MECHANISTIC AND STRUCTURAL STUDIES OF THE NON-HEME IRON OXYGENASE ENZYMES PROTOCATECHUATE 3,4-DIOXYGENASE, CMLA, AND CMLI

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA

BY

CORY JAMES KNOOT

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

ADVISOR: JOHN D. LIPSCOMB

DECEMBER 2015

ACKNOWLEDGEMENTS

I want to thank my advisor John D. Lipscomb. His advice and guidance have made me not only a more rigorous scientist but also a better writer and storyteller. I know that his way of thinking, writing, and solving research problems has left an indelible mark on the way that I approach these things. I have heard some people say that it is almost impossible to get through a graduate career without your advisor imprinting some of his or her traits on you. I can already see that happening to me and I am glad for it.

I would also like to acknowledge and thank all of the members of the Lipscomb group. Without all of you, this journey would have been far more difficult and, perhaps more importantly, a lot less fun. I have been very fortunate to have been able to work with so many talented people that were also a joy to be around. I always knew that I had a place to share my victories and failures. I can’t possibly list all of the people who made this possible, but please know that I am thankful for you.

Finally, I want to thank my friends and family. Throughout all of this, you were always there to support me. I simply could not have done this without you.

To my parents,
Jeffrey and Ulrike Knoot,
for your unending love and support

ABSTRACT

Oxygenase enzymes catalyze a remarkable variety of challenging biological transformations using molecular dioxygen (O_2) as a cosubstrate. Oxygenases make possible the direct reaction of ground-state triplet O_2 with organic singlet substrates, a transformation that is spin-forbidden by quantum mechanics. The set of chemical processes by which enzymes overcome this forbidden reaction are collectively termed 'O_2 activation.' The result of oxygenase reactions is incorporation of one or both atoms of O from O_2 into the organic substrate. Because of the critical role of oxygenases in many fundamental biological processes, an understanding of the catalytic and regulatory mechanisms that underlie O_2 activation is required for the development of new medical, industrial, ecological and agricultural technologies. Non-heme iron oxygenases use mononuclear iron or diiron cofactors that are not coordinated in a porphyrin scaffold. This dissertation focuses on structural and mechanistic studies of three non-heme iron enzymes. The first, protocatechuate 3,4-dioxygenase (3,4-PCD), is an archetypal aromatic ring-cleaving oxygenase from soil bacteria that uses the oxidized form (Fe^{3+}) of the iron cofactor to react with O_2. It is a member of one of only two enzyme families that use Fe^{3+} to activate O_2. CmlA is a diiron cluster-containing oxygenase that is involved in the non-ribosomal peptide synthetase-mediated biosynthesis of the antibiotic chloramphenicol. It catalyzes the first step in this pathway: the β-hydroxylation of an amino acid precursor of the antibiotic. The third enzyme, CmlI, also uses a diiron cluster cofactor and catalyzes the final step in chloramphenicol biosynthesis by converting the arylamine group of the chloramphenicol precursor to its arylnitro analog. Herein, we report the determination of the X-ray crystal structures of the two most critical oxygenated intermediates in the 3,4-PCD catalytic cycle and characterize the reaction of the enzyme with diagnostic substrates. This work confirms mechanistic proposals that have existed since the 1960s and lays the groundwork for future studies. In the last two chapters of the dissertation, we report the X-ray crystal structures of CmlA and CmlI and compare their structures to other diiron cluster-utilizing enzymes. We also discuss the insights gained into their catalytic and regulatory mechanisms.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Metalloenzymes in biological O_2 activation

Roughly one third of all characterized proteins require metal cofactors in order to function; these are referred to as metalloproteins. Metal cofactor-utilizing enzymes (metalloenzymes) are present in all three domains of life[1-3] and serve a staggering variety of critical biological roles including global biogeochemical cycling (N, C, S, H)[4-8], photosynthesis [9], human drug metabolism[10-12], natural products biosynthesis[13], immune response [14], DNA repair[15], oxygen transport, aerobic and anaerobic respiration[5, 16] and more. Metal cofactors can be divided into two very general classes: *(i)* redox-inert metals such as the alkali and alkali earth metals and *(ii)* redox-active transition metals. In most cases, metals of the former group serve structural, signaling, and hydrolytic catalytic roles and will not be discussed in detail here. The focus will instead be on enzymes that utilize the latter group of transition metals as catalytic cofactors to facilitate reactions between molecular dioxygen (O_2) and organic substrates.

Oxygenase enzymes and O_2 activation. All of terrestrial life is bathed in a sea of O_2, but the extreme reactivity of dioxygen (as witnessed in an uncontrolled fire or explosion) is held in check by the ground-state electronic structure of the O_2 molecule. This potent reactivity is a double-edged sword—although dangerous when activated in an uncontrolled fashion to "reactive oxygen species" *in vivo*, the tremendous oxidizing power also lends itself to fantastic chemistries when harnessed and controlled. Controlled O_2 activation is the hallmark of oxygenase enzymes. Oxygenases are defined as those enzymes in which O_2 is the electron acceptor in the catalytic reaction, *and* one or both O atoms of O_2 are incorporated into an organic substrate. Oxygenases belong to several subclasses of the oxidoreductase group of enzymes, which catalyze a wide variety of oxidation:reduction reactions. Oxygenase enzymes that integrate a single O atom or both O atoms into another substrate are termed monooxygenases and dioxygenases, respectively. Oxygenases catalyze tremendously challenging biological oxidations including site- and stereo-specific mono- and dihydroxylations, desaturation, and carbon-carbon bond-cleavage reactions. The structural and catalytic mechanisms by which oxygenases activate O_2 and control the subsequent reactions is the focus of this dissertation.

Some of the most important early studies of oxidative bacterial catabolism and oxygenases were performed by the researchers Osamu Hayaishi, Roger Stanier, Stanley Dagley, and others[17-20]. The discovery and biochemical characterization of the first oxygenase enzyme, catechol 1,2-dioxygenase (then called pyrocatechase), dates back to the early 1950s[17]. Prior to the detailed characterization of this enzyme, it was generally believed that O_2 could never be directly incorporated into organic substrates, and functioned solely as an oxidizing agent[21]. Instead, it was hypothesized that water was invariably integrated into the product, following the oxidation of substrate by O_2. Early experiments with $^{18}O_2$ isotopic labelling using the enzymes catechol 1,2-dioxygenase[22] and mushroom phenol hydroxylase[23] definitively showed that O atoms from molecular O_2, not solely water, could be directly incorporated into organic substrates.

O_2 activation by metal and organic enzyme cofactors. Why do nearly all oxygenases utilize transition metal or redox-active organic cofactors to activate O_2? The chemical logic behind this strategy is explained by looking at the electronic structure of O_2. The ground-state configuration of the molecule has two unpaired electrons in the highest occupied π^* molecular orbitals, leading to a triplet di-radical ground state with a net spin of $S = 1$ (Fig. 1, left). The triplet nature of the O_2 molecule is the key factor which results in its reactive inertness under normal conditions. The direct reaction of triplet O_2 with singlet organic molecules—which have only spin-paired electrons and an overall spin of $S = 0$—is spin-forbidden by quantum mechanics. Thus, although the overall reaction of O_2 with most organic substrates is thermodynamically favorable, the reactions are kinetically controlled by the incompatibility of the electronic ground-states of the reactants. In order to catalyze these reactions, oxygenases "activate" the O_2 molecule and/or organic substrate to a more reactive state.

Several general strategies for activating O_2 to react with organic molecules exist[21, 24] (Fig. 1, right): *(i)* electronic excitation of the O_2 from a triplet state to the higher-energy singlet state which can then directly react with other singlet molecules, *(ii)* one electron (e^-) reduction of the O_2 molecule by a metal or an external reductant to a reactive singlet superoxide radical state ($O_2^{\bullet-}$), *(iii)* two e^- reduction of O_2 and proton transfer to generate a singlet (hydro)peroxo (O_2^{2-} / HO_2^-), or *(iv)* direct transfer of electrons from an organic substrate to O_2 *via* a metal cofactor with vacant d orbitals (here the metal acts as a spin-capacitor and inverter to circumvent the spin-

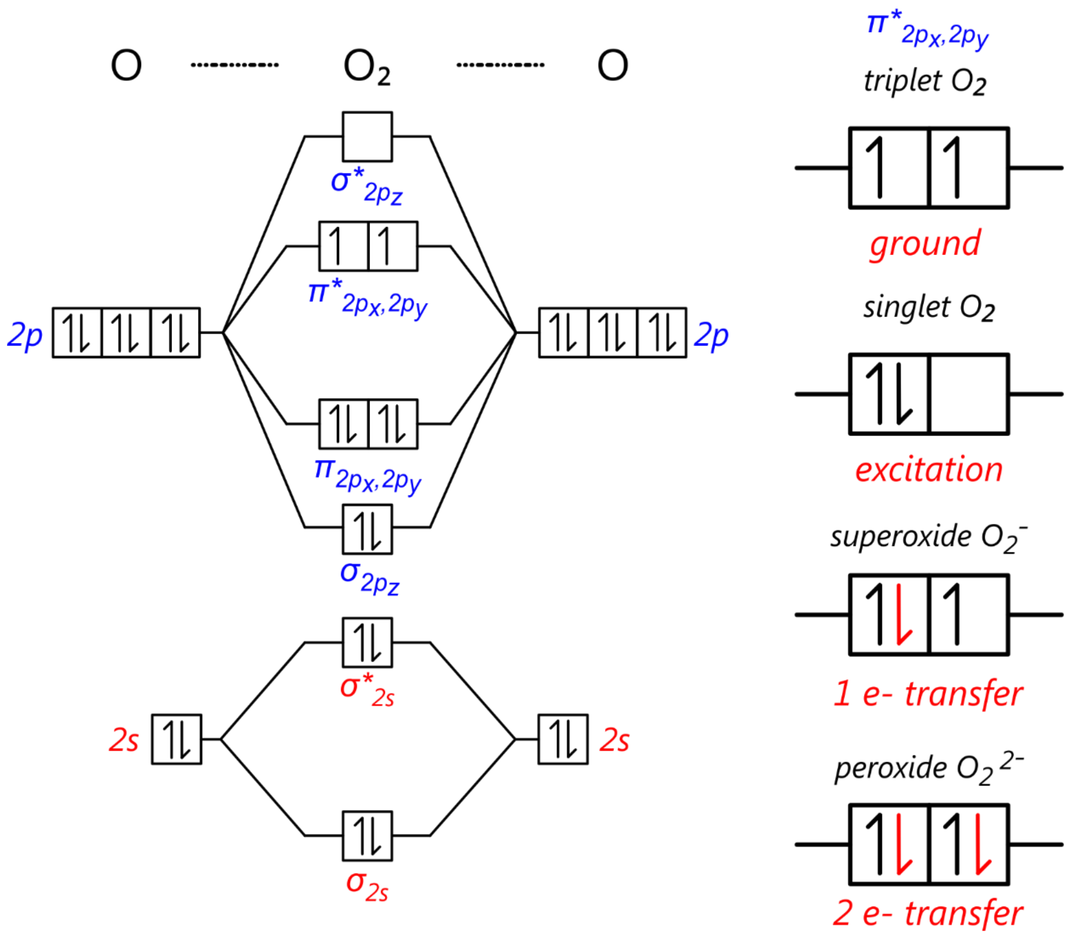

Fig. 1: Molecular orbital diagram of O_2 (left) and various activated forms of O_2 (right). The $\pi^*_{2px,\ 2py}$ antibonding orbitals are the highest occupied molecular orbitals of O_2.

forbidden reaction between the substrates)[25]. Strategy *(i)* is essentially unavailable to most biological catalysts as the singlet excited state lies an intractable 22 kcal mol^{-1} higher in energy. However, strategies *(ii)*, *(iii)*, and *(iv)* are utilized by oxygenases. This makes apparent the need for a metal cofactor: *electron transfer to O_2 is a critical step in the activation of the molecule.* Transition metals such as iron are able to fill this role because their partially filled d orbitals allow them to act as electron transfer conduits and in a redox capacity. Reduced metals such as Fe^{2+} are also able to directly bind O_2.

The source of the e^- for O_2 reduction varies widely depending on the family of enzymes. For instance, in the case of oxygenases with dinuclear metal cluster cofactors and Rieske enzymes, electrons are exogenously delivered by a cognate redox-partner protein such as a ferredoxin (Fd) or flavo-iron-sulfur protein, with the electrons ultimately deriving from

NAD(P)H[26-31]. These proteins deliver two reducing equivalents to the oxygenase component of the redox chain. In contrast, in the mononuclear α-ketoglutarate (α-KG)-dependent or aromatic ring-cleaving families of enzymes and others, organic (co)substrate is itself the source of the electrons[32, 33]. Most classes of oxygenases utilize a transition metal to facilitate the reaction between O_2 and organic substrates. However, some enzymes with redox-active organic cofactors such as flavins also catalyze oxygenase reactions[34]. There are also some oxygenases that do not require any metal or organic cofactor, although their substrate range is considerably more limited[35]. In these systems, the organic substrates themselves are sufficiently redox-active to activate O_2 for reaction.

The non-heme iron family of oxidases and oxygenases

Before exploring the variety of non-heme iron enzymes, it is important to briefly mention one of the largest families of oxygenases: the heme-containing cytochrome P450s. These enzymes are monooxygenases and catalyze a tremendous variety of different biological transformations in nearly all avenues of aerobic life[10]. These enzymes were among the first oxygenase systems to be characterized in biochemical detail[36, 37] and are still being studied today[38, 39]. Cytochromes utilize a heme cofactor, where the metal is coordinated by four modified pyrrole groups that are part of a porphyrin scaffold. The term ‘non-heme’, then, refers to enzymes that utilize iron that is not coordinated by a porphyrin prosthetic group.

Mononuclear non-heme iron oxygenase and oxidase enzymes. Mononuclear non-heme iron enzymes are those that utilize a single iron ion in the active site as a catalytic cofactor. This broad class of enzymes includes the aromatic ring-cleaving enzymes[32, 40], cysteine dioxygenases[41], lipoxygenases[42], aliphatic carbon-carbon bond-cleaving enzymes[43, 44], Rieske oxygenases[29, 45], α-KG and pterin-dependent hydroxylases and halogenases[33, 46], and some oxidases[47, 48]. The coordination mode of the metal ion varies within this family, but by far the most commonly encountered motif is the ‘facial triad,’ which is discussed in more detail in subsequent sections. The ferrous cofactor is the reactive form in all mononuclear non-heme iron enzymes, with two notable exceptions being the intradiol ring-cleaving dioxygenases and lipoxygenases[25].

Dinuclear iron-cluster containing oxygenase and oxidase enzymes. As suggested by the name, the diiron non-heme enzymes have active sites containing a catalytic dinuclear iron cluster. The diiron family of enzymes is broad, containing the bacterial multicomponent monooxygenases (BMMs)[49-53], the R2 subunit of class I ribonucleotide reductases[54], acyl-acyl carrier protein (acyl-ACP) desaturases[55], arylamine oxygenases[56, 57], and various other hydroxylases[58-62]. In most of these enzymes, the diiron cluster is coordinated in a "ferritin-like" 4-helix bundle structure. The metal-coordinating motifs nearly invariably contain clusters bridged by μ-oxo or hydroxo ions and/or protein carboxylates and the metal ions are coordinated by protein-derived His and carboxylate ligands, though the number, ratio, and geometric arrangement of these ligands varies dramatically between enzymes. With the exception of *myo*-inositol oxygenase [59], the diferrous form of the cluster is the state that reacts with O_2.

Common O_2 activation strategies and redox logic. The fundamental mechanism by which O_2 is activated can be reduced to a series of controlled electron and proton transfer steps. Protons play a critical part in this process, as they neutralize the buildup of negative charge on the O_2 molecule as it becomes reduced and thus more anionic. The timing and site of proton delivery is critical in driving catalytic progression and determining the nature of the activated O_2 species that is formed. O_2 reduction occurs in discrete one- or two-electron steps and, as noted above, the source of these electrons varies. Although the details of the oxygenation mechanism and the catalytic role of the protein environment differ greatly between the classes of non-heme iron enzymes, O_2 activation conforms to one of several canonical pathways[63]. This section will illustrate some of these pathways and provide examples as observed in biological systems.

(a) Iron-superoxide intermediates (Fig. 2)

One e^- reduction of the O_2 molecule results in formation of a superoxide anion radical ($O_2^{\bullet -}$). The most common mechanism by which superoxide is formed is by direct reaction of O_2 with a ferrous metal center, resulting in a Fe^{3+}-superoxide species (Fe^{3+}- $O_2^{\bullet -}$). Such species are commonly encountered as transient intermediates on the way to a peroxo or high-valent iron-oxo intermediate, and are only rarely kinetically stable enough to be characterized in detail[59, 64, 65]. Only in the diiron enzyme *myo*-inositol oxygenase (MIOX) has direct evidence for the involvement of a Fe^{3+}-superoxide species been obtained[59], but several mononuclear enzymes[65-68] have been proposed to utilize iron-superoxides in a similar fashion. MIOX catalyzes the four e^- oxidation of *myo*-inositol substrate with concomitant O incorporation to generate D-glucuronate[59]. This enzyme utilizes the mixed-valence cluster in the Fe^{2+}/Fe^{3+} state (Fig. 2A-*(i)*)

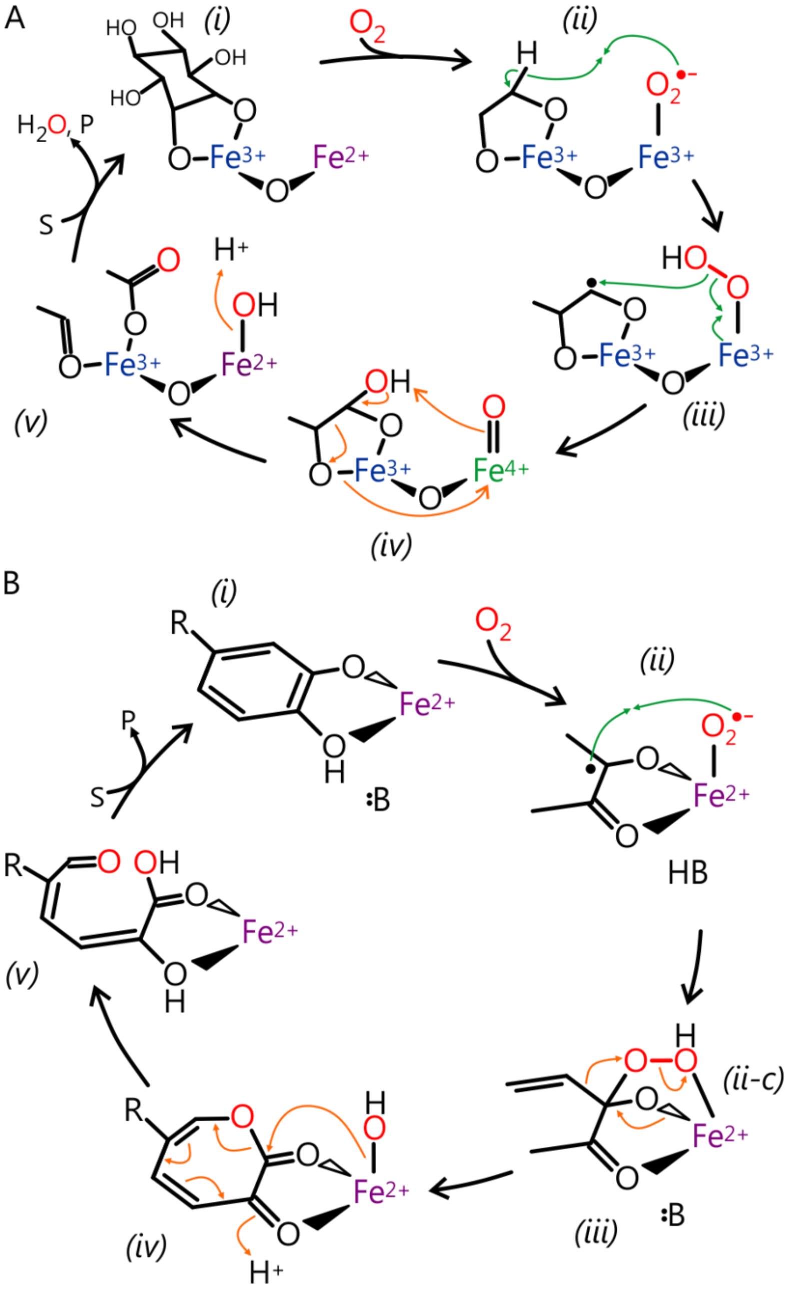

Fig. 2: Proposed catalytic mechanisms for selected enzymes that use a superoxo intermediate for substrate attack. (A) *Myo*-inositol dioxygenase (MIOX)[59, 69]. (B) Homoprotocatechuate 2,3-dioxygenase (2,3-HPCD)[64, 71, 72, 125]. The orange arrows show 2 e^- transfers and green arrows show 1 e^- transfers or radical recombinations.

to react with O_2 to form a Fe^{3+}/Fe^{3+}-superoxide species (Fig. 2A-*(ii)*) which then reacts directly with the organic substrate by hydrogen atom abstraction to generate a hydroperoxo species and substrate radical (Fig. 2A-*(iii)*)[59]. The hydroperoxo is proposed to then undergo homolytic scission and recombination with the substrate radical, resulting in either substrate hydroperoxylation or hydroxylation depending on where the scission occurs[69]. The hydroxylation path is shown in Figure 2A-*(iii)* to *(v)*. To date, this is the single characterized example of a diiron enzyme that uses the Fe^{3+}/Fe^{3+}-superoxide state to catalyze C-H bond activation.

There are several examples of mononuclear oxidases and oxygenases that use a Fe^{3+}- or Fe^{2+}-superoxide species to react with substrate. Two examples are the oxidases 1-aminocyclopropane-1-carboxylic acid oxidase[67] and isopenicillin N-synthase[66]. These enzymes are proposed to use the species to directly oxidize substrate to a radical *via* hydrogen atom abstraction, followed by further oxidation and (de)cyclization of the substrate to generate product. Fe^{3+}-superoxide species are also proposed to be utilized by cysteine dioxygenases, hydroxyethylphosphonate dioxygenase and the Rieske enzyme benzoate dioxygenase to perform the initial oxidation of substrate[65, 68, 70]. The Fe^{2+}-utilizing ring-cleaving dioxygenase homoprotocatechuate 2,3-dioxygenase (2,3-HPCD) generates a Fe^{2+}-superoxide species that reacts with substrate. This Fe^{2+}-superoxide is generated by an electron transfer from substrate to O_2 in which catechol is oxidized to semiquinone and O_2 reduced to superoxide (Fig. 2B-*(ii)*)[71]. The superoxo radical then recombines with the semiquinone radical to generate an Fe^{2+}-alkyl(hydro)peroxo (Fig. 2B-*(ii)* to *(iii)*) that drives the ring-cleavage reaction following heterolytic O-O bond scission (Fig. 2B-*(iv)* to *(v)*)[64, 72]. A Fe^{3+}-superoxide species has also been trapped in 2,3-HPCD but is only stabilized in a mutant of the enzyme when it is reacted with the alternative substrate 4-nitrocatechol[64]. However, this reaction results in hydrogen peroxide production and substrate oxidation to a quinone rather than ring-cleavage. If such a species is generated as a precursor to the Fe^{2+}-superoxide, it may simply be too short-lived to detect[71].

(b) Iron-peroxo intermediates (Fig. 3)

The two e^- reduced form of O_2, peroxide, is also a commonly encountered intermediate on the pathway to more potent oxidants. However, iron-peroxide species are also themselves recognized as capable of acting as organic oxidants in a variety of enzyme systems. The key step leading to the formation of peroxo-level intermediates is reduction of the O_2 molecule by two e^-. As such, the reactive Fe^{2+}/Fe^{2+} cluster state of diiron enzymes provides a convenient mechanism by which to transfer the two reducing equivalents necessary to form the peroxo intermediate. The result is a Fe^{3+}/Fe^{3+}-peroxo (or “diferric-peroxo”) species. Not all peroxos are created equal,

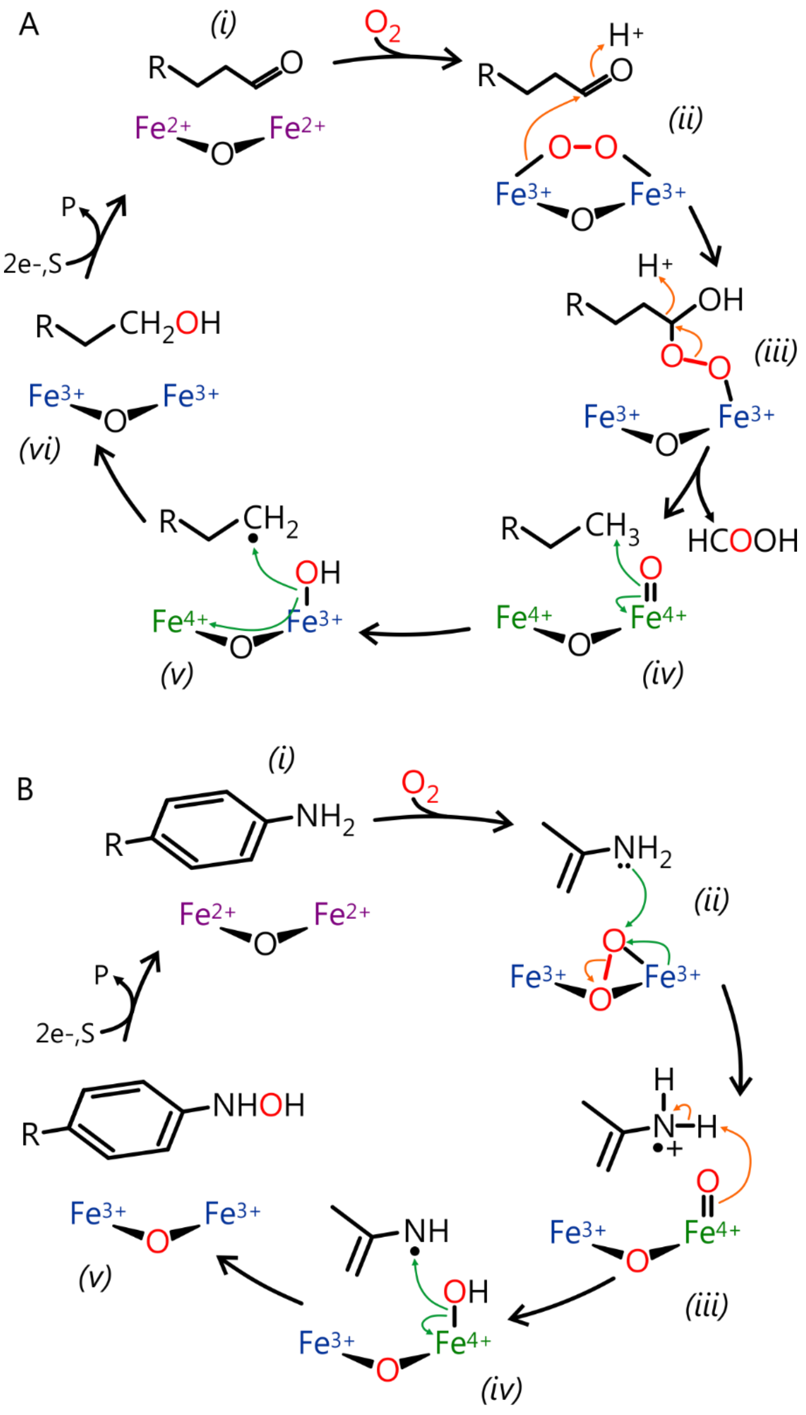

Fig. 3: Proposed catalytic mechanisms for selected enzymes that use a peroxo intermediate for substrate attack. (A) Aldehyde-deformylating oxygenase (ADO)[61]. An alternative mechanism to the one shown has also been proposed[77]. (B) Arylamine oxygenases[56, 57, 73]. The orange arrows show 2 e^- transfers and green arrows show 1 e^- transfers or radical recombinations.

however, and the coordination geometry and electronic properties of the resulting species varies between enzymes[57, 60, 73-75], as does their reactivity.

Several diiron enzyme systems have been described that generate a peroxo-level species that reacts directly with organic substrates. Deoxyhypusine hydroxylase (DOHH), an enzyme involved in regulating cell proliferation, is purified in the Fe^{3+}/Fe^{3+}-peroxo state. The peroxo state of the enzyme is highly stable and has been crystallographically and spectroscopically determined to be a *cis* μ-1,2-peroxo[60, 76]. The structure of a *cis* μ-1,2-peroxo is shown in Fig. 3A-*(ii)*. The DOHH peroxo is proposed to directly react with the protein substrate. The cyanobacterial aldehyde deformylating oxygenases (ADOs) generate a peroxo-level intermediate (Fig. 3A-*(ii)*) that is proposed to attack the terminal electrophilic carbon of the fatty aldehyde substrate, a process that leads to deformylation and alk(a/e)ne formation (Fig. 3A-*(iii)*)[61, 74, 77]. In order to explain the observed ADO product distribution, the formation of a high-valent intermediate (Fig. 3A-*(iv) to (v)*) has been theorized to form following deformylation[61]. The arylamine-oxidizing enzymes AurF and CmlI[57, 73] can generate relatively long-lived Fe^{3+}/Fe^{3+}-peroxo species upon exposure of the reduced Fe^{2+}/Fe^{2+} state to O_2 under certain conditions (Fig. 3B-*(ii)*). The proposed catalytic pathway for the first cycle of arylamine oxygenation is shown in Fig. 3B. Arylamine substrates have been shown to react directly with the CmlI peroxo[57], suggesting that it is the species that effects the initial oxidation, or that substrate binding induces conversion of the peroxo to a more reactive species. The peroxo in these enzymes is proposed to have a geometry distinct from the μ-1,2-type peroxos encountered in several other diiron enzyme systems[78-81]. A μ-$\eta1{:}\eta2$ peroxo has been proposed for CmlI and is shown in Fig. 3B-*(ii)*. These latter two enzyme examples elegantly showcase the ability of the peroxo to act as both a nucleophilic (ADOs) and electrophilic reactant (arylamine oxygenases).

Peroxo intermediates are also formed in the mononuclear iron enzymes but most often they are generated following O_2 activation and after substrate has already been oxidized. Peroxos have only been directly trapped and characterized in extradiol ring-cleaving dioxygenases[71, 72, 82] and the Rieske enzyme benzoate dioxygenase[83]. They have also been proposed to occur in the catalytic cycles of various other mononuclear enzymes[65, 84-88]. In many of these cases, peroxos would be generated following the initial generation of a Fe^{3+}-superoxo followed by substrate oxidation. Hydrogen atom abstraction from substrate would result in a metal-hydroperoxo (see Fig. 2A-*(iii)*)[68, 89, 90]. Alternatively, direct attack of activated O_2 on organic substrate generates an alkylperoxo or peroxohemiketal[72, 86]. The alkylperoxos generated in ring-cleaving enzymes decay to lactone or anhydride intermediates[25, 72, 91, 92] whereas many of the peroxos in other systems have been observed to (or are proposed to) undergo O-O bond cleavage to generate high-valent intermediates that effect substrate oxygenation[86-88, 93, 94]. In some cases, direct transfer of the activated hydroperoxo moiety to substrate ("hydroperoxylation") has been proposed as an alternative oxygenation mechanism[69, 95].

(c) High-valent Fe^{4+}*- or* Fe^{5+}*-oxo/hydroxo intermediates (Fig. 4)*

One of the key differences in the enzyme systems that access high-valent iron intermediates is that these systems utilize a total of four reducing equivalents to activate O_2 *prior* to O_2 attack on the substrate[96, 97]. This mechanism results in O-O bond scission before the substrate is oxygenated and the formation of Fe^{4+}- or Fe^{5+}-oxo/hydroxo species. For example, in the diiron methane monooxygenase hydroxylase, a total four reducing equivalents are injected into the O_2 moiety, breaking the O-O bond and converting the reactive Fe^{2+}/Fe^{2+} cluster to Fe^{4+}/Fe^{4+}. The analogous process in the mononuclear enzymes—which lack a second metal to store the additional electrons—requires 2 e^- oxidation of an organic cosubstrate or an exogenous source of 2 e^- to generate the full 4 e^-. During O_2 activation, the mononuclear iron cofactor cycles from the Fe^{2+} state to Fe^{4+}.

A critical step leading to the formation of high-valent intermediates is scission of the O-O bond of the metal-bound peroxo moiety. The cleavage mechanism varies between enzymes and several versions have been identified. The first is reductive scission of the O-O bond driven by electron donation/transfer to the peroxide moiety from the metals (e.g. the Fe^{3+}/Fe^{3+} cluster converts to Fe^{4+}/Fe^{4+})[81, 98, 99]. Alternatively, cleavage can be driven by nucleophilic attack of one O atom of the peroxo on an electrophilic substrate[61, 74]. Decay can also be induced by protonation of the peroxo itself[100] or a first-sphere iron ligand[101, 102]. The latter could engender a ligand and/or peroxo rearrangement that drives O-O scission[98]. The result of the scission is a dinuclear Fe^{4+}/Fe^{4+} or Fe^{3+}/Fe^{4+} cluster with terminal or bridging oxo ligand(s). In the case of most mononuclear enzymes, the reactive substrate-oxidizing species contains a single ferryl (Fe^{4+}=O) or Fe^{4+}-hydroxo unit. Such species are among the most potent chemical oxidants generated in all of biology and are tasked with performing the most demanding of oxygenation reactions. Fe^{4+}/Fe^{4+}-oxo/hydroxo and mononuclear Fe^{4+} intermediates have been characterized in a few enzyme systems and biomimetic inorganic model compounds[86, 103-108]. High-valent intermediates are also proposed in ADO and the arylamine oxygenases following initial substrate attack by the peroxo (see Fig. 3A-*(iv)* to *(v)* and 3B-*(iii)* to *(iv)*) but have not been trapped for study.

In the diiron-cluster containing enzymes, the most remarkable of the high-valent intermediate is the bis-μ-oxo Fe^{4+}/Fe^{4+} cluster of soluble methane monooxygenase hydroxylase

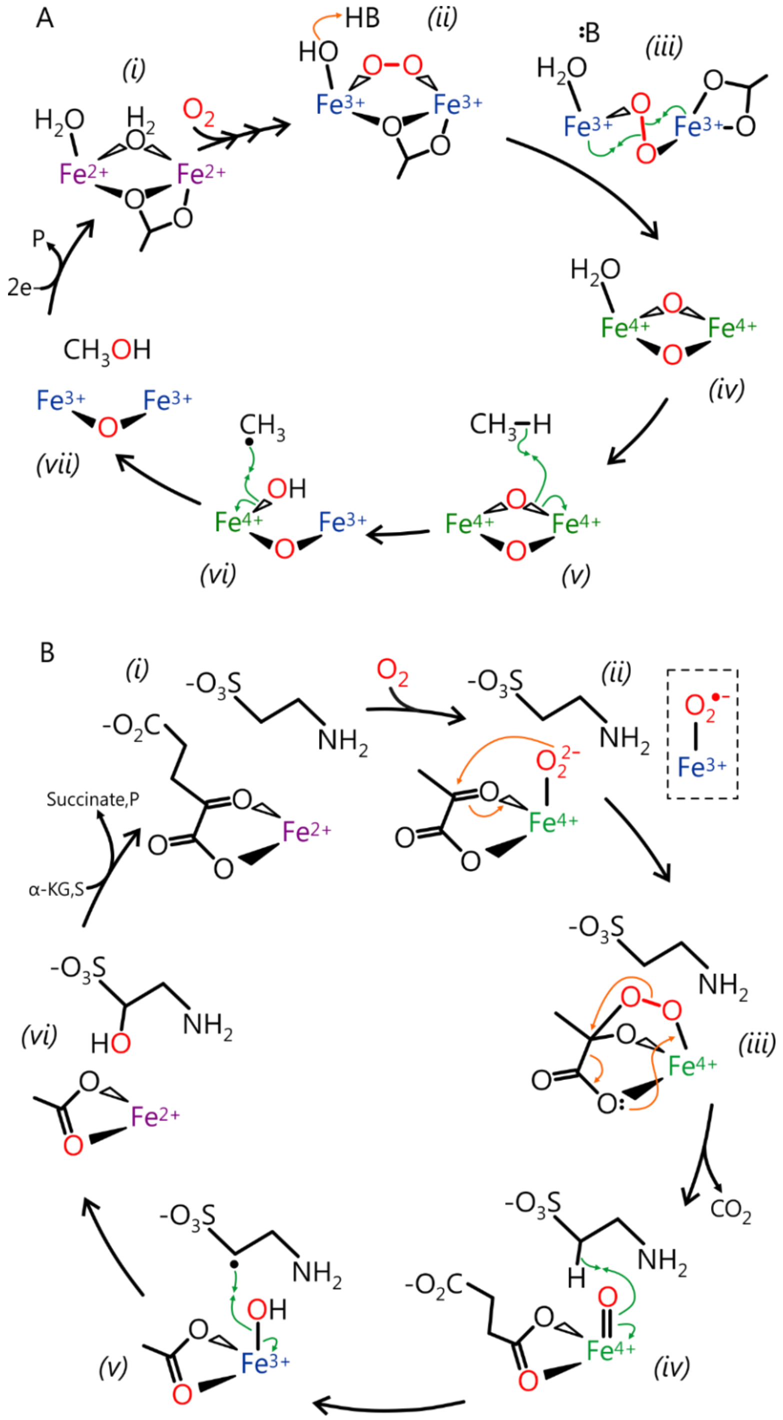

Fig. 4: Proposed catalytic mechanisms for selected enzymes that use high-valent iron species for substrate attack. (A) Soluble methane monooxygenase hydroxylase (sMMOH)[75, 98, 103, 104, 111-113]. (B) Taurine:α-ketoglutarate dioxygenase (TauD)[86, 126]. The orange arrows show 2 e$^-$ transfers and green arrows show 1 e$^-$ transfers or radical recombinations.

(sMMOH), termed intermediate **Q** (Fig. 4A-*(iv)*)[98, 103, 104]. This species is responsible for effecting the oxidation of methane to methanol in a variety of methanotrophic bacteria, thereby allowing these organisms to utilize methane as the sole source of carbon. The kinetics of the catalytic cycle of sMMOH have been extensively studied for nearly thirty years[109, 110]. The enzyme is proposed to pass through several intermediates prior to formation of **Q**, including a *cis* μ-1,2 peroxo complex termed **P** (Fig. 4A-*(ii)*)[75, 79]. Although the conversion of **P** to **Q** was known to be proton-dependent[101], it was only recently that the details of the mechanism of the conversion have emerged, following determination of the long-sought vibrational spectrum of **Q**[98]. These studies indicate that formation of **Q** occurs by homolytic cleavage of the peroxo O-O bond, triggered by protonation of a first-sphere carboxylate ligand or metal-bound solvent molecule (Fig. 4A-*(iii)* to *(iv)*). The conversion of **P** to **Q** may proceed *via* a transient *trans* μ-1,2 peroxo which is depicted in Fig. 4A-*(iii)*[98]. **Q** is proposed to react with the methane substrate by hydrogen atom abstraction and subsequent radical recombination (Fig. 4A-*(v)* to *(vii)*)[111-113]. Although **Q** is currently the only Fe^{4+}/Fe^{4+} cluster that has been directly characterized, **Q**-like species are proposed be generated in the catalytic cycle of other diiron enzymes[58, 61].

Ribonucleotide reductase (RNR) is an $(\alpha\beta)_2$ enzyme that catalyzes the reduction of ribonucleotides to the deoxyribonucleotides required for DNA synthesis. The R2 (β_2) subunit of RNR (RNR-R2) binds a diiron cluster that is necessary for generation of a quasi-stable tyrosine radical cofactor (Y122 in the *E. coli* enzyme) that propagates to the R1 subunit where ribonucleotide reduction occurs[114-117]. RNR-R2 is technically an oxidase but is noteworthy because it uses a high-valent intermediate to generate the catalytic Tyr radical. Upon exposure of the apo RNR-R2 to O_2 and free Fe^{2+}, the Tyr radical cofactor is spontaneously assembled along with the diiron cluster. During assembly, a short-lived Fe^{3+}/Fe^{3+}-peroxo intermediate is formed which subsequently decays to a high-valent intermediate termed **X** following injection of one electron into the system[99, 118]. The peroxo intermediate has been characterized in wild-type (WT) enzyme and in first-sphere mutants of this enzyme and found to be a *cis* μ-1,2 Fe^{3+}/Fe^{3+}-peroxo[119, 120]. Compound **X** contains a high-valent Fe^{3+}/Fe^{4+}-oxo bridged cluster that is competent to perform the oxidation of Y122 to the catalytic radical with coincident decay of the cluster to the oxo-bridged diferric state[105, 121, 122]. However, in certain mutants of the enzyme, the reaction partitions between self-hydroxylation of the protein and radical generation, showing that the cluster is also competent to oxidize aromatic C-H bonds[99, 123].

Ferryl species are generated in many mononuclear iron enzymes. As noted above, these enzymes generate ferryl species at the cost of an organic cosubstrate that is oxidized by two e^- to break the O-O bond[124, 125]. Commonly encountered organic cofactors include α-KG and tetrahydrobiopterin. During the catalytic cycle, α-KG- and pterin-dependent oxygenases oxidize these cosubstrates to succinate/CO_2 and dihydrobiopterin, respectively[86, 93, 106, 124]. The consensus mechanism for the α-KG-dependent enzymes posits formation of a short-lived Fe^{4+}-peroxo species (or an isoelectronic Fe^{3+}-superoxo) as shown in Fig. 4B-*(ii)*[86]. The terminal oxygen atom of the metal-bound activated O_2 species rapidly attacks the α-keto carbon of the organic cofactor leading to oxidative decarboxylation and heterolytic O-O bond cleavage (Fig. 4B-*(iii)*). This process generates a ferryl species (Fig. 4B-*(iv)*) that subsequently oxidizes the substrate using hydrogen atom abstraction (Fig. 4B-*(iv)* to *(vi)*). This results in a Fe^{3+}-hydroxide and organic radical. Substrate is then hydroxylated *via* radical recombination, a process that also regenerates the Fe^{2+} cofactor (Fig. 4B-*(vi)*)[126]. Homologous α-KG-utilizing halogenases are also proposed to use this mechanism but with the key difference being rebound of a metal-bound halogen on the organic radical instead of the Fe^{3+}-hydroxide[107]. The pterin-dependent aromatic amino acid hydroxylases also generate a reactive ferryl species but the steps leading to O_2 activation and the mechanism of O-O bond scission are different[106, 124].

Rieske oxygenases contain a mononuclear iron cofactor where the O_2 reaction takes place and also bind a Fe_2S_2 cofactor (the eponymous “Rieske” cluster)[29, 45, 127]. The Rieske cluster stores a single reducing equivalent and is located near the mononuclear iron in the protein. Based on work with the enzymes naphthalene dioxygenase and benzoate dioxygenase, the reaction with O_2 is proposed to proceed through a Fe^{3+}-(hydro)peroxo intermediate[128]. This species is isoelectronic to an Fe^{5+}-oxo hydroxo which is proposed to catalyze the dearomatization and *cis*-dihydroxylation of the aromatic substrates[83, 127, 128]. The overall reaction oxidizes both the Rieske and mononuclear iron. Subsequent reduction of the latter cofactor is required for product release[128]. Recently, further studies with benzoate dioxygenase have suggested the possibility that a Fe^{3+}-superoxo species may be responsible for catalyzing the initial attack on substrate instead of a high-valent species[70]. This assignment was based on the observation that the rate-determining step is substrate attack/oxidation, and that this event precedes Rieske cluster oxidation[70]. Formation of a reactive Fe^{3+}-(hydro)peroxo or Fe^{5+} species before attack on the substrate requires a net donation of two e^- to the oxygen, one of which would necessarily derive from oxidation of the Rieske cluster before the attack on substrate.

Regulatory strategies used by oxygenases to control O_2 activation. One of the hallmarks of O_2-activating enzymes is the regulatory mechanism used to prevent the adventitious formation of potent oxidizing species in the absence of the cognate substrate. Almost invariably, enzymes that catalyze such reactions have been found to utilize some form of catalytic regulation that ensures O_2 activation only proceeds rapidly in the presence of substrate(s). The need for such mechanisms is obvious: uncontrolled generation of reactive oxygen species is detrimental to the enzyme function and the cell itself, should the activated O_2 species be released from the active site of the enzyme[129]. There are numerous examples of self-hydroxylation of enzymes or uncoupling of the reaction in mutated enzymes where the regulatory mechanism is abolished or where enzymes in the reduced Fe^{2+} state are incubated with O_2 for extended periods in the absence of substrate[123, 130-132]. Of positive note, self-hydroxylation has been proposed as a convenient probe for the formation of high-valent oxidizing intermediates[133]. This section will explore common mechanisms by which oxygenases control O_2 reactivity, again using several enzyme families as examples.

(a) Regulation in mononuclear enzymes with a facial triad iron-binding motif (Fig. 5A)

This enzyme class, though incredibly diverse in terms of catalytic capabilities, utilizes the same iron-binding motif: the “facial triad” in which the metal is an octahedral coordination environment where three sites are coordinated by protein-derived ligands in a facial mode and the other three sites by solvent[96, 125, 134]. The most common ligand set contains two His and one carboxylate (Asp/Glu), though several enzymes with alternative sets have been characterized[135]. The regulatory logic for these enzymes can be reduced to a relatively simple paradigm: the Fe^{2+} coordination sphere remains saturated, i.e., there are no open coordination sites on the metal for O_2 to bind, until substrate has bound to the metal or in the active site but not on the metal. The facial coordination mode of the protein ligands facilitates this mechanism because it offers three available sites in a *cis*-arrangement on the metal. Most often, two are reserved for a substrate chelate, and O_2 reacts in the remaining site. A common pathway is *(i)* direct chelation of cosubstrate (if required) and substrate to the metal (or in some cases in the active site but not directly to the metal), *(ii)* dissociation/displacement of the metal-bound solvent molecules to open a site on the metal *cis* to the substrate or cosubstrate and *(iii)* binding of O_2 to the metal followed by electron transfer and O_2 activation. This type of mechanism has been observed in the α-KG- and pterin-dependent oxygenases[86, 136, 137], cysteine dioxygenase[84, 138], extradiol dioxygenases[91, 139], Rieske oxygenases[29, 140], aliphatic C-C bond-cleaving oxygenases[85, 141], and oxidases[47, 48, 142, 143]. It should be noted that in the Rieske enzyme naphthalene dioxygenase (NDO) and likely other members of the family, substrate must be bound in the active site *and* both the mononuclear iron and the Rieske Fe_2S_2 cluster must be in the reduced state, implicating a second layer of regulation[29, 144]. In addition to controlling the *timing* of O_2 activation, it is believed that direct coordination of the substrate to the metal has two more major effects on the fate of the reaction[66, 91, 138, 145-149]: *(i)* it properly aligns substrate for the reaction with O_2, ensuring regio- and stereospecificity of the reaction product and *(ii)* direct coordination or chelation of the substrate changes the electronic properties of the metal to make it more reactive towards O_2.

(b) Regulation in diiron bacterial multicomponent hydroxylases (Fig. 5B)

The members of the BMM enzyme family catalyze some of the most potent oxygenase chemistries in nature. It is no surprise, then, that they have developed some of the most sophisticated regulatory mechanisms. The most extensively studied regulatory mechanism is that utilized by the sMMOH enzyme system from *Methylosinus trichosporium* OB3b and *Methylococcus capsulatus* (Bath)[109]. O_2 activation in the hydroxylase component is controlled by a small (16 kDa) regulatory protein named MMOB[150]. Although the hydroxylation reaction of a variety of unactivated C-H bonds still occurs in the absence of MMOB[28, 151, 152], addition of MMOB to the system has dramatic effects on the reaction product distribution and maximal enzyme activity[150, 153], the diiron cluster electronic properties[154-156], and the rate constants of several discrete steps in the catalytic cycle[157, 158]. The primary roles of MMOB have been determined to be *(i)* to bind to MMOH, causing a structural change that generates a molecular protein "sieve" that preferentially gates access for small hydrocarbon substrates to the active site and promotes quantum tunneling of the methane hydrogen atom during the reaction with **Q**[158-160] and *(ii)* to increase the reactivity of the active diferrous cluster of MMOH towards O_2 1000-fold by inducing a structural rearrangement of the cluster ligands and lowering the iron redox potential[157, 161, 162]. Hence, the role of MMOB is multi-fold: to ensure entry of only the proper physiological substrate to the active site and permit rapid O_2 reaction/activation only when the protein "gate" is in place. Similar regulatory mechanisms have been studied in other BMM enzymes including toluene-4-monooxygenase[53], phenol hydroxylase[51], and toluene/*o*-xylene monooxygenase[52].

(c) Regulation by binding of a substrate-carrier protein (acyl-ACP desaturase) (Fig. 5C)

The acyl-acyl carrier protein (acyl-ACP) desaturases are diiron enzymes with a protein fold and active-site cluster architecture similar to the BMMs but have a modified regulatory mechanism. The acyl-ACP desaturases oxidize ACP-tethered fatty acids to their corresponding desaturated analogues. The fatty acid substrates are tethered to an ACP *via* a covalent phosphopantetheine linker[163, 164]. These enzyme systems exemplify the third regulatory strategy: systems in which the

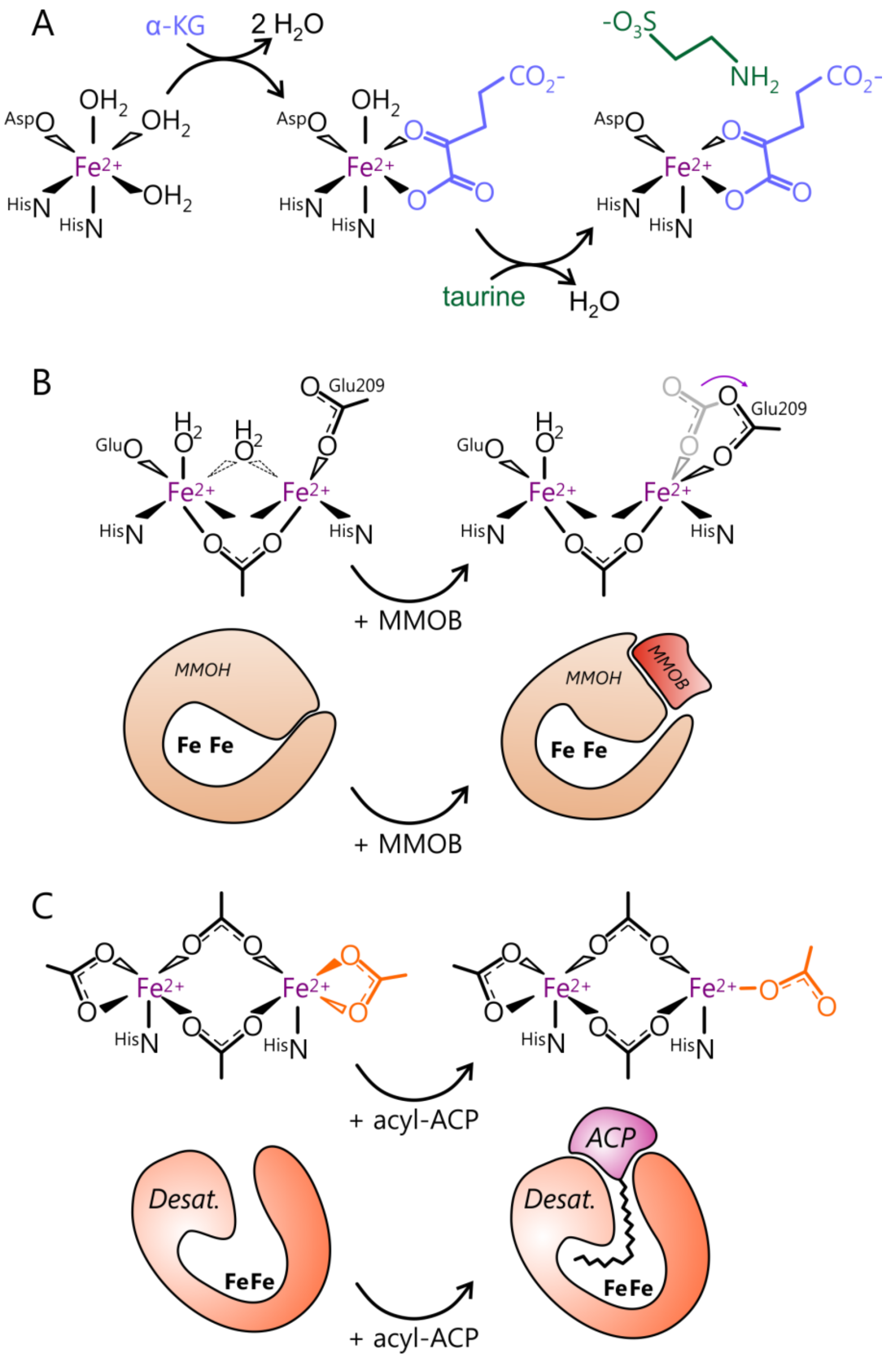

Fig. 5: Examples of strategies used by different oxygenase families to control O_2 activation. (A) Taurine:α-ketoglutarate dioxygenase (TauD) and other α-KG utilizing oxygenases[86]. (B) Soluble methane monooxygenase hydroxylase (sMMOH)[150, 153-158, 160]. (C) Acyl-acyl carrier protein (ACP) desaturase[81, 166-168].

substrate is tethered to another protein. The most studied desaturase is stearoyl-ACP Δ9-desaturase from *Ricinus communis* (castor plant)[55, 165]. Binding of the acyl-ACP to the desaturase enzyme induces structural changes in the cluster that increase the reactivity of the cluster towards O_2[81, 166, 167], similar to the other mechanisms discussed above. The details of the structural changes induced in the desaturase upon acyl-ACP binding have not been definitely determined, but spectroscopic studies have suggested that the cluster ligands undergo a rearrangement causing a five-coordinate to four-coordinate transition of one of the metal ions[166]. X-ray crystallographic studies of Δ9-desaturase have shown that the protein-protein interaction is a critical component that controls O_2 activation as well as the regiospecificity of the desaturation reaction[168]. Thus, the regulatory mechanism and the product specificity are controlled to a significant extent by the interaction with the protein carrier and structural changes induced at the protein-protein interface[168]. This regulatory theme will emerge again when discussing the non-ribosomal peptide synthetase-associated oxygenases in subsequent sections.

