

Atomic and Electronic Structure Determinants Distinguish between Ethylene Formation and L-Arginine Hydroxylation Reaction Mechanisms in the Ethylene-Forming Enzyme

Shobhit S. Chaturvedi,^a Rajeev Ramanan,^a Jian Hu,^{b,c} Robert P. Hausinger,^{b,d} Christo Z. Christov^{a,*}

^a Department of Chemistry, Michigan Technological University, Houghton, MI 49931, USA.

^b Department of Biochemistry and Molecular Biology, ^c Department of Chemistry, and

^d Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing,
MI 48824, USA.

*Corresponding Author: Christo Z. Christov: christov@mtu.edu

Abstract

The ethylene-forming enzyme (EFE) is a non-heme Fe(II), 2-oxoglutarate (2OG), and L-arginine (L-Arg)-dependent oxygenase that catalyzes dual reactions: the generation of ethylene from 2OG and the C5 hydroxylation of L-Arg. Using an integrated molecular dynamics (MD) and quantum mechanics/molecular mechanics (QM/MM) approach that references previous experimental studies, we tested the hypothesis that synergy between the conformation of L-Arg and the coordination mode of 2OG directs the reaction toward ethylene formation or L-Arg hydroxylation. The dynamics of EFE·Fe(III)·OO[·]·2OG·L-Arg show L-Arg can exist in conformation A (productive for hydroxylation) and conformation B (unproductive for hydroxylation). QM/MM calculations show that when 2OG is bound in an off-line mode and L-Arg is present in conformation A, the Fe(III)-OO[·] intermediate undergoes the standard O₂ activation mechanism involving ferryl-

dependent hydroxylation. With the same off-line 2OG coordination, but with conformation B of L-Arg, a unique pathway produces a half-bond ferric-bicarbonate intermediate that decomposes to ethylene, two CO₂, and a ferrous-bicarbonate species. The results demonstrate that when 2OG is coordinated in off-line mode to the Fe center, the L-Arg conformation acts as a switch that directs the reaction toward ethylene formation or hydroxylation. Analysis of the electronic structure shows the L-Arg conformation defines the precise location of an unpaired β electron in the Fe(III)-OO[•] complex, either in a $\pi^*_{||}$ orbital that triggers ethylene formation or a $\pi^*_{\perp}$ orbital that cascades to L-Arg hydroxylation. A change in 2OG coordination from off-line to in-line reduces stabilization of the 2OG C1 carboxylate such that neither conformation of L-Arg produces the ethylene-forming half-bond ferric-bicarbonate intermediate. *Thus, L-Arg conformation-dependent changes in the electronic structure of the Fe(III)-OO[•] orbitals, together with the 2OG binding mode-associated stabilization of the C1-carboxylate, distinguish whether the EFE reaction proceeds via the ethylene-forming pathway or catalyzes a hydroxylation mechanism.*

Keywords: Ethylene forming Enzyme, 2OG-dependent enzyme, non-heme Fe enzymes, QM/MM, Molecular Dynamics.

1. Introduction

The ethylene-forming enzyme (EFE) is a non-heme Fe(II), 2-oxoglutarate (2OG)-, and L-arginine (L-Arg)-dependent oxygenase that catalyzes two reactions (Figure 1).^{1,2} The less prominent EFE-catalyzed transformation involves oxidative decarboxylation of 2OG, forming succinate and CO₂, coupled with C5 hydroxylation of L-Arg, where the hydroxylated L-Arg spontaneously decomposes to guanidine and L- Δ 1-pyrroline-5-carboxylate. The major EFE-

catalyzed reaction involves the decomposition of 2OG to ethylene plus three molecules of CO₂/bicarbonate. EFE's hydroxylation reaction is typical of many other Fe(II)/2OG oxygenases in this superfamily,^{3–6} however, the generation of ethylene by cleaving three C-C bonds of 2OG is unique to this enzyme. L-Arg hydroxylation is suggested to follow a well-established mechanism of non-heme Fe(II) and 2OG dependent enzymes, i.e., dioxygen activation to form an active ferryl species that hydroxylates the substrate.^{7,8} The ethylene-forming mechanism is suggested to diverge from the hydroxylation mechanism during the dioxygen activation process,^{7,8} as also shown by a recent computational study on EFE.⁹ Regardless of the current success in experimental and computational studies of EFE, the atomistic and electronic structure determinants for the L-Arg hydroxylation vs. the ethylene-formation mechanism remain unclear.

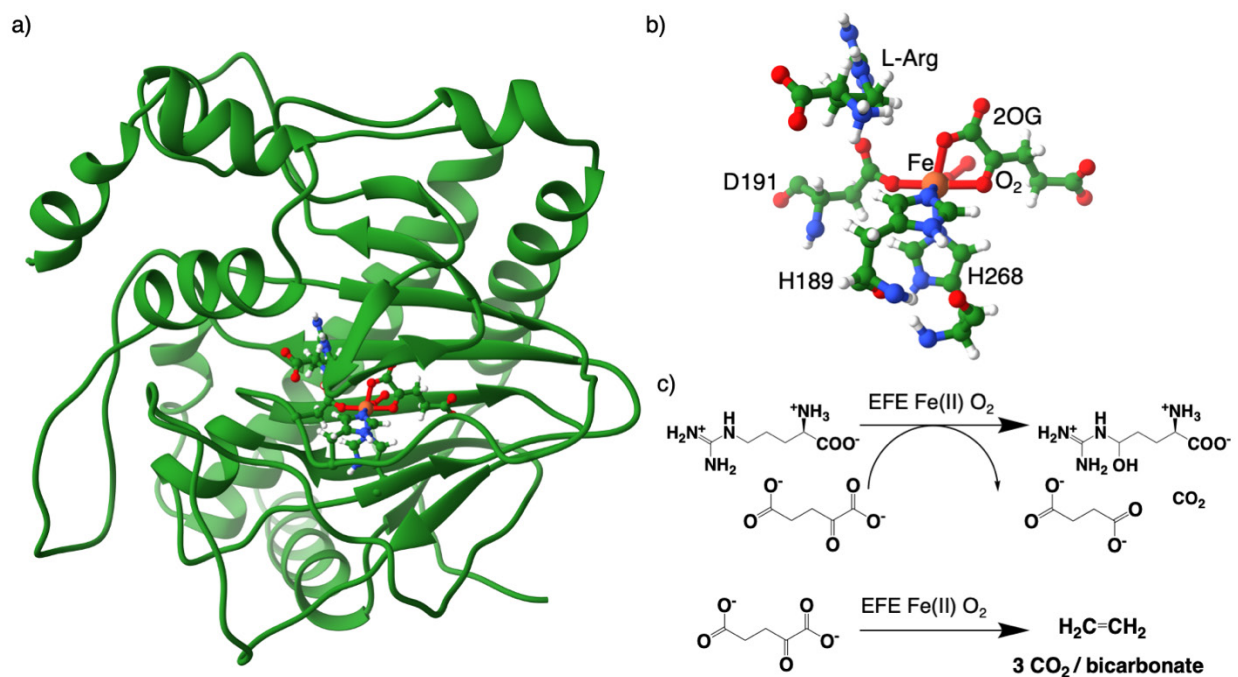
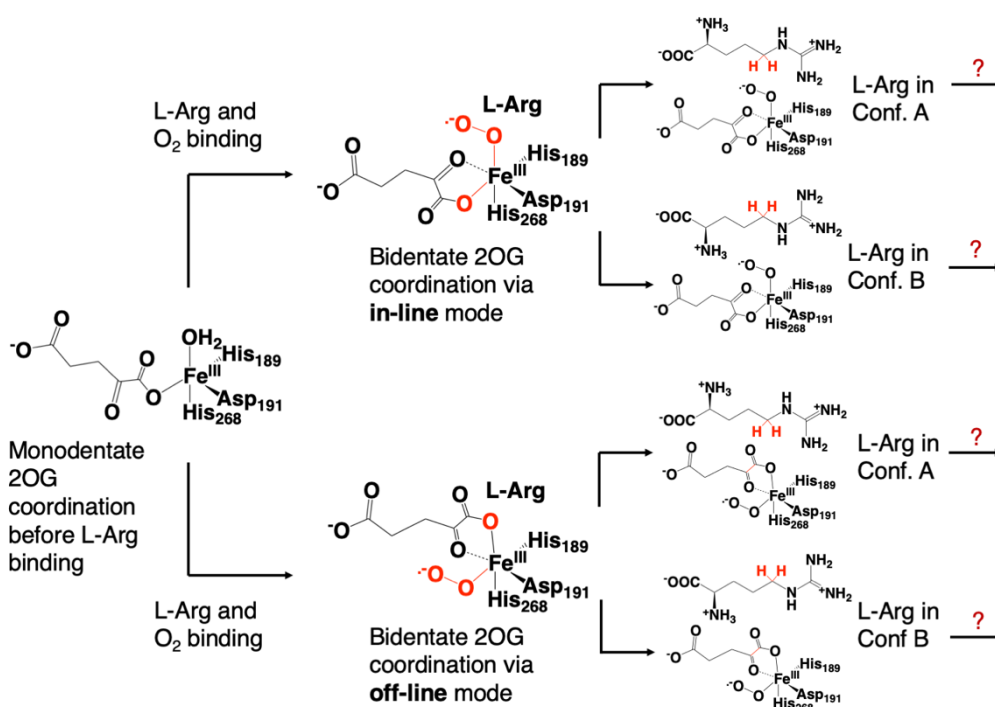


Figure 1. (a) Protein structure for the EFE·Fe(III)·OO··2OG·L-Arg complex modeled using PDB ID 5V2Y with b) highlighted active site shown in ball and stick format. c) Reactions catalyzed by EFE.

Even though L-Arg is not directly involved in ethylene formation, biochemical studies have shown that EFE loses its ethylene-forming capacity in the absence of L-Arg.² Structural studies suggest the binding of L-Arg at the EFE active site plays a critical role in the unique ethylene-forming reaction of this enzyme.^{7,8} In particular, L-Arg binding induces conformational changes at the active site that may be necessary for ethylene formation. Interestingly, in the crystal structure of EFE·Fe·NOG·L-Arg, where NOG is the 2OG analog *N*-oxalylglycine, L-Arg is observed in two binding conformations.⁸ The key difference in these species is the positioning of the C5-methylene, the target for hydroxylation, and the guanidino group. In conformation A, the η^2 NH₂ of the L-Arg guanidinium group forms a hydrogen bond with D191, a ligand of the Fe center, and C5 faces towards the metallocenter (Scheme 1). In conformation B, upon rotation of the η^2 NH₂, the ϵ NH of the guanidinium group of L-Arg forms hydrogen bonds with D191. The rotation of the guanidinium group causes C5 to be pushed away from the Fe center (Scheme 1). Based on the proximity of the Fe center to C5, the target for hydroxylation, different roles were proposed for the two L-Arg conformations (Scheme 1).⁸ Conformation A, where the distance is shorter, was proposed to promote hydroxylation. In contrast, conformation B, where the distance is larger, was suggested to play a role in the ethylene formation pathway. If the two L-Arg conformations are relevant after O₂ binding, their roles in the reaction mechanisms of EFE need further exploration.



Scheme 1. Exploring the proposed role of L-Arg conformations and 2OG coordination modes on ethylene formation and hydroxylation mechanisms carried out by EFE.

Additional features that may have a role in the ethylene-forming ability of EFE are the 2OG bidentate coordination modes to the Fe center, the hydrophobic environment surrounding 2OG, the 2OG C5 carboxylate conformational planarity, and a unique twisted peptide bond at the enzyme active site.^{7,8} 2OG exhibits two types of bidentate coordination to the Fe center of 2OG-dependent oxygenases (Scheme 1). The in-line mode, with the 2OG C1 carboxylate positioned *trans* to His₁ (His₁₈₉ in EFE), leaves the position *trans* to His₂ (His₂₆₈) available for O₂ binding.⁵ Such binding would lead to a reactive ferryl intermediate being proximal to the substrate C-H bond. By contrast, in the off-line mode, the 2OG C1 carboxylate binds *trans* to His₂, leaving the position *trans* to His₁ accessible for O₂ binding; this site is approximately perpendicular to the substrate.⁵ Such binding results in a ferryl pointing perpendicular to the substrate and is therefore not

positioned to perform hydroxylation. In this case, a ferryl flip may occur to align the ferryl in the direction of the substrate for the subsequent hydrogen abstraction step.¹⁰ The 2OG binding groove of EFE is mostly hydrophobic, except for two arginines, R177 and R277, stabilizing the C1 and C5 carboxylates of 2OG.^{7,8} Similar to the two conformations seen for L-Arg, the same EFE crystal structure shows variations in planarity of the distal carboxylate for NOG and its interactions with R277.⁸ A quantum mechanics (QM) model study shows that the planarity of the 2OG C5 carboxylate affects the energy required to transform 2OG to ethylene.⁸ Another notable structural feature of the enzyme active site with bound L-Arg and 2OG is the twisted peptide bond involving the D191 metal-ligand and the following Y192 residue (i.e., D191-C α , D191-C, Y192-N, Y192-C torsion angle of only $\sim 147^\circ$).⁷ The preferred 2OG coordination mode before O₂ binding, the importance of the 2OG hydrophobic environment, the relevance of the 2OG C5 planarity, and the effects of the twisted peptide bond on the two modes on EFE reactivity remain unknown (**Scheme 1**).

Although EFE has been extensively studied experimentally, there remains a substantial gap in our understanding of why EFE produces ethylene from 2OG, whereas other 2OG-dependent oxygenases only decompose 2OG to succinate plus CO₂ as they modify their primary substrates. This knowledge gap is a barrier to success in improving enzymatic ethylene production. In the present manuscript, we apply an integrated molecular dynamics (MD) and combined quantum mechanics, and molecular mechanics (QM/MM) approach that references previous experimental studies, and we *test the hypothesis that a synergy between a specific conformation of the main substrate L-Arg and the coordination mode of cosubstrate 2OG defines the reaction path toward ethylene formation or L-Arg hydroxylation*. We explore the atomistic

and electronic determinants for the hydroxylation vs. ethylene-production reactions of EFE. We test whether L-Arg exhibits the two conformations in the dynamics of the catalytically important EFE·Fe(III)·OO⁻·2OG·L-Arg complex. We investigate the role of the two L-Arg conformations and the two 2OG coordination modes on the dual reactions of EFE. We also consider the importance of the hydrophobic environment of 2OG, the 2OG C5 planarity, and the twisted peptide bond. Finally, using molecular orbitals analysis of the Fe center, we explore the electronic structure variation in the metallocenter that leads to ethylene formation vs. L-Arg hydroxylation. Overall our results highlight the atomistic and electronic structure factors that can be influenced to improve the ethylene yield in EFE.

