Supporting Information for

An Iron(IV)-oxo Intermediate Initiating L-Arginine Oxidation but not Ethylene Production by the 2-Oxoglutarate-dependent Oxygenase, Ethylene-Forming Enzyme

Rachelle A. Copeland,[†] Katherine M. Davis,[†,§] Tokufu Kent C. Shoda,[†] Elizabeth J. Blaesi,[†,&] Amie K. Boal,[†,⊥,*] Carsten Krebs,[†,⊥,*] J. Martin Bollinger, Jr.[†,⊥,*]

[†]Department of Chemistry and [⊥]Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, Pennsylvania 16802, United States

[§]Present Address: Department of Chemistry, Emory University, Atlanta, Georgia 30322, United States

[&]Present Address: GlaxoSmithKline, Collegeville, Pennsylvania, 19426, United States

[*]To whom correspondence should be addressed. Email: akb20@psu.edu, ckrebs@psu.edu, jmb21@psu.edu

Table of Contents

MATERIALS AND METHODS

Materials. 2-Oxoglutaric acid, 1,2,3,4,5-[$^{13}C_5$]-2-oxoglutaric acid ($^{13}C_5$-2OG), and 2,2,3,3-[$^{2}H_4$]-succinic acid (d_4-succinate) were purchased from Sigma-Aldrich (St. Louis, MO). Ethylene was obtained from Air Liquide America Specialty Gases (Plumsteadville, PA). L-Arginine was obtained from Alfa Aesar (Tewksbury, MA). 2,3,3,4,4,5,5-[$^{2}H_7$]-L-arginine (d_7-L-Arg) was purchased from Cambridge Isotope Laboratories, Inc. (Tewksbury, MA), and 2,5,5-[$^{2}H_3$]-L-proline (d_3-proline) was purchased from C/D/N Isotopes Inc (Point-Claire, Quebec). 5,5-d_2-L-Arginine was synthesized by a previously reported procedure.[1]

Preparation of EFE. A plasmid to direct over-expression of the ethylene-forming enzyme (EFE) from *Pseudomonas savastanoi pv. phaseolicola* (GenBank Accession D13182.1) was constructed. The plasmid specifies an N-terminal extension of 21 amino acids containing the His_6 metal-ion-affinity tag to facilitate purification. The gene sequence was codon optimized for expression in *Escherichia coli* (*Ec*), purchased from Life Technologies, and ligated between the *Nde*I and *Xho*I restriction sites of the pET-28a(+) vector. The full coding sequence is given below. To obtain untagged EFE for x-ray crystallography experiments, a second plasmid, encoding a fusion of the His_6-containing affinity tag, a SUMO tag, and a TEV protease cleavage site before the EFE sequence, was constructed. The coding sequence is given below. Variants of each construct encoding the D191E substitution were obtained via site-directed mutagenesis using Q5® High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA). All coding sequences were verified at the Genomics Core Facility at Penn State University.

His_6-EFE

ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAG
CCATATGATGACCAATCTGCAGACCTTTGAACTGCCGACCGAAGTTACCGGTTGTGC
AGCAGATATTAGCCTGGGTCGTGCACTGATTCAGGCATGGCAGAAAGATGGTATTTT
TCAGATTAAAACCGATAGCGAGCAGGATCGTAAAACCCAAGAAGCAATGGCAGCAA
GCAAACAGTTTTGTAAAGAACCGCTGACCTTTAAAAGCAGCTGTGTTAGCGATCTGA
CCTATAGCGGTTATGTTGCAAGCGGTGAAGAAGTGACCGCAGGTAAACCGGATTTTC
CTGAAATCTTTACCGTTTGCAAAGATCTGAGCGTTGGTGATCAGCGTGTTAAAGCAG
GTTGGCCGTGTCATGGTCCGGTTCCGTGGCCGAATAATACCTATCAGAAAAGCATGA
AAACCTTTATGGAAGAACTGGGTCTGGCAGGCGAACGTCTGCTGAAACTGACCGCA
CTGGGTTTTGAGCTGCCGATTAATACATTTACCGATCTGACACGTGATGGTTGGCAT
CACATGCGTGTTCTGCGTTTTCCGCCTCAGACCAGCACCCTGAGCCGTGGTATTGGT
GCACATACCGATTATGGTCTGCTGGTTATTGCAGCACAGGATGATGTTGGTGGTCTG
TATATTCGTCCGCCTGTTGAAGGTGAAAAACGTAATCGTAATTGGCTGCCTGGTGAA
AGCAGCGCAGGTATGTTTGAACATGATGAACCGTGGACCTTTGTTACCCCGACACCG
GGTGTTTGGACCGTTTTTCCGGGTGATATTCTGCAGTTTATGACCGGTGGTCAGCTGC
TGAGCACACCGCATAAAGTTAAACTGAATACCCGTGAACGTTTTGCCTGTGCCTATT
TTCATGAACCGAATTTTGAAGCAAGCGCCTATCCGCTGTTTGAACCGAGCGCAAATG
AACGTATTCATTATGGTGAACACTTTACCAACATGTTCATGCGTTGTTATCCGGATCG
TATTACCACCCAGCGTATTAACAAAGAAAATCGTCTGGCACATCTGGAAGATCTGAA
AAAATACAGCGATACCCGTGCAACCGGTAGC

His_6-SUMO-TEV-EFE

ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCATGTCGGACTCAGAAGT
CAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACATCA
ATTTAAAGGTGTCCGATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTC
CTTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAAGGAAATGGACTCC
TTAAGATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTG
GACATGGAGGATAACGATATTATTGAGGCTCACAGAGAACAGATTGGTGGGATCGA
GGAAAACCTGTACTTCCAATCCATGACCAATCTGCAGACCTTTGAACTGCCGACCGA
AGTTACCGGTTGTGCAGCAGATATTAGCCTGGGTCGTGCACTGATTCAGGCATGGCA
GAAAGATGGTATTTTTCAGATTAAAACCGATAGCGAGCAGGATCGTAAAACCCAAG
AAGCAATGGCAGCAAGCAAACAGTTTTGTAAAGAACCGCTGACCTTTAAAAGCAGC
TGTGTTAGCGATCTGACCTATAGCGGTTATGTTGCAAGCGGTGAAGAAGTGACCGCA
GGTAAACCGGATTTTCCTGAAATCTTTACCGTTTGCAAAGATCTGAGCGTTGGTGAT
CAGCGTGTTAAAGCAGGTTGGCCGTGTCATGGTCCGGTTCCGTGGCCGAATAATACC
TATCAGAAAAGCATGAAAACCTTTATGGAAGAACTGGGTCTGGCAGGCGAACGTCT
GCTGAAACTGACCGCACTGGGTTTTGAGCTGCCGATTAATACATTTACCGATCTGAC
ACGTGATGGTTGGCATCACATGCGTGTTCTGCGTTTTCCGCCTCAGACCAGCACCCT
GAGCCGTGGTATTGGTGCACATACCGATTATGGTCTGCTGGTTATTGCAGCACAGGA
TGATGTTGGTGGTCTGTATATTCGTCCGCCTGTTGAAGGTGAAAAACGTAATCGTAA
TTGGCTGCCTGGTGAAAGCAGCGCAGGTATGTTTGAACATGATGAACCGTGGACCTT
TGTTACCCCGACACCGGGTGTTTGGACCGTTTTTCCGGGTGATATTCTGCAGTTTATG

