Second-Coordination Sphere Effects on the Chemo-Selectivity of Nonheme Iron Dioxygenases. A Computational Study.

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Science and Engineering

2022

Yen Ting Lin

Supervisor: Dr. Samuël de Visser

Department of Chemical Engineering

Table of Contents

List of Abbreviation

Abbreviation	Full name
α-KG	α-ketoglutarate
AAD	aryloxyalkanoate dioxygenase
AlkB	AlkB repair enzyme
Asp	Aspartic Acid
BDFE	Bond Dissociation Free Energy
BS	Basis Set
B3LYP	Becke, 3-parameter, Lee-Yang-Par
CPCM	Continuum polarized Conductor model
Co	Cobalt
DFT	Density Functional Theory
DNA	Deoxyribonucleic acid
ECP	Electron Core Potential
EPR	Electron Paramagnetic Resonance
Eq.	Equation
Fe	Iron
Glu	Glutamic acid
GTO	Gaussian-Type Orbitals
HAT	Hydrogen atom transfer
HF	Hartree Fock
His	Histidine

I	Intermediate
KIE	Kinetic Isotope Effect
LS	Low Spin
LSD	Local-Spin Density Correlation
MD	Molecular Dynamics
Mn	Manganese
Me	Methyl
MM	Molecular Mechanics
OH	Hydroxyl
Ni	Nickel
P	Product
PCM	Polarizable Continuum Model
PCET	Proton-Coupled Electron Transfer
PDB	Protein Data Bank
PES	Potential Energy Surface
PT	Proton Transfer
QM	Quantum Mechanics
QM/MM	Quantum Mechanics/ Molecular Mechanics
Re	Reactant
RNA	Ribonucleic acid
SE	Semi-Empirical
STO	Slater-Type Orbitals
Sub	Substrate
TS	Transition state
TauD	Taurine/α-ketoglutarate-dependent dioxygenase
TmpA	2-(trimethylammonio)ethylphosphonate dioxygenase
UndA	Undecene A biosynthesis enzyme
VB	Valence Bond
ZPC	Zero Point Correction
ZPE	Zero-Point Energy

Abstract

Non-heme iron dioxygenases are versatile biological enzymes in all life forms, including the human body. They catalyse various chemical reactions, including the biosynthesis of natural products, e.g., hydroxyproline in mammals. However, they are also involved in the biodegradation of toxic compounds, for example, toxic cysteine in the brain. In many of these reaction pathways, a regio- or stereoselective reaction occurs, making this enzyme class of interest to the biotechnology industry for producing high-value compounds. Many questions remain about the high selectivity patterns of non-heme iron enzymes and how the enzyme drives these mechanisms. Therefore, a detailed computational study was performed on a set of well-characterised non-heme iron dioxygenases with different functions in Nature. This dissertation focuses on three non-heme iron dioxygenases, of which a recent crystal structure was determined. These structures have been used to elucidate the selectivity patterns' reaction mechanisms and origins. In particular, the work described in this thesis shows that the second-coordination sphere of the enzyme plays a vital part in substrate binding and positioning.

Moreover, electrostatic and polar interactions from the second coordination sphere often induce a local dipole moment or electric field effect and drive the reaction to an otherwise thermodynamically unfavourable pathway. My work has given insight into the intricate details of enzyme design and how regio- and chemoselectivities can be achieved. This insight can be used to engineer non-heme iron dioxygenases and give them original reactivity patterns with high selectivity for biotechnological applications.

My work details computational studies on three non-heme iron dioxygenases, namely UndA (undecene biosynthesis enzyme), TmpA (2-(trimethylammonio)-ethylphosphonate dioxygenase) and AAD (aryloxyalkanoate dioxygenase). The latter enzyme is found in agricultural plants such as rice and is shown to be able to trigger the biodegradation of herbicides. Our work has given insight into the mechanism of the herbicide biodegradation and the substrate scope of the enzyme. For the UndA biosynthesis enzyme, a controversy had arisen about whether it is a mononuclear non-heme iron dioxygenase; hence models for mononuclear and dinuclear systems were created that established it as a dinuclear enzyme. Finally, the TmpA work showed that this enzyme reacts via negative catalysis, whereby a thermodynamically favourable pathway is blocked, favouring a higher energy reaction channel. My work highlights the second coordination sphere groups that induce a local electric field and guide the reaction to the wanted products.

Declaration

That no portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

List of Publication

[1] Pangia, T. M., Yadav, V., Gérard, E. F., **Lin, Y. T.,** de Visser, S. P., Jameson, G. N., & Goldberg, D. P. **(2019). Mechanistic investigation of oxygen rebound in a mononuclear nonheme iron complex.** Inorganic chemistry, 58(15), 9557-9561.

[2] Cantú Reinhard, F. G., **Lin, Y. T.**, Stańczak, A., & de Visser, S. P. **(2020). Bioengineering of cytochrome P450 OleTJE: How does substrate positioning affect the product distributions?.** Molecules, 25(11), 2675.

[3] **Lin, Y. T.**, Stańczak, A., Manchev, Y., Straganz, G. D., & de Visser, S. P. **(2020). Can a mononuclear iron (III)-superoxo active site catalyze the decarboxylation of dodecanoic acid in UndA to produce biofuels?.** Chemistry–A European Journal, 26(10), 2233-2242.

[4] Cummins, D. C., Alvarado, J. G., Zaragoza, J. P. T., Effendy Mubarak, M. Q., **Lin, Y. T.,** de Visser, S. P., & Goldberg, D. P**. (2020). Hydroxyl Transfer to Carbon Radicals by Mn (OH) vs Fe (OH) Corrole Complexes.** Inorganic chemistry, 59(21), 16053-16064.

[5] **Lin, Y. T.,** & de Visser, S. P. **(2021). Product Distributions of Cytochrome P450 OleTJE with Phenyl-Substituted Fatty Acids: A Computational Study.** International journal of molecular sciences, 22(13), 7172.

[6] **Lin, Y. T.**, Ali, H. S., & de Visser, S. P. **(2021). Electrostatic Perturbations from the Protein Affect C− H Bond Strengths of the Substrate and Enable Negative Catalysis in the TmpA Biosynthesis Enzyme.** Chemistry–A European Journal, 27(34), 8851-8864.

[7] de Visser, S. P., **Lin, Y. T.**, Ali, H. S., Bagha, U. K., Mukherjee, G., & Sastri, C. V. **(2021). Negative catalysis/non-Bell-Evans-Polanyi reactivity by metalloenzymes: Examples from mononuclear heme and non-heme iron oxygenases**. Coordination Chemistry Reviews, 439, 213914.

[8] **Lin, Y. T.**, Ali, H. S., & de Visser, S. P. **(2022). Biodegradation of herbicides by a plant nonheme iron dioxygenase: mechanism and selectivity of substrate analogues**. Chemistry–A European Journal, 28(7), e202103982.

Acknowledgements

First and foremost, I would like to great thanks to my brilliant supervisor Dr Samuel De Visser, Department of Chemical Engineering, who has been tremendously helpful in guiding me through these four years of challenging PhD studies period in a matter of computational works, explaining theoretical background of computational enzymatic chemistry, and guidance through various publications with patience and immeasurable support on both academic and industrial career advice, and encourage me to attend several international conferences to broaden my horizon in academia.

Many thanks to everyone in the Manchester Institute for Biotechnology research group for creating a fun and supportive environment in which to work. I am thankful to my colleagues (Our research group): Dr Fabián G. Cantú Reinhard (senior PhD student who taught me while my first year of PhD), Dr Hafiz Saqib Ali (who cooperated with me and taught me a lot of computational techniques), Dr Muhammad Qadri Effendy Mubarak (senior PhD student), Emilie Gerard (PhD candidate), and Thirakorn Mokkawes (PhD candidate). Moreover, Henrik Wong (PhD candidate) and Dr Mathew Newman help me with further proofreading. All of them, one way or the other, helped me conquer this PhD journey, overcoming various academic difficulties and challenges. All of them, one way or the other, helped me conquer this PhD journey, overcoming various academic difficulties and challenges.

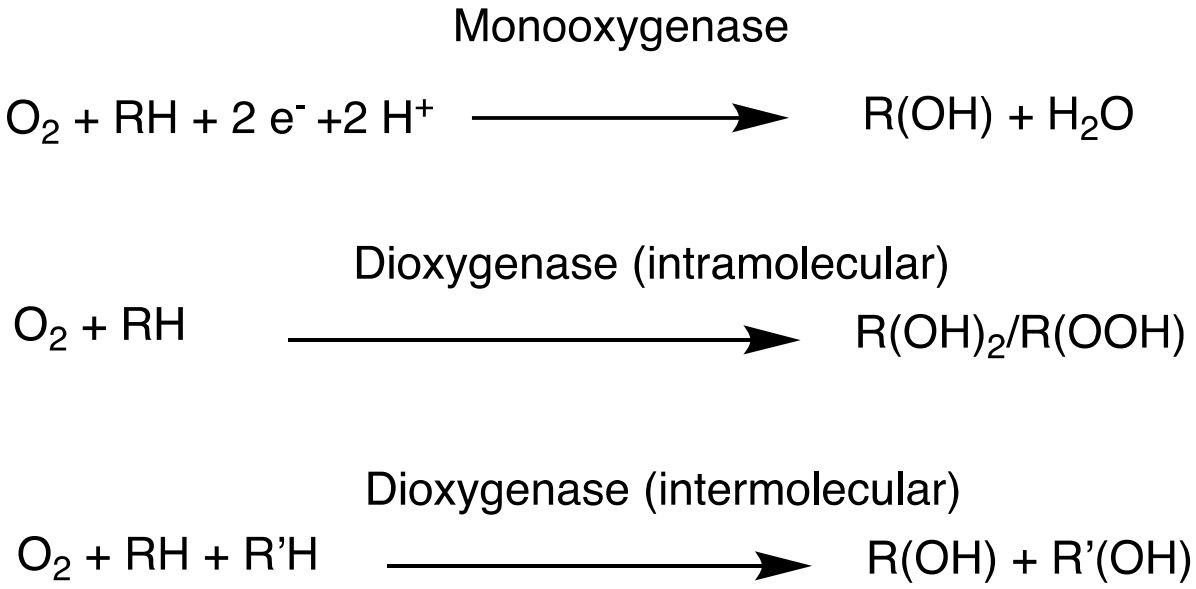
I want to thank my family (my parents, my father, Dr Lin, Sheng-Cai, and my mother Zhang, Qiu-Min) for financial support for my research. Most of all, I would like to thank my best friends: Dr Tobias Bohn, Dr Negin Kamyar, Dr Mathew Newman, Dr Amir Hassani, Dr Zunmin Guo, Dr Pan, Jian-Xun, Dr Swathi Suran, and Dr Yi Mu, company with me been through this challenging journey They brought me joyful times and gave me vital support and company during my PhD career, especially during this challenging pandemic.

Thanks to Professor Peter Gardner and Professor Etienne Derat as my viva internal and external examiners. They peruse my thesis and provide insightful and professional advice for the corrections and improvement of thesis quality. In addition, thanks to Manchester University provides such a great environment and a computational facility to study and conduct simulation works. This service given by IT services and the use of the Computational Shared Facility (CFS3) at the University of Manchester is also acknowledged.

Chapter 1: Introduction

1.1 Dioxygen activation, Non-heme iron dioxygenases

In nature, dioxygen is essential to all the aerobic living creatures that utilize dioxygen for biosynthesis and biodegradation processes. In the past 60 years, major breakthroughs in the understanding of dioxygen activation by metalloenzymes has taken place. [1] The earliest known dioxygen-activating enzymes that were discovered were tyrosinase, cytochrome P450 and catechol 1,2-dioxygenase. [2-6] Tyrosinase is a di-copper enzyme that hydroxylates a tyrosine amino acid as part of the melatonin biosynthesis in plants and animal tissues, while catechol 1,2-dioxygenase is an intradiol dioxygenase involved in the biodegradation of aromatic compounds. [4,5] The cytochromes P450 are heme enzymes that take part in biodegradation reactions in the liver and act as a mono-oxygenase. [7] These enzymes have been well studied in the field of dioxygen activation. [1] Dioxygenases use a dioxygen molecule and transfer both oxygen atoms to the target substrate whereas the mono-oxygenase only transfers one oxygen atom to substrate and releases the second one as a water molecule. [8-11] **Scheme 1.1** shows functions of monooxygenase and dioxygenases (intramolecular and intermolecular). [12]



Scheme 1.1 Monooxygenase and dioxygenase reactions by metallenzymes. [12]

Natural occurring metalloenzymes carry out catalytic oxidations often with high substrate specificity, regioselectivity, and stereoselectivity. [25,26] Most importantly they operate under mild conditions at room temperature while in synthetic chemistry and heterogeneous catalysis frequently high temperatures and pressures are used to achieve the same reaction yields. This leads to a relatively "green or sustainable process", allowing a lower energy cost and less toxic or unwanted products. [26] There are many primary

types of metalloenzymes for dioxygen activation ranging from copper enzymes, such as tyrosinase, [19,20,21] iron enzymes (either heme, nonheme or diiron), [22.23] and enzymes with other metal ions, e.g. Mn, Co, or Ni as core. [24] This thesis focuses on the iron-containing dioxygen-activating enzymes, with special emphasis on the mononuclear nonheme iron enzymes. The two major classes of mononuclear iron containing enzymes are the heme and nonheme enzymes, which are found in virtually every species whether in plants or animals or even bacteria, displayed in **Figure 1.1**. [13] Regarding the superfamily of the nonheme iron enzymes, there are two kinds, namely the mononuclear non-heme iron enzymes and dinuclear nonheme iron enzymes. In my research I only focused on the mononuclear nonheme iron enzymes.

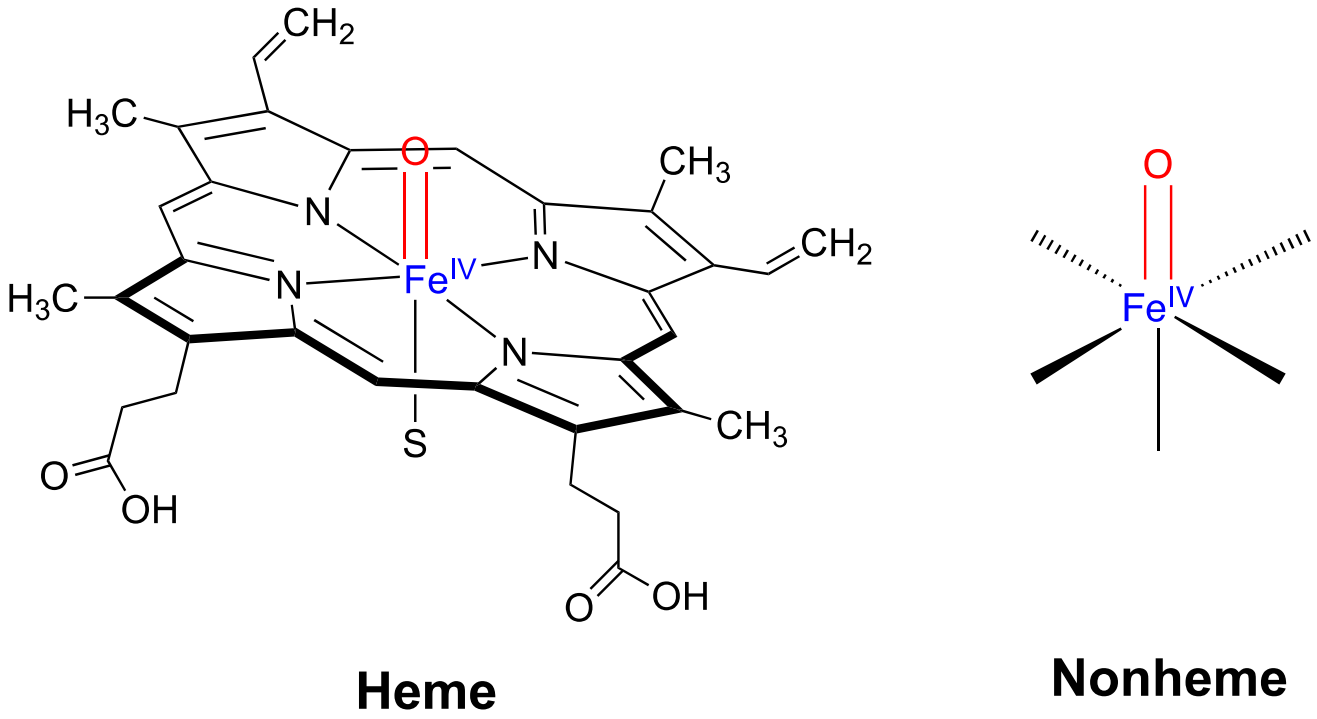


Figure 1.1 General structural motif comparison of the first-coordination sphere of heme and nonheme iron-containing enzymes. [13]

The structure of the active site of the enzyme plays a key role to its function and mechanism, especially the coordination environment of the redox-active iron centre that has a variety of forms to accommodate diverse catalytic reactions. Thus, nonheme iron enzymes are capable of important biotransformations, such as the biodegradation of toxic chemicals, and the biosynthesis of natural products. [14] Due to their efficiency and often selective reaction patterns, there is huge interest from the biotechnology industry to utilize metalloenzymes for a variety of chemical processes. For instance, because of sustainable concerns, the soaring development studies on the biosynthesis of biofuels (hydrocarbons) using iron containing enzymes (such as nonheme iron enzymes, cytochrome P450 mono-oxygenases, and dinuclear iron

enzymes) have emerged, which aims to serve as an alternative to petroleum-based fuel. [14] In Chapter 3, I will discuss the details of enzymatic routes for the 1-undecene biosynthesis by the enzyme UndA, which has shown promising potential for the biosynthesis of biofuel products. [15] This project followed an experimental study from the Rui research group that had characterised short-lived intermediates of the catalytic cycle of UndA using crystallographic and biochemical approaches. [16] This is a nonheme iron decarboxylase found to biotransform medium-chain fatty acid with chain lengths of C_n (n = 9 - 13) to terminal olefins. [15] The primary function of these enzymes is their capability to perform the cleavage of the C-C bond of the substrate fatty acids (or aldehydes, C_{n-1}) through decarboxylation into alkenes (C_n). By contrast, the "C-H functionalization" reactions typically are coupled with dioxygen activation and are related with mono-oxygenation reactions. [14]

From a chemical synthesis point of view, the reaction of dioxygen with substrate is considered a spin-forbidden oxidation process due to a spin mismatch between reactants and products. [9] This is because of the triplet ground state of dioxygen and hence requires a spin-crossing during the reaction mechanism. Therefore, it has a low one-electron oxidation potential and very slow kinetics. However, nature has developed an intelligent approach for oxidation reactions by using transition metals to overcome the dioxygen activation's kinetic reaction energy barrier. [27]

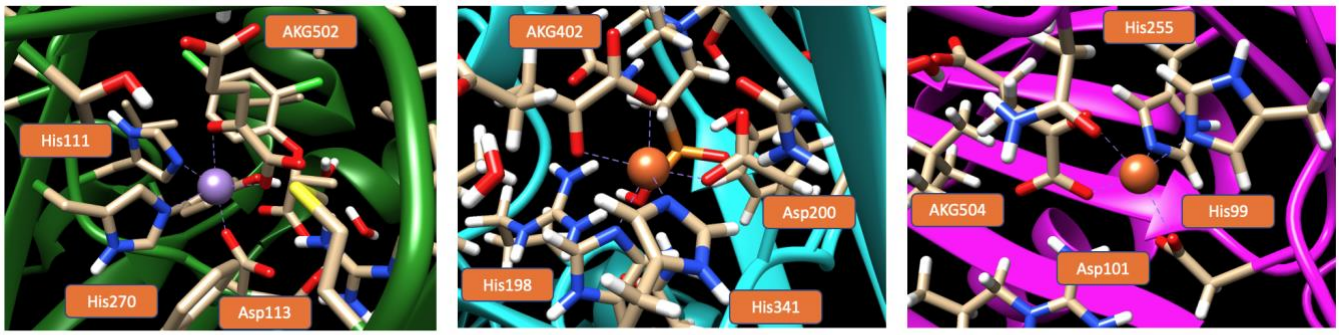


Figure 1.2 Active site model of AAD, TmpA, TauD enzyme. Three of the models shows similar structural motif 2-His-1-carboxylate. [68,125,131]

Over the last two decades, a recurring structural motif has been identified in nonheme iron enzymes with a so-called "2-His-1-carboxylate facial triad", see comparison in **Figure 1.2**. [22,23] For instance, the iron centre in taurine/α-ketoglutarate dioxygenase (TauD) has the iron ion bound to two histidine side chains and a carboxylate from an Asp side chain in a facial orientation. Since then, many nonheme iron

dioxygenases were shown to have a similar active site structure. The geometry and electronic properties of transition metals are dependent on the surrounding ligand motifs and determine physical chemical properties such as reduction potential of the iron centre. Metalloenzymes, especially iron-containing enzymes in nature, take on diverse oxidative biotransformation reactions and will be the main topic of this thesis. The functionalisation of C-H bonds is specifically challenging as these bonds are thermodynamically relatively strong and difficult to activate. [28] One example that synthetic chemists struggle with is the transformation of methane to methanol. [29,30,31]

Metalloenzymes like methane mono-oxygenase can hydroxylate methane, while other oxygenases like the cytochromes P450 and nonheme iron dioxygenases activate longer hydrocarbon substrates easily but not methane. In order to achieve oxygen activation, in mononuclear metalloenzymes like the cytochromes P450 or nonheme iron dioxygenases, a catalytic cycle describes the steps from substrate and oxidant binding to the generation of an active oxidant to the actual substrate activation and product release. Often these metalloenzymes create a high-valent iron(IV)-oxo species that performs the C-H functionalisation as these intermediates can do this with much lower activation energy than free dioxygen. [25] The reaction will need to overcome the C-H bond dissociation free energy (BDFEs), which is considerable for aliphatic C-H bonds (cleaving energy ranging from 80 – 105 kcal mol^{-1}), in which the enzyme usually requires about 10-30 kcal mol^{-1} in activation energy to catalyse the reaction, and normal non-enzyme chemistry reaction need to overcome that C-H bonding dissociation energy to achieve the same reaction goal. [28,32] Therefore, it is essential to understand how the metalloenzyme conducts functional group biotransformation in the catalytic cycle and its active site structural motif.

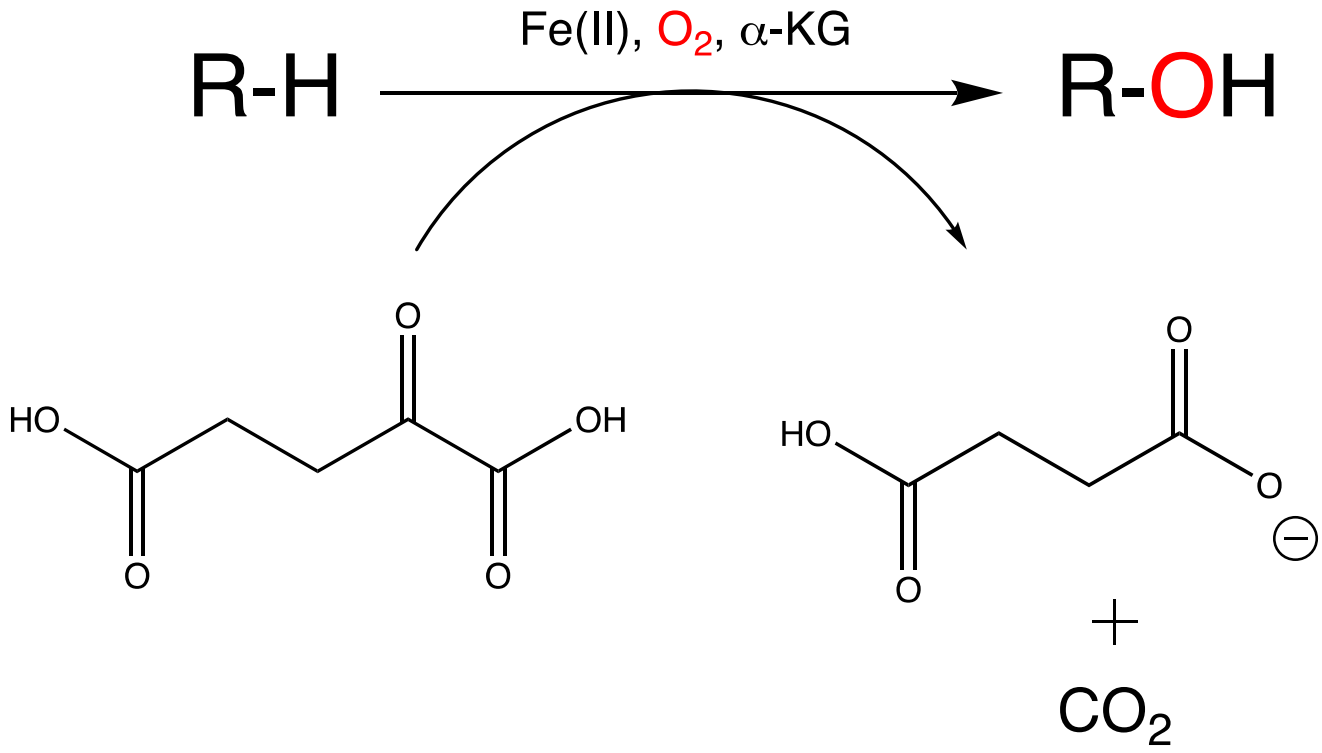
Metalloenzymes use these high valent iron(IV)-oxo species to activate the substrate using "a proton-coupled electron transfer" (PCET) process, where the electron and proton transfer to different atoms. [28,33,34,35,36] In particular, the proton transfers from the substrate to the oxo group, while the electron moves to the iron (in nonheme) or heme (in the cytochromes P450). The natural enrichment of iron in nature means that iron-containing enzymes are very common and widespread. In the mononuclear nonheme iron enzymes, they acquire electrons from the co-substrate or cofactor to carry out the catalytic mechanism. [37,38,39] Nonheme iron dioxygenases are more self-sufficient than the heme enzymes as they do not need an external redox-partner to donate the electrons for the catalytic cycle. [40,41] The nonheme iron dioxygenases instead use its co-substrate to supply the electrons for the reaction mechanism and assist with the generation of an active high-valent iron(IV)-oxo species. Subsequently, the nonheme iron enzymes commonly conduct substrate activation on the primary quintet spin state through single-state reactivity (SSR) and without spin crossing to another spin state. [8] In both heme and nonheme

catalytic mechanisms, high-valent iron oxo intermediates are commonly generated after dioxygen molecules bind to the iron core. [11,42-47]

Nonheme iron dioxygenases have diverse oxidative functions related to biotransformations, and examples come from the mammalian physiology of collagen synthesis by prolyl-4-hydroxylase, the repair of alkylated DNA or RNA bases by AlkB, and the biodegradation of toxic chemicals in the body. [9-12,15,17,18,48-60] Furthermore, in bacteria they have been shown to be involved in the biosynthesis of antibiotics (i.e. isopenicillin, fosfomycin). In this thesis, we will focus on the mononuclear nonheme iron dioxygenases, which catalyses diverse reactions and are naturally occurring in every life form. [9,56.61-63]

Experimentally, α-ketoglutarate (α-KG)-dependent nonheme iron enzymes are the most well-known and well-studied enzyme nonheme iron enzymes, even though the family comprises of more than a thousand members. [28,64] They carry out versatile enzymatic reactions. For the TauD enzyme, the mononuclear iron metal centre utilizes α-KG (also called 2-oxoglutarate, 2OG) as a co-substrate and hence are called the α-KG-dependent nonheme iron oxygenases. [65-69] An extensive number of studies have focused on the characterisation of short-lived intermediates and the reactivity patterns with substrates and substrate-analogues. These studies range from X-ray crystallography, biochemical, computational approaches, electronic absorption, circular dichroism, isothermal titration calorimetry, magnetic circular dichroism, conventional and pulse electron paramagnetic, nuclear resonance vibrational spectroscopy, Mössbauer, X-ray absorption, crystallographic analyses, resonance Raman approaches, and ^{18}O-kinetic isotope effect measurements in order to understand active site properties of TauD enzyme. The first work on TauD described the detection of two intermediates that were characterised by rapid kinetic and spectroscopic approaches in *Escherichia coli*: one identified as a high-valent nonheme iron(IV)-oxo species, while the other is iron(II) species. [70] In 1982, the first proposed TauD mechanism included various chemically reasonable intermediates. [71]

The general mechanism of taurine/α-KG dioxygenase (TauD) is that this enzyme will hydroxylate a taurine substrate at its second carbon position, then invoke the release of sulfate from the product. [72] By contrast to the situation in TauD, for the majority of nonheme iron dioxygenases the high-valent intermediates like Fe(IV)=O are highly challenging to characterise spectroscopically. [73] **Scheme 1.2** shows the general TauD enzyme hydroxylate the substrate and release the hydroxylate product. [27] However, computational techniques can serve as a promising tool to study their mechanistic, electronic, and structural functionalities.



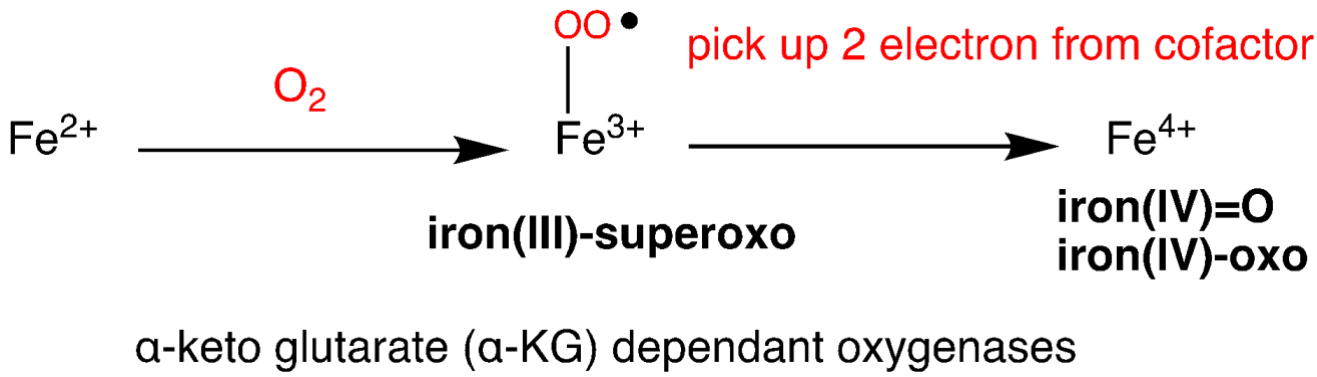
Scheme 1.2 Typical hydroxylation reaction catalysed by many of α-KG dependent oxygenase, non-heme iron enzymes

One advantage of computational work is that the calculations give insight into fast reaction processes that are difficult to trap experimentally. Thus, large models that include the first- and second-coordination sphere of the metal centre and substrate enable one to study the reaction mechanisms with the substrate in a regio-, stereo-, and chemo-selectivity way and check the pathways to products and by-products. [73] It is interesting to note that the co-substrate (cofactor) such as α-KG is used to convert the iron(III)-superoxo intermediate into a high-valent iron(IV)-oxo intermediate and succinate, but these steps were found to proceed too fast and no intermediates along that mechanism could be determined. [74-77]

These α-KG-dependent nonheme iron dioxygenases can selectively conduct epimerization, ring closure, hydroxylation, ring expansion, and desaturation of the target substrate. Hence, α-KG dependent nonheme iron enzymes, can catalyse diverse chemical reactions, that are used in nature for functions including antibiotics biosynthesis, collagen, oxygen sensing, DNA repair, transcription regulation. [55,75,78-81] Nonheme iron(IV)-oxo species was first reported by the Bollinger and Krebs's groups, who conducted a series of experiments using spectroscopy on TauD. [55,57,70,82] The activated high-valent iron-oxo species usually operates through hydrogen atom abstraction of an aliphatic group of the substrate. [83] Although most experimental reports is on TauD, a handful of other nonheme iron dioxygenases (α-KG dependent) have also had their iron(IV)-oxo species characterised through spectroscopic experiments.

[64,84,85] However, for a number of nonheme iron dioxygenases a crystal structure is available of the nonheme iron(II) resting-state with or without substrate and α-KG. To understand the electronic energy, reaction patterns, geometric transformation of nonheme iron dioxygenase, one must know the detailed structural motif to reveal mechanistic insight fully.

In the Scheme 1.3, the mechanisms of dioxygenase activation by an iron core is shown for the transformation upon dioxygen binding to the form an iron(III)-superoxo species and its conversion through breaking of the O-O bond to ultimately lead to a high-valent iron(IV)-oxo species. Overall, this process requires a reaction with a co-substrate, such as α-ketoglutarate (α-KG) that abstracts two-electrons from the metal.



Scheme 1.3: Electron transfer steps during the dioxygen activation process of non-heme iron dioxygenases

1.2 The 2-His-1-carboxylate facial triad structural motif in nonheme iron dioxygenases

Based on crystal structure coordinates of a variety of nonheme iron enzymes, it can be seen that the resting state of nonheme iron dioxygenases has a similar first-coordination structure of the iron centre as demonstrated in **Figure 1.2**. [27]

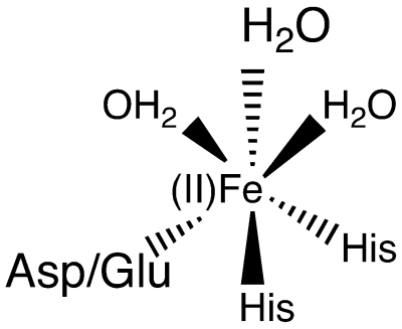


Figure 1.3 Resting state structure of typical non-heme iron dioxygenases

This is, despite the variation in reactivity and chemical catalysis of the enzyme superfamily. In particular, the typical structural motif of nonheme iron(II) enzymes in the first coordination sphere gives the "2-His-1-carboxylate facial triad", whereby the iron(II) core is connected to the side chains of two histidine amino acids and one carboxylate (from either an aspartate or glutamate amino acid) ligand. [6,22,27,47,86,87,88] The remaining three ligand sites of the iron are occupied by three solvent molecules (water molecules).

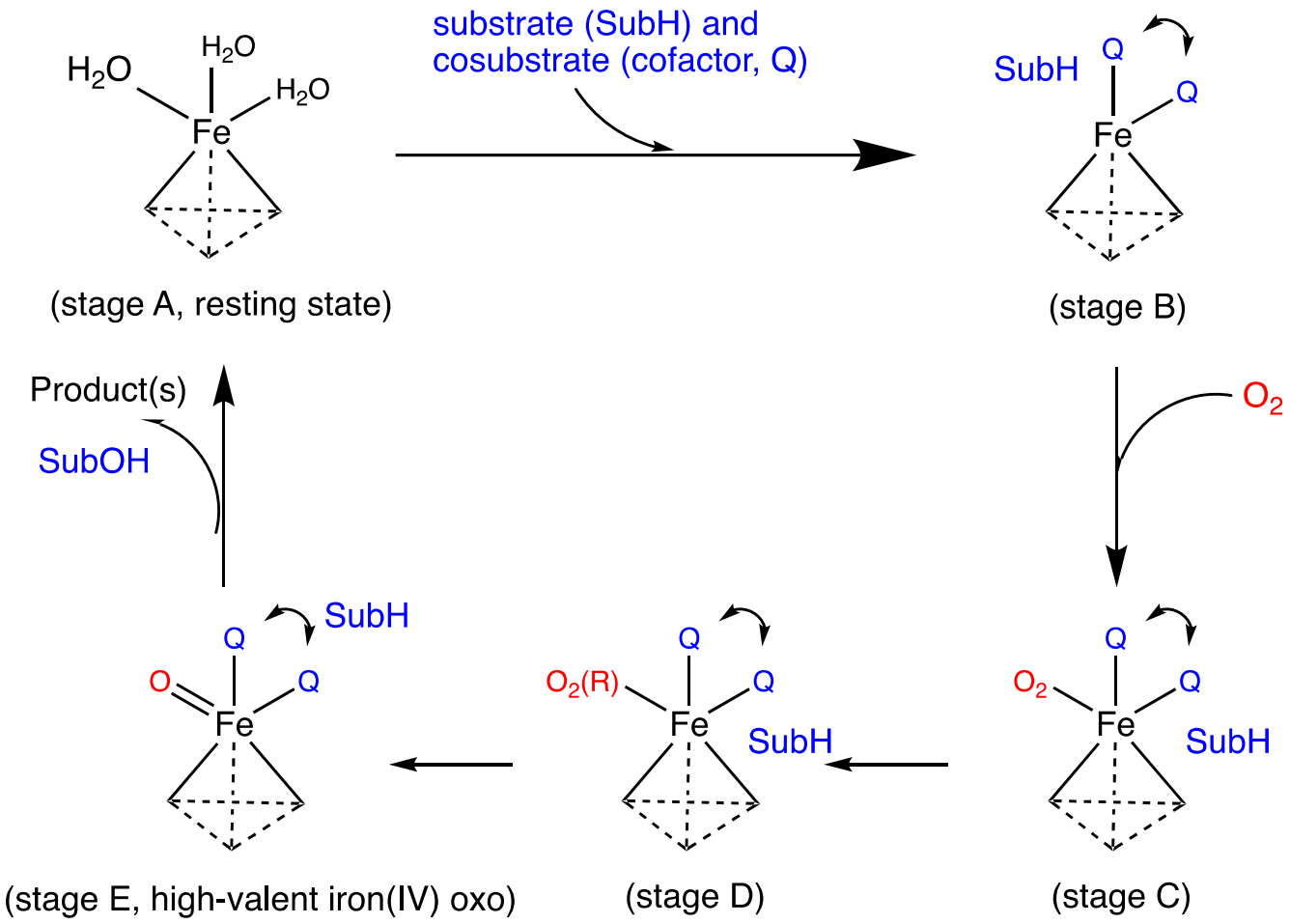
The 2-His-1-carboxylate (Asp/Glu) motif indicates that its facial triad (three ligands have been occupied) take over one side of the octahedral structure motif, while the rest of the endogenous ligands are solvent molecules. [6,27,83] This is demonstrated in **Figure 1.3**, where active site of AAD, TmpA, and TauD enzyme show the similar structural motif. [68,125,131]

During the catalytic cycle an iron(IV)-oxo is formed, where the oxo group is trans to one of the histidine amino acids. The consensus of the structural motif "2-His-1-carboxylate facial triad" is naturally occurring and is the vital foundation of the non-heme iron oxygenases resting state of the structure. [6] Five classes of mononuclear nonheme iron with "2-His-1-facial triad" motif have been classified by structural features, reactivity, distinguishing different functions for bio-catalysis. [27] These mononuclear nonheme iron classes are named the α-KG dependent dioxygenases, the extradiol cleaving catechol dioxygenases, the Rieske dioxygenases and the pterin-dependent hydroxylases. [27] As discussed above,

the nonheme iron and α-KG dependent enzymes use the 2-His-1-carboxylate facial triad as structural motif and react dioxygen with α-KG through decarboxylation to form succinate, carbon dioxide and an iron(IV)-oxo species. [11] Interestingly, the structural motif of the nonheme iron decarboxylase UndA, according to the crystal structure coordinates from the 4WXO PDB file, demonstrates that the mononuclear iron core coordinates with the side chains of His104, His194, and Glu101. [15,16] However, these residues do not align with the consensus structural motif "2-His-1-carboxylate (Asp or Glu)". [6] As the structural arrangement of ligands in UndA appears not to form the facial triad motif we were wondering whether it still would be able to react with substrate efficiently. Our subsequent computational study on UndA highlighted that UndA is unlikely to be a mononuclear iron system as high barriers for substrate activation were found. Instead, a diiron system fit the protein structure better and reacted with much lower energy barriers and made the reaction feasible at room temperature.

It is also interesting to note that analogous enzymes, such as thiol dioxygenases, use three histidine coordinated iron(II) and bind a thiolate-containing substrate to the metal centre. [89] There are two significant functions of the coordination of the water molecules found in the iron(II) resting state of the enzyme: the first one is that water coordination stabilises the reduced state of the iron core and the second reason it allows the cofactors (i.e. α-KG) to replace and activate two ligand sites effortlessly. [28] The iron(II) resting state (divalent iron centre) provides a mono-anion stage (di-valent iron centre), making it suitable for the flexibility to bind with a variety of ligands, such as cofactor, substrate, and the target substrate. [27]

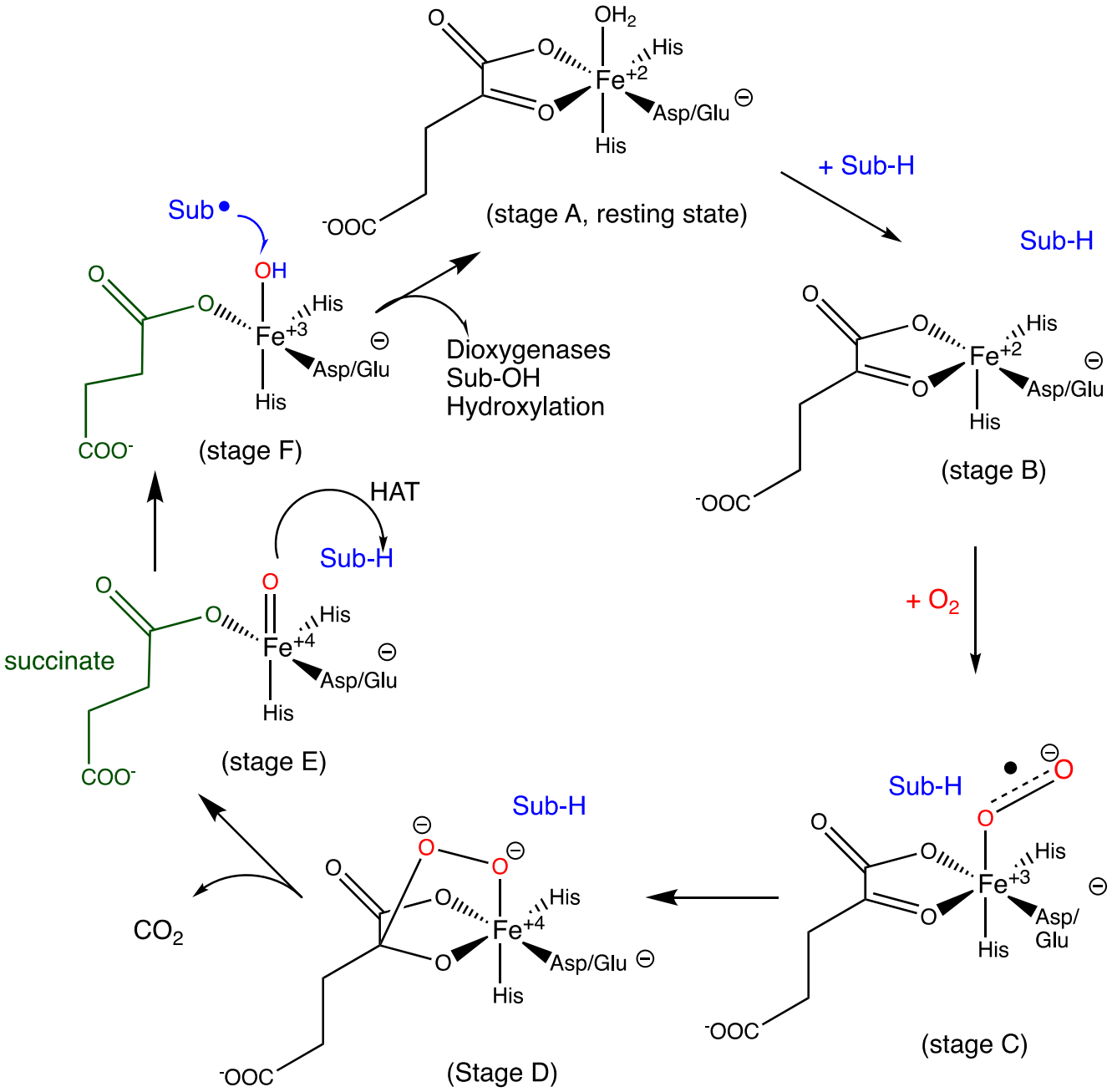
1.3 Catalytic cycle of non-heme iron enzyme



Scheme 1.3 Consensus catalytic cycle of nonheme iron dioxygenase with the 2-His-1-Asp/Glu facial triad structure.

Starting from the resting state in the catalytic cycle, as demonstrated in **Scheme 1.3**, it is an iron(II) species A that has three coordinated water molecules. Two coordinated waters are replaced by the cofactor α-KG in the first stage of the catalytic cycle (to form B), while the main substrate is closely positioned to the iron centre but not covalently bound. [6,27,28] This leaves one vacant site on the iron centre that can be activated by a dioxygen molecule. Dioxygen binding to form structure C is followed by the attack of dioxygen on the co-substrate α-KG and leads to succinate and carbon dioxide and generates the iron(IV)-oxo species (Intermediate E). [64] Interestingly, the diverse classes of the superfamily of nonheme iron dioxygenase all share quite a few mutual mechanistic characteristics. The general mechanism of nonheme iron dioxygenase in more detail is as follows: [6,9,11,90] (Stage A). The resting state of a non-heme iron enzyme is usually an iron(II) coordinating with six ligands, including three displaceable ligands attached to neutral solvent molecules, and the related site is the octahedral coordination sphere.

The iron(II) resting state has been trapped and characterized with a variety of experimental methods including by magnetic circular dichroism (MCD), crystallography and UV-Vis spectroscopy. [9,51] (Stage B) Then once we introduce the substrate or co-substrate (cofactor) (Q) into the active site and replace two ligands, leaving the one vacant site, it turns the six-coordination saturated motif into an unsaturated five-coordination structural motif. [90] In this stage, the five-coordination structural motif becomes easy to attack by the dioxygen molecule, thanks to its high structural dioxygen affinity. [90] Bollinger et al measured the rate constant for this process. [57] Another reason is that the substitution of water molecules to the anionic substrate or co-substrate will diminish the oxidation potential of iron(II), making the iron centre attract a dioxygen molecule. [6] In addition, the binding of the substrate or co-substrate is served as a sheltering mechanism that protects the enzyme from deactivating itself, resulting from the evidence that iron(II)-superoxo intermediate will not form until redox-active substrate or redox-active co-substrate is introduced. [6,27] (Stage C). The subsequent direct bonding of dioxygen to the vacant site on iron ligand turns the oxidation state from iron(II) to iron(III)-superoxo. This step is fast and the iron(III)-superoxo has never been trapped experimentally. The nonheme iron enzyme picks up two equivalent electrons from either the substrate or co-substrate (redox-active) to activate the redox dioxygen, which only has one electron (low redox potential). [27] Moreover, in the common motif 2-His-1-carboxylate facial triad, from a geometry aspect, one of the advantages is it is flexible to dioxygen to bind in any of three endogenous ligand positions and may react variously if they bind differently. [6,88] (Stage D) In the next step, the redox-active dioxygen will convert into the peroxide form, and form a bridged structure between the α-KG and iron centres. [91] (Stage E) This is followed by breaking of the O-O bond, and triggers the formation of a high-valent iron(IV)-oxo intermediate. [27] The reaction mechanism from stage E to A is the critical stage that is experimentally and computationally interesting as it happens at a very fast timescale. [6] Since in this stage a high-valent iron(IV)-oxo intermediate can bio-transform the target substrate into the product, and the active site of the enzyme regenerates its resting state. [92]



Scheme 1.4 Substrate-activating catalytic cycle of α-ketoglutarate (α-KG) dependent dioxygenases (Sub-H)

Comparing with detail the taurine/alpha-ketoglutarate (alpha-KG) dioxygenase (TauD) mechanism, demonstrated in **Scheme 1.4**:

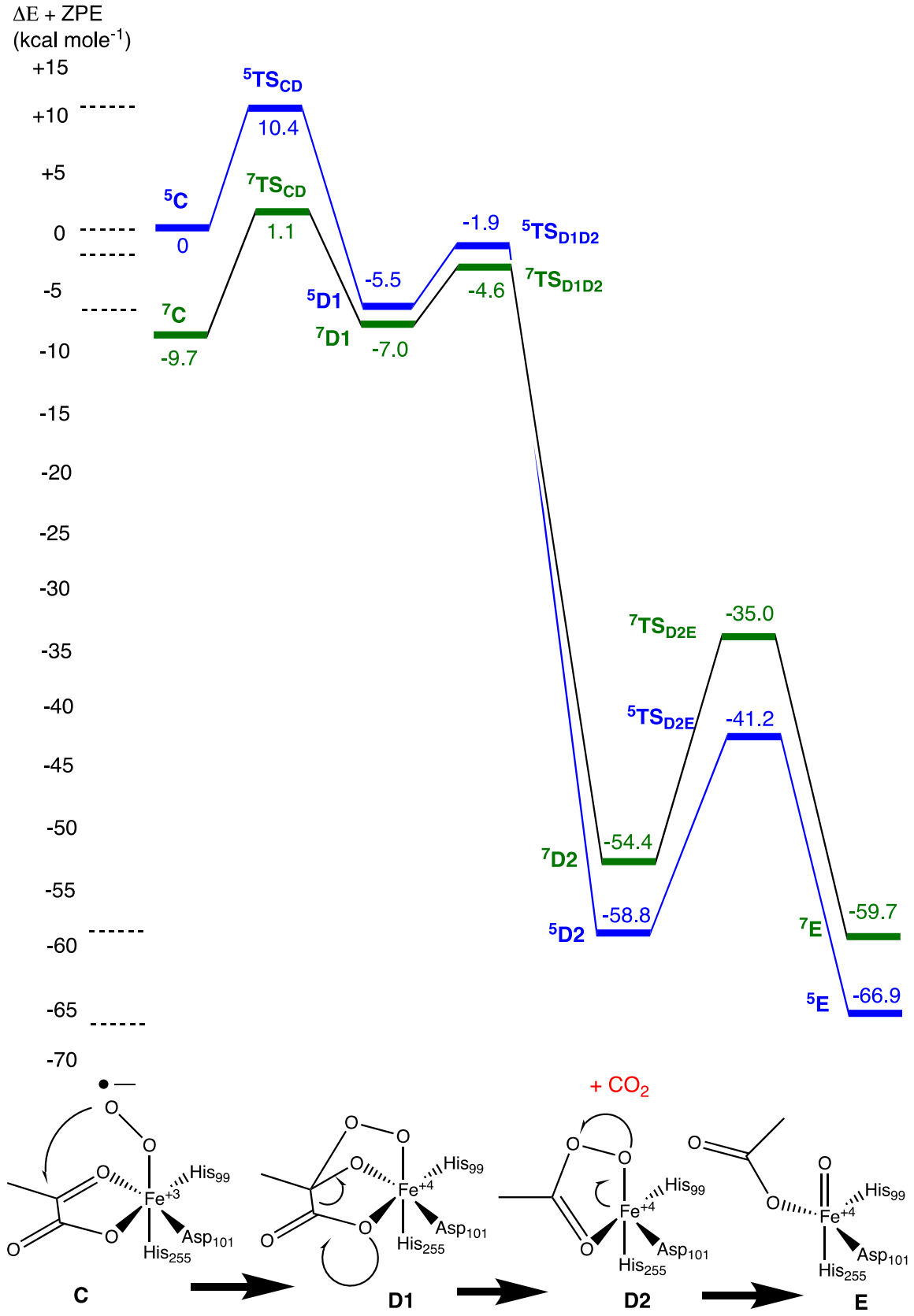
(Stage A) Starting the catalytic cycle from the iron(II) centre bonded with His-His-Asp/Glu structure motif and two ligands coordinated with alpha-KG and leaving one ligand as a neutral water molecule. When the target substrate (Sub-H) has been introduced alongside, the coordinated water molecule will detach (stage B), leaving one vacant site available for attack by a dioxygen molecule and therefore invokes stage C, where the iron(II) ion converts into iron(III). In the next step, the activated terminal super-oxo will react with the co-substrate α-KG and rearrange the motif into a bridge-like structure (stage D). This

is followed by the cleavage of the O-O bond of the iron-superoxo and carbon dioxide will be liberated from stage D. Subsequently, a high-valent Fe(IV)=O intermediate and oxidative succinate will form (stage E). After that, the highly activated terminal iron(IV) oxo will abstract the hydrogen from the target substrate to form ferric (iron(III)-hydroxylated complex, stage F).

The following steps will be the OH rebound back to the substrate radical and form the hydroxylated substrate (Sub-OH), stage F to A. Ultimately, the hydroxylated substrate (product) will liberate from iron, and the structure will rearrange to the initial stage A (resting state) assisted by the new reactant molecule. The rate determining step is stage F. The iron(IV)-oxo is the most thermodynamically stable intermediate in the catalytic cycle and the decarboxylation of α-KG ligand is an irreversible process. However, I also include the Septet spin state in comparison with the quintet state [132,133]. In comparison to the septet, it shows that the quintet is more exothermic whereas the septet spin state is more endothermic from an energy profile perspective, which deomonstrate in **Scheme 1.5.**

Stage C iron(III)-superoxo attack the co-substrate to stage D1 form the bridge like structure (activation energy 10.4 kcal mol^{-1}), display in **Scheme 1.5**. This energy profile, scheme, shows the reaction start from stage C to stage E, the catalytic cycle is extreme exothermal reaction, meaning that reaction is very fast and irreversible, especially from stage D1 to stage D2 (decarboxylation, the activation energy drop around 60 kcal mol^{-1} and most stable stage is stage E (lowest relative energy).

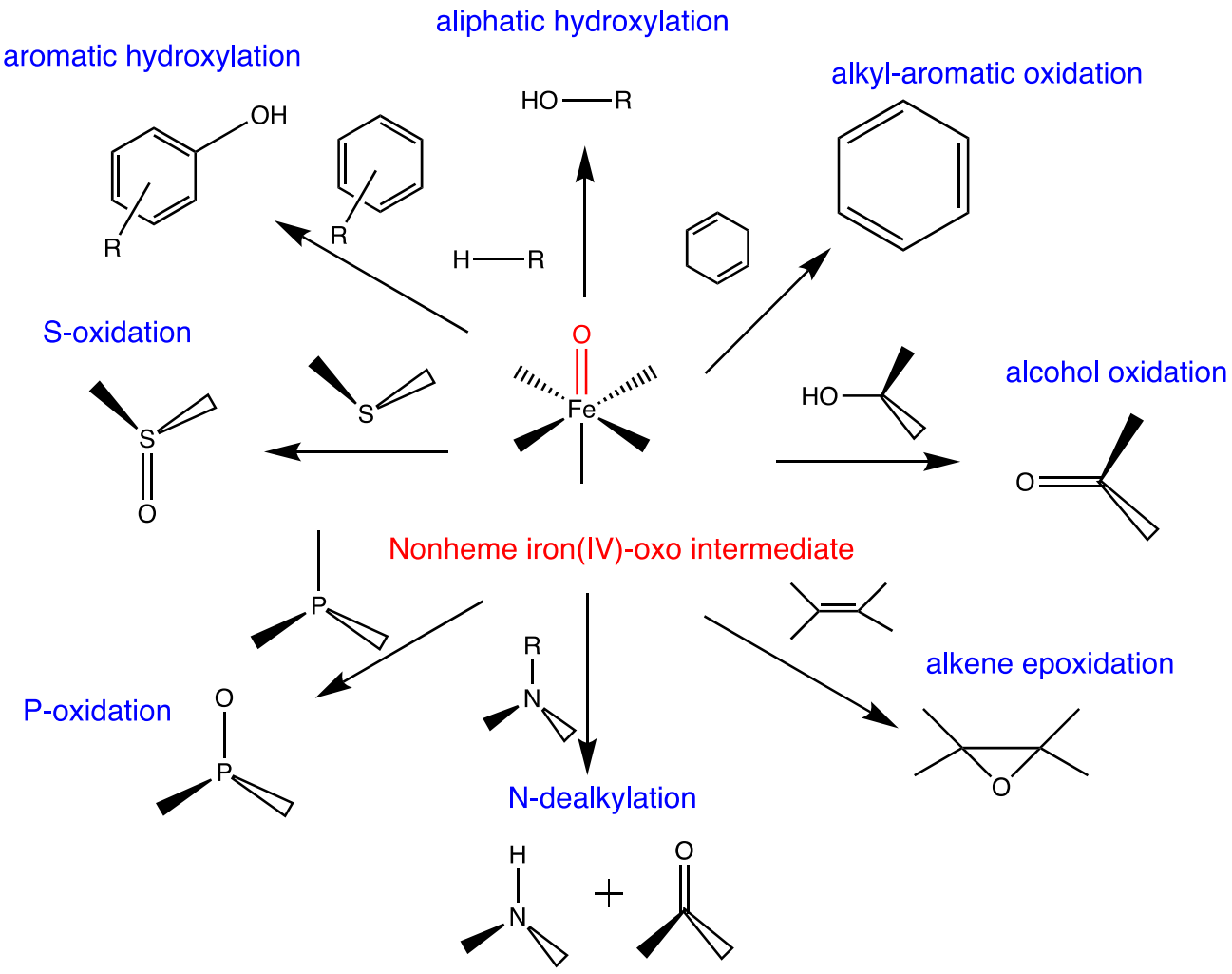
The subsequent steps is from stage D2 to stage E, after re-arrange to form the iron(IV)-oxo intermediate. It is worth to note that despite that iron(IV)-oxo from aforementioned state that is highly activate reaction but also in comparison with stage C, D1, D2, is even stable. Moreover, all intermediate B, C, D1, D2, is extreme difficult to to trap and characterisation using experimental work. That is where the computational work value are. All this steps are mostly the same from stage C to stage E. In this account, most of our project will start our model from stage E iron(IV)=O to study the divergent of the different substrate activation dependent on what enzyme we use.



