

Supplementary Information

Mechanism and Dynamics of Photoswitchable Flavoprotein Charge-Transfer Complexes

Bo Zhuang,^{*a} Guangliu Ran,^b Wenkai Zhang^{*b} and Feng Gai^{*a}

^aBeijing National Laboratory for Molecular Sciences, College of Chemistry and Molecular Engineering, Peking University, Beijing, 100871, China

^bSchool of Physics and Astronomy, Applied Optics Beijing Area Major Laboratory, Center for Advanced Quantum Studies, Beijing Normal University, Beijing 100875, China

***Corresponding authors:** zhuangbo@pku.edu.cn; wkzhang@bnu.edu.cn; fgai@pku.edu.cn

Content

1. Methods
2. Electronic properties of MXA:FAD in MSOX_{WT}
3. TA data of CT complexes in MSOX_{WT} and MTOX_{WT}
4. Determination of the rate constants for the MSeA:FAD complex in MSOX_{WT}
5. Steady- and excited-state properties of MXA:FAD in other protein variants
6. Additional results of MD simulations and QM/MM potential energy scans

1. Methods

Sample Preparation

MSA and PCA were purchased from Shanghai Adamas Reagent Co., Ltd. MSeA was synthesized as previously described.^{1,2}

MSOX_{WT} from *Bacillus sp.* was purchased from Nanjing Duly Biotechnology Co., Ltd. and used without further purification. The genes coding for MTOX_{WT}, MTOX_{T239M}, and MTOX_{T239K} from *E. coli*, as well as MSOX_{M245T} and MSOX_{M245K} from *Bacillus sp.*, carrying a hexahistidine (6xHis)-tag at their carboxyl terminus, were inserted into plasmid pET-28a(+) (Novagen) via the *NcoI* and *XhoI* restriction sites. The expression constructs were used to transform *E. coli* strain BL21(DE3). For the production of each protein variant, a single colony of the bacteria was inoculated into 45 mL of LB media supplemented with kanamycin (50 µg/mL), and cultured at 37°C for 5 h. It was later transferred into 1 L LB media with the same antibiotic and grown at 37°C until the culture reached OD₆₀₀ = 0.6. The expression of protein was induced by the addition of isopropyl-β-D-1-thiogalactopyranoside (1 mM), and the bacteria were cultivated at 30°C for 8 h. Cells were harvested by centrifugation at 8000 rpm for 20 min, resuspended in cold lysis buffer (100 mM Tris, 150 mM NaCl, 10% (v/v) glycerol, pH 8), and then lysed by sonication (Cole-Parmer 130-Watt Ultrasonic Processors) in ice bath applying 20 cycles of 5 s pulses followed by a 5 s cool down. Unless indicated otherwise, all the following purification steps were performed at 4°C. The lysates were clarified by centrifugation at 12000 rpm for 30 min. The resulting supernatants were filtered through a 0.45 µm syringe filter and loaded onto a HisTrap FF column (5 ml, Cytiva). Purification of the expressed protein was performed using the Äkta Explorer FPLC system (GE Healthcare; binding buffer: 20 mM Tris, 200 mM NaCl, 10 mM imidazole, pH 8.0; elution buffer: 20 mM Tris, 200 mM NaCl, 500 mM imidazole, pH 8.0). The obtained protein solution was then concentrated and exchanged to a buffer for storage (100 mM Tris, 150 mM NaCl, 10% v/v glycerol, pH 8), using 10 kDa cut-off centrifuge concentrators at 4000 rpm. Protein purity was evaluated by SDS PAGE.

Spectroscopic Measurements

Steady-state absorption spectra were recorded with an Agilent Technologies Cary 60 UV–Vis spectrophotometer.

The setup for TA measurements was based on the commercial HARPIA-TA transient absorption spectrometer (Light Conversion). Briefly, a Yb:KGW laser (1030 nm, 54 kHz, Light Conversion) is split into two fundamental light beams. One of the light beams is transferred to the optical parametric amplifier (Orpheus, Light Conversion) to generate a high-intensity pulse of a specific wavelength as the pump beam. Meanwhile, the other is focused on a 5-mm sapphire (EKSMA Optics) to generate low-intensity continuum light in the range of ca. 500–900 nm, employed as the probe beam. Alternatively, the 1030-nm beam was passed through a BBO crystal to generate a 515-nm beam, which was then focused on the 5-mm sapphire to generate the continuum probe light in the range of ca. 390–500 nm.³ The pump and probe beams were set at the magic angle (54.7°) and spatially overlapped at the sample. The time delays between pump and probe beams were achieved using a delay stage monitor, and the transmitted probe light is collected using a charge-coupled device. The instrument response function has a 20-to-80% rise time of ca. 110 fs, as determined from experiments on a reference sample, i.e., a polymeric thin film named D18 (Fig. S10, *inset*).⁴ The optical path of the sample cell was 2 mm, and the optical density of each sample was in the range of ca. 0.1–0.4 at the excitation wavelengths. The pump fluence was kept at ca. 25 $\mu\text{J}/\text{cm}^2$, which was confirmed to be in the linear regime for the sample with highest optical density. The spot size of the pump beam was ca. 160 μm . The sample cells were regularly moved at a rate of ca. 1 mm/s during the accumulation of data, to avoid possible thermal damage due to the optical pumping. No noticeable sample degradation was observed, as confirmed by steady-state absorption measurements before and after the TA measurements. The Glotaran program⁵ was used to fit the TA kinetics.

