SET from TDAE (**1**) to the substrate (**2a**) by forcing one electron to move in an overall singlet spin multiplicity. The process was computed to be thermodynamically favorable by -24.9 kcal/mol (Figure 2.3). This result does not immediately seem consistent with known expectation for reversible electron transfer involving TDAE.[11] To understand the role that solvent environment plays on the SET process, MD sampling was performed to construct the free energy curves in Figure 2.4. According to equations 4-7, this data gives ΔG of 16.9 kcal/mol, λ of 43.8 kcal/mol, and an effective barrier (TS_{SET}) of 21.0 kcal/mol for SET. These results are counter to what is

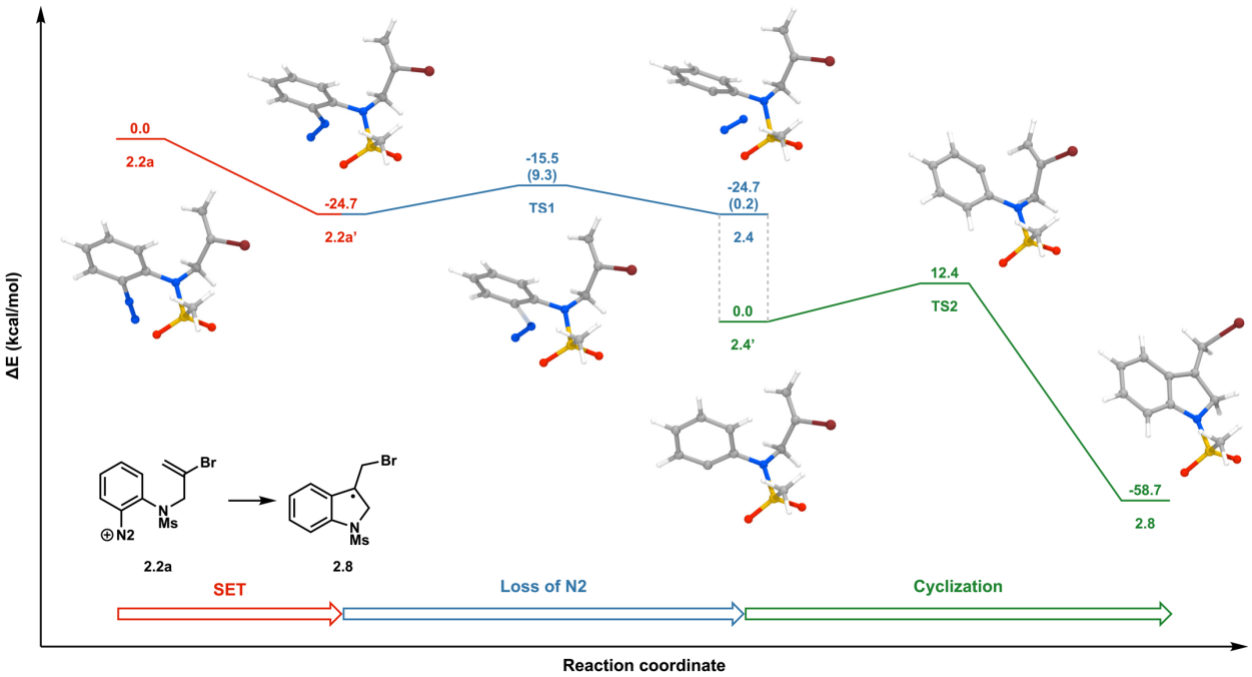

Figure 2.3. QM/MM reaction profile of the TDAE system showing overall thermodynamic favorability for tertiary radical (**2.8**) formation. Energies are in parentheses are relative to the reaction step, and all energies are in kcal/mol. QM region contained TDAE and substrate, but only transformations for substrate are shown for clarity. Gray dashed line represents removal of N_2 from our model. All simulations and calculations were performed at B3LYP 6-31G*/CHARMM36 in a DMF solvent environment.

expected for this system, as it suggests SET is not feasible. Likely, this is the case because the force fields for the MD sampling do not capture the substrate-solvent interactions accurately, and reparameterization of the force fields is needed to address this issue.

After SET, CDFT/MM indicated that N_2 did not immediately dissociate (Figure 2.3), instead, **2a** changed geometry to the aryldiazenyl radical intermediate **2a'**, and that TDAE (**1**) converted into the expected twisted[11] form (**5**). Intermediate **2a'** possesses a decreased C–N–N angle from 179.1° to 125.7°, and increased C–N and N–N bond lengths from 1.36 Å to 1.45 Å and 1.13 Å to 1.19 Å, respectively. These geometric changes reflect the presence of an additional electron on the external nitrogen of the N_2 group. Analysis of the spins of the two nitrogen atoms shows an increase from zero to 0.143 e^- and -0.716 e^- for the internal and external nitrogen atoms, respectively. Taken with the increase in bond length, this indicates a change to a double bond between the

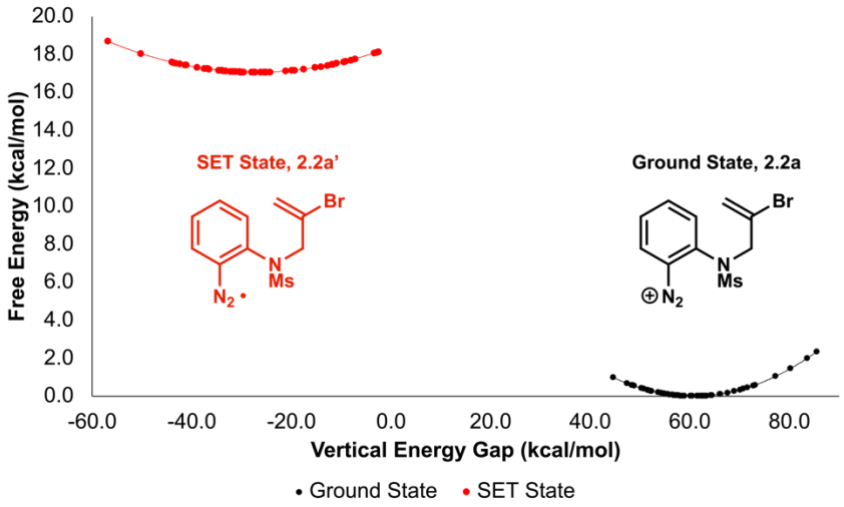

Figure 2.4. Free energy profiles obtained from MD simulations. Black is the ground state MD sampling and red is the SET MD sampling. Current subset of results indicates an unfavorable SET. Ground state simulations performed with TDAE (**2.1**) and substrate (**2.2a**) MM parameters. SET simulations performed with $TDAE^{\cdot+}$ (**2.3**) and aryldiazenyl radical MM parameters obtained from HF 6-31G*. Simulations were performed at using CHARMM36 in a DMF solvent environment and single points were taken at B3LYP 6-31G*/CHARMM36.

nitrogen atoms. This poises the substrate to undergo loss of N_2 through homolysis of the C–N bond, by providing an N-centered radical that will be a part of the N–N triple bond. To finish activation of the substrate and form aryl radical intermediate **4**, dissociation of N_2 occurs with a barrier of 9.4 kcal/mol (**TS1**). This barrier is in agreement with experimental barriers falling between 9.8 kcal/mol to 11.2 kcal/mol.[90-92] The formation of **4** is isoelectronic due to being uphill by only 0.2 kcal/mol. Since the nitrogen gas will bubble out of solution, this step is effectively irreversible and was removed from further simulations (**4** to **4'**).

The CDFT/MM model was then used to examine 5-exo-trig cyclization to form the putative primary radical intermediate **2.6**. This intermediate, however, was found to be unstable, and instead a 1,2-bromo shift led to tertiary radical intermediate **2.8** (Figure 2.3). This is a deviation from the experimental hypothesis, as the 1,2-bromo shift was not suggested as part of the mechanism. However, formation of intermediate **2.8** is likely driven by the delocalization of the radical across the aromatic ring, making this discovered step appear reasonable. Analysis of the geometric structures and spin states along the reaction path provided insights into the barrier and thermodynamics of this step.

To understand the pathway towards formation of intermediate **2.8**, it should first be recognized that the significantly downhill energy of reaction, -58.7 kcal/mol, provides a strong driving force. There is a barrier, however, of 12.4 kcal/mol, which reflects the energetic cost for obtaining a more reactive conformer of the substrate, in which aryl ring and alkene portions of the substrate are rotated to properly align them for C–C bond formation (Figure 2.5, **TS2**). Up until reaching the TS geometry, the spin on the aryl radical carbon is 1.0 e^-, meaning no changes in bonding occurred during the rotation. Following

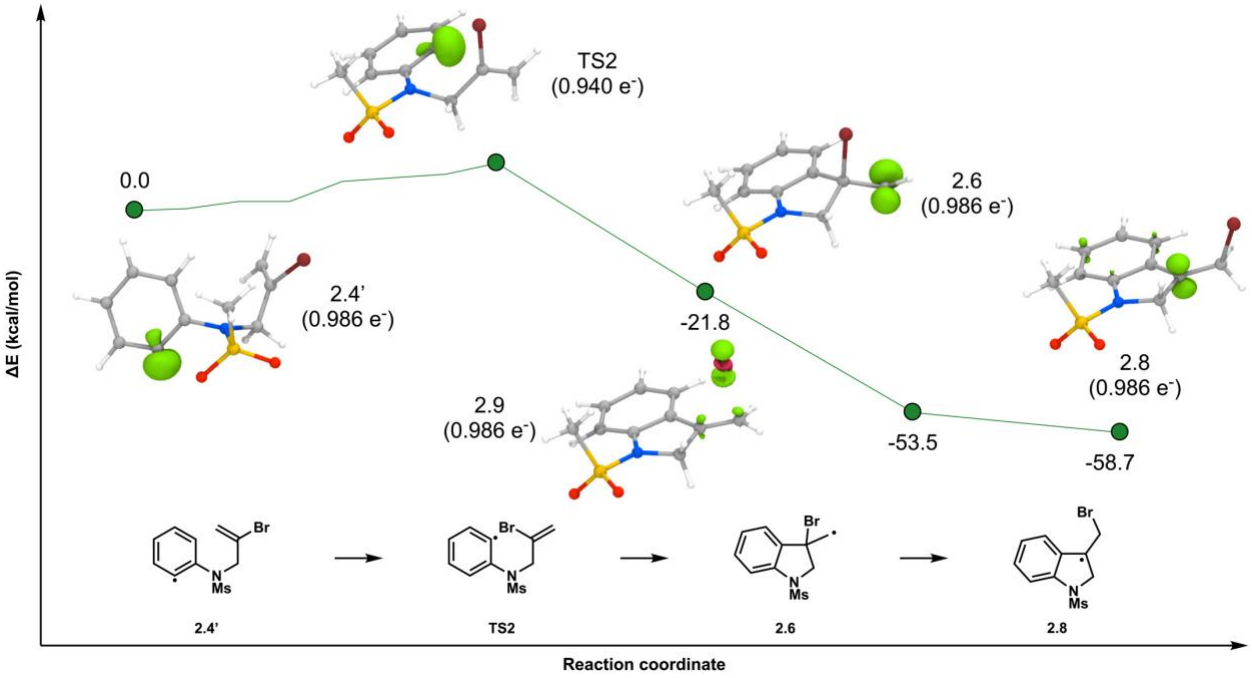

Figure 2.5.Energy profile from CDFT/MM GSM simulation for the cyclization step. 3D structures are the aryl radical (**4'**), TS (**TS2**), primary radical (**6**), 1,2-bromo shift (**9**), and tertiary radical (**8**) intermediates. QM region contained TDAE and substrate, but only transformations for substrate are shown for clarity. Green and orange reflect negative and positive spins respectively. Spins for selected atoms referenced in text are in parentheses. 2D structures highlight the transformation from the **4'** to **8**. All energies are in kcal/mol. Simulations were performed using B3LYP 6-31G*/CHARMM36 in a DMF solvent environment.

the TS, a transient structure analogous to **2.6** of Figure 2.2 was located, with a spin of 1.1 e^- on the external carbon. This intermediate immediately undergoes a 1,2-bromo shift (**2.9**) to give **2.8**. At this point, delocalization of benzylic radical within the ring provides significant electronic stabilization of the radical and the product of this step (**2.4'** to **2.8**) is reached.

Following formation of tertiary radical intermediate **2.8**, bromine extraction and tautomerization to the corresponding indole **2.3a** are the expected next steps. These steps were found to be favorable by -26.2 and -40.0 kcal mol, respectively (full details of these steps, performed using a QM model, are in section 2.7.2). The full, simulated reaction mechanism is provided in Figure 2.6, which illustrates an update to the

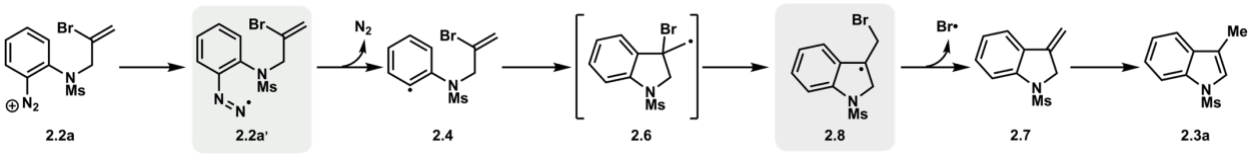

Figure 2.6. Proposed mechanism for formation of indole **2.3a** based on CDFT/MM results. The highlighted regions (boxes) show the formation of the aryldiazenyl radical intermediate **2.2a'** after initial SET and a new tertiary radical intermediate (**2.8**) that is formed following a transient primary radical intermediate (**2.6**).

mechanism proposed in Figure 2.2. Specifically, the mode of activation by SET (**2.2a'**) and introducing the pathway **2.4'** to **2.8** as an unexpected concerted, asynchronous step for ring closure are detailed.

2.4 SET-initiated Alcohol Synthesis

Tetrathifulvalene (TTF, **2.10**) initiates radical-polar crossover reactions between the aryldiazonium salts **2.11a**-**c** and water to produce 2,3-dihydrobenzofuranyl alcohol products (**2.17b-c**, Figure 2.7).[93] Like the TDAE system, the proposed mechanism begins by SET-initiated loss of N_2 from **2.10** to **2.11** to produce oxidized TTF (**2.12**) and an aryl radical intermediate (**2.13**).[94] **2.13** undergoes 5-*exo-trig* cyclization to give a primary radical intermediate (**2.14**). At this step the mechanism diverges from the TDAE pathway, and a radical coupling between **2.12** and **2.14** occurs to form a sulfonium intermediate (**2.15**). Rather than termination of the radical via a radical leaving group (bromine radical in TDAE system), a two-electron elimination and regeneration of **2.10** would be needed to be in line with experimental observations of catalytic TTF.[94] To achieve this, it was proposed that an intramolecular cyclization occurs to form an oxocarbenium intermediate **2.16**, which would then undergo S_N1 solvolysis to form the alcohol product, **2.17**. However, formation of the sulfonium intermediate **2.15a** and not the alcohol **2.17a** were observed, which was ascribed to the instability of a putative primary carbocation. To better

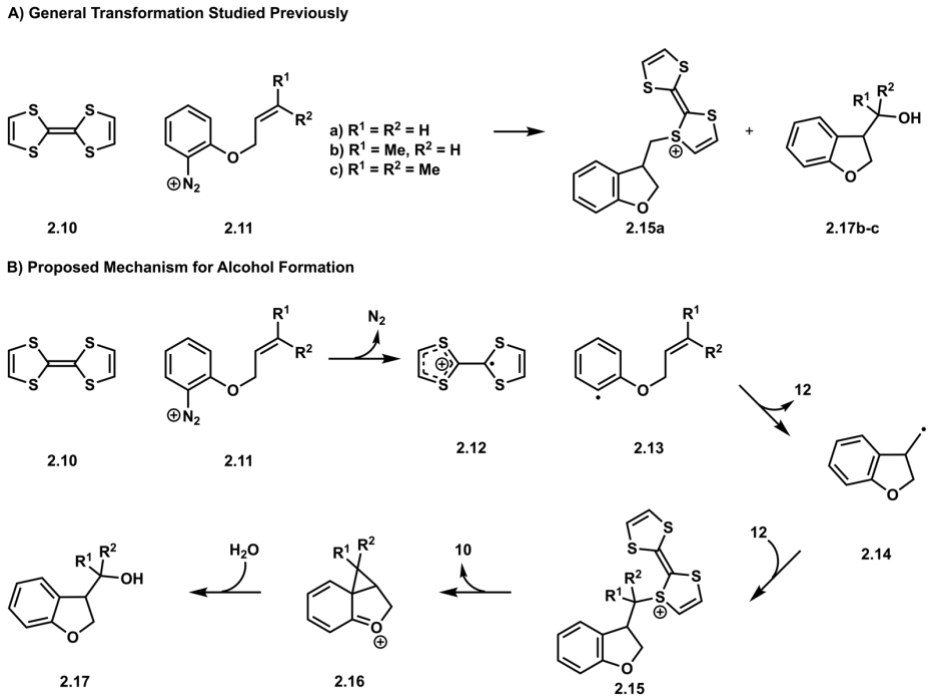

Figure 2 7. Overview of TTF (**2.10**) system. A) Transformation studied by Murphy[94] and coworkers involving **2.10** and the aryldiazonium salts **2.11a-c**. B) Previously proposed mechanisms for SET-initiated formation of sulfonium intermediate **2.15a** and the secondary and tertiary 2,3-dihydrobenzofuranyl alcohols **2.17b** and **2.17c**.

understand this mechanism, CDFT/MM was therefore applied to the elementary steps leading to the formation of 2,3-dihydrobenzofuranyl alcohol products.

Simulations of the reaction mechanism began from substrate **2.11a** (Figure 2.8). After SET, **2.12** and an aryldiazenyl radical (**2.11a'**) intermediate were formed, similar to the activation pathway for the TDAE reaction (Figure 2.3). SET activates the substrate for loss of N_2 by not only decreasing the C–N–N bond angle from 178.8° to 127.1°, but also lengthening the C–N bond from 1.36 Å to 1.44 Å, and the N–N bond from 1.13 Å to 1.19 Å. This step was also calculated to be strongly thermodynamically favored by -32.2 kcal/mol. After activation the aryl radical intermediate (**2.13a**) forms through loss of N_2 with a barrier of 11.5 kcal/mol (**TS3**), with an overall energetic cost of 6.4 kcal/mol. As with the reaction of Figure 2.3, this barrier agrees with prior experimental measurements.[96-98] Although **2.13a** appears to be unfavorable, bubbling out of N_2 from

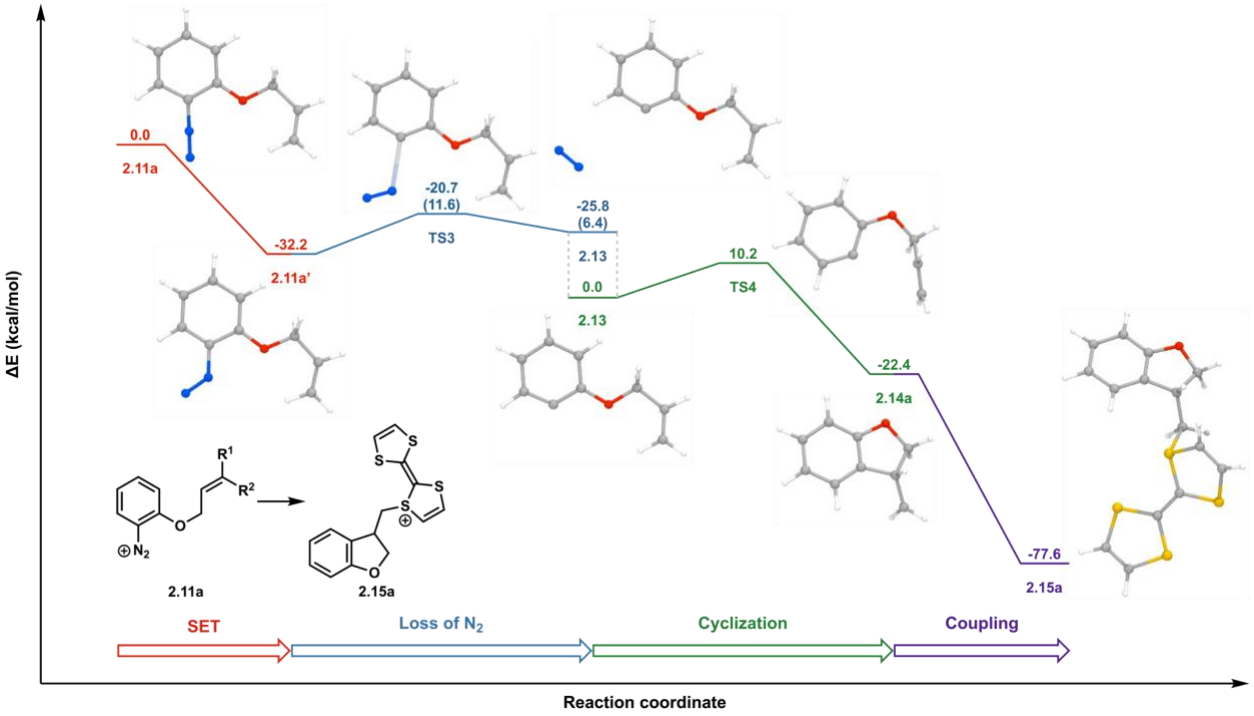

Figure 2.8. QM/MM reaction profile of the TTF system showing overall thermodynamic favorability for sulfonium intermediate (**2.15a**) formation. Energies in parentheses are relative to the reaction step, and all energies are in kcal/mol. Only transformations for substrate are shown for clarity. The gray dashed line represents removal of N_2 from the model. All simulations were performed using B3LYP 6-31G*/CHARMM36 in a DMF solvent environment.

solution (**2.13a** to **2.13a'**) makes this step effectively irreversible.

After SET and N_2 loss, 5-exo-trig cyclization leads to primary radical intermediate **2.14a**, which then goes on to couple with oxidized TTF (**2.12**). This transformation has a strong thermodynamic driving force of -22.2 kcal/mol, and a barrier of 10.4 kcal/mol (**TS4**). This barrier mostly reflects the change to a more reactive conformer to achieve proper alignment for C–C bond formation, with an associated energy cost due to ring strain. Radical coupling between **2.12** and **2.14a** then produces the sulfonium intermediate **2.15a**. This C–S coupling was previously assumed to occur over C–C coupling (based on semi-empirical electronic structure models[94]) and is confirmed (see section 2.7.5) via QM/MM simulations that found barrierless, strongly downhill C–S coupling.

From sulfonium intermediate **2.15**, pathways for the generation of oxocarbenium intermediate (**2.16**) were investigated. Based on chemical intuition and experimental observations, there should be a relationship between substitution at the C_1-carbon and stability. A trend was indeed observed between the energy of activation (E_a) and substitution for the formation of **2.16a-c** (Figure 2.9): as the substitution on **2.15a-c** increased; the barrier decreased. As expected, the barrier for formation of unsubstituted **2.16a** was inaccessible at 26.9 kcal/mol (**TS5a**), which supports experimental observations in that only **2.15a** could be isolated.[95] The barriers for **2.16b** and **2.16c** were accessible at 21.8 kcal/mol (**TS5b**) and 15.6 kcal/mol (**TS5c**), respectively. There was a relationship between the initial C_1–S bond length and E_a (Figure 2.9), and also the C_1–C_2 **2.15a**-**c** and **2.16a**-**c** (see section 2.7.6). Stabilization of the positive charge on the sulfur

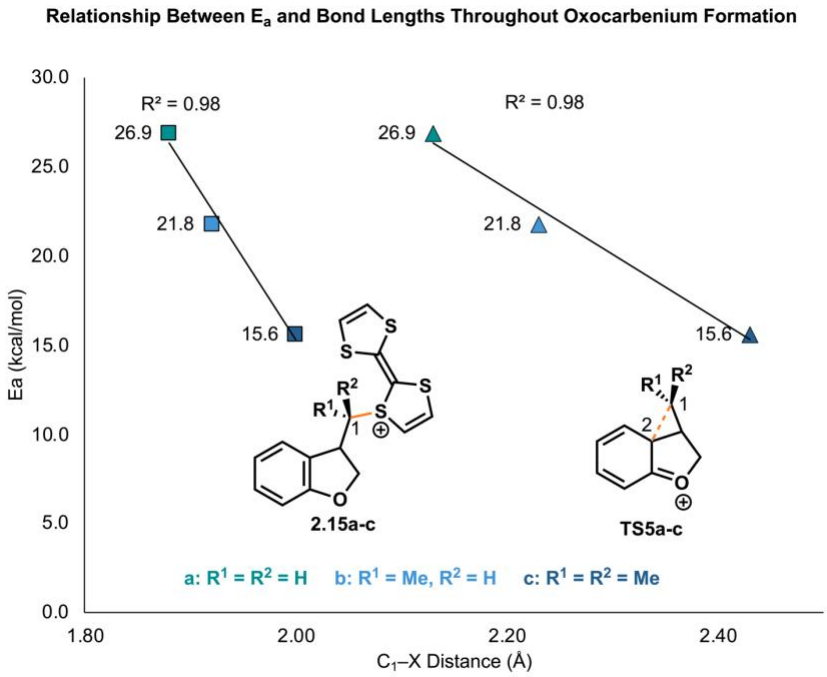

Figure 2.9. Relationship between E_a and the C_1–X bond lengths. **2.15a-c** C_1–S bond lengths (left, squares) and the C_1–C_2 bond lengths for **TS5a-c** (right, triangles) both have a correlation of $R^2 = 0.98$ when compared to E_a. All energies are in kcal/mol.

bond length and E_a at the TS structures (**TS5a**-**c**, Figure 2.9) as well as for intermediates or carbon atoms, as well as steric repulsions between the methyl groups and the substrate, result in lengthening of the C_1–S and C_1–C_2 bond lengths. This lengthening corresponds to weaker bonds and substate destabilization that lower the reaction barriers as substitution increases. Thermodynamically, **2.16b** and **2.16c** are -7.8 kcal/mol and -4.6 kcal/mol, respectively, compared to species **2.15b** and **2.15c**. **2.16a** was observed to be 8.8 kcal/mol uphill, making it thermodynamically disfavored.

The accessibility of **2.16b** and **2.16c** means that the formation of the alcohol products **2.17b** and **2.17c** is plausible via solvolysis. These simulations support the experimental hypothesis that the more substituted substrates proceed through an oxocarbenium intermediate prior to solvolysis. Figure 2.10 depicts the simulation-supported mechanism for forming secondary and tertiary alcohols (**2.17b**-**c**) from an initial SET between **2.10** and **2.11**. Here, one step along the path is qualitatively updated compared to the originally proposed mechanism: the aryldiazenyl radical intermediate is included.

