Supporting Information

Fundamental Insights into Proton-Coupled Electron Transfer in Soybean Lipoxygenase from Quantum Mechanical/Molecular Mechanical Free Energy Simulations

Pengfei Li, Alexander V. Soudackov, and Sharon Hammes-Schiffer*

Department of Chemistry, University of Illinois at Urbana−Champaign, 600 South Mathews Avenue, Urbana, Illinois 61801, United States

Department of Chemistry, Yale University, 225 Prospect Street, New Haven, Connecticut 06520, United States

Corresponding author: sharon.hammes-schiffer@yale.edu

TABLE OF CONTENTS

Protocol for Simulation of the WT System

A. System preparation

1) The PDB entry 3PZW[1] (resolution: 1.4 Å, R-factor: 0.149, R-free: 0.169) was used for the modeling. Residues 1–5, 21–30, and 117–120, together with the sidechain of residue 6 are missing in the original structure. The iron center in the crystal structure is presumed to have an oxidation state of +2, and residue 870 is the water molecule coordinated to the Fe^{2+}, corresponding to the product of PCET. For modeling of the reactant, the oxidation state of Fe was +3, and a hydroxyl group was coordinated to it. The 1,2-ethanediol molecules and acetate ions in the original PDB structure were not considered in the current modeling.

2) The profix program in the Jackal software package[2] was utilized to add the missing heavy atoms (without consideration of the metal ion and crystal waters), and then the conref program in the Jackal software package was used to refine the structures of residues 1–6, 21–30, and 117–120 in three sequential steps.

3) The hydrogen atoms of the protein were added with the H++ program[3] under pH=8. Because the H++ program does not consider the metal ion, ligands, or water molecules in this process, the protonation states of the metal ligating residues (HID499, HID504, HID690, ASN694, ILE839) were manually corrected afterward as necessary. These protonation states are consistent with the work of Ref. [4].

4) The AutodockTools[5] program was used to prepare the docking input files for the protein and ligand. The charges of the Fe^{3+}, oxygen, and hydrogen atoms in the coordinated OH group were chosen to be +3, −1, and 0, respectively, for the docking procedure. The position of the H atom in the coordinated OH group was determined by QM optimization of the small model generated by the MCPB.py program[6] in the AmberTools software package,[7] as discussed in more detail in Step 6. The carboxylic acid group of the linoleic acid was deprotonated, leading to a negatively charged substrate. The AutodockVina[8] program was used to dock the negatively charged linoleic acid into the protein binding pocket. Only the protein, including the iron, was considered in this process -- water molecules, including crystal waters, were not considered here to prevent potential steric conflicts with the ligand, as the 3PZW structure was crystalized without the substrate. During the docking process, the protein was treated as rigid, while the linoleic acid was represented without hydrogen atoms and treated as flexible (i.e., all the C—C single bonds, except the C—C bond in the tail, were treated as rotatable). A 20 Å × 20 Å × 20 Å box centered at the oxygen atom of the coordinating OH group was used for docking. After completing the docking process, a structure with the C_{11}—O distance 2.90 Å, the C_{11}—H_S—O angle (where H_S denotes the pro-S hydrogen) 164.69° (determined after adding hydrogen atoms to the linoleic acid with tleap), and the "carboxylate-in" orientation (with the carboxylate oriented toward the R707 residue) was chosen as the initial structure for the modeling below.

5) The crystal water molecules were kept except for the following cases: (1) any crystal water molecule with its oxygen atom within 2 Å of any added heavy atoms (which were missing in the original PDB file) was deleted — 4 crystal water molecules of this type were deleted; (2) any crystal water molecule with its oxygen within 3 Å of at least one heavy atom of residues 495, 707, and the linoleic acid was deleted because the "carboxylate-in" linoleic acid can potentially form hydrogen bonds with these two residues — 13 crystal water molecules of this type were deleted.

B. Force field parameterization

6) The MCPB.py program[6] in the AmberTools software package[7] was used to facilitate the metal site modeling. This program generates a small model and a large model used to determine the force field parameters of a metal site. Potential energy scans of the small model of the metal site were computed at the DFT B3LYP/6-31G* level of theory based on the optimized structure (at the same level of theory). These scans were used to generate the bond force constants for the ligands relative to the metal ion. The angle force constants involving the metal were obtained from the general AMBER force field (GAFF),[9] treating the metal as a heavy atom. The equilibrium bond lengths and angles involving the metal ion were obtained based on the H++ generated structure except that the equilibrium Fe—O—H angle was obtained from the QM optimized structure of the small model. The RESP charges for the metal site were determined based on DFT B3LYP/6-31G* calculations for the large model. The partial charge on Fe^{3+} was determined to be +0.93e. These force field parameters are given below.

7) The negatively charged linoleic acid structure was generated using the Graphic antechamber program.[10] The HF/6-31G* level of theory was used to optimize the structure and to calculate the electrostatic potential. Then the antechamber program in the AmberTools software package[7] was used to generate the RESP charges. These force field parameters are given below.

C. Equilibration of water and ions

8) The topology and coordinate files were generated using the tleap program in the AmberTools software package.[7] The AMBER ff14SB force field[11] was used for modeling the protein system, while GAFF[9] was used for modeling the linoleic acid. The protein-substrate complex was solvated with a rectangular TIP3P[12] water box (setting the minimum distance between the protein surface and box edge as 12 Å), and 103 Na^+ and 93 Cl^- ions were added to obtain ~0.150 M NaCl solution after the NPT equilibration. The force field parameters for Na^+ and Cl^- were obtained from Ref [13].

9) Two stages of minimization were performed. Each stage contains 2500 steps of minimization with the steepest descent algorithm followed by 2500 steps of minimization with the conjugate gradient algorithm. In the first stage of minimization, the heavy atoms of the protein and ligand were restrained with force constant 500 kcal/(mol•$Å^2$), and the positions of the other atoms (i.e., the water molecules, ions, and hydrogen atoms in the

protein and substrate) were free to move. In the second stage of minimization, the heavy atoms of the protein were restrained with force constant 500 kcal/(mol•Å^2), and the positions of the other atoms (i.e., the substrate, water molecules, ions, and hydrogen atoms in the protein) were free to move. In the second stage of minimization, two one-sided restraints were applied to restrain the linoleic acid structure into a physically bound region: when the C_{11}—O distance was greater than 3.1 Å, a harmonic restraint with force constant 50 kcal/(mol•Å^2) was applied, and when the C_{11}—H_S—O angle was less than 140°, a harmonic restraint with force constant 50 kcal/(mol•rad^2) was applied.

10) After these optimizations, 500 ps NVT equilibration was performed at 300 K, followed by 500 ps NPT equilibration at 300 K and 1 atm to further equilibrate the structure. As in the second stage of minimization, restraints with force constants 500 kcal/(mol•Å^2) were applied to the heavy atoms in the protein, as well as the two one-sided restraints on the C—O bond and C—H—O angle, for both the NVT equilibration and NPT equilibration simulations.

D. Preparation of string

11) The final snapshot of the NPT equilibration was used for preparation of the initial structures in the string simulations. Water molecules, Na^+, and Cl^- ions without any atoms within 11 Å of the protein and linoleic acid were removed. Subsequently, a QM/MM geometry minimization was performed with the MM region fixed to generate the reactant structure.

12) The product structure was created based on the optimized reactant structure (only changing the coordinates of the transferring proton) using GaussView,[14] followed by QM/MM geometry optimization with the MM region fixed.

13) A linear interpolation between the QM/MM minimized reactant and product structures was performed to create an initial transition state structure. The O—H_S and C_{11}—H_S distances in the structure were then modified to be similar (only changing the coordinates of the transferring proton) using GaussView. Subsequently, a QM/MM geometry minimization was performed, imposing a restraint with force constant 500 kcal/(mol•Å^2) to ensure equality of the O—H_S and C_{11}—H_S distances (while keeping the MM region fixed) to obtain the transition state structure.

14) The initial string was generated with 19 images defined by three reaction coordinates, R_{CH}, R_{OH}, and R_{CO}, based on a quadratic interpolation between the QM/MM optimized reactant, transition state, and product structures.

15) The MD equilibration of the MM core region, which was defined to include the MM residues with at least one atom within 18 Å of the iron, was performed for each image for 10 ps while fixing the QM region and the MM outer region, which was defined as the MM region excluding the MM core region.