MD Simulations

MD simulations with classical force fields were performed using the NAMD program (version 3.0b6).⁶ The structure of MSOX_{WT} from *Bacillus sp.* with bound MSA in *Conf 1* (PDB entry: 1EL9; resolution: 2.0 Å), and the structure of MTOX_{WT} (PDB entry: 2UZZ; resolution: 3.2 Å) were taken from the Protein Data Bank. The CHARMM36m force field was used for the protein residues⁷ and the TIP3P model for water;⁸ the force field parameters for the covalently bound FAD

cofactors were adapted from the reported parameters of flavin derivatives.⁹ The force field parameters for the MSA ligand were generated by the CGenFF program (interface: version 1.0.0; force field: version 3.0.1).^{10,11} The initial structures of the mutated proteins were generated using the CHARMM-GUI web server¹² based on the crystal structures of the wild-type proteins, and the MSA ligand was added to the protein active sites in a conformation similar to that in the experimental crystal structure of MSA-bound MSOX_{WT} (*Conf 1*). The protonation states of titratable residues were assigned based on a PROPKA 3.1 analysis¹³ and verified by ideal stereochemistry taking into account of steric effects and potential hydrogen-bonding interactions. The system was centered in a cubic box of aqueous solvent, at least 12 Å away from each of the box edges; the final system contained around 20,000 water molecules and an appropriate number of K⁺ was included to neutralize the net charge of the systems. Periodic boundary conditions were assumed with long-range electrostatic interactions computed using the particle mesh Ewald method.¹⁴ The integration time step was set at 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 For MSOX_{WT}, the production runs were performed for 300 ns, collecting coordinates of the system every 100 ps. For the other protein variant, as the initial structures were artificially generated, the production runs were performed for 500 ns, and analysis was performed based on the trajectories of the last 300 ns.

QM/MM Calculations

QM/MM Calculations were carried out using the ORCA package (version: 5.0.4).¹⁶ QM–MM electrostatic interactions were treated using the electrostatic embedding. Unless indicated otherwise, the QM region included the isoalloxazine ring moiety of FAD, the MXA ligand, and the side chain of the Met245 (MSOX_{WT}) or Thr239 (MTOX_{WT}) residue, which was described by DFT with the hybrid exchange correlation functional cam-B3LYP,¹⁷ and the ma-def2-SVP basis set.¹⁸ For the MTeA ligand, effective core potentials implemented in ORCA were applied for the core electrons of the heavy element Te. The remaining protein components and water molecules (ca. 384 residues and 20,000 water molecules) were included in the MM region and described with the same force fields as in the MD simulations, which were kept fixed in the QM/MM calculations.

Protein geometries relaxed from the experimental crystal by MD simulations in an NVT ensemble for 50 ps were used as the input for geometry optimizations at the QM level. To describe the vertical excitation and obtain the absorption spectra of the MXA:FAD systems in MSOX_{WT}, the excitation energies and oscillator strengths of 30 excited states were then estimated using TDDFT, where were broadened with a Gaussian line shape (FWHM: 0.5 eV) to obtain the absorption spectra. Estimations of the excited-state geometries were also performed based on TDDFT at the cam-B3LYP/ma-def2-SVP level. In the relaxed scans, the MSeA ligand was placed in a conformation similar to *Conf 1* in the experimental crystal structure of MSeA-bound MSOX_{WT} (PDB entry: resolution: 2.0 Å) as the initial geometry. Calculations of the atomic dipole corrected Hirshfeld atomic charge,¹⁹ IGMH analysis,²⁰ and hole–electron analysis of the excited states²¹ were performed using the Multiwfn program (version 3.8(dev)).²² Visualization of the results were carried out with the aid of the VMD programs (version 1.9.4a53).²³

2. Electronic properties of MXA:FAD in MSOX_{WT}

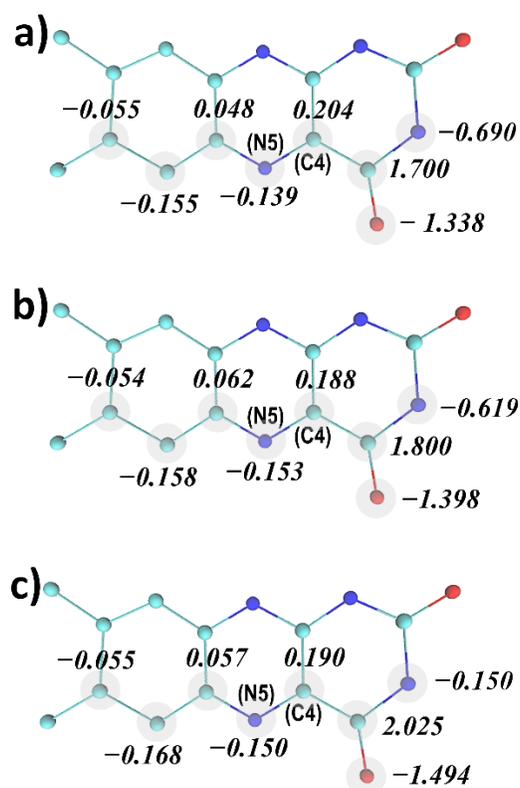


Fig. S1 Atomic charges of the flavin ring of the MSA:FAD (a), MSeA:FAD (b) and MTeA:FAD (c) complexes in *Conf 2* in MSOX_{WT}, computed at their QM/MM-optimized ground-state geometries.

