# Supplementary Information

## Mechanism of Ultrafast Flavin Photoreduction in the Active Site of Flavoenzyme LSD1 Histone Demethylase

Bo Zhuang,[*,a] Rivo Ramodiharilafy,[b] Alexey Aleksandrov,[*,b] Ursula Liebl,[b] and Marten H. Vos[*,b]

[a]*Beijing National Laboratory for Molecular Sciences, College of Chemistry and Molecular Engineering, Peking University, 100871 Beijing, China*

[b]*LOB, CNRS, INSERM, École Polytechnique, Institut Polytechnique de Paris, 91120 Palaiseau, France*

[*]Corresponding authors: zhuangbo@pku.edu.cn; alexey.aleksandrov@polytechnique.edu; marten.vos@polytechnique.edu

**Sample Preparations.** Plasmid pET15b-LSD, carrying human LSD1 transcript variant 2, was a kind gift from Prof. Yang Shi (now Ludwig Cancer Research, Oxford, UK). Site-directed mutants W751F and Y761F LSD1 were synthesized by Genecust (Genecust Europe, Luxembourg) and verified by sequencing. LSD1 proteins were expressed in *E. coli* BL21DE3, induced by 0.2 mM isopropyl β-D-1-thiogalactopyranoside and incubated at 16 ℃ overnight. The cell suspension was sonicated according to the manufacturer's instructions. The His-tagged LSD1 proteins were purified on Protino Ni-TED columns (Macherey-Nagel), followed by imidazole removal on Econo-Pac 10DG columns (BioRad). The eluted yellow fraction was concentrated (Amicon ultra-centrifugal filters), and glycerol was added to a final concentration of 10%. The sample was then flash frozen in liquid nitrogen and stored at −80 °C.

**Time-Resolved Spectroscopic Measurements.** The setup for time-resolved fluorescence employing a Kerr gate was described previously.[1,2] Briefly, part of the 780 nm output from the Ti:sapphire laser/amplifier system (Quantronix Integra-C) operating at 500 Hz was passed through a BBO crystal, yielding an excitation pulse centered at 390 nm. The remaining 780 nm beam was directed through a motorized delay-line and focused into the Kerr medium where it spatially overlapped the fluorescence emitted by the sample. Kerr media with different time-resolution/sensitivity compromises were used, i.e., suprasil (response time ca. 200 fs), and $CS_2$ (response time ca. 1 ps).

Multicolor time-resolved absorption spectra were recorded by the pump–probe technique on an instrument operating at 500 Hz, as described previously.[3,4] Pump pulse centered at 390 nm were obtained by frequency-doubling the fundamental beam and continuum broadband pulses were used as the probe. Pump and probe beams were set at the magic angle (54.7°) to record the isotropic spectra. For all the time-resolved measurements, the excitation power was adapted such that each shot of the pump beam excited less than 10% of the sample, and all the measurements were carried out at 10 °C to prevent protein precipitation. Global analysis of the data was performed using the Glotaran program.[5]

**Computer Simulations.** For molecular dynamics (MD) simulations, the structure of WT LSD1 from *Homo sapiens* was taken from the Protein Data Bank (PDB entry: 2DW4; resolution: 2.3 Å).[6] The CHARMM36m force field was used for the protein residues[7] and the modified TIP3P model for water;[8] the force field model from ref. 9 was used to describe the FAD cofactor. The

protonation states of all titratable residues were assigned based on an analysis PROPKA 3.1,[10] and verified by ideal stereochemistry. For the simulations of the W751F LSD1 and Y761F LSD1 variants, the initial conformations of the corresponding phenylalanine were predicted by the SCWRL4 program.[11] The systems were set up using the CHARMM program (version 47b1),[12] and the MD simulations were carried out using the NAMD programs (version 2.14).[12] Protein residues within a 24 Å sphere centered at the isoalloxazine ring N5 atom of flavins were placed in a 75-Å cubic box of water. Non-hydrogen atoms between 20 and 24 Å from the sphere center were harmonically restrained to their experimentally determined positions in the crystal structure. Periodic boundary conditions were assumed with long-range electrostatic interactions computed using the particle mesh Ewald method,[13] and an appropriate amount of potassium counter ions were included to render the system electrically neutral. The bond length between each hydrogen and the parent atom to which it is bonded was constrained using the SHAKE method[14] for the integration time step of 2 fs. After energy minimization, the system was equilibrated first in an NVT ensemble for 50 ps, followed by a 500 ps simulation in the NPT ensemble, at 295 K and 1.0 atm pressure. The Berendsen thermostat and barostat were employed, with a relaxation time of 500 fs and four timesteps between position rescalings for constant pressure simulations.[15] The production runs were then performed for 200 ns, collecting coordinates of the system every 100 ps.

