

Supporting Information

Sequential oxidation of L-lysine by a non-heme hydroxylase

Elizabeth S. Reynolds, Thomas G. Smith, Anoop R. Damodaran*, Ambika Bhagi-Damodaran*

Department of Chemistry, University of Minnesota, Twin Cities, Minneapolis, MN 55455, USA

*Corresponding Authors Email: ambikab@umn.edu, rdanoop@umn.edu

Table of Contents

Materials and methods	S3
Commercial materials.....	S3
Protein expression.....	S3
Protein purification.....	S3
Construction of sequence similarity network.....	S4
LDO activity assays and product derivatization.....	S5
Time-based LDO reaction assay.....	S5
HPLC detection of reaction products.....	S5
Nuclear magnetic resonance sample preparation and data collection.....	S6
Molecular Dynamics Simulations.....	S6
Supplementary Tables	S8
Table S1. Accurate Mass Determination of LDO enzymatic products.....	S8
Supplementary Figures	S9
Figure S1. Structural comparison of L-lysine 4-hydroxylases.....	S9
Figure S2. ^1H - ^{13}C -HSQC of starting material and succinate standards.....	S10
Figure S3. NMR characterization of LDO enzymatic products from reaction mixture.....	S11
Figure S4. Deuterium-hydrogen exchange time-course of 4-oxo-L-lysine α -hydrogens.....	S12
Figure S5. General consensus mechanism for 2OG-dependent hydroxylases.....	S13
Figure S6. Substrate binding and abstractable hydrogen distances across MD simulations.....	S14
Figure S7. Abstractable hydrogen distances for L-lysine MD simulations.....	S15
Figure S8. Abstractable hydrogen distances for 4 <i>R</i> -OH-L-lysine MD simulations.....	S16
Figure S9. Epoxide formation by non-heme iron enzymes.....	S17
References	S18

Materials and Methods

Commercial materials

Imidazole (ReagentPlus®), 2-mercaptoethanol (BME; BioUltra), Magnesium chloride hexahydrate (BioXtra), Amicon Ultra-centrifugal filter units, Sodium Tetraborate Decahydrate (ACS reagent), Ferrous ammonium sulfate hexahydrate (FAS; BioUltra), Ammonium acetate (HPLC grade – LiChropur®), 2-oxoglutarate (2OG; >99.0%), Sodium ascorbate (BioXtra), Acetonitrile (HPLC grade), and Sodium phosphate dibasic (ReagentPlus®) were purchased from Millipore Sigma. BL21-Gold (DE3) and BL21 (DE3) cells were purchased from Agilent and Thermo Scientific, respectively. 6-Aminoquinolyl-N-hydroxysuccinimidyl Carbamate (AQC; >95%) was purchased from Cayman Chemicals. Kanamycin (USP grade), Isopropyl β-D-1-thiogalactopyranoside (IPTG, >99%), RNase A, and Dithiothreitol (DTT; >99.0%) were purchased from GoldBio. Tryptone (BioReagent), Yeast extract (BioReagent), Calcium chloride (ACS reagent), DNase I solution, Sodium chloride (ACS reagent), Ethylenediaminetetraacetic acid disodium salt dihydrate (EDTA; BioReagent), L-lysine monohydrochloride (BioReagent), and Tris Hydrochloride (Molecular Biology) were purchased from Fisher Scientific. HEPES (>99.5%) was purchased from DOT Scientific. D₂O (99.9%) was purchased from Cambridge Isotope Laboratories. Reagent grade or purity is given in parentheses if applicable. HisTrapFF-5 ml columns were purchased from Cytiva. NMR tubes were purchased from VWR. All HPLC solvents were either HPLC grade or filtered through 0.22 μm filters.

Protein Expression

A previously constructed plasmid¹ for His10-tagged LDO from *Streptomyces roseifaciens* (referred to as Hydrox in previous work²) was used for subsequent studies. A plasmid containing a His-tagged human rhinovirus 3C protease (also commercially known as PreScission™ Protease when fused with glutathione S-transferase) with N-terminal domain of the spider silk protein construct (NT*-HRV3CP) was obtained from Addgene.³ Plasmids containing LDO and NT*-HRV3CP were transformed into BL21-Gold (DE3) and BL21 (DE3) competent cells, respectively. For expression of LDO, colonies of transformed cells from an overnight LB-agar plate ([kanamycin] = 50 μg/mL) were inoculated into primary 2XYT cultures (50 mL, [kanamycin] = 50 μg/mL) and incubated at 37 °C overnight at 200 RPM. Culture flasks with 2XYT media (1.5 L, [kanamycin] = 50 μg/mL) were inoculated with overnight cultures and grown at 37 °C at 200 RPM until cells reached an OD₆₀₀ of 0.6–0.8. After reaching suitable optical densities, protein expression was induced with IPTG (0.5 mM), and then cultures were incubated at 18 °C for 24 hours. Expression for NT*-HRV3CP was performed following similar methodology to LDO expression except TB media was used instead of 2XYT and secondary cultures were incubated for 20 hours after induction with IPTG. Cells were collected by centrifugation at 6000 RPM at 4 °C, flash-frozen, and stored at –20 °C for further use.

Protein purification

For purification of LDO, frozen cells were resuspended in lysis buffer (300 mM NaCl, 10 mM imidazole, 10% (vol/vol) glycerol, 5 mM CaCl₂, 5 mM MgCl₂, 0.1 mg/mL RNase, and 0.05 U/mL DNase 1, 50 mM HEPES (pH=7.5)) and then lysed by sonication. The lysed cells were centrifuged at 20,000 RPM at 4 °C, and the resulting supernatant was filtered through 0.45 μm filters. Filtered supernatant was applied to a 5mL HisTrapFF column using an AKTA Start protein purification system. After sample application, the

column was washed with 10 column volumes (CVs) of wash buffer (300 mM NaCl, 10 mM imidazole, 10% (vol/vol) glycerol, 20 mM 2-mercaptoethanol, 50 mM HEPES (pH= 7.5)), and then the His-tagged LDO was eluted using a two-step isocratic method first with 28% elution buffer (300mM NaCl, 300mM imidazole, 10% glycerol (vol/vol), 20 mM 2-mercaptoethanol, 50mM HEPES (pH 7.5)) for 8 CVs followed by 8 CVs of 100% elution buffer. Fractions from the second elution phase were collected and dialyzed in 100 mM NaCl, 1 mM DTT, 1 mM EDTA, 50 mM HEPES (pH =7.5) for 2 hours, followed by another 2-hour dialysis in 1 mM DTT, 100 mM NaCl, 50 mM HEPES (pH =7.5). After this second round of dialysis, the protein solution was diluted to a concentration of 0.02–0.05 mM and then incubated with NT*-HRV3CP protease (1 mg protease to 100 mg protein), dialyzing against 1 mM DTT, 100 mM NaCl, 50 mM HEPES (pH = 7.5) overnight. The dialyzed protein solution was then concentrated to approximately 10 mL using an Amicon Ultra-15 centrifugal filter (10 kDa MWCO). Imidazole was added to the protein solution at a final concentration of 10 mM, and then protein solution was applied to a 5 mL HisTrapFF column using an AKTA Start protein purification system. His-tagged cleaved LDO was recovered in the flow through during sample application and 4 CVs application of wash buffer. The collected solution was buffer exchanged into 50 mM HEPES (pH = 7.5) through 5 washes in an Amicon Ultra-15 centrifugal filter (10 kDa MWCO). After the final wash, the protein solution was concentrated, and glycerol was added to the protein solution for a final concentration of 10% (vol/vol). After concentration of the protein was assessed at 280 nm using extinction coefficient calculated by ExPASy⁴, the protein was aliquoted, flash-frozen, and stored at –80 °C for further use.

