Supporting Information for

Self-oxidation of cysteine to sulfinic acid in an engineered T67C myoglobin: structure and reactivity

Wei Dai [a†], Hong Yuan [b†], Xiao-Juan Wang [a], Shu-Qin Gao [c], Xiangshi Tan [b], and Ying-Wu Lin[a,c*]

[a] *School of Chemistry and Chemical Engineering, University of South China, Hengyang 421001, China;*

[b]*Department of Chemistry & Institute of Biomedical Science, Fudan University, Shanghai 200433, China;*

[c] *Key Lab of Protein Structure and Function of Universities in Hunan Province, University of South China, Hengyang 421001, China;*

[†] These authors contributed equally.

Corresponding author:
*E-mail: ywlin@usc.edu.cn

1 Experimental Section

1.1 Protein preparation

Wild-type (WT) sperm whale Mb was expressed in BL21 (DE3) cells using the Mb gene of pMbt7-7, as reported by Springer and Sligar.[1] The gene of T67C Mb was constructed by using the QuikChange Site-Directed Mutagenesis Kit (Stratagene) using the WT Mb gene as a template, and the mutation was confirmed by DNA sequencing assay. The T67C Mb was expressed and purified using the same procedure reported previously,[2] which yielded T67C Mb (Cys-SH&SO_2H). The T67C Mb (Cys-SO_2H) form was obtained by restoring the purified T67C Mb (Cys-SH&SO_2H) at room temperature for 2~3 days. To obtain the T67C Mb (Cys-SH) form, the protein solution after centrifugation from the lysate was quickly subjected to an anion exchange chromatography, followed by concentration in an ultrafiltration cup under an N_2 atmosphere. The purified protein was used directly or stored at -80 ℃. T67C-FL Mb was prepared by incubation of T67C Mb (Cys-SH) (0.1 mM) with 3 eq. fluorescein-5-maleimide (FL) dissolved in 100 mM potassium phosphate buffer, pH 7.0 for 2 h. The obtained solution was loaded onto a Sephadex G75 column and equilibrated with 100 mM potassium phosphate buffer (100 mM, pH 7.0) to remove the excess fluorescein-5-maleimide.

1.2 Mass spectrometry

Protein mass spectrometry was performed using a G2-XS QToF mass spectrometer (Waters). The desalted protein solution was mixed with 1 % formic acid and injected into the mass spectrometry chamber for measurement in positive mode. Multiple m/z peaks were converted into the relative molecular mass of the protein using MaxEnt1 software. For Cys oxidation experiment, 50 μL T67C Mb (Cys-SH&SO_2H) (~0.1 mM) was mixed with the 1 eq H_2O_2. The mixture was cultured at room temperature for 1 h and then passed a Sephadex G25 column (PD 10, GE Healthcare) before the mass measurement. The mass spectra of the protein after reaction with diisopropyl azodicarboxylate (DIAD), iodoacetamide (IAM) or

fluorescein-5-maleimide (FL) were determined using the same procedure.

1.3 UV-Vis spectroscopy

UV-Vis spectra were recorded on a Hewlett-Packard 8453 diode array spectrometer. Protein concentration was determined with an extinction coefficient of $\varepsilon_{409\ nm}$ = 157 mM^{-1} cm^{-1} for T67C Mb and T67C-FL Mb, $\varepsilon_{428\ nm}$ = 142 mM^{-1} cm^{-1} for ZnPP-T67C-FL Mb as reported previously for WT Mb and ZnPP-WT Mb.[3] The concentration of fluorescein-5-maleimide was determined with the extinction coefficient of $\varepsilon_{495\ nm}$ = 68 mM^{-1} cm^{-1}.[4]

1.4 X-ray crystallography

T67C Mb (Cys-SO_2H) with high purity (A_{409nm}/A_{280nm}>4.0) was exchanged into 20 mM potassium phosphate buffer (pH 7.0) and concentrated to ~2.0 mM. The vapor diffusion hanging drop technique was used to crystallize the protein. Typically, 250 μL Crystal Screen reagent (Hampton Research), No. 28 (0.2 M Sodium acetate trihydrate, 0.1 M Sodium cacodylate trihydrate pH 6.5, and 30% w/v PEG 8000) was added to each cell in the 24-well plate. The protein solution (2 μL) was mixed with an equal volume of the reservoir solution and placed on a siliconized coverslip. The coverslip was inverted over the plate with vaseline and allowed to stand for three days to one week in an incubator at 16 ℃.