16) In the next step, 100 fs QM/MM equilibration was carried out for each image with the MM outer region fixed, imposing a restraint with 100 kcal/(mol•Å^2) force constant on R_{CO} for all of the images, as well as a restraint with 100 kcal/(mol•Å^2) force constant on R_{CH} and R_{OH} for images 1–8 and 13–19 and 150 kcal/(mol•Å^2) force constant on R_{CH} and R_{OH} for images 9–12 to ensure adequate sampling of the transition state region. We have confirmed that the RMSD is stable after these restrained QM/MM equilibrations.

E. Generation of total string

17) After updating the string based on the QM/MM equilibration, the first cycle of QM/MM sampling was carried out for 100 fs for each image with the MM outer region fixed. In order to better sample the transition-state region, 4 more windows were added to the string after the first cycle of QM/MM sampling. They were generated through linear interpolation between images 8 and 9, 9 and 10, 10 and 11, and 11 and 12, respectively, resulting in 23 images in the new string. During the QM/MM sampling, a restraint with 100 kcal/(mol•Å^2) force constant was imposed on R_{CO} in all images, as well as a restraint with 100 kcal/(mol•Å^2) force constant imposed on R_{CH} and R_{OH} for images 1–8 and 17–23 and 250 kcal/(mol•Å^2) force constant imposed on R_{CH} and R_{OH} for images 9–16. Again, this assignment was intended to ensure adequate sampling of the transition state region.

18) In each subsequent iteration, the string was updated based on the sampling of the former iteration. The next iteration of QM/MM sampling was performed based on the updated string with evenly spaced images along the string. 25 iterations of this type were performed, yielding a total of 57.5 ps sampling for the string simulation. The convergence criteria are discussed in Step 21 below.

19) The binless WHAM procedure[15] was used to obtain the weights for the unbiasing procedure. The sampling data from all iterations were used, and the convergence criterion was set to 0.001 kcal/mol for this procedure. Subsequently, a binning process using the weights from the WHAM procedure was carried out to create the 2D free energy surfaces. The free energy of each image along the final string (i.e., the MFEP) was determined from a bilinear interpolation of the free energies at the four neighbouring bins surrounding the bin containing this image.

F. Extra sampling of the reactant region

20) In order to better characterize the free energy profile along R_{CO}, an extra sampling set was performed. This sampling was performed with 18 images using the same three reaction coordinates as in the string simulations. All of these images used the last snapshot of image 1 in cycle 11 of the string calculations as the initial structure, thereby representing the reactant structure. For this set of images, a restraint was imposed on R_{CO} spanning 2.1 to 3.8 Å with an interval of 0.1 Å, and all images had R_{CH} restrained to 1.09 Å, along with a restraint on R_{OH} to maintain approximate linearity of C—H—O, such that $R_{CO}-R_{CH} = R_{OH}$. The force constant was 100 kcal/mol•Å^{-2} for R_{CO}, 250 kcal/mol•Å^{-2} for

R_{CH}, and 50 kcal/mol•Å^{-2} for R_{OH}. Each image was simulated with 200 fs equilibration and 1 ps sampling, leading to a total of 18 ps sampling in the extra sampling set.

G. Convergence criteria

21) Two criteria were used for checking the convergence of the string calculation. The first criterion is the root-mean-square deviation (RMSD) of each reaction coordinate. The RMSD of each reaction coordinate R_k for a specified iteration was calculated relative to its value averaged over the previous five iterations according to the following prescription:

$$\mathrm{RMSD}\left(R_k\right)=\sqrt{\frac{1}{N}\sum_{i}^{\text{images}}\left(R_k^{(i)}-\bar{R}_k^{(i)}\right)^2}\ ,$$

where N is the number of images in the string, $R_k^{(i)}$ is the average value of the reaction coordinate R_k for the image i for the specified iteration, and $\bar{R}_k^{(i)}$ is this value averaged over the previous five iterations. When the RMSD of each reaction coordinate was less than 0.1 Å, the string was considered to be converged. The second criterion is the free energy profile along the string, comparing this free energy profile while including data from the five most recent data sets (i.e., data from iterations 1–21, 1–22, 1–23, 1–24, or 1–25). If the difference of the free energy barrier (i.e., free energy of activation) was less than 0.5 kcal/mol among these five profiles, the free energy profile was considered to be converged. Both of these convergence criteria must be satisfied. Illustration of the convergence of the string for the WT system is shown in Figure S1.

The initial minimizations (Step 9) and NVT and NPT equilibrations (Step 10) were performed with pmemd.cuda in the AMBER software package,[7] employing the particle mesh Ewald[16] method for long-range electrostatics. The QM/MM minimization with frozen MM region (Steps 11–13) was performed using the AMBER/Q-Chem interface.[17] The MM core region equilibration (Step 15), QM/MM equilibration (Step 16), QM/MM sampling for the string (Steps 17–18), and extra QM/MM sampling (Step 20) were performed with the CHARMM/Q-Chem interface.[18] During all of the QM/MM free energy simulations, the electrostatic embedded scheme with the EXGR (which indicates "exclusion of QM/MM electrostatic interactions of the MM host group") approach was employed to treat the QM/MM interface. The total charge of the QM region was 0, and the multiplicity was 6 throughout all simulations. A large cut-off of 999 Å (i.e., no cut-off for the system investigated) was used for the electrostatic and van der Waals interactions for the QM-QM, QM-MM, and MM-MM interactions. The time step was 1 fs for all classical and QM/MM MD simulations. All of the force constants mentioned above correspond to the k value in the potential energy function $U=k(r-r_{eq})^2$ for bonds or $U=k(\theta-\theta_{eq})^2$ for angles.

Protocol for Simulation of the DM System

The procedures for structure preparation and molecular simulation for the DM system were similar to those for the WT system except for the aspects mentioned below.

In Step 1, PDB entry 4WHA[19] (Resolution: 1.7 Å, R-Value Free: 0.179, R-Value Work: 0.137) was used. Residues 1–5, 22–30, 119–121 were missing in the original PDB file. Residue 1352 is the water molecule coordinated to the iron center.

In Step 2, the heavy atoms of residues 1–5, 22–30, and 119–121 were added with the profix program. These residues were refined in three sequential steps using the conref program.

In Step 4, the docked structure of linoleic acid after docking to the WT protein was used for the docking of the linoleic acid to the DM. The docking for the DM system was performed while keeping both the protein and ligand rigid. A structure with the C_{11}—O distance 2.93 Å, the C_{11}—H_S—O angle 170.64°, and the "carboxylate-in" orientation was chosen for further modeling.

In Step 5, the same criteria were used to delete the crystal water molecules. But in the two steps discussed, 9 and 20 water molecules were deleted, respectively.

In Step 6, the equilibrium bond lengths and angles of the bonds and angles involving the metal were obtained from the DM crystal structure except that the equilibrium Fe—O—H angle was obtained from the QM optimized structure of the small model for WT.

In Step 9, in the second stage of minimization, the metal site residues (HID499, HID504, HID690, ASN694, ILE839) were allowed to move because these coordinates changed somewhat during the profix procedure used to add the missing residues.

In Step 10, the metal site residues (HID499, HID504, HID690, ASN694, ILE839) were allowed to move in the NVT and NPT equilibrations.

In Step 14, the initial string of the DM had 23 rather than 19 images.

In Step 16, during the QM/MM equilibration procedure, restraints with force constants 100 kcal/(mol•Å^2) were imposed on R_{CO}, R_{CH}, and R_{OH} for images 1–8 and 17–23, while a restraint with force constant 100 kcal/(mol•Å^2) was imposed on R_{CO} and restraints with force constants 250 kcal/(mol•Å^2) were imposed on R_{CH} and R_{OH} for images 9–16.

In Step 17, during the QM/MM string sampling for the DM, the above restraints were used as well.

In Step 18, a total of 22 iterations were performed for the string simulation. The same three convergence criteria were satisfied for the DM as well, as shown in Figure S2. Illustration of the convergence of the string for the DM system was shown in Figure S2.

Data Analysis

1. 2D free energy surfaces and free energy profile along the MFEP

The string simulation data were used to generate the 2D free energy surfaces as a function of $R_{CH}-R_{OH}$ and R_{CO} and the free energy profiles along the MFEPs using the procedure explained in Step 19 of the protocol for simulation of the WT system. These data included 57.5 ps sampling for WT and 50.6 ps sampling for DM. The 2D free energy surfaces were generated using 20 bins in each dimension.