Protocatechuate 3,4-dioxygenase

Protocatechuate 3,4-dioxygenase belongs to the intradiol dioxygenase (IDO) family of aromatic ring-cleaving enzymes. It is a mononuclear non-heme iron-containing enzyme that uses a Fe^{3+} cofactor to catalyze the ring-cleavage reaction, and a member of one of only two families of enzymes that utilize the oxidized form of the cofactor to initiate catalysis[25, 169]. IDOs function in the final step in the *ortho*-cleavage pathway of aromatic compounds in many species of soil bacteria and fungi, and allow the linearized products to be funneled into the organism’s central metabolism *via* the β-ketoadipate pathway (Fig. 6, bottom)[32, 40, 170-172]. The 3,4-PCD from the soil bacterium *Pseudomonas putida* is the most thoroughly characterized IDO. Its physiological role is to cleave protocatechuate (PCA, 3,4-dihydroxybenzoate) to the tricarboxylic ß-carboxy-*cis*,*cis*-muconate product (Fig. 6)[19]. PCA is a common catabolite of an enormous variety of aromatic and polyaromatic monomers that are encountered by soil-dwelling organisms, including those formed during the depolymerization and breakdown of lignin (Fig. 6, green box) and the biosynthesis of aromatic amino acids (Fig. 6, red box)[173]. Because of their fundamental role in the metabolism of environmental aromatics including lignin catabolites and anthropogenic pollutants (Fig. 6 orange, black, and blue boxes)[171, 173-176], dissection of the catalytic mechanism of these enzymes is a prerequisite to engineering the proteins (or entire host organisms) for environmental cleanup and

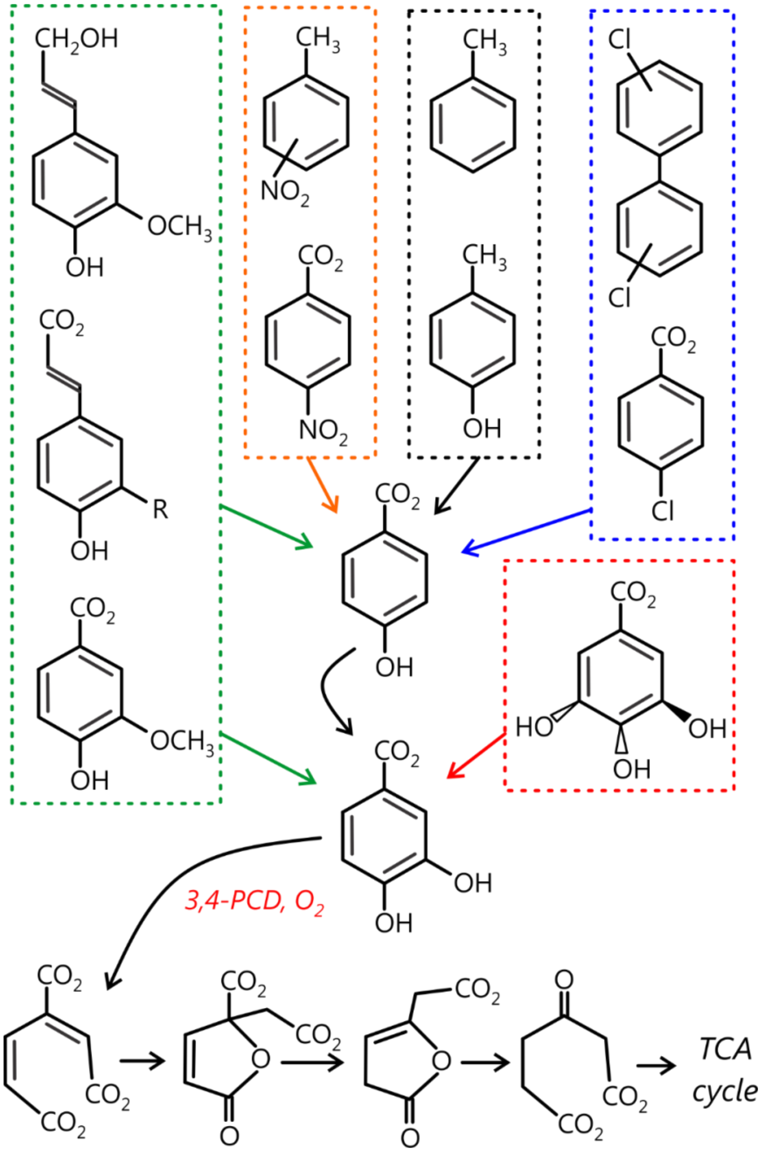

Fig. 6: Metabolic funneling of aromatic compounds into the *ortho*-cleavage pathway initiated by 3,4-PCD. The β-ketoadipate pathway is shown at the bottom of the figure[18, 19, 170].

bio-based industrial applications[177, 178]. Studies of 3,4-PCD and its relatives have revealed a wealth of information, including the kinetic steps in substrate binding and product release[179-184], the electronic properties of the iron cofactor in several states[185-189], and insights into the oxygen activation mechanism[25, 32, 189-195].

IDO protein fold, substrate specificities, and active site architecture. IDOs are characterized by a core fold composed of two sets of mixed-type β-sheets. This fold has only been observed in members of the family[40, 196]. The structurally characterized members fall into three distinct categories based on their overall organization. The first class is typified by the catechol 1,2-dioxygenases (1,2-CTDs) which preferentially cleave unsubstituted or methyl-substituted substrates[197-199] as well as those members which preferentially cleave hydroxyquinol[200] or halo-substituted catechols[201, 202]. This class dimerizes *via* an α-helical N-terminal domain first identified in the *Acinetobacter sp.* ADP1 enzyme[197]; the iron cofactor is located in the C-terminal catalytic domain (Fig. 7A). The second class contains the protocatechuate dioxygenases[196, 203], which are αβ heterodimers (Fig. 7B) and lack the explicit dimerization domain observed in the first class. Members of this group most often exist as large oligomeric $(\alpha\beta)_{2\text{-}12}$ assemblies of the fundamental αβ-Fe^{3+} catalytic unit[184, 193, 196, 204, 205]. The 3,4-PCD from *P. putida* is a $(\alpha\beta)_{12}$ multimer (Fig. 7C)[196, 206]. These α and β subunits are structurally homologous to one another. Interestingly, the 1,2-CTD from *P. arvilla* exists in three dimeric isoforms αα, αβ, ββ, highlighting a possible evolutionary linkage between 1,2-CTDs and 3,4-PCDs[198]. The final structural class of IDOs currently consists of only a single, recently characterized member that preferentially cleaves catecholic derivatives of lignin biosynthetic intermediates[176]. Unlike all other characterized IDOs, it is a monomer (Fig. 7D) and contains a C-terminal carbohydrate binding motif, which is implicated in targeting the enzyme to lignin surfaces. The active site of this enzyme is much more solvent exposed, lacking the tight substrate "pocket" observed in other IDOs. This is consistent with its apparent preference for large catecholic substrates.

All IDOs have a conserved metal-ligation sphere. The resting ferric cofactor is coordinated by a 2-His 2-Tyr motif and a solvent-derived hydroxide (Fig. 8-*(i)*)[176, 196, 197]. This resting ligand set leads to a charge-neutral iron center, a property that has profound relevance to the catalytic cycle of the enzyme[179, 194]. The resting ferric ion is bound in a distorted 5-coordinate trigonal bipyramidal coordination environment. Y447 and H462 (residue numbering for 3,4-PCD from *P. putida*) are the axial ligands and Y408 and H460 along with solvent hydroxide are equatorial ligands. IDOs are characterized by a broad optical absorbance centered near 450 nm which confers a burgundy color to the enzyme. The chromophore derives from ligand-to-metal charge-transfer (LMCT) bands from both Tyr ligands to Fe^{3+} and provides a convenient probe for monitoring changes in the coordination sphere and binding of substrate[185, 186]. The presence of two Tyr ligands in the IDOs also dramatically lowers the redox potential of the metal (-0.5 to -1.1 V)[207, 208]. As a final note, alongside IDOs, there are only a few characterized enzyme

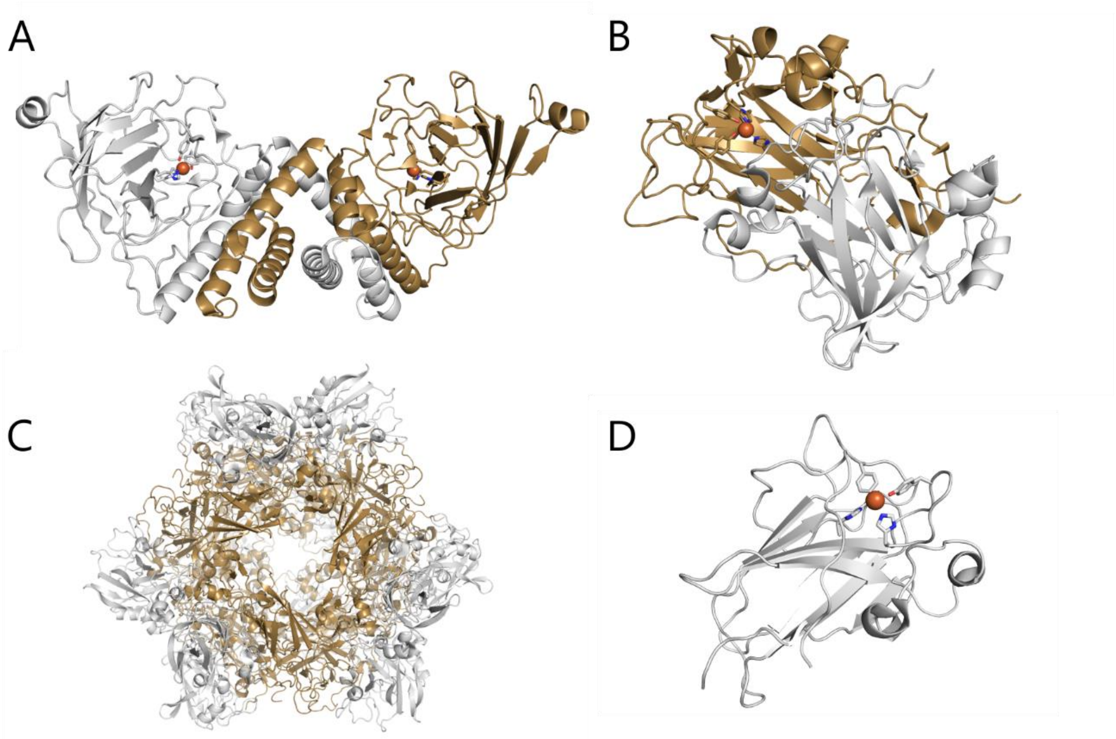

Fig. 7: Crystal structures of various intradiol dioxygenases. (A) Homodimer of catechol 1,2-dioxygenase from *Acinetobacter* sp. ADP1 (PDB structure 1DLM) . (B) The catalytic αβ heterodimer unit of protocatechuate 3,4-dioxygenase from *Pseudomonas putida* (2PCD). (C) The $(\alpha\beta)_{12}$ biological assembly of 3,4-PCD. (D) Dioxygenase domain of SACTE_2871 from *Streptomyces* sp. SirexAA-E (4ILV). In (B) and (C) the α-subunit is shown in white and β-subunit in brown. The iron cofactor is shown as a brown sphere.

families that have Tyr-metal ligands; these include the purple acid phosphatases and galactose oxidases[209, 210].

Proposed catalytic mechanism of 3,4-PCD. The reaction mechanism of 3,4-PCD is ordered sequential, with PCA binding before O_2. Binding of PCA is proposed to occur in several steps, with PCA initially coordinating the Fe^{3+} as a $4O_{PCA}$ monoanion, followed by a shift to a chelating mode that displaces Tyr447. This process results in coordination of the $4O_{PCA}$ *trans* to H462 and $3O_{PCA}$ *trans* to Tyr408. In this enzyme-substrate (E_{PCA}) complex (Fig. 8-*(ii)*), the plane of the aromatic ring of PCA is aligned with the axial direction of the iron coordination. The symmetry of the iron shifts to octahedral with one open site *trans* to H460, which is proposed as the site for eventual reaction with oxygen[181, 182, 203]. However, as O_2 has no affinity for Fe^{3+}, it is important to realize that there has never been any direct spectroscopic evidence for an Fe^{3+}-O_2 complex or

Fe^{2+}-semiquinone radical character in the E_{PCA} complex which might allow direct O_2 binding to the iron[189, 191, 192]. This proposal is supported by several computational studies[25, 188, 211]. The E_{PCA} complex is characterized as a Fe^{3+}-catecholate complex with significant covalent bond character[188], consistent with the spectroscopic properties observed for the complex[189, 212]. However, some dispute on this matter remains, as other calculations have suggested that E_{PCA} does indeed have significant Fe^{2+} character[213, 214]. However, the extremely low redox potential of the metal, and the lack of any spectroscopic evidence for Fe^{2+} in the E_{PCA} complex makes this assignment unlikely to be correct[25, 189, 192, 193, 207]. A key feature of the E_{PCA} complex is an asymmetry in the binding mode of the catecholate: the Fe^{3+}-$O3_{PCA}$ bond is considerably longer than the -$O4_{PCA}$ bond[181]. This apparent weakening of the Fe^{3+}-$O3_{PCA}$ bond is proposed to facilitate later dissociation of the ketonized form of $O3_{PCA}$ from the metal during O_2 activation.

Spectroscopic studies on 3,4-PCD have provided evidence that tyrosine coordination is a necessary feature in the O_2 activation step[185, 186, 188]. Upon formation of the E_{PCA} complex, additional low-energy charge transfer bands (at 500+ nm) arise from multiple catecholate-to-Fe^{3+} LMCTs[212, 215], with the lowest energy CT band proposed to provide a pathway for delivery of the β-spin electron from substrate to the iron during O_2 activation, *vide infra*[188]. The reaction with O_2 is proposed to lead to a bridging Fe^{3+}-alkylperoxo-PCA species (Fig. 8-*(iii)*)[25, 181, 189, 216]. Formation of the alkylperoxo induces a shift of PCA back to the equatorial plane (Fig. 8-*(iv)*)[25, 213]. This shift is facilitated by a ketonization of $3O_{PCA}$ and a lengthening of the bond to Fe^{3+} in a step that is essentially the reverse of the substrate binding process that initially resulted in the PCA chelate. This leads to the formation of a 5-coordinate complex with the axial site *trans* to His462 vacant (where Y447 coordinates Fe^{3+} in the resting state) and $4O_{PCA}$ *trans* to Y408 (Fig. 8-*(iv)*). Formation of the alkylperoxo intermediate is proposed to occur *via* a concerted two e^- transfer process with the Fe^{3+} acting as an electron conduit and spin inverter leading to the simultaneous formation of both the Fe^{3+}-O_{O2} (proximal) and $C4_{PCA}$-O_{O2} (distal) bonds (Fig. 9)[25, 188]. This net two-electron transfer process occurs through three simultaneous single-electron transfer processes: *(i)* transfer of a β-spin electron from the PCA $\pi_{op\text{-}sym}$ HOMO to the iron d_{xz} orbital *via* the lowest energy PCA-Fe^{3+} LMCT bands (Fig. 9 red arrow), *(ii)* simultaneous delivery of an α-spin electron from the iron d_{z2} orbital to the HOMO of O_2 (Fig. 9 green arrow), and *(iii)* transfer of an α-electron from the substrate to the O_2 HOMO with the second unpaired β-electron (Fig. 9 blue arrow). This electron transfer process leads to a net $S = 3/2$ spin system. This state is stabilized by ketonization and dissociation of $O3_{PCA}$ from the iron, resulting in a monodentate coordination mode of the substrate *via* $O4_{PCA}$. The open sixth coordination site on

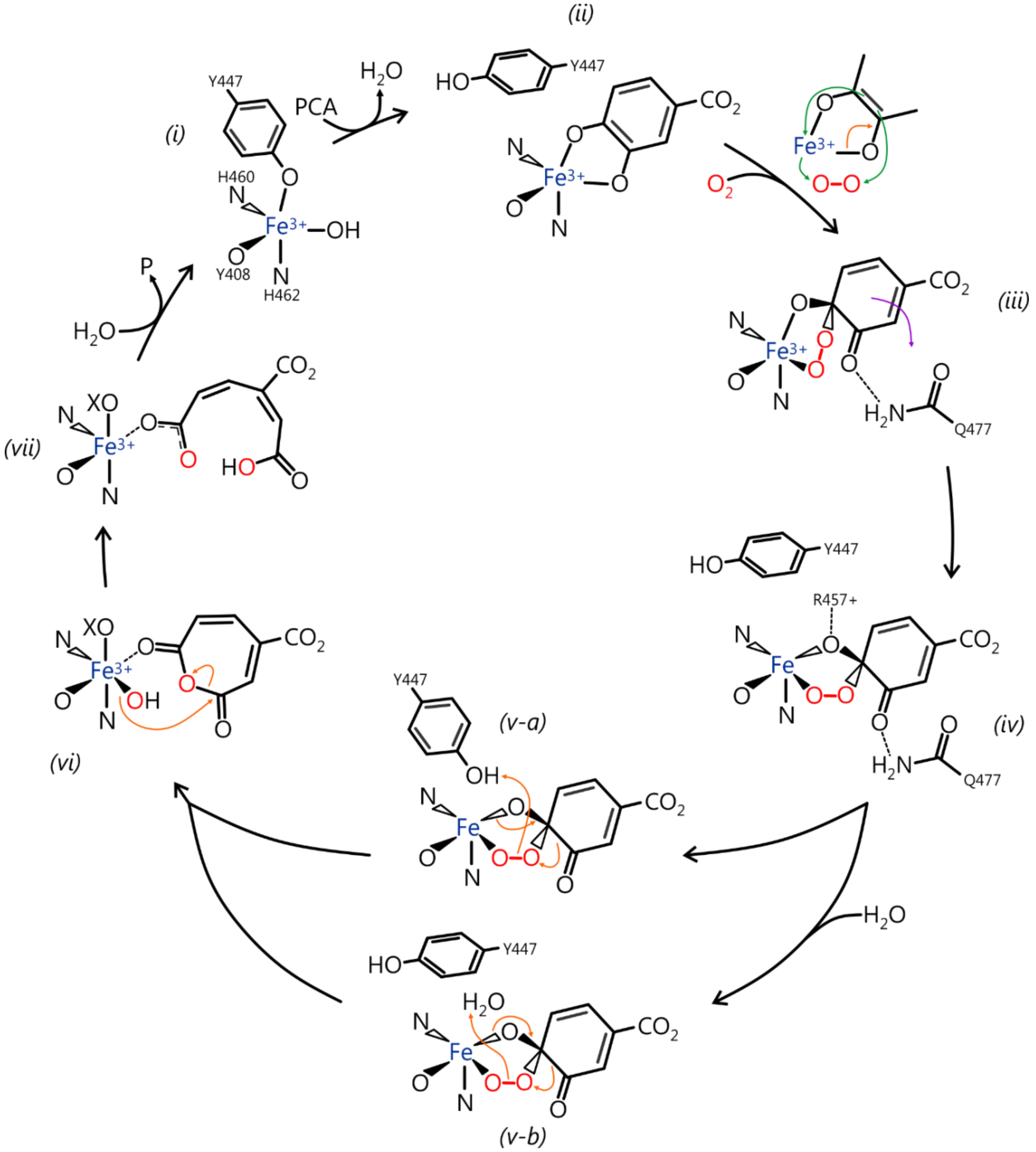

Fig. 8: Proposed catalytic mechanism for cleavage of PCA by 3,4-PCD[25, 179, 181, 189, 213]. The orange arrows show 2 e$^-$ transfers and green arrows show 1 e$^-$ transfers. The purple arrow illustrates the rearrangement of the alkylperoxo intermediate.

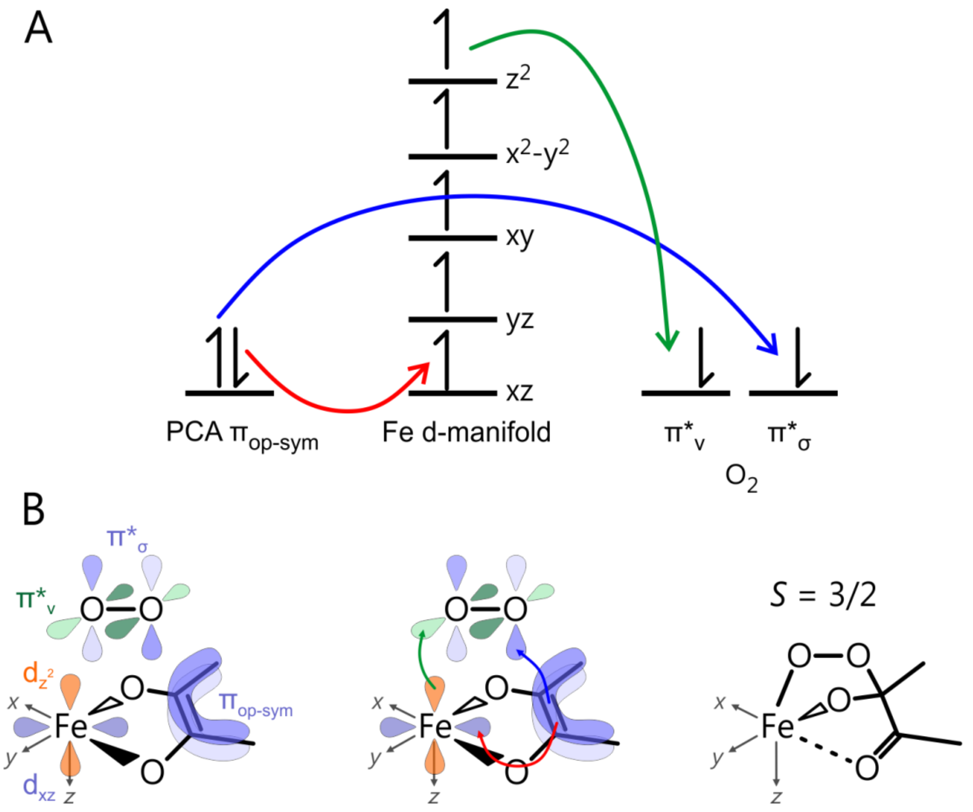

Fig. 9: Proposed electron transfer mechanism for O_2 activation in 3,4-PCD[25]. The Fe^{3+} oxidation state is proposed to be maintained throughout. (A) Diagram showing the reactive orbitals involved in the electron transfer process. (B) Cartoon depicting the location and arrangement of the reactive orbitals.

the iron is the key to continuing the reaction. Protonation of the proximal oxygen in the alkylperoxo is proposed to lead to O-O bond scission. Either Y447 (Fig. 8-*(v-a)*) or solvent (Fig. 8-*(v-b)*) have been proposed as possible proton donors [25, 213, 217]. In either case, the proton donor is proposed to bind in the vacant sixth site on the metal. This return to a six-coordinate center favors intersystem spin crossing back to the $S = 5/2$ system, the spin state of the resting enzyme[189].

In 3,4-PCD, scission of the O-O bond has been theorized to result in the formation of an anhydride intermediate (Fig. 8-*(vi)*). Heterolytic O-O bond cleavage is facilitated by protonation of the proximal O atom of the peroxo moiety[213]. This heterolytic scission results in an Fe^{3+}-bound OH while the distal O atom migrates into the ring *via* a Criegee-type mechanism (Fig. 8-*(vi)*) [181, 190, 194, 206]. This process incorporates the first O atom of O_2 into the substrate ring. After anhydride formation, the iron-bound hydroxide attacks the ring to hydrolyze the anhydride, leading to cleavage and formation of the muconate product (Fig. 8-*(vii)*). Finally, either re-association of Y447 or binding of solvent (to regenerate the resting coordination sphere) could protonate the product leading to muconate dissociation and release[179]. Alternatively, a homolytic O-O cleavage mechanism has also been proposed based on evidence from reactions of a catechol 1,2-dioxygenase with mechanistic probes[195, 218]. This path requires the formation of a ferryl intermediate and acyl radical which recombine and rearrange to generate the anhydride.

Roles of specific active site residues during the catalytic cycle. Mutations of the axial Y447 ligand to His greatly diminishes the rate of formation of the dianionic PCA chelate and reduces the turnover number of the enzyme 600-fold relative to WT[179]. The slowing of chelate formation is likely due to the mismatch in side-chain pK_as between His (pK_a ~ 6) and Tyr (pK_a ~ 10.5). These results implicate the residue in assisting substrate binding by facilitating the deprotonation of the second substrate hydroxyl, thereby driving formation of the asymmetric chelate[179, 181]. Y447 was also assigned as having a role in facilitating product release by protonating one of the carboxylate functions on the muconate product[181]. This was based on the observation that product release in the H447 mutant was slowed relative to wild-type, providing further evidence for its role as an *in situ* acid/base.

The function of the second Tyr ligand, Tyr408, has also been probed by mutagenesis. Various mutants of this residue result in O_2 attack becoming rate-limiting because Y447 fails to dissociate from the iron, so that the substrate cannot progress rapidly from the monodentate to the chelate binding mode in the E_{PCA} complex[194]. The overall turnover number in each of the mutants is reduced to less than 0.1% of WT[194]. These experiments suggest a role for Y408 in maintaining charge balance throughout the cycle and ensuring rapid ligand exchange during substrate binding and product release. *trans*-ligand effects from Y408 on the $3O_{PCA}$ are also proposed to account for a weakening of the $O3_{PCA}$-Fe^{3+} bond driving ketonization of $O3_{PCA}$ after O_2 activation[181, 188, 192]. The presence of one hydroxyl anion and two tyrosinate ligands in the resting coordination sphere also tunes the metal to strongly and preferentially bind the dianionic form of the substrate as two anionic ligands dissociate.

Roles for several other residues have been suggested, but conclusive evidence for their role in the cycle has not been reported. The second-sphere residues R457 and Q477 are proposed to facilitate the shift of the alkylperoxo-PCA species into the equatorial plane by providing a stabilizing interaction with $O4_{PCA}$ and the ketonized $O3_{PCA}$ upon dissociation, respectively (Fig. 8-*(iv)*)[181, 206, 213]. R133 and Y324 are positioned near the entrance of the active site (near the carboxylate function of metal-bound PCA), and have been proposed to help guide the PCA substrate into the active site and orient it properly for the reaction with O_2[181].

Intradiol vs. extradiol dioxygenase cleavage specificity. IDOs represent one of the two major classes of catecholic ring-cleaving dioxygenases along with extradiol dioxygenases (EDOs). The most striking difference between these two families is the oxidation state of the iron. IDOs use an Fe^{3+} cofactor to cleave between the vicinal hydroxyl groups of the catecholic substrate (*ortho* cleavage), whereas EDOs use an Fe^{2+} cofactor and cleave adjacent to the hydroxyl groups (*meta* cleavage). Intradiol and extradiol dioxygenases differ in their protein structure, active site architecture and the oxidation state of the cofactor, but the differences in their catalytic mechanisms are more subtle. Indeed, one of the primary goals of studying these enzyme families has been to understand the mechanism by which the site of ring-cleavage is controlled. The determining mechanism is more complex than simply the cofactor oxidation state (Fe^{3+} vs. Fe^{2+}) or protein fold and ligand set, as a mutant enzyme of the EDO 2,3-HPCD can cleave certain substrates in an *ortho* mode[219] and there have also been a few reports of IDOs performing *meta* cleavage on certain substrates[220-222]. Several mechanisms determining the site of ring cleavage have been proposed, including differences in the protonation site and timing of the intermediates formed during ring cleavage[219, 223-225], the alignment of the alkylperoxo O-O bond relative to the substrate's aromatic ring[72, 226], the coordination mode (facial vs. meridional) of the iron cofactor[227, 228], and differences in the O-O bond scission event controlled by the different iron oxidation states[195, 225, 229].

3,4-PCD: unanswered questions and current research directions. 3,4-PCD is a member of one of only two enzyme families that use the ferric state of the cofactor to activate O_2[25]. Unfortunately, owing to the rapid chemistry of the WT enzyme ($k_{cat} \sim 70\ s^{-1}$ at 25 °C), oxygenated intermediates have never been trapped for detailed characterization. Studies with various mutants of the enzyme have thus far not allowed characterization of the steps of the catalytic cycle in which O_2 is activated[179, 194]. This has prevented the various mechanistic hypotheses from being directly tested[25, 195, 213]. In contrast, the ability to stabilize and trap reactive oxygenated intermediates in the Fe^{2+}-utilizing EDO 2,3-HPCD has allowed us to assign specific roles to many of the active site residues and gain detailed insights into the myriad subtleties of the O_2 activation mechanism[64, 71, 72, 230-233]. Similarly detailed information about the Fe^{3+}-utilizing enzymes would provide fundamental lessons into how nature uses an otherwise inert cofactor to catalyze O_2 activation and facilitate comparisons to the Fe^{2+}-utilizing enzymes. Thus, the first major challenge in 3,4-PCD is a detailed kinetic characterization of the O_2 activation steps and a spectroscopic and structural interrogation of the oxygenated intermediates.

Secondly, the roles of several of the second-sphere residues, including R457, Q477, Y324, and R133 are unknown. These residues are conserved in all 3,4-PCDs but diverge in more distant IDO relatives, such as the 1,2-CTDs[181, 206]. This implies that the residues may play a role in controlling substrate specificity[181, 200, 202] and site-directed mutagenesis could provide a means to increase the enzyme's reactivity toward alternative substrates. As 3,4-PCD and its relatives are positioned at a critical metabolic juncture in aromatic catabolism, these studies are a valuable precursor for engineering of the enzyme for various environmental applications[234, 235]. It also has a direct application to co-opting these enzymes for "green" biosynthesis of valuable chemical feedstocks[177, 236]. The factors that control intradiol *vs.* extradiol ring-cleavage specificity are still not well understood. A direct comparison of the structures of the intermediates in 3,4-PCD to those characterized in 2,3-HPCD[72, 231] would be instrumental in understanding the role of the protein in controlling this site-specificity. This is of fundamental importance to understanding the mechanism of enzymatic catalysis, but also has applications in the design of new biomimetic model compounds or enzymes that catalyze desirable ring-cleavage reactions[228, 237, 238]. The recent characterization of an IDO involved in cleaving lignin precursors may also make these studies applicable to the field of renewable energy[174, 176].

CmlA

The diiron enzyme CmlA catalyzes hydroxylation of the β-carbon of the non-canonical amino acid L-*p*-aminophenylalanine (L-PAPA). The enzyme is part of the non-ribosomal peptide synthetase (NRPS) pathway for chloramphenicol biosynthesis in *Streptomyces venezuelae* (Fig. 10, red)[58]. Based on operon sequences, more than 40 homologs of CmlA exist in a variety of other natural product biosynthetic pathways, but these enzymes have yet to be cloned and biochemically characterized[239-242]. A variety of oxygenase enzymes are involved in NRPS-mediated natural product biosynthesis[13], most often cytochrome P450-type monooxygenases[243], flavin-dependent oxygenases[244], and α-KG-dependent mononuclear iron enzymes[245, 246]. CmlA is the first characterized NRPS-associated oxygenase enzyme to utilize a diiron cluster cofactor.

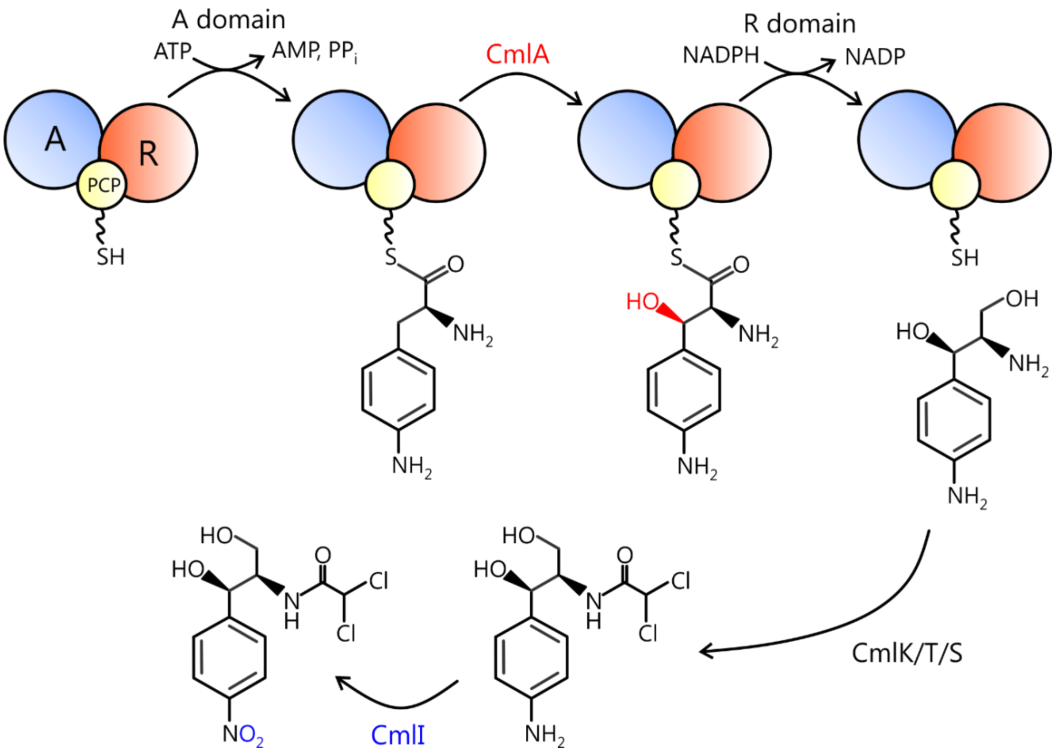

Fig. 10: Non-ribosomal peptide synthetase (NRPS) biosynthetic pathway for chloramphenicol[58, 264]. The NRPS adenylation, peptidyl carrier protein, and reductase domains are shown as blue, yellow and orange spheres, respectively.

Non-ribosomal peptide synthetase pathways and tailoring enzymes. NRPSs are complex, high-fidelity biosynthetic enzyme assemblies that catalyze the formation of many types of secondary metabolites in a variety of bacterial and fungal organisms[247, 248]. The enzymatic machinery of NRPSs relies on a modular biosynthetic logic in which the natural products are generated from amino acid monomers that are sequentially activated using ATP, covalently attached to the NRPS *via* an acyl-thioester linkage and condensed together to form and extend the polymer (Fig. 11A)[247-249]. NRPSs are composed of single or multiple modules that each contain several domains with distinct biosynthetic roles. Multi-domain NRPSs are commonly expressed as single polypeptides from a single gene resulting in immense sizes for some NRPSs[250]. In other pathways, multiple NRPSs expressed from different genes work in tandem to synthesize the final product. In each module, the amino acid monomers are selected and activated using ATP by action of a module's adenylation domain (A domain) (Fig. 11B) [251, 252]. The A domain selects and sequesters the substrates for each of the NRPS domains, ensuring that the correct amino acid monomer is activated for addition to the growing peptide. A second role of this domain is to

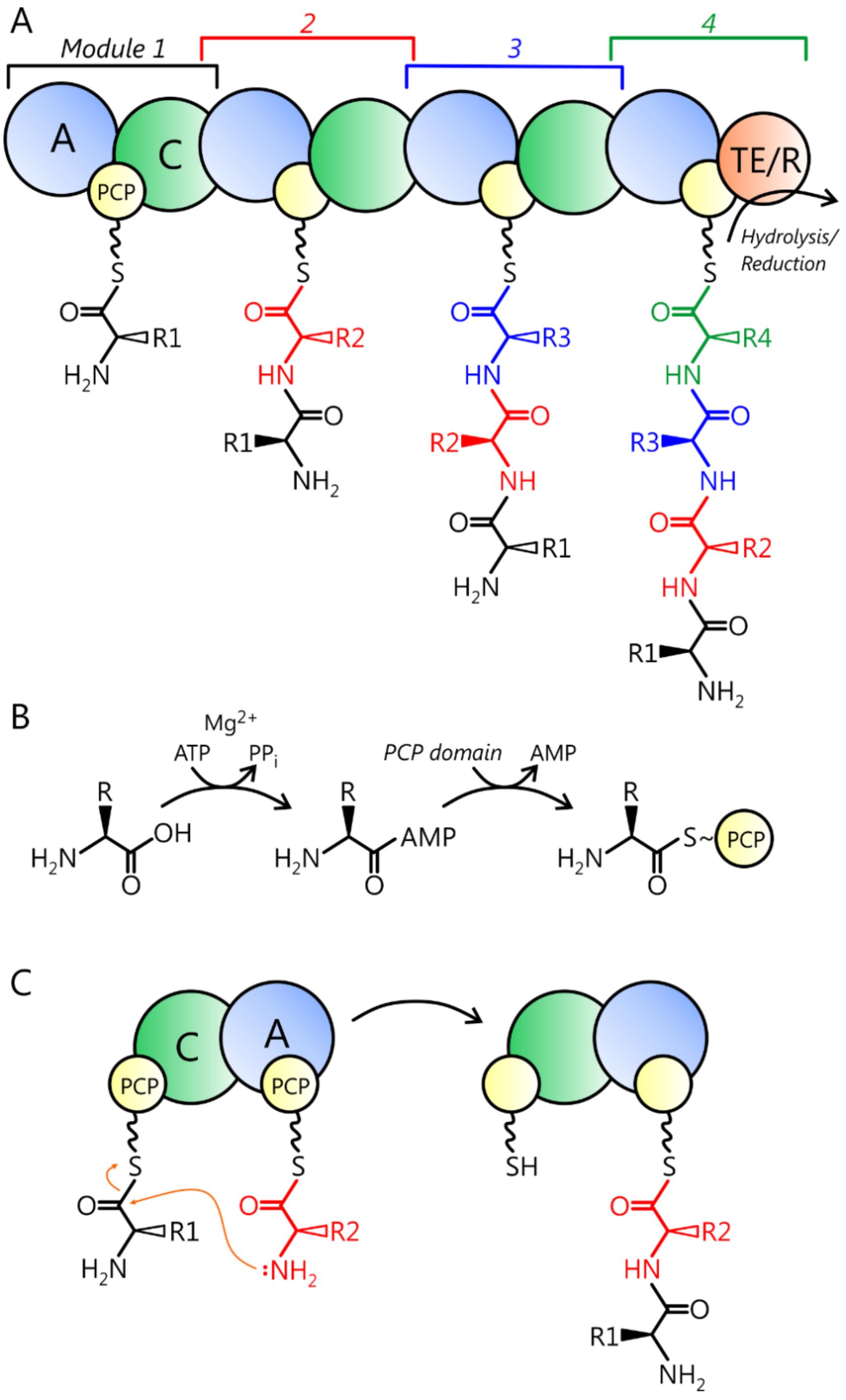

Fig. 11: Biosynthetic logic of NRPS machinery[247]. (A) Example of a four-module NRPS and the synthesis of a tetrapeptide. (B) Steps in the activation of amino acids to the AMP-adenylates in the NRPS A domain. (C) Peptide condensation and extension catalyzed by the NRPS C domain. The various NRPS domains are colored as follows. Adenylation (blue), peptidyl carrier protein (yellow), condensation (green) and thioesterase/reductase (orange).

covalently attach the activated amino acyl-AMP adenylate to a second NRPS domain called the peptidyl carrier protein domain (PCP domain)[253]. The amino acid is attached as an acyl-thioester to a coenzyme A-derived phosphopantetheine group that is itself linked to a serine residue on the PCP domain[163, 254]. Condensation domains (C domains) catalyze condensation reactions between the monomers of sequential modules of the NRPS, forming a peptide bond to extend the nascent peptide (Fig. 11C). Reductase and thioesterase domains (R and TE domains) are found at the C-terminus of NRPS systems and act to reductively or hydrolytically, respectively, release the peptide from the NRPS[255]. There are also *trans*-acting TE enzymes that work to release stalled NRPS assembly lines[256].

The amino acid monomers that are utilized to synthesize the peptide are often themselves chemically modified either prior to or after covalent attachment to a NRPS PCP domain. These modifications are generated by so-called "tailoring enzymes," which may either be part of the NRPS polypeptide, i.e., the same gene (*cis* enzymes), or free enzymes that are expressed in a separate gene (*trans* enzymes)[257]. Tailoring enzymes catalyze an enormous variety of modifications on the amino acid monomers including N- , C-, and O-methylation[258], racemization and epimerization[259, 260], hydroxylation [13, 261], cyclization[255], and glycosylation[240, 262]. Such modifications are critical to the bioactivity of the final polypeptide produced by the NRPS. β-carbon hydroxylated amino acids are common building-blocks of NRPS-synthesized polypeptides[261]. These β-hydroxy moieties are sites for future modification including glycosylation, cyclization, and further oxidation or retro-aldol cleavage.

Using the nomenclature just described, CmlA is a *trans*-acting hydroxylase-tailoring enzyme that β-hydroxylates the L-PAPA amino acid substrate while it is tethered to the PCP domain of the mono-modular NRPS CmlP[58]. CmlP consists of only a single three-domain module with A, PCP, and TE/R domains. The A domain is selective for L-PAPA which it activates and covalently tethers to the PCP domain[263]. The L-PAPA substrate itself is synthesized by the organism *via* a shunt from the shikimate pathway[264].

CmlA protein structure and active-site architecture. CmlA is the first discovered and characterized O_2-activating enzyme to utilize a metallo-β-lactamase-like (MBL) fold, in contrast to the 4-helix bundle present in the BMMs, acyl-ACP desaturases, and RNR-R2. Related enzymes that reduce O_2 in obligate anaerobes or catalyze nitric oxide reduction use the same fold, but they do not generate reactive oxygen species and have a different cluster ligand set[265-267]. The active site has been characterized by EPR, Mössbauer, EXAFS and resonance Raman (rRaman) studies[58, 268]. The diiron cluster of CmlA exhibits some distinct properties, including a high exchange coupling constant (J > 90 cm^{-1}) and an EPR silent $S = 0$ diferrous state[58]. In contrast, diferric MMOH has a lower exchange coupling ($J \sim 14$ cm^{-1}) and a diferrous cluster with $S = 4$ that arises from ferromagnetic coupling of two $S = 2$ ions[109]. The CmlA cluster is proposed to be coordinated by three His ligands and carboxylates (one glutamate and two aspartate) and bridged by an oxo ion. The absence of protons on the bridging moiety was determined using rRaman studies which showed that the oxygen-iron-linked vibrations are unaltered when the enzyme in incubated in H_2O or ^2H_2O[268]. EXAFS measurements showed that both metals are six-coordinate in the resting form and the Fe-Fe distance is 3.3-3.4 Å. Currently, there is no structural information about the changes that occur upon reduction of the cluster or upon NRPS binding, but the exchange coupling of the diferrous cluster is reduced to 12 cm^{-1} suggesting a change in the bridging moiety[58]. This constant is much larger and has the opposite sign from that reported for diferrous MMOH ($J \sim -0.7$ cm^{-1})[49, 109].

CmlA catalytic mechanism and regulation. As this family of enzymes was only recently discovered, very little is known about the mechanism of O_2 activation or its catalytic regulation. It is known that diferrous CmlA is only reactive with CmlP-tethered L-PAPA, implying a regulatory system that must be linked to the protein-protein interaction with the NRPS[58]. Thus, although CmlA uses a completely different fold than other hydroxylases, the cluster reactivity may also be gated by a structural rearrangement induced by interaction with another protein, like that of the MMOH-MMOB system[150, 157, 158, 269-271] or the acyl-ACP desaturases[168]. Similar to the ACP in the latter group, CmlP is likely serving dual roles: *(i)* as a substrate carrier that targets the NRPS to the correct tailoring enzyme thereby ensuring overall product specificity of the synthetase and *(ii)* as a regulatory partner that controls the timing of O_2 activation. No intermediates of the CmlA reaction have yet been trapped for characterization, and it remains to be determined if CmlA has a catalytic cycle analogous to that of the BMMs. Due to the strength of the Cβ-H bond (BDE ~ 85 kcal mol^{-1}) a high-valent Q-like intermediate has been evoked as the oxidizing species[58].

CmlA: unanswered questions and current research directions. There are currently considerably more mysteries than answers regarding CmlA. Some of the major areas of interest have already been alluded to in the above sections. Of primary interest is the determination of the structure of the enzyme. As noted, CmlA has a unique fold among all other characterized oxygenase enzymes. Determination of the structure of the overall fold would permit insightful comparisons to other oxygenases that use different protein architectures to house and coordinate the cofactor. This could provide an explanation for the unique spectroscopic properties observed for the enzyme[58]. The elucidation of the as-of-yet uncharacterized N-terminal domain would also allow us to obtain the first structural clues as to the biological role of this novel fold in CmlA and in the homologs from other biosynthetic pathways. A structure is also a requisite for rational site-directed mutagenesis of the enzyme's cluster ligands.

A second priority is the determination of the nature of the oxygenated intermediate that catalyzes the oxidation of the Cβ-H bond. Since CmlA belongs to an entirely new family of diiron hydroxylases, it is of interest to know whether it generates the same type of oxidizing intermediate as MMOH or if nature has devised an entirely different type that is exclusive to this protein fold and cluster architecture. Thus far, MMOH remains the only enzyme where a Fe^{4+}/Fe^{4+} species has been directly characterized[272]. Having the structure of CmlA available could also be used to answer some basic questions about whether the cluster can support a Q-like intermediate and if it is likely to go through an analogous catalytic cycle to that of sMMOH. Finding evidence of a high-valent Fe^{4+}/Fe^{4+} species in CmlA would show that other protein folds and cluster architectures can support the generation of such an intermediate. This is directly relevant to the design of new catalysts with industrial applications and of basic interest to understanding the nature of activated O_2 species.

Finally, the mechanism of catalytic regulation in CmlA is of interest. As discussed in detail above, nature has devised a wide variety of elegant safeguards to control O_2 activation. The regulatory mechanisms are as diverse as the reactions themselves. The unique protein fold of CmlA (amongst other oxygenases) and its role in an NRPS pathway suggests that it has a new type of regulatory mechanism. A detailed study would provide a new perspective on the methods nature uses to control O_2 activation. The determination of the structure of CmlA would again go a long way towards answering some of these questions. A related topic is studying the mechanism by which CmlA is targeted to its cognate NRPS, which may be tied into the mysterious N-terminal domain of the enzyme. Altering this targeting specificity and/or engineering the enzyme to react with alternative amino acids could one day be used to engineer and harness this system for the biosynthesis of new bioactive compounds such as antibiotics[273-275].'

CmlI

CmlI is a diiron-cluster containing amine-oxidizing enzyme (N-oxygenase) that converts the aminoaryl group of the chloramphenicol precursor (D-threo-1-(4-aminophenyl)-2-dichloroacetylamino-1,3-propanediol, NH_2-CAM) to the corresponding nitroarene, thereby generating chloramphenicol (CAM)[276, 277]. CmlI is in the same genetic operon as the CmlA β-hydroxylase and is the last enzyme in the biosynthetic pathway for chloramphenicol in *S. venezuelae* (Fig. 10, blue)[57, 264]. Unlike CmlA, CmlI is not reactive toward NRPS-tethered amino acids. The primary structure of CmlI suggests that it coordinates the cluster in a 4-helix bundle[278, 279].

Arylamine dioxygenases. The mechanism of N-oxygenation is of considerable interest due to the presence of the nitro function in several antibiotics and other medicinally important compounds[172, 280]. Although nitro groups have been recognized in several bacterial and fungal secondary metabolites, only a few biosynthetic N-oxygenases besides CmlI have been characterized in detail. One is AurF, a diiron arylamine oxygenase in the biosynthetic pathway for the antibiotic aureothin in *Streptomyces thioluteus*[281, 282]. Its structure and function are homologous to those of CmlI. The AurF substrate is *p*-aminobenzoate which the enzyme converts to *p*-nitrobenzoate. The second N-oxygenase that has been studied in detail is PrnD, a Rieske-type mononuclear iron enzyme involved in the biosynthesis of the antifungal antibiotic pyrrolnitrin[283].

CmlI catalytic mechanism. CmlI is a recent isolate and the details of the catalytic mechanism of the enzyme are relatively unknown. The mechanism is likely analogous to that catalyzed by AurF, a reaction that has been studied in more detail. However, many questions regarding the catalytic mechanism of CmlI and AurF are still under dispute, with various groups proposing different mechanisms for how the arylamine group is converted to the nitro[56, 282, 284]. The overall conversion of the arylamine group to the nitro is a six e^- oxidation and the general catalytic mechanism invokes three sequential two e^- oxidative steps for this conversion. The O_2 stoichiometry for this process is still under dispute. Some data suggests that three equivalents of O_2 (along with six exogenous e^-) are required per molecule of amino substrate for the full oxidation[57, 73]. Other data imply that only two O_2 equivalents are necessary and that the third oxidation step proceeds by oxidation of an intermediate by the diferric cluster[284]. One proposed mechanism for the first of the three oxidative cycles is shown in Figure 3B.

Like all other diiron enzymes with the exception of MIOX, the diferrous state is reactive towards O_2. Both CmlI and AurF, under certain conditions, generate long-lived Fe^{3+}/Fe^{3+}-peroxo species upon exposure of the diferrous cluster to O_2[57, 73]. The peroxo species in CmlI has been characterized in more detail than the AurF peroxo. This study showed that the species has a novel geometry when compared to peroxo intermediates previously studied in other diiron systems[57, 73]. The kinetic data have supported a model in which the peroxo reacts directly with the substrate and the oxidation event occurs by nucleophilic attack of the amine (or hydroxylamine) group on the cluster-bound peroxo moiety[57, 73, 284]. How the unique geometry of the peroxo species relates to its ability to oxidize arylamine groups is still under investigation. In the arylamine oxygenases, the role of the diiron cluster is to activate O_2 to a peroxo state, preparing the molecule for attack *by* the amino substrates. This mechanistic proposal also explains why the enzyme is unable to activate C-H bonds, which are unlikely to perform a nucleophilic attack on the peroxo moiety.

CmlI: unanswered questions and current research directions. As with CmlA, the lack of a structure of the enzyme and knowledge of its active site architecture is one of the largest gaps in our current knowledge. A structure would facilitate a comparison to other diiron oxygenases and to the CmlI homologs in other natural product biosynthetic pathways. It would also provide the first insight into how the cluster may direct the formation of the biologically unique peroxo species. The recent spectroscopic characterization of the CmlI peroxo species also presents an opportunity for comparisons to other diiron peroxo modes and the correlation between geometry and reactivity[60, 74, 75, 81, 119]. Due to the long-lived nature of the intermediate, X-ray crystallographic characterization may be possible once crystallization conditions for CmlI are obtained[72, 76, 82]. The availability of a crystal structure would also provide an advantage in resolving some of the current disputes about the catalytic mechanism of the enzyme. Co-crystallization with substrates is also key to determining the mechanism of substrate specificity of the enzyme. This, in turn, could open up the potential for site-directed mutagenesis of the enzyme to perform N-oxygenation reactions on alternative molecules.

Overview of the Chapters

The first two chapters of this dissertation focus on the recent advances made in our understanding of the catalytic mechanism of 3,4-PCD. Chapter 1 presents the X-ray crystallographic trapping of the two most critical oxygenated intermediates in the 3,4-PCD catalytic cycle using an alternative substrate. Chapter 2 presents the kinetic characterization of the reaction of 3,4-PCD with various other alternative substrates and the discusses their relevance to the native reaction cycle. Chapters 3 and 4 focus on the determination of the crystal structures of CmlA and CmlI, respectively. Their structures are compared to those of other diiron enzymes and the myriad insights into their catalytic mechanism and regulation are discussed.

CHAPTER 1

Crystal structures of alkylperoxo and anhydride intermediates in an intradiol ring-cleaving dioxygenase

Note: This chapter is reprinted (adapted) with permission from Knoot, C. J., Purpero, V. M., and Lipscomb, J. D. (2015) Crystal structures of alkylperoxo and anhydride intermediates in an intradiol ring-cleaving dioxygenase, Proc Natl Acad Sci U S A 112, 388-393. Copyright (2015) PNAS.

Abbreviations:

3,4-PCD: protocatechuate 3,4-dioxygenase
4FC: 4-fluorocatechol
PCA: protocatechuate (3,4-dihydroxybenzoate)
IDO: intradiol dioxygenase
E-4FC: the complex of 3,4-PCD with 4FC without O_2
E-PCA: the complex of 3,4-PCD with PCA without O_2
2,3-HPCD: homoprotocatechuate 2,3-dioxygenase
RRT: reciprocal relaxation time
LMCT: ligand-to-metal charge-transfer

Summary

Intradiol aromatic ring-cleaving dioxygenases utilize an active site, nonheme Fe^{3+} cofactor to activate O_2 and catecholic substrates for reaction. The inability of Fe^{3+} to directly bind O_2 presents a mechanistic conundrum. The reaction mechanism of protocatechuate 3,4-dioxygenase is investigated here using the alternative substrate 4-fluorocatechol (4FC). 4FC is found to slow the reaction at several steps throughout the mechanistic cycle, allowing the intermediates to be detected in solution studies. When the reaction was initiated in an enzyme crystal, it was found to halt at one of two intermediates depending on the pH of the surrounding solution. The X-ray crystal structure of the intermediate at pH 6.5 revealed the key Fe^{3+}-alkylperoxo species, while the Fe^{3+}-anhydride intermediate was found for a crystal reacted at pH 8.5. Intermediates of these types have not been structurally characterized for intradiol dioxygenases, and they validate four decades of spectroscopic, kinetic, and computational studies. In contrast to our similar *in crystallo* crystallographic studies of a Fe^{2+}-containing extradiol dioxygenase, no evidence for a superoxo or peroxo intermediate preceding the alkylperoxo was found. This observation and the lack of spectroscopic evidence for an Fe^{2+} intermediate that could bind O_2 are consistent with concerted formation of the alkylperoxo followed by Criegee rearrangement to yield the anhydride and ultimately ring-opened product. Structural comparison of the alkylperoxo intermediates from the intra- and extradiol dioxygenases provides a rationale for site specificity of ring cleavage.