QM/MM calculations on LSD1 variants were performed with the ORCA package (version 5.0.3).[16] QM-MM electrostatic interaction were treated using the electrostatic embedding, and the QM–MM boundary was described by a hydrogen-like link atom. The QM region, which was described by DFT, included the isoalloxazine ring moiety of FAD and the side chains of the residues Y761 and W761, or F751/F761 in the corresponding variants. The remaining protein components and water molecules were included in the MM region and described with the same CHARMM36m force fields as in the MD simulations. 20 equally spaced snapshots along the stabilized MD trajectories of 200 ns were used as input for the QM/MM calculations, which were first subjected to geometry optimization at the B3LYP/ma-def2-SVP[17] level (with the MM atoms fixed); excitation energies and oscillator strengths of 25 excited states were then estimated using time-dependent DFT at the cam-B3LYP[18]/ma-def2-SVP level. The Multiwfn programs (version: 3.8)[19] was used to perform the hole–electron analysis to identify the nature of each excited state.

Visualization of the results were carried out with the aid of the VMD programs (version 1.9.4a53).[20]

The electronic coupling matrixes between the donor LE state and donor–acceptor CT states were estimated along the MD simulations using the two-state generalized Mulliken-Hush (GMH) approximation,[21,22] based on the results of TDDFT calculations. For weak donor–acceptor interactions, the electronic coupling matrixes ($H_{ab}$) between two adiabatic states A and B is given by

$$H_{ab} = \frac{\mu_{ab}\Delta E_{ab}}{\sqrt{(\mu_{aa} - \mu_{bb})^2 + 4(\mu_{ab})^2}} \qquad (eq.\ s1)$$

where $\mu_{ab}$ and $\Delta E_{ab}$ are the transition dipole moment and the energy gap between the states A and B, respectively. $\mu_{aa}$ and $\mu_{bb}$ are the permanent dipole moments of the two adiabatic states. The values of the involved dipole moments were calculated from the TDDFT wavefunctions using the Multiwfn programs.

To estimate the effects of molecular configurations on coupling between the LE and CT states, a model system composed of only the flavin ring (lumiflavin) and sidechain of tyrosine (*p*-methylphenol), adopting either a stacking or non-stacking configuration at a separation of ca. 4 Å, was subjected to TDDFT calculations at the cam-B3LYP/ma-def2-TZVP level. A conductor-like polarizable continuum model (CPCM) with an $\varepsilon$ value of 6 was used to mimic the weak dielectric field of a protein environment.[23,24] The coupling matrixes were then evaluated for the two configurations after identifying the relevant LE and CT states.