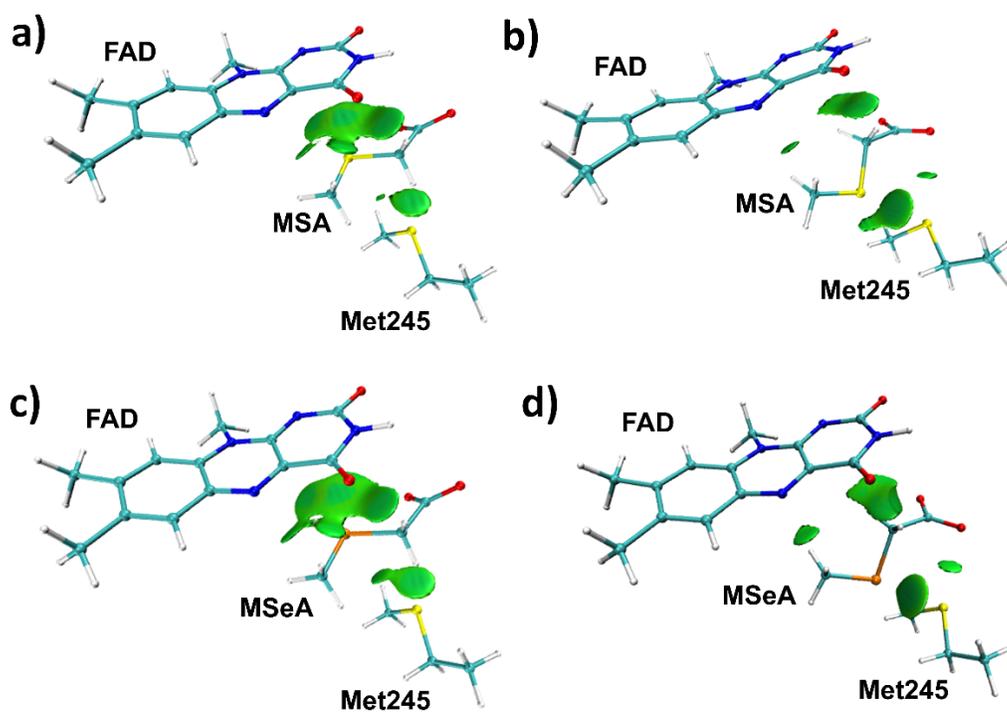


Fig. S2 Visualization of the intermolecular interactions between FAD, MSA, and Met245 (a, b) and those between FAD, MSeA, and Met245 (c, d) in MSOX_{WT} by IGMH analysis, with the ligand bound in *Conf 1* (a, c) and *Conf 2* (b, d). The colored isosurfaces correspond to $\delta g^{\text{inter}} = 0.006 \text{ a.u.}$

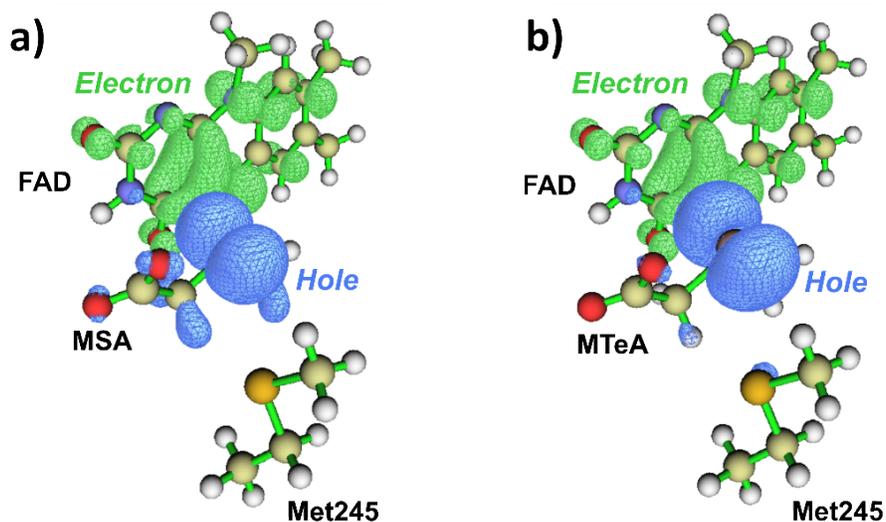


Fig. S3 Hole–electron distributions of the CT excited state of the MSA:FAD (a) and MTeA:FAD (b) complexes in MSOX_{WT} (isovalue: 0.002).

Table S1. Energies of the experimentally determined absorption maxima (λ_{CT}), the corresponding extinction coefficients (ϵ_{CT}) and full width at half maximum (FWHM) of the CT bands, as well as the theoretically calculated vertical transition energies (E_{CT}) and oscillator strengths (f_{osc}) of the MSA, MSeA and MTeA complexes in MSOX_{WT}.

	Experiment ^[a]			Calculation ^[b]	
	λ_{CT} (eV)	ϵ_{CT} (cm ⁻¹ M ⁻¹)	FWHM ^[c] (nm)	E_{CT} (eV)	f_{osc}
MSA	0.241	9217	121	0.266	0.2154
MSeA	0.220	8455	142	0.244	0.2491
MTeA	0.185	6707	191	0.219	0.3245

^[a]Taken from ref. 24.

^[b]Estimated for *Conf 1*.

^[c]Obtained by fitting the absorption bands in Fig. 1B of the main text with overlapping Gaussian functions.

3. TA data of CT complexes in MSeA:WT and MSeA:WT

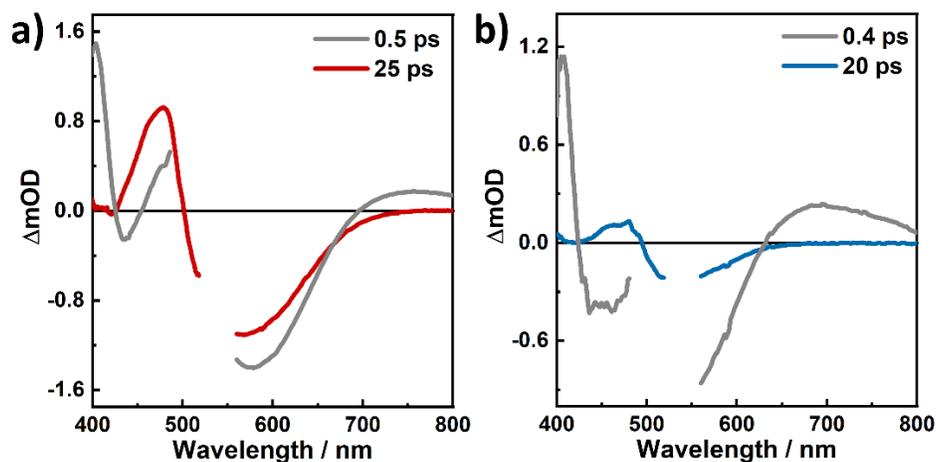


Fig. S4 TA spectra of MSeA:FAD in MSeA:WT (a) and MSeA:WT (b) at selected time delays, with 550-nm excitation.

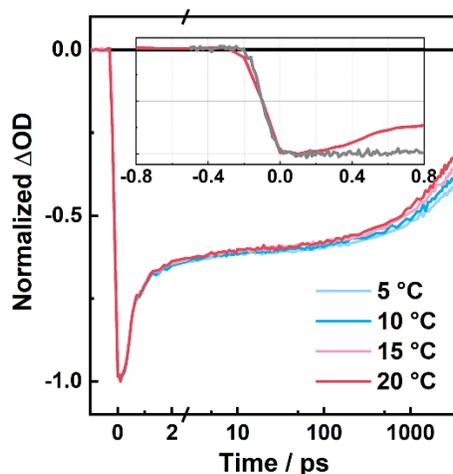


Fig. S5 TA kinetics of MSeA:FAD in MSeA:WT, probed at 570 nm and recorded at different temperatures, with 550-nm excitation. The inset shows a zoom-in of the 20 °C data on the sub-picosecond timescale (red), as well as an instrument response-limited rise of a negative signal (grey).