2. Potential of mean force along R_{CO}

All of the data from the string simulation and the extra sampling set were used to generate the free energy profile along R_{CO}. These profiles were obtained by generating the 2D free energy surface as a function of R_{CH} and R_{CO} (with a bin size of 0.1 Å for each) and evaluating the slice along R_{CO} for which R_{CH} = 1.09 Å (i.e., the approximate equilibrium CH distance). This slice was obtained through linear interpolation between the two closest slices on the 2D free energy surface.

3. Electrostatic field analysis

The extra sampling sets were used for the electrostatic field analysis for WT and DM. These data sets were unbiased with weights obtained from the binless WHAM procedure for each of them. The weighted properties on the 2D surfaces as a function of R_{CH} and R_{CO} were calculated with a bin size of 0.1 Å along R_{CH} and 0.2 Å along R_{CO}. Subsequently, the plots corresponding to R_{CH} = 1.09 Å were obtained through linear interpolation between the two closest slices on the 2D surfaces.

Force Field Parameters for the WT System

The force field parameters that were not obtained directly from AMBER ff14SB and GAFF are given below.

1. Atom type and charge parameters for HID499

ATOM NAME	ATOM TYPE	ATOMIC CHARGE	ATOM NAME	ATOM TYPE	ATOMIC CHARGE
N	N	-0.415700	HD1	H	0.338624
H	H	0.276302	CE1	CR	-0.003757
CA	CX	0.018800	HE1	H5	0.188746
HA	H1	0.126430	NE2	Y1	-0.103715
CB	CT	-0.017388	CD2	CV	-0.125977
HB2	HC	0.052480	HD2	H4	0.092614
HB3	HC	0.052480	C	C	0.597300
CG	CC	-0.010180	O	O	-0.567900
ND1	NA	-0.179198			

2. Atom type and charge parameters for HID504

ATOM NAME	ATOM TYPE	ATOMIC CHARGE	ATOM NAME	ATOM TYPE	ATOMIC CHARGE
N	N	-0.415700	HD1	H	0.311870
H	H	0.354326	CE1	CR	0.027984
CA	CX	0.018800	HE1	H5	0.170806
HA	H1	0.056469	NE2	Y2	-0.313349
CB	CT	-0.064295	CD2	CV	-0.173314
HB2	HC	0.029012	HD2	H4	0.171493
HB3	HC	0.029012	C	C	0.597300
CG	CC	0.124418	O	O	-0.567900
ND1	NA	-0.164102			

3. Atom type and charge parameters for HID690

ATOM NAME	ATOM TYPE	ATOMIC CHARGE	ATOM NAME	ATOM TYPE	ATOMIC CHARGE
N	N	-0.415700	HD1	H	0.289375
H	H	0.306099	CE1	CR	0.039008
CA	CX	0.018800	HE1	H5	0.168562
HA	H1	0.132188	NE2	Y3	-0.201557
CB	CT	-0.069235	CD2	CV	-0.280150
HB2	HC	0.069975	HD2	H4	0.217237
HB3	HC	0.069975	C	C	0.597300
CG	CC	0.063513	O	O	-0.567900
ND1	NA	-0.129844			

4. Atom type and charge parameters for ASN694

ATOM NAME	ATOM TYPE	ATOMIC CHARGE	ATOM NAME	ATOM TYPE	ATOMIC CHARGE

N	N	-0.415700	CG	C	0.387691
H	H	0.307233	OD1	Y4	-0.423627
CA	CX	0.014300	ND2	N	-0.670873
HA	H1	0.084969	HD21	H	0.380659
CB	2C	0.095243	HD22	H	0.318863
HB2	HC	-0.026526	C	C	0.597300
HB3	HC	-0.026526	O	O	-0.567900

5. Atom type and charge parameters for ILE839

ATOM NAME	ATOM TYPE	ATOMIC CHARGE	ATOM NAME	ATOM TYPE	ATOMIC CHARGE
N	N	-0.454405	CG1	2C	-0.047906
H	H	0.219946	HG12	HC	0.031226
CA	CX	0.013321	HG13	HC	0.031226
HA	H1	0.103682	CD1	CT	-0.105021
CB	3C	0.074332	HD11	HC	0.031356
HB	HC	0.068305	HD12	HC	0.031356
CG2	CT	-0.174699	HD13	HC	0.031356
HG21	HC	0.023543	C	C	0.570959
HG22	HC	0.023543	O	O2	-0.571432
HG23	HC	0.023543	OXT	Y5	-0.558225

6. Atom type and charge parameters for the iron

ATOM NAME	ATOM TYPE	ATOMIC CHARGE
FE	M1	0.931471

7. Atom type and charge parameters for the hydroxyl group

ATOM NAME	ATOM TYPE	ATOMIC CHARGE
O	Y6	-0.636765
H	HO	0.463747

8. Atom type and charge parameters for the linoleic acid

ATOM NAME	ATOM TYPE	ATOMIC CHARGE	ATOM NAME	ATOM TYPE	ATOMIC CHARGE
C1	c	0.894951	H7	hc	0.020918
C2	c3	-0.239299	H8	hc	0.020918
O1	o	-0.843408	H9	hc	-0.004524
O2	o	-0.843408	H10	hc	-0.004524
C3	c3	0.180355	H11	hc	0.018625
C4	c3	-0.117048	H12	hc	0.018625
C5	c3	-0.113655	H13	hc	0.027930
C6	c3	0.139032	H14	hc	0.027930
C7	c3	-0.147389	H15	ha	0.118892

C8	c3	0.078273	H16	ha	0.133508
C9	c2	-0.202747	H17	hc	0.020956
C10	c2	-0.304976	H18	hc	0.020956
C11	c3	0.231948	H19	ha	0.126853
C12	c2	-0.246769	H20	ha	0.128495
C13	c2	-0.282278	H21	hc	0.002257
C14	c3	0.125355	H22	hc	0.002257
C15	c3	0.123721	H23	hc	-0.035361
C16	c3	-0.087376	H24	hc	-0.035361
C17	c3	0.073514	H25	hc	0.016036
C18	c3	-0.137118	H26	hc	0.016036
H1	hc	0.013490	H27	hc	-0.012163
H2	hc	0.013490	H28	hc	-0.012163
H3	hc	-0.020626	H29	hc	0.029704
H4	hc	-0.020626	H30	hc	0.029704
H5	hc	0.013192	H31	hc	0.029704
H6	hc	0.013192			

8. Mass and nonbonded parameters for the metal ion and ligating atoms

```
MASS
M1 55.85                Fe ion
Y1 14.01      0.530     sp2 N in 5 memb.ring w/LP (HIS,ADE,GUA)
Y2 14.01      0.530     sp2 N in 5 memb.ring w/LP (HIS,ADE,GUA)
Y3 14.01      0.530     sp2 N in 5 memb.ring w/LP (HIS,ADE,GUA)
Y4 16.00      0.434     carbonyl group oxygen
Y5 16.00      0.434     carboxyl and phosphate group oxygen
Y6 16.00      0.465     oxygen in hydroxyl group

NONB
 M1       1.3860  0.0135709700    IOD set for Fe3+ ion in TIP3P water from Li et al. JPCB, 2015, 119, 883
 Y1       1.8240  0.1700          OPLS
 Y2       1.8240  0.1700          OPLS
 Y3       1.8240  0.1700          OPLS
 Y4       1.6612  0.2100          OPLS
 Y5       1.6612  0.2100          OPLS
 Y6       1.7210  0.2104          OPLS
```

9. Metal involving bond and angle parameters for the metal site in WT

```
BOND
M1-Y6  253.6    2.1120
Y1-M1   73.0    2.2370
Y2-M1   69.8    2.3480
Y3-M1   51.6    2.2660
Y4-M1   75.6    2.8480
Y5-M1  119.5    2.3160

ANGL
C -Y4-M1    70.00     119.63
C -Y5-M1    70.00     129.88
CR-Y1-M1    70.00     117.91
CR-Y2-M1    70.00     119.26
CR-Y3-M1    70.00     129.45
M1-Y1-CV    70.00     131.57
M1-Y2-CV    70.00     131.25
M1-Y3-CV    70.00     124.04
```

M1-Y6-HO	50.00	112.07
Y1-M1-Y2	70.00	95.21
Y1-M1-Y3	70.00	100.45
Y1-M1-Y4	70.00	73.34
Y1-M1-Y5	70.00	167.31
Y1-M1-Y6	70.00	86.98
Y2-M1-Y3	70.00	102.17
Y2-M1-Y4	70.00	163.85
Y2-M1-Y5	70.00	94.43
Y2-M1-Y6	70.00	96.86
Y3-M1-Y4	70.00	91.28
Y3-M1-Y5	70.00	85.52
Y3-M1-Y6	70.00	158.77
Y4-M1-Y5	70.00	95.50
Y4-M1-Y6	70.00	71.67
Y5-M1-Y6	70.00	83.69

Force Field Parameters for the DM System

The charge and atom type parameters for the metal site residues in the DM system are the same as those in the WT system. The force field parameters that are different for the DM system than for the WT system are given below.