For NT*-HRV3CP purification, frozen cells were resuspended in lysis buffer (300 mM NaCl, 10 mM imidazole, 5% (vol/vol) glycerol, 5 mM CaCl₂, 5 mM MgCl₂, 0.1 mg/mL RNase, and 0.05 U/mL DNase1, 50 mM Tris-HCl (pH 7.5)). Cells were lysed, centrifuged, filtered, and applied to 5 mL HisTrapFF column as described for LDO. After sample application, the column was washed with 10 column volumes (CVs) of wash buffer (300 mM NaCl, 10 mM imidazole, 5% (vol/vol) glycerol, 50 mM Tris-HCl (pH 7.5)), and then the His-tagged NT*-HRV3CP was eluted using a two-step isocratic method with first 10% elution buffer (300 mM NaCl, 500 mM imidazole, 5% (vol/vol) glycerol, 50 mM Tris-HCl (pH 7.5)) for 3 CVs followed by 8 CVs of 100% elution buffer. Fractions from the second elution phase were collected and dialyzed in 300 mM NaCl, 1 mM DTT, 50 mM Tris-HCl (pH 7.5) for a total of three rounds of two-hour dialysis. The solution from dialysis was concentrated using an Amicon Ultra-15 centrifugal filter (10 kDa MWCO). After concentration, glycerol was added to the protein solution for a final concentration of 10% (vol/vol) and concentration was assessed 280 nm using extinction coefficient calculated by ExPASy.⁴ The protein was then aliquoted, flash-frozen, and stored at –80 °C for further use.

Construction of sequence similarity network

Sequences for LDO² (WP_107105619), Glib⁵ (AKJ29081.1), KDO2⁶ (ACU60313.1), KDO3⁶ (ABQ06186.1), KDO4⁶ (AEV99100.1), KDO5⁶ (J3BZS6), and K4H-4 (EFK34737.1) were used as inputs for NCBI BLAST search against the nonredundant protein sequence database with an expect threshold of 0.05 and max target sequences set to 5000. Resulting similar sequences for all L-lysine 4-hydroxylases were combined into a single set, and sequences with 65% similarity were removed by CD-HIT.⁷ Using this condensed sequence list as input, a sequence similarity network (SSN) was constructed using the EFI-EST web tool with an alignment score cutoff of 40, which corresponded to a sequence identity of ~35%, a common starting threshold for separating isofunctional homologous enzymes.^{8,9} The SSN was visualized with Cytoscape v.3.10.2.¹⁰

LDO activity assays and product derivatization

Before activity assays, LDO was buffer exchanged into 50 mM HEPES (pH 7.5) using 0.5 mL Amicon™ Ultra 10 kDa MWCO centrifugal filter units at 13500 RPM at 4° C for 15-minute wash cycles to remove storage glycerol. For varied substrate-to-enzyme ratio studies, reaction samples consisted of 5 mM L-lysine HCl, 600 μ M FAS, 20 mM 2OG, 5 mM sodium ascorbate with either 5, 50, or 500 μ M LDO. All reaction samples (50 μ l) were incubated in an Eppendorf Thermomixer® at 25° C for 5 hours while mixing (450 RPM).

After incubation, reactions were diluted 10-fold in 50 mM HEPES (pH 7.5) and then were filtered through 0.5 mL Amicon™ Ultra 10 kDa MWCO centrifugal filter units at 13500 RPM for 20 minutes at 4° C. Reaction sample flowthrough (80 μ L) was diluted 2-fold in 15 mM sodium tetraborate (pH 10.5). Two sequential additions (20 μ L) of AQC in acetonitrile (3 mg/ml) were added to each reaction (final volume: 200 μ L), and samples were vortexed to combine after each addition. Derivatized samples were then analyzed immediately or stored at -80 or -20 ° C for future analysis.

Time-based LDO reaction assay

LDO was buffer exchanged in 50 mM HEPES (pH 7.5) as described above. A reaction (350 μ l) was prepared with the following conditions: 2.5 mM L-lysine HCl, 300 μ M FAS, 20 mM 2OG, 5 mM sodium ascorbate, and 50 μ M LDO in 50 mM HEPES (pH 7.5). Aliquots (20 μ L) of reaction sample were quenched by 10x dilution in 200 μ M EDTA in 50 mM HEPES (pH 7.5) solution at indicated time points. Reaction aliquots (200 μ L) were filtered to remove LDO and derivatized with AQC as described in previous section.

HPLC detection of reaction products

Derivatized amino acid samples were analyzed using a Shimadzu Prominence-i LC-2030C 3D Plus system equipped with a Poroshell 120 CS-C18 column (3.0 mm x 150 mm x 2.7 μ m) and a PDA detector. For all samples, the column temperature was 35°C, the flow rate was set to 0.6 mL/min, and sample injection volume was 5 μ L. Solvent A was 5 mM ammonium acetate (pH = 5.0) and Solvent B was 60% acetonitrile in water. Gradient conditions were: 20% B, 0.0 to 1.5 min; 20% B to 75% B, 1.5 to 5.5 min; 75% B to 100% B, 5.5 to 6.0 min; 100% B, 6.0 to 7.0 min; 100% B to 20% B, 7.0 to 7.5 min; 20% B, 7.5 to 10.0 min. All HPLC traces were constructed from absorbance at 254 nm. Peak area integrations were performed in Shimadzu Lab Solution software. Percent analyte conversion was determined by the peak area of the analyte divided by the sum of all analytes peak areas. To determine the identity of reaction products, samples of HPLC chromatographic peaks were collected and diluted in methanol. Mass spectra of analytes were obtained by direct infusion using a syringe pump onto a Sciex X500R QTOF-MS system with electrospray ionization in positive ionization mode. Acknowledgement: Mass spectrometry analysis was performed at The University of Minnesota Department of Chemistry Mass Spectrometry Laboratory (MSL), supported by the Office of the Vice President of Research, CSE, and the Department of Chemistry at the University of Minnesota. The content of this work is the sole responsibility of the authors.

Nuclear magnetic resonance sample preparation and data collection

All reactions were incubated in an Eppendorf Thermomixer® at 25° C for the indicated amount of time while mixing at 450 RPM. For structural characterization of 4-OH-L-lysine and 4-oxo-L-lysine studies, LDO was buffer exchanged into 50 mM sodium phosphate in D₂O (pD 7.5) using 0.5 mL Amicon™ Ultra (10 kDa MWCO) centrifugal filter units at 13500 RPM at 4° C for five 15-minute wash cycles to remove storage glycerol. The reaction was prepared using stock solutions of compounds in 50 mM sodium phosphate in D₂O (pD 7.5) so that final reaction conditions were 5 mM L-lysine HCl, 300 μM FAS, 10 mM 2OG, 10 mM sodium ascorbate with 200 μM LDO in 50 mM sodium phosphate in D₂O (pD 7.5). After 1.5-hour incubation, the reaction sample was filtered through 0.5 mL Amicon™ Ultra (10 kDa MWCO) centrifugal filter units at 13500 RPM for 20 minutes at 4° C to remove LDO. Flowthrough was transferred to 5 mm O.D. NMR tubes. Starting material and succinate standards were made from diluted stock solutions used to make the reaction sample. For hydrogen-deuterium exchange studies, all stock solutions were prepared in 50 mM sodium phosphate in H₂O (pH 7.5), and a reaction sample was prepared with final concentrations of 5 mM L-lysine HCl, 300 μM FAS, 15 mM 2-oxoglutarate, 10 mM sodium ascorbate with 200 μM LDO in 50 mM Sodium phosphate in D₂O (pD 7.5). After a 1.5-hour incubation at 25 °C, the sample was filtered to remove LDO as described previously, and then the sample was lyophilized to remove H₂O. Sample was resuspended in D₂O immediately prior to initial spectra collection to observe hydrogen-deuterium exchange at α-hydrogens of putative 4-oxo-L-lysine. ¹H, ¹H-¹³C HSQC, and ¹H-¹³C HMBC NMR spectra were acquired at 25 °C on a Bruker 500 MHz Avance III HD spectrometer equipped with a 5 mm Prodigy TCI-H&F Cryoprobe.