ACCGGTGGTCAGCTGCTGAGCACACCGCATAAAGTTAAACTGAATACCCGTGAACG
TTTTGCCTGTGCCTATTTTCATGAACCGAATTTTGAAGCAAGCGCCTATCCGCTGTTT
GAACCGAGCGCAAATGAACGTATTCATTATGGTGAACACTTTACCAACATGTTCATG
CGTTGTTATCCGGATCGTATTACCACCCAGCGTATTAACAAAGAAAATCGTCTGGCA
CATCTGGAAGATCTGAAAAAATACAGCGATACCCGTGCAACCGGTAGC

The sequences of the primers used to alter the codon for D191 to specify E were:

Forward: TGCACATACC**GAA**TATGGTCTGC
Reverse: CCAATACCACGGCTCAGG

Ec BL21(DE3) cells were transformed with an EFE over-expression plasmid and selected for kanamycin resistance. A single colony was used to inoculate 300 mL of rich LB medium (35 g/L tryptone, 20 g/L yeast extract, 5 g/L NaCl) containing 50 μg/mL of kanamycin, which was then incubated overnight at 37 °C with shaking. The following day, 25 mL of the overnight culture was used to inoculate 1 L of kanamycin-supplemented (50 mg/L) rich LB medium. Cultures were incubated at 37 °C until they reached an OD_{600} of 0.8, at which point isopropyl β-D-1-thiogalactopyranoside (IPTG) and 1,10-phenanthroline were added to final concentrations of 0.5 mM (each) to induce expression and limit Fe incorporation into the expressed protein, respectively. The culture was then incubated at 16 °C for an additional 16-18 h. Cells were harvested by centrifugation at 8000*g* at 4 °C, flash frozen in liquid N_2, and stored at -20 °C until being used.

Purification of EFE was carried out at 4 °C. In a typical purification, ~ 30 g cell paste was resuspended in 120 mL buffer A (30 mM Tris-HCl, pH 7.5, 150 mM NaCl, 10 mM imidazole, 0.5 mM DTT, and 5% glycerol). After addition of PMSF to a final concentration of 1 mM, the cells were lysed via sonication. The lysate was separated from cellular debris by centrifugation at 22,000*g* before being loaded onto Ni-NTA agarose resin (1 mL resin/6 mL lysate) in buffer A. After collection of the loading eluate, the resin was washed with buffer A, and EFE was then eluted

using buffer A containing 250 mM imidazole. When removal of the appended sequence from the His_6-SUMO-tagged EFE was required, the eluted protein was incubated overnight with TEV protease, the resulting solution was loaded onto Ni-NTA resin in buffer A, and the flow-through containing untagged EFE was collected. In either case, fractions containing the desired protein were identified by SDS PAGE with Coomassie blue staining, combined, concentrated using a 10K MWCO Macrosep® Advance Centrifugal Device (Pall Corporation, Port Washington, NY), and dialyzed overnight against buffer B (30 mM Tris-HCl, pH 7.5, 30 mM NaCl, 1 mM DTT, and 5% glycerol) containing 10 mM EDTA. The EDTA and protein contaminants were then removed by chromatography on a GE HiLoad 16/600 Superdex 200 pg column, with buffer B as eluant. Fractions containing EFE were combined, concentrated, flash frozen in liquid N_2, and stored at -80 °C. Protein purity was determined by SDS-PAGE with Coomassie blue staining, and concentration was determined using a calculated molar absorption coefficient of 53,400 M^{-1} cm^{-1} at 280 nm according to the method of Gill and von Hippel.[2] EFE was made anoxic prior to use in stopped-flow absorption, freeze-quench Mössbauer, or crystallization experiments.

Stopped-flow Absorption Experiments. Stopped-flow absorption (SF-Abs) experiments were performed in an MBraun (Stratham, NH) LABmaster anoxic chamber and using an Applied Photophysics (Leatherhead, Surrey, UK) SX20 stopped-flow spectrophotometer configured for single mixing and a 1-cm optical path length. An O_2-free solution of EFE, Fe(II), 2OG and L-arginine in buffer (40 mM sodium HEPES, pH 7.5, with 5% glycerol) was mixed with an equal volume of the same buffer containing O_2 at 5 °C. Full absorption spectra or single wavelength absorbance data were recorded with either a photodiode array (PDA) detector with polychromatic light or a photomultiplier tube (PMT) detector with monochromatic light, respectively. Absorbance-versus-time data were simulated using KinTek Explorer chemical kinetics software

(KinTek Corporation, Snow Shoe, PA). The protein, cofactor, and substrate concentrations are given in the figure legends.

Freeze-quench ^{57}Fe Mössbauer Experiments. Freeze-quench samples were prepared using an Update Instrument (Madison, WI) freeze/chemical quench device. An O_2-free solution containing EFE, ^{57}Fe(II), 2OG and d_7-L-Arg in 40 mM sodium HEPES buffer (pH 7.5) with 5% glycerol was mixed at 5 °C with an equal volume of O_2-saturated buffer. The solution was allowed to react for the period of time determined by the flow velocity and volume of the reaction hose, and the reaction was terminated by injection into cold (~ -150 °C) 2-methylbutane. Control samples were obtained by mixing the O_2-free protein solution with an equal volume of O_2-free buffer and manually freezing anoxically. ^{57}Fe Mössbauer spectra were recorded at 4.2 K on a SEE Co. (Minneapolis, MN) instrument equipped with a Janis Research Co. (Woburn, MA) cryostat, as described previously.[3] Spectra were recorder either with no external magnetic field or with a 53-mT field applied parallel to the γ beam. Spectra were simulated using WMOSS Mössbauer spectral analysis software.

Enzyme Assays. Assays were carried out in duplicate at 21 °C. In the standard 1 mL reaction, an air-saturated solution of 2OG and L-arginine in 40 mM sodium HEPES buffer (pH 7.5) was added to a 12 mL vial equipped with a magnetic stir bar. The vial was then capped, and the reaction was initiated by injecting a solution of EFE and Fe(II) through the septum of the cap using a Hamilton gastight syringe. After ~10 min, 0.10 mL of 10% formic acid was injected to terminate the reaction. To determine the yield of ethylene, the quenched reaction was allowed to come to equilibrium with the headspace before 2 mL of headspace gas was withdrawn for GC analysis. After GC data were collected, the vial was opened, and 16 μL of a solution containing 250 mM

$NaCNBH_3$ [used to reduce P5C to proline and 2OG to 2-hydroxyglutarate (2HG)] and 12.5 mM each of $^{13}C_5$-2OG, d_4-succinate and d_3-proline (internal standards) was added to the reaction solution. The solution was then filtered through a 10K MWCO Nanosep® Centrifugal Device (Pall Corporation, Port Washington, NY) to remove denatured protein. To convert proline to Fmoc-proline prior to LC-MS analysis, an aliquot of the filtered reaction was mixed with two equivalent volumes of 100 mM Fmoc-Cl in acetonitrile. The remaining filtered reaction was used directly for succinate and 2HG quantification.

GC-FID and LC-MS Analysis. Ethylene was detected using an HP5890 Series II gas chromatograph equipped with a flame ionization detector. Chromatographic separation from other headspace components was accomplished using an Agilent Porapak Q column (0.125 in x 6 ft), and quantification was based on an external ethylene calibration standard.[4] LC-MS determination of the other analytes was performed using an Agilent 1200 series LC system coupled to an Agilent 6410 triple quadrupole mass spectrometer. Table S3 shows the *m/z* values for each analyte. Chromatographic separation was achieved on an Agilent ZORBAX Extend-C18 column (4.6 x 50 mm, 1.8 μm particle size) using a mixture of 0.1% formic acid in water (solvent A) and acetonitrile (solvent B) as mobile phase. Succinate and 2HG were eluted isocratically by 95% solvent A and 5% solvent B at a flow rate of 0.4 mL/min for 5 min. Fmoc-proline was eluted isocratically by 50% solvent A and 50% solvent B at 0.3 mL/min for 8 min.