1. Mass and nonbonded parameters for the metal ion and ligating atoms

```
MASS
M1 55.85                 Fe ion
Y1 14.01      0.530      sp2 N in 5 memb.ring w/LP (HIS,ADE,GUA)
Y2 14.01      0.530      sp2 N in 5 memb.ring w/LP (HIS,ADE,GUA)
Y3 14.01      0.530      sp2 N in 5 memb.ring w/LP (HIS,ADE,GUA)
Y4 16.00      0.434      carbonyl group oxygen
Y5 16.00      0.434      carboxyl and phosphate group oxygen
Y6 16.00      0.465      oxygen in hydroxyl group

NONB
 M1       1.3860  0.0135709700    IOD set for Fe3+ ion in TIP3P water from Li et al. JPCB, 2015, 119, 883
 Y1       1.8240  0.1700          OPLS
 Y2       1.8240  0.1700          OPLS
 Y3       1.8240  0.1700          OPLS
 Y4       1.6612  0.2100          OPLS
 Y5       1.6612  0.2100          OPLS
 Y6       1.7210  0.2104          OPLS
```

2. Metal involving bond and angle parameters for the metal site in the DM

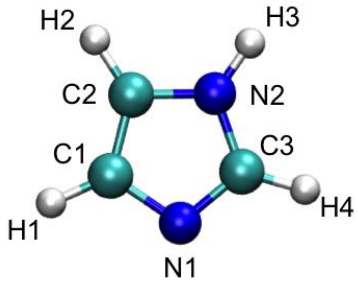
```
BOND
M1-Y6  253.6    2.1360
Y1-M1   73.0    2.1620
Y2-M1   69.8    2.2350
Y3-M1   51.6    2.2600
Y4-M1   75.6    3.1650
Y5-M1  119.5    2.3440

ANGL
C -Y4-M1    70.00    104.49
C -Y5-M1    70.00    132.72
CR-Y1-M1    70.00    115.71
CR-Y2-M1    70.00    123.54
CR-Y3-M1    70.00    126.94
CV-Y1-M1    70.00    135.11
CV-Y2-M1    70.00    127.05
CV-Y3-M1    70.00    124.33
M1-Y6-HO    50.00    112.07
Y1-M1-Y2    70.00     88.81
Y1-M1-Y3    70.00    104.14
Y1-M1-Y4    70.00     77.30
Y1-M1-Y5    70.00    167.17
Y1-M1-Y6    70.00     85.60
Y2-M1-Y3    70.00    102.36
Y2-M1-Y4    70.00    162.78
Y2-M1-Y5    70.00     94.63
Y2-M1-Y6    70.00     93.92
Y3-M1-Y4    70.00     91.00
Y3-M1-Y5    70.00     87.22
Y3-M1-Y6    70.00    161.07
Y4-M1-Y5    70.00     96.85
Y4-M1-Y6    70.00     75.13
Y5-M1-Y6    70.00     81.84
```

Force Field Parameters for the Gas Phase Model System

The partial charge parameters for the metal site in the gas phase model was obtained based on a geometry optimization and the RESP fitting procedure at the B3LYP/6-31G** level of theory. The partial charge parameters for the (2Z,5Z)-2,5-heptadiene ligand were obtained based on a geometry optimization and the RESP fitting procedure at the HF/6-31G* level of theory. The AMBER ff14SB force field was used to model the metal site (i.e., the metal ion with coordinated ligands), while GAFF was used for the (2Z,5Z)-2,5-heptadiene ligand.

1. Atom type and charge parameters for the first imidazole group



ATOM NAME	ATOM TYPE	ATOMIC CHARGE	ATOM NAME	ATOM TYPE	ATOMIC CHARGE
N1	Y1	-0.040648	N2	NA	-0.256228
C1	CV	-0.240415	H3	H	0.348465
H1	H4	0.175619	C3	CR	0.027466
C2	CW	-0.038837	H4	H5	0.159283
H2	H4	0.165177			

2. Atom type and charge parameters for the second imidazole group

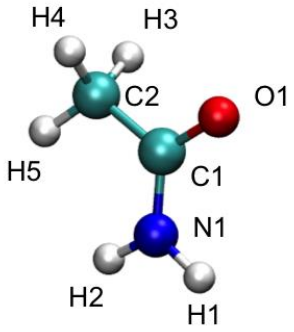
ATOM NAME	ATOM TYPE	ATOMIC CHARGE	ATOM NAME	ATOM TYPE	ATOMIC CHARGE
N1	Y2	-0.040648	N2	NA	-0.256228
C1	CV	-0.240415	H3	H	0.348465
H1	H4	0.175619	C3	CR	0.027466
C2	CW	-0.038837	H4	H5	0.159283
H2	H4	0.165177			

3. Atom type and charge parameters for the third imidazole group

ATOM NAME	ATOM TYPE	ATOMIC CHARGE	ATOM NAME	ATOM TYPE	ATOMIC CHARGE

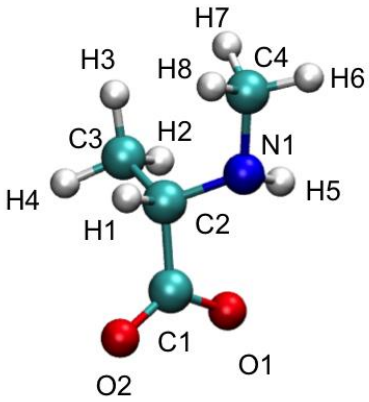
N1	Y3	-0.040648	N2	NA	-0.256228
C1	CV	-0.240415	H3	H	0.348465
H1	H4	0.175619	C3	CR	0.027466
C2	CW	-0.038837	H4	H5	0.159283
H2	H4	0.165177			

4. Atom type and charge parameters for the acetamide group



ATOM NAME	ATOM TYPE	ATOMIC CHARGE	ATOM NAME	ATOM TYPE	ATOMIC CHARGE
O1	Y4	-0.610981	C2	CT	-0.563156
C1	C	0.797172	H3	HC	0.168873
N1	N	-0.785284	H4	HC	0.168873
H1	H	0.398946	H5	HC	0.168873
H2	H	0.398946			

5. Atom type and charge parameters for the N-methylalanine group



ATOM NAME	ATOM TYPE	ATOMIC CHARGE	ATOM NAME	ATOM TYPE	ATOMIC CHARGE
O1	Y5	-0.578446	H4	HC	0.021960

C1	C	0.651028	N1	NT	-0.548544
O2	O2	-0.578446	H5	H	0.293525
C2	CT	0.037878	C4	CT	-0.130519
H1	H1	0.091543	H6	H1	0.085916
C3	CT	-0.049941	H7	H1	0.085916
H2	HC	0.021960	H8	H1	0.085916
H3	HC	0.021960			

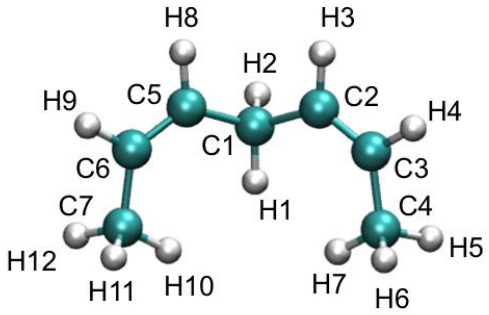
6. Atom type and charge parameters for the metal ion

ATOM NAME	ATOM TYPE	ATOMIC CHARGE
FE	M1	0.555401

7. Atom type and charge parameters for the hydroxyl group

ATOM NAME	ATOM TYPE	ATOMIC CHARGE
O3	Y6	-0.700534
H25	HO	0.377958

8. Atom type and charge parameters for the (2Z,5Z)-2,5-Heptadiene



ATOM NAME	ATOM TYPE	ATOMIC CHARGE	ATOM NAME	ATOM TYPE	ATOMIC CHARGE
H1	hc	-0.074822	H7	hc	0.066348
C1	c3	0.597683	C5	c2	-0.463758
H2	hc	-0.074822	H8	ha	0.162521
C2	c2	-0.463758	C6	c2	-0.057211
H3	ha	0.162521	H9	ha	0.103115
C3	c2	-0.057211	C7	c3	-0.167731
H4	ha	0.103115	H10	hc	0.066348
C4	c3	-0.167731	H11	hc	0.066348
H5	hc	0.066348	H12	hc	0.066348
H6	hc	0.066348			