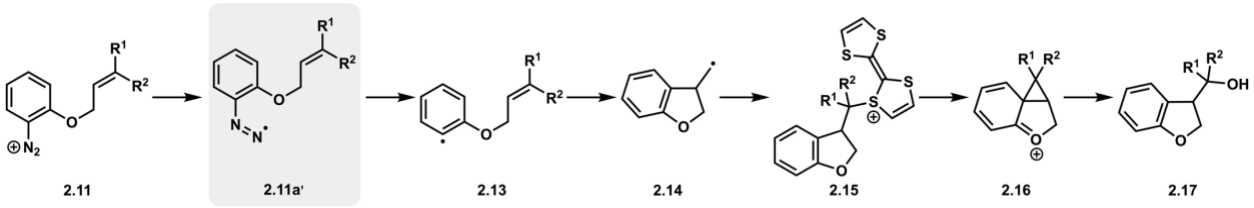

Figure 2.10. Proposed mechanism for formation of secondary and tertiary alcohols based on CDFT/MM. Highlighted region (grey box) shows the simulated aryldiazenyl radical intermediate **11a'** after initial SET.

2.5 Discussion

Simulations of the TDAE (**2.1**) and TTF (**2.10**) systems revealed the importance of solvent modeling in understanding these reactions. For instance, with TTF explicit solvent effects were strong contributors to the energies of key reaction steps. Separation of the QM/MM energies into the QM and MM components provides support to this assertion and gives insights into how either region is impacted by the electronic or geometric changes. The solvent environment (MM region) had a particular impact on the reaction profile for initiation of the reaction via SET. SET from TTF to substrate **2.11** resulted in a thermodynamic favorability of -26.1 kcal/mol, with -14 kcal/mol attributed to the QM region and -13 kcal/mol to the MM region. Following SET there was a loss of an initial H-bond between **2.10** and a solvent molecule due to the sulfur atoms becoming more positive following SET. The MM region then reorganized, leading to a better H-bonding network (see section 2.7.7). This observation highlights the importance of solvent rearrangements following SET, setting up the remaining steps of the reaction path for accurate simulation in the condensed phase solvent environment.

Sampling of the solvent via MD simulations during the ground and SET states also provided insights into the free energies involved with SET. The vertical energy gap was calculated from 2000 snapshots for the TDAE system. Figure 2.4 displays the range of vertical energies that were obtained and ranged within 50 kcal/mol for each electronic state. Free energies for the ground and SET states were obtained from these vertical energy gaps and mostly ranged within 2.0 kcal/mol. When the free energies are plotted against the vertical energy gap, two parabolic Marcus free energy curves are obtained, which give the thermodynamics and kinetics of electron transfer. Importantly, this plot

highlights the relative free energy between the two electronic states, and the activation free energy for the electron transfer, which are not possible to obtain accurately without CDFT/MM sampling.

2.6 Conclusions

CDFT/MM provided considerable insight into the reaction mechanism for two synthetically useful SET-initiated reactions. For the generation of indoles from aryldiazonium salts, CDFT/MM confirmed SET to be reversible and kinetically facile between TDAE (**1**) and the aryldiazonium substrate **2.2a**. The 5-*exo-trig* cyclization step was found to proceed through a previously unproposed intermediate, a tertiary radical (**2.8**) that is favored due to electronic delocalization. For the SET-initiated formation of 2,3-dihydrobenzofuranyl alcohols (**2.17b-c**), CDFT/MM gave atomistic insights into the relationship between lower barriers and increased substitution at the C_1-carbon due to combined electronic and steric effects. A quantitative description of the aryldiazenyl radical intermediates (**2.2a'**, **2.11a'**) was also obtained for both systems. These quantitative and qualitative results are essential for future reaction prediction and modeling of this class of OEDs.

The combined QM and MM treatment therefore captured the full effect of solvent on the reaction profiles. In doing so, the electronic changes between the donor and acceptor molecules and the surrounding solvent could be accurately described. For instance, in treating the solvent effects that stabilize the charge-separated intermediates of the reaction pathways, the thermodynamic impact of solvent was large, -13 kcal/mol attributed to due to the solvent effect. Further, MD sampling of the TDAE system could quantify the free energies for electron transfer, confirming the facility of SET in this

particular system. It will be interesting to study cases where SET is less favorable and use CDFT/MM to identify the root causes of the inhibition of such steps.

The new implementation of the CDFT/MM method connecting pDynamo with QChem extends the capabilities of CDFT into reaction pathway discovery in the presence of explicit solvent. This tool will be useful for various areas in which SET processes are important, such as biological simulations involving redox active enzymes.[22]

2.7 Supporting Information

Additional computational details previously described in section 2.2 include model preparation (2.2.1), MD simulations (2.2.2), free energy surfaces (2.2.3), QM/MM details (2.2.4), and pDynamo interface (2.2.5).

2.7.2 QM Model Details

In addition to the QM/MM models, simulations with gas-phase QM models were also performed (Figures S2.2 and S2.3). Geometry optimizations and frequency calculations for QM models were performed with QChem 5.0 using unrestricted B3LYP and the 6-311++G** basis set[73,-99] for all atoms except Br, which used the 6-311G** basis set[100]. All ground state geometries were confirmed to have no imaginary frequencies, and transition states were confirmed to have one imaginary frequency. Solvent corrections were performed with ORCA 4.2.1[101,102] using the SMD[103-105] implicit solvent mode and the cc-pVTZ[106-108] basis set. Reaction pathways were simulated using the single-ended growing string method.[109] All 3D images were generated using the VMD software program.[68]

2.7.3 QM Simulations of TDAE System

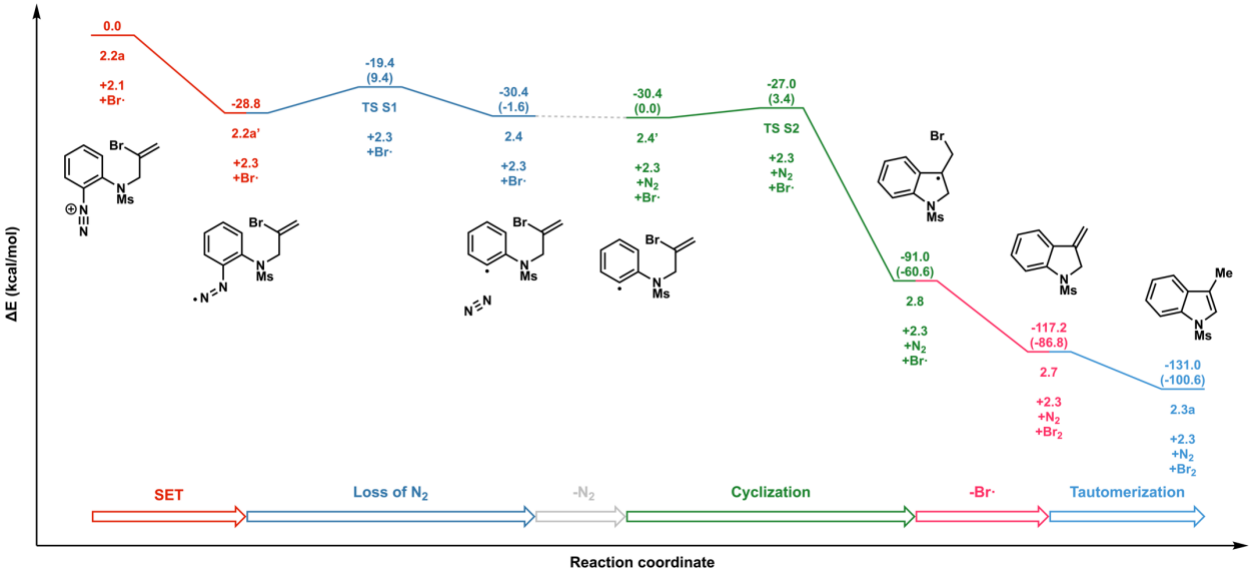

Figure 2S.1**.** QM Energy profile of TDAE (**2.1**) system for formation of indole product **2.3a**. Pathway proceeds through the same tertiary radical intermediate identified by QM/MM simulations. All energies are in kcal/mol.

The reaction profile of Figure S2.2 is largely the same as the QM/MM simulations (Figure 2.3). The aryldiazenyl radical intermediate (**2.2a'**) and the tertiary radical intermediate (**2.8**) therefore are stationary points at the QM as well as QM/MM levels. There is also a strong thermodynamic driving force of -40.0 kcal/mol for formation of indole with respect to the tertiary alcohol intermediate. Different from the QM/MM simulations, however, is minor partial bond formation at the cyclization TS (**TS7**).

2.7.4 QM Simulations of TTF System

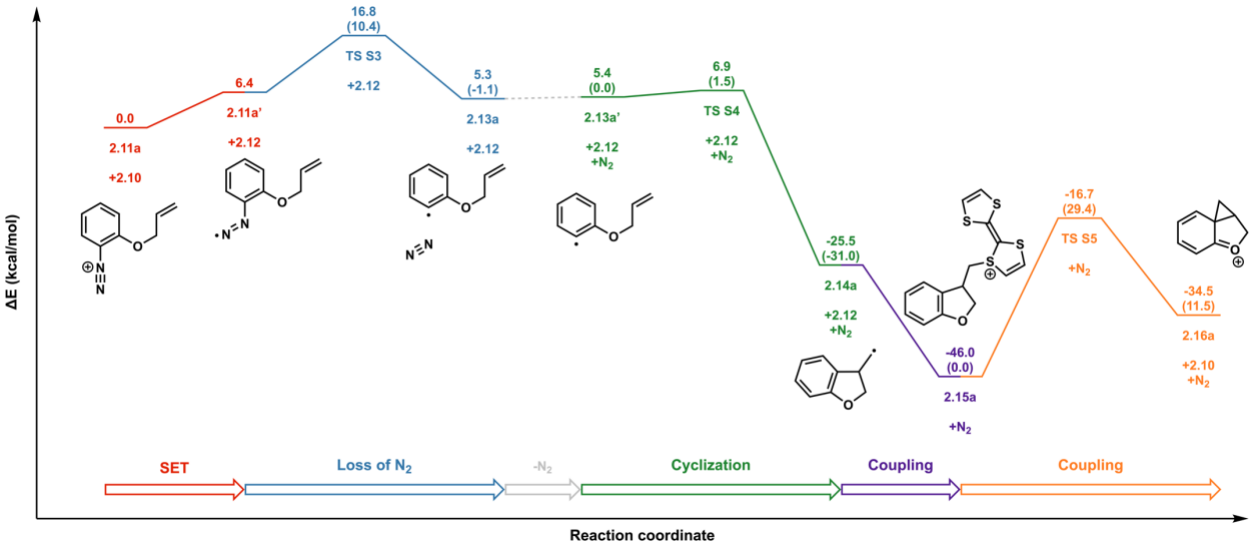

Figure 2S.2. QM Energy profile of TTF (**2.10**) system for formation of 2,3-dihydrobenzofuan (**2.16a**). Pathway proceeds through the same pathway identified by QM/MM simulations. Of importance is the inaccessible barrier of 29.4 kcal/mol during formation of the oxocarbenium intermediate **2.16a**. All energies are in kcal/mol.

As was observed in the TDAE system (section 2.7.3), the profile consists of the same intermediates and pathways as the QM/MM simulations (Figure 2.8). QM and QM/MM predict highly different thermodynamics for SET from **2.10** to the aryldiazonium substrate **2.11**. QM/MM simulations produced a strong thermodynamic favorability of -32.2 kcal/mol, whereas QM gives an uphill result, 6.4 kcal/mol. This result further highlights the need for including a solvent explicit environment to accurately model SET processes.

2.7.5 Spin Density Plot of $TTF^{\bullet+}$

A B3LYP cc-pVTZ gas phase single point calculation was performed on the B3LYP 6-311++G** optimized geometry of the oxidized TTF intermediate **12** in ORCA. The ORCA orca_plot module was used to obtain the CUBE files of spin densities.

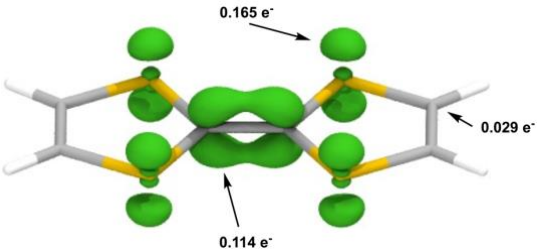

Figure 2S.3**.** Spin density plot of the oxidized TTF intermediate (**2.12**). Spin density is greatest on the sulfur atoms, followed by the internal carbon atoms, and least on the external carbon atoms.

Spin density results confirm the previous qualitative semi-empirical results. Spin density is greatest on the sulfur atoms (0.165 e^-), giving a slight preference for C–S coupling over C–C coupling.

2.7.6 Relationship Between Ea and Intermediate C–C Bond Lengths

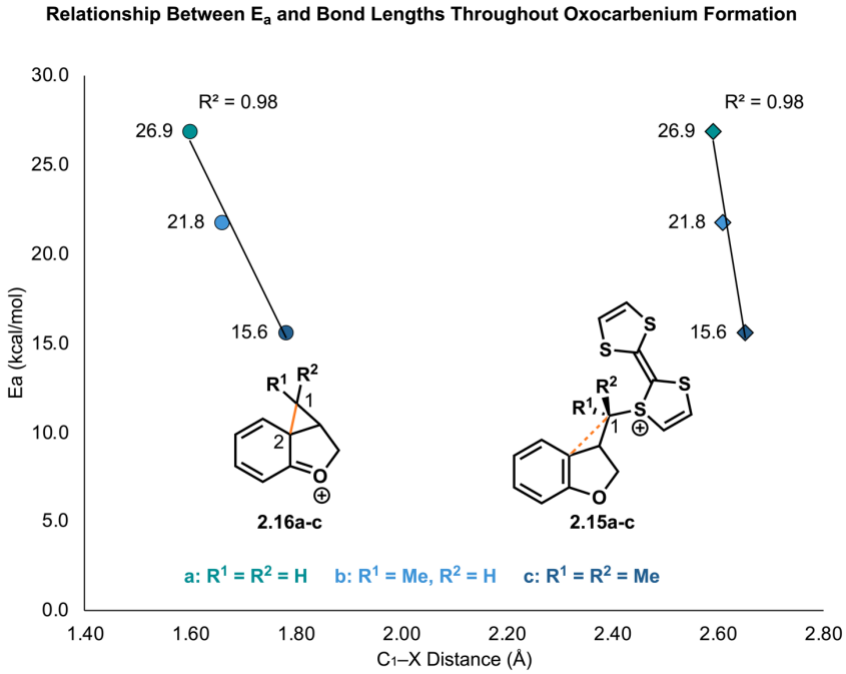

Figure 2S.4. Relationship between E_a and the C_1–C_2 bond lengths. C_1–C_2 bond lengths for the oxocarbenium intermediates **2.16a-c** (left, circles) and the coupled intermediates **2.15a-c** (right, diamonds) both have a correlation of $R^2 = 0.98$ when compared to E_a. All energies are in kcal/mol.

2.7.7 Solvent Reorganization Following SET for TTF System

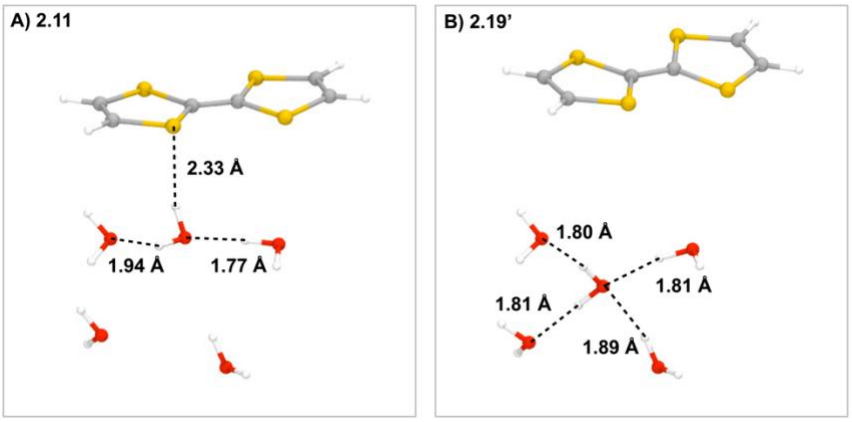

Figure 2S.5**.** Comparison of H-bond network of TTF (**2.10**) system following SET. A) before) and B) after SET. Solvent molecule does not H-bond with **2.11a'**, and instead forms a stronger H-bonding network with the surrounding solvent.

Table S2.1. Number of H-bonds between TTF (**10**) and the solvent, and relevant solvent-solvent H-bonds for the ground and SET states. At the cost of losing one H-bond, there is a gain of two H-bonds.

	11	**11a'**
TTF-Solvent	1	0
Solvent-Solvent	2	4

Following SET from TTF (**10**) to the aryldiazonium substrate **11a**, there is significant solvent reorganization. For example, a water molecule originally H-bonding to a sulfur atom of **10** no longer H-bonds with the oxidized TTF intermediate **12**. This is reasoned as being due to an increased positive charge. Overall, there is a thermodynamic gain the MM region of -13 kcal/mol, in part due to H-bond reorganization seen in Figure S4B.

2.8 References

[1] Twilton, J.; Le, C.; Zhang, P.; Shaw, M. H.; Evans, R. W.; Macmillan, D. W. C. The Merger of Transition Metal and Photocatalysis. *Nat. Rev. Chem.* **2017**, *1*, 0052.

[2] Romero, N. A.; Nicewicz, D. A. Organic Photoredox Catalysis. *Chem. Rev.* **2016**, *116*, 10075–10166.

[3] Ravelli, D.; Fagnoni, M.; Albini, A. Photoorganocatalysis. What For? *Chem. Soc. Rev.* **2013**, *42*, 97–113.

[4] Zhou, Q.-Q.; Zou, Y.-Q.; Lu, L.-Q.; Xiao, W.-J. Visible-Light-Induced Organic Photochemial Reactions Through Energy-Transfer Pathways. *Angew. Chem. Int. Ed.* **2019**, *58*, 1586–1604.

[5] Shaw, M. H.; Twilton, J.; Macmillan, D. W. C. Photoredox Catalysis in Organic Chemistry. *J. Org. Chem.* **2016**, *81*, 6898–6926

[6] Prier, C. K.; Rankic, D. A.; Macmillan, D. W. C. Visible Light Photoredox Catalysis With Transition Metal Complexes: Applications in Organic Synthesis. *Chem. Rev.* **2013**, *113*, 5322–5363.

[7] Xuan, J.; Xiao, W.-J. Visible-Light Photoredox Catalysis. *Angew. Chem. Int. Ed.* **2012**, *51*, 6828–6838.

[8] Narayanam, J. M. R.; Stephenson, C. R. J. Visible Light Photoredox Catalysis: Applications in Organic Synthesis. *Chem. Soc. Rev.* **2011**, *40*, 102–113.

[9] Sun, C.-L.; Shi, Z.-J. Transition-Metal-Free Coupling Reactions. *Chem. Rev.* **2014**, *114*, 9219–9280.

[10] Wang, X.; Studer, A. Iodine(III) Reagents in Radical Chemistry. *Acc. Chem. Res.* **2017**, *50*, 1712–1724.

[11] Broggi, J.; Terme, T.; Vanelle, P. Organic Electron Donors as Powerful Single-electron Reducing Agents in Organic Synthesis. *Angew. Chem. Int. Ed.* **2014**, *53*, 384–413.

[12] Doni, E.; Murphy, J. A. Evolution of Neutral Organic Super-Electron-Donors and Their Applications. *Chem. Comm.* **2014**, *50*, 6073–6087.

[13] Murphy, J. A. Discovery and Development of Organic Super-Electron-Donors. *J. Org. Chem.* **2014**, *79*, 3731–3746.

[14] Beckwith, A. L. J. Regio-selectivity and Setero-selectiviy in Radical Reactions. *Tetrahedron*, **1981**, *37*, 3073–3100.

[15] Blumberger, J. Recent Advances in the Theory and Molecular Simulation of Biological Electron Transfer Reactions. *Chem. Rev.* **2015**, *115*, 11191–11238.

[16] Marcus, R. A. On The Theory of Oxidation-Reduction Reactions Involving Electron Transfer. *J. Chem. Phys.* **1956**, *24*, 966–978.

[17] Van Voorhis, T.; Kowalczyk, T.; Kaduk, B.; Wang, L.-P.; Cheng, C.-L.; Wu, Q. The Diabatic Picture of Electron Transfer, Reaction Barriers, and Molecular Dynamics. *Annu. Rev. Phys. Chem.* **2010**, *61*, 149–170.

[18] López-Estrada, O.; Laguna, H. G.; Barrueta-Flores, C.; Amador-Bedolla, C. Reassessment of the Four-Point Approach to the Electron-transfer Marcus-Hush Theory. *ACS Omega*, **2018**, *3*, 2130–2140.

[19] Nelsen, S. F.; Blackstock, S. C.; Kim, Y. Estimation of Inner Shell Marcus Terms for Amino Nitrogen Compounds by Molecular Orbital Calculations. *J. Am. Chem. Soc.* **1987**, *109*, 677–682.

[20] Kaduk, B.; Kowalczyk, T.; Van Voorhis, T. Constrained Density Functional Theory. *Chem. Rev.* **2012**, *112*, 321–370.

[21] Dederichs, P. H.; Blügel, S.; Zeller, R.; Akai, H. Ground States of Constrained Systems: Application To Cerium Impurities. *Phys. Rev. Lett.* **1984**, *53*, 2512–2515.

[22] Wu, Q.; Van Voorhis, T. Direct Optimization Method to Study Constrained Systems Within Density-Functional Theory. *Phys. Rev. A* **2005**, *72*, 024502.

[23] Ku, C.; Sit, P. H.-L. Oxidation-State Constrained Density Functional Theory for the Study of Electron-transfer Reactions. *J. Chem. Theory Comput.* **2019**, *15*, 4781–4789.

[24] Carmona-Espínfola, J.; Núñez-Rojas, E.; García-Melgarejo, V.; Gázquez, J. L.; Alejandre, J. Constrained Dipole Moment Density Functional Theory for Charge Distributions in Force Fields for the Study of Molecular Fluids. *J. Chem. Phys.* **2020**, *152*, 124116.

[25] Plaisance, C. P.; Van Santen, R. A.; Reuter, K. Constrained-Orbital Density Functional Theory. Computational Method and Applications to Surface Chemical Processes. *J. Chem. Theory Comput.* **2017**, *13*, 3561–3574.

[26] Ramos, P; Pavanello, M. Constrained Subsystem Density Functional Theory. *Phys. Chem. Chem. Phys.* **2016**, *18*, 21172–21178.

[27] Ziegler, T.; Krykunov, M. On the Calculation of Charge Transfer Transitions With Standard Density Functionals Using Constrained Variational Density Functional Theory. *J. Chem. Phys.* **2020**, *133*, 074104.

[28] Evangelista, F. A. Shushkov, P.; Tully, J. C. Orthogonality Constrained Density Functional Theory for Electronic Excited States. *J. Phys. Chem. A* **2013**, *117* 7378–7392.

[29] Oberhofer, H.; Blumberger, J. Electronic Coupling Matrix Elements From Charge Constrained Density Functional Theory Calculations Using a Plane Wave Basis Set. *J. Chem. Phys.* **2010**, *133*, 244105.

[30] Guo, Z.; Liu, Y.; Zhang, X. Constrained Lattice Density Functional Theory and Its Applications On Vapor–Liquid Nucleations. *Sci. Bull.* **2015,** *60*, 320–327.

[31] Wu, Q.; Van Vorhis, T. Extracting Electron Transfer Coupling Elements From Constrained Density Functional Theory. *J. Chem. Phys.* **2006**, *125*, 164105.

[32] Kaduk, B.; Tsuchimochi, T.; Van Voorhis, T. Analytic Energy Gradients for Constrained DFT-Configuration Interaction. *J. Chem. Phys.* **2014**, *140*, 18A503.

[33] Kubas, A.; Hoffmann, F.; Heck, A.; Oberhofer, H.; Elstner, M.; Blumberger, J. Electronic Couplings for Molecular Charge Transfer: Benchmarking CDFT, FODFT, and FODBTB Against High-Level *Ab Initio* Calculations. *J. Chem. Phys.* **2014**, *140*, 104105.

[34] Kubas, A.; Gajdos, F.; Heck, A.; Oberhofer, H.; Elstner, M.; Blumberger, J. Electronic Couplings For Molecular Charge Transfer: Benchmarking CDFT, FODFT And FODFTB Against High-Level *Ab Initio* Calculations. II. *Phys. Chem. Chem. Phys.* **2015**, *17*, 14342–14354.

[35] Kim, H.; Goodson, T., III; Zimmerman, P. M. Density Functional Physicality in Electronic Coupling Estimation: Benchmarks and Error Analysis. *J. Phys. Chem. Lett.* **2017**, *8*, 3242–3248.

[36] Joutsuka, T.; Ando, K. Constrained Density Functional Theory Molecular Dynamics Simulation of Deprotonation in Aqueous Silic Acid. *J. Phys. Chem. B* **2020**, *124*, 8323–8330.

[37] Krenz, M.; Gerstmann, U.; Schmidt, W. G. Photochemical Ring Opening of Oxirane Modeled by Constrained Density Functional Theory. *ACS Omega*, **2020**, *5*, 24057–24063.

[38] Wu, Q.; Kaduk, B.; Van Voorhis, T. Constrained Density Functional Theory Based Configuration Interaction Improves the Prediction of Reaction Barrier Heights. *J. Chem. Phys.* **2009**, *130*, 034109.

[39] Aierken, Y.; Agrawal, A.; Sun, M.; Melander, M.; Crumlin, E. J.; Helms, B. A.; Pendergast, D. Revealing Charge-transfer Dynamics at Electrified Sulfur Cathodes Using Constrained Density Functional Theory. *J. Phys. Chem. Lett.* **2021**, *12*, 739–744.

[40] Goldey, M. B.; Brawand, N. P.; Vörös, M.; Giulia, G. Charge Transport in Nanostructured Materials: Implementation and Verification of Constrained Density Functional Theory. *J. Chem. Theory Comput.* **2017**, *13*, 2581–2590.

[41] Řezáč, J.; De La Lande, A. On the Role of Charge Transfer in Halogen Bonding. *Phys. Chem. Chem. Phys.* **2017**, *19*, 791–803.

