

Accepted manuscript; appeared in the Journal of American Chemical Society,
2016, 138(44), 14623, DOI: 10.1021/jacs.6b06987

Pathways for Arene Oxidation in Non-Heme Diiron Enzymes: Lessons from Computational Studies on Benzoyl-Coenzyme A Epoxidase

Tibor András Rokob

rokob.tibor.andras@ttk.mta.hu, +36-1-3826-980

Institute of Organic Chemistry, Research Centre for Natural Sciences, Hungarian Academy of Sciences, 1117 Budapest, Magyar
Tudósok körútja 2., Hungary

ABSTRACT

Oxygenation of aromatic rings using O₂ is catalyzed by several non-heme carboxylate-bridged diiron enzymes. In order to provide a general mechanistic description for these reactions, computational studies were carried out at the ONIOM(B3LYP/BP86/Amber) level on the non-heme diiron enzyme benzoyl coenzyme A epoxidase BoxB. The calculations revealed four possible pathways for attacking the aromatic ring: a) electrophilic (2e⁻) attack by a bis(μ-oxo)-diiron(IV) species (**Q** pathway); b) electrophilic (2e⁻) attack via the σ* orbital of a μ-η²:η²-peroxo-diiron(III) intermediate (**Pσ*** pathway); c) radical (1e⁻) attack via the π*-orbital of a superoxo-diiron(II,III) species (**Pπ*** pathway); d) radical (1e⁻) attack of a partially quenched bis(μ-oxo)-diiron(IV) intermediate (**Q'** pathway). The results allowed earlier work of de Visser on olefin epoxidation by diiron complexes and QM-cluster studies of Liao and Siegbahn on BoxB to be put into a broader perspective. Parallels with epoxidation using organic peracids were also examined. Specifically for the BoxB enzyme, the **Q** pathway was found to be the most preferred, but the corresponding bis(μ-oxo)-diiron(IV) species is significantly destabilized and not expected to be directly observable. Epoxidation via the **Pσ*** pathway represents an energetically somewhat higher lying alternative; possible strategies for experimental discrimination are discussed. The selectivity toward epoxidation is shown to stem from a combination of inherent electronic properties of the thioacyl substituent and enzymatic constraints. Possible implications of the results for toluene monooxygenases are considered as well.

1. Introduction

Nature has long been a source of inspiration for the development of more efficient, cheaper, or greener chemical transformations. This effort has been an important motivation for the study of enzymatic reaction mechanisms, because the gained understanding can greatly facilitate synthetic applications of enzymes, modified enzymes, and artificial catalysts mimicking their function.¹ In particular, chemo-, regio-, and stereoselective oxidative functionalization of aromatic rings using O₂ or H₂O₂ is a valuable synthetic transformation, which is accomplished by members of several enzyme classes, such as flavoenzymes, cytochrome P450s, non-heme monoiron and diiron, as well as dicopper, proteins.² Accordingly, significant amount of work has been done to develop enzymatic and bioinspired approaches,³ an effort inevitably intertwined with the endeavor of elucidating the mechanism of the attack on the aromatic ring. Theory and experiment have gone hand in hand^{4,5} to uncover a spectacular mechanistic diversity, yet with a number of recurring strategies and features; nevertheless, several open questions still remain.

In flavoprotein hydroxylases (Figure 1A),⁶ the flavin cofactor (**a1**) activates molecular O₂ to produce⁷ C4a-hydroperoxyflavin (**a2**), a quasi-stable intermediate. Upon interaction with a π -system, this organic hydroperoxo species can undergo O–O bond heterolysis and carry out electrophilic attack as an OH⁺-equivalent.⁸ In many cases, the substrate is a phenol derivative, with its hydroxyl group deprotonated in the enzymatic environment, which facilitates the attack and leads directly to a neutral hydroxydienone product (**a3**), subject to subsequent tautomerization (**a4**). Nevertheless, this viewpoint has been recently challenged, and homolytic O–O breakage and transfer of OH radical was suggested instead.⁹

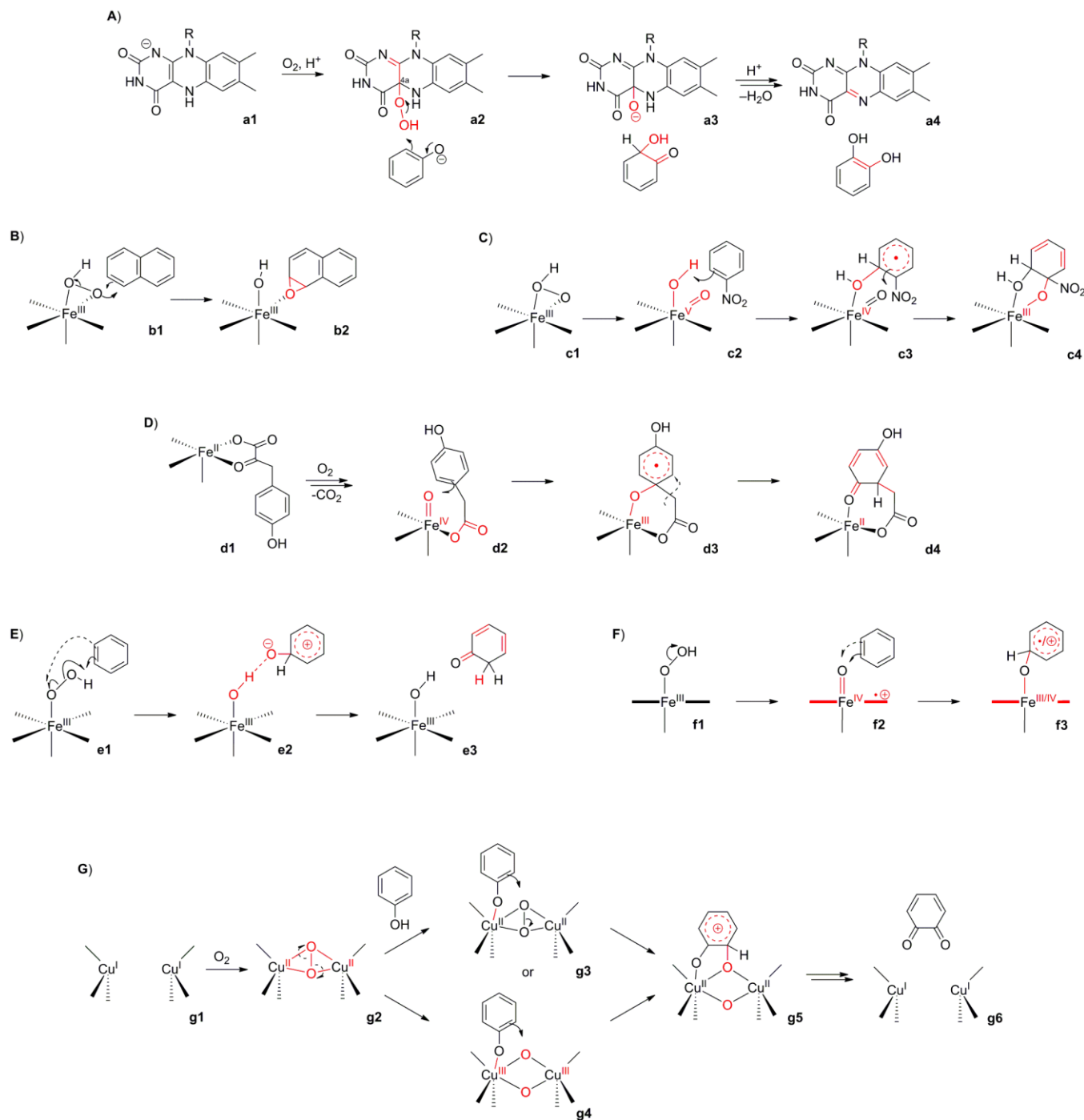


Figure 1 Some possible mechanisms for the oxidation of aromatic rings by enzymes and synthetic complexes. A) electrophilic attack by "OH⁺" in flavoprotein hydroxylases; B) direct attack by hydroperoxo-iron(III) via O-O heterolysis in naphthalene dioxygenase; C) formation and stepwise reactivity of a transient iron(V) complex in nitrobenzene dioxygenase; D) radical attack by an iron(IV)-oxo species in hydroxyphenylpyruvate dioxygenase; E) electrophilic attack with transient radical character by a synthetic hydroperoxo-iron(III) complex containing a pentadentate ligand; F) mixed radical-electrophilic attack in heme-iron systems; G) electrophilic attack by peroxo or bis-oxo-dicopper complexes in tyrosinase. In each step, changing parts of the structures are highlighted in red.

For Rieske-type non-heme monoiron oxygenases¹⁰ hydroxylating aromatic rings, naphthalene dioxygenase for example, no conclusive mechanism has been obtained yet. The presence of two open *cis*-aligned

coordination sites on the iron center allows the formation of a side-on bound Fe(III)–OOH complex, which has long been considered a key intermediate. Still, the further steps of the conversion of this complex to the dihydroxylated product are intensively debated. Possibilities include direct attack of the Fe(III) species on the substrate or initial O–O cleavage yielding a transient Fe(V) oxidizing species.¹¹ Computations on naphthalene dioxygenase suggested the first alternative (Figure 1B),¹² corresponding to a formal heterolysis of the peroxy ligand (**b1**) leading to Fe(III)–OH and a neutral “O” atom, which, in a concerted reaction, produces the iron-bound naphthalene-1,2-epoxide as the first intermediate (**b2**). Key role for the iron(III) complex is also supported by a study highlighting the electrophilic character of the high-spin iron(III)-OOH system.¹³ At the same time, for the related nitrobenzene dioxygenase (Figure 1C), calculations predicted favorable O–O cleavage in the hydroperoxy complex (**c1**) to yield an Fe(V)O(OH) complex (**c2**).¹⁴ Attack on the substrate via the bound OH leads to one-electron oxidation and produces an iron(IV)-oxo species with bound substrate radical (**c3**); in the subsequent step, the second C–O bond formation leads to the Fe(III)-bound dihydroxy product (**c4**). Isotope labeling experiments on naphthalene dioxygenase show some label incorporation from added water, only consistent with the latter mechanism, although only for the peroxide shunt reaction.¹⁵ Related bioinspired non-heme iron catalysts also react with double bonds via Fe(V) or Fe(IV)+radical active species.¹⁶ Notwithstanding the above, for benzoate dioxygenase, recent experimental evidence suggests involvement of a species at the Fe(III)-superoxo, instead of peroxy, level in substrate oxidation.¹⁷

Pterin-dependent aromatic hydroxylases,¹⁸ and the few known α -ketoglutarate-dependent aromatic hydroxylases as well (see Figure 1D),¹⁹ employ their cosubstrate and O₂ (**d1**) to form an iron(IV)-oxo active intermediate (**d2**). Radical attack on the arene yields a substrate radical complexed to iron(III) (**d3**); migration of a hydrogen atom or a carboxymethyl group then produces a ketone (**d4**), which eventually regains aromaticity by tautomerization.

In synthetic non-heme iron model complexes of pentadentate ligands, i.e., with only one open coordination site, end-on Fe(III)–OOH species can be formed (Figure 1E, **e1**). These were also found to react directly with aromatics. While the first identifiable intermediate corresponds to electrophilic attack formally by “OH⁺”, electronic structure studies show that the initial stage of the process is better described as O–O bond *homolysis* and radical attack on the ring, with the second electron being transferred to the incipient Fe(III)-oxyl radical.²⁰ The resulting cationic intermediate (**e2**) preferentially undergoes hydride shift to form ketone products (**e3**).²¹

Heme Fe(III)–OOH complexes in cytochrome P450 enzymes (Figure 1F, **f1**), on the other hand, were shown to be rather sluggish oxidants, poised instead toward heterolytic O–O bond breakage to form OH[–] and porphyrin radical-Fe(IV)=O species (**f2**).^{20,22} These latter complexes, in turn, can easily effect oxidation of aromatic rings, via a mixed electrophilic-radical attack (**f3**).²³ The electrophilic character was found to be higher in the presence of electron-donating ring substituents.^{23b,d}

The dicopper enzyme tyrosinase is also capable of hydroxylating arenes (Figure 1G).²⁴ O₂ reacts directly with the *deoxy* form of the enzyme (dicopper(I), **g1**) to yield a side-on $\mu\text{-}\eta^2\text{:}\eta^2\text{-peroxo-dicopper(II)}$ complex (**g2**). The incoming phenolic substrate, supposed but not yet proven to be bound to the copper center, is then hydroxylated via an electrophilic aromatic substitution mechanism.²⁵ In model systems, direct reaction of the $\mu\text{-}\eta^2\text{:}\eta^2\text{-peroxo-dicopper(II)}$ complex via concerted O–O cleavage and C–O bond formation (**g3**) or O–O cleavage to a bis($\mu\text{-oxo}$)-dicopper(III) species (**g4**) followed by attack on the ring were both shown to be viable pathways; it is not yet clear which mechanism is employed by the enzyme.²⁵ In subsequent steps, the dicopper(I) center is regenerated with concomitant formation of an *ortho*-quinone.

