Supporting Information

Structures and Mechanisms of the Non-Heme Fe(II)- and 2-Oxoglutarate-Dependent Ethylene-Forming Enzyme: Substrate Binding Creates a Twist

Salette Martinez,[†,‖] Matthias Fellner,[‡,‖] Caitlyn Q Herr,[‡] Anastasia Ritchie,[†] Jian Hu,[‡,#,*] and Robert P. Hausinger[†,‡,*]

[†]*Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, Michigan 48824*

[‡]*Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, Michigan 48824*

[#]*Department of Chemistry, Michigan State University, East Lansing, Michigan 48824*

Contents:

1. Experimental Procedures.

2. Figure S1: Overall structure of EFE and regions that undergo the most notable structural changes.

3. Figure S2: LIGPLOTs of bound 2OG and L-Arg, and L-Arg interactions with surrounding residues.

4. Figure S3: Alignment of EFE sequences.

5. Figure S4: Comparison of EFE•Mn•2OG bound models.

6. Figure S5: Comparison of EFE•Mn•2OG•L-Arg and EFE•Fe•NOG•L-Arg models.

7. Figure S6: Comparisons among off-line active site structures.

8. Figure S7: Prior variant substitutions mapped onto the EFE structure.

9. Figure S8: Metal identification at the active site of EFE

10. Tables S1-S3: Crystal statistics.

11. Table S4: Ethylene- and P5C-forming activities of EFE variants.

12. Table S5: Dihedral angles ω of the D191-Y192 peptide bond in EFE samples.

13. Table S6: Ethylene-forming activities of previously described variants.

14. Table S7: Primers used for mutagenesis.

15. Table S8: Crystallization conditions.

16. References

EXPERIMENTAL PROCEDURES

Site-Directed Mutagenesis. Ethylene-forming enzyme (EFE) variants were created by mutagenesis of the *efe* gene located in plasmid pET28a+[1] using the primers listed (Table S7) and the Q5® Site-Directed Mutagenesis Kit (New England BioLabs, Inc.) according to the manufacturer's instructions. We confirmed the changes in each mutated plasmid by DNA sequencing (RTSF Genomics Core at Michigan State University) and transformed the plasmids into *E. coli* BL21 Gold (DE3) cells (Agilent Technologies).

Protein Production and Purification. The genes encoding *N*-terminally H_6-tagged wild-type (WT) and variant forms of EFE were expressed and the corresponding apoproteins were purified with cleavage of the H_6-tag as previously described.[1] During our analysis, we found that the concentration of the protein was previously underestimated;[1] the corrected specific activity for EFE was 612 ± 66 μmole min^{-1} mg^{-1}, in close agreement with the earlier published value of 650-660 U/mg.[2-3]

Assays for Quantifying Product Formation. We performed the standard assay at 25 ± 1°C in 10 × 16 mm Vacutainer glass tubes (Becton Dickinson). WT and variant forms of EFE (1 μM) were incubated in 2 mL of 25 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer, pH 7.5, containing 0.5 mM 2-oxoglutarate (2OG), 0.5 mM L-Arg, 0.2 mM $(NH_4)_2Fe(SO_4)_2$, and 0.4 mM L-ascorbic acid. After 80 min the reactions were terminated with 0.1 mL of 3.6 M hydrochloric acid. Products were quantified by GC (ethylene) and 2-aminobenzaldehyde derivatization of L-Δ-1-pyrroline-5-carboxylate (P5C) as previously described.[1]

Protein crystallization. For crystallization trials, we isolated the EFE apoprotein using our standard procedures and further purified the samples by size-exclusion chromatography using a Superdex HiLoad 16/600 75 prep grade column (GE Healthcare Life Sciences). We equilibrated the column with 25 mM HEPES, pH 8.0, containing 100 mM NaCl and 1 mM tris(2-carboxyethyl)phosphine (TCEP). EFE-containing fractions were pooled and buffer exchanged into 25 mM HEPES, pH 8.0, containing 1 mM TCEP. Subsequently, EFE was concentrated (53-72 mg/mL) and supplemented with 1-9 mM of $CoCl_2$, $MnCl_2$, $NiSO_4$, or $(NH_4)_2Fe(SO_4)_2$ and incubated on ice for 1 h. For co-crystallization experiments, the substrates (2OG, 2-oxoadipate (2OA), and/or L-Arg and L-Arg analogs) were added to the concentrated protein in the range of 0.6-10 mM, then incubated on ice for >1 h.

We carried out the initial crystallization screens at 4 °C in 96-well plates by the sitting-drop vapor diffusion technique and using the mosquito crystallization robot (TTP Labtech). Screening was performed against Index HT (Hampton Research), Crystal Screen HT (Hampton Research), Wizard 1&2 (Rigaku Reagents), and Wizard 3&4 (Rigaku Reagents) sparse-matrix screens. Crystals grew within 7-21 days and were optimized further. Table S8 lists a summary of conditions for all 11 crystal structures. EFE crystallized very well when incubated without a metal or in the presence of Mn. EFE also grew with Ni,

however crystals were of poorer quality (>3 Å and processing issues) and were not optimized. We obtained a Ni-bound structure by soaking an apoprotein crystal in $NiCl_2$ (see details in Table S8). The 2-h soak dehydrated the crystal, resulting in solvent content of 32.1% with a Matthews coefficient of 1.81. In addition, the soaked crystal suffered from high anisotropy, making data processing challenging. We therefore focused on Mn crystals and did not proceed with Ni-bound EFE. EFE did not crystallize in the presence of Co. Most attempts to obtain EFE crystals with Fe failed, with only very poorly diffracting crystals obtained (>8 Å grown at 21 °C), and we were unable to optimize these conditions further. In addition, soaking crystals with Fe, similar to Ni, abolished diffraction from these crystals.

Phase determination. For single-wavelength anomalous diffraction (SAD) experiments, selenomethionine (SeMet)-substituted EFE was produced using the protocol described for inhibition of methionine biosynthesis by Doublié[4] with minor modifications. The protein was purified similarly to the Met-containing EFE, but with buffers containing 2 mM TCEP. For details, see Table S8.

Data collection, processing, and structure determination. X-ray diffraction data were collected at the Advanced Photon Source LS-CAT beamlines (21-ID-D, 21-ID-F and 21-ID-G). For details about dataset collection, see Tables S1-S3. Datasets were indexed and integrated with xdsapp2.0[5] or iMosflm,[6] and merging and scaling were done using aimless.[7] Data quality was assessed using Phenix.xtriage.[8] The phase of a selenomethionine-substituted crystal was solved through SAD in Phenix.autosol,[9] and Phenix.phaser molecular replacement[8] was used for all subsequent datasets. Refinement was done in Phenix[8] utilizing Phenix.refine, Phenix.readyset, Phenix.eLbow, and COOT.[10] In our EFE•Mn•2OG•HO-L-Arg structure there were hints of a second active site conformation at low occupancy. The changes represent a conformation free of HO-L-Arg, with a second Asp191 orientation and with 2OG bound monodentate to the metal (matching our EFE•Mn•2OG structure). Attempts to model the low occupancy species did not provide satisfactory results and this second minor conformation was therefore omitted in the final model. Datasets were validated using MolProbity[11] and uploaded to the protein data bank (PDB) with IDs identified for apoprotein, 5V2U; Ni, 5V2V; Mn•2OG, 5V2X; Mn•2OG•L-Arg, 5V2Y; Mn•2OA•L-Arg, 5V2Z; Mn•2OG•HO-L-Arg, 5VKA; Mn•2OG•L-Arg-amide, 5VKB; Mn•malate•Arg, 5V34; Mn•Arg, 5V31; Mn•tartrate, 5V2T and Mn•malate, 5V32. Statistics for the deposited datasets are listed in the crystallographic details, Tables S1-S3. These tables also list the details for datasets used to determine the identity of the EFE bound metal. Structure figures were created with UCSF Chimera.[12] The identity of either Ni or Mn was demonstrated by dataset collection on a single EFE•Ni or EFE•Mn•tartrate crystal above and below the corresponding absorption edges (Figure S8 and Table S1 and S3).