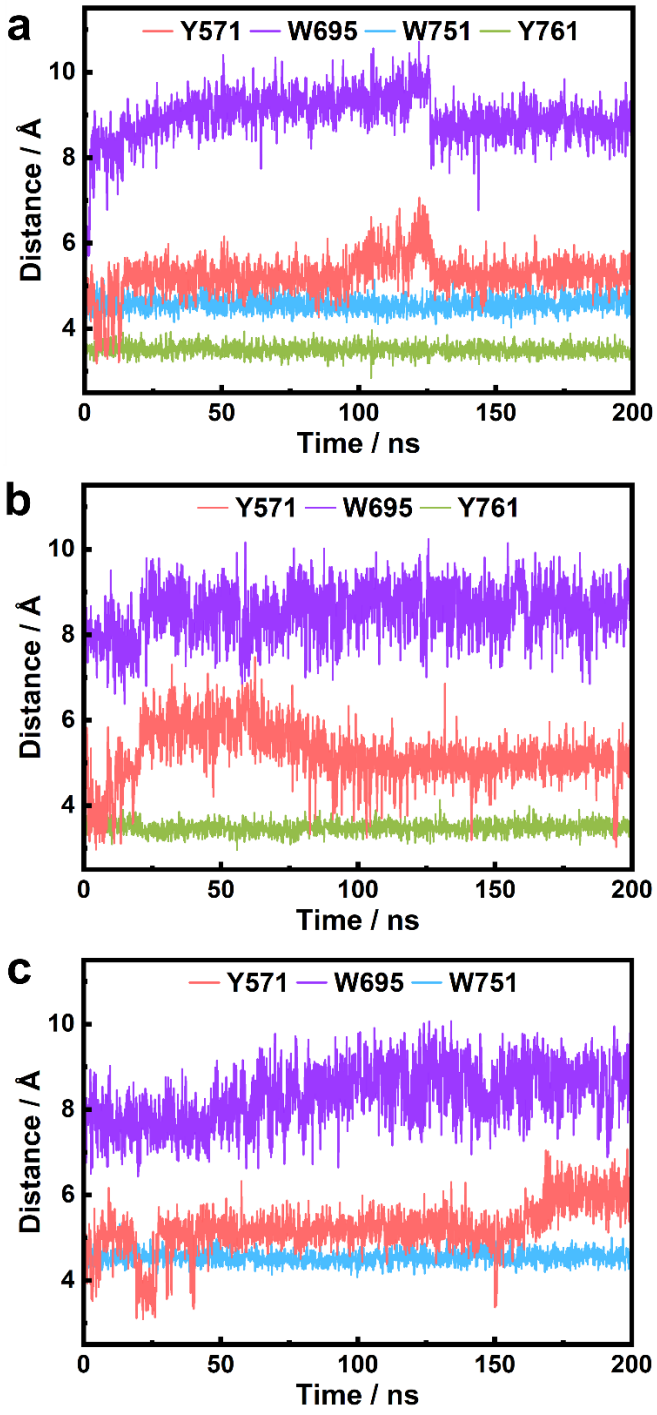


**Fig. S1** Dynamics of distances between FAD and close-by tyrosine and tryptophan residues in 200-ns MD simulations in WT (a), W751F (b), and Y761F (c) LSD1.

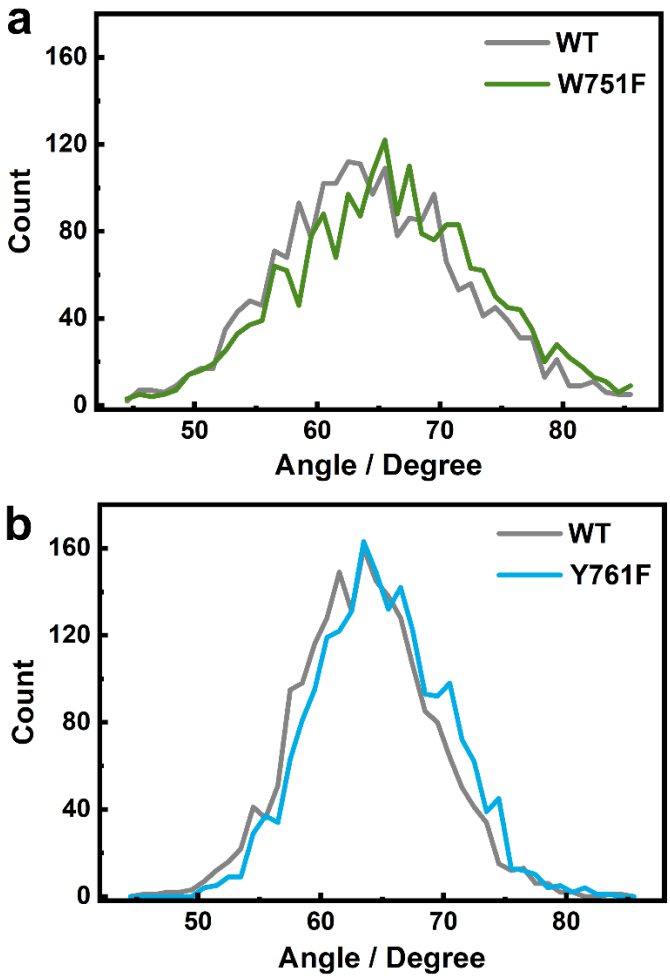


**Fig. S2** Distributions of the angles between the ring planes of the isoalloxazine moiety of FAD and the indole or phenol moiety of Y761 (a) or W751 (b) during the MD simulations of WT, W751F and Y761F LSD1, as indicated.