2. Computational Methods

System Preparation: A crystal structure of EFE in complex with L-Arg, 2OG, and Mn, an analog of Fe (PDB ID: 5V2Y), was used for the current study.⁷ The crystal structure has 2OG in off-line bidentate metallocenter coordination, along with monodentate coordination by two histidines (H189 and H268) and an aspartate (D191). To generate parameters for the EFE·Fe(III)·OO⁻·2OG_{off-line}·L-Arg complex, the Mn ion in the PDB file was changed to Fe, and an O₂ molecule was bound to Fe using GaussView 5.0.¹¹ The protonation states of the ionizable side chains of the protein were assessed using Amber routines.¹² The Amber topology for 2OG and O₂ were developed utilizing the GAFF tool in Antechamber.¹³ The Amber parameters for the Fe(III)-OO⁻ active site complex, containing high spin (HS) Fe(III) ($S=2$, $M=5$) and the coordinating ligands 2OG (bidentate), O₂ (bonded end-on), H189, H268, and D191 (all monodentate), were prepared with the Metal Centre Parameter Builder (MCPB) using MCPB.py v3.0.¹⁴ Bond and angle force

constants in MCPB were derived at the B3LYP/6-31G* level of theory. The remainder of the protein was modeled using the Amber FF14SB force field.¹⁵ The same procedure was repeated to make parameters for the $\text{EFE}\cdot\text{Fe(III)}\cdot\text{OO}^{\cdot-}\cdot 2\text{OG}_{\text{in-line}}\cdot\text{L-Arg}$ complex where 2OG is bound with an in-line coordination mode to Fe.

MD Simulations: The Leap module in AMBER 16 was used to add counter ions (Na^+) for the neutralization of the protein systems. The system was immersed into a rectangular box with TIP3P water molecules present up to a minimum of 10 Å from the farthest protein boundary.¹⁶ Periodic boundary conditions were employed in all simulations. Long-range electrostatic interactions were calculated using the particle mesh Ewald method with a direct space and van der Waals cut-off of 10 Å.¹⁷ With a restraint of $100 \text{ kcal mol}^{-1} \text{ Å}^2$ on solute molecules; the systems were subjected to energy minimization, first using the steepest descent (5000 steps) followed by the conjugate gradient (5000 steps). This sequence was followed by full minimization of the entire systems, again using steepest descent (5000 steps) and a conjugate gradient (5000 steps). The systems were then subjected to controlled heating from 0 to 300K at an NVT ensemble using a Langevin thermostat with a collision frequency of 1 ps^{-1} for 250 ps.¹⁸ The solute molecules were restrained using a harmonic potential of $50 \text{ kcal mol}^{-1} \text{ Å}^2$ during the heating process. The SHAKE algorithm was used to constrain bonds involving hydrogen.¹⁹ A weakly restrained MD simulation for 1 ns was performed to achieve a uniform density after heating dynamics under periodic boundary conditions. All systems were then equilibrated at 300K in an NPT ensemble without restraints for 3 ns; the pressure was maintained at 1 bar using the Berendsen barostat.²⁰ A production MD run from the equilibrated structure with an explicit solvent for 1 μs each with a time step of 2 fs was performed in an NPT ensemble with a target pressure set at 1 bar and

constant pressure coupling of 2 ps. All productive MD simulations were performed with the GPU version of the AMBER 16 package.²¹ Hydrogen bond analysis was done using CPPTRAJ.²² The principal component analysis (PCA) and dynamic cross-correlation analysis (DCCA) was performed on the backbone C α -atoms on the 500 ns-1000 ns range of the trajectory with Bio3D.²³

QM/MM calculations: Structures from the well-equilibrated part of the MD trajectory (>500 ns) were used for the QM/MM calculations from productive MD simulations. For QM/MM calculations with L-Arg conformation A, structures based on close Op to C2 of 2OG distance were selected at ~725 and 761 ns for calculations, and for L-Arg conformation B, structures based on close Op to C2 of 2OG distance were selected at ~960 and 896 ns. Excess water molecules were truncated using CPPTRAJ, such that the protein retained a water solvation layer maximum of 12 Å surrounding it.²² The resulting total size of the systems was ~22500 atoms. QM/MM calculations were performed using ChemShell software,²⁴ with a combination of Turbomole²⁵ for the QM part and DL_POLY²⁶ for the MM part. The non-heme Fe center, the first coordination sphere residues, and the substrate-bound in the active site were included in the QM region. An unrestricted B3LYP functional was used to represent the QM region as it has been used for EFE⁹ and several non-heme and 2OG-enzymes with better accuracy in comparison to other functionals for calculation of the reaction mechanism.^{27–30} The protein region within 8 Å from the QM region was defined as the MM region, and the rest of the system was fixed. The Amber FF14SB force field was used for the MM region. QM/MM boundaries were capped with a hydrogen link atom, and a charge shift model was used.³¹ The polarizing effect of the protein environment (MM region) was accounted for in the QM region using an electrostatic embedding scheme.³² For

geometry optimization and frequency calculations, the def2-SVP basis set [QM(B1)/MM] was used.³³

Optimized reactant complexes (RC) were used to search for transition states (TS) along with the reaction coordinate by performing a relaxed potential energy scan (adiabatic mapping) with a step size of 0.1 Å.³⁴ The highest energy geometry in the potential energy surface was optimized using the P-RFO optimizer in HDLC code without any constraints.³⁵ Frequency calculations of all optimized geometries were carried out to confirm the minima and transition states. To improve the calculated energies, single-point energy calculations on the optimized geometries were conducted using the large all-electron def2-TZVP basis set [QM(B2)/MM].³³ The zero-point energy (ZPE) calculations were performed for all geometries, and the energies are reported as QM(B2+ZPE)/MM. All the reported energies in the Results and Discussion section use this same level of theory. The energy decomposition analysis (EDA) was performed using a Fortran90 program by the Cisneros research group.^{36,37} Experimental studies have shown a quintet spin state ($S=2$) as the ground spin state for studied non-heme enzymes.^{38–40} Several computational studies on O₂ activation and the hydrogen atom transfer (HAT) pathway in non-heme 2OG dependent enzymes have shown the reaction proceeds through an HS quintet spin state.^{9,27–28,30,41} Therefore, we performed the reaction mechanism calculation on the quintet ($S=2$, $M=5$) spin state of the EFE·Fe(III)·OO^{·-}·2OG_{off-line/in-line}·L-Arg complex. The geometry coordinates, spin densities, and charges of all intermediates and transition states are presented in Supplementary Information (SI) Tables S1-S20, and SI pages S27-S72.

3. Results and Discussion

3.1 Dynamics of EFE·Fe(III)·OO⁻·2OG·L-Arg with 2OG in off-line coordination

A previously reported crystal structure of EFE·Fe·NOG·L-Arg (PDB: 5LUN) identified dual binding conformations for the L-Arg and NOG.⁸ To begin to explore whether L-Arg and 2OG exhibit multiple conformations after O₂ binding, we performed a 1 μ s MD simulation of the catalytically important EFE·Fe(III)·OO⁻·2OG_{off-line}·L-Arg intermediate (Figure S1) generated by using an EFE·Mn·2OG_{off-line}·L-Arg structure (PDB: 5V2Y) as described in Computational Methods. Indeed, our MD simulation revealed two binding modes for L-Arg, with the most prominent distinction between the conformations being the orientation of the L-Arg guanidinium group and the C5 methylene carbon atom (the hydroxylation site) (Figure 2). EFE residue R171 specifically stabilizes conformation A of L-Arg via a strong electrostatic interaction with the guanidinium group (Figure 2c). The same electrostatic interaction is weakened throughout the dynamics when L-Arg changes from conformation A to B (Figure 2c). Similarly, the amino group of L-Arg strongly interacts with E84 in conformation B but not in conformation A. The two L-Arg conformations are observed during the dynamics of the catalytically important EFE·Fe(III)·OO⁻·2OG·Arg complex with different occurrences – conformation A in 11.7 % of the trajectory and conformation B – in 88.3%. The population of L-Arg conformation B in the dynamics of the EFE·Fe(III)·OO⁻·2OG_{off-line}·L-Arg complex is different from that in the crystal structure of EFE·Fe·NOG·L-Arg, where both L-Arg conformations are distributed equally, and that in the crystal structures of EFE·Mn·2OG_{off-line}·L-Arg where conformation A of L-Arg predominates. The potential reasons for having a preference of L-Arg conformation B in the EFE·Fe(III)·OO⁻·2OG·Arg complex dynamics include changes in the Fe center and second-sphere residues upon O₂ binding or that the conformation A L-Arg might

be more easily crystallizable than conformation B. Dynamics shows the η^2 NH₂ of the L-Arg guanidinium group forms a hydrogen bond with D191 (44% of the MD snapshots), a ligand of the Fe center. For some snapshots, the ϵ NH (68%) of L-Arg also forms a hydrogen bond with the D191. These findings raise important questions about the roles of the two conformations in the enzyme mechanisms. The proposed 2OG C5 carboxylate conformations, based on results for the NOG-containing crystal structure,⁸ were not maintained during the long-range dynamics; instead, we observed a torsion angle ranging from 0 to 180 between the O3-C5-O4 and C3-C4-C5 planes of 2OG (Figure S2). The 2OG C5 carboxylate forms hydrogen bonds with R277 (92%), while the C1 carboxylate forms a hydrogen bonding interaction with R171 (72%). The 2OG C1 carboxylate stabilization by R171 is improved in L-Arg conformation B (Figure S3). The twisted peptide bond between D191 and Y192 that is present in the crystal structure upon L-Arg binding (measured between D191-C α , D191-C, Y192-N, Y192-C)⁷ relaxes during the dynamics to an average value of 162.1°, occasionally visiting the twisted peptide state (Figure S4). This relaxation suggests the twisted peptide bond is unlikely to be essential for EFE reactivity.

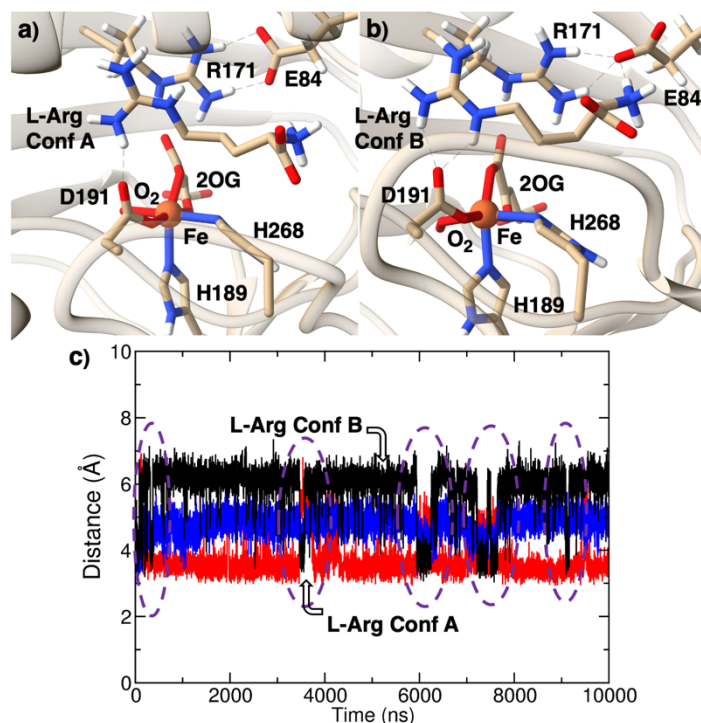


Figure 2. L-Arg binds to EFE·Fe(II)·2OG in two conformations, denoted Conf A and Conf B, in the off-line 2OG coordination mode. The distance between the C5 hydrogen atom of L-Arg and Fe (black trace) is used as a measure of conformational change. The distance between the CZ atom of L-Arg and R171 (blue trace) is associated with the stabilization of conformation A. The distance between the amino nitrogen of L-Arg and the oxygen of the E84 carboxyl group (red trace) is an index for stabilization of conformation B. The purple ovals indicate a change in the conformation of L-Arg.

When the dynamics of EFE·Fe(III)·OO⁻·2OG·L-Arg were assessed by PCA (Figure 3a), three regions were identified as the most flexible: (i) $\beta 4$, $\beta 5$, and the loops connecting them (residues 80-93), (ii) $\beta 11$ and its connecting loops (residues 211-245), and (iii) $\beta 15$ and the loop connecting it to $\alpha 8$ (residues 291-303). The high flexibility of these regions is interesting given their specific roles. $\beta 4$, $\beta 5$, and their loops are involved in L-Arg stabilization. $\beta 11$ and its loops shield the active site from solvent exposure and provide some L-Arg stabilization. The motions of $\beta 15$ and its loop

affect $\beta 14$, which forms the unique hydrophobic environment of the EFE active site. PCA also shows the direction of motion of the region made up of $\beta 4$, $\beta 5$, and its loops with that of $\beta 11$ are towards each other, potentially increasing their interactions. The interdependent motions of these regions are further confirmed by the DCCA, which shows region (iii) is dynamically anti-correlated with regions (i) and (ii), suggesting the motion of region (iii) affects the motions of regions (i) and (ii) (Figure 3b). DCCA also shows region (ii) exhibits anti-correlated motion with the region (i). These findings are significant as the effect of such long-range influences is experimentally shown in EFE. For example, the E215A variant that is altered in the region (ii) nearly abolishes the activity of EFE.⁷ R171, which stabilizes the C1 carboxylate of 2OG shows strong synchronous (positive correlated) motions with nearby residues in $\beta 6$ and $\beta 14$ (residues 94-100 and 279-284). The DCCA results also demonstrate the anti-correlated motion of region (iii) with the EFE active site hydrophobic residues such as F283, F250, and A281 that interact with dioxygen (when 2OG is bound in the off-line mode) or with the C1 carboxylate of 2OG (when 2OG is bound in the in-line mode). The results thus indicate that $\beta 4$, $\beta 5$, $\beta 11$, $\beta 15$ and their loop regions are flexible and have inter-dependent motions in the $\text{EFE} \cdot \text{Fe(III)} \cdot \text{OO} \cdot \cdot \text{2OG}_{\text{off-line}} \cdot \text{L-Arg}$ complex that might influence the active-site interactions, including L-Arg and 2OG stabilization via long-range correlated motions.