Molecular Dynamics Simulations

The starting structure of LDO was taken from the protein data bank (PDB: 7JSD; Chain C).² Missing residues were added using the MODELLER module¹¹ in UCSF ChimeraX¹² and hydrogen atoms were added to the protein using a pH of 7.5 with the H++ Web server.^{13–15} The crystal structure active site was modified using PyMOL¹⁶ and Avogadro¹⁷ to generate the four ferryl intermediate models. LDO structures with 4*R*-OH-L-lysine were made by modifying the bound crystallized L-lysine structure *in-silico* using Avogadro¹⁷. Force field parameters describing covalent bonds to ferryl iron for each of these models were generated with the MCPB.py module of AmberTools23 utilizing Gaussian16 at UB3LYP/6-31g(d) level of theory using a procedure adapted from Li et. al.^{18,19} Geometry optimizations and frequency calculations were performed for the small cluster model which contained iron in addition to the directly coordinating protein residues truncated at the α-carbon and coordinating small molecules. The Seminario method²⁰ was used to generate missing force field parameters. The large model incorporated all atoms in the small model in addition to an amide backbone region that extended to one full amide bond in either direction of the coordinating residue. Only hydrogen positions were optimized in the large cluster model before performing a Merz-Kollman population analysis²¹ and the output of these calculations were used for RESP charge fitting according to default settings in MCPB.py module. Protein residues in LDO were parameterized with the AMBER ff19SB force field.^{22,23} Succinate and substrates were treated with the generalized amber force field (GAFF) with missing parameters developed using the Antechamber package using the AM1-BCC charge model.^{24,25}

The protein was solvated in an OPC water octahedron cell with 12.0 Å between the protein and the closest edge of the unit cell. Counterions Na^+ were added to the cell at a concentration of 0.15 M, followed by the addition of Cl^- to neutralize the system.²⁶ The GPU-accelerated *pmemd* code of Amber20 was used to perform all steps of MD.²² For minimization and heating portions of the simulations, the non-bonded cutoff was 8 Å and was increased to 10 Å for density equilibration and unrestrained MD. Protein was minimized using a seven-stage minimization process where each minimization stage utilized 1000 steps of steepest descent minimization followed by 5000 steps of conjugate gradient minimization. Additionally, at each stage solute heavy atoms were subject to restraints starting at 10.0 kcal/mol/Å² which were decreased geometrically until reaching 0.0 kcal/mol/Å². For all subsequent processes, the SHAKE algorithm was employed for bonds involving hydrogen, a timestep of 2 fs utilized throughout simulations, and temperature was controlled by the Langevin thermostat. The system was gently heated linearly from 10 K to 300 K over 2 ns at a constant volume and all solute atoms were subjected to a restraint of 10.0 kcal/mol/Å² with a collision frequency of 5.0 ps⁻¹ and a random seed. The system was then density equilibrated for 3.5 ns at a constant pressure of 1.0 bar and a constant temperature of 300 K with a collision frequency of 1.0 ps⁻¹ using a seven-stage process where solute heavy atoms were subject to restraints starting at 10.0 kcal/mol/Å² which were decreased geometrically every 0.5 ns until reaching 0.0 kcal/mol/Å² for the final 0.5 ns. After equilibration, three unrestrained independent 250 ns simulations were performed at a constant pressure of 1.0 bar and a constant temperature of 300.0 K with a collision frequency of 1.0 ps⁻¹ for each ferryl intermediate and substrate combination. The CPPTRAJ²⁷ module from AmberTools23 was used to perform all trajectory analyses with H-bond donor/acceptor distance cutoff set to 3.2 Å.

Supplementary Tables

Table S1. Accurate Mass Determination of LDO enzymatic products

Analyte	HPLC retention time (min)	Chemical formula	Theoretical $[M+H]^+$	Observed $[M+H]^+$	Mass error (ppm)
AQC(2x)-OH-L-lysine	4.34	$C_{26}H_{26}N_6O_5$	503.2037	503.2069	6.36
AQC(2x)-oxo-L-lysine	4.17	$C_{26}H_{24}N_6O_5$	501.1881	501.1886	1.00

Supplementary Figures

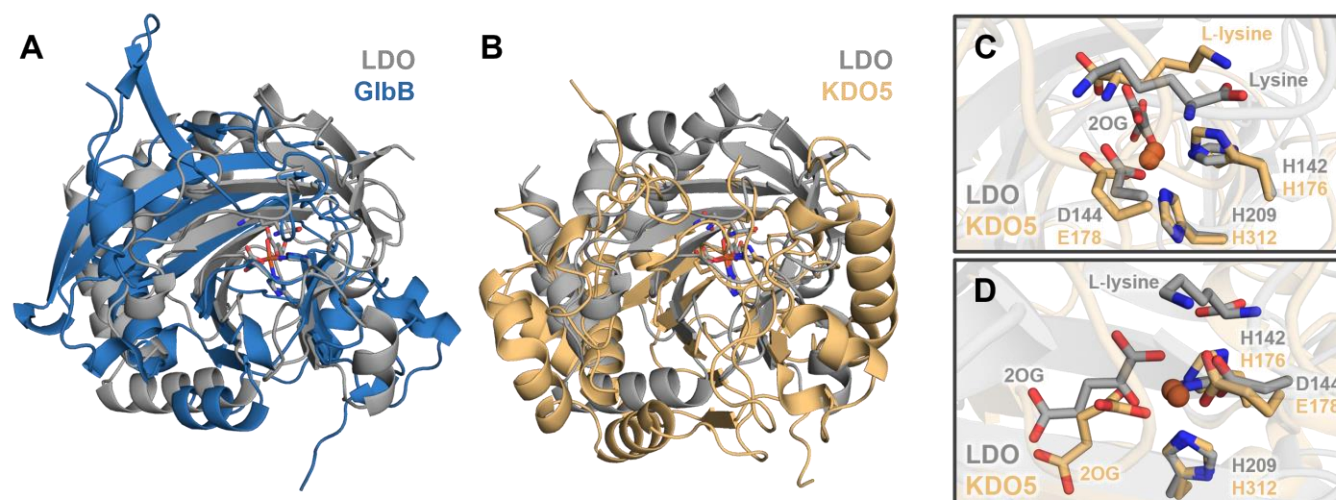


Figure S1. Structural comparison of L-lysine 4-hydroxylases. **(A)** AlphaFold²⁸ structure of GlbB overlaid with crystal structure of LDO (7JSD - chain a) aligned by residues of the facial triad (GlbB: H188, D190, and H251; LDO: H142, D144, and H209) due to lack of structural similarity. **(B)** Crystal structure of KDO5 (6EXF - chain B) overlaid with crystal structure of LDO (7JSD - chain A) aligned by histidines of the facial triad and crystallized Fe (KDO5: H176 and H312; LDO: H142 and H209) due to lack of structural similarity. **(C)** Overlaid active site structures of KDO5 (6EXF - chain B) and LDO (7JSD - chain A) showcasing difference in L-lysine binding orientation **(D)** Overlaid active site structures of KDO5 (6EUR - chain B) and LDO (7JSD - chain A) showcasing differences in 2OG orientation.