As apparent from the above, oxidation of aromatic rings by flavin, monoiron, and dicopper systems has been investigated extensively, and much knowledge has been gathered about the active intermediates. On the other hand, our understanding of aromatic oxidation by non-heme diiron enzymes has remained more limited. Non-heme diiron enzymes contain a pair of iron atoms in their active site, coordinated by carboxylate and histidine ligands, and they utilize molecular O₂ to accomplish a broad range of challenging reactions using diverse reactive species.²⁶ The most extensively studied members of this family carry out C–H abstraction, and even for this specific type of reaction, the active sites are fine-tuned to produce a series of different key intermediates. For example, in methane monooxygenase (MMO) oxidizing methane to methanol,²⁷ a bis($\mu\text{-oxo}$)-diiron(IV) species (termed intermediate **Q**) has been found to be responsible for methane C–H activation.²⁸ In contrast, soluble Δ^9 desaturase,²⁹ inserting a double bond into an alkyl chain by double hydrogen abstraction, seems to operate via a 1,1- $\mu\text{-hydroperoxo-diron(III)}$ intermediate,³⁰ whereas for *myo*-inositol oxygenase,³¹ a superoxo-diron(III) species is held responsible for the initial C–H cleavage.^{31,32}

Methane monooxygenase can oxidize π -systems of olefins and aromatics,³³ and so can several synthetic model non-heme diiron systems.³⁴ On the other hand, non-heme diiron enzymes natively hydroxylating aromatic substrates also exist; long known examples include phenol hydroxylase,³⁵ toluene 4-monooxygenase (T4MO)³⁶ and toluene/*o*-xylene monooxygenase (ToMO).³⁷ For T4MO/ToMO, experiments could not identify a **Q**-like diiron(IV) intermediate, in variance with MMO. Conformational changes of a threonine residue near the active site (Thr201), and the resulting changes in the hydrogen bonding network were identified as key reasons for a behavior different from MMO.³⁸ Instead of **Q**, studies of T4MO and ToMO point to a kinetically competent peroxo-type intermediate in the oxidation.^{37b,39,40}

Concerning the mechanism and the possibility of peroxo-based oxidants, de Visser carried out a computational study on propene epoxidation by a $\mu\text{-oxo-}\mu\text{-peroxo-diron(III)}$ complex.⁴¹ He concluded that the starting 1,2- $\mu\text{-peroxo}$ species first converts to a 1,1- $\mu\text{-peroxo}$ structure, which then attacks the substrate. Still, the barrier was found to be quite high, and the author postulated that other oxygenated forms must be responsible for the experimentally observed epoxidation reactivity.

Recently, two new members of the non-heme diiron enzyme family have been identified, benzoyl coenzyme A epoxidase (BoxB, see Figure 2)^{42,43} and the related phenylacetyl coenzyme A epoxygenase (PaaA).⁴⁴ These enzymes convert their coenzyme A-bound aromatic substrates to epoxides, which represents a key step in one of the routes of the bacterial degradation of aromatic compounds.⁴⁵ This route, only requiring molecular oxygen in the aromaticity-breaking, epoxide-forming step, is operative under low or fluctuating O₂ concentrations, and thereby complements alternative routes via ring-cleaving oxygenases (when O₂ is abundant)⁴⁶ or via the “biochemical Birch-reduction” (when O₂ is absent).⁴⁷

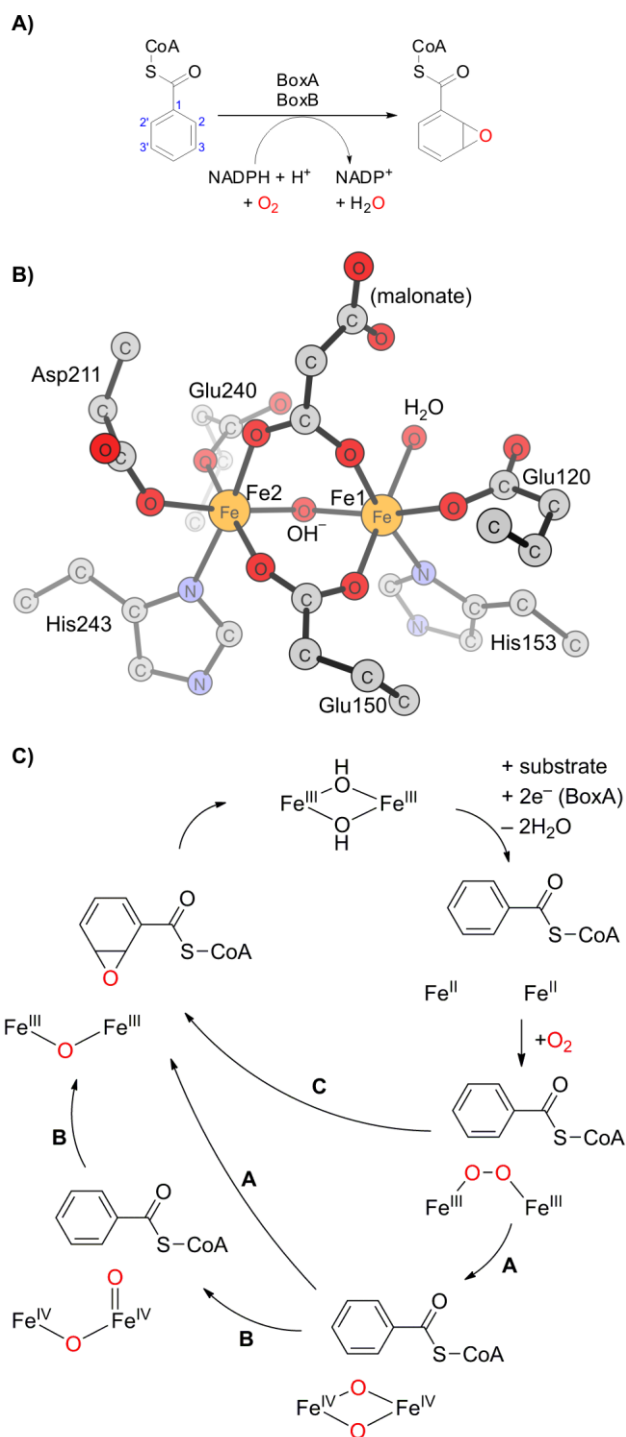


Figure 2 A) Reaction catalyzed by the BoxAB enzyme system, where BoxB is an aromatic epoxidase and BoxA is a NADPH:BoxB oxidoreductase. B) Active site of the BoxB enzyme as seen in the X-ray crystal structure (PDB 3PF7). Hydrogen atoms are omitted; malonate is a crystallization ligand. C) Catalytic cycle proposed in the literature for benzoyl coenzyme A epoxidation.

For the BoxB enzyme, on the basis of the available experimental data and analogies with other non-heme diiron enzymes, Ermler and coworkers suggested the following mechanism (Figure 2C).⁴² The resting state was proposed to be the bis(μ -hydroxo)-diiron(III) form. It is converted to a diiron(II) species upon

substrate binding and reduction by the reductase component (BoxA) of the enzyme complex, which then reacts with O₂ to yield a peroxo-diiron(III) intermediate. The O–O cleavage reaction was supposed to provide a bis(μ -oxo)-diiron(IV) “diamond core” species (path A), which may also convert further to an oxo- μ -oxo-diiron(IV) “open core” complex (path B). Both intermediates may carry out a radical attack on the aromatic ring. Alternatively, the possibility of a concerted attack by a peroxo-level species was also formulated (path C). Nucleophilic attack to an aromatic ring was suggested to be improbable, and due to the electron withdrawing nature of the thioester substituent, electrophilic attack was suggested to be disfavored over the radical pathway.

Liao and Siegbahn undertook a detailed computational investigation of the epoxidation mechanism in BoxB.⁴⁸ Using a ~200 atom cluster model of the enzyme, they came to the conclusion that a μ - η^2 : η^2 -peroxo-diiron(III) complex is the most stable form of the O₂-reacted enzyme. They found that this species can directly attack the aromatic ring, and the O–O cleavage and both C–O bond formation steps proceed in concerted way (essentially corresponding to path C above). Importantly, they claimed that no stable bis(μ -oxo)-diiron(IV) intermediate can be located, which points to an important difference from MMO where this species is fairly stable and forms spontaneously,²⁸ and even from T4MO/ToMO where it is unstable yet identifiable in computations.³⁸ Several other aspects of BoxB reactivity (e.g., product isomerization to oxepin and phenol; product deoxygenation) were also covered.

As apparent, several questions about arene oxidation via non-heme diiron systems have been left unanswered by the computational studies published so far. For the BoxB active site and for a μ -oxo- μ -peroxo-diiron model complex, conceptually different epoxidation reaction pathways have been located, and the reasons for the difference have not been discussed. Reactivity of a Q-like intermediate toward aromatic rings has not been studied, nor has been any explanation given why it does not exist at all in BoxB. The main goal of the present work is to address these issues, so that a generalized picture of this facet of non-heme diiron chemistry can be established.

To this end, we describe herein our findings from a QM/MM computational modeling of substrate epoxidation process via BoxB. Upon searching for possible pathways for the substrate attack, we were able to identify and electronically characterize several routes, including ones via the 1,1- μ -peroxo and μ - η^2 : η^2 -peroxo-diiron(III) species, but we show that the lowest energy pathway involves the formation of a bis(μ -oxo)-diiron(IV) species. We discuss what determines the electrophilic or radical nature of the attack and the factors directing the reaction toward an epoxide instead of ketone or phenol species. We demonstrate how these results provide a framework into which the earlier computational studies^{41,48} on non-heme diiron epoxidation can be easily fit, and we point out possible mechanistic conclusions for toluene monooxygenase or other related enzymes.

2. Computational Details

Full details of protein preparation and of the applied computational protocol are provided in the Supporting Information. Here, a brief summary of the key points is given. The crystal structure of the substrate-bound monomeric BoxB enzyme (PDB 3PM5, 2.3 Å resolution, $R = 0.174$) was employed as the starting point of our studies.⁴² Coordinates for some residues were taken from the oxidized but substrate-free crystal structure (PDB 3PF7, 1.9 Å resolution, $R = 0.182$).⁴² Crystallization ligands were removed, and two hydroxy bridges between the irons were added to prepare the resting state dihydroxo-diiron(III) form of the enzyme.⁴² The protein was protonated using *reduce*⁴⁹ on the basis of PropKa 3.1 results⁵⁰ for a pH of 7.8. Considering irons in the +3 oxidation state, the resulting structure had a charge of -15 . Na^+ ions for neutralization and a 6 Å spherical water shell were added by *leap*.^{51,52} *leap* may not be able to fully solvate the inside cavities of the protein; nevertheless, as the active site is solvent-excluded (*vide infra*), this is not supposed to yield very significant errors.

We chose the two iron ions, the coordinating histidines and carboxylates, the hydroxide bridges, the substrate aromatic ring with the thioester linkage, and two water molecules as the QM region (82 atoms, see Figure 3). The “relaxed MM (RMM) region”, involved in subsequent geometry optimizations, was chosen to contain all residues within 6.0 Å of the QM region (1008 atoms in total). Heavy atoms beyond this region were fixed to their crystallographic positions. In the prepared protein, the QM+RMM region, all added waters (but not crystal waters) and Na^+ ions, and all hydrogen atoms were first minimized, then annealed in a short molecular dynamics run, and minimized again to yield the final structure for the resting state [bis(μ -hydroxo)-diiron(III)] enzyme, containing 36360 atoms. The oxygenated intermediates [e.g., bis(μ -oxo)-diiron(IV)] of interest were prepared by removing the two hydroxo hydrogen atoms, followed by geometry optimization. The resulting QM regions thus contained 80 atoms and had a net charge of zero. Antiferromagnetic coupling and high-spin state for the irons were assumed throughout, consistently with what has been found for oxo- or hydroxo-bridged iron(III)/iron(IV) structures in non-heme diiron enzymes.⁵³

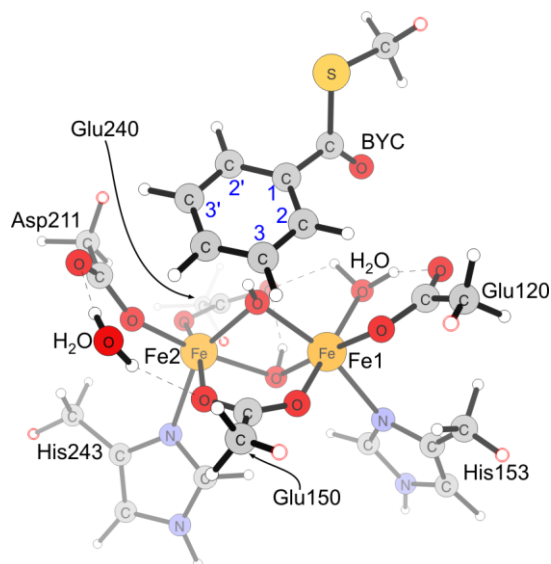


Figure 3 QM region of the resting state bis(μ -hydroxo)-diiron(III) structure. Link hydrogen atoms are denoted by red circles. The benzoyl coenzyme A substrate is denoted by BYC.

Geometry optimizations of the minima and transition states (TSs) were carried out using the mechanical embedding ONIOM method,⁵⁴ with part of the MM charges of the QM atoms iteratively updated,^{55,56} using a quadratically convergent microiterative algorithm.⁵⁷ In these runs, only residues within 6.0 Å of the QM+RMM region were included (1008 optimized atoms, 4309 atoms in total for the resting state). The QM region was treated using the B3LYP-D3 functional^{58,59} (with functional form V of the VWN local correlation^{58a}) and the def2-SV(P) basis set.⁶⁰ In the MM region, the Amber ff12SB force field⁶¹ and the gaff force field⁶² were employed for the protein and substrate, respectively. Following the geometry optimization, QM/MM analytical frequency calculations for the QM+RMM region were carried out on these models (again, 1008 atoms involved in the frequency calculations, 4309 atoms present in total for the resting state). All stationary points had the appropriate number of imaginary frequencies (0 for minima, 1 for TSs). From the transition states, intrinsic reaction coordinate⁶³ (IRC) calculations using the local quadratic approximation⁶⁴ were run to confirm their connection with the corresponding minima and to analyze reaction pathways. The MM charges of the QM atoms were not updated during the IRC runs.

Energies were then further refined by single-point calculations on the full solvated protein using a three-level⁶⁵ electrostatic-embedding ONIOM (QM/QM/MM) protocol, with B3LYP-D3/def2-TZVPP^{58,59,60} for the above shown QM atoms, BP86-D3/def2-SV(P)^{58c,60,66} as a middle layer containing 995 atoms, employing the multipole-accelerated resolution-of-identity approximation,⁶⁷ and ff12SB+gaff as the MM level. Beyond the explicitly included water molecules, continuum solvation effects were estimated using the MM representation of the system and a Poisson–Boltzmann model.⁶⁸ Contributions of the vibrational motions to the Gibbs free energy for $T = 298.15$ K were estimated from the analytic QM/MM frequencies, with the assumption that the contributions outside the QM+RMM region remain constant throughout the reaction pathways. No contributions from overall translation and rotation of the protein or of the active site

were included. The same approach was used to estimate kinetic isotope effects. Mulliken and natural population analyses⁶⁹ in the def2-SV(P) basis set were employed in several cases to examine the computed wavefunctions.

