



**Michigan
Technological
University**

Michigan Technological University
Digital Commons @ Michigan Tech

Michigan Tech Publications, Part 2

4-5-2024

Structural, Spectroscopic, and Computational Insights from Canavanine-Bound and Two Catalytically Compromised Variants of the Ethylene-Forming Enzyme

Shramana Chatterjee

Department of Microbiology, Genetics, & Immunology

Matthias Fellner

Department of Biochemistry and Molecular Biology

Joel A. Rankin

Department of Microbiology, Genetics, & Immunology

Midhun George Thomas

Michigan Technological University, mgeorget@mtu.edu

Simahudeen Bathir Jaber Sathik Rifayee

Michigan Technological University, sbjabers@mtu.edu

Follow this and additional works at: <https://digitalcommons.mtu.edu/michigantech-p2>


 Part of the [Chemistry Commons](#)

Recommended Citation

Chatterjee, S., Fellner, M., Rankin, J., Thomas, M., Rifayee, S., Christov, C., Hu, J., & Hausinger, R. (2024). Structural, Spectroscopic, and Computational Insights from Canavanine-Bound and Two Catalytically Compromised Variants of the Ethylene-Forming Enzyme. *Biochemistry*, 63(8), 1038-1050. <http://doi.org/10.1021/acs.biochem.4c00031>

Retrieved from: <https://digitalcommons.mtu.edu/michigantech-p2/678>

Follow this and additional works at: <https://digitalcommons.mtu.edu/michigantech-p2>

 Part of the [Chemistry Commons](#)

Authors

Shramana Chatterjee, Matthias Fellner, Joel A. Rankin, Midhun George Thomas, Simahudeen Bathir Jaber Sathik Rifayee, Christo Christov, Jian Hu, and Robert P. Hausinger

Structural, Spectroscopic, and Computational Insights from Canavanine-Bound and Two Catalytically Compromised Variants of the Ethylene-Forming Enzyme

Published as part of *Biochemistry virtual special issue* “A Tribute to Christopher T. Walsh”.

Shramana Chatterjee,[∇] Matthias Fellner,[∇] Joel A. Rankin, Midhun G. Thomas, Simahudeen Bathir J S Rifayee, Christo Z. Christov,^{*} Jian Hu,^{*} and Robert P. Hausinger^{*}



Cite This: *Biochemistry* 2024, 63, 1038–1050



Read Online

ACCESS |



Metrics & More

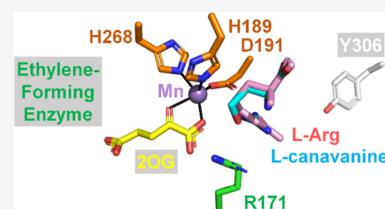


Article Recommendations



Supporting Information

ABSTRACT: The ethylene-forming enzyme (EFE) is an Fe(II), 2-oxoglutarate (2OG), and L-arginine (L-Arg)-dependent oxygenase that either forms ethylene and three CO₂/bicarbonate from 2OG or couples the decarboxylation of 2OG to C5 hydroxylation of L-Arg. L-Arg binds with C5 toward the metal center, causing 2OG to change from monodentate to chelate metal interaction and OD1 to OD2 switch of D191 metal coordination. We applied anaerobic UV–visible spectroscopy, X-ray crystallography, and computational approaches to three EFE systems with high-resolution structures. The ineffective L-Arg analogue L-canavanine binds to the EFE with O5 pointing away from the metal center while promoting chelate formation by 2OG but fails to switch the D191 metal coordination from OD1 to OD2. Substituting alanine for R171 that interacts with 2OG and L-Arg inactivates the protein, prevents metal chelation by 2OG, and weakens L-Arg binding. The R171A EFE had electron density at the 2OG binding site that was identified by mass spectrometry as benzoic acid. The substitution by alanine of Y306 in the EFE, a residue 12 Å away from the catalytic metal center, generates an interior cavity that leads to multiple local and distal structural changes that reduce L-Arg binding and significantly reduce the enzyme activity. Flexibility analyses revealed correlated and anticorrelated motions in each system, with important distinctions from the wild-type enzyme. In combination, the results are congruent with the currently proposed enzyme mechanism, reinforce the importance of metal coordination by OD2 of D191, and highlight the importance of the second coordination sphere and longer range interactions in promoting EFE activity.

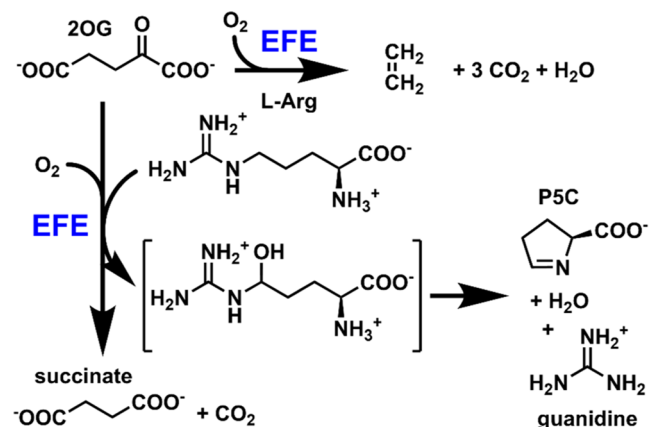


INTRODUCTION

Some bacteria and fungi possess an Fe(II), 2-oxoglutarate (2OG), and L-arginine (L-Arg)-dependent oxygenase known as the ethylene-forming enzyme (EFE)¹ that (1) cleaves three C–C bonds of 2OG to form ethylene and three molecules of CO₂/bicarbonate and (2) catalyzes the oxidative decarboxylation of 2OG to form CO₂ and succinate coupled to the C5 hydroxylation of L-Arg, which decomposes to guanidine and L-Δ¹-pyrroline-5-carboxylate (P5C) (Scheme 1).^{2,3} Computational and biochemical studies suggest that the bifurcated pathway splits at the level of an Fe(III)·superoxo complex, with the major reaction, ethylene generation, requiring dioxygen insertion into the C1–C2 bond of 2OG, formation of a propionyl-3-yl radical and an Fe(III)-bound carbonate, and their coupling followed by fragmentation to yield the target product.^{4–7} In contrast, the minor reaction of guanidine/P5C formation is associated with the well-known hydroxylation chemistry of 2OG-dependent oxygenases that involves a ferryl intermediate.^{1,8}

The best characterized EFE is that from the *Pseudomonas savastanoi* (formerly *Pseudomonas syringae*) pv *phaseolicola* strain PK2, for which several crystal structures have been reported.^{8–10} This enzyme binds mononuclear Fe(II) using

Scheme 1. Major Reactions of the EFE

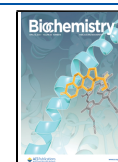


Received: January 19, 2024

Revised: March 16, 2024

Accepted: March 22, 2024

Published: April 5, 2024



three side chains (H189, D191, and H268) with three water molecules completing the six-coordinate geometry. Unlike the situation reported for several other 2OG-dependent oxygenases where 2OG displaces two coordinating waters and chelates the Fe(II) using both its C1 carboxylate and C2 keto group, 2OG binds to Fe(II) of the EFE predominantly in a monodentate mode with just its C1 carboxylate.^{3,9} L-Arg addition induces 2OG to adopt the conventional chelate binding mode, while also causing dissociation of the remaining coordinated water molecules and formation of a twisted peptide bond between the metal ligand D191 and the neighboring residue Y192.⁹ The resulting open coordination site for oxygen binding is directed away from L-Arg (i.e., an “off-line” binding mode).

Here, we extend our understanding of EFE catalysis by carrying out biochemical, structural, and computational studies of the enzyme in complex with a substrate analogue and using two enzyme variants. First, we consider the binding of L-canavanine, a chemical analogue of L-Arg in which the C5 methylene group is replaced by an oxygen atom. The EFE cannot transform this analogue into guanidine and P5C, but we also showed that the L-canavanine-bound enzyme fails to induce ethylene generation.³ To understand this lack of reactivity, we used anaerobic ultraviolet–visible (UV–vis) spectroscopy to examine the effect of L-canavanine on the metal-to-ligand charge-transfer (MLCT) transitions of EFE·Fe(II)·2OG and determined the structure of the EFE·Mn(II)·2OG·L-canavanine complex. Second, we examined the R171A variant of the EFE,⁹ a substitution of a second-sphere residue (i.e., just beyond the immediate metal ligands) that eliminates ethylene-forming activity likely by destabilizing 2OG binding. We carried out MLCT analysis and elucidated the R171A EFE crystal structure to provide evidence for significant changes in the 2OG binding pocket that explain the abolishment of activity, and we used electrospray ionization–mass spectrometry (ESI-MS) to identify an unanticipated ligand at the 2OG binding site. Third, we focused on the nearly inactive Y306A variant of the EFE for which the substitution occurs distant from the active site.⁹ We noted distinct behavior when investigating the MLCT transitions associated with Y306A EFE·Fe(II)·2OG and obtained insights from the structure of the Y306A EFE·Mn(II)·2OG complex that was crystallized in the presence of unbound L-Arg. Crystallographic structures can provide only partial insights into flexibility¹¹ even though enzymes exhibit a broad range of complex motions that can be functionally important.^{12–14} Many examples illustrate the role of the second coordination sphere in such long-range interactions,^{15–20} and the flexibility and stability of different parts of the enzyme can be crucial to understand changes in the catalytic reactivity.^{14,21–25} Thus, we complemented the biochemical and structural results by applying molecular dynamics (MD) simulations^{26–29} to explore the conformational changes and substrate or substrate analogue interactions for all three EFE systems.

MATERIALS AND METHODS

Chemicals. L-Canavanine, tris(2-carboxyethyl)phosphine (TCEP), L-Arg, sodium phosphate monobasic, and imidazole were acquired from Sigma (St. Louis, Missouri). Kanamycin, isopropyl β -D-1-thiogalactopyranoside (IPTG), 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), and Ni-NTA agarose beads were purchased from GoldBio (St. Louis, Missouri). 2OG was purchased from Fluka. EDTA

was purchased from Invitrogen. 50% PEG 3350, PEG 400, PEG MME 550, and 50% ethylene glycol were purchased from Rigaku or Hampton. All other chemicals were reagent grade or better.