9. Additional parameters for the metal site ligands in the gas phase model

```
BOND
CV-CW  518.00   1.371       same as CC-CW, penalty score=  0.1

ANGLE
CW-CV-NB   70.000     120.000   same as CC-CV-NB, penalty score=  0.1
CV-CW-H4   50.000     120.000   same as CC-CW-H4, penalty score=  0.1
CV-CW-NA   70.000     120.000   same as CC-CW-NA, penalty score=  0.1
CW-CV-H4   50.000     120.000   same as CC-CV-H4, penalty score=  0.1
C -CT-NT   80.000     111.200   same as C -CT-N3, penalty score=  0.1

DIHE
NB-CV-CW-H4   4   21.500       180.000          2.000      same as X -CC-CW-X , penalty score= 86.0
NB-CV-CW-NA   4   21.500       180.000          2.000      same as X -CC-CW-X , penalty score= 86.0
H4-CV-CW-H4   4   21.500       180.000          2.000      same as X -CC-CW-X , penalty score= 86.0
H4-CV-CW-NA   4   21.500       180.000          2.000      same as X -CC-CW-X , penalty score= 86.0

IMPROPER
CW-H4-CV-NB        1.1         180.0         2.0          Using the default value
CV-H4-CW-NA        1.1         180.0         2.0          Using the default value
CR-CW-NA-H         1.0         180.0         2.0          Using general improper torsional angle  X- X-NA- H, penalty score=  6.0)
H5-NA-CR-NB        1.1         180.0         2.0          Same as X -X -CR-H5, penalty score=  6.0 (use general term))
CT-N -C -O         10.5          180.0         2.0          Using general improper torsional angle  X- X- C- O, penalty score=  6.0)
C -H -N -H         1.0         180.0         2.0          Using general improper torsional angle  X- X- N- H, penalty score=  6.0)
CT-O2-C -O2        10.5          180.0         2.0          Using general improper torsional angle  X-O2- C-O2, penalty score=  3.0)
```

10. Mass and nonbonded parameters for the metal ion and ligating atoms

```
MASS
M1 55.85       Fe ion
Y1 14.01       sp2 N in 5 memb.ring w/LP (HIS,ADE,GUA)
Y2 14.01       sp2 N in 5 memb.ring w/LP (HIS,ADE,GUA)
Y3 14.01       sp2 N in 5 memb.ring w/LP (HIS,ADE,GUA)
Y4 16.00       carbonyl group oxygen
Y5 16.00       carboxyl and phosphate group oxygen
Y6 16.00       oxygen in hydroxyl group

NONB
 M1        1.3860  0.0135709700      IOD set for Fe3+ ion in TIP3P water from Li et al. JPCB, 2015, 119, 883
 Y1        1.8240  0.1700            OPLS
 Y2        1.8240  0.1700            OPLS
 Y3        1.8240  0.1700            OPLS
 Y4        1.6612  0.2100            OPLS
 Y5        1.6612  0.2100            OPLS
 Y6        1.7210  0.2104            OPLS
```

11. Bond and angle parameters involving the metal in the gas phase model

The bond/angle force constants in the gas phase model were obtained from the bond/angle parameters determined for the metal site in WT SLO. Because the three imidazole groups are the same in the gas phase model, their bond parameters with iron were obtained by averaging the bond force constants for the three Histidine residues with iron. The bond/angle equilibrium values were obtained based on the QM optimized geometry of the metal site at the B3LYP/6-31G** level of theory.

```
BOND
M1-Y1   64.8    2.2070
M1-Y2   64.8    2.2120
```

M1-Y3 64.8 2.1850
M1-Y4 75.6 2.1640
M1-Y5 119.5 2.0800
M1-Y6 253.6 1.9470

ANGL
M1-Y1-CR 50.0 124.72
M1-Y1-CV 50.0 128.48
M1-Y2-CR 50.0 124.04
M1-Y2-CV 50.0 129.68
M1-Y3-CR 50.0 122.12
M1-Y3-CV 50.0 130.44
M1-Y4-C 50.0 135.87
M1-Y5-C 50.0 126.92
M1-Y6-HO 35.0 112.70
Y2-M1-Y1 50.0 91.41
Y3-M1-Y1 50.0 90.42
Y3-M1-Y2 50.0 90.93
Y4-M1-Y1 50.0 85.09
Y4-M1-Y2 50.0 84.10
Y4-M1-Y3 50.0 173.21
Y5-M1-Y1 50.0 174.89
Y5-M1-Y2 50.0 84.22
Y5-M1-Y3 50.0 87.02
Y5-M1-Y4 50.0 97.05
Y6-M1-Y1 50.0 92.70
Y6-M1-Y2 50.0 173.04
Y6-M1-Y3 50.0 94.65
Y6-M1-Y4 50.0 90.65
Y6-M1-Y5 50.0 91.91

Coordinates of the Docked Ligand for the WT System*

ATOM NAME	*x*	*y*	*z*	ATOM NAME	*x*	*y*	*z*
C1	14.379	48.738	8.464	C11	23.267	48.234	12.033
C2	14.265	50.232	8.039	C12	24.227	48.390	13.189
C3	14.892	51.228	9.010	C13	25.506	48.725	13.159
C4	16.404	51.395	8.821	C14	26.353	49.042	11.952
C5	17.238	51.213	10.096	C15	26.467	50.545	11.647
C6	18.559	50.466	9.888	C16	26.214	51.002	10.201
C7	19.698	51.021	10.746	C17	27.456	51.419	9.389
C8	20.663	49.949	11.277	C18	28.762	51.512	10.183
C9	22.101	50.200	10.900	O1	14.846	48.527	9.588
C10	23.178	49.501	11.214	O2	14.001	47.934	7.608

*Only the coordinates of the heavy atoms are shown because the docking was performed using the united atom representation for the linoleic acid. Hydrogen atoms were added prior to the QM/MM simulations.

Coordinates of the Docked Ligand for the DM System*

ATOM NAME	*x*	*y*	*z*	ATOM NAME	*x*	*y*	*z*
C1	13.945	47.919	8.148	C11	22.791	48.547	11.800
C2	13.702	49.353	7.591	C12	23.717	48.891	12.943
C3	14.222	50.484	8.473	C13	24.960	49.338	12.891
C4	15.715	50.772	8.279	C14	25.792	49.623	11.665
C5	16.545	50.780	9.570	C15	25.774	51.098	11.228
C6	17.931	50.141	9.437	C16	25.500	51.401	9.746
C7	19.003	50.871	10.249	C17	26.711	51.856	8.907
C8	20.053	49.943	10.879	C18	27.992	52.138	9.698
C9	21.468	50.290	10.489	O1	14.413	47.852	9.289
C10	22.599	49.723	10.871	O2	13.653	47.012	7.364

* Only the coordinates of the heavy atoms are shown because the docking was performed using the united atom representation for the linoleic acid. Hydrogen atoms were added prior to the QM/MM simulations.