Introduction

Most ring-cleaving dioxygenases belong to the non-heme mononuclear iron-containing enzyme family, many members of which play critical roles in the aerobic biodegradation of natural and synthetic aromatic compounds[32, 40, 175, 177, 285]. Ring-cleavage by these enzymes involves fission of the O-O bond of dioxygen and incorporation of both atoms into the product[17]. Two major classes of catechol ring-cleaving dioxygenases have been described. These differ in the mode of cleavage and the oxidation state of the iron cofactor: the Fe^{2+}-utilizing extradiol dioxygenases (EDOs) cleave adjacent to the vicinal hydroxyl functions of catechols (*meta*-cleavage) while intradiol dioxygenases (IDOs) use a Fe^{3+} cofactor and cleave between the hydroxyl functions (*ortho*-cleavage). The Fe^{3+} in the active site of IDOs will not bind O_2 and no direct spectroscopic evidence for redox cycling has been obtained[32]. This means that IDOs must activate O_2 using a unique mechanism which differs from that of the much more common Fe^{2+}-containing oxygenases that readily form an Fe-O_2 complex after the substrate binds in the active site[64, 125, 286].

The archetypal IDO is protocatechuate 3,4-dioxygenase (3,4-PCD), which catalyzes cleavage of its native substrate (3,4-dihydroxybenzoate, PCA) to yield β-carboxy-*cis,cis*-muconate, thereby funneling aromatic substrates into the β-ketoadipate pathway[170]. Much of our mechanistic knowledge of IDOs comes from kinetic, spectroscopic and crystallographic studies of 3,4-PCD, catechol 1,2-dioxygenase[25, 179, 184, 185, 194, 196], and bio-inspired model compounds[237, 238]. 3,4-PCD coordinates the Fe^{3+} in a 2-His, 2-Tyr ligand set which is conserved in IDOs[196]. The tyrosine ligands contribute several ligand-to-metal-charge transfer bands (LMCTs) in the visible which give IDOs a burgundy color that can be used to monitor the reaction[186]. In 3,4-PCD, the Fe^{3+} is coordinated by axial ligands Y447 and H462 and equatorial ligands H460, Y408 and solvent-derived hydroxide (Fig. 1)[196].

The proposed intermediates for the reaction of 3,4-PCD with PCA and O_2 are shown in Fig. 12 and designated **1-8**[25, 32, 179, 181, 185, 195, 213]. First, PCA binds to the metal of the resting enzyme **1** in a multi-step process leading to a complex **2** where both hydroxyls are ionized and the ring of PCA is axially oriented. Formation of **2** results in dissociation of Y447 from the metal, forming a 5-coordinate (5C) complex with a vacant site *trans* to H460. O_2 reacts with **2** in a concerted process **3**, generating a Fe^{3+}-alkylperoxo species **4**. Spectroscopic and computational studies suggest that catalysis requires a physical shift of a transient six-coordinate (6C) alkyperoxo **4** to a 5C species **5**, opening the axial site *trans* to H462. Protonation of the peroxo moiety **6** drives O-O bond scission and migration of one oxygen atom into the substrate ring to form an anhydride **7**. Attack of the iron-bound hydroxide on the anhydride cleaves the substrate ring to form product complex **8**, which, upon release, regenerates state **1**. Transient kinetic studies have detected intermediates throughout the catalytic cycle, but the kinetics of their interconversion have not allowed any intermediate between the substrate and product complexes to be trapped for detailed characterization[179, 183, 184, 287].

The use of alternative substrates has been a successful means to detect and trap intermediates in enzyme reaction cycles in which the native reaction proceeds too quickly or with unfavorable interconversion rate constants[126]. A powerful extension of this approach involves performing catalysis in enzyme crystals to trap reactive species for characterization using X-ray crystallography, which has been successfully applied to a few systems[72, 82, 231, 288, 289]. This approach was used by our laboratory to identify and characterize four distinct oxygenated intermediates in the catalytic cycle of the EDO homoprotocatechuate 2,3-dioxygenase (2,3-HPCD)[72, 231]. Here, we have extended this *in crystallo* approach to the IDO 3,4-PCD as it turns over the alternative substrate 4-fluorocatechol (4FC), allowing determination of the structures of two intermediates following O_2 addition. These intermediates provide a direct view of the key steps in the oxygen activation and reaction cycle of IDOs. Insight is also gained into one of the longstanding questions in oxygenase research: the basis for cleavage site specificity in the ring cleaving dioxygenase family.

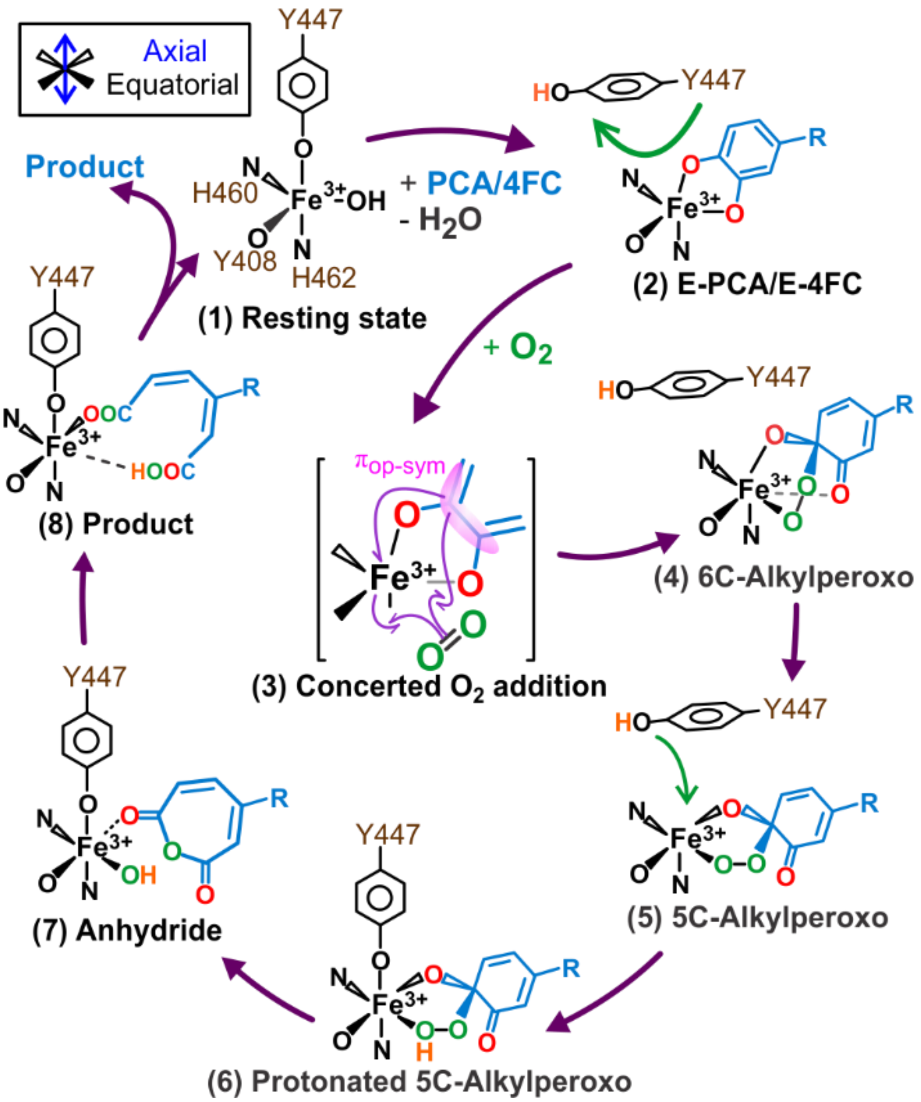

Fig. 12: Proposed intermediates in the reaction of 3,4-PCD with substrate and O_2. R represents either COOH (PCA) or F (4FC). 4FC is bound in the equatorial plane rather than the axial plane as shown in **2**. It may rotate to the axial plane to react. The π_{op}-sym HOMO of substrate is proposed to be the source of the reducing equivalents[25].

Materials and methods

Protein expression and purification. 3,4-PCD was purified as previously described[194]. In addition, we generated an N-terminally 8x-His-tagged construct of the enzyme for rapid purification. The coding sequence for the *pcaH* and *pcaG* genes from *Pseudomonas putida* was amplified using polymerase chain-reaction (PCR) from a pCE120 expression template[194], with the addition of 5’ PmeI and 3’ SgfI restriction sites encoded on the primer. The amplified PCR product was moved into a pCR-4 TOPO TA cloning vector (Invitrogen) and transformed into competent DH5α cells and plated on LB media supplemented with 100 µg/ml carbenicillin. Individual clones were transferred to fresh liquid media. The TOPO-Pme1-*pcaHG*-Sgf1 plasmid was purified and digested with the relevant restriction enzymes. This fragment was ligated into a similarly digested pVP91A vector (DNASU plasmid repository, Arizona State University), transformed, and clones selected on carbenicillin-supplemented media. Sequencing of the expression construct at UMN Biomedical Genomics Center confirmed that the coding sequence was error-free.

The His-tagged 3,4-PCD was expressed in BL21-DE3 *E. coli*. Standard LB growth media was supplemented with 10 mg/L ferrous ammonium sulfate during IPTG induction. The tagged enzyme was purified on a Ni-nitrilotriacetic acid (NTA) HisPur resin (Thermo Scientific) using the recommended procedures. 3,4-PCD was dialyzed into 20 mM HEPES pH 7.5, supplemented with 5% glycerol as cryoprotectant for the storage at -80 °C. The specific activity and electron paramagnetic resonance spectrum of the His-tagged resting enzyme were identical to those previously reported for untagged enzyme, suggesting there was no perturbation of the Fe^{3+} active site.

Metal analysis. Total enzyme was quantitated using the calculated extinction coefficient at 280 nm. The enzyme was digested in 4 M nitric acid while stirring at 90 °C for 3 hours. The digested sample was then centrifuged for 10 minutes, the supernatant quantitatively transferred to a 5 ml volumetric flask and diluted with nitric acid. The metal was quantitated using inductively-coupled plasma mass-spectrometry (ICP-MS) at the UMN Earth Sciences Department. Samples containing only buffer were run as controls. Purified 3,4-PCD was bound with roughly 0.75 mol iron per mol enzyme based on ICP-MS.

Steady-state kinetic assays. 4-fluorocatechol (4FC) was purchased from TCI Chemicals and used without further purification. The extinction coefficient of 4FC at 282 nm was determined by nuclear magnetic resonance (NMR) spectroscopy using an aromatic proton inventory calibrated to an internal tryptophan standard[290] and found to be 3400 ± 80 $M^{-1}cm^{-1}$ at pH 8.5. At pH 7.5, the extinction coefficient at 282 nm was calculated to be 3370 ± 50 $M^{-1}cm^{-1}$. For PCA, the extinction coefficient at 290 nm is 3870 $M^{-1}cm^{-1}$ at pH 8.5[179] and determined here at pH 7.5 to be 3090 ± 40 $M^{-1}cm^{-1}$. These values were used to quantify the concentration of 4FC or PCA used in the assays.

Extinction coefficients for the 4FC reaction product were determined by incubating a known concentration of 4FC with enzyme and 1.5-fold excess of dissolved O_2. Once all spectral changes had completed, the UV absorbance at λ_{max} along with the input substrate concentration was used to calculate the extinction coefficient ε. Kinetic parameters k_{cat} and K_m were determined using an oxygen electrode (Hansatech) to monitor the dissolved O_2 consumption rate. The k_{cat} values were determined in the presence of saturating 4FC (2 mM) and a 3-fold excess of dissolved O_2 over the K_m (750 µM). The high O_2 concentrations were obtained by bubbling O_2 gas through the assay buffers in a sealed vial. The buffers used for the steady-state assays contained 0.3 M Tris-HCl (pH 8.5), 0.3 M Hepes (pH 7.5), and 0.3 M MES (pH 6.5). All buffer solutions were brought to the same conductance using NaCl.

Transient kinetics. Transient kinetic data were collected using an Applied Photophysics SX.18MV stopped-flow instrument in the single wavelength or spectra kinetic mode. For anaerobic binding experiments, Ar-sparged 100 µM 3,4-PCD (Fe^{3+}-loaded active sites) was rapidly mixed with anaerobic buffer (0.3 M HEPES pH 7.5) containing various concentrations of Ar-sparged 4FC substrate. For reactions with O_2, the enzyme-4FC complex was pre-formed in an anaerobic bag, transferred to the stopped-flow and mixed with oxygenated HEPES buffer at 4 °C, pH 7.5. Due to the reversibility of the 4FC binding, a stoichiometric complex of enzyme and 4FC could not be formed with equal concentrations of 3,4-PCD and 4FC. The reaction was studied under two alternative conditions: (1) single turnover was done with excess enzyme and limiting 4FC (900 µM 3,4-PCD, 250 µM 4FC, ~200 µM E-4FC) reacted with 1.9 mM O_2 or (2) multiple turnover conditions (150 µM 3,4-PCD, 4 mM 4FC, ~150 µM E-4FC) to study the O_2 dependence of the RRTs. Under the latter condition, only the first 6 s of the reaction were mathematically fit to avoid complications by multiple-turnover. RRTs were determined while monitoring in single-wavelength mode at 540 nm for the aerobic reactions and separately at 325 nm, 425 nm and 660 nm for the anaerobic binding.

Mathematical fitting of the transient kinetic reactions and extraction of the optical spectra of the intermediates. Nonlinear regression fitting of the stopped-flow 4FC binding and O_2 reaction traces was performed using the Applied Photophysics Pro-Data Software Suite (version 4.2.12). The equations used for the fits to the observed absorbance ($A_{t,obs}$) were sums of exponential functions having the form:

$$A_{t,obs} = \sum_{i=1}^{n} A_i e^{(-\frac{t}{\tau i})} + A_\infty$$

where A_i is the observed amplitude of phase i (of n) in absorbance units, $1/\tau_i$ is the reciprocal relaxation time of the phase (RRT, s^{-1}), t is time (s) and A_∞ is the absorbance at the end of the reaction. Each point and the corresponding error in the plots shown in Fig. 14 was based on three separate experiments using different batches of purified enzyme.

For the 4FC binding experiments, the hyperbolic dependence of the observed RRT on 4FC concentration suggested the following kinetic model for binding:

$$\text{E+4FC} \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} \text{E·4FC} \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} \text{E-4FC}$$

The plot was fit using a hyperbolic function with the form:

$$\frac{1}{\tau_{obs}} = \frac{k_2[\text{4FC}]}{(\frac{k_{-1}}{k_1}) + [\text{4FC}]} + k_{-2}$$

Extraction of the rate constants from the sum and product plots of the RRTs for the aerobic reaction was performed using previously described methods[230, 291, 292]. Three relaxations are observed in the transient reaction implying that there are at least 3 steps in the reaction. The first two phases have a clear dependence on O_2 concentration (Fig. 13G and H). The nonlinear dependence on O_2 concentration of the second relaxation is consistent with a sequential rather than parallel sequence of steps. Moreover, a nonlinear O_2 dependence on one phase would only be observed if the first step in the reaction in which O_2 binds is reversible. Based on the kinetic fit alone, it is possible that all three steps in the reaction are reversible. However, because the optical spectrum of the reaction following complete turnover of substrate is nearly indistinguishable from that of resting enzyme, the final step is unlikely to be reversible. This leads to the following initial model for the reaction with O_2:

$$\text{E-4FC+O}_2 \underset{k_{-3}}{\overset{k_3}{\rightleftarrows}} \text{Int1} \underset{k_{-4}}{\overset{k_4}{\rightleftarrows}} \text{Int2} \xrightarrow{k_5} \text{E+P}$$

The rate constants for the individual steps cannot be determined directly from the exponential fits due to mixing of the relaxations through the reversible step(s). Under the assumption that the final irreversible step is slow, the analysis must be performed by solving a series of differential rate equations describing the first two steps of the reaction. This analysis shows that the first and second RRTs correspond to the negative roots of a quadratic equation

$$(\mathbf{1})\quad \lambda^2 + (k_3[\mathrm{O_2}] + k_{-3} + k_4 + k_{-4})\lambda + \ (k_3[\mathrm{O_2}](k_4 + k_{-4}) + k_{-3}k_{-4})$$

$$(\mathbf{2})\quad \frac{1}{\tau 1}, \frac{1}{\tau 2} = \frac{1}{2}\{(k_3[O_2] + k_{-3} + k_4 + k_{-4}) \pm \sqrt{}[(k_3[\mathrm{O_2}] + k_{-3} + k_4 + k_{-4})^2 - 4(k_3[\mathrm{O_2}](k_4 + k_{-4}) + k_{-3}k_{-4})]\}$$

The rate constants may thus be extracted using the following relationships:

$$(\mathbf{3})\quad \frac{1}{\tau 1} + \frac{1}{\tau 2} = k_3[\mathrm{O_2}] + k_{-3} + k_4 + k_{-4}$$

$$(\mathbf{4})\quad \frac{1}{\tau 1 \tau 2} = k_3(k_4 + k_{-4})[\mathrm{O_2}] + k_{-3}k_{-4}$$

The plot of the products of the RRTs vs. O_2 concentration was revealed to have an intercept near zero (Fig. 14C), implying that the second step is also effectively irreversible (i.e. k_{-4} ~ 0 in the above equations). This being the case, the assumption that the final step is slow is not required and k_5 is given directly by the slowest RRT.

Global kinetic analysis and extraction of the optical spectra of the intermediates shown in Fig. 14A was performed using the Applied Photophysics Pro-K Software Package (version 1). The data for the analysis was generated by monitoring the reaction at 13 discrete wavelengths between 340 nm and 700 nm. These data were analyzed using singular value decomposition in Pro-K using the kinetic model shown on page 51. The full experiment was repeated twice with similar results.

Crystallization conditions and model nomenclature. Crystallography-grade reagents were purchased from Hampton Research. Untagged 3,4-PCD was crystallized using hanging-drop vapor diffusion in mother liquor containing 100 mM buffer (Tris-HCl for pH 8.5, HEPES for pH 7.5, and MES for pH 6.5 soaks), 2.5 mM 2-mercaptoethanol and 1 to 1.4 M ammonium sulfate. The drops contained a 2:1 μl ratio of well solution to 40 mg/ml enzyme solution. Crystals typically formed over the course of two to three days. Nomenclature for the protein models is as follows: Chains A and B corresponds to the alpha and beta subunits of the protomer A, chains C and D to the alpha and beta subunits of the protomer B and chains E and F to the alpha and beta subunits of the protomer C. His-tagged 3,4-PCD could not be crystallized.

Crystal soaking. To perform the *in crystallo* reactions, single crystals (0.3-0.5 mm measured along the longest axis) were removed from the crystallization drop and transferred to drops containing 100 mM 4FC in freshly prepared mother liquor at pH 6.5, 7.5 or 8.5. Crystals were soaked for 30 min with 4FC in mother liquor containing 25% (v/v) glycerol as cryoprotectant and directly frozen in liquid nitrogen. For anaerobic soaks, crystals were grown anaerobically prior to soaking for 10 min with Ar-sparged 50 mM 4FC in mother liquor supplemented with cryoprotectant.

Diffraction data collection. Data were collected at the Structural Biology Center at the Advanced Photon Source, Argonne National Labs on beamline 19-ID at 100 K. To test for possible photoreduction, an in-house data set was collected using a RigakuMSC Micromax 007 X-ray generator with a rotating Cu-anode and a Saturn 944+ CCD detector. Diffraction data were indexed, integrated and scaled using the HKL2000 software package and phased using molecular replacement with a model of 3,4-PCD in space group C2. The final model was prepared using iterative cycles of modeling using Coot and refinement with Refmac5[293, 294]. All ligands were generated using JLigand 1.0.40. The final models were evaluated using Sfcheck and Rampage. All structure figures were produced using PyMOL Molecular Graphics System, Version 1.5.0.4, Schrödinger, LLC.

Ligand refinement and modelling. In order to mitigate ligand bias of the calculated phases, the active site of 3,4-PCD was always the last site modeled, following refinement of the entire protein polypeptide, solvent molecules, mother liquor components, and 4FC molecules which were not bound in the active site. The occupancy of the intermediates was refined until the temperature factors were close to those of the surrounding protein residues and the iron and the residual |Fo|-|Fc| density following refinement with the ligand was minimized. The final percent occupancies were refined at 75% for the iron cofactor (based on temperature factors and ICP-MS analysis), 70% for 4FC (35% for each superimposed orientation), 65% for the alkylperoxo intermediates, and 65% for the anhydride and iron-bound hydroxide. Following refinement with the alkylperoxo intermediate at 65% occupancy, there was evidence for residual superposition of unreacted substrate in the |Fo|-|Fc| map. A superimposed model of equatorially-bound 4FC at 10% occupancy satisfied the density, but 4FC alone failed to account for the all of the density. The observation of some residual unreacted substrate is unsurprising based on the slow O_2 reaction rate of the E-4FC complex. No evidence of residual E-4FC was observed in the anhydride maps. To generate the ligand-omit maps shown in the figures, the ligand was deleted from the final model and this model refined for an additional five cycles in Refmac5. The resulting structure factors were used to calculate the |Fo|-|Fc| difference map (“ligand-omit”), which agreed well with the modeled orientation of the ligands. “Side-chain omit” maps were generated using the same approach.

Refinement of the Y447 side-chain occupancies. In both the E-4FC and alkylperoxo structures, it was clear that some fraction of the active sites still had Y447 in the metal-bound orientation. In the E-4FC complex, the |Fo|-|Fc| density for the dissociated orientation of Y447 was found to be much stronger than for the metal-bound orientation. We have thus modeled this orientation at 75% occupancy (equivalent to the metallated 4FC-bound fraction of active sites) and the substrate-free orientation at 25% occupancy (the metal-free fraction). This assignment is supported by the observation that, in metal-free crystals of 3,4-PCD (based on ICP-MS and UV/vis spectrometry), Y447 was not observed in the dissociated orientation. This suggests that in metal-free active sites, the Y447 side-chain still assumes the “metal-bound” orientation. The assignment of side-chain occupancies in the alkylperoxo complex was more difficult and we were unable to resolve differential occupancies with certainty. Based on the similar residual densities for the two occupancies, each orientation was modeled at 50% occupancy. In contrast, in the anhydride structure, the side-chain of Y447 was unambiguously found in the metal-bound orientation. No density for the dissociated orientation was observed.

Results

Characterization of the reaction of 3,4-PCD, 4FC and O_2 in solution. The k_{cat} for turnover of 4FC is reduced 100-fold relative to that of PCA at 22 °C, pH 7.5, and the K_m values for the organic substrate and O_2 are increased 8-fold and 5-fold, respectively (Table 1). The reaction of 3,4-PCD with 4FC and O_2 yielded a UV-absorbing product with a λ_{max} of 256 nm, $\varepsilon = 17 \pm 0.8$ $mM^{-1}cm^{-1}$ as expected for the *ortho* cleavage product 3-fluoromuconate (Fig. 13A)[295]. Accordingly, acidification of the reaction product to pH 2 converted it to the characteristic lactonization product of 3-fluoromuconate (Fig. 13B)[296].

A comparison of the optical features of 3,4-PCD and anaerobic E-4FC is shown in Fig. 13C. IDO-catechol complexes exhibit additional catecholate-to-Fe^{3+} LMCT bands. Although binding of 4FC elicited changes in the LMCT bands (most notably at wavelengths >500 nm), the absorbance of the complex decreased relative to that of E-PCA (Fig. 13C). These results suggest that 4FC directly coordinates the metal, but this complex has different electronic properties than E-PCA. The structural basis for this change is described below. The anaerobic binding reaction of 3,4-PCD and 4FC at pH 7.5 at 22 °C was studied using stopped-flow. The time courses at 660 nm over a range of 4FC concentrations could each be fit by a single exponential expression (Fig. 13D), consistent with a one-step reaction. However, the plot of reciprocal relaxation time (RRT) versus 4FC concentration is hyperbolic (Fig. 13E). This indicates that the binding process actually occurs in at least two steps where a fast, reversible initial complex is followed by a slower ligand rearrangement or conformational change that gives rise to the chromophoric change. The K_d of the initial complex (k_{-1}/k_1) is given by the apparent K_d (7.5 ± 0.4 mM) for the hyperbolic fit, while the forward ($k_2 = 34 \pm 0.9$ s^{-1}) and reverse ($k_{-2} \sim 0.4$ s^{-1}) rate constants for the second step are determined by analysis of the maximum and intercept values of the hyperbola (see Materials and methods)[292]. The overall K_d for substrate binding is thus predicted to be ~88 μM, which compares favorably to the value of 75 μM determined by a direct optical titration (Fig. 13F). The weak initial binding without a chromophoric change may indicate that 4FC first binds in the hydrophobic active site away from the iron. These experiments show that none of the steps in the 4FC binding process correspond to the rate-limiting step of the reaction ($k_{cat} = 0.71$ s^{-1} at 22 °C).

To study the reaction with O_2, anaerobic E-4FC was rapidly mixed with oxygenated buffer at 4 °C on the stopped-flow. When 100 μM E-4FC (see Materials and methods) was

Table 1. Steady-state kinetic parameters for the 3,4-PCD reaction with PCA and 4FC.

	k_{cat}, pH 7.5, 22 °C (4 °C) (s^{-1})*	k_{cat}, pH 8.5, 22 °C (s^{-1})*	k_{cat}, pH 6.5, 22 °C (s^{-1})*	$K_{m,organic}$ (μM) ‡	$K_{m,O2}$ (μM) ‡	$k_{cat}/K_{m,organic}$ ($M^{-1}s^{-1}$)	$k_{cat}/K_{m,O2}$ ($M^{-1}s^{-1}$)
PCA	81 ± 20 (30†)	106 ± 10	54 ± 9	29†	56†	2.8×10^6	1.5×10^6
4FC	1.2 ± 0.09 (0.18 ± 0.05)	0.51 ± 0.02	0.70 ± 0.04	240 ± 20	270 ± 100	5.0×10^3	4.4×10^3

All values were determined at 22 °C unless stated otherwise. Errors represent one standard deviation. All values normalized to the number of Fe^{3+}-containing active sites.

* The PCA assays contained 2 mM substrate and 750 µM dissolved O_2. 4FC assays contained 2 mM substrate. For 4FC, due to the high K_m for O_2, the k_{cat} values and the corresponding error were extracted from the hyperbolic fit to the initial velocity vs. dissolved O_2 concentration plots.

† value is from reference[179].

‡ values for 4FC determined at pH 7.5.

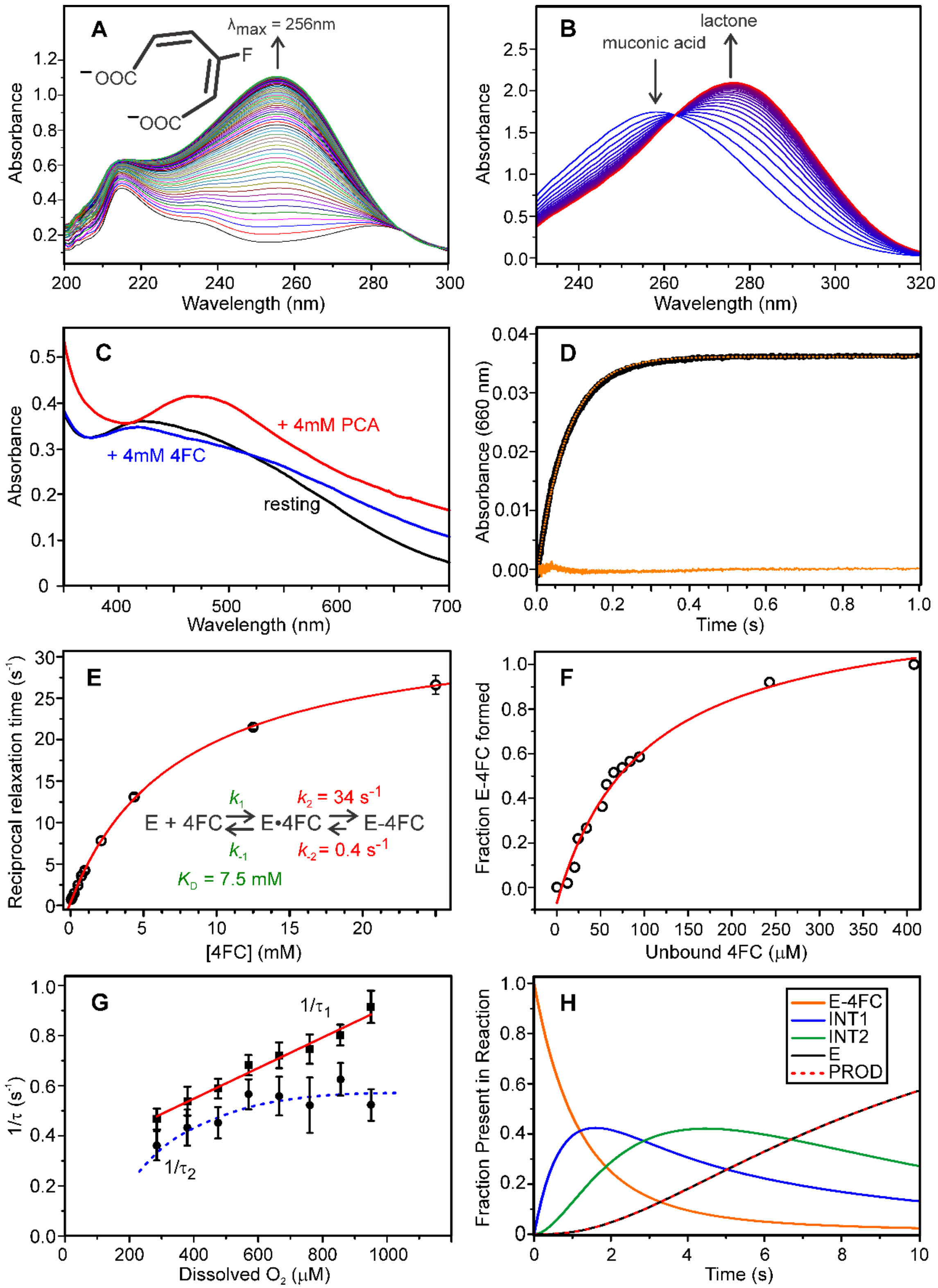
A
λmax = 256nm
−OOC
F
−OOC
Absorbance
Wavelength (nm)
B
lactone
muconic acid
Absorbance
Wavelength (nm)
C
+ 4mM PCA
+ 4mM 4FC
resting
Absorbance
Wavelength (nm)
D
Absorbance (660 nm)
Time (s)
E
Reciprocal relaxation time (s-1)
k1
k2 = 34 s-1
E + 4FC ⇄ E•4FC ⇄ E-4FC
k-1
k-2 = 0.4 s-1
KD = 7.5 mM
[4FC] (mM)
F
Fraction E-4FC formed
Unbound 4FC (μM)
G
1/τ1
1/τ2
1/τ (s-1)
Dissolved O2 (μM)
H
E-4FC
INT1
INT2
E
PROD
Fraction Present in Reaction
Time (s)

Fig. 13 (previous page): Transient kinetics of the binding reaction of 4FC to 3,4-PCD and the reaction of E-4FC with O_2. (A) Reaction product formation in the ultraviolet range. 5 μM 3,4-PCD in oxygenated 0.3 M HEPES pH 7.5 was mixed with 4FC (final conc. 75 μM). Spectra are shown at 10 s intervals. (B) Acidification of the product shown in panel A to pH 2 with HCl resulted in conversion to the corresponding lactone product with λ_{max} 277 nm[296]. (C) Optical spectra of resting 3,4-PCD (black), the anaerobic E-4FC complex (blue), and the anaerobic E-PCA complex (red) at 100 μM concentration. (D) Representative single wavelength data (black) for the reaction of 50 μM 3,4-PCD with 4 mM 4FC monitored at 660 nm. The dotted orange trace is a 1-exponential fit with RRT 13 s^{-1} (residuals in the solid orange trace). (E) Substrate concentration dependence of the single observed RRT for the reaction of 50 μM 3,4-PCD with 4FC at 22 °C and pH 7.5. Error bars represent one standard deviation. The red trace shows a hyperbolic fit for the data with y-intercept 0.4 s^{-1} (k_{-2}) and an extrapolated maximum of 34.4 s^{-1} ($k_2 + k_{-2}$). The K_d for the fit gives the K_d for the preceding unobserved binding step (K_d = 7.5 mM). The extracted kinetic model for 4FC binding is shown as an inset. (F) Adair plot for the anaerobic titration of 100 μM 3,4-PCD with 4FC at pH 7.5. The optical changes were monitored at 700 nm. The K_d for 4FC determined from the hyperbolic fit to the data (red trace) is 75 μM, close to that calculated from the transient kinetics (88 μM). (G) Plot of the dependence of $1/\tau_1$ and $1/\tau_2$ in the reaction of 50 μM E-4FC with O_2 on the dissolved O_2 concentration. Error bars represent one standard deviation. The red trace is a linear fit to the $1/\tau_1$ data. The blue dashed trace illustrates the observed nonlinear dependence of $1/\tau_2$. This data is further analyzed in Figure 14C. (H) Speciation plot of the single-turnover reaction of 100 μM E-4FC with 950 μM O_2 modeled using the kinetic parameters shown on page 51 (Tenua version 2.1).

reacted with 950 μM O_2 at 4 °C, the time course of this pseudo first order reaction as monitored at 540 nm could be fit well by a sum of three exponential functions with RRTs of 0.92 s^{-1}, 0.50 s^{-1} and 0.16 s^{-1} (Fig. 14B). Although the magnitudes of these RRTs are similar, they are resolvable because the phases have opposite amplitudes at 550 nm. Three phases imply that there are at least three steps in this portion of the reaction cycle. The slowest phase has a RRT very similar to k_{cat} from steady-state experiments (0.18 s^{-1} at 4 °C). In all previous studies, we and others have found that product release is the rate-limiting step for IDOs, and the preceding ring-opening step is irreversible[180, 183, 184, 287]. This irreversible step allows the slowest RRT to be associated with the rate constant for a specific step in the reaction, presumably product release. This assignment is supported by the observation that the spectrum of the resting enzyme is observed at the completion of the reaction. Only the fastest two phases show a dependence on O_2 concentration, as expected if the slow product release is preceded by the irreversible ring-opening step (Fig. 13G). Assuming that the fastest two phases arise from two sequential steps in which at least the first is reversible, the RRTs will be described by the two roots of a quadratic equation. Consequently, the plots of the sum and products of the RRTs vs. O_2 will be linear and the all of

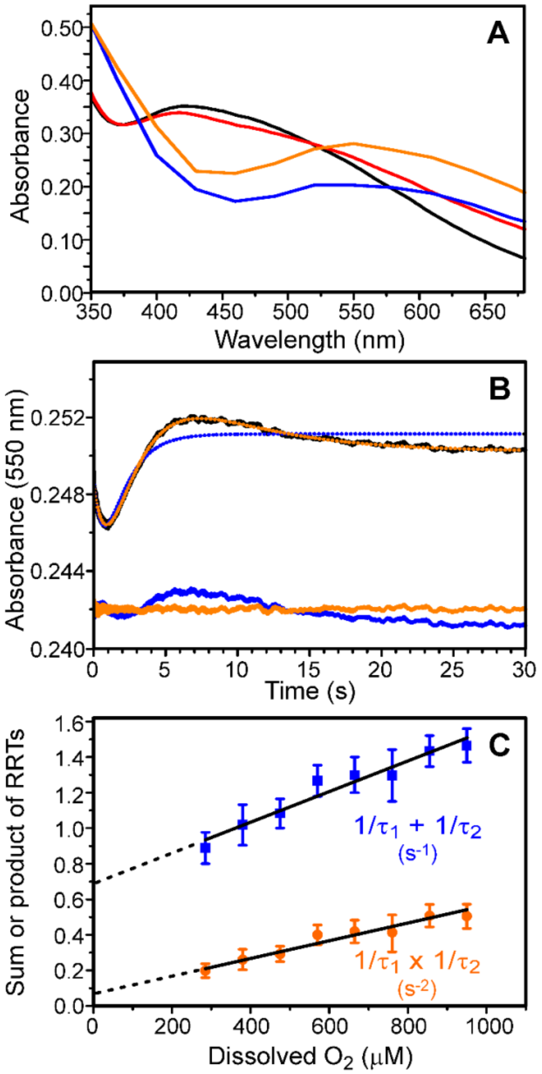

Fig. 14: Transient kinetics of the reaction of E-4FC with O_2. (A) Optical spectra of resting 3,4-PCD (black), the anaerobic E-4FC complex (red), and the extracted spectra of Int1 (blue) and Int2 (orange) shown at 100 μM concentration. (B) Representative single wavelength data (black) for the reaction of 100 μM E-4FC with 950 μM O_2 monitored at 550 nm. The blue and orange dotted traces represent 2- and 3-exponential fits to the data, respectively. Residuals for the fits are shown in their respective colors. (C) Plots of the dependence of the sum (blue) and product (orange) of the two largest RRTs for the reaction of 50 μM E-4FC on the dissolved O_2 concentration. Error bars represent one standard deviation.

the rate constants for the steps can be extracted (see Fig. 14C and Materials and methods). The second fastest step is effectively irreversible as expected. The forward rate constant for the fastest step (8.7 x 10^2 ± 0.8 $M^{-1}s^{-1}$) appears to be the bimolecular rate constant for O_2 binding[184].

This value is 600-fold slower than that for O_2 binding to E-PCA. The derived minimum kinetic model for the reaction is:

$$\mathrm{E{+}4FC} + \mathrm{O_2} \underset{k_{-3} = 0.11\ \mathrm{s^{-1}}}{\overset{k_3 = 8.7\mathrm{x}10^{2}\mathrm{M^{-1}s^{-1}}}{\rightleftarrows}} \mathrm{Int1} \xrightarrow{k_4 = 0.58\ \mathrm{s^{-1}}} \mathrm{Int2} \xrightarrow{k_5 = 0.16\ \mathrm{s^{-1}}} \mathrm{E{+}P}$$

Using global analysis of the single-turnover reaction monitored at 13 wavelengths, we extracted the pure component optical spectra of the intermediates formed during the reaction with O_2 as shown in Fig. 14A. The intermediates have optical spectra with characteristics similar to those reported for transient oxygenated intermediates in related enzymes[183, 184]. The presence of relatively intense optical features for both of the observed intermediates shows that the metal remains in the +3 oxidation state, because reduction of the metal would bleach the optical spectrum almost entirely. The rate constants for the interconversions preclude trapping of the intermediates at high yield in solution (Fig. 13H). However, the slow rate of the reaction between E-4FC and O_2 and of each subsequent step proved to be a great advantage in trapping O_2-bound intermediates *in crystallo*.

Crystal structure of 3,4-PCD with 4FC bound. 3,4-PCD was crystallized in a new space group I222 with three αβ-protomers in the asymmetric unit, differentiated as protomers A, B and C in the model hereafter. No significant differences were noted in the resting active site structure of I222 crystals of 3,4-PCD and those of the monoclinic C2 crystals previously studied, but the I222 crystals generally diffracted to higher resolution (1.3-1.5 Å limit). Data collection and refinement parameters for each of the structures solved for this study can be found in Table 2. The iron occupancy in all structures was refined at 75% based on ICP-MS analysis of the metal content of the enzyme batches used for the crystallizations.

Electron density maps from 3,4-PCD crystals soaked with 4FC at pH 8.5 under anaerobic conditions revealed direct binding of 4FC to the metal in all three protomers (Fig. 15A and B). In contrast to the axial binding mode observed in the PCA complex, 4FC bound in the equatorial plane, *trans* to H460 and Y408. This amounts to a roughly 100° rotation of the aromatic plane relative to the plane of PCA in the E-PCA complex (Fig. 15C). Two conformations of 4FC at roughly equal occupancy were observed, evidenced by the residual density apparent for two positions of the fluorine. The only difference in the binding mode is in the position of fluorine; the iron bond distances and position of the ring are essentially identical. The Fe-O_{4FC} bond lengths are found to be 2.1 Å for both $O1_{4FC}$ and $O2_{4FC}$ in all of the protomers. There is no

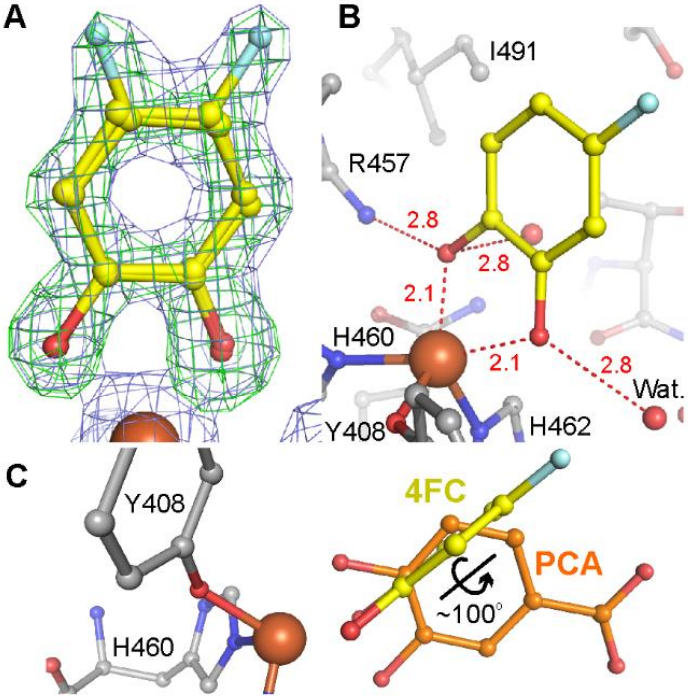

Fig. 15: Structure of the anaerobic complex of 3,4-PCD with 4FC (E-4FC) at pH 8.5. (A) Electron density maps of the active site of the E-4FC complex. The blue mesh is the 2|Fo|-|Fc| map contoured at 1.0 σ after refinement with 4FC. The green mesh is the |Fo|-|Fc| ligand-omit map contoured at +3.8 σ. (B) Bonding environment of 4FC. Only one orientation of 4FC is shown for clarity. Possible interactions are shown as red dashed lines and the distances are given in Å. (C) Comparison of E-4FC and E-PCA (PDB ID: 3PCA). The ring of 4FC is rotated 100° relative to that of PCA. Atoms are shown as spheres. Carbon atoms are gray (protein residues), yellow (4FC) or orange (PCA). Oxygen atoms and solvent atoms are colored red, nitrogen atoms blue and the iron atom brown.

asymmetry in the Fe-hydroxyl bonds as observed for PCA[181]. In each protomer, Y447 is dissociated from the metal and interacts with the phenol group of Y16; its orientation is superimposable with that in the E-PCA complex. The details of the ligand and Y447 refinement strategies and model occupancies are given in the Materials and methods section.

***In crystallo* reaction of E-4FC with O_2 at pH 6.5 and 7.5.** Crystals of 3,4-PCD were transferred to drops supplemented with 100 mM 4FC and reacted under aerobic atmosphere at 22 °C for 15-25 minutes before freezing. Following the incubation, 4FC ring-cleaved product was detectable in the drop. Electron density maps calculated from data sets at pH 7.5 suggested an averaging of multiple species bound to the metal in each protomer. No distinct species could be modeled into these active sites. In contrast, when the experiment was repeated at pH 6.5, the resulting electron density maps had strong residual positive density for a metal-bound species clearly distinct from the 4FC present in the E-4FC complex. Interpretable electron density was only observed in active site A; solvent was modeled into active sites B and C to account for the missing density. The residual density in protomer A suggested the presence of an axially-oriented ring with connected density that "wrapped" around the metal in the equatorial plane to the coordination site *trans* to H460 (Fig. 16A). This site is where the peroxo has been proposed to form during the reaction of O_2, and we accounted for this density by modelling an alkylperoxo species in active site A (Fig. 16A and B). A 4FC-semiquinone-Fe-$O_2^{\bullet-}$ was also tested, but the refined models strongly supported the existence of a covalent bond between $C2_{4FC}$ and $O1_{O2}$.

The peroxo bridges between the metal and the substrate (Fe-$O1_{O2}$ = 1.8 Å, Fe-$O2_{O2}$ = 2.5 Å, Fe-O-O angle = 100°). The refined peroxo O-O bond distance was 1.5 Å. The carbon of 4FC attacked by O_2 is sp^3-hybridized based on the observed bond angles and the ring is puckered, consistent with oxidation of 4FC. $O2_{4FC}$ remains coordinated to the metal (1.9 Å), but it also interacts with the guanidinium group of R457 (2.8 Å) (Fig. 16B). In contrast, the $O1_{4FC}$ ketone is dissociated from the metal (3.8 Å) and interacts with the Q477 side-chain. Y447 is found in both metal-bound and dissociated orientations at roughly equal occupancy. As noted, the ring of 4FC is axially oriented, indicating that 4FC rotates during (or prior to) the reaction with O_2. The Fe-ligand bond distances are considerably shorter than in the 2,3-HPCD alkylperoxo[72], consistent with the expected higher metal oxidation state.

***In crystallo* reaction of E-4FC with O_2 at pH 8.5.** The dramatic effect of pH on stabilization of the alkylperoxo intermediate prompted us to repeat the experiment at pH 8.5. The ligand-omit maps of the crystals reacted at pH 8.5 clearly indicated splitting of the peroxo moiety and formation of two separate metal-bound species. In the proposed mechanism, heterolytic O-O bond scission of the peroxo leads to formation of an anhydride intermediate *via* a Criegee-type rearrangement where one O atom of O_2 migrates into the ring of substrate, while the second O atom of O_2 remains metal-bound as hydroxide. Based on the clear evidence of O-O bond scission, a 4FC-derived anhydride and iron-hydroxide were modeled into the density (Fig. 16C and D). A lactone intermediate (present in the extradiol pathway) was also tested, but the anhydride provided a better fit. No intermediates could be modeled into sites B and C. The bond distance between the metal-bound hydroxide and the bridging O-atom of anhydride is 2.3 Å. The

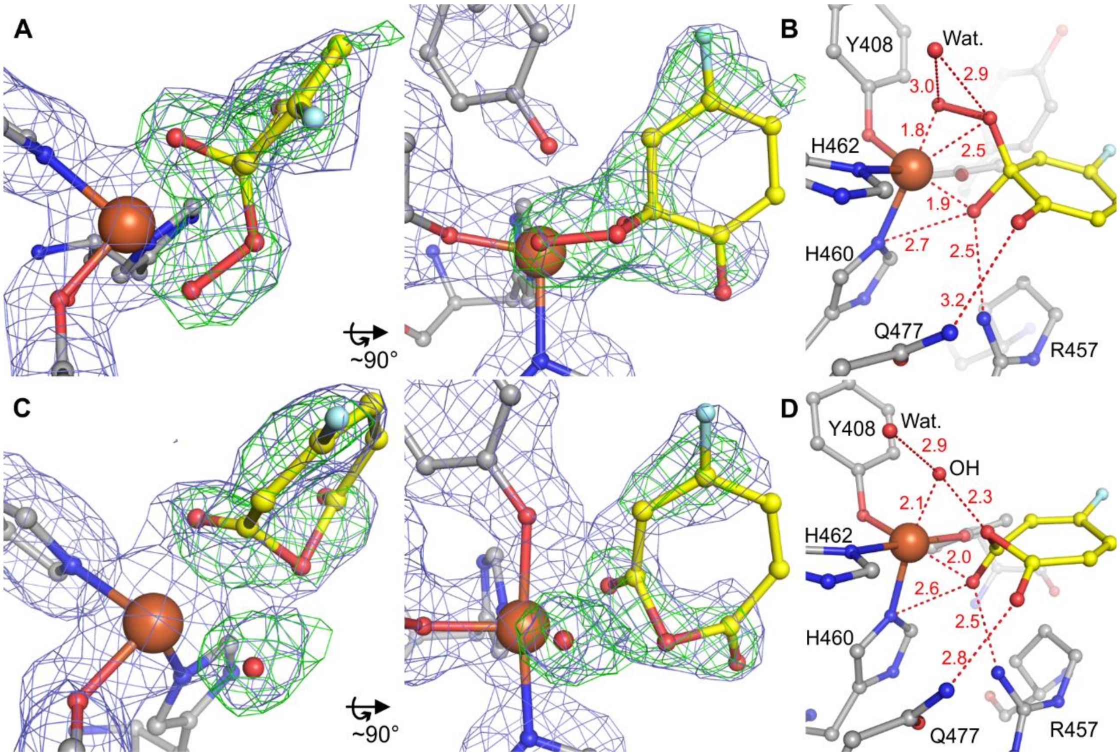

Fig. 16: Structures of the intermediate species observed upon reacting 3,4-PCD crystals with 4FC and O_2 at pH 6.5 and 8.5. (A) Two views of the electron density observed at pH 6.5 and the 4FC-alkylperoxo model (yellow). (B) Bonding environment of the 4FC-alkylperoxo intermediate. (C) Two views of the electron density observed at pH 8.5 and the 4FC-anhydride model (yellow) and hydroxide (red sphere). (D) Bonding environment of the 4FC-anhydride intermediate. In panels A and C the blue mesh is the 2|Fo|-|Fc| map contoured at 1.0 σ after refinement with the ligands and the green mesh is the |Fo|-|Fc| ligand-omit map contoured at +3.5 σ. In panels B and D the possible interactions are shown as dashed red lines and the distances are given in Å. Atom color code is as in Figure 15.

anhydride remains metal-bound and interacts through the ketonized $O2_{4FC}$ (2.0 Å) while the ring remains in the axial orientation. The position of the hydroxide is roughly the same as that of the proximal oxygen of the alkylperoxo moiety, but the Fe bond distance has increased to 2.1 Å. Critically, Y447 is found exclusively in the metal-bound orientation.

Determination of whether the Fe^{3+} is photoreduced to Fe^{2+} during data collection. Two lines of evidence suggest that the Fe^{3+} active site was not photoreduced to Fe^{+2} during data collection on the synchrotron beamline. First, a model of the alkylperoxo derived from a data set collected on an in-house diffractometer instrument resulted in nearly identical iron-ligand bond distances to those observed in the model produced from the synchrotron data (Fig. 17A). The crystal soaking conditions for this sample were the same as described above. Photoreduction to Fe^{2+} would have resulted in considerably longer bond distances in the synchrotron-derived model. Second, the crystals retained the characteristic burgundy color of the oxidized enzyme throughout data collection. Reduction of the cofactor to Fe^{2+} would have bleached the chromophore entirely, resulting in a colorless crystal. Optical spectra of single crystals obtained after diffraction data collection were also consistent with retention of the Fe^{3+} oxidation state (Fig. 17B). The optical spectra of the crystals with bound intermediates are generally consistent with those of the intermediates in solution. However, definitive assignments cannot be made at this time due to the high sensitivity of the spectra on the orientation of the crystal in the micro-spectrophotometer beam and the mixing of optical features from different intermediate species trapped in the active sites of the crystal.

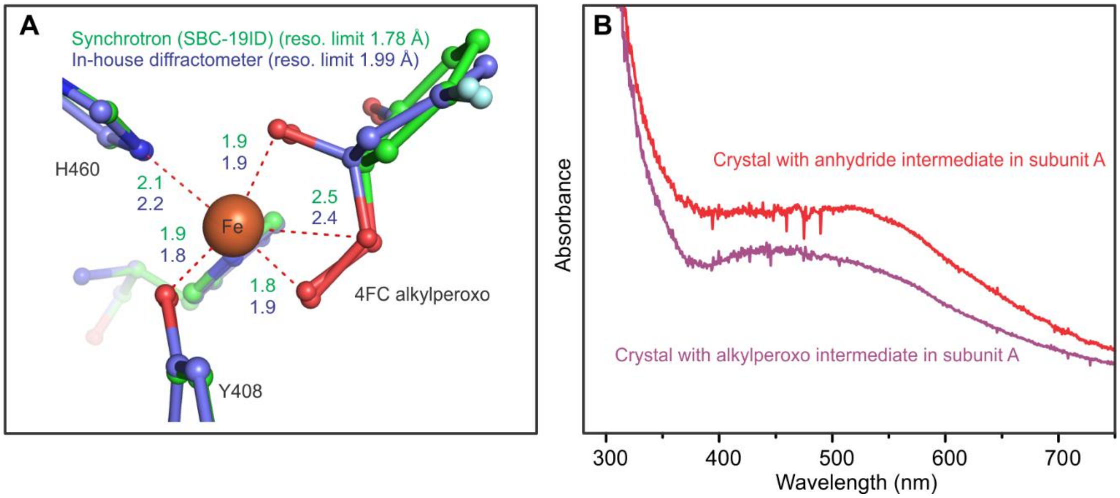

Fig. 17: The active site iron is not reduced by the synchrotron X-ray beam. (A) Comparison of the structure of the alkylperoxo intermediate solved using synchrotron data with that using data from a much lower X-ray intensity in-house generator. The Fe-ligand bond lengths (in Å) are the same within error and are characteristic of Fe^{3+}. (B) Optical spectra recorded of single enzyme crystals after data collection at the synchrotron. The spectra are averages of spectra recorded at 6 crystal orientations. The intense chromophore shows that the metal center is Fe^{3+}. The spectral line-shapes are consistent with Int1 and Int2 observed in solution but a direct assignment cannot be made because only one subunit in the crystal is fully occupied by the intermediate but all subunits contribute to the optical spectrum. Also, the spectra are sensitive to the orientation of the crystal in the light path. The absorbance values on the y-axis are arbitrary.