Diffraction data were collected from a single crystal at Shanghai Synchrotron Radiation Facility (SSRF) BL17U beamline, China, using a MAR mosaic 225 CCD detector with a wavelength of 0.9793 Å at 100 K. The diffraction data were processed and scaled with HKL-2000.[5] The structure was solved by the molecular replacement method and the 1.6 Å structure of WT Mb (PDB entry 1JP[6]) was used as the starting model. Manual adjustment of the model was carried out using the program COOT[7] and the models were refined by PHENIX and Refmac. Stereochemical quality of the structures was checked by using PROCHECK. All of the residues located in the favored and allowed region and none in the disallowed region.

1.5 CD spectrometry

Circular dichroism (CD) spectroscopy was performed at Jasco 1500 spectropolarimeter, equipped with MCB-100 mini-circulating. The far-UV CD spectra were recorded in the range of 190-250 nm for T67C Mb (Cys-SO_2H) and WT Mb (2 μM in 1 mM potassium phosphate buffer, pH 7.0) in 1.0 cm path quartz cuvette.

1.6 Reconstitution of Apo-Mb with ZnPP

ApoT67C-FL Mb was obtained by the conventional method.[8] To reconstitute ZnPP into apoT67C-FL Mb, approximately 5 mg of ZnPP was dissolved in 1 mL of dimethyl formamide and added dropwise to the apoT67C-FL Mb solution with gentle stirring. The obtained solution was loaded onto a Sephadex G75 column, and equilibrated with 100 mM potassium phosphate buffer (100 mM, pH 7.0), to remove the free ZnPP.

1.7 Fluorescence spectroscopy

The fluorescence spectrum was performed on the F-7000 Fluorescence Spectrometer. The voltage of the photomultiplier was set at 350 V. The width of the inner and outer slits was fixed at 10 nm. The emission spectra of fluorescein-5-maleimide, T67C Mb (Cys-SH), T67C-FL Mb, apoT67C-FL Mb, and ZnPP-T67C Mb in the range of 500-700 nm were measured at 494 nm excitation wavelength. All sample concentrations were 5 μM in 100 mM potassium phosphate buffer, pH 7.0.