Figures

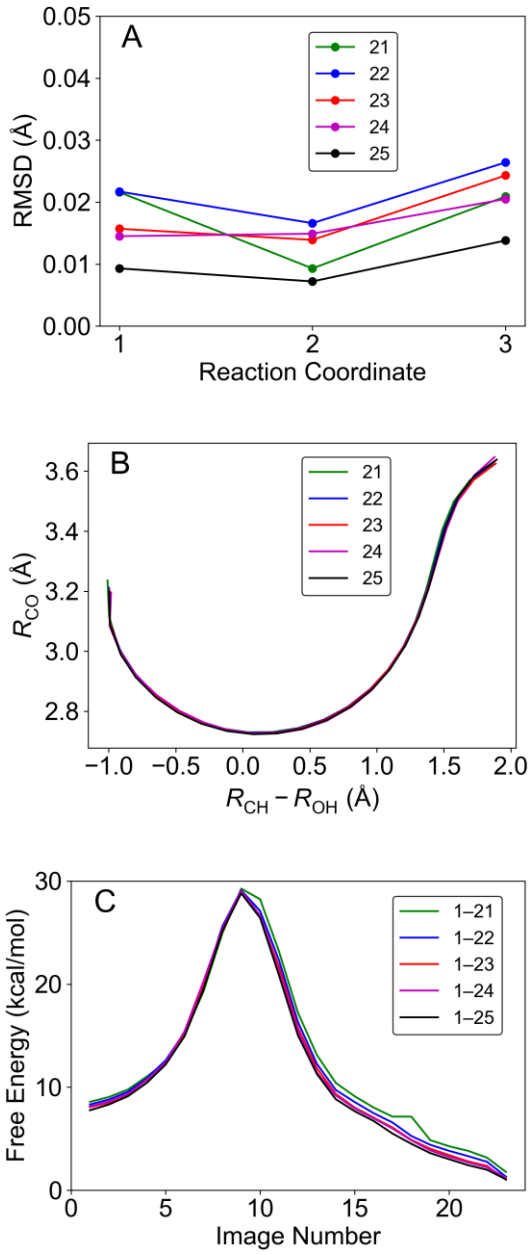


Figure S1. Illustration of convergence of the string for the WT system. (A) The root-mean-square deviation (RMSD) of each reaction coordinate for the last five iterations, with each one compared to the average value of the previous five iterations; reaction coordinates 1, 2, and 3 correspond to R_{CO}, R_{OH}, and R_{CH}, respectively. (B) The projection of the string onto the R_{CH}–R_{OH} and R_{CO} axes for the last five iterations. (C) Free energy profiles along the MFEPs for the data sets including iterations 1–21, 1–22, 1–23, 1–24, and 1–25.

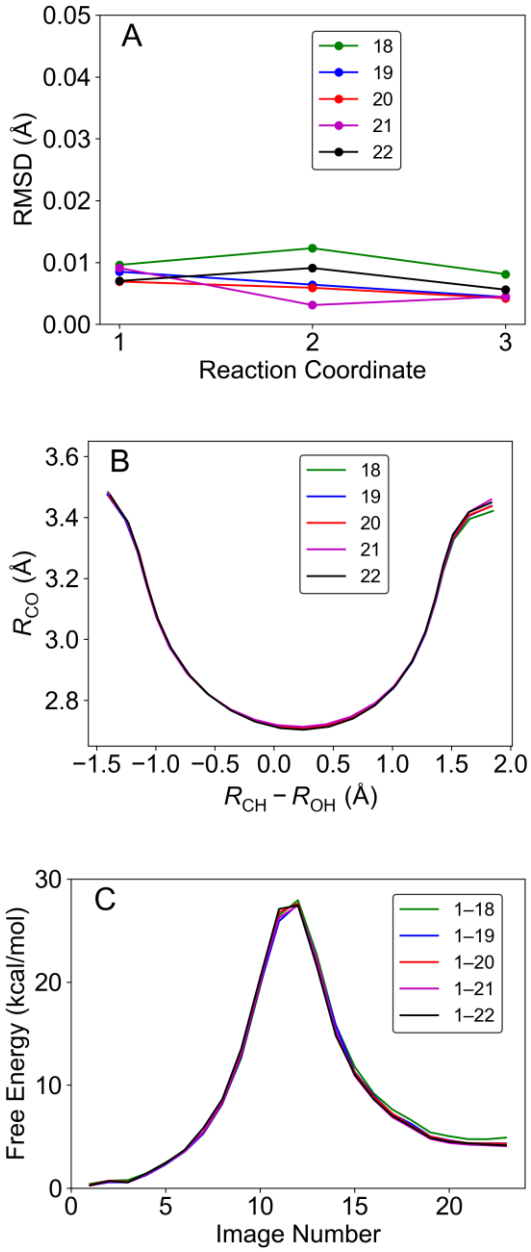


Figure S2. Illustration of convergence of the string for the DM system. (A) The root-mean-square deviation (RMSD) of each reaction coordinate for the last five iterations, with each one compared to the average value of the previous five iterations; reaction coordinates 1, 2, and 3 correspond to R_{CO}, R_{OH}, and R_{CH}, respectively. (B) The projection of the string onto the R_{CH}–R_{OH} and R_{CO} axes for the last five iterations. (C) Free energy profiles along the MFEPs for the data sets including iterations 1–18, 1–19, 1–20, 1–21, and 1–22.

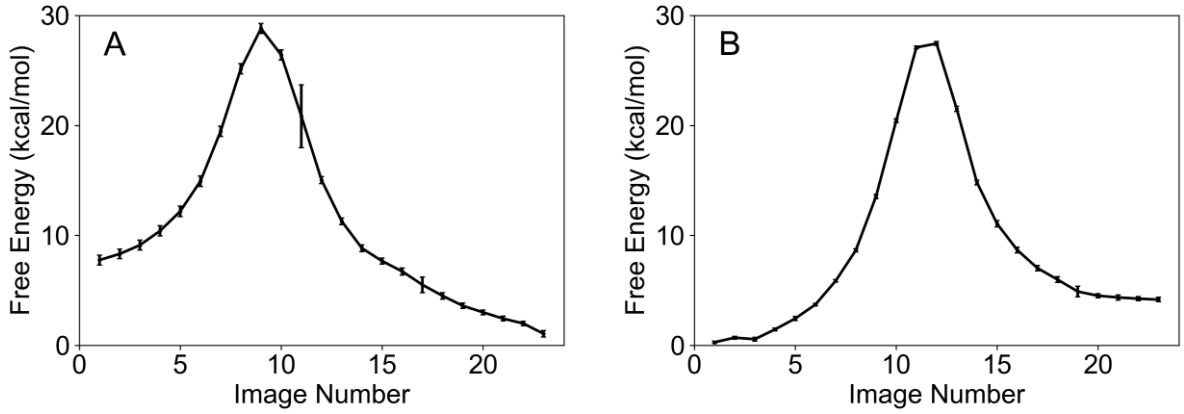


Figure S3. Free energies along the MFEPs with error bars for (A) WT and (B) DM SLO. These error bars correspond to the statistical errors and were obtained using bootstrapping error analysis[20] with nine "fake" data sets. Note that these error bars do not account for errors arising from the level of theory used to generate the potential energy surface and the limited conformational sampling.

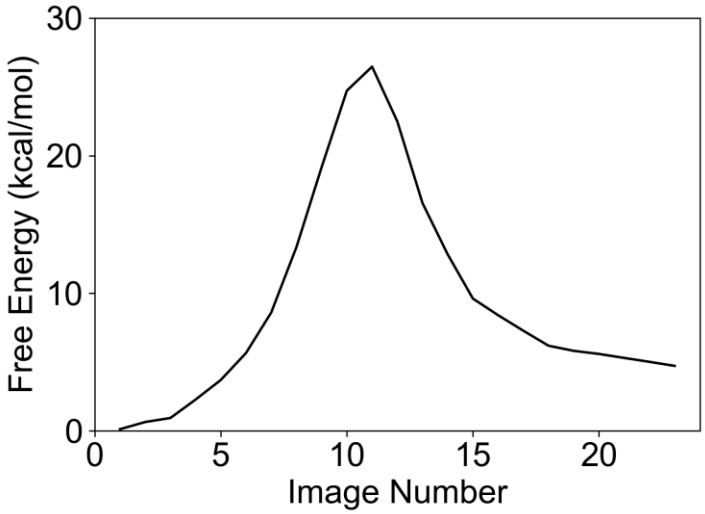


Figure S4. The free energy profile along the MFEP for the DM obtained from an independent string simulation using an alternative protocol. In this protocol, the DM system was built from the WT crystal structure, namely PDB entry 3PZW, using the same substrate coordinates as in the WT system, and mutating the two residues L564 and L754 to Ala using the tleap program in the AmberTools package. This free energy profile is based on 13.8 ps QM/MM sampling and illustrates that the endoergicity of the PCET reaction for the DM (inset of Figure 3B) is reproducible for an independent calculation starting from a completely different initial structure. Note that the string simulation for the DM presented in the main paper started with a DM crystal structure and docking of the linoleic acid substrate to this structure.