Table 2. X-ray data collection and refinement statistics for the various 3,4-PCD-4FC crystal complexes

	3,4-PCD resting pH 6.5	**3,4-PCD resting pH 8.5**	**3,4-PCD + 4FC pH 8.5**	**3,4-PCD + 4FC + O_2 pH 6.5**	**3,4-PCD + 4FC + O_2 pH 8.5**
PDB ID	4WHP	4WHO	4WHS	4WHQ	4WHR
Data collection					
Space group	I222	I222	I222	I222	I222
Cell dimensions					
a, b, c (Å)	127.92, 140.77, 168.19	128.09,140.64, 168.66	128.01, 140.66, 168.14	128.26, 140.72, 167.82	128.44, 140.64, 168.29
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Wavelength (Å)	0.97918	0.97918	0.97918	0.97918	0.97918
Resolution (Å)	50-1.54	50-1.83	50-1.35	50-1.78	50-1.58
Total/Unique reflections	1026896/ 206314	559978/ 121908	1617767/ 309614	689265/ 140429	756615/ 188155
R_{merge} (%) *,†	12.4 (47.5)	16.2 (45.3)	9.9 (50.6)	7.4 (52.8)	14.2 (35.1)
I/σI*	30 (3.7)	25 (5.1)	34 (3.3)	30 (2.3)	24 (2.2)
Completeness (%) *	93.0 (100)	89.3 (99.9)	94.6 (100)	97.5 (96.1)	91.3 (93.7)
Redundancy*	5.0 (3.7)	4.6 (4.6)	5.2 (5.1)	4.8 (3.8)	4.0 (2.1)
Model Refinement					
Resolution (Å)	39.0-1.54	32.5-1.83	29.5-1.35	39.0-1.78	31.6-1.58
R_{work}/R_{free} (%)‡	0.140/0.172	0.180/0.216	0.136/0.167	0.164/0.195	0.148/0.192
Average B, all atoms (Å^2)	21	22	19	27	25
ESU (Å) §	0.033	0.068	0.024	0.065	0.047
RMSDs					
Bond lengths (Å)	0.0072	0.012	0.008	0.011	0.0090
Bond angles (°)	1.3	1.5	1.4	1.5	1.4
Ramachandran analysis					
Favored (%)	97.8	97.7	98.1	97.3	97.6
Allowed (%)	2.2	2.3	1.9	2.7	2.4

All data collected on synchrotron beamline APS SBC-CAT 19ID-D

* Highest resolution shell is shown in parentheses

† $R_{sym} = \Sigma_{hkl}\Sigma_i |I_{hkl,i} - \langle I\rangle_{hkl}|/\Sigma_{hkl}\Sigma_i|I_{hkl,i}|$, where I_{hkl} is the intensity of a reflection and $\langle I\rangle_{hkl}$ is the average of all observations of the reflection

‡ R_{free}, R-factor calculated from 5% of the data excluded from refinement

§ Estimated overall coordinate error (ESU) based on maximum likelihood

RMSD, root mean square deviation from ideal geometry

Discussion

The current results show that 3,4-PCD catalyzes ring cleavage of 4FC to yield the expected intradiol product, but also that the reaction occurs at a greatly decreased rate relative to that with the native substrate, PCA. Importantly, the rate decrease is seen throughout the reaction cycle, allowing intermediates in the key steps following O_2 binding to be detected in solution and *in crystallo*. The visualization of the alkylperoxo and anhydride intermediates from the IDO reaction cycle provides direct insight into the reaction mechanism of this enzyme class at a level that has not previously been possible. Moreover, it allows the results of four decades of kinetic, mutational, spectroscopic, and computational studies to be considered in a single framework. These aspects of the study are discussed here.

Relationship to spectroscopic and computational studies of the IDO mechanism. The structural model of the alkylperoxo intermediate reported here is geometrically nearly identical to the 5C alkylperoxo species proposed from computational studies that were based on spectroscopic and structural studies of the E-PCA complex[25, 213]. This convergence suggests that the mechanistic insight into the ring cleavage mechanism for 4FC is relevant to that for PCA. Several important aspects of the predicted mechanism are confirmed by the structure. First, the initial attack on the substrate by O_2 is computed to be a concerted process (Fig. 12, **3**) that directly forms the bridging peroxo moiety. This is the species we observe (Fig. 16A). In our *in crystallo* studies of 2,3-HPCD, a substrate radical-Fe^{2+}-superoxo intermediate precedes the alkylperoxo species[72]. Similar discrete precursor complexes, such as a substrate radical-Fe^{2+} or substrate radical-Fe^{3+}-superoxo species, have been proposed in previous studies of the IDOs and models[213, 218, 226]. Although not definitive, we find no evidence for these species in the IDO crystal. More significantly, the failure to observe any fully bleached intermediates in the solution studies makes any Fe^{2+} species with a significant lifetime unlikely. These observations are consistent with the proposed concerted attack mechanism **3**[25]. Second, while the initial concerted O_2 attack results in a 6C iron species **4**, a rapid rearrangement to a 5C iron species **5** is computationally predicted to be required for stability and reestablishment of the $S = 5/2$ spin state of the iron[25, 213]. The observed alkylperoxo species is in the predicted shifted position (Fig. 16A). Third, a subsequent shift in the position of the bridging peroxo moiety is predicted to increase the reactivity of the species and establish a geometry **6** that facilitates protonation of the proximal peroxo oxygen[25,

[213]. This proton is proposed to drive O-O bond cleavage and insertion. It may derive from a metal-bound solvent or deprotonation of Y447 as it returns to the iron as seen in **6**. The shifts in the positions of the substrate and peroxo would open the exchangeable axial iron binding site for Y447 or solvent. In the structures reported here, the axial binding site is open in the alkylperoxo intermediate (Fig. 16A), but Y447 is bound following O-O bond cleavage in the anhydride intermediate (Fig. 16C). These observations strongly support both the computed mechanism and the role of Y447 as the source of the critical proton.

The stopped flow studies described here reveal two intermediates following addition of O_2 to the E-4FC complex. While it is possible that these two intermediates are the species found *in crystallo*, we cannot make this correlation at present.

Intradiol versus extradiol dioxygenase ring cleaving specificity. The IDO and EDO families have each been studied extensively with the goal of understanding their individual O_2 activation mechanisms and the basis for their ring cleaving specificity. It is now possible to directly compare representative structures of the alkylperoxo intermediates from these families. EDOs and IDOs mechanistically converge at an alkylperoxo intermediate, and both families have been proposed to utilize a concerted Criegee-like mechanism for oxygen insertion into substrate[32, 297]. While the current study does not address the detailed mechanism of oxygen insertion, it does provide insight into the key question: What determines the position of oxygen insertion when ostensibly the same alkylperoxo intermediate is formed in each enzyme class[195, 226-228, 298]? One possibility supported by computational studies is that the alignment of the O-O bond relative to the scissile bonds of the substrate determines whether acyl (intradiol) or alkenyl (extradiol) migration occurs during the rearrangement. Migration proceeds preferentially into the co-planar bond[213, 218, 299]. Comparison of the geometries of the EDO and IDO alkylperoxo intermediates shows that they differ substantially in the O-O and C-C bond alignments primarily because the substrate bonding to the iron is not the same. In the EDO, the peroxo and both substrate ring oxygen substituents coordinate to the iron, while only one hydroxyl coordinates the iron in the IDO (Fig. 18). In 3,4-PCD this coordination mode comes about as a result of the rearrangement described above and is stabilized by interaction with second sphere conserved residues Q477 and R457. As shown in Figure 18, the co-planar O-O and substrate C-C bond alignment favors formation of the anhydride in 3,4-PCD while alkenyl migration to form the lactone would be favored by the alternative alignment in 2,3-HPCD alkylperoxo.

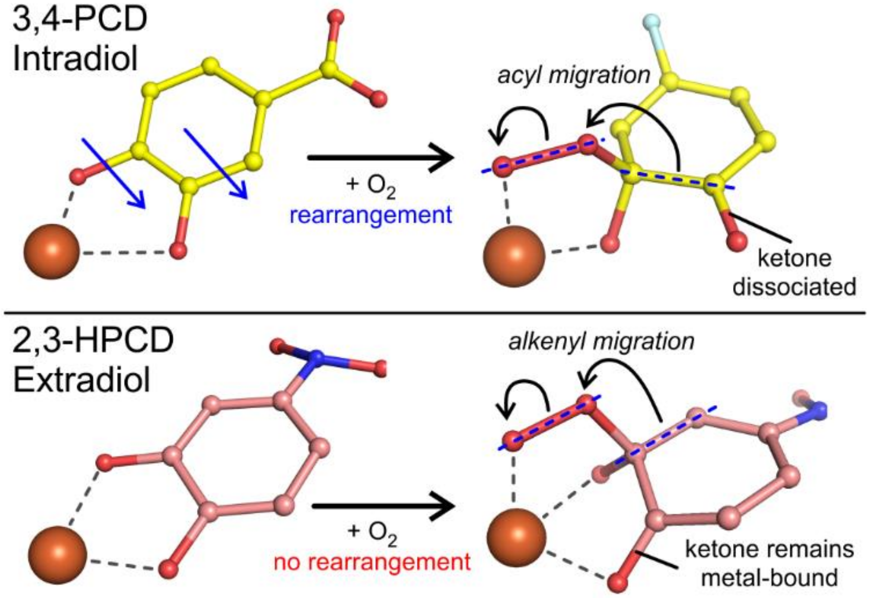

Fig. 18: Comparison of the geometries of the alkylperoxo intermediates in 3,4-PCD and 2,3-HPCD. The coordination mode of the intermediate may dictate acyl versus alkenyl migration and thereby cleavage specificity. The 2,3-HPCD and 3,4-PCD-PCA models are PDB structures 2IGA and 3PCA, respectively.

Equatorial binding of 4FC and its effects on the reaction with O_2. Axial binding of catecholic substrates is a characteristic of all classes of IDOs[180, 181, 197, 199]. The alternative equatorial 4FC binding mode provides insights into the driving force behind the substrate binding-mediated dissociation of Y447. Two possible origins for this rare endogenous ligand dissociation are direct displacement by a substrate hydroxyl function and/or a requirement to retain a constant charge at the metal center[179, 194]. In the current case, Y447 dissociates when 4FC is bound in the equatorial plane, and thus no direct competition for the iron coordination site is possible. On the other hand, coordination of 4FC in the equatorial plane as a dianion adds two negative charges while only displacing one negative hydroxide. The charge would be rebalanced by dissociating Y447 to its alternative site, supporting charge maintenance as the basis for ligand reorganization.

It is unknown whether 4FC can react with O_2 while bound in the equatorial plane. Rotation to the axial plane clearly occurs at some point during the generation of the alkylperoxo intermediate. Intriguingly, although 4FC is rotated ~100° relative to PCA in the substrate complex, the two substrates experience analogous metal environments, i.e. 4FC has a *trans* His ligand (H460) and *cis* phenolate ligand (Y408) like PCA (*trans* H462 and *cis* Y408). Thus, one might expect the rates of O_2 attack to be comparable. Two steric factors might prevent rapid reaction with O_2 while 4FC in bound in the equatorial plane. First, the equatorial binding mode of 4FC does not create a hydrophobic pocket for O_2 to bind near the iron as we have noted for the axially-bound E-PCA complex[181, 203]. Second, the "slot-like" active site is constructed to promote axial binding. As a result, equatorial substrate binding creates tight van der Waals contacts with walls of the active site. This would prevent the shift in substrate orientation required during the formation of the 5C alkylperoxo intermediate. No such restrictions are present for axial substrate binding. Given these restrictions, it seems likely that 4FC is actually bound as an equilibrium mixture of axial and equatorial orientations with only a small fraction bound in the axial orientation, making it undetectable in the electron density. If only the axial fraction is reactive towards O_2, then the 600-fold reduction in the bimolecular reaction rate may be, in part, representative of the necessity of the substrate to reorient before it can react with O_2.

Conclusions

Together with comparable studies of the EDO reaction intermediates, the structures of the alkylperoxo and anhydride intermediates of the IDO reaction cycle make available a broad view of the factors that promote O_2 activation and specificity in ring-cleaving dioxygenases. The formation of alkylperoxo intermediates is widely proposed, especially in the large 2-His-1-carboxylate family of nonheme Fe^{2+} oxygenases. Although the mechanism of formation differs in the IDO case, the resulting complex is similar, and thus its structural parameters and the basis for its reactivity are significant. Finally, this study adds to a small but growing list that show enzyme intermediates can be trapped in high yields when the reactions are carried out in a crystal. This approach may provide an experimental inroad into a fundamental question in enzymology concerning the mechanism of efficient transit through multistep reaction cycles.

Acknowledgements

Diffraction data were collected at Argonne National Laboratory, Structural Biology Center at the Advanced Photon Source. Argonne is operated by UChicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357. We thank Prof. Carrie Wilmot and Dr. Ed Hoeffner at the University of Minnesota Kahlert Structural Biology Laboratory for many discussions and data collection.

CHAPTER 2

Kinetic characterization of the reaction of protocatechuate 3,4-dioxygenase with alternative catecholic substrates

Abbreviations:

3,4-PCD: protocatechuate 3,4-dioxygenase
PCA: protocatechuate (3,4-dihydroxybenzoate)
CAT: catechol
4FC: 4-fluorocatechol
4CC: 4-chlorocatechol
4MC: 4-methylcatechol
ES complex: enzyme-substrate complex, used as a generic term for the complex of 3,4-PCD with one of the catecholic substrates
IDO: intradiol dioxygenase
RRT: reciprocal relaxation time
LMCT: ligand-to-metal charge-transfer
1,2-CTD: catechol 1,2-dioxygenase
CCD: chlorocatechol dioxygenase

Summary

Protocatechuate 3,4-dioxygenase (3,4-PCD) is an intradiol dioxygenase enzyme that catalyzes the oxygenolytic cleavage of catecholic substrates using a non-heme Fe^{3+}-cofactor and O_2 as a cosubstrate. The enzyme uses an ordered sequential mechanism in which the catecholic substrate binds to the iron ion first, then reacts with O_2 to afford the ring-cleaved products. The native substrate protocatechuate (3,4-dihydroxbenzoate, PCA) is turned over too rapidly by the enzyme to study the transient kinetics of the reaction with O_2 in detail. Alternative catecholic substrates provide a means to study these catalytic steps, provided the reaction is slowed sufficiently. Here, the steady-state and transient kinetic characterization of the reaction of 3,4-PCD with a series of alternative substrates is reported. The alternative substrates used include catechol, 4-fluorocatechol, 4-chlorocatechol, and 4-methylcatechol. The kinetic study of these substrates facilitates comparisons among the alternative reactions and with the native substrate PCA. Insights into the rate-limiting steps and the effects of the catechol substituents on the catalytic reaction are gained. These results have implications for the mechanisms of substrate specificity of this family of enzymes and the role of the protein fold in orienting the catechol for the reaction with O_2.

Introduction

Intradiol ring-cleaving dioxygenases (IDOs) catalyze oxygen-dependent aromatic ring-cleavage reactions in a variety of soil-dwelling bacteria and fungi[32]. This class of enzymes uses a non-heme Fe^{3+} cofactor to activate catecholic substrates and O_2 for reaction, one of only two enzyme families to use the ferric form of the cofactor[25, 238]. The site of ring-cleavage catalyzed by IDOs is between the vicinal hydroxyl groups of the catechol and results in incorporation of both atoms of molecular O_2 into the product (dioxygenases) to generate a substituted muconic acid product[19]. There are several families of IDO enzymes that are differentiated by their substrate preference and their overall protein fold. Prominent examples include the protocatechuate 3,4-dioxygenases (3,4-PCDs) that preferentially cleave 3,4-dihydroxybenzoate (protocatechuate, PCA)[196, 203], methylcatechol and catechol 1,2-dioxygenases (1,2-CTDs)[197], hydroxyquinol 1,2-dioxygenases (1,2-HQDs)[200], chlorocatechol 1,2-dioxygenases (1,2-CCDs)[201, 202, 300], and a single, recently discovered member that cleaves catecholic intermediates of lignin biosynthesis[176]. All members use the Fe^{3+} oxidation state of the cofactor and have a conserved 2-His 2-Tyr iron-coordination ligand set that is found only in this enzyme class. In the resting state, the iron also binds a hydroxide anion resulting in a five-coordinate trigonal bipyramidal complex.

The most-studied IDO is the 3,4-PCD from *Pseudomonas putida* (reclassified from *Pseudomonas aeruginosa*)[204]. The biological role of 3,4-PCD is to cleave the native substrate PCA to β-carboxy muconic acid with concomitant incorporation of both atoms of O_2 into product. This acid is then converted to central metabolic intermediates via the β-ketoadipate pathway[301]. In the *P. putida* enzyme, the resting Fe^{3+} is coordinated by Y447 and H462 in the axial orientation and Y408 and H460 and hydroxide in the equatorial plane. The catalytic mechanism of ring-cleavage has been studied since the 1960s[187, 189, 190, 204, 287], but only recently have oxygenated intermediates been trapped for characterization (see Chapter 1). The proposed mechanism is shown in Figure 8 for PCA[25, 32, 179, 181, 184, 188, 192, 194]. The catecholic substrate binds directly to the metal as a dianionic chelate, leading to dissociation of the hydroxide ion and one of the endogenous Tyr ligands (Y447 in the *P. putida* enzyme). The electron transfer that leads to O_2 activation is proposed to occur after O_2 binds in a hydrophobic pocket adjacent to the metal. Because the metal is in a +3 oxidation state and has no affinity for O_2, O_2 is not proposed to bind directly to the metal before the electron transfer. This necessitates a unique O_2 activation mechanism in which two electrons are transferred in a concerted step to O_2 from catechol, one directly and the other *via* the iron d orbital manifold[25]. This results in the formation a Fe^{3+}-alkylperoxo intermediate without a spin violation. After protonation of the peroxo moiety, the Fe^{3+}-alkyl(hydro)peroxo intermediate decays to an anhydride species where one atom of O_2 is incorporated into the substrate ring. The other O atom remains metal-bound as hydroxide. Nucleophilic attack by this hydroxide opens the anhydride ring, generating the ring-cleaved product. The resting active site is regenerated by influx of a water molecule and by the return of Y447 to the metal during O-O bond cleavage.

Although the most critical reactive intermediates have been trapped in this enzyme (Chapter 1, Fig. 16), several important questions about the enzyme's catalytic cycle and the mechanism of substrate selectivity remain unanswered. In particular, it is uncertain how the protein architecture determines a preference for PCA over other catecholic substrates that are un-substituted or have halo or methyl-substituents. Secondly, it is unclear what determines the reactivity of specific IDO enzyme-catechol complexes (ES complexes) towards O_2. As discussed in Chapter 1, reactivity may be directly tied into the binding mode of the catechol substrate on the iron. However, it may also be that reactivity is related to the electronic characteristics of the substrate. Both factors may come into play, but it is unclear which one is more determinant in controlling the reaction rate with O_2.

In order to investigate these and other aspects of the catalytic cycle, we characterized the steady-state and transient kinetic reaction of 3,4-PCD with a series of alternative non-carboxylic substrates: catechol (CAT), 4-methylcatechol (4MC), and 4-chlorocatechol (4CC). The reaction with 4-fluorocatechol (4FC) has been characterized in detail in Chapter 1, but it will be discussed again in context of these other catecholic substrates. Our findings provide lessons into the role of the enzyme fold in producing a highly O_2-reactive ES complex and into the mechanism of substrate specificity. We also find evidence for ways that the enzyme maximizes the capture efficiency and turnover number of substrate. This work has implications on the general mechanisms that enzymes use to stabilize (or de-stabilize) their substrate binding mode to induce catalysis and may be useful for engineering 3,4-PCD and its relatives to accommodate a wider substrate range.

Materials and methods

Determination of catechol extinction coefficients and identification of reaction products. The catecholic substrates were purchased from either TCI America (4FC) or Sigma Aldrich. 4MC was purified by recrystallizing the solid in hexane and removing the solvent by evaporation. The other catechols were used without further purification. The extinction coefficients of the alternative substrates were determined by nuclear magnetic resonance (NMR) spectroscopy using an aromatic proton inventory calibrated to an internal tryptophan standard[290]. The values are reported at pH 8.5 and λ_{max} for each molecule. CAT $\varepsilon_{276\ nm}$ 2500 ± 660 $M^{-1}\ cm^{-1}$. 4CC $\varepsilon_{287\ nm}$ 3200 ± 40 $M^{-1}\ cm^{-1}$. 4MC $\varepsilon_{281\ nm}$ 2400 ± 210 $M^{-1}\ cm^{-1}$. The 4FC coefficient is reported in Chapter 1. Products were identified by first separating the molecule from the enzyme using an Amicon centrifugal filter (30 MWCO) and then comparing of their UV/vis spectra to those reported previously[295] and by liquid-chromatography mass-spectrometry (LC-MS). Mass spectrometry was performed at the UMN Chemistry Department Mass Spectrometry Laboratory and analyzed using a Waters Acquity UPLC/TQD instrument. The parent peak of the products showed a mass increase of 32 atomic mass units (amu) over substrate-only controls, consistent with incorporation of two oxygen atoms. The fragmentation pattern of the product also clearly evidenced successive losses of two carboxylate functions (Δ44 amu) as expected for the di-carboxylic muconic acid products.

Steady-state kinetic characterization. Kinetic studies were performed at either pH 7.5 in HEPES buffer or pH 8.5 in Bicine buffer at 0.1 M concentration. Steady-state experiments were performed either using an oxygen polarimeter (Hansatech) or by monitoring of the reaction-product formation using a UV/vis spectrometer. These methods produced comparable results. Extinction coefficients for the products were determined by incubating a known amount of substrate with enzyme and by assuming full conversion of substrate to product. Because the substrate and product UV spectra overlap, the product-formation rate was determined by using difference in extinction coefficients for the product and substrate at the same wavelength. For the organic K_m, the dissolved O_2 concentration was held nearly constant between 0.7 – 1.0 mM by performing the reaction in diluted O_2-bubbled buffer while the catechol concentration was varied. For the O_2 K_m determination, catechol was maintained under saturating conditions, typically 2 mM, while the O_2 concentration was varied. K_m values were determined by fitting the plots of initial velocity vs. catechol or O_2 concentration using the Michaelis-Menten equation. Due to the high K_m for O_2 for many of the substrates (Table 3), we were unable to attain truly saturating conditions (5-fold over K_m) in some cases because of limited O_2 solubility and technical limitations of the oxygen polarimeter. This meant we were not able to determine the initial velocity at some of the higher O_2 concentrations that would be on the asymptotic portion of the hyperbolic plot. The k_{cat} values could still be extracted from the fit, although the absence of data at higher concentrations resulted in significantly higher errors for the derived K_m values.

Crystal structure determination of E-CAT complex. 3,4-PCD was crystallized at pH 7.5 under anaerobic conditions as described in the Chapter 1 Materials and methods. Single crystals were soaked in mother liquor supplemented with 25% glycerol as cryoprotectant and 50 mM catechol. The crystal was soaked for roughly ten minutes before the crystal was flash-frozen directly in liquid N_2. The structure was solved by direct Fourier phasing with a model of resting WT 3,4-PCD followed by rigid body and restrained refinement. The final model was generated using Coot and refined in Refmac5[293, 294].

Transient kinetic characterization. An Applied Photophysics SX.18MV stopped-flow spectrophotometer was used to study the transient reactions. The enzyme-substrate complexes were pre-formed in an anaerobic glove bag (Coy) and reacted with O_2-containing buffer at 4.5 °C. Since the high K_m for the organic substrates precluded true single-turnover conditions, the experiments were performed under multiple-turnover conditions with excess catechol. Typical experiments contained 100-120 μM enzyme and 2 mM catechol in one syringe and anaerobic or oxygenated buffer in the other. Due to the occurrence of a very slow step near the end of the reaction cycle, multi-exponential fitting allowed two reciprocal relaxation times (RRTs) to be directly determined at early reaction time-points before the enzyme entered steady-state turnover. The RRT value for a third kinetic phase was determined by fitting the decay of the enzyme from steady-state turnover back to ES at long time-points when O_2 was exhausted. This decay was well-fit by a single exponential function. Mathematical fitting of the reaction traces was performed using the Applied Photophysics Pro-Data Viewer software and is described in detail in the Chapter 1 Materials and methods.

Results

Steady state characterization of the reaction of 3,4-PCD with the alternative substrates and comparison to PCA. Each of the alternative catecholic substrates tested are turned over to product at a significantly lower rate than the native substrate (>50-fold). All of the determined steady-state parameters are shown in Table 3. In order of highest to lowest k_{cat}, the following trend was observed at 22 °C: CAT (1.5 s^{-1}) > 4FC (1.2 s^{-1}) > 4CC (0.35 s^{-1}) > 4MC (0.05 s^{-1}). K_m values for the organic substrate are increased relative to PCA (29 μM). This is most noticeable with CAT, 4FC and, 4MC, where the values are increased to 100 μM, 190 μM and 240 μM, respectively. The K_m for 4CC is only two-fold higher than that of PCA, 63 μM. K_m values for O_2 for each of the alternative substrates are also considerably higher than for PCA (56 μM): CAT (130 μM), 4FC (270 μM), 4CC (440 μM), and 4MC (170 μM). Each of these catechols is converted into the expected intradiol ring-cleaved product (see Materials and methods). The decrease in both k_{cat}/K_m ratios relative to that of PCA is indicative of a greatly reduced catechol- and O_2-capture efficiency by the resting enzyme and enzyme-catechol complex, respectively[302].

Optical spectra of the anaerobic enzyme-catechol substrate complexes. The enzyme-substrate (ES) complexes were prepared under anaerobic conditions with 100 μM enzyme and excess (2 mM) catechols. The spectra for these complexes at pH 7.5 are shown in Figure 19 along with that of the resting enzyme (black trace) and the PCA complex (E-PCA, magenta). Each species exhibits the additional ligand-to-metal charge-transfer bands in the 500-750 nm range that are characteristic of the formation of a catechol-iron complex[185, 188, 205]. However, there are substantial differences in the λ_{max} and overall intensities for the complexes. The 4CC complex (E-4CC, green) exhibits the greatest intensity in the 575-750 nm range while being considerably less intense than the other complexes at lower wavelengths. The E-4CC λ_{max} is near 450 nm, but the peak is broader than that those of the other complexes. The 4FC (E-4FC, red) and CAT (E-CAT, blue) complexes are the most similar and differ only modestly at longer wavelengths whereas E-4FC absorbs more intensely. Both complexes absorb maximally around 425 nm but have the lower intensities at longer wavelengths than E-4CC. The spectrum with 4MC (E-4MC) is shown in orange and has a λ_{max} of 500 nm and lower overall intensity at lower wavelengths than E-4FC and E-CAT. Compared to E-PCA (magenta), all of the other ES complexes are considerably less

Table 3. Steady-state kinetic parameters for the reaction of 3,4-PCD with PCA, catechol, 4-methyl, 4-fluoro, and 4-chlorocatechol and O_2.

	k_{cat}* (s^{-1})	$K_{m,organic}$ (μM)	$K_{m,O2}$ (μM)	$k_{cat}/K_{m,organic}$ ($M^{-1}s^{-1}$)	$k_{cat}/K_{m,O2}$ ($M^{-1}s^{-1}$)
PCA†	81 ± 20	29§	56§	2.8×10^6	1.5×10^6
CAT†	1.5 ± 0.31	100 ± 10	130 ± 72	1.5×10^4	1.2×10^4
4MC‡	0.05 ± 0.002	190 ± 10	170 ± 73	2.6×10^2	2.9×10^2
4FC†	1.2 ± 0.09	240 ± 24	270 ± 110	5.0×10^3	4.4×10^3
4CC†	0.35 ± 0.03	63 ± 5.8	440 ± 160	5.6×10^3	8.0×10^2

Errors represent one standard deviation. All values normalized to the number of Fe^{3+}-containing active sites.

* due to the high K_m for O_2 for these substrates, the reported k_{cat} values and the corresponding error were extracted from the hyperbolic fits to the initial velocity vs. dissolved O_2 concentration plots

† values determined at 22 °C in 0.1 M Hepes buffer at pH 7.5.

‡ values determined at 22 °C in 0.1 M Bicine buffer at pH 8.5.

§ values are from reference[179].

intense at all wavelengths. These spectra may indicate that CAT, 4CC, and 4MC, like 4FC, bind to the metal in an altered manner relative to PCA (Fig. 15 shows the binding mode of 4FC).

Crystal structure of the E-CAT complex. Crystals of 3,4-PCD were soaked anaerobically with CAT substrate in order to generate the E-CAT complex for structural characterization. The structure of the complex was solved at 1.67 Å resolution with a model R_{work} and R_{free} of 16.5% and 18.9%, respectively. The data collection and modelling statistics are given in Table 4. The electron density at the active site evidenced clear binding of a catechol molecule directly to the iron cofactor (Fig. 20A). CAT binds to the metal in a rotated orientation and its position is superimposable with the 4FC substrate in the E-4FC crystal structure (Fig. 15). This binding mode amounts to a rotation of roughly 100° of the CAT aromatic ring into the equatorial plane relative to PCA in the E-PCA complex. The two hydroxyl/alkoxide groups of CAT coordinate to the metal *trans* to H460 and Y408 with Fe-O_{CAT} bond distances of 2.1 Å and 2.3 Å, respectively (Fig. 20B). Y447, which fully dissociates from the metal in E-PCA[181], is modeled at a 50/50

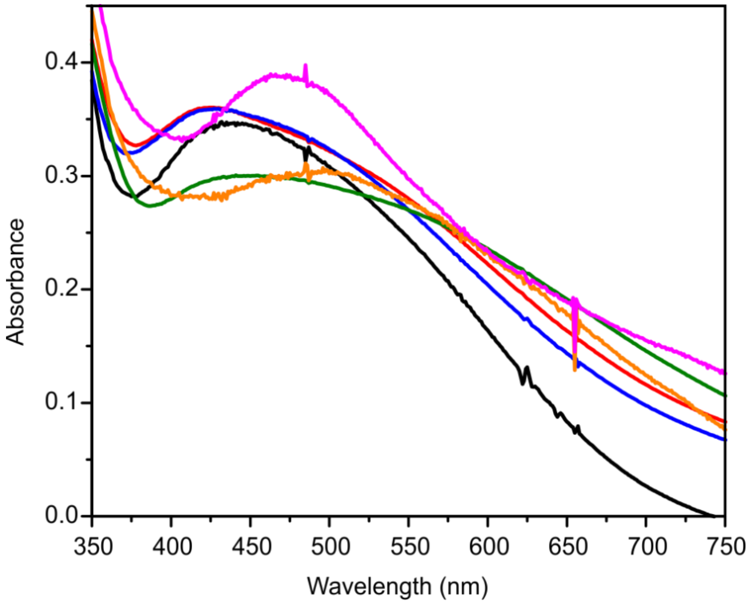

Fig. 19: Optical spectra of the anaerobic ES complexes. Samples contained 100 μM 3,4-PCD and 2 mM catechol substrate at pH 7.5. The figure shows the resting enzyme (black), E-PCA (magenta), E-CAT (blue), E-4FC (red), E-4MC (orange), and E-4CC (green). The spectra were obtained anaerobically in 0.1 M HEPES pH 7.5.

metal-bound and dissociated occupancy. The ability of CAT to cause dissociation of Y447 even though it does not physically displace the residue's side-chain was also observed in E-4FC and is consistent with the propensity of the metal center to maintain charge-balance during the catalytic cycle (see Chapter 1)[179, 194]. The iron is in a distorted octahedral or square pyramidal environment when Y447 is metal-bound or dissociated, respectively.

Transient kinetic characterization of the reaction of the enzyme-substrate complexes with O_2. In order to characterize the reactivity of the ES complexes with O_2, we used stopped-flow spectroscopy to monitor the reaction optically and extract the rate constants for the reaction. Due to the relatively high organic K_m for each of these substrates the transient kinetic analysis was performed under multi-turnover conditions containing saturating catechol concentrations. Under these conditions, the initial portion of the reaction time course of E-CAT, E-4MC, or E-4CC with O_2 could each be fit by a sum of two exponentials. Exponential fitting was appropriate in these cases because a very slow step occurs near the end of the reaction cycle thereby kinetically isolating the initial steps of the reaction from the steady state turnover. Longer time-courses

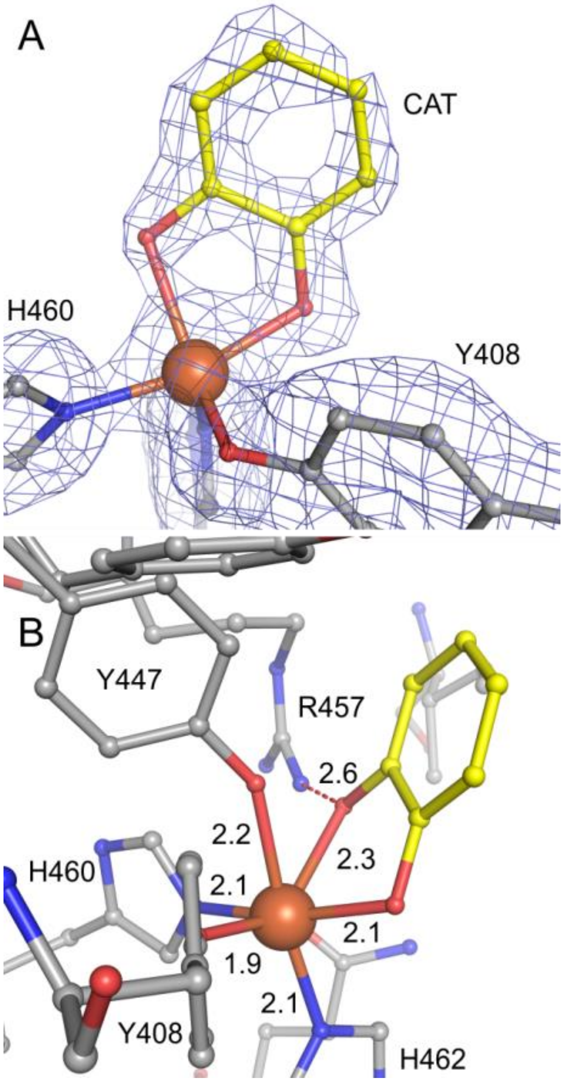

Fig. 20: Crystal structure of the E-CAT complex at pH 7.5. (A) Electron density for one of the active sites in the asymmetric unit of the complex. The mesh represents the 2|Fo|-|Fc| map contoured at 1σ. (B) The interactions of CAT with the iron and nearby residues with bond distances given in Å. In both panels, CAT is shown in yellow (carbon) and protein residues in gray (carbon). Oxygen atoms are red and nitrogen atoms blue. Iron is shown as a brown sphere.

evidenced the existence of a third process that logically corresponds to the decay of a steady-state intermediate back to ES. An example of such a long time-course is shown for a 4MC reaction in Figure 21. Overall, this kinetic behavior suggested a kinetic model with at least three steps and two intermediates where decay of the second intermediate was rate-limiting under

Table 4. E-CAT X-ray diffraction data collection and model statistics.

	3,4-PCD catechol complex
PDB ID	3T67
Data collection	
Space group	I222
Cell dimensions	
a, b, c (Å)	128.01, 140.78, 168.19
α, β, γ (°)	90, 90, 90
Wavelength (Å)	0.979
Resolution (Å)	50-1.67
Total/Unique reflections	163821/ 28740
R_{merge} (%) *,†	5.5 (57.5)
Completeness (%) *	99.3 (99.0)
Redundancy*	5.7 (5.7)
Model Refinement	
Resolution (Å)	30.68-1.67
R_{work}/R_{free} (%)‡	16.5/18.9
Average B, all atoms (Å^2)	21
ESU (Å) §	0.053
RMSDs	
Bond lengths (Å)	0.013
Bond angles (°)	1.42
Ramachandran analysis	
Allowed (%)	99.9
Outlier (%)	0.1

All data collected on synchrotron beamline APS SBC-CAT 19-BM-D

* Highest resolution shell is shown in parentheses

† $R_{sym} = \Sigma_{hkl}\Sigma_i |I_{hkl,i} - <I>_{hkl}|/\Sigma_{hkl}\Sigma_i|I_{hkl,i}|$, where I_{hkl} is the intensity of a reflection and $<I>_{hkl}$ is the average of all observations of the reflection

‡ R_{free}, R-factor calculated from 5% of the data excluded from refinement

§ Estimated overall coordinate error (ESU) based on maximum likelihood

RMSD, root mean square deviation from ideal geometry

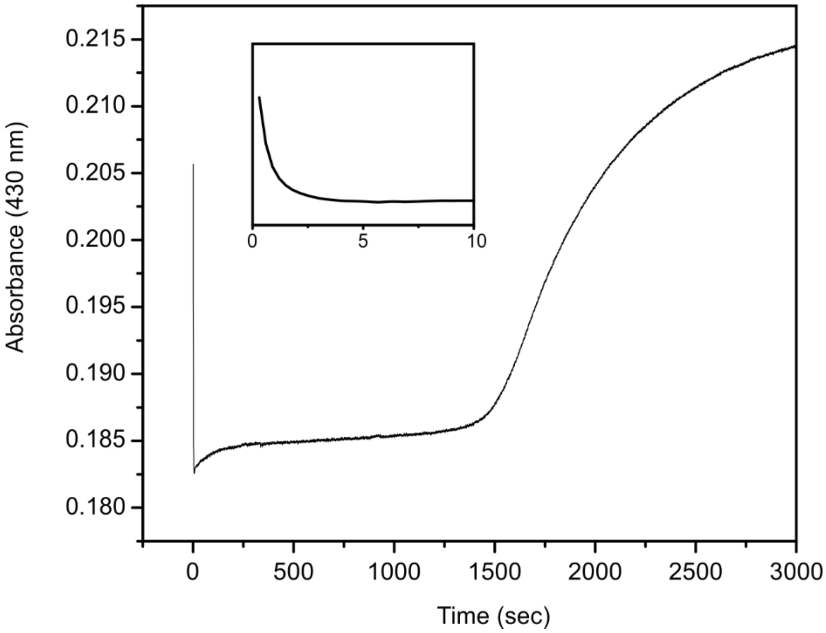

Fig. 21: Single-wavelength trace (430 nm) of the reaction of E-4MC with 950 μM O_2 as monitored using stopped-flow spectroscopy. The inset shows the same data but zoomed on the first 10 seconds of the reaction. The reaction was performed at 4.5 °C in 0.1 M Bicine buffer at pH 8.5.

steady-state conditions. This same behavior was observed for the reaction of E-4FC with O_2 and is discussed in Chapter 1.

The pre-steady state region of the reaction time-courses of E-CAT, E-4MC or E-4CC with O_2 was fit by a sum of two exponentials in order to extract the RRTs (Fig. 22). This same analysis was performed for 4FC (Fig. 14B and C) and discussed in Chapter 1. The sum and product plots of the RRTs and their dependence on the dissolved O_2 concentration are shown for E-CAT in Figure 22A, E-4MC in Figure 22B and E-4CC in Figure 22C. Analysis of the slope and intercepts of these plots yielded the rate constants for the steps which are shown in Table 5 (see the Materials and methods section of Chapter 1 for the details of the methodology). This analysis indicated that the reaction of all three ES complexes with O_2 is consistent with a model with three steps where the first step is reversible whereas the second and third are irreversible (k_{-2} and k_{-3} ~ 0):

$$\mathrm{ES} + \mathrm{O_2} \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} \mathrm{Int1} \xrightarrow{k_2} \mathrm{Int2} \xrightarrow{k_3} \mathrm{E} + \mathrm{P}$$

The same type of kinetic model was derived for 4FC as shown in Chapter 1. The magnitude of k_3 is comparable to the turnover number for the substrates, once the value is corrected for the temperature difference. This is consistent with decay of Int2 being rate-limiting under steady-state conditions. The relative magnitude of the bimolecular rate constant (k_1) for the reaction with O_2 was found to follow the trend E-4MC (6.3×10^3 M^{-1} s^{-1}) > E-CAT (8.9×10^2 M^{-1} s^{-1}) ~ E-4FC (8.7×10^2 M^{-1} s^{-1}) > E-4CC (7.8×10^2 M^{-1} s^{-1}). The equivalent rate constant for E-PCA is 5.0×10^5 M^{-1} s^{-1}, nearly two orders of magnitude greater than that for E-4MC, the most reactive alternative complex. In order to determine whether the observed bimolecular rate constant correlated with the electronic nature of the substituent, we plotted log k vs. the Hammett substituent coefficient for the *meta* (σ_m) and *para* (σ_p) position, as shown in Figure. 23. As evidenced by the plots, there is no clear correlation between the observed rate constant and the electron-withdrawing or donating characteristics of the substituent. Factors other than electronics appear to be determinant of O_2 reactivity.

Table 5. Transient kinetic parameters for the reaction of 3,4-PCD with PCA, catechol, 4-methyl, 4-fluoro, and 4-chlorocatechol and O_2.

	k_1 ($M^{-1}s^{-1}$)	k_{-1} (s^{-1})	k_2 (s^{-1})	k_3 (s^{-1})
PCA*	5.0×10^5		450	36
CAT†	8.9×10^2	0.68	0.51	0.38
4MC‡	6.3×10^3	0.45	1.6	0.01
4FC†	8.7×10^2	0.11	0.58	0.16
4CC†	7.8×10^2	0.13	0.06	0.09

All listed values were determined at 4.5 °C.
* values for PCA are from reference[184].
† values determined at 22 °C in HEPES buffer at pH 7.5.
‡ values determined at 25 °C in Bicine buffer at pH 8.5.

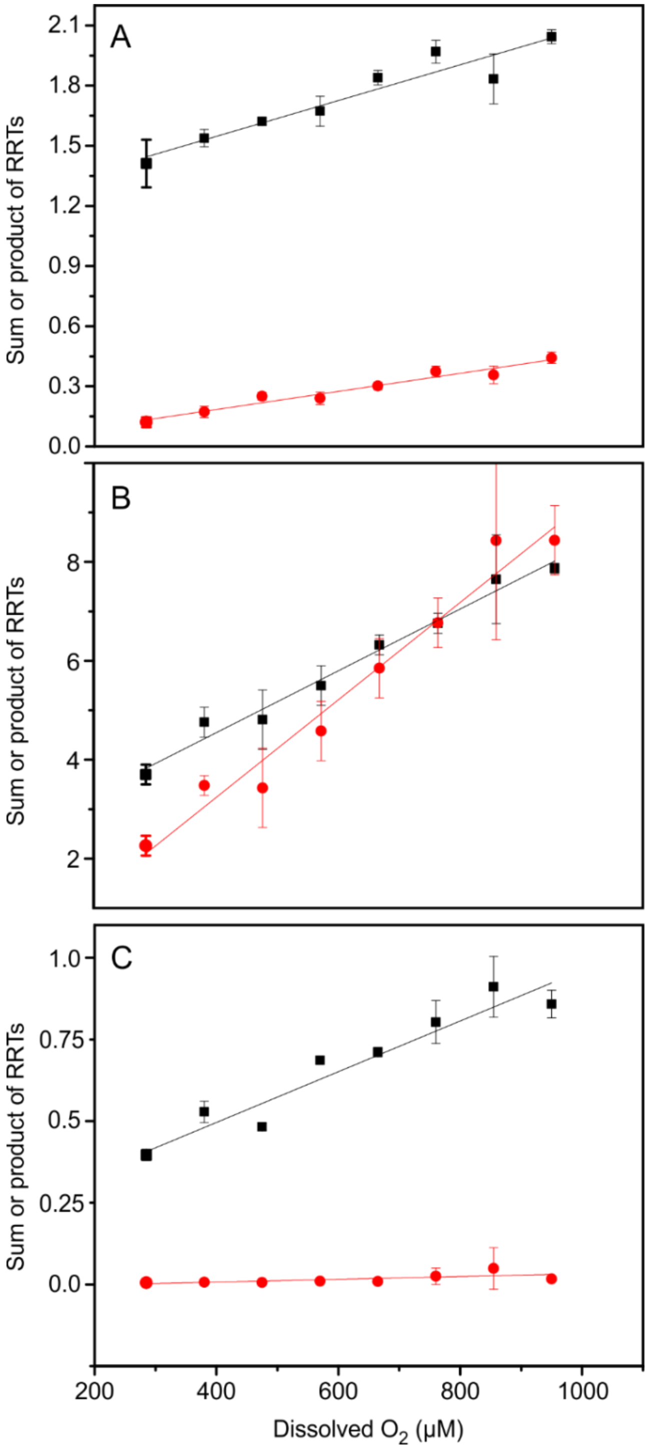

Fig. 22: Sum and product plots of the observed RRT plotted vs. dissolved O_2 concentration. In all panels, the sum plot is shown in black and the product plot in red. Error bars represent on standard deviation. (A) E-CAT vs. O_2. (B) E-4MC vs. O_2. (C) E-4CC vs O_2. Details of the experimental setup and conditions are discussed in the Materials and methods.

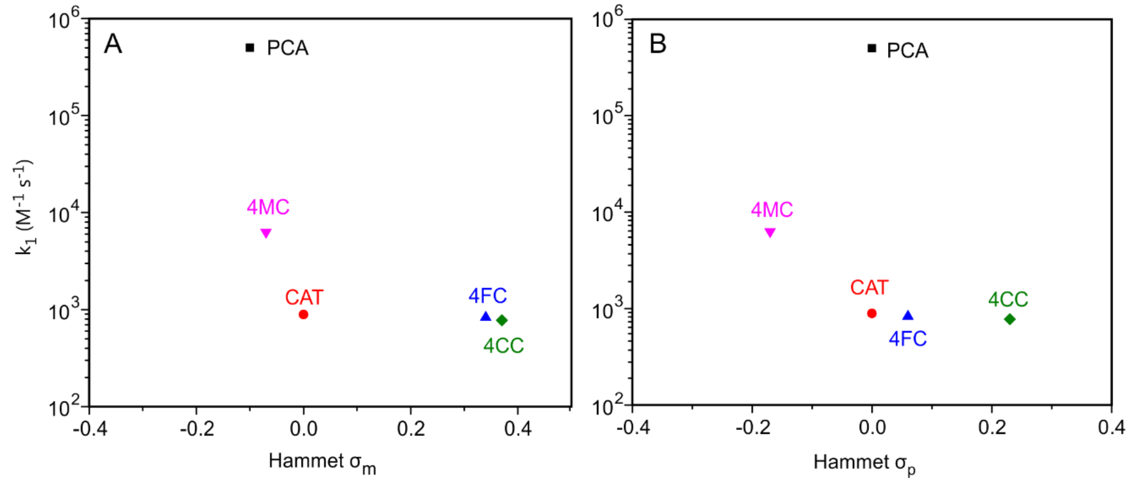

Fig. 23: Plots of the observed forward bimolecular rate constant (log scale) for the reaction with O_2 for the ES complexes vs. the Hammett coefficient for the aromatic substituent. (A) Hammett *meta* coefficient (σ_m). (B) Hammett *para* coefficient (σ_p).

Discussion

It is shown here that the reactions of 3,4-PCD with the alternative substrates makes possible the observation of the catalytic step(s) in which the reaction with O_2 takes place. The kinetic analysis allows us to investigate the effects of substrate electronics and steric bulk on the O_2 reactivity of the ES complexes. Here, we discuss the implications that these findings have on the mechanism of substrate selectivity in IDOs and the factors that determine O_2 reactivity in this enzyme family. We also propose a mechanism by which the enzyme modulates the binding mode of substrate to effect an increase in the substrate-capture efficiency and product-release rate.

Comparison of the reaction of the ES complexes with O_2. All of the reactions of the various ES complexes with O_2 exhibit a similar kinetic pathway. The concentration dependence of the reaction kinetics is consistent with O_2 binding in the first observable step. In this step, the reaction with O_2 is reversible and the forward bimolecular rate constant varies considerably between the substrates. The other steps are essentially irreversible for all substrates and k_3 is always the smallest and rate-limiting. In the O_2 reaction of E-PCA and the 1,2-CTD-substrate complex, two intermediates have also been observed, in which the decay of the second species is rate-limiting[180, 183, 184]. In 1,2-CTD, the first species could be a peroxo intermediate but is too short-lived to characterize in detail[183]. The second species is very likely a product complex (EP), suggesting that product release is rate-limiting with the substrates tested[205, 287]. Indeed, product-release from 3,4-PCD has been determined to be rate-limiting for nearly all substrates previously studied[179, 180, 184, 287]. The assignment of Int2 as a product complex in the reactions studied here is supported by the optical similarity of Int2 to that for the previously characterized EPs (see Figure 14A, orange trace, for the spectrum of Int2 in the 4FC reaction)[179, 183, 205]. This may indicate that product release (k_3) is also rate-limiting with CAT, 4MC, 4FC and 4CC.

It is interesting to note that while the E-4MC complex is the most reactive towards O_2, it has the smallest k_{cat} of all the alternative substrates. The increased O_2 reactivity may be due to the electron-donating character of the methyl substituent which is expected to increase the electron density on the ring and accelerate the reaction with O_2[25, 183, 190]. Although there is no clear linear trend in the Hammett plots for the non-PCA substrates, there is a correlation between the electron-donating/withdrawing character of the substituent and the observed bimolecular rate constant, namely 4MC > CAT ~ 4FC > 4CC. This order correlates with the substituents donating to withdrawing character: methyl > proton > fluoro > chloro (although the carboxylate substituent of PCA does not fit in this series). However, the k_{cat} shows no such trend. Decay of Int2 is rate-limiting for all substrates and the rate constant for this process for 4MC is an order of magnitude lower than that with the most electron-withdrawing substrate, 4CC. If this species corresponds to an enzyme-product complex, it is possible that the presence of a methyl group somehow sterically impedes product release to such an extent that it becomes slower than for all of the other substrates. The absence of a polarizable substituent may also slow egress if the enzyme uses electrostatic interactions to drive the product out of the active site. The methyl group is the largest substituent in terms of van der Waals radius of those tested in this series. However, the carboxylate of the native substrate PCA is larger than a methyl group, but product release is 9,000-times faster ($36\ s^{-1}$ vs. $0.004\ s^{-1}$ at 4° C). An explanation would be that the anionic carboxylate could form interactions with some of the conserved second-sphere residues in 3,4-PCDs. The candidate residues Y324 and R133 were previously identified to be conserved only in 3,4-PCDs[181, 206] but diverge in the 1,2-CTDs, which preferentially cleave unsubstituted and methyl-substituted catechols. Non-covalent interactions of the PCA carboxylate may orient the catecholic substrate (Fig. 24), oxygenated intermediates and product in a way that facilitates rapid chemical interconversion and release from the active site.

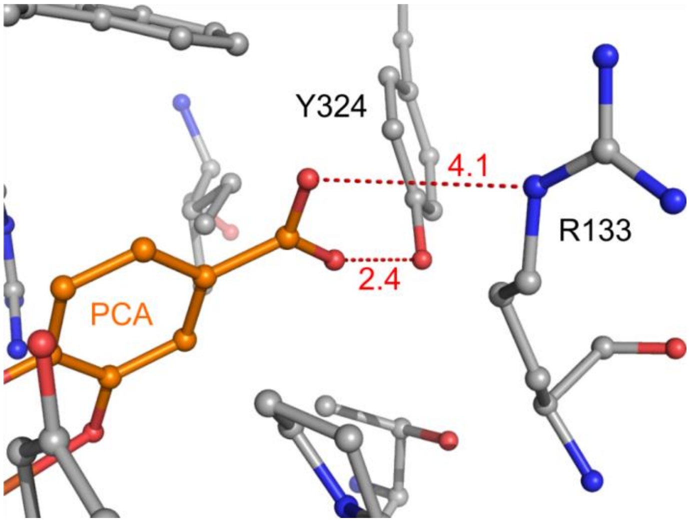

Fig. 24: Non-covalent interactions of the protein with the PCA carboxylate that are proposed to stabilize this substrate in the more reactive axial orientation. Bond distances are given in Å. PCA is in orange and the protein in gray. The image was generated from PDB structure 3PCA.

Possible explanation for the dramatically increased reaction rate of E-PCA towards O_2 vs. the other ES complexes. The bimolecular rate constant for E-PCA reaction with O_2 is orders of magnitude faster than those with all of the other substrates. Substrate electronics are simply unable to account for this dramatic difference as indicated by the Hammett plots of Fig. 23. The crystal structure of E-4FC (Fig. 15) shows that 4FC binds in a dramatically altered orientation relative to PCA: a 100° rotation into the equatorial plane of the metal coordination sphere. This same binding mode is observed in E-CAT (Fig. 20). This provides a possible explanation for the comparable O_2 reactivity of E-4FC and E-CAT and the optical similarity of these two ES complexes. 4MC and 4CC may also bind in an altered mode relative to PCA that may be similar to that of 4FC and CAT. In Chapter 1, the equatorial orientation of 4FC was hypothesized to be responsible for the reduced O_2 reactivity of the ES complex because the active site architecture is not optimized to drive the interconversion of equatorially-oriented substrates and oxygenated intermediates. Because of the similarity of the optical spectra of all the ES complexes (lower intensity and different shape than E-PCA) and the overall low O_2 reactivity, it is plausible that the same mechanism(s) is responsible for slowing the reaction with all of these alternative substrates. A more speculative interpretation of the data is that the substrates are bound to the metal as a mixture of equatorially- and axially-oriented modes and only the axial mode is reactive. The axial mode would need to be present at a very low occupancy, based on our inability to observe this mode in the crystals. The divergence of the observed bimolecular rate constants from a linear Hammett trend could then be representative of the different relative ratios of the axial/equatorial orientation of the alternative substrates.

As discussed in Chapter 1, the reaction of E-PCA with O_2 is rapid because the substrate is already bound in the reactive axial orientation and there is a clear site for O_2 to bind near the iron. Thus, the binding mode of PCA is apparently more important in determining the reaction rate with O_2 than the substrate electronics. Other IDOs which preferentially cleave un-substituted and methyl- or halo-substituted catechols stabilize these molecules in an axial metal-bound orientation in ES[197, 202]. The reactivity of these ES complexes towards O_2 is comparable to that of 3,4-PCD-PCA with O_2[183]. We propose that the preferential stabilization of this binding mode has two major effects on catalysis in IDOs: *(i)* it increases the O_2 capture efficiency ($k_{cat}/K_{m,O2}$) and the reactivity of ES towards O_2 and *(ii)* it may increase rate-limiting product release step, thereby increasing k_{cat}. Both effects are expected to greatly enhance the flux through this ring-cleavage pathway *in vivo*.

Conclusion

3,4-PCD retains intradiol-cleavage specificity even with the alternative methyl- and halo-substituted catechols. It is likely that the product-release step is rate-limiting for all substrates. The steps in the catalytic cycle where the reaction with O_2 occurs are slowed more than 100-fold with the alternative substrates relative to PCA. Based on the trends in O_2 reactivity and the previous observation of an alternative binding mode for the 4FC substrate, we conclude that the reactivity of the ES complex may be primarily controlled by the binding mode of the substrate to the iron. The evidence suggests that 3,4-PCD has evolved to selectively bind PCA in a more reactive orientation than other substrates, conferring a preference for the carboxylated catechol over the other substrates. It may do this by interacting with the carboxylate function of substrate using a set of second-sphere residues that are only conserved in 3,4-PCDs but not in relatives that preferentially cleave methyl or chloro-substituted substrates. Such interactions may also help orient the PCA oxygenated intermediates and cleaved product for rapid progression through the catalytic cycle.

Acknowledgements

We could like to thank Dr. Vincent Purpero for assistance in collecting the transient kinetic data and for solving the structure of the E-CAT complex.