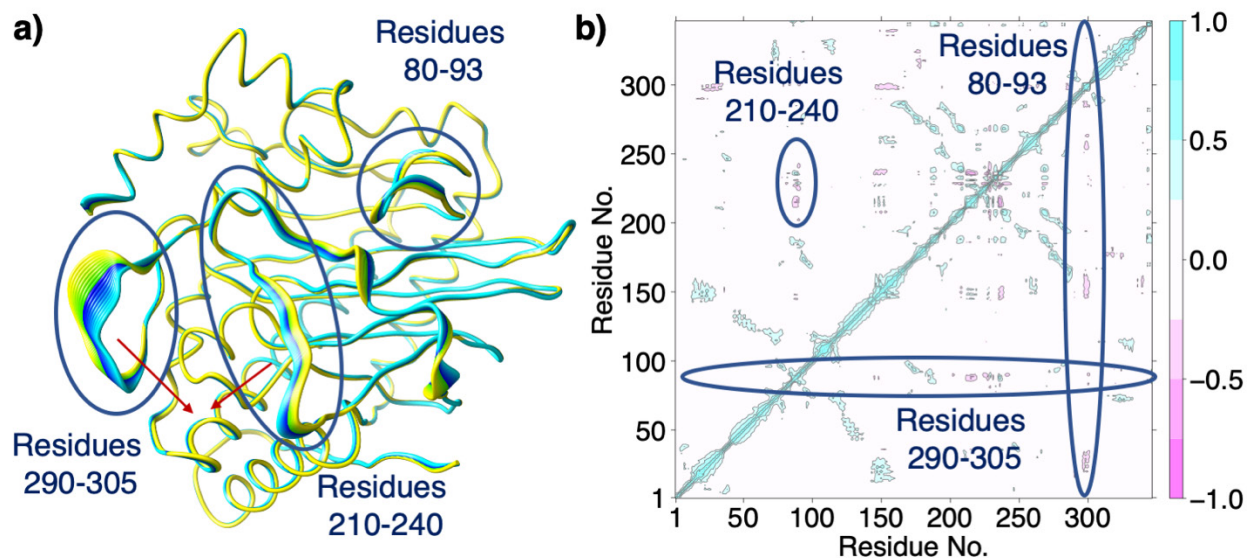


Figure 3. Overall protein dynamics in the $\text{EFE} \cdot \text{Fe(III)} \cdot \text{OO} \cdot \cdot 2\text{OG}_{\text{off-line}} \cdot \text{L-Arg}$ complex. (a) Principal component analysis shows the regions containing residues 80-93, 210-240, and 290-305 to be among the most dynamic regions. The color gradient, yellow to blue, indicates the direction of motion, also highlighted by the red arrow. (b) Dynamic cross-correlation analysis shows the motion of regions containing residues 80-93, 210-240, and 290-305 are correlated with each other. Positive values of correlation (light blue) indicate the protein regions move together in the same direction. Negative values of correlation (pink) indicate the protein regions move in the opposite direction.

3.2.1 Reaction mechanism from $\text{EFE} \cdot \text{Fe(III)} \cdot \text{OO} \cdot \cdot 2\text{OG}_{\text{off-line}} \cdot \text{L-Arg}$ using L-Arg conformation A

Prior crystallographic evidence for two L-Arg conformations in EFE led to the proposal that conformation A facilitates substrate hydroxylation while conformation B plays a role in ethylene formation.⁸ To explore whether these L-Arg binding modes account for the branched mechanism, we initiated QM/MM potential energy surface calculations using the two substrate conformations obtained from the MD simulation of the $\text{EFE} \cdot \text{Fe(III)} \cdot \text{OO} \cdot \cdot 2\text{OG}_{\text{off-line}} \cdot \text{L-Arg}$ complex.

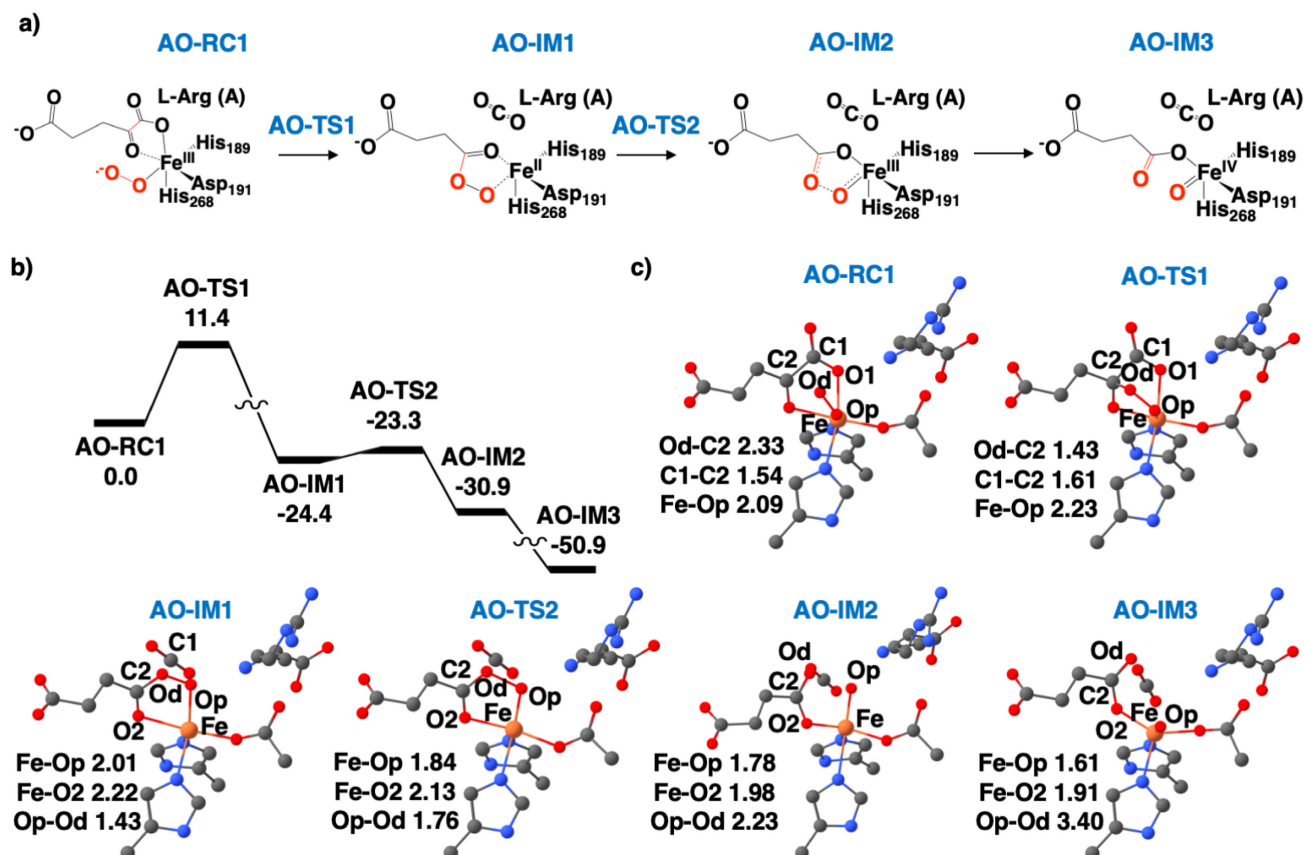


Figure 4. O₂ activation mechanism of EFE with L-Arg in conformation A (**A**) and 2OG in off-line (**O**) coordination mode. a) The proposed mechanistic pathway. b) Potential energy surface for the O₂ activation pathway. c) QM geometries of QM/MM optimized intermediates and transition states. The hydrogen atoms are hidden for clarity. Bond lengths are labeled in angstroms.

As shown in Figure 4, the EFE·Fe(III)·OO[•]·2OG_{off-line}·L-Arg reactant complex with L-Arg in conformation A (**A**) and 2OG in the off-line (**O**) mode (**AO-RC1**) was QM/MM optimized. The reaction proceeds with the attack of the distal oxygen (Od) on C2 of 2OG, leading to decarboxylation. The activation barrier for the decarboxylation is 11.4 kcal/mol (all energy barriers are reported at the QM(B2+ZPE)/MM level of theory). **AO-TS1** has an Od-C2 distance of 1.43 Å, and its C1-C2 bond is elongated to 1.61 Å. EDA identified residues stabilizing **AO-TS1** with

respect to **AO-RC1**, including R171 by -1.0 kcal/mol (Figure S5). The Fe-coordinating H189 plays a destabilizing role for **AO-TS1** by 1.73 kcal/mol. Other residues surrounding **AO-TS1** include F283, F250, and A281. The product of the reaction is a well-established succinyl-peroxide intermediate (**AO-IM1**). The activation barrier for decarboxylation matches well with values obtained in a previous computational study of EFE⁹ and with other 2OG-dependent enzymes like PHF8, AlkB, AlkBH2.^{27,28,30,41,42,43} The next step, cleavage of the O-O bond occurs with a negligible energy barrier from the succinyl-peroxide intermediate (**AO-IM1**) leading, via an unstable half-bond intermediate (**AO-IM2**), to a ferryl species (**AO-IM3**). Since we started with off-line bound 2OG, the generated ferryl points away from the substrate. It is proposed that a flip towards a catalytically productive in-line orientation occurs at the stage of the off-line ferryl or at an earlier intermediate. The mechanistic pathways for such a flip were previously explored, for example, in PHF8, AlkBH2, and AsqJ.^{27,42,43} In order to explore the L-Arg hydroxylation step, an in-line ferryl (**AO-IM4**) was QM/MM optimized (note: the designation of this intermediate retains the nomenclature of the above species from which it is derived even though off-line no longer applies).

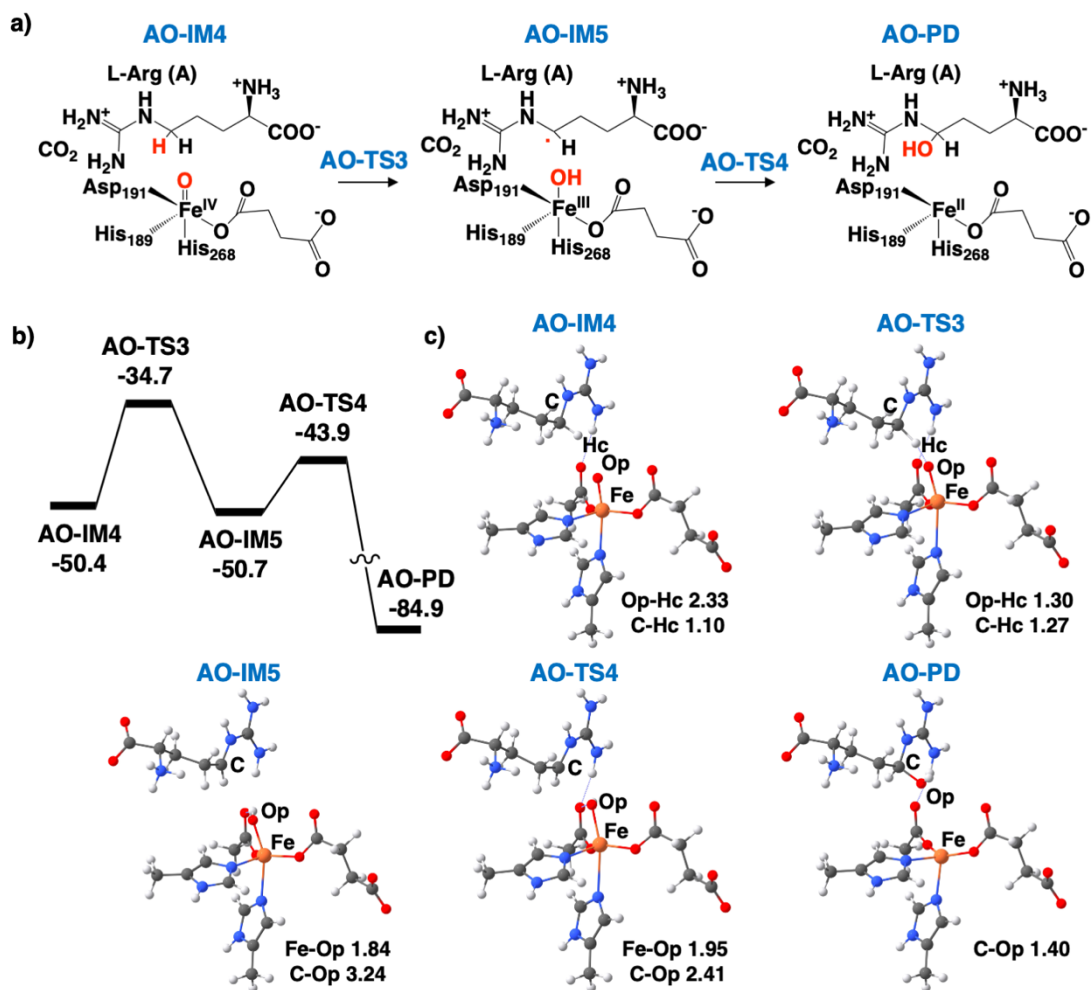


Figure 5. The hydrogen atom transfer and rebound hydroxylation mechanism in EFE with L-Arg in conformation A (**A**) and an in-line ferryl species (the designation **O** is retained because the starting 2OG was off-line). a) The proposed mechanistic pathway. b) Potential energy surface for HAT and rebound hydroxylation. c) QM geometries of QM/MM optimized intermediates and transition states. The hydrogen atoms and CO₂ molecule are hidden for clarity. Bond lengths are labeled in angstroms.