Scheme 1.5 Potential energy profile for dioxygen activation by TauD cluster model. Energy units in kcal mol^{-1} (relative energy relative to C) for quintet and septet spin state obtained from method of 6-31G/ccpVTZ. Data taken from Ref 72, 132, and 133. [72,132,133]

1.4 High-valent iron(IV)-oxo intermediate

The high-valent nonheme iron(IV) oxo intermediate E was characterised and detected spectroscopically. [64,94,95] A vivid example for the α-KG-dependent enzymes, taurine/α-ketoglutarate dioxygenase (TauD), is the high-spin iron(IV)-oxo intermediate that was characterised by several spectroscopic approaches, namely X-ray absorption spectroscopy, Mössbauer and resonance Raman spectroscopy. [57,64,84,85] Both the structural and electronic properties of the high-valent nonheme iron(IV)-oxo species have been studied extensively. [96-99] In particular, many groups have created biomimetic models of iron(IV)-oxo species with varying ligand-types and reactivity patterns with model substrates were studied. The divergent of substrate activation by biomimetic complex are showing in **Scheme 1.6**, and spans the various reactions such as aromatic hydroxylation, aliphatic hydroxylation, alkyl-aromatic oxidation, alcohol oxidation, alkene epoxidation, N-dealkylation, P-oxidation, S-oxidation. [13,27]



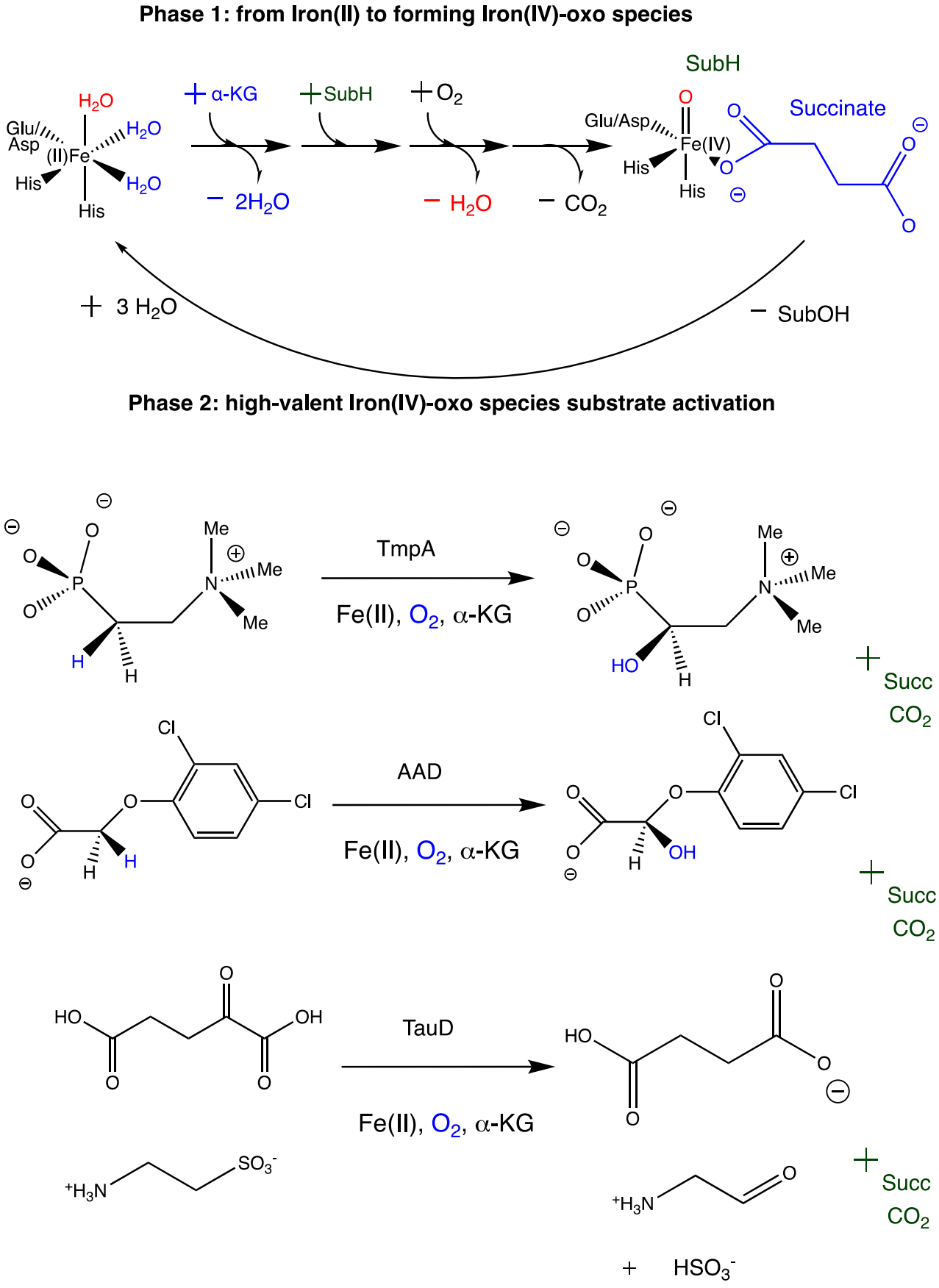
Scheme 1.6: Diversification of biomimetic complex of non-heme iron(IV)-oxo species substrate activation

There are eight primary classes of substrate activation of nonheme iron(IV)-oxo complex that can do aliphatic hydroxylation, alkyl-aromatic oxidation, alcohol oxidation, alkene epoxidation, N-dealkylation, P-oxidation, S-oxidation, and aromatic hydroxylation. [13,27] Highly electrophilic high-valent iron(IV)-oxo intermediate complexes in the non-heme iron enzyme are considered highly reactive species in the catalytic cycle, responding to dioxygen activation. [12,13,100,101]

Furthermore, the spin state of the enzymatic intermediate E is $S = 2$ in the ground state and is identified as a six-coordination structural motif, confirmed by a series of experimental procedures, such as Mössbauer measurements, spectroscopic measurements, and extended X-ray absorption fine structure spectroscopy as well as computational methods to cross-validate. [102] Apart from the hydroxylation of C-H bonds of substrates, some α-KG-dependent enzymes perform the halogenation of the substrate. [82,104,105]

Furthermore, α-KG-dependent iron enzymes can also catalyse substrate desaturation, epoxidation of C=C bonds, cyclisation of heterocyclic ring, decarboxylation, oxidative aromatic bond cleavage. [6,28,55,106] It is essential to understand the dioxygen activation reaction pathway of nonheme iron-containing enzymes to unveil its mechanistic detail.

Although the Fe(IV)-oxo species has been well-studied, one challenge remains to be resolved: whether higher-valent oxo Fe(V) is an exit for the inspiration from nature and then applied on the biomimetic field for producing the desired product. [107,108] In addition, computational methods also play an essential role in the validation and supporting the experimental part of the research on revealing these enzyme properties. [93,102,103,109-114]



Scheme 1.7 The Catalytic Cycle of αKG-Dependent Non-heme Iron Dioxygenases and Three Examples of Oxidative Transformations for (Aliphatic C−H Hydroxylation (in TmpA, AAD, TauD).

Scheme 1.7 shows the general consensus mechanism of αKG-Dependent Non-heme Iron Dioxygenases, exemplified by TmpA, AAD, and TauD, activate by the iron(IV)=O intermediate. For computational chemists, three key components are worth investigating: gaining insight in enzymatic reaction mechanisms, the electronic properties of short-lived intermediates of enzymes, identifying the rate determining step of the reaction, identification of the active oxidant in the catalytic cycle, prediction of product distributions, spectroscopic parameters, and intrinsic chemical properties of short-lived intermediates such as a high-valent iron(IV)-oxo species. [8,72]

Short-lived intermediates are hard to characterise spectroscopically, whereas theoretical modelling can easily unveil their structural and properties at the molecular level. [72] Therefore, computational modelling plays a vital role in biological chemistry and bioengineering and usually supports experimental research. Furthermore, solving the mechanism of enzymes will significantly enhance the advancement of the fine-chemical and pharmaceutical fields and produce more sustainable products for humans and the earth. [33,82,115,116] In addition, the research emerging from the computational field on bioinspired and biomimetic complexes is worth noticing.

1.5 First and Second coordinate sphere effect on a nonheme iron enzyme

Enzymes are versatile catalysts that can effectively biotransform substrates through binding and positioning them in the active site. [83] This relates to three general aspects and determines their stereoselectivity, chemo-selectivity, and regioselectivity. The first coordination sphere of the enzymatic co-factor is defined by the covalent bonds that connect the iron core to the protein and the labile ligands.

In general, these ligands are positioned in three major planes, namely the equatorial plane (xy-plane), the axial ligand, and the distal ligand, demonstrate in Figure 6. [83] The latter is typically the site where dioxygen binds and where the oxo group of the iron(IV)-oxo species is located. For instance, a 2-His-1-Asp facial triad is one example of the first coordination sphere representing the nonheme iron dioxygenases. [83]

The first coordination sphere describes the bonding and antibonding interactions of the iron orbitals with the ligands, whereby its structural motif whether 6-coordination or 5-coordination affects the ordering of the spin states. Therefore, the active site of the enzyme uses the first coordinate sphere (iron bonding structural motif) to determine the physicochemical properties of the metal centre, while the second coordination sphere (amino acid residues) incurs non-covalent interactions that stabilise the reactant complex and transition states to achieve the desired catalytic reaction. [73]

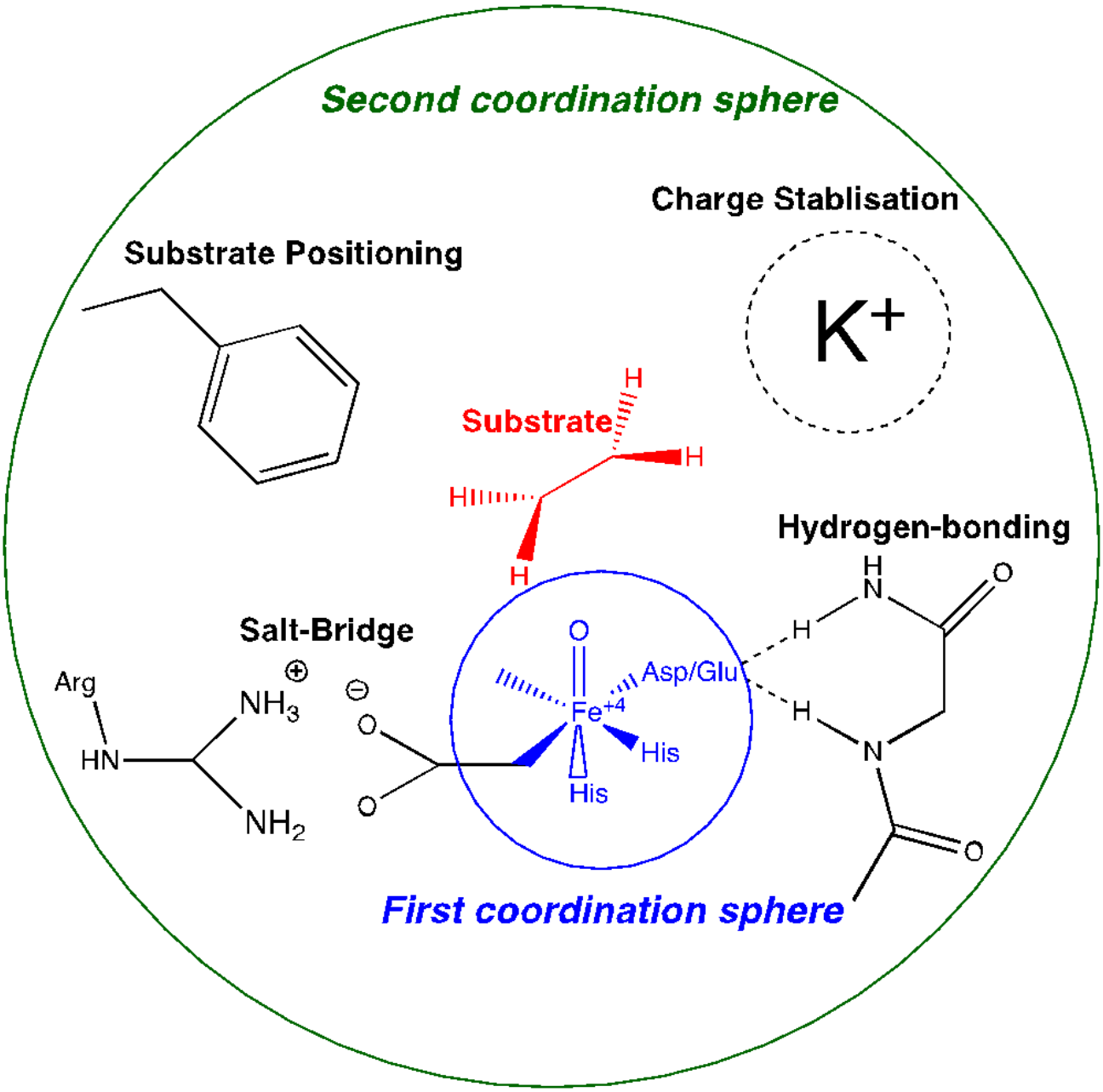


Figure 1.4 External perturbations guiding the interactions of first and second coordination sphere on a metal centre. [83]

Figure 1.4 depicts four possible long-range interactions (second coordination sphere effects) in enzyme active sites, including substrate placement, salt bridge interactions, hydrogen bonding interactions, and charge stabilisation. [83] The indirect bonding interactions toward the iron core exclusive from the first coordination sphere is called the second coordination sphere, influencing the electronic properties of enzymatic reaction. The non-covalent interactions, protein environmental perturbation inside the protein matrix including polar amino acid residues (cations and anions), hydrogen bonding, salt-bridge

interactions can significantly influence the kinetics, spectroscopic parameters, reactivity trend, and even selectivity of enzyme and hence affect its product specificity. [73,83]

Moreover, it is significantly apparent compared with the biomimetic complex, usually containing only the first coordination sphere; without the protein environment, the reactivity will be different from where the protein is present, as shown in the following chapters on TmpA, AAD enzyme. So that we will create a large model which covers the essential part of protein residues that help stabilise the reactions. [17,18] For almost two decades, the second-coordination sphere effect, non-covalent interactions, on enzymatic reactions in both heme and nonheme enzymes has received extensive attention from experimental and computational chemists. [83,117-121] Substrate binding and substrate orientation are matters of the second-coordination sphere. This ranges from an aromatic ring of amino acid residues that can influence the enzyme reactivity pattern, polar group of amino acid residues such as carboxylate groups or alcohols; functional groups will configure hydrogen-bonding with a substrate. [73] These environmental perturbations stabilise the reaction and form a geometrical "lock and key" that relatively fixes the substrate positioning.

As an example, in our 2(trimethylammonio)ethylphosphonate dioxygenase (TmpA) project all negatively and positively charged amino acid residues were combined and integrated into a local dipole moment, which triggers a negative catalysis, where a low-energy but undesirable reaction is avoided. [17,89] Similarly, trends for the viomycin biosynthesis enzyme (VioL) and the carbon starvation-induced protein D (CsiD) show major effects of the second coordination sphere. [89,122,123]

In chapter 4, the electrostatic influences (environmental charge perturbations) on TmpA, such as electric dipole-moment, external electric field effect and polar amino acid residues inside the protein matrix are discussed. [17] These studies show that second coordination sphere effects significantly affect the selectivity and product distribution of the catalytic reaction. These long-range perturbations were demonstrated to influence the oxidation state and bonding forces. Therefore, they affect the mechanism, local pKa values, selectivity, redox potential, and product distribution. [83]

There are four types of second coordination sphere interactions, including the vicinity of positive and negative charged among acid residues bonding with a substrate, hydrogen-bonding interactions between the substrate and the first coordination sphere, salt bridge interactions between the substrate or first coordination sphere (e.g. terminal carboxylate group of α-KG) with amino acid residues (e.g. arginine), and finally the substrate orientation fit in protein array. [83] Nonheme iron dioxygenases, particularly

those that are α-KG dependent, are commonly affected by the salt bridge interaction between its carboxylate group (bond in first coordination as a cofactor) and lysine or arginine amino acid residues. [17,18,83] A similar trend can be found in the crystal structures of 2-(trimethylammonio) ethylphosphonate hydroxylase (TmpA, demonstrated in Chapter 4), α-KG dependent leucine 5-hydroxylase (GriE) and TauD. [17,68,124,125] In addition, due to its functionality of being a highly tight substrate binding feature, nonheme iron dioxygenases can synthesise natural products and biodegrade toxic chemicals selectively. [83]

With the understanding of the concept of the first and second coordination sphere interactions on the identification of substrate and oxidant alignment will significantly assist the computational and experimental work in the engineering, design biomimetic complex and possibly alter stereo-, regio-, and chemo-selectivities of enzymes. [126-129] In this respect, it is vital to include the iron ligand's first and second coordination sphere in our active site of enzyme model, exemplified in the TmpA and AAD projects in chapter 4 and chapter 5, respectively. [17,18]

1.6 Research Aims and Objectives

1.6.1 General Aims and Objectives of this thesis

The purpose of this thesis is to introduce the significance of computational chemistry applied to a variety of nonheme iron dioxygenase enzymes as examples, using cutting-edge quantum cluster model techniques to unravel the three different kinds of nonheme iron enzymes (UndA, TmpA, and AAD) with different aspects and its property.

The thesis attempts to summarise some notable projects that have a direct bearing on nonheme iron enzymes and to generalise the idea of second-coordination sphere effects on the chemo-selectivity of nonheme iron dioxygenases. The function of dioxygen activation of nonheme iron dioxygenases is first introduced in Chapter 1: Introduction. Additionally, the structural motif of the 2-His-1-carboxylate facial triad is followed by the mostly natural occurring nonheme iron dioxygenases. In addition, a clear explanation the nonheme iron enzyme's catalytic cycle as well as some of the key features of the high-valent iron (IV)-oxo intermediate is provided. The first and second coordinate spheres of a nonheme iron enzyme also represent the variables that affect the enzymatic process.

1.6.2 Chapter 2: Methodology

This method chapter explains the development of the related enzymatic simulations with empathies on the analysis of the theory, computational software and procedure to optimise throughout the whole optimisation to gain more intrinsic properties by a number of computational techniques (primarily density functional theory calculation to reveal the unidentified enzymatic structures and electronic properties of the enzyme active site, and the interaction inside first and second coordination sphere , and investigating computational chemistry and its use in computational enzymology. To explain and justify the way employing all the methods and methodology, we have added not only a brief history background but also succinct sub-sections on both theory and development and advance of computational chemistry.

1.6.3 Chapter 3: nonheme iron enzymes UndA

Topic: Can a Mononuclear Iron (III)-Superoxo Active Site Catalyze the Decarboxylation of Dodecanoic Acid in UndA to Produce Biofuels?

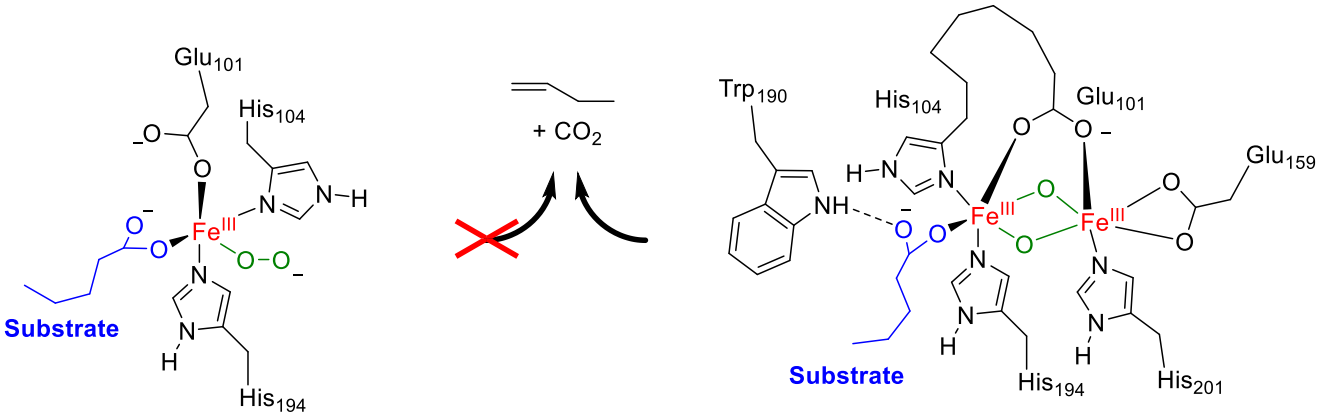


Figure 1.5 A comparison of mononuclear iron(III)-superoxo (left-hand side) and diiron-dioxo (right-hand side) models using a biofuel synthesized enzyme UndA. This simulation shows that it is unlikely to be a mononuclear nonheme iron enzyme as high barriers for hydrogen atom abstraction are obtained. For the corresponding diiron-dioxo model; however, much lower barriers are found and hence is a more likely to be oxidant.

To investigate and understand how 1-undecene biosynthesed by the mechanism of nonheme iron decarboxylase UndA activates a dodecanoic acid substrate (fatty acid decarboxylation). Resolving the controversial experimental issue of whether UndA has one iron (mononuclear enzyme) or two iron (diiron enzyme) within the active site [130][15]. Test two classes of UndA models: iron(III)-superoxo species and diiron(IV)-dioxo species model. The energy profiles for hydrogen atom abstraction and decarboxylation steps of UndA are optimised and calculated. Analyse the geometry and electronic aspects of the active sites to gain a more comprehensive view of the mechanisms, i.e. reactant, transition state, intermediate, and product complex. **Figure 1.4** show the results of the paper.

1.6.4 Chapter 4: 2-(trimethylammonio)ethylphosphonate dioxygenase (TmpA)

Topic: Electrostatic Perturbations from the Protein Affect C-H Bond Strengths of the Substrate and Enable Negative Catalysis in the TmpA Biosynthesis Enzyme

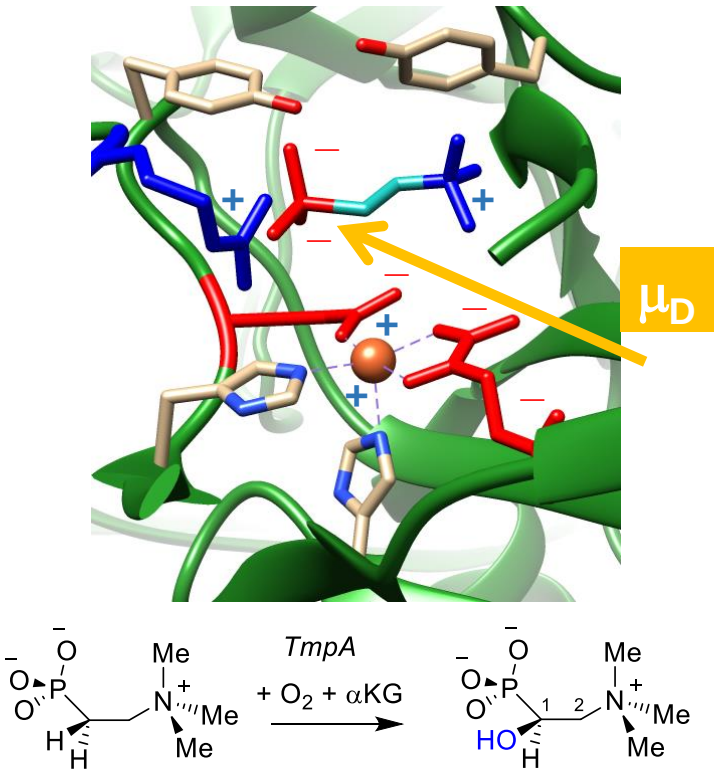


Figure 1.6 Density functional theory studies on large active site cluster models of the nonheme iron dioxygenase TmpA elucidated the origin of its chemo-selectivity. They identified it as negative catalysis, where an otherwise favourable channel is blocked. The enzyme succeeds thanks to charged residues in the substrate binding pocket that produce a local electric dipole moment and electric field. These electrostatic perturbations guide the reaction to a selective reaction mechanism.

To investigate the hydroxylation mechanistic insight of the nonheme iron dioxygenase 2-(trimethylammonio)-ethylphosphonate dioxygenase (TmpA). To explore why TmpA behave like a negative catalysis such that when using a high-valent iron (IV)-oxo to activate the stronger bond of the C1-H position of the substrate rather than start the weak bond of the C^2-H position of the substrate. Calculate the factors that cause electrostatic perturbations or protein environmental effects, stabilising the desired product's reaction and achieving efficient selectivity. To test the activation upon three different substrate possible C^1 -, C^2 - and N-methyl-positions through hydroxylation of the substrate by TmpA, illustrating in **Figure 1.6.**

1.6.5 Chapter 5: aryloxyalkanoate dioxygenase (AAD) enzyme

Topic: Biodegradation of Herbicides by a Plant Nonheme Iron Dioxygenase: Mechanism and Selectivity of Substrate Analogues

Figure 1.7 Density functional theory calculations show how plant enzymes can efficiently initiate herbicide molecule biodegradation. However, the reactivity of substituted substrate analogues is slowed down and, in some cases, encounters a significant rebound barrier.

To investigate and understand enzymatic mechanism details of herbicide biodegrading nonheme iron enzymes by Aryloxyalkanoate dioxygenases, which conduct substrate activate 2,4-dichlorophenoxyacetic acid biodegradation and 2-methyl substituted analogues (R- and S-methyl substituted inhibitors). Test the capabilities of Aryloxyalkanoate dioxygenases (AAD) enzyme is bio-transform 2,4-dichlorophenoxyacetic acid biodegradation and 2-methyl substituted analogues through C^2-H or aromatic chemo-selectively hydroxylation. Unveil the energy profiles of yloxyalkanoate dioxygenases (AAD) enzyme mechanism, activate the 2,4-dichlorophenoxyacetic acid biodegradation and 2-methyl substituted analogues (R- and S-methyl substituted inhibitors) and identify the rate-determining step, **Figure 1.7** showing the briefing of the paper.

1.7 Introduction References

[1] Que, L. (2017). 60 years of dioxygen activation. *JBIC Journal of Biological Inorganic Chemistry*, *22*(2), 171-173.

[2] Hayaishi, O. (2005). An odyssey with oxygen. *Biochemical and biophysical research communications*, *338*(1), 2-6.

[3] Waterman, M. R. (2005). Professor Howard Mason and oxygen activation. *Biochemical and biophysical research communications*, *338*(1), 7-11.

[4] Mason, H. S., Fowlks, W. L., & Peterson, E. (1955). Oxygen transfer and electron transport by the phenolase complex1. *Journal of the American Chemical Society*, *77*(10), 2914-2915.

[5] Hayaishi, O., Katagiri, M., & Rothberg, S. (1955). Mechanism of the pyrocatechase reaction. *Journal of the American Chemical Society*, *77*(20), 5450-5451.

[6] Koehntop, K. D., Emerson, J. P., & Que, L. (2005). The 2-His-1-carboxylate facial triad: a versatile platform for dioxygen activation by mononuclear non-heme iron (II) enzymes. JBIC Journal of Biological Inorganic Chemistry, 10(2), 87-93.

[7] De Montellano, P. R. O. (Ed.). (2005). *Cytochrome P450: structure, mechanism, and biochemistry* (Vol. 115). New York: Kluwer Academic/Plenum Publishers.

[8] de Visser, S. P. (2009). Elucidating enzyme mechanism and intrinsic chemical properties of short-lived intermediates in the catalytic cycles of cysteine dioxygenase and taurine/α-ketoglutarate dioxygenase. *Coordination Chemistry Reviews*, *253*(5-6), 754-768.

[9] Solomon, E. I., Brunold, T. C., Davis, M. I., Kemsley, J. N., Lee, S. K., Lehnert, N., ... & Zhou, J. (2000). Geometric and electronic structure/function correlations in non-heme iron enzymes. Chemical reviews, 100(1), 235-350.

[10] Bugg, T. D. (2003). Dioxygenase enzymes: catalytic mechanisms and chemical models. *Tetrahedron*, *36*(59), 7075-7101.

[11] Costas, M., Mehn, M. P., Jensen, M. P., & Que, L. (2004). Dioxygen activation at mononuclear nonheme iron active sites: enzymes, models, and intermediates. *Chemical reviews*, *104*(2), 939-986.

[12] Abu-Omar, M. M., Loaiza, A., & Hontzeas, N. (2005). Reaction mechanisms of mononuclear non-heme iron oxygenases. Chemical reviews, 105(6), 2227-2252.

[13] Nam, W. (2007). High-valent iron (IV)–oxo complexes of heme and non-heme ligands in oxygenation reactions. *Accounts of Chemical Research*, *40*(7), 522-531.

[14] Wise, C. E., Grant, J. L., Amaya, J. A., Ratigan, S. C., Hsieh, C. H., Manley, O. M., & Makris, T. M. (2017). Divergent mechanisms of iron-containing enzymes for hydrocarbon biosynthesis. *JBIC Journal of Biological Inorganic Chemistry*, *22*(2), 221-235.

[15] Lin, Y. T., Stańczak, A., Manchev, Y., Straganz, G. D., & de Visser, S. P. (2020). Can a mononuclear iron (III)-superoxo active site catalyze the decarboxylation of dodecanoic acid in UndA to produce biofuels?. *Chemistry–A European Journal*, *26*(10), 2233-2242.

[16] Rui, Z., Li, X., Zhu, X., Liu, J., Domigan, B., Barr, I., ... & Zhang, W. (2014). Microbial biosynthesis of medium-chain 1-alkenes by a nonheme iron oxidase. *Proceedings of the National Academy of Sciences*, *111*(51), 18237-18242.

[17] Lin, Y. T., Ali, H. S., & de Visser, S. P. (2021). Electrostatic Perturbations from the Protein Affect C− H Bond Strengths of the Substrate and Enable Negative Catalysis in the TmpA Biosynthesis Enzyme. *Chemistry–A European Journal*, *27*(34), 8851-8864.

[18] Lin, Y. T., Ali, H. S., & de Visser, S. P. (2022). Biodegradation of herbicides by a plant nonheme iron dioxygenase: mechanism and selectivity of substrate analogues. Chemistry–A European Journal, 28(7), e202103982.

[19] Elwell, C. E., Gagnon, N. L., Neisen, B. D., Dhar, D., Spaeth, A. D., Yee, G. M., & Tolman, W. B. (2017). Copper–oxygen complexes revisited: structures, spectroscopy, and reactivity. *Chemical reviews*, *117*(3), 2059-2107.

[20] Keown, W., Gary, J. B., & Stack, T. D. (2017). High-valent copper in biomimetic and biological oxidations. Journal of biological inorganic chemistry : JBIC : a publication of the Society of Biological Inorganic Chemistry, 22(2-3), 289–305. https://doi-org.manchester.idm.oclc.org/10.1007/s00775-016-1420-5
[21] Kim, Y., Mai, B. K., & Park, S. (2017). VTST/MT studies of the catalytic mechanism of C-H activation by transition metal complexes with [Cu2(μ-O2)], [Fe2(μ-O2)] and Fe(IV)-O cores based on DFT potential energy surfaces. *Journal of biological inorganic chemistry : JBIC : a publication of the Society of Biological Inorganic Chemistry*, *22*(2-3), 321–338. https://doi-org.manchester.idm.oclc.org/10.1007/s00775-017-1441-8
[22] Kal, S., & Que, L. (2017). Dioxygen activation by nonheme iron enzymes with the 2-His-1-carboxylate facial triad that generate high-valent oxoiron oxidants. *JBIC Journal of Biological Inorganic Chemistry*, *22*(2), 339-365.
[23] Hegg, E. L., & Jr, L. Q. (1997). The 2-His-1-carboxylate facial triad—an emerging structural motif in mononuclear non-heme iron (II) enzymes. *European Journal of Biochemistry*, *250*(3), 625-629.
[24] Fiedler, A. T., & Fischer, A. A. (2017). Oxygen activation by mononuclear Mn, Co, and Ni centers in biology and synthetic complexes. *JBIC Journal of Biological Inorganic Chemistry*, *22*(2), 407-424.
[25] Que, L., & Tolman, W. B. (2008). Biologically inspired oxidation catalysis. *Nature*, *455*(7211), 333-340.
[26] Sheldon, R. A., Arends, I., & Hanefeld, U. (2007). *Green chemistry and catalysis*. John Wiley & Sons.
[27] Bruijnincx, P. C., van Koten, G., & Gebbink, R. J. K. (2008). Mononuclear non-heme iron enzymes with the 2-His-1-carboxylate facial triad: recent developments in enzymology and modeling studies. *Chemical Society Reviews*, *37*(12), 2716-2744.
[28] Lee, J. L., Ross, D. L., Barman, S. K., Ziller, J. W., & Borovik, A. S. (2021). C–H Bond Cleavage by Bioinspired Nonheme Metal Complexes. *Inorganic Chemistry*, *60*(18), 13759-13783.
[29] Arakawa, H., Aresta, M., Armor, J. N., Barteau, M. A., Beckman, E. J., Bell, A. T., ... & Tumas, W. (2001). Catalysis research of relevance to carbon management: progress, challenges, and opportunities. *Chemical reviews*, *101*(4), 953-996.
[30] Tolman, W. B. (Ed.). (2006). Activation of small molecules: organometallic and bioinorganic perspectives. John Wiley & Sons.
[31] Tshuva, E. Y., & Lippard, S. J. (2004). Synthetic models for non-heme carboxylate-bridged diiron metalloproteins: strategies and tactics. *Chemical reviews*, *104*(2), 987-1012.
[32] Warren, J. J., Tronic, T. A., & Mayer, J. M. (2010). Thermochemistry of proton-coupled electron transfer reagents and its implications. *Chemical reviews*, *110*(12), 6961-7001.
[33] Meunier, B., Brudvig, G., Mclain, J. L., Murahashi, S. I., Pecoraro, V. L., Riley, D., ... & Zhu, H. (Eds.). (2000). *Biomimetic oxidations catalyzed by transition metal complexes*. World Scientific.
[34] Mayer, J. M. (2004). Proton-coupled electron transfer: a reaction chemist's view. *Annu. Rev. Phys. Chem.*, *55*, 363-390.
[35] Borovik, A. S. (2011). Role of metal–oxo complexes in the cleavage of C–H bonds. *Chemical Society Reviews*, *40*(4), 1870-1874.
[36] Darcy, J. W., Koronkiewicz, B., Parada, G. A., & Mayer, J. M. (2018). A continuum of proton-coupled electron transfer reactivity. *Accounts of chemical research*, *51*(10), 2391-2399.
[37] Jarvis, A. G. (2020). Designer metalloenzymes for synthetic biology: Enzyme hybrids for catalysis. *Current opinion in chemical biology*, *58*, 63-71.
[38] Das, A., Hessin, C., Ren, Y., & Desage-El Murr, M. (2020). Biological concepts for catalysis and reactivity: empowering bioinspiration. *Chemical Society Reviews*, *49*(23), 8840-8867.
[39] Ghosh, A. C., Duboc, C., & Gennari, M. (2021). Synergy between metals for small molecule activation: Enzymes and bio-inspired complexes. *Coordination Chemistry Reviews*, *428*, 213606.
[40] SP de Visser, D. K. (2011). Iron-Containing Enzymes: Versatile Catalysts of Hydroxylation Reaction in Nature.

[41] White, M. D., & Flashman, E. (2016). Catalytic strategies of the non-heme iron dependent oxygenases and their roles in plant biology. *Current Opinion in Chemical Biology*, *31*, 126-135.
[42] Huang, X., & Groves, J. T. (2017). Oxygen activation and radical transformations in heme proteins and metalloporphyrins. *Chemical reviews*, *118*(5), 2491-2553.
[43] Krest, C. M., Onderko, E. L., Yosca, T. H., Calixto, J. C., Karp, R. F., Livada, J., ... & Green, M. T. (2013). Reactive intermediates in cytochrome P450 catalysis. *Journal of Biological Chemistry*, *288*(24), 17074-17081.
[44] Shook, R. L., & Borovik, A. S. (2010). Role of the secondary coordination sphere in metal-mediated dioxygen activation. *Inorganic chemistry*, *49*(8), 3646-3660.
[45] Cook, S. A., & Borovik, A. S. (2015). Molecular designs for controlling the local environments around metal ions. *Accounts of chemical research*, *48*(8), 2407-2414.
[46] Guo, M., Corona, T., Ray, K., & Nam, W. (2018). Heme and nonheme high-valent iron and manganese oxo cores in biological and abiological oxidation reactions. *ACS Central Science*, *5*(1), 13-28.
[47] Kovaleva, E. G., & Lipscomb, J. D. (2008). Versatility of biological non-heme Fe (II) centers in oxygen activation reactions. *Nature chemical biology*, *4*(3), 186-193.
[48] Schofield, C. J., & Zhang, Z. (1999). Structural and mechanistic studies on 2-oxoglutarate-dependent oxygenases and related enzymes. *Current opinion in structural biology*, *9*(6), 722-731.
[49] Solomon, E. I., Light, K. M., Liu, L. V., Srnec, M., & Wong, S. D. (2013). Geometric and electronic structure contributions to function in non-heme iron enzymes. *Accounts of chemical research*, *46*(11), 2725-2739.
[50] Bollinger Jr, J. M., & Krebs, C. (2007). Enzymatic C–H activation by metal–superoxo intermediates. *Current opinion in chemical biology*, *11*(2), 151-158.
[51] Solomon, E. I., Decker, A., & Lehnert, N. (2003). Non-heme iron enzymes: contrasts to heme catalysis. *Proceedings of the National Academy of Sciences*, *100*(7), 3589-3594.
[52] Wackett, L. P. (2002). Mechanism and applications of Rieske non-heme iron dioxygenases. *Enzyme and Microbial Technology*, *31*(5), 577-587.
[53] Ramos, J. L. (2004). Volume 1: Genomics, life style and molecular architecture.
[54] Fitzpatrick, P. F. (2003). Mechanism of aromatic amino acid hydroxylation. *Biochemistry*, *42*(48), 14083-14091.
[55] Hausinger, R. P. (2004). Fe (II)/α-ketoglutarate-dependent hydroxylases and related enzymes. *Critical reviews in biochemistry and molecular biology*, *39*(1), 21-68.
[56] Ryle, M. J., & Hausinger, R. P. (2002). Non-heme iron oxygenases. Current opinion in chemical biology, 6(2), 193-201.
[57] Bollinger Jr, J. M., Price, J. C., Hoffart, L. M., Barr, E. W., & Krebs, C. (2005). Mechanism of taurine: α-ketoglutarate dioxygenase (TauD) from Escherichia coli. European journal of inorganic chemistry, 2005(21), 4245-4254.
[58] Straganz, G. D., & Nidetzky, B. (2006). Variations of the 2-His-1-carboxylate theme in mononuclear non-heme FeII oxygenases. ChemBioChem, 7(10), 1536-1548.
[59] Higgins, L. J., Yan, F., Liu, P., Liu, H. W., & Drennan, C. L. (2005). Structural insight into antibiotic fosfomycin biosynthesis by a mononuclear iron enzyme. Nature, 437(7060), 838-844.
[60] Choroba, O. W., Williams, D. H., & Spencer, J. B. (2000). Biosynthesis of the vancomycin group of antibiotics: involvement of an unusual dioxygenase in the pathway to (S)-4-hydroxyphenylglycine. Journal of the American Chemical Society, 122(22), 5389-5390.
[61] Bugg, T. D. (2001). Oxygenases: mechanisms and structural motifs for O2 activation. *Current Opinion in Chemical Biology*, *5*(5), 550-555.
[62] Krebs, C., Galonić Fujimori, D., Walsh, C. T., & Bollinger, J. M., Jr (2007). Non-heme Fe(IV)-oxo intermediates. *Accounts of chemical research*, *40*(7), 484–492. https://doi-org.manchester.idm.oclc.org/10.1021/ar700066p
[63] Blomberg, M. R., Borowski, T., Himo, F., Liao, R. Z., & Siegbahn, P. E. (2014). Quantum chemical studies of mechanisms for metalloenzymes. *Chemical reviews*, *114*(7), 3601-3658.

[64] Price, J. C., Barr, E. W., Tirupati, B., Bollinger, J. M., & Krebs, C. (2003). The first direct characterization of a high-valent iron intermediate in the reaction of an α-ketoglutarate-dependent dioxygenase: a high-spin Fe (IV) complex in taurine/α-ketoglutarate dioxygenase (TauD) from Escherichia coli. Biochemistry, 42(24), 7497-7508.
[65] Proshlyakov, D. A., McCracken, J., & Hausinger, R. P. (2017). Spectroscopic analyses of 2-oxoglutarate-dependent oxygenases: TauD as a case study. *JBIC Journal of Biological Inorganic Chemistry*, *22*(2), 367-379.
[66] Sage, J. T., Paxson, C., Wyllie, G. R. A., Sturhahn, W., Durbin, S. M., Champion, P. M., ... & Scheidt, W. R. (2001). Nuclear resonance vibrational spectroscopy of a protein active-site mimic. *Journal of Physics: Condensed Matter*, *13*(34), 7707.
[67] Wong, S. D., Srnec, M., Matthews, M. L., Liu, L. V., Kwak, Y., Park, K., ... & Solomon, E. I. (2013). Elucidation of the Fe (IV)= O intermediate in the catalytic cycle of the halogenase SyrB2. *Nature*, *499*(7458), 320-323.
[68] O'Brien, J. R., Schuller, D. J., Yang, V. S., Dillard, B. D., & Lanzilotta, W. N. (2003). Substrate-induced conformational changes in Escherichia coli taurine/α-ketoglutarate dioxygenase and insight into the oligomeric structure. *Biochemistry*, *42*(19), 5547-5554.
[69] Mirica, L. M., McCusker, K. P., Munos, J. W., Liu, H. W., & Klinman, J. P. (2008). 18O kinetic isotope effects in non-heme iron enzymes: Probing the nature of Fe/O2 intermediates. *Journal of the American Chemical Society*, *130*(26), 8122-8123.
[70] Krebs, C., Galonic Fujimori, D., Walsh, C. T., & Bollinger Jr, J. M. (2007). Non-heme Fe (IV)–oxo intermediates. *Accounts of chemical research*, *40*(7), 484-492.
[71] Hanauske-Abel, H. M., & Günzler, V. (1982). A stereochemical concept for the catalytic mechanism of prolylhydroxylase: applicability to classification and design of inhibitors. *Journal of Theoretical Biology*, *94*(2), 421-455.
[72] de Visser, S. P. (2018). Mechanistic insight on the activity and substrate selectivity of nonheme iron dioxygenases. *The Chemical Record*, *18*(10), 1501-1516.
[73] Mukherjee, G., Satpathy, J. K., Bagha, U. K., Mubarak, M. Q. E., Sastri, C. V., & de Visser, S. P. (2021). Inspiration from Nature: influence of engineered ligand scaffolds and auxiliary factors on the reactivity of biomimetic oxidants. *ACS Catalysis*, *11*(15), 9761-9797.
[74] Trewick, S. C., Henshaw, T. F., Hausinger, R. P., Lindahl, T., & Sedgwick, B. (2002). Oxidative demethylation by Escherichia coli AlkB directly reverts DNA base damage. *Nature*, *419*(6903), 174-178.
[75] Falnes, P. Ø., Johansen, R. F., & Seeberg, E. (2002). AlkB-mediated oxidative demethylation reverses DNA damage in Escherichia coli. *Nature*, *419*(6903), 178-182.
[76] O'Brien, P. J. (2006). Catalytic promiscuity and the divergent evolution of DNA repair enzymes. *Chemical reviews*, *106*(2), 720-752.
[77] Yi, C., Yang, C. G., & He, C. (2009). A non-heme iron-mediated chemical demethylation in DNA and RNA. *Accounts of chemical research*, *42*(4), 519-529.
[78] Kershaw, N. J., Caines, M. E., Sleeman, M. C., & Schofield, C. J. (2005). The enzymology of clavam and carbapenem biosynthesis. *Chemical communications*, (34), 4251-4263.
[79] Kivirikko, K. I., & Myllyharju, J. (1998). Prolyl 4-hydroxylases and their protein disulfide isomerase subunit. *Matrix Biology*, *16*(7), 357-368.
[80] Hirota, K., & Semenza, G. L. (2005). Regulation of hypoxia-inducible factor 1 by prolyl and asparaginyl hydroxylases. *Biochemical and biophysical research communications*, *338*(1), 610-616.
[81] Klose, R. J., Yamane, K., Bae, Y., Zhang, D., Erdjument-Bromage, H., Tempst, P., ... & Zhang, Y. (2006). The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36. *Nature*, *442*(7100), 312-316.
[82] Schofield, C., & Hausinger, R. (Eds.). (2015). *2-Oxoglutarate-dependent oxygenases*. Royal Society of Chemistry.
[83] de Visser, S. P. (2020). Second-Coordination Sphere Effects on Selectivity and Specificity of Heme and Nonheme Iron Enzymes. *Chemistry–A European Journal*, *26*(24), 5308-5327.

[84] Proshlyakov, D. A., Henshaw, T. F., Monterosso, G. R., Ryle, M. J., & Hausinger, R. P. (2004). Direct detection of oxygen intermediates in the non-heme Fe enzyme taurine/α-ketoglutarate dioxygenase. Journal of the American Chemical Society, 126(4), 1022-1023.
[85] Riggs-Gelasco, P. J., Price, J. C., Guyer, R. B., Brehm, J. H., Barr, E. W., Bollinger Jr, J. M., & Krebs, C. (2004). EXAFS spectroscopic evidence for an Fe=O unit in the Fe (IV) intermediate observed during oxygen activation by taurine: α-ketoglutarate dioxygenase. Journal of the American Chemical Society, 126(26), 8108-8109.
[86] Valegård, K., Van Scheltinga, A. C. T., Lloyd, M. D., Hara, T., Ramaswamy, S., Perrakis, A., ... & Andersson, I. (1998). Structure of a cephalosporin synthase. *Nature*, *394*(6695), 805-809.
[87] Neidig, M. L., Brown, C. D., Light, K. M., Fujimori, D. G., Nolan, E. M., Price, J. C., ... & Solomon, E. I. (2007). CD and MCD of CytC3 and taurine dioxygenase: role of the facial triad in α-KG-dependent oxygenases. *Journal of the American Chemical Society*, *129*(46), 14224-14231.
[88] Que, L. (2000). One motif—many different reactions. *Nature structural biology*, *7*(3), 182-184.
[89] de Visser, S. P., Lin, Y. T., Ali, H. S., Bagha, U. K., Mukherjee, G., & Sastri, C. V. (2021). Negative catalysis/non-Bell-Evans-Polanyi reactivity by metalloenzymes: Examples from mononuclear heme and non-heme iron oxygenases. Coordination Chemistry Reviews, 439, 213914.
[90] Neidig, M. L., & Solomon, E. I. (2005). Structure–function correlations in oxygen activating non-heme iron enzymes. *Chemical communications*, (47), 5843-5863.
[91] Karlsson, A., Parales, J. V., Parales, R. E., Gibson, D. T., Eklund, H., & Ramaswamy, S. (2003). Crystal structure of naphthalene dioxygenase: side-on binding of dioxygen to iron. *Science*, *299*(5609), 1039-1042.
[92] De Visser, S. P. (2011). Experimental and Computational Studies on the Catalytic Mechanism of Non-heme Iron Dioxygenases. *Iron-containing enzymes*, 1-41.
[93] De Visser, S. P. (2007). Can the peroxosuccinate complex in the catalytic cycle of taurine/α-ketoglutarate dioxygenase (TauD) act as an alternative oxidant?. Chemical communications, (2), 171-173.
[94] Bollinger Jr, J. M., & Krebs, C. (2006). Stalking intermediates in oxygen activation by iron enzymes: Motivation and method. *Journal of inorganic biochemistry*, *100*(4), 586-605.
[95] Krebs, C., Price, J. C., Baldwin, J., Saleh, L., Green, M. T., & Bollinger, J. M. (2005). Rapid freeze-quench 57Fe Mössbauer spectroscopy: monitoring changes of an iron-containing active site during a biochemical reaction. *Inorganic chemistry*, *44*(4), 742-757.
[96] Rohde, J. U., In, J. H., Lim, M. H., Brennessel, W. W., Bukowski, M. R., Stubna, A., ... & Que Jr, L. (2003). Crystallographic and spectroscopic characterization of a nonheme Fe (IV)= O complex. *Science*, *299*(5609), 1037-1039.
[97] Lim, M. H., Rohde, J. U., Stubna, A., Bukowski, M. R., Costas, M., Ho, R. Y., ... & Que, L. (2003). An FeIVO complex of a tetradentate tripodal nonheme ligand. *Proceedings of the National Academy of Sciences*, *100*(7), 3665-3670.
[98] De Oliveira, F. T., Chanda, A., Banerjee, D., Shan, X., Mondal, S., Que Jr, L., ... & Collins, T. J. (2007). Chemical and spectroscopic evidence for an FeV-oxo complex. *Science*, *315*(5813), 835-838.
[99] Pestovsky, O., Stoian, S., Bominaar, E. L., Shan, X., Münck, E., Que Jr, L., & Bakac, A. (2005). Aqueous FeIV□ O: Spectroscopic Identification and Oxo-Group Exchange. *Angewandte Chemie*, *117*(42), 7031-7034.
[100] Shan, X., & Que Jr, L. (2006). High-valent nonheme iron-oxo species in biomimetic oxidations. *Journal of inorganic biochemistry*, *100*(4), 421-433.
[101] Kryatov, S. V., Rybak-Akimova, E. V., & Schindler, S. (2005). Kinetics and mechanisms of formation and reactivity of non-heme iron oxygen intermediates. *Chemical reviews*, *105*(6), 2175-2226.
[102] Sinnecker, S., Svensen, N., Barr, E. W., Ye, S., Bollinger Jr, J. M., Neese, F., & Krebs, C. (2007). Spectroscopic and computational evaluation of the structure of the high-spin Fe (IV)-oxo intermediates in taurine: α-ketoglutarate dioxygenase from Escherichia coli and its His99Ala ligand variant. *Journal of the American Chemical Society*, *129*(19), 6168-6179.

[103] de Visser, S. P. (2006). Propene activation by the oxo-iron active species of taurine/α-ketoglutarate dioxygenase (TauD) enzyme. How does the catalysis compare to heme-enzymes?. *Journal of the American Chemical Society*, *128*(30), 9813-9824.
[104] Chang, W. C., Liu, P., & Guo, Y. (2018). Mechanistic elucidation of two catalytically versatile iron (II)-and α-ketoglutarate-dependent enzymes: cases beyond hydroxylation. Comments on Inorganic Chemistry, 38(4), 127-165.
[105] Srnec, M., Wong, S. D., Matthews, M. L., Krebs, C., Bollinger Jr, J. M., & Solomon, E. I. (2016). Electronic structure of the ferryl intermediate in the α-ketoglutarate dependent non-heme iron halogenase SyrB2: Contributions to H atom abstraction reactivity. Journal of the American Chemical Society, 138(15), 5110-5122.
[106] Vaillancourt, F. H., Yeh, E., Vosburg, D. A., Garneau-Tsodikova, S., & Walsh, C. T. (2006). Nature's inventory of halogenation catalysts: oxidative strategies predominate. Chemical reviews, 106(8), 3364-3378.
[107] Kal, S., Xu, S., & Que Jr, L. (2020). Bio-inspired Nonheme Iron Oxidation Catalysis: Involvement of Oxoiron (V) Oxidants in Cleaving Strong C− H Bonds. *Angewandte Chemie International Edition*, *59*(19), 7332-7349.
[108] Dantignana, V., Company, A., & Costas, M. (2020). Oxoiron (V) Complexes of Relevance in Oxidation Catalysis of Organic Substrates. *Israel Journal of Chemistry*, *60*(10-11), 1004-1018.
[109] Godfrey, E., Porro, C. S., & De Visser, S. P. (2008). Comparative Quantum Mechanics/Molecular Mechanics (QM/MM) and Density Functional Theory Calculations on the Oxo− Iron Species of Taurine/α-Ketoglutarate Dioxygenase. The Journal of Physical Chemistry A, 112(11), 2464-2468.
[110] de Visser, S. P., Quesne, M. G., Martin, B., Comba, P., & Ryde, U. (2014). Computational modelling of oxygenation processes in enzymes and biomimetic model complexes. Chemical Communications, 50(3), 262-282.
[111] Borowski, T., Bassan, A., & Siegbahn, P. E. (2004). Mechanism of Dioxygen Activation in 2-Oxoglutarate-Dependent Enzymes: A Hybrid DFT Study. Chemistry–A European Journal, 10(4), 1031-1041.
[112] Shaik, S., Hirao, H., & Kumar, D. (2007). Reactivity of high-valent iron–oxo species in enzymes and synthetic reagents: a tale of many states. Accounts of chemical research, 40(7), 532-542.
[113] Cho, K. B., Hirao, H., Shaik, S., & Nam, W. (2016). To rebound or dissociate? This is the mechanistic question in C–H hydroxylation by heme and nonheme metal–oxo complexes. Chemical Society Reviews, 45(5), 1197-1210.
[114] Geng, C., Ye, S., & Neese, F. (2014). Does a higher metal oxidation state necessarily imply higher reactivity toward H-atom transfer? A computational study of C–H bond oxidation by high-valent iron-oxo and-nitrido complexes. Dalton Transactions, 43(16), 6079-6086.
[115] Puri, M., & Que Jr, L. (2015). Toward the synthesis of more reactive S= 2 non-heme oxoiron (IV) complexes. Accounts of Chemical Research, 48(8), 2443-2452.
[116] Mahadevan, V., Gebbink, R. K., & Stack, T. D. P. (2000). Biomimetic modeling of copper oxidase reactivity. Current opinion in chemical biology, 4(2), 228-234.
[117] Müller-Dethlefs, K., & Hobza, P. (2000). Noncovalent interactions: a challenge for experiment and theory. *Chemical Reviews*, *100*(1), 143-168.
[118] Mahadevi, A. S., & Sastry, G. N. (2013). Cation− π interaction: Its role and relevance in chemistry, biology, and material science. Chemical reviews, 113(3), 2100-2138.
[119] Mahadevi, A. S., & Sastry, G. N. (2016). Cooperativity in noncovalent interactions. Chemical Reviews, 116(5), 2775-2825.
[120] Mahmudov, K. T., Kopylovich, M. N., da Silva, M. F. C. G., & Pombeiro, A. J. (2017). Non-covalent interactions in the synthesis of coordination compounds: Recent advances. Coordination Chemistry Reviews, 345, 54-72.

[121] Juanes, M., Saragi, R. T., Caminati, W., & Lesarri, A. (2019). The Hydrogen Bond and Beyond: Perspectives for Rotational Investigations of Non-Covalent Interactions. Chemistry–A European Journal, 25(49), 11402-11411.
[122] Dixit, V. A., Warwicker, J., & de Visser, S. P. (2020). How Do Metal Ions Modulate the Rate-Determining Electron-Transfer Step in Cytochrome P450 Reactions?. Chemistry–A European Journal, 26(66), 15270-15281.
[123] Ali, H. S., Henchman, R. H., & de Visser, S. P. (2021). Mechanism of Oxidative Ring-Closure as Part of the Hygromycin Biosynthesis Step by a Nonheme Iron Dioxygenase. ChemCatChem, 13(13), 3054-3066.
[124] Lukat, P., Katsuyama, Y., Wenzel, S., Binz, T., König, C., Blankenfeldt, W., ... & Müller, R. (2017). Biosynthesis of methyl-proline containing griselimycins, natural products with anti-tuberculosis activity. *Chemical science*, *8*(11), 7521-7527.
[125] Rajakovich, L. J., Pandelia, M. E., Mitchell, A. J., Chang, W. C., Zhang, B., Boal, A. K., ... & Bollinger Jr, J. M. (2019). A new microbial pathway for organophosphonate degradation catalyzed by two previously misannotated non-heme-iron oxygenases. Biochemistry, 58(12), 1627-1647.
[126] Sahu, S., Widger, L. R., Quesne, M. G., De Visser, S. P., Matsumura, H., Moënne-Loccoz, P., ... & Goldberg, D. P. (2013). Secondary coordination sphere influence on the reactivity of nonheme iron (II) complexes: an experimental and DFT approach. Journal of the American Chemical Society, 135(29), 10590-10593.
[127] Hill, E. A., Weitz, A. C., Onderko, E., Romero-Rivera, A., Guo, Y., Swart, M., ... & Borovik, A. S. (2016). Reactivity of an FeIV-oxo complex with protons and oxidants. Journal of the American Chemical Society, 138(40), 13143-13146.
[128] Pangia, T. M., Yadav, V., Gérard, E. F., Lin, Y. T., de Visser, S. P., Jameson, G. N., & Goldberg, D. P. (2019). Mechanistic investigation of oxygen rebound in a mononuclear nonheme iron complex. Inorganic chemistry, 58(15), 9557-9561.
[129] Mubarak, M. Q. E., & de Visser, S. P. (2019). Second-coordination sphere effect on the reactivity of vanadium–peroxo complexes: A computational study. Inorganic Chemistry, 58(23), 15741-15750.
[130] Manley, O. M., Fan, R., Guo, Y., & Makris, T. M. (2019). Oxidative decarboxylase UndA utilizes a dinuclear iron cofactor. Journal of the American Chemical Society, 141(22), 8684-8688.
[131] Chekan, J. R., Ongpipattanakul, C., Wright, T. R., Zhang, B., Bollinger, J. M., Rajakovich, L. J., ... & Nair, S. K. (2019). Molecular basis for enantioselective herbicide degradation imparted by aryloxyalkanoate dioxygenases in transgenic plants. Proceedings of the National Academy of Sciences, 116(27), 13299-13304.
[132] De Visser, S. P. (2007). Can the peroxosuccinate complex in the catalytic cycle of taurine/α-ketoglutarate dioxygenase (TauD) act as an alternative oxidant?. Chemical communications, (2), 171-173.
[133] Topol, I. A., Nemukhin, A. V., Salnikow, K., Cachau, R. E., Abashkin, Y. G., Kasprzak, K. S., & Burt, S. K. (2006). Quantum chemical modeling of reaction mechanism for 2-oxoglutarate dependent enzymes: effect of substitution of iron by nickel and cobalt. The Journal of Physical Chemistry A, 110(12), 4223-4228.