To analyze the effects of the enzyme environment, QM-only calculations were carried out for several structures by reoptimizing the isolated QM part in vacuo without constraints at the B3LYP-D3/def2-SV(P) level of theory and refining energies using single-point B3LYP-D3/def2-TZVPP computations, again in vacuo. Similarly to the enzymatic case, thermal contributions to free energies were calculated with excluding the overall translation and rotation.

The Amber package⁵¹ was employed for the molecular mechanics minimization and annealing runs, and Gaussian 09⁷⁰ and Turbomole⁷¹ were used for the (QM)/QM/MM calculations. During system preparation, use was made of the WHAT CHECK program,⁷² as well as some of the tools in the ComQum package.⁷³ Interconversion between Amber and Gaussian data files, QM/MM partitioning, and iterative charge updating were carried out using an in-house developed code called ‘gaumber’, available from the author upon request.

3. Results and Discussion

Oxygenated intermediates of BoxB. The first goal of this work was to study the possible oxygenated diiron intermediates and identify the one(s) responsible for the epoxidation of the aromatic ring of the benzoyl coenzyme A substrate. To this end, we optimized the geometries of several candidates and tested their reactivity toward the substrate. In accordance with the crystal structure of reduced BoxB,⁴² as well as with the consensus catalytic cycle of most of the non-heme diiron enzymes,²⁶ we assumed that these intermediates only contain the protein-derived ligands, the water molecule bound to Fe1, and the species derived from O₂, and that they correspond to an oxidation level diiron(III) + peroxide. This assumption is further corroborated by the fact that, as highlighted in the analysis of the crystal structure,⁴² the reaction is shielded from the bulk solvent once the substrate is bound. In several related enzymes, there is a conserved threonine residue near the active site (Thr201 in T4MO/ToMO, Thr213 in MMO), which is postulated to modulate the stability of the oxygenated intermediates^{38,74,75} and to play a role as a proton relay.⁴⁰ Importantly, this threonine is missing from BoxB; the amino acid in the analogous position is Gly214. Furthermore, in ToMO, a pore connects the active site with the solvent, delineated by polar amino acids (Asn202, Gln228),⁷⁶ while in BoxB, several apolar amino acids can be found in that region (Leu237, Leu244, Ala241). It seems therefore reasonable to suppose that the substrate oxidation itself proceeds without the involvement of exogenous protons.

The most relevant oxygenated diiron intermediate species identified in our computations, together with their relative Gibbs free energies and Mulliken spin populations are shown in Figures 4 and 5. As apparent from Figure 4, the peroxo moiety can adopt a variety of bridging coordination modes toward the irons,

including $\mu\text{-}\eta^2\text{:}\eta^2\text{-peroxo}$, $1,2\text{-}\mu\text{-peroxo}$, and $1,1\text{-}\mu\text{-peroxo}$.⁷⁷ For most species, two structures could be located, differing in the position of Glu240. This carboxylate is usually coordinated to Fe2 in a monodentate way, and its other oxygen may point toward the Fe1-bound water molecule (class W) or toward the Fe1-coordinated His153 (class H), with the potential to become bidentate (e.g., in H-1,2- $\mu\text{-peroxo}$). Similar behavior of the analogous carboxylate was also observed for MMO (Glu243),²⁶ but importantly, the enzyme framework seems to impose a larger iron–iron distance here, and BoxB Glu240 cannot occupy the bridging position as was observed in MMO. The (His)C–H \cdots O(Glu) interaction in the H isomers varies from strong ($d_{\text{H}\cdots\text{O}} = 1.97 \text{ \AA}$, $\text{CHO}\angle 176^\circ$ in H- $\mu\text{-}\eta^2\text{:}\eta^2\text{-peroxo}$) to weak ($d_{\text{H}\cdots\text{O}} = 2.21 \text{ \AA}$, $\text{CHO}\angle 127^\circ$ in H-oxo- $\mu\text{-oxo}$).

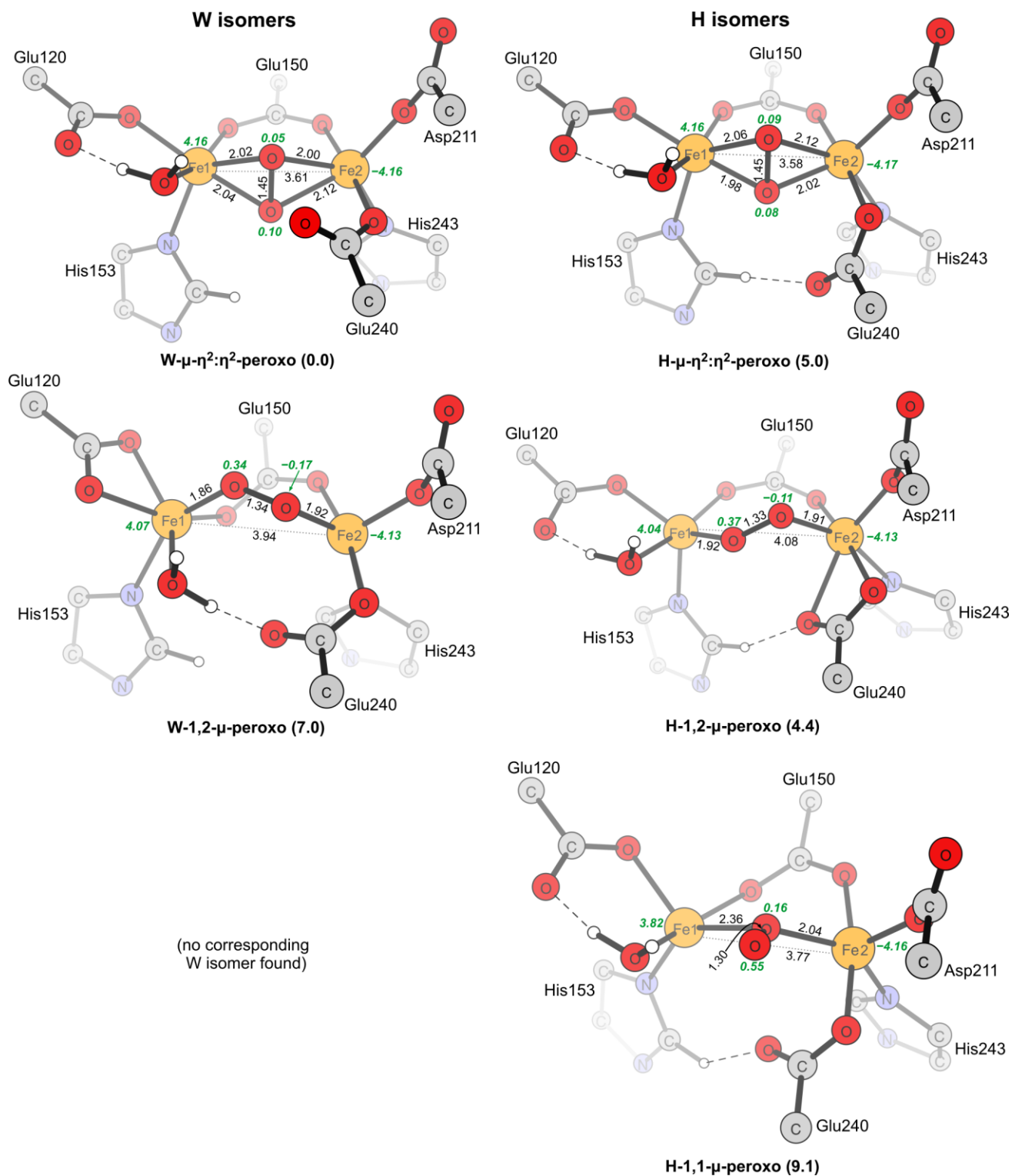


Figure 4 Structure of the most relevant oxygenated intermediates in BoxB containing an O–O bond. W and H isomers are classified on the basis of Glu240 orientation (toward Fe1-bound water or His153, respectively). Only a selection of the QM region atoms is shown. Computed relative free energies are given in parentheses in kcal/mol. Distances are given with normal font in Å; Mulliken atomic spin populations are shown in bold italic green font. Note that the view in Figures 4 and 5 is rotated by 180° with respect to the other figures in the paper to allow better depiction of Glu240 position.

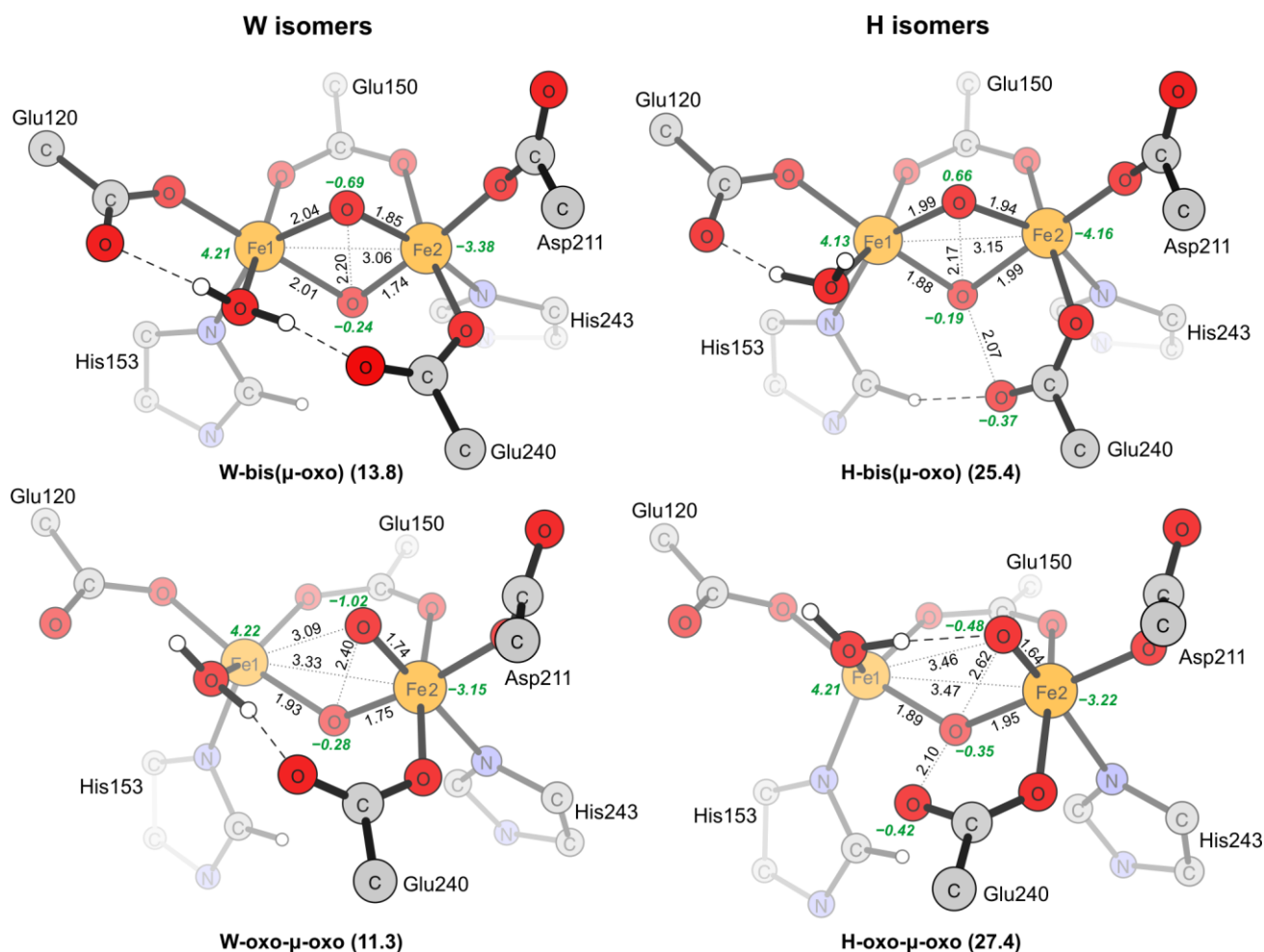


Figure 5 Structure of the most relevant oxygenated intermediates in BoxB without O–O bond. W and H isomers are classified on the basis of Glu240 orientation (toward Fe1-bound water or His153, respectively). Only a selection of the QM region atoms is shown. Computed relative free energies are given in parentheses in kcal/mol. Distances are given with normal font in Å; Mulliken atomic spin populations are shown in bold italic green font. Note that the view in Figures 4 and 5 is rotated by 180° with respect to the other figures in the paper to allow better depiction of Glu240 position.