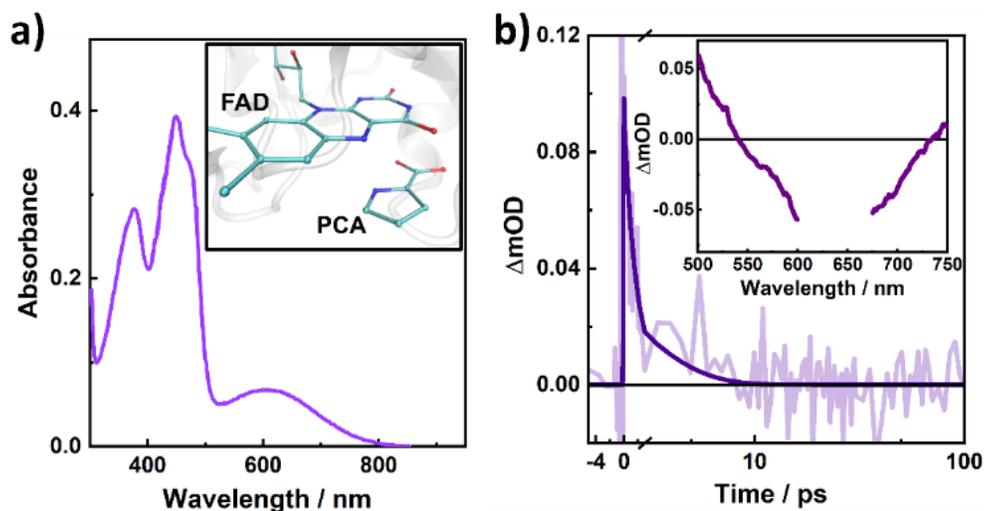


Fig. S6 (a) Steady-state absorption spectrum of the PCA:FAD CT complex in MSOX_{WT}. The inset shows the active site of MSOX_{WT} complexed with PCA (PDB entry: 1ELI). (b) TA kinetics of the PCA:FAD complex at 502 nm with 640-nm excitation. The inset shows the corresponding TA spectrum recorded at 500 fs after the excitation.

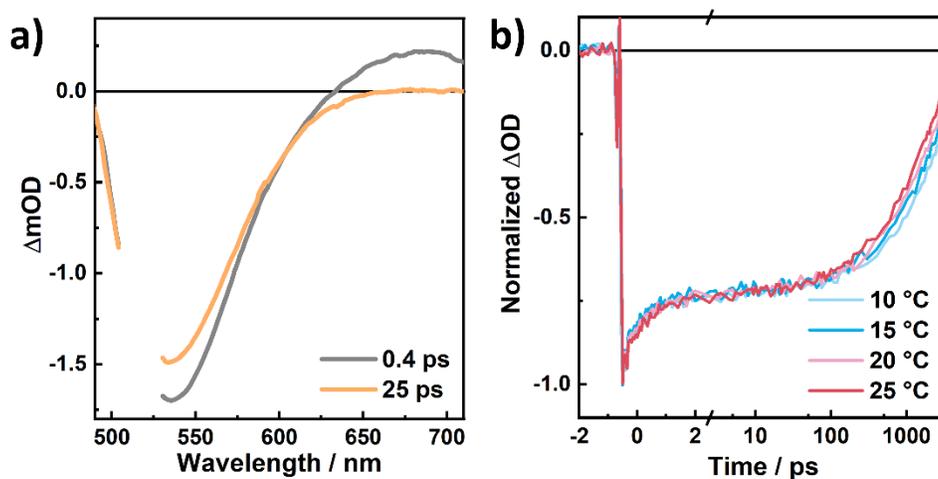


Fig. S7 TA spectra at selected time delays (a) and kinetics probed at 534 nm and recorded at different temperatures (b) of MSA:FAD in MSOX_{WT}, with 520-nm excitation (see also ref. 2).

4. Determination of the rate constants for the MSeA:FAD complex in MSOX_{WT}

Following the reaction scheme given in Fig. 4D of the main text, for the photoswitching of the MSeA:FAD complex in MSOX_{WT}, which only involves two distinct intermediate states (*State A* and *State B*), the kinetic processes give rise to the following rate equations:

$$\frac{d[A]}{dt} = -(k_1 + k_2)[A], \quad (\text{S1})$$

and

$$\frac{d[B]}{dt} = k_2[A] - k_3[B]. \quad (\text{S2})$$

where $[A]$ and $[B]$ are the apparent concentrations of *State A* and *State B*. Solving these rate equations, we get:

$$[A] = [A]_0 e^{-(k_1 + k_2)t}, \quad (\text{S3})$$

and

$$[B] = \frac{k_2[A]_0}{-k_1 - k_2 + k_3} [e^{-(k_1 + k_2)t} - e^{-k_3t}], \quad (\text{S4})$$

where $[A]_0$ corresponds to the initial concentration of the excited-state species. Meanwhile, the experimental observed TA signals are given by:

$$\Delta OD = \epsilon_A[A] + \epsilon_B[B], \quad (\text{S5})$$

where ϵ_A and ϵ_B are the extinction coefficients of *State A* and *State B*, respectively, which can be evaluated as:

$$\epsilon_A = \epsilon_{MSeA:FAD}^{ex} - \epsilon_{MSeA:FAD}^{gs}, \quad (\text{S6})$$

$$\epsilon_B = \epsilon_{FAD} - \epsilon_{MSeA:FAD}^{gs}, \quad (\text{S7})$$

where $\epsilon_{MSeA:FAD}^{ex}$ is the extinction coefficients of the initial excited state, $\epsilon_{MSeA:FAD}^{gs}$ is the ground-state extinction coefficients and ϵ_{FAD} is the extinction coefficients of the oxidized FAD cofactor. Although $\epsilon_{MSeA:FAD}^{ex}$ is unknown and cannot be experimentally determined, ϵ_B can be approximated by the steady-state difference spectra (Fig. 4A of the main text). Moreover, the initial

concentration of the excited state species $[A]_0$ can be determined by using $[\text{Ru}(\text{bpy})_3]^{2+}$ as an external reference:^{25–27}

$$[A]_0 = \frac{[Ru]_0(1 - T_{Ru})}{1 - T_{MSeA:FAD}}, \quad (\text{S8})$$

where T_{Ru} and $T_{MSeA:FAD}$ are the steady-state transmittance of the actual samples of $[\text{Ru}(\text{bpy})_3]^{2+}$ and the CT complexes, respectively; $[Ru]_0$ is the concentration of excited-state $[\text{Ru}(\text{bpy})_3]^{2+}$, which can be estimated via a reference TA measurement with the same experimental conditions and the known excited-state extinction coefficients of $[\text{Ru}(\text{bpy})_3]^{2+}$.²⁸