Enzyme Purification. Wild-type (WT) *P. savastanoi* strain PK2 EFE (Uniprot entry P32021) and its R171A and Y306A variants were obtained as previously described.⁹ Briefly, the genes encoding N-terminal His₆-tagged proteins were expressed in recombinant *Escherichia coli* cells that were grown in Luria broth or Terrific broth and induced with IPTG; the cells were disrupted by use of a French pressure cell or sonication, the proteins were purified by use of nickel-nitrilotriacetic acid (NTA) resin, the tag was cleaved by using His₇-tagged tobacco etch virus (TEV) protease, and the proteins were rechromatographed on the nickel-NTA column to provide the final samples.

Enzyme Assays. The production of ethylene was monitored by gas chromatography, and the formation of P5C was assessed by reaction with 2-aminobenzaldehyde as previously described.³

UV–vis Spectroscopy. Anaerobic UV–vis spectroscopy was used to probe the MLCT transitions associated with 2OG coordinated to the EFE Fe(II) center via the chelate mode.^{3,30} All solutions were prepared anaerobically in 25 mM HEPES buffer that was adjusted to pH 8.0. Spectra were recorded at room temperature using a Shimadzu UV-2600 spectrophotometer.

Crystallization. The WT EFE (0.5 μ L of 64 mg of protein mL⁻¹) was incubated for 1 h on ice with 25 mM HEPES (pH 8.0) buffer containing 1 mM MnCl₂, 1 mM TCEP, 3 mM L-canavanine, and 2 mM 2OG before setting up crystallization drops with 0.5 μ L of reservoir solution at 4 °C. The sitting drop reservoir of 200 μ L contained 25% w/v poly(ethylene glycol) 3350, 0.1 M Bis-Tris (pH 6.5), and 0.2 M sodium chloride. The resulting crystals were soaked for about 1 min in 25% w/v poly(ethylene glycol) 400 and 75% reservoir solution before freezing.

The R171A variant of the EFE in 25 mM HEPES (pH 8.0) buffer containing 1 mM TCEP was crystallized in a similar manner with a sitting drop reservoir containing 20% w/v poly(ethylene glycol) 3350 and 0.2 M sodium citrate tribasic dihydrate. The crystals were briefly soaked in 25% poly(ethylene glycol) 550 monomethyl ether and 75% reservoir solution before freezing.

The Y306A variant of the EFE (0.5 μ L of 64 mg mL⁻¹) was incubated in 25 mM HEPES (pH 8.0) buffer containing 1 mM MnCl₂, 1 mM TCEP, 3 mM L-Arg, and 3 mM 2OG before mixing with 0.5 μ L of reservoir solution. The sitting drop reservoir of 200 μ L contained 25% w/v poly(ethylene glycol) 3350, 0.1 M Bis-Tris (pH 6.5), and 0.2 M sodium chloride. The resulting crystals were soaked for 5 min in 25% w/v ethylene glycol and 75% reservoir solution before freezing.

Structure Determinations. X-ray diffraction data were collected at the Advanced Photon Source LS-CAT beamline 21-ID-D. For details, see Table S1. Datasets were indexed and integrated with HKL-2000 or iMosflm,³¹ and merging and scaling were done using Aimless.³² Molecular replacement and refinement were done in Phenix³³ with model building in COOT.³⁴ Datasets were uploaded to the Protein Data Bank (PDB) with IDs of 6CBA for EFE·Mn(II)·2OG·L-canavanine, 6CF3 for the Y306A variant of EFE·Mn(II)·2OG, and 8UC2 for the R171A variant of EFE·Ni(II)·benzoate.

ESI-MS. R171A EFE crystals were recovered from the drop, cleaned, and dissolved in 10 mM HEPES, pH 8.0. A fourfold volume of acetonitrile was added to precipitate the protein, which was removed by centrifugation, and the samples were dried and dissolved in water containing 0.1% formic acid. Standards (benzoic acid, nicotinic acid, and salicylic acid) were similarly prepared from ACS reagents in water containing 0.1% formic acid. The samples and standards were injected onto a cyano-chemistry HPLC column that was equilibrated in 0.1% formic acid and eluted with an increasing gradient of acetonitrile. The fractions were analyzed by ESI-MS using a XEVO G2-XS instrument in negative ionization mode.

System Preparation for MD Simulations. As starting structures for the EFE-Fe(III)·OO^{•-} complexes of R171A and Y306A variants, the crystal structure of the WT EFE complex with Mn(II), 2OG, and L-Arg (PDB: 5V2Y) was used.⁹ Mn(II) was substituted with Fe(II), and dioxygen was coordinated to the Fe(II) center. The crystal structure of the EFE variant Y306A (PDB: 6CF3) with Mn(II) and 2OG was used as a starting structure for the Y306A EFE·Fe(II)·2OG complex. The structure was modified by replacing the Mn(II) with Fe(II) using GaussView 5.0.³⁵ The R171A and Y306A substitutions were generated using the Leap module of AMBER 18.³⁶ The starting structures for the two EFE variants were generated similarly to that of the WT EFE.⁵ Hydrogen atoms were added according to the protonation states using Amber routines.³⁶ The Antechamber module containing the GAFF tool was used to develop topology files of the nonstandard components such as 2OG and O₂.^{37,38} Parameters for Fe(II) (high spin $S = 2$ and $M = 5$), coordinating 2OG, H189, D191, H268, and WAT503 and parameters for the Fe(III) center (high spin $S = 2$, $M = 5$ after ferromagnetic coupling between the iron and dioxygen) coordinating 2OG, H189, D191, H268, and dioxygen were generated using Metal Center Parameter Builder (MCPB v3.0).³⁹ The rest of the protein was modeled using the FF14SB force field.⁴⁰ The system was solvated with a 10 Å water box from the farthest point of the residue, and counterions (Na⁺) were introduced to neutralize the system.

For the simulation of the EFE containing L-canavanine, we used the crystal structure of the WT EFE containing Mn(II), 2OG, and L-Arg (PDB: 5V2Y)⁹ and replaced the L-Arg coordinates with those for L-canavanine obtained from the EFE·Mn(II)·2OG·L-canavanine crystal structure (PDB: 6CBA). The two crystal structures are highly superimposable; therefore, the WT structure with L-canavanine was used for the simulations. The remaining steps of system preparation for the EFE·Fe(III)·OO^{•-}·L-canavanine complex followed the same procedure as used for the two variants.

MD Simulations. The protonation states were assigned considering the pK_a of the residues at neutral pH with the tLeAP module of AMBER.³⁶ In particular, E84 is deprotonated, Arg171 is protonated, and the guanidino groups of L-canavanine and L-Arg are protonated, while their carboxyl groups are deprotonated. The procedure for MD simulations was the same as that used for the WT EFE.⁵ Initially, the modeled systems were subjected to minimizations in two steps. The first minimization was performed with a restraint weight of 100 kcal mol⁻¹ on the solute molecule with 5000 steps of the steepest descent and 5000 steps of the conjugate gradient. The second minimization was performed similarly on the full system without any restraint. The minimizations were performed using the CPU version of AMBER 18 (SAND-

ER).³⁶ Minimizations were followed by the gradual heating of the system at the NVT ensemble from 0 to 300 K using a Langevin thermostat for 250 ps.⁴¹ The collision frequency of the Langevin thermostat was set to 1 ps⁻¹. During the heating, the solute molecules were restrained using a harmonic potential of 50 kcal mol⁻¹. After heating, 1 ns simulations with weak restraint on the protein were run to achieve uniform density. At a constant pressure of 1 bar, MD was equilibrated for all systems for 3 ns. Constant pressure was achieved by employing a Berendsen barostat. From the equilibrated structure, 1 μs of MD simulation was performed with a 2 fs time step. The particle mesh Ewald (PME) method⁴² was employed to compute long-range interactions within a 10 Å cutoff in all simulations. The simulations were conducted utilizing the GPU-accelerated version of AMBER-PMEMD. Hydrogen bonding interactions were analyzed with CPPTRAJ.⁴³ Principal component analysis (PCA), dynamic cross-correlation analysis (DCCA), and differential DCCA (d-DCCA) (for analyzing the differences in correlated motions between two systems) were conducted on the Cα carbons of the protein residues using the last 500 ns trajectories, and these analyses were executed with the Bio3D module implemented in the R software package.⁴⁴ Binding energies were calculated using molecular mechanics, the generalized Born model, and the solvent accessibility method (MMGBSA).^{45,46}

RESULTS AND DISCUSSION

Experimental Studies on the L-Canavanine-Bound EFE. As previously reported, the addition of 2OG to an anaerobic solution of EFE·Fe(II) generates a weak difference spectrum with a maximum absorption near 525 nm that was attributed to a mixture of monodentate and bidentate binding of 2OG to the metal, where only the chelated species provides the broad MLCT transitions.³ Subsequent addition of L-Arg shifts the equilibrium to favor the chelate binding mode that is essential for activity, as noted by the increased absorbance in this region.³ Very much like the situation for added L-Arg, the addition of L-canavanine to the EFE·Fe(II)·2OG sample leads to an intensification of the calculated difference absorption from ~240 to ~450 M⁻¹ cm⁻¹ (Figure 1). These results are consistent with the binding of L-canavanine leading to a shift from monodentate 2OG binding to Fe(II) to the chelate

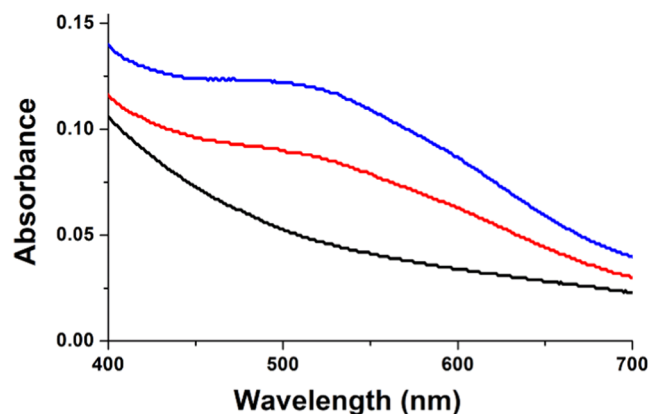


Figure 1. Anaerobic UV-vis spectra of EFE·Fe(II) (black), EFE·Fe(II)·2OG (red), and EFE·Fe(II)·2OG·L-canavanine (blue) complexes. The anaerobic samples contained 242 μM EFE, 1 mM Fe(II), and (when present) 2.5 mM 2OG and 2.5 mM L-canavanine. The spectra were adjusted for dilution.

binding mode. In this case, the chelate binding of 2OG is insufficient for promoting the ethylene-forming activity of the enzyme.