Protein Crystallization. Crystallization of EFE was accomplished in an anoxic glove box (Coy Laboratory Products) containing an atmosphere of approximately 97% N_2 and 3% H_2. All solutions were rendered anoxic prior to use. To generate the reactant complex (EFE•Fe^{II}•L-Arg•2OG) in solution, a stock solution of untagged EFE at 42 mg/mL was first diluted in 30 mM Tris-HCl buffer (pH 7.5) containing 30 mM NaCl and 5% glycerol. 2OG, L-Arg, and

$Fe(NH_4)_2(SO_4)_2$ were then added to yield a final protein concentration of 10 mg/mL (0.25 mM), with 4 molar equivalents of Fe and 40 molar equivalents of the substrates. The resulting solution was incubated for a minimum of 30 min at 5 °C before being mixed with one or two equivalent volumes of a precipitant solution containing 0.1 M bis-tris propane (pH 6.4-6.7), 0.1-0.2 M NaCl, and 26-30% (w/v) PEG 6000. Drops were allowed to come to equilibrium with a 500 μL reservoir of precipitant solution in a hanging-drop vapor-diffusion experiment. Thin, rod-like crystals appeared within two days and grew to maximum size over the following few days. Upon looping, the crystals were soaked in a cryo-protectant generated by adding 25% (v/v) ethylene glycol to the reservoir solution before they were flash frozen in liquid N_2.

X-ray Diffraction Data Collection and Processing. Diffraction data for the wild type (wt) EFE reactant complex were collected at beamline 21-ID-F of the Advanced Photon Source at Argonne National Laboratory on a Rayonix MX300 CCD. Images were collected sequentially ($\Delta\phi$ = 0.5°) at an incident wavelength of 0.97872 Å. Diffraction data for the D191E variant were collected at the Advanced Light Source at Lawrence Berkeley National Laboratory on an ADSC Q315r CCD at beamline 8.2.2. These images were also collected sequentially with $\Delta\phi$ = 1° at an incident wavelength of 1.00001 Å. In all cases, the crystals were maintained at 100 K to minimize x-ray-induced damage. Indexing, merging and scaling of the data were carried out in the HKL2000 software package.[5] The structure of the reactant complex of the wt enzyme was solved via molecular replacement with PHASER using a previously-determined structure of EFE (PDB accession code 5LSQ) as the initial search model.[6,7] Subsequent model building and refinement were performed using Coot[8] and the Phenix software suite,[9] respectively. As the other structures are clearly isomorphous, this model was refined against all new data, maintaining the same R_{free} flags, while PDB_REDO was used to confirm an unbiased test set.[10]

All structures were solved in the $P12_11$ space group and contain four molecules in the asymmetric unit. Residues not modeled due to disorder are listed in Table S2. Although the metal-coordinating residues of the wt structure refined nicely in position, restraints enforcing bond lengths of 2.1 ± 0.15 Å helped clean up difference density surrounding the iron cofactor. Due to potential perturbations to the Fe coordination sphere in the D191E variant, restraints on the binding mode of the facial triad were not enforced. The final model for the wt reactant complex (PDB accession code: 6VP4) was refined to 1.83 Å resolution and contains one iron ion, one molecule of 2OG, and one molecule of L-Arg per monomer. The final model for the D191E reactant complex (PDB accession code: 6VP5) was refined to 1.97 Å resolution and contains one iron ion, one molecule of 2OG, and one molecule of L-Arg per monomer. Selected data processing and refinement statistics are presented in Table S1.

Simulation of Stopped-Flow Absorption Data. We simulated the single-wavelength absorbance-versus-time traces from SF-Abs experiments on wt and D191E EFE with L-Arg and d_7-L-Arg according to two branched-pathway kinetic models. The models differ in terms of the point at which the two pathways bifurcate. That they account for the kinetic data equally well is a reflection of the fact that intermediate states expected to be on one or both pathways (e.g., the initial Fe(III)–$O_2^{-\bullet}$ adduct) do not accumulate, thus obscuring steps of both pathways that are reasonable or necessary to invoke to account for the known outcomes. As such, neither mechanism can be taken as a validated model for the overall EFE reaction; the available data simply do not allow a distinction between them to be made. The point of the analysis was to quantitatively test the notion, apparent from the semi-quantitative considerations presented in the main text, that the differing behavior of the wt and D191E EFE proteins can be accounted for by differences in the partition ratios between the two pathways and the fact that the ferryl intermediate that dominates

the absorbance-versus-time traces (at least at 318 nm) does not form in one of the two pathways (EF).

In the first kinetic model (Scheme S1), two rapidly-equilibrating conformations of the EFE•Fe^{II}•2OG•L-Arg reactant complex (R_1 and R_2) independently undergo irreversible, bimolecular reactions with O_2. Addition of O_2 to R_1 forms intermediate I_1, which is converted to product state P_1 along the EF pathway, whereas addition of O_2 to R_2 forms ferryl intermediate I_2 en route to product state P_2 along the RO pathway. In such a branched reaction, the relative magnitudes of the rates of the competing steps are reflected in the final quantities of the products of each pathway. In this model, competing O_2 addition steps are linked via the reversible equilibration of the reactant complexes. Therefore, in fitting the data to the model, k_1 and k_{-1}, the forward and reverse rate constants for the rapid interconversion of the reactant complex conformations, were made large, equal and fixed, and k_2 and k_5, the second-order rate constants for the bimolecular reactions with O_2 along the RO and EF branches, respectively, were covaried in ratios defined by the relative yields of succinate and ethylene (1:2.5 for wt and 19:1 for D191E). Alternatively, k_1 and k_{-1} could have been made different from one another to give an equilibrium distribution of reactant states equivalent to the final [succinate]:[ethylene] ratio, with k_2 and k_5 then held equal.

The second kinetic model that we evaluated (Scheme S2) involves a single bimolecular reaction of O_2 with reactant complex (R′) to generate an intermediate (I_1′) that is the branchpoint of the RO and EF pathways. This intermediate, which need not accumulate to detectable levels, decays to form a product state (P_1′) along the EF pathway and a ferryl intermediate (I_2′), which is converted to product state (P_2′), along the RO pathway. The first-order rate constants for competing decay of I_1′ to form I_2′ along the RO branch (k_2) and P_1′ along the EF branch (k_5) were co-varied

in a ratio reflecting the final [succinate]:[ethylene] yield, as described above. Results of biochemical assays reveal that deuterium substitution at C5 of L-Arg does not alter the relative final concentrations of succinate and ethylene, thereby obviating examination of a third model, in which the ferryl intermediate decays by a competition between HAT to effect L-Arg oxidation and succinate oxidation to yield ethylene (Scheme 2).

In fitting each of the two kinetic models considered, rates for all steps except decay of the RO ferryl intermediate (k_3 in Schemes S1 and S2) were constrained to be identical for both L-Arg isotopologues with a given enzyme. To determine whether the observed spectroscopic differences between the wt and D191E reactions are adequately explained by pathway partitioning, we constrained the molar absorptivity (ε) for each absorbing species to be identical for both enzymes. Initial estimates of ε for the reactant and ferryl complexes were guided by previously reported values, and product complexes were assumed to be non-absorbing.[3,11] For the model depicted in Scheme S1, in which pathway bifurcation occurs at the reactant state, we considered scenarios in which ε_{318} and ε_{510} for the intermediate formed upon O_2 addition along the EF pathway (I_1) were either variable or fixed at zero. Likewise, we considered both possibilities for the branchpoint intermediate (I_1'), which forms upon O_2 addition in the alternative model (Scheme S2).