From equations (S3) and (S4), it is evident that the experimentally determined time constants

from the TA data are given as $\tau_1 = \frac{1}{k_1 + k_2}$, $\tau_2 = \frac{1}{k_3}$. For MSeA:FAD in MSOX_{WT}, $k_1 + k_2 \gg k_3$ (τ_2 is greater than τ_1 by three orders of magnitude). Thus, on the ps-ns timescale, the TA signal can be approximated by:

$$\Delta OD = \epsilon_B \frac{k_2[A]_0}{k_1 + k_2 - k_3} e^{-k_3 t} \quad (\text{S9})$$

Combining equations (S7) – (S9) allows us to solve for the absolute QY of the photoswitching, which is $k_2/(k_1 + k_2)$, as well as the underlying rates k_1 , and k_2 , from the TA data.

5. Steady- and excited-state properties of MSeA:FAD in other protein variants

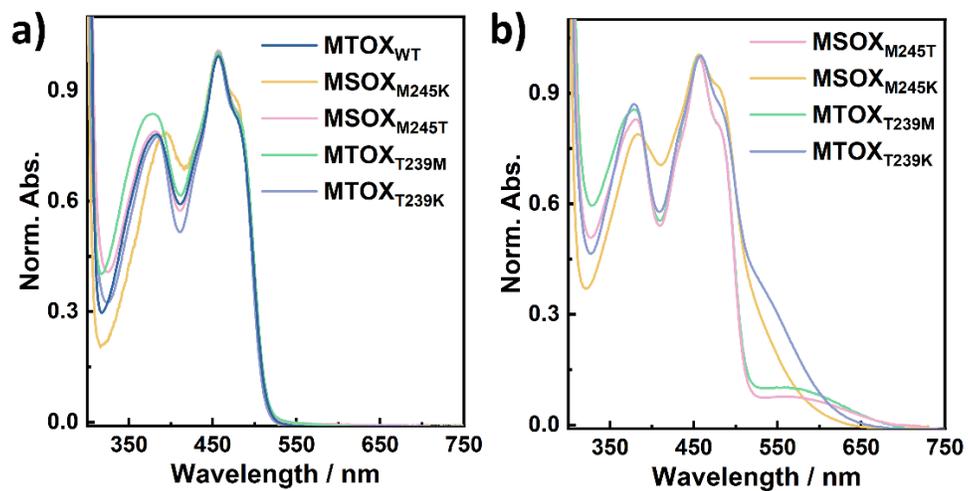


Fig. S8. Steady-state absorption spectra of MSeA:FAD in various protein variants without a ligand (a) and in the presence of ca. 100 mM MSeA (b).

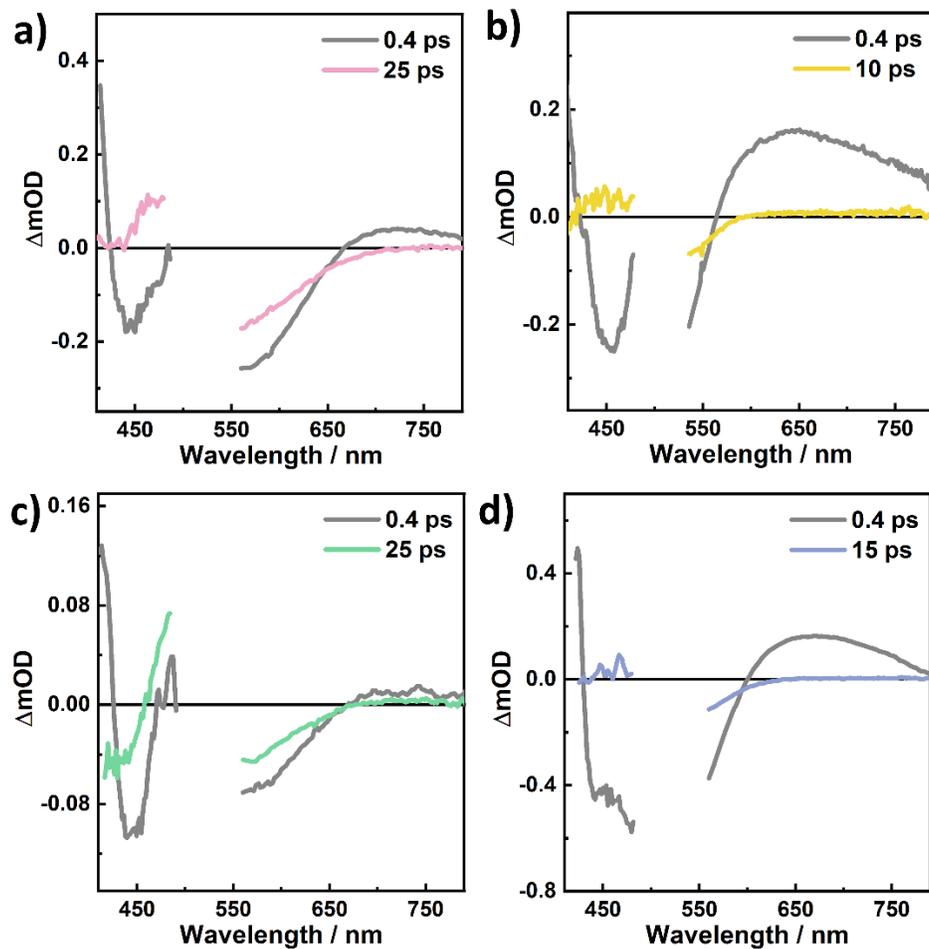


Fig. S9 TA spectra of MSeA:FAD in MSOX_{M245T} (a), MSOX_{M245K} (b), MTOX_{T239M} (c) and MTOX_{T239K} (d) at selected time delays, recorded near room temperature and upon excitation at the CT bands.