To investigate why *L*-canavanine does not promote ethylene formation despite its ability to bind much like *L*-Arg according to the anaerobic spectroscopy evidence, we examined the structure of EFE·Mn(II)·2OG with bound *L*-canavanine. We previously reported that six *L*-Arg analogues significantly reduced the ethylene-forming activity of the EFE when substituted for *L*-Arg.³ In particular, the relative amount of ethylene generated was reduced to ~15% for *N*⁷-hydroxy-*L*-Arg and 0.1–1% when using citrulline, homoarginine, agmatine, *L*-argininamide, and *L*-canavanine. We obtained no evidence for processing of these compounds (e.g., hydroxylation at C3 or C4 of canavanine was not detected). Our earlier studies revealed similar structures for EFE·Mn(II)·2OG·*L*-Arg (PDB: 5V2Y), EFE·Mn(II)·2OG·*N*⁷-hydroxy-*L*-Arg (PDB: 5VKA), and EFE·Mn(II)·2OG·*L*-argininamide (PDB: 5VKB), despite the inability of the latter compounds to generate ethylene.⁹

The EFE·Mn(II)·2OG·*L*-canavanine structure was solved at 1.13 Å resolution (PDB: 6CBA) and shown to be highly superimposable (*C*_α RMSD of 0.07 Å) with the structure of the EFE·Mn(II)·2OG·*L*-Arg complex (PDB: 5V2Y), except that the bound *L*-canavanine adopts a distinct conformation from *L*-Arg (Figure 2). In the *L*-Arg-bound structure, C5 of *L*-Arg points toward Mn; in contrast, O5 of *L*-canavanine, which is equivalent to C5 of *L*-Arg, points away from the metal and toward E84. Close inspection of the density map excluded the possibility of the formation of an O5–E84 hydrogen bond because E84 is very likely in the deprotonated state. Nevertheless, *L*-canavanine adopts an orientation that leads

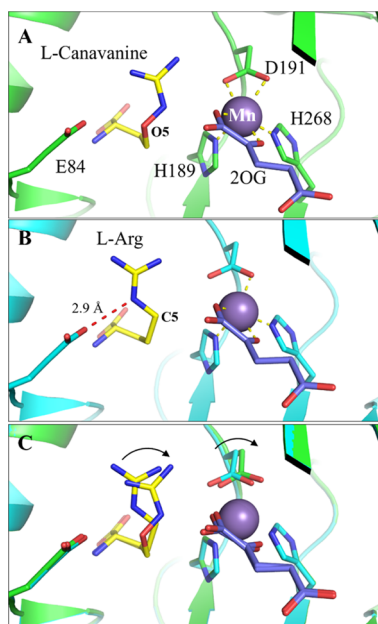


Figure 2. Structural comparison of the EFE·Mn·2OG·*L*-canavanine complex and the EFE·Mn·2OG·*L*-Arg complex. (A) Active site of the EFE with bound *L*-canavanine (PDB: 6CBA). (B) Active site of the EFE with bound *L*-Arg (PDB: 5V2Y). (C) Superimposition of the two structures. The curved arrows indicate the conformational changes from the *L*-Arg-bound structure to the *L*-canavanine-bound structure. The yellow dashed lines indicate metal coordination with Mn(II), whereas the red dashed lines show the hydrogen bonds between the bound ligand and E84.

to a substantial displacement of its guanidine group and accordingly a shift of the side chain of D191. This shift causes a change in the geometry of the metal coordination in the *L*-Arg-bound structure; the distances between Mn and OD1 and OD2 of D191 are 2.70 and 2.17 Å, respectively; in the *L*-canavanine-bound structure, however, the distances are 2.36 and 2.43 Å, respectively. As Mn exhibits monodentate coordination with OD1 in the absence of *L*-Arg (i.e., the EFE·Mn·2OG complex structure, PDB: 5V2X), the *L*-canavanine-bound structure represents an intermediate state between the *L*-Arg-unbound OD1 chelation and the *L*-Arg-bound OD2 chelation (Figure 3).

The *L*-canavanine-bound structure resembles the previously determined structures of the enzyme with *L*-Arg, *N*⁷-hydroxy-*L*-

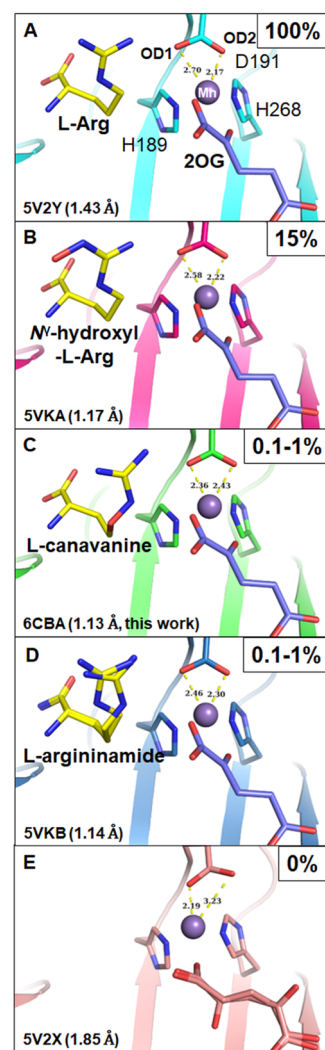


Figure 3. Metal coordination in the EFE without and with *L*-Arg or other *L*-Arg analogues. (A) EFE·Mn(II)·2OG·*L*-Arg complex. (B) EFE·Mn(II)·2OG·*N*⁷-hydroxy-*L*-Arg complex. (C) EFE·Mn(II)·2OG·*L*-canavanine complex. (D) EFE·Mn(II)·2OG·*L*-argininamide complex. (E) EFE·Mn(II)·2OG complex. The relative ethylene producing activity is shown in the upper right corner of each panel, and the PDB ID and resolution for each structure are shown in the lower left corner. The yellow dashed lines indicate the coordination distances (labeled in angstroms) between OD1/2 of D191 and Mn(II) (purple spheres). 2OG is illustrated in sticks with blue carbon atoms when chelating the metal and with pink carbon atoms when coordinating Mn in a monodentate manner.

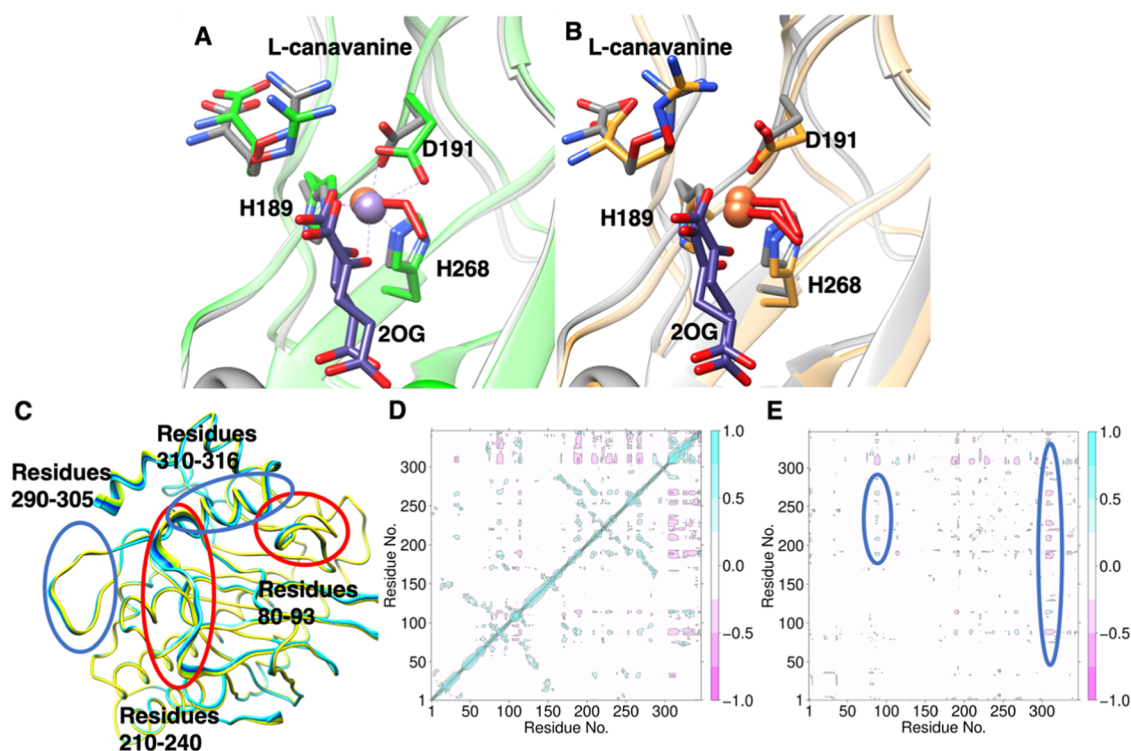


Figure 4. MD analysis of EFE/L-Can. (A) MD average snapshot of EFE-Fe(III)-OO[•]-2OG-L-canavanine with L-canavanine bound in conformation B (gray) and the crystal structure of EFE-Mn-2OG-L-canavanine (green) (PDB: 6CBA). Fe(II) and Mn(II) are shown as orange and purple spheres, respectively, and the coordinated dioxygen is illustrated by red sticks. (B) MD average snapshots of EFE-Fe(III)-OO[•]-2OG-L-canavanine comparing L-canavanine conformations A (orange) and B (gray). (C) PCA of the MD for EFE-Fe(III)-OO[•]-2OG-L-canavanine: circled regions depict the parts of the enzyme that differ (blue) or show similarity (red) in flexibility compared to EFE-Fe(III)-OO[•]-2OG-L-Arg (Figure S4). (D) DCCA showing the anticorrelated (pink) and correlated (blue) motions of EFE/L-Can. (E) d-DCCA, showing the loss of anticorrelation (blue) between β 4 and β 5 (positions 80–93) and the increased anticorrelation (pink) of β 8 residues with other regions of the enzyme comparing EFE/L-Can versus EFE/L-Arg.