CHAPTER 3

Structure of a dinuclear iron cluster-containing β-hydroxylase active in antibiotic biosynthesis

Note: Parts of this chapter are reprinted (adapted) with permission from Makris, T. M., Knoot, C. J., Wilmot, C. M., and Lipscomb, J. D. (2013) Structure of a dinuclear iron cluster-containing beta hydroxylase active in antibiotic biosynthesis, *Biochemistry*, 52 (38), 6662-6671. Copyright (2013) American Chemical Society.

Note: Parts of this chapter are reprinted (adapted) with permission from Knoot, C. J., Makris, T. M., and Lipscomb, J.D. Dinuclear iron cluster-containing oxygenase CmlA, Encyclopedia of Inorganic and Bioinorganic Chemistry, Online © 2011–2015 John Wiley & Sons, Ltd. This article is © 2015 John Wiley & Sons, Ltd. This article was published in the Encyclopedia of Inorganic and Bioinorganic Chemistry in 2015 by John Wiley & Sons, Ltd.

Abbreviations:

NRPS: non-ribosomal peptide synthetase
CmlA: β-hydroxylase of the chloramphenicol biosynthetic pathway
CmlP: NRPS of the chloramphenicol biosynthetic pathway
sMMO: soluble form of methane monooxygenase
MMOB: regulatory component B of the sMMO system
L-PAPA: L-*p*-aminophenylalanine
PPant: 4'-phosphopantetheine cofactor
PCP: peptidyl-carrier protein. This CmlP domain contains the serine attachment site for PPant and the amino-acid substrates that are subsequently covalently linked to PPant. PCP is alternatively termed the thiolation (T) domain.
CmlP$_{AT}$: truncated version of full-length CmlP lacking the C-terminal reductase domain

Summary

A family of dinuclear iron cluster-containing oxygenases was recently described that catalyze β-hydroxylation tailoring reactions in natural product biosynthesis by nonribosomal peptide synthetase (NRPS) systems. Here, the 2.17 Å X-ray crystal structure of the archetypal enzyme from the family, CmlA, is reported. CmlA catalyzes β-hydroxylation of L-*p*-aminophenylalanine during chloramphenicol biosynthesis. The fold of the N-terminal domain of CmlA is unlike any previously reported, but the C-terminal domain has the αββα-fold of the metallo-β-lactamase (MBL) superfamily. The diiron cluster bound in the C-terminal domain is coordinated by an acetate, three His, two Asp, one Glu and a bridging oxo moiety. One of the Asp ligands forms an unusual monodentate bridge. No other oxygen-activating diiron enzyme utilizes this ligation or the MBL protein fold. The N-terminal domain facilitates dimerization, but using computational docking and a sequence-based structural comparison to homologs, we hypothesize that it likely serves additional roles in NRPS recognition and the regulation of O_2 activation.

Introduction

Biosynthesis of many complex secondary metabolites by bacteria and fungi is effected by nonribosomal peptide synthetases (NRPS). These large, modular enzymes act as 'assembly lines' in which each module plays a role in extending or modifying the growing natural product[247]. Generally, a parent amino acid or adduct is covalently tethered to the 4'-phosphopantetheine (PPant) moiety of a modified serine residue on a peptidyl-carrier protein (PCP) domain of the first module of the NRPS by the action of a preceding adenylation (A) domain. The resulting S-acyl amino acid intermediates are either chemically modified by following NRPS domains or transferred to the S-acyl-linked amino acid adduct on the PCP domain of the next module by the action of a condensation (C) (or similar) domain, forming a new peptide bond. When the synthesis is complete, the natural product is released by the action of a thioesterase (TE) or a reductase (R) domain.

In some cases, the chemical modifications of the developing natural product are catalyzed by "tailoring" enzymes that are not modules of the NRPS but are rather encoded elsewhere, often in the same genetic operon[247, 257]. A common modification for tethered amino acids is β-hydroxylation, which is crucial for the function of many natural products including the vancomycin and coumarin antibiotic families as well as the antitumor compounds of the bleomycin family[239-241, 261, 303, 304]. The newly introduced hydroxyl function can serve as a site for glycosylation or act as a nucleophile during release of the natural product from the NRPS to form large macrocyclic molecules. Study of the structural and regulatory mechanisms that control the specificity and spatiotemporal activity of these accessory enzymes is necessary in order to understand the biosynthetic origins of complex natural products. Learning how to harness these capabilities may enable the production of new antibiotics and pharmaceuticals[274, 275, 305].

Sequencing of the chloramphenicol biosynthetic operon[264] led to the discovery of a tailoring enzyme (CmlA) that catalyzes β-hydroxylation of the precursor molecule L-*p*-aminophenylalanine (L-PAPA) to form L-*p*-aminophenylserine[58]. This is a key step in chloramphenicol synthesis as illustrated in Figure 25[306]. The amino acid sequence of CmlA suggests that it has two domains. The N-terminal domain has no known homologs, but the C-terminal domain contains the signature sequence of the αββα-type metallo-β-lactamase (MBL) superfamily[58, 307]. The vast majority of MBLs catalyze hydrolysis reactions utilizing a dizinc cluster coordinated by the signature His-X-His-X-Asp-His scaffold, although MBLs with both

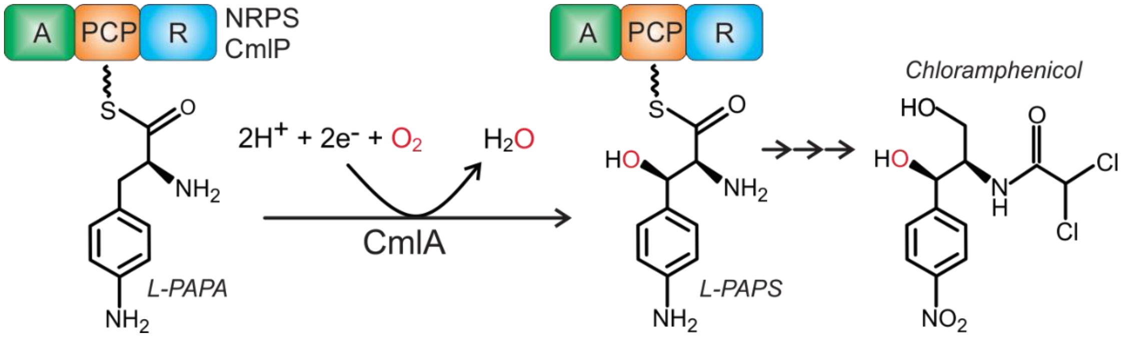

Fig. 25: Steps in the biosynthesis of chloramphenicol. The β-hydroxyl shown in red is derived from O_2 and is incorporated through the catalytic activity of CmlA.

mono- and di-metal clusters containing other metals are known[265]. In contrast, spectroscopic characterization of CmlA showed that it contains a diiron cluster[58, 268].

Prior to the characterization of CmlA, no MBL-like enzymes catalyzing O_2 activation and hydroxylation reactions were known to exist, although a branch of the MBL family that utilizes a diiron cluster to reduce O_2 or NO in obligate anaerobes has been described[265, 266]. Typically, β-hydroxylation reactions in NRPS pathways are catalyzed by α-ketoglutarate-dependent oxygenases or cytochrome P450s[245, 261, 303]. CmlA represents a third class, and it is the only known diiron enzyme to catalyze β-hydroxylation of an amino acid-like substrate. Oxygen activation and hydroxylation reactions involving diiron clusters are generally carried out by members of the bacterial multicomponent monooxygenase (BMM) family[308], which includes enzymes such as the soluble methane monooxygenase (sMMO)[109]. In contrast to CmlA, BMMs coordinate the diiron cluster in a 4-helix bundle. The typical cluster coordination consists of four carboxylate (Asp/Glu) and two histidine ligands[309]. Numerous kinetic, spectroscopic, and structural studies have shown that both the diiron cluster and the protein scaffold play important roles in orchestrating and regulating the oxygenation chemistry[109, 110, 310].

Genome database searching based on the CmlA gene sequence has now shown that there are many enzymes serving similar roles in natural product biosynthesis that are homologous to both the N- and C-terminal regions of CmlA[58]. To date, no enzyme from this group of homologous enzymes has been structurally characterized, and no role for the N-terminal domain has been forthcoming.

Here, we present the crystal structure of CmlA at 2.17 Å resolution. The structure confirms the predicted MBL C-terminal fold and reveals the structure of the N-terminal domain.

The detailed structure of the dinuclear iron cluster differs from those of previously characterized oxidoreductases from the MBL family, as well as those from the O_2-activating BMM family. The overall organization of the enzyme suggests different roles for the N- and C-terminal regions as well as the basis for regulation of O_2 activation through interaction with the NRPS from the chloramphenicol biosynthetic pathway, CmlP.

Materials and methods

Crystallization of CmlA. CmlA and CmlP were expressed, purified and quantified as previously described[58]. Crystals of resting CmlA were acquired by the hanging drop vapor diffusion method at 277 K by addition of 3 μl of a 15 mg/ml solution of CmlA in 50 mM HEPES pH 7.5 to an equal volume of mother liquor containing 100 mM HEPES pH 7.5, 10-15 % polyethylene glycol (F.W. 20,000), 100 mM potassium acetate, and 10 % glycerol. Square-bipyramidal CmlA crystals typically grew to 0.2 – 0.6 mm over the course of several days. Selenomethionine-labeled CmlA was generated by feedback inhibition during expression[311]. Crystals were soaked in mother liquor supplemented with 25 % glycerol prior to looping and flash-freezing in liquid nitrogen. For the amino acid co-crystallization, crystallization and freezing conditions were identical to the non-co-crystallized samples except the mother liquor and cryoprotectant contained L-phenylalanine, L-tyrosine or L-PAPA at final concentrations of 10-20 mM.

Data collection and modeling. Data were collected at the Structural Biology Center at the Advanced Photon Source, Argonne National Labs on beamline 19-ID at 100 K. Diffraction data were indexed, integrated and scaled using the HKL2000 software package[312] and phased using single-wavelength anomalous dispersion of SeMet-labeled enzyme with Phenix[313]. The initial model was prepared with Phenix followed by manual rebuilding. The preliminary model was then used in direct Fourier phasing of a higher-resolution native data set using Refmac5[294] from the CCP4 suite[314]. The final model was prepared using iterative cycles of modeling using Coot[293] and refinement with Refmac5. Iron-ligand bond distances were not restrained during any of the refinement cycles. Model geometries were evaluated using SFCHECK and PROCHECK[314]. X-ray data processing and refinement statistics are summarized in Table 6. All structure figures were produced using PyMOL Molecular Graphics System, Version 1.5, Schrödinger, LLC.

Computational docking of the CmlP PCP domain and PPant amino acid substrate. Homology modeling of the CmlP PCP domain was performed using Swiss-Model[315] and docked to CmlA using the HADDOCK online server[316]. In order to guide the docking of the PCP domain, we defined several residues to be part of the interface based on studies of our CmlA model and prior sequence information. For CmlA we input W438, D496, and H501. These surface residues were chosen based on their proximity to the active site channel. For the CmlP T domain homology model, we input only the serine identified as the putative PPant attachment from sequence comparisons to other structurally characterized PCP domains, with the rationale that this residue must be poised above the channel upon binding of the NRPS to CmlA.

Docking of L-PAPA-PPant to CmlA was performed using the Sybyl-X 2.0 Surflex-Dock Suite [Sybyl-X 2.0 Tripos International, 1699 South Hanley Rd., St. Louis, Missouri, 63144, USA]. The final CmlA model was prepared by deleting all water molecules and non-protein ligands, except the Fe atoms, followed by addition of hydrogen to residues. The structure was then minimized using the Amber7 FF99 package in the docking suite, but Fe atoms and ligand positions were unaltered. The substrate was modeled manually in Sybyl-X and atomic partial charges for both the protein and substrate were automatically calculated and assigned, with the exception of Fe atoms which were manually assigned as +3 and +2 for Fe1 and Fe2, respectively. The parameters for defining the active site were as follows: threshold 0.65, bloat 0. The active site was automatically identified by the program. The protein was not allowed to flex during docking.

Approaches to determine whether CmlA is partially photoreduced during data collection. In order to determine the effects of radiation dosage on the resulting models, we compared models of the active site that were produced from data collected on comparable crystals with the synchrotron beam attenuation set to 5, 50 and 100, corresponding roughly to a total radiation dose of 33 MGy, 3.3 MGy and 1.7 MGy, respectively. Dosage was calculated for the exposed region of the crystal using RADDOSE-3D[317]. The resulting models had identical cluster geometries and bond distances within coordinate error, and there were no significant shifts in the orientation of side-chains or the position of the oxo atom following refinement. However, crystals exposed at high X-ray dosages showed increases up to 10 Å^2 in the B factor in the resulting models, making direct comparisons between these data and those from crystals exposed at lower X-ray doses difficult. The final deposited structure was collected with the beam attenuation set to 50. Data processed in slices and merged also showed no significant changes in the bond distances or coordinates of the ligands. Additionally, soaking of crystals in cryoprotectant supplemented with 10-20 mM sodium ascorbate as an electron scavenger prior to freezing also showed no significant changes in the iron-oxo bond distances or iron ligand orientations. Microspectrophotometer UV-vis absorbance data on single CmlA crystals collected pre- and post-exposure were inconclusive due to the low extinction coefficient (3.5 mM^{-1} cm^{-1} at 340 nm per diiron cluster) of the charge-transfer chromophore and detection limitations of the instrument. CmlA crystals are square-bipyramidal in shape and samples used for the microspectrophotometer experiments were typically 0.3-0.5 mm long (all dimensions). In all X-ray diffraction datasets, the Fe2-oxo distance consistently appeared longer than that of Fe1-oxo. Based on these experiments, we could not definitively identify the redox state of the cluster.

Cloning and mutagenesis. The expression construct for the E430A mutant of CmlA was produced using an Agilent Quickchange II Site Directed Mutagenesis kit. The following primer pair was used to introduce the mutation into a pET28a expression plasmid for wild-type (WT) CmlA:

CGTCTTCATCGGCATGGAGCTCATCGGCGCGG (forward*)*
CCGCGCCGATGAGCTCCATGCCGATGAAGACG (reverse)

Mutated plasmid was transformed into chemically competent XL1-Blue *E. coli* and plated onto media containing 50 µg/ml kanamycin. Individual colonies were screened *via* colony PCR and sequenced to confirm the presence of the mutation. E430A enzyme was expressed and purified as described above. Mass spectrometry of CmlA was done with a QSTAR XL Quadrupole TOF MS from AB Sciex at the University of Minnesota Center for Mass Spectrometry and Proteomics and used to confirm the presence of the desired mutation in purified enzyme. The deconvoluted mass spectrum showed a major peak at 62,204 Da and 62,146 Da for the WT and E430A, respectively, the difference corresponding to the expected loss of a carboxylate and CH_2 group in the mutant.

Stopped flow kinetics. All stopped-flow experiments were performed with an Applied Photophysics model SX.18MV stopped flow device at 4 °C. Stopped-flow experiments and CmlP amino acid loading were performed as previously described[58], with the exception that reduced CmlA was pre-mixed with L-PAPA loaded CmlP_{AT} in an anaerobic chamber and rapidly mixed with O_2 saturated buffer on the instrument. Excess CmlA reducing reagents and CmlP loading reagents were removed prior to the experiment using a PD10 desalting column.

Results and discussion

Overall structure of CmlA. The structure of CmlA was determined by single-wavelength anomalous dispersion of selenomethionine-enriched enzyme. The enzyme crystallized in spacegroup $P4_3 2_1 2$ with a single monomer in the asymmetric unit. The CmlA model was refined to 2.17 Å resolution with a R_{work} of 20.3% and R_{free} of 23.0% with 99.6% of residues in the allowed regions of the Ramachandran plot (Table 6). Overall, 522 residues and 151 solvent molecules were modeled into the electron density map. The polypeptide can be roughly divided into two structural domains of equal length. As expected from the primary structure, the N-terminal half (residues 1-235, Fig. 26, gray surface) has a previously unclassified architecture. The second half (residues 249-532), however, belongs to the metallo β-lactamase (MBL) superfamily (Fig. 26, blue surface)[307]. The two domains are connected by a linker region (residues 236-248) that is unobserved in the electron density. Overall, the secondary structure of the enzyme is 34% α-helical and 21% β-sheet and contains a total of 20 α-helices (α1-20) and 23 β-strands (β1-23) (Fig. 26). The active site is housed entirely within the MBL catalytic domain. Members of the MBL family have a distinct αββα-motif which is also observed in CmlA. Prior to the discovery of CmlA, nearly all characterized diiron hydroxylases bound the active site cluster within a 4-helix bundle motif.

The N-terminal domain of CmlA is relatively flat and roughly triangular in shape and contains a mixture of α-helices and β-sheets (Fig. 27, left). One face of the triangle packs against the MBL-like catalytic domain. Interactions between secondary structural elements is limited to packing of several pairs of α-helices and three sets of antiparallel β-sheets each composed of 2-4 individual strands. A long, kinked helical region (residues 68-107, α5) runs the nearly the length of one of the sides of the triangle. One of the vertices of the domain contains a winding three-turn motif (residues 108-142) which is stabilized by packing of α5 and α6 and a pair of antiparallel β-strands (β4, β5). The surface residues of the vertex engage in electrostatic and hydrophobic interactions with a 2-fold symmetry-related CmlA protomer in the unit cell. The interface area buried between the protomers is 2700 $Å^2$ – 13% of the total accessible monomer surface area - and probably corresponds to the interface of the biological dimeric assembly. This feature is here termed the dimerization arm, as it mediates most of the inter-residue contacts with the second CmlA protomer (Fig. 28 and Fig. 27, left). The dimeric quaternary structure of CmlA was confirmed using native-PAGE gel chromatography. The second vertex of the N-terminal domain

Table 6. CmlA X-ray diffraction data collection and refinement statistics.

	Native	Se
Data collection		
Space group	$P4_3 2_1 2$	$P4_3 2_1 2$
Cell dimensions		
a, b, c (Å)	153.2, 153.2, 93.4	153.3, 153.3, 92.7
α, β, γ (°)	90, 90, 90	90, 90, 90
Wavelength (Å)	0.97857	0.97926
Resolution (Å)*	38.7-2.17 (2.21-2.17)	20.9-2.60 (2.64-2.60)
Unique reflections	56972	35553
R_{merge} (%) *,†	6.3 (43.5)	10.8 (45.0)
I/σI*	28 (3.1)	27 (4.5)
Completeness (%) *	96.4 (81.6)	99.9 (100)
Redundancy*	7.3 (7.1)	7.2 (7.4)
Overall figure of merit		0.405
Cruickshank's Diffraction Precision Index (DPI)	0.154	
Model refinement		
Resolution, Å	38.7-2.17	
Number of reflections	54097	
R_{work}/R_{free} (%)‡	20.3/23.0	
Averaged B factor ($Å^2$)	35.1	
RMSDs		
Bond lengths (Å)	0.0151	
Bond angles (°)	1.78	
Ramachandran analysis		
Favored regions (%)	95.6	
Allowed regions (%)	4.0	
Disallowed regions (%)	0.4	

All data collected on synchrotron beamline APS SBC-CAT 19ID-D

* Highest resolution shell is shown in parentheses

† $R_{sym} = \Sigma_{hkl}\Sigma_i |I_{hkl,i} - \langle I\rangle_{hkl}|/\Sigma_{hkl}\Sigma_i|I_{hkl,i}|$, where I_{hkl} is the intensity of a reflection and $\langle I\rangle_{hkl}$ is the average of all observations of the reflection

‡ R_{free}, R-factor calculated from 5% of the data excluded from refinement.

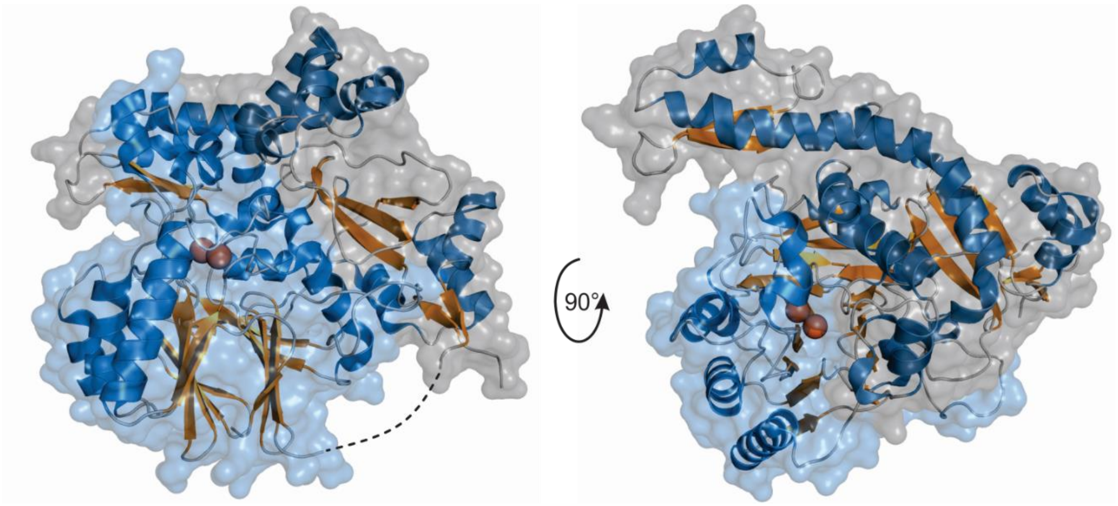

Fig. 26: Crystal structure of the CmlA monomer shown from two views. The surface of the N-terminal domain is shown in gray and that of the C-terminal metallo-β-lactamase-like domain in blue. The dashed line represents a stretch of disordered residues that connects the two domains for which no electron density was observed. α-helices are colored blue and β-strands orange.

contains an L-shaped α-helical extension (Fig. 27, α2- α4). This feature, along with a protrusion from the catalytic domain of CmlA, forms a groove which runs along the surface of CmlA, see below. The channel to the active site is positioned at the base of this groove. The final vertex of the domain contains the N-terminus of the enzyme and is stabilized by a pair of packed helices (α10, α11) and two antiparallel β-strands (β1, β8).

The C-terminal MBL-like catalytic domain houses the diiron active site. The fold is characterized by a central β-sandwich motif surrounded by two α-helical faces. One of these faces packs against the N-terminal domain of the same protomer and the other against the dimerization arm of the symmetry-related protomer. The core is composed of a β-sandwich with each half containing seven β-strands, each in a mixed antiparallel-parallel sheet (Fig. 26 and 27, right). The α-helices which surround the core are in the loop regions between these sheets, most often in αβα-motifs. The loop between β12 and β13 contains the conserved MBL metal coordination motif His-X-His-X-Asp-His which sits above the strands of the β-sandwich. This motif contributes four ligands to the diiron cluster; and the other first-sphere ligands are furnished by loops extending between the other β-strands. A final notable feature of this domain is the presence of an extended, partially helical loop between β19 and β20 which contains α16-α18. This protrusion points away from the core of the domain and, along with the L-shaped arm from the N-terminal domain, forms

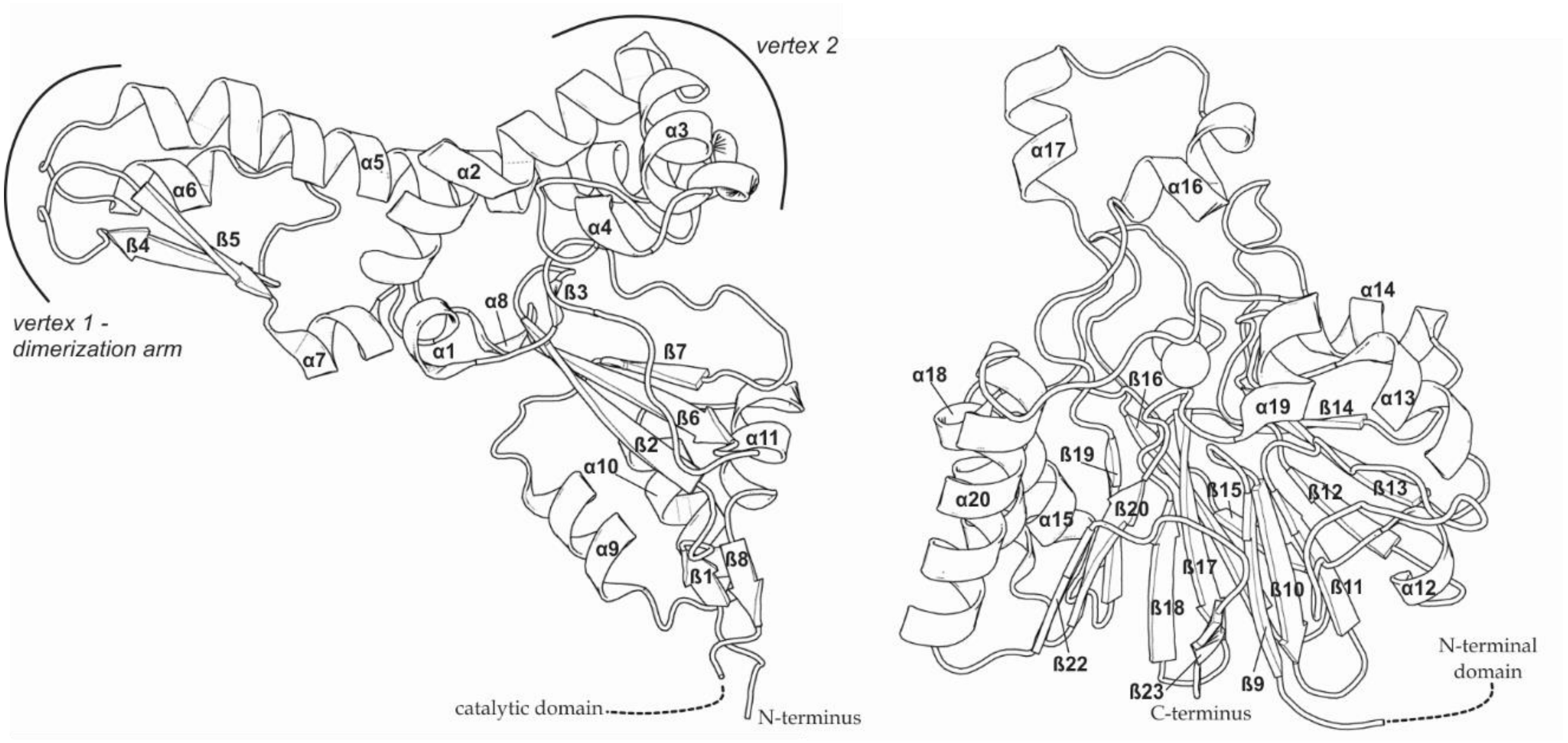

Fig. 27: Schematic depicting the secondary structure of the two domains of CmlA. Notable tertiary structure elements of the N-terminal domain (left) are also labeled and are discussed in the text. The C-terminal catalytic domain is shown on the right.

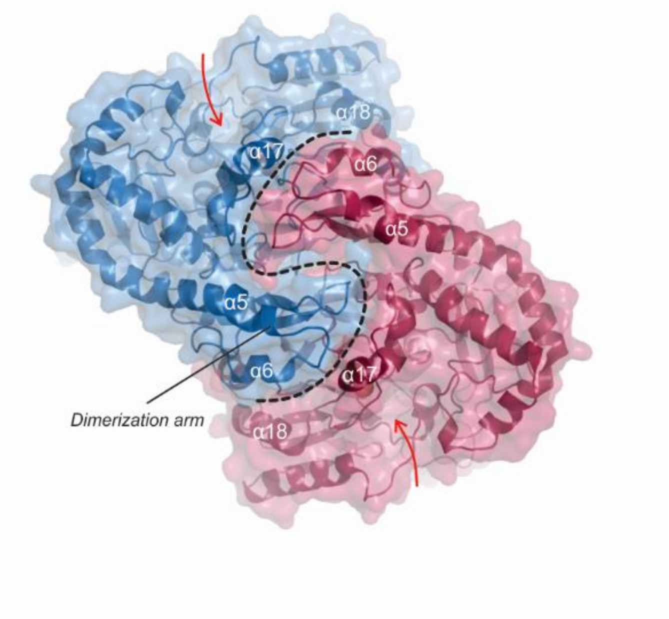

Fig. 28: The dimer interface of CmlA. One monomer is shown in blue and the other in maroon. Notable α-helices are labeled and the shape of the interface is demarcated by the dashed black line. The location of the channel to the active site in each protomer is indicated by the red arrows.

the other side of the groove that runs along the surface of CmlA. It is stabilized in this extended orientation by packing against the dimerization arm of the second CmlA monomer (Fig. 28).

Structure of the CmlA diiron cluster. The active site of CmlA is situated above the β-sheets of the αββα-core of the catalytic domain. Structural alignments of this domain show that the active sites of other members of the MBL-superfamily are found in similar spatial positions relative to the rest of the fold. The iron atoms are ligated in part by four ligands derived from a loop containing the conserved metal binding motif, which extends upward between the fourth β-strand and second α-helix of the catalytic domain. The only apparent access to the active site is from the channel above, although some solvent-occupied caverns can be seen extending from the substrate binding pocket.

The two iron atoms in the enzyme active site are in distorted octahedral environments with six ligands each (Fig. 29A). Herein, iron atoms are referred to as Fe1 and Fe2 for Fe600 and Fe601 in the model, respectively. Iron atom occupancies refined to 1.0 for Fe1 and 0.75 for Fe2 based on an analysis of the model B-factors after refinement. The iron-iron distance was

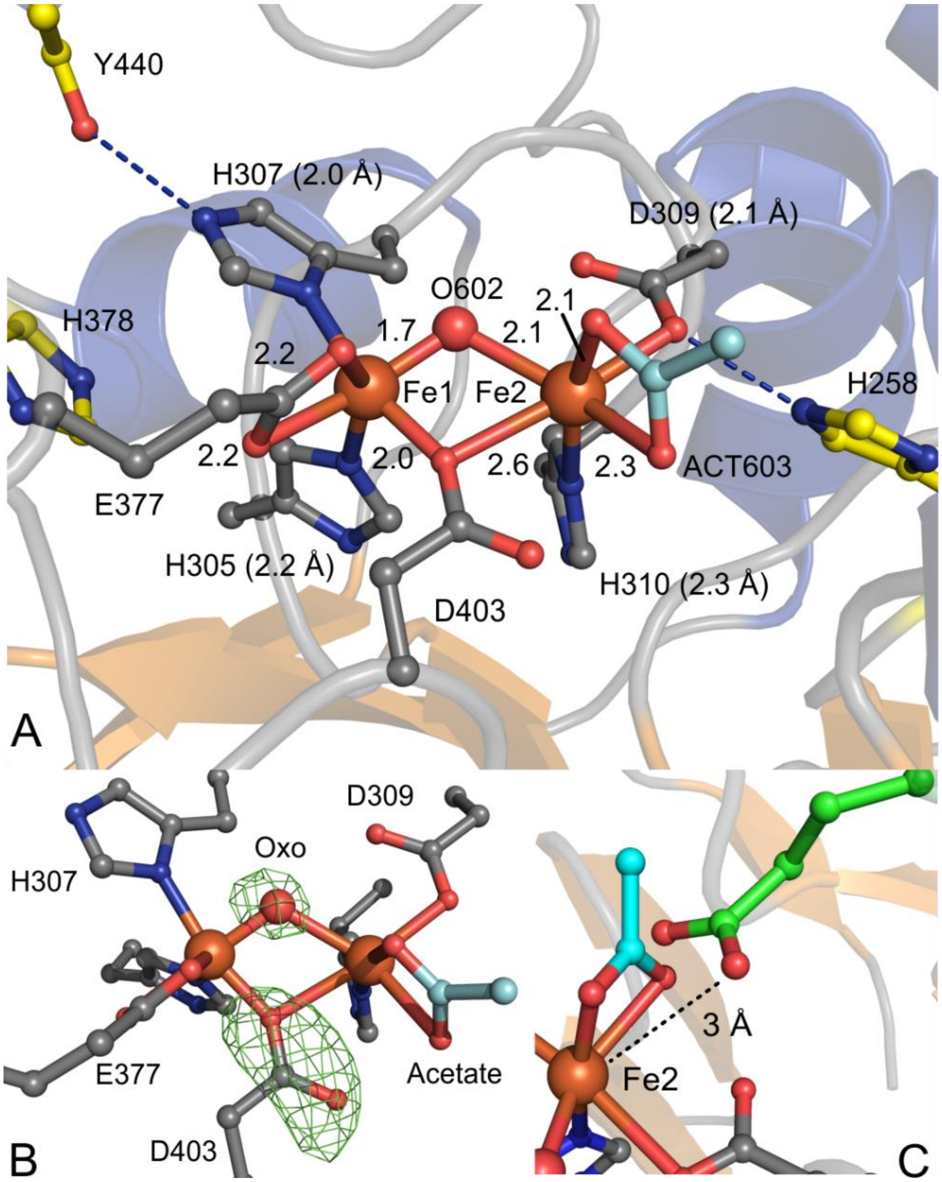

Fig. 29: Structure of the CmlA active site and nearby residues. (A) Cluster geometry and bond distances in Å, with iron ligands in grey, second-sphere conserved residues in yellow and the bridging oxo atom as a red sphere. Blue dashes indicate hydrogen bonds. The exogenous acetate ligand (Act) is shown in cyan. When not directly indicated, distances are given in parenthesis. (B) Positive F_o-F_c difference electron density (green) following removal of the oxo bridge atom and trimming of D403 back to Ala, followed by 5 cycles of refinement. F_o-F_c difference electron density contoured at + 4σ for the oxo and at + 7σ for D403. The map was calculated using a resolution range of 38.7-2.17 Å. (C) One possible rotameric conformation of the disordered E430 side chain (green) places the carboxylate within 3 Å of Fe2, overlapping partially with the acetate (cyan) binding site.

determined to be 3.39 Å, in close agreement with EXAFS studies[268]. Strong density for a bridging atom is also apparent in the difference map following refinement of the cluster (Figure 29B) and, based on previous spectroscopic studies of the diferric state of CmlA[58, 268], we have modeled this density with an oxo atom (O602). The iron-oxo bond distances refined to 1.77 Å and 2.15 Å for Fe1 and Fe2, respectively. EXAFS studies [268] suggested the average iron-oxo distance is ~1.8 Å, raising the possibility that Fe2 may have been reduced to the ferrous state during data collection, forming a mixed-valent $Fe^{3+}Fe^{2+}$ cluster. Consistent with this, the overall average ligand distances refined to 2.0 Å for Fe1 and 2.3 Å for Fe2. We investigated the possibility of reduction during data collection by several standard approaches as detailed in the Materials and methods. However, none of these methods provided definitive evidence for a mixed-valence cluster.

The cluster in CmlA is coordinated by one less endogenous amino acid carboxylate and one more histidine ligand than found for most other diiron hydroxylases. All nitrogen ligands derive from the conserved MBL metal-binding scaffold. The N-terminal domain contributes no cluster ligands. Fe2 is coordinated by D309, H310, a monatomic bridging D403, an acetate molecule (ACT603) derived from the crystallization buffer, and the bridging O atom. Fe1 is coordinated by a chelating E377, bridging D403, H305, H307 and the bridging O atom. It is likely that in solution the coordinating acetate is replaced by solvent because no acetate is added to the purification buffer. Exposure of CmlA to 100 mM potassium acetate in buffered solution results in no detectable changes in the chromophore.

Monodentate diiron cluster bridging structure. The monodentate μ-η^2 binding mode of D403 in CmlA was unexpected, because in most oxygen-activating diiron enzymes and model complexes the carboxylate forms a μ-1,3 bridge. The monodentate bridging mode has been observed in purple acid phosphatase, the reduced form of the D84E mutant of *E. coli* ribonucleotide reductase β subunit, and reduced sMMO hydroxylase component[209, 318-321]. However, in the latter two cases, the bridging mode is μ-(η^1,η^2) where the second oxygen of the carboxylate coordinates to one of the irons. Also, in both cases, a traditional μ-1,3 bridging carboxylate is also present. It is unlikely that D403 in CmlA can assume a μ-1,3 orientation without significant rearrangements of the backbone due to several constraints: *(i)* D403 is sterically restrained and essentially limited to rotation around the Cβ-Cγ bond, *(ii)* D403 needs to rotate towards Fe1 to bring the second oxygen close enough to coordinate Fe2; not only would this introduce severe strain on the sp^3 hybridization of the D403 Cα, but it would result in steric clashes with E377 and a too-short iron-oxygen bond to Fe1. Finally, *(iii)* if the backbone rearranges, it would be required to undergo a nearly 90° rotation to bring the carboxylate into an allowed conformation close enough to coordinate both metals. Indeed, rubredoxin:oxygen oxidoreductase has a bidentate Asp bridge and the backbone runs perpendicular to the iron-iron axis[265], whereas in CmlA and purple acid phosphatase the backbone runs more parallel[209]. Thus, the bridge in CmlA is essentially “locked” into a single orientation during the catalytic cycle in the absence of significant structural rearrangement.

Structure of the substrate binding pocket. The diiron cluster is located 10 Å from the surface of the enzyme. An opening (~6 Å in diameter) and a short channel are located above the cluster. This entrance is positioned at the base of the surface groove that is formed by the protrusions from the catalytic and N-terminal domains of the enzyme (Fig. 30A). The active site cavity is elongated and extends partially into the interior of the protein toward residue Q308 (Fig. 30B). The diiron cluster sits at the end of the pocket that is closer to the entrance. The sides of this pocket are lined by hydrophobic residues W20, F37, W438, I439, L458, V492, and M493.

In all models, additional areas of residual positive $|F_o|$-$|F_c|$ density were observed in the part of the pocket above the cluster. We attempted to model this density with solvent molecules, but this resulted in significant residual positive density after refinement. Co-crystals of CmlA with 10-20 mM amino acids L-tyrosine, L-phenylalanine or L-PAPA resulted in isomorphous crystals, but insufficient density to model any amino acid substrates in the binding pocket. This density appeared nearly identical to that observed in crystals that were not co-crystallized with amino acids and does not correspond well with components of the crystal mother liquor, thus its source remains unknown. The absence of significant electron density for any amino acids in these co-crystallized samples demonstrates that free amino acids have little affinity for binding in the CmlA active site. This behavior was also observed in a cytochrome P450 tailoring system where no amino acid binding was detected even with a 500-fold excess over protein. In contrast, the NRPS-tethered substrate had a binding affinity in the low micromolar range[322]. The possible binding mode of the NRPS-bound substrate in CmlA is modeled and discussed in context below.

The monooxygenase chemistry catalyzed by CmlA requires delivery of two protons and two electrons to the reactive center during turnover, leading to release of a water molecule. We searched for possible endogenous proton donors in CmlA. One candidate is H258 which is positioned only 3 Å from Fe2 in the same plane as the cluster (Fig. 29A). Ramachandran analysis shows H258 to be in a strained orientation, likely because the residue engages in a hydrogen bond

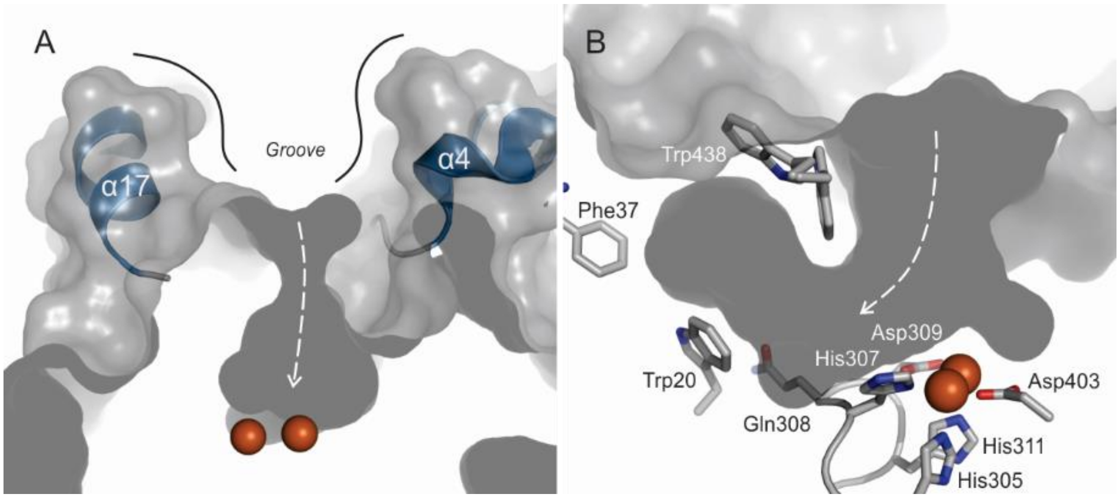

Fig. 30: Structure of substrate-binding cavity of CmlA. (A) The channel to the active site (dark gray) is located at the base of a groove formed by protrusions from both domains of CmlA. The helices α4 and α17 are important in forming the sides of this groove and are indicated in the panel. (B) The substrate-binding cavity extends into the interior of CmlA (dark gray). The side-chains of residues which form the cavity are shown and labeled, along with the diiron cluster. The white dashed arrows illustrate a possible path for the tethered substrate as it is delivered by CmlP.

with D309. On the opposite side, another His residue H378 is positioned 5 Å from Fe1 but points away from the cluster. Alternatively, the surface-exposed E430 could act as a proton shuttle from bulk solvent.

Conserved structural features of CmlA and its homologs. The structure of CmlA permits the sequence-based structural comparison of its family members. CmlA homologs are encoded in the genomes of a wide variety of mostly soil-dwelling bacteria and some cyanobacteria but the biosynthetic roles of most of these enzymes have not been determined. The biosynthetic operons for the glycopeptide antibiotics teicoplanin (from *A. teichomyceticus*)[240, 323], A47934 (from *Streptomyces toyocaensis* strain NRRL 15009)[239], and A40926 (from *Nonomurea* sp.)[241] contain homologs that accept L-Tyr as the NRPS-tethered amino acid substrate. A L-His-accepting homolog may also be involved in bleomycin biosynthesis in *Streptomyces verticillus*[304]. The biosynthetic pathway for lysobactin from *Lysobacter sp.* ATCC 53042 contains three β-hydroxylated amino acids (Phe , Leu, and Asn), but only one hydroxylase is in the operon, suggesting that a broader substrate specificity may be accommodated[324]. Some

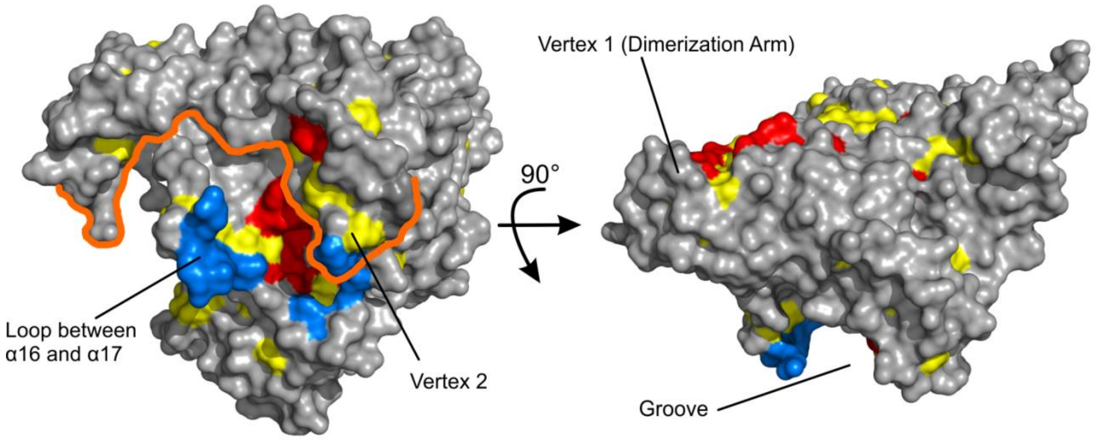

Fig. 31: Surface diagram showing locations of conserved residues based on the alignment of CmlA with 13 homologs. Colors indicate conservation of chemical character (yellow) or absolute (red) conservation. Residues which are not conserved but are at the interface between CmlA and the docked CmlP PCP domain are shown in blue (see Fig. 33A). One side of the N-terminal domain of CmlA is outlined in orange for clarity. Most of the surface of CmlA is highly variable, including the protruding arms that form the surface groove.

other notable species that encode a homolog and are known to produce antibiotics or other bioactive compounds using NRPS systems include *Teredinibacter turnerae*[325], *Nodularia spumigena*[326], and *Streptomyces tsukubaensis*[242].

In general, CmlA homologs are of similar length (500-540 residues) and typically have 30-45% sequence identity. In some cases, a C-terminal thioester-reductase domain is present; this may functionally replace the reductase domain of the NRPS that releases the hydroxylated product. Conserved residues in the N-terminal domain are primarily found to stabilize the tertiary structural elements, particularly the dimerization arm and the kinked, L-shaped arm on vertex 2 of the domain (Fig. 31). The characteristic His-X-His-X-Asp-His motif is present in all of the homologs. The first-sphere iron ligands are conserved in all homologs, along with several second-sphere residues including Y440, H258, H378, and E430 whereas Q308 is replaced by His in many cases. Most of the hydrophobic residues that line the substrate binding cavity (F37, I439, L458, V492, M493) are not conserved but are often substituted by residues with a similar chemical character. In contrast, W20 and W438 are entirely conserved.

Putative electron delivery routes have been identified in other diiron enzymes[327], and we attempted to find such a route in CmlA, as well as identify a possible ferredoxin (Fd) or flavo-

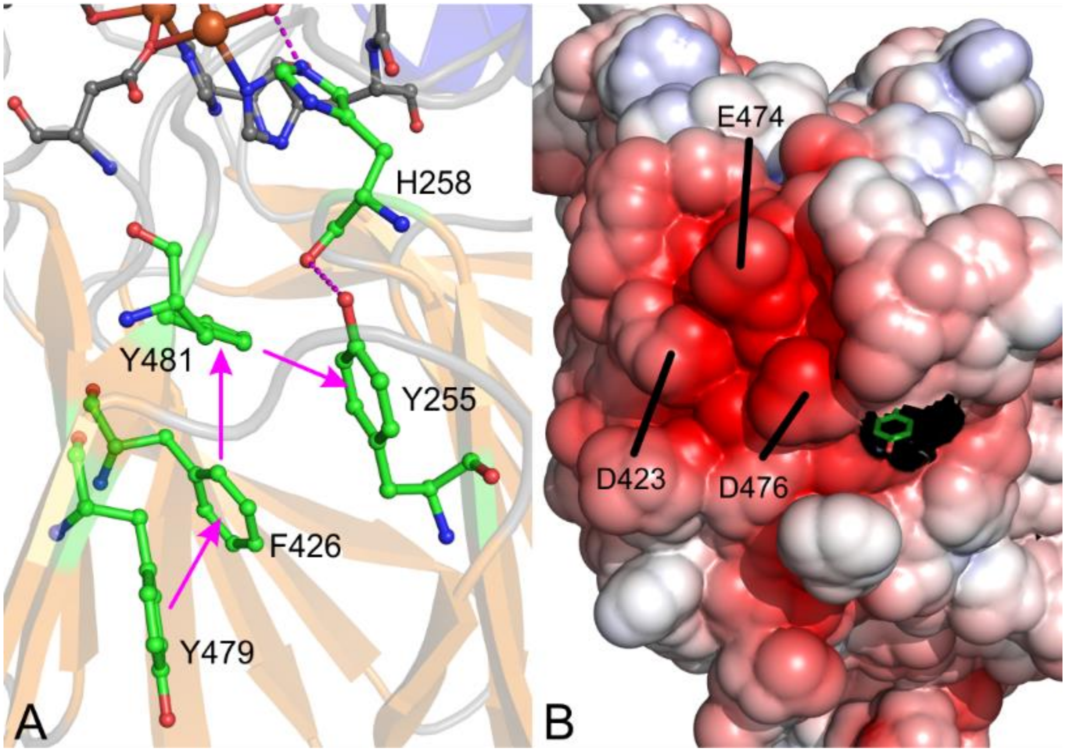

Fig. 32: Proposed electron delivery network and reductase binding site in CmlA. (A) Conserved aromatic network that extends from solvent-exposed Y479 to the diiron cluster. (B) Surface electrostatics map of CmlA showing a possible acidic binding site at the base of catalytic domain. Y479 is drawn in stick colored by atom (carbon, green). The map was generated using APBS [http://www.poissonboltzmann.org/] and contoured at 5 kT/e.

iron-sulfur reductase interaction site. From the sequence alignments, we noted a partially conserved buried aromatic network starting at the solvent-exposed Y479 that runs to underneath the cluster through the interior of the catalytic domain terminating at Y255, which is hydrogen bonded to the carbonyl group of H258 (Figure 32A). In more distant homologs, some positions are substituted by other aromatic residues. CmlA surface electrostatics reveal a strongly acidic patch (residues D423, E474, D476 and E477) at the base of the catalytic domain (Fig. 32B) that might form a binding site for an electron transfer partner. Y479 is next to this patch and engages in a hydrogen bond with the carboxylate of E477. The acidic nature of this region would be more compatible with direct electron transfer from a flavo-iron-sulfur reductase[30] than a Fd[31], which is likely to carry a strong negative charge. Unfortunately, a specific reductase protein for CmlA has

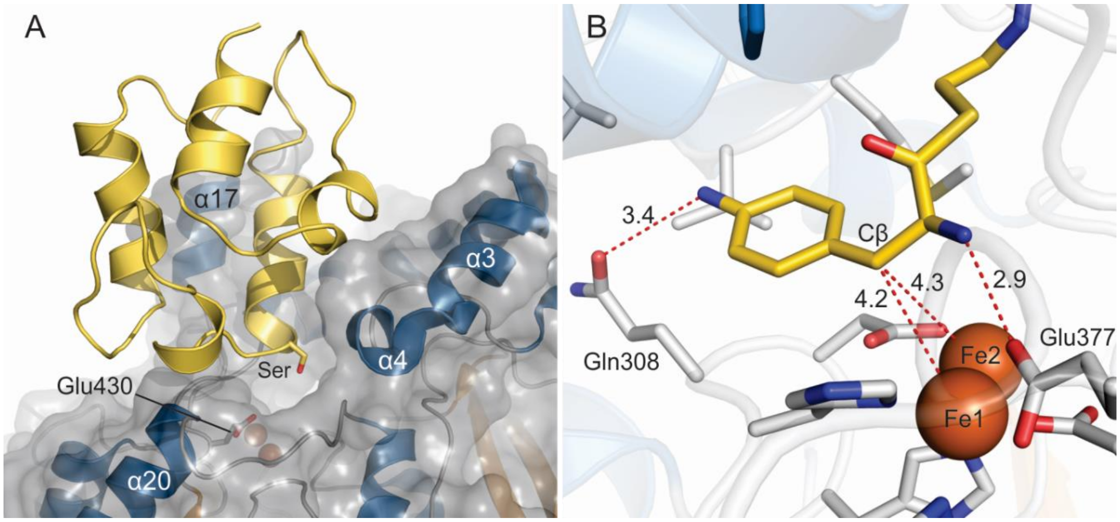

Fig. 33: Structures of the CmlP PCP domain and the amino acid substrate computationally docked to CmlA. (A) Docked structure of a homology model of the CmlP PCP domain (gold) to the surface of CmlA (gray). The conserved serine residue on the PCP domain that is the site of PPant attachment is shown as sticks and is located above the substrate channel. Nearby α-helices on CmlA are labeled. The location of the CmlA residue Glu430 is shown as gray sticks. (B) L-PAPA-PPant docked (gold) in the active site CmlA. The β-carbon of L-PAPA is positioned roughly 4.3 Å from the cluster. Bond distances are given in Å.

not been identified, and the Cml operon appears not to encode such a protein[264]. Thus, a definitive test of this hypothesis through mutagenesis is not currently possible.

Interaction of CmlA with the NRPS. CmlA hydroxylates only L-PAPA tethered to the carrier PCP domain of CmlP via a thioester bond to a PPant linker, minimally necessitating a specific interaction between the holo-PCP domain and CmlA. In order to identify the binding site of the PCP domain of CmlP and elucidate possible features conferring NRPS specificity, we docked a homology model of the highly conserved CmlP PCP domain[253, 328] to the crystal structure of CmlA (Fig. 33A). In this model, the carrier protein binds directly above the channel to the active site, as observed in similar systems[329] and the serine residue where the PPant cofactor attaches is poised above the channel to the active site. The PCP domain interacts primarily with the loop between α16 and α17 that extends from the catalytic domain of CmlA and the residues around the channel to the active site. The binding site is minimally conserved (Fig. 31) consistent with the high specificity of the hydroxylation reaction.

Sequence-based structural comparisons of CmlA and homologs show that the surface composition of the N-terminal domain is highly variable. Nevertheless, there are 12 conserved residues, and these appear to be required to maintain the overall folded structure of each of the protruding arms on CmlA. This suggests that each arm is functionally relevant, but the small size of the CmlP carrier domain would limit its interaction to a portion of the arms identified in the docked model. Analysis of CmlA surface electrostatics reveal that the tips of the arms, which form the groove between domains, have a net positive charge relative to the rest of the surface suggesting that it may be used as a mechanism to electrostatically target the correct NRPS. One interpretation is that NRPS specificity of the *trans*-acting tailoring enzymes from the CmlA family is conferred by interactions with the entire NRPS module and not just the CmlP carrier domain. Resolution of this possibility awaits further study.

NRPS-loaded substrate binding in the active site of CmlA. In the P450$_{biol}$-ACP crystal complex, PPant-tethered substrate is funneled into a long U-shaped cavity that positions the hydroxylation site above the heme for oxidation[329]. We propose that a similar mechanism functions in CmlA, and that substrate positioning is primarily directed by the shape of the pocket and the length restraints of the linker, resulting in the positioning of the β-carbon of L-PAPA above the cluster. The shape of the pocket suggests an orientation where the tethered amino acid substrate is pushed past the cluster, towards Q308 and between the side chains of M493 and I439 (Fig. 30B). The length of the linker requires that the substrate be pushed farther down towards Q308, with the aromatic portion extending beyond the cluster. The Q308 side-chain stabilize L-PAPA in a specific orientation through hydrogen bonding to the *p*-NH_2 group of L-PAPA. In CmlA homologs, Q308 is substituted with other residues capable of H-bonding to substrate. In Tyr-accepting homologs (from *A. teichomyceticus*, *Nonomurea sp.*, *and S. toyocaensis*) involved in synthesis of teicoplanins and other glycopeptides, Q308 is substituted with Ser. In a putative His-accepting homolog from *S. verticillus* involved in bleomycin synthesis, Q308 is replaced by His. The β-carbon of L-PAPA is the site of hydroxylation, and must be positioned near the cluster after binding of substrate. In our docked CmlA-PCP domain model, the distance between the PPant attachment site and the cluster is 13 Å, which matches well with observations from other systems[329, 330]. Upon binding, this would position the L-PAPA amide somewhere above the cluster and β-carbon slightly beyond it, near the site occupied by the oxo bridge in the resting state structure.