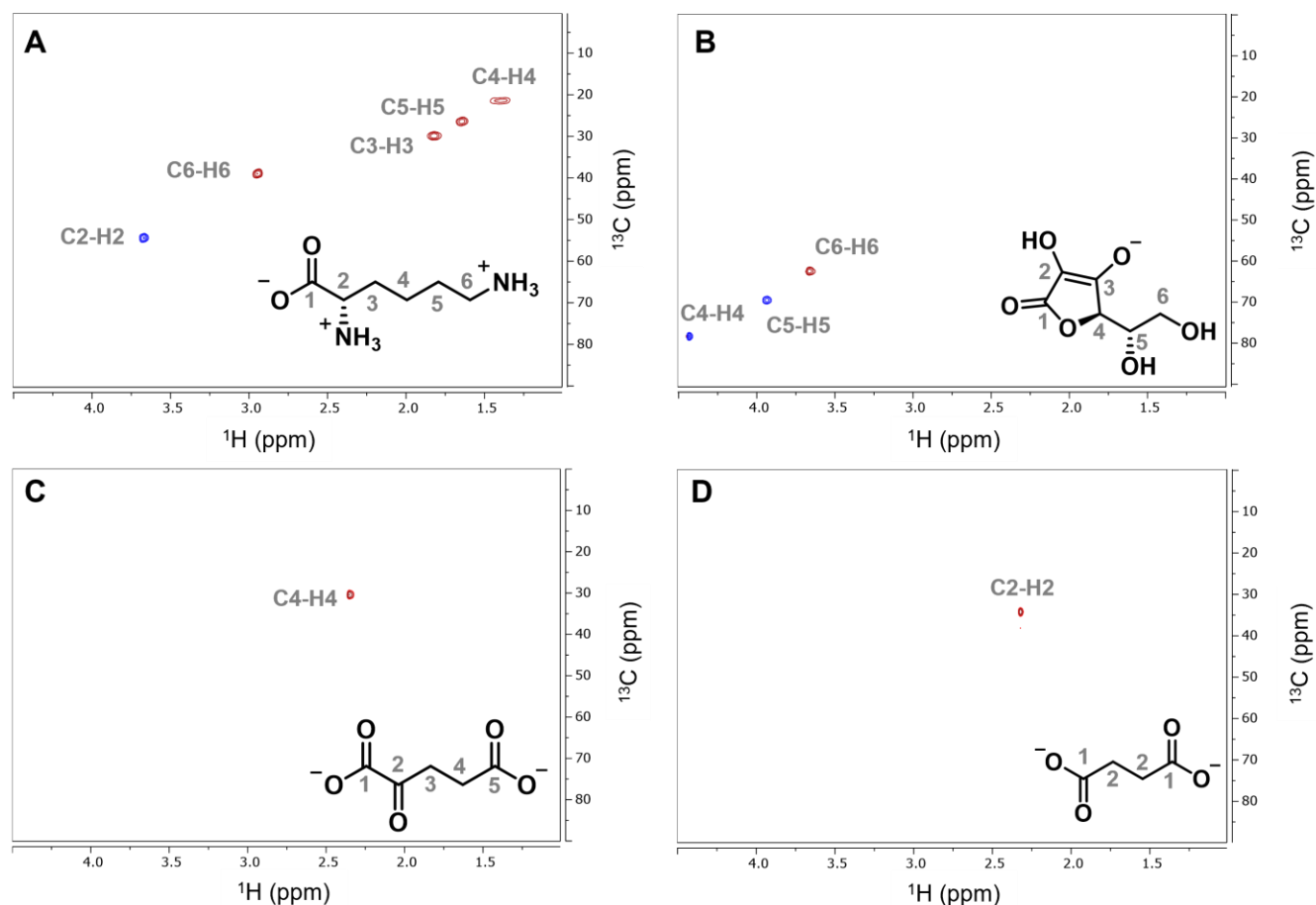


Figure S2. ^1H - ^{13}C HSQC of starting material and succinate standards. **(A)** 5 mM L-lysine HCl, **(B)** 10 mM Sodium Ascorbate, **(C)** 10 mM 2-oxoglutaric acid (C3-H3 signal is eliminated due to hydrogen-deuterium exchange of α -hydrogens of ketone), and **(D)** 10 mM Sodium succinate in 50 mM Sodium Phosphate (pD = 7.5) (500.15 MHz, D₂O). For HSQC experiments, -CH and -CH₃ signals are in-phase (blue) while -CH₂ signals are anti-phase (red).

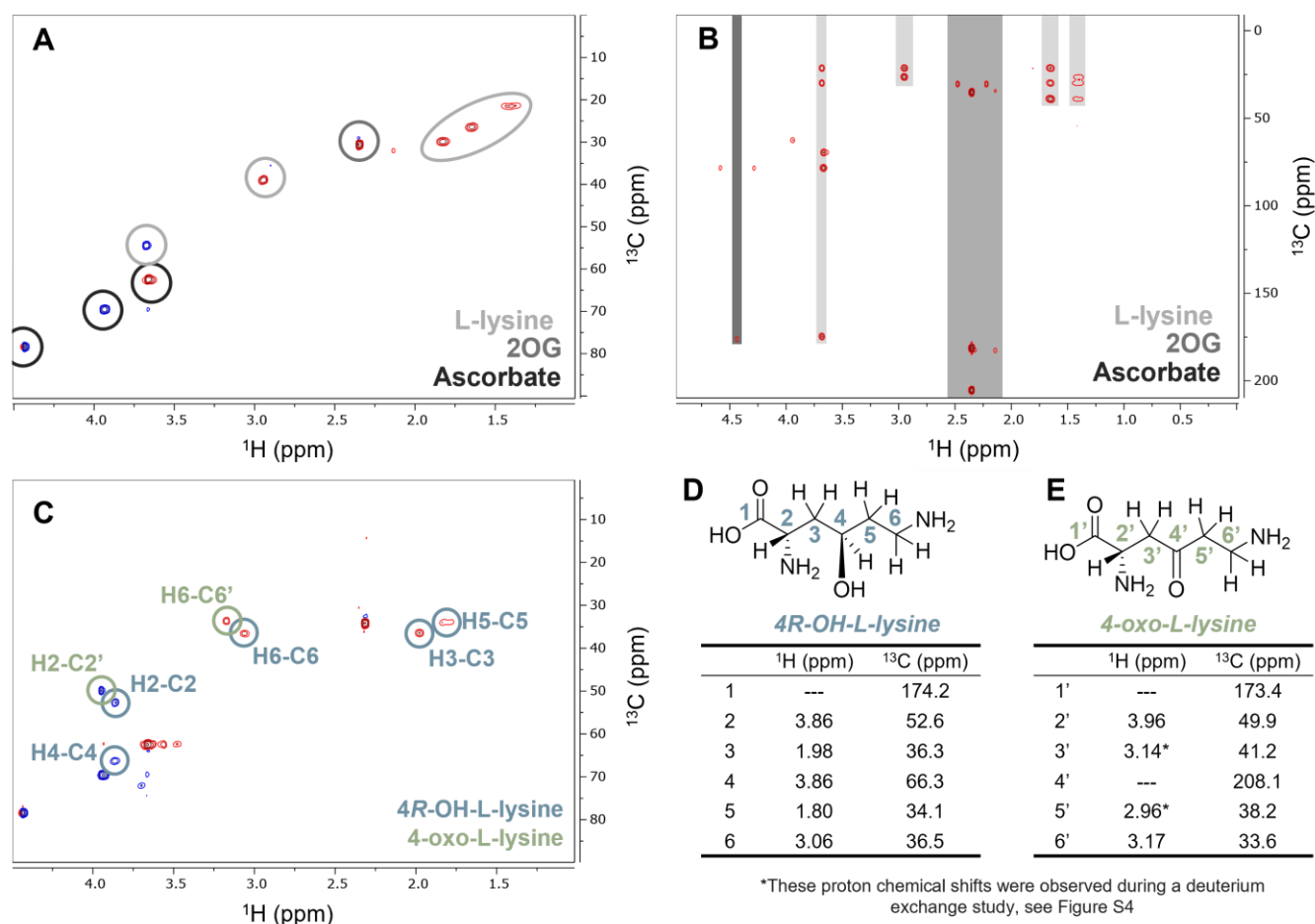


Figure S3. NMR characterization of LDO enzymatic products from reaction mixture. **(A)** ^1H - ^{13}C HSQC of reaction starting materials without iron (5 mM L-lysine-HCl, 10 mM 2-oxoglutarate, 10 mM Sodium Ascorbate in 50 mM sodium phosphate buffer (pD = 7.5)) (500.15 MHz, D₂O) **(B)** ^1H - ^{13}C HMBC of reaction starting materials without iron (5 mM L-lysine HCl, 10 mM 2-oxoglutarate, and 10 mM Sodium Ascorbate in 50 mM sodium phosphate buffer (pD = 7.5)) (500.15 MHz, D₂O). **(C)** ^1H - ^{13}C HSQC of enzymatic reaction after 1.5 hours with LDO removed (500.15 MHz, D₂O). Reaction conditions: 200 μM LDO, 300 μM FAS, 5 mM L-lysine-HCl, 10 mM 2-oxoglutarate, and 10 mM Sodium Ascorbate in 50 mM Sodium Phosphate buffer (pD = 7.5) in D₂O. Chemical shifts determined from the NMR data for **(D)** 4R-OH-L-lysine and **(E)** 4-oxo-L-lysine. 4R-OH-L-lysine chemical shifts match those determined previously in literature.² For HSQC experiments, -CH and -CH₃ signals are in-phase (blue) while -CH₂ signals are anti-phase (red).