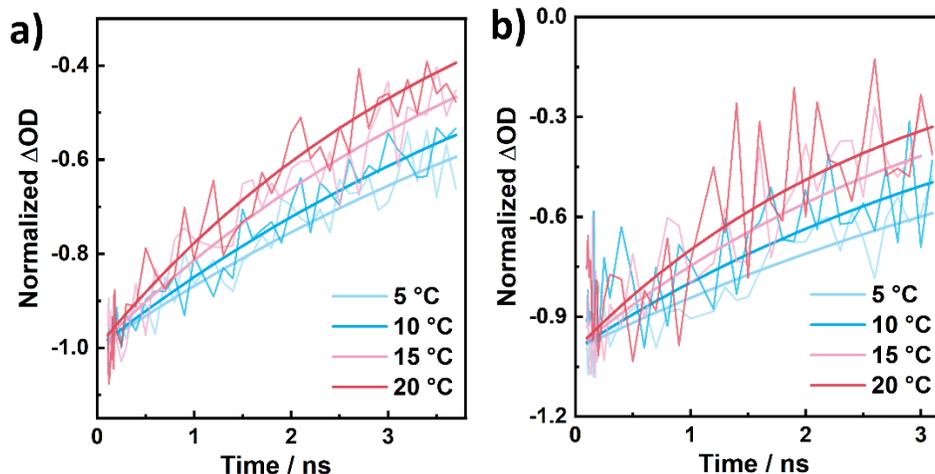


Fig. S10 Temperature dependence of the TA kinetics of the MseA:FAD complex in MSOX_{M245T} (a) and MTOX_{T239M} (b), probed at 573 and 560 nm, respectively, upon excitation at the CT bands. Only the kinetics on the ps–ns timescales are shown for clarity, and the smooth lines represent the best single exponential fits of the data.

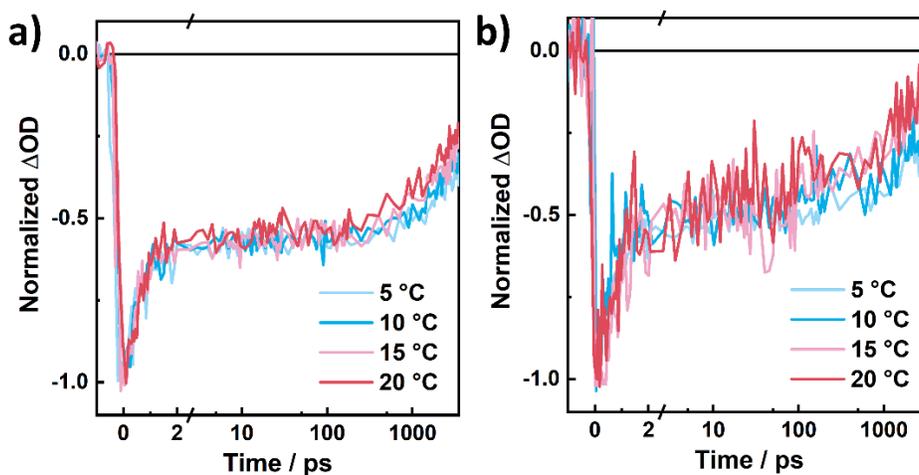


Fig. S11 Full TA kinetics of the the MseA:FAD complex in MSOX_{M245T} (a) and MTOX_{T239M} (b), probed at 573 and 560 nm, respectively, at different temperatures, upon excitation at the CT bands.

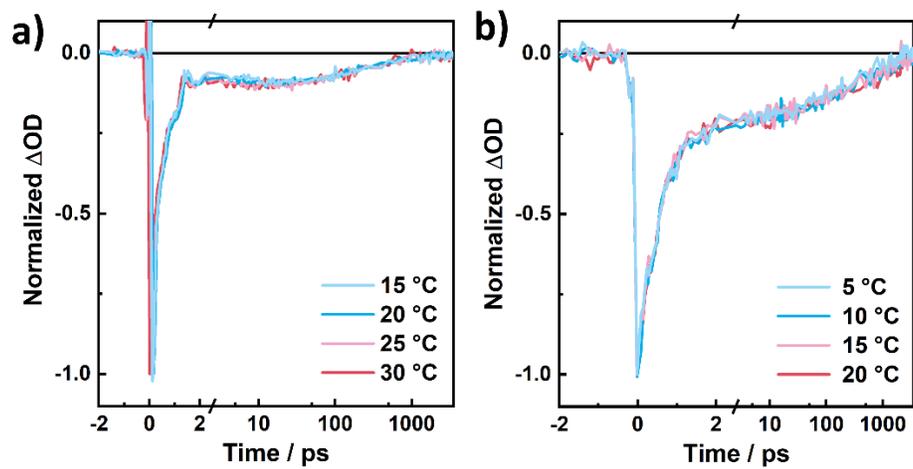


Fig. S12 Temperature dependence of the TA kinetics of the MSeA:FAD complex in MSOX_{M245K} (a) and MTOX_{T239K} (b), probed at 542 nm and 563 nm, respectively, upon excitation at the CT bands.

Table S2. Decay time constants (τ_1 , τ_2 , τ_3) obtained from the exponential fits of the TA kinetics at the selected wavelength (λ_{probe}) of the MXA:FAD complexes in MSOX and MTOX variants at room temperature; the normalized pre-exponential factor of each component is given in parentheses. Also listed are the estimated rate constant (k_3) and activation energy (E_a) of the back recovery process, as well as the estimated quantum yield of the forward switching process (QY).

Protein	Ligand	λ_{probe} (nm)	τ_1 (ps)	τ_2 (ps)	τ_3 (ps)	k_3 (ps ⁻¹)	E_a (kJ/mol)	QY ^[a]
MSOX _{WT}	MSA	536	0.5 (-0.29)	2200 (-0.71)		4.55×10^{-4}	21.9 ± 2.0	0.8
MSOX _{WT}	MSeA	590	0.6 (-0.49)	5500 (-0.51)		1.82×10^{-4}	22.2 ± 2.1	0.6
MTOX _{WT}	MSeA	570	0.8 (-0.77)	37 (-0.10)	690 (-0.13)			
MSOX _{M245T}	MSeA	573	0.7 (-0.50)	4000 (-0.50)		2.50×10^{-4}	33.6 ± 2.6	0.6
MSOX _{M245K}	MSeA	542	0.4 (-0.88)	388 (-0.11)	4100 (-0.01)			
MTOX _{T239M}	MSeA	560	0.9 (-0.54)	2800 (-0.46)		3.57×10^{-4}	26.8 ± 3.8	0.5
MTOX _{T239K}	MSeA	563	0.6 (-0.83)	124 (-0.11)	1600 (-0.06)			

^[a]Estimated according to the method described in Section 4 above. For MSOX_{M245T} and MTOX_{T239M}, as the extinction coefficients of *state B* (Fig. 4D of the main text) could not be accurately determined, and those of MSOX_{WT} were used as approximations.