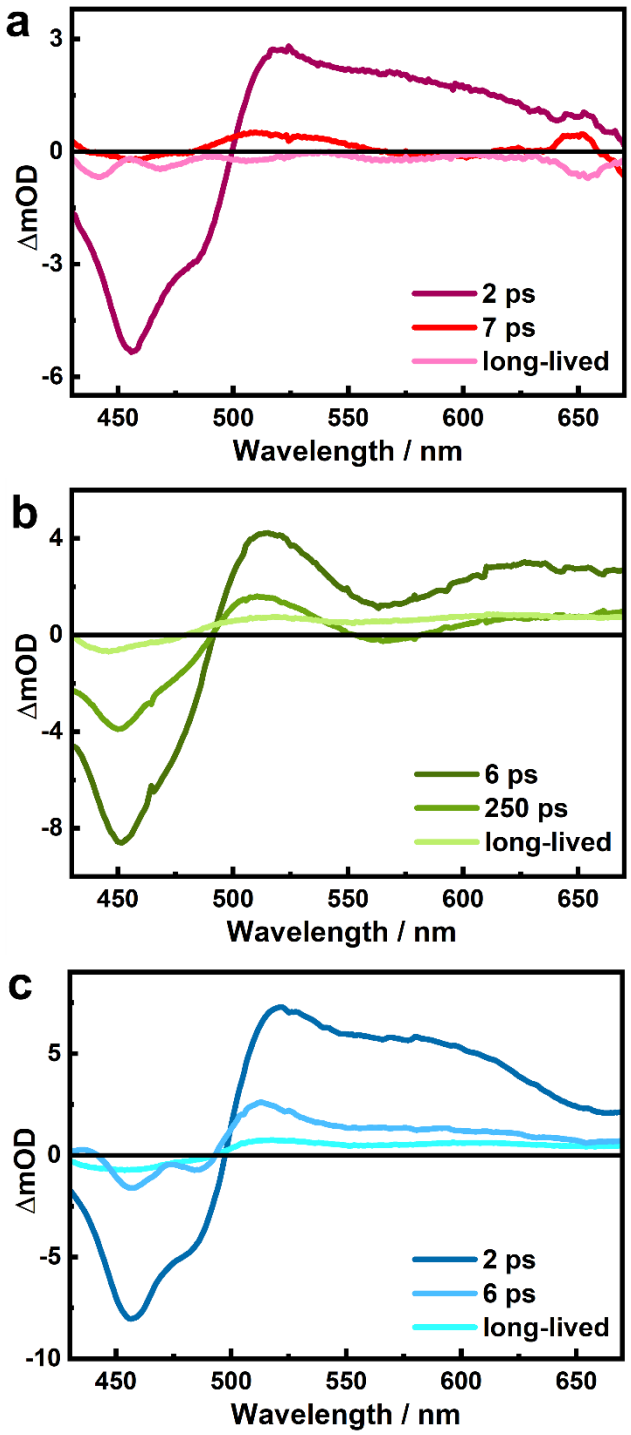


**Fig. S3** EAS obtained from global analysis of transient absorption data of WT (a), W751F (b) and Y761F (c) LSD1.

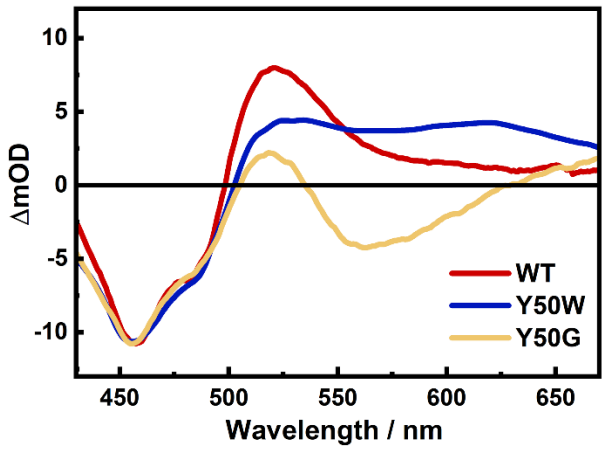


**Fig. S4** EAS obtained from global analysis of transient absorption data of WT, Y50W and Y50G *Bs*FNR, reproduced from ref. 4. In addition to the GSB of FAD, the spectrum of WT *Bs*FNR comprises the induced absorption of $Y^{\bullet+}/FAD^{\bullet-}$ radical pair, the spectrum of Y50W *Bs*FNR comprises the induced absorption of $W^{\bullet+}/FAD^{\bullet-}$ radical pair, and the spectrum of Y50G *Bs*FNR involves the absorption of $FAD^{*}$ and SE.