[42] Roychoudhury, S.; Motta, C.; Sanvito, S. Charge Transfer Energies of Benzene Physisorbed on a Graphene Sheet From Constrained Density Functional Theory. *Phys. Rev. B* **2016**, *93*, 045130.

[43] Difley, S.; Van Voorhis, T. Exciton/Charge-transfer Electronic Couplings in Organic Semiconductors. *J. Chem. Theory Comput.* **2011**, *7*, 594–601.

[44] De La Lande, A.; Salahub, D. R. Derivation of Interpretative Models for Long Range Electron Transfer From Constrained Density Functional Theory. *THEOCHEM* **2010**, *943*, 115-120.

[45] Wu, Q.; Van Voorhis, T. Constrained Density Functional Theory and Its Application in Long-Range Electron Transfer. *J. Chem. Theory Comput.* **2006**, *2*, 765–774.

[46] Phillips, J. J.; Peralta, J. E. Magnetic Exchange Couplings From Constrained Density Functional Theory: An Efficient Approach Utilizing Analytic Derivatives. *J. Chem. Phys.* **2011**, *135*, 184108.

[47] Rudra, I.; Wu, Q.; Van Voorhis, T. Predicting Exchange Coupling Constants in Frustrated Molecular Magnets Using Density Functional Theory. *Inorg. Chem.* **2007**, *46*, 10539–10548.

[48] Rudra, I.; Wu, Q.; Van Voorhis, T. Accurate Magnetic Exchange Couplings in Transition-Metal Complexes From Constrained Density-Functional Theory. *J. Chem. Phys.* **2006**, *124*, 024103.

[49] Yeganeh, S.; Van Voorhis, T. Triplet Excitation Energy Transfer With Constrained Density Functional Theory. *J. Phys. Chem. C* **2010**, *114*, 20756–20763.

[50] Kaduk, B.; Van Voorhis, T. Communication: Conical Intersections Using Constrained Density Functional Theory–Configuration Interaction. *J. Chem. Phys.* **2010**, *133*, 061102.

[51] Otsuka, T.; Sumita, M.; Izawa, H.; Morihashi, K. Theoretical Study on Photo-Induced Processes of 1-Methyl-3-(N-(1,8-Naphthalimidyl)Ethyl)Imidazolium Halide Species: An Application of Constrained Density Functional Theory. *Phys. Chem. Chem. Phys.* **2018**, *20*, 3911–3917.

[52] Otsuka, T.; Sumita, M.; Izawa, H.; Morihashi, K. Intermolecular Electron Transfer States of 1-Methyl-3-(N-(1,8-Naphthalimidyl)Ethyl)Imidazoliumiodide Obtained by Constrained Density Functional Theory. *Phys. Chem. Chem. Phys.* **2016**, *18*, 17795–17798.

[53] Derricotte, W. D.; Evangelista, F. A. Simulation of X-Ray Absorption Spectra With Orthogonality Constrained Density Functional Theory. *Phys. Chem. Chem. Phys.* **2015**, *17*, 14360–14374.

[54] Wu, Q.; Van Voorhis, T. Direct Calculation of Electron Transfer Parameters Through Constrained Density Functional Theory. *J. Phys. Chem. A* **2006**, *110*, 9212–9218.

[55] Kowalczyk, T.; Wang, L.-P.; Van Voorhis, T. Simulation of Solution Phase Electron Transfer in a Compact Donor–Acceptor Dyad. *J. Phys. Chem. B* **2011**, *115*, 12135–12144.

[56] Difley, S.; Lee-Ping, W; Yeganeh, S.; Yost, S. R.; Van Voorhis, T. Electronic Properties of Disordered Organic Semiconductors via QM/MM Simulations. *Acc. Chem. Res.* **2010**, *43*, 995–1004.

[57] Oberhofer, H.; Blumberger, J. Charge Constrained Density Functional Molecular Dynamics for Simulation of Condensed Phase Electron Transfer Reactions. *J. Chem. Phys.* **2009**, *131*, 064101.

[58] Holmberg, N.; Laasonen, K. Efficient Constrained Density Functional Theory Implementation for Simulation of Condensed Phase Electron Transfer Reactions. *J. Chem. Theory Comput.* **2017**, *13*, 587–601.

[59] Řezáč, J.; Lévy, B.; Demachy, I.; De La Lande, A. Robust and Efficient Constrained DFT Molecular Dynamics Approach for Biochemical Modeling. *J. Chem. Theory Comput.* **2012**, *8*, 418–427.

[60] De La Lande, A.; Alvarez-Ibarra, A.; Hasnaoui. K.; Cailliez, F.; Wu, X.; Mineva, T.; Cuny, J.; Calaminici, P.; López-Sosa, L.; Geudtner, G.; Navizet, I.; Iriepa, C. G.; Salahub, D. R.; Köster, A. M. Molecular Simulations With in-Demon2k QM/MM, a Tutorial-Review. *Molecules* **2019**, *24*, 1653.

[61] Gillet, N.; Lévy, B.; Moliner, V.; Demachy, I.; De L Lande, A. Electron and Hydrogen Atom Transfers in the Hydride Carrier Protein Emob. *J. Chem. Theory Comput.* **2014**, *10*, 5036–5046.

[62] Cailliez, F.; Müller, P.; Firmino, T.; Pernot, P.; De La Lande, A. Energetics of Photoinduced Charge Migration Within the Tryptophan Tetrad of an Animal (6−4) Photolyase. *J. Am. Chem. Soc.* **2016**, *138*, 1904–1915.

[63] Firmino, T.; Mangaud, E.; Cailliez, F.; Devolder, A.; Mendive-Tapia, D.; Gatti, F.; Meier, C.; Desouter-Lecomte, M.; De La Lande, A. Quantum Effects in Ultrafast Electron Transfers Within Cryptochromes. *Phys. Chem. Chem. Phys.* **2016**, *18*, 21442–21457.

[64] Brunk, E.; Rothlisberger, U. Mixed Quantum Mechanical/Molecular Mechanical Molecular Dynamics Simulations of Biological Systems in Ground and Electronically Excited States. *Chem. Rev.* **2015**, *115*, 6217–6263.

[65] Chung. L. W.; Sameera, W. M. C.; Ramozzi, R.; Page, A. J.; Hatanaka, M.; Petrova, G. P.; Harris, T. V.; Li, X.; Ke, Z.; Liu, F.; Li, H.-B.; Ding, L.; Morokuma, K. The ONIOM Method and Its Applications. *Chem. Rev.* **2015**, *115*, 5678–5796.

[66] De La Lande, A.; Gillet, N.; Chen, S.; Salahub, D. R. Progress and Challenges in Simulating and Understanding Electron Transfer In Proteins. *Arch. Biochem. Biophys.* **2015**, *582*, 28–41.

[67] Van Der Kamp, M. W.; Mullholland, A. J.; Combined Quantum Mechanics/Molecular Mechanics (QM/MM) Methods in Computational Enzymology. *Biochemistry* **2013**, *52*, 2708–2728.

[68] Humphrey, W.; Dalke, A; Schulten, K., VMD - Visual Molecular Dynamics. *J. Molec. Graphics* **1996**, *14*, 33–38.

[69] Brooks, B. R.; Brooks III, C. L.; Mackkerell Jr, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, M.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Pos, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcosk, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545–1614.

[70] Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell Jr, A. D. CHARMM General Force Field: A Force Field for Drug-Like Molecules Compatile With the CHARMM All-Atom Additive Biological Force Field. *J. Comput. Chem.* **2010**, *31*, 671–690.

[71] Yu, W.; He, X.; Vanommeslaeghe, K.; Mackerell Jr, A. D. Extension of The CHARMM General Force Field to Sulfonyl-Containing Compounds and Its Utility in Biomolecular Simulations. *J. Comput. Chem.* **2012**, *33*, 2451–2468.

[72] Slater, J. C. A Simplification of The Hartree-Fock Method. *Phys. Rev.* **1951**, *81*, 385–390.

[73] Ditchfield, R.; Hehre, W. J.; Pople, J. A. Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1971**, *54*, 724–728.

[74] Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; Defrees, D. J.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XXIII. A Polarization-Type Basis Set for Second-Row Elements. *J. Chem. Phys.* **1982**, *77*, 3654–3665.

[75] Gordon, M. S.; Binkley, J. S.; Pople, John A.; Pietro, William J.; Hehre, W. J. Self-Consistent Molecular-Orbital Methods. 22. Small Split-Valence Basis Sets for Second-Row Elements. J. Am. Chem. Soc. **1982**, *104*, 2797–2803.

[76] Hariharan, P. C.; Pople, J. A. The Influence of Polarization Functions On Molecular Orbital Hydrogenation Energies. Theor. Chim. Acta **1973**, *28*, 213–222.

[77] Hehre, W. J.; Ditchfield, R.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. J. Chem. Phys. **1972**, *56*, 2257–2261.

[78] Rassolov, Vitaly A.; Ratner, Mark A.; Pople, John A., Redfern, Paul C., Curtiss, Larry A. 6-31G* Basis Set for Third-Row Atoms. J. Comput. Chem. **2001**, *22*, 976–984.

[79] Jo, S.; Iyer, V. G.; Im, W. CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM. *J. Comput. Chem.* **2008**, *29*, 1859–1865.

[80] Martinez, L.; Andrade, R.; Birgin, E. G.; Martínez, J. M. Packmol: A Package for Building Initial Configurations for Molecular Dynamics Simulations. *J. Comput. Chem.* **2009**, *30*, 2157–2164.

[81] Huang, J.; Mackerell Jr, A. D., CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *J. Comput. Chem.* **2013,** *34*, 2135–2145.

[82] Eastman, P.; Swails, J.; Chodera, J. D.; Mcgibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L.-P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; Wiewiora, R. P.; Brooks, B. R.; Pande, V. S. Openmm 7: Rapid Development of High Performance Algorithms for Molecular Dynamics. *Plos Comput. Biol.* **2017**, *13*, E1005659.

[83] Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

[84] Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-energy Formula Into a Functional of the Electron Density. *Phys. Rev. B* **1998**, *37*, 785–789.

[85] Field, M. J. The Pdynamo Program for Molecular Simulations Using Hybrid Quantum Chemical and Molecular Potentials. *J. Chem. Theory Comput.* **2008**, *4*, 1151–1161.
[86] Neese, F., The ORCA Program System. *Wires Comput. Molec. Sci.* **2012**, *2*, 73–78.

[87] Epifanovsky, E.; et. al. Software for the Frontiers of Quantum Chemistry: An Overview of Developments in the Q-Chem 5 Package. *J. Chem. Phys.* **2021**, *155*, 084801.

[88] O'Boyle, N. M.; Tenderholt, A. L.; Langer, K. M. Cclib: A Library for Package-Independent Computational Chemistry Algorithms. *J. Comp. Chem.* **2008**, *29*, 839–845.

[89] Mahesh, M.; Murphy, J. A.; Lestrat, F.; Wessel, H. P. Reduction Of Arendiazonium Salts by Tetrakis(Dimethylamino)Ethylene (TDAE): Efficient Formation Of Products Derived From Aryl Radicals. *Beilstein J. Org. Chem.* **2009**, *5*, 1.

[90] Abdulla, H. O.; Scaringi, S.; Amin, A. A.; Mella, M.; Protti, S.; Fagoni, M. Aryldiazenyl Radical From Arylazo Sulfones: Visible Loght-Driven Diazenylation of Enol Silyl Ethers. *Adv. Synth. Catal.* **2020**, *362*, 2150–2154.

[91] Daasbjerg, K.; Sehested, K. Reduction of Substituted Benzenediazonium Salts by Solvated Electrons in Aqueous Neutral Solution Studied by Pulse Radiolysis. *J. Phys. Chem. A* **2002**, *106*, 11098–11106.

[92] Tadashi, S.; Seiichi, M.; Ryuichi, N.; Masamichi, T.; Atsushi, M.; Akemi, K.; Munehiro, D. Decay Reaction of Aryldiazenyl Radicals in Solution. *Bull. Chem. Soc. Jpn.* **1987**, *60*, 3321–3330.

[93] Lampard, C.; Murphy, J. A.; Lewis, N. Tetrathiafulvalene as a Catalyst for Radical–Polar Crossover Reactions. *J. Chem. Soc., Chem. Commun.* **1993**, 295–297.

[94] Zahradnik, R.; Carsky, P.; Hunig, S.; Kieslich, G.; Scheutzoy, D. Conjugated Radicals. VII* Tetrathiofulvalene and a Note on Sulfur-Containing Conjugated Radicals. *Int. J. Sulfur Chem. Part C* **1971**, *6*, 109–122.

[95] Callaghan, O.; Franck, X.; Murphy, J. A. Unprecedented Electrophilic Behaviour of Tetrathiafulvalenium Salts. *Chem. Commun.* **1997**, 1923–1924.

[96] Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. R. Efficient Diffuse Function-Augmented Basis Sets for Anion Calculations. III. The 3-21+G Basis Set for First-Row Elements, Li-F. *J. Comput. Chem.* **1983**, *4*, 294–301.

[97] Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XX. A Basis Set for Correlated Wave Functions. *J. Chem. Phys.* **1980**, *72*, 650–654.

[98] Mclean, A. D.; Chandler, G. S. Contracted Gaussian Basis Sets for Molecular Calculations. I. Second Row Atoms, Z=11-18. *J. Chem. Phys.* **1980**, *72*, 5639–5648.

[99] Spitznagel, G. W.; Clark, T.; Schleyer, P. V. R.; Hehre, W. J. An Evaluation of The Performance of Diffuse Function-Augmented Basis Sets for Second Row Elements, Na-Cl. *J. Comput. Chem.* **1987**, 8, 1109–1116.

[100] Curtiss, L. A.; Mcgrath, M. P.; Blaudeau, J.-P.; Davis, N. E.; Binning, R. C.; Radom, L. Extension of Gaussian-2 Theory to Molecules Containing Third-Row Atoms Ga-Kr. *J. Chem. Phys.* **1995**, *103*, 6104–6113.

[101] Neese, F., The ORCA Program System. *Wires Comput. Molec. Sci.* **2012**, *2*, 73–78.

[102] Neese, F. Software Update: The ORCA Program System, Version 4.0. *Wires Comput. Molec. Sci.* **2017**, *8*, E1327.

[103] Kendall, R. A.; Dunning, Jr., T.H.; Harrison, R. J. Electron Affinities of the First-Row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys.* **1992**, *96*, 6796–6806.

[104] Woon, D. E.; Dunning, Jr., T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. III. The Atoms Aluminum Through Argon. *J. Chem. Phys.* **1993**, *98*, 1358–1371.

[105] Woon, D. E.; Dunning, Jr., T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. IV. Calculation of the Static Electrical Response Properties. *J. Chem. Phys.* **1994**, *100*, 2975–2988.

[106] Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron Through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

[107] Wilson, A. K.; Woon, D. E.; Peterson, K. A.; Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. IX. The Atoms Gallium Through Krypton. *J. Chem. Phys.* **1999**, *110*, 7667–7676.

[108] Woon, D. E.; Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. III. The Atoms Aluminum Through Argon. *J. Chem. Phys.* **1993**, *98*, 1358–1371.

[109] Zimmerman, P. M. Single-Ended Transition State Finding With the Growing String Method. *J. Comput. Chem.* **2015**, *36*, 601–611.

Chapter III

Quantum Chemical Insights into the Ring Expansion Catalyzed by TropC

Portions of this chapter appear in "Radical Tropolone Biosynthesis" by Doyon, T. J.; Skinner, K. C.; et. al. ChemRxiv, **2020**, DOI: 10.26434/chemrxiv.12780044.v1

Abstract

TropC is an α-ketoglutarate (α-KG) dependent non-heme iron (NHI) enzyme that utilizes a highly reactive iron(IV)–oxo species. We considered three different pathways toward the product tropolone: rebound hydroxylation and semipinacol rearrangement (Pathway 1), radical-based ring expansion (Pathway 2), and single-electron transfer (SET) followed by polar ring expansion (Pathway 3). In this chapter, quantum chemical models used to gain insights into the proposed routes are described. Pathway 1 did not lead to tropolone, but instead yielded the trihydroxyorcinaldehyde shunt product (Pathway 4). The dominant pathway was determined to be Pathway 3, and throughout the SET step an unexpected ketene intermediate was formed. The radical-based ring expansion pathway was also found to have a similar barrier and thus a similar feasibility. These results suggest tropolones generated by TropC could be achieved through a one-electron pathway, and that the shunt product is a result of rebound hydroxylation and fragmentation. These small models provide initial insights into the possible pathways for but do not replace the enzyme active site, which likely restricts movement of the iron(IV)–oxo and the substrate and would be key in discriminating between the pathways.

3.1 Introduction

Tropolones are a class of aromatic compounds characterized by a cycloheptatriene core containing an α-hydroxyketone structure (Figure 3.1A, **3.1**) and are important scaffolds with known bioactivity.[1-8] Although chemical methods for synthesis of tropolones exist[6-13], they are often laborious to achieve the desired outcome, requiring extensive prefunctionalization[6,7,9,10], lengthy syntheses[6,7,9], and expensive metals[6,9,10], which result in tropolones with limited substitution patterns[7]. Nature, however, has evolved enzymes to efficiently access the tropolone core. TropC is an α-KG NHI enzyme from the stipitatic acid (**3.2**) biosynthetic pathway in *Talaromyces stipitatus* that is known to catalyze the oxidative ring expansion of **3.3** to form the tropolone product stipitaldehyde (**3.4**, Figure 3.1B).[14] Along with TropC, two additional enzymes have since been identified to catalyze the same oxidative ring expansion, XenC[15] and EupC[16] from the xenovulene A and eupenifeldin biosynthetic pathways, respectively, and have been assigned the same mechanism. Gaining insight into the mechanism of this unprecedented ring expansion will allow for better understanding of how these enzymes can be manipulated through protein engineering to increase production of tropolone product while suppressing shunt product formation and offer broader access to tropolone scaffolds.

First characterized in 2012, the mechanism for TropC proposed by Cox and coworkers was a hydroxylation followed by a semipinacol rearrangement[17,18] (Figure 3.1C, Pathway 1) to form **3.4**. Researchers first proposed an initial H-atom abstraction to produce the primary radical intermediate **3.5** and an iron(III)-OH species. Next, rebound hydroxylation would regenerate the iron(II) center and the hydroxylated intermediate **3.6**.

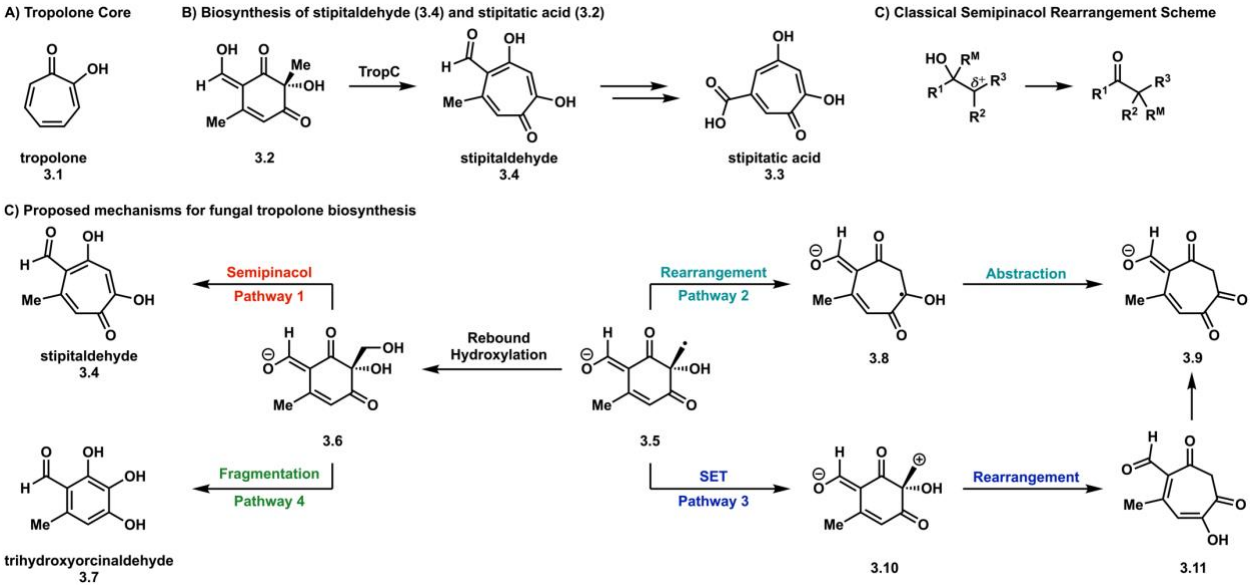

Figure 3.1. Overview of tropolone formation in TropC. A) Core of tropolone natural products. B) Biosynthetic pathway of stipitatic acid (3.2) catalyzed by TropC. C) Scheme of semipinacol rearrangements. D) Mechanisms on route to ring expansion in TropC and related enzymes.

Following formation of **3.6**, a semipinacol rearrangement facilitated by residues withisn the enzyme was proposed to produce the tropolone product **3.4**. Support for this mechanism was derived from the production of the trihydoxyorcinaldehyde shunt product (**3.7**), an irreversibly made product. **3.7** would arise from a retro-aldol fragmentation and rearomatization of **3.6**, akin to a demethylation mechanism in α-KG NHI enzymes.

Understanding the fate of the radical intermediate in these enzymes is critical to identifying the underlying mechanism of catalyzed transformations. To this end, we envisioned two additional mechanisms for tropolone biosynthesis (Figure 3.1D). First, from the primary radical intermediate (**3.5**), rather than a rebound hydroxylation a radical-based ring expansion would produce a tertiary-like radical intermediate **3.8**. **3.8** could then undergo a second H-atom abstraction from the iron(III)-OH species to reform the iron(II) center and produce a tautomer of the tropolone product (**3.9**). Also from **3.5**, we envisioned a pathway in which an SET to the iron(III)-OH species would result in the

zwitterionic intermediate **3.10**. **3.10** could then undergo the ring expansion to produce the neutral intermediate **3.11**. Deprotonation and tautomerization of **3.11** would produce stipitaldehyde (**3.4**). These mechanistic possibilities we explored both experimentally and computationally in collaborative effort with Dr. Tyler Doyon under the guidance of Prof. Alison Narayan and Dr. Leena Mallik under the guidance of Prof. Markos Koutmos. This chapter will provide evidence that a one-electron pathway is favorable over the semipinacol rearrangement through the use of 3D enzyme and active site mimic models. The following chapter (4) will explore the mechanism within the enzyme using QM/MM with the model generated herein.

3.2 Synthetic Studies on the Proposed Ring Expansion

Dr. Tyler Doyon obtained experimental evidence to support a one-electron pathway over the semipinacol rearrangement for tropolone formation via TropC. Using a two-enzyme cascade (Figure 3.2A), intermediate **3.6** was produced. CitB[19] was first used to produce the benzylic alcohol **3.12** from the CitB native substrate **3.13** via enzymatic hydroxylation, after which TropB performed an oxidative dearomatization. In solution, instead of observing **3.6**, only the trihydoxyorcinaldehyde shunt product **3.7** was detected. Since no tropolone was detected, this indicated that the fragmentation pathway to the shunt product was dominant over the semipinacol rearrangement. An alternative approach of using a combination of TropB and TropC was used in an attempt to produce stipitaldehyde from the benzylic alcohol **3.12** (Figure 3.2B). TropB enabled oxidative dearomatization to produce **3.6** and TropC was hypothesized to perform the semipinacol rearrangement to yield **3.4**. However, only the trihydoxyorcinaldehyde **3.7** shunt product

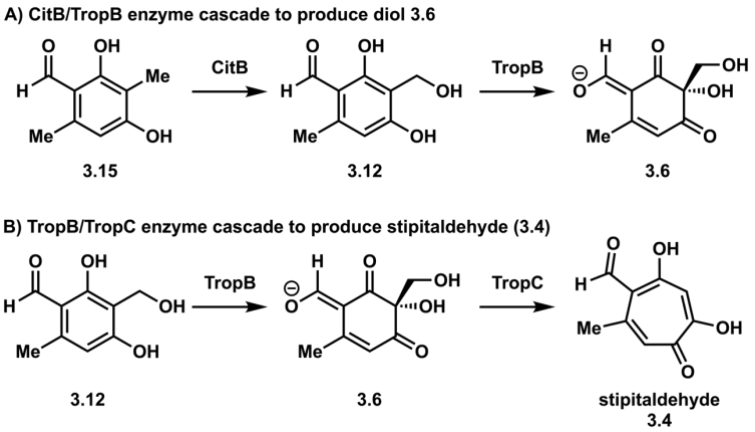

Figure 3.2. Experimental two-enzyme cascades used to form reactive intermediates. A) Diol intermediate **3.6** and B) stipitaldehyde (**3.4**).

was detected, further providing evidence that the fragmentation pathway is dominant. These results also suggest that the proposed semipinacol rearrangement mechanism is not the mechanism mediated by TropC. This indicates that one of our proposed alternative pathways could led to tropolone product.

3.3 Mutagenesis Studies on the Proposed Ring Expansion

A collaboration between Dr. Tyler Doyon and Dr. Leena Mallik produced a partial crystal structure for TropC using a sitting drop vapor diffusion method (PDB ID: 6XJJ). The α-KG NHI enzyme thymine-7-hydroxylase (T7H)[20] was used as a search model, due to its highly similar sequence and structure[21], for the molecular replacement method to solve the crystal structure of TropC with a resolution of 2.7 Å. Analysis of the crystal structure of TropC consisted of a double-stranded β-helix (DSBH or jelly-roll) with 10 anti-parallel β-strands forming the major (β1-5, β8, β10) and minor (β-6, β-7, β9) β-sheets, with the major β-sheet surrounded by α-helices to form a compact globular structure. These features are common for the architecture of α-KG NHI enzymes. Comparison between the TropC and T7H structures provided initial insight into the active site location and

residues that may be important for catalysis. The phenylalanine residues Phe284 and Phe213 in TropC aligned well with phenylalanine Phe292 and tyrosine Tyr217 in T7H. In T7H these residues π-stack with the thymine substrate and position it near the iron center, which indicates the corresponding residues in TropC play a similar role and are important for substrate positioning. Also shared were a 2-His-1-Asp facial triad of iron-coordinating residues. To gain further insight into residues important for substrate positioning or catalysis, a full substrate-bound model was desired. Therefore, a 3D model was constructed using the computational programs MUSCLE[22,23] and MODELLER[24] to construct the missing C-terminus helix, and CHARMM[25] to place the substrate and iron in the active site and equilibrate the model (see section 3.8 for full details).