Arg, and L-argininamide (Figure 3).⁹ As for the case of bound L-Arg, all these analogues induce a bidentate binding mode of 2OG, congruent with the anaerobic spectroscopy results. In addition, L-canavanine introduces the same side chain changes seen with the other analogues, as discussed in detail in our earlier work.⁹ For EFE-Mn(II)-2OG-*N*⁷-hydroxy-L-Arg, the most active analogue (15% activity), the distances between Mn and OD1 and OD2 of D191 are 2.58 and 2.22 Å, respectively. This structure is nearly identical to that of the L-Arg-bound structure except for the extra *N*⁷-hydroxyl group that prevents C5 hydroxylation, leading to a decrease of activity likely due to the subtle conformation change. The complex containing L-argininamide revealed two conformations of the substrate analogue, one matching L-Arg/*N*⁷-hydroxy-L-Arg and the second shifting the C5 and the *N* ϵ atom with the guanidine group still forming a similar hydrogen bond to OD1 of D191. Only one conformation for D191 is observed in the structure with distances of 2.46 and 2.30 Å. When comparing these high-resolution structures with bound L-Arg analogues as well as the L-Arg-bound and L-Arg-absent structures, we note a very nice transition of the metal chelation, which is influenced by the position and orientation of the guanidine nitrogen of the ligand (Figure 3).

Combined with the changes at the C5 atom, these observations explain the reduced activity when L-Arg is replaced by L-canavanine. In particular, the orientation of L-canavanine with its O5 atom positioned away from the metal site leads to an inability of the first-shell residue D191 to bind

the metal using OD2 that now appears to be critical for ethylene generation. Prior studies with the D191E variant of the EFE also emphasize the importance of the carboxylate metal ligand.⁸ The conservative alteration of D191 led to reduced activity, a smaller ratio of ethylene production versus L-Arg oxidation, and an increased concentration of the ferryl reaction intermediate. Structural studies of the D191E variant enzyme revealed two conformations for E191, one that mimics the orientation of D191 and one that significantly alters the coordination geometry around the Fe(II). Notably, the D191E EFE is the only structurally characterized variant thus far reported for this enzyme; below, we expand structural characterization to the R171A variant, eliminating a residue that stabilizes C1 of 2OG, and Y306A, a variant with an altered residue that lies distant from the active site in the WT EFE.⁹

MD Simulations of the L-Canavanine-Bound EFE. To explore the conformational flexibility of the EFE with bound L-canavanine, we performed MD simulations of the O₂-bound EFE-Fe(III)-OO[•]-2OG-L-canavanine complex (EFE/L-Can) (Figure S1).⁵ Our previous computational study explained the conformational motions of the EFE-Fe(III)-OO[•]-2OG-L-Arg complex (EFE/L-Arg) in extensive detail.⁵ The present study reveals that as in EFE/L-Arg, the substrate analogue in EFE/L-Can exhibits two conformations: one with the O5 atom facing toward the metal (conformation A, Figure 4B) (46.3% of the 1 μ s trajectory) and the other with O5 facing away from the metal (conformation B, Figure 4A,B) (53.6% of the 1 μ s trajectory) (Figure S2). Although the population of con-

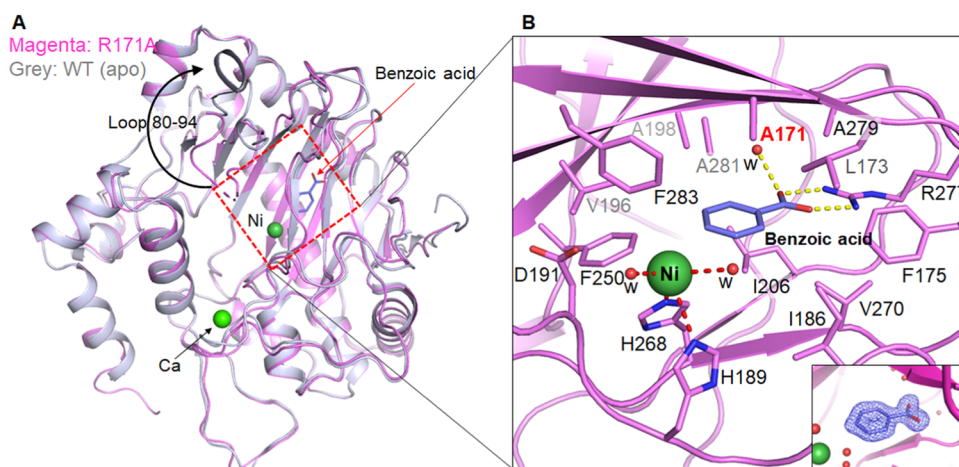


Figure 5. Structural comparison of the R171A variant (PDB: 8UC2) and WT EFE in the apoprotein (apo) state (PDB: 5V2U). (A) Structural superimposition of the two structures. A large conformational change of the loop containing residues 80–94 is indicated by the black arrow. (B) Close-up view of the active site of the R171 variant. An unexpected ligand, benzoic acid, was identified in the crystal structure. The $2F_o - F_c$ map ($\sigma = 1$) of benzoic acid is shown in the inset. The yellow dashed lines indicate the polar interactions of benzoic acid with R277 and a water molecule (w). The hydrophobic residues lining the 2OG binding site are shown and labeled. Ni(II) with 40% occupancy was identified at the metal binding site, and the red dashed lines indicate its coordination with H189, H268, and two water molecules.

formation B in EFE/*L*-Can is reduced compared to that in the EFE/*L*-Arg system (11.7% conformation A and 88.3% conformation B),⁵ state B remains the most populated conformation, in agreement with the crystal structure. The key hydrogen bonds and salt bridge interactions between the C5 carboxylate of 2OG and R277 (92%) and between the C1 carboxylate of 2OG and R171 (70.4%) are stable through the MD simulations, consistent with the crystal structure of the EFE with *L*-canavanine (Figure 4A) and the MD simulations of EFE/*L*-Arg. In snapshots with *L*-canavanine conformation A, the guanidine group ϵ N forms a stable hydrogen bond with nonbonded OD1 of D191 (46.3%), while such an interaction is less prominent in snapshots with *L*-canavanine in conformation B. Importantly, the oxygen atom at the fifth position of *L*-canavanine has van der Waals interactions with E84 in the case of conformation B (25%, with distances varying between 2.7 and 3.6, 3.3 Å between OS of *L*-canavanine and the carboxylate oxygens of E84), similar to the crystal structure (Figure S3), which disappears in conformation A where the oxygen faces toward the metal, moving away from E84. The reason that bidentate D191 is observed in the crystal structure might be because *L*-canavanine in conformation B does not lead to a hydrogen bond between D191 and *L*-canavanine. The transition from intermediate bidentate-coordinated D191 to monodentate coordination might occur upon the binding of dioxygen.

PCA is a tool that helps analyze the motions of the flexible regions of proteins and discriminates relevant conformational changes.⁴⁷ Compared to the PCA of EFE/*L*-Arg (Figure S4), the EFE/*L*-Can system had reduced flexibility (Figure 4C). The flexibility observed in residues 210–240 and 290–305 was substantially reduced. Unlike the EFE/*L*-Arg simulation, the EFE/*L*-Can system showed increased flexibility in the β 8 strand (residues 310–316). This change of flexibility is reflected in the interdependent correlated motions as well. Based on the d-DCCA (Figure 4E) between simulations of the protein with *L*-Arg and *L*-canavanine, the simulations with *L*-canavanine showed a substantial loss of anticorrelated motion between the loop connecting β 4 and β 5 (80–93) and the β 11 loop (210–240) and reduced anticorrelated motion between

β 4 and β 5 (80–93) and the loop containing β 15 and α 8 (290–305). There is, however, an increased anticorrelated motion between β 8 (residues 310–316) and the three loops mentioned above, which is not present in the case of *L*-Arg (Figure 4C). *L*-Canavanine exhibits correlated motion with β 5 and anticorrelated motion with β 8. The motion of these loops is important as they form the hydrophobic environment of the active site, which determines the substrate analogue and 2OG conformations. The DCCA of EFE/*L*-Can also revealed correlation between β 15 and the loop connecting to α 8, associated with active site hydrophobic residues including F283, F250, and A282 that interact with dioxygen (Figure 4D). The above residues exhibited anticorrelated motion in EFE/*L*-Arg. This combination of changes in the correlated motions of the EFE in the presence of the *L*-canavanine, in addition to the differences in the substrate analogue conformation and second-sphere hydrogen bonding interactions, could contribute to structural changes and ultimately to the loss of ethylene forming activity of the enzyme compared to the *L*-Arg-bound EFE. Overall, the results suggest the importance of the second coordination sphere and long range interactions in promoting EFE activity.

Experimental Studies of the R171A Variant of the EFE. Anaerobic UV–vis spectroscopy demonstrated a lack of MLCT transitions when Fe(II) and 2OG were added to the R171A EFE, regardless of whether *L*-Arg was also provided (data not shown). This result suggested that either Fe(II) or 2OG fails to bind to the enzyme or they do not form the chelate structure that is typically associated with Fe(II)/2OG-dependent oxygenases.