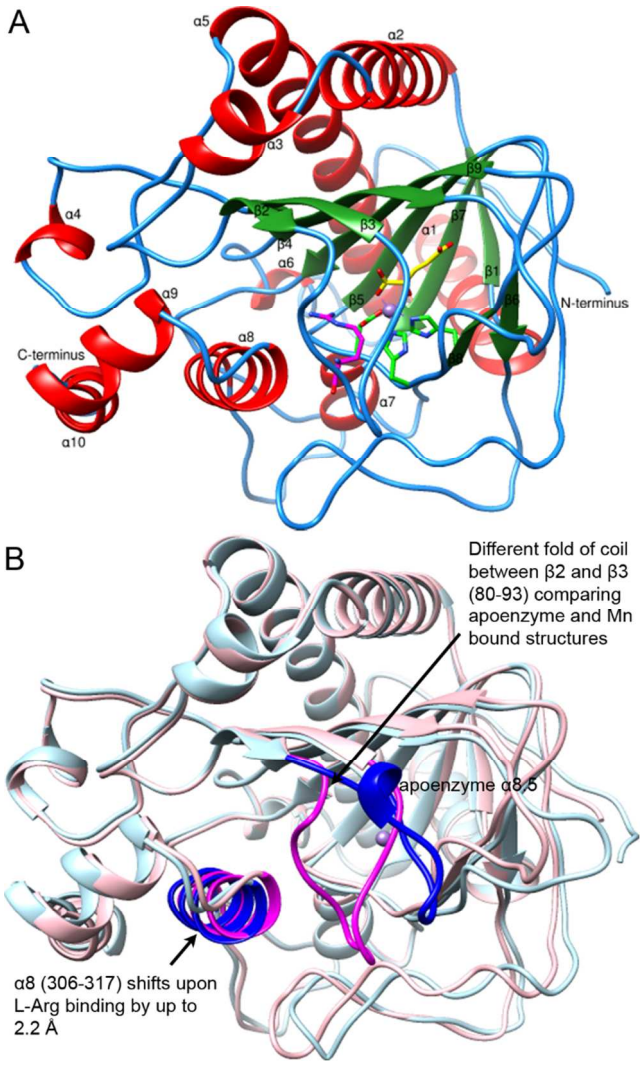
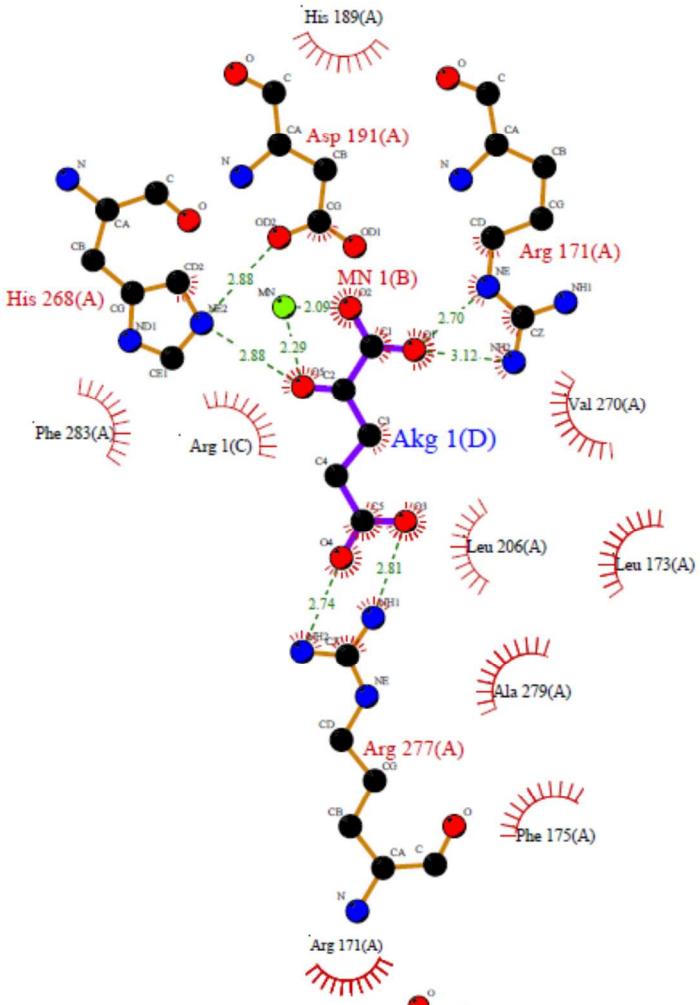
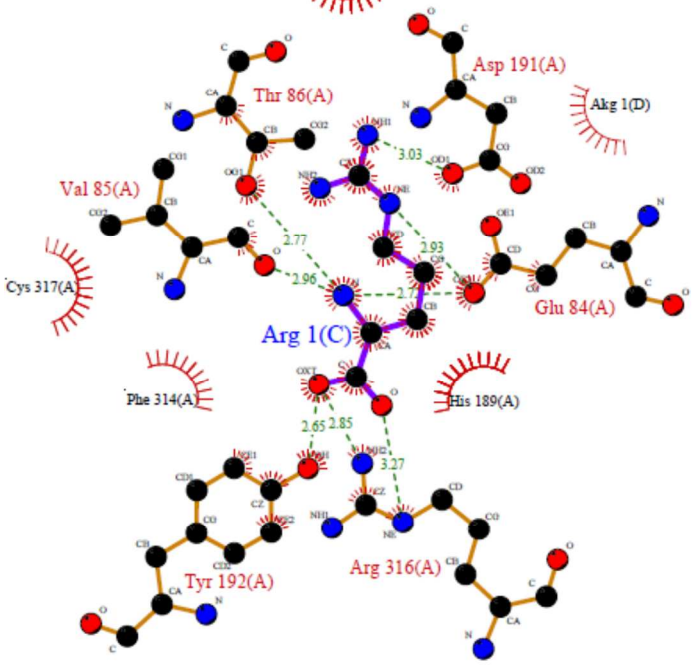


Figure S1. Overall structure of EFE from *P. syringae* pv. *phaseolicola* PK2 and regions that undergo the most notable structural changes. (A) The structure of EFE•Mn•2OG•L-Arg with helices depicted in red and labeled α1-α10, sheets shown in green and labeled β1-β9, and coil regions in light blue. Residues chelating the bound Mn (purple sphere) are shown with green carbon atoms, and the bound ligands, 2OG and L-Arg, are illustrated with carbon atoms in yellow and magenta, respectively. (B) The backbone of EFE•Mn•2OG•L-Arg is shown in faint pink with residues 80-93 and 306-317 depicted in dark pink, while EFE apoenzyme is shown in faint blue with the same residue ranges shown in dark blue. The short helix labeled α-8.5 is formed only in the apoprotein and occurs between α-8 and α-9 of EFE•Mn•2OG•L-Arg.

A


His 189(A)
Asp 191(A)
Arg 171(A)
MN 1(B)
His 268(A)
2.88
2.09
2.70
2.88
2.29
3.12
Val 270(A)
Phe 283(A)
Arg 1(C)
Akg 1(D)
Leu 206(A)
Leu 173(A)
2.81
2.74
Ala 279(A)
Arg 277(A)
Phe 175(A)
Arg 171(A)

B


Asp 191(A)
Thr 86(A)
Akg 1(D)
3.03
Val 85(A)
2.77
2.93
2.96
Cys 317(A)
Glu 84(A)
Arg 1(C)
Phe 314(A)
His 189(A)
2.65
2.85
3.27
Tyr 192(A)
Arg 316(A)

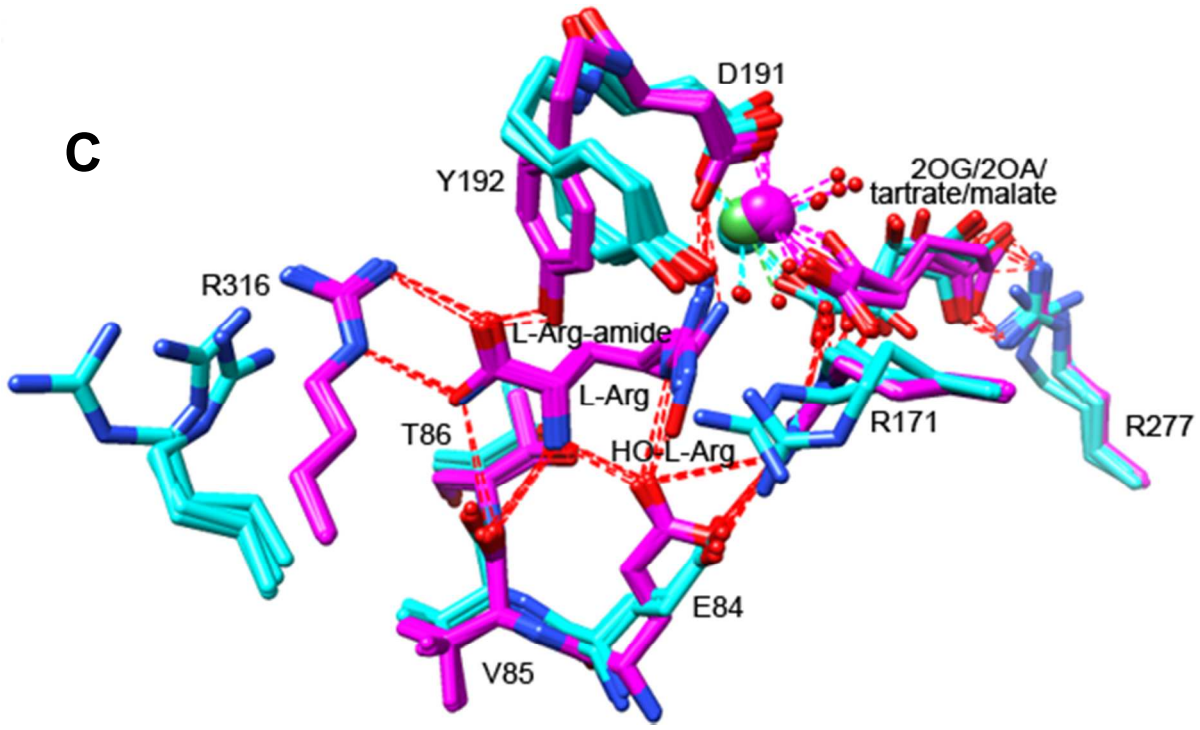


Figure S2. Side chain interactions of EFE with 2OG and L-Arg, and comparison of changes for the L-Arg binding site among the 11 structures. Interactions of (A) 2OG and (B) L-Arg with the surrounding active site residues in Mn•EFE•2OG•L-Arg as depicted by LIGPLOT.[13] (C) Comparison of the 11 EFE structures obtained here. All structures containing L-Arg or an L-Arg analog are colored in magenta; the other structures are colored in cyan. Hydrogen bonds are shown as red dashed lines. The Mn atom and its coordination sphere are also colored magenta or cyan, accordingly. The Ni atom and its coordination sphere are colored green. Water molecules bound to the metal are shown as red spheres.