As demonstrated by the solid lines in Figure 2 and Figure S2, both kinetic models can account for the absorbance-versus-time traces (at 318 and 510 nm) reasonably well. The ε_{510} values obtained for the reactant and ferryl complexes (~90 M^{-1} cm^{-1} and ~200 M^{-1} cm^{-1}, respectively) are in reasonable agreement with previous reports, as are the ε_{318} values of ~1500-2000 M^{-1} cm^{-1} for the ferryl complexes (Tables S4 and S6). The small ε_{318} values for the reactant complexes (~20-40 M^{-1} cm^{-1}) are consistent with the gain of absorbance upon conversion of product complex(es) to reactant complex(es) observed here and with other Fe/2OG enzymes.[12] Although we report

positive ε_{318} and ε_{510} values for the EF intermediate (I_1) in Scheme S1 and branchpoint intermediate (I_1') in Scheme S2 in Tables S4 and S6, respectively, satisfactory fits could also be obtained when these parameters were fixed at zero. The realization that these intermediates could contribute either modestly or negligibly to the absorbance changes – and, moreover, the more fundamental fact that two different kinetic models can suitably account for the data – again highlights the inherent limitations in the information content of the data.

The ε_{318} values obtained for the ferryl complex upon fitting the SF-Abs data reasonably predict the amount of ferryl observed in the Mössbauer results. The most important consequence of fitting the SF-Abs data using a global set of ε values for wt and D191E and without having to invoke a significantly smaller active fraction of wt enzyme relative to D191E is the demonstration that the diminished amplitudes of ΔA_{318} for the wt in comparison to D191E enzymes are attributable to kinetic pathway partitioning, such that flux through the RO pathway changes from only ~ 30% of the total in the wt enzyme to 95% in D191E EFE.

Simulation of Multiple-Turnover Reactions of wild-type EFE. The invariable ~2.5:1 ethylene-to-P5C (or succinate) product ratio observed upon complete conversion of 2OG under multiple-turnover conditions indicates that EF:RO pathway partitioning is insensitive to the presence of deuterium at C5 of L-Arg and thus precedes the ferryl complex. This conclusion is corroborated in the single-turnover regime by Mössbauer results that show simultaneous formation of an Fe(II) product state of the EF pathway and the RO ferryl intermediate in a similar ratio when EFE is reacted with deuterated L-Arg. What might not immediately be obvious from both sets of results, however, is the fact that during the transient state, the product partition ratio will evolve and can be sensitive to a D-KIE on HAT if the products being considered are not formed concomitantly, as is the case for ethylene and P5C. To illustrate this point more clearly, we

simulated multiple-turnover reactions according to the kinetic model in which a ferryl precursor is the branchpoint of the EF and RO pathways (Scheme S2) and probed the effects of varying k_{HAT} (Figure S7). In each cycle, ~ 70% of active sites form the EF products while the other ~ 30% form the RO ferryl complex. At smaller values of k_{HAT} (denoted as k_3 in Scheme S2), the rate at which the enzyme cycles through the RO pathway becomes smaller, resulting in the accumulation of a larger fraction of enzyme in the RO ferryl state (Figure S7, *panel A*). Because ethylene forms *before* and P5C forms *after* the isotope-sensitive HAT step, the product partition ratio, when expressed as [ethylene]/([ethylene] + [P5C]) (*panel C, solid lines*), is close to unity at early time points, and evolves to a final value of ~ 0.7; correspondingly, [P5C]/([ethylene] + [P5C]) (*panel C, dashed lines*) evolves from almost 0 to ~ 0.3 (Figure S7). These partition ratios do not reach their final values until all of the ferryl complex has decayed. This fact implies that, if the D-KIE were large enough (e.g. compare *red* and *green* traces in *Panel C*), different partition ratios could be observed with unlabeled L-Arg (*red traces*) and deuterated L-Arg (*green traces*) upon quenching the reactions before all of the 2OG was converted and all the ferryl complex decayed. In contrast, because ethylene and succinate are both formed before the isotope-sensitive step (and concomitantly at the branchpoint in this model), the [ethylene]/([ethylene] + [succinate]) partition ratio and the corresponding [succinate]/([ethylene] + [succinate]) ratio (*solid* and *dotted lines* in inset of *Panel C*) have constant values of ~ 0.7 and ~ 0.3, respectively, throughout the transient and steady states and are unaffected by any D-KIE.

Table S1. Crystallographic data processing and refinement statistics.

	wild type	D191E
PDB ID	**6VP4**	**6VP5**
Data Collection[a]		
Space group	$P12_11$	$P12_11$
Unit cell (Å)	a = 49.73, b = 78.81, c = 195.91 $\alpha = \gamma = 90, \beta = 92.67$	a = 49.54, b = 78.50, c = 195.49 $\alpha = \gamma = 90, \beta = 92.34$
Wavelength (Å)	0.97872	1.00001
Resolution range (Å)	40.76 – 1.83 (1.89 – 1.83)	49.52 – 1.97 (2.04 – 1.97)
Total observations	912947 (86734)	790232 (75007)
Total unique observations	132130 (12759)	105982 (9842)
I/σ_I	18.82 (2.52)	16.65 (2.38)
Completeness (%)	98.74 (96.18)	99.21 (93.79)
R_{merge}	0.07705 (0.7406)	0.1027 (0.6984)
R_{pim}	0.03098 (0.3007)	0.04034 (0.2801)
Redundancy	6.9 (6.8)	7.5 (7.1)
Refinement Statistics		
Resolution range (Å)	40.76 – 1.83	49.5 – 1.97
Reflections (total)	132010 (12753)	105152 (9841)
Reflections (test)	2005 (201)	1592 (142)
Total atoms refined	11761	11986
Solvent	1107	1251
R_{work} (R_{free})	17.44 (19.73)	17.36 (20.64)
rmsd		
Bond lengths (Å)/ angles (°)	0.006/0.73	0.003/0.63
Ramachandran plot		
Favored/allowed (%)	97.90/1.65	97.97/1.58
Mean B values (Å^2)		
Protein chains A/B/C/D	24.82/21.62/25.60/25.47	18.44/17.37/20.02/18.04
Fe A/B/C/D	18.06/15.74/19.29/18.72	11.55/10.35/15.23/9.69
L-Arg A/B/C/D	21.86/17.60/21.54/22.36	16.28/14.74/18.08/13.87
2OG A/B/C/D	23.45/21.13/24.75/25.33	22.18/20.15/23.02/21.35
NHA A/B/C/D	-	-
Solvent	32.79	28.32

[a] Values in parentheses refer to the high-resolution shell.

Table S2. Missing residues for each structure.

PDB ID	wild type 6VP4	D191E 6VP5
Chain A	334-350	299, 337-350
Chain B	341-350	343-350
Chain C	1-2, 299, 343-350	1-2, 298-300, 343-350
Chain D	1-2, 334-350	1-2, 299, 335-350

Table S3. Target *m/z* ratios of analytes detected via electrospray ionization and single-ion monitoring.