We solved the structure of the presumed R171A variant apoprotein of the EFE to a resolution of 1.60 Å (PDB: 8UC2). Although no metal ions had been provided during crystallization, we observed electron density that was consistent with the bound metal at two sites (Figure 5A). The active site bound a metal at approximately 40% occupancy that was assigned to nickel, suspected to be derived during purification of the enzyme using the Ni-NTA resin. Ni(II) was bound to H189 and H268 but not to D191, which along with F250 exhibited distinct conformations when compared to other EFE

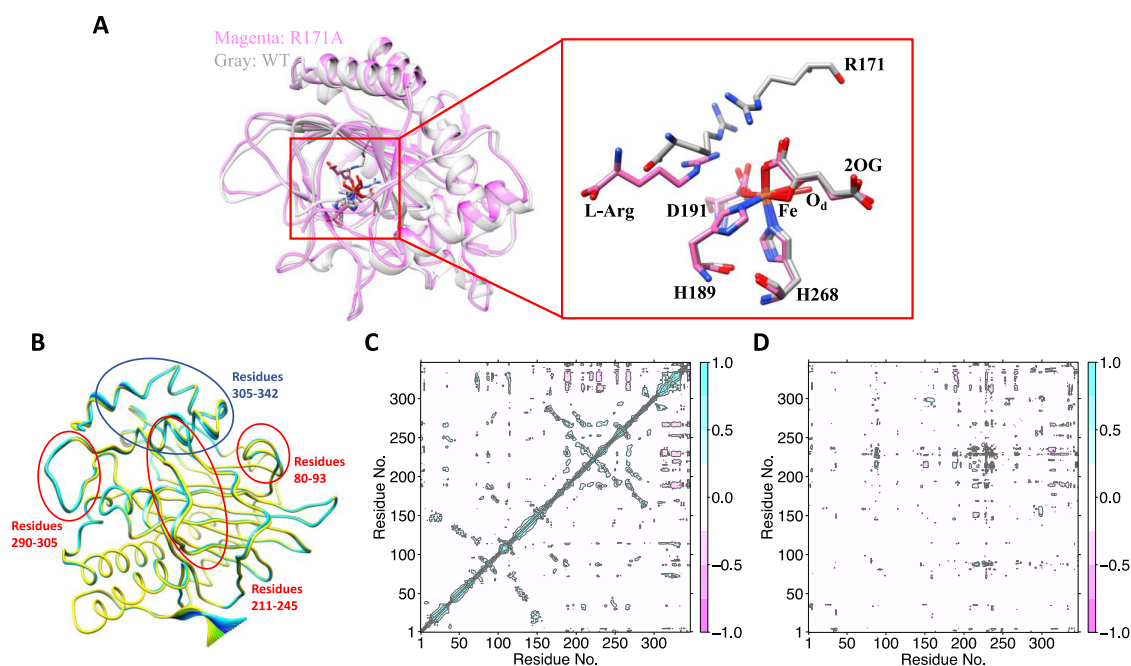


Figure 6. MD simulation of the R171A EFE. (A) Overlay of the average MD snapshots for the WT and R171A variant EFE. The active site residues are expanded for clarity on the right. The WT EFE is shown in gray, and the R171A variant is magenta. (B) PCA showing the flexible regions in the R171A EFE. (C) DCCA and (D) d-DCCA comparing the correlated motions in the WT and R171A EFE.

structures. The second metal was bound to the backbone carbonyl atoms of N220, L222, and E225; this metal was tentatively assigned as Ca based on the electron density and geometry. The source of this metal ion is unknown, and it is not further discussed. Unexpectedly, electron density was present at the 2OG binding pocket, but this density did not match the structure of 2OG. Analysis of the crystals by LC-MS yielded a feature with the same mass and retention time (4.61 min) as benzoic acid (Figure S5), in agreement with the appearance of the electron density (Figure 5B, inset). By contrast, we eliminated nicotinic acid and salicylic acid as possible ligands based on their distinct elution positions (1.21 and 4.81 min, respectively) and masses. The carboxylate group of benzoic acid forms a salt bridge with R277, which normally forms a salt bridge with the C5 carboxylate of 2OG in the WT protein. Surrounding the aromatic ring of this molecule is a cage of hydrophobic residues: L173, I186, V196, A198, L206, F250, V270, A279, A281, and F283 (Figure 5B). The source of benzoic acid in the protein is unknown, but it must be obtained from the medium during growth of the recombinant cells or during variant enzyme isolation. Benzoic acid was found to not inhibit the WT EFE when tested at 5 mM concentration, indicating that it cannot outcompete 2OG binding even when at 10-fold excess. We postulate that the reason why this compound was able to bind to the R171A variant but not to the WT EFE when 2OG was not added in the crystallization buffer is that the deletion of the bulky and charged side chain of R171 leaves an open and hydrophobic space for this compound to enter (Figure 5B). Overall, these studies reveal that substitution of a substrate-binding, second-sphere residue can lead to profound changes at the active site.

MD Simulations of the R171A EFE. Prior experimental and computational studies of the WT EFE had revealed the role of R171 in stabilizing the off-line conformation of 2OG,^{5,9} i.e., an orientation with an oxygen-binding site that is not directed toward L-Arg. Further, R171 plays a decisive role in

Arg positioning since the guanidinium group of R171 forms an electrostatic interaction with this substrate.⁵ Loss of these interactions might disrupt the substrate orientation in the active site. We performed a 1 μ s MD simulation of the ferric-superoxo complex of the R171A variant to compare the substrate interactions to the WT enzyme (Figure 6A). The stability of the MD simulation was determined based on the RMSD of the C α -carbons in the EFE variant (Figure S6). Previous studies had shown that in the WT EFE, η^2 of the NH₂ and ϵ NH in L-Arg forms 44 and 68% hydrogen bond interactions with D191, respectively. Furthermore, E84 makes a hydrogen bond interaction with the ammonium group of the L-Arg in the WT EFE.⁵ Notably, substrate interactions with D191 and E84 are absent in the R171A variant, indicating substrate reorientation in the active site. Instead of E84, the backbone carbonyl of V85 interacts with the ammonium group of the L-Arg (45%), and the guanidinium group flips between the C1 carboxylate oxygens of the 2OG and D191 (Figure S7). The time-dependent fluctuations of the distance between Fe and the C5 hydrogen in the variant demonstrate the existence of both conformations A and B (Figure S8). At the same time, the position of the guanidinium group varies in the WT and variant enzymes (Figure 6A). The salt bridge interaction of R277 with the C5 carboxylate of 2OG is retained in the R171A variant (85%). These interactions illustrate the substitution's impact on the enzyme's active site. Additionally, MMGBSA calculations indicate weak L-Arg binding in the variant enzyme with a binding free energy of -17.0 kcal mol⁻¹ in comparison to the -35.2 kcal mol⁻¹ for the WT EFE.

PCA demonstrated the motion of the flexible regions in the R171A EFE (Figure 6B). Three regions in the WT enzyme were shown to be considerably flexible; however, the flexibility of the region involving residues 291–303 (a loop that connects with α 8 and β 15) is completely lost in the variant. The other flexible regions in the WT EFE involve (i) residues from 80 to

93 (a loop connecting $\beta 4$ and $\beta 5$) and (ii) a loop connecting $\beta 11$ (residues 211–245). The flexibility of these regions in R171A was appreciably lost compared to the WT enzyme; these changes might affect the binding of 2OG and L-Arg. Strikingly, regions of $\alpha 7$, $\alpha 8$, $\alpha 9$, and connecting residues (positions 305–342) in the R171A EFE were shown to move away from the active site. There was also a slight increase in the flexibility of the loop connecting $\beta 8$ and $\beta 9$ (residues 189–195) in the variant enzyme compared to the WT EFE. Thus, the PCA results reveal substantial changes in the motion of the flexible region as an implication of the substitution of R171.

DCCA elucidates the differences in correlated motions between the R171A and WT EFE (Figure 6C,D). Previous studies have shown that R171 positively correlated with the residues of $\beta 6$ and $\beta 14$.⁵ The variant showed no remarkable correlation with $\beta 6$ and $\beta 14$ but demonstrated anticorrelated motions with residues of $\alpha 8$ and $\alpha 9$; the loss of positive correlation might affect the L-Arg orientation in the active site. Furthermore, the correlated motion of the region encompassing the loop connecting $\alpha 8$ and $\beta 15$, along with the loop connecting $\beta 4$ and $\beta 5$, and the loop linking $\beta 11$, which played a crucial role in L-Arg binding in the WT EFE, is disrupted in the variant. The loss of these correlated motions with the loop connecting $\alpha 8$ and $\beta 15$ might have destabilized L-Arg binding in the active site. The region consisting of residues 305–342 ($\alpha 7$, $\alpha 8$, $\alpha 9$, and its connecting residues) in the variant makes an asynchronous (anticorrelated) motion with the regions connecting $\beta 4$ and $\beta 5$, $\beta 8$ and $\beta 9$, and the loop joining $\beta 11$. Hydrophobic residues (F250, A281, and F283) interacting with active site residues showed no anticorrelated motion in the variant enzyme, as seen in the WT enzyme. The increased flexibility and correlated motions of the region constituting residues 305–342 might be due to the loss of interaction of R171 with the residues of $\alpha 7$, which in turn might affect the L-Arg stability in the active site. The MD study of the R171A EFE reveals a significant alteration in the L-Arg conformation and long-range correlated motion, potentially contributing to the reduced catalytic efficiency of this variant.

Experimental Studies of the Y306A Variant of the EFE. Y306 is a conserved residue that we proposed to be important because of its hydrophobic interactions.⁹ Compared to the WT enzyme, Y306A and Y306F possess ethylene-forming activities of 4 and 6% and have P5C-forming activities of 5 and 10%, respectively. The addition of 2OG to Y306A EFE·Fe(II) led to a very small increase in absorption at ~ 525 nm (calculated to be ~ 130 M⁻¹ cm⁻¹), consistent with only partial chelation of Fe(II) by 2OG. For WT EFE·Fe(II)·2OG, the addition of L-Arg or L-canavanine led to an increase in these MLCT transitions (Figure 1); however, L-Arg addition to Y306A EFE·Fe(II)·2OG appeared to eliminate this absorption peak and generated an increase in light scattering (Figure 7). We conclude that the Y306A variant of the EFE does not bind L-Arg in the same manner as seen for the WT enzyme and fails to promote bidentate binding of 2OG.

The Y306A variant of EFE was cocrystallized with MnCl₂, 2OG, and L-Arg, but the structure solved at 1.12 Å (PDB: 6CF3) showed a Y306A EFE·Mn(II)·2OG complex, lacking L-Arg, which is in contrast to all WT EFE structures cocrystallized with similar compounds. As a result, the Y306A EFE·Mn(II)·2OG complex exhibits structural features that are similar to those of the EFE·Mn(II)·2OG complex (PDB: 5V2X). In particular, the 2OG adopts two conformations—it is coordinated in a bidentate manner with the

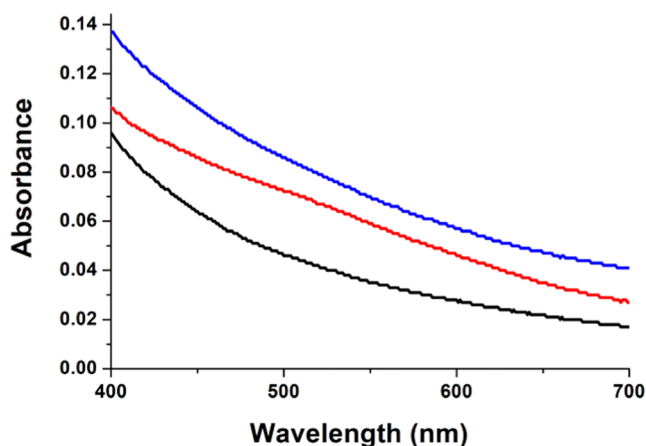


Figure 7. Anaerobic UV–vis spectra of Y306A EFE·Fe(II) (black), Y306A EFE·Fe(II)·2OG (red), and Y306A EFE·Fe(II)·2OG·L-Arg (blue).

metal in a productive conformation (with an occupancy of 63%) or in a monodentate manner with the metal in an opposite orientation (with an occupancy of 37%); the structural observation of this mixture matches results from the UV–vis studies. The latter conformation is observed only when L-Arg is absent in the crystal structure for the WT EFE. When compared to the L-Arg-bound structure of the WT EFE, several structural elements are in different conformations (Figure 8A). These include the loop containing residues 80–94, which form multiple hydrogen bonds with L-Arg in the WT enzyme but flip away from the L-Arg-binding site in the variant structure; the loop containing residues 177–184, which fold into a one-turn helix not seen in the L-Arg-bound structure; and the helix containing residues 306–318 and 321–330 that shift away from the L-Arg binding site.