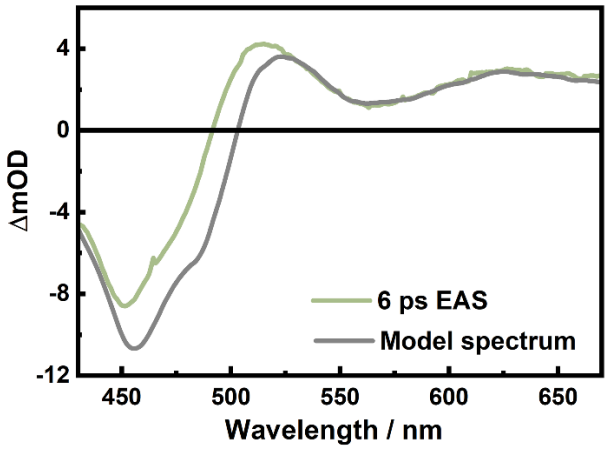


**Fig. S5** Comparison of the 6 ps EAS of W751F LSD1 (Fig. S3b) with a model TA spectrum that comprises 30% $FAD^{*}$ (including both induced absorption and SE) and 70% $W^{\bullet+}/FAD^{\bullet-}$ radical pair, constructed based on the TA spectra of Y50G and Y50W *Bs*FNR (Fig. S4), respectively.

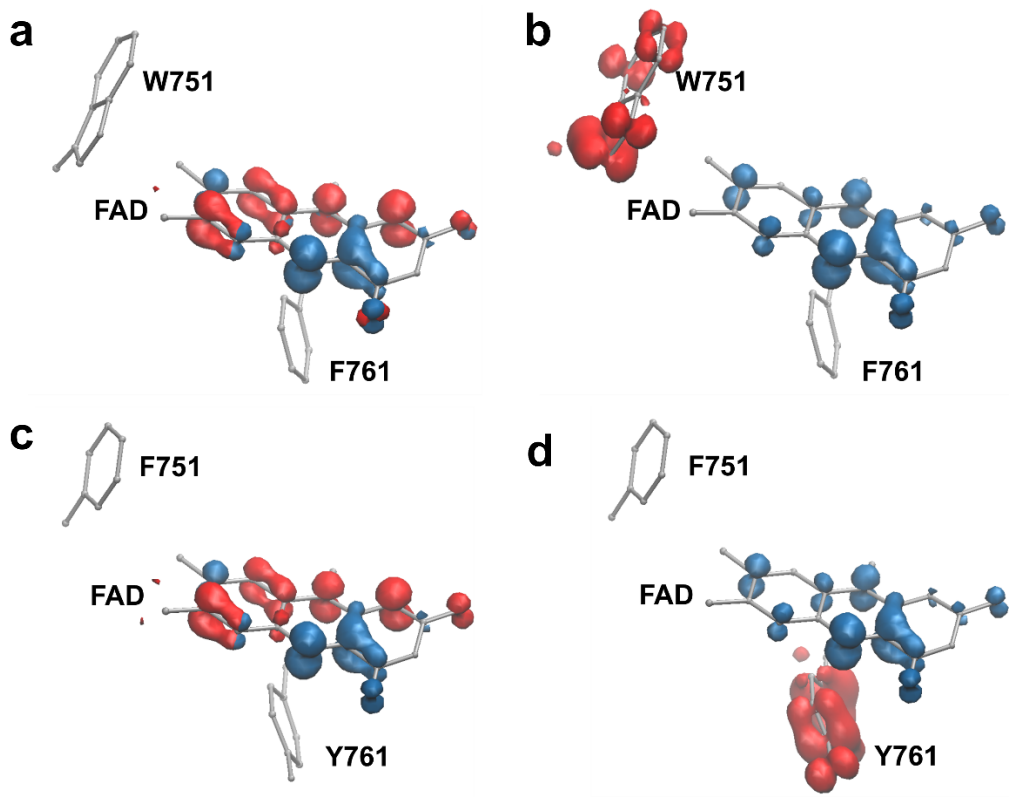


**Fig. S6** Hole–electron distributions of $LE_{FAD}$ (a) and $CT_{Y761}$ (b) in Y751F LSD1, and those of $LE_{FAD}$ (c) and $CT_{W751}$ (d) in Y761F LSD1. Red and blue regions denote the hole and electron distributions, respectively (isovalue = 0.004). Hydrogen atoms are not shown for clarity.