Compound	***m/z***	**Ion Mode**
succinate	117	neg
d_4-succinate	121	neg
2HG	147	neg
$^{13}C_5$-2HG	152	neg
Fmoc-proline	338	pos
d_2-Fmoc-proline	340	pos
d_3-Fmoc-proline	341	pos
d_6-Fmoc-proline	344	pos

Table S4. Molar absorptivities (ε) obtained upon global fitting of SF-Abs data shown in Figure 2 according to the kinetic model depicted in Scheme S1, in which EF and RO pathways branch at the reactant state.

State	ε_{318nm} (M^{-1} cm^{-1})	ε_{510nm} (M^{-1} cm^{-1})
R_1, R_2	39	90
I_1	101	37
I_2	1,530	188
P_1 [a]	0	0
P_2 [a]	0	0

[a]Values were fixed prior to fitting.

Table S5. Rate constants obtained upon global fitting of SF-Abs data shown in Figure 2 according to the kinetic model depicted in Scheme S1, in which EF and RO pathways branch at the reactant state.

Rate Constant	Wild Type	D191E
k_1[a], k_{-1}[a]	10,000 s^{-1}	10,000 s^{-1}
k_2[b]	7,600 M^{-1} s^{-1}	50,500 M^{-1} s^{-1}
k_3	27 s^{-1} *(L-Arg)* 1 s^{-1} *(d_7-L-Arg)*	47 s^{-1} *(L-Arg)* 1.6 s^{-1} *(d_7-L-Arg)*
k_4	0.6 s^{-1}	0.3 s^{-1}
k_5[b]	19,100 M^{-1} s^{-1}	2,700 M^{-1} s^{-1}
k_6	1.3 s^{-1}	20 s^{-1}
k_7	1.2 s^{-1}	0.04 s^{-1}

[a]Values were fixed prior to fitting.
[b]Values were covaried in a defined ratio.

Table S6. Molar absorptivities obtained upon global fitting of SF-Abs data shown in Figure S2 according to the kinetic model depicted in Scheme S2, in which an intermediate complex is the branchpoint of the EF and RO pathways.

State	ε_{318nm} (M^{-1} cm^{-1})	ε_{510nm} (M^{-1} cm^{-1})
R'	27	94
I_1'	1190	128
I_2'	1550	196
$P_1'^{a}$	0	0
$P_2'^{a}$	0	0

[a]Values were fixed prior to fitting.

Table S7. Rate constants obtained upon global fitting of SF-Abs data shown in Figure S2 according to the kinetic model depicted in Scheme S2, in which an intermediate complex is the branchpoint of the EF and RO pathways.

Rate Constant	Wild Type	D191E
k_1	6,600 M^{-1} s^{-1}	27,900 M^{-1} s^{-1}
k_2[a]	10 s^{-1}	288 s^{-1}
k_3	42 s^{-1} *(L-Arg)* 1.3 s^{-1} *(d_7-L-Arg)*	53 s^{-1} *(L-Arg)* 1.6 s^{-1} *(d_7-L-Arg)*
k_4	0.8 s^{-1}	0.3 s^{-1}
k_5[a]	25 s^{-1}	15 s^{-1}
k_6	0.9 s^{-1}	0.04 s^{-1}

[a]Values were covaried in a defined ratio.

Table S8. Parameters used to simulate the quadrupole doublet components of the zero-field 340-ms-*minus*-anoxic-control (wt) and 140-ms-*minus*-anoxic-control (D191E) Mössbauer difference spectra.

	Species	**δ (mm/s)**	**$\lvert\Delta E_Q\rvert$ (mm/s)**	**Γ (mm/s)**
wt	Reactant complex	1.10	2.47	0.24
wt	Fe(IV)=O	0.26	0.96	0.25
wt	EF complex	1.17	3.10	0.54
D191E	Reactant complex	1.17	3.38	0.26
D191E	Fe(IV)=O	0.32	0.88	0.27

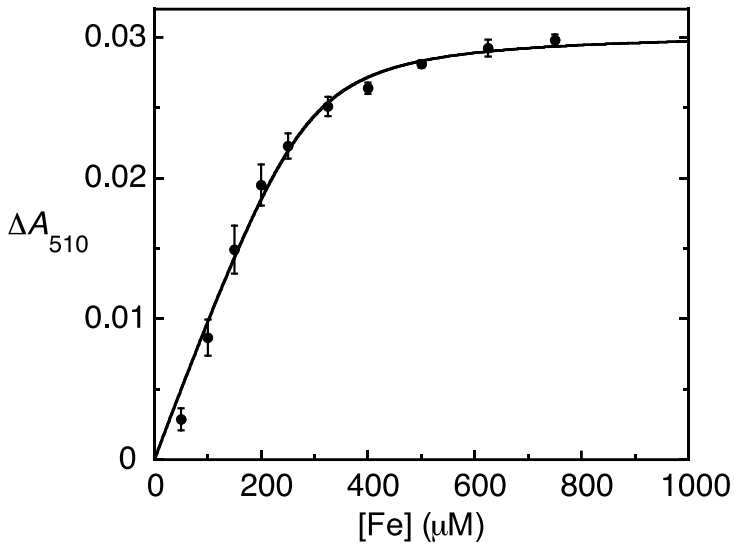


Figure S1. Amplitude-versus-concentration data obtained upon anoxic titration of Fe(II) into a solution containing wild-type EFE, 2OG and L-Arg at 5 °C. The absorbance change at 510 nm was measured and amplitudes fit to the equation $\Delta A = \Delta A_{max} \cdot \frac{(K_d+E+S)-\sqrt{(K_d+E+S)^2+4\cdot E\cdot S}}{2\cdot E}$. The points plotted represent averages of 2 or 3 trials. Final concentrations: 0.25 mM EFE, 1 mM 2OG, 1 mM L-Arg, and 0.05 to 0.75 mM Fe(II). K_d = 17 ± 12 μM, ε_{510} = 110 ± 10 $M^{-1}cm^{-1}$.