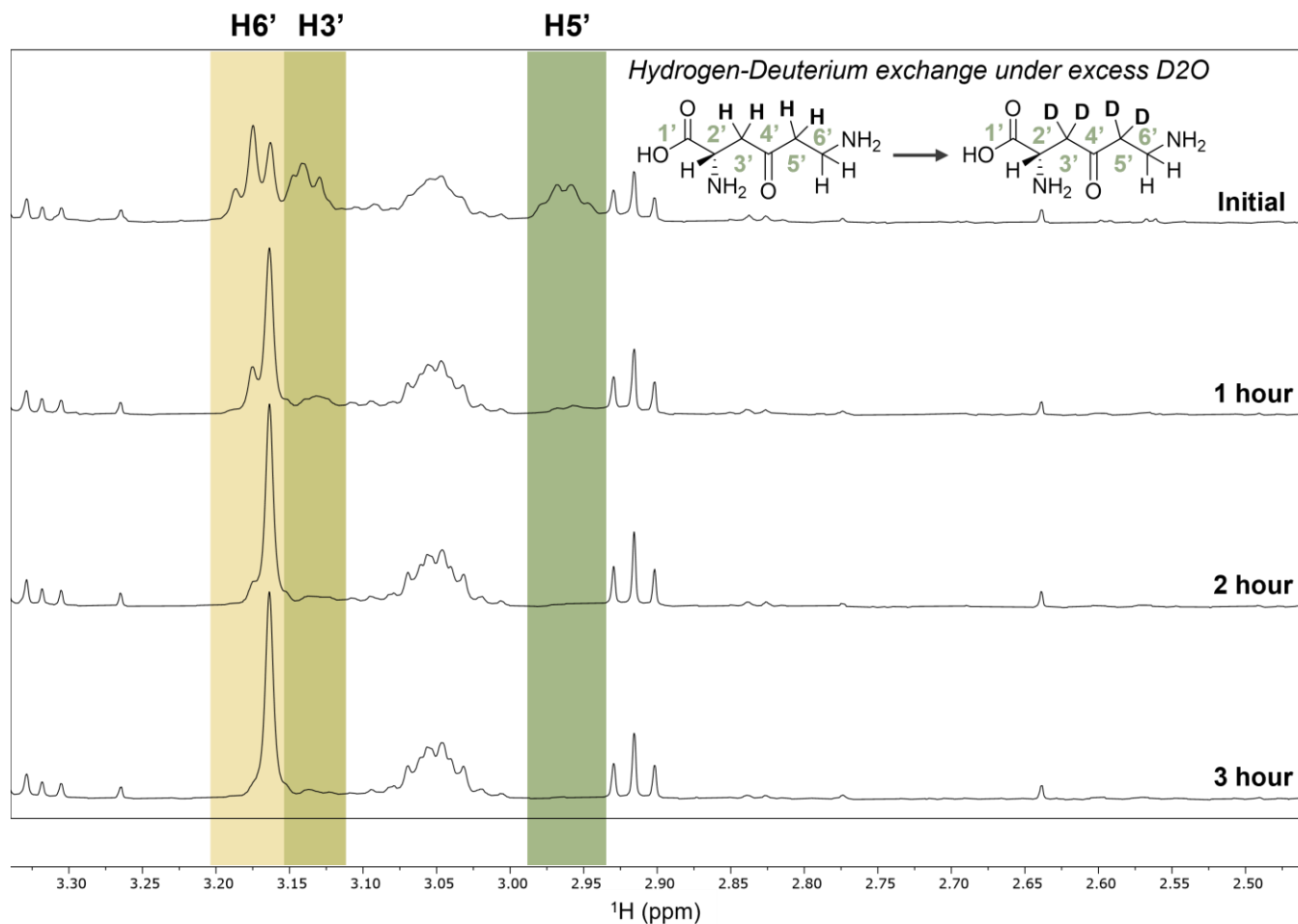


Figure S4. Deuterium-hydrogen exchange time-course of 4-oxo-L-lysine α -hydrogens. ¹H-NMR of enzymatic reaction after 1.5 hours with LDO removed (500.15 MHz, D₂O). Reaction conditions: 200 μ M LDO, 300 μ M ferrous ammonium sulphate, 5 mM L-lysine-HCl, 15 mM 2-oxoglutarate, and 10 mM Sodium Ascorbate in 50 mM sodium phosphate buffer (pH = 7.5) in H₂O. After reaction, LDO was removed, and sample was lyophilized. Sample was reconstituted in D₂O immediately prior to initial spectra collection and measurements were taken every hour after that point.

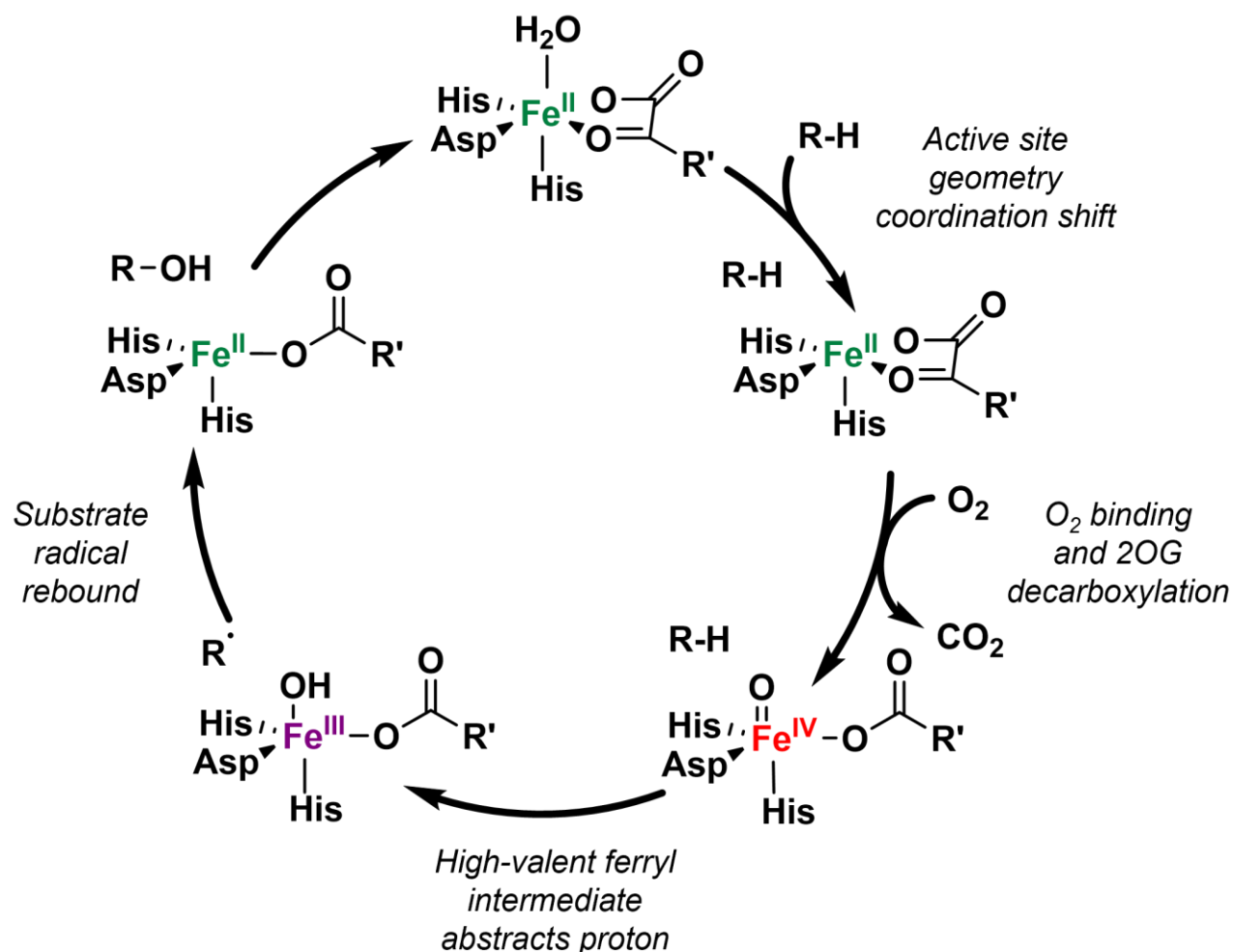


Figure S5. General consensus mechanism for 2OG-dependent hydroxylases. In the resting state of the enzyme, the iron is coordinated by an aqua ligand, 2OG, and protein derived histidine and aspartate ligands. Upon substrate binding, the aqua ligand is displaced, and the active site coordination shifts from a 6-coordinate environment to a 5-coordinate environment to allow space for efficient oxygen binding. Upon O_2 binding, co-substrate 2OG is decarboxylated, enabling the formation of the high valent ferryl intermediate. This highly reactive intermediate then abstracts a hydrogen atom from the substrate, forming a substrate radical which then rebounds with the hydroxyl ligand to form the hydroxylated product.