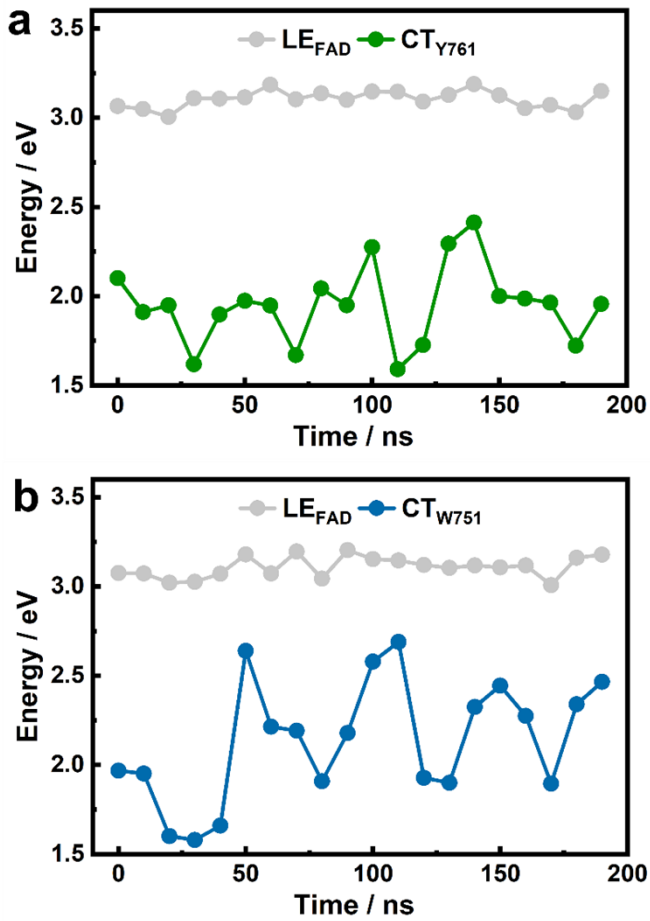


**Fig. S7** Energy levels of $LE_{FAD}$ and $CT_{Y761}$ in W751F LSD1 (a) and those of $LE_{FAD}$ and $CT_{W751}$ in Y761F LSD1 along the MD trajectories.