6. Additional results of MD simulations and QM/MM potential energy scans

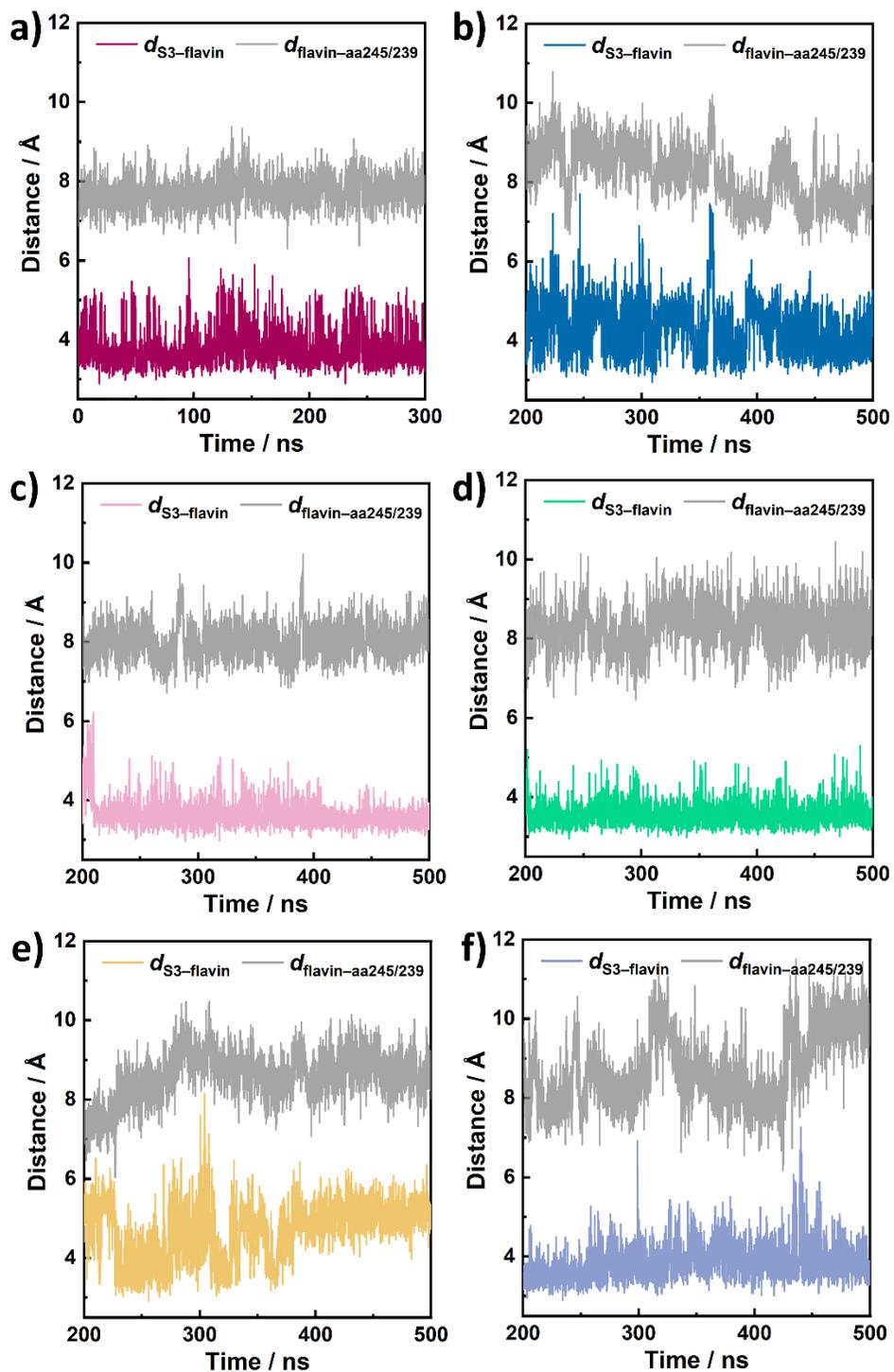


Fig. S13 Dynamics of $d_{S3-flavin}$ and $d_{flavin-aa245/239}$ (defined in the main text) in the MD simulations of MSOX_{WT} (a), MTOX_{WT} (b), MSOX_{M245T} (c), MTOX_{T239M} (d), MSOX_{M245K} (e), and MTOX_{T239K} (f).