The close-up view of the superimposed structures reveals the structural basis of the loss of L-Arg binding capacity for the Y306A variant of the EFE (Figure 8B). Substitution of the bulky Y306 by alanine leads to a cavity in the structure, which is occupied by F310 and Y192 that flip down their side chains from the positions in the L-Arg-bound state. The conformational change of Y192 appears to have a profound effect on L-Arg binding because it not only forms a hydrogen bond with the carboxylic acid group of L-Arg but also defines the shape of the L-Arg binding pocket and limits the flexibility of the bound ligand. The empty space created by the flipping of Y192 is only partially occupied by an ethylene glycol molecule (from the cryoprotectant), which is very unlikely to be able to compensate for the loss of the bulky and rigid side chain of Y192. The defect in L-Arg binding is further exaggerated by the loss of additional key interactions, including R316 in the helix associated with residues 306–318, which is shifted away from the L-Arg-binding pocket, and residues 84–86 in the loop involving positions 80–94 that flips away from the L-Arg-binding site. We and others have shown by substitution of the L-Arg-binding residues E84 and T86 that residues in this loop are crucial for ethylene formation.^{9,10} E84 is particularly important in that it not only forms hydrogen bonds with the amino and guanidine groups of L-Arg but also plays a role in orienting R171 via a salt bridge. R171 is a key residue in 2OG binding, and the reoriented R171 likely accounts for the nonproductive binding of 2OG. In the Y306A EFE structure, R171 adopts a conformation that would clash with the

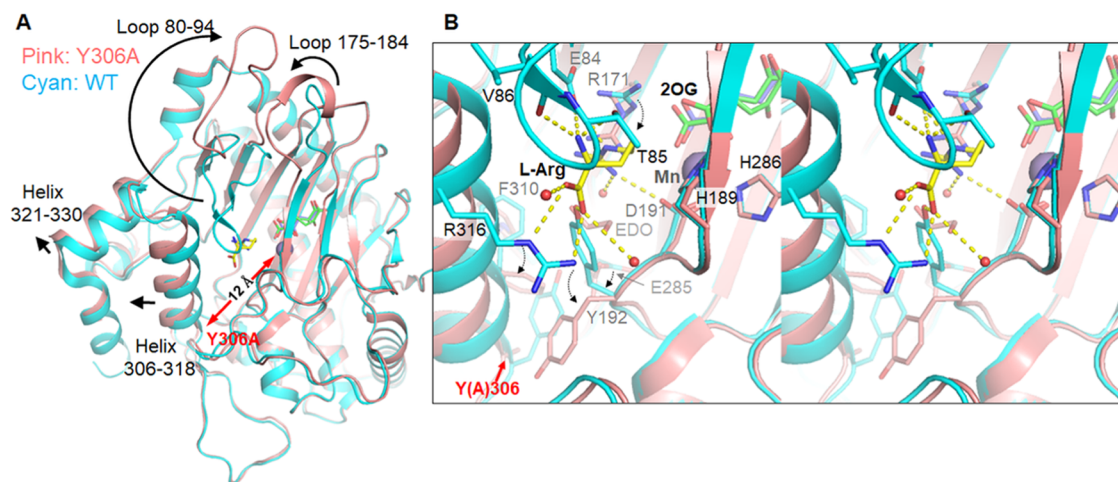


Figure 8. Structural comparison of the Y306A variant EFE (PDB: 6CF3) and WT EFE (PDB: 5V2Y). (A) Structural superimposition of the Y306A EFE·Mn(II)·2OG complex and the EFE·Mn(II)·2OG·L-Arg complex. Significant conformational changes were identified for four secondary structural elements, which are indicated by the black arrows. (B) Stereo-view of the active sites of the WT and the Y306A variant EFE. The black curved arrows indicate the significant structural changes of four residues of the Y306A variant when compared to the corresponding residues in the WT EFE. The yellow dashed lines indicate the polar interactions of L-Arg (sticks with yellow carbon atoms) with the residues in the L-Arg-binding site and water molecules. An ethylene glycol (EDO) molecule was identified in the Y306A structure.

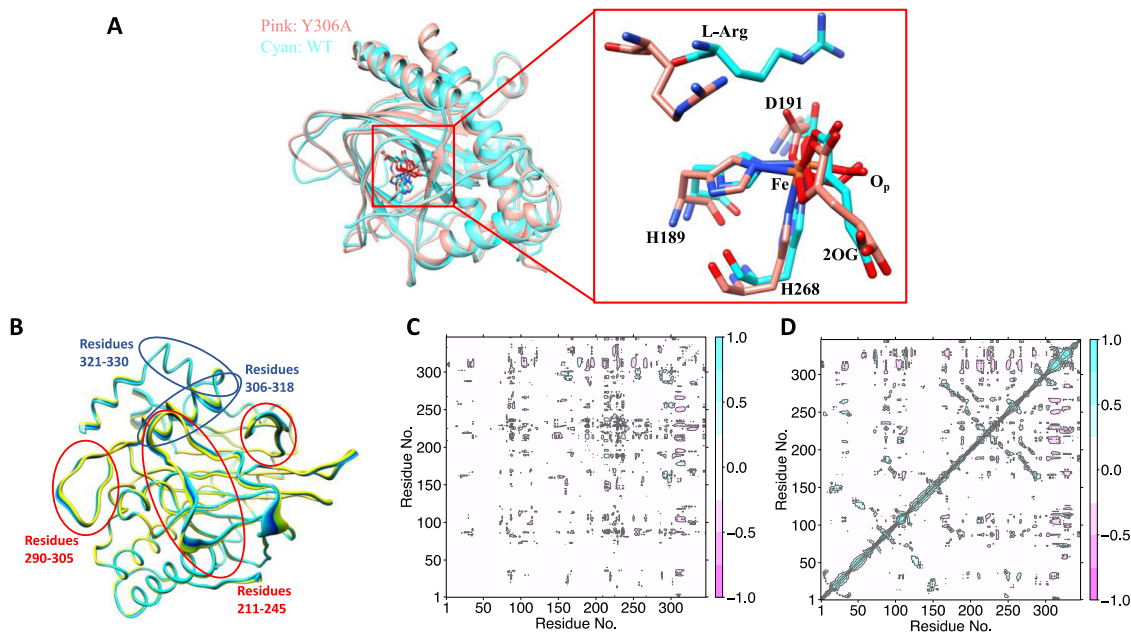


Figure 9. MD simulation of Y306A EFE·Fe(III)·OO^{•−}·2OG·L-Arg. (A) Overlay of MD snapshots of the WT and Y306A variant EFE. The active site residues are expanded on the right. The WT EFE is shown in cyan, and the variant is shown in pink. (B) PCA showing the flexible regions in Y306A. The regions in red are flexible in the WT and variant enzymes, whereas the region in blue shows flexibility only in the variant EFE. (C) d-DCCA and (D) DCCA comparing the correlated motions in the Y306A and WT EFE.

guanidine group of L-Arg if the latter was bound in the same way as in the WT enzyme. In addition, an altered orientation of the side chain of E285, which directly interacts with Y306, may also contribute to the inability of the variant to bind L-Arg, as this residue is connected to the ligand-binding site via ordered water molecules. Taken together, this study demonstrates a case where an interior cavity created by a single amino acid substitution 12 Å away from the catalytic metal center leads to local and distal structural changes that significantly reduce the enzyme's activity.

MD Simulations of the Y306A EFE. We explored the conformational flexibility of the Y306A variant of the EFE in

the presence and absence of the L-Arg substrate. A 1 μs MD simulation was performed utilizing the crystal structure of the Y306A EFE to explore the conformational changes in the absence of L-Arg (Figure S9). Like what was found in the crystal structure, a salt bridge interaction between R277 and the CS carboxylate oxygen and a hydrogen bonding interaction stabilize the binding of 2OG. The MD simulation showed that the side chains of F310 and Y192 occupy the vacancy created by the substitution of Y306, as in the crystal structure. However, these residues are dynamic with high flexibility and occupy different orientations (Figure S10); hence, they may

not contribute to the L-Arg binding as effectively as in the WT EFE. Similarly, the side chain of R171 is flexible and can influence the L-Arg binding (Figure S11). The loop 80–94 and helical residues 306–318, which move away from the active site in the crystal structure, are also highly flexible in the MD (Figure S9B). The combined effect of the flexibility of the loop and helical regions, along with the openness of the side chains, potentially destabilizes L-Arg binding in the active site, which might have prevented the crystallization of the substrate.