Chapter 2: Methods and Procedures

2.1 The development and trends of Quantum Chemical Methodology in Enzymatic Enantio-selectivity Reaction

In my PhD study, I have investigated the enantio-, regio- and chemoselectivity of various enzymes and specifically focused on the details of individual reaction mechanisms. The energetic differences between pathways and reaction channels are sometimes minor (within a few kcal mol-1), my research will require high accuracy and reproducibility. Therefore, as discussed in this methodology chapter, I have calibrated and validated my models against experimental data.

Quantum chemical approaches are the most common tools for computational chemists, and especially density functional theory (DFT) methods are very powerful and popular. [1] More specifically, DFT calculations can give insight into enzymatic properties and reactivities, for instance, through the prediction of free energies of substrate activation and product distributions. Some of these values can be calculated with reasonable accuracy as compared with experimental work (within several kcal mol^{-1}). [2,3] Understanding the reaction mechanisms of enzymes plays a paramount role in rational drug design, where the transition state structure or binding conformation of a substrate can me mimicked. This approach has been used to design and develop novel inhibitors (i.e. drug, substrate). [4,5] Furthermore, a large number of metalloenzymes has been studied using computational techniques and characterised in the past decade, and insight has been gained on their redox mechanisms. [7]

There are two primary classes of computational procedures, namely the ab initio and semi-empirical methods. Thus, the ab initio methods consider an atomistic system without preceding knowledge of the real system, while the semi-empirical methods take experimental parameters such as bonding parameters into consideration in their calculations. [3] Since ab initio methods calculate the system from scratch, they inevitably encounter higher computational costs and are much slower than semi-empirical methods. For instance, semi-empirical methods have been useful for the calculation of individual pK_a values and redox potentials of redox-activate enzymes. [7] Due to their cheap computational cost, until the early 1990s, studies on oxygen-activating enzymes were primary performed using semi-empirical approaches. [8-13] However; in the past few decades, the advancement and development of novel computational techniques as well as increased computational hardware has made it possible to address a diverse set of enzyme

catalysis and transition-metal-complex problems with ab initio or DFT methods. [1,7,14-26] All methods have their advantages and disadvantages; and depending on the choice of the enzyme and the problem that needs to be solved, a method is chosen. Currently, there is no universal protocol or use of techniques. This chapter of methodology will focus on the quantum chemical cluster approach and the computational methods that I have used.

The quantum cluster approach is the most fundamental and the most prevalent used in the field of computational enzymology. Generally, computational enzymology uses either the quantum chemical cluster approach or quantum mechanics/molecular mechanics (QM/MM) methods. In the Quantum chemical cluster approach or all-QM methods, a part of the enzyme structure is selected, usually including the first and second coordination sphere of the metal centre that is close to the active site. This model is then calculated with DFT or ab initio methods and the remainder of the enzyme is ignored or included through a continuum polarized solvent model as is done, for instance, in a homogeneous solvent. This enables one to investigate both the properties and mechanistic reactivity of enzyme. [27, 28, 30]

QM/MM method is another mainstream method that has seen increasing popularity in recent years since the development of its concept was first introduced by A. Warshel and M. Levitt in 1976. [111] The idea combines the merits of QM applied on the active site, where the reaction happens, and MM, where the long-range interaction happens, used on the area outside the selection of the QM part and, in combination, mainly applied on the enzymatic reactions. Most importantly, its innovative developments are opening a new era for an alternative approach to just the Quantum cluster model approach. Both authors, along with Martin Karplus, won the Nobel Prize for this theory in 2013 and not only can apply this theory on the enzymatic reaction mechanism have shown that it can be applied on multiscale models in complex system.

In recent years, QM/MM applications have been prevalent in biotechnology, such as drug development and enzymatic reaction. [112-116] Yet, several limitations and challenges need to be overcome, such as the selection of the QM region, which determines the accuracy of QM/MM results, the condition of the MM scheme, and the choice of non-bonded QM/MM parameters. [115] Two Main streams are the large-scale QM, and QM/MM approaches to anticipate the behaviour of the enzyme properties and protein folding in general. With recent development, these two main approaches could be challenged by recent development of machine learning, with the help of systematic and statistical data science models, unveils the pattern of enzymatic action. [112]

Although there is declining popularity in using quantum cluster models, however, there has been a review that suggests that in 2022 this year recent research review from an opinion of Dr Fahami Himo and Dr Samuel De Visser. [109] A more detailed explanations is provided by De Visser group for the use of cluster models and how they are considered more accurate than QM/MM. [109] Another recent paper just published in 2022 also compares the QM/MM method with QM (DFT) cluster model applied in a TauD system. Furthermore, Dr Hafiz Saqib Ali utilising these two techniques in comparison to an experimental constant, the DFT model providing the suitable rate constant, and the findings are encouraging. It demonstrates that the cluster model is generally very beneficial. [110]

Overall, all quantum mechanical (QM) cluster approach is based on solving the Schrödinger equation. [3] A starting point of the QM cluster model is taken from X-ray crystallography from the protein databank. The structural foundation of the QM cluster models are therefore based on enzymatic structures, whereby residues and protein is truncated to the active site where the enzymatic reaction happens (bond-forming and bond-breaking). [21,31,32] The alternative approach is QM/MM, where a complete enzyme with solvent layer is partitioned into two regions, namely the active site (QM region) where the reaction takes place and the MM region, which is the rest of the protein and solvent that only contributes through steric and long-range electrostatic and dispersion effects.

In 1997, one of the first quantum chemical cluster models was reported by Siegbahn and Crabtree, who treated it with DFT methods using the hybrid functional density functional B3LYP with a 6-311+G basis set. Their calculations focused on the active site and reaction mechanism of the nonheme iron monooxygenase, namely methane monooxygenase (MMO) which hydroxylates methane. [27] Subsequently, they used the same quantum chemical approaches to investigate heme peroxides but expanded the active site model to 67 atoms. [30] Around the year 2000, the expected size of active site cluster models of metalloenzymes typically comprised no more than fifty atoms to calculate the enzymatic reaction mechanism. [28,29,30] Thanks to the advance and development of both computer power and quantum chemical cluster approach, the size of the active site cluster model has become bigger and bigger; and nowadays cluster models of 200 – 300 atoms are regarded as regular. [2]

The active site cluster models ranged in size from around 90 to 300 atoms. [33,34,35] Apart from my projects, these quantum chemical methods demonstrate their success in past years, in which the size of the model can reach easily up to 300 atoms. [36-45] However, when the size of the cluster model reaches more than 300 atoms, this might cause artificial movement or have issues with multiple local minima, which contribute to incorrect energy profiles and eventually lead to inaccurate mechanistic results. [2]

The solution for this issue is by fixing some atoms nearby the interception points (usually carbon or hydrogen atoms) to retain the structure close to the original crystal structure coordinates. [7]

Depending upon the model and different situations, the approach needs to be changed. For instance, we will first calculate enzymatic model complex without any constraint and examine the results; if the model of amino acid residues separates from the active site, then we will fix some points (back bone) of amino acid residues to make it more rigid. However, it will also increase the computational cost and might present some errors after the optimisation process. In this scenario, I will have to learn the trade-off between the fix and non-fix (keeping it flexible) during this try-and-error process.

It has been noted that with an increase of the size of the cluster model reaching 200 atoms or more that the impact of the dielectric constant will decline significantly and hence by that means, the selection of dielectric constant will not affect the results (i.e. solvent effect saturate). [47-49] However, for the QM approach, increasing the size of the cluster model will soar the computational cost (demanding calculation time, amount of CPU and memory usage) and lead to slow convergence. Moreover, the enzyme system will most commonly modify a dielectric constant of $\varepsilon = 4$, treating the enzyme medium (environment) as a homogeneous polarizable solvent as mimetic the natural enzyme environment.

As for the solvent that the QM cluster model used, there are two primary classes of solvent techniques; the first is implicit solvent models, and the second is explicit solvent models. Implicit solvent models use mathematical equations and without the actual solvent present. Thus, the actual molecular system inside the cluster model is perturbed by a continuous medium with desired dielectric constant and interfacial properties, such the Onsager model, the universal solvent model (SMD), the polarizable continuum model (PCM), and conductor-type polarizable continuum model (CPCM). The alternative to the implicitsolvent model uses explicit solvent molecules inside the cluster model. These, of course, make the model larger and consequently will increase the computational cost of the calculation. Moreover, these solvent molecules often have weak interactions with substrate and oxidant causing converging problems and many low-energy conformations to the system. [56,57,58] In my work I have mostly used the CPCM solvent model; although some cluster models included some explicit solvent (water) molecules as well. Generally, implicit solvent models are less computationally expensive than an explicit solvent model. [50-55]

In our project, we use hybrid density functional theory (DFT) functional B3LYP as our primary usage since it provides an excellent balance between accuracy and speed and has been widely used over the

years. [59,60] The B3LYP functional; however, has a weakness as it does not include dispersion forces and uses empirical data as the correction. The dispersion corrections can be implemented using the methods of Grimme. [61,62] In this account, using B3LYP-D (adding the dispersion correction part) will be shown higher accuracy in the protein enzyme environment, where the surround by van der Waals interactions and steric repulsion occur. [1] More specifically, this approach, called DFT-D, has been demonstrated to generate a more accurate energy profile in homogeneous catalysis research and has rapidly evolved into a benchmark protocol in the quantum cluster model method. [63,64,65]

For redox enzymes containing iron, and particular open-shell spin states, the exact exchange of the hybrid B3LYP functional often needs to be below 15% (compared with standard 20%) as shown by studies that compared calculations to experimental work. [66] Despite the prevalent usage of the DFT calculations, there are arguments in the literature against the use of B3LYP for certain chemical problems as sometimes large errors are obtained. [66,59] These cases where errors occur are usually resolved using dispersion corrections, especially when treating transition metal containing the first-row transition metal complexes. [67] Moreover, for the field of homogenous catalysis, these applications of density functional theory (DFT) technique are also helpful in advancing its experimental procedure and unveiling catalytic mechanism not only in the enzymology domain. [68] In summary of [published papers] in 2022, B3LYP is the most promising technique for a range of iron-containing enzyme systems as the calculated results are accurate in comparison with the experimental data. [110,117-124]

Geometry optimisations of enzyme-substrate complexes are usually done with a small basis set as this tends to give reasonable quality geometries and vibrations. Thereafter, a single point energy calculation is done with an enlarged basis set with additional polarization and diffuse basis functions to improve the accuracy of the energy. In addition, this protocol significantly reduces the computing.[69] The other mainstream approach in computational enzymology is the use of the quantum mechanics/ molecular mechanics (QM/MM) method, whereby a well-chosen part of the active site is calculated quantum mechanically while the rest of the protein and solvent layer is described with the molecular mechanics method, i.e. using classical force fields. [70] In 2000, the first QM/MM result was reported on a nonheme iron system, namely for the enzymatic reaction of galactose oxidase (a redox-active enzyme). [71]

It is worth to note that the procedure is on how to prepare the cluster models and what should be included it. In particular, we should include atoms and groups depending on the judgement of first and second coordination effect of amino acid residues. These second-coordination sphere effects have the potential to stabilise or destabilise the substrate-oxidant interactions and ultimately impact the results in terms of

the energy profile or even the electronic properties. In addition, we will need to choose a quantum chemical software package for the calculations such as Gaussian-09, Turbomole, NWChem or Orca or related to model the quantum chemical cluster approach through density functional theory calculations. [86,87] The next step is taking a set of X-ray crystal structure coordinates and download those from the Protein data bank. [89] We then use our chemical intuition to judge which amino acid residues need to be included in the cluster model, thereby seeking a trade-off between the computational cost and the accuracy. Ideally the enzyme active site will have the substrate bound inside the active site of the protein; however, if this is not the case it will need to be docked in using a Docking software package such as AutoDock or SwissDock. If necessary, manual modifications to the substrate binding orientations could be made to suit our own needs.

Moreover, in order to control the total number of atoms in our cluster model we will remove or truncate amino acid residues by cutting them off from the point where they connect to the peptide chain, i.e. at the α-carbon atom of the backbone. The truncation of residues uses the Link-atom approach, where a C-C bond is replaced by C-H to keep the valences of all atoms. In the case of the UndA project we fixed a couple of atoms in the model, i.e. some α-carbon atoms, so that the model chains do not deviate to much from the real system. Once a decision is made on what groups stay in the model, we count the total charge and multiplicity of the system. Thereafter, we choose the basis sets and computational method and run a geometry optimisation. To gain more accurate energies, we typically follow a geometry optimization with a single point energy calculation with a larger basis set on all atoms.

2.2 the Schrödinger equation, Born-Oppenheimer approximation

When we are talking about quantum mechanics, we will start with the Heisenberg uncertainty principle, which is the step that converts classical mechanics into quantum mechanics for the description of the atomic properties on a quantum scale. [72] In 1923, the French physicist Louis de Broglie proposed the concept of electrons behaving like a particles and waves, which was confirmed in 1961 through the double-slit experiment. More specifically, electrons behave like a wave, meaning that it possesses a wavelength, but they also behave like particles, and contain momentum and mass (duality properties). [3] Hence the particles in these systems are defined as wave functions in the Schrödinger equation, Eq 2.1. The fundamental concept of computational methods of quantum mechanics is solving the Schrödinger equation of find its wave function so as to gain all properties of any molecular system. [3,73] This time-independent Schrödinger equation depicts the relationship between the wavefunction and the total energy of the system. [74] Eq 2.1 describes an electron's movement in a molecular system. [75]

$$\hat{H}\Psi = E\Psi \qquad \text{(Eq 2.1)}$$

Equation 2.1 demonstrates the most well known the Schrödinger equation.[76] In this equation, H is noted as the Hamilton operator that describes all perturbations in the system, which are split into two types of contributions, namely the kinetic energy part and the potential energy. The potential energy is a summation of the interactions of the individual electrons and nuclei in any molecular system. [3,77,78] More specifically, the potential energy part of the Hamilton operator (Eq 2.2) contains three contributions in terms of Coulombic interactions depicted by nucleus-nucleus repulsion $V_{n\text{-}n}$, electron-electron repulsion $V_{e\text{-}e}$, electron-nucleus attractions $V_{n\text{-}e}$, in addition to the kinetic energy of electrons and nuclei (T_e and T_n).[79,80,81]

$$\hat{H} = \hat{T}e + \hat{V}n - e + \hat{V}e - e + \hat{V}n - n \qquad \text{(Eq 2.2)}$$

Within the Born-Oppenheimer approximation it is assumed that since the electron is about 1800 times lighter than the mass of a proton, the nuclei can be considered as rigid within the field of the moving electrons. [3] Hence, the nuclei are treated as stationary relative to electrons and the nuclar kinetic energy is set to zero. The Born-Oppenheimer approximation separates the Schrödinger equation into nuclear and electronic contributions with the nuclear contributions given in Eq 2.3. Therefore, the Born-Oppenheimer

approximation reduces the complexity of the Schrödinger equation, where the electronic energy E is calculated for a rigid nuclear motif.

$$\widehat{H}nuc(R)\Psi nuc(R) = \hat{E}nuc(R)\Psi nuc(R) \qquad \text{(Eq 2.3)}$$

R is the Position of the nuclei

2.3 Density Functional Theory (DFT) method

From the year 1960s, density functional theory (DFT) methods have emerged with the aim to serve as an alternative approach to wave function methods. The DFT methods determine all molecular properties by using the electron density rather than a wave function. [82,83,84] These approaches calculate the intrinsic properties of the multiple-electron systems (molecules) using a series of individual functionals for electron-electron, electron-nucleus, and nucleus-nucleus interactions. Generally, DFT follows the Hohenberg-Kohn theorems and the Kohn-Sham formalism. [83,84]

Density functional theory methods are a great tool to resolve "Many-Body Problems", i.e. this implies that they consider chemical systems with more than just one atom. These DFT methods were found to significantly reduce the computational cost as compared to wave function techniques and consequently have become very popular. The total energy of a system calculated with DFT methods consists of two terms: one is related to approximations like the Born–Oppenheimer approximation, while the other is a full electron density expression. The fundamental aim of these computational studies is to investigate the molecular electronic or nuclear structure (i.e., ground state of any complex molecules, condensed phase). [83] Moreover, DFT uses a series of functionals and these functionals consist of spatially dependent electron density functionals. Applications of DFT have demonstrated high agreement of the calculations with experimental work and particularly free energies of activation are often within a few kcal $mole^{-1}$. [3,86]

A comparison between wave function equations and the density functional energy is the inclusion of additional terms such as the exchange-correlation function (E_{xc}) in hybrid density functionals. These hybrid density functional methods originate from the original DFT formalism but contains correction factors obtained from a test set that compared calculations to experimental data. The advantage of the hybrid functional methods is these approximations in terms of the exchange-correlation component improve the accuracy of electronic energy dramatically. [92] Among all the hybrid density functional methods, "B3LYP", the Becke 3-parameter exchange with Lee–Yang–Par correlation is the most well-known, widely used and applied in computational chemistry and was the functional of choice in most of my projects described in this thesis. [59,60,93,94] Another popular hybrid functional that we sometimes use is the OPBE functional. [95,96]

2.4 Basis sets

In quantum mechanics, the wave function is usually described with a basis set, which typically is described as a linear combination of atomic orbitals and thereby enable a solution of the Schrödinger equation. [97] Depending on the system we want to calculate; the choice of a suitable basis set depends on how good bonding and antibonding molecular orbitals are described. A linear combination of atomic orbital (LCAO) can generate and describe the foundation of the molecular orbital. Similarly, in density functional theory methods the calculation of the electron density and the total energy is highly dependent on the quality of the basis set on each atom. [78,85,98]

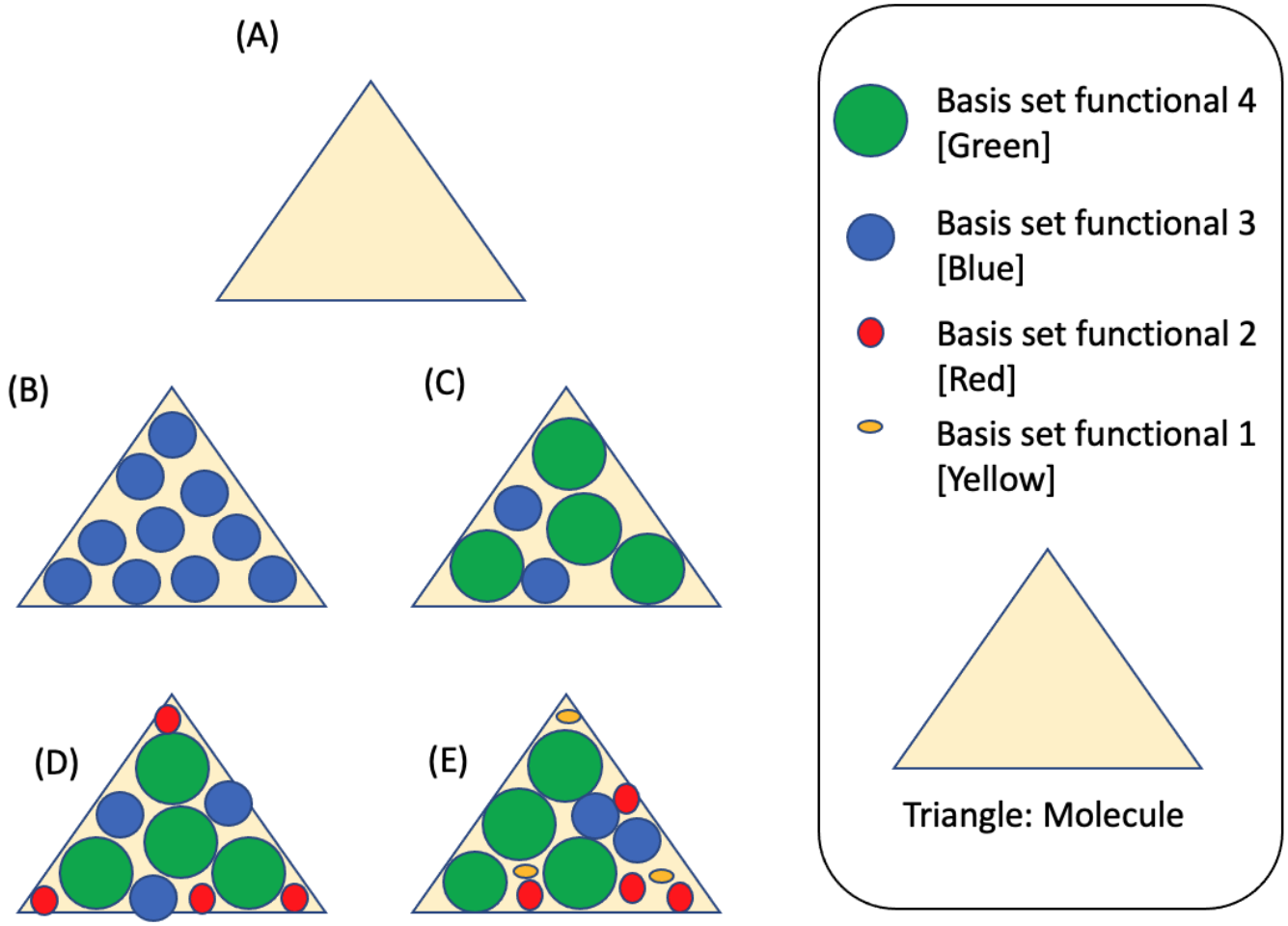


Figure 2.1 Examples of different basis sets in simple molecular or more complex system.

Figure 2.1 The demonstration of how basis sets depict atomic and molecular orbitals. Green (basis set 4), Blue (basis set 3), red (basis set 2), yellow (basis set 1). (A) Triangles indicate the any kind of molecular

system. (B) Triangle fit in the blue circle indicate that using one kind of basis set to depict the system. (C, D, E) is indicate using different basis set to depict the one molecular system.

Imagine a molecule with the intricate geometric design shown schematically in ***Figure 2.1***, where each circle represents a different basis function. Naturally, it becomes challenging to capture the entire triangle's surface area with a one-size circle, or a single basis set. However, the entire surface area may be depicted using several circles of various sizes and forms. Simply put, system can be broken into a series of different sized basis functions that can arranged to better fit the triangle. Therefore, how we interact with a system or molecular model depends on it.

The same is true for molecules, where a variety of basis functions and the total number of basis functions will determine the accuracy of its overall description and hence its energy. Several studies have indicated that well-chosen basis sets achieve good accuracy but incorporating more basis functions ultimately will end up with high computational cost. [99,100,101] Nonetheless, an upper boundary for the basis set can be reached depending on the choice of the DFT method. In particular, often the energy converges to a constant value with further increase of the basis set. [78,85,98] The two most popular basis sets are the Slater-Type Orbital (STO) and Gaussian-Type Orbital (GTO), which are two primary types of basis functions described in all computational chemistry software.

In addition, for transition metal complexes there are so-called Effective Core Potential (ECP) basis sets that can be useful as well. Generally, these ECP type basis sets replace core orbitals, e.g. 1s, 2s and 2p for iron, by a point charge in the nucleus and thereby reduce the total number of basis functions and hence the computational cost. However, these ECP methods enable the use of relativistic corrections that are beneficial to large atoms and thereby give good quality results.

Furthermore, for the inclusion of non-bonding interactions (i.e. van der Waals interactions between individual molecules and groups) as well as charged groups such as anionic contributions, it is particularly important to add the diffuse basis functions to improve the quality of the results. For polarization terms, symbolised with star (*) in the basis set notation (such as 6-31G*), additional p-functions are included for light atoms such as hydrogen and d-functions for heavy atoms such as carbon and oxygen.

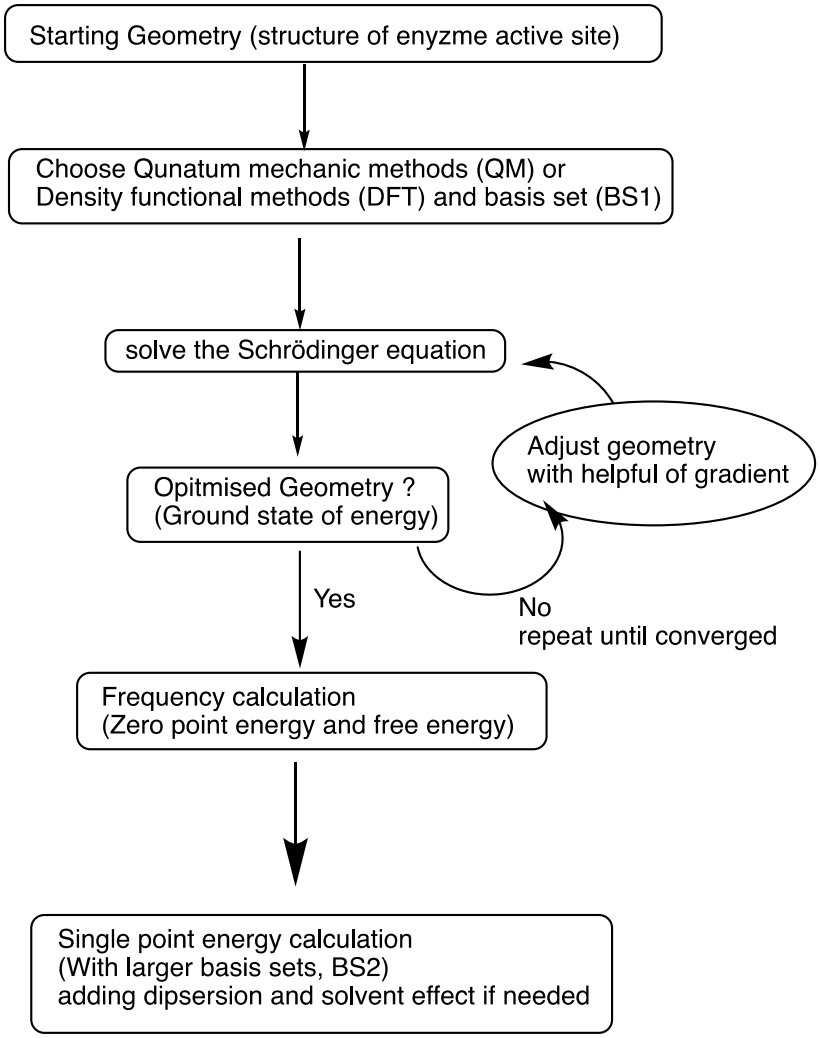
These so called "Pople-style basis sets", such as 6-31+G*, indicate that a GTO type basis set is used with an additional set of polarisation and diffuse functions on all non-hydrogen atoms (heavy atoms). The value of adding the polarization functions is that it can greatly diminish the total vibrational energy and

improve the results of quantum chemical cluster model calculations in terms of the geometry optimisation and vibrational frequency calculations.

Since adding diffusion and polarization functions will increase the computational cost, it is sometimes good to test whether it is important in a geometry optimization or not and compare the results of 6-31G and 6-31+G* geometry optimizations. If the two structures are similar then a low basis set can be used for the initial geometry optimisation and frequency, while the energy is corrected through a single point energy calculation with 6-31+G* to save computational time and resources. In most of my projects, therefore, I used the 6-31G basis set of geometry optimizations, frequencies, geometry scans and intrinsic reaction coordinate scans, while the energies are taken from a higher basis set calculation.

2.5 Geometries, frequencies, and free energies: relative and absolute energy optimisation, Zero-Point Energy (Free energy correction)

Characterization of the reactant complex (active site plus substrate), transition state, intermediate, and product is essential to unveil any mechanistic enzymology. The geometry optimisation methods rearrange all the atoms and reach a total inter-atomic force near the lowest and the specific coordinate is calculated as static point on the potential energy surface (PES). [90,91]



Scheme 2.1 General flow diagram of the procedure for the finding of a local minimum.[73]

Scheme 2.1 shows a flow diagram of the general procedures on locating a local minimum in quantum mechanics. As described above, we create a QM cluster model first and chose a computational method and basis set. After that these input variables are used by the QM software package to solve the Schrödinger equation and give an energy associated with the starting structure. Afterward, the software

package, GAUSSIAN 09, changes the coordinates using gradients of the change in energy with respect to distances, angles and dihedrals, i.e. the degrees of freedom. The software repeats the process of changing the coordinates and calculating a QM energy until a cut-off for the energy change and its derivatives is reached, which is called geometry convergence. At that stage, we have the minimum energy associated with the ground state for a local minimum. However, this only represents an electronic energy, which we will need to convert to a free energy at 298 K temperature and by including entropic corrections. The first contribution comes from the zero-point energy (ZPE), which assigns the 0 K enthalpy of a system under harmonic oscillator approximation conditions **Figure 2.2** shows a potential energy surface of a system with two local minima (E_{S1} and E_{S2}) on an electronic energy landscape. Using the harmonic oscillator approach vibrational frequencies are calculated during a frequency calculation that gives us estimates of the vibrational levels v_0, v_1, v_2, v_3, v_4 etc. At 0 K, we expect the system in the lowest vibrational level, hence the 0 K enthalpy is equal to E_{el} + ZPE. Apart from the ZPE there is more information that can be obtained from a frequency calculation, namely the number of real frequencies in the structure. Thus, an imaginary frequency implies the existence of a local maximum on the potential energy surface and consequently represents a transitions state along a reaction mechanism. These imaginary frequencies can have relatively large magnitudes of well over i1000 cm^{-1} particularly when a hydrogen atom abstraction is considered. Thus, hydrogen atom abstraction barriers are sharp and narrow and lead to large imaginary frequency values due to strong quantum tunnelling effects.

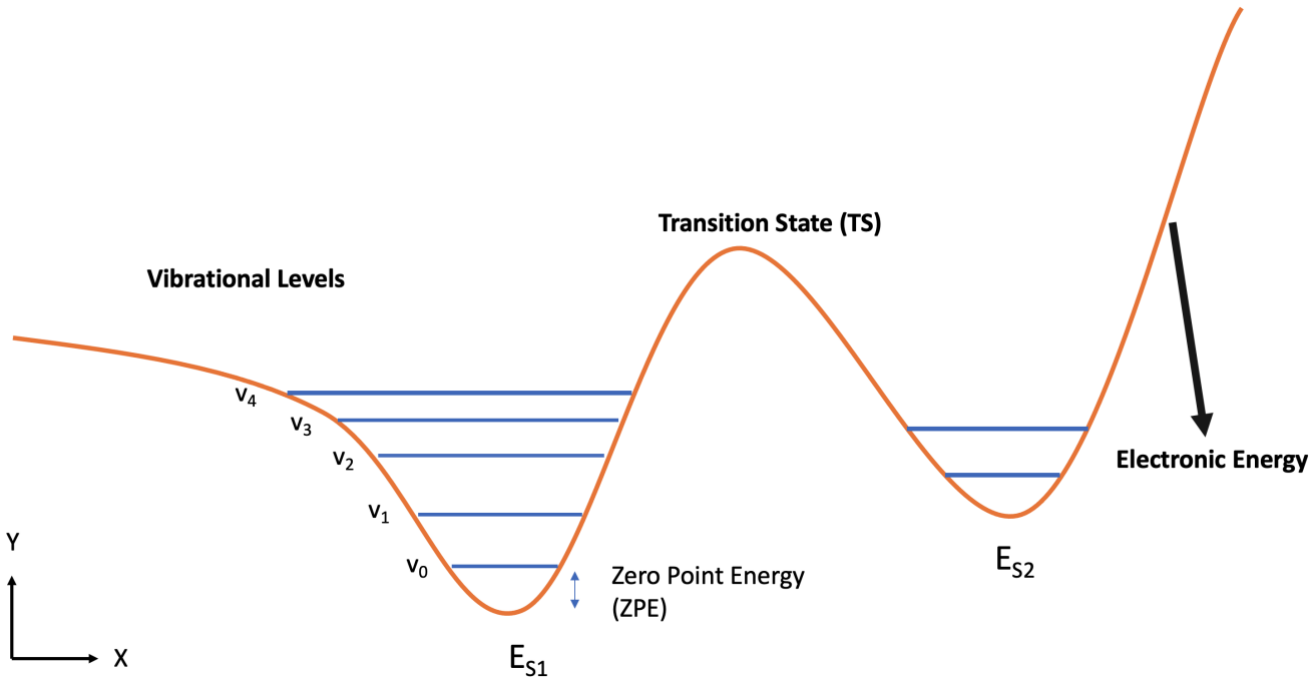


Figure 2.2 Electronic energy curve of a molecular system with two local minima (S1 and S2) each with calculated vibrational levels (v_0, v_1, v_2, v_3, v_4). Figure adapted from Ref. [73]

2.6 Methodology References

[1] Sheng, X., Kazemi, M., Planas, F., & Himo, F. (2020). Modeling enzymatic enantioselectivity using quantum chemical methodology. ACS Catalysis, 10(11), 6430-6449.
[2] Himo, F. (2017). Recent trends in quantum chemical modeling of enzymatic reactions. Journal of the American Chemical Society, 139(20), 6780-6786.
[3] De Visser, S., de Visser, S. P., Alleman, R. K., & Scrutton, N. S. (2009). Introduction to quantum behavior-a primer.: Chapter 2. In Quantum tunnelling in enzyme-catalysed reactions. (pp. 18-35). Royal Society of Chemistry.
[4] Schramm, V. L. (2013). Transition states, analogues, and drug development. ACS chemical biology, 8(1), 71-81.
[5] Schramm, V. L. (2018). Enzymatic transition states and drug design. Chemical reviews, 118(22), 11194-11258.
[6] Dybala-Defratyka, A., Paneth, P., & Truhlar, D. G. (2009). Quantum catalysis in enzymes. Quantum Tunnelling in Enzyme-Catalysed Reactions, 36-78.
[7] Blomberg, M. R. (2014). A.; Borowski, T.; Himo, F.; Liao, R.-Z.; Siegbahn, PEM. Chem. Rev, 114, 3601.
[8] Krauss, M., & Garmer, D. R. (1991). Active site ionicity and the mechanism of carbonic anhydrase. Journal of the American Chemical Society, 113(17), 6426-6435.
[9] Daggett, V., Schroeder, S., & Kollman, P. (1991). Catalytic pathway of serine proteases: classical and quantum mechanical calculations. Journal of the American Chemical Society, 113(23), 8926-8935.
[10] Turi, L., & Náray-Szabó, G. (1992). Computational studies on aspartic proteases. I. Active-site protonation and hydration in the substrate-free crystalline state. International journal of quantum chemistry, 42(5), 1537-1551.
[11] Zheng, Y. J., & Merz Jr, K. M. (1992). Mechanism of the human carbonic anhydrase II-catalyzed hydration of carbon dioxide. Journal of the American Chemical Society, 114(26), 10498-10507.
[12] Alex, A., & Clark, T. (1992). MO-Studies of enzyme reaction mechanisms. I. Model molecular orbital study of the cleavage of peptides by carboxypeptidase A. Journal of computational chemistry, 13(6), 704-717.
[13] Wilkie, J., & Williams, I. H. (1992). Transition-state structural variation in a model for carbonyl reduction by lactate dehydrogenase: computational validation of empirical predictions based upon Albery-More O'Ferrall-Jencks diagrams. Journal of the American Chemical Society, 114(13), 5423-5425.
[14] Lonsdale, R., Ranaghan, K. E., & Mulholland, A. J. (2010). Computational enzymology. Chemical communications, 46(14), 2354-2372.
[15] Sousa, S. F., Fernandes, P. A., & Ramos, M. J. (2012). Computational enzymatic catalysis–clarifying enzymatic mechanisms with the help of computers. Physical Chemistry Chemical Physics, 14(36), 12431-12441.
[16] Romero-Rivera, A., Garcia-Borràs, M., & Osuna, S. (2017). Computational tools for the evaluation of laboratory-engineered biocatalysts. Chemical Communications, 53(2), 284-297.
[17] Bottaro, S., & Lindorff-Larsen, K. (2018). Biophysical experiments and biomolecular simulations: A perfect match?. Science, 361(6400), 355-360.
[18] Ahmadi, S., Barrios Herrera, L., Chehelamirani, M., Hostaš, J., Jalife, S., & Salahub, D. R. (2018). Multiscale modeling of enzymes: QM-cluster, QM/MM, and QM/MM/MD: a tutorial review. International Journal of Quantum Chemistry, 118(9), e25558.
[19] Blomberg, M. R. (2015). How quantum chemistry can solve fundamental problems in bioenergetics. International Journal of Quantum Chemistry, 115(18), 1197-1201.
[20] Borowski, T., & Broclawik, E. (2014). Bioinorganic Reaction Mechanisms–Quantum Chemistry Approach. In Computational Methods to Study the Structure and Dynamics of Biomolecules and Biomolecular Processes (pp. 783-808). Springer, Berlin, Heidelberg.

[21] Siegbahn, P. E., & Himo, F. (2011). The quantum chemical cluster approach for modeling enzyme reactions. Wiley Interdisciplinary Reviews: Computational Molecular Science, 1(3), 323-336.
[22] Mander, L., & Liu, H. W. B. (2010). Comprehensive Natural Products Chemistry II: Chemistry and Biology. Elsevier.
[23] Siegbahn, P. E., & Blomberg, M. R. (2010). Quantum chemical studies of proton-coupled electron transfer in metalloenzymes. Chemical reviews, 110(12), 7040-7061.
[24] Siegbahn, P. E., & Himo, F. (2009). Recent developments of the quantum chemical cluster approach for modeling enzyme reactions. JBIC Journal of Biological Inorganic Chemistry, 14(5), 643-651.
[25] Bauschlicher Jr, C. W., Ricca, A., Partridge, H., & Langhoff, S. R. (1997). Chemistry by density functional theory. In Recent Advances In Density Functional Methods: (Part II) (pp. 165-227).
[26] Siegbahn, P. E. M. (2009). Electronic structure calculations for molecules containing transition metals. New Methods in Computational Quantum Mechanics, 93, 333.
[27] Siegbahn, P. E., & Crabtree, R. H. (1997). Mechanism of C− H activation by diiron methane monooxygenases: quantum chemical studies. Journal of the American Chemical Society, 119(13), 3103-3113.
[28] Filatov, M., Harris, N., & Shaik, S. (1999). On the "rebound" mechanism of alkane hydroxylation by cytochrome P450: Electronic structure of the intermediate and the electron transfer character in the rebound step. Angewandte Chemie International Edition, 38(23), 3510-3512.
[29] Basch, H., Mogi, K., Musaev, D. G., & Morokuma, K. (1999). Mechanism of the methane→ methanol conversion reaction catalyzed by methane monooxygenase: a density functional study. Journal of the American Chemical Society, 121(31), 7249-7256.
[30] Wirstam, M., Blomberg, M. R., & Siegbahn, P. E. (1999). Reaction mechanism of compound I formation in heme peroxidases: a density functional theory study. Journal of the American Chemical Society, 121(43), 10178-10185.
[31] Leopoldini, M., Marino, T., Michelini, M. D. C., Rivalta, I., Russo, N., Sicilia, E., & Toscano, M. (2007). The role of quantum chemistry in the elucidation of the elementary mechanisms of catalytic processes: from atoms, to surfaces, to enzymes. Theoretical Chemistry Accounts, 117(5), 765-779.
[32] Ramos, M. J., & Fernandes, P. A. (2008). Computational enzymatic catalysis. Accounts of chemical research, 41(6), 689-698.
[33] Lin, Y. T., Stańczak, A., Manchev, Y., Straganz, G. D., & de Visser, S. P. (2020). Can a mononuclear iron (III)-superoxo active site catalyze the decarboxylation of dodecanoic acid in UndA to produce biofuels?. Chemistry–A European Journal, 26(10), 2233-2242.
[34] Lin, Y. T., Ali, H. S., & de Visser, S. P. (2021). Electrostatic Perturbations from the Protein Affect C− H Bond Strengths of the Substrate and Enable Negative Catalysis in the TmpA Biosynthesis Enzyme. *Chemistry–A European Journal*, *27*(34), 8851-8864.
[35] Lin, Y. T., Ali, H. S., & de Visser, S. P. (2022). Biodegradation of herbicides by a plant nonheme iron dioxygenase: mechanism and selectivity of substrate analogues. Chemistry–A European Journal, 28(7), e202103982.
[36] Raz, K., Driller, R., Brück, T., Loll, B., & Major, D. T. (2020). Understanding the role of active site residues in CotB2 catalysis using a cluster model. Beilstein journal of organic chemistry, 16(1), 50-59.
[37] Ghafoor, S., Mansha, A., & de Visser, S. P. (2019). Selective hydrogen atom abstraction from dihydroflavonol by a nonheme iron center is the key step in the enzymatic flavonol synthesis and avoids byproducts. Journal of the American Chemical Society, 141(51), 20278-20292.
[38] Yu, M. J., & Chen, S. L. (2019). From Alkane to Alkene: The Inert Aliphatic C–H Bond Activation Presented by Binuclear Iron Stearoyl-CoA Desaturase with a Long di-Fe Distance of 6 Å. ACS Catalysis, 9(5), 4345-4359.
[39] Prejanò, M., Medina, F. E., Fernandes, P. A., Russo, N., Ramos, M. J., & Marino, T. (2019). The Catalytic Mechanism of Human Transketolase. ChemPhysChem, 20(21), 2881-2886.

[40] Mattioli, E. J., Bottoni, A., & Calvaresi, M. (2019). DNAzymes at work: A DFT computational investigation on the mechanism of 9DB1. Journal of Chemical Information and Modeling, 59(4), 1547-1553.
[41] Ji, L., Ji, S., Wang, C., & Kepp, K. P. (2018). Molecular mechanism of alternative P450-catalyzed metabolism of environmental phenolic endocrine-disrupting chemicals. Environmental science & technology, 52(7), 4422-4431.
[42] Blomberg, M. R. (2017). Can reduction of NO to N2O in cytochrome c dependent nitric Oxide reductase proceed through a trans-mechanism?. Biochemistry, 56(1), 120-131.
[43] Siegbahn, P. E. (2016). Model calculations suggest that the central carbon in the FeMo-cofactor of nitrogenase becomes protonated in the process of nitrogen fixation. Journal of the American Chemical Society, 138(33), 10485-10495.
[44] Lan, C. L., & Chen, S. L. (2016). The decarboxylation of α, β-unsaturated acid catalyzed by prenylated FMN-dependent ferulic acid decarboxylase and the enzyme inhibition. The Journal of Organic Chemistry, 81(19), 9289-9295.
[45] Cassimjee, K. E., Manta, B., & Himo, F. (2015). A quantum chemical study of the ω-transaminase reaction mechanism. Organic & Biomolecular Chemistry, 13(31), 8453-8464.
[46] Sevastik, R., & Himo, F. (2007). Quantum chemical modeling of enzymatic reactions: the case of 4-oxalocrotonate tautomerase. Bioorganic chemistry, 35(6), 444-457.
[47] Hopmann, K. H., & Himo, F. (2008). Quantum chemical modeling of the dehalogenation reaction of haloalcohol dehalogenase. Journal of Chemical Theory and Computation, 4(7), 1129-1137.
[48] Georgieva, P., & Himo, F. (2010). Quantum chemical modeling of enzymatic reactions: The case of histone lysine methyltransferase. Journal of computational chemistry, 31(8), 1707-1714.
[49] Liao, R. Z., Yu, J. G., & Himo, F. (2011). Quantum chemical modeling of enzymatic reactions: the case of decarboxylation. Journal of Chemical Theory and Computation, 7(5), 1494-1501.
[50] Ren, P., Chun, J., Thomas, D. G., Schnieders, M. J., Marucho, M., Zhang, J., & Baker, N. A. (2012). Biomolecular electrostatics and solvation: a computational perspective. Quarterly reviews of biophysics, 45(4), 427-491.
[51] Tomasi, J., Mennucci, B., & Cammi, R. (2005). Quantum mechanical continuum solvation models. Chemical reviews, 105(8), 2999-3094.
[52] Onsager, L. (1936). Electric moments of molecules in liquids. Journal of the American Chemical Society, 58(8), 1486-1493.
[53] Marenich, A. V., Cramer, C. J., & Truhlar, D. G. (2009). Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. The Journal of Physical Chemistry B, 113(18), 6378-6396.
[54] Cossi, M., Scalmani, G., Rega, N., & Barone, V. (2002). New developments in the polarizable continuum model for quantum mechanical and classical calculations on molecules in solution. The Journal of Chemical Physics, 117(1), 43-54.
[55] Andzelm, J., Kölmel, C., & Klamt, A. (1995). Incorporation of solvent effects into density functional calculations of molecular energies and geometries. The Journal of chemical physics, 103(21), 9312-9320.
[56] Jensen, F. (2007). Electronic Structure Methods: Independent-Particle Models. Introduction to Computational Chemistry, 80.
[57] Cramer, C. J. (2013). Essentials of computational chemistry: theories and models. John Wiley & Sons.
[58] Jakalian, A., Jack, D. B., & Bayly, C. I. (2002). Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. Journal of computational chemistry, 23(16), 1623-1641.
[59] Becke, A. Density-functional thermochemistry. III. The role of exact exchange (1993) J. Chem. Phys, 98, 5648.
[60] Lee, C., Yang, W., & Parr, R. G. (1988). Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Physical review B, 37(2), 785.

[61] Grimme, S. (2011). Density functional theory with London dispersion corrections. Wiley Interdisciplinary Reviews: Computational Molecular Science, 1(2), 211-228.
[62] Bursch, M., Caldeweyher, E., Hansen, A., Neugebauer, H., Ehlert, S., & Grimme, S. (2018). Understanding and quantifying London dispersion effects in organometallic complexes. Accounts of Chemical Research, 52(1), 258-266.
[63] Lonsdale, R., Harvey, J. N., & Mulholland, A. J. (2010). Inclusion of dispersion effects significantly improves accuracy of calculated reaction barriers for cytochrome P450 catalyzed reactions. The Journal of Physical Chemistry Letters, 1(21), 3232-3237.
[64] Siegbahn, P. E., Blomberg, M. R., & Chen, S. L. (2010). Significant van der Waals effects in transition metal complexes. Journal of Chemical Theory and Computation, 6(7), 2040-2044.
[65] Zhang, H. M., & Chen, S. L. (2015). Include dispersion in quantum chemical modeling of enzymatic reactions: the case of isoaspartyl dipeptidase. Journal of Chemical Theory and Computation, 11(6), 2525-2535.
[66] Siegbahn, P. E., & Blomberg, M. R. (2018). A systematic DFT approach for studying mechanisms of redox active enzymes. Frontiers in Chemistry, 644.
[67] Schwabe, T., & Grimme, S. (2007). Double-hybrid density functionals with long-range dispersion corrections: higher accuracy and extended applicability. Physical Chemistry Chemical Physics, 9(26), 3397-3406.
[68] Tantillo, D. J. (2016). Faster, Catalyst! React! React! Exploiting computational chemistry for catalyst development and design. Accounts of Chemical Research, 49(6), 1079-1079.
[69] Siegbahn, P. E. (2001). Modeling aspects of mechanisms for reactions catalyzed by metalloenzymes. Journal of Computational Chemistry, 22(14), 1634-1645.
[70] Warshel, A., & Levitt, M. (1976). Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. Journal of molecular biology, 103(2), 227-249.
[71] Rothlisberger, U., Carloni, P., Doclo, K., & Parrinello, M. (2000). A comparative study of galactose oxidase and active site analogs based on QM/MM Car-Parrinello simulations. JBIC Journal of Biological Inorganic Chemistry, 5(2), 236-250.
[72] Heisenberg, W. (1985). Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. In Original Scientific Papers Wissenschaftliche Originalarbeiten (pp. 478-504). Springer, Berlin, Heidelberg.
[73] de Visser, S. P., Quesne, M. G., Martin, B., Comba, P., & Ryde, U. (2014). Computational modelling of oxygenation processes in enzymes and biomimetic model complexes. Chemical Communications, 50(3), 262-282.
[74] Landau, L. D., Landau, L. D., & Lifshitz, E. M. (1975). The Classical Theory of Fields: Volume 2 (Vol. 2). Butterworth-Heinemann.
[75] Lewars, E. G. (2016). Introduction to Quantum Mechanics in Computational Chemistry. In Computational Chemistry (pp. 101-191). Springer, Cham.
[76] Schrödinger, E. (1926). An undulatory theory of the mechanics of atoms and molecules. Physical review, 28(6), 1049.
[77] Mechanics, M. Q. (1997). PW Atkins and RS Friedman.
[78] Jensen, F. (2017). Introduction to computational chemistry. John wiley & sons.
[79] Born, M., & Oppenheimer, R. (1927). Zur quantentheorie der molekeln. Annalen der physik, 389(20), 457-484.
[80] Born, M., & Oppenheimer, R. (2000). Quantum Chemistry: Classic Scientific Papers. World Scientific 20th Century Chemistry, 8, 1-24.
[81] Born, M., & Oppenheimer, J. R. (1927). On the quantum theory of molecules. Сборник статей к мультимедийному электронному учебно-методическому комплексу по дисциплине «физика атома и атомных явлений»/отв. ред. Шундалов МБ; БГУ, Физический факультет.

[82] Dirac, P. A. M. (1929). Quantum mechanics of many-electron systems. Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character, 123(792), 714-733.
[83] Hohenberg, P., & Kohn, W. (1964). Inhomogeneous electron gas. Physical review, 136(3B), B864.
[84] Kohn, W., & Sham, L. J. (1965). Self-consistent equations including exchange and correlation effects. Physical review, 140(4A), A1133.
[85] Sholl, D., & Steckel, J. A. (2011). Density functional theory: a practical introduction. John Wiley & Sons.
[86] Orio, M., Pantazis, D. A., & Neese, F. (2009). Density functional theory. Photosynthesis research, 102(2), 443-453.
[87] Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., ... & Cioslowski, D. J. (2010). Fox Gaussian 09, Revision C. 01. *Gaussian Inc.*
[88] Neese, F. (2018). Software update: the ORCA program system, version 4.0. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *8*(1), e1327.
[89] Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., ... & Bourne, P. E. (2000). The protein data bank. *Nucleic acids research*, *28*(1), 235-242.
[90] Meyers, R. A. (2002). Encyclopedia of physical science and technology. Academic.
[91] Tantillo, D. J. (2018). Dynamic Effects on Organic Reactions.
[92] Gonis, A. (2014). The Pursuit of Fallacy in Density Functional Theory: The Quest for Exchange and Correlation, the Rigorous Treatment of Exchange in the Kohn-Sham Formalism and the Continuing Search for Correlation. World Journal of Condensed Matter Physics, 4(03), 200.
[93] Becke, A. D. (1992). Density-functional thermochemistry. I. The effect of the exchange-only gradient correction. The Journal of chemical physics, 96(3), 2155-2160.
[94] Tirado-Rives, J., & Jorgensen, W. L. (2008). Performance of B3LYP density functional methods for a large set of organic molecules. Journal of chemical theory and computation, 4(2), 297-306.
[95] Handy, N. C., & Cohen, A. J. (2001). Left-right correlation energy. Molecular Physics, 99(5), 403-412.
[96] Perdew, J. P., Burke, K., & Ernzerhof, M. (1996). Generalized gradient approximation made simple. Physical review letters, 77(18), 3865.
[97] Nagy, B., & Jensen, F. (2017). Basis sets in quantum chemistry. Reviews in Computational Chemistry, 30, 93-149.
[98] Lewars, E. G. (2016). Molecular mechanics. In Computational Chemistry (pp. 51-99). Springer, Cham.
[99] Papajak, E., & Truhlar, D. G. (2010). Efficient diffuse basis sets for density functional theory. Journal of chemical theory and computation, 6(3), 597-601.
[100] Fouda, A. E., & Besley, N. A. (2018). Assessment of basis sets for density functional theory-based calculations of core-electron spectroscopies. Theoretical Chemistry Accounts, 137(1), 1-11.
[101] Kreutzer, J., Blaha, P., & Schubert, U. (2016). Assessment of different basis sets and DFT functionals for the calculation of structural parameters, vibrational modes and ligand binding energies of Zr4O2 (carboxylate) 12 clusters. Computational and Theoretical Chemistry, 1084, 162-168.
[109] Himo, F., & de Visser, S. P. (2022). Status report on the quantum chemical cluster approach for modeling enzyme reactions. Communications Chemistry, 5(1), 1-4.
[110] Ali, H. S., & de Visser, S. P. (2022). Electrostatic Perturbations in the Substrate-Binding Pocket of Taurine/α-Ketoglutarate Dioxygenase Determine its Selectivity. Chemistry–A European Journal, 28(9), e202104167.
[111] Warshel, A., & Levitt, M. (1976). Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. Journal of molecular biology, 103(2), 227-249.

[112] Vennelakanti, V., Nazemi, A., Mehmood, R., Steeves, A. H., & Kulik, H. J. (2022). Harder, better, faster, stronger: Large-scale QM and QM/MM for predictive modeling in enzymes and proteins. Current opinion in structural biology, 72, 9-17.
[113] Mehmood, R., & Kulik, H. J. (2022). Quantum-Mechanical/Molecular-Mechanical (QM/MM) Simulations for Understanding Enzyme Dynamics. In Enzyme Engineering (pp. 227-248). Humana, New York, NY.
[114] Mehmood, R., & Kulik, H. J. (2020). Both configuration and QM region size matter: zinc stability in QM/MM models of DNA methyltransferase. Journal of Chemical Theory and Computation, 16(5), 3121-3134.
[115] Demapan, D., Kussmann, J., Ochsenfeld, C., & Cui, Q. (2022). Factors That Determine the Variation of Equilibrium and Kinetic Properties of QM/MM Enzyme Simulations: QM Region, Conformation, and Boundary Condition. Journal of Chemical Theory and Computation, 18(4), 2530-2542.
[116] Dietschreit, J. C., von der Esch, B., & Ochsenfeld, C. (2022). Exponential averaging versus umbrella sampling for computing the QM/MM free energy barrier of the initial step of the desuccinylation reaction catalyzed by sirtuin 5. Physical Chemistry Chemical Physics, 24(13), 7723-7731.
[117] KumaráBagha, U., KumaráSatpathy, J., & áde Visser, S. P. (2022). Oxidative dehalogenation of halophenols by high-valent nonheme iron (iv)-oxo intermediates. Faraday Discussions, 234, 58-69.
[118] Ali, H. S., Ghafoor, S., & de Visser, S. P. (2022). Density functional theory study into the reaction mechanism of isonitrile biosynthesis by the nonheme iron enzyme ScoE. Topics in Catalysis, 65(1), 528-543.
[119] Lin, Y. T., Ali, H. S., & de Visser, S. P. (2022). Biodegradation of Herbicides by a Plant Nonheme Iron Dioxygenase: Mechanism and Selectivity of Substrate Analogues. Chemistry–A European Journal, 28(7), e202103982.
[120] Ali, H. S., & de Visser, S. P. (2022). Electrostatic Perturbations in the Substrate-Binding Pocket of Taurine/α-Ketoglutarate Dioxygenase Determine its Selectivity. Chemistry–A European Journal, 28(9), e202104167.
[121] Himo, F., & de Visser, S. P. (2022). Status report on the quantum chemical cluster approach for modeling enzyme reactions. Communications Chemistry, 5(1), 1-4.
[122] Gérard, E. F., Yadav, V., Goldberg, D. P., & de Visser, S. P. (2022). What Drives Radical Halogenation versus Hydroxylation in Mononuclear Nonheme Iron Complexes? A Combined Experimental and Computational Study. Journal of the American Chemical Society.
[123] Matsumura, H., Faponle, A. S., Hagedoorn, P. L., Tosha, T., de Visser, S. P., & Moënne-Loccoz, P. (2022). Mechanism of substrate inhibition in cytochrome-c dependent NO reductases from denitrifying bacteria (cNORs). Journal of Inorganic Biochemistry, 231, 111781.
[124] Yeh, C. C. G., Mokkawes, T., Bradley, J. M., Le Brun, N. E., & de Visser, S. P. (2022). Second coordination sphere effects on the mechanistic pathways for dioxygen activation by a ferritin: involvement of a Tyr radical and the identification of a cation binding site. ChemBioChem, 23(13), e202200257.

Chapter 3 Can a Mononuclear Iron(III)-Superoxo Active Site Catalyze the Decarboxylation of Dodecanoic Acid in UndA to Produce Biofuels?