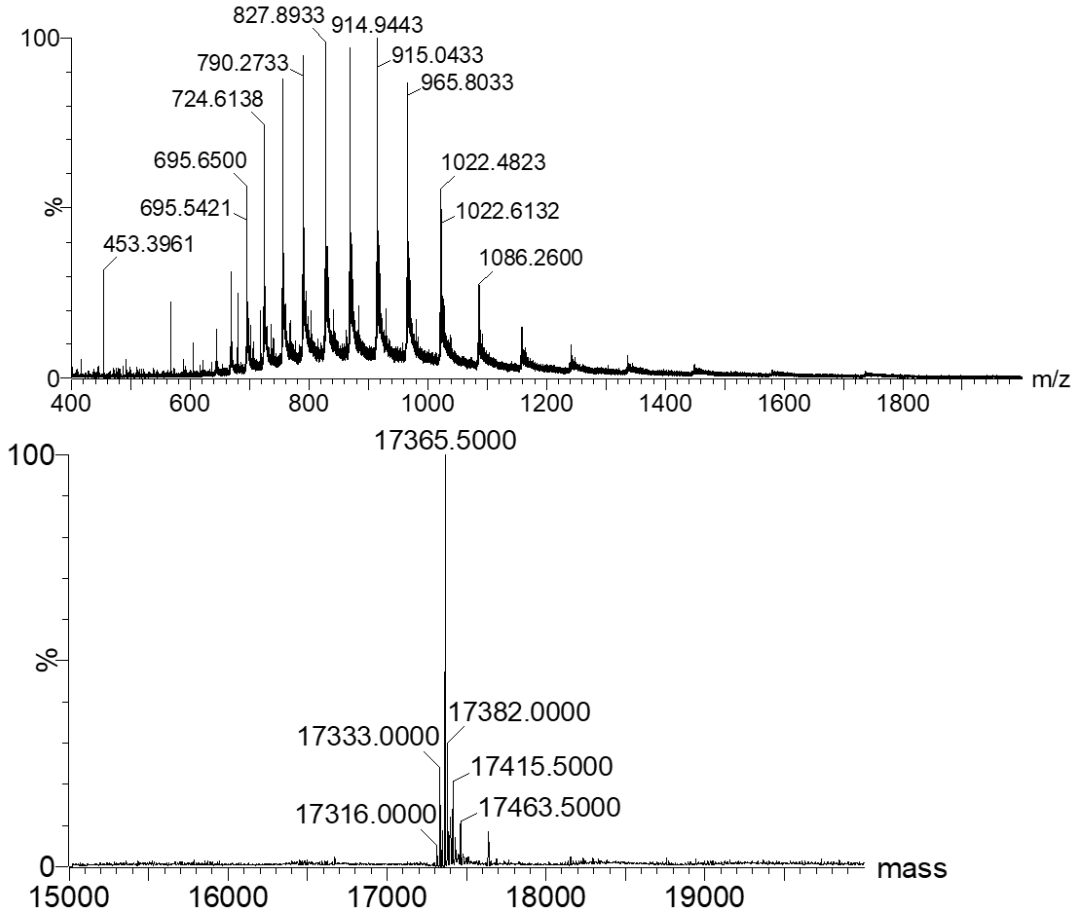

Figure S1. ESI-MS spectra of T67C Mb (Cys-SO_2H).

Table S1. Data collection and refinement statistics for T67C Mb.

PDB ID 7XCF	**T67C Myoglobin**
Wavelength	0.9791
Resolution range	24.35 - 1.7 (1.761 - 1.7)
Space group	P 21 21 21
Unit cell	39.703 48.697 78.755 90 90 90
Total reflections	186651 (20025)
Unique reflections	16888 (1689)
Multiplicity	11.1 (11.8)
Completeness (%)	97.01 (99.94)
Mean I/sigma(I)	21.73 (8.39)
Wilson B-factor	20.33
R-merge	0.07379 (0.2261)
R-meas	0.07757 (0.2363)
R-pim	0.02331 (0.06762)
CC1/2	0.997 (0.99)
CC*	0.999 (0.997)
Reflections used in refinement	16884 (1689)
Reflections used for R-free	847 (93)
R-work	0.1717 (0.1909)
R-free	0.1984 (0.2218)
CC(work)	0.940 (0.731)
CC(free)	0.953 (0.660)
Number of non-hydrogen atoms	1454
macromolecules	1229
ligands	43
solvent	182
Protein residues	153
RMS(bonds)	0.010
RMS(angles)	1.20
Ramachandran favored (%)	97.30
Ramachandran allowed (%)	2.70
Ramachandran outliers (%)	0.00
Rotamer outliers (%)	0.00
Clashscore	4.31
Average B-factor	22.57
macromolecules	21.29
ligands	17.34
solvent	32.42

Statistics for the highest-resolution shell are shown in parentheses.