The most stable isomer is the W- μ - η^2 : η^2 -peroxo structure, in accordance with the QM-cluster calculations by Liao and Siegbahn.⁷⁸ All free energies in this paper will be referenced to this structure. On the basis of the spin populations and the O–O distance, it can be characterized as a true peroxo-diiron(III) complex. The other identified coordination modes of the peroxo species (1,2- μ -peroxo and 1,1- μ -peroxo) are also low enough in free energy to be easily accessible as reactive intermediates. In these structures, the electron transfer from dioxygen to the irons is incomplete, and the resulting electronic structures are closer to superoxo-diiron(II,III).⁷⁹

Beyond the peroxo-type intermediates, we were also able to identify well-defined local minima with the O–O bond broken for both W and H classes (see Figure 5). Importantly, bis(μ -oxo) “diamond-core” intermediates, analogous to **Q** in MMO, can be located. The corresponding W isomer lies at $\Delta G = 13.8$ kcal/mol, which means that its formation is notably endergonic, but it may still be relevant as an active

species. Keeping in mind that the B3LYP functional tends to understabilize the **Q**-type structure,^{80,81} it can still be stated that W-bis(μ -oxo) (“BoxB-**Q**”) is destabilized with respect to the W- μ - η^2 : η^2 -peroxo structure in comparison with the analogous species of MMO, where the analogous energy difference was computed to be 0.7 kcal/mol.⁸⁰ The destabilization by the enzyme is also evident if we consider our unconstrained QM-only optimizations on BoxB, where the **Q**-like active site is only 7.8 kcal/mol higher in free energy than the W- μ - η^2 : η^2 -peroxo structure. One important structural feature of W-bis(μ -oxo), responsible for approximately half of this effect, is that the QM/MM optimized Fe \cdots Fe distance (3.06 Å) is larger than both the distance obtained in unconstrained QM-only optimization of the isolated active site (2.94 Å) and the QM/MM Fe \cdots Fe distance for MMO **Q** (2.84 Å)^{80,82}. As a consequence, the present QM/MM electronic structure is best described as μ -oxo- μ -oxyl-diiron(III,IV), in contrast to the fully oxidized diiron(IV) state of the unconstrained geometry as well as that of MMO **Q**.⁸³ For MMO, such mixed-valent structure was only observed when the C–H abstraction TS is approached.⁸⁴

Among the identified BoxB bis(μ -oxo) structures, another interesting feature is exhibited by the H-bis(μ -oxo) isomer: apparently, the negative charge on the pendant oxygen of Glu240 cannot coexist with the strong oxidant in the absence of the hydrogen bond with the Fe1-bound water, and a partial O–O bond between the oxo ligand and the carboxylate oxygen is formed as indicated by the fairly short O–O distance (2.07 Å) and the significant spin population gained by the carboxylate.⁸⁵ Importantly, in the synthetic non-heme monoiron system Fe(*S,S*-PDP),⁸⁶ a computational study⁸⁷ identified analogous geometric and electronic structure for an unstable local minimum as well as for a transient geometry en route to C–H activation.

Besides the bis(μ -oxo) isomers, open-core oxo- μ -oxo structures also exist as local minima; apart from the missing Fe1–O bond, their structural and electronic features are very similar to the corresponding bis(μ -oxo) intermediates. Open-core structures with oxo on Fe1 could not be located as stable minima; instead, they undergo barrierless substrate attack.

As apparent from the computed free energies, most of the presented oxygenated structures have low enough free energies to be possibly relevant for substrate oxidation. Interconversion among all these structures was not studied in detail; we only investigated two key processes. One of them is the interconversion between class W and class H isomers, which can obviously proceed via many transition states. We located the one connecting W- and H- μ - η^2 : η^2 -peroxo; the TS corresponds to the rotation of the carboxylate moiety of Glu240 and lies at a free energy of $\Delta G^\ddagger = 7.8$ kcal/mol with respect to W- μ - η^2 : η^2 -peroxo, demonstrating the feasibility of the interconversion. The other process we investigated was the O–O bond breakage. We found that both W- μ - η^2 : η^2 -peroxo and H- μ - η^2 : η^2 -peroxo can undergo O–O bond scission and yield the corresponding bis(μ -oxo) isomer; the transition states lie at $\Delta G^\ddagger = 17.6$ and 26.6 kcal/mol, respectively, with reference to W- μ - η^2 : η^2 -peroxo.⁸⁸ Their geometric and electronic structures do not possess any unexpected features; they are shown in Figure S8 in the Supporting Information.

All located oxygenated intermediates were tested to assess their capability of attacking the aromatic ring of the substrate and converting it to the epoxide. On the basis of these calculations, we could identify four electronically different pathways. These will be discussed in detail in the subsequent sections.

Epoxidation pathways: the Q pathway. The smallest free energy barrier for the attack on the benzoyl ring was found for the W-bis(μ -oxo) species (overall $\Delta G^\ddagger = 20.9$ kcal/mol from W- μ - η^2 : η^2 -peroxo), in line with the original suggestion of Ermler et al.⁴² In order to get insight into the mechanism of the epoxidation, the intrinsic reaction coordinate (IRC) has been followed from the TS to the minima. Importantly, the energy profile obtained in such calculations refers to the electronic (potential) energy at the level used for geometry optimization. These values thus differ from our best estimates for the free energy, which contain, e.g., large basis set, three-level electrostatic embedding, thermal, and continuum solvation corrections. Table S2 in the Supporting Information contains a comparison of the electronic energies and free energies for the species relevant to the discussion of IRC results. It is also to be noted that potential energies of the minima from IRC calculations may differ by 1-2 kcal/mol from the fully optimized values.⁸⁹

For the epoxidation reaction of the W-bis(μ -oxo) species, the energy profile along the IRC with key structural features, Mulliken spin populations and natural atomic charges at selected points of the path is depicted in Figure 6. Considering the profile and the structural data, it is apparent that the reaction has a single energetic barrier but proceeds in two stages: starting from the reactant (point 1), the C3–O bond is formed first (via the transition state at point 2), and the shoulder in the energy vs. reaction coordinate plot points to the existence of a “hidden intermediate” (point 3).⁹⁰ This hidden intermediate then transforms further via synchronous formation of the C2–O bond and breakage of the Fe2–O bond (via point 4) to yield the Fe1-bound epoxide product (point 5). A closer look at the spin population and charge data unveils that the W-bis(μ -oxo) reactant (point 1) has a pronounced oxyl character on the attacking oxygen atom (*vide supra*), and around the transition state (point 2), this is partially transferred to the substrate. However, upon formation of the first C–O bond and approaching the hidden intermediate (point 3), this radical character essentially disappears. In the hidden intermediate, both irons are already reduced to iron(III) and the substrate is present as an arenium cation. The second step is thus a heterolytic ring closure (via point 4) to yield the product (point 5) without any further change in oxidation states. The reaction can thus be described as an electrophilic attack with a small amount of transient radical character. The W-bis(μ -oxo) species can effect a two-electron oxidation of the aromatic substrate in spite of the electron-withdrawing thioester substituent, but the two electrons are removed in a slightly asynchronous fashion. Consistently with this picture, the α - and β -LUMOs of the W-bis(μ -oxo) species (see Figure S2 in Supporting Information) exhibit a clear *p*-contribution on the attacking oxygen atom, with somewhat larger coefficient in the α case. These orbitals closely resemble those computed for the related species in MMO.⁸⁴ From the involved intermediate, we term this direct two-electron oxidation of the aromatic ring

by the W-bis(μ -oxo) species the “Q pathway”. We note that all our computations searching routes from the open-core W-oxo- μ -oxo structure converged to TSs connecting the product with W-bis(μ -oxo).

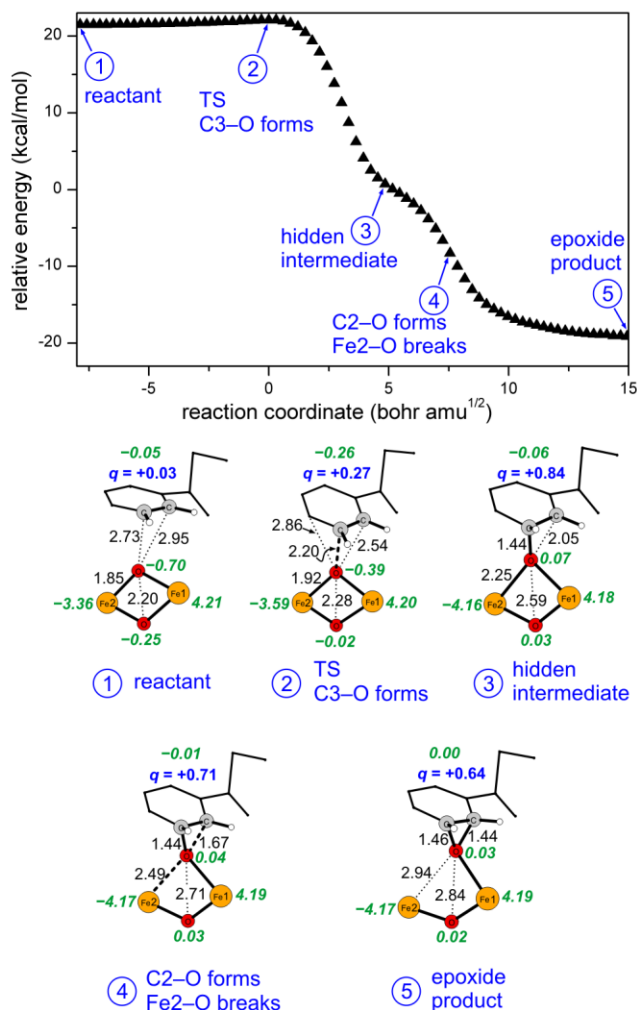


Figure 6 Energy profile along the intrinsic reaction coordinate for the aromatic oxidation by the W-bis(μ -oxo) species (Q pathway). Reported energy values are electronic (potential) energies relative to W- μ - η^2 : η^2 -peroxo computed at the level employed for geometry optimization (small-basis-set QM/MM). Selected structures on the path, labeled and denoted by blue numbers in circles, are also shown. O–O, Fe2–O, and C–O distances are given with normal font in Å; Mulliken atomic spin populations on the irons, oxygens, and on the whole substrate are shown in bold italic green font. The atomic charge (q) for the whole substrate (excluding the epoxide oxygen) is given in bold blue.

Epoxidation pathways: the $P\sigma^*$ pathway. In line with the report of Liao and Siegbahn,⁴⁸ we found that a pathway involving concerted O–O breakage and epoxide formation is also feasible, with an activation free energy of 23.9 kcal/mol for the reaction of W- μ - η^2 : η^2 -peroxo.⁹¹ As revealed by the energy profile and the structures along the intrinsic reaction coordinate (see Figure 7), this epoxidation reaction is also concerted and asynchronous. Starting from the reactant (point 1), O–O breakage and formation of the C3–O bond occur simultaneously (via TS point 2), yielding again a “hidden intermediate” (point 3), essentially

common with the **Q** pathway. Analogously to that case, the second C–O bond formation synchronously with Fe2–O bond breakage (via point 4) yield the epoxide product coordinated to Fe1 (point 5). A quick glance at the Mulliken spin populations of the irons shows that they remain in the +3 oxidation state throughout, in accordance with Liao’s and Siegbahn’s calculations.⁴⁸ There is only a small radical character on the oxygen atoms or on the substrate, and the hidden intermediate can be identified as an arenium cation, resulting from the electrophilic attack of the peroxo species.^{92,93} The geometry and the analysis of the orbitals of the transition state clearly show that the σ^* orbital of the peroxo moiety is responsible for the attack (see Figures S3 and S4 in Supporting Information); we thus term this pathway the **P** σ^* pathway. Importantly, the σ^* orbital lies quite high in energy in the stable W- μ - η^2 : η^2 -peroxo species (LUMO+10), but the observed elongation of the O–O distance (compare points 1 and 2) lowers its energy, and it becomes the LUMO of the peroxo-diiron moiety distorted to the TS geometry. Actually, the resulting O–O σ^* orbitals bear some similarity to the oxygen-centered *p*-type LUMOs of W-bis(μ -oxo) on the **Q** pathway. However, the iron-oxygen distances (particularly to Fe2) remain longer on the **P** σ^* pathway, which leads to a notably smaller participation of the iron *d* orbitals, i.e., a much less significant role in terms of redox chemistry.

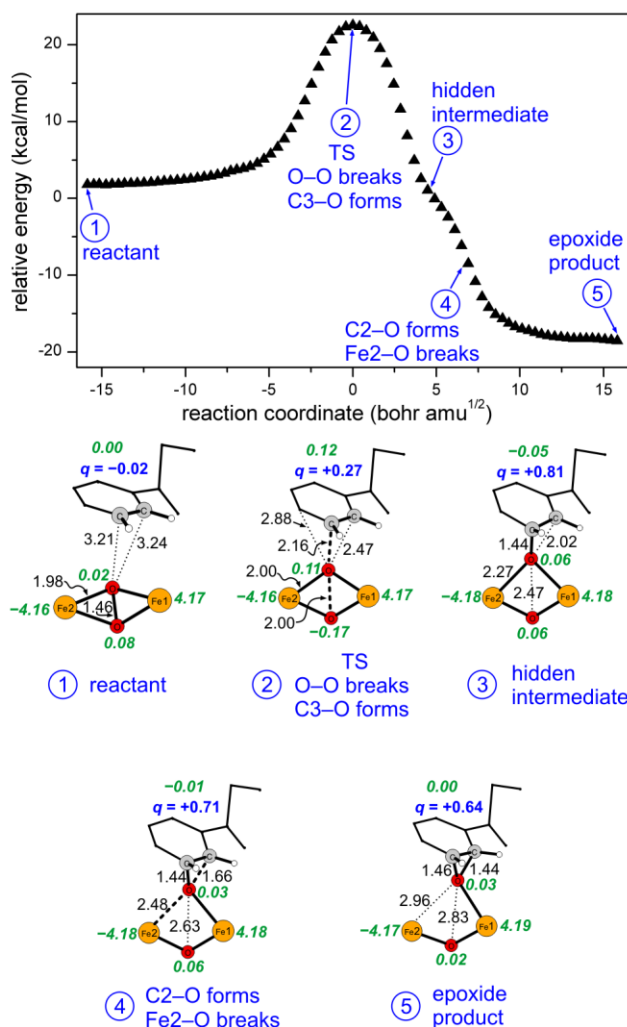


Figure 7 Energy profile along the intrinsic reaction coordinate for the aromatic oxidation by the $W-\mu-\eta^2:\eta^2$ -peroxo species ($P\sigma^*$ pathway). Reported energy values are electronic (potential) energies relative to $W-\mu-\eta^2:\eta^2$ -peroxo computed at the level employed for geometry optimization (small-basis-set QM/MM). Selected structures on the path, labeled and denoted by blue numbers in circles, are also shown. O–O, Fe2–O, and C–O distances are given with normal font in Å; Mulliken atomic spin populations on the irons, oxygens, and on the whole substrate are shown in bold italic green font. The atomic charge (*q*) for the whole substrate (excluding the epoxide oxygen) is given in bold blue.