The properly directed ferryl intermediate (**AO-IM4**) abstracts hydrogen from C5 of L-Arg to produce a substrate-based radical and an Fe(III)-OH intermediate (**AO-IM5**) as shown in Figure 5. The energy required to cross the transition state (**AO-TS3**) for hydrogen abstraction (the rate-

determining step for the hydroxylation reaction) is 15.7 kcal/mol. **AO-TS3** has an Op-Hc distance of 1.30 Å and an Hc-C distance of 1.27 Å. The residues stabilizing **AO-TS3** include R171 and F283. The hydroxylation of L-Arg is completed by a rebound of the hydroxyl group to C5 (**AO-PD**) with an activation barrier of 6.8 kcal/mol. The energetics and structural features of the transition states and intermediates for hydroxylation match well with previous computational studies on EFE and similar non-heme iron/2OG-dependent enzymes like PHF8, AlkB, and AlkBH2.^{9,30,42-47}

We also explored the possibility of the five-coordinated (5C) ferryl species (**AO-IM4**) converting to a six-coordinated (6C) ferryl intermediate (**AO-IM4-6coord**), where the succinate coordinates to the iron in a bidentate fashion, before performing subsequent reaction mechanism steps (Figure S6 and S7). The conversion from 5C to 6C ferryl possesses an energy barrier of 5.4 kcal/mol, and the 6C ferryl **AO-IM4-6coord** is 3.1 kcal/mol higher in energy in comparison to the 5C ferryl (**AO-IM4**). The hydrogen atom abstraction starting from **AO-IM4-6coord** has an energy barrier of 14.3 kcal/mol, passing through a 6C TS (**AO-TS3-6coord**) to give a substrate-based radical and a 6C Fe(III)-OH intermediate **AO-IM5-6coord**. The energy barrier for 6C HAT is lower by 1.4 kcal/mol than the one for 5C HAT. However, the *overall path of the conversion from 5C ferryl to 6C ferryl and subsequent hydrogen abstraction* exhibits a total energy barrier of 17.4 kcal/mol, which is 1.7 kcal/mol higher than the energy required for a direct HAT mechanism from the 5C ferryl **AO-IM4**. Recent spectroscopic studies of 2OG-dependent oxygenase TauD showed the ferryl to be present in a five-coordination geometry.⁴⁸ We, however, note that in our calculations, the difference in the activation barriers is relatively small to make a rigorous conclusion. The rebound step that completes L-Arg hydroxylation from the 6C Fe(III)-

OH intermediate **AO-IM5-6coord** is characterized by an energy barrier of 9.1 kcal/mol, which is 2.3 kcal/mol higher than the one for the 5C rebound reaction.

Earlier experimental studies proposed that F283 might prevent the ferryl flip so that the off-line ferryl might lead to ethylene formation.⁷ Therefore, we further tested the proposal that intermediates along with the standard O₂ activation mechanism (with L-Arg bound in conformation A and 2OG coordinated off-line), including the off-line ferryl, can produce ethylene.^{7,8,49} We calculated the reaction paths for producing ethylene with a reaction coordinate combining an increase in the C2-C3 bond length and a reduction in the C3-C4 bond length (with numbering retained from 2OG), starting from each of the above intermediates (succinyl-peroxide **AO-IM1**, half-bond **AO-IM2**, and off-line ferryl **AO-IM3**). The activation energies required to decompose 2OG to form ethylene from **AO-IM1**, **AO-IM2**, and **AO-IM3** are 68.1, 19.2, and 43.3 kcal/mol, respectively. Therefore ethylene formation from the succinyl-peroxide **AO-IM1** and the off-line ferryl **AO-IM3** intermediates seems unrealistic because of the high energy requirement, in agreement with earlier studies of EFE.⁹ The ethylene formation pathway from the half-bond **AO-IM2** intermediate has a reasonable barrier, but the stability of **AO-IM2** is a concern as it very quickly rearranges to **AO-IM3**. Thus ethylene formation from any of the standard O₂ activation intermediates for the starting sample with L-Arg in conformation A and an off-line 2OG seems unrealistic.

3.2.2 Reaction mechanism from EFE·Fe(III)·OO··2OG_{off-line}·L-Arg using L-Arg conformation B

We next performed QM/MM optimization of the $\text{EFE}\cdot\text{Fe(III)}\cdot\text{OO}\cdot\cdot\text{2OG}\cdot\text{Arg}$ intermediate with L-Arg in conformation B (**B**) and 2OG bound in the off-line (**O**) mode (**BO-RC1**) to explore how the change of the L-Arg conformation might influence the direction of the chemical transformations (Figure 6). Despite using the same reaction coordinate, i.e., with the Od oxygen attacking C2 of 2OG, the succinyl peroxide intermediate was not produced. Rather, oxygen attack using this substrate conformation led to breakage of the C1-C2 bond in 2OG with the formation of bonds between C1 and the proximal oxygen (Op) and between Od and C2 (**BO-IM1**). In contrast to the reaction forming the succinyl-peroxide intermediate **AO-IM1**, carbon dioxide is not produced. The activation barrier to cross **BO-TS1** was found to be 10.0 kcal/mol. The Od-C2 and C1-C2 distances in **BO-TS1** are 1.42 Å and 1.62 Å, respectively, (Figure 6) with an imaginary frequency of -85.69 cm^{-1} along the C1-C2 bond axis (Figure S8). EDA identified residues stabilizing **BO-TS1** with respect to **BO-RC1**, including R171 by -1.2 kcal/mol (Figure S9). The Fe-coordinating H189 plays a destabilizing role for **BO-TS1** by 0.96 kcal/mol. The other residues surrounding **BO-TS1** include F283, A281, and F250. R171 exhibits correlated motions with β_6 and β_{14} , while the hydrophobic region containing residues F283, F250, and A281 has correlated motions with β_{15} .

The peroxic anhydride species (**BO-IM1**) was previously proposed as an intermediate for the ethylene-forming mechanism,⁹ but in that case, the species was suggested to arise by a very different process – an insertion of CO_2 into the Fe-peroxysuccinate bond. The addition of CO_2 in the previous study was forced by the reaction coordinate in the presence of conformation A of L-Arg.⁹ To explore in detail the formation of **BO-IM1** in our calculations, we repeated the relaxed potential energy scan by including the C1-C2 bond in the reaction coordinate with the attack of Od on C2 of 2OG as we could not locate any structure between **BO-TS1** and **BO-IM1** in previous

scans (Figure S10), even with a smaller step size of 0.025 Å. The relaxed potential energy scan located a **pre-BO-IM1** structure (Figure S11) and suggests that after **BO-TS1**, the CO₂ separates and then forms a bond with Op, as it is oriented close to the Op (2.58 Å between Op and C1 of CO₂). However, we are unable to freely optimize the **pre-BO-IM1** structure, as it converges to the geometry of **BO-IM1**, indicating the formation of the Op-C1 bond is highly favorable in structures with L-Arg in conformation B. Furthermore, it is important to note that while a relaxed potential energy scan allows structures to relax, it cannot comprehensively account for the flexibility of the active site and the entire enzyme. It might be possible that if such flexibility is accounted for more comprehensively within the reaction path calculations, we could see a more favorable positioning of C1 with respect to Op. Such insight can be done either by i) performing the reaction path calculations using restrained QM/MM MD (e.g., QM/MM Potential of Mean Force or QM/MM Metadynamics); or ii) performing a series of potential energy scans using multiple (five or more) initial structures conditions taken from MD. However, both approaches are very computationally demanding and would be the subject of future studies.

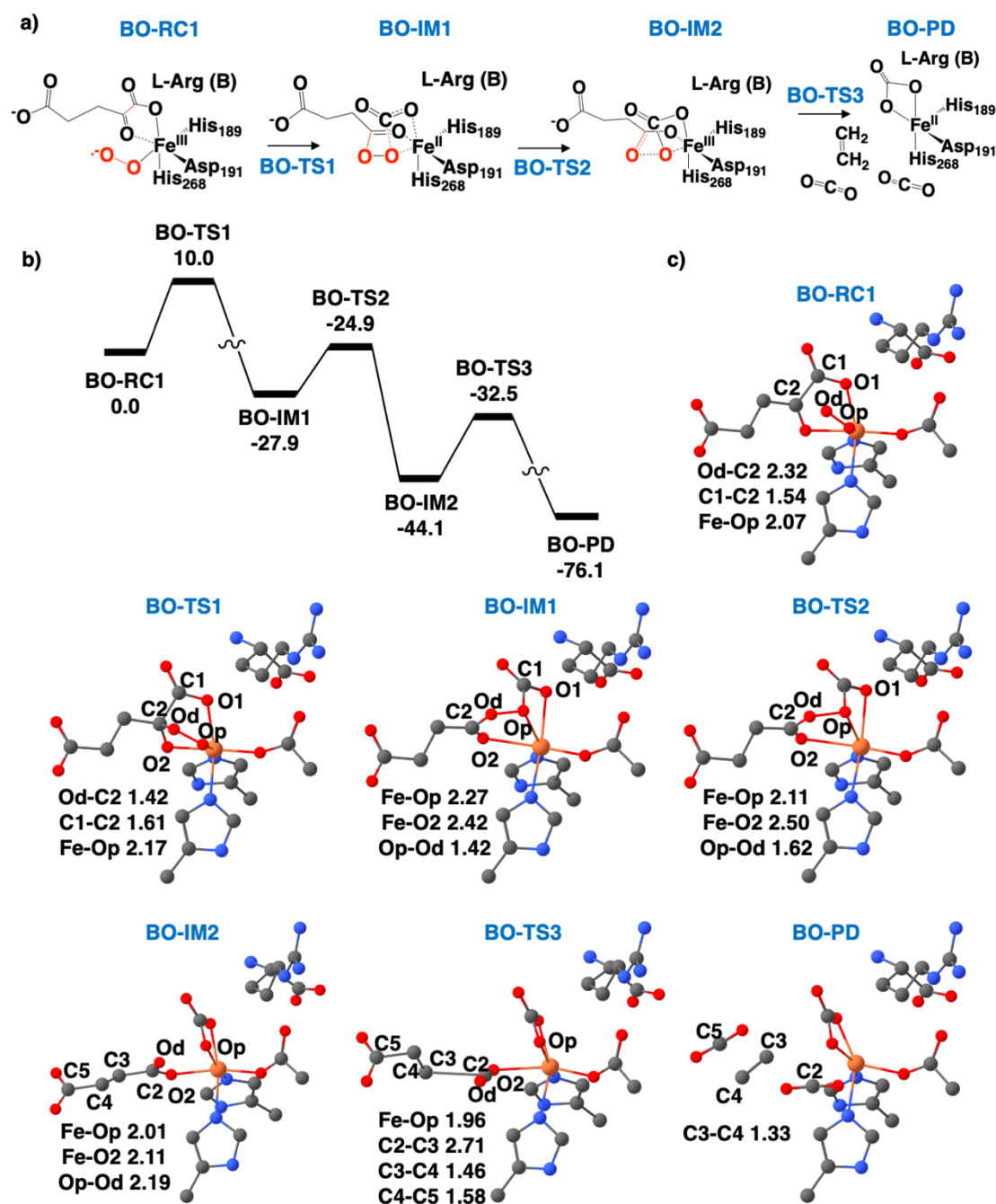


Figure 6. The ethylene-forming mechanism of EFE with L-Arg in conformation B (**B**) and 2OG in off-line (**O**) coordination mode. a) The proposed mechanistic pathway. b) Potential energy surface for the ethylene-forming pathway. c) QM geometries of QM/MM optimized intermediates and transition states.

The hydrogen atoms are hidden for clarity. Bond lengths are labeled in angstroms.

Similar to the O-O cleavage step in the standard O₂ activation mechanism (Figure 4), where two electrons are transferred from Fe in a stepwise fashion, forming an Fe(III) half-bond intermediate (**AO-IM2**), and then an Fe(IV) ferryl intermediate (**AO-IM3**),²⁸ the **BO-IM1** intermediate undergoes two one-electron transfers. The reaction passes through a transition state (**BO-TS2**) with an energy barrier of 3.0 kcal/mol and forms the **BO-IM2** intermediate with a partial O-O bond, labeled as a half-bond ferric-bicarbonate intermediate with bound succinate. The distance between the two oxygen atoms is 2.2 Å with a radical shared between them (spin density of -0.54 and -0.22 on O_d and O_p). A finer relaxed potential energy surface scan with a smaller step size (0.025 Å) shows two new intermediates with energies very close to **BO-IM2** (Figure S12)- an Fe(III)-bicarbonate intermediate with completely cleaved O-O bond and a radical on the C5 carboxylate of succinate (**BO-IM3**, retaining the numbering from 2OG, with the geometry shown in Figure S13) and another intermediate (**BO-IM4**) with a delocalized radical over the succinate carbon chain (Figure S14). The **BO-IM4** undergoes C2-C3 bond decomposition (with an energy barrier of 11.6 kcal/mol from **BO-IM2**), to form a propionate radical and carbon dioxide (**pre-BO-PD**, Figure S15). Upon free optimization, **pre-BO-PD** decomposes to ferrous-bicarbonate, ethylene, and two carbon dioxide molecules (**BO-PD**).

The reaction path calculations were repeated using snapshots of the L-Arg conformations A and B from molecular dynamics, and similar results were obtained (Figure S16-S20).

3.3 Electronic structure analysis of O₂ activation and the ethylene-forming mechanism of EFE

To explore how the different L-Arg conformations and interactions influence the electronic structure of the EFE Fe center that leads to ethylene formation vs. the hydroxylation

mechanism, we performed a detailed molecular orbital analysis. 2OG is coordinated to the Fe center in an off-line mode;^{7,8} hence the orbital axes are labeled accordingly. The z-axis is directed along with the Fe-O1 bond (where O is part of C1 carboxylate of 2OG), the y-axis is along Fe-O2 (the keto oxygen atom of 2OG), and the x-axis is Fe-O₂ (Figure 7). The reactant complex with L-Arg in conformation A (**AO-RC1**) is an Fe(III)-superoxo complex with a high-spin Fe(III) center ($S=5/2$) that is antiferromagnetically coupled to the superoxo anion radical ($S = 1/2$). The electronic configuration of **AO-RC1** is $1d_{yz} 1d_{xy} 1d_{xz} 1d_{z^2} 1d_{x^2-y^2} 1\pi^*_{\perp} 1\pi^*_{\parallel}$. The bonding between the iron and dioxygen originates from singlet coupling between the Fe d_{xz} and the $\pi^*_{\perp}$ orbital of the O₂, as seen in the natural orbitals (NO) and the spin natural orbitals (SNO) (Figure 7a). As the oxygen attacks C2 of 2OG, resulting in decarboxylation, two electrons from the σ_{C1-C2} bond orbital of 2OG are transferred to the Fe-O bond orbitals. An α electron is transferred to the $\pi^*_{\perp}$ orbital of the dioxygen, while a β electron is transferred to the d_{xz} orbital of Fe. The resulting electronic structure for the succinyl-peroxide intermediate (**AO-IM1**) is $1d_{yz} 1d_{xy} 1d_{xz} 1d_{z^2} 1d_{x^2-y^2} 1\pi^*_{\perp} 1\pi^*_{\parallel}$. The subsequent homolytic cleavage of the Op-Od bond in the succinyl-peroxide intermediate is effectively achieved by another two-electron transfer, this time from Fe (II) in a two-step process. The first electron is transferred from the d_{xz} orbital of Fe to the σ^*_{Op-Od} orbital of dioxygen. This transfer results in the half-bond intermediate (**AO-IM2**). Then the second electron is transferred from the $d_{x^2-y^2}$ orbital of Fe to the σ^*_{Op-Od} orbital of dioxygen to complete the Op-Od bond cleavage. This step yields the final intermediate from the oxygen activation mechanism - the off-line ferryl (**AO-IM3**) with an electronic configuration of $1d_{yz} 1d_{xy} 1d_{xz} 1d_{z^2} 0d_{x^2-y^2}$. Upon rearrangement of ferryl to the in-line ferryl (**AO-IM4**), the electronic configuration shifts to $1d_{yz} 1d_{xy} 1d_{xz} 1d_{x^2-y^2} 0d_{z^2}$. HAT then occurs through a σ^* transfer from the σ_{CH} orbital of C5 in L-Arg to the d_{z^2} orbital of the ferryl,

leading to an electronic structure of $1d_{yz}1d_{xy}1d_{xz}1d_{z^2}1d_{x^2-y^2}$ in the Fe(III)-OH intermediate (**AO-IM5**) and a substrate based unpaired β electron.