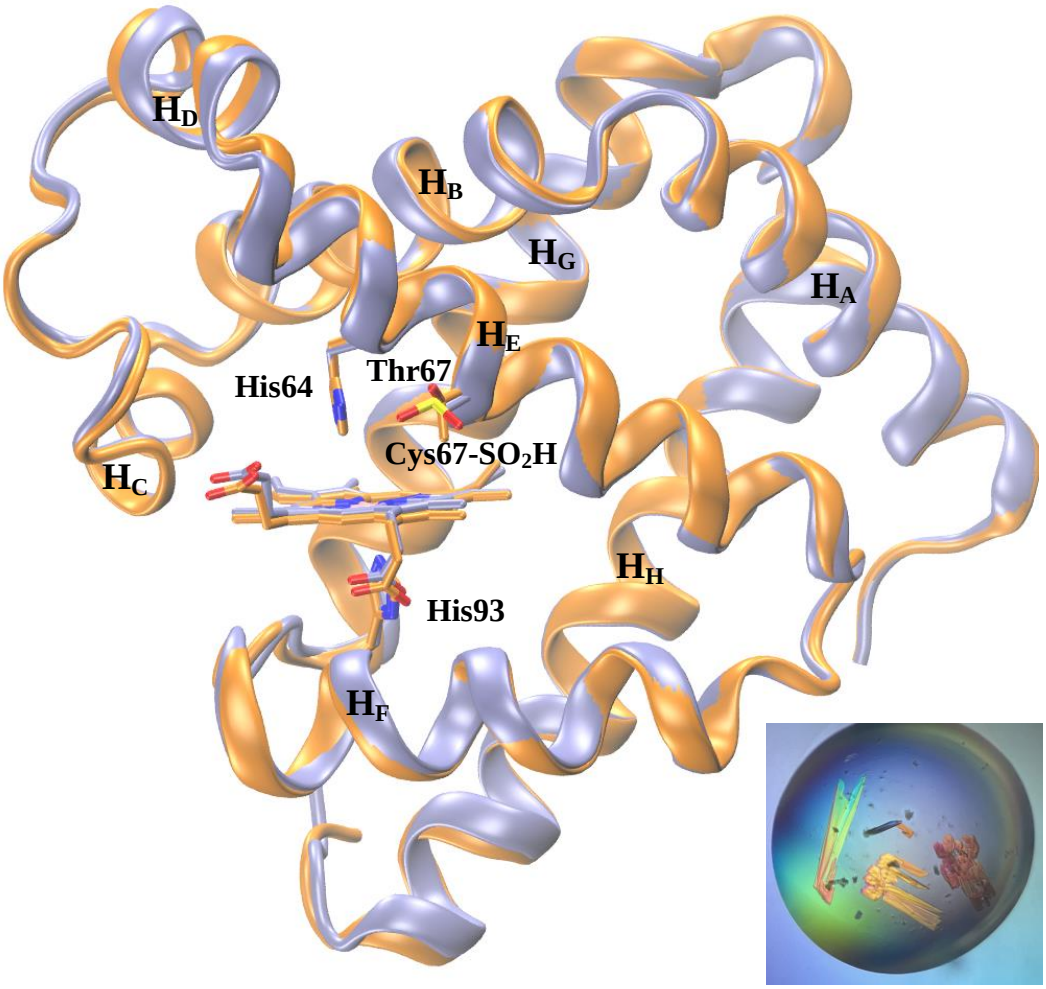

Figure S2. Overlay of the X-ray crystal structure of T67C Mb (blue) with that of WT Mb (orange). The heme group, Thr67 and the sulfinic acid Cys67-SO_2H, His64 and His93, and the secondary structure of helices A-H are labeled. The digital microscope image of the protein crystals is shown at the right bottom.