The Y306A variant exhibits very low rates of ethylene production and L-Arg hydroxylation (i.e., 3–5% of that for the WT enzyme), but it still has slight activity. To further understand how the Y306A substitution influences L-Arg binding, we performed a 1 μ s MD simulation of Y306A EFE·Fe(III)·OO[•]·2OG·L-Arg that was generated from the WT EFE crystal structure with bound 2OG and L-Arg after performing *in silico* substitution of Y306 by alanine (Figure S12). The simulation demonstrates substantial changes in the orientation of the substrate in the active site (Figure 9A) with respect to the WT EFE/L-Arg.⁵ The hydrogen bonding interaction of the C1 carboxyl groups of 2OG with R171 (47%) is weakened, and the salt bridge interaction of R277 with the C5 oxygen atoms is retained in the simulation for the variant enzyme. The guanidium group of L-Arg also makes weaker electrostatic interactions with R171 (Figure S13). Previous simulation of the WT EFE/L-Arg revealed that stronger interactions between the guanidium group of the L-Arg and R171 favor conformation A, while weaker interactions favor conformation B.⁵ Therefore, in the Y306A variant, L-Arg is displaced to conformation B (Figure S14), but the guanidinium group of L-Arg moves further away than in WT EFE/L-Arg as shown in Figure S13. The larger shift in the guanidinium group of L-Arg in the Y306A variant protein might be a factor in the low reactivity of this variant enzyme. The interaction of E84 with the amino group of L-Arg stabilizes conformation B in the WT EFE/L-Arg;⁵ this interaction is not present in Y306A, whereas the amino group interacts with the backbone of V85 (66%) (Figure S15). Further, the carboxylate group of the L-Arg does not interact with Y192 as in the WT enzyme since it is flexible due to the alanine substitution. Overall, the MD analysis shows that L-Arg exhibits weaker interactions with the residues in the variant protein compared to the WT EFE. Consistently, the calculated binding energy of the substrate using MMGBSA displayed weaker binding of L-Arg in the Y306A variant (-17.1 kcal mol⁻¹) than in the WT enzyme (-35.2 kcal mol⁻¹).

PCA showed that helix-forming residues 306–318 and 321–330 have high flexibility in the Y306A variant EFE, which was not observed in the WT enzyme (Figure 9B). Crystallographic studies have also demonstrated that this helical region shifts away from the L-Arg-binding site; therefore, the flexibility of these helical regions might destabilize binding of the substrate in the active site. d-DCCA of the WT and variant enzymes shows anticorrelated motion between the loop connecting β 15 and α 8 (residues 290–305) with the loop connecting β 4 and β 5 (80–93) and β 11 loop (210–240). The lack of L-Arg in the Y306A EFE crystal structure is consistent with a significant decrease in binding affinity compared to the WT enzyme, implying that the long-range motions have a role in substrate binding (Figure 9C). Also, helical residues 306–318 and 321–330 show anticorrelated motions with the remaining regions of the protein, further indicating the influence of long-range correlations in stabilizing L-Arg binding (Figure 9D). The

different correlated motions and binding interactions in the variant and WT EFE might reflect the destabilization of the substrate, thereby reducing the catalytic turnover of the Y306A EFE.

CONCLUSIONS

L-Canavanine binds to the EFE with its O5 atom pointing away from the metal center, whereas C5 of L-Arg points toward the iron atom, consistent with the ability of the enzyme to react with only the latter compound. Congruent with the structural results, MD simulations predict that the most favored L-canavanine conformation in the EFE·Fe(III)·OO[•]·2OG·L-canavanine complex has its O5 pointed away from the metal center and facing toward E84. Anaerobic UV–vis spectroscopy of EFE·Fe(II)·2OG·L-canavanine and structural studies of EFE·Mn(II)·2OG·L-canavanine confirm that binding of the substrate analogue leads to bidentate coordination of 2OG to Fe(II). Critical insights into why the L-canavanine-bound enzyme is ineffective for ethylene generation were obtained from the crystal structure of EFE·Mn(II)·2OG·L-canavanine that revealed that OD1 of D191 coordinated the metal instead of the catalytically productive OD2 coordination noted for bound L-Arg.

Alanine substitution of R171, a residue that forms a hydrogen bond with the C1 carboxylate of 2OG, leads to a significant shift of the loop containing residues 80–94, a loss in activity, and an inability to form the chromophore associated with chelate binding of Fe(II) by 2OG. These changes cause reduced stability of binding 2OG, leading to the adventitious binding of benzoate at the 2OG binding site. The source of benzoate is unclear, analogous to the situation for crystals of canavalin, a jack bean vicilin protein, that was purified from *E. coli* cells and shown to have bound benzoate.⁴⁸ Both the UV–vis studies and MD simulations suggest a weakened binding of L-Arg for this variant protein.

Y306 is a generally conserved residue located 12 Å from the metal site, with the active site residue Y192 located in between. We had proposed that Y306 was important because of its hydrophobic interactions and its role in forming a hydrogen bond between its hydroxyl group and E285.⁹ Substitution of E285 to alanine or glutamine abolished ethylene formation while maintaining 15 or 20% of P5C formation, respectively. The Y306 and E285 side chain orientations were the same in all WT enzyme structures, regardless of the presence of metal, 2OG, or L-Arg. However, E285 also forms hydrogen bonds to water molecules that connect to the backbone of D191 and Y192. A water-to-water hydrogen bond links E285 to the guanidine group of L-Arg and L-Arg analogues as well as to H169. These results indicate that there is a connection of Y306 to the active site that could explain the reduction in activity for the Y306A variant. The structure of this protein now offers an even more clear and surprising picture of major changes introduced by a single-point substitution. Substitution of Y306 by alanine leads to a large conformational change of one loop with additional propagated changes to affect several residues important for L-Arg binding, accounting for the inactivity of this variant. UV–vis and structural studies both reveal that 2OG binds to the protein to yield a mixture of the monodentate- and bidentate-coordinated species. The loss of the MLCT transitions upon addition of L-Arg is consistent with a deficiency in L-Arg binding as revealed by its absence of this amino acid in the crystal structure despite being present in the crystallization buffer. In line with experimental studies, the

binding energy calculated from the MD simulations using MMGBSA revealed a lower binding affinity of L-Arg in the Y306A variant enzyme.

In combination, the results presented here are compatible with the currently proposed enzyme mechanism,^{4–8} emphasize the necessity for OD2 of D191 to coordinate the metal, and highlight the importance of the second coordination sphere and longer range interactions in promoting EFE activity. Such interactions must be kept in mind when investigating variant forms of the EFE for enhanced production of ethylene as a biofuel or for making increased amounts of guanidine as an agricultural fertilizer using recombinant microorganisms.^{49–53}

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.biochem.4c00031>.

Crystal statistics for structures of 6CBA, 6CF3, and 8UC2, MD analysis for the three samples along with conformation analyses, predicted hydrogen bonding interactions, and PCA, and ESI-MS analysis of the organic compound extracted from R171A crystals (PDF)

Accession Codes

Uniprot accession code for EFE: P32021

■ AUTHOR INFORMATION

Corresponding Authors

Christo Z. Christov – Department of Chemistry, Michigan Technological University, Houghton, Michigan 49931, United States; orcid.org/0000-0002-4481-0246; Email: christov@mtu.edu

Jian Hu – Department of Biochemistry and Molecular Biology and Department of Chemistry, Michigan State University, East Lansing, Michigan 48824, United States; orcid.org/0000-0001-6657-9826; Email: hujian1@msu.edu

Robert P. Hausinger – Department of Microbiology, Genetics, and Immunology, Michigan State University, East Lansing, Michigan 48824, United States; Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, Michigan 48824, United States; orcid.org/0000-0002-3643-2054; Email: hausinge@msu.edu

Authors

Shramana Chatterjee – Department of Microbiology, Genetics, and Immunology, Michigan State University, East Lansing, Michigan 48824, United States

Matthias Fellner – Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, Michigan 48824, United States; Present Address: Biochemistry Department, School of Biomedical Sciences, University of Otago, Dunedin 9054, New Zealand; orcid.org/0000-0003-3192-6984

Joela Rankin – Department of Microbiology, Genetics, and Immunology, Michigan State University, East Lansing, Michigan 48824, United States; Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, Michigan 48824, United States; Present Address: Department of Biochemistry Molecular Biology and Biophysics, University of Minnesota, Minneapolis, Minnesota 55108, United States

Midhun G. Thomas – Department of Chemistry, Michigan Technological University, Houghton, Michigan 49931, United States

Simahudeen Bathir J S Rifayee – Department of Chemistry, Michigan Technological University, Houghton, Michigan 49931, United States; orcid.org/0000-0001-8557-078X

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.biochem.4c00031>

Author Contributions

[▽]S.C. and M.F. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank Salette Martinez for assistance with crystallization of the L-canavanine-bound and Y306A variant forms of the EFE, Dr. Tony Schillmiller for assistance with mass spectrometry, and the National Science Foundation (grants 1904215 and 2203630 to C.Z.C. and 1904295 and 2203472 to R.P.H. and J.H.).

■ DEDICATION

This article is dedicated to the memory of Christopher T. Walsh, an outstanding mechanistic enzymologist and post-doctoral mentor to RPH.