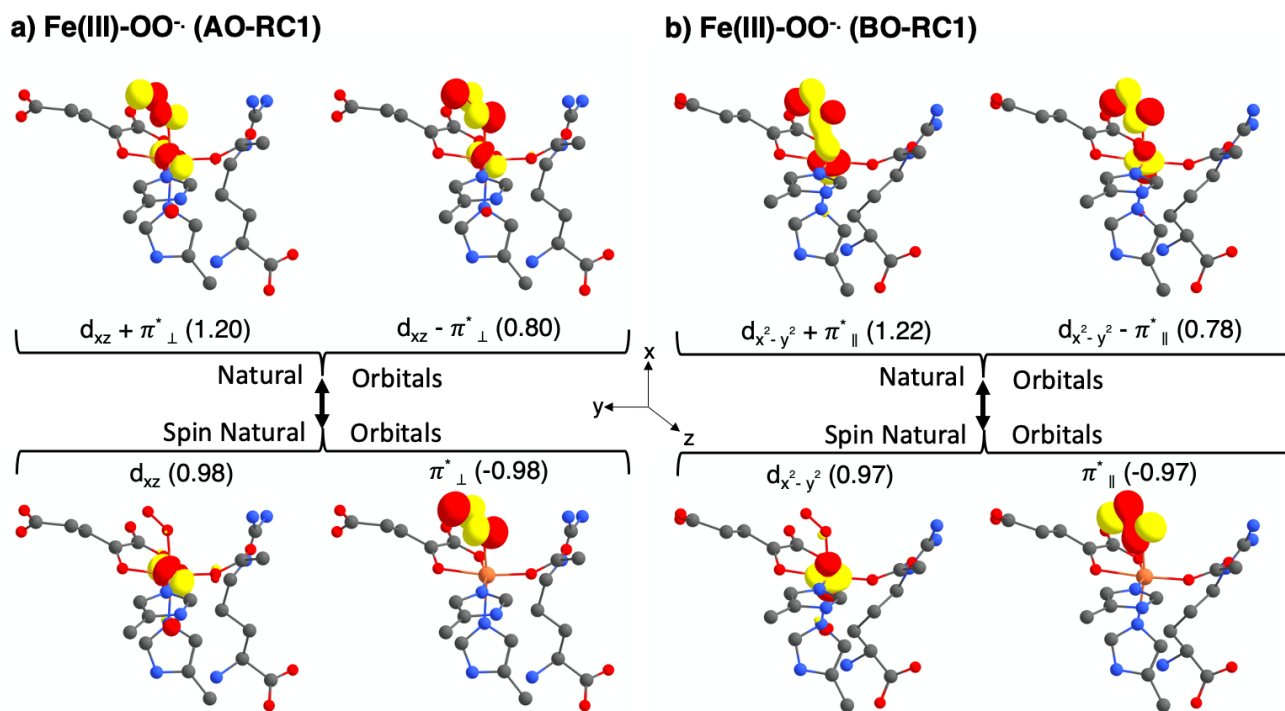


Figure 7. Change in bonding character of the Fe-O bond shown using natural orbitals and spin natural Orbitals in Fe-OO $\cdot$ upon changing the L-Arg conformation from A to B. The hydrogen atoms are hidden, and the orientation of the Fe center is rotated for clarity in depicting orbitals.

The electronic structure analysis of the Fe center in the reactant complex with L-Arg in conformation B (**BO-RC1**) shows a different electronic configuration of $1d_{xy}1d_{xz}1d_{yz}1d_{x^2-y^2}1d_{z^2}1\pi^*_{\perp}1\pi^*_{\parallel}$. Importantly, the unpaired β electron of the superoxo anion radical is present in the $\pi^*_{\parallel}$ dioxygen orbital instead of the $\pi^*_{\perp}$ orbital as for **AO-RC1**. A previous QM calculation for a non-heme quintet model system showed a switching of the dioxygen orbital with the unpaired β electron, as seen in the electronic configuration of **AO-RC1** and **BO-RC1**, which are isoenergetic

in the absence of a protein environment.⁵⁰ Based on current calculations, we hypothesize that one possible reason for the unpaired β electron switching in EFE can be the destabilization of the $\pi^*_{||}$ dioxygen orbital by exchange repulsion with F283. The side chain of F283 can be seen to come slightly closer towards Op for conformation B than for conformation A, which may potentially destabilize the $\pi^*_{||}$ dioxygen orbital in conformation B (Figure S21 and S22). However, further calculations using restrained QM/MM MD (e.g., QM/MM Potential of Mean Force or QM/MM Metadynamics) that can comprehensively account for the flexibility around the active-site of EFE may be required to confirm this hypothesis. This finding may also explain why substituting residues in the hydrophobic region of the EFE active site completely abolishes the ethylene-forming reaction while retaining varying levels of hydroxylation activity⁷ – by potentially changing the hydrophobic interactions with respect to the O₂ and disrupting the electronic structure changes that lead to ethylene formation. As the unpaired β electron is present in the $\pi^*_{||}$ dioxygen orbital, the Fe-O₂ bond is formed by a singlet coupling between the Fe $d_{x^2-y^2}$ and the $\pi^*_{||}$ orbital of the dioxygen as seen in NO and SNO (Figure 7b). The Op-C1 bond that is present in **BO-IM1**, but not in **AO-IM1**, is perpendicular to the Fe-O bond and is formed with the overlap of the non-bonding $\pi^*_{\perp}$ orbital of dioxygen and a C1 orbital. The O₂ attack on C2 of 2OG transfers two electrons from the σ_{C1-C2} bond orbital of 2OG with one electron going to the $\pi^*_{||}$ orbital of superoxo and the other electron to the $d_{x^2-y^2}$ orbital of Fe, giving an electronic configuration of $1d_{xy} 1d_{xz} 1d_{yz} 1d_{x^2-y^2} 1d_z^2 1\pi^*_{\perp} 1\pi^*_{||}$ for **BO-IM1**. Subsequently, a half-bond ferric-bicarbonate intermediate (**BO-IM2**) is generated as a β electron from the Fe $d_{x^2-y^2}$ orbital is transferred to the σ^*_{Op-Od} orbital of dioxygen. In contrast to the standard O-O bond cleavage (from **AO-IM2** to **AO-IM3**), here the second electron required to complete breakage of the Op-Od bond is transferred

from the C5 carboxylate of succinate (using the original 2OG numbering) to the $\sigma^*_{\text{Op-Od}}$ orbital. Therefore, the resulting intermediate **BO-IM3** is a ferric Fe(III)-bicarbonate intermediate with a radical on C5 carboxylate of succinate. The succinate radical becomes delocalized over the succinate carbons to elongate C2-C3 and C4-C5 bonds while slightly reducing the C3-C4 bond length in BO-IM4 before decomposing to form ferrous-bicarbonate, ethylene, and two molecules of carbon dioxide (**BO-PD**).

The switching of the dioxygen orbital containing the unpaired β electron seems to be a function of the L-Arg conformation in $\text{EFE}\cdot\text{Fe(III)}\cdot\text{OO}^\cdot\cdot\text{2OG}_{\text{off-line}}\cdot\text{L-Arg}$. The location of the unpaired β electron determines which O_2 π^* and Fe d-orbital constitute the $\sigma_{\text{Fe-O}}$ bond. In the case of L-Arg conformation A, where the unpaired β electron is present in the $\pi^*_{\perp}$ dioxygen orbital, the $\sigma_{\text{Fe-O}}$ bond is formed by the d_{xz} orbital of Fe and $\pi^*_{\perp}$ dioxygen orbital leaving $\pi^*_{\parallel}$ as a doubly occupied non-bonding orbital. For L-Arg conformation B, the unpaired β electron is present in the $\pi^*_{\parallel}$ dioxygen orbital, and the $\sigma_{\text{Fe-O}}$ bond is formed by the $d_{x^2-y^2}$ orbital of Fe and $\pi^*_{\parallel}$ dioxygen orbital with $\pi^*_{\perp}$ as a doubly occupied non-bonding orbital. The C1 carboxylate approaches Op from the direction perpendicular to the Fe-O bond and thus requires a non-bonding doubly occupied $\pi^*_{\perp}$ orbital to form the Op-C1 bond. **BO-RC1** has a non-bonding doubly occupied $\pi^*_{\perp}$ orbital, and thus an Op-C1 bond is formed in **BO-IM1** while the non-bonding doubly occupied orbital is $\pi^*_{\parallel}$ in **AO-RC1** and hence an Op-C1 bond is not formed in **AO-IM1**. This result was confirmed by repeating the electronic structure analysis on two other snapshots (one for each of the two substrate conformations); the data are shown in Figure S23. Thus, the difference in the Fe-O₂ bonding orbitals and the resulting non-bonding dioxygen orbital leads to the

formation of a chemically different **BO-IM1** intermediate with different orbital occupations, which separates the ethylene pathway from hydroxylation.

3.4 Rearrangement between the coordination modes of 2OG prior to O₂ binding

The EFE·Mn·2OG crystal structure (PDB ID 5V2X) shows 2OG in monodentate coordination to Fe, whereas the EFE·Fe·NOG·L-Arg and EFE·Mn·2OG·L-Arg crystal structures (PDB ID 5LUN and 5V2Y) show bidentate off-line binding of 2OG.^{7,8} The crystal structures and QM/MM calculations on similar non-heme Fe 2OG-dependent enzymes like PHF8, KDM4A, AlkB, and AlkBH2 showed the 2OG could exist in both off-line and in-line bidentate coordination to Fe.^{42,43,51-55} To explore the 2OG coordination mode in EFE prior to O₂ binding, we started by optimizing the five-coordinate EFE·Fe·2OG·L-Arg complex. The QM/MM optimized enzyme-substrate complex of EFE·Fe·2OG_{off-line}·L-Arg (**AO-ES**) has the L-Arg in conformation A and 2OG in the off-line coordination mode (Figure S24). The torsion angle derived from N_{H189}-Fe-O2-O1 can be used to differentiate the coordination modes of 2OG, with around 90° representing the off-line coordination mode and around 180° representing the in-line coordination mode. **AO-ES** has Fe-O1 and Fe-O2 bond lengths of 2.04 Å and 2.27 Å, respectively, and an N_{H189}-Fe-O2-O1 torsion angle of 109.8°. In order to reveal the barrier for the transition from off-line to in-line 2OG coordination, we performed a scan of the potential energy surface by gradually rotating the dihedral angle N_{H189}-Fe-O2-O1 from 109.9° towards 180.0° to generate the in-line bound 2OG structure. The energy barrier for the rotation is 21.8 kcal/mol, and the in-line bound 2OG five-coordinate EFE enzyme-substrate complex (**AI-ES**) is higher in energy by 20.9 kcal/mol compared to the off-line bound complex (**AO-ES**). **AI-ES** has Fe-O1 and Fe-O2 bond lengths of 1.99 Å and 2.31 Å with a dihedral angle N_{H189}-

Fe-O2-O1 of 168° (Figure S24). The much stronger stabilization of **AO-ES** with respect to **AI-ES**, by 20.9 kcal/mol, suggests that the off-line coordination of 2OG to the Fe-center is highly favored over the in-line 2OG coordination before O₂ binds to EFE. The energy barrier of the transformation from **AO-ES** to **AI-ES** is 21.8 kcal/mol. The calculations using conformation B of L-Arg at the five-coordination stage show the same results (Figure S25). Even after O₂ binding in EFE, the comparison of the QM/MM optimized structures energies show a trend similar to the **AO-ES** and **AI-ES** energies, the off-line 2OG binding mode in EFE is more stable by 22.6 kcal/mol as compared to the in-line 2OG binding mode. In contrast, similar QM/MM calculations in PHF8 and AlkBH2 show the transition from off-line 2OG to in-line 2OG before O₂ binding is 1.4 and 2.9 kcal/mol, respectively, with the in-line coordinated 2OG slightly more stable than the off-line species.^{42,43} The preference for an off-line coordination mode in **AO-ES** of EFE can be due to the strong C1 carboxylate stabilization by R171, which is not maintained in the in-line coordination mode. QM/MM level analysis of an R171A variant shows an energy barrier for transforming **AO-ES** to **AI-ES** of only 9.6 kcal/mol compared to 21.8 kcal/mol calculated for the wild-type protein. Importantly, the R171A variant makes the **AI-ES** slightly more stable than **AO-ES** (by 2.95 kcal/mol). This results in a dramatic change from the wild-type enzyme where the **AO-ES** was strongly stabilized by 20.9 kcal/mol versus **AI-ES**. Furthermore, EFE exhibits a significantly hydrophobic environment around the C1 carboxylate for the in-line coordinated 2OG, which is not the case for PHF8, AlkB, and AlkBH2. Thus, R171 stabilization in the off-line coordination mode and the presence of a highly hydrophobic environment in the in-line coordination mode of 2OG might explain the preference for off-line coordination to the Fe center in EFE before O₂ binding.