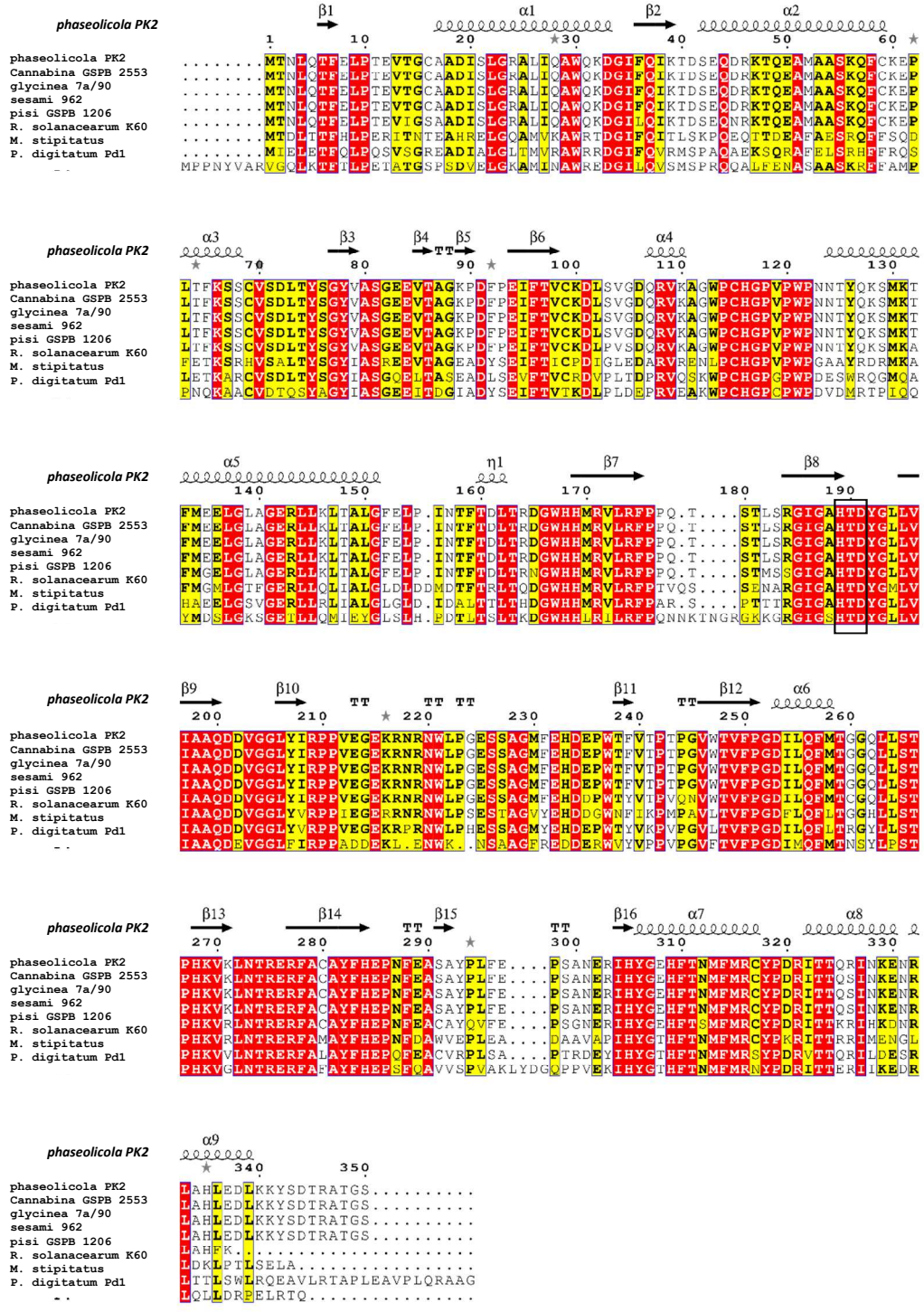


Figure S3. Alignment of EFE sequences for strains known to produce ethylene as shown by assays using cell-free extracts supplemented with 2OG and L-Arg (*Ralstonia solanacearum* K60 and the following pathovars of *P. syringae*: *cannabina* GSPB2553, *glycinea* 7a/90, *sesami* 962, and *pisi* GSPB1206) or by analysis of recombinant yeast cells containing *efe* from *Penicillium digitatum* Pd1 or *Myxococcus stipitatus*.[14-15] The Figure was created with EsPript 3.[16]

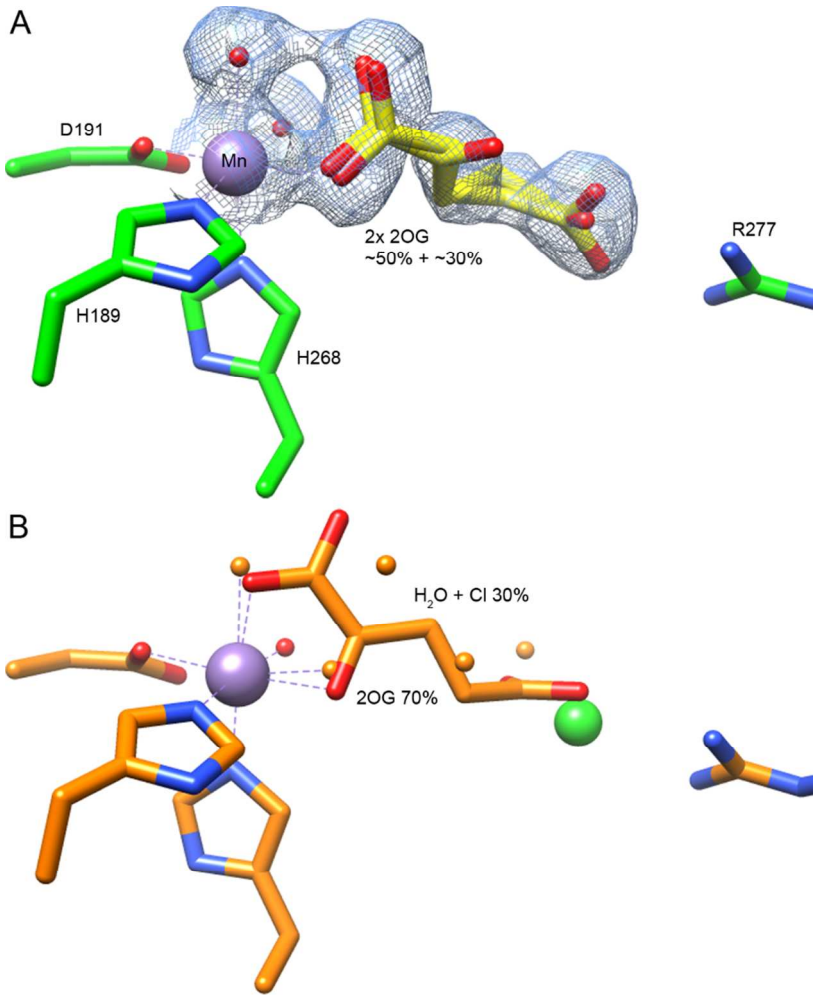


Figure S4. Comparison of EFE•Mn•2OG bound models. (A) Our EFE•Mn•2OG model (5V2X) with its $2F_O$ - F_C map shown as blue meshes at 1 σ. The metal is chelated by two water molecules and a monodentate bound 2OG depicted in two conformations at ~50 and ~30% occupancy. (B) Model by Zhang *et al*. (5MOF)[17] with carbon atoms in orange accounting for very similar electron density with the metal chelated by one water molecule (100%, red sphere) and either (at 30% occupancy) two chelating water molecules, additional water molecules (orange spheres), and a chloride (green sphere) towards R277 or (at 70% occupancy) bidentate bound 2OG that includes a single hydrogen bond with R277.