In order to test this possibility, we docked a model of the PPant-L-PAPA substrate to our structure of CmlA (Fig. 33B). In this model, the aromatic group of L-PAPA is positioned in the pocket identified in Figure 29, and *p*-NH_2 group interacts with nearby carbonyl groups from V492 and the side-chain of Q308. The β-carbon is positioned above the cluster, with the correct pro-chiral hydrogen placed within 3.4 Å of both Fe atoms. These results support our proposed binding orientation and are consistent with a role for Q308 in orienting substrate in the pocket.

Regulation of O_2 activation. Intermediates formed during O_2 activation are potentially dangerous oxidants to living systems[129, 133]. BMMs have evolved sophisticated regulatory mechanisms to only generate these oxidizing species in the presence of substrate and/or regulatory protein binding partners in order to mitigate undesired side-reactions (see Figure 5 in the Introduction)[53, 74, 109, 166, 270]. The triggering of O_2 activation by CmlA only upon binding to substrate-loaded CmlP is a salient example, and it implies a regulatory role for the NRPS[58]. The mechanism of regulation of O_2 activation in the BMM family often involves structural changes. For example, the binding of the regulatory protein MMOB to the hydroxylase component of sMMO causes structural changes near the diiron cluster[162, 271], and similar perturbations may also occur in CmlA. Comparison of the predicted and observed cluster ligand structures of CmlA suggests a possible regulatory mechanism.

In our earlier study, the cluster ligands reported here were correctly predicted based on sequence comparisons to other members of the MBL superfamily and alignments to CmlA homologs[58]. However, E430 was also identified as a likely fourth carboxylate ligand to the diiron cluster. Analysis of the electron density maps shows that this residue is positioned nearby on a loop (430-437) that runs along the base of the groove between the domains above the acetate binding site of Fe2. It does not coordinate the cluster and the side chain of E430 appears disordered. One possible rotameric conformation of E430 places the carboxylate group within ~3 Å of Fe2 (green in Figure 29C), an orientation that partially overlaps with the acetate ligand (and the solvent(s) that presumably replaces the acetate when the enzyme is in solution). The residue is far enough away that it cannot coordinate to the cluster without a concomitant shift in the backbone. Coordination of E430 to Fe2 would build the carboxylate rich diiron cluster that differentiates O_2 carrier proteins from O_2-activating enzymes[58, 268, 331]. Consequently, a conformational change that allows E430 to coordinate might serve as the switch to promote O_2 activation. E430 is located on a loop directly below the tentative binding site of the CmlP PCP domain identified in the docked structure (Fig. 33A), so that the docking of CmlP may shift E430 close enough to coordinate to Fe2. Conservation of the nearby G433 residue could indicate it is required in order to maintain the backbone flexibility required for this change.

An initial examination of the hypothesized role of E430 was made using the E430A variant. A single turnover of fully reduced E430A in complex with L-PAPA-loaded CmlP$_{AT}$ resulted in a 25-fold decrease in the maximal rate constant for reaction with O_2 as shown in Figure 34. In contrast, the autooxidation rate in the absence of L-PAPA-loaded CmlP$_{AT}$ was unchanged from that observed for wild type CmlA. Total iron loading of the active site of the E430A variant was comparable to that of wild type CmlA, but exhibited a blue-shifted optical absorption spectrum. This demonstrates that E430 interacts strongly with the diiron cluster despite its remote position from Fe2. As expected, the optical spectra of diferrous CmlA and E430, with or without L-PAPA loaded CmlP$_{AT}$ bound, were bleached and featureless in the visible region[58]. These results support the hypothesis that E430 has a role in establishing a diiron cluster structure that is capable of rapidly activating O_2. However, the fact that E430A still accelerates the rate of O_2 reaction over autooxidation suggests that the complex with the loaded NRPS has additional structural and/or electronic effects on this reaction. Initial attempts to crystallize E430A were successful, but the resulting crystal forms did not diffract to high resolution.

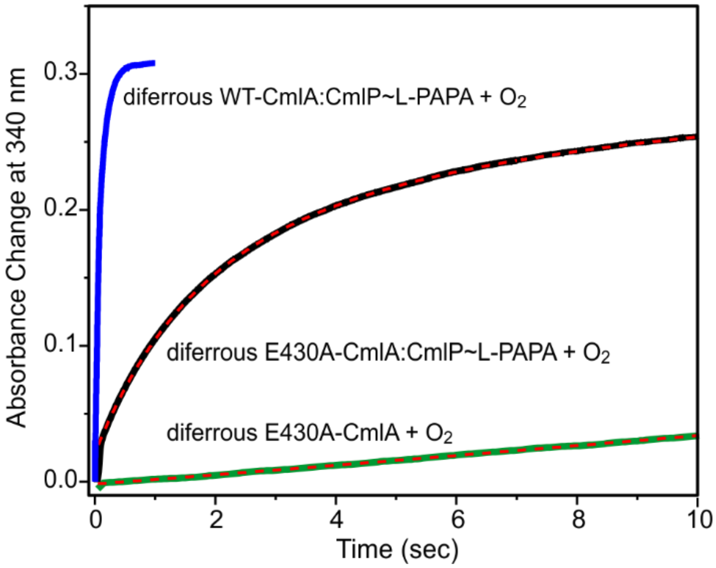

Fig. 34: Stopped flow reaction between 100 µM chemically reduced wild type or E430A CmlA in complex with 2.5-fold excess L-PAPA-loaded CmlPAT rapidly mixed with O_2 saturated buffer at 4 °C at pH 7.5. The reaction of wild type CmlA (blue trace) is complete within 1 s ($1/\tau_1 = 17\ s^{-1}$ and $1/\tau_2 = 5.2\ s^{-1}$). The E430A reaction (black trace) can be fit by a sum of three exponentials (red dashed line) with $1/\tau_1 = 0.68 \pm 0.09\ s^{-1}$, $1/\tau_2 = 0.18 \pm 0.04\ s^{-1}$ and $1/\tau_3 = 0.020 \pm 0.006\ s^{-1}$. The basis for the multiple exponential time course is unknown. Auto-oxidation of the cluster upon exposure of reduced E430A to O_2 in the absence of CmlP (green trace) can be fit (red dashed line) to a single exponential with $1/\tau_1 = 0.014 \pm 0.007\ s^{-1}$, the same as that observed for wild type CmlA autooxidation. The reaction of wild type CmlA (blue trace) is from reference[58].

Conclusion

Tailoring of natural products during biosynthesis in NRPS-based systems is highly specific. The X-ray crystal structure of the archetypal diiron cluster-containing β-hydroxylase from an NRPS system described here provides insight into the basis for this specificity. It is likely to involve interactions directly with the substrate, L-PAPA, in the CmlA active site and also with the NRPS to which L-PAPA is covalently attached. The two structural domains identified in CmlA may serve different roles in establishing specificity. The structure demonstrates directly that the C-terminal domain houses the active site in which the diiron cluster responsible for oxygen activation is bound. The full role of the N-terminal domain with its previously uncharacterized structural motif remains open to speculation. While the structure shows that this domain facilitates dimerization in this family of enzymes, its proximity to the channel into the active site and the structural conservation of CmlA homologs also imply a role in establishing the NRPS complex. Finally, the NRPS may also act as a regulator of O_2 activation, delivering substrate and engendering conformational changes in CmlA that ensure that the oxidative chemistry only occurs in the presence of the tethered amino acid.

Acknowledgements

We thank Drs. Lawrence Que, Jr. and Erik Yukl for valuable discussions. Diffraction data were collected at Argonne National Laboratory, Structural Biology Center at the Advanced Photon Source. Argonne is operated by UChicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357. We are grateful for resources from the Supercomputing Institute and the Kahlert Center for Structural Biology at the University of Minnesota.

CHAPTER 4

Crystal structure of CmlI, the arylamine oxygenase from the chloramphenicol biosynthetic pathway

Abbreviations:

CmlI: full length wild-type CmlI enzyme

CmlIΔ33: 33 amino acid N-terminally truncated variant of CmlI used for crystallization

CAM: chloramphenicol

NH_2-CAM: D-threo-1-(4-aminophenyl)-2- dichloroacetylamino-1,3-propanediol, the arylamine precursor of CAM, native CmlI substrate

Fe-AurF: the functional diiron variant of the aureothin arylamine oxygenase, AurF

Mn-AurF: the catalytically inactive dimanganese variant of AurF

RNR-R2: the R2 subunit of ribonucleotide reductase which houses the diiron cluster

sMMO: soluble form of methane monooxygenase

Summary

Nitroaromatic compounds are found in a variety of important industrial, medical, and agricultural chemicals. Natural biosynthetic pathways for nitro compounds are thus of considerable interest due to their potential importance in devising green synthetic routes for these chemicals. Only a few enzymes involved in the biological formation of nitro compounds have been studied. The chloramphenicol biosynthetic operon in *Streptomyces venezuelae* encodes an arylamine oxygenase enzyme that catalyzes the final step in the pathway. This enzyme, CmlI, converts the arylamine group of the chloramphenicol precursor to the equivalent nitro, thereby generating the final bioactive antibiotic. CmlI uses a diiron cluster cofactor to activate molecular oxygen (O_2) for the reaction with organic substrate. In this chapter, we report the X-ray crystal structure of an N-terminally-truncated variant of CmlI in the as-isolated Fe^{3+}/Fe^{3+} and chemically-reduced Fe^{2+}/Fe^{2+} cluster states. The structures allow us to compare CmlI to other arylamine oxygenases and more distantly related diiron enzymes. We also find evidence for geometric flexibility at the cluster in the form of a carboxylate shift of one of the cluster ligands. This flexibility may represent an as-of-yet uncharacterized regulatory mechanism.

Introduction

Nitro-substituted aromatic compounds (nitroaromatics) are used as synthetic building blocks for various chemicals including explosives, pesticides, and pharmaceuticals. Although relatively rare in nature, several natural pathways for the biosynthesis of nitroaromatics have been discovered and characterized[172, 280]. Many of these compounds are produced by soil bacteria of the *Pseudomonas* and *Streptomyces* genera and have antibiotic and medicinal properties[264, 332, 333]. The primary route to the generation of biological nitroaromatics is by direct enzymatic nitration of substrates[280, 334]. An alternative pathway involves oxygenase enzymes which catalyze the direct incorporation of the O atoms of molecular dioxygen (O_2) into the substrate to form the nitro group[56, 277, 283]. One of the most well-known and widely used natural nitroaromatic compounds is the antibiotic chloramphenicol (CAM) which is synthesized by *Streptomyces venezuelae*[58, 264, 277, 306]. The biosynthesis of CAM uses a non-ribosomal peptide synthetase (NRPS) pathway and several tailoring enzymes to chemically modify an aromatic amino acid precursor to the final product[58, 247, 257, 263, 277, 335]. The final step in CAM biosynthesis is oxidation of the arylamine function of the CAM precursor D-threo-1-(4-aminophenyl)-2- dichloroacetylamino-1,3-propanediol (NH_2-CAM) by the N-oxygenase CmlI (Fig. 10, blue)[277].

CmlI is part of a diverse family of enzymes that use a carboxylate-bridged diiron cluster cofactor to activate O_2 for oxidation reactions or the hydroxylation of organic substrates[109]. Some of the enzymes found in this broad family include: the hydroxylase component of the soluble methane monooxygenase (sMMOH) [28, 309], the R2 subunit of class I ribonucleotide reductase (RNR-R2)[54], plant fatty-acid desaturases[165, 336], aldehyde-deformylating oxygenase[337] and many more. All of the aforementioned enzymes coordinate the diiron cluster in a highly conserved 4-helix bundle fold and are part of the “ferritin-like” superfamily[331]. The as-isolated CmlI cluster is diferric and only reacts with O_2 upon reduction of the cluster to the diferrous state[57]. The proposed catalytic cycle of CmlI is shown in Figure 3B. Upon exposure of the cluster to O_2, it generates a diferric peroxo species which reacts directly with NH_2-CAM (Fig. 3B-*(ii)*). Most of the peroxo species that have been characterized in enzymes have a *cis* μ-1,2 geometry in which each O atom of the peroxo coordinates a different iron[60, 75, 81, 120]. In contrast, the CmlI peroxo species is proposed to have a μ-η^1:η^2 geometry in which one O atom coordinates to both metals but the other to only one. Prior to CmlI, such a species had not been observed or studied in an O_2-activating enzyme. At pH 9.0, the CmlI peroxo intermediate was found to be highly stable, having a $t_{1/2}$ of nearly 3 hours at 4 °C. The exceptional longevity of the species makes it amenable to various structural techniques, possibly including X-ray crystallography. It is of interest to determine how the unique geometry of the CmlI peroxo relates to its ability to perform N-oxygenations.

A homolog of CmlI called AurF has been characterized biochemically and structurally[56, 278, 279, 284]. This enzyme was isolated from *Streptomyces thioluteus* and catalyzes the conversion of *p*-aminobenzoate to *p*-nitrobenzoate in the biosynthesis of aureothin[333]. Crystal structures of this enzyme have been reported in the diiron (Fe-AurF)[279] and dimanganese form (Mn-AurF)[278]. The dimanganese form is catalytically inactive, but its structure was solved because of early disputes over the identity of the functional metal[338]. AurF is composed of a core 4-helix bundle motif, which is used to house a μ-oxo- and carboxylate-bridged diiron cluster[278]. The structure of Fe-AurF with the reaction product showed that the substrate binds in a cavity above the cluster but not directly to the metals[279]. The crystal structure of CmlI has not yet been solved, but it is of considerable interest as it catalyzes the final step in the synthesis of one of the most widely-used antibiotics. Its structure would also facilitate comparisons to AurF, which has a different substrate specificity. The structure could also be compared to other diiron enzymes in this diverse family which would generate insights into how the architecture of the fold and active site directs the chemistry that takes place.

Here we report the crystal structure of an N-terminally truncated variant of CmlI in the as-isolated state at 1.97 Å resolution and the chemically reduced state at 1.96 Å resolution. This truncated form of the enzyme exhibits the same reactivity as the full-length enzyme and can generate CAM product. Although the crystal structures reveal that CmlIΔ33 has an overall fold that is similar to that of AurF, there are marked differences in some of the secondary structure elements, the diiron cluster architecture, and substrate channel between these two enzymes. The differences in the clusters provide insights into the geometric variability of diiron cofactors and permits structural comparisons to other family members like sMMOH, RNR-R2 and the fatty acid desaturases. The structural malleability of the CmlIΔ33 cluster may represent a type of catalytic regulation found in the arylamine N-oxygenases.

Materials and methods

Mutagenesis, overexpression, and purification. Polymerase chain-reaction (PCR) was used to generate the 33 amino acid N-terminally truncated version of CmlI (CmlIΔ33) in a pVP91A expression plasmid (Arizona State University plasmid repository). The primers used to generate the truncated coding sequence from the full-length pVP template were as follows (gene coding sequence underlined and restriction sites bolded)

5’ AAAA**GCGATCGC**GGAGAACGCGGTCATC 3’ (forward)

5’ AAAA**GTTTAAAC**TCATCGGGTCACCGTC 3’ (reverse)

This amplicon was cloned into a pCR4-TOPO vector (Thermo Fisher Scientific) and transformed into XL1 Blue cells. The CmlIΔ33-TOPO vector was isolated and digested with SgfI and PmeI restriction enzymes (New England Biolabs) and ligated into a similarly-digested pVP91A vector before transforming into XL1 Blue cells for plasmid propagation. Sequence confirmation of this plasmid was performed at the University of Minnesota Biomedical Genomics Center. Overexpression and purification of the truncated enzyme were accomplished using the same methods as previously described for the full-length enzyme[57].

Kinetic characterization of CmlIΔ33 and CAM product detection. NH$_2$-CAM was purchased from Toronto Research Chemicals and stored as a concentrated stock in methanol at -20 °C prior to use. All assays were performed in 50 mM Bicine buffer at pH 9.0 in order to facilitate comparisons to the full-length enzyme. The CmlI peroxo was generated using the methods described previously[57]. Transient kinetic experiments were performed using an Applied Photophysics SX.18MV stopped-flow spectrophotometer at 4.5 °C. For these experiments, the CmlIΔ33 diferric-peroxo species was generated beforehand, loaded into the instrument and reacted with NH$_2$-CAM in aerobic Bicine buffer. These reactions were monitored at 500 nm to follow the decay of the peroxo species. Product analysis was performed by incubating pre-formed CmlI peroxo (200 µM) with 0.5 equivalents of NH$_2$-CAM substrate on ice for 30 minutes. The reaction product was separated from the enzyme using an Amicon centrifugal filter (3 kDa MWCO). The flow-through was acidified by addition of trifluoroacetic acid to 1% (vol/vol) and residual precipitant pelleted by centrifugation. High-performance liquid chromatography (HPLC) was performed using a Waters 1525 model binary HPLC pump and elution monitored using a Waters 2487 detector. The acidified samples were injected into a Phenomenex Luna hexyl-phenyl column (5 µm particle size, 150 x 4.6 mm) and eluted using a gradient of 5% to 95% acetonitrile in 0.1% formic acid at a flow-rate of 2 ml/min. Under these conditions NH_2-CAM had a retention time (t_R) of 1.14 min and CAM a t_R of 5.81 minutes. The elution was monitored at 280 nm.

Crystallization of CmlIΔ33 and molecular replacement. We screened for initial crystallization condition hits of CmlIΔ33 at the UMN Nanoliter Crystallization Facility. Initial hits were obtained from conditions containing 10-20% polyethylene glycol (PEG) 1,500 and 0.2 M ammonium tartrate, which produced small (5-10 µm) barrel-shaped crystals. The conditions were optimized for hanging-drop crystallization and modified to contain 10-20% PEG 1,500, 0.15 M ammonium tartrate, and 0-2% glycerol. Typical drop ratios contained 1-1.5 µl mother liquor and 1 µl CmlIΔ33 at 15-20 mg/ml concentration with a well reservoir volume of 0.5-1.0 ml. The pH of these drops was near 6.5. Crystals were obtained at 20-22 °C and at 4 °C. The space group obtained at the different temperatures was the same, but the morphology of the actual crystals was different. Those grown at room-temperature were triangular with longest dimensions of up to 200 µm whereas the 4 °C conditions yielded hexagonal plate-like crystals that were 50-100 µm. The 4 °C crystals required a longer time to grow—between three to seven days—as compared to overnight for those grown at room temperature. Crystals of CmlIΔ33 were chemically reduced by placing single crystals in a drop containing Ar-sparged anaerobic mother liquor with 25% glycerol, 40 mM sodium dithionite and a catalytic amount of methyl viologen. The crystals were incubated for 10 minutes before flash-freezing in liquid N_2.

Initial diffraction data of the crystals were collected in-house at the University of Minnesota Kahlert Structural Biology Laboratory at 100 K. Crystals were soaked in mother liquor supplemented with 25% glycerol as cryoprotectant and flash-frozen in liquid N_2. Data reduction of the initial data sets indicated that the crystals belonged to a trigonal P3 2 1 space group. This is an enantiomorphic space group which could possess either a $P3_2$ or $P3_1$ screw-axis symmetry. As such, we processed the diffraction data separately in both space groups. The $P3_2$ 2 1 enantiomorph determined to be the correct space group during the molecular replacement process. Analysis of the solvent content and Matthews coefficient for the crystals suggested the most likely scenario contained a single CmlI monomer in the asymmetric unit (solvent content 38% and Matthews coefficient 1.98 $Å^3$/ Da). The structure was solved by molecular replacement with the Phaser MR software[339] using a trimmed model of an AurF monomer[279]. For the molecular replacement, most of the loops near the AurF surface were removed prior to inputting the search model. Following the initial solution, CmlI was iteratively modeled using Coot[293] and the model refined using Refmac5[294]. Synchrotron-derived diffraction data for the reported structures were collected at the Structural Biology Center at Argonne National Laboratories, Beamline 19-ID at 100K. All structure figures were produced using PyMOL Molecular Graphics System, Version 1.5, Schrödinger, LLC.

Results

Rationale for generation of the N-terminal CmlI truncation variant. Attempts to crystallize the full-length wild-type (WT) CmlI with or without the N-terminal 8x-His purification tag were unsuccessful. In order to identify the possible reasons for our inability to crystallize this form of the enzyme, we generated a sequence alignment of CmlI with its closest structurally-characterized homolog, AurF. AurF has 37% sequence identity with CmlI and has been crystallized in two different space groups[278, 279]. The two enzymes are of similar overall length (Fig. 35), but we noted that CmlI contained a stretch of ten amino acids near the N-terminus that are not present in AurF. Analysis of the available crystal structures of AurF shows that the N-terminus is positioned near a protomer packing interface in the unit cell (PDB structures 3CHU and 3CHI)[279]. We theorized that the additional ten amino acid stretch of residues in CmlI could be interfering with the crystal packing and nucleation. To test this possibility, we generated a truncated variant of the enzyme where the first 33 amino acids were removed (CmlIΔ33, Fig. 35). This site was chosen based on the observation that in two of the three AurF crystal structures the first 23 amino acids are not visible presumably because of conformation flexibility, and this was interpreted as meaning that the region was not critical to the AurF crystallization[279].

Biochemical characterization of CmlIΔ33. The reaction of CmlIΔ33 was studied with the native substrate NH_2-CAM and the results compared to that reported for the full-length enzyme[57]. CmlI is able to generate a long-lived diferric-peroxo species ($t_{1/2}$ ~ 3 hours, 4 °C) at pH 9.0 after exposure of the chemically reduced enzyme to O_2[57]. This species is characterized by a weak optical band centered at 500 nm ($\varepsilon \sim 0.5$ mM^{-1} cm^{-1}). We were able to generate this species in CmlIΔ33 using the same methods and conditions as for full-length enzyme (Fig. 36A, black trace). Titration of this peroxo with NH_2-CAM indicated that roughly 3.3 equivalents peroxo/substrate were sufficient to fully consume the species (Fig. 36B). A similar 3:1 stoichiometry was observed for full-length CmlI[57]. CmlIΔ33 was also able to generate the CAM product in the expected yield (three peroxos consumed/CAM molecule produced) as determined using high-performance liquid chromatography and comparison to a CAM standard (Fig. 36C). Substrates have been shown to react with the CmlI peroxo in what appears kinetically to be a second order reaction[57]. We used stopped-flow spectroscopy to determine whether the CmlIΔ33 peroxo exhibited similar kinetic behavior. Reaction of pre-formed CmlIΔ33 peroxo with different

```
                                                  →
CMLI            MRDHTDEKSEAAGNDDGHVRIGGLPAFDPDDPAENAVINRLVGNWHRRAAVK--REEPDV 58
AurF            MREEQPHLATTWAARG----------WVEEEGIGSATLGRLVRAWPRRAAVVNKADILDE 50
                **:.  . : : .  .          :  ::   .*.:.***  * *****    :  *

                                                                  *
CMLI            YALFDPGRPDFREDMIPFRGHPIWERLSDETRSRLLSWGWVAYNRNTVLIEQRIANPAFE 118
AurF            WADYDTLVPDYPLEIVPFAEHPLFLAAEPHQRQRVLTGMWIGYNERVIATEQLIAEPAFD 110
                :* :*.  **:  :::**  **::   . . *.*:*:  *:.**...:  ** **:***:

                                         *  *
CMLI            LVIGGAYPGLGGQQLELAVAQAMVDEQYHTLMHINGSAVTRRMRRSDFSDRVLPDSHITT 178
AurF            LVMHGVFPGSDDPLIRKSVQQAIVDESFHTYMHMLAIDRTRELRKISERP-PQPELVTYR 169
                **: *.:** ..  :. :* **:***.:** **: .   **.:*: .      *:

                                          *                          *   *
CMLI            IHQEHLDRCEEPWQRSLTTLGFATVAEISINAYLDLLADDQEIQVVNSTTVKLHNRDEYC 238
AurF            RLRRVLADMPEQWERDIAVLVWGAVAETCINALLALLARDATIQPMHSLITTLHLRDETA 229
                  :. *    * *:*.::.* :.:*** .*** * *** *  ** ::*  ..** *** .

                *
CMLI            HASISGEMMKQVYEALPADRRRFLLEKVVAGLEPSSPR-LHDLGVDRRVRGRPRWGEAAA 297
AurF            HGSIVVEVVRELYARMNEQQRRALVRCLPIALEAFAEQDLSALLLELNAAGIRGAEEIVG 289
                *.**  *:::::*  :  ::** *:. :  .**. : : *  * :: .. *     * ..

CMLI            EVRQAQGGTHLVQDHSGIHTLLTEMDVLDQVEFG------WGTTVTR 338
AurF            DLRSTAGGTRLVRDFSGARKMVEQLGLDDAVDFDFPERPDWSPHTPR 336
                ::*.: ***:**:*.** :.:: ::.: * *:*.      *.. ..*|
```

Fig. 35: Alignment of the CmlI amino acid sequence from *S. venezuelae* (Genbank CCA54211.1) and AurF from *S. thioluteus* (Genbank CAE02601.1). The alignment was generated using the ClustalW server. The red arrow shows the start of the truncated CmlI variant (ENAV…) and starred residues are the diiron cluster ligands, as based on the AurF crystal structures.

concentrations of NH$_2$-CAM substrate in buffer suggested that the truncated variant also reacts directly with substrate, as evidenced by the linearity of the peroxo decay rate constant vs. NH$_2$-CAM concentration plot (Fig. 36D). The zero-intercept and linearity of the plot means that the slope of the fit gives the apparent bimolecular rate constant for the reaction of peroxo with substrate[57]. This value was comparable between CmlIΔ33 and the full-length enzyme: 12 mM^{-1} s^{-1} vs. 7.1 mM^{-1} s^{-1}, respectively. In both cases, the zero intercept of the plot shows that the reaction is irreversible. The experiments performed here demonstrate that CmlIΔ33 is functionally identical to the full-length enzyme and that the N-terminal truncation does not alter the reactivity of the enzyme significantly. Thus, the findings derived from a structure of the truncated mutant are likely to be applicable to the natural form of CmlI.

Crystallization of CmlIΔ33 and its overall structure. Crystals of CmlIΔ33 were obtained in conditions containing 10-20 % PEG 1,500, 0.15 M ammonium tartrate and 0-2 % glycerol. The enzyme only crystallized near pH 6.5; attempts to generate crystals at higher pH values 7.5 - 9.0

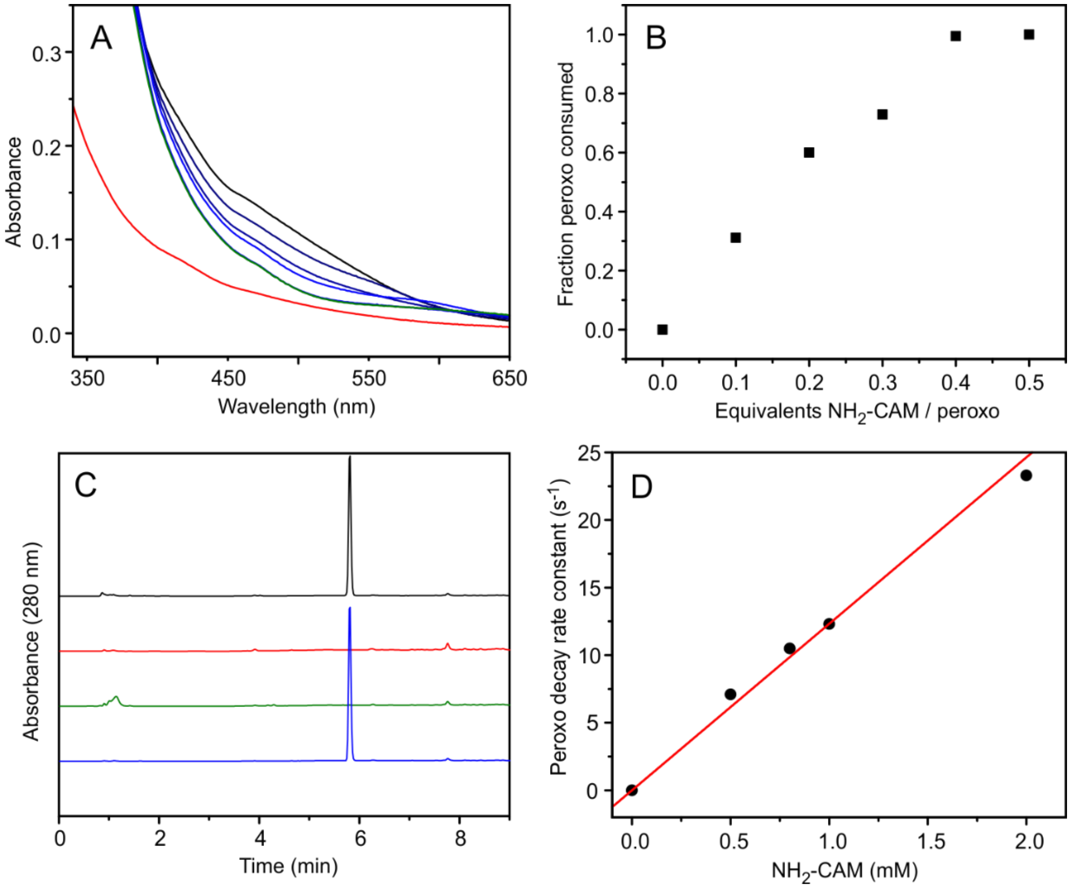

Fig. 36: Results of the biochemical characterization of CmlIΔ33. (A) Optical spectra of 200 µM CmlIΔ33 in the chemically reduced state (red), the diferric peroxo state (black) and upon titration with 0.1 molar equivalents of NH_2-CAM substrate (blue traces). The green trace is the final spectrum of diferric CmlI after the peroxo is consumed. (B) Consumption of the CmlIΔ33 peroxo when titrated with NH_2-CAM based on the observed intensity at 500 nm. (C) HPLC analysis of the CmlIΔ33 reaction products. Extracted CmlIΔ33 reaction product from reaction of 200 µM diferric peroxo with 100 µM NH_2-CAM (black), buffer-only negative control (red), 60 µM NH_2-CAM substrate standard (green), 60 µM CAM standard (blue). (D) Transient kinetic analysis of the dependence of peroxo decay on NH_2-CAM concentration as monitored at 500 nm. The red line shows a linear fit to the data. All of the reactions in this figure were performed at pH 9.0 in 50 mM Bicine buffer. The transient kinetic reactions in panel D were completed at 4.5 °C.

were unsuccessful even at 50 % higher precipitant concentrations. Data collection and model refinement statistics for the as-isolated and chemically reduced structures are given in Table 7. The enzyme crystallized in a trigonal $P3_2$ 2 1 space group with one CmlI monomer in the asymmetric unit (Fig. 37A). In all, 283 residues were modeled in the as-isolated structure and 290 residues in the reduced structure. The two structures are nearly superimposable (atomic

Table 7. CmlI X-ray diffraction data collection and model statistics.

	CmlIΔ33 as-isolated	CmlIΔ33 chemically reduced
Data collection		
Space group	$P3_2 2 1$	$P3_2 2 1$
Cell dimensions		
a, b, c (Å)	56.73, 56.73, 150.28	56.96, 56.96, 151.10
α, β, γ (°)	90, 90, 120	90, 90, 120
Wavelength (Å)	0.97926	0.97933
Resolution (Å)	50-1.97	50-1.95
Total/Unique reflections	71031/ 17711	145686/ 21534
R_{merge} (%) *,†	8.2 (50.5)	8.1 (55.3)
I/σI*	33 (2.9)	33 (3.1)
Completeness (%) *	85.7 (91.8)	99.6 (99.5)
Redundancy*	4.0 (4.2)	6.8 (6.9)
Model Refinement		
Resolution (Å)	26.54-1.97	35.24-1.95
R_{work}/R_{free} (%)‡	0.205/0.264	0.191/0.240
Average B, all atoms (Å^2)	41	37
ESU (Å) §	0.140	0.112
RMSDs		
Bond lengths (Å)	0.0108	0.0118
Bond angles (°)	1.43	1.47
Ramachandran analysis		
Allowed (%)	100	100
Outlier (%)	0	0

All data collected on synchrotron beamline APS SBC-CAT 19ID-D

* Highest resolution shell is shown in parentheses

† $R_{sym} = \Sigma_{hkl}\Sigma_i |I_{hkl,i} - \langle I\rangle_{hkl}|/\Sigma_{hkl}\Sigma_i|I_{hkl,i}|$, where I_{hkl} is the intensity of a reflection and $\langle I\rangle_{hkl}$ is the average of all observations of the reflection

‡ R_{free}, R-factor calculated from 5% of the data excluded from refinement

§ Estimated overall coordinate error (ESU) based on maximum likelihood

RMSD, root mean square deviation from ideal geometry

coordinate RMSDs 0.14 Å). Density for the N-terminal purification tag and a portion of the C-terminal section of the protein was missing and not able to be modeled in either structure. In both structures, a disordered stretch of residues between K52 and L61 could also not be modeled (dotted line in Figure 37A). CmlI is entirely composed of α-helices and the intervening loop regions. In total, the protein contains ten α-helices (α1-α10) which are labeled in Figure 37A. Like many other diiron-cluster containing enzymes, it coordinates the diiron cluster in a 4-helix bundle. His and Glu residues in helices α2, α3, α5, and α6 are ligands for the cluster. A structural search of the CmlI protein using the DALI server shows that the closest structural hit is Mn-AurF (RMSD 1.5 Å)[278] followed by Fe-AurF (RMSD 2.1 Å)[279]. The next closest hits are the R2-subunits of class I ribonucleotide reductases and the hydroxylase component of sMMO (9-12% sequence identity, RMSD 3.0-3.4 Å)[54, 309]. These enzymes also use a core 4-helix bundle fold to house the diiron cluster. The biological unit of CmlI is a dimer and the dimer interface lies along one of the crystallographic symmetry axes in the unit cell (Fig. 37B). The CmlI dimer interface buries a surface area of 1090 Å^2 and the interaction is mediated by residues in helices α2 and α3. This results in a 22 Å separation of the two diiron clusters, the same as in AurF.

In the proposed catalytic model for CmlI, the arylamine substrate binds near the diiron cluster for reaction with activated O_2[57]. The crystal structures of Fe-AurF revealed a clear channel for substrate ingress to the interior of the protein and to the buried cluster[279]. In contrast, in Mn-AurF this channel is collapsed with no clear open path for substrate to access the active site[278]. A putative substrate channel was also identified in CmlI (Fig. 38A). The entrance is located between helices α2, α5, α7, and α9 (Fig. 38B and orange arrow in 38A). This channel is lined by primarily hydrophobic amino acid side-chains: F69, P75, M73, Y101, I108, I208, Y211, F273, F278 and W281. These residues also encompass a relatively large cavity that is just inside the entrance to the substrate channel (starred in Fig. 38A). Above the diiron cluster, this channel contracts to a diameter that is too small to accommodate the relatively bulky NH_2-CAM.

Comparison of the overall structure to Fe-AurF and Mn-AurF. An alignment of the overall structure of CmlI with Fe-AurF shows that there are differences in the arrangements of some of the α-helices. The most notable of these is a break in the secondary structure of one of the four core helices (equivalent to α5 in CmlI). In AurF, there is a stretch of residues numbered 197 to 201 that are not α-helical and are pushed out away from the substrate channel and the cluster (Fig. 39A, purple cartoon). In contrast, in CmlI, this region remains α-helical (Fig. 39A, gray cartoon)

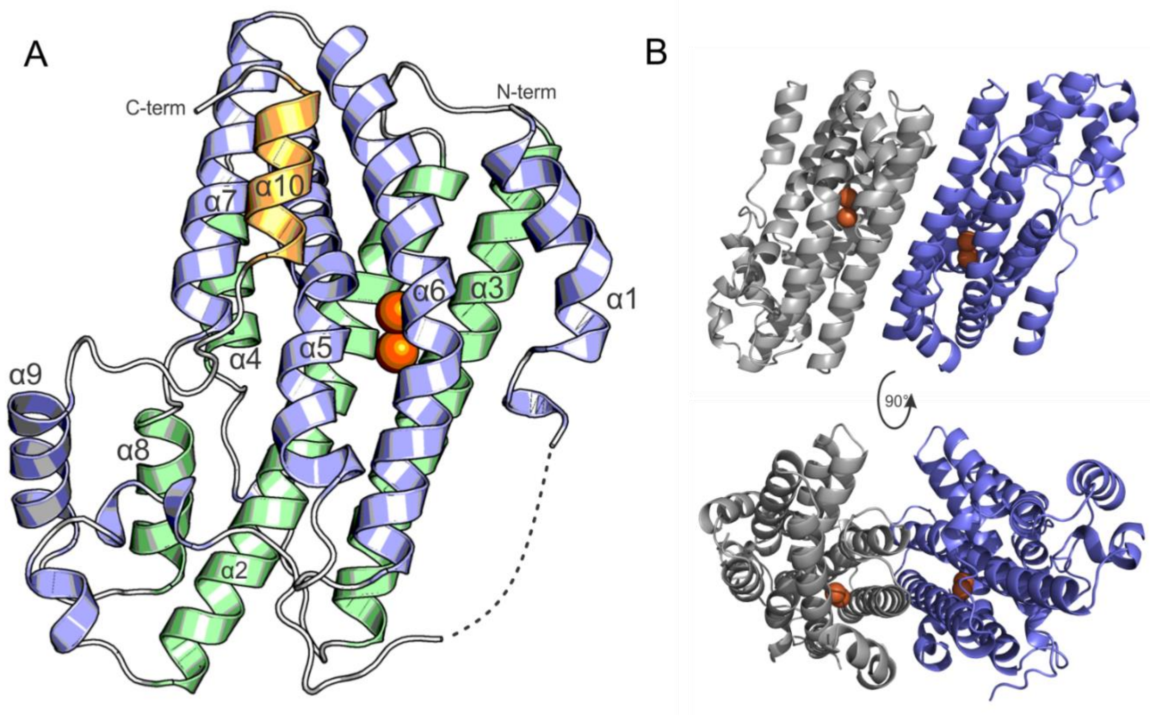

Fig. 37: Overall structure of the CmlIΔ33 monomer and dimer. (A) The CmlIΔ33 monomer with secondary structures labeled. The helices are labeled and referred to throughout the text. The dotted line represents a disordered stretch of residues that was not modeled. (B) Two views of the CmlIΔ33 dimer. One monomer is in gray and the other in blue. In both panels the iron ions are shown as brown spheres.

and is pushed in closer toward the putative substrate channel and the cluster. This has the effect of collapsing the channel above the cluster in CmlIΔ33 whereas it remains open in AurF (Fig. 39B and C). In particular, the side-chain of residue I208 (I199 in AurF) occludes the space above the CmlI cluster but not in AurF, where the side-chain is pulled away from the cluster. Interestingly, the structures of CmlIΔ33 and Mn-AurF are nearly identical in this respect. Namely, in both structures the region shown in Figure 39A remains α-helical and I208/I199 is pushed toward the channel and presumably occludes access to the active site. It will be shown in the next section that the position of I208 also has an important effect on one of the diiron cluster ligands.

In an attempt to crystallize an alternate form of CmlI where the channel was open as in Fe-AurF, we tried to co-crystallize CmlIΔ33 with various substrates including NH_2-CAM. Unfortunately, this did not yield any alternate crystal forms, nor was there any evidence for substrate binding in crystals prepared under these conditions. The substrate channel was still collapsed similar to those prepared without any substrate present in the mother liquor. Attempts to incubate the enzyme with substrate prior to adding it to the mother liquor also did not result in any detectable differences. Chemical reduction of the cluster had no observable effect on the

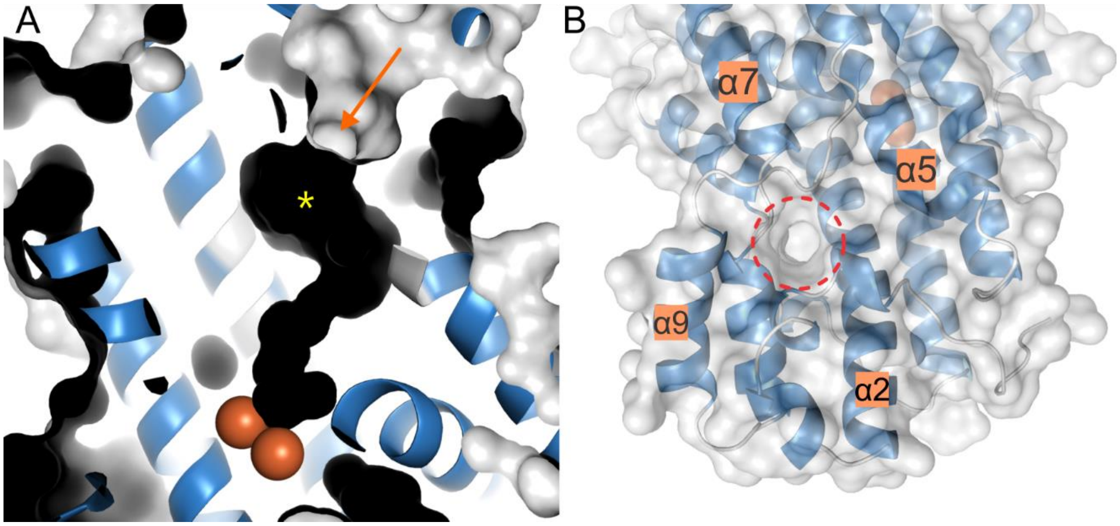

Fig. 38: A putative substrate channel and entrance found in the CmlIΔ33 crystal structure. (A) A channel in the interior of CmlIΔ33 that leads to the diiron cluster (orange spheres). The orange arrow shows the entrance to this channel that is on the surface of the protein. Above the cluster, the channel is collapsed and too small to accommodate the NH_2-CAM substrate. (B) Surface-view of CmlIΔ33 showing the entrance to the channel with the surrounding helices labeled. The α-helices of CmlIΔ33 are shown in blue and intervening loop regions in white.

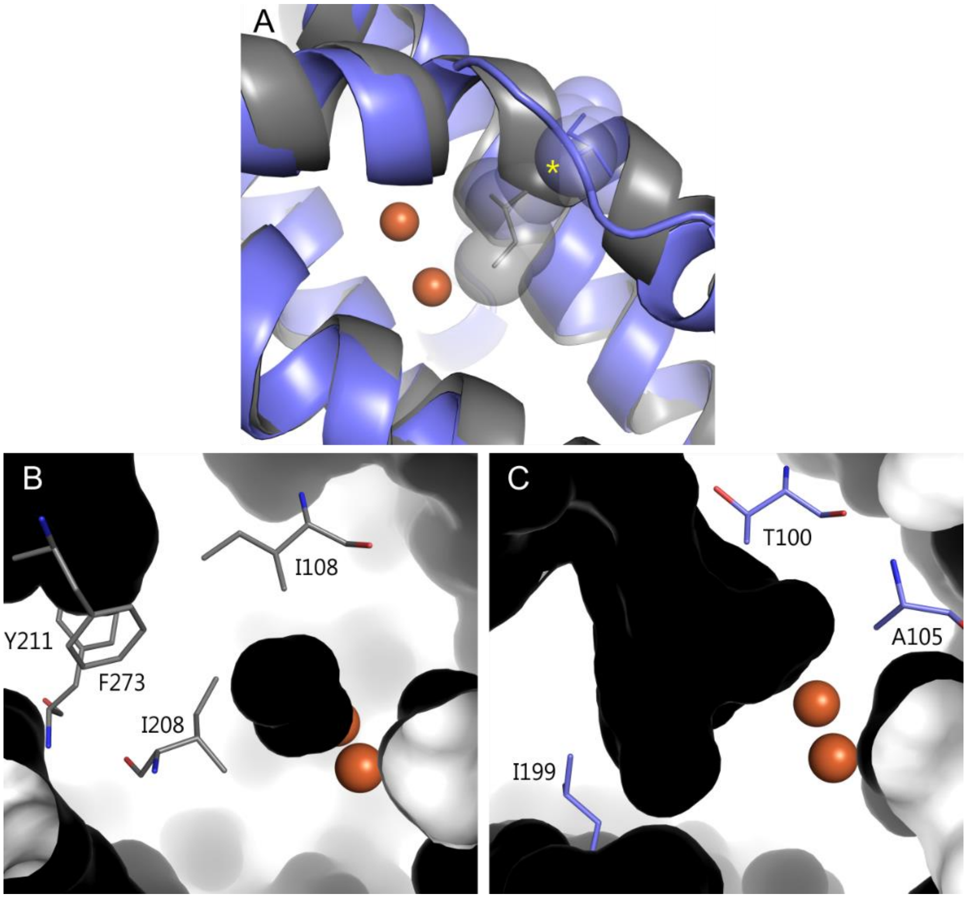

Fig. 39: Comparison of the substrate channels in CmlI and Fe-AurF and the structural difference observed in helix α5. (A) Alignment of the structure of CmlIΔ33 (gray) and Fe-AurF (purple, PDB ID 3CHH) showing the region of helix α5 which differs between the structures. In CmlI this region remains helical but in Fe-AurF there is a break in the loop as some of the residues lose their secondary structure (yellow star). The side-chains of residue I201 (CmlIΔ33, gray) and I199 (Fe-AurF, purple) are shown as sticks and the atomic radii as transparent spheres. I199 points away from the channel and cluster in Fe-AurF. (B) The collapsed substrate channel in CmlIΔ33. Important residues are shown as sticks and labeled. (C) The open substrate channel in Fe-AurF with important residues shown as sticks and labeled. Iron ions are shown as brown spheres. The channel in CmlIΔ33 appears different from that in Figure 37A because a slightly larger solvent-probe radius was used to calculate the cavity volume. This made it easier to compare to the Fe-AurF channel.

size of the substrate channel or the orientation of the residues that line the channel. Due to the similarity of the CmlIΔ33 structure to Mn-AurF, we checked to ensure that the active site was not loaded with other transition metals such as manganese. Inductively-coupled plasma mass spectrometry (ICP-MS) of the acid-digested enzyme used for the crystallization revealed that that iron was present in roughly stoichiometric ratio (1.9:1) relative to the amount of protein, as expected. Iron was also present at nearly 20-fold higher concentration than any of the other metals, making it unlikely that these other metals were present in the active site above a trace level. It should also be noted that if CmlIΔ33 were loaded with non-cognate metals it would not be expected to exhibit the reactivity or be able to generate product at the yield that we observe.

Structure of the as-isolated CmlIΔ33 diiron cluster. The diiron cluster is coordinated by four Glu and three His ligands (Fig 40A). All cluster ligands are derived from the core 4-helix bundle: E109 (from helix α2), E144 and H147 (α3), E205 (α5), and H232, E236, and H239 (α6). One metal ion (Fe1) is in a six-coordinate environment with five protein-derived ligands: E109, H232 (Nε mode), H147 (Nδ mode) and E144 and E236 which both bridge between the iron ions in a μ-1,3 mode. The other metal (Fe2) is five coordinate with four protein-derived ligands: the two bridging carboxylates E144 and E236, H239 (Nδ mode) and monodentate E205. A weakly-coordinated solvent molecule (2.4-2.5 Å bond distance) is found in the putative substrate binding cavity and interacts with both metals. The Fe-Fe distance is 3.3 Å which is similar to that seen in several other diiron enzymes in the oxidized diferric state[54, 336, 337] but is shorter than in the Fe-AurF homolog (3.45 Å)[279]. No evidence for a μ-oxo ligand is observable even though the spectroscopic characterization of the enzyme suggested the presence of such a moiety in the as-isolated state[57]. Assuming that this structure represents the diferric form of the enzyme, the absence of a dianionic μ-oxo bridge also implies that the cluster is not charge-neutral, as there are only four anionic ligands to neutralize the six positive charges.

The cluster overlays well with that of AurF with one important exception: the carboxylate of E236 is rotated in CmlIΔ33 and occupies the site where the μ-oxo bridge is found in the Fe-AurF structure (Fig. 41A)[279]. As noted, this residue bridges between the metals in a μ-1,3 mode and the carboxylate function lies in the same plane as the Fe-Fe vector. This same coordination mode of the Glu residue is observed in Mn-AurF, for which there was also no evidence of a μ-oxo bridge in the crystal structure[278]. In AurF, the equivalent residue to CmlI's E236 is E227, which is instead found in a monodentate mode, coordinated only to Fe2. The difference in orientation between the proteins stems from a projection of the side-chain of I201 towards the cluster (Fig. 39A, sticks and spheres), which was discussed in a previous section. This positioning of I201 forces the carboxylate and Cγ of E236 into nearly the same plane as the Fe-Fe vector of the cluster. In AurF, E227 is presumably able to assume the monodentate orientation because the

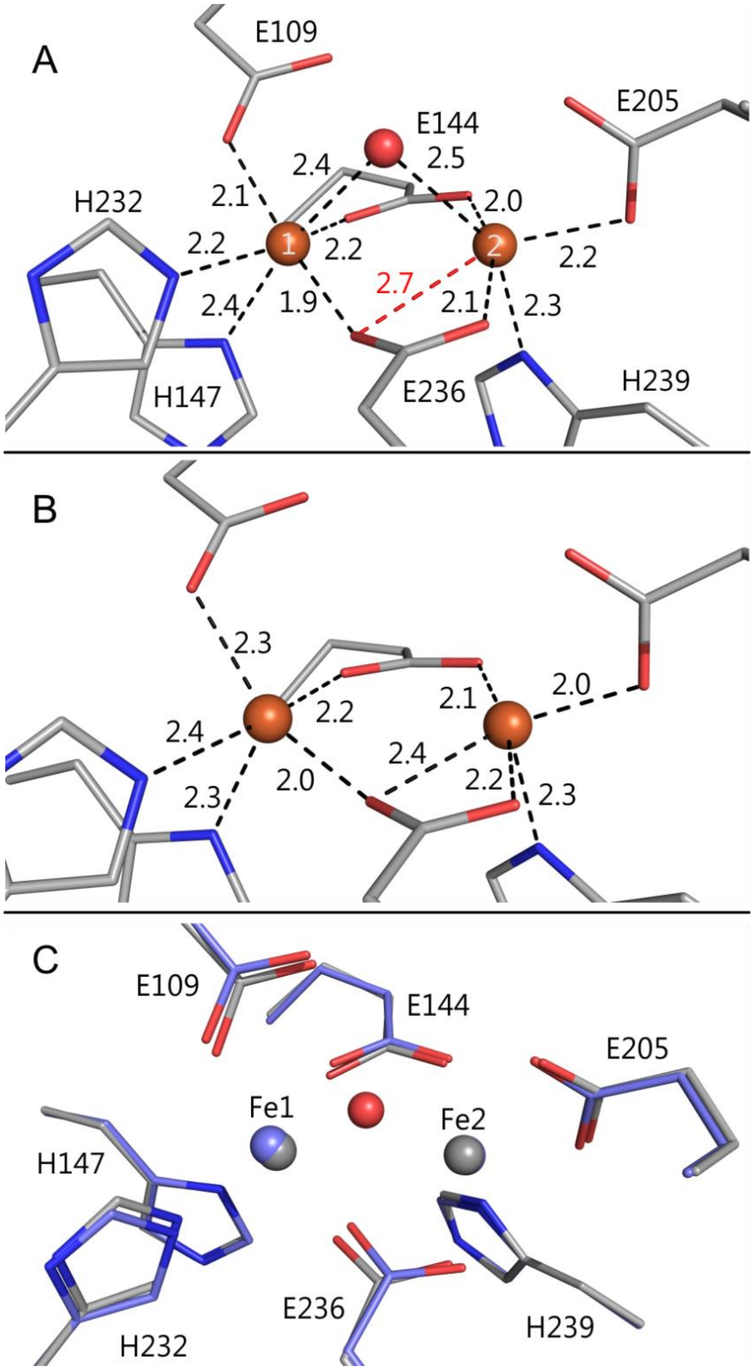

Fig. 40: Structure of the CmlIΔ33 diiron cluster. (A) As-isolated structure. (B) Chemically reduced structure. (C) Overlay of the as-isolated cluster (gray) and reduced cluster (blue). Bond distances are given in Å. Oxygen atoms are red and nitrogen atoms blue. Iron atoms in panels A and B are colored brown.

side-chain of the equivalent I199 residue is pointing away from the cluster (Fig. 39A, purple spheres). The bidentate bridging mode of E236 means that CmlI has two μ-1,3 carboxylate bridges (E144 and E236) whereas Fe-AurF only has one. Clusters with this type of ligand arrangement are also present in the diferrous RNR-R2 subunit and diferrous plant desaturases[165, 340]. A structural alignment of the CmlIΔ33 cluster with that of Fe-AurF and Mn-AurF is shown in Figure 41A and B, respectively.

Inadvertent photo-reduction of oxidized diiron clusters by the high-intensity X-ray beams generated on synchrotron beamlines is a commonly encountered phenomenon when studying this class of enzymes[341, 342]. One could argue that the absence of a μ-oxo bridge evidences reduction of the cluster. We used several approaches to test for and mitigate reduction of the CmlIΔ33 cluster during data collection and crystallization. First, the initial data sets were collected on a low photon-flux in-house diffractometer which generates X-rays orders of magnitude less intense than synchrotron radiation. Structures from these data sets did not indicate any notable differences from the structures solved at the synchrotron and were also lacking a μ-oxo bridge. For some data sets, we co-crystallized or soaked crystals with a low-millimolar concentration of the oxidizing agent potassium hexachloroiridate(IV), in order to absorb the solvent-derived electrons generated during data collection. This also did not have a detectable effect on the cluster architecture. Although we cannot rule out reduction of the cluster during data collection or crystallization, explicit chemical reduction of the crystals did result in some observable differences from the as-isolated cluster.