Analogy with organic epoxidation. The minor importance of radical character in both the Q and $P\sigma^*$ pathways, and innocence of the irons in the $P\sigma^*$ pathway encourage one to seek analogies in “purely organic” systems. Indeed, the epoxidation of double bonds with organic peroxy acids⁹⁴ easily lends itself to making parallels. Experimentally, the reagent of choice is often *meta*-chloroperbenzoic acid for olefins⁹⁵ and peroxytrifluoroacetic acid for aromatics.⁹⁶ Nevertheless, with respect to the electronic structure, it is very instructive to consider some hypothetical model reactions as shown in Figure 8. Reaction A between ethylene and peroxoformic acid proceeds via a single concerted transition state, referred to as the “butterfly” TS in the literature.⁹⁷ In this reaction, the O–O bond is broken and the two C–

O bonds are formed in a synchronous manner (together with a proton transfer, which is irrelevant for the present discussion). In this “butterfly” transition state, the O–O bond is located over the midpoint of the double bond in a perpendicular end-on orientation. Natural bond orbital analysis confirms that the reactivity, as well as the specific TS geometry, is the result of two simultaneous electron donation processes: the C–C π^* accepts electrons from an appropriately oriented O lone-pair (LP; see part D of the figure), while the π -orbital of ethylene donates electrons to the O–O σ^* orbital (see part E). The balance of the electron donation processes can be altered toward both constituents. In Reaction B, when peroxyformate is substituted for peroxyformic acid, the O LP \rightarrow C–C π^* donation becomes dominating, which is clearly borne out by the geometry showing attack at one of the ethylene carbons, with a short O–O bond, the O lone pair in an sp^3 -like orientation, and the O–C–C angle of 99.7° being close to that of an ideal nucleophilic attack. On the other hand, in reaction C between peroxyformic acid and phenol, the C–C $\pi \rightarrow$ O–O σ^* donation becomes more important, resulting in a fairly long O–O distance, and an attack perpendicular to the aromatic ring, almost directly over one carbon atom instead of the bond midpoint to maximize the π/σ^* overlap.

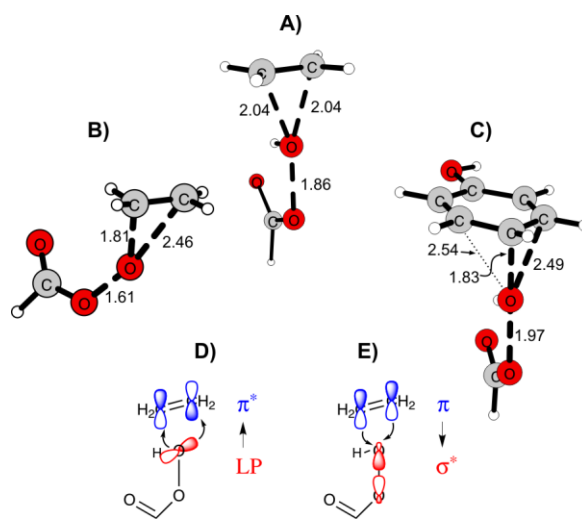


Figure 8 B3LYP-D3/def2-SV(P) optimized TSs for epoxidation of ethylene by peroxyformic acid (A), ethylene by peroxyformate (B), and phenol by peroxyformic acid (C), together with key electron donation processes (D and E).

In terms of geometry and key electron donation processes, the enzymatic **Q** and **P σ^*** pathways with the natural substrate fall somewhere between cases C and A. The dominating interaction is the donation of aromatic π electrons to the oxygen- p (and partially iron- d) type orbital in the **Q** pathway or to the O–O σ^* orbital in the **P σ^*** pathway; hence, the attack is electrophilic. Nevertheless, the donation in the reverse direction, from an O lone pair to the aromatic π^* , is also non-negligible, which is borne out by the notably unequal distances between the attacking O atom and the two aromatic carbons neighboring the attacked

one (2.86/2.54 Å for **Q**; 2.88/2.47 Å for **Pσ***; see point 2 in Figures 6 and 7, respectively, and compare with 2.54/2.49 Å for phenol/peroxoformic acid in Figure 8C).⁹⁸

Epoxidation pathways: the **Pπ* pathway.** De Visser showed⁴¹ that in a μ-oxo-μ-peroxo-bridged diiron model complex, propene epoxidation can be carried out by the 1,1-μ-peroxo isomer and produces a radical intermediate. For BoxB, we were also able to identify the analogous H-1,1-μ-peroxo structure, and as discussed above, it is best described as a superoxo-diiron(II,III) complex. We found that a direct attack of this species on the aromatic ring can be realized. The route with the lowest activation free energy for this isomer ($\Delta G^\ddagger = 35.8$ and 36.0 kcal/mol for the two involved steps from W-μ-η²:η²-peroxo, *vide infra*) is still notably disfavored over the above discussed pathways, and it eventually produces the 1,2-epoxide (in contrast to the experimentally observed 2,3-isomer).⁹⁹ Nevertheless, it serves well as an example for the discussion of the electronic features of the superoxo-type structures in aromatic oxidation. The energy profile and the most relevant structures are shown in Figure 9.¹⁰⁰ As apparent from the Figure, the reaction proceeds in two distinct steps. In the first step, the H-1,1-μ-peroxo reactant (point 1) attacks the substrate, and C–O bond formation occurs (TS at point 2, intermediate at point 3). The second step corresponds to the concerted, synchronous formation of the second C–O bond and cleavage of the O–O bond (TS at point 4) to yield the epoxide product, not bound to the irons (point 5). The depicted geometrical parameters and spin populations confirm that the reactant (point 1) is a superoxo-diiron(II,III) structure, which attacks the substrate via its single unfilled superoxo π*-orbital (point 2, see also Figure S5 in Supporting Information for a visualization of the orbital). As revealed by the increase in the O–O distance, the superoxo moiety is converted to peroxo, but the irons do not change their oxidation states, and an aryl radical-peroxo-diiron(II,III) intermediate is formed (point 3). This intermediate must then cross another barrier (point 4), where the second C–O bond is formed in concert with the O–O bond breakage. In this step, the O–O σ* orbital is filled with two electrons, one coming from the aryl radical, and the other one from Fe1 which is oxidized from iron(II) to iron(III), eventually yielding the diiron(III) product (point 5). No evidence of any arenium cation intermediate, hidden or real, was found in this case. In our analysis of de Visser's analogous transition state,⁴¹ we found identical features, indicating that the reactions proceed via a common mechanism; we term this pathway the **Pπ*** pathway.

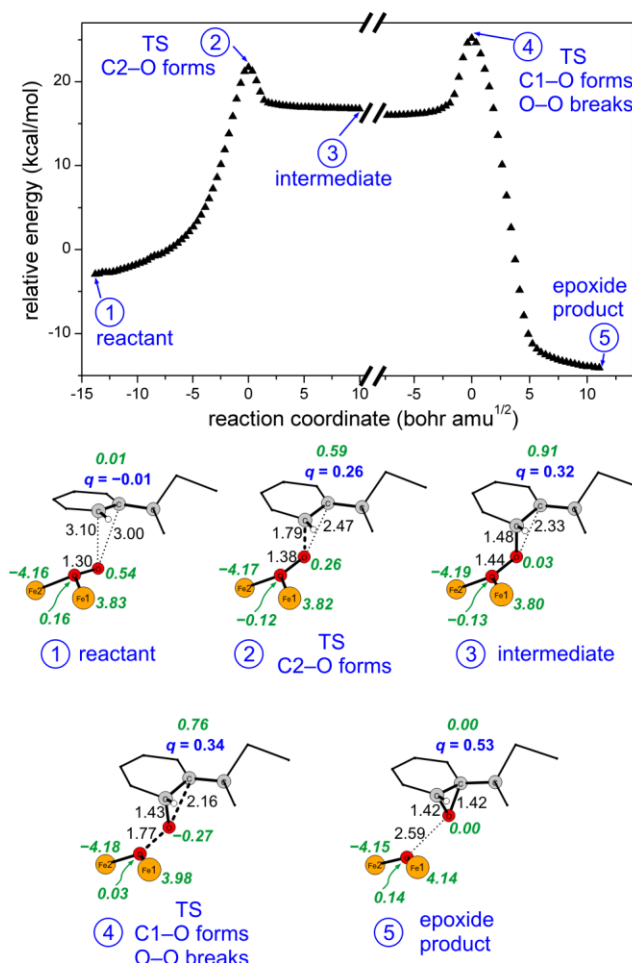


Figure 9 Energy profile along the intrinsic reaction coordinate for the aromatic oxidation by the H-1,1- μ -peroxo species ($\mathbf{P}\pi^*$ pathway). Reported energy values are electronic (potential) energies relative to W- μ - η^2 : η^2 -peroxo computed at the level employed for geometry optimization (small-basis-set QM/MM). Note that on this energy scale, unlike in free energy, the H-1,1- μ -peroxo reactant falls somewhat below the reference W- μ - η^2 : η^2 -peroxo. Selected structures on the path, labeled and denoted by blue numbers in circles, are also shown. O–O, and C–O distances are given with normal font in Å; Mulliken atomic spin populations on the irons, oxygens, and on the whole substrate are shown in bold italic green font. The atomic charge (q) for the whole substrate (excluding the epoxide oxygen) is given in bold blue.

As noted above, the BoxB active site can support several geometries with superoxo-diiron(II,III) electronic structure. Indeed, for all of them, we were able to find the corresponding transition states for C–O bond formation via the $\mathbf{P}\pi^*$ pathway. The lowest of those barriers was found for C2 attack by H-1,2- μ -peroxo ($\Delta G^\ddagger = 34.2$ kcal/mol), but it is by a mere 1.6 kcal/mol lower than for the above discussed H-1,1- μ -peroxo and still notably higher than for the \mathbf{Q} or $\mathbf{P}\sigma^*$ pathways.¹⁰¹ It is interesting to note that in the model complex studied by de Visser,⁴¹ the 1,2- μ -peroxo isomer has no superoxo character. Hence, the π^* orbital is not available to accept electrons, and the σ^* orbital is unreachable due to the 1,2-coordination mode; it is thus understandable why he did not find any favorable propene epoxidation route involving the 1,2- μ -peroxo complex.

Epoxidation pathways: the Q' pathway. In the H-bis(μ -oxo) and H-oxo- μ -oxo structures, as discussed above, partial quenching of the oxidizing power occurs due to interaction of one of the oxo bridges with the dangling Glu240 carboxylate. In accordance with this notion, these species were found to be reactive toward the substrate, but through barriers higher than those for W structures, the lowest lying at $\Delta G^\ddagger = 33.7$ kcal/mol for the C2-attack by H-bis(μ -oxo). A quick look at the alpha and beta LUMOs of H-bis(μ -oxo) (see Figure S6 in Supporting Information) indicates that one of them has significantly decreased p -character on the attacking oxygen atom; instead, it has the characteristics of a carboxylate-oxo σ^* orbital. Correspondingly, the reaction by H-bis(μ -oxo) starts with one-electron oxidation of the substrate, producing a radical intermediate where the carboxylate-oxo interaction still persists. In a second reaction step, cleavage of the carboxylate-oxo interaction⁸⁷ triggers the transfer of the second electron from the substrate, yielding an arenium cation hidden intermediate, ultimately collapsing to the oxo-diiron(III) species with the Fe1-bound epoxide. On the basis of several common features with the Q pathway, we term this possibility the Q' pathway and discuss it in more detail in the Supporting Information. We note that H-oxo- μ -oxo behaves analogously, but through quite high barriers (>45 kcal/mol).

Role and experimental relevance of the Q and P σ^* pathways. Our results, presented thus far, show that in BoxB, the P π^* and Q' pathways are quite high in energy; hence, they do not play a role in substrate epoxidation. In the most probable reaction route, the O–O bond of the lowest energy oxygenated intermediate W- μ - η^2 : η^2 -peroxo is broken first ($\Delta G^\ddagger = 17.6$ kcal/mol) to yield W-bis(μ -oxo), and the substrate is then oxidized by the W-bis(μ -oxo) intermediate via the Q-pathway ($\Delta G^\ddagger = 20.9$ kcal/mol, the highest point along the whole reaction route). The computed overall free energy barrier of 20.9 kcal/mol is not unrealistically high; unfortunately, no experimental kinetic data have been reported yet for a comparison. Direct reaction from W- μ - η^2 : η^2 -peroxo via the P σ^* pathway, also located by Liao and Siegbahn,⁴⁸ is higher in energy by a small amount (~ 3 kcal/mol).¹⁰² On the other hand, W-bis(μ -oxo) lies 13.8 kcal/mol above W- μ - η^2 : η^2 -peroxo and its back-conversion to that structure is predicted to be faster than the reaction with the substrate. Hence, experimentally, one could only identify W- μ - η^2 : η^2 -peroxo as the kinetically competent intermediate for substrate oxidation, regardless of whether it oxidizes the substrate via the Q or the P σ^* pathways. As discussed in the introduction, in T4MO/ToMO, a peroxo-type intermediate was indeed shown to be responsible for the oxidation; this raises the interesting question whether an analogous, destabilized diiron(IV) species is also present along the pathway of those enzymes.

Strategies for experimentally corroborating the role of W-bis(μ -oxo) or similar intermediates in BoxB or T4MO/ToMO could involve the employment of different substrates, for which various levels of oxidizing power are required.¹⁰³ On the other hand, properties of the rate-determining transition state could also be probed by experimental techniques. In this respect, our calculations predict that the primary ¹⁸O₂ kinetic isotope effect is appreciably higher for the direct P σ^* -type reaction from W- μ - η^2 : η^2 -peroxo (1.073) than for the Q-type reaction (1.033), and the (inverse) secondary ²H kinetic isotope effect for a pentadeuterated

benzoyl ring is also more significant on the $\text{P}\sigma^*$ pathway (0.886) than on the Q pathway (0.954), which both might serve as a diagnostic difference.

On the other hand, the existence of the low-energy $\text{P}\sigma^*$ pathway with redox-innocent iron(III) centers suggests that epoxidation of electron-deficient aromatics might be feasible using H_2O_2 activated by appropriately placed and tuned Lewis acidic centers. Although reconstitution of the enzyme with other metals may be non-trivial, synthesis of small-molecule aromatic oxidation catalysts based on redox-inactive metals seems to be an appealing possibility.¹⁰⁴ Such systems might have improved chemoselectivity due to suppressing side reactions mediated by iron(IV)/iron(V) or similar high-valent species.