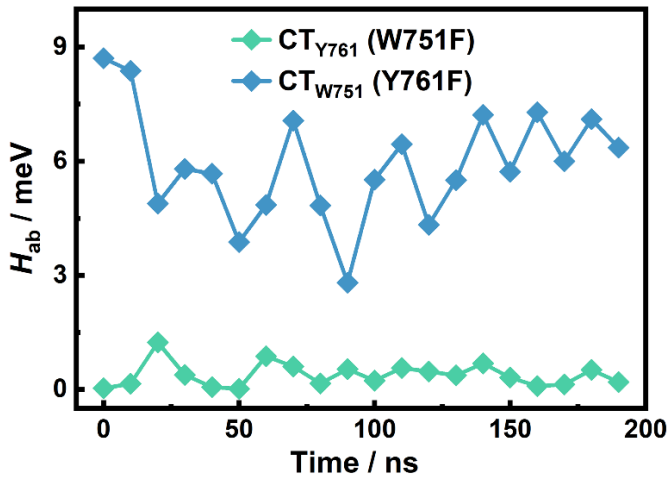


**Fig. S8** Electronic couplings between the LE and CT states in W751F and Y761F LSD1 along the MD trajectories.

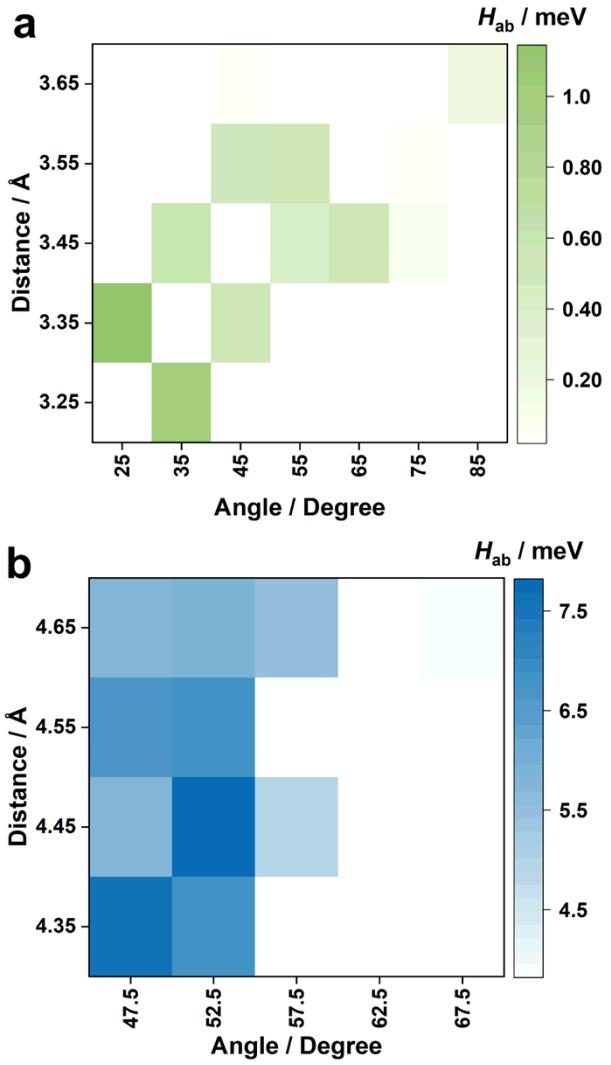


**Fig. S9** Two-dimensional heatmap of the electronic couplings between $LE_{FAD}$ and $CT_{Y761}$ (a) and those between $LE_{FAD}$ and $CT_{W751}$ (b) in WT LSD1, along the MD trajectories, plotted against the distance and angles between the aromatic residues and the flavin ring (after QM/MM geometry optimization).

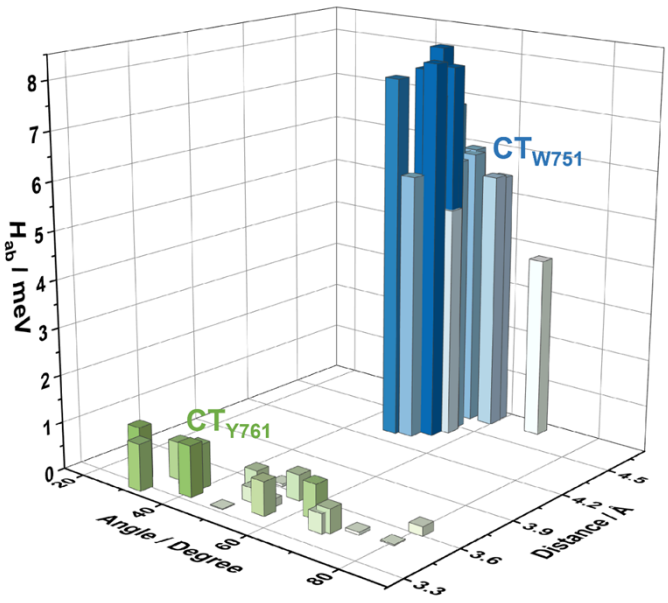


**Fig. S10** Three-dimensional bar chart of the electronic couplings between $LE_{FAD}$ and $CT_{Y761}$ and those between $LE_{FAD}$ and $CT_{W751}$, as indicated, along the MD trajectories of WT LSD1, plotted against the distance and angles between the aromatic residues and the flavin ring (after QM/MM geometry optimization).

**Table S1.** Time constants from the global fits of time-resolved fluorescence decays of LSD1 variants.[a]

| Protein | $\tau_1$ / ps | $\tau_2$ / ps | $\tau_3$ / ps |
|---|---|---|---|
| WT | 0.18 (0.89) | 5 (0.11) | |
| W751F | 0.27 (0.86) | 7 (0.07) | > 1000 (0.07) |
| Y761F | 9 (0.76) | 200 (0.24) | |

[a]The pre-exponential factor of each component is given in parentheses.

**Table S2.** Averaged excitation energy ($E_{ex}$), oscillator strength ($f_{osc}$) of $LE_{FAD}$, $CT_{Y761}$ and $CT_{W751}$, as well as the electronic coupling ($H_{ab}$) between the CT and LE states, in LSD1 variants along the MD trajectories.[a] N/A, not applicable.

| Protein | State | $E_{ex}$ / eV | $f_{osc}$ | $H_{ab}$ / meV |
|---|---|---|---|---|
| WT | $LE_{FAD}$ | 3.08 (0.04) | 0.19 (0.03) | N/A |
| | $CT_{Y761}$ | 2.02 (0.19) | 0.00 (0.00) | 0.46 (0.39) |
| | $CT_{W751}$ | 2.25 (0.23) | 0.00 (0.00) | 6.42 (1.21) |
| W751F | $LE_{FAD}$ | 3.10 (0.05) | 0.18 (0.03) | N/A |
| | $CT_{Y761}$ | 1.94 (0.26) | 0.00 (0.00) | 0.37 (0.31) |
| Y761F | $LE_{FAD}$ | 3.11 (0.06) | 0.18 (0.02) | N/A |
| | $CT_{W751}$ | 2.14 (0.34) | 0.00 (0.00) | 5.92 (1.47) |

[a]Statistical uncertainty in parentheses estimated as the standard deviation.

**Supporting references**


(1) Laptenok, S. P.; Bouzhir-Sima, L.; Lambry, J. C.; Myllykallio, H.; Liebl, U.; Vos, M. H. Ultrafast Real-Time Visualization of Active Site Flexibility of Flavoenzyme Thymidylate Synthase ThyX. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (22), 8924–8929.