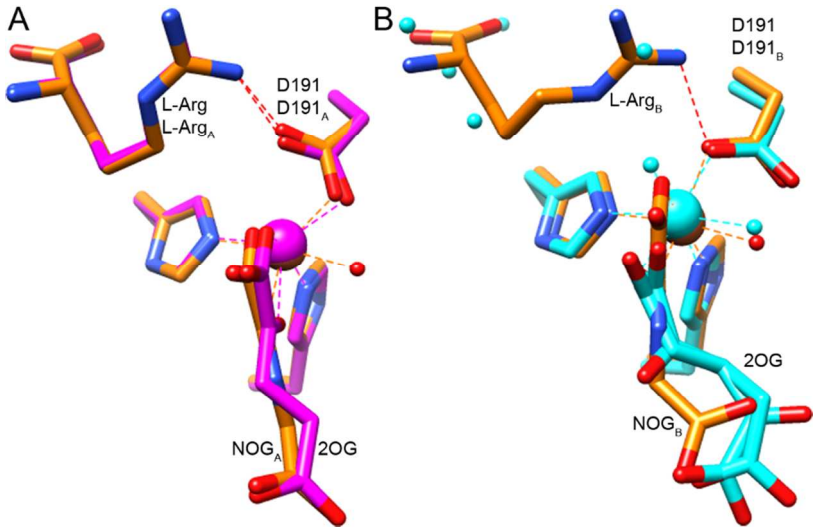


Figure S5. Comparison of EFE•Mn•2OG•L-Arg and EFE•Fe•NOG•L-Arg models. (A) Comparison of our EFE•Mn•2OG•L-Arg (5V2Y, carbon and Mn atoms in magenta) and the A conformation of EFE•Fe•NOG•L-Arg by Zhang *et al*. (5LUN, carbon and Fe atoms in orange).[17] (B) Comparison of one of our L-Arg free structures (EFE•Mn•malate; 5V32, carbon, water oxygen and Mn atoms in cyan) and the B conformation of 5LUN.

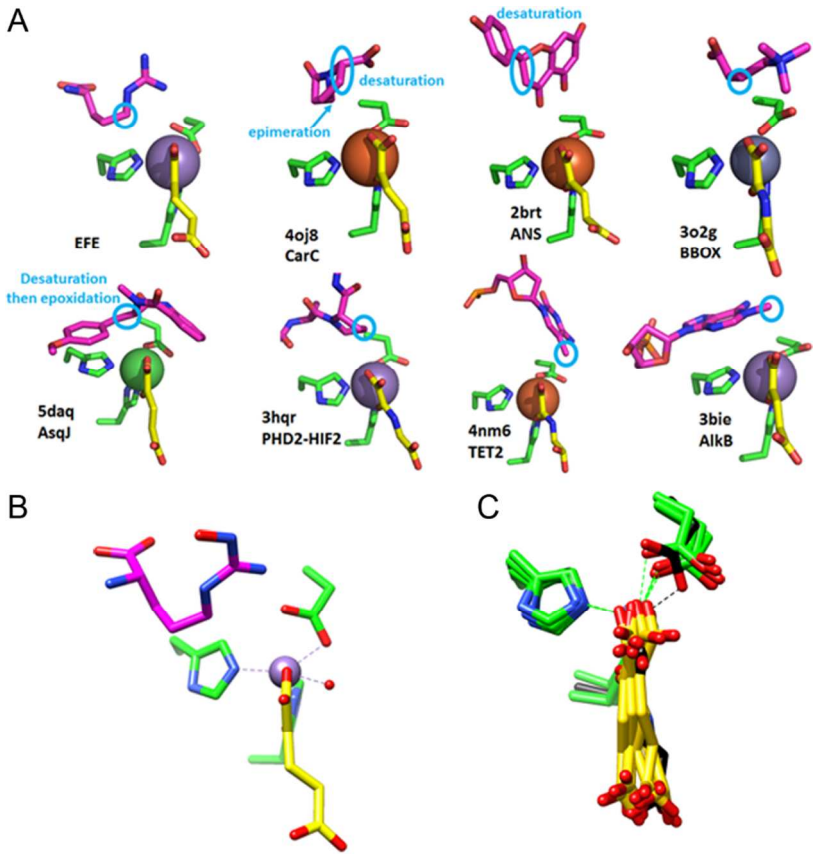


Figure S6. Comparison of the EFE•Mn•2OG•L-Arg active site geometry to that of selected other structurally characterized Fe(II)/2OG oxygenases which exhibit off-line geometry. (A) In each case, the metallocenters (spheres), 2-histidine-1-carboxylate side chains (green carbons), and 2OG or NOG (yellow carbons) are shown in similar orientations, with the substrates (magenta carbons) positioned away from the apparent dioxygen-binding site to the right of the metal. The sites of oxygenation (hydroxylation or desaturation) on the substrates are highlighted by the blue ovals. The enzyme active sites depicted include our EFE•Mn•2OG•L-Arg (5V2Y) structure, as well as carbapenem synthase (CarC, PDB access code: 4OJ8),[18] anthocyanidin synthase (ANS, 2BRT),[19] γ-butyrobetaine hydroxylase (BBOX, 3O2G), AsqJ (5DAQ),[20] HIF hydroxylase (PHD2, 3HQR),[21] 5-methylcytosine hydroxylase (TET2, 4NM6),[22] and AlkB demethylase (3BIE).[23] (B) Offline representation of our EFE•Mn•2OG•HO-L-Arg (5VKA) structure showing a water molecule at the oxygen binding site. (C) Offline comparison of structures shown in (A) with our EFE•Mn•2OG•L-Arg colored in black. Note that the carboxylate metal chelating ligand in EFE adopts a distinct conformation from the other Fe(II)/2OG oxygenases.

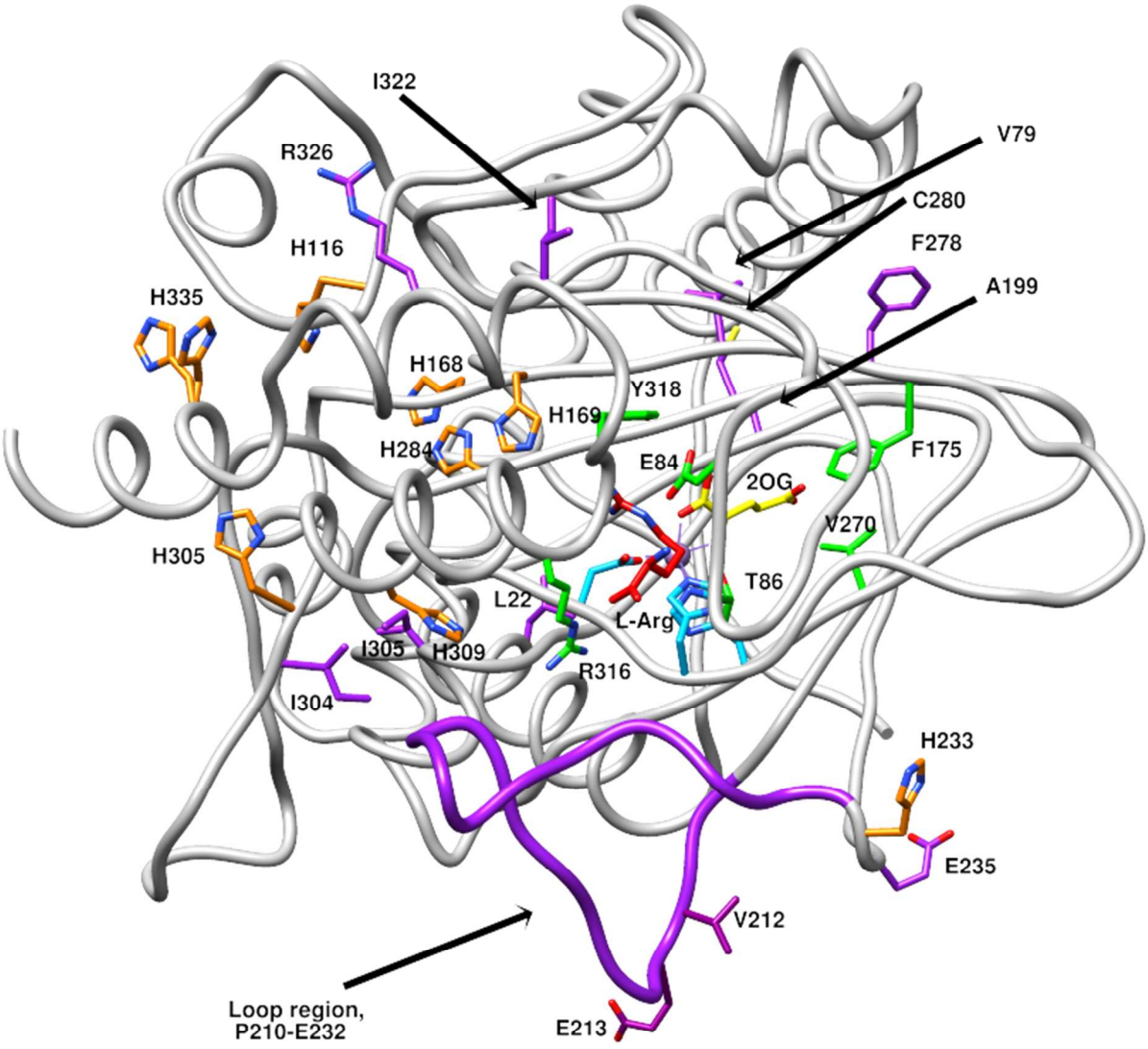


Figure S7. Structure of EFE•Mn•2OG•L-Arg showing the location of residues previously substituted by others, excluding the active site residues. Residues colored in orange were studied by Nagahama *et al*.,[24] those shown in purple were investigated by Johansson *et al*.,[15] and residues in green were characterized by Zhang *et al*.[17] 2OG and L-Arg are shown with yellow and red carbon atoms, respectively. Metal-binding residues are shown in cyan.