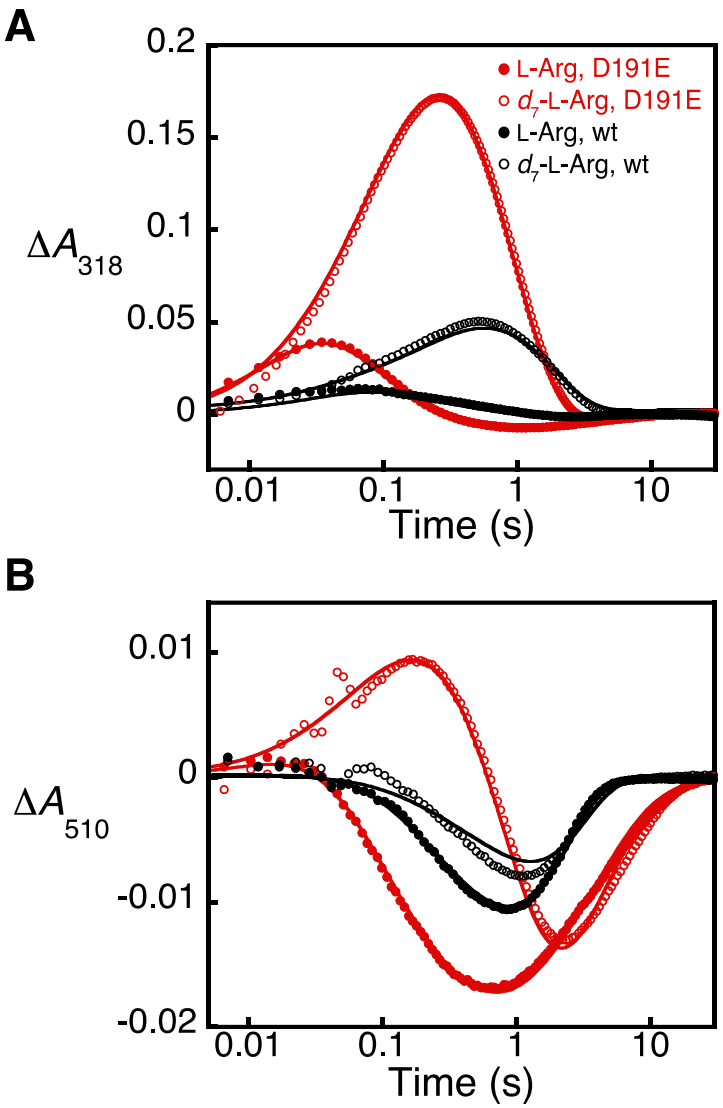


Figure S2. Stopped-flow kinetic traces at 318 nm (***A***) and 510 nm (***B***) obtained upon mixing of an anoxic solution containing 0.7 mM wt (*black*) or D191E (*red*) EFE, 0.6 mM Fe(II), 8 mM 2OG, and 8 mM L-Arg (*solid points*) or d_7-L-Arg (*open points*) with an equal volume of 40% room-temperature O_2-saturated buffer (O_2 concentration of ~ 0.48 mM) at 5 °C. The solid lines are the results of global fitting of the data to a kinetic model in which the reaction pathway bifurcates via decay of an intermediate, as shown in Scheme S2 and described above.

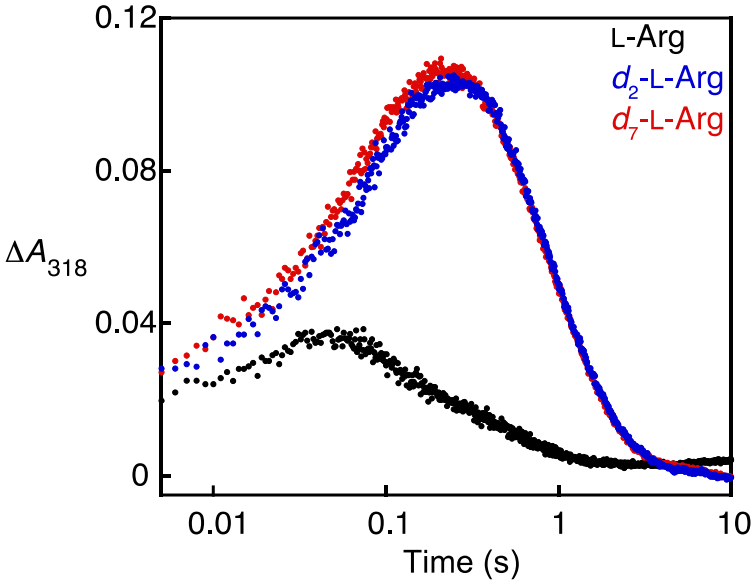


Figure S3. The effect of deuteration of either only C5 or C2-C5 on kinetics of decay of the ferryl intermediate monitored by its absorption at 318 nm. In the stopped-flow experiment, O_2-saturated buffer (~1.8 mM O_2) was mixed with an equal volume of an anoxic solution containing 1 mM EFE, 0.8 mM Fe(II), 0.5 mM 2OG, and 5 mM of unlabeled L-Arg (*black circles*), perdeuterated L-Arginine (d_7-L-Arg, *red circles*) or 5,5-d_2-L-Arginine (d_2-L-Arg, *blue circles*). Reactions were carried out at 5 °C and monitored with the PDA detector.

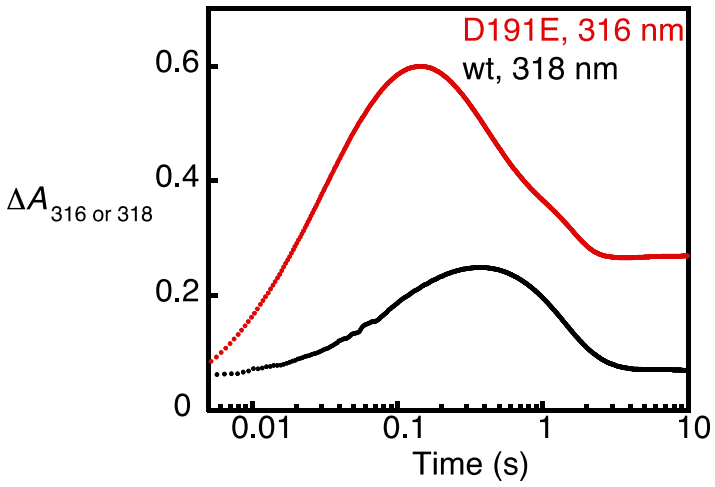


Figure S4. Single-wavelength Δ*A*-versus-time data from stopped-flow experiments monitoring the reaction of EFE•Fe^{II}•2OG•d_7-L-Arg with O_2 under conditions corresponding to the FQ-Möss experiments (1.8 mM EFE, 1.4 mM Fe(II), 5 mM 2OG, 5 mM *d*$_7$-L-Arg in 40 mM HEPES, pH 7.5, mixed with an equal volume of buffer containing 1.8 mM O_2. Wild-type EFE was monitored at 318 nm (*black circles*) and D191E EFE was monitored at 316 nm (*red circles*). Reactions were carried out at 5 °C and monitored at using the PMT detector.