3.5 Dynamics of the EFE·Fe(III)·OO⁻·2OG·L-Arg with 2OG in the in-line coordination mode

The crystal structures of EFE·Fe·NOG·L-Arg and EFE·Mn·2OG·L-Arg (PDB ID 5LUN and 5V2Y) and the results of 2OG rearrangement before O₂ binding indicate the enzyme prefers to bind 2OG in the off-line mode in the presence of L-Arg.^{7,8} However, experiments on the non-heme Fe and 2OG-dependent KDM2A show that binding nitric oxide (an analog of O₂) can influence other binding modes of Fe center.⁵⁶ We wanted to investigate if O₂ binding by EFE leads to an in-line bound 2OG and then determine whether this species can produce ethylene. We carried out a 1 μ s MD simulation starting with the EFE·Fe(III)·OO⁻·2OG_{in-line}·Arg complex (Figure S26). Similar to the MD trajectory of the off-line bound 2OG intermediate, the in-line bound 2OG dynamics show L-Arg exists in both conformations A and B (Figure S27). In contrast to the 7.5-fold preference for conformation B during off-line 2OG dynamics, this preference was less than two-fold during in-line 2OG dynamics – conformation A occurs in 36.7% of the trajectory and conformation B in 63.3%. Also, the switching between the L-Arg conformations is not as frequent as in the off-line bound 2OG dynamics. A hydrogen bond between η^2 NH₂ of the L-Arg guanidinium and the D191 carboxylate was present in 68% of the MD snapshots, and a hydrogen bond between the ϵ NH and D191 carboxylate was seen in 25% of the MD snapshots. The C5 carboxylate, as found for the off-line bound 2OG dynamics, does not possess a strained torsion angle for the O3-C5-O4 and C3-C4-C5 planes (Figure S28). The C5 carboxylate exhibits a hydrogen bonding interaction with the R277 (90%). Shifting the 2OG coordination from off-line to in-line leads to a loss of the C1 carboxylate stabilization by R171, and the carboxylate becomes surrounded by a set of hydrophobic residues including F283, A281, V196, A198, and F250 (Figure

S29). In contrast to the off-line bound 2OG dynamics, in the simulation with the in-line 2OG, the peptide bond relaxes to an average value of 186.1° , without much deviation, indicating the changes in the Fe center might affect the backbone orientations of D191 and Y192 (Figure S30) and again suggesting the twisted peptide bond is not critical to EFE catalysis.

The overall dynamics of the EFE·Fe(III)·OO^{·-}·2OG_{in-line}·Arg complex also shows a notable difference from the dynamics of the reactant complex with off-line bound 2OG (Figure 8). PCA shows only the region containing $\beta 4$, $\beta 5$ and the loops connecting them (residues 80-93) show some flexibility while the motions in the regions $\beta 11$, $\beta 15$, and their loops are lost. Also, the region with $\beta 4$, $\beta 5$, and their loops move in the direction away from $\beta 11$, which is opposite to the direction of motion in the off-line 2OG dynamics. DCCA of the dynamics shows the loss of correlated motions from the region of $\beta 11$, $\beta 15$, and their loops. Thus, the results highlight how the change in the coordination mode of the 2OG has the potential to influence the long-range correlated motions of EFE.

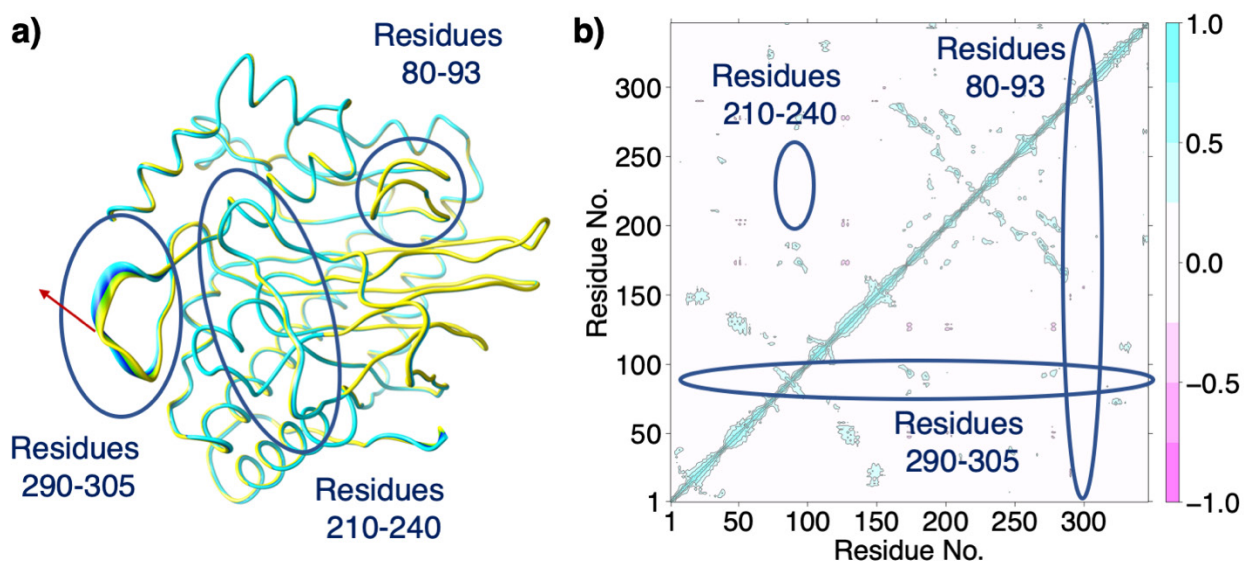


Figure 8. The overall protein dynamics in the EFE·Fe(III)·OO^{·-}·2OG_{in-line}·L-Arg complex. (a) PCA shows a lack of motion in the regions containing residues 210-240 and 80-93 compared to what was found for the off-

line 2OG dynamics, whereas the dynamics are reduced with an altered direction of motion for residues 290-305. The color gradient yellow to blue indicates the direction of motion, also highlighted by the red arrow. (b) DCCA shows loss of the correlated motions in the regions containing residues 290-305, 210-240, and 80-93 in comparison to off-line 2OG dynamics. Positive values of correlation (light blue) indicate the region of protein moves together in the same direction. Negative values of correlation (pink) indicate the regions of protein move either in the same direction to each other or away from each other.

3.6 Reaction mechanism from $\text{EFE}\cdot\text{Fe(III)}\cdot\text{OO}^{\cdot-}\cdot\text{2OG}_{\text{in-line}}\cdot\text{L-Arg}$ using two L-Arg conformations

To explore whether the in-line coordination mode of 2OG affects ethylene formation for the two L-Arg conformations, structures of each substrate conformation were used from the dynamics for calculations of the potential energy surface. For one snapshot, the reactant complex $\text{EFE}\cdot\text{Fe(III)}\cdot\text{OO}^{\cdot-}\cdot\text{2OG}_{\text{in-line}}\cdot\text{Arg}$ with L-Arg in conformation A (**A**) and 2OG bound using the in-line (**I**) mode (**AI-RC1**) was QM/MM optimized (Figure 9). The reaction coordinate of the Od oxygen attack on the C2 atom of 2OG led to decarboxylation and formation of a far more stable succinylperoxide intermediate (**AI-IM1**) than in the off-line mode with conformation A. The activation barrier for decarboxylation with in-line bound 2OG is higher at 15.2 kcal/mol as compared to 11.4 kcal/mol for off-line bound 2OG with the same L-Arg conformation A. EDA shows the reason for higher energy required to cross the **AI-TS1** is in part due to destabilization by the Fe-coordinating H268 (by 2.43 kcal/mol) on the **AI-TS1** in comparison to the **AI-RC1** (Figure S31). From **AI-IM1**, a half-bond intermediate (**AI-IM2**) is formed with a negligible barrier and quickly rearranges to give a ferryl intermediate (**AI-IM3**). Since the formation of a ferryl intermediate (**AI-IM3**) is from an in-line bound 2OG, no rearrangement is required, and this intermediate can directly proceed to L-Arg hydroxylation.

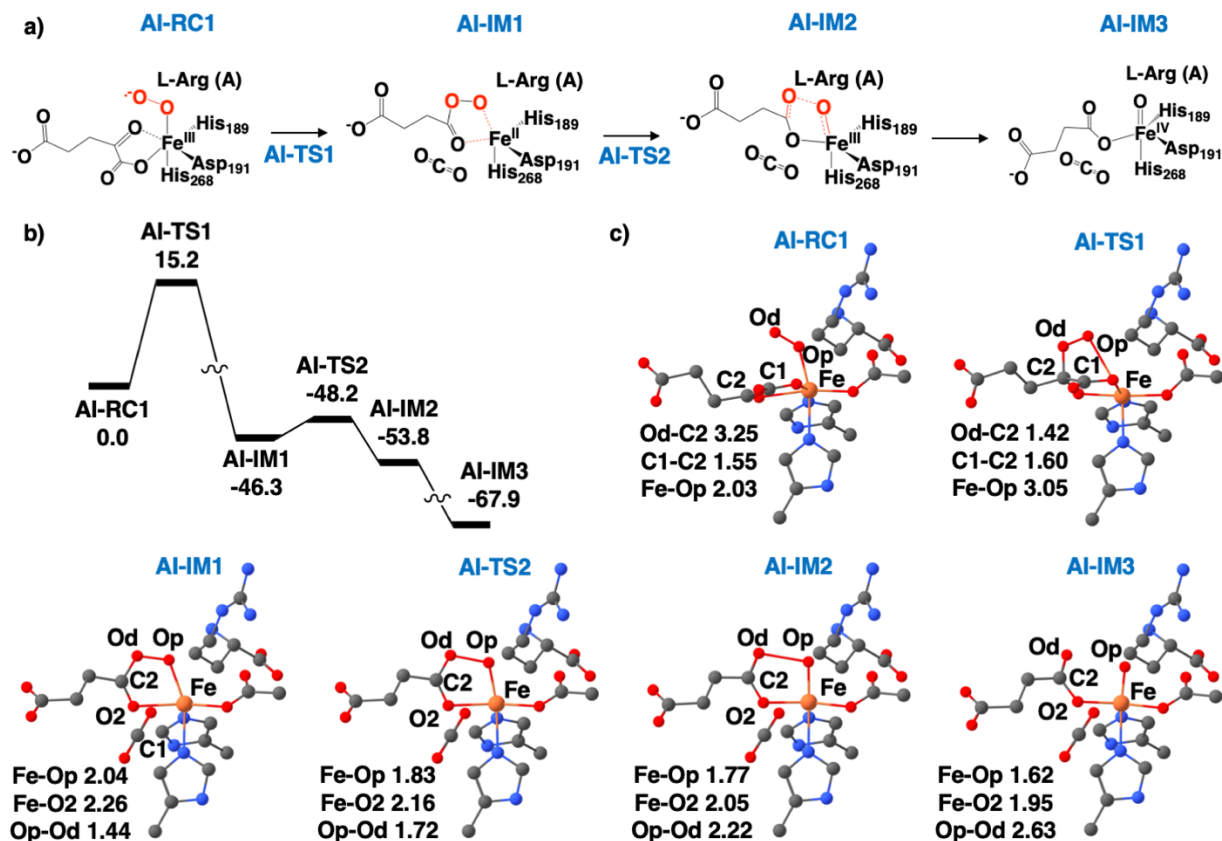


Figure 9. The O₂ activation mechanism of EFE with L-Arg in conformation A (**A**) and 2OG in the in-line (**I**) coordination mode. a) The proposed mechanistic pathway. b) Potential energy surface for the O₂ activation pathway. c) QM geometries of QM/MM optimized intermediates and transition states. The hydrogen atoms are hidden for clarity. Bond lengths are labeled in angstroms.

The ferryl intermediate (**AI-IM3**) abstracts a hydrogen atom from the C5 methylene of L-Arg with an energy barrier of 12.0 kcal/mol to generate a ferric-hydroxyl intermediate (**AI-IM4**) (Figure S32). The HAT transition state (**AI-TS3**) is stabilized by 3.7 kcal/mol as compared to the off-line 2OG bound snapshot's HAT transition state (**AO-TS3**) with the same A conformation of the substrate. The conversion from 5C ferryl (**AI-IM3**) to 6C ferryl and then performing hydrogen atom abstraction involves an energy barrier of 17.8 kcal/mol (Figure S33) which is higher than

direct hydrogen abstraction step from the 5C ferryl intermediate by 5.8 kcal/mol - a more substantial difference than for off-line 2OG and L-Arg conformation A calculations.

The ferric hydroxyl intermediate (**AI-IM4**) completes hydroxylation of the substrate via a rebound mechanism with an energy barrier of 2.0 kcal/mol.

A snapshot from the dynamics for the reactant complex $\text{EFE} \cdot \text{Fe(III)} \cdot \text{OO}^{\cdot-} \cdot 2\text{OG}_{\text{in-line}} \cdot \text{Arg}$ (**BI-RC1**) with L-Arg in conformation B (**B**) and 2OG bound in in-line (**I**) mode was QM/MM optimized (Figure 10). From **BI-RC1**, the reaction coordinate for the Od attack on C2 of 2OG led to the formation of a succinyl-peroxide intermediate (**BI-IM1**). The energy barrier for the process was quite high, at 21.8 kcal/mol. EDA shows that, in contrast to the stabilizing role of R171 in the off-line bound 2OG **BO-TS1**, R171 plays a highly destabilizing role (by 6.0 kcal/mol) in **BI-TS1** with respect to **BI-RC1** (Figure S34). Subsequently, the half-bond intermediate (**BI-IM2**) is formed with a very low barrier and quickly rearranges to give a ferryl intermediate (**BI-IM3**). Since L-Arg is in conformation B, which is not very energetically efficient for hydrogen abstraction, the energy barrier for the step is 22.1 kcal/mol (Figure S35). The conversion from 5C ferryl to 6C ferryl for this case, and the 6C ferryl (**BI-IM3-6coord**) performing the hydrogen abstraction step slightly reduces the barrier to 20.7 kcal/mol (Figure S36). The ferric hydroxyl intermediate (**BI-IM4**) obtained from HAT undergoes rebound to complete the hydroxylation of L-Arg with an energy barrier of 9.1 kcal/mol, which is a lot higher than with conformation A. Thus, the calculations show the in-line bound 2OG cannot produce ethylene from either the A or B conformations of L-Arg. The two L-Arg conformations may undergo hydroxylation even though the process is not energetically efficient for conformation B.