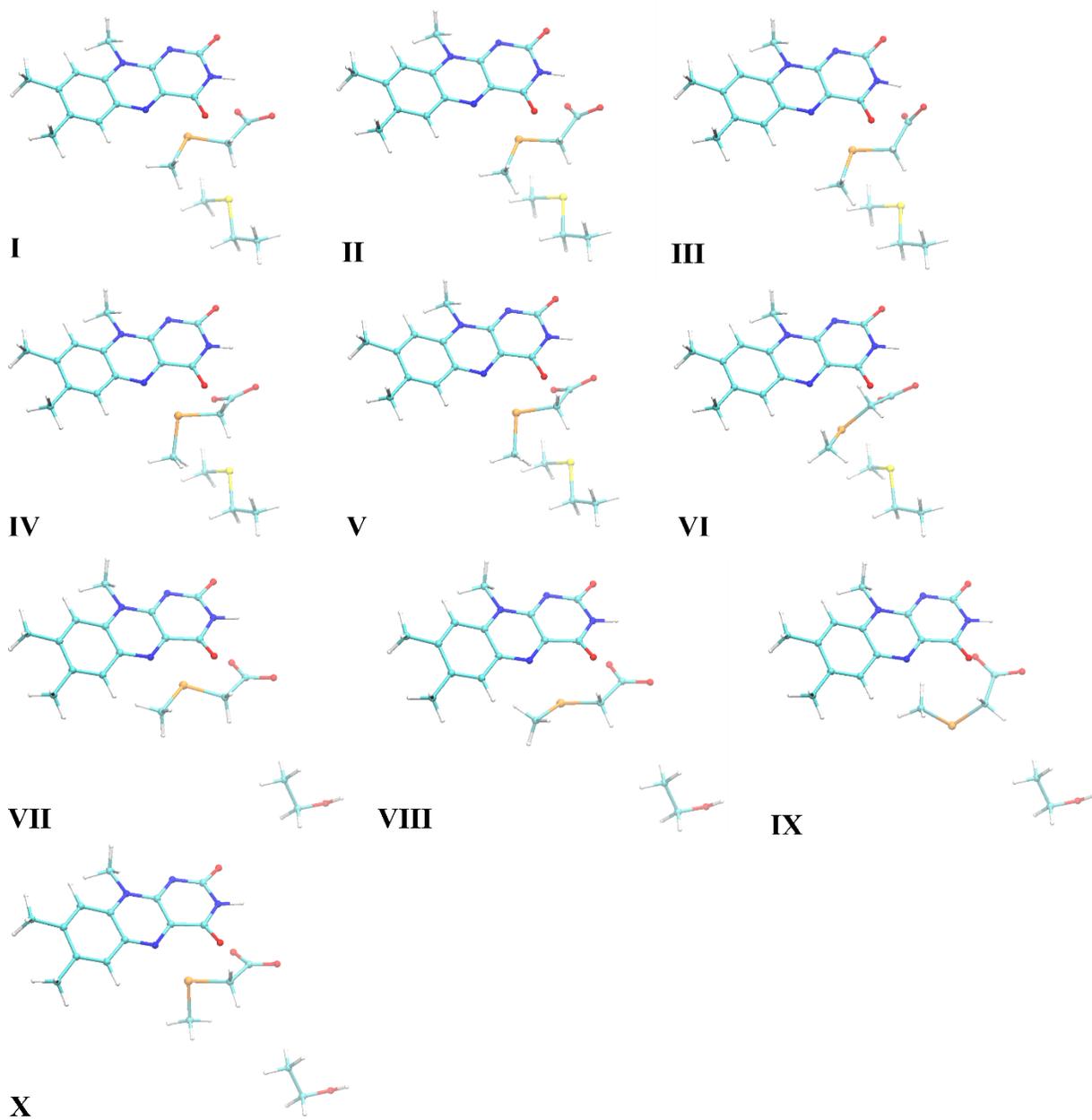


Fig. S14 Geometries (QM region) of the minima and transition states on the S_0 and S_1 potential energy surfaces obtained from the QM/MM potential energy scans in MSOX_{WT} and MTOX_{WT} (see Fig. 8 of the main text).

Supporting references

- 1 M. A. Arrica and T. Wirth, *European J. Org. Chem.*, 2005, **2005**, 395–403.
- 2 B. Zhuang and M. H. Vos, *J. Am. Chem. Soc.*, 2022, **144**, 11569–11573.
- 3 A. Dubietis, G. Tamošauskas, R. Šuminas, V. Jukna and A. Couairon, *Lith. J. Phys.*, 2017, **57**, 113–157.
- 4 Q. Liu, Y. Jiang, K. Jin, J. Qin, J. Xu, W. Li, J. Xiong, J. Liu, Z. Xiao, K. Sun, S. Yang, X. Zhang and L. Ding, *Sci. Bull.*, 2020, **65**, 272–275.
- 5 J. J. Snellenburg, S. P. Liptonok, R. Seger, K. M. Mullen and I. H. M. van Stokkum, *J. Stat. Softw.*
- 6 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé and K. Schulten, *J. Comput. Chem.*, 2005, **26**, 1781–1802.
- 7 J. Huang and A. D. Mackerell, *J. Comput. Chem.*, 2013, **34**, 2135–2145.
- 8 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.
- 9 A. Aleksandrov, *J. Comp. Chem.*, 2019, **40**, 2834–2842.
- 10 K. Vanommeslaeghe and A. D. MacKerell, *J. Chem. Inf. Model.*, 2012, **52**, 3144–3154.
- 11 K. Vanommeslaeghe, E. P. Raman and A. D. MacKerell, *J. Chem. Inf. Model.*, 2012, **52**, 3155–3168.
- 12 S. Jo, T. Kim, V. G. Iyer and W. Im, *J. Comput. Chem.*, 2008, **29**, 1859–1865.
- 13 M. H. M. Olsson, C. R. SØndergaard, M. Rostkowski and J. H. Jensen, *J. Chem. Theory Comput.*, 2011, **7**, 525–537.
- 14 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089–10092.
- 15 H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684.
- 16 F. Neese, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2012, **2**, 73–78.
- 17 T. Yanai, D. P. Tew and N. C. Handy, *Chem. Phys. Lett.*, 2004, **393**, 51–57.
- 18 J. Zheng, X. Xu and D. G. Truhlar, *Theor. Chem. Acc.*, 2011, **128**, 295–305.
- 19 T. Lu and F. Chen, *J. Theor. Comput. Chem.*, 2012, **11**, 163–183.
- 20 T. Lu and Q. Chen, *J. Comput. Chem.*, 2022, **43**, 539–555.
- 21 Z. Liu, T. Lu and Q. Chen, *Carbon N. Y.*, 2020, **165**, 461–467.
- 22 T. Lu and F. Chen, *J. Comput. Chem.*, 2012, **33**, 580–592.
- 23 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph.*, 1996, **14**, 33–38.

- 24 M. A. Wagner, P. Trickey, Z. W. Che, F. S. Mathews and M. S. Jorns, *Biochemistry*, 2000, **39**, 8813–8824.
- 25 P. Müller and K. Brettel, *Photochem. Photobiol. Sci.*, 2012, **11**, 632–636.
- 26 F. Lacombat, P. Plaza, M. A. Plamont and A. Espagne, *J. Phys. Chem. Lett.*, 2017, **8**, 1489–1495.
- 27 B. Zhuang, D. Seo, A. Aleksandrov and M. H. Vos, *J. Am. Chem. Soc.*, 2021, **143**, 2757–2768.
- 28 P. Müller and K. Brettel, *Photochem. Photobiol. Sci.*, 2012, **11**, 632–636.