It is interesting to note that, as mentioned above, the enzyme environment notably increases the free energy difference between W -bis(μ -oxo) and W - μ - η^2 : η^2 -peroxo (13.8 kcal/mol in the enzyme vs. 7.8 kcal/mol in the isolated, unconstrained model), but the effect on the free energy difference between the Q and $\text{P}\sigma^*$ TSs is much smaller (-3.1 kcal/mol in the enzyme vs. -0.6 kcal/mol in the isolated model). This behavior might be part of a strategy for suppressing side reactions by avoiding the formation of large amounts of unnecessarily powerful oxidants.¹⁰⁵ Hence, the actual reasons for this phenomenon might warrant further study, which is, however, beyond the scope of the present paper.

Regio- and stereoselectivity. In line with Liao's and Siegbahn's report,⁴⁸ we found that the kinetically favored product is the $2S,3R$ -epoxide on both the Q and $\text{P}\sigma^*$ pathways (with the already mentioned free energy barriers of 20.9 and 23.9 kcal/mol, respectively). Our calculations mapping all possible routes toward other isomers only provided three more transition states, corresponding to the Q and $\text{P}\sigma^*$ pathway toward the $2R,3S$ -epoxide (24.0 and 29.0 kcal/mol), and to the Q pathway toward the $3S,4R$ -epoxide (24.6 kcal/mol). The selectivity toward the 2,3-isomer is in line with the available experimental information.^{43a}

Origin of chemoselectivity toward the epoxide. As the result of the electrophilic attack of an oxygen species to an aromatic ring, an arenium ion is first formed. The hydrogen on the attacked carbon can then be removed as a proton, leading directly to a phenolate, or it can undergo a 1,2-hydride shift, known as the NIH shift,¹⁰⁶ which produces a cyclohexadienone, later regaining aromaticity by tautomerization. Alternatively, formation of a second C–O bond can lead to an epoxide. For the BoxB enzyme with the natural benzoyl coenzyme A substrate, we above found that the arenium ion intermediate cannot be located as a local minimum; instead, it spontaneously collapses into the product epoxide. In contrast to an arenium ion, production of phenol via proton transfer or hydride shift from the epoxide is significantly more difficult because one of the C–O bonds must be at least partially broken first. The barrierless collapse to the epoxide is thus an important contribution to achieve chemoselective oxygenation, and its origin is worth investigating.

Isolated, unconstrained active site structures with the model for natural substrate (PhC(=O)SMe) behave analogously to the full enzyme for both the **Q** and **P σ^*** pathways. In contrast, with other substrates (benzene, nitrobenzene, phenol), the arenium cation does not immediately collapse to the epoxide, but instead undergoes geometry relaxation, and subsequently, it is spontaneously deprotonated by one of the coordinating carboxylates to yield a phenolate complexed to the irons. To identify the reason behind this behavior, it is useful to consider the geometry of the transition states for substrate attack (see Figure 10). As revealed from the O–C(neighboring) distances for both the **Q**-type and **P σ^*** -type TSs, the reaction of the natural substrate proceeds in the most asymmetric geometry. In light of Figure 8 and its discussion above, a large asymmetry corresponds to a large contribution from O lone pair (LP) \rightarrow aromatic π^* donation. It thus seems that the thioacyl substituent is very effective in making the π^* system more susceptible to nucleophilic attack, even more so than the nitro group. The marked O LP \rightarrow π^* electron donation quickly turns into bond formation along the reaction pathway.

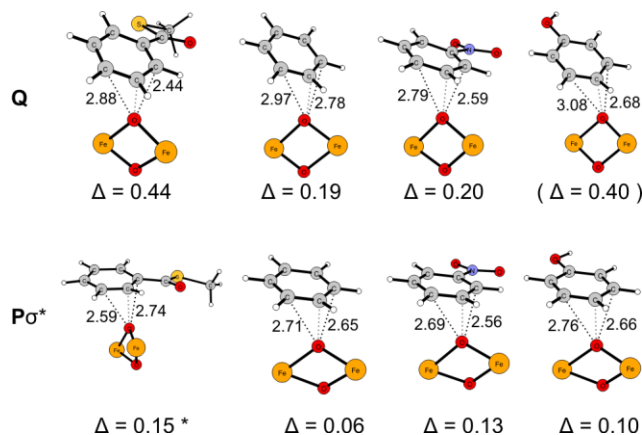


Figure 10 O–C(neighboring) distances and their difference (denoted by Δ , all in Å) in optimized TS geometries of isolated, unconstrained active site models with (from left to right) natural substrate, benzene, nitrobenzene, and phenol, presented for **Q**-type (top) and **P σ^*** -type (bottom) attacks. Note that **Q**/phenol is essentially barrierless, and a real TS could hardly be located, the geometry may thus not reflect electronic structure very accurately (Δ value given in parentheses). Note furthermore that in the isolated **P σ^*** /natural substrate TS, C2–O bond forms first, in contrast to C3–O in the enzyme (Δ value marked by asterisk).¹⁰⁷

The enzyme enforces the asymmetry, and thus high relative importance of the O LP \rightarrow π^* donation, also by presenting the substrate in a tilted geometry. Hence, it can be concluded that the formation of the epoxide becomes favored due to a combination of the electronic properties of the thioacyl substituent and the orientation of the substrate with respect to the active site. On the other hand, formation and deprotonation of the arenium ion intermediate might thus be the major pathway for other substrates such as toluene in T4MO/ToMO, where the occurrence of a crucial arenium cation intermediate was indeed suggested on the basis of docking calculations;¹⁰⁸ although epoxide ring opening mechanism is supported by isotope effect studies.¹⁰⁹ It is interesting to note that the phenylacetyl coenzyme A epoxidase enzyme (Paa),⁴⁴ closely related to Box, produces a 1,2-epoxide with no evidence for phenol formation. Our preliminary studies using HCOOOH as the model oxidant indicate that the O LP \rightarrow π^* donation can be

also quite significant for the PhCH₂C(=O)SMe model substrate as well if the initial electrophilic attack occurs in the C1 (*ipso*) position, but it seems weak for the C2 (*ortho*) attack. Clearly, a more detailed study of the Paa enzyme is warranted to identify the reasons for its epoxide selectivity.

4. Conclusions

QM/MM calculations at the B3LYP-D3/Amber-ff12 level were carried out to characterize the oxygenated intermediate(s) responsible for substrate epoxidation in the non-heme diiron enzyme benzoyl coenzyme A epoxygenase BoxB. From the analysis of the geometric and electronic features of the identified structures and reaction pathways, both in the enzyme and on model complexes, a description of the routes for attacking aromatic rings by diiron enzymes emerged, whose validity is probably not limited to the BoxB enzyme. We found four general pathways for oxidizing aromatic rings:

1. The **Q** pathway involves a bis(μ -oxo)-diiron(IV) species. This intermediate can attack the ring in an almost barrierless process. The transition state corresponds to the first C–O bond formation, and in this process, two electrons are transferred, somewhat asynchronously, from the substrate to the irons. The attack is thus electrophilic with some transient radical character; the resulting arenium cation may or may not correspond to a local minimum, and it may undergo epoxide ring closure, deprotonation, or hydride shift depending on the substituents and the constraints imposed by the enzyme.
2. The **P σ^*** pathway corresponds to the direct electrophilic attack of a μ - η^2 : η^2 -peroxo (or μ - η^2 : η^1 -peroxo) species on the aromatic ring. Synchronous C–O bond formation and O–O bond breakage lead again to an arenium cation, whose properties and further possible transformations are common with the **Q** pathway. The iron atoms remain redox-neutral throughout the process, which suggests the interesting possibility of catalysis by redox-inactive Lewis acids.
3. The **P π^*** pathway can be operative from a superoxo species, having a single unfilled π^* orbital available for a radical attack on the aromatic ring. In the obtained aryl radical-peroxo-diiron(II,III) intermediate, the peroxo moiety is then cleaved, oxidizing the iron(II) and the substrate to finally yield the epoxide and a μ -oxo-diiron(III) species. No arenium cation seems to play a role in this case.
4. The **Q'** pathway is characteristic for a variant of the bis(μ -oxo) intermediate where some of the oxidizing power is quenched by partial O–O bond formation between the oxo bridge and a dangling carboxylate. First, the aromatic ring of the substrate is oxidized in a one-electron process, yielding an aryl radical intermediate bound to the oxo bridge. Upon cleavage of the partial O–O bond, the second electron from the substrate is removed, and the produced arenium ion may then be transformed to the epoxide or phenol product.

This classification also helped put earlier studies into a broader perspective. De Visser's work,⁴¹ describing propene epoxidation by a 1,1- μ -peroxo- μ -oxo-diiron species, shows an example of the **P** π^* pathway, while Liao and Siegbahn⁴⁸ presented results for the BoxB **P** σ^* pathway. Furthermore, the related dicopper systems showing tyrosinase activity via the bis(μ -oxo)-dicopper(III) and μ - η^2 : η^2 -peroxo-dicopper(II) species bear many analogies with the **Q** and **P** σ^* pathways, respectively.^{24,25}

In addition, the calculations allowed to draw the following conclusions concerning the BoxB enzyme specifically:

1. The most stable oxygenated intermediate is the W- μ - η^2 : η^2 -peroxo structure, but several other isomers, including the **Q**-like W-bis(μ -oxo) have only somewhat higher free energy and are thus accessible as possible reactive intermediates.
2. The most favorable route for epoxidation of the benzoyl coenzyme A substrate proceeds via the W-bis(μ -oxo) species, thus via the **Q** pathway, and requires a total activation free energy of 20.9 kcal/mol. W-bis(μ -oxo) is destabilized by the enzyme with respect to W- μ - η^2 : η^2 -peroxo and easily interconverts with it; experiments would therefore identify the latter as the kinetically competent epoxygenating species. While no experiments are available for BoxB, such observations have indeed been made for T4MO/ToMO.
3. W- μ - η^2 : η^2 -peroxo can also directly epoxidize the substrate via the **P** σ^* pathway, but the barrier is somewhat higher than for the **Q** pathway, consistently within the enzyme and on active site models.¹⁰²
4. **P** π^* and **Q'** pathways are possible but disfavored for BoxB.
5. In the course of epoxygenation via the electrophilic **Q** and **P** σ^* pathways within BoxB or its active site models, the arenium cation only appears as a hidden intermediate, and it collapses without barrier to the epoxide. Electronic properties of the thioacyl substituent favoring electron donation from the attacking oxygen lone pair to the aromatic π^* , further enhanced by enzyme-imposed constraints, are responsible for this behavior, ensuring chemoselectivity of the oxygenation. Preliminary studies indicate that this effect might also be operative in the case of the related phenylacetyl coenzyme A epoxygenase. Reactions of enzymes oxygenating other aromatic substrates, such as T4MO/ToMO, may involve more stable arenium cation intermediates that undergo deprotonation instead of C–O bond formation, thus avoiding epoxide formation.

Although many aspects of the oxygenation of aromatic rings have been considered, more work has to be done in particular to understand the factors leading to the preference of the given pathways, and to uncover the mechanistic details and selectivity-determining factors in the related T4MO/ToMO and Paa enzymes. Computational research along these lines is in progress in our laboratory.

Acknowledgements

I am indebted to Dr. Imre Pápai and Dr. Lubomír Rulišek for their continuous support in starting my independent research career. Valuable comments on the manuscript from them and from Dr. Martin Srnec are also gratefully acknowledged. This work was supported by the Hungarian Scientific Research Fund (OTKA, grant no. PD 108955). The NIIF Institute is acknowledged for granting access to computational resources at Szeged, Hungary.

Supporting Information

Details of protein preparation, and computational protocol; O–O cleavage transition states; potential vs. free energy comparison relevant for IRC calculations; description of **Q'** pathway; LUMO figures; description and Cartesian coordinates of all stationary points from QM/MM and QM-only calculations.