Using the 3D model of the enzyme, nine additional residues outlining the active site were identified based on proximity to the substrate and subjected to an experimental alanine scan to identify their role and importance in catalysis (Figure 3.3B). This scan is

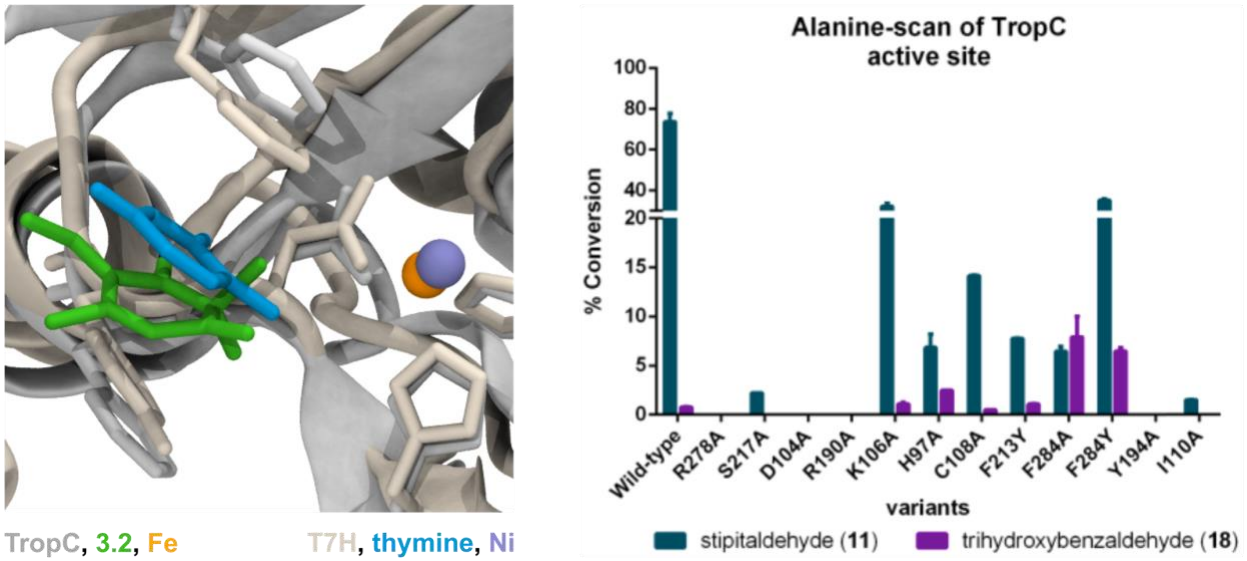

Figure 3.3. Mutagenesis study performed for TropC. A) Overlay of TropC model (gray) and T7H (tan). TropC native substrate (**3.2**) is shown in green, and thymine in blue. Iron from TropC in orange, and nickel from T7H in purple. B) Alanine scan results highlighting changes in activity. Only a switch in activity due to F284A mutation likely disrupting binding of **3.2**.

based on that idea that residues surrounding the substrate likely influence reactivity by binding and orientating the substrate and can be directly involved in catalysis, for instance via protonation or deprotonation.[26] The alanine scan—which replaces the residues of interest with alanine—allows probing of the role of a residue by removing steric and electronic parameters but maintaining the overall structure of the enzyme and observing the impact on catalysis. A residue was deemed important when the selectivity changed, for instance the major product switched from tropolone to shunt product or when they were made in equal quantities. Using this method, Dr. Tyler Doyon was able to make two key observations that support a one-electron pathway.

Of the 11 residues probed in the alanine scan, most were deemed not important due to a lack of change in chemistry. Many of the substitutions resulted in a loss of catalytic activity, likely due to structural changes within the enzyme that render it incapable of binding or performing catalysis. Among the catalytically active variants, only a loss in activity was observed, which would indicate the residues play a minor secondary role in catalysis. Included in these were polar residues Ser217, Asp104, Lys106, and His97 that were hypothesized to aide in the semipinacol rearrangement by protonating and deprotonating the substrate. Since the polar residues could be replaced without eliminating catalytic activity, these results suggest that a semipinacol rearrangement is unlikely.

Although most of the residues in the alanine scan were not identified to be important, a single exception is the residue Phe284. The F284A variant shifted conversions from ~75% tropolone and ~1% shunt product, to ~7% for each product. Based on the crystal structure comparison, Phe284 was hypothesized to be important for

substrate positioning. Specifically, Phe284 could guide the methyl group towards the iron center via steric repulsion, where otherwise improper positioning would shut down or greatly decrease catalysis. Consistent with this hypothesis, the F284Y variant, which is isoelectronic (aromatic) and isosteric (phenyl ring) with respect to Phe284, resulted in catalytic activity in favor of the tropolone. Based on these observations and computational results in section 3.4-3.6 we propose substrate positioning is critical for the reaction to undergo the one-electron pathway rather than rebound hydroxylation, akin to how halogenation pathways have been rationalized.[27] These results suggest that positioning within the active site can lead to two different reaction outcomes based on the proximity of the substrate to the iron center.

3.4 Computational Studies on the Proposed Ring Expansion

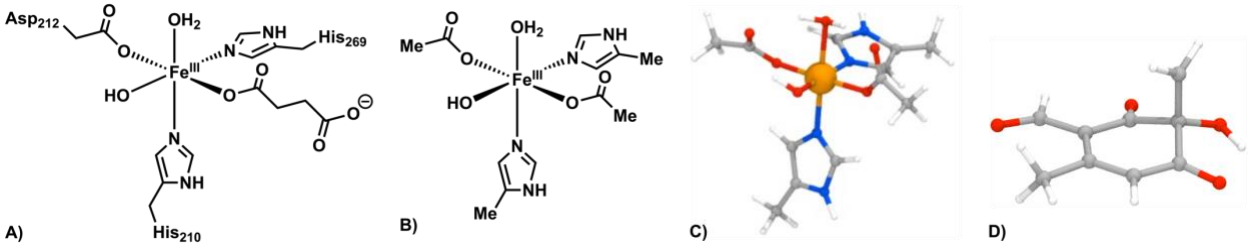

Figure 3.4**.** QM models used to study proposed mechanisms for TropC. Structures of iron in the A) enzyme active site B) QM model, and C) 3D model. Histidine was truncated to methylimidazole, aspartate and succinate to acetic acid, and retention of water molecule. D) Primary radical intermediate **3.5** produced after an initial H-atom abstraction.

Using an active site mimic, a 3D model containing the substrate, iron, and ligands (Figure 3.4) was used to perform QM simulations to explore the initially proposed polar ring expansion mechanism, Pathway 1 (see section 3.8 for full computational details). Specifically, the two histidine residues from the 3D enzyme model (His210 and His269) were replaced with methylimidazoles, the aspartate residue Asp212 replaced with acetic

acid, and the hydroxyl and the water molecule were retained. Using this model and beginning from the primary radical intermediate **3.5**, rebound hydroxylation was found to be a barrierless process. This finding is due to the QM model having complete translational and rotational freedom, though we would expect a non-zero barrier within the enzyme active site (see Chapter 4). Nonetheless, from **3.5** the semipinacol rearrangement was simulated with water molecules assisting protonation and deprotonation events. This rearrangement was found to be inaccessible, with a barrier of 48.8 kcal/mol (Figure 3.5, **TS1**). Analysis of **TS1**'s geometry provided context for why the barrier is high (Figure 3.6A). First, the forming C_1–C_3 bond had a length of 2.30 Å, which is only a small change from the initial length of 2.47 Å. Taken with the observation of complete C_1–O_1 bond breakage (2.06 Å), this indicated the presence of a fully, rather than partially, formed primary carbocation. Second, the geometry of the primary carbon atom itself is more representative of the breaking C_1–O_1 bond than the forming C_1–C_3 bond, as the hydrogens were pointed away from the oxygen atom. These observations can explain the high barrier, as the geometry is highly strained while balancing the loss of water and ring expansion in a concerted, asynchronous pathway.

The high barrier suggests that the semipinacol mechanism is not operative, and it also resembles prior results on a related system. Also using an active site mimic, Siegbahn and coworkers simulated a similar semipinacol ring expansion on the diol intermediate **3.13**.[28] Their model used the substrate coordinated to the iron, which would be expected to act as a Lewis acid and therefore better facilitate the rearrangement. Despite this benefit, a barrier of 48.4 kcal/mol was observed, again indicating an inaccessible pathway. Siegbahn and coworkers suggested an alternative mechanism

involving a second iron(IV)-oxo species, but it was observed experimentally that only one equivalent of α-KG was consumed throughout the production of stipitaldehyde (**3.4**). These results suggest that the semipinacol rearrangement is not feasible, and an alternative mechanism is required to access tropolones.

Although the rebound hydroxylation and subsequent semipinacol rearrangement cannot explain tropolone formation, the first step (rebound hydroxylation) can still be used to rationalize the formation of the shunt product (Pathway 4). The simulation of fragmentation and rearomatization of **3.6** to form the trihydoxyorcinaldehyde shunt product **3.7** was performed using the substrate and a different configuration of water molecules. The barrier for this process is 18.2 kcal/mol (**TS2**), with a thermodynamic favorability of -28.7 kcal/mol. The geometry of **TS2** is reflective of a concerted pathway,

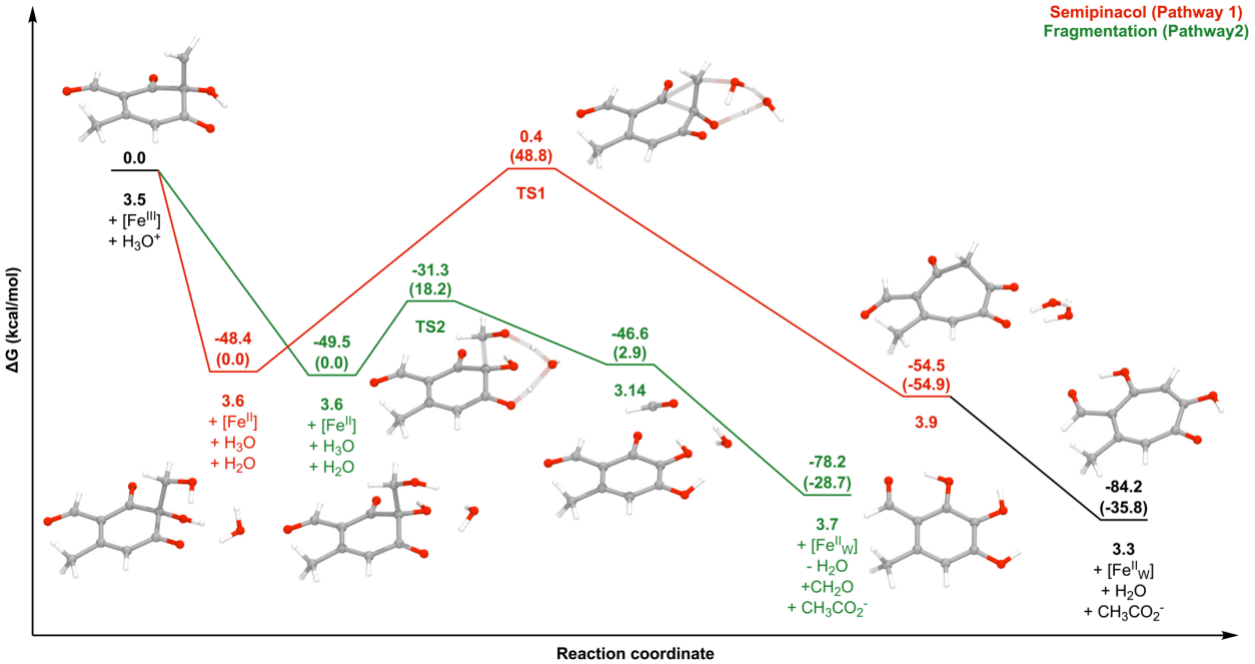

Figure 3.5. Energy profiles of the semipinacol and fragmentation pathways. Translucent lines at the TS are bonds forming or breaking. Results show that if rebound hydroxylation occurs, then the preferred pathway is fragmentation to generate the triohydroxyorcinaldehyde shunt product **3.7**.

where there is partial bond formation and breakage to form weakened bonds (Figure 3.6B). The breaking C_1–C_2 bond length at the TS is 1.70 Å, and the C_1–O_1 bond 1.33 Å, which is between the lengths of 1.41 Å and 1.21 Å for the initial and final structures, respectively. Finally, the hydrogens on C_1 were pointed away from C_2 rather than being in plane with O_2. These results indicate that if rebound hydroxylation were to occur, then the preferred pathway would be fragmentation and rearomatization. This finding provides more evidence that a one-electron mechanism is responsible for the formation of the tropolone natural product.

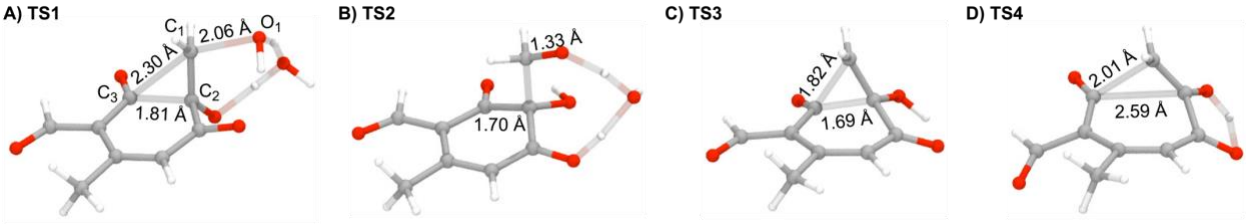

Figure 3.6. Transition state geometries. Geometries from A) the semipinacol (**TS1**), B) shunt product (**TS2**), C) radical-based rearrangement (**TS3**), and SET (**TS4**) pathways. Bond lengths in **TS1** is reflective of a concerted, asynchronous pathway, and concerted for **TS2, TS3**, and **TS4**. Translucent lines at the TS are bonds forming or breaking.

3.5 Computational Studies on the Radical Rearrangement Pathway

The first one-electron pathway that we envisioned was a radical-based ring expansion from **3.5** to **3.8** (Pathway 2). We predicted a kinetically and thermodynamically favorable ring expansion due to two factors. The first being a thermodynamic driving force due to formation of a more stable radical. Second, literature precedent has shown production of tropolones from *ortho*-quinol radical precursors.[29] To test the hypothesis of a radical-based ring expansion (Figure 3.7), simulations were started from **3.5**. A feasible barrier of 12.7 kcal/mol (**TS3**) was found, with a thermodynamic favorability of -16.9 kcal/mol.

Analysis of **TS3** (Figure 3.6C) provides insight into the favorability of this transformation over the proposed semipinacol rearranement pathway. The forming C_1–C_3 bond has a length of 1.82 Å, which is significantly shorter than the initial length of 2.43 Å and that of the semipinacol pathway (Figure 3.6A, 2.30 Å). The breaking C_2–C_3 bond has a small increase in length from 1.57 Å to 1.69 Å, which indicated only a partial, rather than full, bond breaking event as expected. The geometry of the primary carbon atom is again able to explain the kinetic favorability as the hydrogens are properly pointing away from the carbon atom. These three observations can be used to explain the favorability of this radical pathway over the semipinacol like pathway (Pathway 1), in that an unstrained transition state was formed.

To determine if the radical pathway is achievable, the second H-atom abstraction step was examined. This step was calculated to be barrierless, but we would again expect

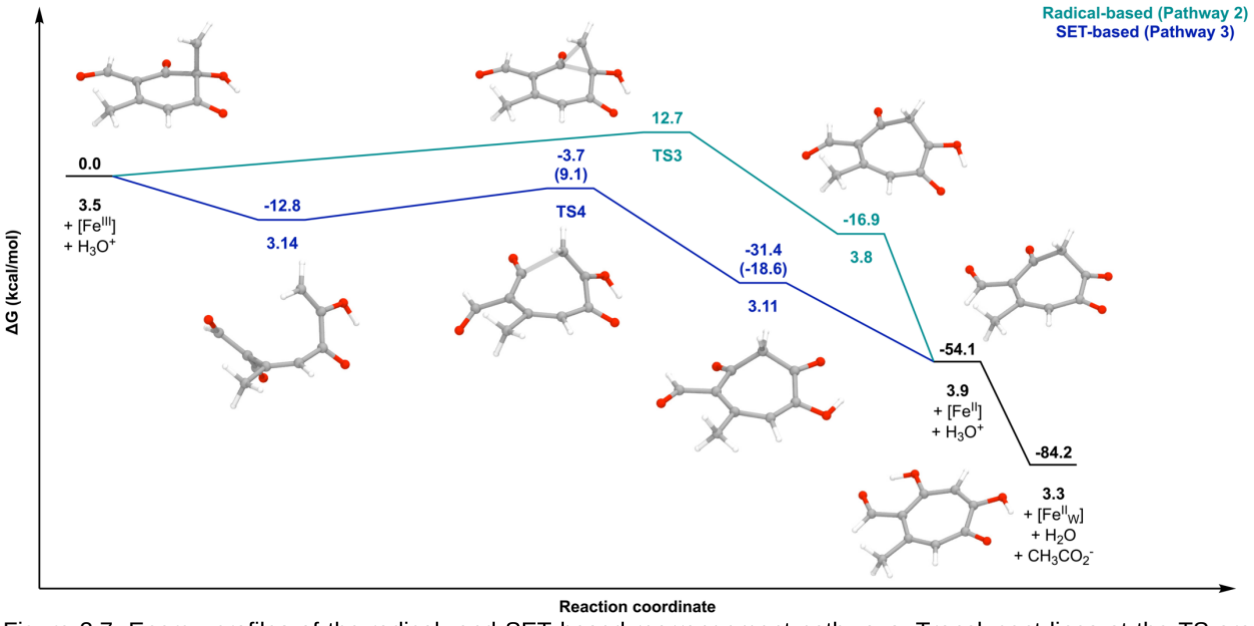

Figure 3.7. Energy profiles of the radical- and SET-based rearrangement pathways. Translucent lines at the TS are bonds forming or breaking. Results show that both pathways are kinetically and thermodynamically accessible.

a non-zero barrier within an enzyme active site due to positional limitations. Regardless, these results indicate that a one-electron pathway is viable to afford tropolone product. These results also highlight how products from two different pathways are accessible, where the radical-based ring expansion produces tropolones and rebound hydroxylation yields the shunt product. These results support the hypothesis of substrate positioning being critical for TropC to guide the reaction towards the tropolone product, as without this the shunt product is within reach. As discussed in Section 3.3, the Phe284 residue likely provides the key guiding factor for tropolone formation.

3.6 Computational Studies on the SET Pathway

An alternative to the radical-based ring expansion mechanism discussed in section 3.5 involves SET from the primary radical anion intermediate **3.5** to the iron(III) center (Pathway 3), which would result in the formation of the ground state, zwitterion intermediate **3.10** and an iron(II) center. The SET process was found to be thermodynamically favorable by -12.8 kcal/mol (Figure 3.7). To determine if this pathway is kinetically feasible (k_{et}), the Marcus theory parameters of ΔH, λ, and H_{ab} were gathered. k_{et} was calculated using eq. 1 and found to be 9.57×10^5 s^{-1}, or 9.3 kcal/mol (see section 3.8 for full details). These results indicated SET was kinetically and thermodynamically accessible and suggested it could compete with radical rearrangement. Based on the barrier of 12.7 kcal/mol for the radical rearrangement pathway, the effective barrier for SET transfer is lower by 3.4 kcal/mol and therefore the kinetically favorable pathway.

Interestingly, during optimization of the zwitterionic intermediate **3.10**, instead of ring expansion to the neutral intermediate **3.11** to quench a transient tertiary-like

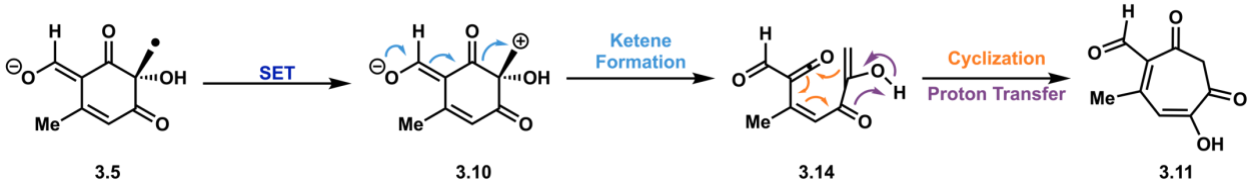

Figure 3.8. Mechanism of formation of the neutrally expanded intermediate **3.11** from the primary radical intermediate **3.5** following SET to the iron species.

carbocation, the ketene intermediate **3.14** was formed because of the instability of **3.10**. Due to the presence of **3.14** after SET, an alternative pathway to the tropolone tautomer **3.9** was investigated. Specifically, concerted intramolecular cyclization and proton transfer leads to **3.11** (Figure 3.8), and this occurs with a low barrier due to the **3.14** being a highly reactive intermediate (Figure 3.7). A small barrier of 9.1 kcal/mol (**TS4**) and a thermodynamic favorability of -21.2 kcal/mol were found. Like **TS2** and **TS3**, **TS4** is a concerted pathway, with forming C_1–C_3 and breaking C_2–C_3 bond lengths of 2.01 Å and 2.59 Å, respectively. This makes **TS4** representative of a late TS, where the geometry more reflects the product than reactant (Figure 3.6D). Deprotonation of **3.11** by the iron(II)–OH species would result in **3.9**, and this process was computed to have no barrier.

Based on the kinetic and thermodynamic favorability of SET, and a low barrier pathway to formation of the neutral ring expanded intermediate **3.11**, the results herein suggest that a SET pathway is feasible for tropolone production. These small active site models, however, have considerably more translational and rotational freedom than the active site embedded within the enzyme. Therefore, these results can only provide insights into what could be possible, and not to draw concrete conclusions without consideration of the enzyme active site. This topic will be reexamined in Chapter 4.

3.7 Conclusions

Based on both experimental and computational observations, the dominant pathway for tropolone formation mediated by TropC (and related enzymes XenC and EupC) is likely a one-electron pathway (Pathways 2/3). The originally proposed, Pathway 1, rebound hydroxylation and semipinacol rearrangement mechanisms are not favored by the models of this chapter. Although the precursors to stipitaldehyde (**3.4**) were produced by two different means via enzyme cascades, only the trihydroxyorcinaldehyde shunt product **3.7** was observed. Throughout the alanine scan, substitution of polar residues proposed to facilitate ring expansion only diminished activity, suggesting they do not play a role in the formation of **3.4**. However, the phenylalanine residue Phe284 was deemed critical for proper substrate positioning since the F284A mutation switched activity to favor shunt product formation. The dominance of shunt product formation is supported by simulations that show the semipinacol rearrangement having an inaccessible barrier of 48.8 kcal/mol. Instead, if rebound hydroxylation occurs, then fragmentation and rearomatization is dominant with a reasonable barrier of 18.2 kcal/mol. Exploration of two distinct one-electron pathways through quantum chemical models suggest that both are feasible. Radical rearrangement of the primary radical intermediate **3.5** to form the tertiary-like radical intermediate **3.8** (Pathway 2) was calculated to have a barrier of only 12.7 kcal/mol, and a favorable SET from **3.5** to the iron(III)-OH species (Pathway 3) would result in a ketene intermediate with a low barrier (9.1 kcal/mol) to the neutral expanded intermediate **3.11**. Our collaborative work is also supported by literature precedent. The first is previous simulations show a high barrier pathway for the semipinacol rearrangement, even with iron acting as a Lewis acid.[28] Second, it has been shown that

ortho-quinol substrates undergo a radical rearrangement to form tropolones.[29] Given these observations, it is possible that TropC and related enzymes may operate on a one-electron pathway for ring expansion towards tropolone production, but a traditional rebound hydroxylation of fragmentation pathway for shunt product formation. However, due to limitations mentioned in this chapter of QM models, a full QM/MM study would be required to make conclusions on what pathway is most dominant within the active site of TropC. Understanding the influence of the active site has on the reaction pathways is key to understanding the mechanisms and will be discussed in the next chapter. The results in this chapter seek to clarify the mechanistic possibilities and add to known radical rearrangements catalyzed by α-KG NHI enzymes.

3.8 Computational Details

All 3D images were generated using the Visual Molecular Dynamics software program.[30] All isovalues for spin densities were set to +/- 0.3.

3.8.1 Molecular Dynamics (MD) Simulations

Model Construction

To build the 3D model of TropC, the partial crystal structure, and the MUSCLE and MODELLER software packages were used to construct the C-terminal region for which no density was observed in X-ray crystallography experiments.[31] To develop a homology model for this region, the TropC amino acid sequence was submitted to the RCSB PDB[32] and JPred4[33] servers for secondary structure prediction. These programs indicated that the thebaine 6-O-demethylase crystal structure (PDB ID: 5O7Y[34], 27% sequence identity,

Pymol[35] superalign Cα-RMSD: 4.378) was most closely related to TropC. Based on these structure prediction results, residues 307-323 of the C-terminal region were modeled as a helix.

To generate the substrate- and cofactor-bound model for MD simulations, further analysis of the RCSB PDB structure revealed that the α-KG NHI enzyme T7H possessed structural similarity to the *apo* crystal structure of TropC (PDB ID: 5C3Q, 28% sequence identity, Pymol superalign Cα-RMSD: 3.139). The deposited crystal structure of T7H was co-crystallized with substrate, nickel, and α-KG. In addition, similarities between the native substrates of TropC and T7H are evident. Both are small, cyclic substrates with a similar substitution pattern of a methyl group ortho to a hydroxyl group. To further establish the model for MD simulations, the TropC coordinates were overlaid and aligned with chain A of the T7H structure, and then transferred the nickel, substrate, and α-KG coordinate to the TropC model. The thymine was replaced with the TropC native substrate **3.2** and the nickel atom changed to an iron atom.