Ferryl intermediate

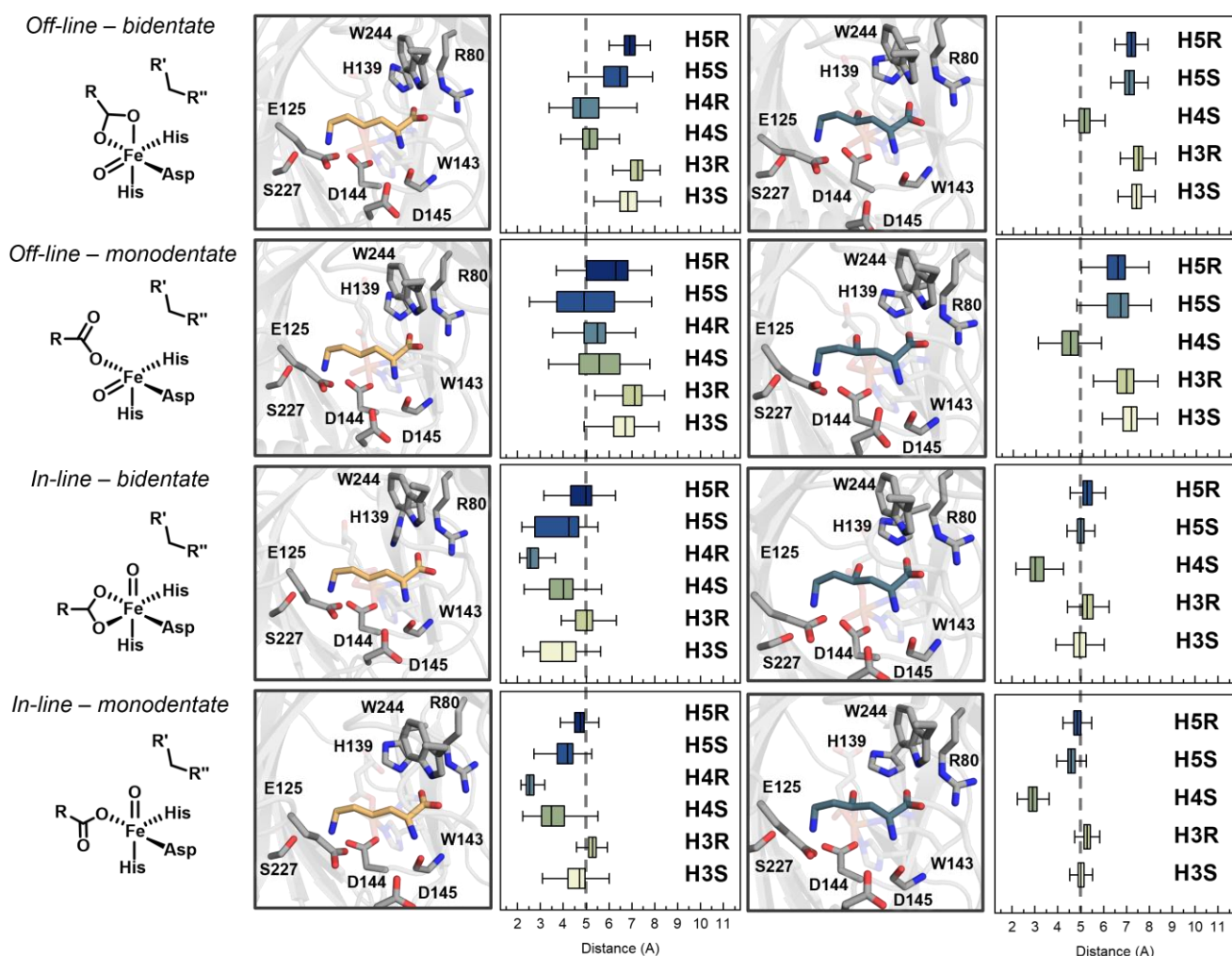


Figure S6. Abstractable hydrogen distances for MD simulations. Distance from the protons attached to C3-5 carbons of L-lysine (yellow) or 4R-OH-L-lysine (gray blue) to the oxo-group of the ferryl intermediate for the four different models. Structures shown are representative frames from the most populated cluster generated by the k-means clustering algorithm based on substrate atom positions and show similar lysine binding positions between L-lysine and 4R-OH-L-lysine simulations. Dashed line is shown at 5 Å which is approximately the threshold for facile H-atom abstraction.

Ferryl intermediate

Lysine-bound triplicate simulations over time

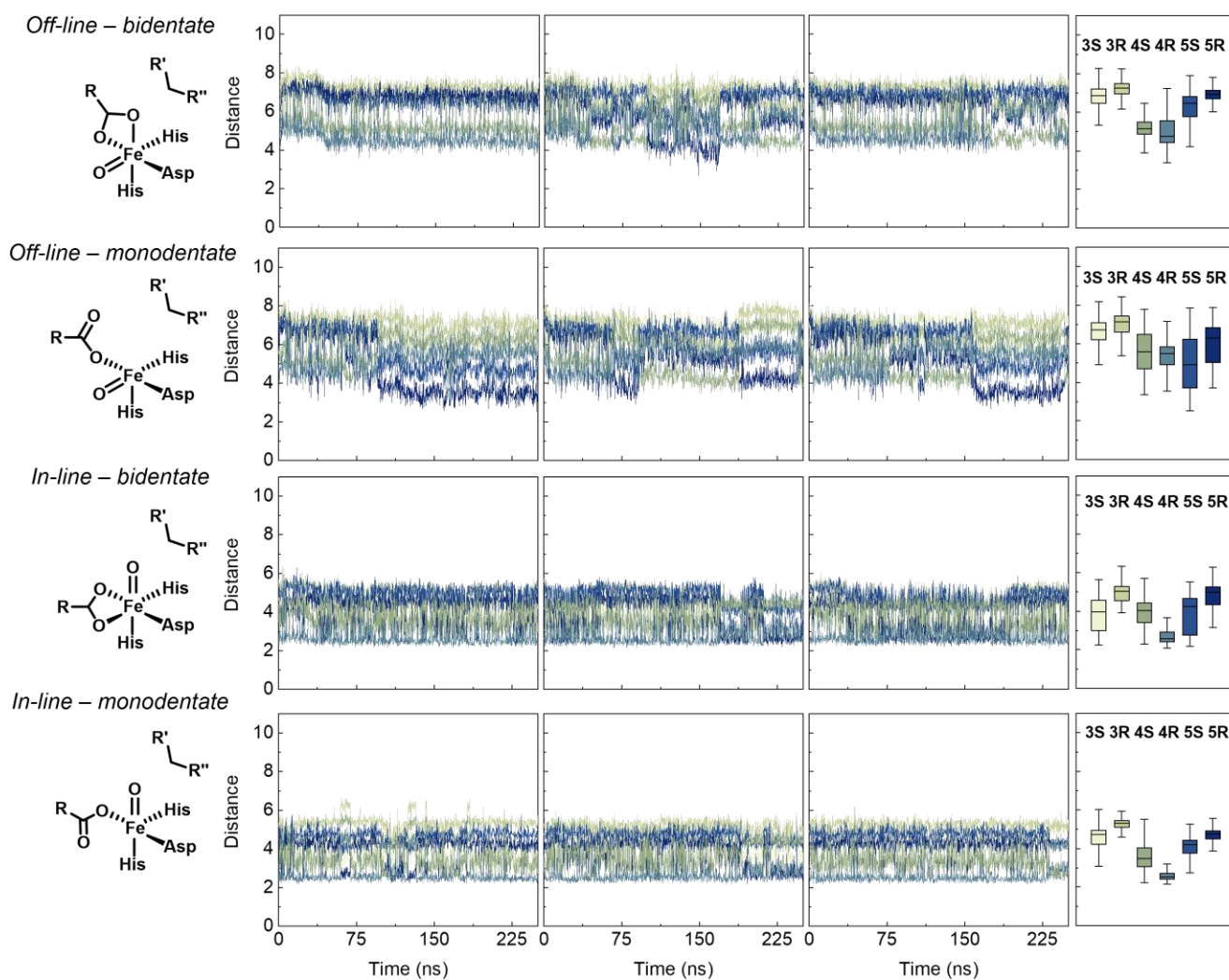


Figure S7. Abstractable hydrogen distances for MD simulations over time for L-lysine containing simulations. Boxplots to the far right represent compiled distances for the four different models across replicate simulations which are reproduced from **Figure S6** for clarity.