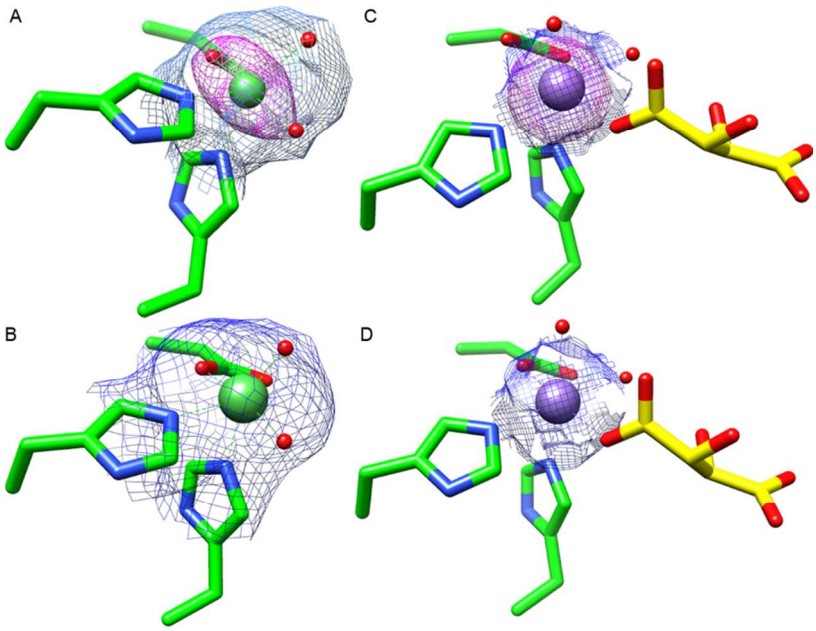


Figure S8. Metal identification at the active site of EFE. Carbon atoms of the metal-chelating residues (H189, D191 and H268) are shown in green; tartrate is depicted in yellow. Ni and Mn atoms are shown as green and purple spheres, respectively. The $2F_O$ - F_C maps for metal atoms are shown as blue meshes at 1 σ. The anomalous maps for metal atoms are shown as magenta meshes at 5 σ. (A) EFE•Ni examined above the Ni edge at 8.39 keV. (B) EFE•Ni examined below the Ni edge at 8.28 keV. (C) EFE•Mn•tartrate examined above the Mn edge at 6.59 keV. (D) EFE•Mn•tartrate examined below the Mn edge at 6.48 keV.

Table S1. Crystal statistics for EFE apoprotein; EFE in complex with Ni(II); EFE with bound Mn(II) and 2OG; EFE with Mn(II), 2OG, and L-Arg; and EFE with Mn(II), 2OA, and L-Arg.

EFE crystals	apoprotein	Ni	Mn•2OG	Mn•2OG •L-Arg	Mn•2OA •L-Arg
Data collection					
Beamline	LS-CAT 21-ID-F	LS-CAT 21-ID-D	LS-CAT 21-ID-F	LS-CAT 21-ID-G	LS-CAT 21-ID-D
Wavelength (Å)	0.979	1.478	0.979	0.979	1.127
Detector distance (mm)	240	150	240	190	100
Number of frames	90 at 1°	164 at 1°	140 at 1°	150 at 1°	310 at 1°
Space group	$P2_12_12_1$	$P2\ 2_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell a, b, c (Å) α, β, γ (°)	43, 87, 92 90, 90, 90	42, 79, 88 90, 90, 90	74, 98, 102 90, 90, 90	49, 82, 88 90, 90, 90	49, 82, 88 90, 90, 90
[a]Resolution (Å)	45.91 – 2.06 (2.11 – 2.06)	87.75 – 3.04 (3.22 – 3.04)	45.33 – 1.85 (1.89 – 1.85)	43.80 – 1.43 (1.45 – 1.43)	48.45 – 1.23 (1.25 – 1.23)
Unique reflections	21,539 (1,652)	5,959 (941)	64,351 (3,789)	65,410 (3,159)	100,282 (4,164)
[a]Redundancy	3.7 (3.6)	4.8 (5.1)	5.8 (5.6)	5.8 (3.9)	10.1 (4.8)
[a]Completeness (%)	98.1 (98.2)	99.3 (99.7)	99.6 (95.9)	99.7 (96.3)	98.7 (84.1)
[a]$I/\sigma I$	9.6 (2.8)	7.0 (3.3)	11.6 (2.5)	15.9 (2.2)	14.6 (2.4)
[a]R_{merge}	0.114 (0.413)	0.132 (0.350)	0.119 (0.675)	0.060 (0.532)	0.080 (0.560)
[a]R_{pim}	0.093 (0.343)	0.094 (0.246)	0.083 (0.478)	0.041 (0.441)	0.037 (0.419)
[b]$CC_{1/2}$	0.989 (0.778)	0.990 (0.941)	0.997 (0.830)	0.999 (0.773)	0.998 (0.816)
Refinement					
Protein atoms	2,739	2,316	5,507	2,805	2,921
Hydrogen atoms	0	0	0	2,718	2,842
Metal atoms	0	1	2	1	1
2OG/2OA molecules	0	0	4	1	1
L-Arg molecules	0	0	0	1	1
H_2O molecules	278	2	994	353	439
[c]R_{work}/R_{free}	0.186 / 0.245	0.199 / 0.268	0.163 / 0.201	0.132 / 0.161	0.141 / 0.160
B-factors (Å^2)	18.9	47.7	18.2	18.5	16.7
Protein	18.5	47.7	16.1	16.0	14.2
Hydrogens	-	-	-	19.8	17.6
Metal	-	52.01	13.4	11.3	8.5
2OG/2OA molecules	-	-	18.8	17.5	14.3
L-Arg molecules	-	-	-	14.7	12.0
H_2O	22.2	44.13	30.0	29.4	27.7
R.m.s. deviation in bond lengths (Å)	0.003	0.005	0.007	0.007	0.011
R.m.s. deviation in bond angles (°)	0.616	0.533	0.750	0.949	1.110
Ramachandran plot (%) favored	97.95	95.27	97.81	97.94	97.98
Ramachandran plot (%) outliers	0.29	0.34	0	0.59	0.58
Rotamer outliers (%)	0	0	0	0.33	0
PDB ID	5V2U	5V2V	5V2X	5V2Y	5V2Z

[a]Highest resolution shell is shown in parentheses.

[b]$CC_{1/2}$ is the correlation coefficient of the half datasets.

[c]$R_{work} = \sum_{hkl} | |F_{obs}| - |F_{calc}| | / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} is the observed and the calculated structure factor, respectively. R_{free} is the cross-validation R factor for the test set of reflections (5% of the total, except for Ni 15%) omitted in model refinement.

Table S2. Crystal statistics for EFE in complex with Mn, 2OG, and HO-L-Arg; Mn, 2OG, and Arg-amide; Mn and L-Arg; and Mn and tartrate.