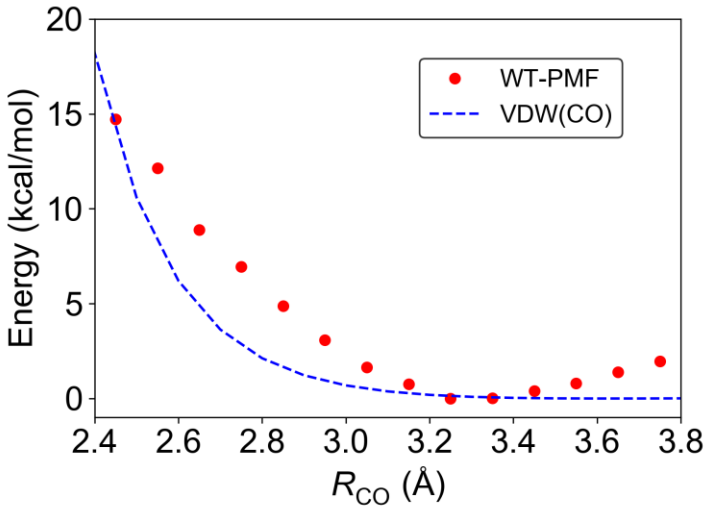


Figure S5. The potential of mean force (i.e., free energy) for WT SLO and the van der Waals (VDW) interaction energy between a carbon and oxygen atom as a function of R_{CO}. Both curves are shifted vertically so that the minimum corresponds to zero energy. The van der Waals parameters for C and O are from the atom types CT ($R_{min}/2$ = 1.908 Å, ε = 0.1094 kcal/mol) and OH ($R_{min}/2$ = 1.721 Å, ε = 0.2104 kcal/mol) in the AMBER ff14SB force field. Here $R_{min}/2$ is the van der Waals radius.

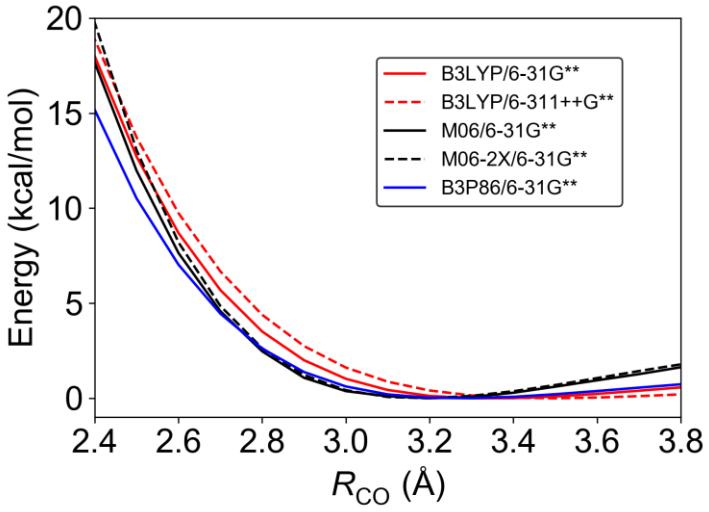


Figure S6. Rigid potential energy scans along R_{CO} for the gas phase model depicted in Figure 5 obtained at different levels of theory. The energies were computed by starting with the transition state geometry for this complex and retaining the geometries of the individual substrate and Fe complex components as R_{CO} is changed.

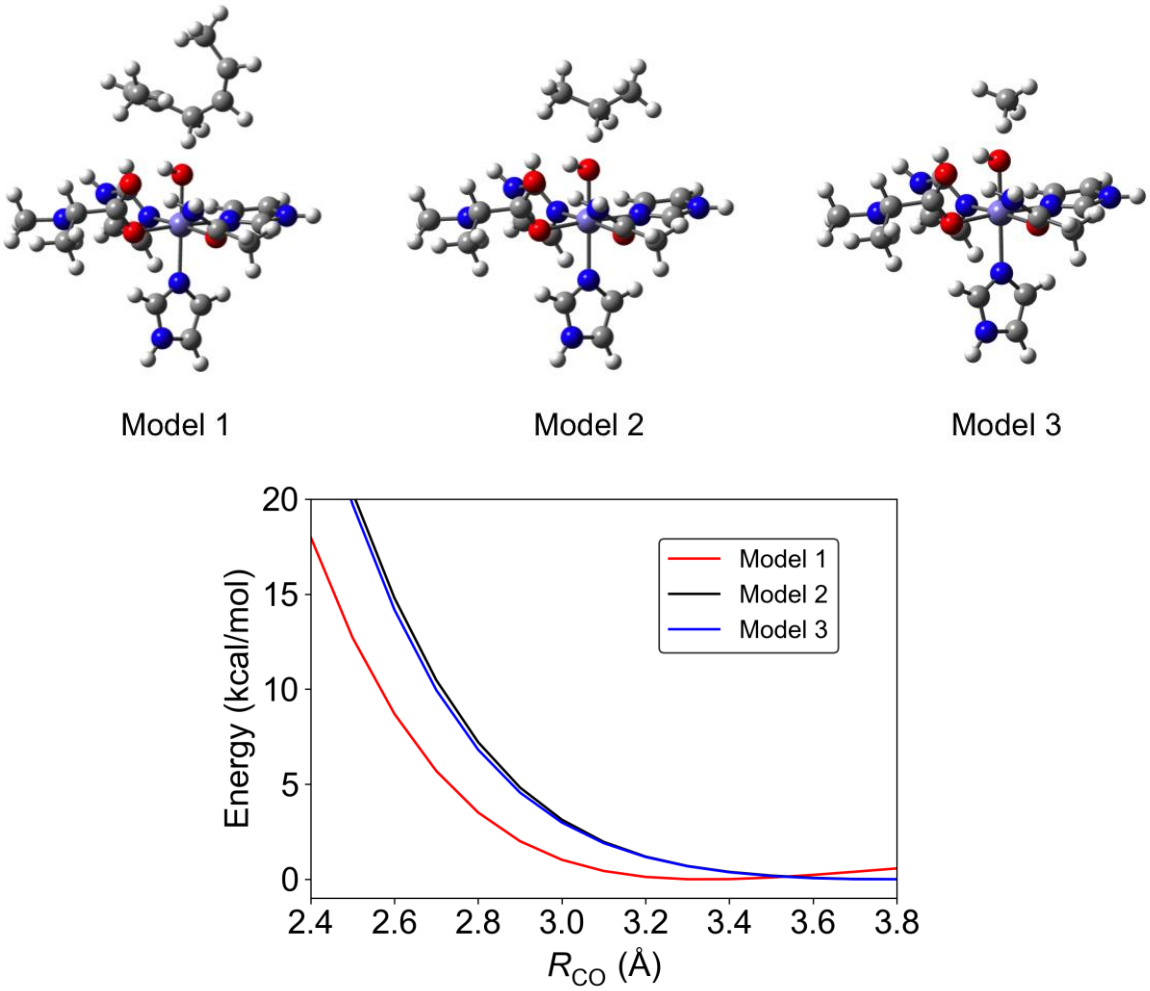


Figure S7. Rigid potential energy scans along R_{CO} for the three gas phase models depicted at the top of the figure. Model 1 is the cluster used to generate the data in Figure 5 and Figure S4, while Models 2 and 3 are the same as this cluster except that C_3H_8 or CH_4 replace C_7H_{12} to represent the linoleic acid substrate. This figure illustrates that the π-backbone of the substrate significantly softens the potential energy surface.

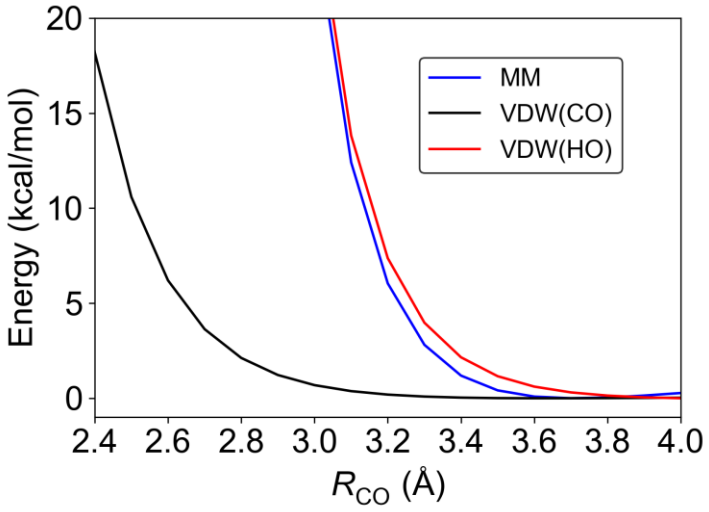


Figure S8. Rigid potential energy scans along R_{CO} for the gas phase model depicted in Figure 5 using an MM force field (blue curve here, which is identical to the blue dashed curve in Figure 5), as compared to the van der Waals (VDW) interaction energy between the carbon and oxygen atom (black curve) and the VDW interaction energy between the H and O atoms at the CH---O interface (red curve). This figure illustrates that the strong repulsive interaction for the MM force field potential at shorter distances is mainly due to the VDW interaction between the H and O atoms at the proton transfer interface.