Ferryl intermediate

4R-OH-Lysine-bound triplicate simulations over time

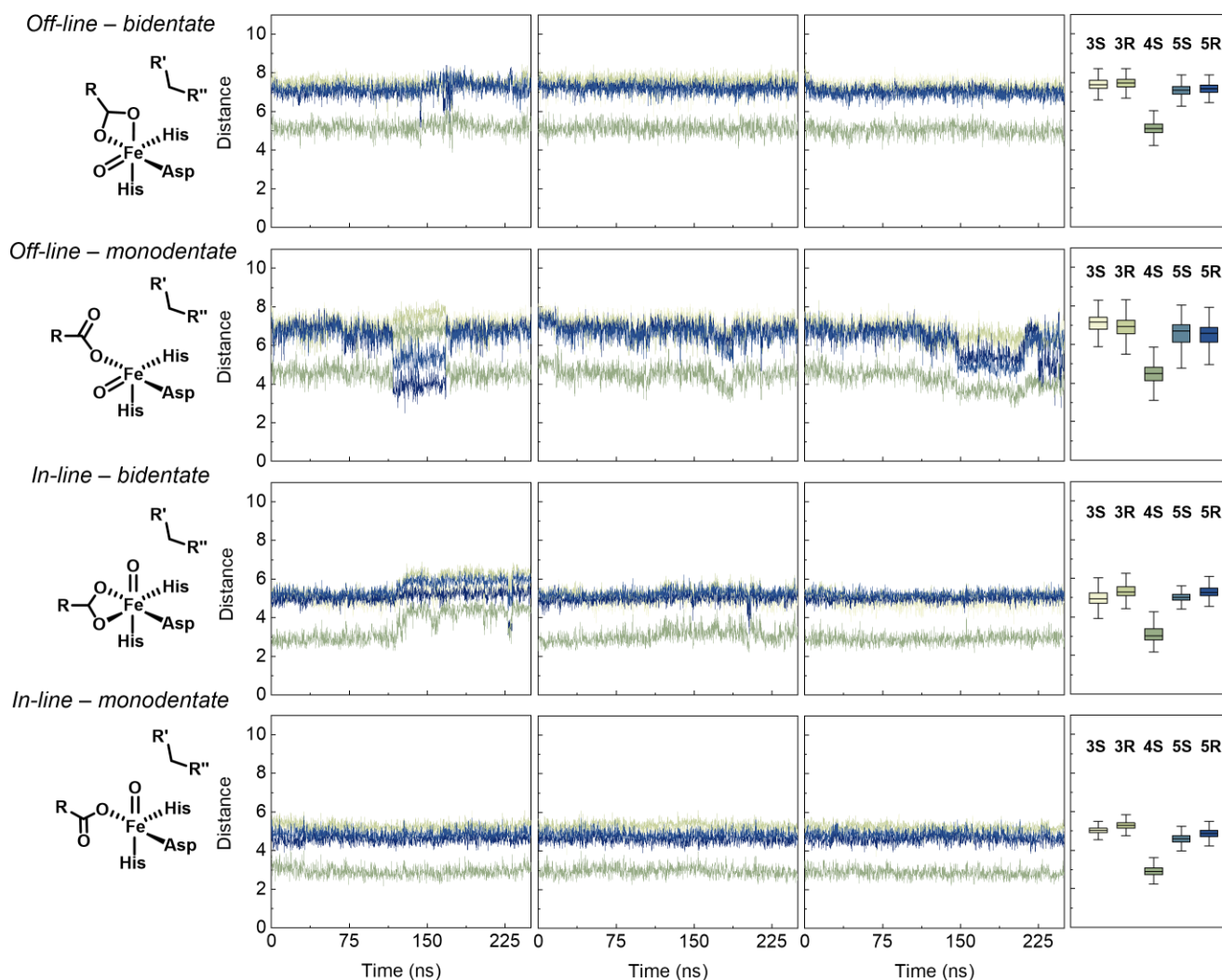


Figure S8. Abstractable hydrogen distances for MD simulations over time for 4R-OH-L-lysine containing simulations. Boxplots to the far right represent compiled distances for the four different models across replicate simulations which are reproduced from **Figure S6** for clarity.

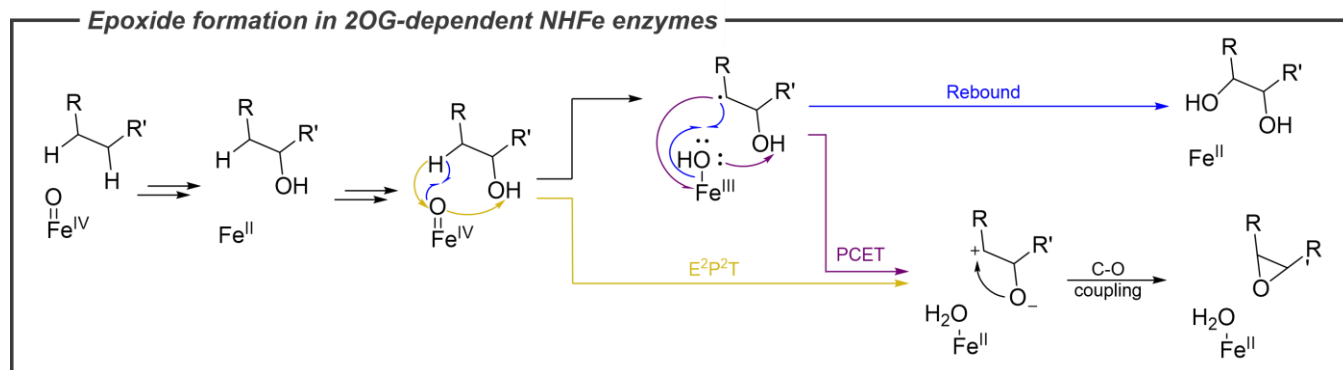


Figure S9. Epoxide formation by non-heme iron enzymes through a mono-hydroxylated intermediary product as proposed for Hyoscyamine 6 β -hydroxylase (H6H).²⁹ Other routes to epoxide formation for 2OG-dependent non-heme iron enzymes involve an initial alkene product, as in the case of AsqJ.³⁰