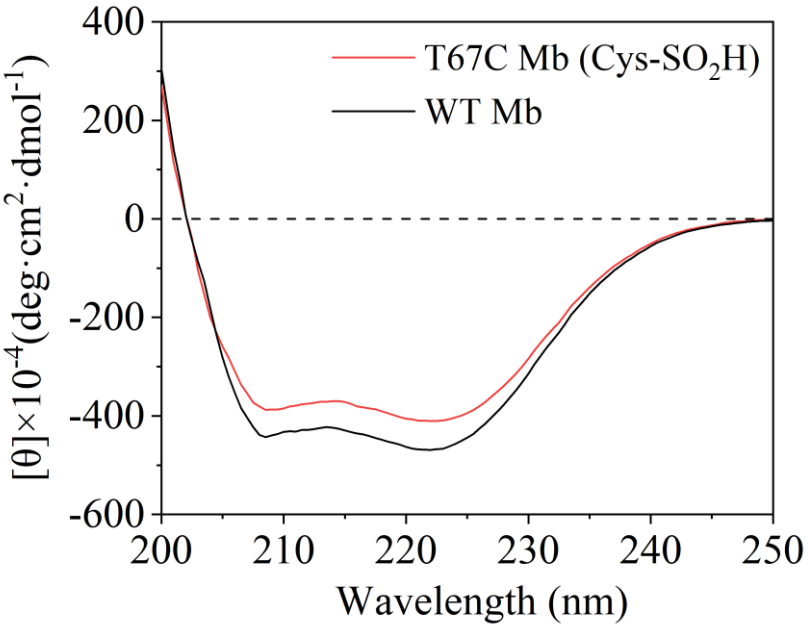

Figure S3. CD spectra of 2 μM T67C Mb (Cys-SO_2H) and WT Mb in potassium phosphate buffer (1 mM, pH 7.0). (T67C Mb (Cys-SO_2H): α-helix, 72.57% ; β-strand, 1.15%; WT Mb: α-helix, 80.76%; β-strand, 0.53%, as estimated by the K2D3 server.[9]

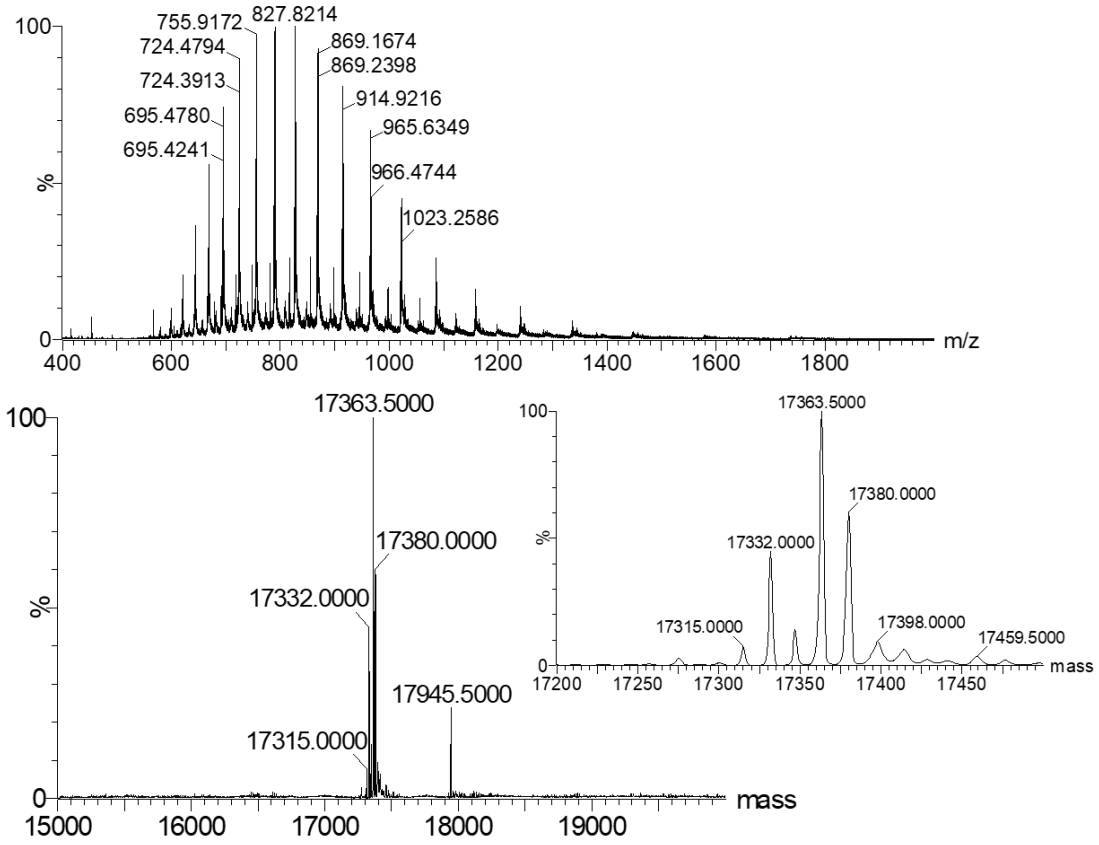

Figure S4. ESI-MS spectra of T67C Mb (Cys-SH&SO_2H) (0.1 mM) after incubating with 1 eq H_2O_2 for 1 hour in 100 mM potassium phosphate buffer, pH 7.0, at 25 ℃.