■ REFERENCES

- (1) Hausinger, R. P.; Rifayee, S. B. J. S.; Thomas, M. G.; Chatterjee, S.; Hu, J.; Christov, C. Z. Biological formation of ethylene. *RSC Chem. Biol.* **2023**, *4*, 635–646.
- (2) Fukuda, H.; Ogawa, T.; Tazaki, M.; Nagahama, K.; Fujii, 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*, 483–489.
- (3) 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*, S989–S999.
- (4) Xue, J.; Lu, J.; Lai, W. Mechanistic insights into a non-heme 2-oxoglutarate-dependent ethylene-forming enzyme: Selectively of ethylene-formation versus L-Arg hydroxylation. *Phys. Chem. Chem. Phys.* **2019**, *21*, 9957–9968.
- (5) Chaturvedi, S. S.; Ramanan, R.; Hu, J.; Hausinger, R. P.; Christov, C. Z. Atomic and electronic structure determinants distinguish between ethylene formation and L-arginine hydroxylation reactions mechanisms in the ethylene-forming enzyme. *ACS Catal.* **2021**, *11*, 1578–1592.
- (6) Copeland, R. A.; Zhou, S.; Shaperdorth, I.; Shoda, T. K.; Bollinger, J. M., Jr.; Krebs, C. Hybrid radical-polar pathway for excision of ethylene from 2-oxoglutarate by an iron oxygenase. *Science* **2021**, *373*, 1489–1493.
- (7) Burke, E. J.; Copeland, R. A.; Dixit, Y.; Krebs, C.; Bollinger, J. M., Jr. Steric perturbation of the Grob-like final step of ethylene-forming enzyme enables 3-hydroxypropionate and propylene production. *J. Am. Chem. Soc.* **2024**, *146*, 1977–1983.
- (8) Copeland, R. A.; Davis, K. M.; Shoda, T. K. C.; Blaesi, E. J.; Boal, A. K.; Krebs, C.; Bollinger, J. M., Jr. An iron(IV)-oxo intermediate initiating L-arginine oxidation but not ethylene production by the 2-oxoglutarate-dependent oxygenase, ethylene-forming enzyme. *J. Am. Chem. Soc.* **2021**, *143*, 2293–2303.
- (9) 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*, 11980–11988.
- (10) 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*, 4667–4672.
- (11) Kruschel, D.; Zagrovic, B. Conformational averaging in structural biology: Issues, challenges and computational solutions. *Mol. Biosyst.* **2009**, *5*, 1606–1616.
- (12) Karplus, M.; Kuriyan, J. Molecular dynamics and protein function. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6679–6685.
- (13) Karplus, M.; McCammon, J. A. Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.* **2002**, *9*, 646–652.
- (14) Corbella, M.; Pinto, G. P.; Kamerlin, S. C. L. Loop dynamics and the evolution of enzyme activity. *Nat. Rev. Chem.* **2023**, *7*, 536–547.
- (15) Stripp, S. T.; Duffus, B. R.; Fourmond, V.; Léger, C.; Leimkühler, S.; Hirota, S.; Hu, Y.; Jasniowski, A.; Ogata, H.; Ribbe, M. W. Second and outer coordination sphere effects in nitrogenase, hydrogenase, formate dehydrogenase, and CO dehydrogenase. *Chem. Rev.* **2022**, *122*, 11900–11973.
- (16) de Visser, S. P. Second-coordination sphere effects on selectivity and specificity of heme and nonheme iron enzymes. *Chem. - Eur. J.* **2020**, *26*, 5308–5327, DOI: 10.1002/chem.201905119.
- (17) Wojdyla, Z.; Borowski, T. Properties of the reactants and their interactions within and with the enzyme binding cavity determine reaction selectivities. The case of Fe(II)/2-oxoglutarate dependent enzymes. *Chem. - Eur. J.* **2022**, *28*, No. e202104106.
- (18) Zoi, I.; Suarez, J.; Antoniou, D.; Cameron, S. A.; Schramm, V. L.; Schwartz, S. D. Modulating enzyme catalysis through mutations designed to alter rapid protein dynamics. *J. Am. Chem. Soc.* **2016**, *138*, 3403–3409.
- (19) Schafer, J. W.; Zoi, I.; Antoniou, D.; Schwartz, S. D. Optimization of the turnover in artificial enzymes via directed evolution results in the coupling of protein dynamics to chemistry. *J. Am. Chem. Soc.* **2019**, *141*, 10431–10439.
- (20) Rabe, P.; Kamps, J. J. A. G.; Sutherland, K. D.; Linyard, J. D. S.; Aller, P.; Pham, C. C.; Makita, H.; Clifton, I.; McDonough, M. A.; Leissing, T. M.; Shutin, D.; Lang, P. A.; Butryn, A.; Brem, J.; Gul, S.; Fuller, F. D.; Kim, I.-S.; Cheah, M. H.; Fransson, T.; Bhowmick, A.; Young, I. D.; O'Riordan, L.; Brewster, A. S.; Pettinati, I.; Doyle, M.; Joti, Y.; Owada, S.; Tono, K.; Batyuk, A.; Hunter, M. S.; Alonso-Mori, R.; Bergman, U.; Owen, R. L.; Sauter, N. K.; Claridge, T. D. W.; Robinson, C. V.; Yachandra, V. K.; Yano, J.; Kern, J. F.; Orville, A. M.; Schofield, C. J. X-ray free-electron laser studies reveal correlated motion during isopenicillin N synthase catalysis. *Sci. Adv.* **2021**, *7*, No. eabh0250.
- (21) Light, K. M.; Hangasky, J. A.; Knapp, M. J.; Solomon, E. I. Spectroscopic studies of the mononuclear non-heme Fe^{II} enzyme FIH: Second-sphere contributions to reactivity. *J. Am. Chem. Soc.* **2013**, *135*, 9665–9674.
- (22) Goldberg, N. W.; Knight, A. M.; Zhang, R. K.; Arnold, F. H. Nitrene transfer catalyzed by a non-heme iron enzyme and enhanced by non-native small-molecule ligands. *J. Am. Chem. Soc.* **2019**, *141*, 19585–19588.
- (23) Miller, G. P.; Wahnou, D. C.; Benkovic, S. J. Interloop contacts modulate ligand cycling during catalysis by *Escherichia coli* dihydrofolate reductase. *Biochemistry* **2001**, *40*, 867–875.
- (24) Hanoian, P.; Liu, C. T.; Hammes-Schiffer, S.; Benkovic, S. Perspectives on electrostatics and conformational motions in enzyme catalysis. *Acc. Chem. Res.* **2015**, *48*, 482–489.
- (25) Watney, J. B.; Agarwal, P. K.; Hammes-Schiffer, S. Effect of mutation on enzyme motion in dihydrofolate reductase. *J. Am. Chem. Soc.* **2003**, *125*, 3745–3750.
- (26) Hospital, A.; Goñi, J. R.; Orozco, M.; Gelpi, J. L. Molecular dynamics simulations: Advances and applications. *Adv. Appl. Bioinf. Chem.* **2015**, *8*, 37–47, DOI: 10.2147/AABC.S70333.
- (27) Orozco, M. A theoretical view of protein dynamics. *Chem. Soc. Rev.* **2014**, *43*, S051–S066.
- (28) Adcock, S. A.; McCammon, J. A. Molecular dynamics: Survey of methods for simulating the activity of proteins. *Chem. Rev.* **2006**, *106*, 1589–1615.
- (29) Wang, W. Recent advances in atomic molecular dynamics simulation of intrinsically disordered proteins. *Phys. Chem. Chem. Phys.* **2021**, *23*, 777–784.
- (30) Pavel, E. G.; Zhou, J.; Busby, R. W.; Gunsior, M.; Townsend, C. A.; Solomon, E. I. Circular dichroism and magnetic circular dichroism spectroscopic studies of the non-heme ferrous active site in clavamate synthase and its interaction with α -ketoglutarate cosubstrate. *J. Am. Chem. Soc.* **1998**, *120*, 743–753.
- (31) Battye, T. G. G.; Kontogiannis, L.; Johnson, O.; Powell, H. R.; Leslie, A. G. iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 271–281, DOI: 10.1107/S0907444910048675.
- (32) Evans, P. R.; Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2013**, *69*, 1204–1214, DOI: 10.1107/S0907444913000061.
- (33) Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 213–221, DOI: 10.1107/S0907444909052925.
- (34) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and development of *Coot*. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 486–501, DOI: 10.1107/S0907444910007493.
- (35) Dennington, R.; Ketih, T. A.; Millam, J. M. *GaussView*, Version 6.1; Semichem Inc.: Shawnee Mission, KS, 2016.
- (36) Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T. E., III; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Greene, D.; Harris, R.; Homeyer, N.; Huang, Y.; Izadi, S.; Kovalenko, A.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D. J.; Merz, K. M.; Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Smith, J.; Salomon-Ferrer, R.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. *AMBER*; University of California: San Francisco, 2018.
- (37) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graphics Modell.* **2006**, *25*, 247–260, DOI: 10.1016/j.jmgm.2005.12.005.
- (38) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general AMBER force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.
- (39) Li, P.; Merz, K. M., Jr. MCPB.py: A python based metal center parameter builder. *J. Chem. Inf. Model.* **2016**, *56*, 599–604.
- (40) 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*, 3696–3713.
- (41) Davidchack, R. L.; Handel, R.; Tretyakov, M. V. Langevin thermostat for rigid body dynamics. *J. Chem. Phys.* **2009**, *130*, No. 234101.
- (42) 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*, 10089–10092.
- (43) 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*, 3084–3095.
- (44) 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*, 2695–2696.

(45) Miller, B. R., III; McGee, T. D., Jr.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E. *MMPBSA.py**: An efficient program for end-state free energy calculations. *J. Chem. Theory Comput.* **2012**, *8*, 3314–3321.

(46) Wang, E.; Sun, H.; Wang, J.; Wang, Z.; Liu, H.; Zhang, J. Z. H.; Hou, T. End-point binding free energy calculation with MM/PBSA and MM/GBSA: Strategies and applications in drug design. *Chem. Rev.* **2019**, *119*, 9478–9508.

(47) Balsera, M. A.; Wriggers, W.; Oono, Y.; Schulten, K. Principal component analysis and long time protein dynamics. *J. Phys. Chem. A* **1996**, *100*, 2567–2572.

(48) McPherson, M. J. Binding of benzoic acid and anions within the cupin domains of the vicilin protein canavalin from jack bean (*Canavalia ensiformis*): Crystal structures. *Biochem. Biophys. Res. Commun.* **2020**, *524*, 268–271.

(49) Xiong, W.; Morgan, J. A.; Ungerer, J.; Wang, B.; Maness, P.-C.; Yu, J. The plasticity of cyanobacterial metabolism supports direct CO₂ conversion to ethylene. *Nat. Plants* **2015**, *1*, No. 15053, DOI: 10.1038/nplants.2015.53.

(50) Lynch, S.; Eckert, C.; Yu, J.; Gill, R.; Maness, P.-C. Overcoming substrate limitations for improved production of ethylene in *E. coli*. *Biotechnol. Biofuels* **2016**, *9*, 3 DOI: 10.1186/s13068-015-0413-x.

(51) Zavr̃el, T.; Knoop, H.; Steuer, R.; Jones, P. R.; Cervený, J.; Trtiek, M. A quantitative evaluation of ethylene production in the recombinant cyanobacterium *Synechocystis* sp. PCC 6803 harboring the ethylene-forming enzyme by membrane inlet mass spectrometry. *Bioresour. Technol.* **2016**, *202*, 142–151.

(52) Zhu, T.; Xie, X.; Li, Z.; Tan, X.; Lu, X. Enhancing photosynthetic production of ethylene in genetically engineered *Synechocystis* sp. PCC 6803. *Green Chem.* **2015**, *17*, 421–434.

(53) Wang, B.; Xu, X.; Wang, X.; Yuan, J. S.; Johnson, C. H.; Young, J. D.; Yu, J. A guanidine-degrading enzyme controls genomic stability of ethylene-producing cyanobacteria. *Nat. Commun.* **2021**, *12*, No. 5150.