EFE crystals	Mn•2OG• HO-L-Arg	Mn•2OG• L-Arg-amide	Mn•L-Arg	Mn•tartrate
Data collection				
Beamline	LS-CAT 21-ID-F	LS-CAT 21-ID-F	LS-CAT 21-ID-D	LS-CAT 21-ID-G
Wavelength (Å)	0.979	0.979	1.127	0.979
Detector distance (mm)	140	130	200	150
Number of frames	150 at 1°	220 at 1°	99 at 1°	290 at 1°
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	I2 2 2
Unit cell a, b, c (Å) α, β, γ (°)	49, 82, 88 90, 90, 90	49, 82, 88 90, 90, 90	49, 82, 87 90, 90, 90	80, 98, 98 90, 90, 90
[a]Resolution (Å)	48.58 – 1.17 (1.19 – 1.17)	48.48 – 1.14 (1.16 – 1.14)	87.16 – 2.45 (2.55 – 2.45)	48.99 – 1.23 (1.25 – 1.23)
Unique reflections	118,142 (5,146)	127,532 (5,998)	12,891 (1,451)	111,277 (5,275)
[a]Redundancy	5.7 (3.6)	8.4 (5.4)	3.5 (3.6)	11.1 (7.1)
[a]Completeness (%)	99.2 (87.9)	99.6 (95.4)	97.3 (98.0)	99.7 (95.6)
[a]$I/\sigma I$	12.3 (2.2)	17.1 (2.0)	5.4 (2.4)	20.3 (2.3)
[a]R_{merge}	0.071 (0.515)	0.064 (0.822)	0.163 (0.514)	0.066 (0.905)
[a]R_{pim}	0.048 (0.401)	0.034 (0.565)	0.132 (0.425)	0.030 (0.530)
[b]$CC_{1/2}$	0.997 (0.752)	0.999 (0.747)	0.972 (0.718)	0.999 (0.739)
Refinement				
Protein atoms	2,863	2,906	2,706	2,805
Hydrogen atoms	2,782	2,829	0	2,706
Mn atoms	1	1	1	1
2OG/tartrate molecules	1	1	0	1
L-Arg like molecules	1	2	1	0
H_2O molecules	505	472	128	466
[c]R_{work}/R_{free}	0.172 / 0.188	0.147 / 0.160	0.180 / 0.247	0.139 / 0.153
B-factors (Å^2)	16.2	19.1	25.0	18.7
Protein	13.6	18.3	25.0	15.8
Hydrogens	16.8	20.1	-	19.4
Manganese	14.0	11.2	25.6	10.2
2OG/tartrate molecules	16.2	17.8	-	19.2
L-Arg like molecules	13.5	19.2	30.7	-
H_2O	27.7	29.4	25.5	31.5
R.m.s. deviation in bond lengths (Å)	0.009	0.008	0.016	0.008
R.m.s. deviation in bond angles (°)	0.961	0.957	0.915	0.928
Ramachandran plot (%) favored	97.69	97.97	97.63	97.95
Ramachandran plot (%) outliers	0.58	0.58	0.59	0.29
Rotamer outliers	0.33	0	0	0
PDB ID	5VKA	5VKB	5V31	5V2T

[a]Highest resolution shell is shown in parentheses.

[b]$CC_{1/2}$ is the correlation coefficient of the half datasets.

[e]$R_{work} = \sum_{hkl} | |F_{obs}| - |F_{calc}| | / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} is the observed and the calculated structure factor, respectively. R_{free} is the cross-validation R factor for the test set of reflections (5% of the total) omitted in model refinement.

Table S3. Crystal statistics for EFE in complex with Mn and malic acid; Mn, malic acid, and L-Arg; as well as datasets collected for the metal determination of Ni and Mn.

EFE crystals	**Mn•malate**	**Mn•malate •L-Arg**	**Ni 8.28 keV**	**Mn•tartrate 6.59 keV**	**Mn•tartrate 6.48 keV**
Data collection					
Beamline	LS-CAT 21-ID-G	LS-CAT 21-ID-D	LS-CAT 21-ID-D	LS-CAT 21-ID-D	LS-CAT 21-ID-D
Wavelength (Å)	0.979	1.127	1.497	1.881	1.913
Detector distance (mm)	180	120	150	100	100
Number of frames	140 at 1°	131 at 1°	151 at 1°	200 at 1°	170 at 1°
Space group	I2 2 2	$P3_22\ 1$	$P2_12_12_1$	I2 2 2	I2 2 2
Unit cell a, b, c (Å) α, β, γ (°)	79, 98, 98 90, 90, 90	87, 87, 104 90, 90, 120	42, 76, 85 90, 90, 90	80, 98, 98 90, 90, 90	80, 98, 98 90, 90, 90
[a]Resolution (Å)	38.35 – 1.49 (1.51 – 1.49)	61.24 – 1.48 (1.51 – 1.48)	88.49 – 2.76 (2.91 – 2.76)	69.40 – 2.25 (2.32 – 2.25)	69.43 – 2.29 (2.37 – 2.29)
Unique reflections	62,625 (2,835)	75,179 (3,605)	7,032 (1,005)	18,082 (1,618)	17,705 (1,695)
[a]Redundancy	5.7 (5.4)	7.2 (7.4)	4.3 (4.1)	7.0 (7.2)	5.8 (6.0)
[a]Completeness (%)	99.5 (91.0)	97.9 (96.4)	91.9 (93.4)	97.6 (96.5)	99.6 (99.5)
[a]$I/\sigma I$	15.7 (2.6)	12.8 (1.9)	5.6 (1.7)	10.2 (3.3)	8.8 (2.4)
[a,b]R_{merge}	0.072 (0.901)	0.078 (0.968)	0.127 (0.427)	0.112 (0.516)	0.119 (0.653)
[a,c]R_{pim}	0.049 (0.608)	0.045 (0.563)	0.080 (0.298)	0.066 (0.303)	0.079 (0.427)
[d]$CC_{1/2}$	0.999 (0.844)	0.998 (0.639)	0.990 (0.881)	0.993 (0.854)	0.994 (0.758)
Refinement					
Protein atoms	2,841	2,797	-	-	-
Hydrogen atoms	2,748	2,717	-	-	-
Metal atoms	1	1	-	-	-
Malic acid molecules	1	1	-	-	-
L-Arg molecules	0	1	-	-	-
H_2O molecules	278	473	-	-	-
[e]R_{work}/R_{free}	0.142 / 0.169	0.135 / 0.162	-	-	-
B-factors (Å^2)	19.9	23.7	-	-	-
Protein	17.2	20.6	-	-	-
Hydrogens	21.2	25.3	-	-	-
Metal	10.4	15.8	-	-	-
Malic acid	20.7	24.2	-	-	-
L-Arg	-	17.7	-	-	-
H_2O	22.2	34.0	-	-	-
R.m.s. deviation in bond lengths (Å)	0.008	0.010	-	-	-
R.m.s. deviation in bond angles (°)	0.911	1.029	-	-	-
Ramachandran plot (%) favored	98.22	97.06	-	-	-
Ramachandran plot (%) outliers	0.30	0.59	-	-	-
Rotamer outliers (%)	0	0	-	-	-
PDB ID	5V32	5V34	-	-	-

[a]Highest resolution shell is shown in parentheses.

[b]$CC_{1/2}$ is the correlation coefficient of the half datasets.

[c]$R_{work} = \sum_{hkl} | |F_{obs}| - |F_{calc}| | / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} is the observed and the calculated structure factor, respectively. R_{free} is the cross-validation R factor for the test set of reflections (5% of the total) omitted in model refinement.

Table S4. Ethylene and P5C production by EFE variants.[a]

Variant	Role	Ethylene, nmol (%)	P5C, nmol (%)	N
WT		561 ±130 (100 ± 23)	253 ± 63 (100 ± 25)	15
L73K	Surface	297 ± 94 (53 ± 17)	145 ± 16 (57 ± 7)	3
L73R	Surface	27 ± 19 (5 ± 3)	86 ± 9 (34 ± 3)	4
S81R	Surface	22 ± 13 (4 ± 2)	15 ± 11 (6 ± 4)	3
S81Y	Surface	7.2 ± 0.2 (1.29 ± 0.04)	16 ± 6 (6 ± 2)	2
E84A	L-Arg binding	0.0 ± 0.0 (0 ± 0)	13 ± 6 (5 ± 2)	4
R171A	2OG binding	0.0 ± 0.0 (0 ± 0)	12 ± 5 (5 ± 2)	2
R184A	Surface	397 ± 82 (71 ± 15)	175 ± 14 (69 ± 6)	2
Y192F	L-Arg binding	69 ± 10 (12 ± 2)	39 ± 2 (15 ± 1)	2
Y192R	L-Arg binding	0.0 ± 0.0 (0 ± 0)	7 ± 8 (3 ± 3)	4
Y192W	L-Arg binding	3 ± 2 (0.6 ± 0.3)	52.7 ± 4.3 (21 ± 2)	3
V196R	Hydrophobic pocket	0.2 ± 0.2 (0.03 ± 0.04)	1 ± 1 (0.3 ± 0.4)	2
V196F	Hydrophobic pocket	N/A[b]	N/A	N/A
A198V	Hydrophobic pocket	14 ± 1 (2.3 ± 0.3)	157 ± 10 (62 ± 4)	3
E213A	Surface	582 ± 216 (104 ± 38)	235 ± 15 (93 ± 6)	2
E213A/E215A	Surface	19 ± 22 (3 ± 4)	19 ± 2 (8 ± 1)	2
E215A	Surface	23 ± 7 (4.0 ± 1.1)	21 ± 0.3 (8.0 ± 0.1)	2
R277A	2OG binding	N/A[b]	N/A	N/A
A281R	Hydrophobic pocket	0.1 ± 0.1 (0.02 ± 0.02)	2 ± 3 (0.7 ± 1.0)	2
A281V	Hydrophobic pocket	15 ± 13 (3 ± 2)	12 ± 5 (5 ± 2)	2
F283A	Hydrophobic pocket	1.0 ± 0.6 (0.12 ± 0.11)	50 ± 8 (20 ± 3)	3
F283R	Hydrophobic pocket	0.0 ± 0.0 (0 ± 0)	48 ± 4 (19 ± 1)	3
F283V	Hydrophobic pocket	2.4 ± 4.2 (0.4 ± 0.8)	48 ± 11 (19 ± 4)	3
F283W	Hydrophobic pocket	1.1 ± 0.8 (0.2 ± 0.1)	57 ± 10 (22 ± 4)	3
F283Y	Hydrophobic pocket	0.0 ± 0.0 (0 ± 0)	0.0 ± 0.0 (0 ± 0)	2
E285A	Surface	0.0 ± 0.0 (0 ± 0)	38 ± 22 (15 ± 9)	3
E285Q	Surface	0.0 ± 0.0 (0 ± 0)	65 ± 1 (26 ± 1)	2
Y306A	Hydrophobic interactions	24 ± 8 (4 ± 1)	14 ± 4 (5 ± 1)	3
Y306F	Hydrophobic interactions	31 ± 8 (6 ± 1)	25 ± 4 (10 ± 2)	2
F310R	Hydrophobic interactions	0.4 ± 0.3 (0.06 ± 0.05)	8 ± 3 (3 ± 1)	3
F310W	Hydrophobic interactions	120 ± 47 (21 ± 8)	76 ± 12 (30 ± 5)	2
C317N	L-Arg binding	154 ± 33 (27 ± 6)	106 ± 19 (42 ± 8)	3
C317S	L-Arg binding	464 ± 4 (83 ± 1)	242 ± 3 (96 ± 1)	2
Δ338-350	C-terminus	314 ± 99 (56 ± 18)	163 ± 9 (64 ± 4)	2