Structure of the chemically reduced diiron cluster. Crystals of CmlIΔ33 were chemically reduced under anaerobic conditions using methyl viologen as mediator dye and sodium dithionite as a source of reducing equivalents. No significant changes in the space group or unit cell parameters were noted compared to the as-isolated crystals. Only relatively subtle changes occur upon reduction of the cluster (Fig. 40B). Both of the metals are five-coordinate in a square planar geometry. There is no evidence of the weakly-coordinated solvent molecule and there is an open site on each iron that is pointing towards the putative substrate channel/ binding cavity. The coordination mode of all of the ligands remains the same but the overall Fe-ligand bond distances are increased relative to the as-isolated cluster (Fig. 40C). One more notable change is a movement of the carboxylate of E236 closer to Fe2 which results in an alteration of its bridging orientation closer to a μ-1,1(η^1,η^2) mode. In this orientation, one of the oxygen atoms coordinates to both irons and the other only to Fe2. This type of ligand coordination is also present in diferrous sMMOH[319] and the RNR-R2 D84E variant[318]. Perhaps the most noticeable change is an increase in the Fe-Fe distance from 3.3 to 3.5 Å upon reduction. The side-chain of I201 remains in the same location, which locks the carboxylate of E236 into the same plane as in the as-isolated state.

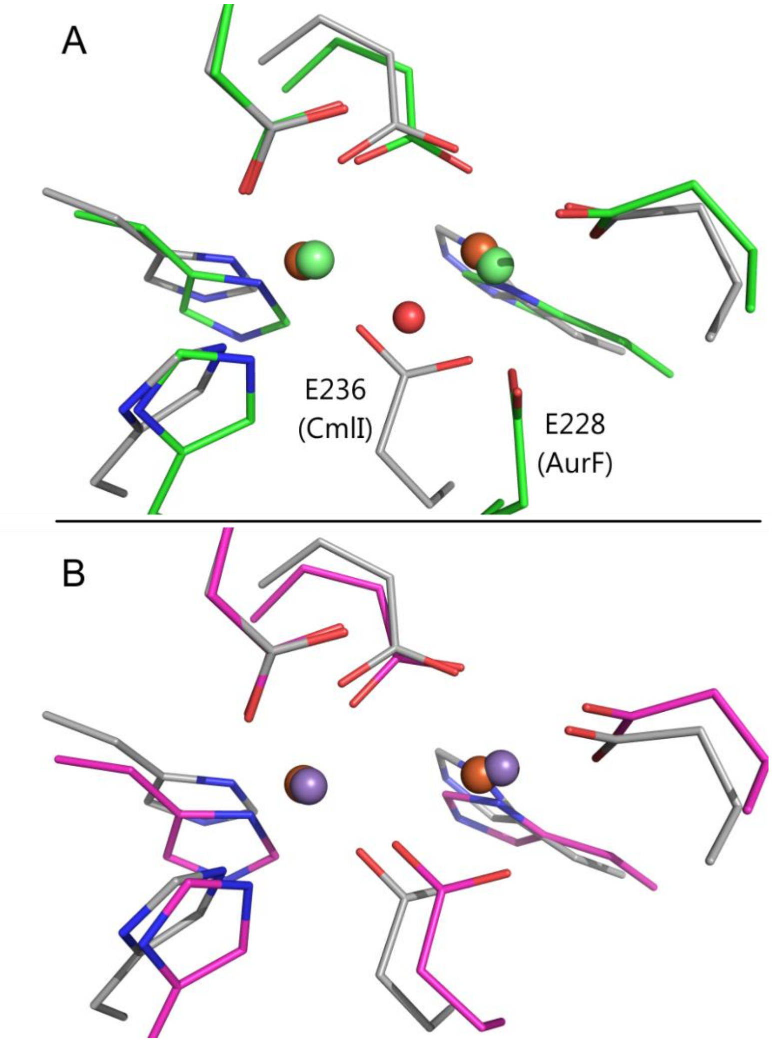

Fig. 41: Structural alignment of the CmlI and AurF clusters. (A) CmlIΔ33 (gray carbon atoms) and Fe-AurF (green, PDB ID 3CHH, 2.0 Å maximal resolution). The green spheres are the iron ions in the Fe-AurF structure and the red sphere is the μ-oxo ligand. (B) CmlIΔ33 (gray) and Mn-AurF (magenta, PDB ID 2JCD, 2.1 Å maximal resolution). The purple spheres are the Mn ions. Other coloring is the same as in Figure 40.

Discussion

A large variety of different diiron cluster ligand architectures have been described, and the crystal structure of CmlI permits another opportunity to compare and contrast the diverse structural motifs that can be found in this class of enzymes. Although the overall structure of the protein and active site is similar to that of the AurF homolog, there were some surprising differences that evidence the conformational flexibility inherent in the protein scaffold and the cofactor ligands. These observations provide an explanation for some of the spectroscopic properties observed for this enzyme in previous studies. Discussed here are the relationship of the CmlI cluster to those of other diiron enzymes as well as the implications that the structure have on the catalytic cycle of the enzyme.

Explanations for the differences in the CmlIΔ33 diiron cluster crystal structure and that of Fe-AurF. There are subtle differences in the secondary structure of CmlIΔ33 and AurF that have a dramatic effect on the cluster geometry and the size of the substrate channel (Fig. 39B and C and Fig. 41A). We considered several possible causes for these differences. One is that the pH of the crystallization buffer may have an effect on the cluster and/or the secondary structure of the protein. Prior spectroscopic studies of CmlI indicated that the diferric enzyme contained a mixture of several different cluster species which were sensitive to the pH at which the enzyme sample was prepared[57]. Diferrous CmlI was only studied at pH 9.0, but the experiments suggested the presence of a single major cluster species. The changes in the diferric cluster observed upon alteration of the pH were noted to be more complex than a simple binary conversion of one species to another. The molecular cause of these differences was not ascertained but could be caused by protonation of the bridging μ-oxo/hydroxo moiety and/or a shift of a cluster ligand. The structure of CmlIΔ33 was solved at pH 6.5, considerably lower than the pH 9.0 used for most of the spectroscopic characterizations. One possible explanation is that the complex pH-dependent behavior of the cluster is caused by loss of the bridging moiety at lower pH and a shift of E236, possibly to the bidentate μ-1,3 mode seen in the crystal. The shift of the Glu to this orientation may drive the rearrangement of α-helix identified in Fig. 39A and also result in a collapse of the substrate channel (Fig. 39B). As a comparison, Fe-AurF crystallized at pH 8.5[279] and Mn-AurF at pH 7.8[278]. The corresponding proton concentrations are 100-fold and 20-fold lower, respectively, than at the pH where CmlI was crystallized. If pH is the sole explanation for the differences, it would imply that a cluster transition takes place at some point between pH 6.5 and 7.8, as Mn-AurF also lacks the μ-oxo bridge.

An alternative explanation could be that crystal-packing forces are responsible for inducing a rearrangement of helix α5 in CmlIΔ33. This, in turn, forces the side-chain of I201 into the active site and induces a rearrangement of E236. Indeed, we find that the tip of helix α7 from one of the crystal symmetry-related protomers in the CmlIΔ33 lattice packs against this region of the protein. If this packing is the cause of the structural difference, it would mean that the transition takes place upon nucleation and, due to the high symmetry of the crystal, is propagated to other CmlI molecules as they become incorporated into the lattice. In orthorhombic Fe-AurF crystals (PDB structure 3CHH) and the monoclinic form (3CHI) there is no nearby packing interface and the residues in this region may be able to relax to the non-helical state, thereby allowing E227 to assume the monodentate binding mode. However, in Mn-AurF (2JCD) there is also no packing interface nearby, but the helix remains intact. Thus, the cause is very likely more complex than simply pH or crystal packing and is probably some combination of these or additional as-of-yet unknown forces. These observations do, however, highlight the important effects that the crystallization conditions can have on the derived enzyme crystal structures.

Comparisons of the CmlI cluster architecture to other diiron enzymes. The as-isolated form of CmlIΔ33 has two μ-1,3 carboxylate bridges which has, to our knowledge, not been observed in other diferric clusters. Clusters with one μ-1,3 bridge are present in a variety of diiron enzymes including sMMOH and its relatives[52, 53, 309, 343], RNR-R2[54], aldehyde deformylating oxygenase[337], and O_2 and nitric oxide reductases[265, 266]. Those with two μ-1,3 bridges have so far only been seen in diferrous clusters as in dithionite-reduced RNR-R2 crystals[340] and diferrous desaturases[165]. However, the Fe-Fe distance in the reduced RNR-R2 and desaturase clusters is considerably longer: 3.9 Å in RNR-R2 and 4.2 Å in the desaturases vs. 3.3 Å in CmlI. Chemical reduction increases the CmlI Fe-Fe distance but also subtly changes the orientation of E236 closer to a μ-$1,1(\eta^1,\eta^2)$ mode which makes direct comparisons more difficult. A possible explanation for the longer Fe-Fe distance in RNR-R2 is that one of the terminal Glu ligands - equivalent to CmlI's E109 - is replaced by a shorter Asp residue. Mutation of this residue in RNR-R2 (D84E) decreases the Fe-Fe separation of the mutant diferrous cluster to 3.4 Å[318]. Thus, the differences in Fe-Fe separation are likely due to a compression of the cluster by the terminal ligands. In the desaturases, the terminal ligands are both Glu, but they chelate the metals and are in a very different orientation than in CmlI, which may allow the metals to relax apart upon reduction.

There are some other differences in the cluster architectures that are worth pointing out. Compared to the clusters of RNR-R2, sMMOH, and the castor bean desaturase, the CmlI cluster has an additional His ligand to Fe1 (H232). Secondly, the orientation of the terminal Fe2 ligand E205 is rotated so that the carboxylate is nearly chelating the metal. In RNR-R2 and sMMOH this carboxylate points away from the cluster. In the castor bean desaturase, the terminal ligands are both Glu but chelate the metals in an orientation rotated roughly 90° relative to that in CmlI[165, 336]. A third difference in RNR-R2 and sMMOH is the orientation of the Glu that is equivalent to E236 in CmlI. In RNR-R2 it is in a monodentate orientation very similar to that seen in Fe-AurF. In oxidized RNR-R2, a μ-oxo bridge is also present in the same site as in Fe-AurF. When the RNR-R2 cluster is reduced to the diferrous state, this Glu ligand assumes the bridging μ-1,3 mode like in CmlI. Similarly, in sMMOH, the Glu residue in this position (E243) undergoes a significant shift upon reduction of the cluster, from an Fe2-monodentate mode (like in Fe-AurF) to $\mu\text{-}1,1(\eta^1,\eta^2)$ (more like reduced CmlI)[319, 343]. In sMMOH, E243 has a major regulatory role in the catalytic cycle of the enzyme that is tied into the residue's ability to undergo a carboxylate shift[109, 162]. It is interesting that although these enzymes are functionally very different, they can exhibit similar ligand rearrangements, a fact that highlights the evolutionary relationship of the 4-helix bundle diiron enzymes[308, 331].

The differences in the clusters discussed above may also underlie the ability of CmlI to generate a stable peroxo species compared to these other enzymes. It is possible that the presence of an additional protein-derived His ligand to Fe1 and the different compressive forces on the diiron cluster are able to prevent the enzyme from accessing the more potent Fe^{4+} intermediates such as the oxidizing bis-μ-oxo Fe^{4+}/Fe^{4+} ("**Q**") in sMMOH or the Fe^{3+}/Fe^{4+} compound "**X**" in RNR-R2[114, 272]. CmlI is unable to catalyze hydrocarbon C-H bond activation, unlike sMMOH and some variants of RNR-R2[123]. The different cluster geometries may also preferentially drive the formation of a $\mu\text{-}\eta^1\text{:}\eta^2$ peroxo in CmlI instead of the *cis* μ-1,2 species generated in the other enzymes[75, 81, 120]. This may be caused by altering the relative positioning of the available coordination sites and/or the spacing between the metal ions.

Speculation on the possible relevance of the carboxylate shift to the CmlI/AurF catalytic cycle. It is common for diiron clusters to undergo ligand rearrangements during reduction and while the enzyme progresses through the catalytic cycle. One form in which these rearrangements are manifested is through carboxylate shifts of first-sphere Asp or Glu ligands[53, 162, 166, 319, 340, 343]. Such structural changes can control the reactivity of the cluster towards O_2[158, 166, 270, 271, 344] and drive interconversion of oxygenated intermediates to generate the reactive species[98, 101]. The conformational flexibility of E236 in CmlI and E327 in AurF is further evidence that the diiron cluster ligands have a significant degree of freedom. It is unknown whether the different coordination modes are relevant to the catalytic cycle, or whether it is simply an artifact of the crystallization. The structure of the CmlIΔ33 cluster reported here would be more consistent with a "traditional" *cis* μ-1,2 peroxo based on the relative orientation of the available binding sites. What is intriguing about the plasticity of E236 is that it appears to also be tied into a change in the protein's secondary structure and a constriction or enlarging of the substrate channel. The fact that nearly the same structural change has been observed in two of these family members may suggest that it is more than a simple artifact, as the packing and crystallization conditions of the enzymes are not identical. A tentative hypothesis is that these structural changes represent a sort of catalytic regulation in this family of enzymes. In this model, rearrangement of E236 or E227 during catalysis induces a concomitant constriction of the substrate channel and thereby act to gate substrate access to the cluster and later eject the oxidized product. The gating of substrate could represent a mechanism by which the enzyme protects the reactive peroxo species until the proper substrate is available. Further experiments will be needed to elucidate if and when these protein motions occur and how they affect substrate and O_2 binding.

Conclusion

The crystal structure of CmlI is an important step in furthering our understanding of biological N-oxygenation reactions. Although the structure is quite similar to that of the aureothin-synthesizing AurF homolog, there are some interesting differences that could represent facets of a regulatory mechanism like those in other oxygenase enzymes. The presence of two bidentate bridging carboxylate ligands in the diferric state of CmlI is relatively unexpected as such bridging structures have only been observed in a few enzymes in the diferrous state. The ability of one of the CmlI carboxylate cluster ligands to assume different orientations has precedents in the diiron enzyme literature. The shifting of carboxylate ligands is intimately tied into catalytic regulation of or progression through the cycle. Further experiments will be needed to test whether the observed carboxylate shift in CmlI is of functional relevance.

Acknowledgements

We would like to thank Prof. Carrie Wilmot and her research group for many helpful discussions. We also thank Dr. Ed Hoeffner at the University of Minnesota Kahlert Structural Biology Laboratory for suggestions in indexing the CmlIΔ33 data and collecting the data sets. We also thank Klaus Lovendahl for generating the expression construct for the truncated enzyme and Anna Komor and Brent Rivard for assistance in setting up the biochemical characterization experiments. Diffraction data were collected at Argonne National Laboratory, Structural Biology Center Beamline 19-ID at the Advanced Photon Source. Argonne is operated by UChicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357. We are thankful for computational resources from the Supercomputing Institute and the facilities at the Kahlert Structural Biology Laboratory at the University of Minnesota.

CONCLUSION

The results presented herein represent a significant step forward in our understanding of the enzymology of the three enzymes and the basic principles underlying biological O_2 activation. The trapping of the two most critical oxygenated intermediates in the 3,4-PCD catalytic cycle validates decades of prior research and opens avenues for future study. The structures show that many of the mechanistic hypotheses forwarded by our group and others are likely to be correct. The next goals for the study of 3,4-PCD are to dissect the structural and catalytic roles of the auxiliary residues identified in Chapters 1 and 2 of this dissertation. Site-directed mutagenesis of these residues may even open the potential to trap intermediates with the native substrate PCA, and alter the enzyme activity and substrate specificity for industrial and environmental applications. Due to the crucial positioning of these enzymes in the global carbon cycle, it is also of import to test the hypothesis that the orientation of the alkylperoxo moiety relative to the substrate ring is determinant of the ring-cleavage site.

The crystal structures of CmlA and CmlI perhaps raise more questions than answers at this point. However, the determination of their structures provides the foundation for future studies of these enzymes. The structure of CmlA has allowed us to produce hypotheses as to the mechanism by which the enzyme is regulated and targeted to a specific NRPS and amino acid substrate. It also allows us to generate site-directed mutants of this enzyme to test many of the hypotheses generated from the analysis of the structure. Perhaps most interestingly, the CmlA structure shows that nature has devised two completely independent solutions to the problem of biological C-H bond activation. The comparison of CmlA and CmlI to the bacterial multicomponent monooxygenases and other diiron enzymes allows us to learn more about the basic principles underlying the mechanism of O_2 activation. This may then allow us to co-opt the enzymes for the synthesis of new chemicals useful to modern medicine, agriculture, and industry.

REFERENCES

[1] Osterberg, R. (1974) Origins of metal ions in biology, *Nature 249*, 382-383.

[2] Anbar, A. D. (2008) Elements and evolution, *Science 322*, 1481-1483.

[3] Dupont, C. L., Butcher, A., Valas, R. E., Bourne, P. E., and Caetano-Anollés, G. (2010) History of biological metal utilization inferred through phylogenomic analysis of protein structures, *Proc. Natl. Acad. Sci. USA 107*, 10567-10572.

[4] Howard, J. B., and Rees, D. C. (1996) Structural basis of biological nitrogen fixation, *Chem. Rev. 96*, 2965-2982.

[5] Ferry, J. G. (1999) Enzymology of one-carbon metabolism in methanogenic pathways, *FEMS Microbiol. Rev. 23*, 13-38.

[6] Banerjee, R., and Ragsdale, S. W. (2003) The many faces of vitamin B12: Catalysis by cobalamin-dependent enzymes, *Annu. Rev. Biochem. 72*, 209-247.

[7] Fauque, G., Peck, H. D., Moura, J. J. G., Huynh, B. H., Berlier, Y., DerVartanian, D. V., Teixeira, M., Przybyla, A. E., Lespinat, P. A., Moura, I., and LeGall, J. (1988) The three classes of hydrogenases from sulfate-reducing bacteria of the genus Desulfovibrio, *FEMS Microbiol. Rev. 4*, 299-344.

[8] Crane, B. R., Siegel, L. M., and Getzoff, E. D. (1995) Sulfite reductase structure at 1.6 Å: evolution and catalysis for reduction of inorganic anions, *Science 270*, 59-67.

[9] Suga, M., Akita, F., Hirata, K., Ueno, G., Murakami, H., Nakajima, Y., Shimizu, T., Yamashita, K., Yamamoto, M., Ago, H., and Shen, J.-R. (2015) Native structure of photosystem II at 1.95 A resolution viewed by femtosecond X-ray pulses, *Nature 517*, 99-103.

[10] Sono, M., Roach, M. P., Coulter, E. D., and Dawson, J. H. (1996) Heme-containing oxygenases [Review], *Chem. Rev. 96*, 2841-2887.

[11] Mitchell, S. C. (2008) Flavin mono-oxygenase (FMO) - The ‘other’ oxidase, *Curr. Drug Metab. 9*, 280-284.

[12] Sato, R., and Omura, T. (1978) *Cytochrome P-450*, Academic Press, New York.

[13] Cochrane, R. V. K., and Vederas, J. C. (2014) Highly selective but multifunctional oxygenases in secondary metabolism, *Acc. Chem. Res. 47*, 3148-3161.

[14] Yeung, Amanda W. S., Terentis, Andrew C., King, Nicholas J. C., and Thomas, Shane R. (2015) Role of indoleamine 2,3-dioxygenase in health and disease, *Clin. Sci. 129*, 601-672.

[15] Fedeles, B. I., Singh, V., Delaney, J. C., Li, D., and Essigmann, J. M. (2015) The AlkB family of Fe(II)/α-ketoglutarate-dependent dioxygenases: Repairing nucleic acid alkylation damage and beyond, *J. Biol. Chem. 290*, 20734-20742.

[16] Gennis, R., and Ferguson-Miller, S. (1995) Structure of cytochrome c oxidase, energy generator of aerobic life, *Science 269*, 1063-1064.

[17] Hayaishi, O., and Hashimoto, K. (1950) Pyrocatecase a new enzyme catalizing oxidative breakdown of pyrocatechin, *Biochem. J. 37*, 371-374.

[18] Stanier, R. Y., Hayaishi, O., and Tsuchida, M. (1951) The bacterial oxidation of tryptophan. I. A general survey of the pathways, *J. Bacteriol. 62*, 355-366.

[19] Stanier, R. Y., and Ingraham, J. L. (1954) Protocatechuic acid oxidase, *J. Biol. Chem. 210*, 799-808.

[20] Dagley, S., Fewster, M. E., and Happold, F. C. (1953) Bacterial oxidation of aromatic compounds, *J. Gen. Microbiol. 8*, 1-7.

[21] Hayaishi (1974) *Molecular mechanisms of oxygen activation*, Academic Press, New York.
[22] Hayaishi, O., Katagiri, H., and Rothberg, S. (1955) Mechanism of the Pyrocatechase Reaction, *J. Am. Chem. Soc. 77*, 5450-5451.
[23] Mason, H. S., Fowlks, W. L., and Peterson, E. (1955) Oxygen transfer and electron transport by the phenolase complex, *J. Am. Chem. Soc. 77*, 2914-2915.
[24] Bugg, T. D. H. (2003) Dioxygenase enzymes: catalytic mechanisms and chemical models, *Tetrahedron 59*, 7075-7101.
[25] Pau, M. Y. M., Lipscomb, J. D., and Solomon, E. I. (2007) Substrate activation for O_2 reactions by oxidized metal centers in biology, *Proc. Natl. Acad. Sci. USA 104*, 18355-18362.
[26] Mason, J. R., and Cammack, R. (1992) The electron-transport proteins of hydroxylating bacterial dioxygenases, *Annu. Rev. Microbiol. 46*, 277-305.
[27] Poulos, T. L., and Raag, R. (1992) Cytochrome $P450_{cam}$: crystallography, oxygen activation, and electron transfer., *FASEB J. 6*, 674-679.
[28] Fox, B. G., Froland, W. A., Dege, J. E., and Lipscomb, J. D. (1989) Methane monooxygenase from *Methylosinus trichosporium* OB3b. Purification and properties of a three-component system with high specific activity from a type II methanotroph, *J. Biol. Chem. 264*, 10023-10033.
[29] Wolfe, M. D., Parales, J. V., Gibson, D. T., and Lipscomb, J. D. (2001) Single turnover chemistry and regulation of O_2 activation by the oxygenase component of naphthalene 1,2-dioxygenase, *J. Biol. Chem. 276*, 1945-1953.
[30] Muller, J., Lugovskoy, A. A., Wagner, G., and Lippard, S. J. (2002) NMR structure of the [2Fe-2S] ferredoxin domain from soluble methane monooxygenase reductase and interaction with its hydroxylase, *Biochemistry 41*, 42-51.
[31] Sobrado, P., Lyle, K. S., Kaul, S. P., Turco, M. M., Arabshahi, I., Marwah, A., and Fox, B. G. (2006) Identification of the binding region of the [2Fe-2S] ferredoxin in stearoyl-acyl carrier protein desaturase: Insight into the catalytic complex and mechanism of action, *Biochemistry 45*, 4848-4858.
[32] Lipscomb, J. D., and Orville, A. M. (1992) Mechanistic aspects of dihydroxybenzoate dioxygenases, *Metal Ions Biol. Syst. 28*, 243-298.
[33] Schofield, C. J., and Zhang, Z. H. (1999) Structural and mechanistic studies on 2-oxoglutarate-dependent oxygenases and related enzymes, *Curr. Opin. Struct. Biol. 9*, 722-731.
[34] van Berkel, W. J. H., Kamerbeek, N. M., and Fraaije, M. W. (2006) Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts, *J. Biotechnol. 124*, 670-689.
[35] Fetzner, S., and Steiner, R. A. (2010) Cofactor-independent oxidases and oxygenases., *Appl. Microbiol. Biotechnol. 86*, 791-804.
[36] Hedegaard, J., and Gunsalus, I. C. (1965) Mixed function oxidation: IV. an induced methylene hydroxylase in camphor oxidation, *J. Biol. Chem. 240*, 4038-4043.
[37] Gunsalus, I. C., Tyson, C. A., and Lipscomb, J. D. (1973) Cytochrome P-450 reduction and oxygenation systems, In *Oxidases and Related Redox Systems* (King, T. E., Mason, H. S., and Morrison, M., Eds.), pp 583-603, University Park Press, Baltimore.
[38] Rittle, J., and Green, M. T. (2010) Cytochrome P450 compound I: capture, characterization, and C-H bond activation kinetics, *Science 330*, 933-937.
[39] Grant, J. L., Hsieh, C. H., and Makris, T. M. (2015) Decarboxylation of fatty Acids to terminal alkenes by cytochrome P450 compound I, *J. Am. Chem. Soc. 137*, 4940-4943.
[40] Vaillancourt, F. H., Bolin, J. T., and Eltis, L. D. (2006) The ins and outs of ring-cleaving dioxygenases, *Crit. Rev. Biochem. Mol. Biol. 41*, 241-267.

[41] Tchesnokov, E. P., Wilbanks, S. M., and Jameson, G. N. L. (2012) A strongly bound high-spin iron(II) coordinates cysteine and homocysteine in cysteine dioxygenase, *Biochemistry 51*, 257-264.

[42] Ivanov, I., Heydeck, D., Hofheinz, K., Roffeis, J., O'Donnell, V. B., Kuhn, H., and Walther, M. (2010) Molecular enzymology of lipoxygenases, *Arch. Biochem. Biophys. 503*, 161-174.

[43] Allpress, C. J., and Berreau, L. M. (2013) Oxidative aliphatic carbon–carbon bond cleavage reactions, *Coord. Chem. Rev. 257*, 3005-3029.

[44] Harrison, P. J., and Bugg, T. D. H. (2014) Enzymology of the carotenoid cleavage dioxygenases: Reaction mechanisms, inhibition and biochemical roles, *Arch. Biochem. Biophys. 544*, 105-111.

[45] Wolfe, M. D., Altier, D. J., Stubna, A., Popescu, C. V., Münck, E., and Lipscomb, J. D. (2002) Benzoate 1,2-dioxygenase from *Pseudomonas putida*: Single turnover kinetics and regulation of a two-component Rieske dioxygenase, *Biochemistry 41*, 9611-9626.

[46] Roberts, K. M., and Fitzpatrick, P. F. (2013) Mechanisms of tryptophan and tyrosine hydroxylase, *IUBMB Life 65*, 350-357.

[47] Rocklin, A. M., Tierney, D. L., Kofman, V., Brunhuber, N. M. W., Hoffman, B. M., Christoffersen, R. E., Reich, N. O., Lipscomb, J. D., and Que, L., Jr. (1999) Role of the nonheme Fe(II) center in the biosynthesis of the plant hormone ethylene, *Proc. Natl. Acad. Sci. USA 96*, 7905-7909.

[48] Chen, V. J., Orville, A. M., Harpel, M. R., Frolik, C. A., Surerus, K. K., Münck, E., and Lipscomb, J. D. (1989) Spectroscopic studies of isopenicillin N synthase. A mononuclear nonheme Fe^{2+} oxidase with metal coordination sites for small molecules and substrate, *J. Biol. Chem. 264*, 21677-21681.

[49] Fox, B. G., and Lipscomb, J. D. (1988) Purification of a high specific activity methane monooxygenase hydroxylase component from a type II methanotroph, *Biochem. Biophys. Res. Commun. 154*, 165-170.

[50] Small, F. J., and Ensign, S. A. (1997) Alkene monooxygenase from Xanthobacter strain Py2 - Purification and characterization of a four-component system central to the bacterial metabolism of aliphatic alkenes, *J. Biol. Chem. 272*, 24913-24920.

[51] Sazinsky, M. H., Dunten, P. W., McCormick, M. S., DiDonato, A., and Lippard, S. J. (2006) X-ray structure of a hydroxylase-regulatory protein complex from a hydrocarbon-oxidizing multicomponent monooxygenase, *Pseudomonas sp.* OX1 phenol hydroxylase, *Biochemistry 45*, 15392-15404.

[52] Sazinsky, M. H., Bard, J., Di Donato, A., and Lippard, S. J. (2004) Crystal structure of the toluene/o-xylene monooxygenase hydroxylase from *Pseudomonas stutzeri* OX1. Insight into the substrate specificity, substrate channeling, and active site tuning of multicomponent monooxygenases, *J. Biol. Chem. 279*, 30600-30610.

[53] Bailey, L. J., McCoy, J. G., Phillips, G. N., Jr., and Fox, B. G. (2008) Structural consequences of effector protein complex formation in a diiron hydroxylase, *Proc. Natl. Acad. Sci. USA 105*, 19194-19198.

[54] Nordlund, P., Sjöberg, B. M., and Eklund, H. (1990) Three-dimensional structure of the free radical protein of ribonucleotide reductase, *Nature 345*, 593-598.

[55] Fox, B. G., Shanklin, J., Somerville, C., and Münck, E. (1993) Stearoyl-acyl carrier protein Δ^9 desaturase from *Ricinus communis* is a diiron-oxo protein, *Proc. Natl. Acad. Sci. USA. 90*, 2486-2490.

[56] Simurdiak, M., Lee, J., and Zhao, H. (2006) A new class of arylamine oxygenases: evidence that p-aminobenzoate N-oxygenase (AurF) is a di-iron enzyme and further mechanistic studies, *ChemBioChem 7*, 1169-1172.

[57] Makris, T. M., Vu, V. V., Meier, K. K., Komor, A. J., Rivard, B. S., Münck, E., Que, L., Jr., and Lipscomb, J. D. (2015) An unusual peroxo intermediate of the arylamine oxygenase of the chloramphenicol biosynthetic pathway, *J. Am. Chem. Soc. 137*, 1608-1617.

[58] Makris, T. M., Chakrabarti, M., Münck, E., and Lipscomb, J. D. (2010) A family of diiron monooxygenases catalyzing amino acid beta-hydroxylation in antibiotic biosynthesis, *Proc. Natl. Acad. Sci. USA 107*, 15391-15396.

[59] Xing, G., Barr, E. W., Diao, Y., Hoffart, L. M., Prabhu, K. S., Arner, R. J., Reddy, C. C., Krebs, C., and Bollinger, J. M., Jr. (2006) Oxygen activation by a mixed-valent, diiron(II/III) cluster in the glycol cleavage reaction catalyzed by myo-inositol oxygenase, *Biochemistry 45*, 5402-5412.

[60] Vu, V. V., Emerson, J. P., Martinho, M., Kim, Y. S., Münck, E., Park, M. H., and Que, L., Jr. (2009) Human deoxyhypusine hydroxylase, an enzyme involved in regulating cell growth, activates O_2 with a nonheme diiron center, *Proc. Natl. Acad. Sci. USA 106*, 14814-14819.

[61] Aukema, K. G., Makris, T. M., Stoian, S. A., Richman, J. E., Münck, E., Lipscomb, J. D., and Wackett, L. P. (2013) Cyanobacterial aldehyde deformylase oxygenation of aldehydes yields n - 1 aldehydes and alcohols in addition to alkanes, *ACS Catal. 3*, 2228-2238.

[62] Mathevon, C., Pierrel, F., Oddou, J.-L., Garcia-Serres, R., Blondin, G., Latour, J.-M., Menage, S., Gambarelli, S., Fontecave, M., and Atta, M. (2007) tRNA-modifying MiaE protein from *Salmonella typhimurium* is a nonheme diiron monooxygenase, *Proc. Natl. Acad. Sci. USA 104*, 13295-13300.

[63] Ray, K., Pfaff, F. F., Wang, B., and Nam, W. (2014) Status of reactive non-heme metal–oxygen intermediates in chemical and enzymatic reactions, *J. Am. Chem. Soc. 136*, 13942-13958.

[64] Mbughuni, M. M., Chakrabarti, M., Hayden, J. A., Bominaar, E. L., Hendrich, M. P., Münck, E., and Lipscomb, J. D. (2010) Trapping and spectroscopic characterization of an Fe^{III}-superoxo intermediate from a nonheme mononuclear iron-containing enzyme, *Proc. Natl. Acad. Sci. USA 107*, 16788-16793.

[65] Crawford, J. A., Li, W., and Pierce, B. S. (2011) Single turnover of substrate-bound ferric cysteine dioxygenase with superoxide anion: Enzymatic reactivation, product formation, and a transient Intermediate, *Biochemistry 50*, 10241-10253.

[66] Brown, C. D., Neidig, M. L., Neibergall, M. B., Lipscomb, J. D., and Solomon, E. I. (2007) VTVH-MCD and DFT studies of thiolate bonding to $\{FeNO\}^7/\{FeO_2\}^8$ complexes of isopenicillin N synthase: Substrate determination of oxidase versus oxygenase activity in nonheme Fe enzymes, *J. Am. Chem. Soc. 129*, 7427-7438.

[67] Rocklin, A. M., Kato, K., Liu, H.-w., Que, L., Jr., and Lipscomb, J. D. (2004) Mechanistic studies of 1-aminocyclopropane-1-carboxylic acid oxidase: single turnover reaction, *J. Biol. Inorg. Chem. 9*, 171-182.

[68] Zhu, H., Peck, S. C., Bonnot, F., van der Donk, W. A., and Klinman, J. P. (2015) Oxygen-18 kinetic isotope effects of nonheme iron enzymes HEPD and MPnS support iron(III) superoxide as the hydrogen abstraction species, *J. Am. Chem. Soc. 137*, 10448-10451.

[69] Bollinger, J. M., Jr., Diao, Y., Matthews, M. L., Xing, G., and Krebs, C. (2009) myo-Inositol oxygenase: a radical new pathway for O_2 and C-H activation at a nonheme diiron cluster, *Dalton Trans.*, 905-914.

[70] Rivard, B. S., Rogers, M. S., Marell, D. J., Neibergall, M. B., Chakrabarty, S., Cramer, C. J., and Lipscomb, J. D. (2015) Rate-determining attack on substrate precedes Rieske cluster oxidation during cis-dihydroxylation by benzoate dioxygenase, *Biochemistry 54*, 4652-4664.

[71] Mbughuni, M. M., Chakrabarti, M., Hayden, J. A., Meier, K. K., Dalluge, J. J., Hendrich, M. P., Münck, E., and Lipscomb, J. D. (2011) Oxy-intermediates of homoprotocatechuate 2,3-dioxygenase: Facile electron transfer between substrates, *Biochemistry 50*, 10262-10274.

[72] Kovaleva, E. G., and Lipscomb, J. D. (2007) Crystal structures of Fe^{2+} dioxygenase superoxo, alkylperoxo, and bound product intermediates, *Science 316*, 453-457.

[73] Korboukh, V. K., Li, N., Barr, E. W., Bollinger, J. M., Jr., and Krebs, C. (2009) A long-lived, substrate-hydroxylating peroxodiiron(III/III) intermediate in the amine oxygenase, AurF, from *Streptomyces thioluteus*, *J. Am. Chem. Soc. 131*, 13608-13609.

[74] Pandelia, M. E., Li, N., Noergaard, H., Warui, D. M., Rajakovich, L. J., Chang, W.-c., Booker, S. J., Krebs, C., and Bollinger, J. M., Jr. (2013) Substrate-triggered addition of dioxygen to the diferrous cofactor of aldehyde-deformylating oxygenase to form a diferric-peroxide intermediate, *J. Am. Chem. Soc. 135*, 15801-15812.

[75] Liu, K. E., Valentine, A. M., Qiu, D., Edmondson, D. E., Appelman, E. H., Spiro, T. G., and Lippard, S. J. (1995) Characterization of a diiron(III) peroxo intermediate in the reaction cycle of methane monooxygenase hydroxylase from *Methylococcus capsulatus* (Bath) (withdrawn), *J. Am. Chem. Soc. 117*, 4997-4998.

[76] Han, Z., Sakai, N., Boettger, L. H., Klinke, S., Hauber, J., Trautwein, A. X., and Hilgenfeld, R. (2015) Crystal structure of the peroxo-diiron(III) intermediate of deoxyhypusine hydroxylase, an oxygenase involved in hypusination, *Structure 23*, 882-892.

[77] Rajakovich, L. J., Nørgaard, H., Warui, D. M., Chang, W.-c., Li, N., Booker, S. J., Krebs, C., Bollinger, J. M., and Pandelia, M.-E. (2015) Rapid reduction of the diferric-peroxyhemiacetal intermediate in aldehyde-deformylating oxygenase by a cyanobacterial ferredoxin: Evidence for a free-radical mechanism, *J. Am. Chem. Soc. 137*, 11695-11709.

[78] Skulan, A., J., Brunold, T., C., Baldwin, J., Saleh, L., Bollinger, J. M., Jr., and Solomon, E., I. (2004) Nature of the peroxo intermediate of the W48F/D84E ribonucleotide reductase variant: implications for O_2 activation by binuclear non-heme iron enzymes, *J. Am. Chem. Soc. 126*, 8842-8855.

[79] Banerjee, R., Meier, K. K., Münck, E., and Lipscomb, J. D. (2013) Intermediate P* from soluble methane monooxygenase contains a diferrous cluster, *Biochemistry 52*, 4331-4342.

[80] Bailey, L. J., and Fox, B. G. (2009) Crystallographic and catalytic studies of the peroxide-shunt reaction in a diiron hydroxylase, *Biochemistry 48*, 8932-8939.

[81] Broadwater, J. A., Ai, J., Loehr, T. M., Sanders-Loehr, J., and Fox, B. G. (1998) Peroxodiferric Intermediate of stearoyl-acyl carrier protein Δ^9 desaturase: Oxidase reactivity during single turnover and implications for the mechanism of desaturation, *Biochemistry 37*, 14664-14671.

[82] Jeoung, J.-H., Bommer, M., Lin, T.-Y., and Dobbek, H. (2013) Visualizing the substrate-, superoxo-, alkylperoxo-, and product-bound states at the nonheme Fe(II) site of homogentisate dioxygenase, *Proc. Natl. Acad. Sci. USA 110*, 12625-12630.

[83] Neibergall, M. B., Stubna, A., Mekmouche, Y., Münck, E., and Lipscomb, J. D. (2007) Hydrogen peroxide dependent *cis*-dihydroxylation of benzoate by fully oxidized benzoate 1,2-dioxygenase, *Biochemistry 46*, 8004-8016.

[84] Blaesi, E. J., Gardner, J. D., Fox, B. G., and Brunold, T. C. (2013) Spectroscopic and computational characterization of the NO adduct of substrate-bound Fe(II) cysteine dioxygenase: Insights into the mechanism of O_2 activation, *Biochemistry 52*, 6040-6051.

[85] Cicchillo, R. M., Zhang, H., Blodgett, J. A., Whitteck, J. T., Li, G., Nair, S. K., van der Donk, W. A., and Metcalf, W. W. (2009) An unusual carbon-carbon bond cleavage reaction during phosphinothricin biosynthesis, *Nature 459*, 871-874.

[86] Price, J. C., Barr, E. W., Tirupati, B., Bollinger, J. M., Jr., and Krebs, C. (2003) The first direct characterization of a high-valent iron intermediate in the reaction of an alpha-ketoglutarate-dependent dioxygenase: a high-spin Fe^{IV} complex in taurine/alpha-ketoglutarate dioxygenase (TauD) from *Escherichia coli*, *Biochemistry 42*, 7497-7508.

[87] Peck, S. C., Cooke, H. A., Cicchillo, R. M., Malova, P., Hammerschmidt, F., Nair, S. K., and van der Donk, W. A. (2011) Mechanism and substrate recognition of 2-hydroxyethylphosphonate dioxygenase, *Biochemistry 50*, 6598-6605.

[88] Wang, C., Chang, W.-c., Guo, Y., Huang, H., Peck, S. C., Pandelia, M. E., Lin, G.-m., Liu, H.-w., Krebs, C., and Bollinger, J. M., Jr. (2013) Evidence that the fosfomycin-producing epoxidase, HppE, Is a non-heme-iron peroxidase, *Science 342*, 991-995.

[89] Huang, H., Chang, W.-c., Pai, P.-J., Romo, A., Mansoorabadi, S. O., Russell, D. H., and Liu, H.-w. (2012) Evidence for radical-mediated catalysis by HppE: a study using cyclopropyl and methylenecyclopropyl substrate analogues., *J. Am. Chem. Soc. 134*, 16171-16174.

[90] Mirica, L. M., McCusker, K. P., Munos, J. W., Liu, H.-w., and Klinman, J. P. (2008) ^{18}O kinetic isotope effects in non-heme iron enzymes: probing the nature of Fe/O_2 intermediates, *J. Am. Chem. Soc. 130*, 8122-8123.

[91] Harpel, M. R., and Lipscomb, J. D. (1990) Gentisate 1,2-dioxygenase from Pseudomonas. Substrate coordination to active site Fe^{2+} and mechanism of turnover, *J. Biol. Chem. 265,* 22187-22196.

[92] Sanvoisin, J., Langley, G. J., and Bugg, T. D. H. (1995) Mechanism of extradiol catechol dioxygenases: Evidence for a lactone intermediate in the 2,3-dihydroxyphenylpropionate 1,2-dioxygenase reaction, *J. Am. Chem. Soc. 117*, 7836-7837.

[93] Hoffart, L. M., Barr, E. W., Guyer, R. B., Bollinger, J. M., Jr., and Krebs, C. (2006) Direct spectroscopic detection of a C-H-cleaving high-spin Fe(IV) complex in a prolyl-4-hydroxylase, *Proc. Natl. Acad. Sci. USA 103*, 14738-14743.

[94] Diebold, A. R., Straganz, G. D., and Solomon, E. I. (2011) Spectroscopic and computational studies of α-keto acid binding to Dke1: Understanding the role of the facial triad and the reactivity of β-diketones, *J. Am. Chem. Soc. 133*, 15979-15991.

[95] Whitteck, J. T., Cicchillo, R. M., and van der Donk, W. A. (2009) Hydroperoxylation by hydroxyethylphosphonate dioxygenase., *J. Am. Chem. Soc. 131*, 16225-16232.

[96] Solomon, E. I., Light, K. M., Liu, L. V., Srnec, M., and Wong, S. D. (2013) Geometric and electronic structure contributions to function in non-heme iron enzymes, *Acc. Chem. Res. 46*, 2725-2739.

[97] Groves, J. T. (2006) High-valent iron in chemical and biological oxidations, *J. Inorg. Biochem. 100*, 434-447.

[98] Banerjee, R., Proshlyakov, Y., Lipscomb, J. D., and Proshlyakov, D. A. (2015) Structure of the key species in the enzymatic oxidation of methane to methanol, *Nature 518*, 431-434.

[99] Parkin, S. E., Chen, S., Ley, B. A., Mangravite, L., Edmondson, D. E., Huynh, B. H., and Bollinger, J. M., Jr. (1998) Electron injection through a specific pathway determines the outcome of oxygen activation at the diiron cluster in the F208Y mutant of *Escherichia coli* ribonucleotide reductase protein R2, *Biochemistry 37*, 1124-1130.

[100] Srnec, M., Rokob, T. A., Schwartz, J. K., Kwak, Y., Rulisek, L., and Solomon, E. I. (2012) Structural and spectroscopic properties of the peroxodiferric intermediate of *Ricinus communis* soluble Δ^9 desaturase, *Inorg. Chem. 51*, 2806-2820.

[101] Lee, S. K., and Lipscomb, J. D. (1999) Oxygen activation catalyzed by methane monooxygenase hydroxylase component: Proton delivery during the O-O bond cleavage steps, *Biochemistry. 38*, 4423-4432.

[102] Tinberg, C. E., and Lippard, S. J. (2009) Revisiting the mechanism of dioxygen activation in soluble methane monooxygenase from *M. capsulatus* (Bath): Evidence for a multi-step, proton-dependent reaction pathway, *Biochemistry 48*, 12145-12158.

[103] Lee, S.-K., Fox, B. G., Froland, W. A., Lipscomb, J. D., and Münck, E. (1993) A transient intermediate of the methane monooxygenase catalytic cycle containing a $Fe^{IV}Fe^{IV}$ cluster, *J. Am. Chem. Soc. 115*, 6450-6451.

[104] Shu, L. J., Liu, Y., Lipscomb, J. D., and Que, L., Jr. (1996) X-ray absorption spectroscopic studies of the methane monooxygenase hydroxylase component from *Methylosinus trichosporium* OB3b, *J. Biol. Inorg. Chem. 1*, 297-304.

[105] Sturgeon, B. E., Burdi, D., Chen, S., Huynh, B.-H., Edmondson, D. E., Stubbe, J., and Hoffman, B. M. (1996) Reconsideration of **X**, the diiron intermediate formed during cofactor assembly in *E. coli* ribonucleotide reductase, *J. Am. Chem. Soc. 118*, 7551-7557.

[106] Eser, B. E., Barr, E. W., Frantom, P. A., Saleh, L., Bollinger, J. M., Jr., Krebs, C., and Fitzpatrick, P. F. (2007) Direct spectroscopic evidence for a high-spin Fe(IV) intermediate in tyrosine hydroxylase, *J. Am. Chem. Soc. 129*, 11334-11335.

[107] Wong, S. D., Srnec, M., Matthews, M. L., Liu, L. V., Kwak, Y., Park, K., Bell, C. B., III, Alp, E. E., Zhao, J., Yoda, Y., Kitao, S., Seto, M., Krebs, C., Bollinger, J. M., Jr., and Solomon, E. I. (2013) Elucidation of the Fe(IV)=O intermediate in the catalytic cycle of the halogenase SyrB2, *Nature 499*, 320-323.

[108] McDonald, A. R., and Que Jr, L. (2013) High-valent nonheme iron-oxo complexes: Synthesis, structure, and spectroscopy, *Coord. Chem. Rev. 257*, 414-428.

[109] Wallar, B. J., and Lipscomb, J. D. (1996) Dioxygen activation by enzymes containing binuclear non-heme iron clusters, *Chem. Rev. 96*, 2625-2657.

[110] Tinberg, C. E., and Lippard, S. J. (2011) Dioxygen activation in soluble methane monooxygenase, *Acc. Chem. Res. 44*, 280-288.

[111] Nesheim, J. C., and Lipscomb, J. D. (1996) Large isotope effects in methane oxidation catalyzed by methane monooxygenase: Evidence for C−H bond cleavage in a reaction cycle intermediate, *Biochemistry 35*, 10240-10247.

[112] Jin, Y., and Lipscomb, J. D. (1999) Probing the mechanism of C-H activation: oxidation of methylcubane by soluble methane monooxygenase from *Methylosinus trichosporium* OB3b, *Biochemistry 38*, 6178-6186.

[113] Choi, S.-Y., Eaton, P. E., Hollenberg, P. F., Liu, K. E., Lippard, S. J., Newcomb, M., Putt, D. A., Upadhyaya, S. P., and Xiong, Y. (1996) Regiochemical variations in reactions of metylcubane with *tert*-butoxyl radical, cytochrome P-450 enzymes, and a methane monooxygenase system, *J. Am. Chem. Soc. 118*, 6547-6555.

[114] Bollinger, J. M., Jr., Edmondson, D. E., Huynh, B. H., Filley, J., Norton, J. R., and Stubbe, J. (1991) Mechanism of assembly of the tyrosyl radical-dinuclear iron cluster cofactor of ribonucleotide reductase, *Science 253*, 292-298.

[115] Ravi, N., Bollinger, J. M., Jr., Huynh, B. H., Edmondson, D. E., and Stubbe, J. (1994) Mechanism of assembly of the tyrosyl radical-diiron(III) cofactor of *E. coli* ribonucleotide reductase. 1. Mössbauer characterization of the diferric radical precursor, *J. Am. Chem. Soc. 116*, 8007-8014.

[116] Bollinger, J. M., Jr., Tong, W. H., Ravi, N., Huynh, B. H., Edmondson, D. E., and Stubbe, J. (1994) Mechanism of assembly of the tyrosyl radical-diiron(III) cofactor of *E. coli* ribonucleotide reductase. 2. Kinetics of the excess Fe^{2+} reaction by optical, EPR, and Mössbauer spectroscopies, *J. Am. Chem. Soc. 116*, 8015-8023.

[117] Bollinger, J. M., Jr., Tong, W. H., Ravi, N., Huynh, B. H., Edmondson, D. E., and Stubbe, J. (1994) Mechanism of assembly of the tyrosyl radical-diiron(III) cofactor of *E. coli*

ribonucleotide reductase. 3. Kinetics of the limiting Fe^{2+} reaction by optical, EPR, and Mössbauer spectroscopies, *J. Am. Chem. Soc. 116*, 8024-8032.

[118] Tong, W. H., Chen, S., Lloyd, S. G., Edmondson, D. E., Huynh, B.-H., and Stubbe, J. (1996) Mechanism of assembly of the diferric cluster-tyrosyl radical cofactor of *E. coli* ribonucleotide reductase from the diferrous form of the R2 subunit, *J. Am. Chem. Soc. 118*, 2107-2108.

[119] Baldwin, J., Krebs, C., Saleh, L., Stelling, M., Huynh, B. H., Bollinger, J. M., Jr., and Riggs-Gelasco, P. (2003) Structural characterization of the peroxodiiron(III) intermediate generated during oxygen activation by the W48A/D84E variant of ribonucleotide reductase protein R2 from *Escherichia coli*, *Biochemistry 42*, 13269-13279.

[120] Yun, D., Garcia-Serres, R., Chicalese, B. M., An, Y. H., Huynh, B. H., and Bollinger, J. M., Jr. (2007) (μ-1,2-peroxo)diiron(III/III) complex as a precursor to the diiron(III/IV) intermediate X in the assembly of the iron-radical cofactor of ribonucleotide reductase from mouse, *Biochemistry 46*, 1925-1932.

[121] Bollinger, J. M., Jr., Stubbe, J., Huynh, B. H., and Edmondson, D. E. (1991) Novel diferric radical intermediate responsible for tyrosyl radical formation in assembly of the cofactor of ribonucleotide reductase, *J. Am. Chem. Soc. 113*, 6289-6291.

[122] Dassama, L. M. K., Silakov, A., Krest, C. M., Calixto, J. C., Krebs, C., Bollinger, J. M., Jr., and Green, M. T. (2013) A 2.8 Å Fe-Fe separation in the Fe2III/IV intermediate (X) from *Escherichia coli* ribonucleotide reductase, *J. Am. Chem. Soc.*

[123] Åberg, A., Ormö, M., Nordlund, P., and Sjöberg, B.-M. (1993) Autocatalytic generation of dopa in the engineered protein R2 F208Y from *Escherichia coli* ribonucleotide reductase and crystal structure of the DOPA-208 protein, *Biochemistry 32*, 9845-9850.

[124] Panay, A. J., Lee, M., Krebs, C., Bollinger, J. M., Jr., and Fitzpatrick, P. F. (2011) Evidence for a high-spin Fe(IV) species in the catalytic cycle of a bacterial phenylalanine hydroxylase, *Biochemistry 50*, 1928-1933.

[125] Kovaleva, E. G., and Lipscomb, J. D. (2008) Versatility of biological non-heme Fe(II) centers in oxygen activation reactions, *Nat. Chem. Biol. 4*, 186-193.

[126] Price, J. C., Barr, E. W., Hoffart, L. M., Krebs, C., and Bollinger, J. M., Jr. (2005) Kinetic dissection of the catalytic mechanism of taurine:alpha-ketoglutarate dioxygenase (TauD) from *Escherichia coli*, *Biochemistry 44*, 8138-8147.

[127] Barry, S., and Challis, G. L. (2013) Mechanism and catalytic diversity of Rieske non-heme iron-dependent oxygenases, *ACS Catal. 3*, 2362–2370.

[128] Wolfe, M. D., and Lipscomb, J. D. (2003) Hydrogen peroxide-coupled *cis*-diol formation catalyzed by naphthalene 1,2-dioxygenase, *J. Biol. Chem. 278*, 829-835.

[129] Kurtz, D. M., Jr. (2006) Avoiding high-valent iron intermediates: superoxide reductase and rubrerythrin, *J. Inorg. Biochem. 100*, 679-693.

[130] Liu, P., Mehn, M. P., Yan, F., Zhao, Z., Que, L., Jr., and Liu, H.-w. (2004) Oxygenase activity in the self-hydroxylation of (S)-2-hydroxypropylphosphonic acid epoxidase involved in fosfomycin biosynthesis, *J. Am. Chem. Soc. 126*, 10306-10312.

[131] Liu, A., Ho, R. Y. N., Que, L., Jr., Ryle, M. J., Phinney, B. S., and Hausinger, R. P. (2001) Alternative reactivity of an α-ketoglutarate-dependent iron(II) oxygenase: enzyme self-hydroxylation, *J. Am. Chem. Soc. 123*, 5126-5127.

[132] Ryle, M. J., Koehntop, K. D., Liu, A., Que, L., Jr., and Hausinger, R. P. (2003) Interconversion of two oxidized forms of taurine/alpha-ketoglutarate dioxygenase, a non-heme iron hydroxylase: Evidence for bicarbonate binding, *Proc. Natl. Acad. Sci. USA. 100*, 3790-3795.

[133] Farquhar, E., R., Koehntop, K., D., Emerson, J., P., and Que, L., Jr. (2005) Post-translational self-hydroxylation: a probe for oxygen activation mechanisms in non-heme iron enzymes, *Biochem. Biophys. Res. Commun. 338*, 230-239.

[134] Koehntop, K. D., Emerson, J. P., and Que, L., Jr. (2005) The 2-His-1-carboxylate facial triad: a versatile platform for dioxygen activation by mononuclear non-heme iron(II) enzymes, *J. Biol. Inorg. Chem. 10*, 87-93.

[135] Buongiorno, D., and Straganz, G. D. (2013) Structure and function of atypically coordinated enzymatic mononuclear non-heme-Fe(II) centers, *Coord. Chem. Rev. 257*, 541-563.

[136] Ho, R. Y. N., Mehn, M. P., Hegg, E. L., Liu, A., Ryle, M. J., Hausinger, R. P., and Que, L., Jr. (2001) Resonance Raman studies of the iron(II)-alpha-keto acid chromophore in model and enzyme complexes, *J. Am. Chem. Soc. 123*, 5022-5029.

[137] Loeb, K. E., Westre, T. E., Kappock, T. J., Mitic, N., Glasfeld, E., Caradonna, J. P., Hedman, B., Hodgson, K. O., and Solomon, E. I. (1997) Spectroscopic characterization of the catalytically competent ferrous site of the resting, activated, and substrate-bound forms of phenylalanine hydroxylase, *J. Am. Chem. Soc. 119*, 1901-1915.

[138] Gardner, J. D., Pierce, B. S., Fox, B. G., and Brunold, T. C. (2010) Spectroscopic and computational characterization of substrate-bound mouse cysteine dioxygenase: Nature of the ferrous and ferric cysteine adducts and mechanistic implications, *Biochemistry 49*, 6033-6041.

[139] Sato, N., Uragami, Y., Nishizaki, T., Takahashi, Y., Sazaki, G., Sugimoto, K., Nonaka, T., Masai, E., Fukuda, M., and Senda, T. (2002) Crystal structures of the reaction intermediate and its homologue of an extradiol-cleaving catecholic dioxygenase, *J. Mol. Biol. 321*, 621-636.