System Preparation

The CHARMM-gui was used to add hydrogen atoms and protons to the TropC model, which were added according to their canonical protonation states. These protonation states were determined based on the use of the Propka-3.1 software.[36,37] The tautomeric forms of histidine were assigned to maximize hydrogen bonds. To obtain an overall electrically neutral charge model, 47 potassium and 42 chlorine atoms were added. A water box was used with periodic boundary conditions, and the particle mesh Ewald method was used to describe long-range electrostatic effects. Dimensions of the water box were established to fit the largest dimension of the protein plus 10 Å. All water

molecules within 2.8 angstroms of heavy atoms were removed. The final model contained 53,819 atoms. Parameters for the substrate and α-KG were obtained using CGenFF 4.0.[38,39] Parameters for iron were taking from the paper by Pang et al.[40] A restrained active site was used (50 kcal/molÅ^2 force constant), in which the restraints were based on crystal structure distances for T7H' and the α-KG NHI VioC (PDB ID: 2WBO[41]). VioC was chosen due to the substrate-arginine interaction. Included in these restraints were a positional restraint on Phe213 residue to more closely resemble that of the complimentary residue in 5C3Q to maintain closure of the active site, and restraints based on H-bonds within the active site. A 5,000 step MM geometry optimization was then was performed using the steepest decent algorithm and the CHARMM36 force field.[39]

MD Simulations

Simulations were performed with the OpenMM/CHARMM[42] interface (version c42b1) using periodic boundary conditions, particle mesh Ewald for long range electrostatic interactions, and a Langevin thermostat (friction coefficient = 5). A three step MD protocol was used to equilibrate the model, which all used a 2 fs timestep. The first phase slowly heated the model from 10K to 298 K over 500 ps and included the restrained active site as well as harmonic restraints of 10.0 kcal mol/Å^{-2} on all heavy atoms of the solute. Next, the harmonic restraints were then removed, and the temperature was kept constant for an additional 500 ps. The final phase removed the Phe213 positional restraint and simulations were performed for 13 nanoseconds (ns), with the first two ns used as equilibration.

3.8.2 QM Simulations

General Computational Details

All geometry optimizations and frequency calculations were performed with QChem 4.0[43] using the unrestricted B3LYP density functional[44,45] and the 6-311++G** basis set[46-50] for all atoms except Fe, which used the 6-311G** basis set[51]. All ground state geometries were confirmed to have no imaginary frequencies, and transition states were confirmed to have one imaginary frequency. Solvent corrections were performed with ORCA 4.0[52] using the SMD[53-55] implicit solvent mode and the cc-pVTZ[56-58] basis set. Reaction pathways were simulated using the single-ended growing string method.[59] Reported energies are Gibbs free energies in solvent with gas-phase entropy and enthalpy corrections (at 298.15 K).

3.8.3 Marcus Theory Parameters

H_{ab} was estimated by comparing the substrate highest occupied molecular orbital (HOMO) and lowest singly occupied molecular orbital (SOMO) of the iron complex. The fragments were manually aligned and H_{ab} was estimated to be 0.788 eV by taking the difference between the orbital energies for the separated and combined fragments and dividing by two. This strong coupling value would suggest that the SET would likely be favorable. To calculate ΔH, the two fragments were optimized post-SET separately, their energies combined, and compared to the pre-SET energy minimums. ΔH for this system was calculated to be -3.9 kcal/mol in favor of SET. λ was calculated to be 61.0 kcal/mol by taking the pre-SET geometries and calculating post-SET single point energies, and

then comparing to the post-SET optimized energies. With all the required Marcus theory parameters now collected, k_{et} was calculated using eq 1 and found to be 9.57×10^5 s^{-1}.

$$k_{et} = \frac{2\pi}{\hbar}|H_{AB}|^2 \frac{1}{\sqrt{4\pi\lambda k_B T}} \exp\left(\frac{(\lambda+\Delta H)^2}{\sqrt{4\lambda k_B T}}\right) \qquad \text{eq 1}$$

3.9 References

[1] Nakamura, H.; Matsuda, Y.; Abe, I. Unique Chemistry Of Non-heme Iron Enzymes in Fungal Biosynthetic Pathways. *Nat. Prod. Rep.* **2018**, *35*, 633–645.

[2] Tang, M.-C.; Zou, Y.; Watanabe, K.; Walsh, C. T.; Tang, Y. Oxidative Cyclization in Natural Product Biosynthesis. *Chem. Rev.* **2017**, 117*,* 5226–5333.

[3] Wu, L.-F.; Meng, S.; Tang, G.-L. Ferrous Iron and α-Ketoglutarate-Dependent Dioxygenases in the Biosynthesis of Microbial Natural Products. *Biochim. Biophys. Acta, Proteins Proteomics* **2016**, *1864*, 453–470.

[4] Cox, R. Oxidative Rearrangements During Fungal Biosynthesis. *Nat. Prod. Rep.* **2014**, *31*, 1405–1424.

[5] Cox, R. J.; Al-Fahad, A. Chemical Mechanisms Involved During the Biosynthesis of Tropolones. *Curr. Opin. Chem. Biol.* **2013**, *17*, 532–536.

[6] Guo, H.; Roman, D.; Beemelmanns, C. Tropolone Natural Products. *Nat. Prod. Rep.* **2019**, *36*, 1137–1155.

[7] Meck, C.; D'Erasmo, M. P.; Hirsch, D. R.; Murelli, R. P. The Biology and Synthesis of α-Hydroxytropolones. *Med. Chem. Commun.* **2014**, *5*, 842–852.

[8] Hirsch, D. R.; Schiavone, D. V.; Berkowitz, A. J.; Morrison, L. A.; Masaoka, T.; Wilson, J. A.; Lomonosova, E.; Zhao, H.; Patel, B. S.; Dalta, S. H.; Hoft, S. G.; Majidi, S. J.; Pal, R. K.; Gallicchio, E.; Tang, L.; Tavis, J. E.; Le Grice, S. F. J.; Beutler, J. A.; Murelli, R. P. Synthesis and Biological Assessment Of 3,7-Dihydroxyytopolones. *Org. Biomol. Chem.* **2018**, *16*, 62–69.

[9] Liu, N.; Song, W.; Schienebeck, C. M.; Zhang, M.; Tang, W. Synthesis of Naturally Occurring Tropones and Tropolones. *Tetrahedron* **2014**, *70*, 9281–9305.

[10] Reisman, S. E.; Nani, R. R.; Levin, S. Buchner and Beyond: Arene Cyclopropanation as Applied to Natural Product Total Syntesis. *Synlett* **2011**, 2437–2442.

[11] Stevens, H. C.; Reich, D. A.; Brandt, D. R.; Fountain, K. R.; Gaughan, E. J.; A New Tropolone Synthesis via Dichloroketene. *J. Am. Chem. Soc.* **1965**, *87*, 5257–5359.

[12] Barbier, M.; Burton, D. H. R.; Devys, M.; Topgi, R. S.; A Simple Synthesis of Tropoones and Related Compounds. *Tetrahedron* **1987**, *43*, 5031–5038.

[13] Barbier, M.; Burton, D. H. R.; Devys, M.; Topgi, R. S.; A Simple Synthesis of the Tropone Nucleus. *J. Chem. Soc. Chem. Commun.* **1984**, 743–744.

[14] Davison, J.; Fahad, A.; Cai, M.; Song, Z.; Yehia, S. Y.; Lazarus, C. M.; Bailey, A. M.; Simpson, T. J.; Cox, R. J. Genetic, Molecular, and Biochemical Basis of Fungal Tropolone Biosynthesis. *PNAS* **2012**, *109*, 7642–7647.

[15] Schor, R.; Schott, C.; Wibberg, D.; Kalinowski, J.; Cox, R. J. Three Previously Unrecognized Classes of Biosynthetic Enzymes Revealed During the Production of Xenovulene A. *Nat. Commun.* **2018**, *9*, 1963.

[16] Chen, Q.; Gao, J.; Jamieson, C.; Liu, J.; Ohashi, M.; Bai, J.; Yan, D.; Liu, B.; Che, Y.; Wang, Y.; Houk. K. N.; Hu, Y. Enzymatic Intermolecular Hetero-Diels-Alder Reaction in the Biosynthesis of Tropolonic Sesquiterpenes. *J. Am. Chem. Soc.* **2019**, *141*, 14052–14056.

[17] Zhang, X.-M.; Li, B.-S.; Wang, S.-H.; Zhang, K.; Zhang, F.-M.; Tu, Y.-Q. Recent Development and Applications of Semipinacol Rearrangement Reactions. *Chem. Sci.* **2021**, *12*, 9262–9724.

[18] Song, Z.-L.; Fan, C.-A.; Tu, Y.-Q. Semipinacol Rearrangement in Natural Product Synthesis. *Chem. Rev,* **2011**, *11*, 7523–7556.

[19] He, Y.; Cox, R. J. The Molecular Steps of Citrinin Biosynthesis in Fungi. *Chem. Sci.* **2016**, *7*, 2119–2127.

[20] Li, W.; Zhang, T.; Ding, J. Molecular Basis for the Substrate Specificity and Catalytic Mechanism of Thymine-7-Hydroxylae in Fungi. *Nucl. Acids Res.* **2015**, *43*, 10026–10038.

[21] Doyon, T. J.; Skinner, K. C.; Yang, D.; Mallik, L.; Wymore, T.; Koutmos, M.; Zimmerman, P. M.; Narayan, A. R. H. Radical Tropolone Biosynthesis. *ChemRxiv*, **2020**, DOI: 10.26434/chemrxiv.12780044.v1

[22] Edgar, R. C. MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797.

[23] Edgar, R. C. MUSCLE: A Multiple Sequence Alignment Method With Reduced Time and Space Complexity. *BMC Bioinform.* **2004**, 113.

[24] Šali, A.; Blundell, T. L. Comparative Protein Modelling by Satisfaction of Spatial Restraints. *J. Mol. Biol.* **1993**, *234*, 234, 779–815.

[25] Brooks, B. R.; Brooks III, C. L.; Mackkerell Jr, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, M.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Pos, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcosk, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545–1614.

[26] Morrison, K. M.; Weiss, G. A. Combinatorial Alanine-Scanning. *Curr. Opin. Chem. Biol.* **2001**, *5*, 302–307.

[27] Mitchell, A. J.; Zhu, Q.; Maggiolo, A. O.; Ananth, N. R.; Hillwig, M. L.; Liu, X.; Boal, A. K. Structural Basis for Halogenation by Iron- and 2-Oxo-Glutarate-Dependent Enzyme Welo5. *Nat. Chem. Biol.* **2016**, *12*, 636–640.

[28] Borowski, T.; Wójcik, A.; Miłaczewska, A.; Georgiev, V.; Blomberg, M. R. A.; Siegbahn, P. E. M., The Alkenyl Migration Mechanism Catalyzed by Extradiol Dioxygenases: A Hybrid DFT Study. *J. Biol. Inorg. Chem.* **2012,** *17*, 881–890.

[29] Cacioli, P.; Reiss, J. Reactions Of 1-Oxaspiro[2.5]Octa-5,7-Diene-4-Ones With Nucleophiles. *Aust. J. Chem.* **1984**, *37*, 2525–2535.

[30] Humphrey, W.; Dalke, A; Schulten, K., VMD - Visual Molecular Dynamics. *J. Molec. Graphics* **1996**, *14*, 33–38.

[31] Martí-Renom, M. A.; Stuart, A. C.; Fiser, A.; Sánchez, R.; Melo, F.; Śali, A. Comparative Protein Structure Modeling of Genes and Genomes. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 291–325.

[32] Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucl. Acids Res.* **2000**, *28*, 235–242.

[33] Drozdetskiy, A.; Cole, C.; Procter, J.; Barton, G. J. Jpred4: A Protein Secondary Structure Prediction Server. *Nucl. Acids Res.* **2015**, *43*, W389–W394.

[34] Kluza, A., Niedzialkowska, E., Kurpiewska, K., Wojdyla, Z., Quesne, M., Kot, E., Porebski, P.J., Borowski, T. Crystal Structure of Thebaine 6-O-Demethylase from the Morphine Biosynthesis Pathway. *J. Struct. Biol.* **2018**, *202*, 229–235.

[35] The Pymol Molecular Graphics System, V., Schrödinger, LLC.

[36] Sondergaard, C R.; Olsson, M. H. M.; Rostkowski, Michal.; Jensen, J. H Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization Of pKa Values. *J. Chem. Theory Comput.* **2011**, *7*, 2284–2295.

[37] Olsson, M. H. M.; Sondergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *J. Chem. Theory. Comput.* **2011**, *7*, 525–537.

[38] Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell Jr, A. D. CHARMM General Force Field: A Force Field for Drug-Like Molecules Compatile With the CHARMM All-Atom Additive Biological Force Field. *J. Comput. Chem.* **2010**, *31*, 671–690.

[39] Yu, W.; He, X.; Vanommeslaeghe, K.; Mackerell Jr, A. D. Extension Of The CHARMM General Force Field To Sulfonyl-Containing Compounds and Its Utility in Biomolecular Simulations. *J. Comput. Chem.* **2012**, *33*, 2451–2468.

[40] Pang, X.; Han, K.; Cui, Q. A Simple but Effective Modeling Strategy for Structural Properties of Non-heme Fe(II) Sites in Proteins: Test Of Force Field Models and Application to Proteins in the AlkB Family. *J. Comput. Chem.* **2013**, *34*, 1620-1635.

[41] Helmetag, V.; Samel, S. A.; Thomas, M. G.; Marahiel, M. A.; Essen, L.-O. Structural Basis for the Erythro-Stereospecificity of the L-Arginine Oxygenase Vioc in Viomycin Biosynthesis. *FEBS J.* **2009**, *276*, 3669–3682.

[42] Eastman, P.; Swails, J.; Chodera, J. D.; Mcgibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L.-P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; Wiewiora, R. P.; Brooks, B. R.; Pande, V. S. Openmm 7: Rapid Development of High Performance Algorithms for Molecular Dynamics. *Plos Comput. Biol.* **2017**, *13*, E1005659.

[43] Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T. B.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X.; Ghosh, D.; Goldey, M.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Khaliullin, R. Z.; Kuś, T.; Landau, A.; Liu, J.; Proynov, E. I.; Rhee, Y. M.; Richard, R. M.; Rohrdanz, M. A.; Steele, R. P.; Sundstrom, E. J.; Woodcock, H. L.; Zimmerman, P. M.; Zuev, D.; Albrecht, B.; Alguire, E.; Austin, B.; Beran, G. J. O.; Bernard, Y. A.; Berquist, E.; Brandhorst, K.; Bravaya, K. B.; Brown, S. T.; Casanova, D.; Chang, C.-M.; Chen, Y.; Chien, S. H.; Closser, K. D.; Crittenden, D. L.; Diedenhofen, M.; Distasio, R. A.; Do, H.; Dutoi, A. D.; Edgar, R. G.; Fatehi, S.; Fusti-Molnar, L.; Ghysels, A.; Golubeva-Zadorozhnaya, A.; Gomes, J.; Hanson-Heine, M. W. D.; Harbach, P. H. P.; Hauser, A. W.; Hohenstein, E. G.; Holden, Z. C.; Jagau, T.-C.; Ji, H.; Kaduk, B.; Khistyaev, K.; Kim, J.; Kim, J.; King, R. A.; Klunzinger, P.; Kosenkov, D.; Kowalczyk, T.; Krauter, C. M.; Lao, K. U.; Laurent, A. D.; Lawler, K. V.; Levchenko, S. V.; Lin, C. Y.; Liu, F.; Livshits, E.; Lochan, R. C.; Luenser, A.; Manohar, P.; Manzer, S. F.; Mao, S.-P.; Mardirossian, N.; Marenich, A. V.; Maurer, S. A.; Mayhall, N. J.; Neuscamman, E.; Oana, C. M.; Olivares-Amaya, R.; O'Neill, D. P.; Parkhill, J. A.; Perrine, T. M.; Peverati, R.; Prociuk, A.; Rehn, D. R.; Rosta, E.; Russ, N. J.; Sharada, S. M.; Sharma, S.; Small, D. W.; Sodt, A.; Stein, T.; Stück, D.; Su, Y.-C.; Thom, A. J. W.; Tsuchimochi, T.; Vanovschi, V.; Vogt, L.; Vydrov, O.; Wang, T.; Watson, M. A.; Wenzel, J.; White, A.; Williams, C. F.; Yang, J.; Yeganeh, S.; Yost, S. R.; You, Z.-Q.; Zhang, I. Y.; Zhang, X.; Zhao, Y.; Brooks, B. R.; Chan, G. K. L.; Chipman, D. M.; Cramer, C. J.; Goddard, W. A.; Gordon, M. S.; Hehre, W. J.; Klamt, A.; Schaefer, H. F.; Schmidt, M. W.; Sherrill, C. D.; Truhlar, D. G.; Warshel, A.; Xu, X.; Aspuru-Guzik, A.; Baer, R.; Bell, A. T.; Besley, N. A.; Chai, J.-D.; Dreuw, A.; Dunietz, B. D.; Furlani, T. R.; Gwaltney, S. R.; Hsu, C.-P.; Jung, Y.; Kong, J.; Lambrecht, D. S.; Liang, W.; Ochsenfeld, C.; Rassolov, V. A.; Slipchenko, L. V.; Subotnik, J. E.; Van

Voorhis, T.; Herbert, J. M.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M., Advances In Molecular Quantum Chemistry Contained in the Q-Chem 4 Program Package. *Molec. Phys.* **2015**, *113*, 184–215.

[44] Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

[45] Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula Into a Functional of the Electron Density. *Phys. Rev. B* **1998**, *37*, 785–789.

[46] Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. R. Efficient Diffuse Function-Augmented Basis Sets for Anion Calculations. III. The 3-21+G Basis Set for First-Row Elements, Li-F. *J. Comput. Chem.* **1983**, *4*, 294–301.

[47] Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; Defrees, D. J.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XXIII. A Polarization-Type Basis Set for Second-Row Elements. *J. Chem. Phys.* **1982**, *77*, 365–43665.

[48] Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XX. A Basis Set for Correlated Wave Functions. *J. Chem. Phys.* **1980**, *72*, 650–654.

[49] Mclean, A. D.; Chandler, G. S. Contracted Gaussian Basis Sets for Molecular Calculations. I. Second Row Atoms, Z=11-18. *J. Chem. Phys.* **1980**, *72*, 5639–5648.

[50] Spitznagel, G. W.; Clark, T.; Schleyer, P. V. R.; Hehre, W. J. An Evaluation of the Performance of Diffuse Function-Augmented Basis Sets for Second Row Elements, Na-Cl. *J. Comput. Chem.* **1987**, *8*, 1109–1116.

[51] Curtiss, L. A.; Mcgrath, M. P.; Blaudeau, J.-P.; Davis, N. E.; Binning, R. C.; Radom, L. Extension of Gaussian-2 Theory to Molecules Containing Third-Row Atoms Ga-Kr. *J. Chem. Phys.* **1995**, *103*, 6104–6113.

[52] Neese, F., The ORCA Program System. *Wires Comput. Molec. Sci.* **2012**, *2*, 73–78.

[53] Kendall, R. A.; Dunning, Jr., T.H.; Harrison, R. J. Electron Affinities of the First-Row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys.* **1992**, *96*, 6796–6806.

[54] Woon, D. E.; Dunning, Jr., T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. III. The Atoms Aluminum Through Argon. *J. Chem. Phys.* **1993**, *98*, 1358–1371.

[55] Woon, D. E.; Dunning, Jr., T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. IV. Calculation of the Static Electrical Response Properties. *J. Chem. Phys.* **1994**, *100*, 2975–2988.

[56] Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron Through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

[57] Wilson, A. K.; Woon, D. E.; Peterson, K. A.; Dunning, T. H. Gaussian Basis Sets for Use In Correlated Molecular Calculations. IX. The Atoms Gallium Through Krypton. *J. Chem. Phys.* **1999**, *110*, 7667–7676.

[58] Woon, D. E.; Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. III. The Atoms Aluminum Through Argon. *J. Chem. Phys.* **1993**, *98*, 1358–1371.

[59] Zimmerman, P. M. Single-ended Transition State Finding With the Growing String Method. *J. Comput. Chem.* **2015**, *36*, 601–611.

Chapter IV

Radical-based Ring Expansion and Potential Single-electron Transfer in TropC

Abstract

TropC is an α-ketoglutarate (α-KG) dependent non-heme iron (NHI) enzyme that catalyzes an unprecedented ring expansion to form a tropolone natural product. Three pathways were studied using a solvent explicit QM/MM model: rebound hydroxylation (Pathway 1), radical-based ring expansion (Pathway 2), and single-electron transfer-(SET) based ring expansion (Pathway 3). Pathways 1 and 2 were found to be viable with a slight kinetic advantage for Pathway 2, and Pathway 3 was found to be potentially viable. Analysis of a near-crossing structure revealed that on the ground state SET from the substrate to the iron is possible. Noncovalent interactions between the substrate and the enzyme and solvent were analyzed throughout the reaction pathways and found to be key for discriminating between the pathways. TropC positions the substrate and reactive oxidant species in a manner that partially inhibits Pathway 1 and gives preference to Pathway 2.

4.1 Introduction

α-Ketoglutarate-dependent (α-KG) non-heme iron (NHI) enzymes catalyze a breadth of reactions, including hydroxylation, epoxidation, halogenation, and skeletal rearrangements.[1-5] Although consensus mechanisms for many NHI transformations exist,

the precise mechanism of skeletal rearrangements is not universal and can differ among NHI enzymes. In particular for these rearrangements, various mechanisms can diverge from the reactive intermediate, leading to considerable uncertainty about the true nature of the reaction.[1-5]

Given the breadth of transformations available to NHI-catalyzed skeletal rearrangements, their underlying mechanisms deserve to be better understood. Skeletal rearrangements occur through indirect (post-enzyme reaction) and direct (radical- or single-electron transfer- based (SET-based) mechanisms, as discussed in Chapter 1. These three mechanistic routes for skeletal rearrangements can lead to important differences in substrate reactivity. The importance of mechanism in enzymatic reactions has been reflected in, for instance, observations that removing an α-heteroatom switches the native reaction from desaturation to hydroxylation in NapI,[6] and replacing D-arginine with L-arginine switches the native reaction of VioC from hydroxylation to oxidative deamination.[7] In another example, a new chemistry (nitrene transfer) was enabled by inhibiting the main reaction pathway using insight garnered from its mechanism.[8,9]

An interesting example of a skeletal rearrangement discussed in the previous chapter (3) is the unprecedented ring expansion in the biosynthesis of the tropolone stipitaldehyde (**4.1)** from a cyclic, nonaromatic six-membered dicarbonyl substrate (**4.2**) (Figure 4.1A). Cox and coworkers originally proposed that the mechanism for tropolone formation proceeded through an indirect mechanism via hydroxylation to the diol (Pathway 1) intermediate **4.3** and a subsequent semipinacol rearrangement (Figure 4.1B).[14] However, based on simulations and experiments, it can be argued that if rebound hydroxylation were to occur the preferred pathway would be fragmentation to produce the

trihydroxyorcinaldehyde shunt product (Pathway 4, **4.4**).[10] There was an inaccessible barrier for the semipinacol rearrangement, but a feasible barrier for shunt product formation. Instead, the skeletal rearrangement was found to be more plausible through a direct process (Figure 4.1C) that can arise from either a radical-based ring expansion (Pathway 2) or one post-SET (Pathway 3).[61117] From the primary radical intermediate **4.5**, a radical ring expansion to produce the tertiary-like radial intermediate (**4.6**) and a subsequent second H-atom abstraction for the formation of the tropolone tautomer (**4.7**) were found to be feasible. Also from **4.5**, simulation of an SET event to produce the transient zwitterion **4.8** and ground state neutral, ring expanded (**4.9**) intermediates were also found to be feasible. However, these models were small active site mimics with translational and rotational freedom, so they are not a substitute for enzyme active site constraint which would limit these freedoms of motion. In this chapter, simulations within the enzyme will provide insights into key interactions between the substrate and surrounding residues important for positioning the substrate within the enzyme active site. This will lead to better understanding of the kinetics and thermodynamics of the proposed mechanisms and allow for identification of the preferred pathway towards tropolone production within TropC.

To provide detailed atomistic descriptions of SET within complex environments, such as solvent phases and proteins, our group developed a combined constrained density functional theory and molecular mechanics method (CDFT/MM) to model electron transfer processes along reaction pathways by constraining the electron density to the acceptor and donor regions (see Chapter II).[18-20] The full environment of the active

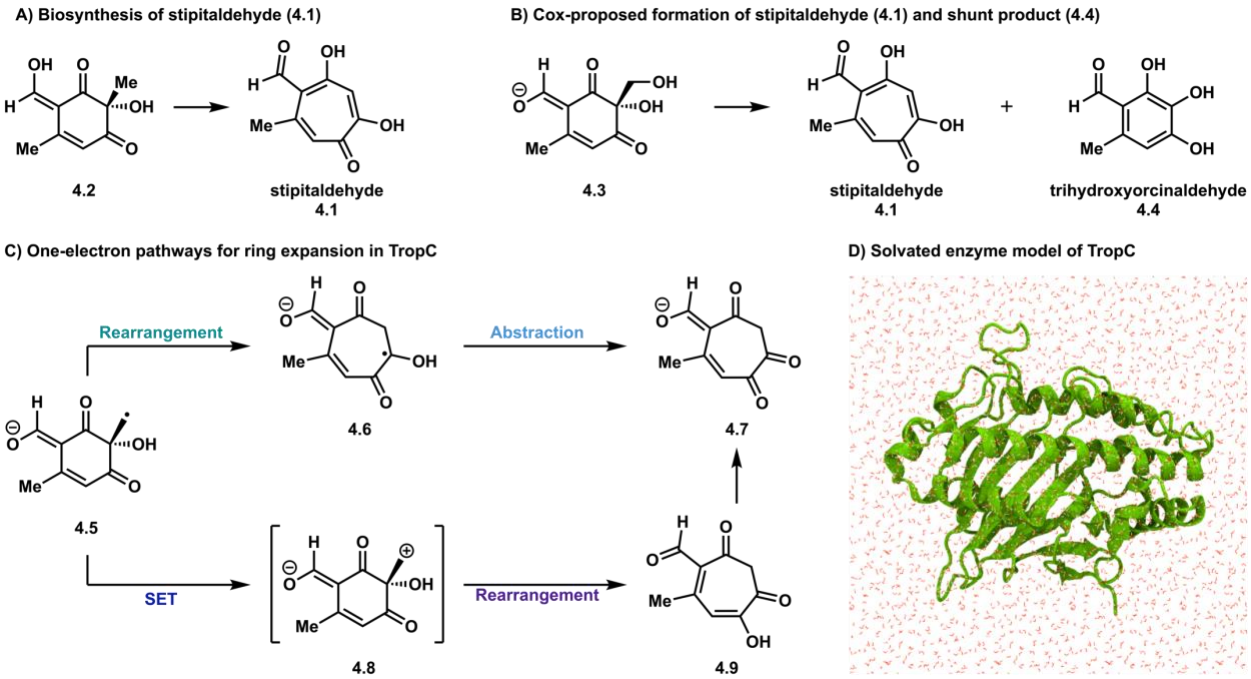

Figure 4.1. Overview of stipitaldehyde formation. Schemes representing A) biosynthesis of stipitaldehyde (**4.1**) b) Cox-proposed mechanism for formation of **4.1** and trihydroxyorcinaldehyde shunt product (**4.4**), C) proposed one-electron pathway for ring-expansion in TropC and related enzymes. D) 3D solvated model of TropC.

site within the enzyme is captured in CDFT/MM, allowing reconfiguration of the entire enzyme in response to ET. CDFT/MM has been achieved before for enzyme systems but were limited to molecular dynamic (MD) simulations to obtain free energies.[21,22] In this full enzyme model, reliable descriptions of regular reaction pathways can be compared to electron-transfer pathways, providing detailed data to delineate the variety of possible reaction mechanisms.[23] This chapter will investigate the potential mechanisms for TropC and related enzymes using solvated enzyme models via QM/MM (Figure 4.1D). Analysis of non-covalent interactions (NCIs) between substrate and active site residues will provide understanding of how TropC positions and constrains the substrate within the active site. Reaction profiles will give insight into the thermodynamics and kinetics of the proposed mechanisms to distinguish between possible mechanisms within TropC and provide support to previous computational and experimental conclusions.