[a]The product concentrations are the average of at least two independent assays and are reported as nmoles ± the standard deviation. Values in parentheses are % product ± the standard deviation. Whereas ethylene production provides a direct readout of that product, the levels of P5C indicate the amount of C5 hydroxylation of L-Arg. N is the number of replicates for that sample. [b]N/A, not available due to localization of the protein in inclusion bodies.

Table S5. Dihedral peptide angle ω of the Asp191-Tyr192 bond (ideal 180° for C_α, C, N_{+1}, $C_{\alpha+1}$)

Near ideal ω 180		twisted >30° from ideal	
Mn•2OG	170.7	Mn•2OG•L-Arg	145.4
apoprotein	169.7	Mn•2OA•L-Arg	145.7
Ni	163.1	Mn•2OG•HO-L-Arg	146.8
Mn•tartrate	169.2	Mn•2OG•L-Arg-amide	147.7
Mn•malate	168.0	Mn•malate•L-Arg	144.5
		Mn•L-Arg	150.7
average	168.1	average	146.8
Mn•2OG[a]	170.4		
Mn•BTP[b]	174.0		
Fe•NOG•L-Arg[c]			
chain A-B	153.7	chain A-A	146.9
chain B-B	151.6	chain B-A	145.8
chain C-B	152.3	chain C-A	147.3
chain D-B	152.5	chain D-A	146.3
average	152.5	average	146.6

[a]PDB 5MFO: EFE•Mn•2OG from Zhang *et al*.[17]

[b]PDB 5LSQ: EFE•Mn•bis-Tris-propane from Zhang *et al*.[17]

[c]PDB 5LUN: EFE•Fe•NOG•L-Arg from Zhang *et al*.[17] (Note that Asp191 was modelled with two conformations of the sidechain; B represents O_{D1} chelation of the metal, A represents O_{D2} chelation that introduces the twisted peptide bond).

Table S6. Ethylene-forming activities of previously described EFE variants. Information shown on a white, light gray, and dark gray backgrounds were for variants investigated by Zhang *et al*.,[17] Johansson *et al*. (estimated from the published graph),[15] and Nagahama *et al*.[24] P5C levels were not monitored in these investigations.

EFE variants	Relative ethylene forming activity (%)	Proposed role
Wild type	100±3.6	
E84Q	ND[a]	L-Arg binding
E84D	ND	L-Arg binding
T86S	31.3 ±0.7	L-Arg binding
T86V	ND	L-Arg binding
R171K	ND	L-Arg binding
R171A[b]	ND	L-Arg binding
Y192F[b]	5.6 ±0.2	L-Arg binding
R316K	13.1 ±0.3	L-Arg binding
R316A	3.7 ±0.1	L-Arg binding
C317S[b]	21.4 ±0.5	L-Arg binding
C317A	34.0 ±0.7	L-Arg binding
Y318F	65.6 ±1.2	L-Arg binding
F175Y	18.6 ±0.5	2OG binding
V270T	4.3 ±0.2	2OG binding
L22M	100.00	Surface
V172T	80.00	Surface
A199G	~38	Surface
VE212-213YS	~50	Loop/Surface
E235D	~90	Surface
I254M	~95	Surface
F278Y	~110	Surface
I304N	ND	Surface/possible hydrophobic interaction with I254
I322V	~110	Surface
Loop deletion, 210-232	ND	Loop/surface/structural stability
C280F	~55	Surface/possible interaction with M54
R326S	~70	Surface
H189A	ND	Metal ligand
D191A	ND	Metal ligand
H268A	ND	Metal ligand
H233A	~125	Surface
H116Q	2.40	Surface
H168Q	3.00	Surface/hydrogen bonds with solvent
H169Q	9.30	Weak hydrogen bonding with L-Arg and π stacking

		interactions with F283
H189Q	ND	Metal ligand
H233Q	ND	Surface/hydrogen bonds with solvent
H268Q	1.80	Metal ligand
H284Q	2.00	Hydrogen bonding with E142 and Y282
H305Q	40.00	Surface/Hydrogen bonds with E308 and solvent
H309Q	3.30	Part of helix that moves toward active site upon bind of L-Arg; hydrogen bonds to R312 and D253
H335Q	60.00	Surface

[a]ND, not detected.

[b]Identical to the variants we report.

Table S7. Oligonucleotides used for EFE mutagenesis. F and R correspond to forward and reverse primers, respectively.

EFE Variant	Oligonucleotides
Deletion Δ338-350	F: TAAGGATCCGAATTCGAGC R: TTCCAGATGGGCTAACCG
C317N	F: GTTTATGCGGaacTACCCCGATC R: ATGTTGGTAAAGTGTTCG
C317S	F: GTTTATGCGGtctTACCCCGATC R: ATGTTGGTAAAGTGTTCG
F310R	F: TGGCGAACACcgtACCAACATGTTTATG R: TAATGAATCCGTTCATTGG
F310W	F: TGGCGAACACtggACCAACATGT R: TAATGAATCCGTTCATTGGC
Y306F	F: ACGGATTCATtttGGCGAACACTTTAC R: TCATTGGCGGAGGGTTCA
Y306A	F: ACGGATTCATgcgGGCGAACACTTTACC R: TCATTGGCGGAGGGTTCA
E285A	F: TTATTTTCACgcgCCCAACTTTGAAGCCTC R: GCACAGGCAAACCGTTCA
E285Q	F: TTATTTTCACcagCCCAACTTTGAAGCCTC R: GCACAGGCAAACCGTTCA
F283R	F: CTGTGCTTATcgtCACGAACCCAACTTTG R: GCAAACCGTTCACGGGTA
F283Y	F: CTGTGCTTATtatCACGAACCCAAC R: GCAAACCGTTCACGGGTA
F283W	F: CTGTGCTTATtggCACGAACCCAAC R: GCAAACCGTTCACGGGTA
F283A	F: CTGTGCTTATgcgCACGAACCCAACTTTG R: GCAAACCGTTCACGGGTA
F283V	F: CTGTGCTTATgtgCACGAACCCAAC R: GCAAACCGTTCACGGGTA
A281V	F: GTTTGCCTGTgtgTATTTTCACGAACC R: CGTTCACGGGTATTCAATTTC
A281R	F: GTTTGCCTGTcgtTATTTTCACGAACC R: CGTTCACGGGTATTCAATTTC
R277A	F: TACCCGTGAAgcgTTTGCCTGTG R: TTCAATTTCACTTTATGGGG
E215A	F: CGTGGAAGGGgcgAAACGGAATC R: GGGGGACGAATGTACAGA
E213A/E215A	F: ggccAAACGGAATCGCAACTGG R: cccgcCACGGGGGGACGAATGTA
E213A	F: TCCCCCCGTGgcgGGGGAAAAAC R: CGAATGTACAGACCGCCAA
A198V	F: GTTGGTGATTgtgGCTCAGGATG R: AGACCATAATCGGTATGG
V196R	F: TGGTCTGTTGcgtATTGCCGCTCAGGATG R: TAATCGGTATGGGCACCA
V196F	F: TGGTCTGTTGtttATTGCCGCTC