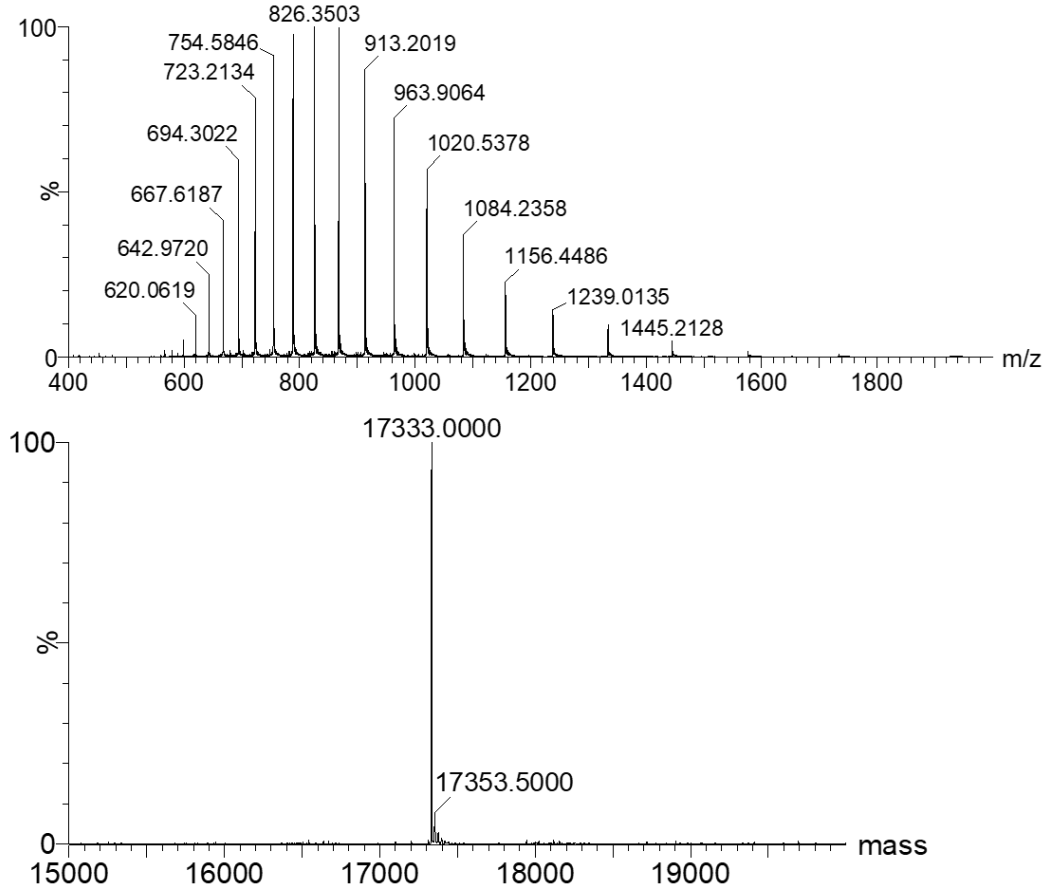

Figure S5. ESI-MS spectra of T67C Mb (Cys-SH).