4.2 Model Preparation

To perform QM/MM simulations a 3D model of the enzyme in a solvent environment was constructed and equilibrated. The construction of the 3D model is described in chapter III. Using this model, equilibration was perform using a three step MD protocol, including a total of 10 ns of sampling. Successful equilibration of the model was determined based on an RMSD plot of the protein backbone showing little fluctuations (see section **4.10.3**, Figure S4.1). Two snapshots from the equilibrated enzyme model were selected based on the reaction coordinate parameters between the substrate and the iron-coordinated oxygen of α-KG, specifically the H_1–O_x distance and C_7–H_1–O_x angle (see section **4.10.3**). MM optimizations were performed on each snapshot to reduce high energy interactions within the model and locate the local minimum. Snapshot 1881 was selected based on an H–O distance of 2.36 Å and a C_7–H_1–O_x angle of 128.1°. Although this is a deviation from ideal parameters, it is relative to the coordinated α-KG rather than the iron(IV)–oxo. Therefore, snapshot 1881 was selected for construction of the 3D QM/MM model.

The 3D model based on snapshot 1881 was subjected to a series of non-constrained QM/MM minimizations. Figure 4.2A shows reaction parameters of the enzyme-substrate complex based on this model. The substrate-complex H_1–O_x distance was 2.06 Å and the C_7–H_1–O_x angle 160.7°, reflecting a significant shift toward ideal alignment relative to the MM minimization results (2.36 Å, 128.1°). In addition to these alignments, the H-bonding network mostly remained intact (Figure 4.2B): the exception being two H-bonds between the substrate C_1-carbonyl and Arg186 are 1.93 Å and 2.47 Å, reflecting a weaking of the latter interaction. Although the second H-bond was weakened, a second weak H-bond with the oxygen of iron(IV)–oxo was observed with a

distance of 2.41 Å. The C_3-carbonyl and Phe209 H-bond distance was 2.12 Å, and the C_4-enolate oxygen and Asn282 H-bond was 2.21 Å. Interestingly, three additional H-bonds were observed to have formed: Phe209 and C_2-hydroxyl (2.17 Å) and the glycine residue Gly210 and the C_3-carbonyl (2.21 Å). There were also two nontraditional H-bonds observed: a C–H hydrogen bond between Asp208 and the C_2-hyroxyl (2.37 Å) and an H-bond between the C_2-hyroxyl and the phenyl ring of Phe209. Existence of these H-bonds

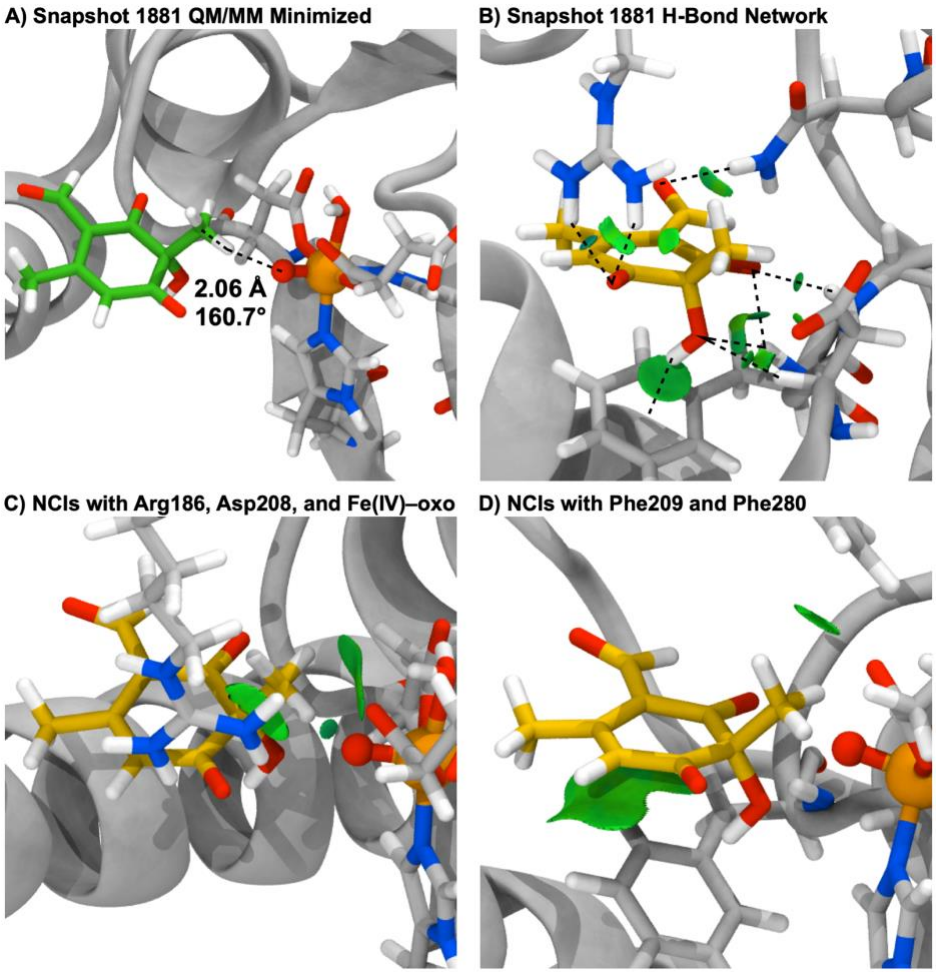

Figure 4.2. QM/MM minimized initial structure. A) reaction coordinate parameters, B) H-bond network, C) NCIs between substrate and Arg186, Asp208, and the iron(IV)–oxo, and D) NCIs between substrate and Phe209 and Phe280. Reaction parameters are more aligned with ideal H-atom abstraction parameters than MD and MM minimization results. H-bond network was maintained and a total of eight interactions were observed. Gray cartoon is the TropC model, gray licorice is α-KG and iron-coordinating residues, green licorice is substrate, yellow licorice is substrate for clarity, and orange and red ball and stick represent the iron(IV)–oxo species. Green surfaces reflect weak interactions and bluish green moderately favorable interactions. Black dashed lines represent H-bonds.

were confirmed via NCI analysis (Figure 4.2B,C). These H-bonds likely play a key role in positioning of the substrate in a reactive pose, so that the methyl group is directed towards the iron(IV)–oxo species. Also observed in the NCI analysis were steric interactions between the phenyl rings of Phe209 and Phe280 and the substrate (Figure 4.2D). Specifically, for Phe280 there is an interaction with the primary methyl group of the substrate, which suggests our hypothesis of its role in substrate positioning is correct. Interestingly there is a favorable interaction between the primary methyl group of the substrate and the iron(IV)–oxo species, and also interactions with Asp208 and Arg186. The combination of steric and H-bond interactions aide in the alignment of the hydrogen for abstraction. Using this 3D model, the first step of H-atom abstraction was investigated.

4.3 Initial H-atom Abstraction

Using the 3D model described in the previous section, simulations examined the H-atom abstraction step (Figure 4.3) leading to the primary radical intermediate **4.5** from the TropC native substrate (**4.2**). The barrier for abstraction was calculated to be 21.7 kcal/mol, and at the transition state (**TS1**) the H_1–O_x length decreased to 1.34 Å from 2.06 Å, and the C7–H_1–O_x angle increased to 167.0° from 160.6°. The QM simulations performed in the previous chapter proposed a barrierless pathway, but the substrate was not restricted by active site residues (see section 4.4 for more detail). Intermediate **4.5** is thermodynamically favored by -2.1 kcal/mol over the reactant complex. Analysis of the structure revealed a slight change in the H-bond network. The most noticeable being the Phe209 to C_2-hydroxyl distance lengthening from 2.17 Å to 2.34 Å and the Phe209 to C_3-

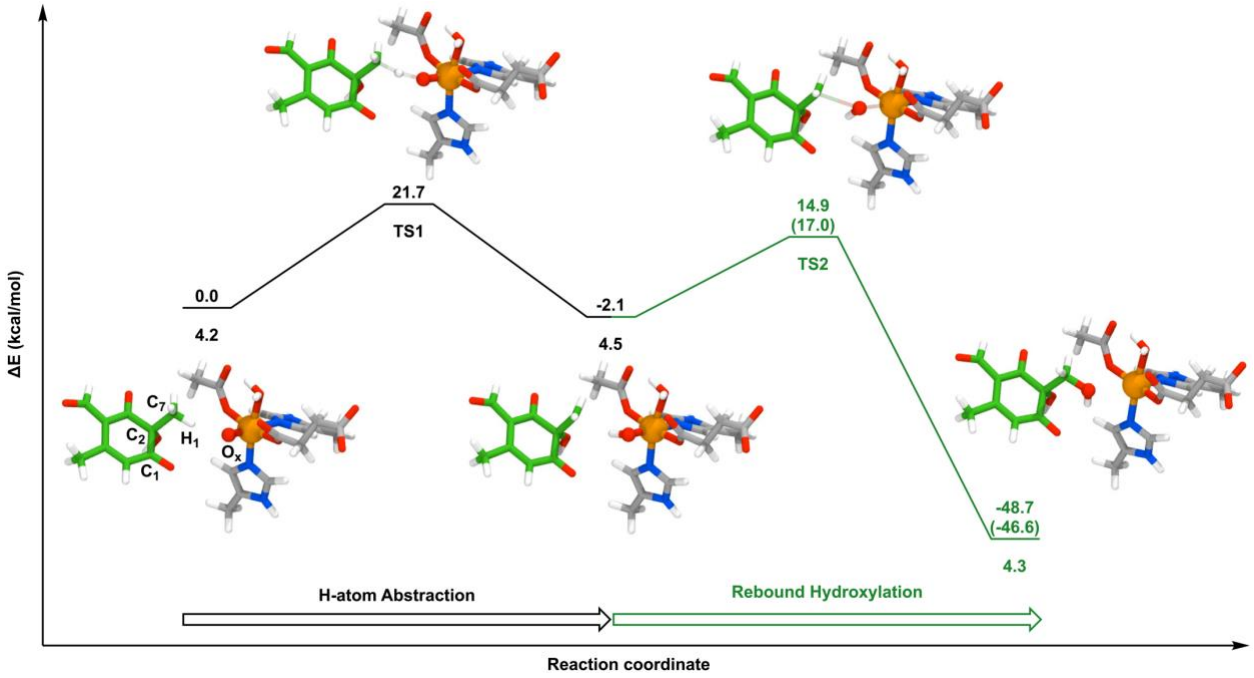

Figure 4.3. Energy diagram for the H-atom abstraction and rebound hydroxylation steps of the Cox-proposed mechanism (Pathway 1). H-atom abstraction is the RDS step of the mechanism. Gray licorice is α-KG and iron-coordinating residues, green licorice is the substrate, orange and red ball and stick represent the iron-oxygen species. Translucent lines reflect bond forming and breaking events at the TSs. Numbers in parentheses are relative elementary step energies.

carbonyl shortening from 2.12 Å to 1.97 Å. This change is reflective of the H-bonds becoming weaker and stronger, respectively. The H-bond between Arg186 and the newly formed iron(III)–OH species decreased from 2.41 Å to 1.79 Å, which is likely restricting the movement of the iron(III)–OH species.

4.4 Rebound Hydroxylation

Following H-atom abstraction the divergence point for the mechanisms discussed in this chapter is formed with primary radical intermediate **4.5**. Pathway 1 is rebound hydroxylation, which follows the mechanism proposed by Cox and coworkers (Figure 4.1B). TS2 was found to have a barrier of 17.0 kcal/mol and a thermodynamic favorability of -48.5 kcal/mol, leading to diol intermediate **4.3** (Figure 4.3). The thermodynamic driving

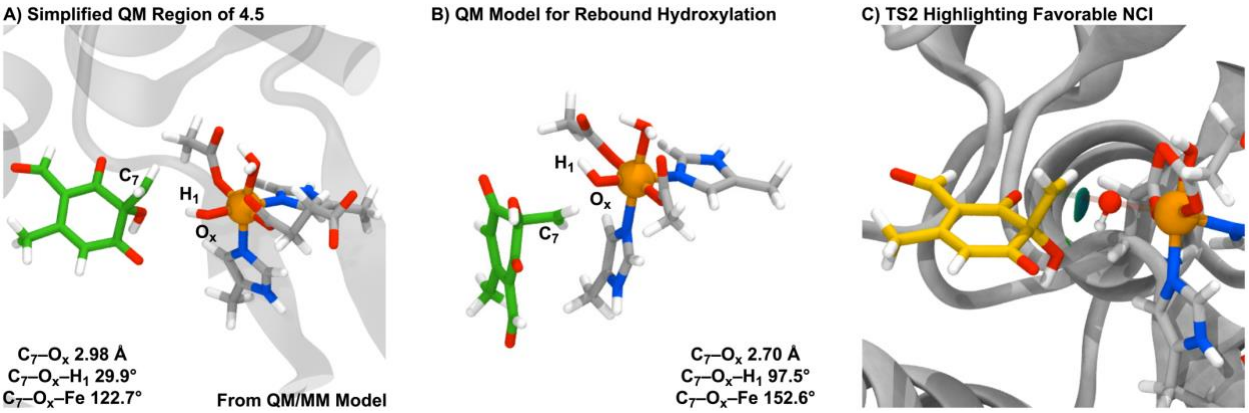

Figure 4.4. 3D models of primary radical intermediate **4.5**. Highlighted are the A) simplified (for clarity) QM region of QM/MM model of **4.5**, B) QM model alignment of primary radical and iron(III)–OH intermediates for rebound hydroxylation, and C) **TS2** highlighting NCI between the substrate and the iron(III)–OH species. Reaction parameters highlighting differences in QM and QM/MM models. There is a favorable interaction between C_7 and the transferring hydroxyl group shown by a bluish-green surface. Gray cartoon is the TropC model, gray licorice is α-KG and iron-coordinating residues, green licorice is substrate, yellow licorice is substrate in presence of NCIs for clarity, and orange and red ball and stick represent the iron(III)–OH species. Bluish green moderately favorable interactions.

force is consistent with the QM simulations in Chapter III, which resulted in -46.6 kcal/mol. The QM/MM model (Figure 4.4A) results contrast the QM results, however, in that the QM model predicts a barrierless pathway. This was attributed to the translation and rotational freedom of the QM model (Figure 4.4B). Within the enzyme active site, the substrate is constrained by residues via H-bonds and steric interactions described in section 4.2. Arg186 H-bonds with the iron(III)–OH, which is likely restricting the movement of the iron(III)–OH and is resulting in a barrier. Observed at **TS2** was expected weakening of the O–Fe bond from 1.80Å to 2.25 Å and partial formation of the C–O bond from 2.98 Å to 2.38 Å. The barrier is attributed in part to the rotation of the hydroxyl group, where in **4.5** the C_7–O_x–H_1 angle was initially 29.9° but increases to 63.1° at **TS2**. Also observed was a favorable interaction between C_7 and the transferring hydroxyl group (Figure 4.4C), which likely contributes to the feasibility of this reaction. There is also an initial decrease in the Arg186 to iron(III)–OH H-bond from 1.79 Å to 1.73 Å, but then an increase to 2.01

Å following formation of **4.3**. The remaining H-bond network remained intact in both **TS2** and **4.3**, but with continued weaking of the H-bond between Phe209 and the C_2-hydroxyl and increased strengthening of C_3-carbonyl. The distances of the H-bonds are 2.49 Å and 2.01 Å respectively in **4.3**. Overall, both steps are kinetically feasible and **4.3** is thermodynamically favorable. These simulations support experimental evidence that rebound hydroxylation can be facilitated by TropC. TropC positions the substrate within the active site via an H-bond network, and the methyl group is directed towards the iron(IV)–oxo via steric interactions with Phe280.

4.5 Radical-based Ring Expansion Pathway

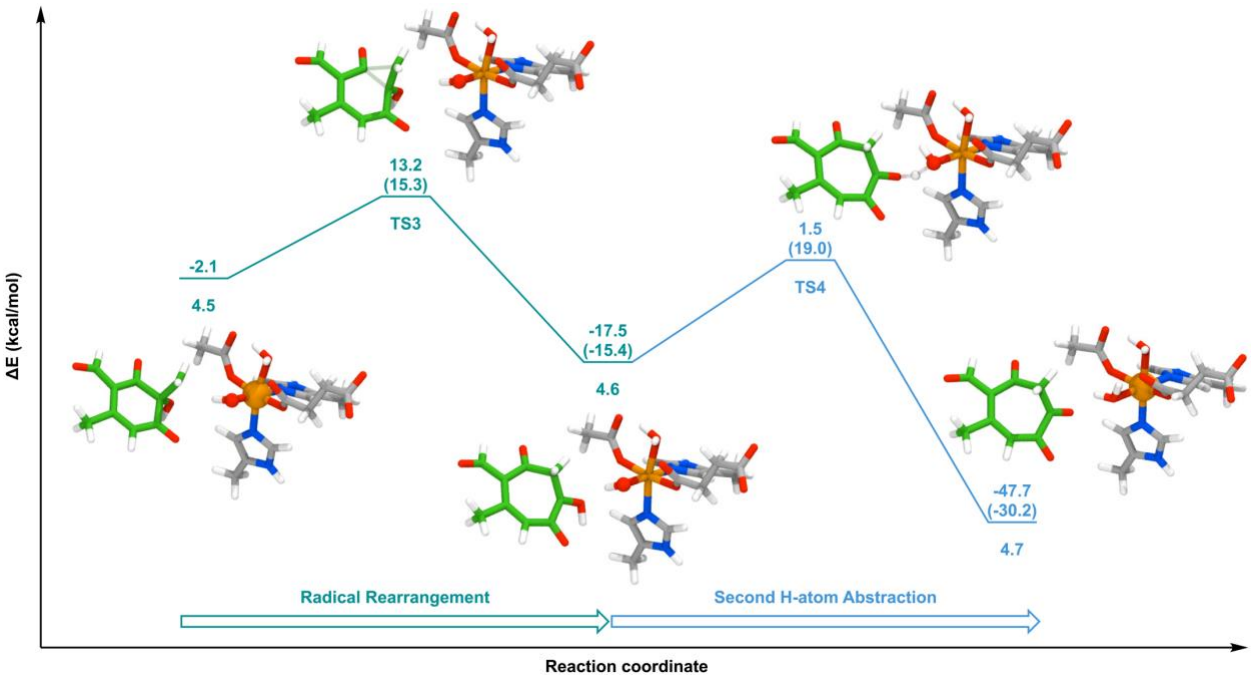

Figure 4.5**.** Energy diagram for the radical-based ring expansion and second H-atom abstraction steps the radical-based ring expansion pathway (Pathway 2). Both activation energies are lower than H-atom abstraction (21.7 kcal/mol). High barrier for second H-atom abstraction is attributed to disruption of H-bonding network. Gray licorice is α-KG and iron-coordinating residues, green licorice is the substrate, orange and red ball and stick represent the iron-oxygen species. Translucent lines reflect bond forming and breaking events at the TSs. Numbers in parentheses are relative elementary step energies.

Pathway 2 diverges from the primary radical intermediate **4.5** via a direct radical-based ring expansion to produce the tertiary-like radical intermediate **4.6** (Figure 4.5). This transformation was simulated to have a barrier of 15.3 kcal/mol (**TS3**). At **TS3** the breaking C_2–C_3 bond and forming C_3–C_7 bonds are stretched to distances of 1.66 Å and 1.81 Å, respectively. A change in H-bonding network was observed between **4.5** and **4.6**. The H-bond between Phe209 and the C_2–hydroxyl group is lost, which is a result of the

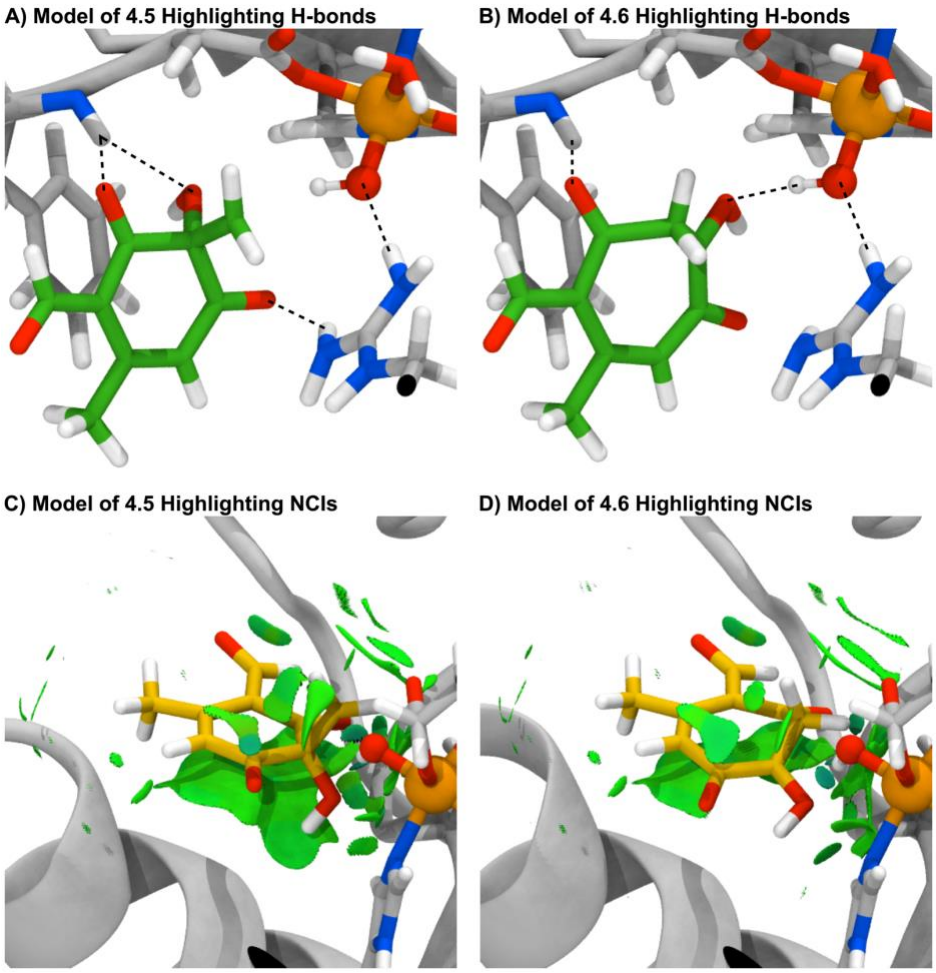

Figure 4.6. 3D models of primary radical intermediate **4.5** and tertiary radical intermediate **4.6**. Highlighted are the H-bond network for A) primary radical (**4.5**) and B) tertiary radical (**4.6**) intermediates, and all NCIs between TropC and C) **4.5** and D) **4.6**. Noticeably less steric interactions along the C_5-methyl portion of the substrate of each image indicating room for ring expansion and larger substrates. H-bonds shown in dashed black lines. Gray cartoon is the TropC model, gray licorice is α-KG and iron-coordinating residues, green licorice is substrate, yellow licorice is substrate in presence of NCIs for clarity, and orange and red ball and stick represent the iron(III)–OH species. Green surfaces reflect weak interactions and bluish green moderately favorable interactions.

ring undergoing expansion and the hydroxyl group moving away from Phe209. However, the H-bond between the C_2–hydroxyl and Phe209 is replaced by an H-bond between the C_2–hydroxyl and iron(III)–OH species (Figure 4.6B). There is also a loss of H-bonding between Arg186 and the C_1–carbonyl group as a result of the ring expansion. Figure 4.6 displays the full NCIs for intermediates **4.5** and **4.6** and TropC. There are considerably less steric interactions on the C_5-portion of the substrate, when compared to the iron-facing portion. This would suggest ample room for ring expansion as well as potential for larger substrates, which was explored in a related system by Dr. Tyler Doyon.[24] Following ring expansion there are decreased steric interactions (**4.6**), which is a result of the substrate becoming more planar.

4.6 Second H-atom abstraction

For the radical-based ring expansion pathway to be viable towards the formation of the tropolone tautomer **4.7**, the second H-atom abstraction must be feasible, wherein the tertiary-like radical is quenched to return the substrate to a closed shell intermediate. H-atom abstraction from the C_2–hydroxyl of **4.6** resulted in quenching of the tertiary-like radical and formation of **4.7** occurred with a barrier of 19.0 kcal/mol (**TS4**). The barrier is attributed to the rotation of the C_2–hydroxyl group and breaking of the newly formed H-bond between C_2–hydroxyl and iron(III)–OH groups. However, once the reaction is complete an H-bond between the newly formed water molecule and C_2–carbonyl is formed, which likely adds to the thermodynamic favorability of -30.3 kcal/mol for the formation of **4.7**.

4.7 SET-based Ring Expansion Pathway

The final postulated pathway from the primary radical intermediate **4.5** involves a SET from the substrate to the iron(III)–OH species (Pathway 3). SET results in the unstable primary zwitterionic intermediate **4.8** that is expected to lead to the neutral, ring expanded intermediate **4.9** and iron(II)–OH (Figure 4.1C). In particular, the presence of a primary carbocation in **4.8** results in a strong thermodynamic driving force for ring expansion to form a more stable tertiary-like carbocation. After ring expansion, intramolecular proton transfer quenches the carbocation and enolate, giving an overall neutral product. Using the CDFT/MM method described in Chapter II, SET from **4.5** gives **4.8** at a vertical (unoptimized geometry) energy cost of +138.0 kcal/mol. A decrease in spin density from 0.986 e^- to -0.019 e- was observed on C_7 of intermediates **4.5** and **4.8**, and an increase in spin density from 3.004 e^- to 3.724 e^- on the iron atom, confirming that SET moved an electron from the substrate **4.5** to iron(II)–OH (Figure 4.7). Relaxation of the geometry of **4.8** leads to barrierless ring expansion, with an energy 3.7 kcal/mol above species **4.5**. Instead of direct formation of **4.9** via this ring expansion, the zwitterionic intermediate **4.10** was produced, which differs from **4.9** in that there is a hydrogen bond between the C_2–protonated carbonyl of **4.10** and a solvent water molecule that prevented immediate proton transfer. As predicted, the primary carbocation of **4.8** was a strong thermodynamic driving force for ring expansion and resulted in a much more thermodynamically favorable tertiary-like carbocation on **4.10**, which is also stabilized by a lone pair of electrons on oxygen. Intramolecular proton transfer of **4.10** resulted in the formation of **4.9** (Figure S4.4).