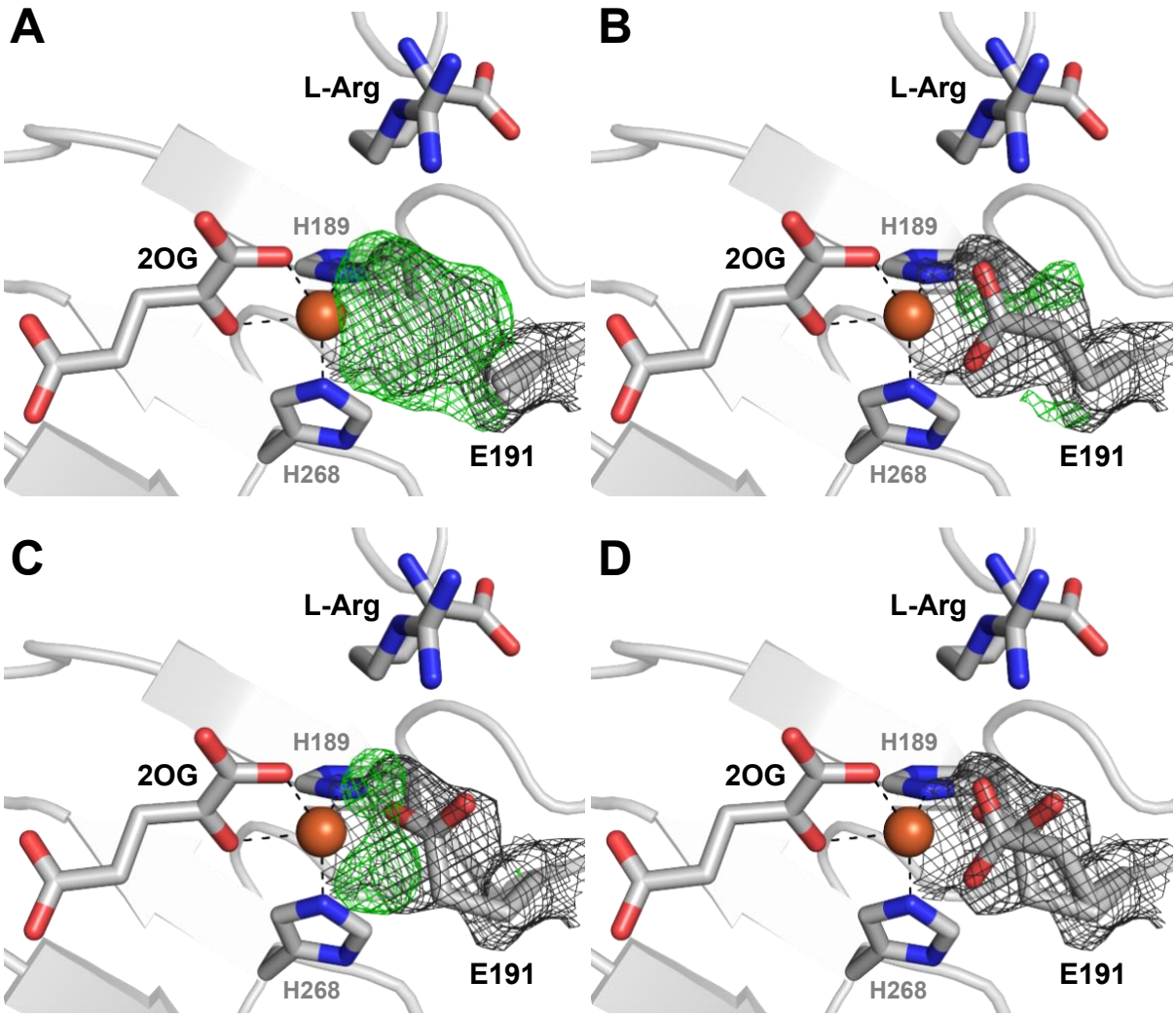


Figure S5. Modeling of the observed electron density for E191 justifies the presence of two side chain conformations. (***A***) As a control, residue 191 was first modeled as alanine. (***B***) E191 was then modeled in a similar orientation to the wild type aspartate. Positive difference density suggests the presence of an alternate conformation. (***C***) Modeling the alternate conformation of E191 alone likewise cannot account for observed density. (***D***) Both conformations of E191 modeled together provides the best agreement with the data. In all panels, $2F_o$-F_c maps are shown in black and contoured to 1.0σ, while F_o-F_c maps are shown in green/red and are contoured to ±3.0σ.

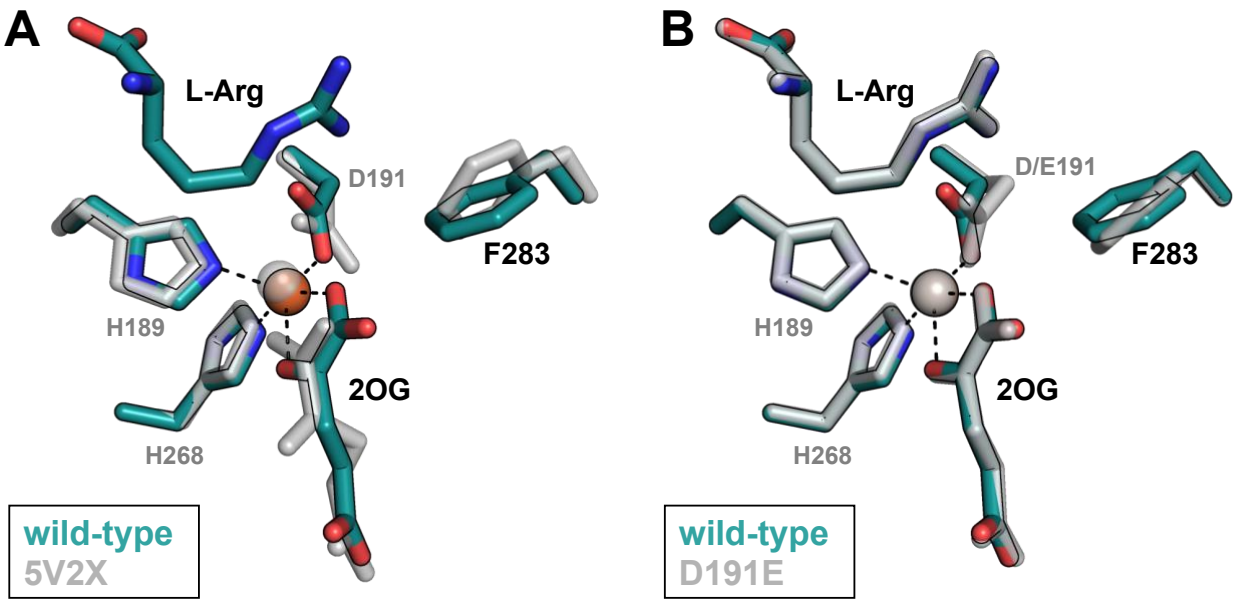


Figure S6. Active-site comparisons of wild-type EFE•Fe•2OG•L-Arg with (***A***) EFE•Mn•2OG (PDB: 5V2X) and (***B***) D191E EFE•Fe•2OG•L-Arg display minor changes in the position of F283 upon L-Arg binding (***A***) and substitution of the Asp ligand with Glu (***B***). For clarity, the 2OG conformation in which C5 ligates the Fe center is omitted from Panel A, and only conformation A of D191E is depicted in Panel B.