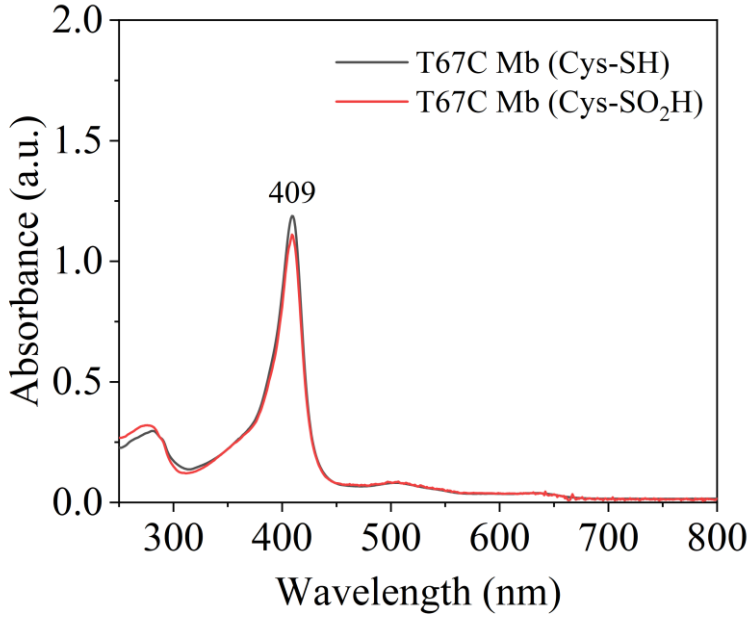

Figure S6. UV-Vis spectra of T67C Mb (Cys-SH) and T67C Mb (Cys-SO_2H) in 100 mM potassium phosphate buffer (pH 7.0).

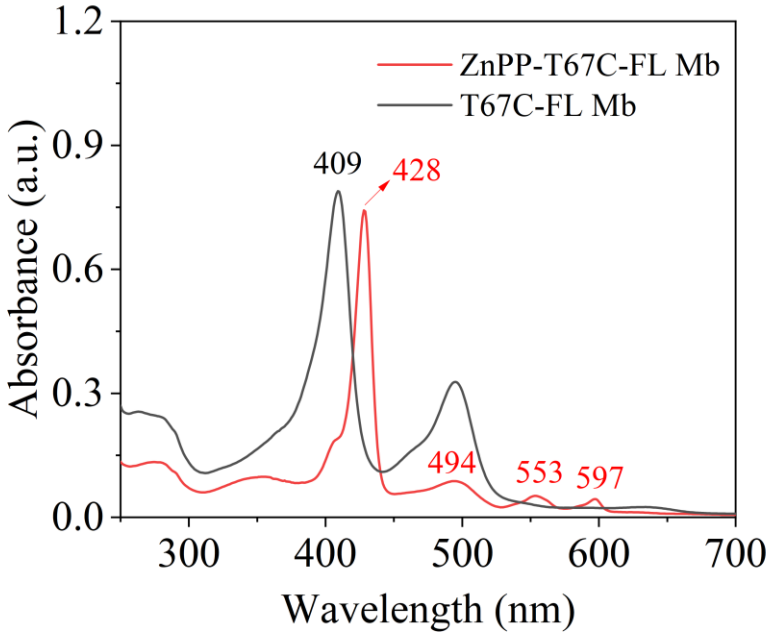

Figure S7. UV-Vis spectra of T67C-FL Mb and ZnPP-T67C-FL Mb (5 μM) in 100 mM potassium phosphate buffer (pH 7.0).

References:

1. B. A. Springer and S. G. Sligar, *Proc. Nat. Acad. Sci. U. S. A.*, 1987, **84**, 8961-8965.
2. J. A. Sigman, B. C. Kwok and Y. Lu, *J. Am. Chem. Soc.*, 2000, **122**, 8192-8196.
3. Z.-H. Shi, K.-J. Du, B. He, S.-Q. Gao, G.-B. Wen and Y.-W. Lin, *Inorg. Chem. Front.*, 2017, **4**, 2033-2036.
4. J. S. Nanda and J. R. Lorsch, *Methods Enzymol*, 2014, **536**, 79-86.
5. Z. Otwinowski and W. Minor, in *Methods in Enzymology*, Academic Press, 1997, vol. 276, pp. 307-326.
6. P. Urayama, G. N. Phillips and S. M. Gruner, *Structure*, 2002, **10**, 51-60.
7. P. Emsley and K. Cowtan, *Acta Crystal. D*, 2004, **60**, 2126-2132.
8. Y. Dong, Y. M. Chen, X. J. Kong, S. Q. Gao, J. J. Lang, K. J. Du and Y. W. Lin, *J. Inorg. Biochem.*, 2022, **235**, 111943.
9. C. Perez-Iratxeta and M. A. Andrade-Navarro, *BMC Struct. Biol.*, 2008, **8**, 25.