-
- ¹ Renata, H.; Wang, Z. J.; Arnold, F. H. *Angew. Chem., Int. Ed.* **2015**, *54*, 3351.
- ² (a) Que Jr, L.; Tolman, W. B. *Nature* **2008**, *455*, 333. (b) Holtmann, D.; Fraaije, M. W.; Arends, I. W. C. E.; Opperman, D. J.; Hollmann, F. *Chem. Commun.* **2014**, *50*, 13180. (c) Fuchs, G.; Boll, M.; Heider, J. *Nat. Rev. Microbiol.* **2011**, *9*, 803. (d) Di Gennaro, P.; Bargna, A.; Sello, G. *Appl. Microbiol. Biotechnol.* **2011**, *90*, 1817. (e) Ullrich, R.; Hofrichter, M. *Cell. Mol. Life Sci.* **2007**, *64*, 271.
- ³ Lindhorst, A. C.; Haslinger, S.; Kühn, F. E. *Chem. Commun.* **2015**, *51*, 17193.
- ⁴ For a recent review of non-heme monoiron and diiron biochemistry from this perspective, see: Rokob, T. A.; Chalupský, J.; Bím, D.; Andrikopoulos, P. C.; Srnec, M.; Rulišek, L. *J. Biol. Inorg. Chem.* **2016**, *21*, 619.
- ⁵ Concerning cytochrome P450 systems, see: (a) Meunier, B.; de Visser, S. P.; Shaik, S. *Chem. Rev.* **2004**, *104*, 3947. (b) De Visser, S. P. *Adv. Inorg. Chem.* **2012**, *64*, 1.
- ⁶ Huijbers, M. M. E.; Montersino, S.; Westphal, A. H.; Tischler, D.; Van Berkel, W. J. H. *Arch. Biochem. Biophys.* **2014**, *544*, 2.
- ⁷ Visitsathawong, S.; Chenprakhon, P.; Chaiyen, P.; Surawatanawong, P. *J. Am. Chem. Soc.* **2015**, *137*, 9363.
- ⁸ (a) Ortiz-Maldonado, M.; Ballou, D. P.; Massey, V. *Biochemistry* **1999**, *38*, 8124. (b) Ridder, L.; Mulholland, A. J.; Rietjens, I. M. C. M.; Vervoort, J. *J. Mol. Graph. Model.* **1999**, *17*, 163. (c) Ridder, L.; Harvey, J. N.; Rietjens, I. M. C. M.; Vervoort, J.; Mulholland, A. J. *J. Phys. Chem. B* **2002**, *107*, 2118.
- ⁹ Bach, R. D. *J. Phys. Chem. A* **2011**, *115*, 11087.
- ¹⁰ Barry, S. M.; Challis, G. L. *ACS Catal.* **2013**, *3*, 2362.
- ¹¹ (a) Costas, M.; Mehn, M. P.; Jensen, M. P.; Que Jr., L. *Chem. Rev.* **2004**, *104*, 939. (b) Abu-Omar, M. M.; Loaiza, A.; Hontzeas, N. *Chem. Rev.* **2005**, *105*, 2227.
- ¹² Bassan, A.; Blomberg, M. R. A.; Siegbahn, P. E. M. *J. Biol. Inorg. Chem.* **2004**, *9*, 439.
- ¹³ Liu, L. V.; Hong, S.; Cho, J.; Nam, W.; Solomon, E. I. *J. Am. Chem. Soc.* **2013**, *135*, 3286.
- ¹⁴ Pabis, A.; Geronimo, I.; Paneth, P. *J. Phys. Chem. B* **2014**, *118*, 3245.
- ¹⁵ Wolfe, M. D.; Lipscomb, J. D. *J. Biol. Chem.* **2003**, *278*, 829.
- ¹⁶ (a) Oloo, W. N.; Que Jr., L. *Acc. Chem. Res.* **2015**, *48*, 2612. (b) Ansari, A.; Kaushik, A.; Rajaraman, G. *J. Am. Chem. Soc.* **2013**, *135*, 4235.
- ¹⁷ Rivard, B. S.; Rogers, M. S.; Marell, D. J.; Neibergall, M. B.; Chakrabarty, S.; Cramer, C. J.; Lipscomb, J. D. *Biochemistry* **2015**, *54*, 4652.
- ¹⁸ Olsson, E.; Teigen, K.; Martinez, A.; Jensen, V. R. *Adv. Inorg. Chem.* **2010**, *62*, 437.
- ¹⁹ (a) Moran, G. R. *Arch. Biochem. Biophys.* **2005**, *433*, 117. (b) Hausinger, R. P. *Crit. Rev. Biochem. Mol. Biol.* **2004**, *39*, 21. (c) Wójcik, A.; Broclawik, E.; Siegbahn, P. E. M.; Lundberg, M.; Moran, G.; Borowski, T. *J. Am. Chem. Soc.* **2014**, *136*, 14472.
- ²⁰ Faponle, A. S.; Quesne, M. G.; Sastri, C. V.; Banse, F.; de Visser, S. P. *Chem. Eur. J.* **2015**, *21*, 1221.
- ²¹ Faponle, A. S.; Banse, F.; de Visser, S. P. *J. Biol. Inorg. Chem.* **2016**, *21*, 453.
- ²² (a) De Visser, S. P. *Adv. Inorg. Chem.* **2012**, *64*, 1. (c) Shaik, S.; Cohen, S.; Wang, Y.; Chen, H.; Kumar, D.; Thiel, W. *Chem. Rev.* **2010**, *110*, 949.
- ²³ (a) Bathelt, C. M.; Mulholland, A. J.; Harvey, J. N. *J. Phys. Chem. A* **2008**, *112*, 13149. (b) Bathelt, C. M.; Ridder, L.; Mulholland, A. J.; Harvey, J. N. *J. Am. Chem. Soc.* **2003**, *125*, 15004. (c) De Visser, S. P.; Shaik, S. *J. Am. Chem. Soc.* **2003**, *125*, 7413. (d) Shaik, S.; Milko, P.; Schyman, P.; Usharani, D.; Chen, H. *J. Chem. Theory Comput.* **2011**, *7*, 327.
- ²⁴ Solomon, E. I.; Heppner, D. E.; Johnston, E. M.; Ginsbach, J. W.; Cirera, J.; Qayyum, M.; Kieber-Emmons, M. T.; Kjaergaard, C. H.; Hadt, R. G.; Tian, L. *Chem. Rev.* **2014**, *114*, 3659.
- ²⁵ For a selection of papers, see: (a) Qayyum, M. F.; Sarangi, R.; Fujisawa, K.; Stack, T. D. P.; Karlin, K. D.; Hodgson, K. O.; Hedman, B.; Solomon, E. I. *J. Am. Chem. Soc.* **2013**, *135*, 17417. (b) Rolff, M.; Schottenheim, J.; Decker, H.; Tuzcek, F. *Chem. Soc. Rev.* **2011**, *40*, 4077. (c) Liu, Y. F.; Yu, J. G.; Siegbahn, P. E. M.; Blomberg, M. R. A. *Chem. Eur. J.* **2013**, *19*, 1942. (d) Mirica, L. M.; Vance, M.; Rudd, D. J.; Hedman, B.; Hodgson, K. O.; Solomon, E. I.; Stack, T. D. P. *Science* **2005**, *308*, 1890. (e)

- Poater, A.; Ribas, X.; Llobet, A.; Cavallo, L.; Solà, M. *J. Am. Chem. Soc.* **2008**, *130*, 17710. (f) Inoue, T.; Shiota, Y.; Yoshizawa, K. *J. Am. Chem. Soc.* **2008**, *130*, 16890. (g) Holt, B. T. O.; Vance, M. A.; Mirica, L. M.; Heppner, D. E.; Stack, T. D. P.; Solomon, E. I. *J. Am. Chem. Soc.* **2009**, *131*, 6421. (h) Pidcock, E.; Obias, H. V.; Zhang, C. X.; Karlin, K. D.; Solomon, E. I. *J. Am. Chem. Soc.* **1998**, *120*, 7841. (i) Naka, H.; Kondo, Y.; Usui, S.; Hashimoto, Y.; Uchiyama, M. *Adv. Synth. Catal.* **2007**, *349*, 595. (j) Sander, O.; Henß, A.; Näther, C.; Würtele, C.; Holthausen, M. C.; Schindler, S.; Tuczek, F. *Chem. Eur. J.* **2008**, *14*, 9714.
- ²⁶ Sazinsky, M. H.; Lippard, S. J. *Acc. Chem. Res.* **2006**, *39*, 558.
- ²⁷ (a) Tinberg, C. E.; Lippard, S. J. *Acc. Chem. Res.* **2011**, *44*, 280. (b) Wang, W.; Liang, A. D.; Lippard, S. J. *Acc. Chem. Res.* **2015**, *48*, 2632.
- ²⁸ Banerjee, R.; Proshlyakov, Y.; Lipscomb, J. D.; Proshlyakov, D. A. *Nature* **2015**, *518*, 431.
- ²⁹ Fox, B. G.; Lyle, K. S.; Rogge, C. E. *Acc. Chem. Res.* **2004**, *37*, 421.
- ³⁰ Chalupský, J.; Rokob, T. A.; Kurashige, Y.; Yanai, T.; Solomon, E. I.; Rulišek, L.; Srncic, M. *J. Am. Chem. Soc.* **2014**, *136*, 15977.
- ³¹ Bollinger Jr., J. M.; Diao, Y.; Matthews, M. L.; Xing, G.; Krebs, C. *Dalton Trans.* **2009**, 905.
- ³² Hirao, H.; Morokuma, K. *J. Am. Chem. Soc.* **2009**, *131*, 17206.
- ³³ Colby, J.; Stirling, D. I.; Dalton, H. *Biochem. J.* **1977**, *165*, 395.
- ³⁴ (a) Gu, E.; Zhong, W.; Ma, H.; Xu, B.; Wang, H.; Liu, X. *Inorg. Chim. Acta* **2016**, *444*, 159. (b) Sekino, M.; Furutachi, H.; Tasaki, K.; Ishikawa, T.; Mori, S.; Fujinami, S.; Akine, S.; Sakata, Y.; Nomura, T.; Ogura, T.; Kitagawa, T.; Suzuki, M. *Dalton Trans.* **2016**, *45*, 469. (c) Kejriwal, A.; Bandyopadhyay, P.; Biswas, A. N. *Dalton Trans.* **2015**, *44*, 17261. (d) Wang, X.; Zhang, T.; Yang, Q.; Jiang, S.; Li, B. *Eur. J. Inorg. Chem.* **2015**, 817. (e) Siewert, I.; Limberg, C. *Chem. Eur. J.* **2009**, *15*, 10316. (f) Yamashita, M.; Furutachi, H.; Toshi, T.; Fujinami, S.; Saito, W.; Maeda, Y.; Takahashi, K.; Tanaka, K.; Kitagawa, T.; Suzuki, M. *J. Am. Chem. Soc.* **2007**, *129*, 2. (g) Avenier, F.; Dubois, L.; Latour, J.-M. *New J. Chem.* **2004**, *28*, 782.
- ³⁵ (a) Cafaro, V.; Izzo, V.; Scognamiglio, R.; Notomista, E.; Capasso, P.; Casbarra, A.; Pucci, P.; Di Donato, A. *Appl. Environ. Microbiol.* **2004**, *70*, 2211. (b) Tinberg, C. E.; Song, W. J.; Izzo, V.; Lippard, S. J. *Biochemistry* **2011**, *50*, 1788. (c) McCormick, M. S.; Lippard, S. J. *Biochemistry* **2011**, *50*, 11058.
- ³⁶ (a) Whited, G. M.; Gibson, D. T. *J. Bacteriol.* **1991**, *173*, 3010. (b) Acheson, J. F.; Bailey, L. J.; Elsen, N. L.; Fox, B. G. *Nat. Commun.* **2014**, *5*, 5009.
- ³⁷ (a) Cafaro, V.; Scognamiglio, R.; Viggiani, A.; Izzo, V.; Passaro, I.; Notomista, E.; Dal Piaz, F.; Amoresano, A.; Casbarra, A.; Pucci, P.; Di Donato, A. *Eur. J. Biochem.* **2002**, *269*, 5689. (b) Liang, A. D.; Lippard, S. J. *J. Am. Chem. Soc.* **2015**, *137*, 10520.
- ³⁸ Bochevarov, A. D.; Li, J.; Song, W. J.; Friesner, R. A.; Lippard, S. J. *J. Am. Chem. Soc.* **2011**, *133*, 7384.
- ³⁹ (a) Murray, L. J.; Naik, S. G.; Ortillo, D. O.; García-Serres, R.; Lee, J. K.; Huynh, B. H.; Lippard, S. J. *J. Am. Chem. Soc.* **2007**, *129*, 14500. (b) Song, W. J.; Lippard, S. J. *Biochemistry* **2011**, *50*, 5391.
- ⁴⁰ Song, W. J.; McCormick, M. S.; Behan, R. K.; Sazinsky, M. H.; Jiang, W.; Lin, J.; Krebs, C.; Lippard, S. J. *J. Am. Chem. Soc.* **2010**, *132*, 13582.
- ⁴¹ de Visser, S. P. *Chem. Eur. J.* **2008**, *14*, 4533.
- ⁴² Rather, L. J.; Weinert, T.; Demmer, U.; Bill, E.; Ismail, W.; Fuchs, G.; Ermler, U. *J. Biol. Chem.* **2011**, *286*, 29241.
- ⁴³ (a) Rather, L. J.; Knapp, B.; Haehnel, W.; Fuchs, G. *J. Biol. Chem.* **2010**, *285*, 20615. (b) Zaar, A.; Gescher, J.; Eisenreich, W.; Bacher, A.; Fuchs, G. *Mol. Microbiol.* **2004**, *54*, 223. (c) Gescher, J.; Zaar, A.; Mohamed, M.; Schägger, H.; Fuchs, G. *J. Bacteriol.* **2002**, *184*, 6301.
- ⁴⁴ (a) Teufel, R.; Friedrich, T.; Fuchs, G. *Nature* **2012**, *483*, 359. (b) Grishin, A. M.; Ajamian, E.; Tao, L.; Zhang, L.; Menard, R.; Cygler, M. *J. Biol. Chem.* **2011**, *286*, 10735. (c) Teufel, R.; Mascaraque, V.; Ismail, W.; Voss, M.; Perera, J.; Eisenreich, W.; Haehnel, W.; Fuchs, G. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 14390. (d) Fernández, C.; Ferrández, A.; Miñambres, B.; Díaz, E.; García, J. L. *Appl. Environ. Microbiol.* **2006**, *72*, 7422.
- ⁴⁵ Fuchs, G.; Boll, M.; Heider, J. *Nat. Rev. Microbiol.* **2011**, *9*, 803.
- ⁴⁶ Vaillancourt, F. H.; Bolin, J. T.; Eltis, L. D. *Crit. Rev. Biochem. Mol. Biol.* **2006**, *41*, 241.
- ⁴⁷ (a) Boll, M. *Biochim. Biophys. Acta* **2005**, *1707*, 34. (b) Kung, J. W.; Löffler, C.; Dörner, K.; Heintz, D.; Gallien, S.; Van Dorsselaer, A.; Friedrich, T.; Boll, M. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 17687.
- ⁴⁸ Liao, R.-Z.; Siegbahn, P. E. M. *Chem. Sci.* **2015**, *6*, 2754.
- ⁴⁹ Word, J. M.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. *J. Mol. Biol.* **1999**, *285*, 1735.
- ⁵⁰ (a) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. *J. Chem. Theory Comput.* **2011**, *7*, 525. (b) Søndergaard, C. R.; Olsson, M. H. M.; Rostkowski, M.; Jensen, J. H. *J. Chem. Theory Comput.* **2011**, *7*, 2284.
- ⁵¹ Case, D. A.; Darden, T. A.; Cheatham III, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Götz, A. W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *Amber 12 and AmberTools 13*; University of California, San Francisco, 2012.
- ⁵² Salomon-Ferrer, R.; Case, D. A.; Walker, R. C. *WIREs Comput. Mol. Sci.* **2013**, *3*, 198.
- ⁵³ Solomon, E. I.; Brunold, T. C.; Davis, M. I.; Kemsley, J. N.; Lee, S.-K.; Lehnert, N.; Neese, F.; Skulan, A. J.; Yang, Y.-S.; Zhou, J. *Chem. Rev.* **2000**, *100*, 235.
- ⁵⁴ Chung, L. W.; Sameera, W. M. C.; Ramozzi, R.; Page, A. J.; Hatanaka, M.; Petrova, G. P.; Harris, T. V.; Li, X.; Ke, Z.; Liu, F.; Li, H. B.; Ding, L.; Morokuma, K. *Chem. Rev.* **2015**, *115*, 5678.
- ⁵⁵ Tao, P.; Fisher, J. F.; Shi, Q.; Vreven, T.; Mobashery, S.; Schlegel, H. B. *Biochemistry* **2009**, *48*, 9839.
- ⁵⁶ Hu, L.; Söderhjelm, P.; Ryde, U. *J. Chem. Theory Comput.* **2011**, *7*, 761.
- ⁵⁷ Vreven, T.; Frisch, M. J.; Kudin, K. N.; Schlegel, H. B.; Morokuma, K. *Mol. Phys.* **2006**, *104*, 701.