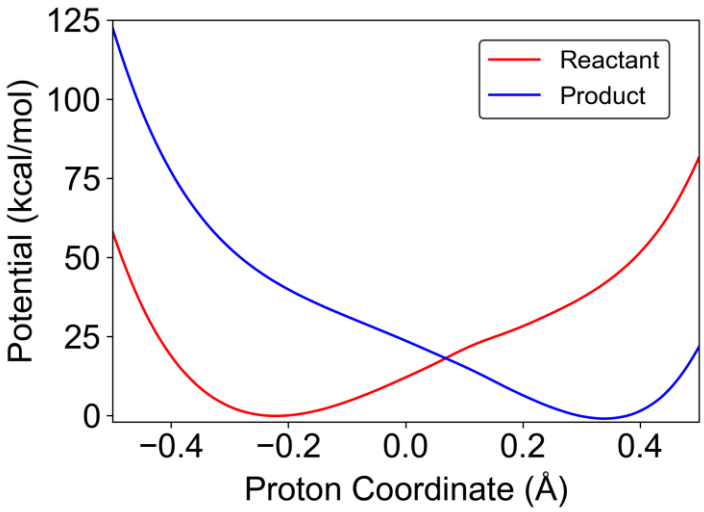


Figure S9. The diabatic proton potential energy curves for the reactant and product states calculated with CDFT-CI/ωB97X/6-31G**. The structures were obtained by moving the proton along a one-dimensional grid for the transition state geometry obtained at the B3LYP/6-31G** level of theory. The C–O distance was 2.61 Å at this geometry, and the proton coordinate is defined as $\left(R_{\mathrm{CH}} - R_{\mathrm{OH}}\right)/2$. These curves were obtained from the data computed in Ref. 21.

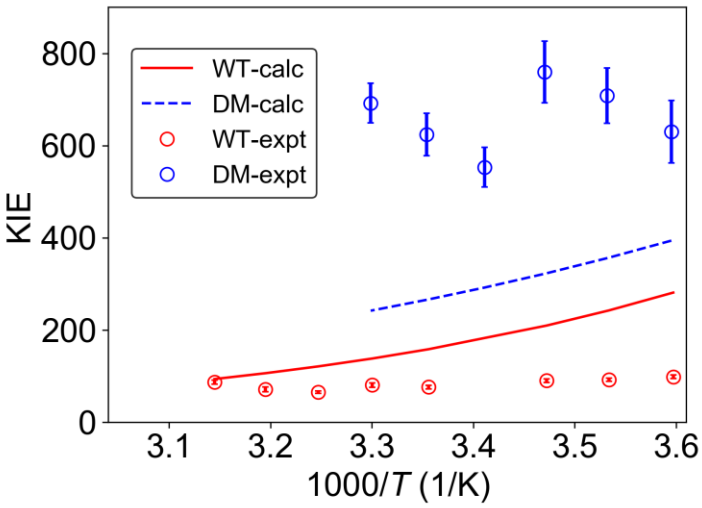


Figure S10. Experimentally measured (data points with error bars) and theoretically calculated (solid and dashed curves) KIEs for WT and DM SLO. The calculated curves were obtained using the raw PMFs obtained from the QM/MM free energy simulations without any fitting to experiment (i.e., no parameterization). The experimental data were obtained from Ref. 22 for WT and Ref. 23 for DM SLO. Note that the trend corresponding to the KIE being greater for DM than for WT SLO is reproduced, but quantitative agreement with the experimental data is lacking. When the WT and DM PMFs are scaled by a factor of 0.95 and 1.15, respectively, the agreement with experiment is significantly improved, as shown in Figure 8, illustrating the sensitivity of the KIEs to the shape of the PMFs.

Tables

Table S1. The equilibrium distances and harmonic force constants proposed for the C—O vibrational mode in WT and the DM SLO in previous work and the present work.

WT			DM		
Reference	R_{eq} (Å)	k_{eq} (kcal/mol•Å^{-2})	Reference	R_{eq} (Å)	k_{eq} (kcal/mol•Å^{-2})
Ref [24]	2.87	148[a]	Ref [25]	2.77–2.9 (3.0–3.3)	763–143[e] (153–66)[f]
Ref [25]	2.77[b] (2.88[c])	~148[b] (~115[c])	Ref 19	2.88–3.2	297–74[g]
Ref [26]	2.83–2.85	77–72[d]	Ref [26]	3.03–3.78	75–3[h]
Current work	3.25	~26	Current work	3.5	~35

[a]Based on M=14 amu, Ω=353 cm^{-1}. [b]Based on M=100 amu, Ω=132 cm^{-1}. [c]Based on M=10 amu, Ω=368.2 cm^{-1}. [d]Based on σ_{eff} in the range of 0.088–0.091 Å. [e]Based on M=100 amu, Ω=300–130 cm^{-1}. [f]Based on M=10 amu, Ω=425–280 cm^{-1}. [g]Based on M=14 amu, Ω=500–250 cm^{-1}. [h]Based on σ_{eff} in the range of 0.089–0.46 Å.

References

1. Chruszcz, M.; Wlodawer, A.; Minor, W., *Biophys. J.* **2008,** 95, 1-9.
2. Xiang, Z.; Honig, B., *J. Mol. Biol.* **2001,** 311, 421-430.
3. Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A., *Nucleic Acids Res.* **2005,** 33, W368-W371.
4. Yu, T.; Soudackov, A. V.; Hammes-Schiffer, S., *J. Phys. Chem. Lett.* **2016,** 7, 3429-3433.
5. Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J., *J. Comput. Chem.* **2009,** 30, 2785-2791.
6. Li, P.; Merz Jr, K. M., *J. Chem. Inf. Model.* **2016,** 56, 599-604.
7. Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz Jr, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J., *J. Comput. Chem.* **2005,** 26, 1668-1688.
8. Trott, O.; Olson, A. J., *J. Comput. Chem.* **2010,** 31, 455-461.
9. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A., *J. Comput. Chem.* **2004,** 25, 1157-1174.
10. Wang, J. https://mulan.swmed.edu/mmfft/antechamber_graphics.php.
11. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C., *J. Chem. Theory Comput.* **2015,** 11, 3696-3713.
12. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., *J. Chem. Phys.* **1983,** 79, 926-935.
13. Li, P.; Song, L. F.; Merz Jr, K. M., *J. Chem. Theory Comput.* **2015,** 11, 1645-1657.
14. Dennington, R.; Keith, T.; Millam, J.; Eppinnett, K.; Hovell, W. L.; Gilliland, R., GaussView. Version: 2009.
15. Souaille, M.; Roux, B. t., *Comput. Phys. Commun.* **2001,** 135, 40-57.
16. Darden, T.; York, D.; Pedersen, L., *J. Chem. Phys.* **1993,** 98, 10089-10092.
17. Götz, A. W.; Clark, M. A.; Walker, R. C., *J. Comput. Chem.* **2014,** 35, 95-108.
18. Woodcock, H. L.; Hodošček, M.; Gilbert, A. T.; Gill, P. M.; Schaefer, H. F.; Brooks, B. R., *J. Comput. Chem.* **2007,** 28, 1485-1502.
19. Hu, S.; Sharma, S. C.; Scouras, A. D.; Soudackov, A. V.; Carr, C. A. M.; Hammes-Schiffer, S.; Alber, T.; Klinman, J. P., *J. Am. Chem. Soc.* **2014,** 136, 8157-8160.
20. Grossfield, A., *WHAM version 2.0.9*.
21. Soudackov, A. V.; Hammes-Schiffer, S., *J. Phys. Chem. Lett.* **2014,** 5, 3274-3278.
22. Knapp, M. J.; Rickert, K.; Klinman, J. P., *J. Am. Chem. Soc.* **2002,** 124, 3865-3874.
23. Hu, S.; Soudackov, A. V.; Hammes-Schiffer, S.; Klinman, J. P., *ACS Catal.* **2017,** 7, 3569-3574.
24. Hatcher, E.; Soudackov, A. V.; Hammes-Schiffer, S., *J. Am. Chem. Soc.* **2007,** 129, 187-196.
25. Soudakov, A. V.; Hammes-Schiffer, S., *Faraday Discuss.* **2016,** 195, 171-189.
26. Salna, B.; Benabbas, A.; Russo, D.; Champion, P. M., *J. Phys. Chem. B* **2017,** 121, 6869-6881.