	R: TAATCGGTATGGGCACCA
Y192F	F: CCATACCGATtttGGTCTGTTGGTG R: GCACCAATGCCCCGACTC
Y192W	F: CCATACCGATtggGGTCTGTTGGTGATTGC R: GCACCAATGCCCCGACTC
Y192R	F: CCATACCGATcgtGGTCTGTTGGTGATTGC R: GCACCAATGCCCCGACTC
R184A	F: CACTCTGAGTgcgGGCATTGGTG R: GAGGTTTGGGGGGGAAAA
R171A	F: GCATCACATGgcgGTGTTACGTTTTCCCC R: CATCCGTCACGGGTCAAA
E84A	F: CTCCGGTGAAgcgGTTACCGCTG R: GCCACGTAGCCAGAATAA
S81R	F: CTACGTGGCCcgtGGTGAAGAAG R: CCAGAATAAGTTAAATCGC
S81Y	F: CTACGTGGCCtatGGTGAAGAAG R: CCAGAATAAGTTAAATCGC
L73R	F: CGTGAGCGATcgtACTTATTCTGG R: CAACTGGATTTAAAGGTCAAG
L73K	F: CGTGAGCGATcgtACTTATTCTGG R: CAACTGGATTTAAAGGTCAAG

Table S8. Crystallization conditions for 4 °C sitting drop vapor diffusion. Protein buffer was always 25 mM HEPES, pH 8.0, with 1 mM TCEP. Crystals were cryoprotected by soaking ~1 min in 25% cryoprotectant and 75% reservoir solution.

Structure ligand	PDB ID	Drop µl	EFE mg/ml	Reservoir	µl	Cryo 25%
apoprotein	5V2U	0.5 + 0.5	72	20% w/v PEG 6,000 0.1 M Tris, pH 8.0 0.2 M LiCl	200	none
Nickel[a]	5V2V	0.5 + 0.5	72	20% w/v PEG 6,000 0.1 M Tris, pH 8.0 0.2 M LiCl	200	none
Mn•2OG 1 mM $MnCl_2$ 0.6 mM 2OG	5V2X	0.18 + 0.18	61	20% w/v PEG 6,000 0.1 M Tris, pH 8.0 0.2 M LiCl	50	glycerol
Mn•2OG•L-Arg 4 mM $MnCl_2$ 2.5 mM 2OG + L-Arg	5V2Y	0.2 + 0.2	64	25% w/v PEG 3,350 0.1 M Bis-Tris, pH 6.5 0.2 M NaCl	50	PEG 400
Mn•2OA•L-Arg 4 mM $MnCl_2$ 3 mM 2OA 3 mM L-Arg	5V2Z	0.5 + 0.5	64	25% w/v PEG 3,350 0.1 M Bis-Tris, pH 6.5 0.2 M NaCl	200	PEG 400
Mn•2OG•HO-L-Arg 9 mM $MnCl_2$ 100 mM 2OG 10 mM HO-L-Arg	5VKA	0.5 + 0.5	72	20% w/v PEG 6,000 0.1 M Tris, pH 8.0 0.2 M LiCl	200	ethylene glycol
Mn•2OG•L-Arg-amide 9 mM $MnCl_2$ 100 mM 2OG 10 mM L-Arg-amide	5VKB	0.5 + 0.5	72	20% w/v PEG 6,000 0.1 M Tris, pH 8.0 0.2 M LiCl	200	PEG 400
Mn•malate•L-Arg 4 mM $MnCl_2$ 2.5 mM L-Arg	5V34	0.2 + 0.2	64	25% w/v PEG 1,500 0.1 M DL-malic acid:MES:Tris buffer, pH 6.5	50	PEG 400
Mn•L-Arg 1 mM $MnCl_2$ 2.5 mM L-Arg	5V31	0.2 + 0.2	64	25% w/v PEG 10,000 0.1 M HEPES, pH 7.5	50	PEG 400
Mn•tartrate[b] 1.2 mM $MnCl_2$	5V2T	1 + 1	59	1 M potassium sodium tartrate 0.1 M MES, pH 6.0	200	glycerol
Mn•malate 4 mM $MnCl_2$	5V32	0.2 + 0.2	64	20% w/v PEG 3,350 0.15 M DL-malic acid, pH 7.0	50	PEG 400

[a]A single apoprotein crystal was soaked for 2 h in a sitting drop containing 0.1 M $NiCl_2$, 0.2 M LiCl, 0.1 M Tris, pH 8.0, 20% w/v PEG 6,000, with the reservoir of 100 µl containing 0.2 M LiCl, 0.1 M Tris-HCl, pH 8.0.

[b]This dataset was used to solve the phases by SAD using EFE containing Se-Met.

REFERENCES

1. Martinez, S.; Hausinger, R. P. *Biochemistry* **2016,** *55*, 5989-5999.
2. Nagahama, K.; Ogawa, T.; Fujii, T.; Tazaki, M.; Tanase, S.; Morino, Y.; Fukuda, H. *Microbiology* **1991,** *137*, 2281-2286.
3. Fukuda, H.; Ogawa, T.; Tazaki, M.; Nagahama, K.; Fujiil, T.; Tanase, S.; Morino, Y. *Biochem. Biophys. Res. Commun.* **1992,** *188*, 483-489.
4. Doublié, S. *Macromolecular crystallography protocols: Volume 1, preparation and crystallization of macromolecules*, Walker, J. M.; Doublié, S., Eds. Humana Press: Totowa, NJ, 2007; pp 91-108.
5. Sparta, K. M.; Krug, M.; Heinemann, U.; Mueller, U.; Weiss, M. S. *J. Appl. Crystallogr.* **2016,** *49*, 1085-1092.
6. Battye, T. G.; Kontogiannis, L.; Johnson, O.; Powell, H. R.; Leslie, A. G. *Acta Crystallogr.* **2011,** *D67*, 271-281.
7. Evans, P. R.; Murshudov, G. N. *Acta Crystallogr.* **2013,** *D69*, 1204-1214.
8. 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. *Acta Crystallogr.* **2010,** *D66*, 213-221.
9. Terwilliger, T. C.; Adams, P. D.; Read, R. J.; McCoy, A. J.; Moriarity, N. W.; Grosse-Kunstleve, R. W.; Afonine, P. V.; Zwart, P. H.; Hung, L.-W. *Acta Crystallogr.* **2009,** *D65*, 582-601.
10. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. *Acta Crystallogr.* **2010,** *D66*, 486-501.
11. Chen, V. B.; Arendall, B., III; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. *Acta Crystallogr.* **2010,** *D66*, 12-21.
12. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004,** *25*, 1605-1612.
13. Laskowski, R. A.; Swindells, M. B. *J. Chem. Inf. Modeling* **2011,** *51*, 2778-2786.
14. Weingart, H.; Völksch, B.; Ullrich, M. S. *Phytopathology* **1999,** *89*, 360-365.
15. Johansson, N.; Persson, K. O.; Larsson, C.; Norbeck, J. *BMC Biochem.* **2014,** *15*, 1-8.
16. Robert, X.; Gouet, P. *Nucl. Acids Res.* **2014,** *42*, W320-W324.
17. 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. *Proc. Natl. Acad. Sci. U.S.A.* **2017,** *114*, 4667-4672.
18. Chang, W.-c.; Guo, Y.; Wang, C.; Butch, S. E.; Rosenzweig, A. C.; Boal, A. K.; Krebs, C.; Bollinger, J. M., Jr. *Science* **2014,** *343*, 1140-1144.
19. Welford, R. W. D.; Clifton, I. J.; Turnbull, J. J.; Wilson, S. C.; Schofield, C. J. *Org. Biomolec. Chem.* **2005,** *3*, 3117-3126.
20. Bräuer, A.; Beck, P.; Hintermann, L.; Groll, M. *Angew. Chemie* **2016,** *55*, 422-426.
21. Chowdhury, R.; McDonough, M. A.; Mecinovíc, J.; Loenarz, C.; Flashman, E.; Hewitson, K. S.; Domene, C.; Schofield, C. J. *Structure* **2009,** *17*, 981-989.
22. Hu, L.; Li, Z.; Cheng, J.; Rao, Q.; Gong, W.; Liu, M.; Shi, Y. G.; Zhu, J.; Wang, P.; Xu, Y. *Cell* **2013,** *155*, 1545-1555.
23. Yang, C.-G.; Yi, C.; Duguid, E. M.; Sullivan, C. T.; Jian, X.; Rice, P. A.; He, C. *Nature* **2008,** *452*, 961-965.
24. Nagahama, K.; Yoshino, K.; Matsuoka, M.; Tanase, S.; Ogawa, T.; Fukuda, H. *J. Ferment. Bioeng.* **1998,** *85*, 255-258.