- ⁵⁸ For the B3LYP functional, see: (a) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200. (b) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785. (c) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098. (d) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648. (e) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.
- ⁵⁹ For the D3 dispersion correction, see: Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.
- ⁶⁰ For the def2 basis sets, see: Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297.
- ⁶¹ (a) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. *J. Chem. Theory Comput.* **2015**, *11*, 3696. (b) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712. (c) Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049. (d) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
- ⁶² Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157.
- ⁶³ Fukui, K. *Acc. Chem. Res.* **1981**, *14*, 363.
- ⁶⁴ (a) Page, M.; McIver Jr., J. W. *J. Chem. Phys.* **1988**, *88*, 922. (b) Page, M.; Doubleday, C.; McIver Jr., J. W. *J. Chem. Phys.* **1990**, *93*, 5634.
- ⁶⁵ (a) Hu, L.; Söderhjelm, P.; Ryde, U. *J. Chem. Theory Comput.* **2013**, *9*, 640. (b) Sumner, S.; Söderhjelm, P.; Ryde, U. *J. Chem. Theory Comput.* **2013**, *9*, 4205.
- ⁶⁶ Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822.
- ⁶⁷ (a) Weigend, F. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057. (b) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. *Chem. Phys. Lett.* **1995**, *242*, 652. (c) Sierka, M.; Hogeckamp, A.; Ahlrichs, R. *J. Chem. Phys.* **2003**, *118*, 9136.
- ⁶⁸ Kaukonen, M.; Söderhjelm, P.; Heimdal, J.; Ryde, U. *J. Phys. Chem. B* **2008**, *112*, 12537.
- ⁶⁹ Glendening, E. D.; Landis, C. R.; Weinhold, F. *WIREs Comput Mol Sci* **2012**, *2*, 1.
- ⁷⁰ Gaussian 09, Revision D.01, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, 2009.
- ⁷¹ (a) *Turbomole v6.6 2014*, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from <http://www.turbomole.com>. (b) Furche, F.; Ahlrichs, R.; Hättig, C.; Klopper, W.; Sierka, M.; Weigend, F. *WIREs Comput. Mol. Sci.* **2014**, *4*, 91.
- ⁷² Hoof, R. W. W.; Vriend, G.; Sander, C.; Abola, E. E. *Nature* **1996**, *381*, 272.
- ⁷³ (a) Ryde, U. *J. Comput. Aided Mol. Des.* **1996**, *10*, 153. (b) Ryde, U.; Olsson, M. H. M. *Int. J. Quantum Chem.* **2001**, *81*, 335.
- ⁷⁴ Elsen, N. L.; Bailey, L. J.; Hauser, A. D.; Fox, B. G. *Biochemistry* **2009**, *48*, 3838.
- ⁷⁵ Song, W. J.; Behan, R. K.; Naik, S. G.; Huynh, B. H.; Lippard, S. J. *J. Am. Chem. Soc.* **2009**, *131*, 6074.
- ⁷⁶ Liang, A. D.; Wrobel, A. T.; Lippard, S. J. *Biochemistry* **2014**, *53*, 3585.
- ⁷⁷ Some further isomers ($\mu\text{-}\eta^2\text{:}\eta^1\text{-peroxo}$, $\mu\text{-}\eta^1\text{:}\eta^2\text{-peroxo}$, and $\eta^1\text{-Fe2-peroxo}$) were also found; see the Supporting Information.
- ⁷⁸ The correspondence between their located stationary points and ours is described in the Supporting Information.
- ⁷⁹ The B3LYP functional employed for the geometry optimization may somewhat underestimate the O–O distance and overestimate the superoxide character. See ref. 30.
- ⁸⁰ For a discussion of this issue in methane monooxygenase, see: Rinaldo, D.; Philipp, D. M.; Lippard, S. J.; Friesner, R. A. *J. Am. Chem. Soc.* **2007**, *129*, 3135.
- ⁸¹ The relatively stability of the peroxo and bis($\mu\text{-oxo}$) isomers of the related Cu_2O_2 core is also known to be difficult to estimate accurately using DFT. For a selection of papers, see: (a) Gherman, B. F.; Cramer, C. J. *Coord. Chem. Rev.* **2009**, *253*, 723. (b) Liakos, D.; Neese, F. *J. Chem. Theory Comput.* **2011**, *7*, 1511. (c) Rohrmüller, M.; Hoffmann, A.; Thierfelder, C.; Herres-Pawlis, S.; Schmidt, W. G. *J. Comput. Chem.* **2015**, *36*, 1672.
- ⁸² The isolated active site model with the $\text{Fe}\cdots\text{Fe}$ distance constrained to the value in the enzyme is 3.5 kcal/mol higher in electronic energy than the unconstrained structure, which is approximately half of the ~ 6.0 kcal/mol destabilization by the enzyme.
- ⁸³ The $\text{Fe}\cdots\text{Fe}$ distance constraint is sufficient in the isolated active site model to shift the electronic structure to diiron(III,IV); however, the oxyl radical character does not localize to one of the oxo bridges as in the enzyme (Mulliken spin populations of the oxo bridges: $-0.50/-0.58$, cf. QM/MM values of $-0.24/-0.69$).
- ⁸⁴ Baik, M.-H.; Gherman, B. F.; Friesner, R. A.; Lippard, S. J. *J. Am. Chem. Soc.* **2002**, *124*, 14608.
- ⁸⁵ There are no second-sphere residues, not included in the QM region, that would provide obvious electrostatic stabilization for the dangling oxygen. Reoptimization of the geometry using an extra *sp* set of diffuse basis functions on the two oxo and two Glu240 carboxylate oxygens, taken from the 6-31+G basis set, negligibly changes the O(oxo)–O(Glu) distance ($+0.02\text{Å}$) and the corresponding spin populations (± 0.03 at most). Furthermore, a comparison of the computed spin populations at the BP86/def2-SV(P) level for the 80-atom and 995-atom QM regions (-0.17 vs. -0.11) indicate that the effect is diminished but not quenched by the environment. We thus consider it unlikely that this partial bond formation would be an artefact of the QM/MM computations.
- ⁸⁶ Chen, M. S.; White, M. C. *Science* **2007**, *318*, 783.
- ⁸⁷ Wang, Y.; Janardanan, D.; Usharani, D.; Han, K.; Que Jr, L.; Shaik, S. *ACS Catal.* **2013**, *3*, 1334.
- ⁸⁸ Our first attempts to locate the W-class O–O bond scission TS only yielded TSs describing concerted O–O bond scission and C–O bond formation with the aromatic ring, in line with what was reported by Liao and Siegbahn (ref. 48). However, a comparison of the geometries of W- $\mu\text{-}\eta^2\text{:}\eta^2\text{-peroxo}$ and W-bis($\mu\text{-oxo}$) showed that the substrate occupies essentially the same

position in these two; we therefore located a transition state for the interconversion with the position of the substrate constrained. At the level employed for the geometry optimization, the energy of the constrained O–O breakage transition state is higher than that of the concerted epoxidation. However, the higher-level single point computations confirm that the O–O breakage *without* substrate attack is notably lower in energy than the concerted epoxidation; it thus seems justified to consider O–O breakage as a separate step. We also note that we encountered no difficulty in locating the O–O TS without constraints for the H isomer.

⁸⁹ A major source of this discrepancy is that the MM iterative charge updating was not employed in IRC runs. Furthermore, IRC calculations locate the minima with smaller accuracy and less tight convergence criteria than do the algorithms employed for geometry optimization.

⁹⁰ Kraka, E.; Cremer, D. *Acc. Chem. Res.* **2010**, *43*, 591.

⁹¹ An alternative $P\sigma^*$ pathway starts from the H- μ - η^2 : η^1 -peroxo structure with a somewhat higher activation free energy (28.0 kcal/mol). See the Supporting Information.

⁹² We located the analogous transition states in the unconstrained isolated active site model with the substrates phenol, benzene, PhC(=O)SMe, and nitrobenzene, and we found that the maximum spin population on the substrate along the IRC increases in this order (0.10, 0.16, 0.24, 0.27, respectively). It thus seems that the radical nature of the attack is somewhat enhanced as the substrate becomes less reactive toward an electrophilic attack. Nevertheless, the hidden intermediate is an arenium cation in all considered cases.

⁹³ The Fe^{III}(H₂O₂) complex in heme systems can bring about sulfoxidation in an analogous manner. See: Wang, B.; Li, C.; Cho, K.-B.; Nam, W.; Shaik, S. *J. Chem. Theory Comput.* **2013**, *9*, 2519.

⁹⁴ Smith, M. B.; March, J. *March's Advanced Organic Chemistry*, 5th edition, Wiley-Interscience, New York, **2001**, pp. 1051.

⁹⁵ Hussain, H.; Al-Harrasi, A.; Green, I. R.; Ahmed, I.; Abbas, G.; Rehman, N. U. *RSC Adv.* **2014**, *4*, 12882.

⁹⁶ Jerina, D. M.; Daly, J. W.; Witkop, B. *Biochemistry* **1971**, *10*, 366.

⁹⁷ (a) Bartlett, P. D. *Rec. Chem. Prog.* **1957**, *18*, 111. (b) Houk, K. N.; Liu, J.; DeMello, N. C.; Condroski, K. R. *J. Am. Chem. Soc.* **1997**, *119*, 10147. (c) Singleton, D. A.; Merrigan, S. R.; Liu, J.; Houk, K. N. *J. Am. Chem. Soc.* **1997**, *119*, 3385. (d) Bach, R. D.; Canepa, C.; Winter, J. E.; Blanchette, P. E. *J. Org. Chem.* **1997**, *62*, 5191. (e) Freccero, M.; Gandolfi, R.; Sarzi-Amadè, M.; Rastelli, A. *J. Org. Chem.* **2000**, *65*, 8948.

⁹⁸ As expected, the $P\sigma^*$ -type reaction between W- μ - η^2 : η^2 -peroxo as an isolated active site model and a propene substrate shows characteristics of case A (the geometry is shown in Figure S7 in Supporting Information).

⁹⁹ The route toward the 2,3-epoxide is characterized by activation free energies of 46.6 kcal/mol and 55.5 kcal/mol.

¹⁰⁰ The IRC calculations started from the two TSs converge to the same minimum (point 3) within the accuracy allowed by the convergence criteria and by the use of different sets of MM charges for the QM atoms.

¹⁰¹ As these barriers are quite high, we did not locate the TSs for the second step (O–O cleavage and second C–O bond formation).

¹⁰² The preference for the **Q** pathway over $P\sigma^*$ is also reproduced if TPSSh or M11L functionals are employed in the inner QM region of the single-point calculations on the enzyme. Moreover, we were also able to locate the TS of the **Q** pathway using the cluster method and the QM level employed by Liao and Siegbahn (ref. 48), and it lies ~0.8kcal/mol below their $P\sigma^*$ -type **TS1_{SR}**. Considering these results and the potential error bars of the calculations, we can state that most probably, the **Q** pathway is operative, but we cannot exclude contribution from the $P\sigma^*$.

¹⁰³ Tinberg, C. E.; Lippard, S. J. *Biochemistry* **2010**, *49*, 7902.

¹⁰⁴ It is interesting to note that Sc³⁺ was found to trigger O₂ activation in a non-heme monoiron(II) complex via a Fe³⁺-(μ - η^2 : η^2 -peroxo)-Sc³⁺ species, analogous to BoxB W- μ - η^2 : η^2 -peroxo. See: (a) Li, F.; Van Heuvelen, K. M.; Meier, K. K.; Münck, E.; Que Jr., L. *J. Am. Chem. Soc.* **2013**, *135*, 10198. (b) Poater, A.; Vummaleti, S. V. C.; Cavallo, L. *Dalton Trans.* **2014**, *43*, 11190.

¹⁰⁵ This aspect was highlighted in connection with the unusual peroxo species structure of the non-heme diiron arylamine oxygenase CmlI. See: Makris, T. M.; Vu, V. V.; Meier, K. K.; Komor, A. J.; Rivard, B. S.; Münck, E.; Que Jr., L.; Lipscomb, J. D. *J. Am. Chem. Soc.* **2015**, *137*, 1608.

¹⁰⁶ Guroff, G.; Daly, J. W.; Jerina, D. M.; Renson, J.; Witkop, B.; Udenfriend, S. *Science* **1967**, *157*, 1524.

¹⁰⁷ Our attempts to locate the $P\sigma^*$ C2 attack transition state invariably led to the TS for the C3 attack within the enzyme, and conversely, optimizations toward the $P\sigma^*$ C3 attack TS on the isolated models converged to C2 attack structures. Nevertheless, model calculations with HCOOOH as oxidant reveal larger asymmetry for the C3 than for the C2 attack; the value in question therefore likely represents an underestimation.

¹⁰⁸ Notomista, E.; Cafaro, V.; Bozza, G.; Di Donato, A. *Appl. Environ. Microbiol.* **2009**, *75*, 823.

¹⁰⁹ Mitchell, K. H.; Rogge, C. E.; Gierahn, T.; Fox, B. G. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3784.

TOC graphic