References

- (1) Wilson, H. R.; Diaz, D. J.; Damodaran, A. R.; Bhagi-Damodaran, A. Machine Learning Guided Rational Design of a Non-Heme Iron-Based Lysine Dioxygenase Improves Its Total Turnover Number. *ChemBioChem* **2024**, 25 (24), e202400495. <https://doi.org/10.1002/cbic.202400495>.
- (2) Neugebauer, M. E.; Kissman, E. N.; Marchand, J. A.; Pelton, J. G.; Sambold, N. A.; Millar, D. C.; Chang, M. C. Y. Reaction Pathway Engineering Converts a Radical Hydroxylase into a Halogenase. *Nature Chemical Biology* **2020**, 18, 171–179. <https://doi.org/10.1038/s41589-021-00944-x>.
- (3) Abdelkader, E. H.; Otting, G. NT*-HRV3CP: An Optimized Construct of Human Rhinovirus 14 3C Protease for High-Yield Expression and Fast Affinity-Tag Cleavage. *Journal of Biotechnology* **2021**, 325, 145–151. <https://doi.org/10.1016/j.jbiotec.2020.11.005>.
- (4) Gasteiger, E. ExPASy: The Proteomics Server for in-Depth Protein Knowledge and Analysis. *Nucleic Acids Research* **2003**, 31 (13), 3784–3788. <https://doi.org/10.1093/nar/gkg563>.
- (5) Amatuni, A.; Renata, H. Identification of a Lysine 4-Hydroxylase from the Glidobactin Biosynthesis and Evaluation of Its Biocatalytic Potential. *Org. Biomol. Chem.* **2019**, 17 (7), 1736–1739. <https://doi.org/10.1039/C8OB02054J>.
- (6) Baud, D.; Saaïdi, P.; Monfleur, A.; Harari, M.; Cuccaro, J.; Fossey, A.; Besnard, M.; Debard, A.; Mariage, A.; Pellouin, V.; Petit, J.; Salanoubat, M.; Weissenbach, J.; de Berardinis, V.; Zapparucha, A. Synthesis of Mono- and Dihydroxylated Amino Acids with New α -Ketoglutarate-Dependent Dioxygenases: Biocatalytic Oxidation of C-H Bonds. *ChemCatChem* **2014**, 6 (10), 3012–3017. <https://doi.org/10.1002/cctc.201402498>.
- (7) Huang, Y.; Niu, B.; Gao, Y.; Fu, L.; Li, W. CD-HIT Suite: A Web Server for Clustering and Comparing Biological Sequences. *Bioinformatics* **2010**, 26 (5), 680–682. <https://doi.org/10.1093/bioinformatics/btq003>.
- (8) Zallot, R.; Oberg, N.; Gerlt, J. A. The EFI Web Resource for Genomic Enzymology Tools: Leveraging Protein, Genome, and Metagenome Databases to Discover Novel Enzymes and Metabolic Pathways. *Biochemistry* **2019**, 58 (41), 4169–4182. <https://doi.org/10.1021/acs.biochem.9b00735>.
- (9) Oberg, N.; Zallot, R.; Gerlt, J. A. EFI-EST, EFI-GNT, and EFI-CGFP: Enzyme Function Initiative (EFI) Web Resource for Genomic Enzymology Tools. *Journal of Molecular Biology* **2023**, 435 (14), 168018. <https://doi.org/10.1016/j.jmb.2023.168018>.
- (10) Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **2003**, 13 (11), 2498–2504. <https://doi.org/10.1101/gr.1239303>.
- (11) Webb, B.; Sali, A. Comparative Protein Structure Modeling Using MODELLER. *CP in Bioinformatics* **2016**, 54 (1). <https://doi.org/10.1002/cpbi.3>.
- (12) Meng, E. C.; Goddard, T. D.; Pettersen, E. F.; Couch, G. S.; Pearson, Z. J.; Morris, J. H.; Ferrin, T. E. UCSF CHIMERAX : Tools for Structure Building and Analysis. *Protein Science* **2023**, 32 (11), e4792. <https://doi.org/10.1002/pro.4792>.
- (13) Anandkrishnan, R.; Aguilar, B.; Onufriev, A. V. H++ 3.0 : Automating pK Prediction and the Preparation of Biomolecular Structures for Atomistic Molecular Modeling and Simulations. *Nucleic Acids Research* **2012**, 40, 537–541. <https://doi.org/10.1093/nar/gks375>.
- (14) Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A. H ++ : A Server for Estimating p K a s and Adding Missing Hydrogens to Macromolecules. *Nucleic Acids Research* **2005**, 33, 368–371. <https://doi.org/10.1093/nar/gki464>.
- (15) Myers, J.; Grothaus, G.; Narayanan, S.; Onufriev, A. A Simple Clustering Algorithm Can Be Accurate Enough for Use in Calculations of pKs in Macromolecules. *PROTEINS: Structure, Function, and Bioinformatics* **2006**, 63, 928–938. <https://doi.org/10.1002/prot>.
- (16) The PyMOL Molecular Graphics System, Version 2.5.0 Schrödinger, LLC.

- (17) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. *J Cheminform* **2012**, 4 (1), 17. <https://doi.org/10.1186/1758-2946-4-17>.
- (18) Li, P.; Merz, K. M. MCPB.Py: A Python Based Metal Center Parameter Builder. *Journal of Chemical Information and Modeling* **2016**, 56 (4), 599–604. <https://doi.org/10.1021/acs.jcim.5b00674>.
- (19) Li, P.; Merz, K. M. Parameterization of a Dioxygen Binding Metal Site Using the MCPB.Py Program. In *Structural Genomics*; Chen, Y. W., Yiu, C.-P. B., Eds.; Methods in Molecular Biology; Springer US: New York, NY, 2021; Vol. 2199, pp 257–275. https://doi.org/10.1007/978-1-0716-0892-0_15.
- (20) Seminario, J. M. Calculation of Intramolecular Force Fields from Second-Derivative Tensors. *Int. J. Quantum Chem.* **1996**, 60 (7), 1271–1277. [https://doi.org/10.1002/\(SICI\)1097-461X\(1996\)60:7<1271::AID-QUA8>3.0.CO;2-W](https://doi.org/10.1002/(SICI)1097-461X(1996)60:7<1271::AID-QUA8>3.0.CO;2-W).
- (21) Besler, B. H.; Merz, K. M.; Kollman, P. A. Atomic Charges Derived from Semiempirical Methods. *J Comput Chem* **1990**, 11 (4), 431–439. <https://doi.org/10.1002/jcc.540110404>.
- (22) Case, D. A.; Belfon, K.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T. E. I.; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Giambasu, G.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Harris, R.; Izadi, S.; Izmailov, S. A.; Kasavajhala, K.; Kovalenko, A.; Krasny, R.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Man, V.; Merz, K. M.; Miao, Y.; Mikhailovskii, O.; Monard, G.; Nguyen, H.; Onufriev, A.; Pan, F.; Pantano, S.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Skrynnikov, N. R.; Smith, J.; Swails, J.; Walker, R. C.; Wang, J.; Wilson, L.; Wolf, R. M.; Wu, X.; Xiong, Y.; Xue, Y.; D.M., Y.; Kollman, P. A. AMBER 2020, 2020.
- (23) Tian, C.; Kasavajhala, K.; Belfon, K. A. A.; Raguette, L.; Huang, H.; Migués, A.; Bickel, J.; Wang, Y.; Pincay, J.; Wu, Q.; Simmerling, C. ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *Journal of Chemical Theory and Computation* **2020**, 16, 528–552. <https://doi.org/10.1021/acs.jctc.9b00591>.
- (24) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *Journal of Molecular Graphics and Modelling* **2006**, 25, 247–260.
- (25) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General AMBER Force Field. *Journal of Computational Chemistry* **2004**, 25, 1157–1174.
- (26) Izadi, S.; Anandakrishnan, R.; Onufriev, A. V. Building Water Models : A Different Approach. *The Journal of Physical Chemistry Letters* **2014**, 5, 3863–3871.
- (27) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *Journal of Chemical Theory and Computation* **2013**, 9, 3084–3095.
- (28) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, 596 (7873), 583–589. <https://doi.org/10.1038/s41586-021-03819-2>.
- (29) Wenger, E. S.; Martinie, R. J.; Ushimaru, R.; Pollock, C. J.; Sil, D.; Li, A.; Hoang, N.; Palowitch, G. M.; Graham, B. P.; Schaperdorth, I.; Burke, E. J.; Maggiolo, A. O.; Chang, W.; Allen, B. D.; Krebs, C.; Silakov, A.; Boal, A. K.; Bollinger, J. M. Optimized Substrate Positioning Enables Switches in the C–H Cleavage Site and Reaction Outcome in the Hydroxylation–Epoxidation Sequence Catalyzed by Hyoscyamine 6 β -Hydroxylase. *J. Am. Chem. Soc.* **2024**, 146 (35), 24271–24287. <https://doi.org/10.1021/jacs.4c04406>.

- (30) Li, J.; Liao, H.-J.; Tang, Y.; Huang, J.-L.; Cha, L.; Lin, T.-S.; Lee, J. L.; Kurnikov, I. V.; Kurnikova, M. G.; Chang, W.; Chan, N.-L.; Guo, Y. Epoxidation Catalyzed by the Nonheme Iron(II)- and 2-Oxoglutarate-Dependent Oxygenase, AsqJ: Mechanistic Elucidation of Oxygen Atom Transfer by a Ferryl Intermediate. *J. Am. Chem. Soc.* **2020**, *142* (13), 6268–6284.
<https://doi.org/10.1021/jacs.0c00484>.