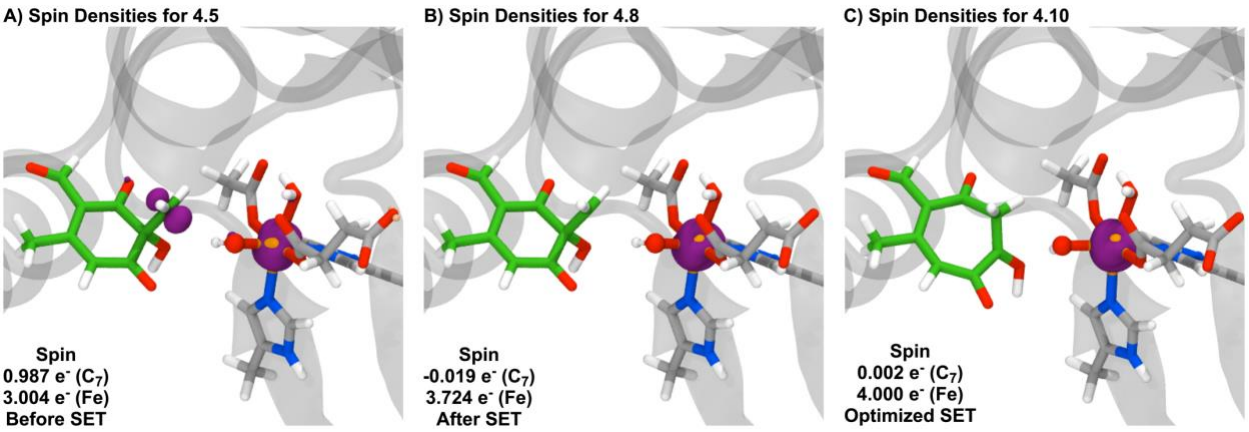

Figure 4.7. QM/MM spin density analysis along SET pathway. There is a decrease in spin on C_7 and increase in spin of Fe atom for A) intermediate **4.5** to form B) **4.8**. Spin is maintained after optimization of **4.8** in formation of C) **4.10**. Gray licorice is α-KG and iron-coordinating residues, green licorice is the substrate, orange and red ball and stick represent the iron-oxygen species.

Starting from species **4.5**, the vertical energy gap of +138.0 kcal/mol for SET is insurmountable and suggests that SET cannot occur from this specific geometry. It remains possible, however, that a geometry between **4.5** and **4.9'** could undergo SET. Therefore attempts were made to find crossing points, where the ground (radical) electronic state became nearly degenerate with the SET electronic state. To do so, double-ended growing string method (GSM) simulations and minimum energy crossing point (MECP) optimizations were performed using a QM cluster model, as described in the computational details. First, radical (Pathway 2) and SET ring expansion (Pathway 3) pathways were optimized starting at the geometry for **4.5**. The radical pathway resulted in a barrier of 23.8 kcal/mol and a thermodynamic favorability of -4.2 kcal/mol (Figure 4.8). Similar to the QM/MM optimization model, the SET pathway to **4.10** has a large vertical energy gap (60.6 kcal/mol) and is uphill by 6.8 kcal/mol with respect to **4.5**. SET was confirmed to have occurred as a decrease in spin from 1.010 e^- to -0.012 e^- on C_7 of species **4.5** and **4.8**, and an increase in spin from 2.883 e^- to 3.805 e^- for the iron atom. See S4.10.6 for details on the differences between the QM cluster and QM/MM models.

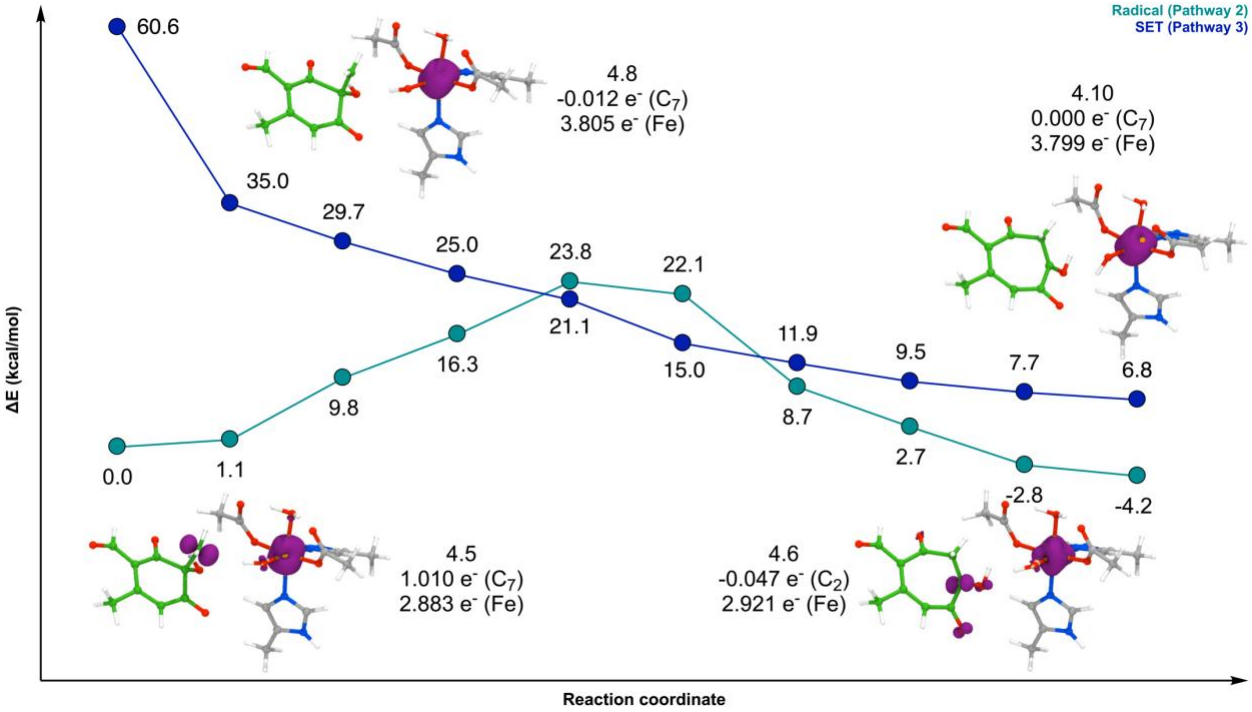

Figure 4.8. QM energy profiles for the radical- and SET-based ring expansion. Radical pathway (cyan) has barrier of 23.8 kcal/mol and SET pathway (blue) is entirely downhill. Nodes four through eight were selected for MECP optimizations. Green ball and sticks are the substrate, and the gray ball and sticks are the iron ligands.

As seen in figure 4.8, structures four through eight of Pathways 2 and 3 are in the vicinity of an electronic state crossing. This suggests there is a possibility for SET in the transition state region, where the radical path is relatively high energy, and the SET pathway is dropping in energy (Figure 4.8). Therefore, the structures of nodes four to eight from the radical string were subjected to MECP optimizations (see S4.10.6 for more details). To see how the radical and SET pathways might interact along the reaction coordinate to the ring-expanded product, the C_3-C_7 distances were frozen for structures four to eight. The MECP optimizations produced a series of structures following the reaction path for radical-based ring expansion. Two of the structures, with C_3-C_7 distance of 1.67 Å and 1.61 Å, showed the radical and SET states being close in energy, specifically within 4 kcal/mol of each other. This indicates—from a qualitative

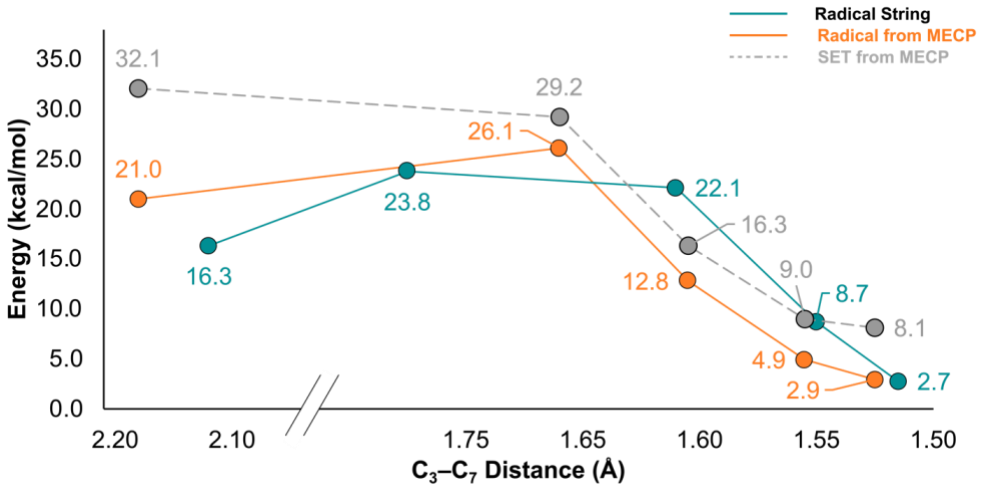

Figure 4.9. Energy analysis of nodes from MECP optimizations and radical string. The smallest energy gap between the radical (orange) and SET (gray) states 3.1 kcal/mol (node 5), followed by 3.4 kcal/mol (node 6). Results show the possibility for MECP following the transition state, but model limitations prevented location of MECP.

perspective—that the SET state may be accessible along the ring-expansion reaction path. One of the near-crossing points that was located was around 13 kcal/mol above the starting structure (**4.5**), corresponding to a C_3–C_7 distance of 1.61 Å, indicating that crossing is energetically accessible. Regardless, these near-crossings occur after the transition state of the radical-based ring expansion. In other words, the current model does not predict that the SET path is competitive with the radical-based ring expansion, as the rate-limiting barrier for ring expansion will be crossed before SET is likely to occur. SET may therefore occur after ring expansion, but it is not a necessary step in the overall pathway.

Analysis of the spin density of the near-crossing geometry with a C_3–C_7 distance of 1.61 Å (Figure 4.10A) provided insights into why the ground and SET states are close in energy. Figure 4.10D shows the geometry and spin density along the radical-based ring expansion pathway, where the 4 unpaired electrons are distributed as 1 on the substrate and 3 on the iron. For the ground state of the near-crossing geometry, the spin

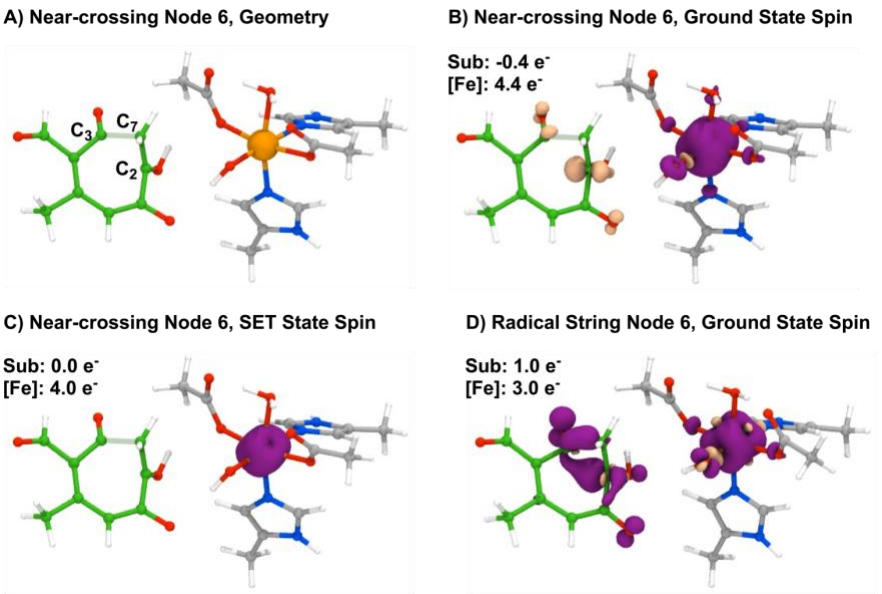

Figure 4.10. Analysis of spin densities at C_3–C_7 distance of 1.61 Å. A) Near crossing geometry, B) ground state spin density, C) SET state spin density, and D) spin density of node six from radical string. Green ball and sticks are the substrate, gray ball and sticks are the iron ligands, and the orange ball is iron atom. Purple density reflect positive spin density and pink density reflect negative spin density. Isovalues for spin density were set to +/- 0.1.

on the substrate is -0.4 e- (beta density) and the spin on the iron center is 4.4 e^-, which corresponds to exchange of electrons between the substrate and the iron complex (Figure 4.10B). The non-unit spin densities indicate a significant interaction between the two moieties, consistent with a strong coupling between the valence electrons. For the SET state at the near-crossing geometry, the expected spin density of 4.0 e^- is seen for the iron complex and 0.0 e^- for the substrate (Figure 4.10C). The two electronic states at the near-crossing geometry (Figure 4.10C and D) therefore resemble one another, providing an explanation for why the two states are close together in energy. While these results are close to the limit of what DFT simulations can delineate, the hypothesis that state crossing is plausible is supported by these results.

4.8 Discussion

Comparison of the optimization results of structure **4.8** from the QM/MM and QM models differ from the gas-phase model discussed in Chapter III. QM/MM and QM models herein resulted in ring expansion following SET, whereas a ketene intermediate was formed upon optimization of **4.8** in Chapter III. This is due to having modeled the active site and how it constrains the substrate in such that formation of the ketene intermediate was not feasible. The difference in QM/MM, QM cluster-like, and gas phase QM models highlights the need to account for an enzyme active site to make concrete proposals on enzyme catalyzed reactions.

Comparison of the ring expansion pathways following rebound hydroxylation (Pathway 1), radical-based ring expansion (Pathway 2), and SET-based ring expansion (Pathway 3) suggests that two of the three pathways are viable within the active site of TropC. The determining factor for deciding the dominant pathway is the distance between the substrate and the iron center, and more precisely the location of the generated radical relative to the iron center. Arguments for the preference of radical rearrangement over rebound hydroxylation are akin to those for why halogenation is preferred over hydroxylation in α-KG NHI enzymes.[25] Positioning of the active site residues restricts the substrate from forming a rebound hydroxylation favored position, such that the Pathway 2 is kinetically favored over Pathway 1. Despite the radical-based ring expansion being kinetically preferred (15.3 kcal/mol), the barrier for rebound hydroxylation (17.0 kcal/mol) supports the experimental conclusion that rebound hydroxylation is possible. Pathway 3 is also potentially viable within TropC and may also subjected to a distance dependance like other another α-KG NHI enzyme.[17] There likely exists a position within TropC in which

the SET is more accessible and favorable over radical-based ring expansion, but this would require the substrate to be positioned closer to the iron center. This closer placement to the iron would mean that the SET would not compete with the radical-based ring expansion, but instead the rebound hydroxylation process. Based on the results of this snapshot, Pathway 2 is suggested to be the dominant pathway.

4.9 Conclusions

Based on the simulations performed and discussed in this chapter, the dominant pathway within TropC is a one-electron, radical-based pathway (Pathway 2) towards the production of the TropC tropolone native product stipitaldehyde (**4.1**). Diverging from the common primary radical intermediate **4.5** the barriers for Pathways 1 and 2 were found to be competitive. Previously, Cox and coworkers proposed rebound hydroxylation followed by a subsequent semipinacol expansion (Pathway 1). Although our simulations support rebound hydroxylation as a viable pathway, previous QM results suggest this pathway leads to production of the trihydroxyorcinaldehyde shunt product (Pathway 4, **4.4**). This would account for experimental observations of both tropolone and shunt product formation (see Chapter III). The SET-based pathway (Pathway 3) was found to be potentially viable, but the current model is limited in its ability to accurate reflect the enzyme active site. A near-crossing structure with an energy gap of 3.4 kcal/mol between the radical and SET states was located, and the small gap was reasoned based on spin densities reflecting electron exchange in the ground state. NCIs were mapped throughout the elementary reaction steps to provide insight into the role of the active site residues. There were steric interactions between both Phe209 and Phe280 that support

experimental observations and our hypothesis that these residues play a role in positioning of the primary methyl group via steric interaction, which results in almost ideal alignment for initial H-atom abstraction (2.06 Å, 160.7°). Additional steric interactions between the C_7 methyl group and Arg186 and Asp208 further position H_1 for abstraction. In addition to these steric interactions, an H-bond network between the substrate and residues Arg186, Phe209, Gly210, and Asn282 were also observed and mostly maintained, and are responsible for further substrate positioning. There was also an H-bond between Arg186 and the iron(IV)–oxo that became stronger as the reaction coordinate proceeded and likely restricts the movement of the oxygen atom. The results in this chapter further support our previous experimental observations and QM simulations and add to the growing literature of radical- and SET-based events within α-KG NHI enzymes.

4.10 Computational Details

All 3D images were generated using the Visual Molecular Dynamics software program.[26] Isovalues for spin densities were set t0 +/- 0.3 unless otherwise stated.

4.10.1 General QM/MM Details

All simulations were performed using the unrestricted B3LYP hybrid density functional[27,28], the 6-31G* basis set[29-34], and D3(BJ) empirical dispersion correction[35]. Initial H-atom abstraction, rebound hydroxylation, radical-based ring expansion, and second H-atom pathways were all simulated using pDynamo[36] 1.9.0 using the conjugate gradient optimizer and ORCA 4.2.1[37] for the QM region. Remaining pathways were

simulated using a development version of pDynamo[36] 3.0.9 utilizing the L-BFGS optimizer and QChem 5.2.[38]

4.10.2 Equilibration Check of Model

RMSD plot of the protein backbone (Figure S4.1), excluding flexible loop regions, shows the enzyme is equilibrated throughout the 10 ns sampling phase relative to the initial solvated enzyme model.

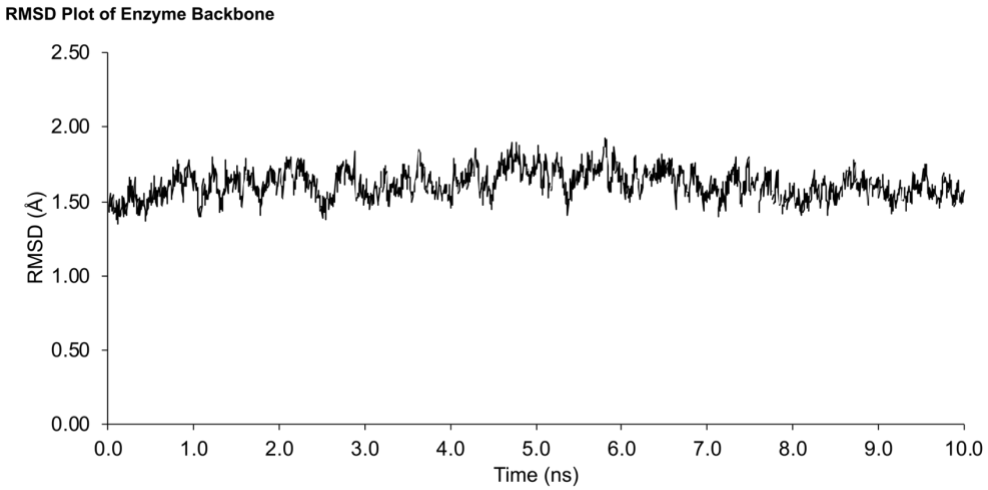

Figure 4S.1. Plot of RMSD backbone relative to initial solvated enzyme model. Plot shows 3D enzyme model is equilibrated throughout the 10 ns sampling phase of MD.

4.10.3 Snapshot Selection

An H-bond between the arginine Arg321 and aspartate Asp100 residues was observed from frame 352 and onwards (Figure S4.2). Therefore, the first 351 frames were not further analyzed. From the remaining frames, two frames were selected based on the reaction coordinate parameters of H_1–O_x distance and C_7–H_1–O_x angle, with respect to the coordinated oxygen of α-KG. Frame 657 had an H_1–O_x distance and C_7–H_1–O_x angle of 2.08 Å and 132.0° respectively (Figure S4.3A), and frame 1881 2.21 Å and 142.2° respectively (Figure S4.3B). These frames were then each subjected to 5000 ns of MM

minimization. Afterwards, the H_1–O_x distance and C_7–H_1–O_x angle for frame 657 were 2.46 Å and 107.4°. Due to this deviation, frame 657 was no longer considered. MM minimization of frame 1881 resulted in a H_1–O_x distance of 2.36 Å and a C_7–H_1–O_x angle of 128.05. Although this is a deviation from ideal parameters, it was not as severe as frame 657. Therefore, frame 1881 was selected for construction of the 3D QM/MM model.

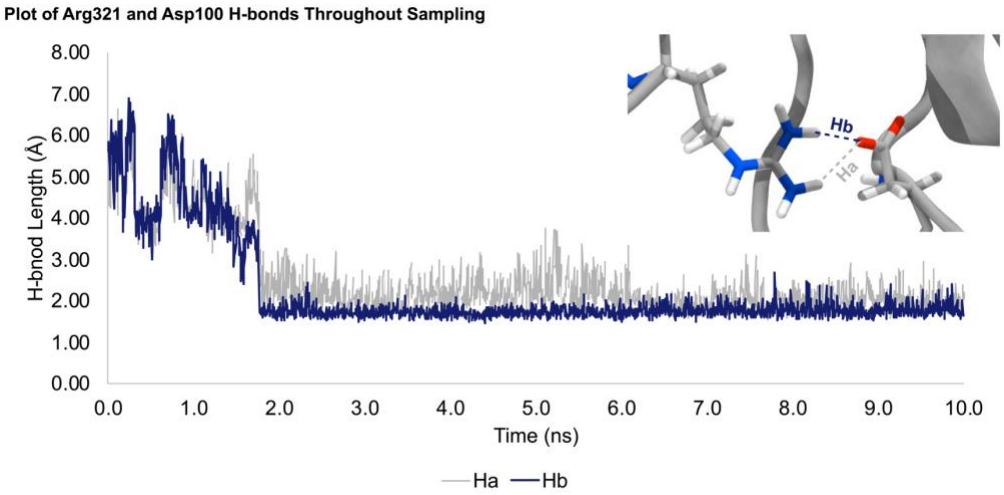

Figure 4S.2. Plot of H-bonds lengths between Arg321 and Asp100. Plot shows contacts made after frame 351 are generally maintained for remainder of sampling.

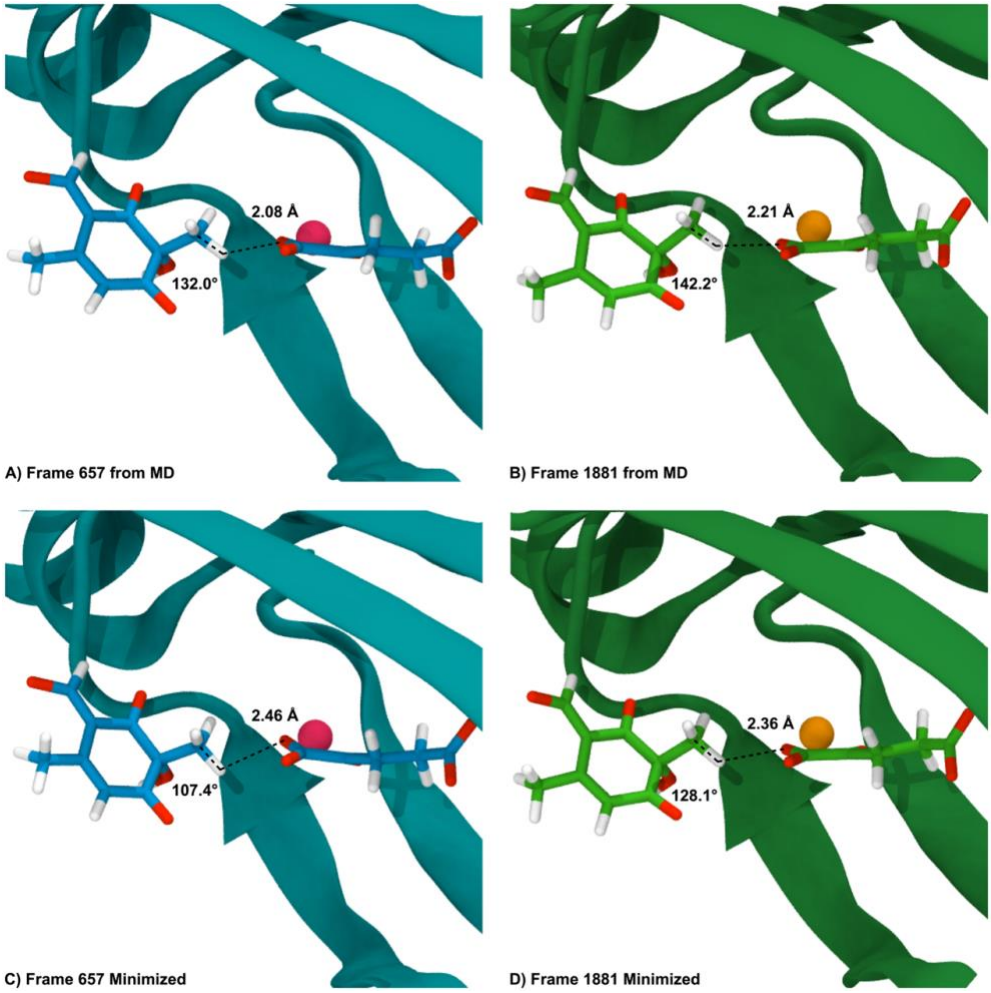

Figure 4S.3. Snapshots selected from MD simulations for QM/MM simulations. Snapshot 657 (blue) from A) MD sampling and B) minimization post-MD. Snapshot 1881 (green) from C) MD sampling and D) minimization post-MD. Both snapshots show a deviation from ideal H-atom abstraction parameters. Snapshot 1881 chosen as deviation was not as severe as snapshot 657. Enzyme shown in cartoon, substrate and α-KG in licorice, and iron atoms in CPK.

4.10.4 QM/MM Model Preparation

Model taken from MD was pruned of all water molecules and counterions to within 30 Å of the iron center. All atoms within 20 Å of the iron center were free to move, and the outer 10 Å shell was frozen. Link atoms were automatically placed by pDynamo. Total atoms in the QM region were 98. Initial BP86[39,40] D3 LANL2DZ[41,42] minimization was performed for all atoms in the QM region to achieve reliable initial structure[43], which was followed by

a B3LYP D3 6-31G* minimization. All atoms in the MM region were treated using the CHARMM36[44] forcefields.

4.10.5 H-bond Network Comparison Between Tropolone Tautomers

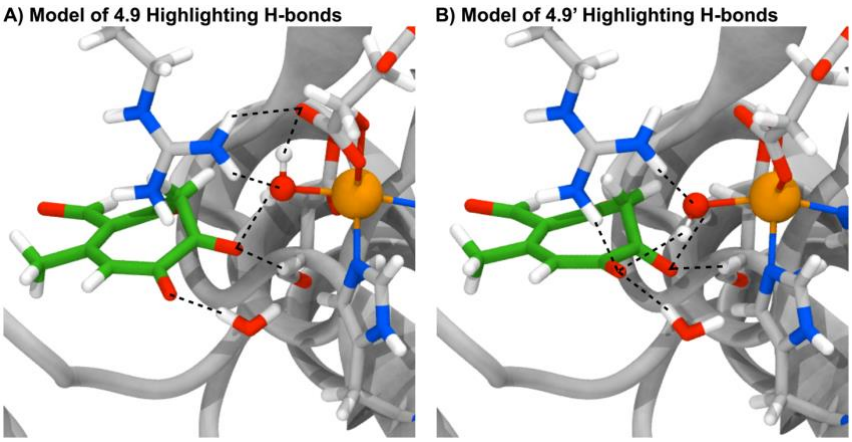

Figure 4S.4. Comparison of tautomer products from radical and SET pathways. A) **4.9** from radical pathway and B) **4.9'** from SET pathway. **4.9'** is higher in energy partially due to a different, less favorable H-bonding network (black dashed lines). Gray licorice is α-KG and iron-coordinating residues, green licorice is the substrate, orange and red ball and stick represent the iron-oxygen species.

The tropolone tautomer (**4.9**) generated from the radical- and SET-based pathways (Pathways 2 and 3, respectively) differ in their H-bond networks. In part due to this difference in H-bond network, **4.9** from Pathway 3 is higher in energy than **4.9** from Pathway 2. The difference in H-bonding network arose from the direction in which the proton was oriented throughout the deprotonation process relative to the H-atom in the second H-atom abstraction process.

4.10.6 QM Simulations

General Computational Details

QM simulations and optimizations with gas-phase, cluster-like QM models were performed within QChem with B3LYP D3(BJ) 6-31G*. The QM region containing the primary radical intermediate **4.5** and iron(III)–OH complex from the radical pathway were obtained from the QM/MM model. The terminal atoms of each ligand were fixed, as well as the amine nitrogen atoms to simulate hydrogen bonding within the enzyme. Along with the fixed atoms on the iron ligands, atoms of the substrate were also selected to be fixed to simulate the active site of TropC. Based on an overlay (Figure S4.5) of **4.5**, the tertiary radical intermediate (**4.6**), and the expanded zwitterionic intermediate **4.10** the atoms with stars were selected to be fixed. In addition to these fixed atoms, external force explicitly included (EFEI) and SMD (water) was used to maintain coordination of the ligands to iron. A force constant of 1.0 nN was used for EFEI.

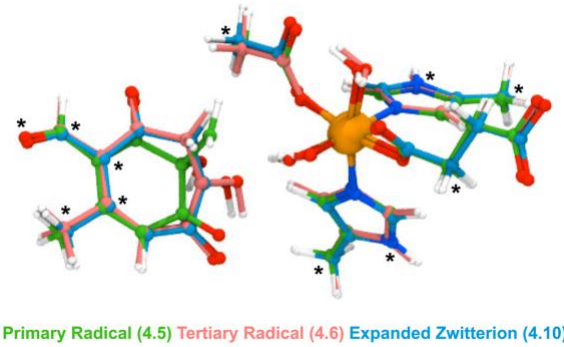

Figure 4S.5. QM region containing primary radical (**4.5**), tertiary radical (**4.6**), expanded zwitterionic intermediate (**4.10**), and iron complexes from QM/MM simulations. Atoms with stars were selected to be fixed. Gray licorice is α-KG and iron-coordinating residues, green balls and sticks are the **4.5** system, light blue balls and sticks are the **4.6** system, pink balls and sticks are the **4.10** system, and the orange ball is the iron atom.

Reaction Pathway Simulations

All reaction pathway simulations and optimization used the fixed atoms, EFEI, and SMD approach. Initial reaction pathways were simulated using the single-ended GSM to obtain **4.6** from **4.5**.[45] CDFT was then used to obtain **4.10** from **4.6**. Double-ended GSM[46,47] simulations were performed between **4.5** and **4.6** for the radical pathway and using CDFT the primary zwitterionic intermediate **4.8** (**4.5** with CDFT constraints) and **4.10** for the SET pathway.

MECP Optimizations

MECP optimizations were performed using code developed within the Zimmerman group.[48] Nodes three through eight from the radical pathway were selected for MECP optimization. In addition to the fixed atoms, EFEI, and SMD, additional fixed atoms were selected along the reaction coordinate to avoid collapse into lower energy MECPs. All nodes were converged to within 0.001 Ha/Bohr RMS gradient. The difference energies for all nodes were found range between within 3.1 kcal/mol to 11.1 kcal/mol (Table S1).

Table S4.1 Energy analysis of MECP optimized nodes four through eight. No nodes resulted in an MECP, but near-crossing are located at nodes five and six.

Structure	Energy (kcal/mol)	Difference Energy (kcal/mol)	ΔE, DFT (kcal/mol)	ΔE, CDFT (kcal/mol)
4.5	-3052.9328594387	-	0.0	-
Node 4	-3052.9328594387	11.1	21.0	32.1
Node 5	-3052.92469385	3.1	26.1	29.2
Node 6	-3052.9458869419	3.4	12.8	16.3
Node 7	-3052.9585972157	4.2	4.9	9.0
Node 8	-3052.9616831612	5.2	2.9	8.1

4.11 References

[1] Gao, S.-S.; Naowarojna, N.; Cheng, R.; Liu, X.; Liu, P. Recent Examples of α-Ketoglutarate-dependent Mononuclear Non-haem Iron Enzymes in Natural Product Biosynthesis. *Nat. Prod. Rep.* **2018**, *35*, 792–837.

[2] Martinez, S.; Hausinger, R. P. Catalytic Mechanisms of Fe(II)-and 2-Oxoglutarate-dependent Oxygenases. *J. Biol. Chem.* **2015**, *290*, 20702–20711.

[3] Kal, S.; Que, L. Dioxygen Activation by Nonheme Iron Enzymes With the 2-His-1-Carboxylate Facial Triad That Generate High-Valent Oxoiron Oxidants. *J. Biol. Inorg. Chem.* **2017**, *22*, 339–365.

[4] Wu, L.-F.; Meng, S.; Tang, G.-L. Ferrous Iron and α-Ketoglutarate-dependent Dioxygenases in the Biosynthesis of Microbial Natural Products. *Biochim. Biophys. Acta* **2016**, *1864*, 453–470.

[5] Tang, M.-C.; Zou, Y.; Watanabe, K.; Walsh, C. T.; Tang, Y. Oxidative Cyclization in Natural Product Biosynthesis. *Chem. Rev.* **2017**, *117*, 5226–5333.

[6] Dunham, N. P.; Chang, W.-.; Mitchell, A. J.; Martinine, R. J.; Zhang, B.; Bergman, J. A.; Rajakovich, L. J.; Wang, B.; Silakov, A.; Krebs, C.; Boal, A. K.; Bollinger, J. M. Jr. Two Distinct Mechanisms For C–C Desaturation By Iron(II)- and 2-(Oxo)Glutarate-dependent Oxygenases: Importance of α-Heteroatom Assistance. *J. Am. Chem. Soc.* **2018**, *140*, 7116–7126.

[7] Dunham, N. P.; Mitchell, A. J.; Del Río Pantoja, J. M.; Krebs, C.; Bollinger, J. M. Jr.; Boal, A. K. A-Amine Desaturation of D-Arginine by the Iron(II)- and 2-(Oxo)Glutarate-dependent L-Arginine 3-Hydroxylase, VioC. *Biochemistry* **2018**, *57*, 6479–6488.

[8] Goldberg, N. W.; Knight, A. M.; Zhang, R. K.; Arnold, F. H. Nitrene Transfer Catalyzed by a Non-heme Iron Enzyme and Enhanced by Non-Native Small-Molecule Ligands. *J. Am. Chem. Soc.* **2019**, *141*, 19585–19588.

[9] Vila, M. A.; Steck, V.; Giordano, S. R.; Carrera, I.; Fasan, R. C–H Amination via Nitrene Transfer Catalyzed by Mononuclear Non-heme Iron-dependent Enzymes. *Chembiochem* **2020**, *21*, 1981–1987.

[10] Doyon, T. J.; Skinner, K. C.; Yang, D.; Mallik, L.; Wymore, T.; Koutmos, M.; Zimmerman, P. M.; Narayan, A. R. H. Radical Tropolone Biosynthesis. *ChemRxiv*, **2020**, DOI: 10.26434/chemrxiv.12780044.v1

[11] Dunham, N. P.; Mitchell, A. J.; Del Río Pantoja, J. M.; Krebs, C.; Bollinger, J. M. Jr.; Boal, A. K. α-Amine Desaturation Of D-Arginine by the Iron(II)- and 2-(Oxo)Glutarate-dependent L-Arginine 3-Hydroxylase, VioC. *Biochemistry* **2018**, *57*, 6479–6488.

[12] Chang, W.-C.; Yang, Z.-J.; Tu, Y.-.; Chien, T.-C. Reaction Mechanism of a Nonheme Iron Catalyzed Oxidative Cyclization via C–C Bond Formation. *Org. Lett.* **2019**, *21*, 228–232.

[13] Huang, J. -L.; Tang, Y.; Yu, Chng, -P.; Sanyl, D. Jia, X.; Guo, Y.; Chang, W. -C. Mechanistic Investigation of Oxidative Decarboxylation Catalyzed by Two Iron(II)- and 2-Oxoglutarate-Dependent Enzymes. *Biochemistry*, **2018**, *57*, 1838–1841.

[14] Chen, T.-Y.; Xue, S.; Tsai, W.-C.; Chien, T.-C.; Guo, Y.; Chang, W.-C. Deciphering Pyyrolidine and Olefin Formation Mechanism in Kainic Acid Biosynthesis. *ACS Catal.* **2021**, *11*, 278–282.

[15] Li, H.; Zhu, W.; Liu, Y. Mechanism of the Uncoupled Carbocylization and Epimerization Catalyzed by Two Non-heme Iron/α-Ketoglutarate Dependent Enzymes. *J. Chem. Inf. Model.* **2019**, *59*, 5086–5098.

[16] Li, H.; Liy, Y. Mechanistic Investigation of Isonitrile Formation Catalyzed by the Nonheme Iron/α-KG-dependent Decarboxylase (ScoE). *ACS Catal.* **2020**, *10*, 2942–2957.

[17] Yan, L. Lui, Y. Insights Into the Mechanism and Enantioselectivity in the Biosynthesis of Ergot Alkaloid Cycloclavine Catalyzed by Aj_EasH From *Aspergillus Japonicus*. *Inorg*. *Chem*. **2019**, *58*, 13771–13781.

[18] Kaduk, B.; Kowalczyk, T.; Van Voorhis, T. Constrained Density Functional Theory. *Chem. Rev.* **2012**, *112*, 321–370.

[19] Skinner, K. C.; Kammeraad, J. A.; Wymore, T.; Narayan, A. R. H.; Zimmerman, P. M. Simulating Electron Transfer Reactions in Solution: Radical-Polar Crossover. In review.

[20] De La Lande, A.; Gillet, N.; Chen, S.; Salahub, D. R. Progress and Challenges in Simulating and Understanding Electron Transfer in Proteins. *Arch. Biochem. Biophys.* **2015**, *582*, 28–41.

[21] Řezáč, J.; Lévy, B.; Demachy, I.; De La Lande, A. Robust and Efficient Constrained DFT Molecular Dynamics Approach for Biochemical Modeling. *J. Chem. Theory Comput.* **2012**, *8*, 418–427.

[22] Gillet, N.; Lévy, B.; Molinear, V.; Demachy, I.; De La Lande, A. Electron and Hydrogen Atom Transfers in the Hydride Carrier Protein EmoB. *J. Chem. Theory Comput.* **2014**, *10*, 5036–5046.

[23] Van Der Kamp, M. W.; Mullholland, A. J.; Combined Quantum Mechanics/Molecular Mechanics (QM/MM) Methods in Computational Enzymology. *Biochemistry* **2013**, *52*, 2708–2728.

[24] Doyon, T. J. Development and Characterization of Non-heme Iron Biocatalysts For Complex Molecule Synthesis. Ph.D. Dissertation, University of Michigan, Ann Arbor, MI, 2020.

[25] Mitchell, A. J.; Zhu, Q.; Maggiolo, A. O.; Ananth, N. R.; Hillwig, M. L.; Liu, X.; Boal, A. K. Structural Basis for Halogenation by Iron- and 2-Oxo-glutarate-dependent Enzyme WelO5. *Nat. Chem. Biol.* **2016**, *12*, 636–640.

[26] Humphrey, W.; Dalke, A; Schulten, K., VMD - Visual Molecular Dynamics. *J. Molec. Graphics* **1996**, *14*, 33–38.

[27] Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

[28] Lee, C.; Yang, W.; Parr, R. G. Development of The Colle-Salvetti Correlation-Energy Formula Into a Functional of the Electron Density. *Phys. Rev. B* **1998**, *37*, 785–789.

[29] Ditchfield, R.; Hehre, W. J.; Pople, J. A. Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1971**, *54*, 724–728.

[30] Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; Defrees, D. J.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XXIII. A Polarization-Type Basis Set for Second-Row Elements. *J. Chem. Phys.* **1982**, *77*, 3654–3665.

[31] Gordon, M. S.; Binkley, J. S.; Pople, John A.; Pietro, William J.; Hehre, W. J. Self-Consistent Molecular-Orbital Methods. 22. Small Split-Valence Basis Sets for Second-Row Elements. *J. Am. Chem. Soc.* **1982**, *104*, 2797–2803.

[32] Hariharan, P. C.; Pople, J. A. The Influence of Polarization Functions on Molecular Orbital Hydrogenation Energies. *Theor. Chim. Acta* **1973**, *28*, 213–222.

[33] Hehre, W. J.; Ditchfield, R.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. J. Chem. Phys. **1972**, *56*, 2257–2261.

[34] Rassolov, Vitaly A.; Ratner, Mark A.; Pople, John A., Redfern, Paul C., Curtiss, Larry A. 6-31G* Basis Set for Third-Row Atoms. J. Comput. Chem. **2001**, *22*, 976–984.

[35] Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32*, 1426–1465.

[36] Field, M. J. The Pdynamo Program for Molecular Simulations Using Hybrid Quantum Chemical and Molecular Potentials. *J. Chem. Theory Comput.* **2008**, *4*, 1151–1161.

[37] Neese, F., The ORCA Program System. *Wires Comput. Molec. Sci.* **2012**, *2*, 73–78.

[38] Epifanovsky, E.; et. al. Software for the Frontiers of Quantum Chemistry: An Overview of Developments in the Q-Chem 5 Package. *J. Chem. Phys.* **2021**, *155*, 084801.

[39] Becke, A. D. Density-Functional Exchange-Energy Approximation With Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100.

[40] Perdew, J. P.; Density-Functional Approximation for the Correlation Energy of the Inhomogeneous Electron Gas. *Phys. Rev. B* **1986**, *33*, 8822–8824.

[41] Dunning, T. H. Jr.; Hay, H. J. Gaussian Basis Sets for Molecular Calculations. in Methods of Electronic Structure Theory; Vol. 3; Schaefer, H. F., III, Ed.; Plenum: New York, **1977**, 1–28.

[42] Hay, P. J.; Wadt, W. R. *Ab Initio* Effective Core Potentials for Molecular Calculations. Potentials for K to Au Including the Outermost Core Orbitals. *J. Chem. Phys.* **1985**, *82*, 299–310.

[43] Hirao, H. Which DFT Functional Performs Well in the Calculation of Methylcobalamin? Comparison of the B3LYP and BP86 Functionals and Evaluation of the Impact of Emperical Dispersion Correction. *J. Phys. Chem. A* **2011**, *115*, 9308–9313.

[44] Yu, W.; He, X.; Vanommeslaeghe, K.; Mackerell Jr, A. D. Extension of the CHARMM General Force Field to Sulfonyl-Containing Compounds and Its Utility in Biomolecular Simulations. *J. Comput. Chem.* **2012**, *33*, 2451–2468.

[45] Zimmerman, P. M. Single-Ended Transition State Finding With the Growing String Method. *J. Comput. Chem.* **2015**, *36*, 601-611.

[46] Zimmerman, P. M. Reliable Transition State Searches Integrated With the Growing String Method. *J. Chem. Theory Comput.* **2013**, *9*, 3043–3050.

[47] Zimmerman, P. M. Growing String Method With Interpolation and Optimization in Internal Coordinates: Methods And Examples. *J. Chem. Phys.* **138**, 184102.

[48] Aldaz, C.; Kammeraad, J.; Zimmerman, P. M. Discovery of Conical Intersection Mediated Photochemistry With Growing String Methods. *Phys. Chem. Chem. Phys.* **2018**, *20*, 27934–27405.

Chapter V

Conclusions and Future Outlooks

5.1 Conclusions

Reactions involving single-electron transfer(s) (SETs) are ubiquitous in (bio)chemical reactions. Studying SET processes and their radical intermediates can be challenging due to the short lifetimes of the radical intermediates. However, computational modeling provides a way to the study of SET process and the subsequent transformations. In particular, the use of Marcus theory to describe the thermodynamics and kinetics of electron transfer has seen great success.[1,2], the latter of which can be calculated by using free energies and reorganization energies. Reorganization energies are broken down into the inner-sphere and outer-sphere components. Inner-sphere reorganization energy refers to the energy associated with the change in electronics and geometry for the donor and acceptor complex, and the outer-sphere reorganization energy reflects the change in the solvent environment to accommodate the new electronic state. Quantum mechanical (QM) models are required to accurately capture the inner-sphere reorganization energy but are too computationally expensive to model the solvent environment. Molecular mechanics (MM) can be used to efficiently model the large solvent environment and are used in molecular dynamics simulations to collect the distribution of solvent configurations needed to calculate free energies. Isn a condensed phase system, such as organic reactions and those taking place within an enzyme, the

outer-sphere reorganization energy is the dominant factor within Marcus Theory.[1] Therefore, the donor and acceptor complex needs to be studied with explicit solvent included in the model. To this end, combined QM/MM methods have been constructed to accurately calculate the electronic and geometric changes of the donor and acceptor fragments, while effectively taking advantage of the efficient modeling of the solvent environment. The content contained within this thesis seeks to add to this area of scientific research of reactions involving SETs in two facets: development of a new computational tool to study SET process and subsequent reactions and insights gained from its applications. The new tool was used to study reactions involving SET for the first time for an organic system in explicit solvent and an enzyme active site.

The two reactions studied in Chapter II utilize TDAE and TTF to generate indoles and benzofuranyl alcohols, respectively, after intramolecular cyclization of their respective aryl radicals. The new combined constrained density functional theory/molecular mechanics (CDFT/MM) method was key to studying these reactions, due to initiation by single-electron transfer (SET). The models and method were successful in achieving SET was between the respective OEDs and aryldiazonium salts. Successful reproduction of known kinetic barriers were achieved. The relationship between C–C and C–S bond distances and activation energy was investigated for oxocarbenium formation, in which there is a combination of electronic stabilization and steric repulsion lengthened the bonds with increasing substitution and resulted in a decrease of barriers for oxocarbenium formation. Finally, the impact of the solvent and its rearrangement on reactivity was studied, where relative free energies for the TDAE system agrees with known chemical reactivity and the impact of solvent rearrangement

in the TTF was observed following SET. The new CDFT/MM method is envisioned to be useful in additional complex environments, such as enzyme active sites.

The new CDFT/MM tool was also applied to the enzyme TropC to study our proposed SET pathway with the enzyme and solvent environments modeled explicitly. Using DFT/MM, the rebound hydroxylation and radical-based rearrangement pathways were both found to be kinetically accessible with barriers of 14.9 kcal/mol and 15.3 kcal/mol respectively. The SET pathway was modeled using the new CDFT/MM method to simulate the transfer of an electron from the substrate to the iron center of the active site. Following SET from the substrate to the iron center, there was a barrierless ring expansion. However, the initial vertical energy gap was calculated to be excessively high at 138.0 kcal/mol. This excessive vertical energy gap suggested that the initial intermediate was not capable of performing SET, but that a structure in-between the two endpoints could be conducive to SET and that a minimum energy crossing point (MECP) could exist between the radical and SET pathways. Following reaction pathway simulations and MECP optimizations, an MECP was not identified but instead a near-crossing point with an energy difference of 3.4 kcal/mol between the ground and SET state. Location of a near-crossing point indicates that SET is potentially viable within TropC. Based on the results of reaction pathway simulations, rebound hydroxylation and radical-based rearrangement suggest that all the first two pathways are feasible and competitive within TropC, and that the discriminating factor between them is substrate positioning within the active site. The closer the substrate is to the iron the more likely rebound hydroxylation, similar to the rationale used for hydroxylases.[3] The further away from the iron the more likely radical-based rearrangement is to occur. There also likely

exists a distance between rebound hydroxylation and radical-based rearrangement in which SET is most favored.

5.2 Future Outlooks

The CDFT/MM tool developed within Chapter II is envisioned to be applicable to a broad audience as SETs are ubiquitous within (bio)chemical reactions, and the insights gained in Chapters III and IV can be used to guide rational enzyme design as well as study other enzymes with interesting electronic systems, such as those containing iron-sulfur clusters[4] and cofactors such as pyrroloquinoline quinone.[5]

The new CDFT/MM tool is useful to those studying SET reactions within condensed phases, where reactions beyond the SET step are rate limiting. This could include additional organic systems like the ones studied in Chapter II, or additionally complex systems such as those containing metals. Reactions involving proposed SETs within α-KG NHI enzymes, and other enzyme classes, could also apply this tool to understand the feasibility of such transformations.

Understanding the relationship between distance and chemical reactivity within TropC can be used to rationally redesign TropC and the related enzymes XenC[6] and EupC[7], in which it was previously found by Dr. Tyler Doyon that XenC has better performance than TropC.[8] Redesign of XenC could increase the substrate scope beyond its current limitations and could include scaffolds for bioactive natural products, such as manicol, sepedonin, malettinin E, and purpurogallin (Figure 5.1).[9,10] Additional analyses of residues near the substrate could lead to target for mutation. Variant enzyme models can be generated to test hypothetical mutations and understand their impact on the

various reaction pathways and identification of beneficial mutations that would decrease shunt product formation and increase tropolone production.

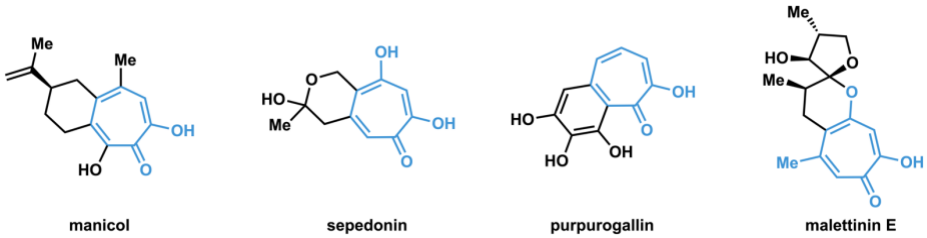

Figure 5.1**.** Potential target tropolones cores for XenC.

5.3 References

[1] Kaduk, B.; Kowalczyk, T.; Van Voorhis, T. Constrained Density Functional Theory. *Chem. Rev.* **2012**, *112*, 321–370.

[2] Blumberger, J. Recent Advances in the Theory and Molecular Simulation of Biological Electron Transfer Reactions. *Chem. Rev.* **2015**, *115*, 11191–11238.

[3] Mitchell, A. J.; Zhu, Q.; Maggiolo, A. O.; Ananth, N. R.; Hillwig, M. L.; Liu, X.; Boal, A. K. Structural basis for halogenation by iron- and 2-Oxo-glutarate-dependent enzyme WelO5. *Nat. Chem. Biol.* **2016**, *12*, 636–640.

[4] Liu, J.; Chakraborty, S.; Hosseinzadeh, P.; Yu, Y.; Tian, S.; Petrik, I.; Bhagi, A.; Lu, Y. Metalloproteins Containing Cytochrome, Iron–Sulfur, or Copper Redox centers. *Chem. Rev.* **2014**, *114*, 4366–4469.

[5] Klinman, J. P.; Bonnot, F. Intrigues and Intricacies of the Biosynthetic Pathways for the Enzymatic Quinocofactors: PQQ, TTQ, CTQ, TPQ, and LTQ. *Chem. Rev.* **2014**, *114*, 4343–4365.

[6] Schor, R.; Schott, C.; Wibberg, D.; Kalinowski, J.; Cox, R. J. Three previously unrecognized classes of biosynthetic enzymes revealed during the production of xenovulene A. *Nat. Commun.* **2018**, *9*, 1963.

[7] Chen, Q.; Gao, J.; Jamieson, C.; Liu, J.; Ohashi, M.; Bai, J.; Yan, D.; Liu, B.; Che, Y.; Wang, Y.; Houk. K. N.; Hu, Y. Enzymatic Intermolecular Hetero-Diels-Alder Reaction in the Biosynthesis of Tropolonic Sesquiterpenes. *J. Am. Chem. Soc.* **2019**, *141*, 14052–14056.

[8] Doyon, T. J. Development and Characterization of Non-heme Iron Biocatalysts for Complex Molecule Synthesis. Ph.D. Dissertation, University of Michigan, Ann Arbor, MI, 2020.

[9] Silber, J.; Ohlendorf, B.; Labes, A.; Wenzel-Storhohann, A.; Näther, C.; Imhoff, J. F. Malettinin E, an Antibacterial and Antifungal Tropolone Produced by a Marine *Cladosporium* Strain. *Font. Mar. Sci.* **2014**, 1–6.

[10] Bentley, R. A Fresh Look at Natural Tropolonoids. *Nat. Prod. Rep.* **2008**, *25*, 118–138.