[140] Karlsson, A., Parales, J. V., Parales, R. E., Gibson, D. T., Eklund, H., and Ramaswamy, S. (2003) Crystal structure of naphthalene dioxygenase: Side-on binding of dioxygen to iron, *Science 299*, 1039-1042.

[141] Straganz, G. D., Hofer, H., Steiner, W., and Nidetzky, B. (2004) Electronic substituent effects on the cleavage specificity of a non-heme Fe^{2+}-dependent β-diketone dioxygenase and their mechanistic implications, *J. Am. Chem. Soc. 126*, 12202-12203.

[142] Roach, P. L., Clifton, I. J., Hensgens, C. M., Shibata, N., Schofield, C. J., Hajdu, J., and Baldwin, J. E. (1997) Structure of isopenicillin N synthase complexed with substrate and the mechanism of penicillin formation, *Nature 387*, 827-830.

[143] Yan, F., Moon, S.-J., Liu, P., Zhao, Z., Lipscomb, J. D., Liu, A., and Liu, H.-w. (2007) Determination of the substrate binding mode to the active site iron of (S)-2-hydroxypropylphosphonic acid epoxidase using ^{17}O-enriched substrates and substrate analogues, *Biochemistry 46*, 12628-12638.

[144] Yang, T.-C., Wolfe, M. D., Neibergall, M. B., Mekmouche, Y., Lipscomb, J. D., and Hoffman, B. M. (2003) Modulation of substrate binding to naphthalene 1,2-dioxygenase by Rieske cluster reduction/oxidation, *J. Am. Chem. Soc. 125*, 2034-2035.

[145] Arciero, D. M., Orville, A. M., and Lipscomb, J. D. (1985) [^{17}O]Water and nitric oxide binding by protocatechuate 4,5-dioxygenase and catechol 2,3-dioxygenase. Evidence for binding of exogenous ligands to the active site Fe^{2+} of extradiol dioxygenases, *J. Biol. Chem. 260*, 14035-14044.

[146] Arciero, D. M., and Lipscomb, J. D. (1986) Binding of ^{17}O-labeled substrate and inhibitors to protocatechuate 4,5-dioxygenase-nitrosyl complex. Evidence for direct substrate binding to the active site Fe^{2+} of extradiol dioxygenases, *J. Biol. Chem. 261*, 2170-2178.

[147] Pierce, B. S., Gardner, J. D., Bailey, L. J., Brunold, T. C., and Fox, B. G. (2007) Characterization of the nitrosyl adduct of substrate-bound mouse cysteine dioxygenase by

electron paramagnetic resonance: Electronic structure of the active site and mechanistic implications, *Biochemistry 46*, 8569-8578.

[148] Yun, D., Dey, M., Higgins, L. J., Yan, F., Liu, H.-w., and Drennan, C. L. (2011) Structural basis of regiospecificity of a mononuclear iron enzyme in antibiotic fosfomycin biosynthesis, *J. Am. Chem. Soc. 133*, 11262-11269.

[149] Straganz, G. D., and Nidetzky, B. (2005) Reaction coordinate analysis for beta-diketone cleavage by the non-heme Fe^{2+}-dependent dioxygenase Dke1, *J. Am. Chem. Soc. 127*, 12306-12314.

[150] Froland, W. A., Andersson, K. K., Lee, S.-K., Liu, Y., and Lipscomb, J. D. (1992) Methane monooxygenase component B and reductase alter the regioselectivity of the hydroxylase component-catalyzed reactions. A novel role for protein-protein interactions in an oxygenase mechanism, *J. Biol. Chem. 267*, 17588-17597.

[151] Dalton, H. (1980) Oxidation of hydrocarbons by methane monooxygenase from a variety of microbes, *Adv. Appl. Microbiol. 26*, 71-87.

[152] Higgins, I. J., Best, D. J., and Hammond, R. C. (1980) New findings in methane-utilizing bacteria highlight their importance in the biosphere and their commercial potential, *Nature 286*, 561-564.

[153] Liu, K. E., and Lippard, S. J. (1995) Studies of the soluble methane monooxygenase protein system: Structure, component interactions, and hydroxylation mechanism, *Adv. Inorg. Chem. 42*, 263-289.

[154] Fox, B. G., Hendrich, M. P., Surerus, K. K., Andersson, K. K., Froland, W. A., Lipscomb, J. D., and Münck, E. (1993) Mössbauer, EPR, and ENDOR studies of the hydroxylase and reductase components of methane monooxygenase from *Methylosinus trichosporium* OB3b, *J. Am. Chem. Soc. 115*, 3688-3701.

[155] Paulsen, K. E., Stankovich, M. T., Liu, Y., Lipscomb, J. D., Fox, B. G., and Münck, E. (1993) EPR-spectroelectrochemical characterization of methane monooxygenase hydroxylase and regulation by components B, *Proc. - Electrochem. Soc. 93-11*, 118-124.

[156] Coufal, D. E., Tavares, P., Pereira, A. S., Hyunh, B. H., and Lippard, S. J. (1999) Reactions of nitric oxide with the reduced non-heme diiron center of the soluble methane monooxygenase hydroxylase, *Biochemistry 38*, 4504-4513.

[157] Liu, Y., Nesheim, J. C., Lee, S.-K., and Lipscomb, J. D. (1995) Gating effects of component B on oxygen activation by the methane monooxygenase hydroxylase component, *J. Biol. Chem. 270*, 24662-24665.

[158] Wallar, B. J., and Lipscomb, J. D. (2001) Methane monooxygenase component B mutants alter the kinetics of steps throughout the catalytic cycle, *Biochemistry 40*, 2220-2233.

[159] Zhang, J., Wallar, B. J., Popescu, C. V., Renner, D. B., Thomas, D. D., and Lipscomb, J. D. (2006) Methane monooxygenase hydroxylase and B component interactions, *Biochemistry 45*, 2913-2926.

[160] Zheng, H., and Lipscomb, J. D. (2006) Regulation of methane monooxygenase catalysis based on size exclusion and quantum tunneling, *Biochemistry 45*, 1685-1692.

[161] Paulsen, K. E., Liu, Y., Fox, B. G., Lipscomb, J. D., Münck, E., and Stankovich, M. T. (1994) Oxidation-reduction potentials of the methane monooxygenase hydroxylase component from *Methylosinus trichosporium* OB3b, *Biochemistry 33*, 713-722.

[162] Lee, S. J., McCormick, M. S., Lippard, S. J., and Cho, U.-S. (2013) Control of substrate access to the active site in methane monooxygenase, *Nature 494*, 380-384.

[163] Lambalot, R. H., Gehring, A. M., Flugel, R. S., Zuber, P., LaCelle, M., Marahiel, M. A., Reid, R., Khosla, C., and Walsh, C. T. (1996) A new enzyme superfamily - the phosphopantetheinyl transferases, *Chem. Biol. 3*, 923-936.

[164] Byers, D. M., and Gong, H. (2007) Acyl carrier protein: structure–function relationships in a conserved multifunctional protein family, *Biochem. Cell Biol. 85*, 649-662.

[165] Lindqvist, Y., Huang, W. J., Schneider, G., and Shanklin, J. (1996) Crystal structure of delta(9) stearoyl-acyl carrier protein desaturase from castor seed and its relationship to other diiron proteins, *EMBO 15*, 4081-4092.

[166] Yang, Y.-S., Broadwater, J. A., Pulver, S. C., Fox, B. G., and Solomon, E. I. (1999) Circular dichroism and magnetic circular dichroism studies of the reduced binuclear non-heme iron site of stearoyl-ACP D^9-desaturase: substrate binding and comparison to ribonucleotide reductase, *J. Am. Chem. Soc. 121*, 2770-2783.

[167] Davydov, R., Behrouzian, B., Smoukov, S., Stubbe, J., Hoffman, B. M., and Shanklin, J. (2005) Effect of substrate on the diiron(III) site in stearoyl acyl carrier protein delta 9-desaturase as disclosed by cryoreduction electron paramagnetic resonance/electron nuclear double resonance spectroscopy., *Biochemistry 44*, 1309-1315.

[168] Guy, J. E., Whittle, E., Moche, M., Lengqvist, J., Lindqvist, Y., and Shanklin, J. (2011) Remote control of regioselectivity in acyl-acyl carrier protein-desaturases., *Proc. Natl. Acad. Sci. USA. 108*, 16594-16599.

[169] Nelson, M. J., and Seitz, S. P. (1994) The structure and function of lipoxygenase, *Curr. Opin. Struct. Biol. 4*, 878-884.

[170] Ornston, L. N., and Stanier, R. Y. (1964) Mechanism of β-ketoadipate formation by bacteria, *Nature 204,* 1279-1283.

[171] Harwood, C. S., and Parales, R. E. (1996) The beta-ketoadipate pathway and the biology of self-identity, *Ann. Rev. Microbiol. 50*, 553-590.

[172] Ju, K.-S., and Parales, R. E. (2010) Nitroaromatic compounds, from synthesis to biodegradation, *Microbiol. Mol. Biol. Rev. 74*, 250-272.

[173] Wood, J. M., Crawford, R. L., Münck, E., Zimmermann, R., Lipscomb, J. D., Stephens, R. S., Bromley, J. W., Que, L., Jr., Howard, J. B., and Orme-Johnson, W. H. (1977) Structure and function of dioxygenases. One approach to lignin degradation, *J. Agric. Food Chem. 25*, 698-704.

[174] Bugg, T. D. H., Ahmad, M., Hardiman, E. M., and Rahmanpour, R. (2011) Pathways for degradation of lignin in bacteria and fungi, *Nat. Prod. Rep. 28*, 1883-1896.

[175] Olaniran, A. O., and Igbinosa, E. O. (2011) Chlorophenols and other related derivatives of environmental concern: Properties, distribution and microbial degradation processes, *Chemosphere 83*, 1297-1306.

[176] Bianchetti, C. M., Harmann, C. H., Takasuka, T. E., Hura, G. L., Dyer, K., and Fox, B. G. (2013) Fusion of dioxygenase and lignin-binding domains in a novel secreted enzyme from cellulolytic Streptomyces sp. SirexAA-E., *J. Biol. Chem. 288,* 18574-18587.

[177] Wells, T., Jr., and Ragauskas, A. J. (2012) Biotechnological opportunities with the β-ketoadipate pathway, *Trends Biotechnol. 30,* 627-637.

[178] Nikel, P. I., Martinez-Garcia, E., and de Lorenzo, V. (2014) Biotechnological domestication of pseudomonads using synthetic biology, *Nat. Rev. Micro. 12*, 368-379.

[179] Frazee, R. W., Orville, A. M., Dolbeare, K. B., Yu, H., Ohlendorf, D. H., and Lipscomb, J. D. (1998) The axial tyrosinate Fe^{3+} ligand in protocatechuate 3,4-dioxygenase influences substrate binding and product release: Evidence for new reaction cycle intermediates, *Biochemistry 37*, 2131-2144.

[180] Elgren, T. E., Orville, A. M., Kelly, K. A., Lipscomb, J. D., Ohlendorf, D. H., and Que, L., Jr. (1997) Crystal structure and resonance Raman studies of protocatechuate 3,4-dioxygenase complexed with 3,4-dihydroxyphenylacetate, *Biochemistry 36*, 11504-11513.

[181] Orville, A. M., Lipscomb, J. D., and Ohlendorf, D. H. (1997) Crystal structures of substrate and substrate analog complexes of protocatechuate 3,4-dioxygenase: endogenous Fe^{3+} ligand displacement in response to substrate binding, *Biochemistry 36*, 10052-10066.

[182] Orville, A. M., Elango, N., Lipscomb, J. D., and Ohlendorf, D. H. (1997) Structures of competitive inhibitor complexes of protocatechuate 3,4-dioxygenase: multiple exogenous ligand binding orientations within the active site, *Biochemistry 36*, 10039-10051.

[183] Walsh, T. A., Ballou, D. P., Mayer, R., and Que, L., Jr. (1983) Rapid reaction studies on the oxygenation reactions of catechol dioxygenase, *J. Biol. Chem. 258*, 14422-14427.

[184] Bull, C., Ballou, D. P., and Otsuka, S. (1981) The reaction of oxygen with protocatechuate 3,4-dioxygenase from *Pseudomonas putida*. Characterization of a new oxygenated intermediate, *J. Biol. Chem. 256*, 12681-12686.

[185] Horsman, G. P., Jirasek, A., Vaillancourt, F. H., Barbosa, C. J., Jarzecki, A. A., Xu, C., Mekmouche, Y., Spiro, T. G., Lipscomb, J. D., Blades, M. W., Turner, R. F. B., and Eltis, L. D. (2005) Spectroscopic studies of the anaerobic enzyme-substrate complex of catechol 1,2-dioxygenase, *J. Am. Chem. Soc. 127*, 16882-16891.

[186] Davis, M. I., Orville, A. M., Neese, F., Zaleski, J. M., Lipscomb, J. D., and Solomon, E. I. (2002) Spectroscopic and electronic structure studies of protocatechuate 3,4-dioxygenase: Nature of tyrosinate-Fe(III) bonds and their contribution to reactivity, *J. Am. Chem. Soc. 124*, 602-614.

[187] Tatsuno, Y., Saeki, Y., Iwaki, M., Yagi, T., Nozaki, M., Kitagawa, T., and Otsuka, S. (1978) Resonance Raman spectra of protocatechuate 3,4-dioxygenase. Evidence for coordination of tyrosine residue to ferric iron, *J. Am. Chem. Soc. 100*, 4614-4615.

[188] Pau, M. Y. M., Davis, M. I., Orville, A. M., Lipscomb, J. D., and Solomon, E. I. (2007) Spectroscopic and electronic structure study of the enzyme-substrate complex of intradiol dioxygenases: substrate activation by a high-spin ferric non-heme iron site, *J. Am. Chem. Soc. 129*, 1944-1958.

[189] Que, L., Jr., Lipscomb, J. D., Zimmermann, R., Münck, E., Orme-Johnson, N. R., and Orme-Johnson, W. H. (1976) Mössbauer and EPR spectroscopy of protocatechuate 3,4-dioxygenase from *Pseudomonas aeruginosa*, *Biochim. Biophys. Acta 452*, 320-334.

[190] Que, L., Jr., Lipscomb, J. D., Münck, E., and Wood, J. M. (1977) Protocatechuate 3,4-dioxygenase. Inhibitor studies and mechanistic implications, *Biochim. Biophys. Acta 485*, 60-74.

[191] Whittaker, J. W., and Lipscomb, J. D. (1984) ^{17}O-Water and cyanide ligation by the active site iron of protocatechuate 3,4-dioxygenase. Evidence for displaceable ligands in the native enzyme and in complexes with inhibitors or transition state analogs, *J. Biol. Chem. 259*, 4487-4495.

[192] Whittaker, J. W., and Lipscomb, J. D. (1984) Transition state analogs for protocatechuate 3,4-dioxygenase. Spectroscopic and kinetic studies of the binding reactions of ketonized substrate analogs, *J. Biol. Chem. 259*, 4476-4486.

[193] Whittaker, J. W., Lipscomb, J. D., Kent, T. A., and Münck, E. (1984) *Brevibacterium fuscum* protocatechuate 3,4-dioxygenase. Purification, crystallization, and characterization, *J. Biol. Chem. 259*, 4466-4475.

[194] Valley, M. P., Brown, C. K., Burk, D. L., Vetting, M. W., Ohlendorf, D. H., and Lipscomb, J. D. (2005) Roles of the equatorial tyrosyl iron ligand of protocatechuate 3,4-dioxygenase in catalysis, *Biochemistry 44*, 11024-11039.

[195] Xin, M., and Bugg, T. D. H. (2008) Evidence from mechanistic probes for distinct hydroperoxide rearrangement mechanisms in the intradiol and extradiol catechol dioxygenases, *J. Am. Chem. Soc. 130*, 10422-10430.

[196] Ohlendorf, D. H., Lipscomb, J. D., and Weber, P. C. (1988) Structure and assembly of protocatechuate 3,4-dioxygenase, *Nature 336*, 403-405.

[197] Vetting, M. W., and Ohlendorf, D. H. (2000) The 1.8 angstrom crystal structure of catechol 1,2-dioxygenase reveals a novel hydrophobic helical zipper as a subunit linker, *Structure 8*, 429-440.

[198] Earhart, C., A., Vetting, M., W., Gosu, R., Michaud-Soret, I., Que, L., Jr., and Ohlendorf, D., H. (2005) Structure of catechol 1,2-dioxygenase from *Pseudomonas arvilla, Biochem. Biophys. Res. Commun. 338*, 198-205.

[199] Matera, I., Ferraroni, M., Kolomytseva, M., Golovleva, L., Scozzafava, A., and Briganti, F. (2010) Catechol 1,2-dioxygenase from the Gram-positive *Rhodococcus opacus* 1CP: quantitative structure/activity relationship and the crystal structures of native enzyme and catechols adducts, *J. Struct. Biol. 170*, 548-564.

[200] Ferraroni, M., Seifert, J., Travkin, V. M., Thiel, M., Kaschabek, S., Scozzafava, A., Golovleva, L., Schloemann, M., and Briganti, F. (2005) Crystal structure of the hydroxyquinol 1,2-dioxygenase from *Nocardioides simplex* 3E, a key enzyme Involved in polychlorinated aromatics biodegradation, *J. Biol. Chem. 280*, 21144-21154.

[201] Ferraroni, M., Solyanikova, I. P., Kolomytseva, M. P., Scozzafava, A., Golovleva, L., and Briganti, F. (2004) Crystal structure of 4-chlorocatechol 1,2-dioxygenase from the chlorophenol-utilizing gram-positive *Rhodococcus opacus* 1CP, *J. Biol. Chem. 279*, 27646-27655.

[202] Ferraroni, M., Kolomytseva, M., Scozzafava, A., Golovleva, L., and Briganti, F. (2013) X-ray structures of 4-chlorocatechol 1,2-dioxygenase adducts with substituted catechols: New perspectives in the molecular basis of intradiol ring cleaving dioxygenases specificity, *J. Struct. Biol. 181*, 274-282.

[203] Vetting, M. W., D'Argenio, D. A., Ornston, L. N., and Ohlendorf, D. H. (2000) Structure of Acinetobacter strain ADP1 protocatechuate 3,4-dioxygenase at 2.2 angstrom resolution: Implications for the mechanism of an intradiol dioxygenase, *Biochemistry 39*, 7943-7955.

[204] Fujisawa, H., and Hayaishi, O. (1968) Protocatechuate 3,4-dioxygenase. I. Crystallization and characterization, *J. Biol. Chem. 243*, 2673-2681.

[205] Fujisawa, H., Uyeda, M., Kojima, Y., Nozaki, M., and Hayaishi, O. (1972) Protocatechuate 3,4-dioxygenase. II. Electron spin resonance and spectral studies on interaction of substrates and enzyme, *J. Biol. Chem. 247*, 4414-4421.

[206] Ohlendorf, D. H., Orville, A. M., and Lipscomb, J. D. (1994) Structure of protocatechuate 3,4-dioxygenase from *Pseudomonas aeruginosa* at 2.15 Å resolution, *J. Mol. Biol. 244*, 586-608.

[207] Hitomi, Y., Yoshida, M., Higuchi, M., Minami, H., Tanaka, T., and Funabiki, T. (2005) A linear correlation between energy of LMCT band and oxygenation reaction rate of a series of catecholatoiron(III) complexes: initial oxygen binding during intradiol catechol oxygenation, *J. Inorg. Biochem. 99*, 755-763.

[208] Kurahashi, T., Oda, K., Sugimoto, M., Ogura, T., and Fujii, H. (2006) Trigonal-bipyramidal geometry induced by an external water ligand in a sterically hindered iron salen complex, related to the active site of protocatechuate 3,4-dioxygenase, *Inorg. Chem. 45*, 7709-7721.

[209] Strater, N., Klabunde, T., Tucker, P., Witzel, H., and Krebs, B. (1995) Crystal structure of a purple acid phosphatase containing a dinuclear Fe(III)-Zn(II) active site, *Science 268*, 1489-1492.

[210] Rogers, M. S., and Dooley, D. M. (2003) Copper-tyrosyl radical enzymes, *Curr. Opin. Chem. Biol. 7*, 189-196.

[211] Nakatani, N., Hitomi, Y., and Sakaki, S. (2011) Multistate CASPT2 study of native iron(III)-dependent catechol dioxygenase and its functional models: electronic structure and ligand-to-metal charge-transfer excitation, *J. Phys. Chem. B 115*, 4781-4789.

[212] Que, L., Jr., and Heistand, R. H., II. (1979) Resonance Raman studies on pyrocatechase, *J. Am. Chem. Soc. 101*, 2219-2221.

[213] Borowski, T., and Siegbahn, P. E. M. (2006) Mechanism for catechol ring cleavage by non-heme iron in intradiol dioxygenases: A hybrid DFT study, *J. Am. Chem. Soc. 128*, 12941-12953.

[214] Nakatani, N., Nakao, Y., Sato, H., and Sakaki, S. (2009) Theoretical study of dioxygen binding process in iron(III) catechol dioxygenase: "oxygen activation" vs "substrate activation". *J. Phys. Chem. B 113*, 4826-4836.

[215] Cox, D. D., and Que, L., Jr. (1988) Functional models for catechol 1,2-dioxygenase. The role of the iron(III) center, *J. Am. Chem. Soc. 110*, 8085-8092.

[216] Orville, A. M., and Lipscomb, J. D. (1989) Binding of isotopically labeled substrates, inhibitors, and cyanide by protocatechuate 3,4-dioxygenase, *J. Biol. Chem. 264*, 8791-8801.

[217] Wojcik, A., Borowski, T., and Broclawik, E. (2011) The mechanism of the reaction of intradiol dioxygenase with hydroperoxy probe. A DFT study, *Catal. Today 169*, 207-216.

[218] Georgiev, V., Noack, H., Borowski, T., Blomberg, M. R., and Siegbahn, P. E. (2010) DFT study on the catalytic reactivity of a functional model complex for intradiol-cleaving dioxygenases, *J. Phys. Chem. B 114*, 5878-5885.

[219] Groce, S. L., and Lipscomb, J. D. (2003) Conversion of extradiol aromatic ring-cleaving homoprotocatechuate 2,3-dioxygenase into an intradiol cleaving enzyme, *J. Am. Chem. Soc. 125*, 11780-11781.

[220] Fujiwara, M., Golovleva, L. A., Saeki, Y., Nozaki, M., and Hayaishi, O. (1975) Extradiol cleavage of 3-substituted catechols by an intradiol dioxygenase, pyrocatechase, from a Pseudomonad, *J. Biol. Chem. 250*, 4848-4855.

[221] Que, L., Jr. (1978) Extradiol cleavage of o-aminophenol by pyrocatechase, *Biochem. Biophys. Res. Commun. 84*, 123-129.

[222] Murakami, S., Okuno, T., Matsumura, E., Takenaka, S., Shinke, R., and Aoki, K. (1999) Cloning of a gene encoding hydroxyquinol 1,2-dioxygenase that catalyzes both intradiol and extradiol ring cleavage of catechol, *Biosci. Biotechnol. Biochem. 63*, 859-865.

[223] Mendel, S., Arndt, A., and Bugg, T. D. H. (2004) Acid-base catalysis in the extradiol catechol dioxygenase reaction mechanism: Site-directed mutagenesis of His-115 and His-179 in *Escherichia coli* 2,3-dihydroxyphenylpropionate 1,2-dioxygenase (MhpB), *Biochemistry 43*, 13390-13396.

[224] Brivio, M., Schlosrich, J., Ahmad, M., Tolond, C., and Bugg, T. D. H. (2009) Investigation of acid-base catalysis in the extradiol and intradiol catechol dioxygenase reactions using a broad specificity mutant enzyme and model chemistry, *Org. Biomol. Chem. 7*, 1368-1373.

[225] Lin, G., Reid, G., and Bugg, T. D. H. (2001) Extradiol oxidative cleavage of catechols by ferrous and ferric complexes of 1,4,7-triazacyclononane: Insight into the mechanism of the extradiol catechol dioxygenases, *J. Am. Chem. Soc. 123*, 5030-5039.

[226] Bugg, T. D. H., and Lin, G. (2001) Solving the riddle of the intradiol and extradiol catechol dioxygenases: How do enzymes control hydroperoxide rearrangements?, *Chem. Commun.*, 941-952.

[227] Jo, D.-H., and Que, L., Jr. (2000) Tuning the regiospecificity of cleavage in FeIII catecholate complexes: tridentate facial versus meridional ligands, *Angew. Chem. Int. Ed. 39*, 4284-4287.

[228] Varadi, T., Pap, J. S., Giorgi, M., Parkanyi, L., Csay, T., Speier, G., and Kaizer, J. (2013) Iron(III) complexes with meridional ligands as functional models of intradiol-cleaving catechol dioxygenases, *Inorg. Chem. 52*, 1559-1569.

[229] Ray, K., Lee, S. M., and Que, L., Jr. (2008) Iron-oxidation-state-dependent O-O bond cleavage of meta-chloroperbenzoic acid to form an iron(IV)-oxo complex, *Inorg. Chim. Acta 361*, 1066-1069.

[230] Groce, S. L., and Lipscomb, J. D. (2005) Aromatic ring cleavage by homoprotocatechuate 2,3-dioxygenase: Role of His200 in the kinetics of interconversion of reaction cycle intermediates, *Biochemistry 44*, 7175-7188.

[231] Kovaleva, E. G., and Lipscomb, J. D. (2008) Intermediate in the O-O bond cleavage reaction of an extradiol dioxygenase, *Biochemistry 47*, 11168-11170.

[232] Mbughuni, M. M., Meier, K. K., Münck, E., and Lipscomb, J. D. (2012) Substrate-mediated oxygen activation by homoprotocatechuate 2,3-dioxygenase: Intermediates formed by a tyrosine 257 variant, *Biochemistry 51*, 8743-8754.

[233] Kovaleva, E. G., and Lipscomb, J. D. (2012) Structural basis for the role of tyrosine 257 of homoprotocatechuate 2,3-dioxygenase in substrate and oxygen activation, *Biochemistry 51*, 8755-8763.

[234] Eltis, L. D., and Kushmaro, A. (2011) Environmental biotechnology for sustainability: unleashing the might of the small, *Curr. Opin. Biotechnol. 22*, 386-387.

[235] Copley, S. D. (2009) Evolution of efficient pathways for degradation of anthropogenic chemicals, *Nat. Chem. Biol. 5*, 559-566.

[236] Polen, T., Spelberg, M., and Bott, M. (2013) Toward biotechnological production of adipic acid and precursors from biorenewables, *J. Biotechnol. 167*, 75-84.

[237] Fujii, H., and Funahashi, Y. (2002) A trigonal-bipyramidal ferric aqua complex with a sterically hindered salen ligand as a model for the active site of protocatechuate 3,4-dioxygenase, *Angew. Chem. Int. Ed. 41*, 3638-3641.

[238] Costas, M., Mehn, M. P., Jensen, M. P., and Que, L., Jr. (2004) Dioxygen activation at mononuclear nonheme iron active sites: enzymes, models, and intermediates, *Chem. Rev. 104*, 939-986.

[239] Pootoolal, J., Thomas, M. G., Marshall, C. G., Neu, J. M., Hubbard, B. K., Walsh, C. T., and Wright, G. D. (2002) Assembling the glycopeptide antibiotic scaffold: the biosynthesis of A47934 from S*treptomyces toyocaensis* NRRL15009, *Proc. Natl. Acad. Sci. USA. 99*, 8962-8967.

[240] Li, T.-L., Huang, F., Haydock, S. F., Mironenko, T., Leadlay, P. F., and Spencer, J. B. (2004) Biosynthetic gene cluster of the glycopeptide antibiotic teicoplanin. Characterization of two glycosyltransferases and the key acyltransferase, *Chem. Biol. 11*, 107-119.

[241] Sosio, M., Stinchi, S., Beltrametti, F., Lazzarini, A., and Donadio, S. (2003) The gene cluster for the biosynthesis of the glycopeptide antibiotic A40926 by *Nonomuraea species*, *Chem. Biol. 10*, 541-549.

[242] Blazic, M., Starcevic, A., Lisfi, M., Baranasic, D., Goranovic, D., Fujs, S., Kuscer, E., Kosec, G., Petkovic, H., Cullum, J., Hranueli, D., and Zucko, J. (2012) Annotation of the modular polyketide synthase and nonribosomal peptide synthetase gene clusters in the genome of *Streptomyces tsukubaensis* NRRL18488, *Appl. Environ. Microbiol. 78,* 8183-8190.

[243] Haslinger, K., Brieke, C., Uhlmann, S., Sieverling, L., Süssmuth, R. D., and Cryle, M. J. (2014) The structure of a transient complex of a nonribosomal peptide synthetase and a cytochrome P450 monooxygenase, *Angew. Chem. Int. Ed. 53,* 8518-8522.

[244] Fujimori, D. G., Hrvatin, S., Neumann, C. S., Strieker, M., Marahiel, M. A., and Walsh, C. T. (2007) Cloning and characterization of the biosynthetic gene cluster for kutznerides, *Proc. Natl. Acad. Sci. USA 104*, 16498-16503.

[245] Neary, J. M., Powell, A., Gordon, L., Milne, C., Flett, F., Wilkinson, B., Smith, C. P., and Micklefield, J. (2007) An asparagine oxygenase (AsnO) and a 3-hydroxyasparaginyl phosphotransferase (HasP) are involved in the biosynthesis of calcium-dependent lipopeptide antibiotics, *Microbiology 153*, 768-776.

[246] Eustáquio, A. S., Janso, J. E., Ratnayake, A. S., O'Donnell, C. J., and Koehn, F. E. (2014) Spliceostatin hemiketal biosynthesis in Burkholderia spp. is catalyzed by an iron/α-ketoglutarate–dependent dioxygenase, *Proc. Natl. Acad. Sci. USA 111*, E3376-E3385.

[247] Fischbach, M. A., and Walsh, C. T. (2006) Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: Logic, machinery, and mechanisms, *Chem. Rev. 106*, 3468-3496.

[248] Weissman, K. J. (2014) The structural biology of biosynthetic megaenzymes, *Nat. Chem. Biol. 11*, 660-670.

[249] Strieker, M., Tanović, A., and Marahiel, M. A. (2010) Nonribosomal peptide synthetases: structures and dynamics, *Curr. Opin. Struct. Biol. 20*, 234-240.

[250] Bode, H. B., Brachmann, A. O., Jadhav, K. B., Seyfarth, L., Dauth, C., Fuchs, S. W., Kaiser, M., Waterfield, N. R., Sack, H., Heinemann, S. H., and Arndt, H.-D. (2015) Structure elucidation and activity of Kolossin A, the D-/L-pentadecapeptide product of a giant nonribosomal peptide synthetase, *Angew. Chem. Int. Ed. 54*, 10352-10355.

[251] Conti, E., Stachelhaus, T., Marahiel, M. A., and Brick, P. (1997) Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S, *The EMBO Journal 16*, 4174-4183.

[252] May, J. J., Kessler, N., Marahiel, M. A., and Stubbs, M. T. (2002) Crystal structure of DhbE, an archetype for aryl acid activating domains of modular nonribosomal peptide synthetases, *Proc. Natl. Acad. Sci. USA 99*, 12120-12125.

[253] Lai, J. R., Koglin, A., and Walsh, C. T. (2006) Carrier protein structure and recognition in polyketide and nonribosomal peptide biosynthesis, *Biochemistry 45*, 14869-14879.

[254] Ehmann, D. E., Gehring, A. M., and Walsh, C. T. (1999) Lysine biosynthesis in Saccharomyces cerevisiae: mechanism of α-aminoadipate reductase (Lys2) involves posttranslational phosphopantetheinylation by Lys5, *Biochemistry 38*, 6171-6177.

[255] Trauger, J. W., Kohli, R. M., Mootz, H. D., Marahiel, M. A., and Walsh, C. T. (2000) Peptide cyclization catalysed by the thioesterase domain of tyrocidine synthetase, *Nature 407*, 215-218.

[256] Yeh, E., Kohli, R. M., Bruner, S. D., and Walsh, C. T. (2004) Type II thioesterase restores activity of a NRPS module stalled with an aminoacyl-S-enzyme that cannot be elongated, *ChemBioChem 5*, 1290-1293.

[257] Walsh, C. T., Chen, H., Keating, T. A., Hubbard, B. K., Losey, H. C., Luo, L., Marshall, C. G., Miller, D. A., and Patel, H. M. (2001) Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines, *Curr. Opin. Chem. Biol. 5*, 525-534.

[258] Ansari, M., Sharma, J., Gokhale, R., and Mohanty, D. (2008) In silico analysis of methyltransferase domains involved in biosynthesis of secondary metabolites, *BMC Bioinform. 9*, 454.

[259] Patel, H. M., Tao, J., and Walsh, C. T. (2003) Epimerization of an l-cysteinyl to a d-cysteinyl residue during thiazoline ring formation in siderophore chain elongation by pyochelin synthetase from *Pseudomonas aeruginosa*, *Biochemistry 42*, 10514-10527.

[260] Hoffmann, K., Schneider-Scherzer, E., Kleinkauf, H., and Zocher, R. (1994) Purification and characterization of eucaryotic alanine racemase acting as key enzyme in cyclosporin biosynthesis, *J. Biol. Chem. 269*, 12710-12714.

[261] Chen, H., Thomas, M. G., O'Connor, S. E., Hubbard, B. K., Burkart, M. D., and Walsh, C. T. (2001) Aminoacyl-S-enzyme intermediates in β-hydroxylations and α,β-desaturations of amino acids in heptide antibiotics, *Biochemistry 40*, 11651-11659.

[262] Losey, H. C., Peczuh, M. W., Chen, Z., Eggert, U. S., Dong, S. D., Pelczer, I., Kahne, D., and Walsh, C. T. (2001) Tandem action of glycosyltransferases in the maturation of vancomycin and teicoplanin aglycones: novel glycopeptides, *Biochemistry 40*, 4745-4755.

[263] Pacholec, M., Sello, J. K., Walsh, C. T., and Thomas, M. G. (2007) Formation of an aminoacyl-S-enzyme intermediate is a key step in the biosynthesis of chloramphenicol, *Org. Biomol. Chem. 5*, 1692-1694.

[264] He, J., Magarvey, N., Piraee, M., and Vining, L. C. (2001) The gene cluster for chloramphenicol biosynthesis in *Streptomyces venezuelae* ISP5230 includes novel shikimate pathway homologues and a monomodular non-ribosomal peptide synthetase gene, *Microbiology 147*, 2817-2829.

[265] Frazão, C., Silva, G., Gomes, C. M., Matias, P., Coelho, R., Sieker, L., Macedo, S., Liu, M. Y., Oliveira, S., Teixeira, M., Xavier, A. V., Rodrigues-Pousada, C., Carrondo, M. A., and Le Gall, J. (2000) Structure of a dioxygen reduction enzyme from *Desulfovibrio gigas*, *Nat. Struct. Mol. Biol. 7*, 1041-1045.

[266] Silaghi-Dumitrescu, R., Kurtz, D. M., Jr., Ljungdahl, L. G., and Lanzilotta, W. N. (2005) X-ray crystal structures of *Moorella thermoacetica* FprA. Novel diiron site structure and mechanistic insights into a scavenging nitric oxide reductase, *Biochemistry 44*, 6492-6501.

[267] Hayashi, T., Caranto, J. D., Wampler, D. A., Kurtz, D. M., Jr., and Moenne-Loccoz, P. (2010) Insights into the nitric oxide reductase mechanism of flavodiiron proteins from a flavin-free enzyme, *Biochemistry 49*, 7040-7049.

[268] Vu, V. V., Makris, T. M., Lipscomb, J. D., and Que, L., Jr. (2011) Active-site structure of a beta -hydroxylase in antibiotic biosynthesis, *J. Am. Chem. Soc. 133*, 6938-6941.

[269] Paulsen, K. E., Orville, A. M., Frerman, F. E., Lipscomb, J. D., and Stankovich, M. T. (1992) The redox properties of electron-transfer flavoprotein ubiquinone oxidoreductase as determined by EPR-spectroelectrochemistry, *Biochemistry 31*, 11755-11761.

[270] Liu, Y., Nesheim, J. C., Paulsen, K. E., Wallar, B. J., Stankovich, M. T., and Lipscomb, J. D. (1995) Roles of component B and reductase in methane monooxygenase (MMO) turnover, *J. Inorg. Biochem. 59*, 368.

[271] Mitić, N., Schwartz, J. K., Brazeau, B. J., Lipscomb, J. D., and Solomon, E. I. (2008) CD and MCD studies of the effects of component B variant binding on the biferrous active site of methane monooxygenase, *Biochemistry 47*, 8386-8397.

[272] Shu, L., Nesheim, J. C., Kauffmann, K., Münck, E., Lipscomb, J. D., and Que, L., Jr. (1997) An $Fe^{IV}{}_2O_2$ diamond core structure for the key intermediate Q of methane monooxygenase, *Science 275*, 515-518.

[273] Zhang, C., Griffith, B. R., Fu, Q., Albermann, C., Fu, X., Lee, I.-K., Li, L., and Thorson, J. S. (2006) Exploiting the reversibility of natural product glycosyltransferase-catalyzed reactions, *Science 313*, 1291-1294.

[274] Williams, G. J. (2013) Engineering polyketide synthases and nonribosomal peptide synthetases, *Curr. Opin. Struct. Biol. 23*, 603-612.

[275] Baltz, R. H. (2014) Combinatorial biosynthesis of cyclic lipopeptide antibiotics: a model for synthetic biology to accelerate the evolution of secondary metabolite biosynthetic pathways, *ACS Synth. Biol. 3*, 748-758.

[276] Siddiqueullah, M., McGrath, R. M., Vining, L. C., Sala, F., and Westlake, D. W. S. (1967) Biosynthesis of chloramphenicol. II. p-Aminophenylalanine as a precursor of the p-nitrophenylserinol moiety, *Can. J. Biochem. 45*, 1881-1889.

[277] Lu, H. G., Chanco, E., and Zhao, H. M. (2012) CmlI is an N-oxygenase in the biosynthesis of chloramphenicol, *Tetrahedron 68*, 7651-7654.

[278] Zocher, G., Winkler, R., Hertweck, C., and Schulz, G. E. (2007) Structure and action of the N-oxygenase AurF from *Streptomyces thioluteus*, *J. Mol. Biol. 373*, 65-74.

[279] Choi, Y. S., Zhang, H., Brunzelle, J. S., Nair, S. K., and Zhao, H. (2008) In vitro reconstitution and crystal structure of p-aminobenzoate N-oxygenase (AurF) involved in aureothin biosynthesis, *Proc. Natl. Acad. Sci. USA. 105*, 6858-6863.

[280] Winkler, R., and Hertweck, C. (2007) Biosynthesis of nitro compounds, *ChemBioChem 8*, 973-977.

[281] He, J., and Hertweck, C. (2004) Biosynthetic origin of the rare nitroaryl moiety of the polyketide antibiotic aureothin: Involvement of an unprecedented N-oxygenase, *J. Am. Chem. Soc. 126*, 3694-3695.

[282] Winkler, R., and Hertweck, C. (2005) Sequential enzymatic oxidation of aminoarenes to nitroarenes via hydroxylamines, *Angew. Chem. Int. Ed. 44*, 4083-4087.

[283] Lee, J., Simurdiak, M., and Zhao, H. (2005) Reconstitution and characterization of aminopyrrolnitrin oxygenase, a Rieske N-oxygenase that catalyzes unusual arylamine oxidation., *J. Biol. Chem. 280*, 36719-36727.

[284] Li, N., Korboukh, V. K., Krebs, C., and Bollinger, J. M., Jr. (2010) Four-electron oxidation of p-hydroxylaminobenzoate to p-nitrobenzoate by a peroxodiferric complex in AurF from *Streptomyces thioluteus*, *Proc. Natl. Acad. Sci. USA 107*, 15722-15727.

[285] Díaz, E., Jiménez, J. I., and Nogales, J. (2013) Aerobic degradation of aromatic compounds, *Curr. Opin. Biotechnol. 24*, 431-442.

[286] Lipscomb, J. D. (2008) Mechanism of extradiol aromatic ring-cleaving dioxygenases, *Curr. Opin. Struct. Biol. 18*, 644-649.

[287] Nakata, H., Yamauchi, T., and Fujisawa, H. (1978) Studies on the reaction intermediate of protocatechuate 3,4-dioxygenase. Formation of enzyme-product complex, *Biochim. Biophys. Acta 527*, 171-181.

[288] Wilmot, C. M., Hajdu, J., McPherson, M. J., Knowles, P. F., and Phillips, S. E. V. (1999) Visualization of dioxygen bound to copper during enzyme catalysis, *Science 286*, 1724-1728.

[289] Schlichting, I., Berendzen, J., Chu, K., Stock, A. M., Maves, S. A., Benson, D. E., Sweet, R. M., Ringe, D., Petsko, G. A., and Sligar, S. G. (2000) The catalytic pathway of cytochrome P450$_{\text{cam}}$ at atomic resolution, *Science 287*, 1615-1622.

[290] Johnson-Winters, K., Purpero, V. M., Kavana, M., and Moran, G. R. (2005) Accumulation of multiple intermediates in the catalytic cycle of (4-hydroxyphenyl)pyruvate dioxygenase from *Streptomyces avermitilis*, *Biochemistry 44*, 7189-7199.

[291] Vergé, D., and Arrio-Dupont, M. (1981) Interactions between apoaspartate aminotransferase and pyridoxal 5'-phosphate. A stopped-flow study, *Biochemistry 20*, 1210-1216.

[292] Groce, S. L., Miller-Rodeberg, M. A., and Lipscomb, J. D. (2004) Single-turnover kinetics of homoprotocatechuate 2,3-dioxygenase, *Biochemistry 43*, 15141-15153.

[293] Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallogr. E60*, 2126-2132.

[294] Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Longa, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures, *Acta Cryst. D67* 355-367.

[295] Dorn, E., and Knackmuss, H. J. (1978) Chemical structure and biodegradability of halogenated aromatic compounds. Substituent effects on 1,2-dioxygenation of catechol, *Biochem. J. 174*, 85-94.

[296] Pieken, W. A., and Kozarich, J. W. (1990) Lactonization of cis,cis-3-halomuconates: influence of pH and halo substituent on the regiochemistry, *J. Org. Chem. 55*, 3029-3035.

[297] Bugg, T. D. H., and Ramaswamy, S. (2008) Non-heme iron-dependent dioxygenases: unravelling catalytic mechanisms for complex enzymatic oxidations, *Curr. Opin. Chem. Biol. 12*, 134-140.

[298] Siegbahn, P. E. M., and Haeffner, F. (2004) Mechanism for catechol ring-cleavage by non-heme iron extradiol dioxygenases, *J. Am. Chem. Soc. 126*, 8919-8932.

[299] Goodman, R. M., and Kishi, Y. (1998) Experimental support for the primary stereoelectronic effect governing Baeyer−Villiger oxidation and Criegee rearrangement, *J. Am. Chem. Soc. 120*, 9392-9393.

[300] Ferraroni, M., Kolomytseva, M. P., Solyanikova, I. P., Scozzafava, A., Golovleva, L. A., and Briganti, F. (2006) Crystal structure of 3-chlorocatechol 1,2-dioxygenase: Key enzyme of a new modified ortho-pathway from the Gram-positive *Rhodococcus opacus* 1CP grown on 2-chlorophenol, *J. Mol. Biol. 360*, 788-799.

[301] Stanier, R. Y., and Ornston, L. N. (1973) The beta-ketoadipate pathway, *Adv. Microb. Phys. 9*, 89-151.

[302] Northrop, D. B. (1998) On the meaning of K_m and V/K in enzyme kinetics, *J. Chem. Ed. 75*, 1153-1157.

[303] Chen, H., and Walsh, C. T. (2001) Coumarin formation in novobiocin biosynthesis: β-Hydroxylation of the aminoacyl enzyme tyrosyl-S-NovH by a cytochrome P450 NovI, *Chem. Biol. 8*, 301-312.

[304] Du, L., Sanchez, C., Chen, M., Edwards, D. J., and Shen, B. (2000) The biosynthetic gene cluster for the antitumor drug bleomycin from *Streptomyces verticillus* ATCC15003 supporting functional interactions between nonribosomal peptide synthetases and a polyketide synthase, *Chem. Biol. 7*, 623-642.

[305] Nguyen, K. T., Ritz, D., Gu, J.-Q., Alexander, D., Chu, M., Miao, V., Brian, P., and Baltz, R. H. (2006) Combinatorial biosynthesis of novel antibiotics related to daptomycin, *Proc. Natl. Acad. Sci. USA 103*, 17462-17467.

[306] Doull, J., Ahmed, Z., Stuttard, C., and Vining, L. C. (1985) Isolation and characterization of *Streptomyces venezuelae* mutants blocked in chloramphenicol biosynthesis, *J. Gen. Microbiol. 131*, 97-104.

[307] Bebrone, C. (2007) Metallo-beta-lactamases (classification, activity, genetic organization, structure, zinc coordination) and their superfamily, *Biochem. Pharmacol. 74*, 1686-1701.

[308] Notomista, E., Lahm, A., Di Donato, A., and Tramontano, A. (2003) Evolution of bacterial and archaeal multicomponent monooxygenases, *J. Mol. Evol. 56*, 435-445.

[309] Rosenzweig, A. C., Frederick, C. A., Lippard, S. J., and Nordlund, P. (1993) Crystal structure of a bacterial non-haem iron hydroxylase that catalyses the biological oxidation of methane, *Nature 366*, 537-543.

[310] Kovaleva, E. G., Neibergall, M. B., Chakrabarty, S., and Lipscomb, J. D. (2007) Finding intermediates in the O_2 activation pathways of non-heme iron oxygenases, *Acc. Chem. Res. 40*, 475-483.

[311] Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L., and Clardy, J. (1993) Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin, *J. Mol. Biol. 229*, 105-124.

[312] Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol. 276*, 307-326.

[313] Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution., *Acta Crystallogr. D66*, 213-221.

[314] Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., and Evans, P. R. (2011) Overview of the CCP4 suite and current developments, *Acta Crystallogr D 67*, 235-242.

[315] Kiefer, F., Arnold, K., Künzli, M., Bordoli, L., and Schwede, T. (2009) The SWISS-MODEL Repository and associated resources, *Nucleic Acids Res. 37*, D387-D392.

[316] de Vries, S. J., van Dijk, M., and Bonvin, A. M. J. J. (2010) The HADDOCK web server for data-driven biomolecular docking., *Nature Protocols 5*, 883-897.

[317] Zeldin, O. B., Gerstel, M., and Garman, E. F. (2013) RADDOSE-3D: time- and space-resolved modelling of dose in macromolecular crystallography, *J. Appl. Crystallogr. 46*, 1225-1230.

[318] Voegtli, W. C., Khidekel, N., Baldwin, J., Ley, B. A., Bollinger, J. M., Jr., and Rosenzweig, A. C. (2000) Crystal structure of the ribonucleotide reductase R2 mutant that accumulates a μ -1,2-peroxodiiron(III) intermediate during oxygen activation, *J. Am. Chem. Soc. 122*, 3255-3261.

[319] Rosenzweig, A. C., Nordlund, P., Takahara, P. M., Frederick, C. A., and Lippard, S. J. (1995) Geometry of the soluble methane monooxygenase catalytic diiron center in two oxidation states, *Chem. Biol. 2*, 409-418.

[320] Whittington, D. A., and Lippard, S. J. (2001) Crystal structures of the soluble methane monooxygenase hydroxylase from *Methylococcus capsulatus* (Bath) demonstrating geometrical variability at the dinuclear iron active site, *J. Am. Chem. Soc. 123*, 827-838.

[321] Wei, P.-p., Skulan, A. J., Mitić, N., Yang, Y.-S., Saleh, L., Bollinger, J. M., Jr., and Solomon, E. I. (2004) Electronic and spectroscopic studies of the non-heme reduced binuclear iron sites of two ribonucleotide reductase variants: comparison to reduced methane monooxygenase and contributions to O_2 reactivity, *J. Am. Chem. Soc. 126*, 3777-3788.

[322] Cryle, M. J., Meinhart, A., and Schlichting, I. (2010) Structural characterization of OxyD, a cytochrome P450 involved in β-hydroxytyrosine formation in vancomycin biosynthesis, *J. Biol. Chem. 285*, 24562-24574.

[323] McCafferty, D. G., Cudic, P., Frankel, B. A., Barkallah, S., Kruger, R. G., and Li, W. (2002) Chemistry and biology of the ramoplanin family of peptide antibiotics, *Biopolymers 66*, 261-284.

[324] Hou, J., Robbel, L., and Marahiel, M. A. (2011) Identification and characterization of the lysobactin biosynthetic gene cluster reveals mechanistic insights into an unusual termination module architecture, *Chem. Biol. 18*, 6555-6664.

[325] Trindade-Silva, A. E., Machado-Ferreira, E., Senra, M. V. X., Vizzoni, V. F., Yparraguirre, L. A., Leoncini, O., and Soares, C. A. G. (2009) Physiological traits of the symbiotic bacterium *Teredinibacter turnerae* isolated from the mangrove shipworm *Neoteredo reynei*, *Genet. Mol. Biol. 32*, 572-581.

[326] Liu, L., Budnjo, A., Jokela, J., Haug, B. E., Fewer, D. P., Wahlsten, M., Rouhiainen, L., Permi, P., Fossen, T., and Sivonen, K. (2014) Pseudoaeruginosins, nonribosomal peptides in *Nodularia spumigena*, *ACS Chem. Biol.*

[327] Stubbe, J., Nocera, D. G., Yee, C. S., and Chang, M. C. Y. (2003) Radical initiation in the class I ribonucleotide reductase: Long-range proton-coupled electron transfer?, *Chem. Rev. 103*, 2167-2201.

[328] Koglin, A., Mofid, M. R., Löhr, F., Schäfer, B., Rogov, V. V., Blum, M.-M., Mittag, T., Marahiel, M. A., Bernhard, F., and Dötsch, V. (2006) Conformational switches modulate protein interactions in peptide antibiotic synthetases, *Science 312*, 273-276.

[329] Cryle, M. J., and Schlichting, I. (2008) Structural insights from a P450 carrier protein complex reveal how specificity is achieved in the P450BioI ACP complex, *Proc. Natl. Acad. Sci. USA 105*, 15696-15701.

[330] Tanovic, A., Samel, S. A., Essen, L.-O., and Marahiel, M. A. (2008) Crystal structure of the termination module of a nonribosomal peptide synthetase, *Science 321*, 659-663.

[331] Andrews, S. C. (2010) The Ferritin-like superfamily: Evolution of the biological iron storeman from a rubrerythrin-like ancestor, *Biochimica et Biophysica Acta (BBA) - General Subjects 1800*, 691-705.

[332] Kirner, S., Hammer, P. E., Hill, D. S., Altmann, A., Fischer, I., Weislo, L. J., Lanahan, M., van Pée, K.-H., and Ligon, J. M. (1998) Functions encoded by pyrrolnitrin biosynthetic genes from *Pseudomonas fluorescens*, *J. Bacteriol. 180*, 1939-1943.

[333] He, J., and Hertweck, C. (2003) Iteration as programmed event during polyketide assembly; Molecular analysis of the aureothin biosynthesis gene cluster, *Chem. Biol. 10*, 1225-1232.

[334] Dodani, S. C., Cahn, J. K. B., Heinisch, T., Brinkmann-Chen, S., McIntosh, J. A., and Arnold, F. H. (2014) Structural, functional, and spectroscopic characterization of the substrate scope of the novel nitrating cytochrome P450 TxtE, *ChemBioChem 15*, 2259-2267.

[335] Piraee, M., White, R. L., and Vining, L. C. (2004) Biosynthesis of the dichloroacetyl component of chloramphenicol in *Streptomyces venezuelae* ISP5230: Genes required for halogenation, *Microbiology (Reading, United Kingdom) 150*, 85-94.

[336] Guy, J. E., Whittle, E., Kumaran, D., Lindqvist, Y., and Shanklin, J. (2007) The crystal structure of the ivy Delta4-16:0-ACP desaturase reveals structural details of the oxidized active site and potential determinants of regioselectivity., *J. Biol. Chem. 282*, 19863-19871.

[337] Schirmer, A., Rude, M. A., Li, X., Popova, E., and del Cardayre, S. B. (2010) Microbial synthesis of alkanes, *Science (Washington, D. C.) 329*, 559-562.

[338] Winkler, R., Zocher, G., Richter, I., Friedrich, T., Schulz, G. E., and Hertweck, C. (2007) A binuclear manganese cluster that catalyzes radical-mediated N-oxygenation, *Angew. Chem. Int. Ed. 46*, 8605-8608.

[339] McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software, *J. Appl. Crystallogr. 40*, 658-674.

[340] Logan, D. T., Su, X. D., Aberg, A., Regnstrom, K., Hajdu, J., Eklund, H., and Nordlund, P. (1996) Crystal structure of reduced protein R2 of ribonucleotide reductase - the structural basis for oxygen activation at a dinuclear iron site, *Structure 4*, 1053-1064.

[341] Berglund, G. I., Carlsson, G. H., Smith, A. T., Szoke, H., Henriksen, A., and Hajdu, J. (2002) The catalytic pathway of horseradish peroxidase at high resolution, *Nature 417*, 463-468.

[342] Karlsson, A., Parales, J. V., Parales, R. E., Gibson, D. T., Eklund, H., and Ramaswamy, S. (2000) The reduction of the Rieske iron-sulfur cluster in naphthalene dioxygenase by X-rays, *J. Inorg. Biochem. 78*, 83-87.

[343] Elango, N., Radhakrishnan, R., Froland, W. A., Wallar, B. J., Earhart, C. A., Lipscomb, J. D., and Ohlendorf, D. H. (1997) Crystal structure of the hydroxylase component of methane monooxygenase from *Methylosinus trichosporium* OB3b, *Protein Sci. 6*, 556-568.

[344] Schwartz, J. K., Wei, P.-p., Mitchell, K. H., Fox, B. G., and Solomon, E. I. (2008) Geometric and electronic structure studies of the binuclear nonheme ferrous active site of toluene-4-monooxygenase: Parallels with methane monooxygenase and insight into the role of the effector proteins in O_2 activation, *J. Am. Chem. Soc. 130*, 7098-7109.