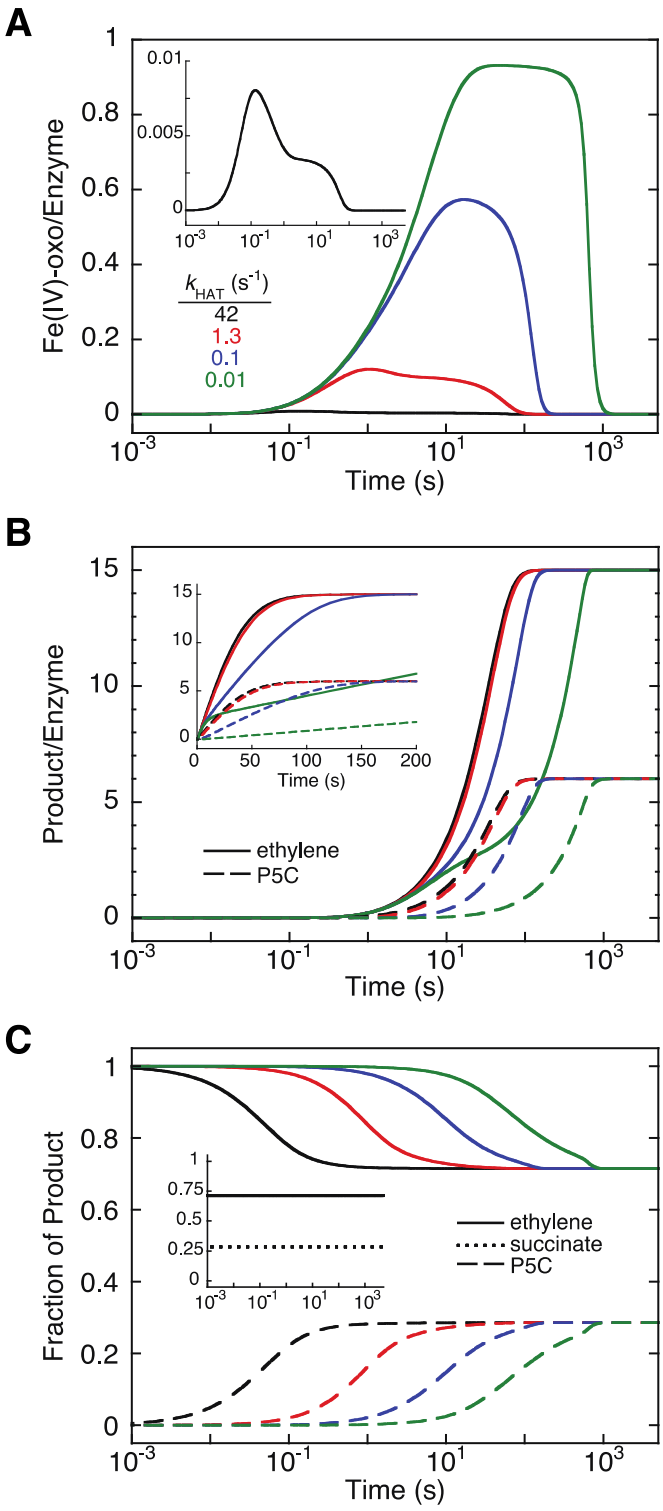
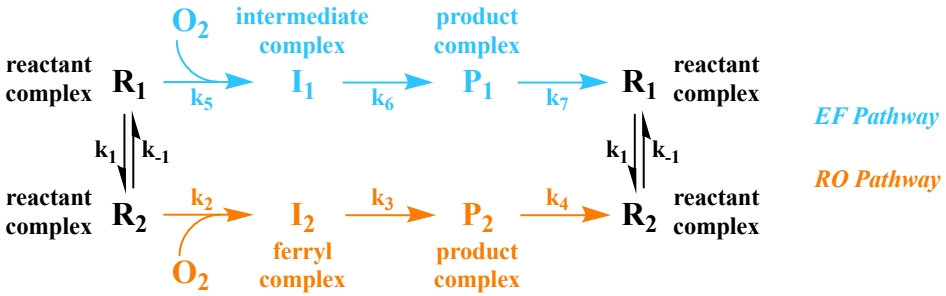
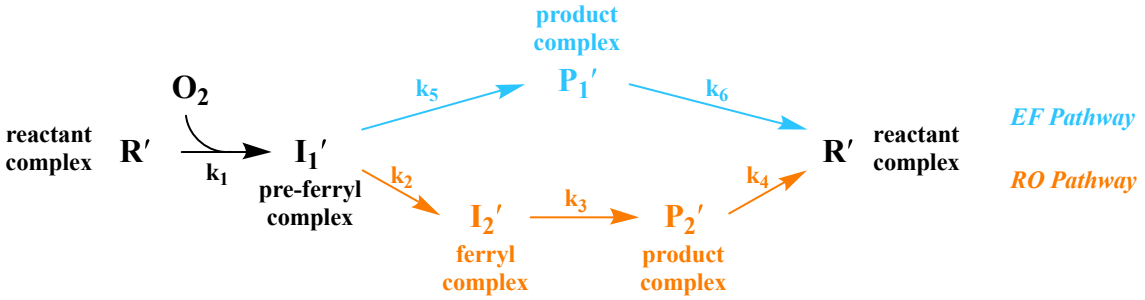


Figure S7. Simulations of multiple-turnover reactions of wild-type EFE according to the kinetic model shown in Scheme S2, using different rate constants for ferryl decay via HAT from L-Arg (k_{HAT}). The two largest k_{HAT} values, 42 s^{-1} (*black lines*) and 1.3 s^{-1} (*red lines*), are obtained – along

with rates for other steps in the model – from the global simulation analysis of single-turnover SF-Abs data shown in Table S7 and Figure S2. The smaller values of 0.1 s^{-1} (*blue lines*) and 0.01 s^{-1} (*green lines*) are hypothetical values that correspond to increasingly large D-KIEs on HAT. Panel A shows the fraction of enzyme in the RO ferryl state over time, with the black trace – corresponding to the largest k_{HAT}, 42 s^{-1} – magnified in the inset. Panel B and its inset illustrate progress curves for products of each pathway, ethylene (*solid lines*) and P5C (*dashed lines*), on logarithmic and linear time scales. Panel C shows pathway partitioning as a function of time. The partitioning is shown in terms of ethylene (*solid lines*) and P5C (*dashed lines*) on the main *y*-axis and in terms of ethylene (*solid line*) and succinate (*dotted line*) in the inset. Concentrations of EFE and substrates (corresponding to reactant complex and O_2, respectively, in Scheme S2) are 10 μM and 210 μM, respectively.



Scheme S1. Kinetic model in which EF and RO pathways branch at the reactant state. Simulations of SF-Abs data using this model are shown in Figure 2 and described above, and the associated ε values and rate constants are given in Tables S4 and S5.



Scheme S2. Alternative kinetic model in which an intermediate complex is the branchpoint of the EF and RO pathways. Simulations of SF-Abs data using this model are shown in Figure S2, and the corresponding ε values and rate constants are given in Tables S6 and S7, respectively.

Supporting References

1. Dunham, N. P.; Chang, W. C.; Mitchell, A. J.; Martinie, R. J.; Zhang, B.; Bergman, J. A.; Rajakovich, L. J.; Wang, B.; Silakov, A.; Krebs, C.; Boal, A. K.; Bollinger, J. M., Two Distinct Mechanisms for C-C Desaturation by Iron(II)- and 2-(Oxo)glutarate-Dependent Oxygenases: Importance of α-Heteroatom Assistance. *Journal of the American Chemical Society* **2018,** *140* (23), 7116-7126.

2. Gill, S. C.; von Hippel, P. H., Calculation of protein extinction coefficients from amino acid sequence data. *Analytical Biochemistry* **1989,** *182* (2), 319-326.

3. Price, J. C.; Barr, E. W.; Tirupati, B.; Bollinger, J. M.; Krebs, C., The first direct characterization of a high-valent iron intermediate in the reaction of an α-ketoglutarate-dependent dioxygenase: A high-spin Fe(IV) complex in taurine/α-ketoglutarate dioxygenase (TauD) from Escherichia coli. *Biochemistry* **2003,** *42* (24), 7497-7508.

4. Kampbell, D. H.; Vandegrift, S. A., Analysis of Dissolved Methane, Ethane, and Ethylene in Ground Water by a Standard Gas Chromatographic Technique. *Journal of Chromatographic Science* **1998,** *36* (5), 253-256.

5. Otwinowski, Z.; Minor, W., [20] Processing of X-ray diffraction data collected in oscillation mode. In *Method Enzymol.*, Academic Press: 1997; Vol. 276, pp 307-326.

6. McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J., Phaser crystallographic software. *J. Appl. Crystallogr.* **2007,** *40* (Pt 4), 658-674.

7. 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. *P Natl Acad Sci USA* **2017,** *114* (18), 4667-4672.

8. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., Features and development of Coot. *Acta Crystallogr. D* **2010,** *66* (4), 486-501.

9. 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. D* **2010,** *66* (Pt 2), 213-221.

10. Joosten, R. P.; Joosten, K.; Murshudov, G. N.; Perrakis, A., PDB_REDO: constructive validation, more than just looking for errors. *Acta Crystallogr. D* **2012,** *68* (4), 484-496.

11. 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.

12. Dunham, N. P.; Mitchell, A. J.; Del Río Pantoja, J. M.; Krebs, C.; Bollinger, J. M.; Boal, A. K., α-Amine Desaturation of D-Arginine by the Iron(II)- and 2-(Oxo)glutarate-Dependent L-Arginine 3-Hydroxylase, VioC. *Biochemistry* **2018,** *57* (46), 6479-6488.