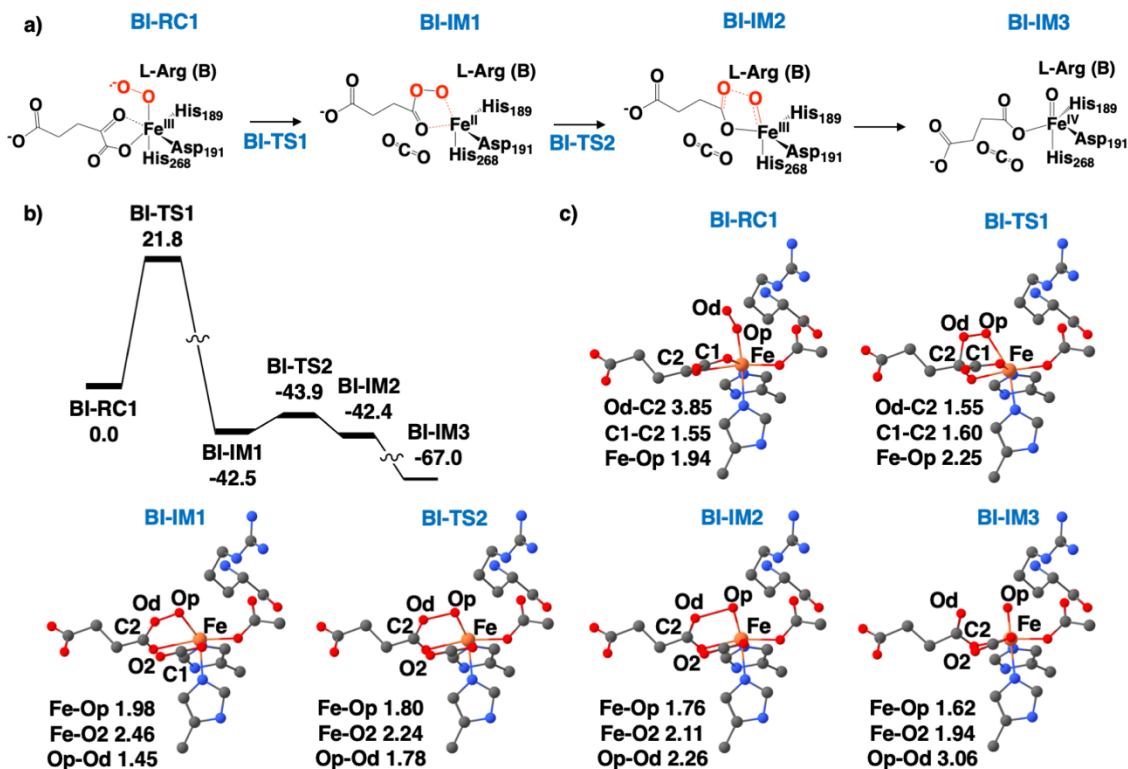


Figure 10. The O₂ activation mechanism of EFE with L-Arg in conformation B (**B**) and 2OG in the in-line (**I**) coordination mode. a) The proposed mechanistic pathway. b) Potential energy surface for the O₂ activation pathway. c) QM geometries of QM/MM optimized intermediates and transition states. The hydrogen atoms are hidden for clarity. Bond lengths are labeled in angstroms.

We next examined the effect of L-Arg conformation on the electronic structure in the systems with in-line 2OG binding to assess whether they correlate with the off-line bound 2OG structures. Because the O₂ and 2OG C1 carboxylate binding positions at the Fe center are swapped, the orbitals were relabeled accordingly. The z-axis is directed along with the Fe-O₂ bond, the y-axis is shown by Fe-O2 (where O is in the C2 keto group of 2OG), and the x-axis is along Fe-O1 (part of the C1 carboxylate of 2OG). The orbital analysis of the reactant with conformation A of L-Arg (**AI-RC1**) shows the unpaired β electron present in the $\pi^*_{\perp}$ oxygen and

an electronic configuration of $1d_{yz} 1d_{xy} 1d_{xz} 1d_{z^2} 1d_{x^2-y^2} 1\pi^*_{\perp} 1\pi^*_{\parallel}$. The Fe-O₂ bonding character originates from the singlet coupling between the Fe d_{xz} and the $\pi^*_{\perp}$ orbital of the O₂ (Figure S37), as seen in the off-line 2OG and L-Arg conformation A structure (**AO-RC1**). The unpaired β electron changes to the $\pi^*_{\parallel}$ orbital of dioxygen in the in-line 2OG and L-Arg conformation B structure (**BI-RC1**). Thus, the electronic configuration of **BI-RC1** is $1d_{xy} 1d_{xz} 1d_{yz} 1d_{x^2-y^2} 1d_{z^2} 1\pi^*_{\perp} 1\pi^*_{\parallel}$, and the Fe-O₂ bonding character for optimum overlap is obtained by a singlet coupling between the Fe d_{z^2} and the $\pi^*_{\parallel}$ orbital of the oxygen (Figure S37). Despite the favorable electronic structure being present in BI-RC1, the half-bond ferric-bicarbonate intermediate (**BO-IM2**) is not generated. The reason for this difference is that the in-line binding mode of 2OG lacks C1 carboxylate stabilization by R171, which is the second important factor (in addition to the electronic structure) that is responsible for the formation of **BO-IM1** (Figure S29). While the C1 carboxylate in the off-line coordination mode is stabilized by strong interaction with R171 at the active site, the C1 carboxylate associated with the in-line mode of 2OG is surrounded by a hydrophobic environment consisting of F283, A281, V196, and F250. The hydrophobic environment around the C1 carboxylate does not stabilize the C1 carboxylate in the active site to form the Op-C1 bond necessary to form the ferric-bicarbonate line intermediate. Thus, even though the switching of the oxygen orbital containing the unpaired β electron is maintained for the in-line 2OG structure with conformation A and B of L-Arg, the absence of C1 carboxylate stabilization hinders the formation of the necessary intermediates to produce ethylene.

4. Conclusions

EFE utilizes a non-heme Fe center to catalyze dual reactions using 2OG, dioxygen, and L-Arg. The hydroxylation of L-Arg follows the mechanism commonly observed in many other 2OG-dependent oxygenases. In contrast, the decomposition of 2OG by cleaving three C-C bonds to produce ethylene and three molecules of CO₂/bicarbonate is unique to EFE. We used a combined MD and QM/MM approach to explore the determinants for hydroxylation vs. ethylene production by EFE. The 1 μ s MD of the catalytically important EFE·Fe(III)·OO^{·-}·2OG_{off-line}·L-Arg intermediate, where 2OG is bound to Fe in an off-line mode, indicates L-Arg exists in two conformations, A and B. Conformation A has a shorter distance between C5 and the Fe center and was proposed to carry out hydroxylation, whereas conformation B with its considerably larger distance was suggested to have a role in ethylene formation. PCA and DCCA show β 4, β 5, β 11, and β 15 along with their loops have high flexibility with anti-correlated motions in the EFE·Fe(III)·OO^{·-}·2OG_{off-line}·L-Arg intermediate. Using the two L-Arg conformation structures from MD, we further explored the role of the substrate conformations and its associated second sphere changes on the mechanism of EFE.

The QM/MM reaction mechanism calculations show that when 2OG is bound in the off-line mode, and L-Arg is present in conformation A, the attack by Fe-coordinated O₂ on 2OG follows the standard O₂ activation mechanism with the formation of an off-line ferryl species which rearranges to in-line ferryl potentially through a ferryl-flip mechanism and is responsible for L-Arg hydroxylation. In contrast, the O₂ attack on 2OG using conformation B of L-Arg forms a half-bond ferric-bicarbonate intermediate that decomposes to ethylene, two CO₂, and a ferrous-bicarbonate species. To explore how this change in L-Arg conformation affects the mechanism, we analyzed the electronic structure of the Fe center for both reactions. Molecular orbital

analysis shows the dioxygen orbital with the unpaired β electron switches from the $\pi^*_{||}$ to the $\pi^*_{\perp}$ orbital upon changing in L-Arg conformation. The switching of the dioxygen orbital containing the unpaired β electron and the strong 2OG C1 carboxylate stabilization by Arg171 promotes the formation of the half-bond ferric-bicarbonate that decomposes to produce ethylene.

QM/MM calculations on the five-coordinated enzyme-substrate complex of EFE show, in contrast to similar 2OG oxygenases like PHF8, AlkBH2, etc., that prior to O₂ binding, the 2OG in EFE preferentially coordinates in the off-line mode. This preference for off-line coordination is likely to be due to the strong C1 carboxylate stabilization by R171 in the off-line mode and the unfavorable hydrophobic environment for the in-line coordinated 2OG. To explore whether O₂ binding in EFE leads to EFE·Fe(III)·OO[·]·2OG_{in-line}·L-Arg that produces ethylene, we performed 1 μ s MD of this species. PCA and DCCA show a loss in flexibility and correlated motions for regions containing residues 290-305, 210-240, and 80-93 in EFE·Fe(III)·OO[·]·2OG_{inline}·L-Arg. However, the change in the 2OG coordination mode does not affect the presence of two L-Arg conformations. Despite swapping the unpaired β electron containing oxygen orbitals, ethylene generation does not seem feasible for the enzyme with in-line bound 2OG because the C1 carboxylate stabilization is lost – a feature that is important for the formation of the half-bond ferric-bicarbonate intermediate.

The overall results identify a synergy between Fe(III)-OO[·] orbital changes associated with the L-Arg conformation and the 2OG coordination mode associated with C1-carboxylate stabilization; the synergy distinguishes the ethylene forming pathway from the hydroxylation mechanism in EFE. This finding indicates one may be able to improve ethylene yield in EFE by

stabilizing the off-line coordination mode of 2OG via modulating the motions in $\beta 4$, $\beta 5$, $\beta 11$, and $\beta 15$ and by stabilizing conformation B of L-Arg by balancing R171 and E84 interactions.

Supporting Information

The QM geometries of QM/MM optimized structures, spin densities, Mulliken charges, and supporting data on QM/MM and MD results are included in Supporting Information.

Acknowledgment

This research was supported by the NSF grants 1904215 and 1904295.

References

- (1) Fukuda, H.; Ogawa, T.; Tazaki, M.; Nagahama, K.; Fujiil, T.; Tanase, S.; Morino, Y. Two Reactions Are Simultaneously Catalyzed by a Single Enzyme: The Arginine-Dependent Simultaneous Formation of Two Products, Ethylene and Succinate, from 2-Oxoglutarate by an Enzyme from *Pseudomonas Syringae*. *Biochem. Biophys. Res. Commun.* 1992, 188 (2), 483–489. [https://doi.org/10.1016/0006-291X\(92\)91081-Z](https://doi.org/10.1016/0006-291X(92)91081-Z).
- (2) Martinez, S.; Hausinger, R. P. Biochemical and Spectroscopic Characterization of the Non-Heme Fe(II)- and 2-Oxoglutarate-Dependent Ethylene-Forming Enzyme from *Pseudomonas Syringae* Pv. *Phaseolicola* PK2. *Biochemistry* 2016, 55 (43), 5989–5999. <https://doi.org/10.1021/acs.biochem.6b00890>.
- (3) Schofield, C. J., Hausinger, R. P., Eds. 2-Oxoglutarate-Dependent Oxygenases; Royal Society of Chemistry: Cambridge, U.K., 2015. <https://doi.org/10.1039/9781782621959>.

- (4) Herr, C. Q.; Hausinger, R. P. Amazing Diversity in Biochemical Roles of Fe(II)/2-Oxoglutarate Oxygenases. *Trends in Biochemical Sciences*. Trends Biochem. Sci. 2018, 43(7), 517–532. <https://doi.org/10.1016/j.tibs.2018.04.002>.
- (5) Hausinger, R. P. Fe(II)/ α -Ketoglutarate-Dependent Hydroxylases and Related Enzymes. *Critical Reviews in Biochemistry and Molecular Biology*. Crit. Rev. Biochem. Mol. Biol. 2004, 39(1), 21–68. <https://doi.org/10.1080/10409230490440541>.
- (6) Hausinger, R. P. Biochemical Diversity of 2-Oxoglutarate-Dependent Oxygenases. In *2-Oxoglutarate-Dependent Oxygenases*. Schofield, C. J., Hausinger, R. P., Eds. Royal Society of Chemistry: Cambridge, U.K., 2015; 1–58. <https://doi.org/10.1039/9781782621959-00001>.
- (7) Martinez, S.; Fellner, M.; Herr, C. Q.; Ritchie, A.; Hu, J.; Hausinger, R. P. Structures and Mechanisms of the Non-Heme Fe(II)- and 2-Oxoglutarate-Dependent Ethylene-Forming Enzyme: Substrate Binding Creates a Twist. *J. Am. Chem. Soc.* 2017, 139 (34), 11980–11988. <https://doi.org/10.1021/jacs.7b06186>.
- (8) Zhang, Z.; Smart, T. J.; Choi, H.; Hardy, F.; Lohans, C. T.; Abboud, M. I.; Richardson, M. S. W.; Paton, R. S.; McDonough, M. A.; Schofield, C. J. Structural and Stereoelectronic Insights into Oxygenase-Catalyzed Formation of Ethylene from 2-Oxoglutarate. *Proc. Natl. Acad. Sci. U. S. A.* 2017, 114 (18), 4667–4672. <https://doi.org/10.1073/pnas.1617760114>.
- (9) Xue, J.; Lu, J.; Lai, W. Mechanistic Insights into a Non-Heme 2-Oxoglutarate-Dependent Ethylene-Forming Enzyme: Selectivity of Ethylene-Formation: Versus l-Arg Hydroxylation. *Phys. Chem. Chem. Phys.* 2019, 21 (19), 9957–9968. <https://doi.org/10.1039/c9cp00794f>.
- (10) Zhang, Z.; Ren, J.; Harlos, K.; McKinnon, C. H.; Clifton, I. J.; Schofield, C. J. Crystal Structure of a Clavamate Synthase-Fe(II)-2-Oxoglutarate-Substrate-NO Complex: Evidence for Metal

Centred Rearrangements. *FEBS Lett.* 2002, 517 (1–3), 7–12. [https://doi.org/10.1016/S0014-5793\(02\)02520-6](https://doi.org/10.1016/S0014-5793(02)02520-6).

(11) Dennington, R.; Keith, T.; Millam, J. GaussView, Version 5. Semichem Inc., Shawnee Mission, KS, 2016.

(12) Gordon, J. C.; Myers, J. B.; Foltz, T.; Shoja, V.; Heath, L. S.; Onufriev, A. H++: A Server for Estimating PKas and Adding Missing Hydrogens to Macromolecules. *Nucleic Acids Res.* 2005, 33 (SUPPL. 2), W368-71. <https://doi.org/10.1093/nar/gki464>.

(13) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *J. Mol. Graph. Model.* 2006, 25 (2), 247–260. <https://doi.org/10.1016/j.jmglm.2005.12.005>.