(2) Laptenok, S. P.; Nuernberger, P.; Lukacs, A.; Vos, M. H. Subpicosecond Kerr-Gate Spectrofluorometry. In *Methods in Molecular Biology, Fluorescence Spectroscopy and Microscopy: Methods and Protocols, Vol. 1076*; Engelborghs, Y., Visser, A. J. W. G., Eds.; Humana Press: New York, 2014; pp 321–336.

(3) Nag, L.; Sournia, P.; Myllykallio, H.; Liebl, U.; Vos, M. H. Identification of the TyrOH•+ Radical Cation in the Flavoenzyme TrmFO. *J. Am. Chem. Soc.* **2017**, *139* (33), 11500–11505.

(4) Zhuang, B.; Seo, D.; Aleksandrov, A.; Vos, M. Characterization of Light-Induced, Short-Lived Interacting Radicals in the Active Site of Flavoprotein Ferredoxin-NADP+ Oxidoreductase. *J. Am. Chem. Soc.* **2021**, *143* (7), 2757–2768.

(5) Snellenburg, J. J.; Laptenok, S. P.; Seger, R.; Mullen, K. M.; van Stokkum, I. H. M. Glotaran: A Java-Based Graphical User Interface for the R Package TIMP. *J. Stat. Softw.* **2012**, *49* (3).

(6) Mimasu, S.; Sengoku, T.; Fukuzawa, S.; Umehara, T.; Yokoyama, S. Crystal Structure of Histone Demethylase LSD1 and Tranylcypromine at 2.25 Å. *Biochem. Biophys. Res. Commun.* **2008**, *366* (1), 15–22.

(7) Huang, J.; Mackerell, A. D. CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *J. Comput. Chem.* **2013**, *34* (25), 2135–2145.

(8) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79* (2), 926–935.

(9) Aleksandrov, A. A Molecular Mechanics Model for Flavins. *J. Comp. Chem.* **2019**, *40* (32), 2834–2842.

(10) Olsson, M. H. M.; SØndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical PKa Predictions. *J. Chem. Theory Comput.* **2011**, *7* (2), 525–537.

(11) Krivov, G. G.; Shapovalov, M. V.; Dunbrack, R. L. Improved Prediction of Protein Side-Chain Conformations with SCWRL4. *Proteins Struct. Funct. Bioinforma.* **2009**, *77* (4), 778–795.

(12) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26* (16), 1781–1802.

(13) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N·log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98* (12), 10089–10092.

(14) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* **1977**, *23* (3), 327–341.

(15) Berendsen, H. J. C.; Postma, J. P. M.; Gunsteren, W. F. van; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81* (8), 3684.

(16) Neese, F. The ORCA Program System. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2* (1), 73–78.

(17) Zheng, J.; Xu, X.; Truhlar, D. G. Minimally Augmented Karlsruhe Basis Sets. *Theor. Chem. Acc.* **2011**, *128* (3), 295–305.

(18) Yanai, T.; Tew, D. P.; Handy, N. C. A New Hybrid Exchange–Correlation Functional Using the Coulomb-Attenuating Method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393* (1–3), 51–57.

(19) Lu, T.; Chen, F. Multiwfn: A Multifunctional Wavefunction Analyzer. *J. Comput. Chem.* **2012**, *33* (5), 580–592.

(20) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol.*

*Graph.* **1996**, *14* (1), 33–38.

(21) Cave, R. J.; Newton, M. D. Generalization of the Mulliken-Hush Treatment for the Calculation of Electron Transfer Matrix Elements. *Chem. Phys. Lett.* **1996**, *249* (1–2), 15–19.

(22) Cave, R. J.; Newton, M. D. Calculation of Electronic Coupling Matrix Elements for Ground and Excited State Electron Transfer Reactions: Comparison of the Generalized Mulliken–Hush and Block Diagonalization Methods. *J. Chem. Phys.* **1997**, *106* (22), 9213–9226.

(23) Li, L.; Li, C.; Zhang, Z.; Alexov, E. On the Dielectric "Constant" of Proteins: Smooth Dielectric Function for Macromolecular Modeling and Its Implementation in DelPhi. *J. Chem. Theory Comput.* **2013**, *9* (4), 2126–2136.

(24) Su, D.; Pabel Kabir, M.; Orozco-Gonzalez, Y.; Gozem, S.; Gadda, G. Fluorescence Properties of Flavin Semiquinone Radicals in Nitronate Monooxygenase. *ChemBioChem* **2019**, *20*, 1646–1652.