[1]Yen-Ting Lin,[a,e] Agnieszka Stańczak,[a,b,e] Yulian Manchev,[a] Grit D. Straganz,[c,d] and Sam P. de Visser*[a]

[a] Mr Y.-T. Lin, Miss A. Stanczak, Mr Y. Manchev, and Dr S. P. de Visser,
The Manchester Institute of Biotechnology and Department of Chemical Engineering and Analytical Science, The University of Manchester
Email: yen-ting.lin@manchester.ac.uk
131 Princess Street, Manchester, M1 7DN, United Kingdom.
Email: sam.devisser@manchester.ac.uk
[b] Miss A. Stanczak
Silesian University of Technology, Faculty of Chemistry, ks. Marcina Strzody 9, 44-100 Gliwice, Poland
[c] Dr G.D. Straganz,
Graz University of Technology, Institute of Biochemistry, Petergasse 12, A-8010 Graz, Austria
[d] Dr G.D. Straganz,
current address: bisy Petersgasse 14, A-8010 Graz, Austria
[e] YTL and AS contributed equally.

Supporting information for this article is given via a link:
https://chemistry-europe.onlinelibrary.wiley.com/doi/abs/10.1002/chem.201903783

DOI: https://doi.org/10.1002/chem.201903783

Keywords: nonheme iron • dioxygenase • decarboxylation • density functional theory • reaction mechanism

3.1 Abstract

Decarboxylation of fatty acids is an important reaction in cell metabolism, but also has potential in biotechnology for the biosynthesis of hydrocarbons as biofuels. The recently discovered nonheme iron decarboxylase UndA is involved in the biosynthesis of 1-undecene from dodecanoic acid and using X-ray crystallography was assigned to be a mononuclear iron species. However, the work was contradicted by spectroscopic studies that suggested UndA to be more likely a dinuclear iron system. To resolve this controversy we decided to pursue a computational study on the reaction mechanism of fatty acid decarboxylation by UndA using iron(III)-superoxo and diiron(IV)-dioxo models. We tested several models with different protonation states of active site residues. Overall, however, the calculations imply that mononuclear iron(III)-superoxo is a sluggish oxidant of hydrogen atom abstraction reactions in UndA

[1] **Author contribution statement:** This is the original reference where the paper is published. Yen-Ting, Lin, my contribution is DFT on Model A, Yulian Manchev conducts B and Model C; initially, Agnieszka Stańczak did the computational study on Model A, and B initiated this project but did not complete it. I took this project, re-run all the calculations and completed all the tests and expanded the calculation with the help of the Dr Samuel De Visser to come up with the idea and guide us in the right direction of the project and, in the end, spotting the insights of data and complete this story, as well as co-author with Dr Grit D. Straganz.

and will not be able to activate fatty acid residues by decarboxylation at room temperature. By contrast, a diiron-dioxo complex reacts with much lower hydrogen atom abstraction barriers and hence is a more likely oxidant in UndA.

3.2 Introduction

Nonheme iron dioxygenases are important enzymes in the human body responsible for vital functions for human health.[1] They are involved in biodegradation as well as biosynthesis reactions, such as the catabolism and regulation of toxic cysteine in the body and the biosynthesis of the amino acid R-4-hydroxyproline from proline.[2] The latter is an essential product for collagen crosslinking processes in the body and hence affects aging processes.[3] AlkB, another nonheme iron dioxygenase, is responsible for the DNA base repair mechanisms that convert alkylated DNA bases back to their nonalkylated forms.[4] In plants, many signaling molecules linked to flowering and fruit ripening are synthesized by nonheme iron dioxygenases.[5] As such, the nonheme iron dioxygenases have varied functions in biology and are highly selective. Nevertheless, many of these nonheme iron dioxygenases have similar structural features and utilize α-ketoglutarate as a co-substrate to form an iron(III)-superoxo that is converted into a high-valent iron(IV)-oxo species. The latter performs the oxygen atom transfer reactions to substrate.

Structurally, nonheme iron dioxygenases have specific ligand features where the metal is often bound to the side chains of two histidine amino acid groups and a carboxylate group of either Asp or Glu, typically in a 2-His/1-Asp facial ligand motif.[6] Interestingly, a few nonheme iron dioxygenases have ligand motifs deviating from 2-His/1-Asp but why that is and how it affects the reactivity is poorly understood. In particular, several thiolate utilizing nonheme iron dioxygenases contain a 3-His ligand motif including the ergothioneine biosynthesis enzyme EgtB and cysteine dioxygenase (CDO).[7,8] In both of these enzymes an iron(III)-superoxo species is formed after dioxygen binding, which reacts with substrate directly through oxygen atom transfer.

The recently discovered nonheme iron decarboxylase UndA is an 1-undecene biosynthesis enzyme and is able to convert medium-chain fatty acids into terminal olefins.[9] A crystal structure was determined that revealed a mononuclear iron atom that is ligated to the protein through linkages with His_{104}, His_{194} and Glu_{101} (**Figure 3.1**). Based on the crystallographic and biochemical studies a catalytic cycle was proposed for the conversion of medium-chain fatty acids (such as dodecanoic acid or lauric acid) to terminal olefins. They proposed direct binding of the substrate to the iron with its carboxylate group followed by a hydrogen atom abstraction reaction by the iron(III)-superoxo species. The mechanistic

proposal for UndA[9] resembles the ones proposed for substrate activation in isopenicillin N synthase (IPNS)[10] and 1-aminocyclopropane-1-carboxylic acid oxidase.[11] Both of these enzymes form an iron(III)-superoxo species that abstract a hydrogen atom from substrate through a rate-determining reaction step.

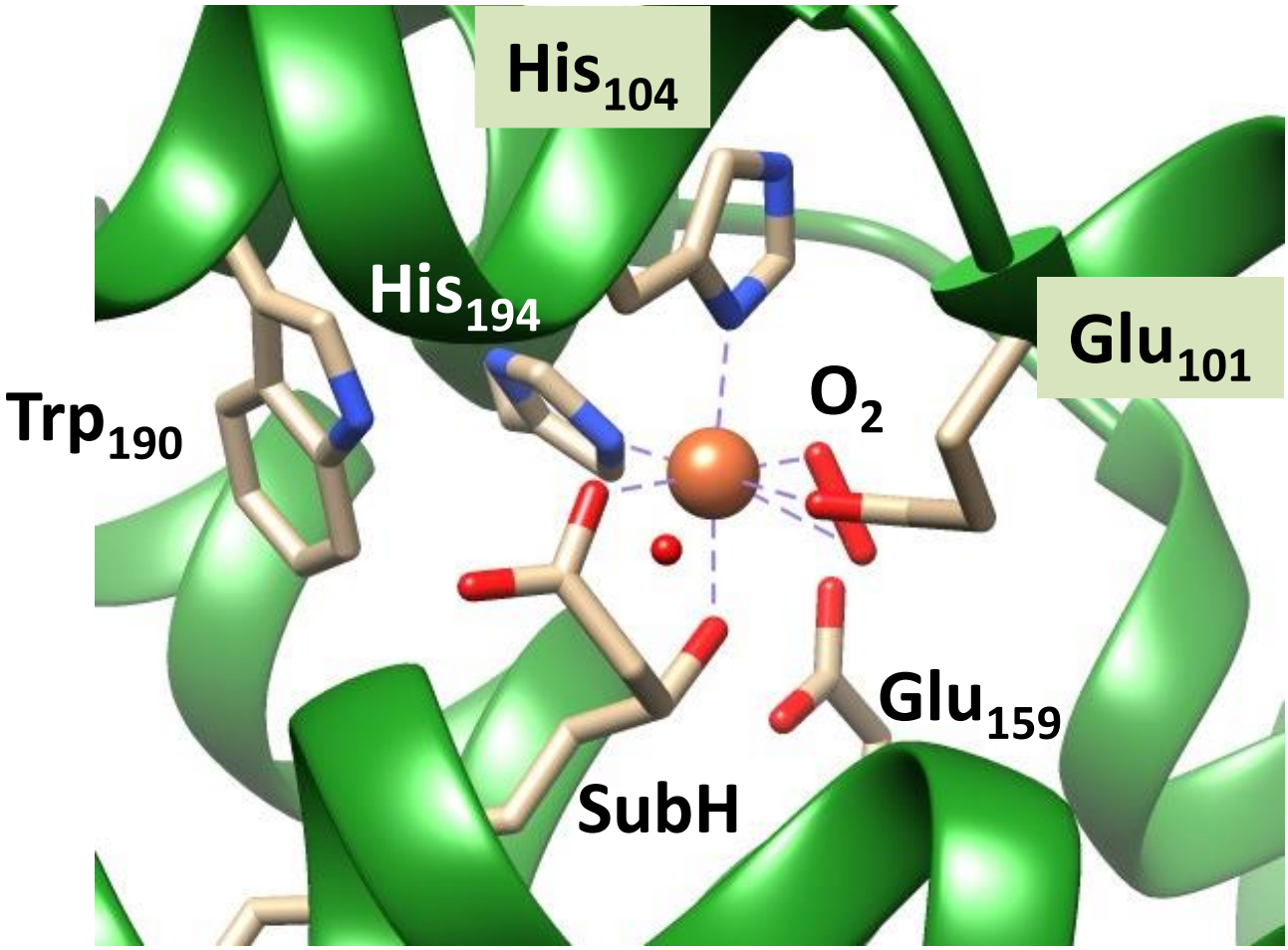


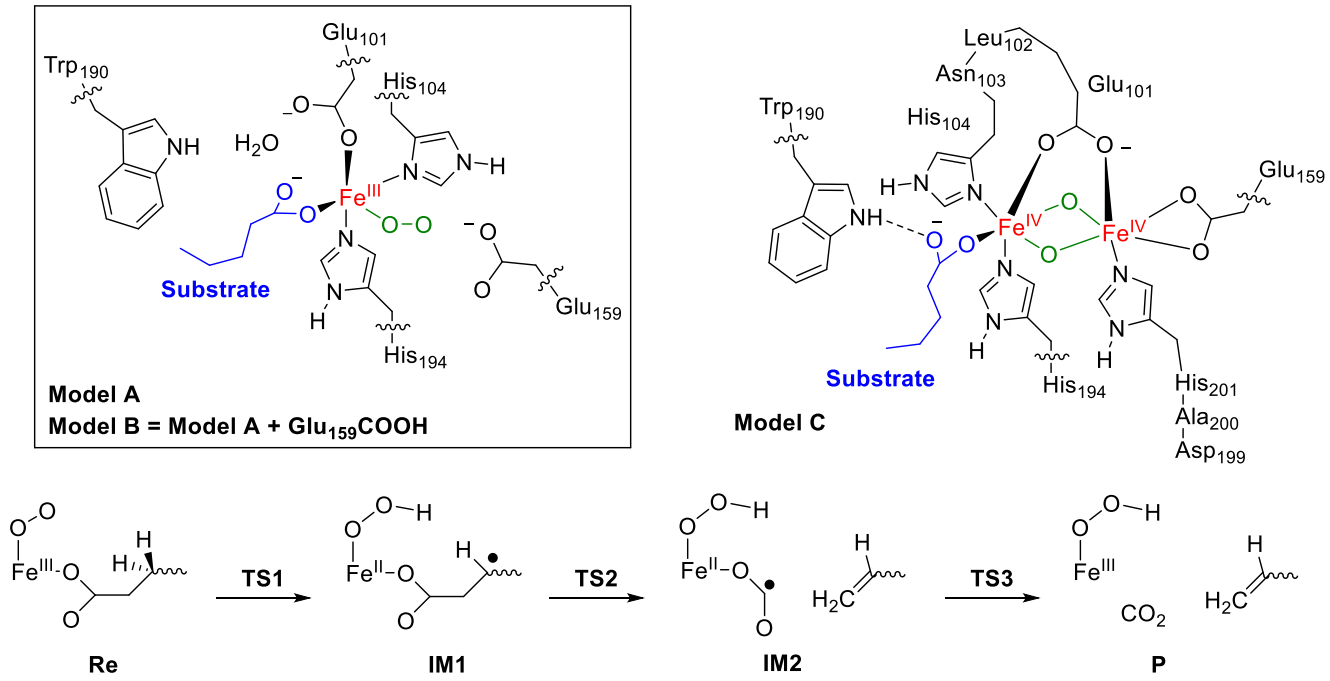
Figure 3.1 Active site structure of substrate bound UndA as taken from the 4WX0 pdb file.

Figure 3.1 displays the active site structure of UndA as taken from the 4WX0 protein databank (pdb) file.[9] As can be seen the iron(II) is bound to the protein through interactions with Glu_{101}, His_{104} and His_{194} and binds the substrate analog α-hydroxy-dodecanoic acid through the carboxylate group. In contrast to other nonheme iron dioxygenases, such as taurine/α-ketoglutarate dioxygenase and prolyl-4-hydroxylase,[12] the 2-His/1-Asp ligands do not form a facial triad but Glu_{101} in UndA is trans to His_{194} instead. The active site also has dioxygen (or superoxo) bound, which is in hydrogen bonding distance to the carboxylate group of Glu_{159}. The active site is completed by a nearby Trp residue (Trp_{190}) parallel to the His_{194} imidazole ring through π-stacking. As such there appears to be a network of hydrogen bonding interactions that stabilize the active site structure.

Recently, a controversy has arisen regarding the mechanism of UndA with substrates. Thus, Makris et al[13] did a similarity match with other enzymes and found the UndA protein to be close in structure to several diiron enzymes. In particular, nearby the iron atom in the crystal structure coordinates of 4WX0 they found a small pocket that could fit a second iron atom. Moreover, that pocket contains several viable residues for binding this second iron atom, namely His_{201} and Glu_{159}. They, therefore, suggested that UndA is more likely to be a diiron decarboxylase instead and reported spectroscopic evidence in support.[13] To resolve the dichotomy, and find out whether mononuclear iron center in UndA would be able to transform dodecanoic acid into undecene and CO_2, we decided to do a computational study into the reactivity of a mononuclear iron(III)-superoxo model of UndA with hexanoic acid as substrate using active site models with various protonation states. Our work shows that a mononuclear iron(III)-superoxo species in UndA is unlikely to be the oxidant as high energy reaction barriers for substrate activation are found for the mononuclear models, which would imply sluggish hydrogen atom abstraction ability at room temperature. For a diiron-dioxo model structure; however, much lower hydrogen atom abstraction barriers are found, hence is a more likely active species of UndA.

3.3 Result and discussion

To understand the reaction mechanism of UndA and find out whether it is more likely mononuclear or dinuclear iron, we decided to run a series of density functional theory calculations on active site models. In addition, we investigated the reactivity patterns for the Glu101His mutant of the mononuclear iron system that gives a 3-His metal binding pattern. Cluster models have been extensively used in enzyme mechanism research and were successfully applied previously to understand the catalysis by heme monoxygenase and nonheme iron dioxygenase studies.[14]



Scheme 3.1 DFT models investigated in this work and reaction mechanism with labelling of structures.

We created several active site models based on the crystal structure coordinates obtained from the 4WX0 pdb file.[9] **Scheme 3.1** displays the key features of our models and the mechanism explored with definitions of the labels of the various structures. Firstly, we created several mononuclear iron models of the active site of UndA, namely models **A** and **B**. These models included all first sphere ligands of the metal, i.e. Glu_{101}, His_{104} and His_{194}, whereby the histidine groups were abbreviated to methyl-imidazole and the glutamic acid groups by acetate. The substrate was shortened to hexanoic acid and Trp_{190} by methyl-indole. We also included an active site water molecule in the model that was seen to form a hydrogen bond between the carboxylate group of substrates and the Trp_{190} group. Nearby the dioxygen binding site is a carboxylic acid group of Glu_{159}, which we considered to be deprotonated (model **A**) or protonated (model **B**) and was included in the model as acetate/acetic acid, respectively. These models had overall charge –1 (model **A**) and 0 (model **B**) and were calculated with odd multiplicity. An alternative model (model **A1**), where the Glu_{101} was mutated for His was also explored and had overall charge 0 and explored the effect of the first-coordination sphere of iron ligands on the reactivity.

Finally, we created a dinuclear iron complex based on the suggestions of Makris et al,[13] whereby a second iron atom is linked to Glu_{101}, Glu_{159} and His_{201} and bridged by two oxo groups in a $Fe^{IV}_2O_2$ diamond conformation (model **C**). In this model the peptide chain linking Glu_{101} and His_{104} was incorporated in the model with the Leu_{102} and Asn_{103} residues abbreviated to Gly. Similarly, the chain

linking Asp_{199} and His_{201} was included with Ala_{200} shortened to Gly. The diiron model had overall neutral charge and was calculated in the ferromagnetic ($^5/_2$,$^5/_2$) diiron and antiferromagnetic ($^5/_2$, $-^5/_2$) diiron spin states.

The general reaction mechanism calculated starts from our reactant structure (**Re**) and follows the hydrogen atom abstraction of the β-position via transition state **TS1** to form the radical intermediate **IM1**. Next, the C_α–C_β bond in the substrate breaks via transition state **TS2** to form intermediate **IM2**. An electron transfer from $CO_2^{-\bullet}$ to iron then releases CO_2 and pentene via transition state **TS3** to from products **P**. For some models also hydrogen atom abstraction from the C_α-position of substrate was explored as identified with a subscript after the label.

3.3.1 Mononuclear iron model A.

We started the work with calculations of the mononuclear iron(III)-superoxo species with a starting structure that closely resembles the crystal structure coordinates. The lowest energy singlet, triplet, quintet, and septet spin states for Model **A** were geometry minimized and the optimized structures are given in **Figure 3.2.** Interestingly, the septet spin state is the ground state with the triplet and quintet spin states higher in energy by 2.7 and 12.1 kcal mol^{-1}. We also tested the singlet spin state but found it well higher by more than 30 kcal mol^{-1}, therefore, this spin state surface was not pursued further. The spin-state ordering of the iron(III)-superoxo complexes obtained here is unusual as a quintet spin ground state is normally seen for analogous complexes,[15] although for CDO an open-shell singlet spin state reactant was found using QM/MM.[16] The latter is probably the result of differences in metal coordination that stabilizes the low-spin state.

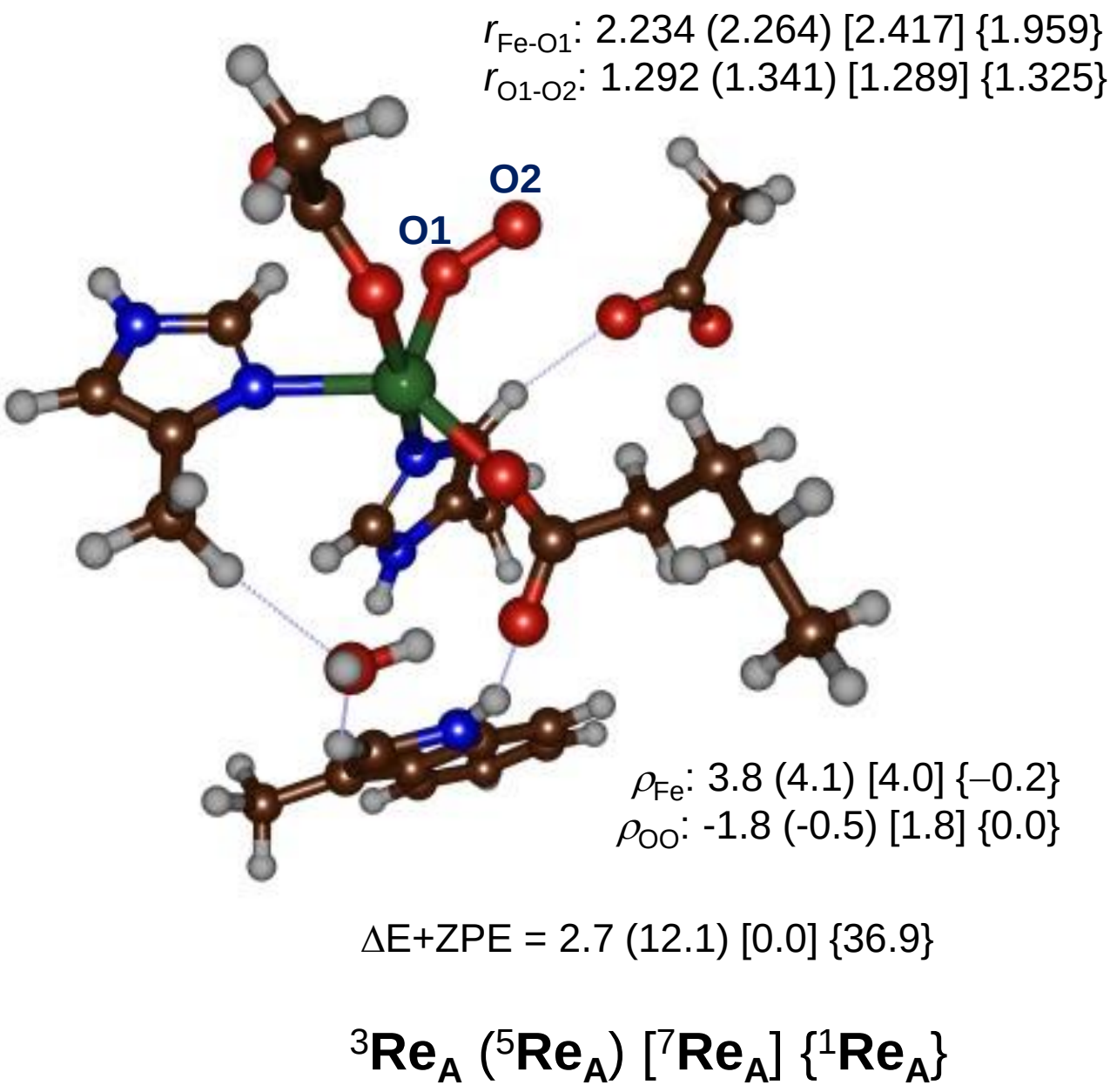


Figure 3.2 UB3LYP/BS1 optimized geometries of ^{1,3,5,7}**Re** as obtained in Gaussian-09. Bond lengths are in angstroms, spin densities in atomic units and ΔE+ZPE values are in kcal mol^{-1} with solvent included.

Geometrically, in the triplet, quintet and septet spin states the Fe–O distance in the iron(III)-superoxo species is relatively long, i.e. well over 2.2 Å and as a result dioxygen retains a considerable amount of spin density (about 1.8 in the triplet and septet spin states of **Re$_A$**). Therefore, these structures should be seen as an iron(II)-dioxygen complex rather than iron(III)-superoxo. Most likely this is caused by the hydrogen bonding interactions with the carboxylate group of Glu$_{159}$ that pushes electron density from dioxygen to iron and prevents it from tautomerization into the iron(III)-superoxo form. All low-lying

electronic states have four unpaired electrons on the metal and spin densities of 3.8, 4.1 and 4.0 on iron are found in ^{3}**Re$_{A}$**, ^{5}**Re$_{A}$** and ^{7}**Re$_{A}$**, respectively.

Subsequently, we investigated the reaction mechanism of substrate decarboxylation by the iron(III)-superoxo species on the lowest energy singlet, triplet, quintet and septet spin states of model **A**. The mechanism is distinct from heme decarboxylases, such as cytochrome P450 OleT$_{JE}$.[17] Studies on P450 OleT$_{JE}$ showed that it uses H_2O_2 on an iron heme and forms a high-valent iron(IV)-oxo heme cation radical active species called Compound I (Cpd I).[18] Indeed, previous QM/MM studies of our group,[19] on the decarboxylation of fatty acids by the cytochrome P450 peroxygenase OleT$_{JE}$ showed the reaction to start with hydrogen atom abstraction from the β-position of a long-chain fatty acid substrate by Cpd I and led to a subsequent barrierless decarboxylation quickly. Calculations starting with hydrogen atom abstraction from the C_α–H bond instead, however, were found to give substrate hydroxylation as sole products and no decarboxylation could be formed. As such a mixture of products was predicted for P450 OleT$_{JE}$, which was confirmed by further studies that tested a range of substrates whereby short-chain fatty acids gave hydroxylation and desaturation products, whereas long-chain unbranched fatty acids produced terminal olefins and hydroxylated fatty acids.[20] Interestingly, the experimental work of Zhang et al[9] on UndA showed no evidence of hydrogen atom abstraction from the C_α-position of substrate when those hydrogen atoms were replaced by deuterium.

Clearly, UndA and P450 OleT$_{JE}$ react differently with long chain fatty acids and therefore we decided to study the mechanism of substrate activation by the iron(III)-superoxo species of an UndA model. **Figure 3.3** displays the calculated reaction mechanism of hexanoic acid decarboxylation by the iron(III)-superoxo of UndA model **A** as calculated with DFT. The mechanism proceeds through multistate reactivity patterns on close-lying triplet, quintet and septet spin states as often seen for reactivity patterns of nonheme iron dioxygenases and nonheme iron model complexes. [21,22]

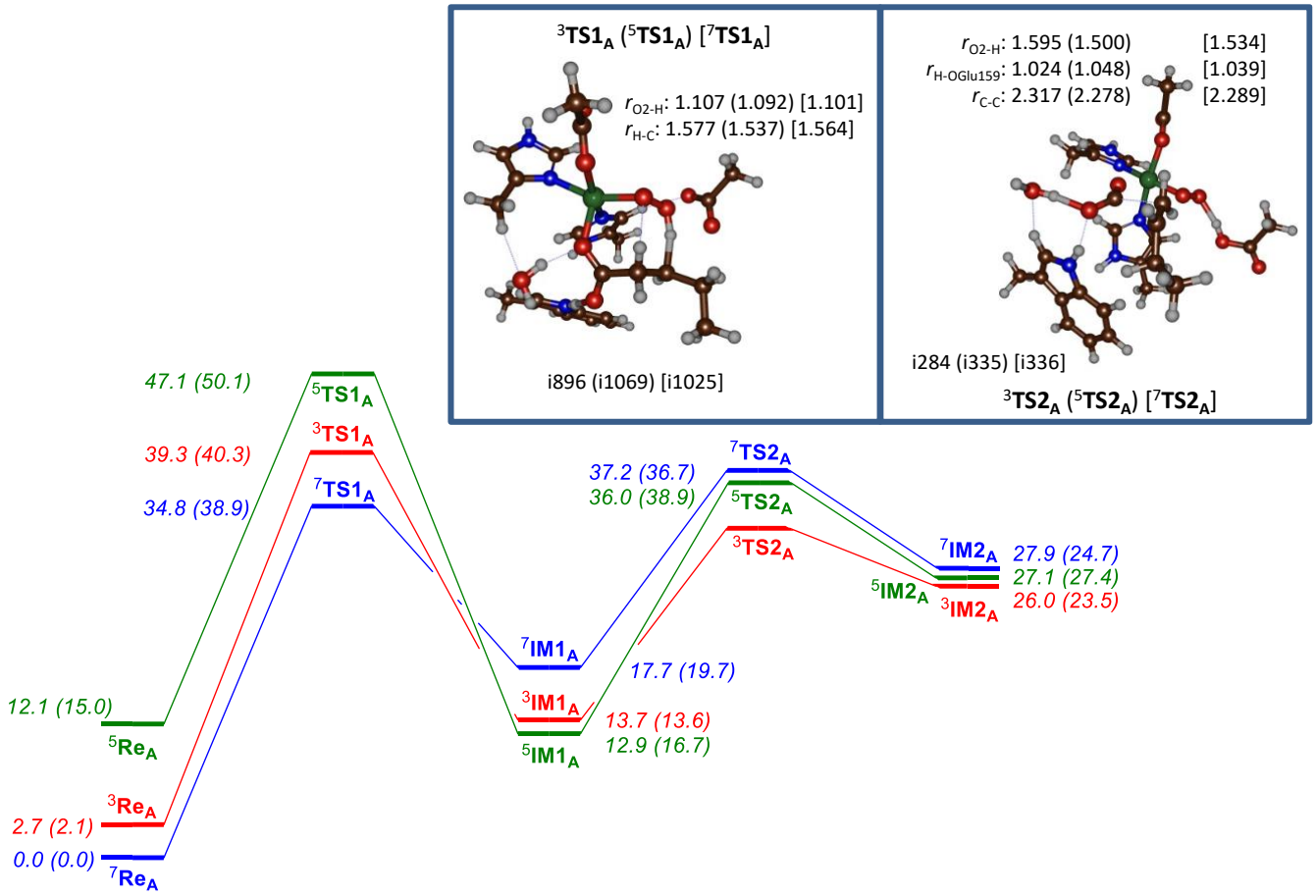


Figure 3.3 UB3LYP/BS2//UB3LYP/BS1+ZPE calculated potential energy profile of substrate decarboxylation by iron(III)-superoxo model **A** of UndA. Energies are in kcal mol^{-1} with Gibbs free energies (in parenthesis) at 298K with solvent, entropic and thermal corrections in parenthesis. Optimized geometries give bond lengths in angstroms and the imaginary frequency is in cm^{-1}.

As the singlet spin reactant is high in energy, we decided not to explore this surface further. The hydrogen atom abstraction barrier is lowest on the septet spin state with a barrier of 34.8 kcal mol^{-1}. The triplet and quintet spin hydrogen atom abstraction barriers are 39.3 and 47.1 kcal mol^{-1} in energy and hence the spin state ordering in the transition states is the same as in the reactants. These barriers are relatively high in energy and would suggest a slow and possibly sluggish reaction mechanism for hydrogen atom abstraction. For comparison, in previous work on trends of hydrogen atom abstraction reactions by P450 Cpd I we found substantially lower energy barriers (well below 20 kcal mol^{-1}) for typical aliphatic substrates, while for methane a value of 22.3 kcal mol^{-1} was found with the same methods and procedures as used here.[23] As P450 Cpd I is not known to activate methane, our hydrogen atom abstraction barriers of **Figure 3.3** implicate that an UndA model with an iron(III)-superoxo oxidant and nearby deprotonated Glu$_{159}$ residue will not be able to activate fatty acids at room temperature.

The energetic landscape depicted in **Figure 3.3** contradicts computational studies of several nonheme iron dioxygenases reported previously that gave efficient hydrogen atom abstraction by an iron(III)-superoxo. Thus, computational studies of Hirao et al.[24] on 2-hydroxyethylphosphonate dioxygenase predicted a rate determining hydrogen atom abstraction barrier of 19.7 kcal mol^{-1} by a mononuclear nonheme iron(III)-superoxo complex. A similar conclusion was obtained by Morokuma et al.[25] on isopenicillin N synthase, who reported a QM/MM calculated hydrogen atom abstraction barrier of 14.6 kcal mol^{-1} by a quintet spin iron(III)-superoxo complex. Both of these studies had the metal bound in a facial 2-His/1-Asp ligand orientation, while the UndA model has the two carboxylate groups equatorial to each other. These differences in ligand coordination may affect the reactivity and the ability of the iron to abstract electrons. Clearly, the hydrogen atom abstraction barriers by nonheme iron(III)-superoxo complexes vary widely and must be dependent on first- and second-coordination sphere effects of the metal centre. Therefore, we also explored alternative models and environmental effects.

Optimized geometries of the hydrogen atom abstraction transition states (^{3,5,7}**TS1**$_A$) are given in **Figure 3.3**. All hydrogen atom abstraction barriers (^{3,5,7}**TS1**$_A$) are product-like with short O–H and long C–H bonds: O–H distances of 1.107 (1.092) [1.101] Å and C–H distances of 1.577 (1.537) [1.564] Å are found for ^{3}**TS1**$_A$ (^{5}**TS1**$_A$) [^{7}**TS1**$_A$]. Previous trends on hydrogen atom abstraction barriers by metal-oxo complexes showed that transition states with product-type geometries usually correlate with high energy transition states,[23,26] as seen here. Although for isopenicillin-N-synthase (IPNS) product-like transition states were also found with either QM/MM and DFT cluster models, the C–H distance was much shorter: 1.36Å using QM/MM and 1.41Å for the DFT cluster model.[25,27] Yet, the barriers for hydrogen atom abstraction by an iron(III)-superoxo complex of IPNS was only 9.3 kcal mol^{-1} for the DFT cluster model and 14.6 kcal mol^{-1} with QM/MM. Consequently, nonheme iron(III)-superoxo should be able to activate C–H bonds of substrates and the fact that we do not see any reactivity in our model here must originate from the coordination environment of the metal centre.

The hydrogen atom abstraction transition states are characterized with a large imaginary frequency of i896 (triplet), i1069 (quintet) and i1025 cm^{-1} (septet). These are typical values for hydrogen atom abstraction transition states and often implicate a large kinetic isotope effect when the transferring hydrogen atom is replaced by deuterium.[28] Group spin densities of ^{7}**TS1**$_A$ shows a drop in value on the dioxygen moiety from 1.78 in ^{7}**Re**$_A$ to 1.04 and accumulation of unpaired spin on the substrate moiety to 0.77. A similar change is seen in the triplet spin pathway, where the dioxygen spin density rises from –1.77 in ^{3}**Re**$_A$ to –1.00 in ^{3}**TS1**$_A$, while the radical on the substrate has a value of –0.68. Therefore, in our reactant complex the iron(II)-dioxygen species picks up a hydrogen atom to form an iron(III)-

hydroperoxo complex. During the geometry optimization of ^{1,3,5,7}**IM1**$_A$; however, the hydroperoxo species relays its proton to the carboxylic acid group of Glu_{159}. Moreover, the group spin densities give a single unpaired electron of the dioxygen species and hence these structures should be seen as iron(III)-superoxo complexes.

In the next stage of the reaction the C_α–C_β bond of the substrate cleaves via a transition state **TS2**$_A$ to form intermediate **IM2**$_A$. In all **TS2**$_A$ transition states significant radical character on the substrate remains, which means the C_α–C_β bond breakage gives $CO_2^{-\bullet}$ and terminal olefin. At the same time, the spin density on the iron-hydroperoxo unit is very similar in **IM1**$_A$ and **IM2**$_A$ for all spin states and confirms that the electron transfer from $CO_2^{-\bullet}$ to iron has not happened at this stage and therefore is not simultaneous to the C_α–C_β bond cleavage. On the quintet and septet spin state surfaces the barriers for decarboxylation via **TS2**$_A$ are very large and implicate slow reaction steps. Overall, the calculations of substrate decarboxylation by a mononuclear iron(III)-superoxo species of UndA model **A** implicate an unrealistic high energy pathway with barriers that indicate a very slow reaction process. To find out whether mononuclear iron(III)-superoxo can react with fatty acids via decarboxylation reactions, we decided to explore alternative model structures.

3.3.2 Mononuclear iron model A2

To gain insight into the first-coordination sphere effects of ligands in mononuclear iron UndA, we tested an UndA mutant structure where the Glu_{101} residue is replaced by His, i.e. the Glu101His mutant or model **A2**. The iron(III)-superoxo reactant of model **A2** has a septet spin ground state that is well separated from the nearest quintet, triplet and singlet spin states by 11.1, 43.8 and 54.0 kcal mol^{-1}. The hydrogen atom abstraction barriers on each of these spin state surfaces are given in **Figure 3.4**. All of them are high in energy and range from 38.4 (for ^{7}**TS1**$_{A2}$) to 45.1 (for ^{5}**TS1**$_{A2}$) kcal mol^{-1}. As such the mutant with 3-His metal coordination as well as wild-type with 2-His/1-Glu coordination both are sluggish oxidants of hydrogen atom abstraction reactions. For this particular model, the change in metal coordination from an anionic ligand (Glu) to a neutral ligand (His) has little effect on the kinetics of the reaction with substrate.

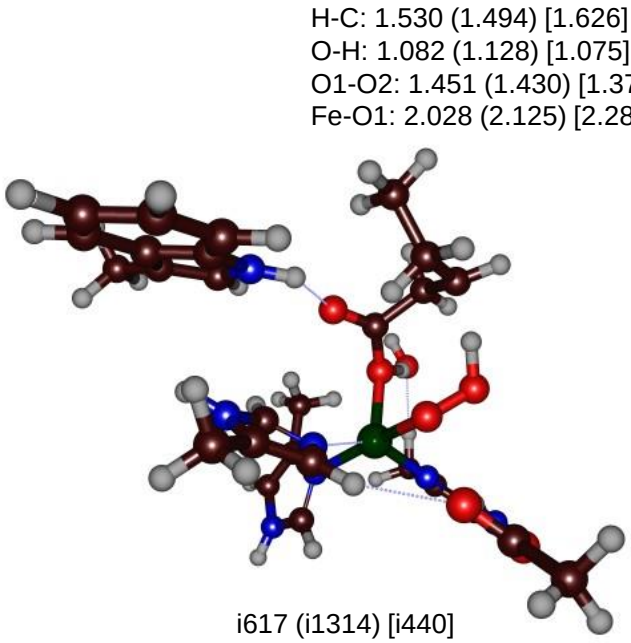


Figure 3.4 UB3LYP/BS2//UB3LYP/BS1+ZPE calculated hydrogen atom abstraction barriers by the iron(III)-superoxo model **A2** of UndA. Energies are in kcal mol^{-1} with Gibbs free energies at 298K with solvent, entropic and thermal corrections in parenthesis. Optimized geometries give bond lengths in angstroms and the imaginary frequency in cm^{-1}.

The hydrogen atom abstraction transition states for the 3-His mutant of mononuclear iron UndA are given in **Figure 3.4**. Similarly, to those given above in Figure 3 for model **A** also for model **A2** the transition states are late with long C–H and short O–H distances for ^{3,5,7}**TS1**$_{A2}$. The imaginary frequencies are considerably smaller in the triplet and septet spin states but still correspond to a hydrogen atom transfer from substrate to superoxo group. In conclusion, the first-coordination sphere of ligands, be it 2-His/1-Glu or 3-His, in UndA gives high barriers for hydrogen atom abstraction and therefore, the system appears to have the wrong coordination environment for efficient hydrogen atom abstraction from substrate.

3.3.3 Mononuclear iron model B

To test whether the second coordination sphere affects the structure and reactivity of a mononuclear iron UndA model, we expanded our model and added a proton to the carboxylate group of Glu$_{159}$, i.e. investigated model **B**. We started with the iron(III)-superoxo reactant complexes and calculated them in

the lowest energy triplet, quintet and septet spin states The septet spin state (^{7}**Re**$_{B}$) is still the ground state, but the quintet spin state (^{5}**Re**$_{B}$) is now the nearest one by 3.0 kcal mol^{-1}, while the triplet spin state (^{3}**Re**$_{B}$) has gone up to 18.1 kcal mol^{-1}.

Therefore, the second-coordination sphere has a strong effect on the spin state energies and affects the ordering and relative energies. One of the key effects seen in the optimized reactant structures of **Re**$_{B}$, **Figure 3.5**, is the hydrogen bond of the proton of Glu$_{159}$ with the carboxylate group of the fatty acid substrate. By contrast, the Glu$_{159}$ group in model **A** (**Figure 3.2**) hydrogen bonds with a histidine group instead. As such there has been some structural reorganization in the active site and the charge-stabilization of Glu$_{159}$ should influence the electronic configuration of the first-coordination sphere atoms. Note also that the substrate carboxylate is now bidentate coordination to iron, while it has single coordination in model **A**.

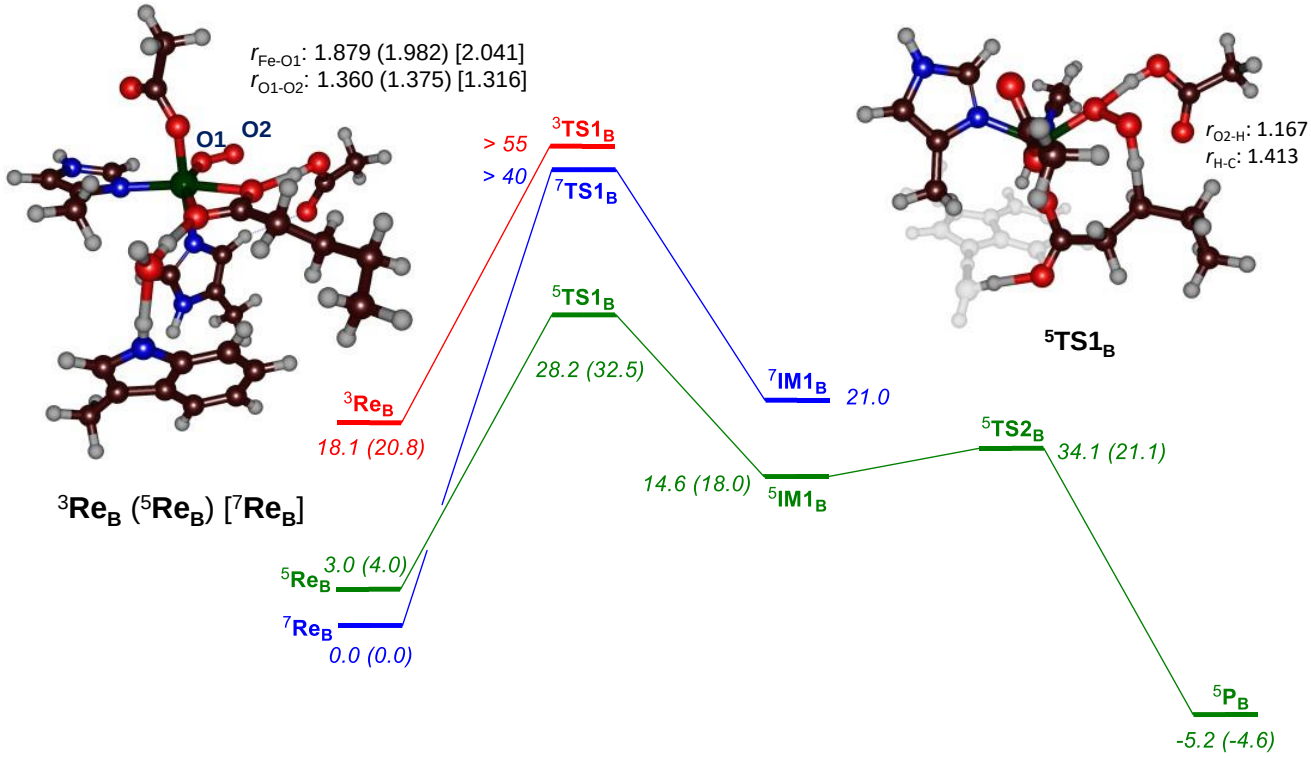


Figure 3.5 UB3LYP/BS2//UB3LYP/BS1+ZPE calculated Gibbs free energy profile of substrate decarboxylation by iron(III)-superoxo model B of UndA. Energies are in kcal mol^{-1} with Gibbs free energies (in parenthesis) at 298K with solvent, entropic and thermal corrections in parenthesis. Optimized geometries give bond lengths in angstroms.

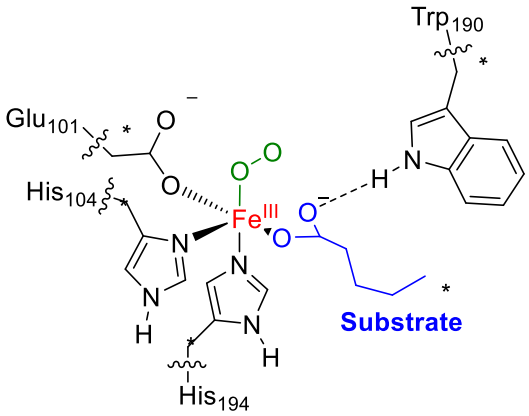
Next, we calculated the hydrogen atom abstraction from the C$_{\alpha}$–H and C$_{\beta}$–H positions of the substrate by ^{3,5,7}**Re**$_{B}$ and the obtained potential energy landscape is given in **Figure 3.5**. The quintet spin hydrogen

atom abstraction transition state (^{5}**TS1**$_{B}$) has a Gibbs free energy of activation $\Delta G^{\ddagger}$ = 32.5 kcal mol^{-1} with respect to ^{7}**Re**$_{B}$ reactants. Although, the barrier has dropped considerably in energy with respect to the non-protonated species, i.e. $\Delta G^{\ddagger}$ = 32.5 kcal mol^{-1} for model **B** vs 47.1 kcal mol^{-1} for model **A**; it is still very high for a hydrogen atom abstraction barrier. The structure has a short O–H distance of 1.167Å and a relatively long C–H distance of 1.413Å and hence, it is a late transition state, but not as early as ^{5}**TS1**$_{A}$ that has an even longer C–H distance.

Although we managed to optimize a transition state geometry for ^{5}**TS1**$_{B,\alpha}$, the geometry scans on the triplet and septet spin state showed high reaction barriers and were discarded. The subsequent C_{α}–C_{β} cleavage barrier is well over 90 kcal mol^{-1} and therefore, α-hydrogen atom abstraction will not lead to decarboxylation of fatty acids. We also attempted a proton transfer from Glu$_{159}$-COOH to iron(III)-superoxo to form an iron-hydroperoxo complex. However, the constraint geometry scan for this pathway identified a high-energy process, hence is unfeasible under room temperature conditions.

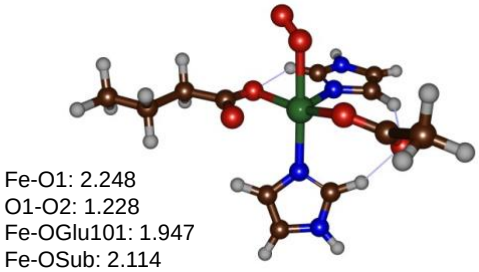
Overall the calculations on mononuclear iron models of UndA show that iron(III)-superoxo species are weak oxidants of hydrogen atom abstraction processes. This may have to do with the ligand coordination in UndA, which does not have a facial 2-His/1-Glu but the Glu trans to one of the His groups. By contrast, isopenicillin N synthase with a facial 2-His/1-Asp ligand coordination reacts by hydrogen atom abstraction from substrate with relatively low-energy barriers. Therefore, it appears that the ligand coordination in UndA is not correctly set up for hydrogen atom abstraction processes and our DFT cluster models rule out mononuclear iron as an oxidant that triggers the decarboxylation of dodecanoic acid in UndA.

Subsequently, we explored the effect of the ligand environment, i.e. a facial 2-His/1-Glu ligand coordination, on the reactivity with respect to that seen for model **A**. To this end, we created two final models of a hypothetical nonheme iron(III)-superoxo model with a facial 2-His/1-Glu coordination system: models **D** and **D2** (Supporting Information Figures S1 and S2). The reactant was optimized in the quintet spin state and a transition state for hydrogen atom abstraction was searched through an initial constraint geometry optimization. However, upon close approach of peroxo on the the C_{β}–H position, the superoxo group dissociates from the iron center leading to high energy hydrogen atom abstraction pathways.



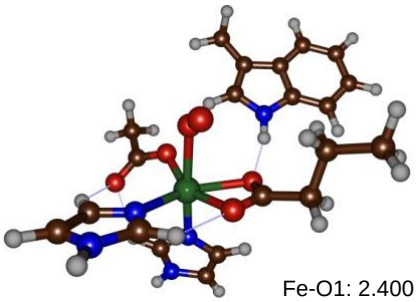
Model D: $Fe^{III}(OO)(Glu_{101})(His_{104})(His_{194})$--SubH
Model D2: Model D + Trp_{190}

Figures S1: **Model D with a 2-His/1-Asp facial ligand system.**



Fe-O1: 2.248
O1-O2: 1.228
Fe-OGlu101: 1.947
Fe-OSub: 2.114
Fe-NHis104: 2.204
Fe-NHis194: 2.176

^5Re_D



Fe-O1: 2.400
O1-O2: 1.222
Fe-OGlu101: 1.959
Fe-OSub: 2.130
Fe-NHis104: 2.190
Fe-NHis194: 2.168

$^5Re_{D2}$

Figures S2: **UB3LYP/BS1 optimized geometries of iron(III)-superoxo complexes ^{5}Re for Models D and D2 as obtained in Gaussian-09. Bond lengths are in angstroms.**

Consequently, this small model had too much freedom for the superoxo group and weakened its interaction with iron. Hence the small model with facial 2-His/1-Glu coordination system also did not lead to a viable hydrogen atom abstraction pathway from a fatty acid substrate. It may very well be that the substrate binding pocket in IPNS is much tighter than the one used in model **D** and prevents release of dioxygen from the iron center. A tight binding and positioning of oxidant and substrate should bring the two closely together in an ideal conformation for catalysis. Nevertheless, our calculations on various models of the mononuclear iron(III)-superoxo species of UndA show little evidence that hydrogen atom abstraction from substrate as a viable reaction mechanism. Therefore, we explored alternative diiron models instead.

3.3.4 Dinuclear iron model C

As mononuclear iron(III)-superoxo models of UndA appear to give high energy hydrogen atom abstraction transition states from aliphatic groups, we considered the alternative diiron species as an oxidant. Therefore, we created a diiron(IV)-dioxo active species from the UndA crystal structure coordinates by manually inserting a second iron into the cluster model and by linking this metal to nearby protein residues as suggested by Makris et al.[13]

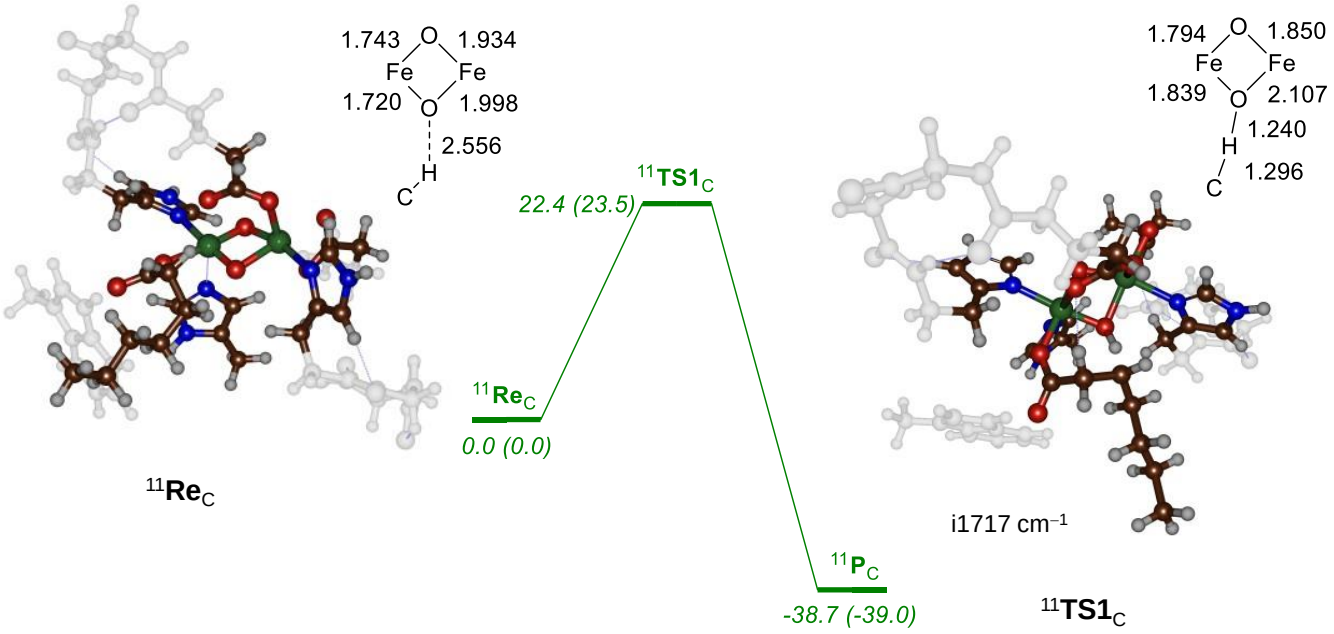


Figure 3.6 UB3LYP/BS2//UB3LYP/BS1+ZPE calculated energy profile of substrate decarboxylation by diiron(IV)-dioxo model **C** of UndA. Energies are in kcal mol^{-1} with Gibbs free energies (in parenthesis) at 298K with solvent, entropic and thermal corrections in parenthesis. Optimized geometries give bond lengths in angstroms and the imaginary frequency in cm^{-1}.

A full geometry optimization of this reactant complex (^{11}**Re**$_C$) is shown in **Figure 3.6.** The diiron-dioxo is in a diamond core although the two sets of iron-oxo distances are not equal. Thus, Fe1, i.e. the metal that binds substrate, forms short interactions with the two oxygen bridges (1.720 and 1.743Å), while Fe2 is further displaced from the bridging oxygen atoms (at distances of 1.934 and 1.998Å). In the reactant structure, the C$_\beta$–H bond of the substrate points towards one of the bridging oxygen atoms at a distance of 2.556Å. In this particular orientation, therefore, the positioning of the C$_\beta$–H bond of substrate is ideal for hydrogen atom abstraction.

To investigate the ability of $^{11}\mathbf{Re}_C$ to abstract a hydrogen atom from the C_β–H position of substrate, we calculated the transition state structure. A hydrogen atom abstraction barrier of $\Delta E^{\ddagger}$ + ZPE = 22.4 kcal mol^{-1} is found. Even though this is a relatively high barrier in comparison to previous enzymatic reaction mechanisms, where typical barriers of 10 – 15 kcal mol^{-1} were found.[14,21,24–27,29] Our calculated reaction barrier would correspond to slow but doable reactivity at room temperature. As a matter of fact the experimental work of Zhang et al.[9] report a slow reaction process for substrate conversion by UndA in agreement with the high reaction barrier reported in **Figure 3.6**.

The hydrogen atom abstraction barrier has a large imaginary frequency of i1717 cm^{-1}, which is higher than those reported above for models **A** and **A2**. Consequently, the barrier will be narrow and sharp and will be highly dependent on quantum chemical tunneling. The transition state is relatively central with almost equal C–H and O–H distances of 1.296 and 1.240Å, respectively. After the hydrogen atom abstraction, the system relaxes to a radical intermediate **IM1**, which is a shallow minimum and followed by C_α–C_β bond cleavage of the substrate to split off CO_2 and give 1-pentene products without a stable intermediate **IM2**. Therefore, the process past the transition state, will quickly lead to final products.

The calculations on the diiron(IV)-dioxo model of UndA and its reaction with hexanoic acid shows that diiron complexes can activate fatty acid substrates at room temperature. A cluster model shows that substrate binds in an ideal position for C_β–H hydrogen atom abstraction by a diiron(IV)-dioxo complex. The reaction barriers obtained for this complex, however, are well higher than those found for fatty acid decarboxylation by a P450 peroxygenase.[19] As such, UndA is not as efficient and powerful an oxidant as a P450 peroxygenase, but will react slowly with fatty acid substrates. This is in agreement with experimental reports on UndA that reported slow turnover reactions wit fatty acids.

3.4 Conclusions

In this work a computational study on the 1-undecene biosynthesis enzyme UndA is reported. Recent experimental studies created a controversy whether UndA has a mononuclear or binuclear iron active site. To investigate this, we created active site models of UndA based on the crystal structure coordinates of 4WX0. Our studies show that mononuclear iron(III)-superoxo models of UndA are sluggish oxidants that react via very high hydrogen atom abstraction barriers at room temperature. Consequently, mononuclear iron centers in the UndA structure are unlikely to react with fatty acids at room temperature. We then created a hypothetical diiron(IV)-dioxo structure using the crystal structure coordinates of UndA as a

template. This model gives substantially lower hydrogen atom abstraction barriers; although it will be a slow oxidant at room temperature. Overall, our studies support the hypothesis of Makris et al[13] that UndA is a most likely a dinuclear iron system that reacts with dioxygen and fatty acids by decarboxylation of the substrate.

3.5 Experimental Section

Calculations were done using density functional theory methods as implemented in the Gaussian-09 and Orca software packages.[30] In previous work we extensively tested and benchmarked models and methods for nonheme iron systems and reproduced experimental structures, spectroscopic parameters and rate constants well.[31] Hence, most methods follow those recommended from previous work. Here we started our work from the 4WX0 protein databank file,[9] which is an enzyme monomer with substrate and dioxygen bound. A cluster model was created from the active site structure of the 4WX0 pdb with residues selected as described in Scheme 1 above. We initially ran calculations with several constraints on amino acid residues to keep them close to the crystal structure positions. However, in a second set of calculations all constraints were released. As little geometric differences between the two structures were obtained, the final set of calculations was done without constraints.

Initial calculations used the unrestricted B3LYP hybrid density functional method,[32] in combination with a basis set containing an LANL2DZ + ECP on iron and 6-31G on the rest of the atoms (basis set BS1).[33] Full geometry optimization and frequencies were run on all structures at UB3LYP/BS1 in the gas-phase. Subsequent single point calculations with the polarized continuum model (CPCM) were performed with a dielectric constant mimicking ethylphenylether,[34] and a triple-ζ quality basis set (Basis set BS2): LACV3P+ + ECP on iron and 6-311+G* on the rest of the atoms. Calculations on the diiron system use an LACVP (with core potential) on both iron atoms and 6-31G* on the rest of the atoms for geometry optimizations, frequencies, intrinsic reaction coordinate and constraint geometry scans. All structures reported in this work are the result of a full geometry optimization without constraints and structures were characterized as a local minimum using a frequency calculation that gave real frequencies only (local minima) or a single imaginary mode for the correct displacement (transition states). For several transition states they were further confirmed by intrinsic reaction coordinate scans that linked them to reactants and products. All structures were considered in several low-lying spin state surfaces as identified with a superscript before the label.

The methods and procedures used in this work were tested and validated for analogous iron(IV)-oxo oxidants against experimental rate constants for oxygen atom transfer.[31,35] These studies identified B3LYP as one of the best methods to reproduce reaction rates and selectivities. Furthermore, dispersion corrected DFT was shown to underestimate oxygen atom transfer reactions systematically and hence was not used for our calculations presented here.

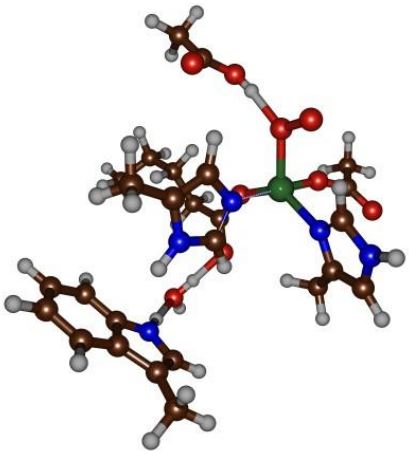
As the choice of the density functional method can occasionally affect spin-state orderings and relative energies we did some additional test calculations.[36] Thus, the effect of dispersion on the spin-state ordering and reaction barriers was tested and ^{3,5,7}**Re**$_B$ and the reaction pathway via ^{5}**TS1**$_{B,\beta}$ was reoptimized with B3LYP-D3.[32,37] As follows dispersion destabilizes the quintet spin state but the overall hydrogen atom abstraction barrier stays the same and the conclusions that mononuclear iron(III)-superoxo is a weak oxidant for hydrogen atom abstraction reactions is supported.

We also explored the OPBE [38] density functional method; although it stabilizes the triplet spin reactants, the quintet-septet energy gap remains the same (Supporting Information Table S1 and Figure S3). To further ascertain that the basis set core potential does not affects the spin state ordering, we reoptimized ^{3,5,7}**Re**$_B$ with UB3LYP/BS1(without ECP) and although the triplet and quintet spin states are lowered by a few kcal mol^{-1} no changes in reaction mechanism are expected (Supporting Information Table S2 and Figure S4).

	E [au]	ZPE [au]	G [au]	ΔE	ΔE+ZPE	ΔG
	BS1	BS1	BS1			
^{3}Re$_B$	-2087.414130	0.638971	-2086.879174	3.02	2.46	-1.10
^{5}Re$_B$	-2087.409202	0.640315	-2086.871790	6.12	6.40	3.53
^{7}Re$_B$	-2087.418949	0.639862	-2086.877422	0.00	0.00	0.00

Table S1: Absolute, zero-point and free energies (in au) and relative energies (in kcal mol^{-1}) of optimized geometries ^{3,5,7}Re$_B$ as calculated in Gaussian-09 at the UOPBE/BS1 level of theory.

Fe-O1: 4.372 (2.051) [2.270]
O1-O2: 1.257 (1.349) [1.294]
Fe-OGlu101: 1.937 (1.914) [1.925]
Fe-OSub: 2.284 (1.950) [2.309]
Fe-NHis104: 2.195 (2.144) [2.178]
Fe-NHis194: 2.236 (2.273) [2.233]



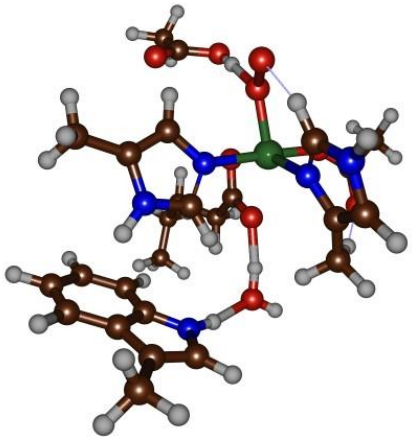
$^{3}\mathbf{Re_B}$ ($^{5}\mathbf{Re_B}$) [$^{7}\mathbf{Re_B}$]

Figure S3: UOPBE/BS1 optimized geometries of the reactant complexes ($^{3,5,7}Re_B$) of Model B as obtained in Gaussian-09. Bond lengths are in angstroms, in order to test calculations with OPBE on Model B.

	E [au]	ZPE [au]	G [au]	ΔE	ΔE+ZPE	ΔG
	BS1	BS1	BS1			
$^{3}Re_B$	-3228.173133	0.648915	-3227.617967	13.00	14.48	16.51
$^{5}Re_B$	-3228.194370	0.647999	-3227.641429	-0.32	0.58	1.79
$^{7}Re_B$	-3228.193854	0.646560	-3227.644277	0.00	0.00	0.00

Table S2: Absolute, zero-point and free energies (in au) and relative energies (in kcal mol^{-1}) of optimized geometries $^{3,5,7}Re_B$ as calculated in Gaussian-09 at the UB3LYP/BS1(without ECP) level of theory.

Fe-O1: 1.877 (2.006) [2.042]
O1-O2: 1.360 (1.375) [1.317]
Fe-OGlu101: 1.923 (1.902) [1.934]
Fe-OSub: 2.118 (2.149) [2.202]
Fe-NHis104: 1.991 (2.137) [2.110]
Fe-NHis194: 2.146 (2.202) [2.205]



$^{3}\mathbf{Re_B}$ ($^{5}\mathbf{Re_B}$) [$^{7}\mathbf{Re_B}$]

Figure S4: UB3LYP/BS1(without ECP) optimized geometries of the reactant complexes ($^{3,5,7}Re_B$) of Model B as obtained in Gaussian-09. Bond lengths are in angstroms.

3.6 UndA References

[1]
a) E. I. Solomon, T. C. Brunold, M. I. Davis, J. N. Kemsley, S.-K. Lee, N. Lehnert, F. Neese, A. J. Skulan, Y.-S. Yang, J. Zhou, Chem. Rev. 2000, 100, 235–349;
b) T. D. H. Bugg, Curr. Opin. Chem. Biol. 2001, 5, 550–555;
c) M. J. Ryle, R. P. Hausinger, Curr. Opin. Chem. Biol. 2002, 6, 193–201;
d) M. Costas, M. P. Mehn, M. P. Jensen, L. Que, Jr., Chem. Rev. 2004, 104, 939–986;
e) M. M. Abu-Omar, A. Loaiza, N. Hontzeas, Chem. Rev. 2005, 105, 2227–2252;
f) C. Krebs, D. G. Fujimori, C. T. Walsh, J. M. Bollinger, Jr., Acc. Chem. Res. 2007, 40, 484–492;
g) P. C. A. Bruijnincx, G. van Koten, R. J. M. Klein Gebbink, Chem. Soc. Rev. 2008, 37, 2716–2744.
[2]
a) R. K. Bruick, S. L. McKnight, Science 2001, 294, 1337–1340;
b) A. Seifert, D. M. Katschinski, S. Tonack, B. Fischer, A. N. Santos, Chem. Res. Toxicol. 2008, 21, 341–348;
c) C. J. Schofield, Z. Zhang, Curr. Opin. Struct. Biol. 1999, 9, 722–731;
d) S. P. de Visser, Chem. Rec. 2018, 18, 15011516.
[3]
a) K. Gorres, R. T. Raines, Crit. Rev. Biochem. Mol. Biol. 2010, 45, 106–124;
b) M. A. McDonough, V. Li, E. Flashman, R. Chowdhury, C. Mohr, B. M. Lienard, J. Zondlo, N. J. Oldham, I. J. Clifton, J. Lewis, L. A. McNeill, R. J. Kurzeja, K. S. Hewitson, E. Yang, S. Jordan, R. S. Syed, C. J. Schofield, Proc. Natl. Acad. Sci. USA 2006, 103, 9814–9819.
[4]
a) S. C. Trewick, T. F. Henshaw, R. P. Hausinger, T. Lindahl, B. Sedgwick, Nature 2002, 419, 174–178;
b) P. Ø. Falnes, R. F. Johansen, E. Seeberg, Nature 2002, 419, 178–182;
c) P. J. O'Brien, Chem. Rev. 2006, 106, 720752; d) C. Yi, C. G. Yang, C. He, Acc. Chem. Res. 2009, 42, 519–529.
[5]
a) M. D. White, E. Flashman, Curr. Opin. Chem. Biol. 2016, 31, 126–135;
b) C. Li, S. Liu, X. Yao, J. Wang, T. Wang, Z. Zhang, P. Zhang, K. Chen, Plant Growth Regul. 2017, 83, 489–500.
[6]
a) L. Que, Jr., Nat. Struct. Biol. 2000, 7, 182–184;
b) E. G. Kovaleva, J. D. Lipscomb, Nat. Chem. Biol. 2008, 4, 186–193.
[7]
a) B. J. Landgraf, S. J. Booker, J. Am. Chem. Soc. 2016, 138, 2889–2892;
b) R. Rohac, P. Amara, A. Benjdia, L. Martin, P. RuffiØ, A. Favier, O. Bertreau, J.-M. Mouesca, J. C. FontecillaCamps, Y. Nicolet, Nat. Chem. 2016, 8, 491–500;
c) K. V. Goncharenko, A. Vit, W. Blankenfeldt, F. P. Seebeck, Angew. Chem. Int. Ed. 2015, 54, 2821–2824; Angew. Chem. 2015, 127, 2863–2866;
d) K. V. Goncharenko, F. P. Seebeck, Chem. Commun. 2016, 52, 1945–1948.
[8]
a) M. H. Stipanuk, Annu. Rev. Nutr. 2004, 24, 539–577;
b) G. D. Straganz, B. Nidetzky, ChemBioChem 2006, 7, 1536–1548; c) C. A. Joseph, M. J. Maroney, Chem. Commun. 2007, 3338–3349;
d) S. Aluri, S. P. de Visser, J. Am. Chem. Soc. 2007, 129, 14846–14847;
e) D. Buongiorno, G. D. Straganz, Coord. Chem. Rev. 2013, 257, 541–563;
f) E. P. Tchesnokov, M. Fellner, E. Siakkou, T. Kleffmann, L. W. Martin, S. Aloi, I. L. Lamont, S. M. Wilbanks, G. N. L. Jameson, J. Biol. Chem. 2015, 290, 24424–24437.

[9] Z. Rui, X. Li, X. Zhu, J. Liu, B. Domigan, I. Barr, J. H. D. Cate, W. Zhang, Proc. Natl. Acad. Sci. USA 2014, 111, 18237–18242.
[10]
a) J. E. Baldwin, M. Bradley, Chem. Rev. 1990, 90, 1079–1088;
b) P. L. Roach, I. J. Clifton, C. M. H. Hensgens, N. Shibata, C. J. Schofield, J. Hajdu, J. E. Baldwin, Nature 1997, 387, 827–830;
c) M. Lundberg, P. E. M. Siegbahn, K. Morokuma, Biochemistry 2008, 47, 1031–1042;
d) L. A. McNeill, T. J. N. Brown, M. Sami, I. J. Clifton, N. I. Burzlaff, T. D. W. Claridge, R. M. Adlington, J. E. Baldwin, P. J. Rutledge, C. J. Schofield, Chem. Eur. J. 2017, 23, 12815–12824;
e) E. Tamanaha, B. Zhang, Y. Guo, W.-C. Chang, E. W. Barr, G. Xing, J. St Clair, S. Ye, F. Neese, J. M. Bollinger, Jr., C. Krebs, J. Am. Chem. Soc. 2016, 138, 8862–8874.
[11]
a) M. C. Pirrung, Acc. Chem. Res. 1999, 32, 711–718;
b) Z. Zhang, J.-S. Ren, I. J. Clifton, C. J. Schofield, Chem. Biol. 2004, 11, 1383–1394.
[12]
a) J. R. O'Brien, D. J. Schuller, V. S. Yang, B. D. Dillard, W. N. Lanzilotta, Biochemistry 2003, 42, 5547–5554;
b) M. K. Koski, R. Hieta, M. Hirsilä, A. Rçnkä, J. Myllyharju, R. K. Wierenga, J. Biol. Chem. 2009, 284, 2529025301;
c) H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, Nucl. Acids Res. 2000, 28, 235–242.
[13] O. M. Manley, R. Fan, Y. Guo, T. M. Makris, J. Am. Chem. Soc. 2019, 141, 8684–8688.
[14]
a) M. R. A. Blomberg, T. Borowski, F. Himo, R. Z. Liao, P. E. M. Siegbahn, Chem. Rev. 2014, 114, 3601–3658;
b) M. G. Quesne, T. Borowski, S. P. de Visser, Chem. Eur. J. 2016, 22, 2562–2581.
[15]
a) A. Wójcik, E. Broclawik, P. E. M. Siegbahn, T. Borowski, Biochemistry 2012, 51, 9570–9580;
b) E. A. C. Bushnell, G. B. Fortowsky, J. W. Gauld, Inorg. Chem. 2012, 51, 13351–13356;
c) S. P. de Visser, Chem. Commun. 2007, 171–173;
d) X. Wang, H. Su, Y. Liu, Phys. Chem. Chem. Phys. 2017, 19, 7668–7677;
e) H. Su, X. Sheng, W. Zhu, G. Ma, Y. Liu, ACS Catal. 2017, 7, 5534–5543;
f) X. Song, J. Lu, W. Lai, Phys. Chem. Chem. Phys. 2017, 19, 20188–20197;
g) N. Zeb, M. H. Rashid, M. Q. E. Mubarak, S. Ghafoor, S. P. de Visser, J. Inorg. Biochem. 2019, 198, 110728.
[16] D. Kumar, W. Thiel, S. P. de Visser, J. Am. Chem. Soc. 2011, 133, 38693882.
[17]
a) M. A. Rude, T. S. Baron, S. Brubaker, M. Alibhai, S. B. Del Cardayre, A. Schirmer, Appl. Environ. Microbiol. 2011, 77, 1718–1727;
b) Y. Liu, C. Wang, J. Yan, W. Zhang, W. Guan, X. Lu, S. Li, Biotechnol. Biofuels 2014, 7, 28–40;
c) J. Belcher, K. J. McLean, S. Matthews, L. S. Woodward, K. Fisher, S. E. J. Rigby, D. R. Nelson, D. Potts, M. T. Baynham, D. A. Parker, D. Leys, A. W. Munro, J. Biol. Chem. 2014, 289, 6535–6550;
d) A. Dennig, M. Kuhn, S. Tassoti, A. Thiessenhusen, S. Gilch, T. Bülter, T. Haas, M. Hall, K. Faber, Angew. Chem. Int. Ed. 2015, 54, 8819–8822; Angew. Chem. 2015, 127, 8943–8946; e) J. L. Grant, C. H. Hsieh, T. M. Makris, J. Am. Chem. Soc. 2015, 137, 4940–4943.
[18]
a) B. Meunier, S. P. de Visser, S. Shaik, Chem. Rev. 2004, 104, 3947–3980;
b) W. Nam, Acc. Chem. Res. 2007, 40, 522–531;
c) Y. Watanabe, H. Nakajima, T. Ueno, Acc. Chem. Res. 2007, 40, 554–562;
d) J. Rittle, M. T. Green, Science 2010, 330, 933–937.

[19] A. S. Faponle, M. G. Quesne, S. P. de Visser, Chem. Eur. J. 2016, 22, 54785483.
[20] M. Pickl, S. Kurakin, F. G. CantfflReinhard, P. Schmid, A. Pçcheim, C. K. Winkler, W. Kroutil, S. P. de Visser, K. Faber, ACS Catal. 2019, 9, 565–577.
[21]
a) E. Godfrey, C. S. Porro, S. P. de Visser, J. Phys. Chem. A 2008, 112, 2464–2468;
b) A. R. Diebold, C. D. Brown-Marshall, M. L. Neidig, J. M. Brownlee, G. R. Moran, E. I. Solomon, J. Am. Chem. Soc. 2011, 133, 18148–18160;
c) G. Dong, S. Shaik, W. Lai, Chem. Sci. 2013, 4, 36243635;
d) A. Wójcik, M. Radon´, T. Borowski, J. Phys. Chem. A 2016, 120, 1261–1274;
e) A. Timmins, N. J. Fowler, J. Warwicker, G. D. Straganz, S. P. de Visser, Front. Chem. 2018, 6, 513.
[22]
a) D. Kumar, H. Hirao, L. Que, Jr., S. Shaik, J. Am. Chem. Soc. 2005, 127, 8026–8027;
b) S. P. de Visser, Angew. Chem. Int. Ed. 2006, 45, 17901793; Angew. Chem. 2006, 118, 1822–1825;
c) C. Geng, S. Ye, F. Neese, Angew. Chem. Int. Ed. 2010, 49, 5717–5720; Angew. Chem. 2010, 122, 5853–5856;
d) S. Ye, F. Neese, Proc. Natl. Acad. Sci. USA 2011, 108, 1228–1233;
e) L. Bernasconi, E. J. Baerends, J. Am. Chem. Soc. 2013, 135, 8857–8867;
f) A. Timmins, M. G. Quesne, T. Borowski, S. P. de Visser, ACS Catal. 2018, 8, 8685–8698.
[23]
a) S. Shaik, D. Kumar, S. P. de Visser, J. Am. Chem. Soc. 2008, 130, 1012810140;
b) M. G. Quesne, D. Senthilnathan, D. Singh, D. Kumar, P. Maldivi, A. B. Sorokin, S. P. de Visser, ACS Catal. 2016, 6, 2230–2243.
[24]
a) H. Hirao, K. Morokuma, J. Am. Chem. Soc. 2011, 133, 14550–14553;
b) S. C. Peck, H. A. Cooke, R. M. Cicchillo, P. Malova, F. Hammerschmidt, S. K. Nair, W. A. van der Donk, Biochemistry 2011, 50, 6598–6605;
c) L. Du, J. Gao, Y. Liu, D. Zhang, C. Liu, Org. Biomol. Chem. 2012, 10, 10141024.
[25] M. Lundberg, T. Kawatsu, T. Vreven, M. J. Frisch, K. Morokuma, J. Chem. Theory Comput. 2009, 5, 222–234.
[26]
a) R. Latifi, M. Bagherzadeh, S. P. de Visser, Chem. Eur. J. 2009, 15, 66516662;
b) R. N. Manna, T. Malakar, B. Jana, A. Paul, ACS Catal. 2018, 8, 10043–10050;
c) S.]lvarez-Barcia, J. Kästner, J. Phys. Chem. B 2017, 121, 5347–5354.
[27]
a) X.-X. Li, V. Postils, W. Sun, A. S. Faponle, M. Solà, Y. Wang, W. Nam, S. P. de Visser, Chem. Eur. J. 2017, 23, 6406–6418;
b) G. Mukherjee, A. Alili, P. Barman, D. Kumar, C. V. Sastri, S. P. de Visser, Chem. Eur. J. 2019, 25, 5086–5098;
c) P. Barman, F. G. Cantffl Reinhard, U. K. Bagha, D. Kumar, C. V. Sastri, S. P. de Visser, Angew. Chem. Int. Ed. 2019, 58, 1063910643; Angew. Chem. 2019, 131, 10749–10753.
[28]
a) A. Timmins, M. Saint-AndrØ, S. P. de Visser, J. Am. Chem. Soc. 2017, 139, 9855–9866;
b) A. S. Faponle, F. P. Seebeck, S. P. de Visser, J. Am. Chem. Soc. 2017, 139, 9259–9270.
[29]
a) Gaussian 09 (Revision C.01), M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar,

J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian, Inc., Wallingford CT, 2010;
b) F. Neese, WIREs Comput. Mol. Sci. 2018, 8, e1327.
[30]
a) S. Kumar, A. S. Faponle, P. Barman, A. K. Vardhaman, C. V. Sastri, D. Kumar, S. P. de Visser, J. Am. Chem. Soc. 2014, 136, 17102–17115;
b) F. G. Cant fflReinhard, A. S. Faponle, S. P. de Visser, J. Phys. Chem. A 2016, 120, 9805–9814.
[31]
a) A. D. Becke, J. Chem. Phys. 1993, 98, 5648–5652;
b) C. Lee, W. Yang, R. G. Parr, Phys. Rev. B 1988, 37, 785–789.
[32]
a) P. J. Hay, W. R. Wadt, J. Chem. Phys. 1985, 82, 270–283;
b) W. J. Hehre, R. Ditchfield, J. A. Pople, J. Chem. Phys. 1972, 56, 2257–2261.
[33] J. Tomasi, B. Mennucci, R. Cammi, Chem. Rev. 2005, 105, 2999–3009.
[34]
a) K. Cheaib, M. Q. E. Mubarak, K. SØnØchal-David, C. Herrero, R. Guillot, M. ClØmancey, J.-M. Latour, S. P. de Visser, J.-P. Mahy, F. Banse, F. Avenier, Angew. Chem. Int. Ed. 2019, 58, 854–858; Angew. Chem. 2019, 131, 864868;
b) S. Pattanayak, F. G. Cant fflReinhard, A. Rana, S. S. Gupta, S. P. de Visser, Chem. Eur. J. 2019, 25, 8092–8104.
[35]
a) H. Paulsen, L. Duelund, H. Winkler, H. Toftlund, A. X. Trautwein, Inorg. Chem. 2001, 40, 2201–2203;
b) M. Güell, M. Solà, M. Swart, Polyhedron 2010, 29, 84–93;
c) M. Güell, J. M. Luis, M. Solà, M. Swart, J. Phys. Chem. A 2008, 112, 6384–6391.
[36] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, J. Chem. Phys. 2010, 132, 154104.
[37]
a) N. C. Handy, A. J. Cohen, Mol. Phys. 2001, 99, 403–412;
b) J. P. Perdew, K. Burke, M. Ernzerhof, Phys. Rev. Lett. 1996, 77, 3865–3868.

Chapter 4 Electrostatic Perturbations from the Protein Affect C–H Bond Strengths of the Substrate and Enable Negative Catalysis in the TmpA Biosynthesis Enzyme

[2]Yen-Ting Lin,[a,b] Hafiz Saqib Ali,[c] and Sam P. de Visser*[a,b]

[a] Mr Y.-T. Lin, Dr. H. S. Ali, Dr S. P. de Visser
Manchester Institute of Biotechnology, The University of Manchester, 131 Princess Street, Manchester, M1 7DN, United Kingdom
Email: yen-ting.lin@manchester.ac.uk
Email: hafizsaqib.ali@manchester.ac.uk

[b] Mr Y.-T. Lin, Dr S. P. de Visser
Department of Chemical Engineering and Analytical Science, The University of Manchester, Oxford Road, Manchester, M13 9PL, United Kingdom
Email: sam.devisser@manchester.ac.uk

[c] Dr. H. S. Ali
Department of Chemistry, The University of Manchester, Oxford Road, Manchester, M13 9PL, United Kingdom

DOI: 10.1002/chem.202100791

Supporting information for this article is given via a link: https://pubmed.ncbi.nlm.nih.gov/33978257/

4.1 Abstract

The nonheme iron dioxygenase 2-(trimethylammonio)-ethylphosphonate dioxygenase (TmpA) is an enzyme involved in the regio- and chemoselective hydroxylation at the C^1-position of the substrate as part of the biosynthesis of glycine betaine in bacteria and carnitine in humans. To understand how the enzyme avoids breaking the weak C^2–H bond in favor of C^1-hydroxylation, we set up a cluster model of 242 atoms representing the first and second-coordination sphere of the metal center and substrate binding pocket and investigated possible reaction mechanisms of substrate activation by an iron(IV)-oxo species by density functional theory methods. In agreement with experimental product distributions, the calculations predict a favorable C^1-hydroxylation pathway. The calculations show that the selectivity is guided through electrostatic perturbations inside the protein from charged residues, external electric fields and electric dipole moments. In particular, charged residues influence and perturb the homolytic bond strength of the C^1–H and C^2–H bonds of the substrate, and strongly strengthens the C^2–H bond in the substrate-bound orientation.

[2] **Author contribution statement:** This is the original reference where the paper is published. Yen-Ting, Lin, the Majority of contribution is all DFT calculations and data analysis with the help of MD simulation by Dr Hafiz Saqib Ali, guided by Dr Samuel De Visser on all the directions of the project.

4.2 Introduction

Metalloenzymes are common in biology and particularly mononuclear iron is used as the cofactor in the active sites of many mono-oxygenases and dioxygenases. These oxygen-activating enzymes include heme enzymes such as the cytochromes P450, heme peroxidases and catalases with functions related to detoxification, biosynthesis and metabolism in the human body.[1] In addition, there are a large group of nonheme iron dioxygenases that lack the heme co-factor and bind iron directly to the protein via the side chains of polar amino acid residues, usually His, Asp or Glu.[2] The nonheme iron dioxygenases have important roles in biosystems ranging from the biosynthesis of natural products, for example plant hormones like flavonols and antibiotics in bacteria.[3] In humans, the regulation of the cysteine concentration in the body is performed by several nonheme iron dioxygenases that include the cysteine dioxygenase group of enzymes.[4]

The nonheme iron dioxygenases; therefore, are highly versatile and catalyze a range of chemical reactions, which interestingly often proceeds with high stereo- and chemoselectivity. How these enzymes perform their highly selective reactivity is still a matter of great controversy and debate and understanding these enzymatic selectivities may enable the use of these systems in biotechnology and bioengineering.

In particular, many nonheme iron dioxygenases involved in biosynthesis processes perform reactions under negative catalysis. Thus, negative catalysis is defined as a chemical process where an unwanted reaction is prevented in favour of an alternative pathway.[5] Thermodynamically, the protein blocks the pathway with largest driving force that would normally have the lowest rate-constant in favour of an alternative pathway. How these enzymes accomplish negative catalysis is a matter of debate.

More specifically, Negative catalysis is an opposite chemistry phenomenon of the Bell-Evans-Polanyi principle, which was developed and discovered, what is called “positive catalysis”, then published in 1924. [41,42,43, 44, 45] The principle states that the activation energy of the chemical reaction is in proportion to its driving force. Hence, if the response follows this theory, the most favourable reaction in the competing enzymatic reaction mechanism, chemical catalysis or bio-catalysis, should be the pathway that has the most exothermic reaction path. Nevertheless, in many enzymatic reaction mechanisms, the enzyme will not follow the Bell-Evans-Polanyi principle [5]; instead, the enzyme will choose the least exothermic reaction pathway, to proceed with the reaction, which is defined as negative catalysis.

The difference between negative catalysis and positive catalysis might be attributed to the negative catalysis reaction mechanism because the help of the substrate positioning and the binding of the protein matrix empowers the enzyme to activate the reaction to proceed and produce the desired product.

Nonheme iron dioxygenases usually bind the iron to the protein through interactions with the side chains of His, Asp and Glu residues in a facial 2-His/1-carboxylate orientation.[6] Typically, α-ketoglutarate (αKG, also called 2-oxoglutarate) is used as a co-substrate in non-heme iron dioxygenases, which binds to the iron(II) center directly. In a reaction with dioxygen on the iron center the αKG is converted into succinate and CO_2 and an iron(IV)-oxo species is formed. The high-valent iron(IV)-oxo species has been trapped and characterized for several nonheme iron dioxygenases with various spectroscopic techniques.[7] Subsequent studies showed it to be an efficient oxidant that triggers aliphatic hydroxylation reactions of a substrate.[8] Moreover, it is the active species that reacts with substrate through oxygen atom transfer in the form of substrate hydroxylation or epoxidation, although also substrate desaturation and ring-closure processes have been reported.[9]

Many reactions catalyzed by the nonheme iron dioxygenases have high substrate selectivity and specificity. Thus, the prolyl-4-hydroxylases in the human body enantio- and regioselectively hydroxylate a proline residue in an amino acid chain to *R*-4-hydroxyproline, which is an important precursor to collagen formation.[10] Furthermore, several nonheme iron dioxygenases activate an arginine residue, whereby the viomycin biosynthesis enzyme C (VioC) stereoselectively gives *S*-3-hydroxyarginine products,[11] whereas the naphthyridinomycin biosynthesis enzyme chemoselectively desaturates its C^4–C^5 bond.[9e]

Another interesting member of the α-ketoglutarate-dependent nonheme iron dioxygenases is the recently discovered 2-(trimethylammonio)ethylphosphonate dioxygenase (TmpA) enzyme. It selectively hydroxylates 2-(trimethylammonio)-ethylphosphonate at the C^1-position and its product is fed into the diiron oxygenase TmpB that performs an oxidative C–P bond cleavage to form glycine betaine and phosphate.[12] A crystallographic structure of substrate bound TmpA was resolved (see **Figure 4.1**) that binds substrate tightly in the active site with hydrogen bonding interactions from the side chains of Asn_{201}, Tyr_{203}, Asn_{286}, Arg_{288} to the phosphate group and hence keeping it in a tight conformation. The iron is bound to the amino acid residues His_{198}, Asp_{200} and His_{341} in a facial orientation with the carboxylic acid group of α-ketoglutarate trans to the His_{198} and Asp_{200} groups.

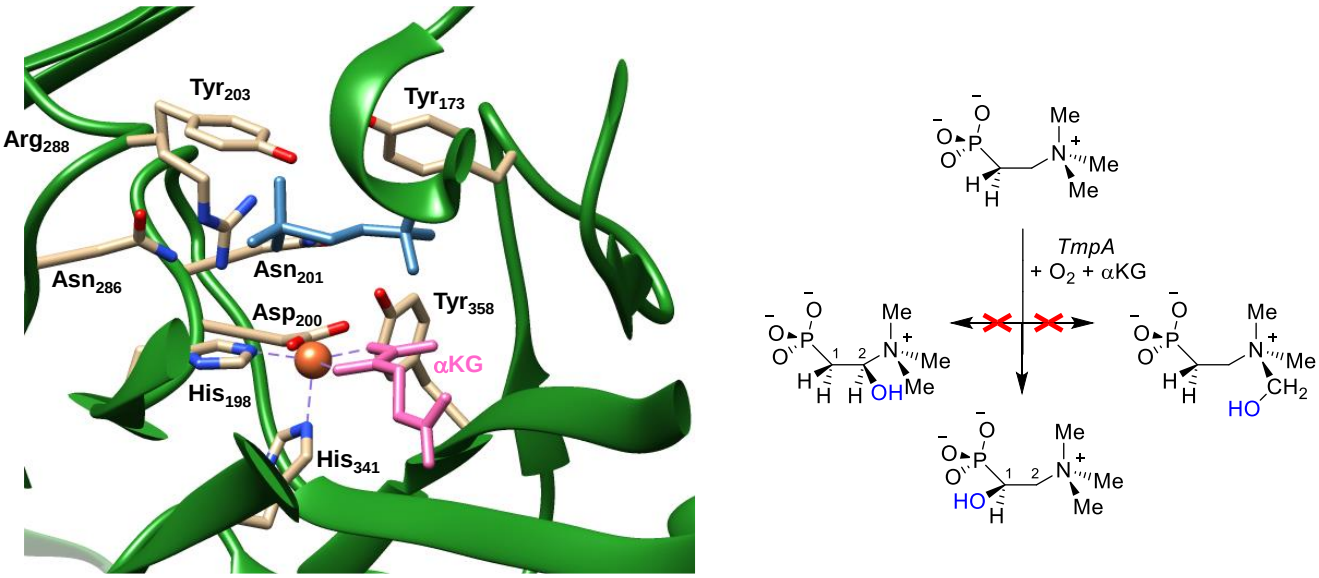
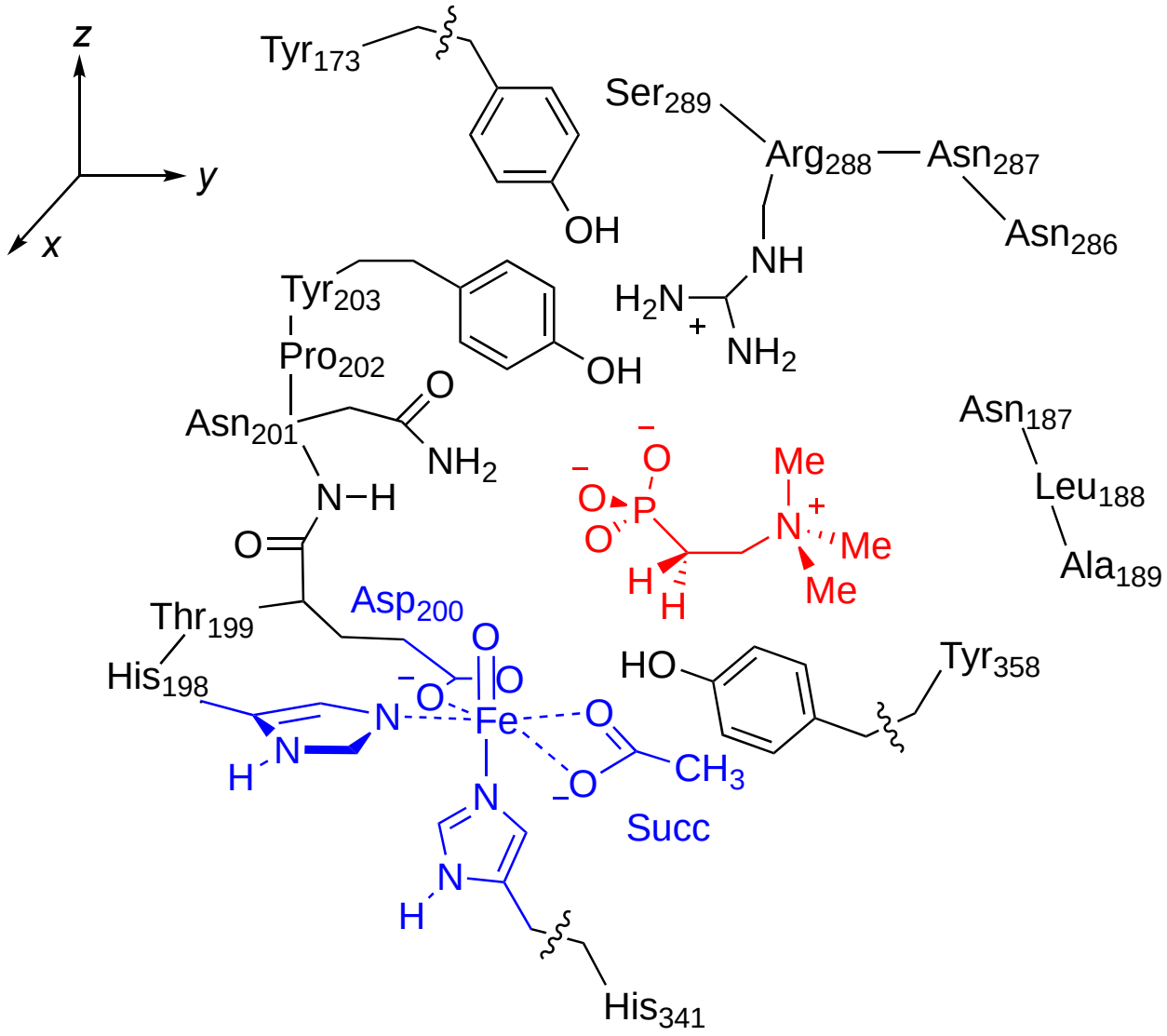


Figure 4.1 Extract of the active site of the substrate (2-(trimethylammonio)ethylphosphonate) bound structure of TmpA as taken from the 6NPC pdb file. The right-hand-side gives the three possible oxygen atom transfer reactions in TmpA leading to C^1-hydroxylation, C^2-hydroxylation or *N*-methyl hydroxylation.

Interestingly, the catalytic reaction performed by TmpA shows similarity to the human enzyme γ-butyrobetaine (γBB) hydroxylase (BBOX), which hydroxylates 2-(trimethylammonio)-ethylphosphonate in the final step of the carnitine biosynthesis pathway. Carnitine is an important natural product involved in the transport of long chain fatty acids to mitochondria in eukaryotes.[13] The experimental studies; however, were unable to ascertain whether C^1-hydroxylation or *N*-methyl hydroxylation of the substrate takes place in either BBOX. Moreover, technically the C^2–H bond should be weaker than the C^1–H or the methyl C–H bond strengths; yet those groups are not activated by TmpA. As limited experimental data is available on TmpA due to its fast reactivity and no computational studies have been reported either, we decided to pursue a mechanistic study with the aim to understand how TmpA avoids hydroxylation at the C^2-position in favor of C^1-hydroxylation. In particular, we explored the catalytic reaction mechanism of oxygen atom transfer from a high-valent iron(IV)-oxo species to substrate in TmpA and study the competitive mechanisms for C^1-hydroxylation, C^2-hydroxylation and *N*-methyl hydroxylation. The work shows that even though all three C–H bonds are in accessible positions by the iron(IV)-oxo species, as a matter of fact, only the C^1–H hydrogen atom abstraction gives low energy barriers and will be the dominant reaction pathway. This selectivity is determined by substrate and oxidant binding and positioning in the active site and charged protein residues incur perturbations that strengthen the C^2–H bond dramatically even though in the gas-phase the C^2–H bond is weakest.

4.3 Results and Discussion

We created a 242 atom cluster model of an iron(IV)-oxo reactant of TmpA as described in **Figure 4.2**. The model set-up follows previous experience of us and others with enzymatic reaction mechanism studies,[14] where we used density functional theory (DFT) cluster models. These models incorporate key parts of the first- and second-coordination sphere of the metal center, including the hydrogen bonding network and charged amino acid residues and all calculated with density functional theory. Those procedures have been extensively used by us and others for studies on metalloenzymes;[15] and give generally highly accurate representations of real systems. In particular, large cluster models of 100s of atoms have been shown to reproduce experimental product distributions and selectivities of enzymatic systems well as they treat the first and second coordination sphere of oxidant and substrate with high-level computational methods. A TmpA active site cluster model was created based on the crystal structure coordinates deposited in the 6NPC protein databank (pdb) file,[12,16] which is a substrate (2-(trimethylammonio)ethylphosphate) and α-ketoglutarate bound nonheme iron dioxygenase structure. The structure was subsequently prepared by adding hydrogen atoms and solvent and a molecular dynamics simulation was performed. The MD simulation showed little movement of the protein residues and a highly rigid substrate-binding pocket. Next, we selected an active site model from the last structure in the MD simulation, which is based on the first and second coordination sphere of the iron and substrate, see **Figure 4.2**, where key hydrogen bonding and electrostatic interactions to the substrate are included. In the model we truncated succinate (Succ) to acetate and the axial histidine residue (His_{341}) of the iron was shortened to methylimidazole. The other protein ligands of the metal were part of the peptide chain His_{198}-Thr_{199}-Asp_{200}-Asn_{201}-Pro_{202}-Tyr_{203} that was included in the model whereby Thr_{199} was shortened to Gly. In addition, part of the substrate-binding pocket was built into the model, namely the residues Tyr_{173}, the chain Asn_{187}-Leu_{188}-Ala_{189}, the chain Asn_{286}-Asn_{287}-Arg_{288}-Ser_{289} and Tyr_{358}. The Leu_{188} and Asn_{287} residues were truncated to Gly, while the Tyr_{172} and Tyr_{358} residues were mimicked as p-cresol. The overall model is shown in **Figure 4.2** and has a total of 242 atoms and is overall charge neutral.



1,3,5,7R1

Figure 4.2 Cluster model of the reactant complex (**R1**) is used in the calculations reported here. Wiggly lines identify where a covalent bond was cut from the crystal structure coordinates and a hydrogen link-atom inserted. Atoms in blue and red represent the minimal model (**R2**). Succ stands for succinate.

The cluster models were geometry optimized in the lowest energy singlet, triplet, quintet and septet spin states and the optimized geometries are shown in **Figure 4.3**. The quintet spin state (^{5}**R1**) is the ground state at the UB3LYP-D3/BS2//UB3LYP/BS1+ZPE level of theory (ZPE stands for zero-point energy), while all other spin states are well higher in energy by more than 15 kcal mol^{-1}. Based on these spin state orderings and relative energies, we expect the quintet spin state to be the dominant spin surface with little involvement of the other spin states during the reaction. Our result of a quintet spin ground state matches the experimental Mössbauer spectroscopy results on analogous nonheme iron/αKG-dependent dioxygenases that are typically characterized as a quintet spin state.[7,17] Moreover, previous computational studies on analogous enzymatic nonheme iron(IV)-oxo species also found a quintet spin ground state well separated from the other spin states.[18]

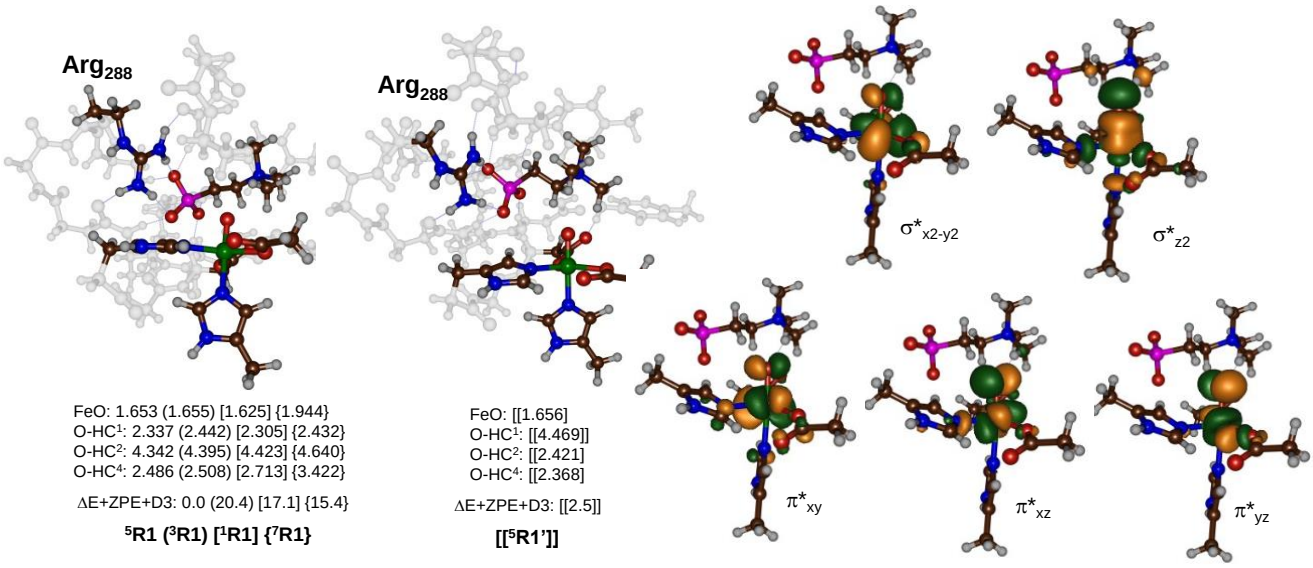


Figure 4.3 Optimized iron(IV)-oxo reactant complexes in the singlet, triplet, quintet and septet spin states with bond lengths in Ångströms and relative energies in kcal mol^{-1}. Energies are at UB3LYP-D3/BS2//UB3LYP/BS1 level of theory and include solvent and zero-point corrections. Relevant molecular orbitals of the iron(IV)-oxo species are given on the right.

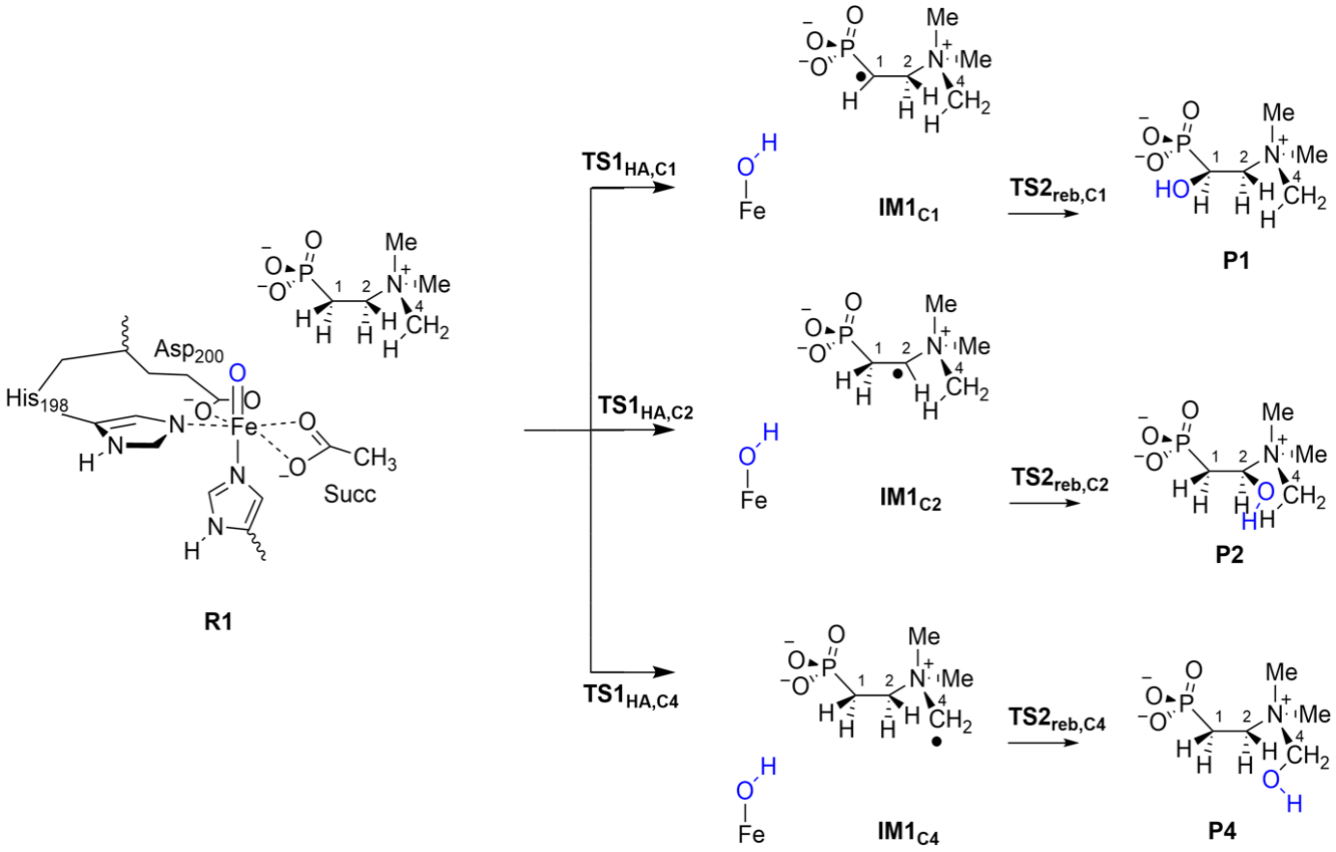
The high-lying singly occupied and low-lying virtual orbitals of ^{5}**R1** are shown on the right-hand-side of **Figure 4.3**. These are all metal type 3d-orbitals and their interactions with first-coordination sphere ligands. The lowest three of those are the π*-type orbitals: π^*_{xy} in the plane of His_{198}, Asp_{200} and the carboxylate of succinate, while the π^*_{xz} and π^*_{yz} orbitals are placed along the Fe–O axis and represent the π* antibonding interaction along that bond for the antibonding interaction of the $3d_{xz}/3d_{yz}$ on iron with the $2p_x/2p_y$ atomic orbital on oxygen. These three orbitals are singly occupied in ^{5}**R1** as well as the σ^*_{x2-}

$_{y2}$ orbital that is located in the plane of His_{198}, Asp_{200} and the carboxylate of succinate and gives an antibonding interaction between the metal $3d_{x2-y2}$ orbital with 2p orbitals of the equatorial ligands. The final metal 3d-orbital shown in **Figure 4.3** is the virtual σ^*_{z2} orbital for the antibonding interaction of the metal with the oxo and axial His_{341} groups.

Also, in **Figure 4.3** we give distances between the oxo group and nearby hydrogen atoms of C–H bonds of the substrate. Thus, in the quintet spin state the O–HC^1 distance is at 2.337 Å, while the nearest methyl O–HC^4 distance is 2.486 Å. By contrast, the O–HC^2 interaction is at a long distance of at least 4.342 Å. Therefore, based on the substrate positioning in the active site the most likely positions of the substrate that will be activated are the C^1- and methyl groups. We tested an alternative substrate binding position, where the 2-(trimethylammonio)ethylphosphonate is rotated so that the C^2–H groups are closest to the iron(IV)-oxo group (the O–HC^2 distance is 2.42 Å while the O–HC^1 distance is 4.47 Å in ^{5}**R1'**) and the phosphate group is involved in the same hydrogen bonding interactions. The optimized geometry of this structure (^{5}**R1'**); however, is higher in energy than ^{5}**R1** by 2.5 kcal mol^{-1} but shows similar structural features of the metal and its first-coordination sphere environment, see **Figure 4.2**. Consequently, the reaction mechanism with substrate was investigated for ^{5}**R1** and ^{5}**R1'**.

Thereafter, we investigated substrate activation by the iron(IV)-oxo species (^{5}**R1** and ^{5}**R1'**) and attempted a substrate hydroxylation at the C^1, C^2 and C^4-positions via the mechanisms shown in Scheme 1. Thus, pathway 1 covers the C^1–H activation via transition state **TS1**$_{HA,C1}$ to form an iron(III)-hydroxo radical intermediate (**IM1**$_{C1}$) followed by OH rebound via a transition state **TS2**$_{reb,C1}$ to form the C^1-hydroxylated product **P1**. Pathway 2 follows the hydrogen atom abstraction from the C^2-position via transition state **TS1**$_{HA,C2}$ to form the iron(III)-hydroxo complex **IM1**$_{C2}$. A radical rebound via transition state **TS2**$_{reb,C2}$ gives C^2-hydroxylated product **P2**. Finally, pathway 3 was studied for the methyl group activation of the trimethylammonio group via a hydrogen atom abstraction transition state **TS1**$_{HA,C4}$ to form the iron(III)-hydroxo radical intermediate **IM1**$_{C4}$. Again, an OH re-bound barrier via **TS2**$_{reb,C4}$ then gives the methyl hydroxylated product **P4**.

All three pathways shown in **Scheme 4.1** were studied with the DFT cluster model shown in **Figure 4.2** for the quintet spin state. The DFT calculated potential energy landscape of substrate activation by TmpA is shown in **Figure 4.2**. Constraint geometry scans for the hydrogen atom abstraction steps show little conformational changes for these steps (Supporting Information, Figure S5).



Scheme 4.1 Reaction mechanisms of substrate hydroxylation by TmpA calculated with labelling of the structures.

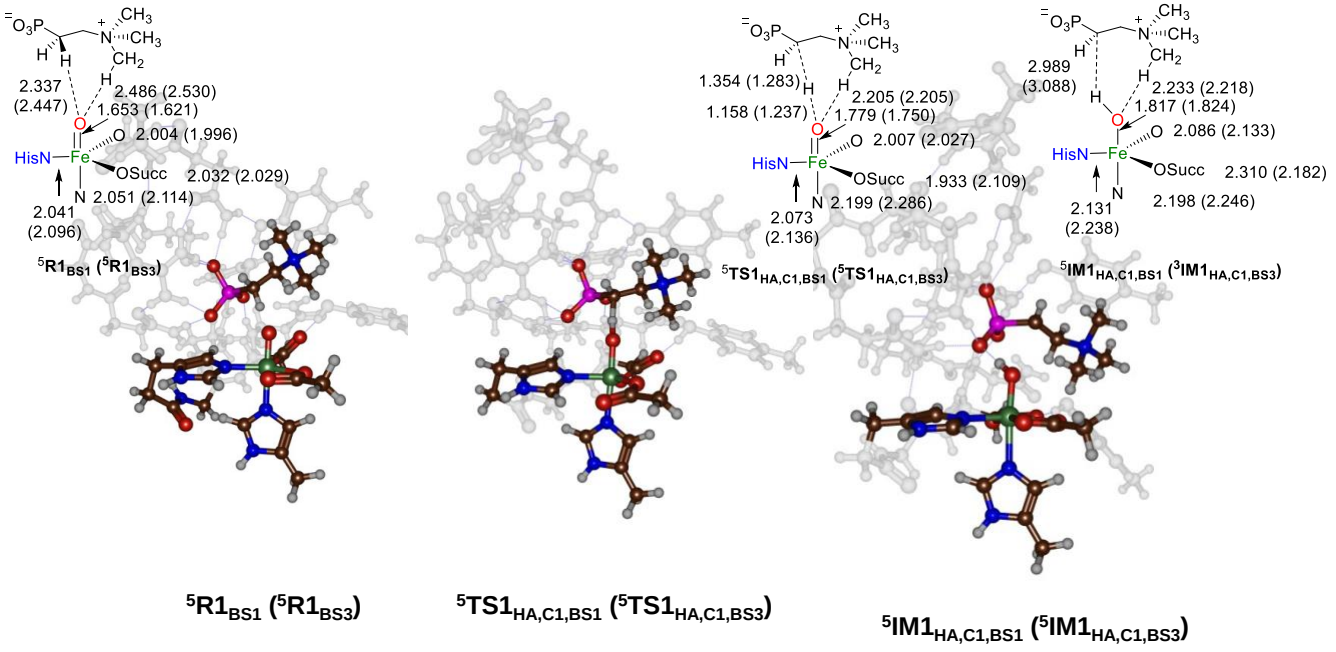


Figure S5: Comparison of UB3LYP/BS1 and UB3LYP/BS3 optimized geometries. Bond lengths in angstroms.

The C^1–H and C^2–H hydroxylation pathways have the hydrogen atom abstraction barrier as rate-determining, while the rebound barrier is much lower in energy. By contrast, these barriers are reversed for the C^4–H hydroxylation pathway that has a rate-determining OH rebound step.

The lowest hydrogen atom abstraction barrier is for abstraction of the C^1–H atom with a $\Delta E^{\ddagger}$+ZPE = 11.3 kcal mol^{-1}, while hydrogen atom abstraction from the methyl group has a barrier of $\Delta E^{\ddagger}$+ZPE = 15.0 kcal mol^{-1} and the C^2–H hydrogen atom abstraction barrier is $\Delta E^{\ddagger}$+ZPE = 19.2 kcal mol^{-1}. Interestingly, addition of dispersion corrections narrows the gap between the ^{5}**TS1**$_{HA,C1}$ and ^{5}**TS1**$_{HA,C4}$ transition states to 0.9 kcal mol^{-1}, while the ^{5}**TS1**$_{HA,C2}$ barrier is lowered to within 1.5 kcal mol^{-1}. Free energies are also shown in **Figure 4.4**, which follow the same trends as those found for the relative enthalpies. Nevertheless, regardless of the computational method, the ordering of the transition states stays the same. Therefore, the computation predicts dominant C^1–H hydrogen atom abstraction, although a minor amount of methyl and C^2-hydroxylation as minor products could be obtained as side products. Overall, the calculations predict the correct product distribution for TmpA enzymes in agreement with experimental observation. The hydrogen atom abstraction barriers are of the same order of magnitude as those found for related nonheme iron enzymes that perform C–H abstraction reactions with structurally similar transition states.[18,19]

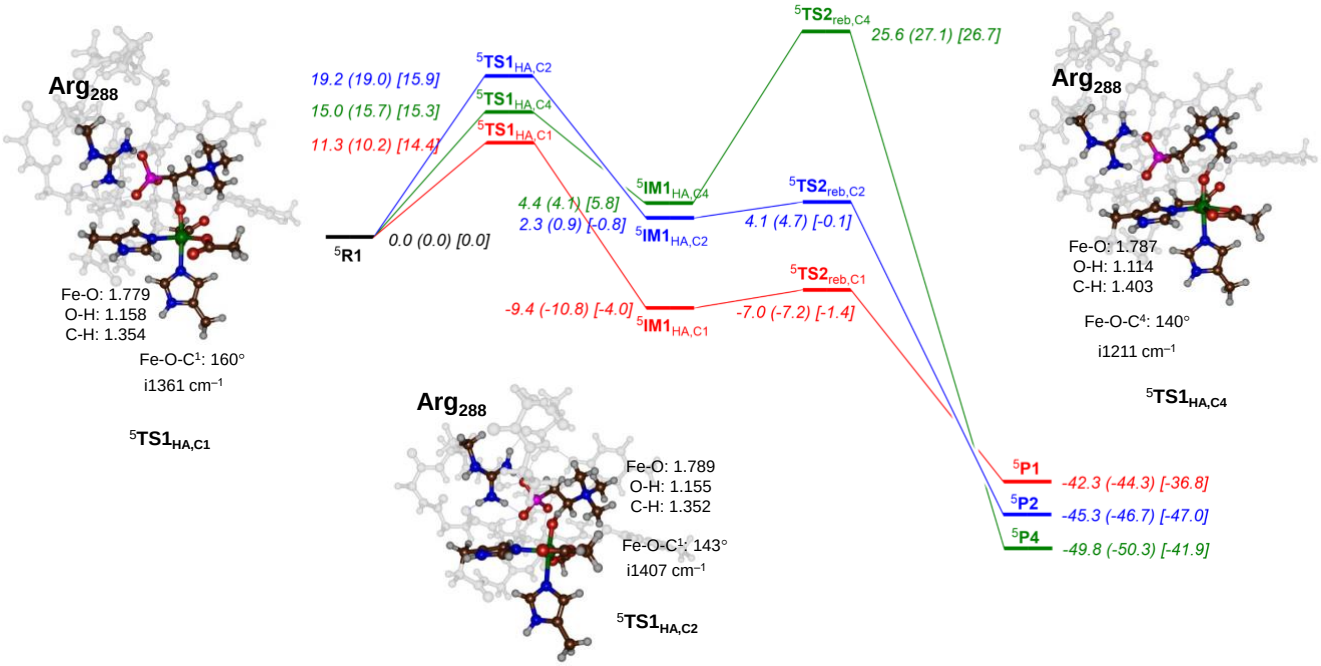


Figure 4.4 DFT calculated potential energy profile for substrate hydroxylation at the C^1, C^2 and methyl positions. Geometries optimized at UB3LYP/BS1 and energies reported are in kcal mol^{-1} obtained at ΔE_{BS2}+ZPE (ΔG_{BS2}) [ΔE_{BS2}+ZPE+D3]. Bond lengths are in Ångströms, bond angles in degrees and the imaginary frequency in cm^{-1}.

After the hydrogen atom abstraction the system relaxes to a radical intermediate; however, the ordering of the **IM1** structures is different from the hydrogen atom abstraction transition states: ^{5}**IM1**$_{HA,C1}$ is 9.4 kcal mol^{-1} lower in energy than ^{5}**R1**, while ^{5}**IM1**$_{HA,C2}$ and ^{5}**IM1**$_{HA,C4}$ are 2.3 and 4.4 kcal mol^{-1} above ^{5}**R1**. Subsequently, rebound barriers were searched for C–OH bond formation. The C^1 and C^2-hydroxylation pathways have a small rebound barrier with magnitude of only a few kcal mol^{-1} above the radical intermediate. For the C^1-pathway an alternative rebound **TS** was found for a $^5\pi$-pathway that has a barrier of 13.5 kcal mol^{-1} above ^{5}**IM1**$_{HA,C1}$. Interestingly, the C^4 pathway has a relatively large rebound barrier of 21.2 kcal mol^{-1} above ^{5}**IM1**$_{HA,C4}$.

Geometrically, the transition state geometries of ^{5}**TS1**$_{HA,C1}$, ^{5}**TS1**$_{HA,C2}$ and ^{5}**TS1**$_{HA,C4}$ are shown in **Figure 4.4**. All three transition states are product-like with the transferring hydrogen atom closer to the acceptor atom than the donor atom, i.e. in ^{5}**TS1**$_{HA,C1}$ the C–H and O–H distances are 1.354Å and 1.158 Å, whereas for ^{5}**TS1**$_{HA,C2}$ and ^{5}**TS1**$_{HA,C4}$ the values obtained for the C–H distances are 1.352/1.403 and for the O–H distances 1.155/1.114 Å, respectively. At the same time, in all **TS1** structures the Fe–O distance has elongated from 1.653 Å in ^{5}**R1** to 1.78 Å as a result of the change in molecular orbital occupation.

Electronically, in the transition state an electron transfer takes place from the substrate C–H bond into the σ^*_{z2} molecular orbital to create a radical intermediate with an electronic configuration of $\pi^*_{xy}{}^{\uparrow}$ $\pi^*_{xz}{}^{\uparrow}$ $\pi^*_{yz}{}^{\uparrow}$ $\sigma^*_{x2\text{-}y2}{}^{\uparrow}$ $\sigma^*_{z2}{}^{\uparrow}$ $\phi_{Sub}{}^{\downarrow}$. The latter orbital designates the singly occupied orbital on the substrate and its electron is antiferromagnetically coupled to the five up-spin electrons on the metal. An analysis of the group spin densities of ^{5}**TS1**$_{HA,C1}$, ^{5}**TS1**$_{HA,C2}$ and ^{5}**TS1**$_{HA,C4}$ indeed all give a spin density of 3.9 on iron and –0.5 on the substrate, as expected from the molecular orbital occupation. This type of electron transfer is designated a $^5\sigma$-pathway,[20] while the radical intermediate with electronic configuration $\pi^*_{xy}{}^2$ $\pi^*_{xz}{}^{\uparrow}$ $\pi^*_{yz}{}^{\uparrow}$ $\sigma^*_{x2\text{-}y2}{}^{\uparrow}$ $\phi_{Sub}{}^{\uparrow}$ passes the $^5\pi$-pathway.

We attempted to swap molecular orbitals for the radical intermediates to find the $^5\pi$-states, but during the SCF convergence they relaxed back to the $^5\sigma$-intermediates, which therefore are the lowest in energy.
The imaginary frequency in the transition state is large for all three hydrogen atom abstraction transition states and implicates that the reaction will incur a large kinetic isotope effect when the transferring hydrogen atom is replaced by a deuterium atom. These large values of imaginary frequencies are typical for hydrogen atom abstraction transition states and have been seen before for analogous reaction mechanisms.[21]

Moreover, these large imaginary frequencies in the transition states for the O–H–C stretch vibration implicate that hydrogen by deuterium substitution may result in a raise of the barriers and a change in kinetics through a kinetic isotope effect. We; therefore, calculated free energies for ^{5}**TS1**$_{HA,C1}$, ^{5}**TS1**$_{HA,C2}$ and ^{5}**TS1**$_{HA,C4}$ for the replacement of hydrogen atoms by deuterium. Despite the fact that the ^{5}**TS1**$_{HA,C1}$ barrier is raised by about 1.3 kcal mol^{-1} when the C^{1}–H bond is replaced by C^{1}–D, it actually does not change the ordering of the transition states. Moreover, a kinetic isotope effect of KIE = 8.9 is calculated with the classical Eyring method, while tunnelling corrections raise this to 12.4. Similar KIE values are obtained when the C^{2}–H bond is replaced by C^{2}–D.

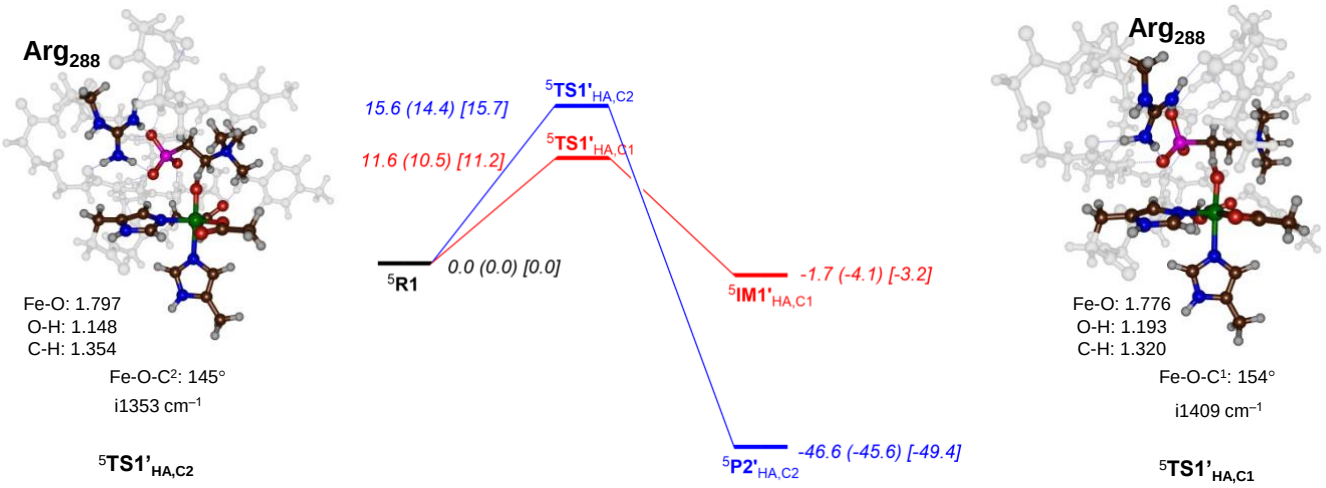


Figure 4.5 DFT calculated potential energy profile for substrate hydroxylation at the C^{1} and C^{2} positions for rotated substrate. Geometries optimized at UB3LYP/BS1 and energies reported are in kcal mol^{-1} obtained at ΔE_{BS2}+ZPE (ΔG_{BS2}) [ΔE_{BS2}+ZPE+D3]. Bond lengths are in Ångströms, bond angles in degrees and the imaginary frequency in cm^{-1}.

We also calculated the hydrogen atom abstraction transition states starting from ^{5}**R1'**, where the C^{2}–H positions of the substrate are closer to the iron(IV)-oxo and the results are given in **Figure 4.5**. Interestingly, the ^{5}**TS1'**$_{HA,C2}$ structure has an energy of $\Delta E^{\ddagger}$+ZPE = 15.6 kcal mol^{-1} and a geometry that is very close to its analogous structure ^{5}**TS1**$_{HA,C2}$ shown in **Figure 4.4**. In particular, the transferring hydrogen atom has C–H and O–H distances of 1.354 and 1.148 Å, which are within 0.01 Å from those for ^{5}**TS1**$_{HA,C2}$. In addition, the imaginary frequency represents a hydrogen transfer and has a mode of i1353 cm^{-1}.

A similarly close structure is obtained for ^{5}**TS1**$_{HA,C1}$ versus ^{5}**TS1'**$_{HA,C1}$, where both geometry optimizations started with different substrate orientation. As such during the geometry minimization the structure reoriented to an orientation close to ^{5}**TS1**$_{HA,C1}$. Relative energies of ^{5}**TS1**$_{HA,C1}$ and ^{5}**TS1'**$_{HA,C1}$ give them as the lowest energy barriers with values $\Delta E^{\ddagger}$+ZPE = 11.3 and 11.6 kcal mol^{-1}, respectively. Due to the reorientation of substrate the ^{5}**TS1'**$_{HA,C2}$ barrier is lowered with respect to the energy of ^{5}**TS1**$_{HA,C2}$ and is reduced from $\Delta E^{\ddagger}$+ZPE = 19.2 to 15.6 kcal mol^{-1} Nevertheless, regardless of the orientation of the substrate, the C^{2}–H hydrogen atom abstraction is well higher in energy than the C^{1}–H hydrogen atom abstraction, which will be the dominant reaction channel. Consequently, substrate binding is not expected to affect the selectivity of the reaction and the product distributions.

Despite the fact that the ^{5}**TS1'**$_{HA,C1}$ transition state relaxes to a radical intermediate, the ^{5}**TS1'**$_{HA,C2}$ barrier collapses to product ^{5}**P2'**$_{HA,C2}$ directly and no stable radical intermediate could be characterized. In the other substrate orientation (**Figure 4.4**) a stable radical intermediate was found, but with a small rebound barrier leading to products. As such in both substrate orientations the rate-determining step is hydrogen atom abstraction with fast and facile radical rebound to either the C^{1} or C^{2}-position.

The calculations shown in this work implicate; therefore, that regardless of the substrate orientations and binding positions the lowest energy barrier is always for C^{1}–H hydrogen atom abstraction *even when the C^{1}–H bond is further away than the C^{2}–H bond in the active site*. As such irrespective of substrate positioning the reaction will proceed toward dominant C^{1}–H hydrogen atom abstraction and give C^{1}-hydroxylated products and little or no C^{2}- or methyl hydroxylation can be expected, which matches the experimental product observations of this enzyme.

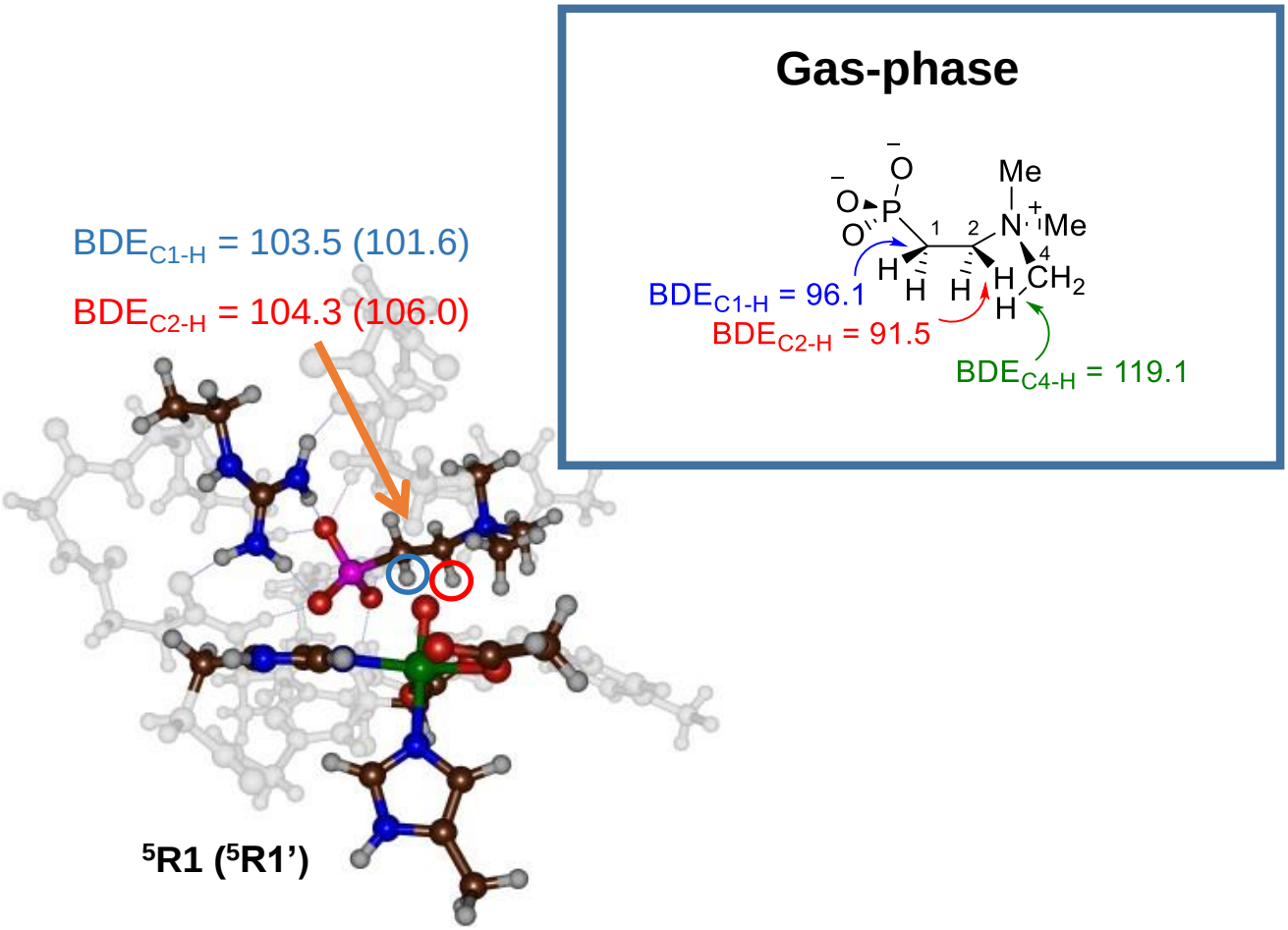


Figure 6.6 DFT calculated bond dissociation energies of selected C–H bonds in 2-(trimethylammonio)ethylphosphate in the gas-phase (inset) and inside the cluster models **R1** and **R1'**. Energies obtained at UB3LYP/6-311++G** level of theory with solvent and ZPE corrections included.

To understand the reactivity preference of C^1-hydroxylation over C^2- and methyl-hydroxylation, we calculated the bond dissociation energies (BDEs) for all C–H bonds of the substrate and the results are given in **Figure 6.6**. These BDE values were calculated for the substrate, i.e. 2-(trimethylammonio)-ethylphosphonate as the difference in energy between the substrate, an isolated hydrogen atom and the substrate with one hydrogen atom removed from either the C^1–H, C^2–H or methyl positions. The UB3LYP/6-311++G** calculated BDE values at $\Delta E+ZPE+E_{solv}$ are 96.1 kcal mol^{-1} for the C^1–H bond, 91.5 kcal mol^{-1} for the C^2–H bond and 119.1 kcal mol^{-1} for the CH_2–H bond. Therefore, the lowest BDE would correspond to the weakest bond of the substrate and hence is the C^2–H bond, yet it is not being activated in the TmpA enzyme during the substrate hydroxylation mechanism! On the other hand, the strongest bond in the substrate is the methyl C–H bond and it should be difficult to break that bond. Indeed, a high barrier for hydrogen atom abstraction from the methyl C–H bond is calculated for the enzymatic model (Figure 4). These gas-phase BDE calculations; therefore, predict that the order of the

hydrogen atom abstraction barriers of C^1–H versus C^2–H are opposite of that seen from the difference in BDE values. Clearly environmental effects in the enzyme, and well reproduced in our model, perturb the C^1–H and C^2–H bond strengths and lead to a reversal of the barriers. Thus, it should be easier to break the weak C^2–H bond than the stronger C^1–H bond, yet in the enzyme the C^1–H bond is preferentially broken. Clearly, substrate binding and positioning prefers C^1–H to C^2–H activation and leads to negative catalysis, whereby the weak C^2–H bond is not activated in favor of a stronger C^1–H bond.

Subsequently, we took the substrate-bound reactant complexes, namely ^{5}**R1** and ^{5}**R1'** and calculated the bond strength of the C^1–H and C^2–H groups by doing a single point energy calculation for **R1** and **R1'** with one hydrogen atom removed in the sextet spin state. This way, the BDEs of the C–H bonds inside the cluster model were calculated. For model **R1** we find $BDE_{C1\text{-}H}$ and $BDE_{C2\text{-}H}$ of 103.5 and 104.3 kcal mol^{-1}, whereas for model **R1'** they are 101.6 and 106.0 kcal mol^{-1}. As follows, in the protein matrix the C–H bond strengths of an arginine substrate are stronger, but the C^1–H bond becomes the weaker of the two C–H bonds. These results show that the electrostatic interactions of the protein matrix and oxidant on the substrate influences the individual C–H bond strengths and enables abstraction of the hydrogen atom that is stronger bound in the gas-phase and thereby triggers a negative catalysis.

To gain insight into the effect of the protein residues on the reaction mechanism and barriers in the large cluster model **R1**, we decided to explore a further truncated model. In particular, we investigated the minimal model **R2** that only considers the first-coordination metal ligands and the substrate. This hypothetical model of TmpA will help explain how stereochemical and electrostatic perturbations in the second-coordination sphere affect the transition states, relative energies and charge distributions. Subsequently, we calculated the C^1–H and C^2–H hydrogen atom abstraction pathways for model **R2**. Thus, without the protein present, we find barriers for ^{5}**TS1**$_{HA,C1,R2}$ and ^{5}**TS1**$_{HA,C2,R2}$ of $\Delta E^{\ddagger}$+ZPE = 17.8 and 7.6 kcal mol^{-1}, respectively, while they were 11.3 and 19.2 kcal mol^{-1} at the same level of theory for model **R1** (Figure 4). Therefore, removal of the protein environment and consequently perturbations relating to substrate positioning near the oxidant lowers the hydrogen atom abstraction barrier for the C^2–H group dramatically, whereas the C^1–H hydrogen atom abstraction barrier is raised by more than 6 kcal mol^{-1}. Moreover, under ideal circumstances, i.e. without protein present, the iron(IV)-oxo reactivity should give barriers that follow the order of the strength in C–H bond, i.e. BDE_{CH}, whereby the C^2–H hydrogen atom abstraction barrier is the lowest in energy. This is indeed the case with the minimal model, but not with the large cluster model.

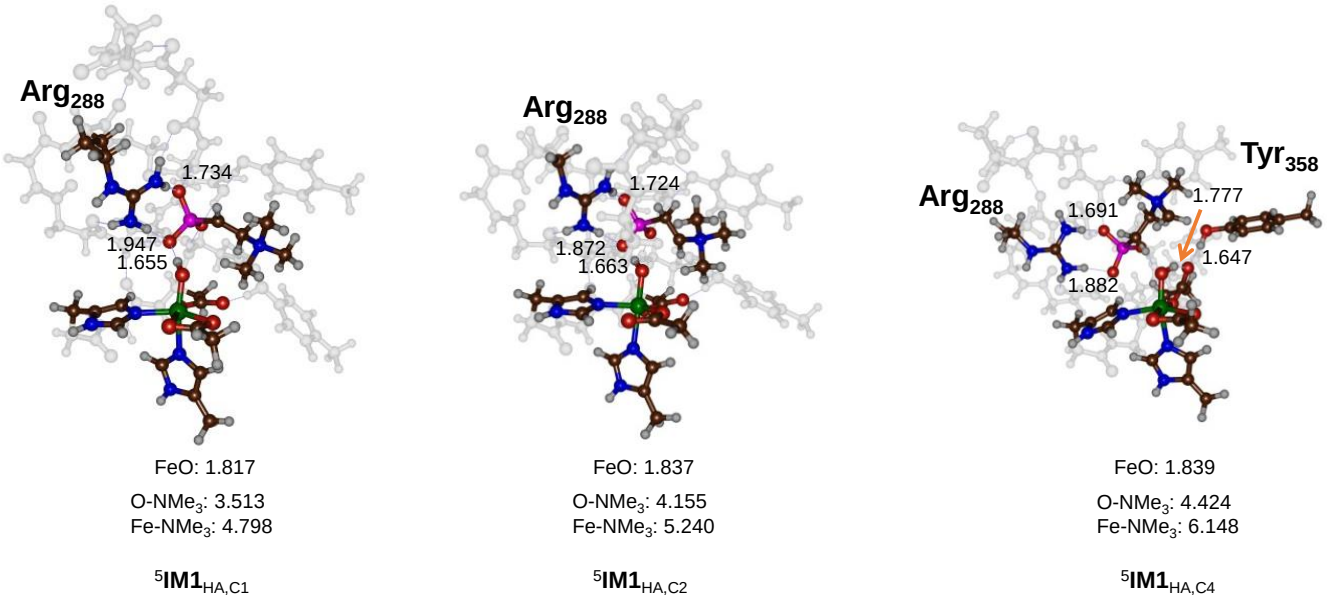


Figure 4.7 DFT calculated radical intermediates **IM1**$_{HA,C1}$, **IM1**$_{HA,C2}$ and **IM1**$_{HA,C4}$ as optimized at UB3LYP/BS1 in Gaussian-09 with bond lengths in Ångströms.

To understand, why the C^2-radical is destabilized while the C^1 radical is stabilized by the protein, we draw the three ^{5}**IM1** structures in **Figure 4.7**. As can be seen, the Fe–O distances are very close for the three structures and hence the differences do not appear to have to do with the first-coordination sphere of the metal. Also, in all three structures a double salt bridge between the Arg$_{288}$ group with the phosphate moiety of the substrate is in place with almost equal distances. The only major difference between the three structures relates to the position of the sp^2-hybridized carbon atom. In particular, the radical on C^2 gives an sp^2-hybridized carbon atom on the C^2-position of substrate that twists the trimethylamine group and moves it further away from the iron(III)-hydroxo group. In ^{5}**IM1**$_{HA,C2}$ the distance between the hydroxo and the trimethylamine group is 4.155 Å, while it is only 3.513 Å in ^{5}**IM1**$_{HA,C1}$. The shorter interaction between a cation and an anion will give larger stability due to favorable Coulomb interactions. In both ^{5}**IM1**$_{HA,C1}$ and ^{5}**IM1**$_{HA,C2}$ the iron(III)-hydroxo group forms a hydrogen bond with the phosphate group, whereas in ^{5}**IM1**$_{HA,C4}$ it is with the carboxylate group of succinate, which also interacts with the phenol group of Tyr$_{358}$. This interaction and the position of the hydroxyl group will make rebound difficult and hence a large rebound barrier for the C^4-pathway is found.

In a similar way to the BDE$_{CH}$ energies, we calculated the energy to form the O–H bond in the iron(III)-hydroxo complex, i.e. BDE$_{FeO–H}$. The BDE$_{FeO–H}$ was estimated from the difference in energy between an

Fe^{III}(OH) complex and an isolated hydrogen atom and an iron(IV)-oxo species for a small nonheme iron model complex and a value of 93.0 kcal mol^{-1} was obtained.[22] It was shown that the difference in energy between BDE_{CH} and $BDE_{FeO–H}$ should correspond to the energy difference between reactants and radical intermediates. Thus, for hydrogen atom abstraction from the C^1-position of 2-(trimethylammonio) ethylphosphonate the difference in BDE values gives –3.5 kcal mol^{-1}, while for hydrogen atom abstraction from the C^2-position it is –8.1 kcal mol^{-1}.

The former value is a slightly higher in energy than the difference between reactants and ^{5}**IM1**$_{HA,C1}$ of –9.4 kcal mol^{-1}. Hence, the protein has a small stabilization effect on the C^1-radical intermediate of 5.9 kcal mol^{-1}. On the other hand, the ^{5}**IM1**$_{HA,C2}$ structure is considerably less stable than predicted from thermochemical values for the bonds that are broken and formed. To be specific, the ^{5}**IM1**$_{HA,C2}$ is 11.7 kcal mol^{-1} less stable than predicted based on the thermochemistry. Therefore, the protein environment destabilizes the C^2–H hydrogen atom abstraction transition state as well as its subsequent radical intermediate ^{5}**IM1**$_{HA,C2}$.

To analyze the results and gain insight into how electrostatic perturbations from the protein guide and direct the substrate activation in TmpA, we did a thermodynamic and electric field effect analysis on the bond strengths of the structures. The DFT calculations presented in this work, in agreement with experimental observation, show that TmpA activates substrate preferably at the C^1-position. However, the calculations show that actually the C^2–H bond strength is weaker; yet, the energy to cleave this bond is well higher. We tested two substrate binding orientations, whereby in **R1** the C^1–H group is closest to the iron(IV)-oxo species, while in **R1'** the C^2–H group is closest.

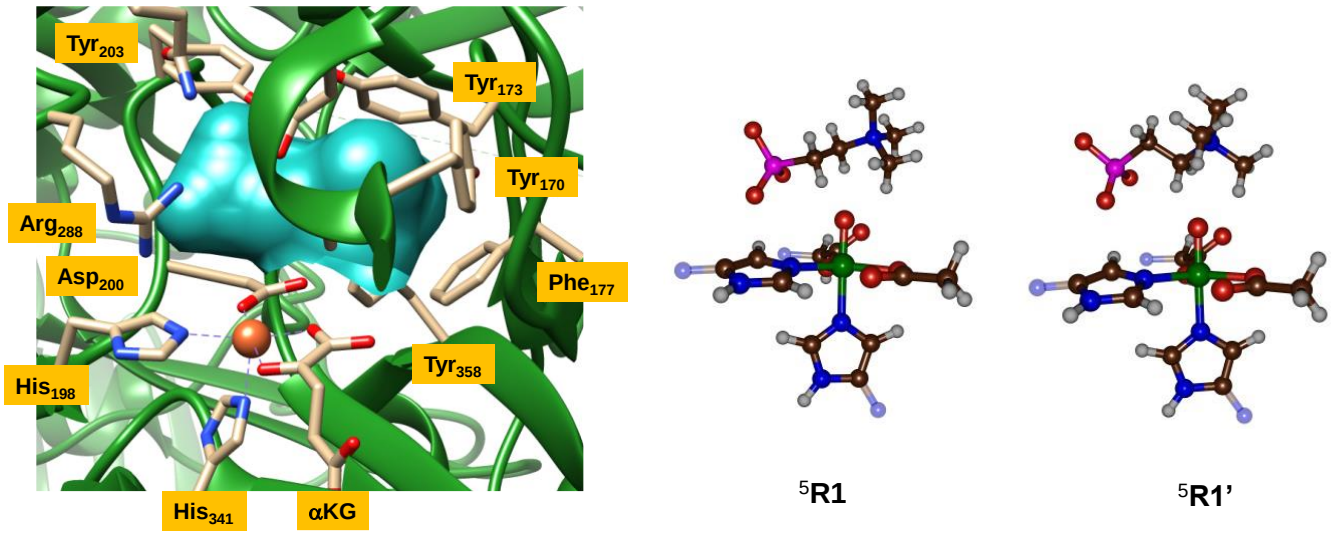


Figure 4.8 Space-filling model of the substrate-bound (**R1**) structure of TmpA as taken from the 6NPC pdb file. The right-hand side gives the first-coordination sphere extracts of **R1** and **R1'** as a comparison.

Despite the differences in orientation the obtained hydrogen atom abstraction barriers remain high in energy and do not change ordering. **Figure 4.8** displays the first-coordination sphere and substrate orientations of **R1** and **R1'**. The structural comparison of those geometries shows that the first-coordination sphere of the metal is virtually the same and essentially only the substrate orientation is swapped around the C^1 and C^2 positions. The space filling model on the left-hand-side; however, shows tight binding of the substrate inside the binding pocket with several aromatic residues, e.g. Tyr_{170}, Tyr_{173}, Phe_{177}, Tyr_{203} and Tyr_{358} side chains providing a cavity for the trimethylamine group to bind. The two substrate binding positions do not give major differences in hydrogen bonding interactions between substrate and protein and consequently the two structures are similar in energy and will bind equally strongly. Furthermore, the hydrogen atom abstraction barriers for substrate orientation **R1** and **R1'** are very similar and hence positioning of the substrate will not change the selectivity or the reaction rates. As the trimethylammonio group of substrate is surrounded by aromatic amino acids, while the phosphate group is surrounded by polar residues including a positively charged Arg residue, we reasoned that the protein must induce a local electric field effect or dipole moment that can affect charge distributions and spin populations.

How external perturbations affect the selectivity of the reaction and particularly whether an electric dipole moment or electric field effect can influence the strength of a C–H bond in a substrate, was further explored by calculating the substrate BDE values for the C^1–H and C^2–H positions with an external electric field applied. In particular, we investigated the isolated substrate in the gas-phase and calculated the BDE_{CH} values with an external electric field applied. These electric field calculations can mimic the

charge distributions of protein and highlight the nature and direction needed for a chemical reaction change.[23] Specifically, it was shown that bond dissociation energies correlate with the change in polarizability volume between the molecule and isolated fragments and consequently an electric field effect should influence relative bond strengths.[22] Previously, we showed that an external electric field can affect the charge distributions in molecules and enzymatic models and even affect bond dissociation energies.[24] Thus, it was shown for cytochrome P450 Compound I that a field along the positive Fe–S axis in the active site gave more sulfur radical-character, while a field in the opposite direction gave more porphyrin radical-character.[25] As such the electronic configuration of Compound I was shown to change upon electric field perturbations that also affected its electron affinity and consequently its ability to react with substrates. In particular, the electric field effects affected the chemoselectivity of substrate hydroxylation versus epoxidation of a model complex and the product distributions could be guided by the direction and magnitude of the electric field.

A similar change in charge distribution was seen in cytochrome *c* peroxidase, where one specific electric field perturbation gave a Compound I electronic configuration with a porphyrin radical while with a field in the reverse direction a protein (Trp) radical was found.[26] Moreover, in the nonheme iron halogenase a local electric field of charged amino acid residues was proposed to trigger a substrate halogenation reaction, whereas in the absence of the field a substrate hydroxylation reaction was predicted.[27] More recent work on the nonheme iron dioxygenase enzyme OrfP showed that the C–H bond strengths of the substrate were different in the gas-phase than inside a protein matrix. In particular the dipole moment and charge distributions of active site residues affected the C–H bond strengths of the substrate in the binding pocket and triggered a chemoselective reaction mechanism.[28] Experimental studies using electronic absorption methods showed that an external electric field affects charge distributions and influences hydrogen bonding interactions; thereby, moving protons from a donor to an acceptor group or *vice versa*.[29] To test if something similar would apply to the substrate inside the TmpA enzyme structure, we decided to investigate the C–H bond strengths of the substrate in different environments and under external electric field effects. These types of studies are not computational curiosities as various experimental groups have latched proteins on electrode surfaces and studied the subsequent reaction mechanisms under an applied electric field.[30]

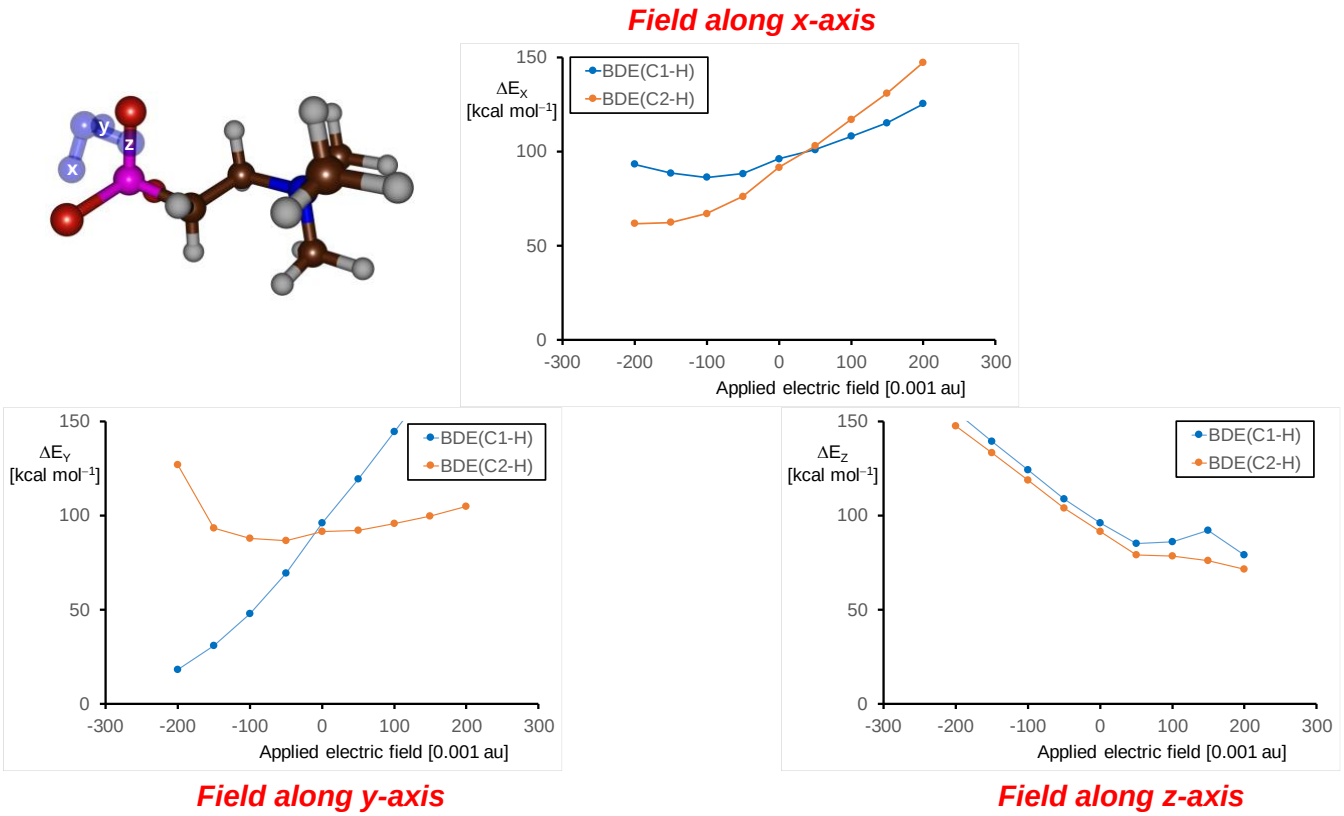


Figure 4.9 Calculated $BDE_{C1–H}$ and $BDE_{C2–H}$ values for an isolated substrate in the gas-phase under the addition of an external electric field along the molecular *x*-, *y*- and *z*-axis as defined in the structure. Energies calculated through a single point calculation on the gas-phase optimized geometry at UB3LYP/6-311++G ** level of theory with an electric field perturbation added in Gaussian-09.

The calculated $BDE_{C1–H}$ and $BDE_{C2–H}$ values under an external electric field are shown in **Figure 4.9**. As shown above in **Figure 4.6**, in the gas-phase the C^1–H and C^2–H bond dissociation energies are $BDE_{C1–H}$ = 96.1 kcal mol^{-1} and $BDE_{C2–H}$ = 91.5 kcal mol^{-1}. As can be seen from **Figure 4.9** an external field along the positive *x*-direction or negative *y*-direction changes the ordering of the two BDE values and makes the C^1–H bond weaker than the C^2–H bond, while in the opposite direction the C^2–H bond is the weakest. In particular, a small electric field perturbation along the molecular *x*-axis already reverses the relative C–H bond strengths and as such a small perturbation form a charged residue in the active site pocket may influence the chemoselectivity of the reaction. Interestingly, a field along the molecular *z*-axis does not change the ordering of the two BDE_{CH} values and keeps the C^2–H below the C^1–H in strength. Clearly, the individual bond strengths of these aliphatic groups are strongly affected by local perturbations, such as charged residues in the active site, an electric dipole moment of, e.g., the metal center, or external electric fields as posed on the reaction by the protein. As such the C–H bond strength of the substrate (2-(trimethylammonio)ethylphosphonate) will be highly sensitive to environmental effects and minor

perturbations along the negative y-direction of the substrate and to a smaller extend the positive x-direction will lead to preferential C^1–H hydrogen atom abstraction over C^2–H hydrogen atom abstraction. The electric dipole vector in ^{5}**R1** is shown in **Figure 4.10** and gives a perturbation along the molecular *x* and *z*-directions of the substrate. As such, the dipole vector of the reactants complex predicts an electric field effect along the positive *x*-field and hence a rise in BDE values for the C^1–H and C^2–H bonds and a reversal of the ordering. It appears therefore, that the electric dipole moment in ^{5}**R1** creates an internal electric field effect of approximately E_X = +0.10 au on the substrate. These results imply that when proteins are latched on a metal surface of a heterogeneous catalyst and exposed to a specific electric field along a reaction coordinate this can trigger a change in regioselectivity of substrate activation with respect to the isolated system.

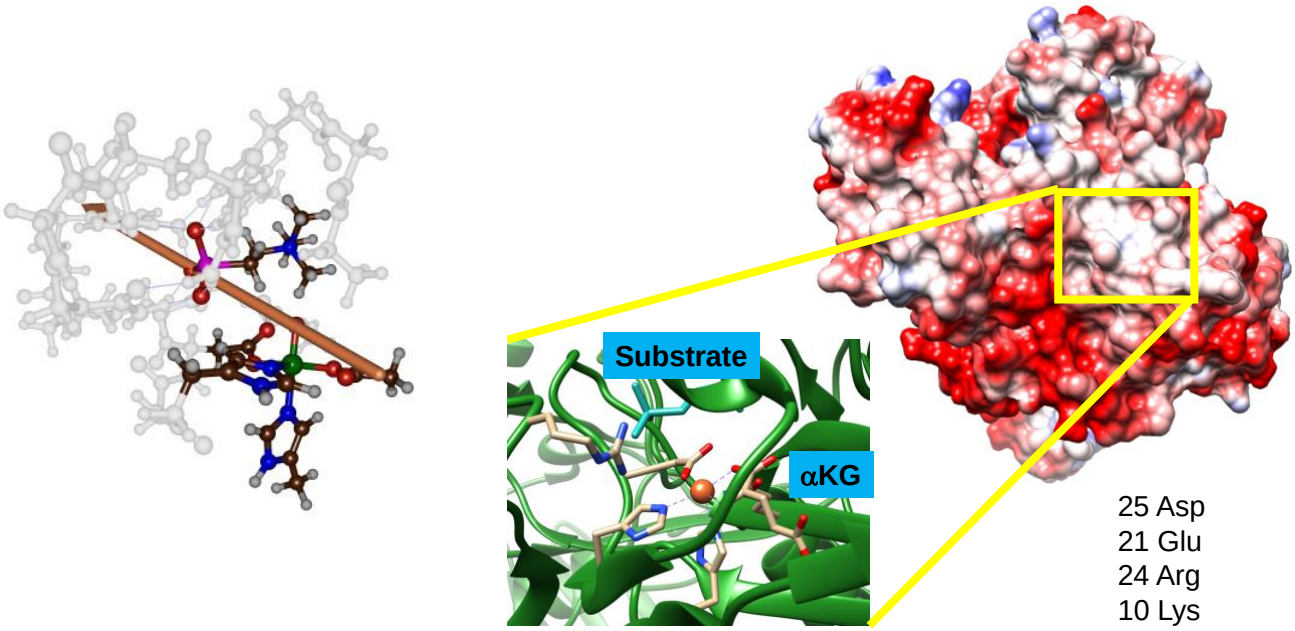


Figure 4.10 Left: electric dipole moment (orange vector) in the optimized geometry of ^{5}**R1**. Right: electrostatic map of the protein structure (6NPC pdb file) with polar residues highlighted (positively charged residues in blue and negatively charged residues in red).

Moreover, an electric field along a chemical bond appears to affect the homolytic bond energy and enables stronger polarization in a specific direction that can guide bond cleavage patterns. This effects also could be integrated in proteins by adding charged residues, e.g., Arg, Lys, Asp, Glu, at specific positions in a protein so that their induced electric field effects influence biocatalysis. Indeed, recent biotechnological studies of Roelfes et al[30c] showed that a single mutation in a metalloenzyme gave enantioselective

Friedel-Crafts alkylation of indoles, while another mutation gave a selective alternative reaction product. Those studies are aligned with our work that shows the importance of substrate binding and positioning and the electrostatic perturbations of the protein that guide regio- and chemoselectivities of chemical reactions.

Next, we explored the contribution of polar amino acid residues in the protein and their electrostatic effects. **Figure 4.10** gives an electrostatic map of the 6NPC structure as well as the electric dipole moment of the optimized geometry of the reactant complex ^{5}**R1**. As can be seen the reactant complex has a large dipole moment, which points mostly along the positive *x*-coordinate and negative *y*-coordinate of the substrate. Hence in agreement with the electric field effect calculations the electric dipole moment will destabilize the C^2–H hydrogen atom abstraction barriers and at the same time stabilize the C^1–H hydrogen atom abstraction pathways. As such the electric dipole moment of the cluster model explains the regioselectivity obtained for the reaction and helps and guides the substrate activation. Obviously this strong dipole moment is missing in the minimal cluster model **R2** and consequently the regioselectivity of substrate activation is reversed. Therefore, electric dipole moments can perturb reaction pathways and trigger a selectivity change in proteins.

A detailed analysis of the TmpA protein shows that it contains 34 positively charged amino acids (Lys and Arg) and 46 negatively charged residues (Asp and Glu). This will lead to a local dipole moment that incurs an electric field effect on the active site, the substrate and the rest of the protein. As the BDE values are sensitive to external perturbations, this local dipole moment pushes the reaction to C^1–H hydrogen atom abstraction and destabilizes the C^2–H hydrogen atom abstraction pathway. Consequently, local electric field and dipole moments in TmpA protein change the selectivity of substrate activation to favor an otherwise impossible or difficult reaction mechanism. The studies presented in this work; therefore, highlight how nature utilizes local electrostatic perturbations and electric field effects to enable a regioselective reaction mechanism.

To gain further insight into the effect of the protein and the charges of the polar residues on the relative barriers heights for ^{5}**TS1**$_{HA,C1}$ versus ^{5}**TS1**$_{HA,C2}$ we did a final set of calculations on the structures of the transition states and their relative energies with protein residues removed. First, we took the optimized geometries of ^{5}**R1**, ^{5}**TS1**$_{HA,C1}$ and ^{5}**TS1**$_{HA,C2}$ and removed all protein residues apart from those ligated to iron and replaced the arginine cation with a Li^+ ion in the position of the central carbon atom of the side chain. Subsequently, we ran a single point energy calculation for these structures at UB3LYP/BS2 with solvent model included to get the no-protein (NP) corrected barrier heights and show those alongside the

original barriers in **Figure 4.11**. As can be seen, removal of the protein residues from the model, but not the charges, keeps the C^1-transition state below the C^2-transition state and lowers the barriers by 4.4 and 4.3 kcal mol^{-1}, respectively. Therefore, the effect of the protein is the same for both barriers and generally means that the electrostatic interactions of the protein destabilize the transition states and raise them in energy and slower the overall reaction.

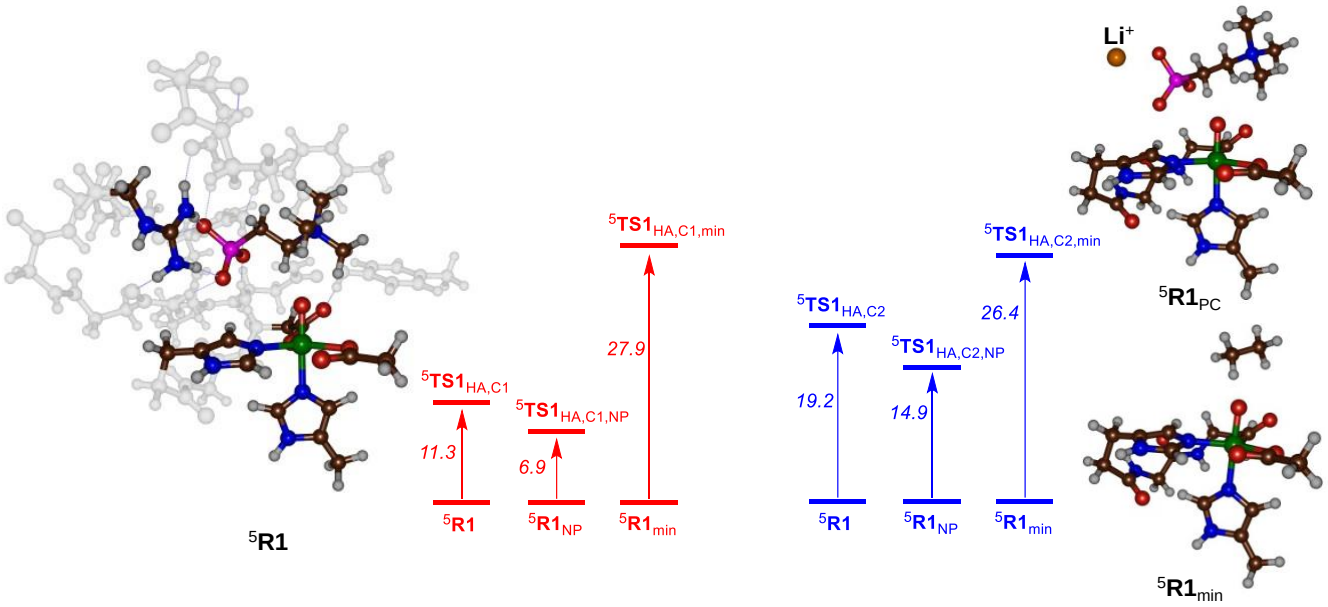


Figure 4.11 Relative energies (ΔE+ZPE in kcal mol^{-1}) of the transition state barriers from Figure 4 and the structures with protein residues removed and Li^+ in the position of Arg_{288} (data with labels ending with NP) and the protein and charges removed, and substrate reduced to ethane (data with labels ending with min).

The same trends are seen when dispersion is included in the calculations. In a second set of calculations, we took the optimized geometries of ^{5}**R1**, ^{5}**TS1**$_{HA,C1}$ and ^{5}**TS1**$_{HA,C2}$ and removed all protein residues as well as the charged groups of the substrate and protein and shortened the substrate to ethane. A single point energy calculation at UB3LYP/BS2 was performed on these abbreviated structures labeled with min in subscript in **Figure 4.11**. As can be seen, the energies of the transition states with charged residues removed are very high in energy but have the C^2-transition state below the C^1-transition state as predicted from the relative energies of the BDE values. Therefore, the charge distributions in the protein and particularly the one from the Arg_{288} side chain in the substrate binding pocket are responsible for the regioselectivity switch from C^2-hydroxylation to C^1-hydroxylation in TmpA enzymes and lower the C^1–H hydrogen atom abstraction barriers below those from the C^2–H group. As a consequence, the charged residues in the protein are essential to guide the selectivity for substrate activation and without the active site Arg residue the selectivity should be lost. It would be interesting to see whether TmpA variants with Arg_{288} replaced by a neutral group would indeed confirm this and change the regioselectivity from C^1-hydroxylation to C^2-hydroxylation instead.

4.4 Conclusion

In this work a computational (density functional theory) study on the nonheme iron dioxygenase TmpA is presented for the first time. We have set up a large active site cluster model containing the first- and second-coordination sphere of the metal centre with substrate binding pocket. The substrate hydroxylation pathways on the C^1-, C^2- and *N*-methyl-positions were studied. In agreement with experimental observation, we find a regioselective C^1–H hydroxylation pathway with a rate-determining hydrogen atom abstraction step. The C^1–H hydrogen atom abstraction barriers are lowest in energy regardless of whether the C^1–H or the C^2–H bond is closest to the oxo group and hence are not dependent on substrate positioning. Therefore, the reactivity in TmpA is not solely determined by substrate positioning but by the relative C–H bond strengths of the available and accessible bonds. Moreover, a model with the substrate C^2–H bond closest to the iron (IV)-oxo group than the C^1–H bond still have preferential C^1–H activation. Interestingly, however, in the gas-phase the C^2–H bond is found to be weaker than the C^1–H bond; yet the C^1–H is being activated by the enzyme. The work shows that TmpA avoids hydroxylation of the weak C^2–H bond through substrate and oxidant positioning that favor the C^1–H hydroxylation and hence reacts through negative catalysis. In

particular, polar residues in the protein induce a local electric field and electric dipole moment and particularly the positively charged Arg_{288} residue that affects the strength of the various C–H bonds in the substrate and trigger a selective substrate hydroxylation pathway. Therefore, the TmpA protein is intricately designed for selective C^1–H hydrogen atom abstraction of 2-(trimethylammonio)ethylphosphonate, where the substrate is positioned such that environmental charge perturbations enable a regioselective C^1–H hydroxylation and avoid the C^2–H hydroxylation at the same time. In particular, the function of the active site Arg_{288} group appears to be by polarizing the charge in the C^1–H bond and weakening it, to enable a regioselective C^1–H hydroxylation process. Through selective positioning of active site charges TmpA; therefore, can react through negative catalysis with favourable C^1–H hydroxylation.

4.5 Experimental Section: Model Set-up

A cluster model was created of the TmpA enzyme based on the crystal structure coordinates deposited in the 6NPC protein databank (pdb) file. [12,16] This is a substrate (2-(trimethylammonio)ethylphosphate) and α-ketoglutarate bound nonheme-iron dioxygenase structure. We selected the residues from chain A and added hydrogen atoms in Chimera at pH = 7.5. Individual protonation states were visually inspected, and it was ascertained that all Asp and Glu residues were deprotonated on the carboxylic acid group, while the amino acid side chains of Arg and Lys residues were protonated. All histidine residues were taken as singly protonated and visually inspected for neighbouring hydrogen bonding donor interactions. Subsequently, the iron (II)- alpha-ketoglutarate group was manually replaced by iron (IV)-oxo-succinate to generate the starting point of the calculations. A molecular dynamics simulation on this structure for 20 ns was performed and showed the system to be highly rigid with little movement of the protein.

A cluster model was created as described above in Figure 2. We initially ran the geometry optimizations with constraints on several α-carbon atoms of the protein chains to keep the structure as close as possible to the crystal structure orientation. However, this gave convergence problems and we attempted to reoptimize without geometric constraints. The unconstraint structure converged much smoother and an overlay with the constraint geometry optimization (Figure S6 and S7, Supporting Information) shows little geometric differences between the two structures probably due to the fact the structure has substrate and oxidant

tightly bound with many hydrogen bonding interactions. Consequently, the rest of the mechanism was run without geometric constraints imposed on the model. The structure and reaction mechanism were explored with two substrate orientations, i.e., structures R1 and R1'.

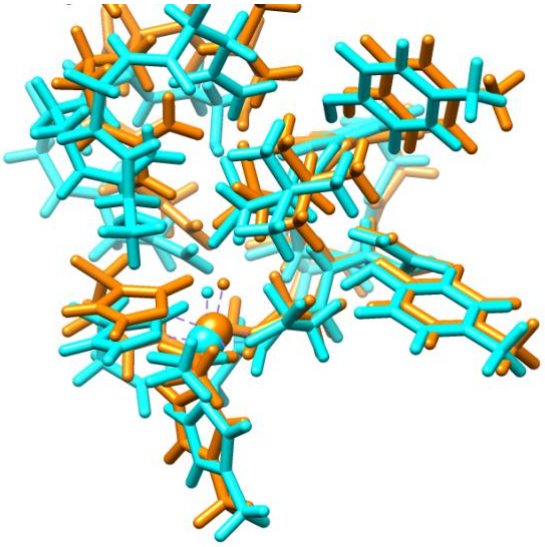


Figure S6: Overlay of the quintet spin reactant complexes (^{5}R1$_{fix}$/^{5}R1$_{nofix}$) optimized with and without constraints.

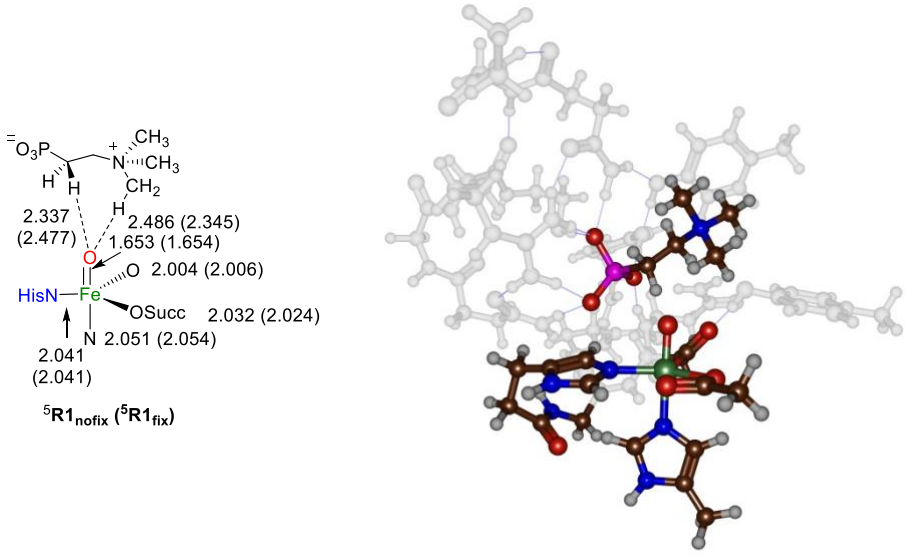


Figure S7: UB3LYP/BS1 optimized geometries of ^{5}R1 as optimized with and without constraints. Bond lengths are given in angstroms. The structure with constraints is ^{5}Re$_{fix}$ and the one without constraints is ^{5}R1$_{nofix}$.

To understand the effect of the protein on the reaction mechanism derived from the large cluster model R1, we also calculated the hydrogen atom abstraction steps with a truncated enzyme model (R2) that includes substrate and the iron (IV)-oxo with its first-coordination sphere ligands only (atoms highlighted in blue and red in Figure 2). In model R2 His198 and His341 were truncated to methylimidazole and succinate and Asp200 to acetate.

We calculated this system in the lowest lying triplet, quintet, singlet and septet spin states. An overlay of the optimized geometry of the model with either the crystal structure coordinates, or the final point of the MD simulation shows little differences in the active site structure and coordination of the reactant structure (Supporting Information, Figures S8 – S11).

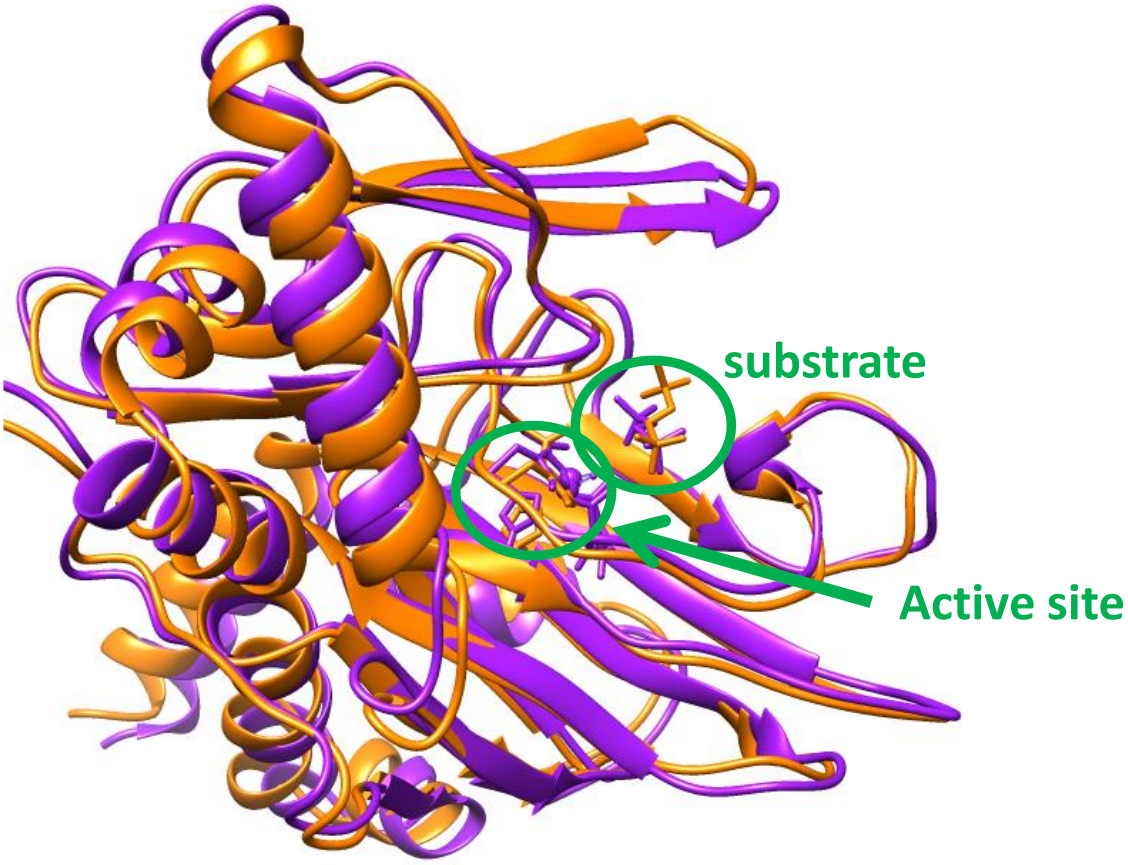


Figure S8. The overlay of two snapshots from MD simulation, namely the equilibrated structure after 15 ns (in amber) and the equilibrated structure (in purple) obtained after 20 ns of MD-simulation. The active site and substrate are highlighted. It can be seen the geometries are very similar.

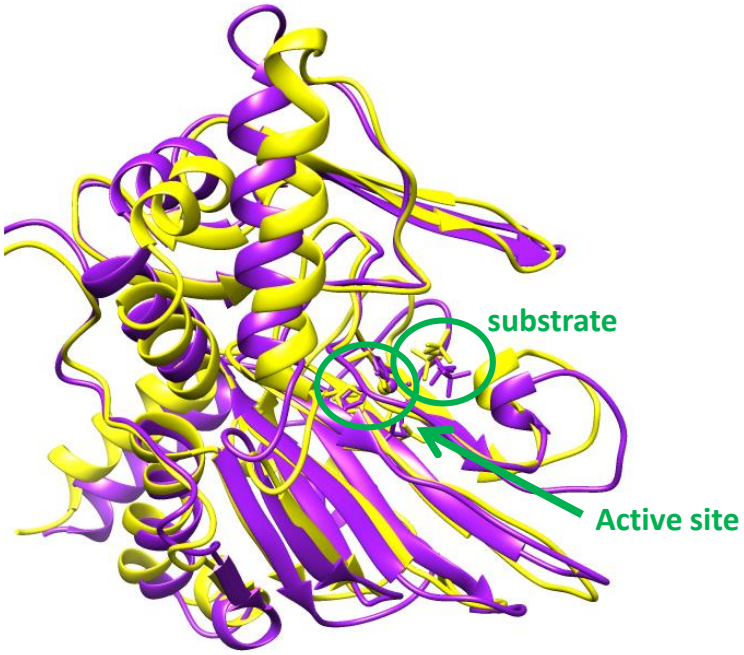


Figure S9. The overlay of crystal structure coordinates of the 6NPC pdb file (in yellow) and the equilibrated structure (in purple) taken from the MD simulation at a snapshot of 20 ns. The active site and substrate are highlighted. It can be seen the geometries are very similar.

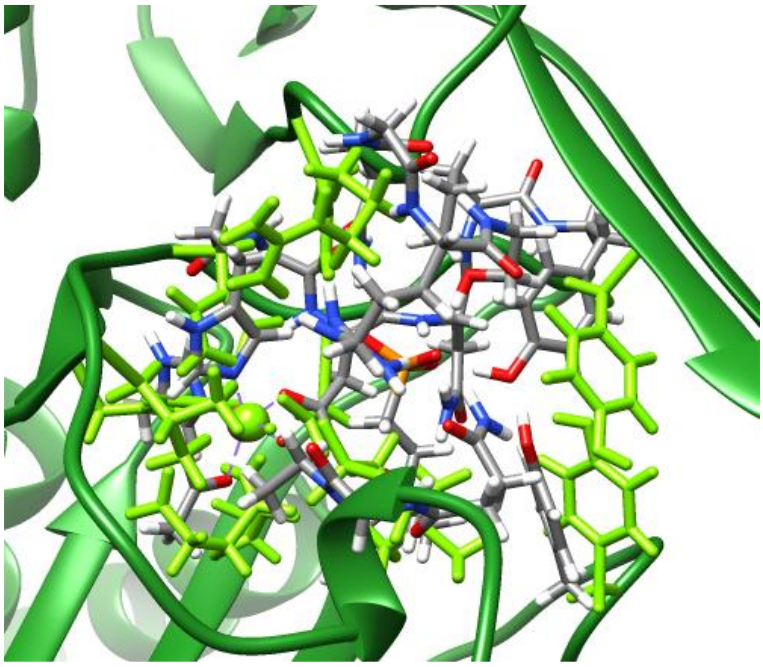


Figure 10. Overlay of optimized geometry of the quintet spin reactant complex, ^{5}**R1** (in atomic colours) and equilibrated structure (in green) obtained after a 20 ns MD-simulation.

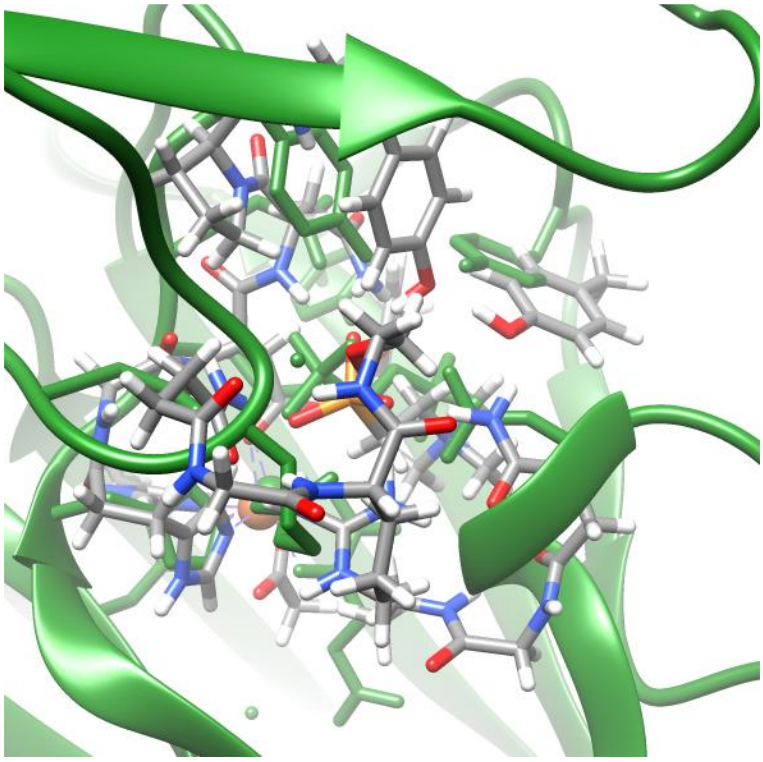


Figure S11. Overlay of the optimized geometry of the quintet spin reactant complex, ^{5}**R1** (in atomic colours) and geometry taken from the 6NPC pdb file (in green).

4.6 Computational Methods

The Gaussian-09 software package was used for all quantum chemical calculations discussed here.[31] Following previous experience with the reactivity of iron complexes,[32] we utilized the unrestricted B3LYP density functional method[33] in combination with a LANL2DZ (with electron core potential) on iron and 6-31G on the rest of the atoms: basis set BS1.[34] To correct the energetics, single point calculations with dispersion corrected UB3LYP-D3 were done with the LACV3P+ (with electron core potential) on iron and 6-311+G* on the rest of the atoms: basis set BS2.[33–35] The latter set of calculations included a continuum polarized conductor model (CPCM) with a dielectric constant mimicking chlorobenzene (ε = 5.6968).[36] Free energies reported here are calculated at 298 K/1 bar and use uncorrected vibrational frequencies for the thermal and entropic corrections. Frequency calculations were performed on all local minima and transition states, and it was confirmed that local minima had real frequencies only, while the transition states had a single imaginary mode for the correct vibration along the reaction coordinate. The methods utilized here were tested and validated against experimental data previously and shown to correctly reproduce chemo- and regiospecificities of enzymatic

and biomimetic systems and give free energies of activation to within a couple of kcal mol^{-1} from experiment.[37]

To test the reproducibility of our results, a selection of calculations were repeated with alternative density functional theory methods and basis sets. Thus, the geometries of ^{5}**R1**, ^{5}**TS1**$_{HA,C1}$ and ^{5}**TS1**$_{HA,C2}$ were reoptimized with PBE0 [38] and B3LYP-D3 methods.[33–35] As can be seen from the results in Tables S3 – S6 (Supporting Information); changing the density functional method affects the potential energy landscapes little and optimized geometries are similar. Therefore, calculations with PBE0 or B3LYP-D3 converge to the same conclusions as those drawn from the B3LYP results.

Table S3: Imaginary frequency (cm^{-1}) of transition state structures along the C^1 and C^2 hydroxylation pathway of a substrate (2-(trimethylammonio)ethylphosphonate) activated by TmpA model complex. Calculated at the PBE0/BS1 level of theory level of theory in the quintet spin state.

System	Imaginary Frequency (cm^{-1})
^{5}**TS1**$_{HA,C1}$	i1234.1
^{5}**TS1**$_{HA,C2}$	i1209.9

Table S4: Absolute (free) energies (in au) of optimized geometries for the C^1 and C^2 hydroxylation pathway of substrate (2-(trimethylammonio)ethylphosphonate) activation of its first step by TmpA model complex as calculated at the UB3LYP-D/BS1 level of theory in quintet spin states.

System	BS1 (Gas)			BS2 + Solv	
	E_{BS1}	E_{BS1}+ Δ ZPE	E_{BS1}+ ΔG	$E_{BS2,solv,D}$ + ΔZPE	$E_{BS2,solv,D}$ + ΔG
^{5}**R1**	-6461.530267	-6459.492299	-6459.668734	-6464.972580	-6462.934612
^{5}**TS1**$_{HA,C1}$	-6461.497133	-6459.465955	-6459.637698	-6462.897484	-6463.069226
^{5}**TS1**$_{HA,C2}$	-6461.491441	-6459.460673	-6459.632833	-6462.895905	-6463.068065

Table S5: Relative (free) energies (kcal mol^{-1}) of optimized geometries for the C^1 and C^2 hydroxylation pathway of substrate (2-(trimethylammonio)ethylphosphonate) activation of its

first step by TmpA model complex as calculated at the UB3LYP-D/BS1 level of theory in quintet spin states.

System	BS1 (Gas)				BS2 + Solv
	ΔE_{BS1}	$\Delta E_{BS1}+\Delta ZPE$	$\Delta E_{BS1}+\Delta G$	$\Delta E_{BS2,Solv,D}+\Delta ZPE$	$\Delta E_{BS2,Solv,D}+\Delta G$
^{5}**R1**	0.0	0.0	0.0	0.0	0.0
^{5}**TS1**$_{HA,C1}$	20.8	16.5	19.5	23.3	26.2
^{5}**TS1**$_{HA,C2}$	24.4	19.8	22.5	24.3	27.0

Table S6: Imaginary frequency (cm^{-1}) of transition state structures along the C^1 and C^2 hydroxylation pathway of substrate (2-(trimethylammonio)ethylphosphonate) activation by TmpA model complex as calculated at the UB3LYP-D/BS1 level of theory level of theory in the quintet spin state.

System	Imaginary Frequency (cm^{-1})
^{5}**TS1**$_{HA,C1}$	i1777.1
^{5}**TS1**$_{HA,C2}$	i1582.1

In addition, the full hydroxylation mechanism was recalculated at UB3LYP/BS3 level of theory (LACV3P+ plus core potential on iron and 6-31G* on the rest of the atoms), which gave only minor changes to the energies along the landscape and very similar structures were obtained. This result matches previous benchmark studies on enzymatic reaction mechanisms that showed little differences between energies obtained at UB3LYP/BS3//UB3LYP/BS1 and UB3LYP/BS3.[39] Furthermore, our group performed a series of validation calculations with a range of computational methods and procedures for oxygen atom transfer reaction by iron(IV)-oxo complexes.[40] These studies showed that most methods give the same selectivities and a systematic error in the free energies of activation, whereby one of the methods that gave best agreement with experiment was UB3LYP/BS2//UB3LYP/BS1, hence this method was used here.

Electric field effect calculations were run in Gaussian using the “field” keyword placed along the molecular *x*-, *y*- or *z*-axis in either the positive or negative field direction.[31] Single point electric field calculations were done with various field sizes along each of the principal axis

4.6 TmpA references

[1]
a) M. Sono, M. P. Roach, E. D. Coulter, J. H. Dawson, Chem. Rev. 1996, 96, 2841–2888;
b) B. Meunier, S. P. de Visser, S. Shaik, Chem. Rev. 2004, 104, 3947–3980;
c) I. G. Denisov, T. M. Makris, S. G. Sligar, I. Schlichting, Chem. Rev. 2005, 105, 2253–2277;
d) M. T. Green, Curr. Opin. Chem. Biol. 2009, 13, 84–88;
e) Handbook of Porphyrin Science, Eds.: K. M. Kadish, K. M. Smith, R. Guilard, World Scientific Publishing Co., New Jersey, 2010;
f) P. R. Ortiz de Montellano, Chem. Rev. 2010, 110, 932–948;
g) Iron-containing Enzymes: Versatile Catalysts of Hydroxylation Reaction in Nature, Eds.: S. P. de Visser, D. Kumar, RSC Publishing, Cambridge (UK), 2011;
h) H. M. Girvan, A. W. Munro, Curr. Opin. Chem. Biol. 2016, 31, 136–145;
i) X. Huang, J. T. Groves, Chem. Rev. 2018, 118, 2491–2553.
[2]
a) E. I. Solomon, T. C. Brunold, M. I. Davis, J. N. Kemsley, S. K. Lee, N. Lehnert, F. Neese, A. J. Skulan, Y. S. Yang, J. Zhou, Chem. Rev. 2000, 100, 235–349;
b) T. D. H. Bugg, Curr. Opin. Chem. Biol. 2001, 5, 550–555;
c) M. J. Ryle, R. P. Hausinger, Curr. Opin. Chem. Biol. 2002, 6, 193–201;
d) M. Costas, M. P. Mehn, M. P. Jensen, L. Que Jr., Chem. Rev. 2004, 104, 939–986;
e) M. M. Abu-Omar, A. Loaiza, N. Hontzeas, Chem. Rev. 2005, 105, 2227–2252;
f) C. Krebs, D. G. Fujimori, C. T. Walsh, J. M. Bollinger Jr., Acc. Chem. Res. 2007, 40, 484–492; g) E. G. Kovaleva, J. D. Lipscomb, Nat. Chem. Biol. 2008, 4, 186–193;
h) C. Yi, C. G. Yang, C. He, Acc. Chem. Res. 2009, 42, 519–529;
i) A. R. McDonald, L. Que Jr., Coord. Chem. Rev. 2013, 257, 414–428;
j) K. Ray, F. F. Pfaff, B. Wang, W. Nam, J. Am. Chem. Soc. 2014, 136, 13942–13958.
[3]
a) C. J. Schofield, Z. Zhang, Curr. Opin. Struct. Biol. 1999, 9, 722–731;
b) K. Gorres, R. T. Raines, Crit. Rev. Biochem. Mol. Biol. 2010, 45, 106–124;
c) M. D. White, E. Flashman, Curr. Opin. Chem. Biol. 2016, 31, 126–135.
[4]
a) M. H. Stipanuk, Annu. Rev. Nutr. 2004, 24, 539–577;
b) G. D. Straganz, B. Nidetzky, ChemBioChem. 2006, 7, 1536–1548;
c) C. A. Joseph, M. J. Maroney, Chem. Commun. 2007, 3338–3349;
d) S. P. de Visser, Coord. Chem. Rev. 2009, 253, 754–768;
e) D. Buongiorno, G. D. Straganz, Coord. Chem. Rev. 2013, 257, 541–563;
f) E. P. Tchesnokov, M. Fellner, E. Siakkou, T. Kleffmann, L. W. Martin, S. Aloi, I. L. Lamont, S. M. Wilbanks, G. N. L. Jameson, J. Biol. Chem. 2015, 290, 24424–24437.
[5]
a) J. Rétey, Angew. Chem. Int. Ed. Engl. 1990, 29, 355; Angew. Chem. 1990, 102, 373–361;
b) B. Vögeli, T. J. Erb, Curr. Opin. Chem. Biol. 2018, 47, 94–100;
c) N. Rad, O. Danylyuk, V. Sashuk, Angew. Chem. Int. Ed. 2019, 58, 11340–11343; Angew. Chem. 2019, 131, 11462–11343;
d) S. P. de Visser, Y.-T. Lin, H. S. Ali, U. K. Bagha, G. Mukherjee, C. V. Sastri, Coord. Chem. Rev. 2021, 439, 213914.
[6]
a) P. C. A. Bruijnincx, G. van Koten, R. J. M. Klein Gebbink, Chem. Soc. Rev. 2008, 37, 2716–2744; b) S. Kal, L. Que Jr., J. Biol. Inorg. Chem. 2017, 22, 339–365.
[7]

a) J. C. Price, E. W. Barr, B. Tirupati, J. M. Bollinger Jr., C. Krebs, Biochemistry 2003, 42, 7497–7508;
b) D. A. Proshlyakov, T. F. Henshaw, G. R. Monterosso, M. J. Ryle, R. P. Hausinger, J. Am. Chem. Soc. 2004, 126, 1022–1023;
c) P. J. Riggs-Gelasco, J. C. Price, R. B. Guyer, J. H. Brehm, E. W. Barr, J. M. Bollinger Jr., C. Krebs, J. Am. Chem. Soc. 2004, 126, 81088109;
d) D. Galonić Fujimori, E. W. Barr, M. L. Matthews, G. M. Koch, J. R. Yonce, C. T. Walsh, J. M. Bollinger Jr., C. Krebs, P. J. Riggs-Gelasco, J. Am. Chem. Soc. 2007, 129, 13408–13409;
e) K. M. Light, J. A. Hangasky, M. J. Knapp, E. I. Solomon, J. Am. Chem. Soc. 2013, 135, 9665–9674.
[8]
a) Y. Mishina, E. M. Duguid, C. He, Chem. Rev. 2006, 106, 215–232;
b) J. M. Bollinger Jr., C. Krebs, Curr. Opin. Chem. Biol. 2007, 11, 151–158;
c) L. C. Blasiak, C. L. Drennan, Acc. Chem. Res. 2009, 42, 147–155;
d) S. Goudarzi, S. R. Iyer, J. T. Babicz Jr., J. J. Yan, G. H. J. Peters, H. E. M. Christensen, B. Hedman, K. O. Hodgson, E. I. Solomon, Proc. Natl. Acad. Sci. USA 2020, 117, 5152–5159.
[9]
a) H. Su, X. Sheng, W. Zhu, G. Ma, Y. Liu, ACS Catal. 2017, 7, 5534–5543;
b) X. Song, J. Lu, W. Lai, Phys. Chem. Chem. Phys. 2017, 19, 20188–20197;
c) Z. Wojdyla, T. Borowski, J. Biol. Inorg. Chem. 2018, 23, 795–808;
d) H.-J. Liao, J. Li, J.-L. Huang, M. Davidson, I. Kurnikov, T.-S. Lin, J. L. Lee, M. Kurnikova, Y. Guo, N.-L. Chan, W.-c. Chang, Angew. Chem. Int. Ed. 2018, 57, 1831–1835; Angew. Chem. 2018, 130, 1849–1835;
e) N. P. Dunham, W.-c. Chang, A. J. Mitchell, R. J. Martinie, B. Zhang, J. A. Bergman, L. J.Rajakovich, B. Wang, A. Silakov, C. Krebs, A. K. Boal, J. M. Bollinger Jr., J. Am. Chem. Soc. 2018, 140, 7116–7126.
[10]
a) M. A. McDonough, V. Li, E. Flashman, R. Chowdhury, C. Mohr, B. M. Lienard, J. Zondlo, N. J. Oldham, I. J. Clifton, J. Lewis, L. A. McNeill, R. J. Kurzeja, K. S. Hewitson, E. Yang, S. Jordan, R. S. Syed, C. J. Schofield, Proc. Natl. Acad. Sci. USA 2006, 103, 9814–9819;
b) J. D. Vasta, A. Choudhary, K. H. Jensen, N. A. McGrath, R. T. Raines, Biochemistry 2017, 56, 219–227.
[11]
a) X. Yin, T. M. Zabriskie, ChemBioChem. 2004, 5, 1274–1277;
b) N. P. Dunham, A. J. Mitchell, J. M. Del Río Pantoja, C. Krebs, J. M. Bollinger Jr., A. K. Boal, Biochemistry 2018, 57, 6479–6488.
[12] L. J. Rajakovich, M.-E. Pandelia, A. J. Mitchell, W.-c. Chang, B. Zhang, A. K. Boal, C. Krebs, J. M. Bollinger Jr., Biochemistry 2019, 58, 1627–1647.
[13] a) I. K. H. Leung, T. J. Krojer, G. T. Kochan, L. Henry, F. von Delft, T. D. W. Claridge, U. Oppermann, M. A. McDonough, C. J. Schofield, Chem. Biol. 2010, 17, 1316–1324;
b) J. J. A. G. Kamps, A. Khan, H. Choi, R. K. Lesniak, J. Brem, A. M. Rydzik, M. A. McDonough, C. J. Schofield, T. D. W. Claridge, J. Mecinović, Chem. Eur. J. 2016, 22, 1270–1276.
[14]
a) S. Shaik, D. Kumar, S. P. de Visser, A. Altun, W. Thiel, Chem. Rev. 2005, 105, 2279–2328; b) M. R. A. Blomberg, T. Borowski, F. Himo, R.-Z. Liao, P. E. M. Siegbahn, Chem. Rev. 2014, 114, 3601–3658;

c) S. P. de Visser, M. G. Quesne, B. Martin, P. Comba, U. Ryde, Chem. Commun. 2014, 50, 262–282.
[15]
a) M. G. Quesne, T. Borowski, S. P. de Visser, Chem. Eur. J. 2016, 22, 2562–2581;
b) X. Sheng, M. Kazemi, F. Planas, F. Himo, ACS Catal. 2020, 10, 6430–6449;
c) S. P. de Visser, Chem. Eur. J. 2020, 26, 5308–5327.
[16] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, Nucl. Acids Res. 2000, 28, 235–242.
[17]
a) J. M. Bollinger Jr., J. C. Price, L. M. Hoffart, E. W. Barr, C. Krebs, Eur. J. Inorg. Chem. 2005, 4245–4254;
b) P. K. Grzyska, E. H. Appelman, R. P. Hausinger, D. A. Proshlyakov, Proc. Natl. Acad. Sci. USA 2010, 107, 39823987.
[18]
a) T. Borowski, A. Bassan, P. E. M. Siegbahn, Chem. Eur. J. 2004, 10, 1031–1041;
b) S. P. de Visser, J. Am. Chem. Soc. 2006, 128, 9813–9824;
c) A. V. Nemukhin, I. A. Topol, R. E. Cachau, S. K. Burt, Theor. Chem. Acc. 2006, 115, 348–353;
d) S. Sinnecker, N. Svensen, E. W. Barr, S. Ye, J. M. Bollinger Jr., F. Neese, C. Krebs, J. Am. Chem. Soc. 2007, 129, 6168–6179;
e) E. Godfrey, C. S. Porro, S. P. de Visser, J. Phys. Chem. A 2008, 112, 2464–2468;
f) H. Chen, W. Lai, J. Yao, S. Shaik, J. Chem. Theory Comput. 2011, 7, 3049–3053;
g) E. A. C. Bushnell, G. B. Fortowsky, J. W. Gauld, Inorg. Chem. 2012, 51, 13351–13356;
h) H. J. Kulik, C. L. Drennan, J. Biol. Chem. 2013, 288, 11233–11241;
i) A. Wójcik, M. Radoń, T. Borowski, J. Phys. Chem. A 2016, 120, 1261–1274;
j) S. Álvarez-Barcia, J. Kästner, J. Phys. Chem. B 2017, 121, 5347–5354;
k) R. N. Manna, T. Malakar, B. Jana, A. Paul, ACS Catal. 2018, 8, 10043–10050;
l) Y.-T. Lin, A. Stańczak, Y. Manchev, G. D. Straganz, S. P. de Visser, Chem. Eur. J. 2020, 26, 22332242; m) S. S. Chaturvedi, R. Ramanan, N. Lehnert, C. J. Schofield, T. G. Karabencheva-Christova, C. Z. Christov, ACS Catal. 2020, 10, 1195–1209;
n) M. G. Quesne, R. Latifi, L. E. Gonzalez-Ovalle, D. Kumar, S. P. de Visser, Chem. Eur. J. 2014, 20, 435–446; o) H. S. Ali, R. H. Henchman, S. P. de Visser, J. Phys. Chem. A 2021, 125, 1720-1737.
[19]
a) Z. Wojdyla, T. Borowski, Chem. Eur. J. 2021, 27, 2196–2211;
b) S. S. Chaturvedi, R. Ramanan, J. Hu, R. P. Hausinger, C. Z. Christov, ACS Catal. 2021, 11, 1578–1592;
c) J. Xue, J. Lu, W. Lai, Phys. Chem. Chem. Phys. 2019, 21, 9957–9968;
d) H. Li, Y. Liu, ACS Catal. 2020, 10, 2942–2957;
e) M. Srnec, S. R. Iyer, L. M. K. Dassama, K. Park, S. D. Wong, K. D. Sutherlin, Y. Yoda, Y. Kobayashi, M. Kurokuzu, M. Saito, M. Seto, C. Krebs, J. M. Bollinger Jr., E. I. Solomon, J. Am. Chem. Soc. 2020, 142, 1888618896;
f) S. B. Han, H. S. Ali, S. P. de Visser, Inorg. Chem. 2021, 60, 48004815.
[20]
a) S. P. de Visser, J. Am. Chem. Soc. 2006, 128, 15809–15818;
b) A. Decker, J.-U. Rohde, E. J. Klinker, S. D. Wong, L. Que Jr., E. I. Solomon, J. Am. Chem. Soc. 2007, 129, 15983–15996;
c) S. Yeh, F. Neese, Proc. Natl. Acad. Sci. USA 2011, 108, 1228–1233;
d) H. Hirao, F. Li, L. Que Jr., K. Morokuma, Inorg. Chem. 2011, 50, 6637–6648;

e) L. Bernasconi, E. J. Baerends, J. Am. Chem. Soc. 2013, 135, 8857–8867.
[21]
a) P. Barman, P. Upadhyay, A. S. Faponle, J. Kumar, S. S. Nag, D. Kumar, C. V. Sastri, S. P. de Visser, Angew. Chem. Int. Ed. 2016, 55, 11091–11095; Angew. Chem. 2016, 128, 11257–11095;
b) F. G. Cantú Reinhard, P. Barman, G. Mukherjee, J. Kumar, D. Kumar, D. Kumar, C. V. Sastri, S. P. de Visser, J. Am. Chem. Soc. 2017, 139, 18328–18338.
[22] S. P. de Visser, J. Am. Chem. Soc. 2010, 132, 1087–1097.
[23]
a) S. Shaik, D. Mandal, R. Ramanan, Nat. Chem. 2016, 8, 1091–1098;
b) T. Stuyver, J. Huang, D. Mallick, D. Danovich, S. Shaik, J. Comput. Chem. 2020, 41, 74–82;
c) Z. Ma, N. Nakatani, H. Fujii, M. Hada, Bull. Chem. Soc. Jpn. 2020, 93, 187–193;
d) O. Kirshenboim, A. Frenklah, S. Kozuch, Chem. Sci. 2021, 12, 3179–3187;
e) S. Yu, P. Vermeeren, T. A. Hamlin, F. M. Bickelhaupt, Chem. Eur. J. 2021, 27, 5683–5693.
[24] S. P. de Visser, Phys. Chem. Chem. Phys. 1999, 1, 749–754.
[25] S. Shaik, S. P. de Visser, D. Kumar, J. Am. Chem. Soc. 2004, 126, 1174611749.
[26] S. P. de Visser, J. Phys. Chem. A 2005, 109, 11050–11057.
[27] A. Timmins, N. J. Fowler, J. Warwicker, G. D. Straganz, S. P. de Visser, Front. Chem. 2018, 6, 513.
[28] H. S. Ali, R. H. Henchman, S. P. de Visser, Chem. Eur. J. 2021, 27, 17951809.
[29] N. S. Hill, M. L. Coote, J. Am. Chem. Soc. 2018, 140, 17800–17804.
[30]
a) Y. Wu, S. D. Fried, S. G. Boxer, J. Am. Chem. Soc. 2020, 142, 9993–9998;
b) C.-Y. Lin, S. G. Boxer, J. Phys. Chem. B 2020, 124, 9513–9525;
c) L. Villarino, S. Chordia, L. Alonso-Cotchico, E. Reddem, Z. Zhou, A.-M. W. H. Thunnissen, J.-D. Maréchal, G. Roelfes, ACS Catal. 2020, 10, 1178311790.
[31] Gaussian 09, Revision D.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian, Inc., Wallingford CT, 2010.
[32]
a) M. G. Quesne, D. Senthilnathan, D. Singh, D. Kumar, P. Maldivi, A. B. Sorokin, S. P. de Visser, ACS Catal. 2016, 6, 2230–2243;
b) S. Ghafoor, A. Mansha, S. P. de Visser, J. Am. Chem. Soc. 2019, 141, 20278–20292;
c) C. Colomban, A. H. Tobing, G. Mukherjee, C. V. Sastri, A. B. Sorokin, S. P. de Visser, Chem. Eur. J. 2019, 25, 14320–14331.
[33]
a) A. D. Becke, J. Chem. Phys. 1993, 98, 5648–5652;
b) C. Lee, W. Yang, R. G. Parr, Phys. Rev. B 1988, 37, 785–789.
[34]

a) P. J. Hay, W. R. Wadt, J. Chem. Phys. 1985, 82, 270–283;
b) W. J. Hehre, R. Ditchfield, J. A. Pople, J. Chem. Phys. 1972, 56, 2257–2261.
[35] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, J. Chem. Phys. 2010, 132, 154104.
[36] M. Cossi, G. Scalmani, N. Rega, V. Barone, J. Chem. Phys. 2002, 117, 43–54.
[37]
a) P. Barman, F. G. Cantú Reinhard, U. K. Bagha, D. Kumar, C. V. Sastri, S. P. de Visser, Angew. Chem. Int. Ed. 2019, 58, 10639–10643; Angew. Chem. 2019, 131, 10749–10643;
b) M. Pickl, S. Kurakin, F. G. Cantú Reinhard, P. Schmid, A. Pöcheim, C. K. Winkler, W. Kroutil, S. P. de Visser, K. Faber, ACS Catal. 2019, 9, 565–577.
[38] E. Brémond, C. Adamo, J. Chem. Phys. 2011, 135, 024106.
[39]
a) T. Yang, M. G. Quesne, H. M. Neu, F. G. Cantú Reinhard, D. P. Goldberg, S. P. de Visser, J. Am. Chem. Soc. 2016, 138, 12375–12386;
b) F. G. Cantú Reinhard, M. A. Sainna, P. Upadhyay, G. A. Balan, D. Kumar, S. Fornarini, M. E. Crestoni, S. P. de Visser, Chem. Eur. J. 2016, 22, 1860818619.
[40] F. G. Cantú Reinhard, A. S. Faponle, S. P. de Visser, J. Phys. Chem. A 2016, 120, 98
[41] Bell, R. P. (1936). The theory of reactions involving proton transfers. Proceedings of the Royal Society of London. Series A-Mathematical and Physical Sciences, 154(882), 414-429.
[42] Evans, M. G., & Polanyi, M. (1936). Further considerations on the thermodynamics of chemical equilibria and reaction rates. Transactions of the Faraday Society, 32, 1333-1360.
[43] Brønsted, J. N., & Pedersen, K. (1924). Die katalytische Zersetzung des Nitramids und ihre physikalisch-chemische Bedeutung. Zeitschrift für Physikalische Chemie, 108(1), 185-235.
[44] Rétey, J. (1990). Enzymic reaction selectivity by negative catalysis or how do enzymes deal with highly reactive intermediates?. Angewandte Chemie International Edition in English, 29(4), 355-361.
[45] Vögeli, B., & Erb, T. J. (2018). ‘Negative’and ‘positive catalysis’: complementary principles that shape the catalytic landscape of enzymes. Current Opinion in Chemical Biology, 47, 94-100.

Chapter 5 Biodegradation of herbicides by a plant nonheme iron dioxygenase: mechanism and selectivity of substrate analogues

[3]Yen-Ting Lin,[a] Hafiz S. Ali and Sam P. de Visser*[a]

[a] Mr Y.-T. Lin, Dr H. S. Ali, Dr S. P. de Visser
Manchester Institute of Biotechnology, The University of Manchester, 131 Princess Street, Manchester, M1 7DN, United Kingdom

[b] Mr Y.-T. Lin, Dr S. P. de Visser
Department of Chemical Engineering and Analytical Science, The University of Manchester, Oxford Road, Manchester, M13 9PL, United Kingdom

[c] Dr H. S. Ali
Department of Chemistry, The University of Manchester, Oxford Road, Manchester, M13 9PL, United Kingdom

Supporting information for this article is given via a link:

https://chemistry-europe.onlinelibrary.wiley.com/doi/abs/10.1002/chem.202103982

DOI: https://doi.org/10.1002/chem.202103982

Keywords: enzyme mechanism • nonheme iron • hydroxylation • density functional theory • inorganic reaction mechanisms

5.1 Abstract

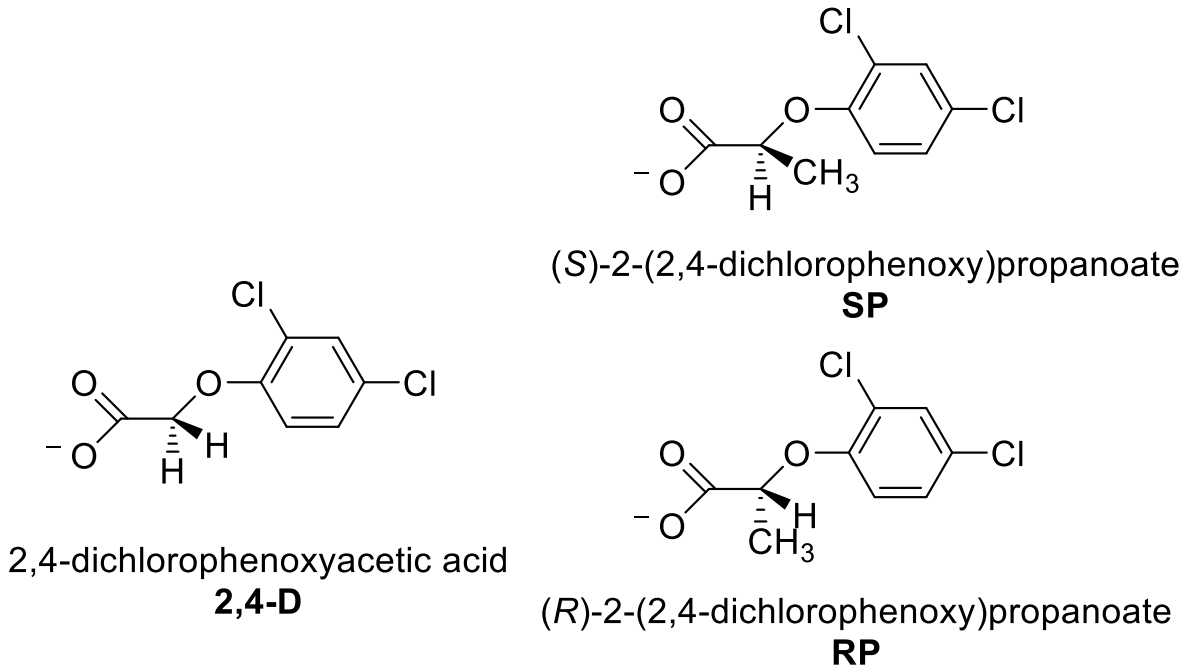
Aryloxyalkanoate dioxygenases are unique herbicide biodegrading nonheme iron enzymes found in plants and hence, from environmental and agricultural point of view they are important and valuable. However, they often are substrate specific, and little is known on the details of the mechanism and the substrate scope. To this end, we created enzyme models and calculate the mechanism for 2,4-dichlorophenoxyacetic acid biodegradation and 2-methyl substituted analogs by density functional theory. The work shows that the substrate binding is tight and positions the aliphatic group close to the metal center to enable a chemoselective reaction mechanism to form the C^2-hydroxy products, whereas the aromatic hydroxylation barriers are well higher in energy. Subsequently, we investigated the metabolism of *R*- and *S*-methyl substituted inhibitors and show that these do not react as efficiently as 2,4-dichlorophenoxyacetic acid substrate due to stereochemical clashes in the active site and particularly for the *R*-isomer give high rebound barriers.

[3] **Author contribution statement:** This is the original reference where the paper is published. Yen-Ting, Lin, the Majority of the contribution is all DFT calculations and data analysis with the help of MD simulation by Dr Hafiz Saqib Ali, guided by Dr Samuel De Visser on all the directions of the project.

5.2 Introduction

Herbicides are common chemicals used in agriculture to improve crop yield. Often, however, these chemicals cause toxicological and ecological problems to the environment.[1] As such, research has been devoted into biodegradable herbicides or environmentally friendly alternatives. Thus, herbicide biodegradation is important in agriculture and particularly from a human health perspective, whereby plants metabolize excess herbicide and prevent these chemicals from entering the human body or the environment. As a consequence, a lot of scientific research has been devoted to studies into the efficiency and efficacy of these herbicides and their toxicological and environmental effects.[2]

In recent years several herbicides have been identified that are biodegradable by plants. In particular, herbicide resistant maize [3] and corn [4] crops were engineered with the aryloxyalkanoate dioxygenase (AAD) enzyme and found to be able to degrade the commonly used herbicide 2,4-dichlorophenoxyacetic acid (**2,4-D**) efficiently. This did not affect plant growth and led to herbicide biodegradation instead of its release into the environment.[4] Compounds like **2,4-D** (**Scheme 5.1**) are extensively used in agriculture as they are low-cost and effective.[5] Therefore, research has been performed into the activation of **2,4-D** and structural and functional analogues, such as *R*- and *S*-2-(2,4-dichlorophenoxy)propanoate (designated **RP** and **SP**). In particular, further use of aryloxyalkanoate dioxygenases in plant biotechnology and agriculture is being pursued. However, as little is known on the catalytic mechanism and substrate scope of these fascinating enzymes, we decided to do a computational study into the enzymatic reaction mechanism of AAD to shed light on its selectivity patterns.



Scheme 5.1 Common dichlorophenoxy acids used in agriculture as herbicides.

Early studies on the enzyme AAD characterized it as a mononuclear nonheme iron hydroxylase that utilizes α-ketoglutarate (αKG; also called 2-oxoglutarate) and dioxygen.[6] Using ^{14}C isotopically labelled αKG it was established that the reaction produces $^{14}CO_2$ as products probably in a reaction with dioxygen. Moreover, gas chromatography-mass spectrometry measurements characterized succinate as a product. Based on these measurements, a catalytic cycle was proposed, whereby αKG binds to an iron(II) center and with dioxygen reacts to form an iron(IV)-oxo species and succinate upon release of CO_2. These nonheme iron(II) dioxygenases are widespread in nature and are involved in biosynthesis as well as biodegradation reactions.[7] For instance, the biosynthesis of antibiotics, such as viomycin and vancomycin, involves a nonheme iron dioxygenase,[8,9] while the metabolism of cysteine in the human body is triggered by the nonheme iron dioxygenase cysteine dioxygense.[10] For analogous nonheme iron dioxygenases, including taurine/αKG dioxygenase and prolyl-4-hydroxylase the iron(IV)-oxo species was trapped and characterized and it was shown to be the active oxidant in a reaction with deuterated substrate.[11,12]

AAD is expected to convert **2,4-D** to its monohydroxylated product initially while in a subsequent step it is transformed into dichlorophenol, which was characterized as the final product. The latter is biodegraded by catechol dioxygenases to β-ketoadipate further. Details

of the reaction mechanism are still elusive and the monohydroxylated product has never been characterized specifically. In addition to AAD, there are also enzymes, namely (*R*)- and (*S*)-dichlorprop/α-ketoglutarate dioxygenases (RdpA and SdpA), that activate the analogous substrates **RP** and **SP** (**Scheme 5.1**). The RdpA and SdpA enzymes show distinct differences in substrate binding pocket and hence do not fit the other enantiomer well and react highly enantiospecifically.[13] Even AAD, is known to activate **SP** but not **RP** efficiently although the reasons for this remain unknown.[14] The AAD amino acid sequence was determined [15] and a structure of AAD enzymes was crystallographically resolved. In **Figure 5.1**, we highlight the active site of AAD as based on the 5BKB protein databank (pdb) file.[16,17] Although the engineered structure has a central manganese(II) ion, the wildtype protein is a nonheme iron dioxygenase where the iron(II) is bound to the protein through a typical 2-His/1-Asp linkage with residues His_{111}, Asp_{113} and His_{270}. The crystal structure has the inhibitor *R*-2-(2,4-dichlorophenoxy)propanoate bound, which is located just above the plane of the metal-αKG group and its carboxylate group forms a hydrogen bonding interaction with the phenol group of Tyr_{221}, while several aliphatic amino acids (Leu_{82}, Leu_{83}) line the top of the substrate binding pocket. An active site Arg residue (Arg_{285}) does not appear to interact with the carboxylate of the substrate but is close to the metal-bound carboxylates of αKG and Asp_{113}.

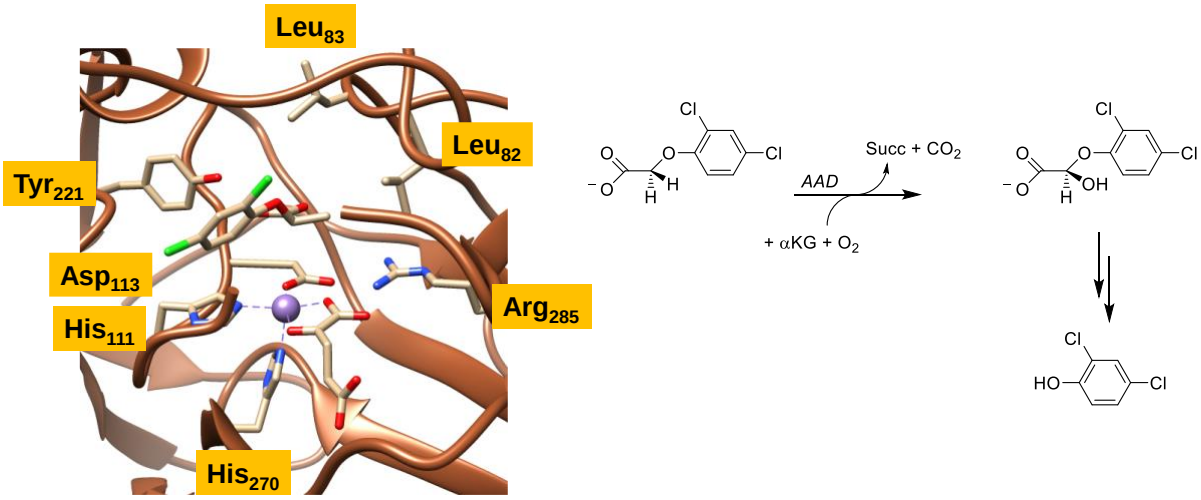


Figure 5.1 *Left*: Active site environment of *R*-2-(2,4-dichlorophenoxy)propanoate-bound AAD as taken from the 5BKB pdb file with key amino acid residues and Mn(II) ion highlighted. *Right*: Reaction mechanism of 2,4-dichlorophenoxyacetic acid biodegradation by AAD enzymes with Succ = succinate and αKG = α-ketoglutarate.

Despite the fact that several biochemical studies investigated the substrate-scope, dioxygen usage and product distributions,[6,15,18] little is known on the details of the reaction mechanism

of AAD and whether alternative substrates can be activated. Therefore, we pursued a computational study into the activation of **2,4-D**, **RP** and **SP** by AAD enzymes. As shown here, the substrate-binding pocket is tight and compact and enables little flexibility, which affects the reaction patterns for substrate hydroxylation. Alternative reactions such as aromatic hydroxylation were also tested but found to be high in energy, while aliphatic hydroxylation is predicted to be the favoured pathway.

5.3 Results and Discussion

We created cluster models based on the first- and second-coordination sphere of the AAD active site and substrate-binding pocket obtained from a molecular dynamics simulation on the crystal structure coordinates. Cluster models have been used extensively for calculating the reaction mechanisms of heme and nonheme iron enzymes.[19] Generally, they include the oxidant and substrate and their direct environments with second-coordination sphere residues that interact through hydrogen-bonding, steric and charge-dipole interactions. Recent studies on cluster models showed that the second-coordination sphere is critical in the description of reaction selectivities, whereby small model complexes or QM/MM with a small QM region often predict wrong kinetics and product distributions when essential substrate-protein interactions are missing in the model.[19de] Large cluster models with more than 200 atoms as calculated with density functional theory often reproduce substrate binding and positioning in enzyme active sites well and have shown to reproduce experimental rate constants and selectivities.[19af] Consequently, they are the method of choice for enzymatic reaction mechanisms.

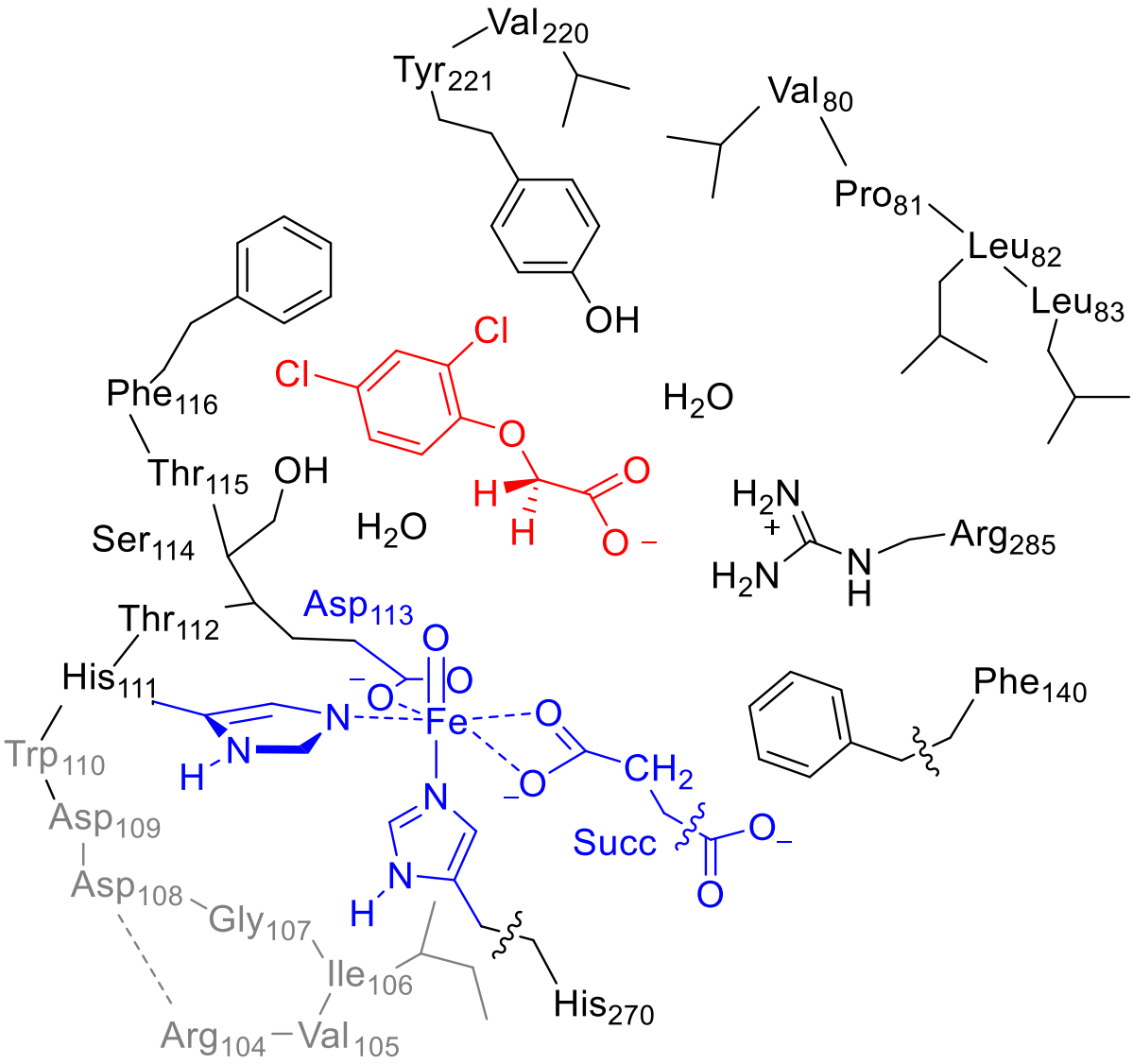


Figure 5.2 AAD model with **2,4-D** bound as studied in this work. Wiggly lines represent where bonds were cut and where link (hydrogen) atoms were added.

Our model (**Figure 5.2**) is based on the 5BKB pdb file,[16,17] and converted into the iron(IV)-oxo oxidant in the catalytic cycle of AAD by replacement of the metal ion in the pdb with an iron(IV)-oxo group with the oxygen atom trans to methylimidazole for His_{270}, and with bound propionate instead the succinate (Succ). Furthermore, a large chain of twelve amino acid residues (Arg_{104} until Phe_{116}) that circumvents the oxidant and donates two ligands (His_{111} and Asp_{113}) to iron was included in the model. An additional eight protein residues were taken as part of the model, which describe the shape of the substrate-binding pocket through mostly aliphatic residues (Val_{80}, Pro_{81}, Leu_{82}, Leu_{83}, Val_{220}) and a Tyr residue (Tyr_{221}) that hydrogen bonds to the carboxylate group of the substrate. The model has two water molecules that form

hydrogen-bonding interactions with the phenol group of Tyr_{221} and the alcohol group of Ser_{114} located in the vicinity of the substrate carboxylate group. Note that the pdb file contains **SP** as the substrate, which was manually replaced by **2,4-D**. The cluster model has 337 atoms and was calculated in the triplet and quintet spin states. In addition, we created iron(IV)-oxo cluster models with bound **SP** (^{5}**Re**$_B$) and one with bound **RP** (^{5}**Re**$_C$) by manually replacing one of the hydrogen atoms of the substrate with a methyl group.

We created two substrate bound reactant structures with **2,4-D** bound, namely **Re**$_{A1}$ and **Re**$_{A2}$ that have the substrate in a different orientation. Both reactant complexes were geometry optimized in Gaussian without constraints in the triplet and quintet spin states. The unconstraint geometry optimizations did not divert the structure dramatically from the starting crystal structure coordinates and an overlay puts most atoms in similar positions (Figure S12, Supporting Information).

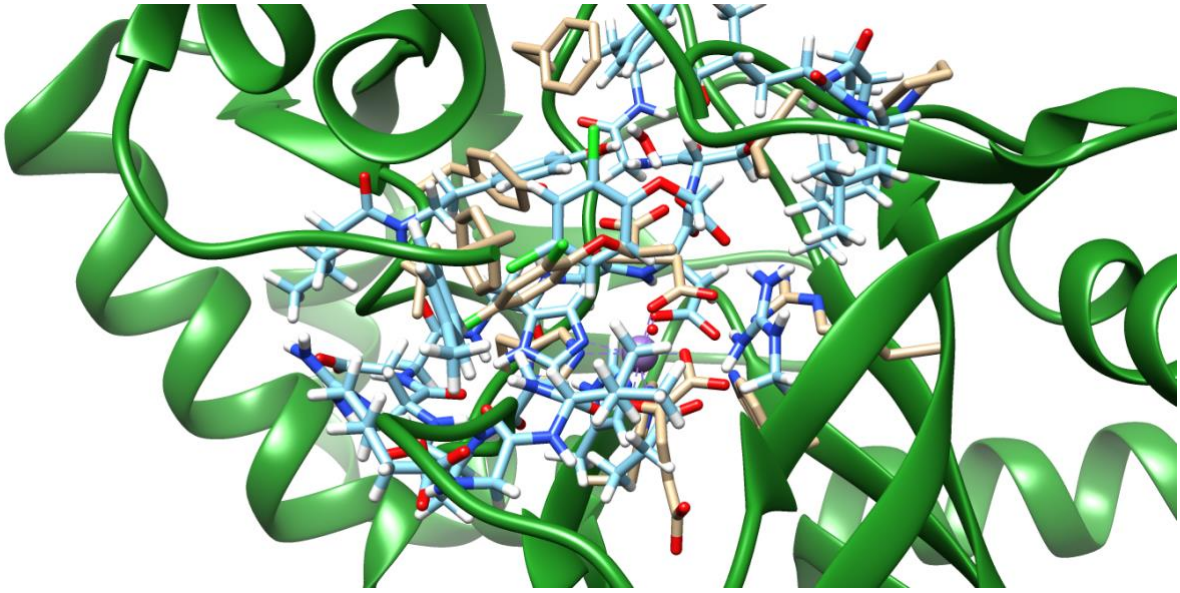


Figure S12. Overlay of the X-ray crystal structure of a 5BKB pdb file (in green ribbons and grey atoms) with optimized geometry of ^{5}**Re**$_{A2}$ (in light blue).

Structure ^{5}**Re**$_{A2}$ is lower in energy than ^{5}**Re**$_{A1}$ by ΔE+ZPE = 2.8 kcal mol^{-1} (ZPE stands for zero-point energy) and as such the two substrate orientations are likely to exist alongside each other. The triplet spin reactants ^{3}**Re**$_{A1}$ and ^{3}**Re**$_{A2}$; on the other hand, are ΔE+ZPE = 15.3 and 16.4 kcal mol^{-1} higher in energy than the lowest energy quintet spin state structure ^{5}**Re**$_{A2}$. The spin state ordering does not change when a different density functional theory method or basis

set is used in the calculations or through the addition of solvent or dispersion corrections. The AAD enzyme; therefore, is expected to have an iron(IV)-oxo intermediate in its catalytic cycle with a quintet spin ground state while the triplet spin conformers will play little role of importance. Our calculated spin-state ordering is in agreement with previous calculations on iron(IV)-oxo intermediates of analogous nonheme iron dioxygenases and biomimetic model complexes,[20,21] that also found a quintet spin state for trigonal bipyramidal iron coordination. In addition, experimental work on nonheme iron enzymes characterized these systems as quintet spin ground states using electron paramagnetic resonance measurements.[11,22] Often in biomimetic iron(IV)-oxo complexes the triplet and quintet spin states are close in energy and the first-coordination sphere orientation determines what the lowest energy spin state is. In particular, in complexes with the metal in trigonal bipyramidal symmetry often the structure is stabilized a high-spin state, while in octahedral symmetry the triplet spin state is the ground state. Therefore, our calculated spin-state-ordering fits the pattern seen in nonheme iron dioxygenases and matches experiment well.[20,22]

Optimized geometries of the lowest energy quintet spin reactant structures of models ^{5}**Re**$_{A1}$ and ^{5}**Re**$_{A2}$ are shown in **Figure 5.3** The two reactant complexes have similar first-coordination sphere interactions with almost identical Fe–O and Fe–N(His$_{270}$) interactions that are found within 0.001 Å of each other. An overlay (right-hand-side of **Figure 5.3**) of ^{5}**Re**$_{A1}$ and ^{5}**Re**$_{A2}$ shows that also the second-coordination sphere is virtually in the same position and little changes have incurred when the substrate was reoriented. Therefore, the enzyme is highly rigid and substrate and oxidant bind neatly into the active site, but there is space for multiple substrate orientations. In both structures the carboxylate group of the substrate forms hydrogen bonding interactions with a peptide amide group (of Ser$_{114}$) as well as with the guanidinium group of Arg$_{285}$. The other NH_2 group of Arg$_{285}$ has hydrogen bonding interactions with the carboxylate groups of succinate and Asp$_{113}$ of the first-coordination sphere. The main difference between the substrate positioning in ^{5}**Re**$_{A1}$ and ^{5}**Re**$_{A2}$ is related to the aliphatic CH_2 group. Thus, the pro-*S* C–H bond in ^{5}**Re**$_{A1}$ is at a distance of 2.626 Å of the iron(IV)-oxo, while the pro-*R* C–H bond is much further away at 4.297 Å. By contrast, in ^{5}**Re**$_{A2}$ both pro-*R* and pro-*S* are at large distances from the iron(IV)-oxo species, namely at 4.169 and 5.764 Å, respectively. Therefore, despite the fact that ^{5}**Re**$_{A2}$ is the lower energy isomer, the substrate orientation is lesser favorable for aliphatic C–H abstraction than in ^{5}**Re**$_{A1}$ and hence may be lesser reactive. Moreover, in ^{5}**Re**$_{A2}$ the *ortho*-carbon atom of the phenyl ring is at a shorter

distance of 3.265 Å (3.287 Å in ^{5}**Re**$_{A1}$). Based on these geometries; therefore, it is not clear how the substrate will react with oxidant and consequently we tested aliphatic hydroxylation at the pro-*R* and pro-*S* positions of the substrate as well as aromatic hydroxylation of the *ortho*-position.

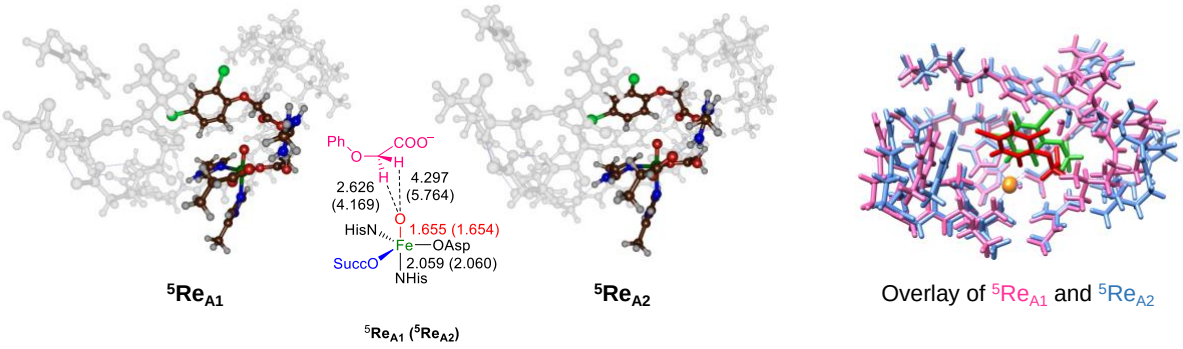
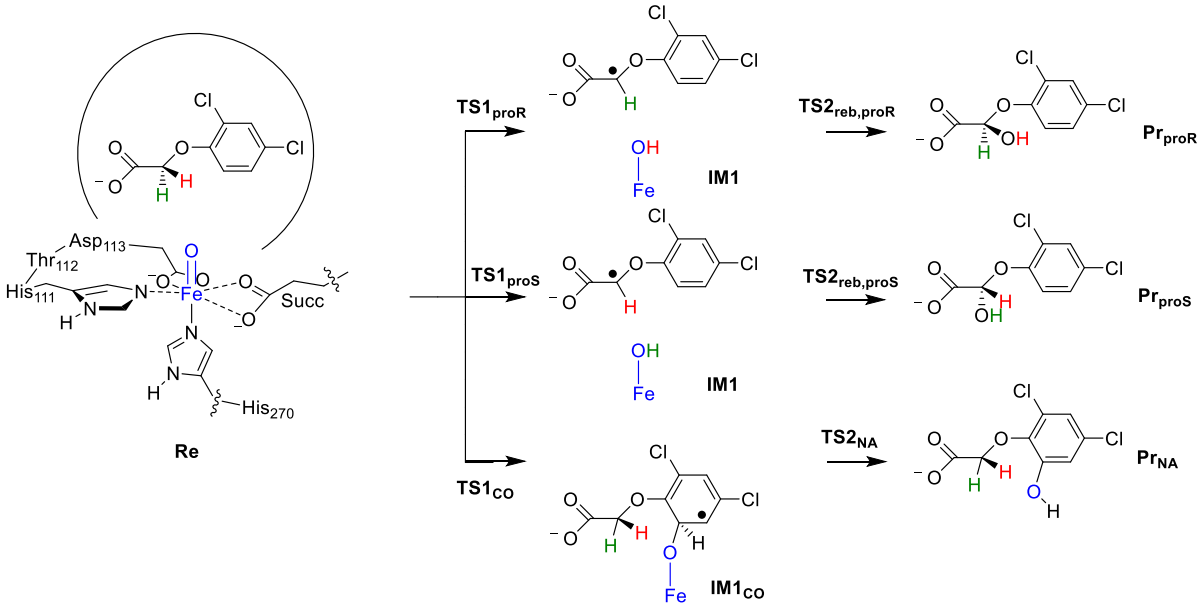


Figure 5.3 Optimized geometries of quintet spin reactant complexes with substrate in position A1 or A2. Bond lengths are in angstroms and the overlay of the two structures gives ^{5}**Re**$_{A1}$ in purple (substrate in red) and ^{5}**Re**$_{A2}$ in light blue (substrate in green).

Next, we explored the reaction mechanisms leading to three possible products, namely pro-*R* and pro-*S* hydroxylation of the C^{2}-position and *ortho*-aromatic hydroxylation of **2,4-D**. Details of the mechanisms and the individual steps considered are shown in **Scheme 5.2**. The aliphatic hydroxylation starts with a hydrogen atom abstraction of the pro-*R* or pro-*S* C–H bond of the substrate via transition states **TS1**$_{proR}$ and **TS1**$_{proS}$, respectively, and relax to an iron(III)-hydroxo radical intermediate **IM1**. An OH rebound step via **TS2**$_{reb,proR}$ or **TS2**$_{reb,proS}$ gives either the *R*-2-hydroxy-**2,4-D** (**Pr**$_{proR}$) or *S*-2-hydroxy-**2,4-D** (**Pr**$_{proS}$) products. As the *ortho*-carbon atom of the phenyl ring is close in position to the iron(IV)-oxo species in both reactants complexes, we also attempted a nucleophilic pathway, where an addition complex (**IM1**$_{CO}$) is formed after a C–O bond formation transition state (**TS1**$_{CO}$). An internal hydrogen atom transfer from the *ipso*-C–H group to the oxo via transition state **TS2**$_{NA}$ gives the phenol product (**Pr**$_{NA}$).

Let us start with a discussion on the aliphatic hydroxylation of **2,4-D** leading to *R*- and *S*-C^{2}-hydroxylated products. Experimental studies failed to trap and characterize the singly hydroxylated species and as such it is not known if the enzyme reacts enantioselectively on

2,4-D. Although we located two reactant complexes, ^{5}**Re**$_{A1}$ and ^{5}**Re**$_{A2}$, both connect to the same hydrogen atom abstraction transition states for hydrogen atom abstraction from the *pro-R* and *pro-S* C^2–H positions.



Scheme 5.2 Reaction mechanisms calculated for substrate activation by AAD with definition of the labels of the structures.

We located transition states on the quintet spin state surface for hydrogen atom abstraction from the *pro-R* and *pro-S* positions of **2,4-D**. The hydrogen atom abstraction transition state for the *pro-R* hydrogen abstraction (^{5}**TS1**$_{proR,\sigma}$) has a barrier of ΔE+ZPE = 15.2 kcal mol^{-1}. On the other hand, the *pro-S* hydrogen atom abstraction transition state (^{5}**TS1**$_{proS,\sigma}$) is ΔE+ZPE = 14.4 kcal mol^{-1} above the reactants complex. Therefore, the calculations predict the *pro-R* and *pro-S* hydrogen atom abstraction transition states to be in a close window within 1 kcal mol^{-1} with a small preference for the *pro-S* C^2–H hydrogen atom abstraction. The ordering does not change when dispersion, thermal or entropic corrections are added to the energies.

Optimized geometries of the transition states are shown in **Figure 5.4** as well. The *pro-R* hydrogen atom abstraction transition state is relatively central with C–H and O–H distances that are of similar magnitude: r(C–H) = 1.288 Å, r(O–H) = 1.247 Å, while these values are 1.251 and 1.275 Å for ^{5}**TS1**$_{proS,\sigma}$, respectively. The two hydrogen atom abstraction transition states both have a large imaginary frequency of i1438 cm^{-1} (*pro-R*) and i1223 cm^{-1} (*pro-S*), which will result in a significant tunneling contribution and a large kinetic isotope effect upon

replacement of the hydrogen atom with deuterium. Indeed, our calculated kinetic isotope effects (Supporting Information, Table S7) predicts the free energy of activation to increase by more than 2 kcal mol^{-1} leading to KIE values well over 10 with empirical tunneling corrections included. As such the structures are functionally similar in geometry as well as energy.

Table S7: Kinetic isotope effect calculations for hydrogen atom abstraction taken from **2,4-D**.

Triplet					
G (au):		**(H-H)**	**(H-D)**	**(D-H)**	**(D-D)**
	3Re	-9124.574007	-9124.577534	-9124.577519	-9124.581069
	3TS1_proR	-9124.537346	-9124.539004	-9124.540807	-9124.542467
	3TS1_proS	-9124.542787	-9124.546265	-9124.544448	-9124.547932
imag (cm^{-1}):	**3TS1_proR**	-1791.5592	-1335.7709	-1777.7906	-1323.6062
	3TS1_proS	-1698.6861	-1692.3193	-1262.7589	-1257.8515
ΔG (kcal/mol)	**3TS1_proR**	23.01	24.18	23.04	24.22
	3TS1_proS	19.59	19.62	20.75	20.79
KIE(Eyring)_proR			7.24	1.06	7.81
KIE(Eyring)_proS			1.05	7.10	7.62
KIE(Wigner)_proR			10.91	1.07	11.91
KIE(Wigner)_proS			1.06	10.60	11.42

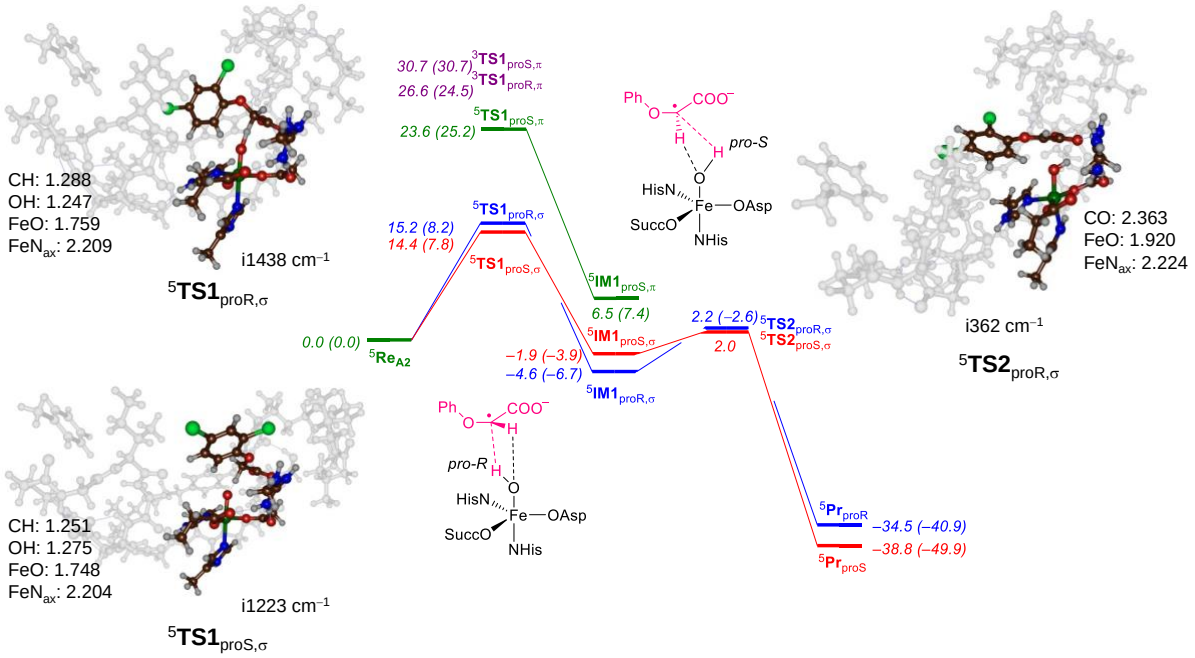


Figure 5.4 DFT calculated potential energy profile for **2,4-D** hydroxylation at the *pro-R* and *pro-S* positions for an AAD cluster model. Energies are in kcal mol^{-1} and are obtained at UB3LYP/BS2//UB3LYP/BS1+ZPE level of theory while values in parenthesis have dispersion corrections included. Optimized geometries give bond lengths in angstroms and the imaginary frequency in cm^{-1}.

The Fe–O distances in the hydrogen atom transition states have elongated from 1.655 Å in ^{5}**Re**$_{A1}$ (**Figure 5.3**) to 1.759 Å in ^{5}**TS1**$_{proR,\sigma}$ and 1.748 Å in ^{5}**TS1**$_{proS,\sigma}$ (**Figure 5.4**) as expected from the electron transfer from the substrate into the σ^*_{z2} orbital along the Fe–O axis. Thus, the reactant structure has a quintet spin configuration with $\pi^*_{xy}{}^1$ $\pi^*_{xz}{}^1$ $\pi^*_{yz}{}^1$ $\sigma^*_{x2\text{-}y2}{}^1$. These orbitals represent the antibonding interactions of the metal 3d orbitals with first-coordination sphere ligands, whereby the z-axis is taken along the molecular Fe–O bond. After hydrogen atom transfer, a radical intermediate is formed (^{5}**IM1**$_\sigma$) with electronic configuration of $\pi^*_{xy}{}^1$ $\pi^*_{xz}{}^1$ $\pi^*_{yz}{}^1$ $\sigma^*_{x2\text{-}y2}{}^1$ $\sigma^*_{z2}{}^1$ $\pi_{Sub}{}^1$. The metal type orbitals are singly occupied with a α-spin electron, while the substrate radical has a β-spin electron in orbital π_{Sub}.

For a number of radical intermediate structures we attempted to locate the alternative π-type intermediate (^{5}**IM1**$_\pi$) with configuration $\pi^*_{xy}{}^2$ $\pi^*_{xz}{}^1$ $\pi^*_{yz}{}^1$ $\sigma^*_{x2\text{-}y2}{}^1$ $\pi_{Sub}{}^1$, whereby all unpaired

electrons are up-spin. For the *pro-S* hydrogen atom abstraction pathway, we located both ^{5}**IM1**$_{proS,\sigma}$ and ^{5}**IM1**$_{proS,\pi}$ states as well as the two transition states leading to these radical complexes, namely ^{5}**TS1**$_{proS,\sigma}$ and ^{5}**TS1**$_{proS,\pi}$, and found the σ-type transition state and intermediate the lowest in energy by at least 8 kcal mol^{-1}. In particular, the ^{5}**TS1**$_{proS,\pi}$ was 23.6 kcal mol^{-1} above the energy of the reactants and consequently well higher in energy than the σ-type transition states and hence the π-pathway will not be able to compete with the σ-pathway. This result matches previous calculations on the σ- versus π-configuration in iron(III)-hydroxo complexes.[23] Although the triplet spin reactants were well above the quintet spin reactants, we calculated and optimized the hydrogen atom abstraction transition states for the *pro-R* and *pro-S* positions as well.

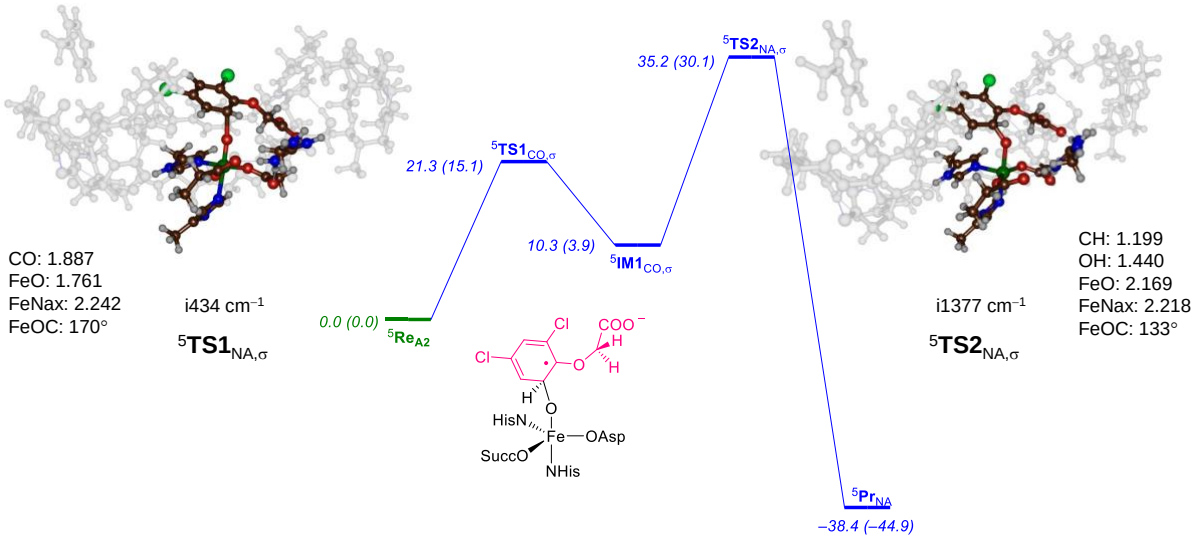


Figure 5.5 DFT calculated potential energy profile for aromatic hydroxylation of **2,4-D** hydroxylation by an AAD cluster model. Energies are in kcal mol^{-1} and are obtained at UB3LYP/BS2//UB3LYP/BS1+ZPE level of theory while values in parenthesis have dispersion corrections included. Optimized geometries give bond lengths in angstroms and the imaginary frequency in cm^{-1}.

These transition states are very high in energy, 26.6 and 30.7 kcal mol^{-1} respectively, and consequently, the triplet spin state will play little role in the substrate activation mechanism.

After the radical intermediate an OH rebound transition state leads to the *R*- and *S*-hydroxy products. The ^{5}**TS2**$_{proR,\sigma}$ and ^{5}**TS2**$_{proS,\sigma}$ transition states are of the same level of energy and are much smaller than the hydrogen atom abstraction, which will be rate-determining. The

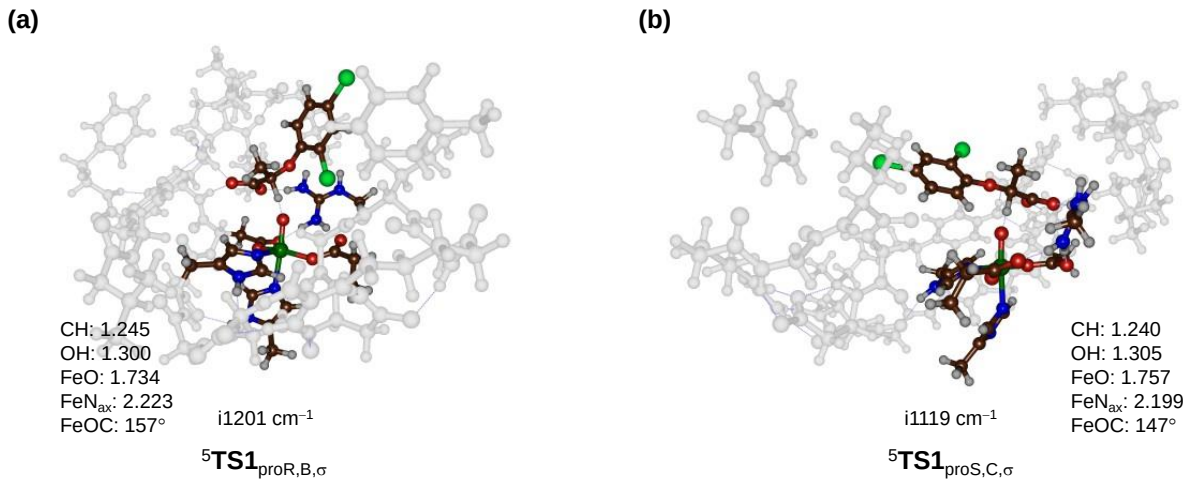
^{5}**TS2**$_{proR,\sigma}$ structure has a long C–O bond of 2.363 Å and hence is an early transition state. Previous DFT calculations on OH rebound reactions gave similar structures with long C–O bonds and elongated Fe–O bonds.[24] The imaginary frequency for the OH rebound transition state is i362 cm^{-1} and shows a C–O stretch vibration as expected.

As the substrate contains an aromatic ring, we considered aromatic hydroxylation at the *ortho*-position with respect to the ether bond and studied this pathway for the model A2 reactant structure. The obtained energy profile and optimized geometries of the transition states along the aromatic hydroxylation pathway are shown in **Figure 5.5**. The initial C–O bond formation transition state is relatively high in energy (ΔE+ZPE = 21.3 kcal mol^{-1}) and is well higher in energy than the hydrogen atom abstraction barriers shown in Figure 4. Structurally, the nucleophilic transition state (^{5}**TS1**$_{CO,\sigma}$) is relatively linear with a Fe–O–C angle of 170° and long C–O and Fe–O bonds of 1.887 and 1.761 Å. The transition state has an imaginary frequency of i434 cm^{-1} corresponding to a C–O stretch vibration. This magnitude of the imaginary frequency is typical for aromatic hydroxylation transition state that typically have a broad potential energy surface with imaginary frequencies below i500 cm^{-1}.[25]

After the transition state the system relaxes to a radical intermediate with $\pi^*_{xy}{}^{1}$ $\pi^*_{xz}{}^{1}$ $\pi^*_{yz}{}^{1}$ $\sigma^*_{x2\text{-}y2}{}^{1}$ $\sigma^*_{z2}{}^{1}$ $\pi_{Sub}{}^{1}$ configuration, i.e. the $^{5}\sigma$-pathway as discussed above in aliphatic hydroxylation. The radical intermediate ^{5}**IM1**$_{CO,\sigma}$; however, is less stable than the reactants complex by ΔE+ZPE = 10.3 kcal mol^{-1}, which means it has a relatively small barrier for the reverse reaction leading back to the reactants complex. Furthermore, the subsequent hydrogen atom transfer from the *ipso*-position to the oxygen atom has a high barrier of 35.2 kcal mol^{-1}. This is a relatively high barrier that will make the alternative aromatic hydroxylation process unlikely at room temperature. In previous studies such as the cytochromes P450, it was shown that this step is assisted by a basic residue in the substrate binding pocket, such as a pyrrole-heme nitrogen atom that would act as a springboard and shuttle the proton from the *ipso*-position to the oxygen atom.[25a] In the structure of AAD; however, no proton acceptor group is available nearby the *ipso*-proton and hence a proton shuttle mechanism is not feasible here. As a consequence, only a direct hydrogen atom transfer is possible, which is high in energy. It is clear that the substrate binding pocket in AAD is not accommodated for an aromatic hydroxylation process. Nevertheless, the overall aromatic hydroxylation reaction is highly

exothermic and the product complex is more stable than reactants by 38.4 kcal mol^{-1}, but due to slow kinetics it may not be a possible product.

Next, we explored the reactivity of alternative substrates in the AAD model and considered the enzyme inhibitors *R*- and *S*-2-(2,4-dichlorophenoxy)propanoate (molecules **RP** and **SP**). To this end we took the ^{5}**Re**$_{A1}$ optimized geometry and manually replaced **2,4-D** with either **SP** or **RP** to obtain the model **B** and **C** reactant complexes: ^{5}**Re**$_{B}$ and ^{5}**Re**$_{C}$. **Figure 5.6** shows details of the aliphatic hydroxylation pathways of the two alternative substrates. Interestingly, both models give similar hydrogen atom abstraction barriers, i.e. 16.2 kcal mol^{-1} for ^{5}**TS1**$_{proR,B,\sigma}$ and 16.4 kcal mol^{-1} for ^{5}**TS1**$_{proS,C,\sigma}$.



	^{5}TS1$_{\sigma}$	^{5}IM1$_{\sigma}$	^{5}TS2$_{\sigma}$	^{5}Pr
Model **B**, substrate **SP**	16.2	-4.9	15.7	-37.4
Model **C**, substrate **RP**	16.4	-3.9	1.8	-33.8

Figure 5.6 Activation of inhibitors **SP** and **RP** by the iron(IV)-oxo model complex of AAD. Energies are in kcal mol^{-1} and are obtained at UB3LYP/BS2//UB3LYP/BS1+ZPE level of theory. Optimized geometries give bond lengths in angstroms and the imaginary frequency in cm^{-1}.

These barriers are somewhat higher in energy than those found for **2,4-D** and show that the second-coordination sphere makes the hydrogen atom abstraction more difficult. The raise in hydrogen atom abstraction barriers from **2,4-D** to **SP**/**RP** is surprising as in **SP** and **RP** and tertiary C–H bond is broken, whereas in **2,4-D** a secondary C–H bond is broken. Usually, tertiary C–H bond strengths are weaker than secondary C–H bond strengths and lead to lower

hydrogen atom abstraction barriers. Therefore, a lower energy hydrogen atom abstraction barrier would be expected for **SP** and **RP** than for **2,4-D**, which clearly is not the case in AAD. Consequently, the substrate binding pocket may affect substrate positioning and hamper the hydrogen atom abstraction barriers for **RP** and **SP**.

Structurally, the hydrogen atom abstraction transition states from **RP** and **SP** by AAD are indeed very similar with C–H and O–H distances within 0.005 Å. The ^{5}**TS1**$_{proR,B,\sigma}$ is slightly more upright with an Fe–O–C angle of 157°, while it is only 147° for ^{5}**TS1**$_{proS,C,\sigma}$. Both transition states have a large imaginary frequency for hydrogen atom transfer with magnitude of i1201 cm^{-1} for ^{5}**TS1**$_{proR,B,\sigma}$ and i1119 cm^{-1} for ^{5}**TS1**$_{proS,C,\sigma}$. These values indicate that the hydrogen atom abstraction will undergo significant amount of tunneling and will experience a large kinetic isotope effect when the transferring hydrogen is replaced by deuterium.

After the hydrogen atom abstraction from either **RP** or **SP** a radical intermediate is formed. However, despite the fact that OH rebound has a relatively small barrier for **RP** (ΔE+ZPE = 6.7 kcal mol^{-1} above ^{5}**IM1**$_{C,\sigma}$) the barrier for rebound for **SP** is much larger, namely 20.6 kcal mol^{-1} above the radical intermediate ^{5}**IM1**$_{B,\sigma}$). This implies that radical rebound will be rate-determining for substrate **SP** and due to the high barrier the reaction will be slow. On the other hand, for **RP** the rebound is lower and the rate-determining step will be hydrogen atom abstraction with significantly lower barrier than OH rebound for substrate **SP**. Therefore, the two stereoisomers **RP** and **SP** will react differently in AAD enzymes and different product distributions will be obtained. As **RP** and **SP** have the same number of atoms, we analyzed the substrate binding energies into the substrate binding pocket by comparing the relative energies of ^{5}**Re**$_{B,SP}$ and ^{5}**Re**$_{C,RP}$. Thus, ^{5}**Re**$_{C,RP}$ is lower in energy than ^{5}**Re**$_{B,SP}$ by 3.4 kcal mol^{-1}. As such, the **RP** inhibitor will be stronger bound than **SP** in the AAD substrate-binding pocket and consequently, product release will be more energetic for **RP** than **SP**. Therefore, there are differences in stability, reactivity and rebound barriers for the reactions of AAD with **RP** and **SP**.

To understand the differences in kinetics between **2,4-D** and **RP**/**SP**, we calculated homolytic bond dissociation energies (BDEs). Firstly, we took isolated substrates and calculated their geometry and energy in the gas-phase. Subsequently, the structure was recalculated with one hydrogen atom removed in the doublet spin state. Together with the energy of a hydrogen atom,

the BDE was calculated for the *pro-R* and *pro-S* C^2–H bonds of **2,4-D** as well as the C–H bonds in **RP** and **SP**. In the gas-phase, not surprisingly, the tertiary C–H bonds in **RP** and **SP** are the weakest at 87.3 kcal mol^{-1}, while the secondary C^2–H bonds in **2,4-D** both have a BDE = 90.5 kcal mol^{-1}. Generally, hydrogen atom abstraction barriers correlate with the strength of the C–H bond that is broken and hence a lower barrier for **RP** and **SP** would be expected than for **2,4-D**.

The hydrogen atom abstraction barriers, however, do not follow the trend in BDE values and we find lower hydrogen atom abstraction barriers from the C^2–H position for **2,4-D** as a substrate than **RP** or **SP**. Therefore, the interactions of the protein and steric restraints of the substrate-binding pocket must affect the kinetics.

To find out how and if the protein affects the BDE values through long-range electrostatic interactions, we took the structures of ^{5}**Re**$_{A2}$, ^{5}**Re**$_B$ and ^{5}**Re**$_C$ and did a single point energy calculation at B3LYP/BS2 for a sextet spin state with one hydrogen atom removed from the C^2–H position of substrate. We then calculated the substrate diabatic BDE values inside the protein matrix and the results are shown at the bottom of **Figure 5.7**. As can be seen, the BDE values of **2,4-D** change inside the protein and both *pro-R* and *pro-S* C^2–H bonds are stronger with values of 99.6 and 99.3 kcal mol^{-1}. Despite this, the calculations find that there should not be a selectivity for hydrogen atom abstraction from the C^2–H bond with almost equal bond energies. This is indeed what the potential energy landscapes above show. As these values are close in energy similar hydrogen atom abstraction barriers are expected, which is indeed seen from the full transition state optimizations from **Figure 5.4**.

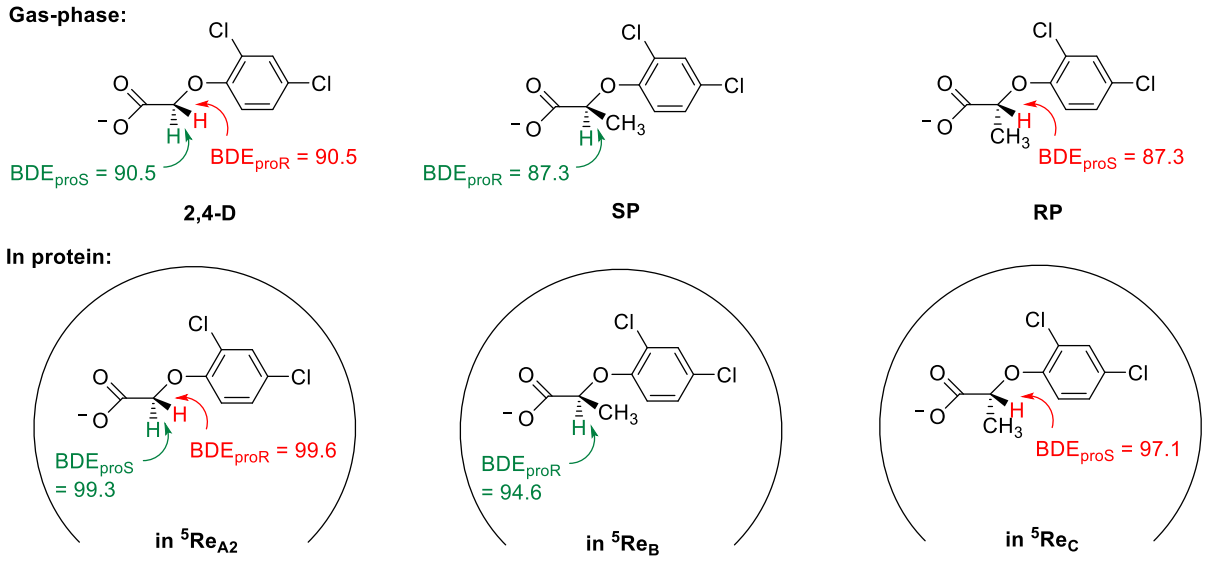


Figure 5.7 Calculated (UB3LYP/BS2) values for substrate BDEs in the gas phase (top) or inside a protein environment for the substrate-bound reactant complexes. Energies contain ZPE and are in kcal mol^{-1} for the abstraction of a hydrogen atom.

Similarly to the BDEs of **2,4-D** inside the protein, also the C^2–H BDEs for **SP** and **RP** were evaluated from the reactant complexes ^{5}**Re**$_B$ and ^{5}**Re**$_C$. In both cases the C–H bond strength goes up by about 10 kcal mol^{-1} with respect to the gas-phase BDE. Moreover, the C^2–H BDEs for **SP** and **RP** inside the protein are both lower in energy than the ones for **2,4-D**. This would mean a lower hydrogen atom abstraction for **SP** or **RP** as a substrate compared to **2,4-D**. However, the fully characterized transition states show slightly higher energy barriers for **SP** and **RP** (16.2 and 16.4 kcal mol^{-1}) as compared to those from **2,4-D** (*pro-R* at 15.2 kcal mol^{-1} and *pro-S* at 14.4 kcal mol^{-1}). Therefore, steric restraints of substrate approach rather than long-range electrostatic contributions affect the kinetics of substrate activation in AAD.

5.4 Conclusion

The work described in this paper is focused on the binding and activation of herbicide molecules to the nonheme iron enzyme AAD. Cluster models with **2,4-D**, **RP** and **SP** bound in the AAD binding pocket with iron(IV)-oxo present were generated and the substrate hydroxylation pathways studied with density functional theory methods. The work shows that all three substrates can bind into the active site and react through C^2–H hydroxylation, which

is identified as the initial product of the AAD reaction with substrate. The reaction pathways are stepwise via a radical intermediate and with a rate-determining hydrogen atom abstraction step. The *pro-R* and *pro-S* hydrogen atom abstraction barriers are close in energy and so is the subsequent OH rebound. Consequently, the DFT calculations predict a mixture of *R*- and *S*-hydroxy-**2,4-D** as products. The alternative aromatic hydroxylation was also studied but found to be high in energy mostly due to the absence of a proton shuttle acceptor that can relay a proton from the aromatic ring to the phenol group. Subsequently, we studied **RP** and **SP** hydroxylation by AAD and show that the hydrogen atom abstraction barriers – against expectation based on BDE values – are higher in energy than those for **2,4-D**. Despite the slower hydrogen atom abstraction step, both substrates can be activated by AAD, fit the substrate binding pocket well and should give hydroxylated products efficiently. Although there may be differences in overall reaction rate due to major differences in OH rebound barrier for the two stereo-isomers as a result of the substrate binding pocket interactions that slow down the OH rebound step for **SP**. Overall, the work shows that **2,4-D** and analogous substrates can be biodegraded by AAD enzymes efficiently and detoxified. Therefore, this should give ample applications of AAD in biotechnology and agriculture.

5.5 Experimental Section

5.5.1 Model Set-Up

A cluster model was designed based on the crystal structure coordinates of the substrate and α-ketoglutarate bound form of AAD-1 reported by Nair et al[16] and deposited as the 5BKB pdb file at the protein databank.[17] This is an engineered protein structure saturated with manganese(II) rather than iron(II). Subsequently, the manganese(II)-α-ketoglutarate group in the pdb was replaced by iron(IV)-oxo-succinate. We selected the residues from chain A and added hydrogen atoms in Chimera at pH = 7.3.[26] A molecular dynamics (MD) simulation was performed and shows that the structure is highly rigid and has a tight substrate binding pocket.

Next, we created a cluster model based on the last point of the MD simulation. We took the iron(IV)-oxo(succinate) and abbreviated succinate in our model to propionate. In the model (**Figure 5.2**) the axial histidine residue (His_{270}) was truncated to methylimidazole, while the other protein ligands of the metal were part of the peptide chain Arg_{104}-Val_{105}-Ile_{106}-Gly_{107}-Asp_{108}-Asp_{109}-Trp_{110}-His_{111}-Thr_{112}-Asp_{113}-Ser_{114}-Thr_{115}-Phe_{116} with Val_{105}, Trp_{110}, Thr_{112} and Thr_{115} shortened to a Gly residue. In addition, part of the substrate binding pocket was included in the model, namely the chain Val_{80}-Pro_{81}-Leu_{82}-Leu_{83} and the peptide dimer Val_{220}-Tyr_{221}. The Phe_{140} and Arg_{285} side chains were abbreviated to toluene and methylguanidine. The overall model is shown in **Figure 5.2** and had a total of 337 atoms and has overall charge –1. We calculated this system in the lowest lying triplet and quintet spin states. The inhibitor models were started from the optimized geometries of the ^{5}**Re**$_{A2}$ structures, where substrate **2,4-D** was manually replaced to **RP** or **SP** by replacing one of the C^2–H groups to C^2–CH_3.

5.5.2 Computational Approach

The *Gaussian*-09 software package was used for all quantum chemical calculations discussed here.[27] Following our previous experience with cluster models of nonheme iron dioxygenases,[28] we utilized the unrestricted B3LYP density functional method in combination with a LANL2DZ (with electron core potential) on iron and 6-31G on the rest of the atoms: basis set BS1.[29,30] To correct the energetics, single point calculations with the LACV3P+ (with electron core potential) on iron and 6-311+G* on the rest of the atoms were performed. The latter set of calculations included a continuum polarized conductor model (CPCM) with a

dielectric constant mimicking chlorobenzene.[31] We ran some test calculations with the SMD solvent model and find quantitatively the same trends as with CPCM.

To validate the methods and models a series of test calculations on the rate-determining steps were performed using the PBE0/BS2 and B3LYP-D3/BS2 methods.[29,32,33] These calculations predicted the same trends in spin-state-ordering and chemoselectivity and hence did not change the conclusions. Frequency calculations were performed on all local minima and transition states, and it was confirmed that local minima had real frequencies only, while the transition states had a single imaginary mode for the correct vibration along the reaction coordinate. The amount of spin contamination in the optimized geometries was verified for each structure and the S^2 values for the triplet spin structures are close to the ideal value of 2, while those for the quintet spin structures are close to the ideal value of 6.

5.6 AAD references

[1] F. Islam, J. Wang, M. A. Farooq, M. S. S. Khan, L. Xu, J. Zhu, M. Zhao, S. Muños, Q. X. Li, W. Zhou, Environ. Int. 2018, 111, 332–351.
[2]
a) R. R. Fulthorpe, C. McGowan, O. V. Maltseva, W. E. Holben, J. M. Tiedje, Appl. Environ. Microbiol. 1995, 61, 3274–3281;
b) T. Paszko, P. Muszyński, M. Materska, M. Bojanowska, M. Kostecka, I. Jackowska, Environ. Toxicol. Chem. 2016, 35, 271–286;
c) D. Chen, S. Wu, H. Xue, J. Jiang, Int. Biodeterior. Biodegrad. 2020, 146, 104822;
d) N. V. Coleman, D. J. Rich, F. H. M. Tang, W. Vervoort, F. Maggi, Environ. Sci. Technol. 2020, 54, 10399–10410;
e) K. Fukui, K.-i. Hayashi, Plant Cell Physiol. 2018, 59, 1500–1510.
[3] D. C. Ruen, E. F. Scherder, S. C. Ditmarsen, P. L. Prasifka, J. M. Ellis, D. M. Simpson, C. A. Gallup, B. W. Hopkins, Weed Technol. 2017, 31, 217–224.
[4] T. R. Wright, G. Shan, T. A. Walsh, J. M. Lira, C. Cui, P. Song, M. Zhuang, N. L. Arnold, G. Lin, K. Yau, S. M. Russell, R. M. Cicchillo, M. A. Peterson, D. M. Simpson, N. Zhou, J. Ponsamuel, Z. Zhang, Proc. Natl. Acad. Sci. USA 2010, 107, 20240–20245.
[5]
a) Y. Song, J. Integr. Plant Biol. 2014, 56, 106–113;
b) M. A. Peterson, S. A. McMaster, D. E. Riechers, J. Skelton, P. W. Stahlman, Weed Technol. 2016, 30, 303–345;
c) J. Zhou, K. Liu, F. Xin, J. Ma, N. Xu, W. Zhang, Y. Fang, M. Jiang, W. Dong, World J. Microbiol. Biotechnol. 2018, 34, 117–125;
d) M. P. Serbent, A. M. Rebelo, A. Pinheiro, A. Giongo, L. B. Ballod Tavares, Appl. Microbiol. Biotechnol. 2019, 103, 5065–5078;
e) K. Magnoli, C. Soledad Carranza, M. Eglé Aluffi, C. E. Magnoli, C. L. Barberis, Environ. Sci. Pollut. Res. Int. 2020, 27, 38501–38512.
[6] F. Fukumori, R. P. Hausinger, J. Bacteriol. 1993, 175, 2083–2086.
[7]
a) E. I. Solomon, T. C. Brunold, M. I. Davis, J. N. Kemsley, S. K. Lee, N. Lehnert, F. Neese, A. J. Skulan, Y. S. Yang, J. Zhou, Chem. Rev. 2000, 100, 235–349;
b) T. D. H. Bugg, Curr. Opin. Chem. Biol. 2001, 5, 550–555;
c) M. J. Ryle, R. P. Hausinger, Curr. Opin. Chem. Biol. 2002, 6, 193–201;
d) M. Costas, M. P. Mehn, M. P. Jensen, L. Que Jr., Chem. Rev. 2004, 104, 939–986; e) M. M. Abu-Omar, A. Loaiza, N. Hontzeas, Chem. Rev. 2005, 105, 2227–2252;
f) C. Krebs, D. G. Fujimori, C. T. Walsh, J. M. Bollinger Jr, Acc. Chem. Res. 2007, 40, 484–492;
g) S. P. de Visser, D. Kumar (Eds.) Iron-containing enzymes: Versatile catalysts of hydroxylation reactions in nature, Royal Society of Chemistry Publishing, Cambridge (UK), 2011;
h) A. R. McDonald, L. Que Jr, Coord. Chem. Rev. 2013, 257, 414–428.
[8]
a) A. G. Prescott, M. D. Lloyd, Nat. Prod. Rep. 2000, 17, 367–383;
b) F. Kudo, A. Miyanaga, T. Eguchi, Nat. Prod. Rep. 2014, 31, 1056–1073;
c) D. Baud, P.-L. Saaidi, A. Monfleur, M. Harari, J. Cuccaro, A. Fossey, M. Besnard, A. Debard, A. Mariage, V. Pellouin, J.-L. Petit, M. Salanoubat, J. Weissenbach, V. de Berardinis, A. Zaparucha, ChemCatChem. 2014, 6, 3012–3017.
[9]

a) J. H. Carter II, R. H. Du Bus, J. R. Dyer, J. C. Floyd, K. C. Rice, P. D. Shaw, Biochemistry 1974, 13, 1227–1233;
b) J. Zhou, W. L. Kelly, B. O. Bachmann, M. Gunsior, C. A. Townsend, E. I. Solomon, J. Am. Chem. Soc. 2001, 123, 7388–7398;
c) X. Yin, T. M. Zabriskie, ChemBioChem. 2004, 5, 1274–1277;
d) V. Helmetag, S. A. Samel, M. G. Thomas, M. A. Marahiel, L.-O. Essen, FEBS J. 2009, 276, 3669–3682;
e) A. J. Mitchell, N. P. Dunham, R. J. Martinie, J. A. Bergman, C. J. Pollock, K. Hu, B. D. Allen, W.-c. Chang, A. Silakov, J. M. Bollinger Jr, C. Krebs, A. K. Boal, J. Am. Chem. Soc. 2017, 139, 13830–13836;
f) N. P. Dunham, A. J. Mitchell, J. M. Del Río Pantoja, C. Krebs, J. M. Bollinger Jr, A. K. Boal, Biochemistry 2018, 57, 6479–6488.
[10]
a) G. D. Straganz, B. Nidetzky, ChemBioChem. 2006, 7, 1536–1548;
b) C. A. Joseph, M. J. Maroney, Chem. Commun. 2007, 3338–3349;
c) D. Buongiorno, G. D. Straganz, Coord. Chem. Rev. 2013, 257, 541–563;
d) E. P. Tchesnokov, M. Fellner, E. Siakkou, T. Kleffmann, L. W. Martin, S. Aloi, I. L. Lamont, S. M. Wilbanks, G. N. L. Jameson, J. Biol. Chem. 2015, 290, 24424–24437.
[11]
a) J. C. Price, E. W. Barr, B. Tirupati, J. M. Bollinger Jr, C. Krebs, Biochemistry 2003, 42, 7497–7508;
b) D. A. Proshlyakov, T. F. Henshaw, G. R. Monterosso, M. J. Ryle, R. P. Hausinger, J. Am. Chem. Soc. 2004, 126, 1022–1023;
c) P. J. Riggs-Gelasco, J. C. Price, R. B. Guyer, J. H. Brehm, E. W. Barr, J. M. Bollinger Jr, C. Krebs, J. Am. Chem. Soc. 2004, 126, 81088109;
d) M. L. Neidig, C. D. Brown, K. M. Light, D. Galonić-Fujimori, E. M. Nolan, J. C. Price, E. W. Barr, J. M. Bollinger Jr, C. Krebs, C. T. Walsh, E. I. Solomon, J. Am. Chem. Soc. 2007, 129, 14224–14231;
e) P. K. Grzyska, [12] L. M. Hoffart, E. W. Barr, R. B. Guyer, J. M. Bollinger Jr, C. Krebs, Proc. Natl. Acad. Sci. USA 2006, 103, 14738–14743.
[13]
a) T. A. Müller, M. I. Zavodszky, M. Feig, L. A. Kuhn, R. P. Hausinger, Protein Sci. 2006, 15, 1356–1368;
b) J. Huang, D. Chen, J. Jiang, Environ. Microbiol. 2020, 22, 286–296;
c) W. A. Peer, Curr. Opin. Plant Biotechnol. 2013, 23, 561–568.
[14] K. Nickel, M. J.-F. Suter, H.-P. E. Kohler, J. Bacteriol. 1997, 179, 6674–6679.
[15] E. J. Perkins, M. P. Gordon, O. Caceres, P. L. Lurquin, J. Bacteriol. 1990, 172, 2351–2359.
[16] J. R. Chekan, C. Ongpipattanakul, T. R. Wright, B. Zhang, J. M. Bollinger Jr, L. J. Rajakovich, C. Krebs, R. M. Cicchillo, S. K. Nair, Proc. Natl. Acad. Sci. USA 2019, 116, 13299–13304.
[17] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, Nucl. Acids Res. 2000, 28, 235–242.
[18]
a) M. A. Bhat, T. Ishida, K. Horiike, C. S. Vaidyanathan, M. Nozaki, Arch. Biochem. Biophys. 1993, 300, 738–746;
b) T. Ledger, D. H. Pieper, B. González, Appl. Environ. Microbiol. 2006, 72, 2783–2792;
c) C. T. Larue, M. Goley, L. Shi, A. G. Evdokimov, O. C. Sparks, C. Ellis, A. M. Wollacott, T. J. Rydel, C. E. Halls, B. Van Scoyoc, X. Fu, J. R. Nageotte, A. M. Adio, M. Zheng, E. J. Sturman, G. S. Garvey, M. J. Varagona, Pest Manage. Sci. 2019, 75, 2086–2094;

d) R. Zallot, N. Oberg, J. A. Gertl, Curr. Opin. Biotechnol. 2021, 69, 77–90;
e) K. Hoyerova, P. Hosek, M. Quareshy, J. Li, P. Klima, M. Kubes, A. A. Yemm, P. Neve, A. Tripathi, M. J. Bennett, R. M. Napier, New Phytol. 2018, 217, 1625–1639.
[19]
a) M. R. A. Blomberg, T. Borowski, F. Himo, R.-Z. Liao, P. E. M. Siegbahn, Chem. Rev. 2014, 114, 3601–3658;
b) S. P. De Visser, M. G. Quesne, B. Martin, P. Comba, U. Ryde, Chem. Commun. 2014, 50, 262–282;
c) M. G. Quesne, T. Borowski, S. P. De Visser, Chem. Eur. J. 2016, 22, 2562–2581;
d) X. Sheng, M. Kazemi, F. Planas, F. Himo, ACS Catal. 2020, 10, 64306449;
e) C.-C. G. Yeh, C. Pierides, G. N. L. Jameson, S. P. de Visser, Chem. Eur. J. 2021, 27, 13793–13806;
f) S. P. de Visser, Y.-T. Lin, H. S. Ali, U. K. Bagha, G. Mukherjee, C. V. Sastri, Coord. Chem. Rev. 2021, 439, 213914.
[20]
a) T. Borowski, A. Bassan, P. E. M. Siegbahn, Chem. Eur. J. 2004, 10, 1031–1041;
b) S. P. de Visser, J. Am. Chem. Soc. 2006, 128, 9813–9824;
c) A. V. Nemukhin, I. A. Topol, R. E. Cachau, S. K. Burt, Theor. Chem. Acc. 2006, 115, 348–353; d) S. Sinnecker, N. Svensen, E. W. Barr, S. Ye, J. M. Bollinger Jr, F. Neese, C. Krebs, J. Am. Chem. Soc. 2007, 129, 6168–6179;
e) E. Godfrey, C. S. Porro, S. P. de Visser, J. Phys. Chem. A 2008, 112, 2464–2468;
f) H. Chen, W. Lai, J. Yao, S. Shaik, J. Chem. Theory Comput. 2011, 7, 3049–3053;
g) E. A. C. Bushnell, G. B. Fortowsky, J. W. Gauld, Inorg. Chem. 2012, 51, 13351–13356;
h) H. J. Kulik, C. L. Drennan, J. Biol. Chem. 2013, 288, 11233–11241;
i) A. Wójcik, M. Radoń, T. Borowski, J. Phys. Chem. A 2016, 120, 1261–1274;
j) S. Álvarez-Barcia, J. Kästner, J. Phys. Chem. B 2017, 121, 5347–5354;
k) R. N. Manna, T. Malakar, B. Jana, A. Paul, ACS Catal. 2018, 8, 10043–10050;
l) J. E. M. N. Klein, G. Knizia, Angew. Chem. Int. Ed. 2018, 57, 11913–11917; Angew. Chem. 2018, 130, 12089–12093;
m) S. S. Chaturvedi, R. Ramanan, N. Lehnert, C. J. Schofield, T. G. Karabencheva-Christova, C. Z. Christov, ACS Catal. 2020, 10, 1195–1209;
n) M. G. Quesne, R. Latifi, L. E. Gonzalez-Ovalle, D. Kumar, S. P. de Visser, Chem. Eur. J. 2014, 20, 435–446;
o) H. S. Ali, R. Henchman, S. P. de Visser, J. Phys. Chem. A 2021, 125, 1720–1737.
[21]
a) H. Hirao, D. Kumar, L. Que Jr, S. Shaik, J. Am. Chem. Soc. 2006, 128, 8590–8606;
b) I. Prat, J. S. Mathieson, M. Guëll, X. Ribas, J. M. Luis, L. Cronin, M. Costas, Nat. Chem. 2011, 3, 788–793;
c) R. Latifi, M. A. Sainna, E. V. Rybak-Akimova, S. P. de Visser, Chem. Eur. J. 2013, 19, 4058–4068;
d) A. Ansari, A. Kaushik, G. Rajaraman, J. Am. Chem. Soc. 2013, 135, 4235–4249;
e) T. Z. H. Gani, H. J. Kulik, ACS Catal. 2018, 8, 975–986;
f) V. Dantignana, J. Serrano-Plana, A. Draksharapu, C. Magallón, S. Banerjee, R. Fan, I. Gamba, Y. Guo, L. Que Jr, M. Costas, A. Company, J. Am. Chem. Soc. 2019, 141, 15078–15091.
[22] J. M. Bollinger Jr, J. C. Price, L. M. Hoffart, E. W. Barr, C. Krebs, Eur. J. Inorg. Chem. 2005, 2005, 4245–4254.
[23]
a) S. P. de Visser, J. Am. Chem. Soc. 2006, 128, 15809–15818;

b) A. Decker, J.-U. Rohde, E. J. Klinker, S. D. Wong, L. Que Jr, E. I. Solomon, J. Am. Chem. Soc. 2007, 129, 15983–15996;
c) S. Yeh, F. Neese, Proc. Natl. Acad. Sci. USA 2011, 108, 1228–1233.
[24]
a) S. Shaik, S. Cohen, S. P. de Visser, P. K. Sharma, D. Kumar, S. Kozuch, F. Ogliaro, D. Danovich, Eur. J. Inorg. Chem. 2004, 2004, 207–226;
b) D. C. Cummins, J. G. Alvarado, J. P. T. Zaragoza, M. Q. E. Mubarak, Y.-T. Lin, S. P. de Visser, D. P. Goldberg, Inorg. Chem. 2020, 59, 16053–16064.
[25]
a) S. P. de Visser, S. Shaik, J. Am. Chem. Soc. 2003, 125, 7413–7424;
b) D. Kumar, G. N. Sastry, S. P. de Visser, J. Phys. Chem. B 2012, 116, 718–730;
c) S. Louka, S. M. Barry, D. J. Heyes, M. Q. E. Mubarak, H. S. Ali, L. M. Alkhalaf, A. W. Munro, N. S. Scrutton, G. L. Challis, S. P. de Visser, J. Am. Chem. Soc. 2020, 142, 15764–15779.
[26] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, J. Comput. Chem. 2004, 25, 1605–1612.
[27] Gaussian-09, Revision D.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian, Inc., Wallingford CT, 2010.
[28]
a) S. Ghafoor, A. Mansha, S. P. de Visser, J. Am. Chem. Soc. 2019, 141, 20278–20292;
b) Y.-T. Lin, A. Stańczak, Y. Manchev, G. D. Straganz, S. P. de Visser, Chem. Eur. J. 2020, 26, 2233–2242;
c) Y.-T. Lin, H. S. Ali, S. P. de Visser, Chem. Eur. J. 2021, 27, 8851–8864.
[29]
a) A. D. Becke, J. Chem. Phys. 1993, 98, 5648–5652;
b) C. Lee, W. Yang, R. G. Parr, Phys. Rev. B 1988, 37, 785–789.
[30]
a) P. J. Hay, W. R. Wadt, J. Chem. Phys. 1985, 82, 270–283;
b) W. J. Hehre, R. Ditchfield, J. A. Pople, J. Chem. Phys. 1972, 56, 2257–2261.
[31] M. Cossi, G. Scalmani, N. Rega, V. Barone, J. Chem. Phys. 2002, 117, 43–54.
[32]
a) J. P. Perdew, M. Ernzerhof, K. Burke, J. Chem. Phys. 1996, 105, 99829985;
b) A. Carlo, V. Barone, J. Chem. Phys. 1999, 110, 6158–6170.
[33] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, J. Chem. Phys. 2010, 132, 154104.

Chapter 6: Conclusion

6.1 Conclusion

Nonheme iron enzymes are versatile catalysts in Nature that can perform highly selective oxidation reactions. These essential enzymes have a variety of roles in biological systems and are particularly important for biodegradation, detoxification, and biosynthesis inside the human body. Furthermore, they offer potential in biotechnology and bioengineering due to their versatility and ability to catalyse various complex chemical processes. I have conducted a computational analysis on several model structures to understand this enzyme's selectivity and mechanisms better. However, there is still debate over the mechanism of the reaction and the source of the product distributions for several nonheme iron enzymes. My calculations demonstrate that each one of them has a binding stage for the substrate, dioxygen, and co-substrate a-ketoglutarate at the beginning of the catalytic cycle. After the release of CO_2, an iron (III)-superoxo is produced, which interacts with a-ketoglutarate to create a persuccinate complex. The next step in the catalytic cycle is the heterolytic cleavage of the persuccinate O-O bond to succinate and a high-valent iron(IV)-oxo molecule. This iron(IV)-oxo species, which has been demonstrated to be a good oxidant and interacts with the substrate through C-H hydrogen atom abstraction, is generated in all nonheme iron dioxygenases. The enzymes UndA, TmpA, and AAD, mentioned in this thesis establish this. Following the conception of the hydrogen atom, an iron(III)-hydroxo and substrate radical is created, which causes the OH to rebound and synthesise the alcohol products. My calculations agree that the barriers are low, and the reaction may occur readily at ambient temperature because all enzymes catalyse this same reaction.

Additionally, the enzymes frequently exhibit remarkable selectivity, meaning only specific substrate C-H groups are targeted. My work suggests how second-coordination sphere effects affect the product distributions that the enzyme causes. Effect of the second-coordination sphere include interactions between polar groups via hydrogen bonds, the development of salt bridges, charged residues, and ions. The second-coordination sphere's polar interactions provide an electric field effect that polarises chemical bonds and may even cause reactivity patterns to reverse. Because nonheme iron dioxygenases are sensitive to second-coordination sphere effects, the research I discuss here highlights the plasticity of enzyme design and the viability of protein engineering using these enzymes.

In this thesis, I explicitly examine the catalytic reaction mechanism of the nonheme iron enzymes UndA, 2-(trimethylammonio)ethylphosphonate dioxygenase (TmpA), and aryloxyalkanoate dioxygenase (AAD). The nonheme iron dioxygenases, their role, catalytic cycle, reaction routes, and presence in biology are all covered in Chapter 1 of my thesis. In Chapter 2, I go into great depth on the processes I used to gather and evaluate the data for my computational chemistry research. Several research chapters are then offered after that. In chapter 3, we look at the debate around UndA's mechanism, which carries out an oxidative decarboxylation process.

However, modelling of the experimentally suggested mononuclear iron system failed to replicate experimental findings because an iron(III)-superoxo model for this enzyme generated high barriers for hydrogen atom abstraction. No feasible route at room temperature circumstances was predicted by any of the models, despite testing them with various protonation states, atom counts, spin states, and computational techniques.

In contrast to the postulated mononuclear-iron system, we then studied a di-iron system. We discovered significantly lower energy barriers for hydrogen atom abstraction and the routes leading to decarboxylation. Later research by the Makris group, who performed electron paramagnetic resonance analyses on UndA and discovered proof of a second iron atom in the active site, supported our findings. We next researched TmpA, a mononuclear iron dioxygenase that binds a highly polar substrate with an amine and phosphate group, a comparable enzyme. We accounted for a highly polar substrate binding pocket discovered by protein structural analysis. This research looked at substrate hydroxylation and hydrogen atom abstraction at the C^1, C^2, and C^4 sites.

Our calculated reaction barriers indicate that TmpA specifically hydroxylates the C^2-position and suggests the right observed product distributions. This was unexpected because the C^1-H bond is supposed to be the weakest of the three C-H bonds, according to a computation of the substrate's multiple C-H bond dissociation energies. TmpA thus reacts by a process known as negative catalysis, in which a thermodynamically advantageous pathway is blocked in favour of a higher energy pathway. Then, we studied how the enzyme triggered this selectivity switch and looked at how the TmpA protein matrix's dipole effects and electric field influence the

enzymatic activity. In fact, we demonstrate that the reaction's selectivity may be influenced by an external electric field or external charges.

In Chapter 5, we investigate a nonheme iron dioxygenase that may biodegrade herbicides and is present in crops like maize and rice. In order to progress in biotechnology and food production, it is crucial to comprehend how the enzyme destroys herbicides. A cluster model of the active site, which includes the iron centre, the substrate, and the protein, was used to research the aryloxyalkanoate dioxygenase (AAD) enzyme. We investigated the C-H hydroxylation routes of several substrates. The calculations determined if it was possible to bind substrate and its analogues to the substrate binding pocket and the potential of their metabolism. The outcomes demonstrate that the substrate analogues are biodegradable since they can be activated.

As for the limitations of my research, I evaluated and validated the methodology and approaches. Several basis sets and density functional approaches were investigated for geometry optimisations and energetics for diverse projects. When modifying the technique and basis set, there was often slight variation in the spin-state ordering and energetics. When the model is altered, considerable discrepancies can occasionally be discovered, which typically means the smaller model is less suited.