(14) Li, P.; Merz, K. M. MCPB.Py: A Python Based Metal Center Parameter Builder. *J. Chem. Inf. Model.* 2016, 56 (4), 599–604. <https://doi.org/10.1021/acs.jcim.5b00674>.

(15) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* 2015, 11 (8), 3696–3713. <https://doi.org/10.1021/acs.jctc.5b00255>.

(16) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* 1983, 79 (2), 926–935. <https://doi.org/10.1063/1.445869>.

(17) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N-log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* 1993, 98 (12), 10089–10092. <https://doi.org/10.1063/1.464397>.

- (18) Davidchack, R. L.; Handel, R.; Tret'yakov, M. V. Langevin Thermostat for Rigid Body Dynamics. *J. Chem. Phys.* 2009, 130 (23), 234101. <https://doi.org/10.1063/1.3149788>.
- (19) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. . Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* 1977, 23 (3), 327–341. [https://doi.org/10.1016/0021-9991\(77\)90098-5](https://doi.org/10.1016/0021-9991(77)90098-5).
- (20) Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* 1984, 81 (8), 3684–3690. <https://doi.org/10.1063/1.448118>.
- (21) Case, D. A.; Betz, R. M.; Cerutti, D. S.; Cheatham, T. E.; Daeden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Homeyer, N.; Izadi, S.; Janowski, P.; Kaus, J.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Luchko, T.; Luo, R.; Madej, B.; Mermelstein, D.; Merz, K. M.; Monard, G.; Nguyen, H.; Nguyen, H. T.; Omelyan, I.; Onufriev, A.; Roe, D. R.; Roitber, A.; Sagui, C.; Simmerling, C. L.; Botello-Smith, W. M.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, L.; Xiao, L.; Kollman, P. A. AMBER 2016. University of California: San Francisco 2016.
- (22) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* 2013, 9 (7), 3084–3095. <https://doi.org/10.1021/ct400341p>.
- (23) Grant, B. J.; Rodrigues, A. P. C.; ElSawy, K. M.; McCammon, J. A.; Caves, L. S. D. Bio3d: An R Package for the Comparative Analysis of Protein Structures. *Bioinformatics* 2006, 22 (21), 2695–2696. <https://doi.org/10.1093/bioinformatics/btl461>.

- (24) Metz, S.; Kästner, J.; Sokol, A. A.; Keal, T. W.; Sherwood, P. ChemShell-a Modular Software Package for QM/MM Simulations. *WIREs Comput. Mol. Sci.* 2014, 4 (2), 101–110. <https://doi.org/10.1002/wcms.1163>.
- (25) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. Electronic Structure Calculations on Workstation Computers: The Program System Turbomole. *Chem. Phys. Lett.* 1989, 162 (3), 165–169. [https://doi.org/10.1016/0009-2614\(89\)85118-8](https://doi.org/10.1016/0009-2614(89)85118-8).
- (26) Smith, W.; Yong, C. W.; Rodger, P. M. DL_POLY: Application to Molecular Simulation. *Mol. Simul.* 2002, 28 (5), 385–471. <https://doi.org/10.1080/08927020290018769>.
- (27) Song, X.; Lu, J.; Lai, W. Mechanistic Insights into Dioxygen Activation, Oxygen Atom Exchange and Substrate Epoxidation by AsqJ Dioxygenase from Quantum Mechanical/Molecular Mechanical Calculations. *Phys. Chem. Chem. Phys.* 2017, 19 (30), 20188–20197. <https://doi.org/10.1039/C7CP02687K>.
- (28) Ye, S.; Riplinger, C.; Hansen, A.; Krebs, C.; Bollinger, J. M.; Neese, F. Electronic Structure Analysis of the Oxygen-Activation Mechanism by Fe II- and α -Ketoglutarate (AKG)-Dependent Dioxygenases. *Chem. - A Eur. J.* 2012, 18 (21), 6555–6567. <https://doi.org/10.1002/chem.201102829>.
- (29) Wang, B.; Usharani, D.; Li, C.; Shaik, S. Theory Uncovers an Unusual Mechanism of DNA Repair of a Lesioned Adenine by AlkB Enzymes. *J. Am. Chem. Soc.* 2014, 136 (39), 13895–13901. <https://doi.org/10.1021/ja507934g>.
- (30) Wójcik, A.; Radoń, M.; Borowski, T. Mechanism of O₂ Activation by α -Ketoglutarate Dependent Oxygenases Revisited. A Quantum Chemical Study. *J. Phys. Chem. A* 2016, 120 (8), 1261–1274. <https://doi.org/10.1021/acs.jpca.5b12311>.

- (31) Sherwood, P.; De Vries, A. H.; Guest, M. F.; Schreckenbach, G.; Catlow, C. R. A.; French, S. A.; Sokol, A. A.; Bromley, S. T.; Thiel, W.; Turner, A. J.; Billeter, S.; Terstegen, F.; Thiel, S.; Kendrick, J.; Rogers, S. C.; Casci, J.; Watson, M.; King, F.; Karlsen, E.; Sjøvoll, M.; Fahmi, A.; Schäfer, A.; Lennartz, C. QUASI: A General Purpose Implementation of the QM/MM Approach and Its Application to Problems in Catalysis. *J. Mol. Struct. THEOCHEM* 2003, 632 (1–3), 1–28. [https://doi.org/10.1016/s0166-1280\(03\)00285-9](https://doi.org/10.1016/s0166-1280(03)00285-9).
- (32) Bakowies, D.; Thiel, W. Hybrid Models for Combined Quantum Mechanical and Molecular Mechanical Approaches. *J. Phys. Chem.* 1996, 100 (25), 10580–10594. <https://doi.org/10.1021/jp9536514>.
- (33) Schäfer, A.; Horn, H.; Ahlrichs, R. Fully Optimized Contracted Gaussian Basis Sets for Atoms Li to Kr. *J. Chem. Phys.* 1992, 97 (4), 2571–2577. <https://doi.org/10.1063/1.463096>.
- (34) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem. Int. Ed.* 2009, 48, 1198–1229. <https://doi.org/10.1002/anie.200802019>.
- (35) Billeter, S. R.; Turner, A. J.; Thiel, W. Linear Scaling Geometry Optimisation and Transition State Search in Hybrid Delocalised Internal Coordinates. *Phys. Chem. Chem. Phys.* 2000, 2 (10), 2177–2186. <https://doi.org/10.1039/a909486e>.
- (36) Graham, S. E.; Syeda, F.; Cisneros, G. A. Computational Prediction of Residues Involved in Fidelity Checking for DNA Synthesis in DNA Polymerase I. *Biochemistry* 2012, 51 (12), 2569–2578. <https://doi.org/10.1021/bi201856m>.
- (37) Andrés Cisneros, G.; Perera, L.; Sehaaper, R. M.; Pedersen, L. C.; London, R. E.; Pedersen, L. G.; Darden, T. A. Reaction Mechanism of the ϵ Subunit of E. Coli DNA Polymerase III: Insights

into Active Site Metal Coordination and Catalytically Significant Residues. *J. Am. Chem. Soc.* 2009, 131 (4), 1550–1556. <https://doi.org/10.1021/ja8082818>.

(38) Krebs, C.; Fujimori, D. G.; Walsh, C. T.; Bollinger, J. M. Non-Heme Fe(IV)-Oxo Intermediates. *Acc. Chem. Res.* 2007, 40(7), 484–492. <https://doi.org/10.1021/ar700066p>.

(39) Que, L. The Road to Non-Heme Oxoferryls and Beyond. *Acc. Chem. Res.* 2007, 40(7), 493–500. <https://doi.org/10.1021/ar700024g>.

(40) Hong, S.; Sutherlin, K. D.; Park, J.; Kwon, E.; Siegler, M. A.; Solomon, E. I.; Nam, W. Crystallographic and Spectroscopic Characterization and Reactivities of a Mononuclear Non-Haem Iron(III)-Superoxo Complex. *Nat. Commun.* 2014, 5 (1), 5440. <https://doi.org/10.1038/ncomms6440>.

(41) Borowski, T.; Bassan, A.; Siegbahn, P. E. M. Mechanism of Dioxygen Activation in 2-Oxoglutarate-Dependent Enzymes: A Hybrid DFT Study. *Chem. - A Eur. J.* 2004, 10 (4), 1031–1041. <https://doi.org/10.1002/chem.200305306>.

(42) Chaturvedi, S. S.; Ramanan, R.; Lehnert, N.; Schofield, C. J.; Karabancheva-Christova, T. G.; Christov, C. Z. Catalysis by the Non-Heme Iron(II) Histone Demethylase PHF8 Involves Iron Center Rearrangement and Conformational Modulation of Substrate Orientation. *ACS Catal.* 2020, 10 (2), 1195–1209. <https://doi.org/10.1021/acscatal.9b04907>.

(43) Waheed, S. O.; Ramanan, R.; Chaturvedi, S. S.; Lehnert, N.; Schofield, C. J.; Christov, C. Z.; Karabancheva-Christova, T. G. Role of Structural Dynamics in Selectivity and Mechanism of Non-Heme Fe(II) and 2-Oxoglutarate-Dependent Oxygenases Involved in DNA Repair. *ACS Cent. Sci.* 2020, 6 (5), 795–814. <https://doi.org/10.1021/acscentsci.0c00312>.

- (44) Quesne, M. G.; Latifi, R.; Gonzalez-Ovalle, L. E.; Kumar, D.; De Visser, S. P. Quantum Mechanics/Molecular Mechanics Study on the Oxygen Binding and Substrate Hydroxylation Step in AlkB Repair Enzymes. *Chem. - A Eur. J.* 2014, 20 (2), 435–446. <https://doi.org/10.1002/chem.201303282>.
- (45) Wang, B.; Cao, Z.; Sharon, D. A.; Shaik, S. Computations Reveal a Rich Mechanistic Variation of Demethylation of N-Methylated DNA/RNA Nucleotides by FTO. *ACS Catal.* 2015, 5 (12), 7077–7090. <https://doi.org/10.1021/acscatal.5b01867>.
- (46) Torabifard, H.; Cisneros, G. A. Insight into Wild-Type and T1372E TET2-Mediated 5hmC Oxidation Using Ab Initio QM/MM Calculations. *Chem. Sci.* 2018, 9 (44), 8433–8445. <https://doi.org/10.1039/C8SC02961J>.
- (47) Fang, D.; Cisneros, G. A. Alternative Pathway for the Reaction Catalyzed by DNA Dealkylase AlkB from Ab Initio QM/MM Calculations. *J. Chem. Theory Comput.* 2014, 10 (11), 5136–5148. <https://doi.org/10.1021/ct500572t>.
- (48) Srnec, M.; Iyer, S. R.; Dassama, L. M. K.; Park, K.; Wong, S. D.; Sutherlin, K. D.; Yoda, Y.; Kobayashi, Y.; Kurokuzu, M.; Saito, M.; Seto, M.; Krebs, C.; Bollinger Jr., J. M.; Solomon, E. I. Nuclear Resonance Vibrational Spectroscopic Definition of the Facial Triad $\text{Fe}^{\text{IV}}=\text{O}$ Intermediate in Taurine Dioxygenase: Evaluation of Structural Contributions to Hydrogen Atom Abstraction. *J. Am. Chem. Soc.* 2020, 142, 44, 18886–18896.
- (49) Li, M.; Martinez, S.; Hausinger, R. P.; Emerson, J. P. Thermodynamics of Iron(II) and Substrate Binding to the Ethylene-Forming Enzyme. *Biochemistry* 2018, 57 (39), 5696–5705. <https://doi.org/10.1021/acs.biochem.8b00730>.

- (50) Chen, H.; Cho, K. Bin; Lai, W.; Nam, W.; Shaik, S. Dioxygen Activation by a Non-Heme Iron(II) Complex: Theoretical Study toward Understanding Ferric-Superoxo Complexes. *J. Chem. Theory Comput.* 2012, 8 (3), 915–926. <https://doi.org/10.1021/ct300015y>.
- (51) Horton, J. R.; Upadhyay, A. K.; Qi, H. H.; Zhang, X.; Shi, Y.; Cheng, X. Enzymatic and Structural Insights for Substrate Specificity of a Family of Jumonji Histone Lysine Demethylases. *Nat. Struct. Mol. Biol.* 2010, 17 (1), 38–44. <https://doi.org/10.1038/nsmb.1753>.
- (52) Yu, L.; Wang, Y.; Huang, S.; Wang, J.; Deng, Z.; Zhang, Q.; Wu, W.; Zhang, X.; Liu, Z.; Gong, W.; Chen, Z. Structural Insights into a Novel Histone Demethylase PHF8. *Cell Res.* 2010, 20 (2), 166–173. <https://doi.org/10.1038/cr.2010.8>.
- (53) Ng, S. S.; Kavanagh, K. L.; McDonough, M. A.; Butler, D.; Pilka, E. S.; Lienard, B. M. R.; Bray, J. E.; Savitsky, P.; Gileadi, O.; Von Delft, F.; Rose, N. R.; Offer, J.; Scheinost, J. C.; Borowski, T.; Sundstrom, M.; Schofield, C. J.; Oppermann, U. Crystal Structures of Histone Demethylase JMJD2A Reveal Basis for Substrate Specificity. *Nature* 2007, 448 (7149), 87–91. <https://doi.org/10.1038/nature05971>.
- (54) Yu, B.; Hunt, J. F. Enzymological and Structural Studies of the Mechanism of Promiscuous Substrate Recognition by the Oxidative DNA Repair Enzyme AlkB. *Proc. Natl. Acad. Sci. U. S. A.* 2009, 106 (34), 14315–14320. <https://doi.org/10.1073/pnas.0812938106>.
- (55) Yi, C.; Chen, B.; Qi, B.; Zhang, W.; Jia, G.; Zhang, L.; Li, C. J.; Dinner, A. R.; Yang, C. G.; He, C. Duplex Interrogation by a Direct DNA Repair Protein in Search of Base Damage. *Nat. Struct. Mol. Biol.* 2012, 19 (7), 671–676. <https://doi.org/10.1038/nsmb.2320>.
- (56) Cheng, Z.; Cheung, P.; Kuo, A. J.; Yukl, E. T.; Wilmot, C. M.; Gozani, O.; Patel, D. J. A Molecular Threading Mechanism Underlies Jumonji Lysine Demethylase KDM2A Regulation of

Methylated H3K36. Genes Dev. 2014, 28 (16), 1758–1771.

<https://doi.org/10.1101/gad.246561.114>.

TOC:

