Electronic Supplementary Material (ESI) for ChemComm.


# Supporting Information for

# Structural and functional regulations by a disulfide bond designed in myoglobin like human neuroglobin

Li-Juan Sun[a†], Hong Yuan[b†], Lu Yu [c], Shu-Qin Gao[a,d], Ge-Bo Wen[a,d], Xiangshi Tan[b], and Ying-Wu Lin[a,d*]

[a] *Hengyang Medical School, University of South China, Hengyang 421001, China;*

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

[c] *High Magnetic Field Laboratory, Chinese Academy of Sciences, Hefei, Anhui 230031, China.*

[d] *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

# Contents

## 1 Experimental Section

### *1.1 Protein preparation*

Wild-type (WT) sperm whale myoglobin (Mb) and its mutants were expressed in BL21 (DE3) cells using the Mb gene of pMbt7-7 and as reported by Springer and Sligar.[1] The genes of F46C Mb and F46C/L61C Mb were constructed by using the QuikChange Site-Directed Mutagenesis Kit (Stratagene) using WT Mb gene as a template, and the mutations were confirmed by DNA sequencing assay.[2] The proteins were expressed and purified using the same procedure reported previously.[3]

### *1.2 Mass spectrometry*

The F46C Mb or F46C/L61C Mb sample was diluted with 0.1 M acetic acid (pH 3.0) to ~20 μM. The protein solution was mixed with 1% formic acid, which was transferred into the mass spectrometer chamber for measurement under positive mode. The data were analyzed using G2-XS QTOF mass spectrometry (Waters) and the multiple m/z peaks were transformed to the protein molecular weight by using the software MaxEnt1. For the mass measurement of the reduced F46C/L61C Mb, the reducing agent tris-(2-carboxyethyl)-phosphine (TCEP) or dithionite was added to the protein solution at a final concentration of 10 mM. The mixture was cultured at room temperature for 30 min and then through a Sephadex G25 column (PD 10, GE Healthcare) before mass measurements.

### *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 $\varepsilon_{408\ nm}$ = 151 ± 5 $mM^{-1}$ $cm^{-1}$ for F46C Mb, $\varepsilon_{413\ nm}$ = 144 ± 5 $mM^{-1}$ $cm^{-1}$ for F46C/L61C Mb, as calculated using the hemochromogen method.[4] Gdn·HCl-induced unfolding studies of F46C Mb and F46C/L61C Mb were performed by using the procedure as described previously.[5]

The autooxidation rates of WT Mb, F46C Mb, and F46C/L61C Mb were performed by monitoring the UV-vis spectra changes for 9, 2 and 1 h at 25 $^{o}$C, respectively. The ferric protein was reduced by excess dithionite, which was removed from the solution by a PD-10 column (GE Healthcare). The reduced protein was then diluted into $O_2$-saturated 100 mM potassium phosphate (pH 7.0) to generate the oxy-form protein. The single wavelength changes in intensity at ~580 nm of the oxy-form were fitted to a single-exponential decay function.[6]

### *1.4 X-ray crystallography*

Ferric F46C Mb and F46C/L61C Mb with high purity ($A_{\text{soret band}}/A_{280\text{nm}}$>4.0) were exchanged into 20 mM potassium phosphate buffer (pH 7.0) and concentrated to ~4.0 mM. The vapor diffusion hanging drop technique was used to crystallize the protein. Diffraction data were collected from a single crystal at Shanghai Synchrotron Radiation Facility (SSRF) BL02U1 beamline, China, using a MAR mosaic 225 CCD detector with a wavelength of 0.9791 Å at 100 K. The diffraction data were processed and scaled with HKL-2000.[7] The structure was solved by the molecular replacement method and the 1.6 Å structure of WT Mb (PDB entry 1JP6[8]) was used as the starting model. Manual adjustment of the model was carried out using the program COOT[9] and the models were refined by PHENIX[10] and Refmac5[11]. Stereochemical quality of the structures was checked by using PROCHECK[12]. All of residues locate in the favored and allowed region and none in the disallowed region.

### *1.5 EPR studies*

Electron paramagnetic resonance (EPR) spectra of ferric WT Mb, F46C Mb and F46C/L61C Mb (0.3 mM) were collected at the high magnetic field laboratory of Chinese Academy of Science, Hefei, China. The sample was analyzed by X-band EPR on a Bruker EMX plus 10/12 spectrometer. A standard Bruker cavity (ER4119hs TE011) was used in conjunction with an Oxford Instrument EPR910 liquid helium continuous-flow cryostat for low-temperature analysis. The spectra were measured at

a low temperature of 10 K, with a frequency of 9.395 GHz, center field 2750 G and sweep width 4500 G, microwave power 1 mW, gain 1000, and modulation amplitude 5.0 G.

### *1.6 Stopped-flow spectroscopy*

The reactions of F46C Mb and F46C/L61C Mb with hydrogen peroxide ($H_2O_2$) were determined using a dual mixing stopped-flow spectrophotometer (SF-61DX2 Hi-Tech KinetAsyst$^{TM}$). Typically, one syringe contains 10 μM of protein (in 100 mM potassium phosphate buffer, pH 7.0), and the second syringe contains $H_2O_2$ with concentrations ranging from 0.25 to 2 mM. The reaction was started by mixing an equal volume of solutions from both syringes. 100 time-dependent spectra were collected over 5 and 20 sec for F46C Mb and F46C/L61C Mb, respectively, from 350 to 700 nm at 25 $^{o}$C. The apparent rate constant, $k_1$, was obtained by linear regression fitting the plot of the observed rate constants, $k_{obs}$, versus the concentration of $H_2O_2$.

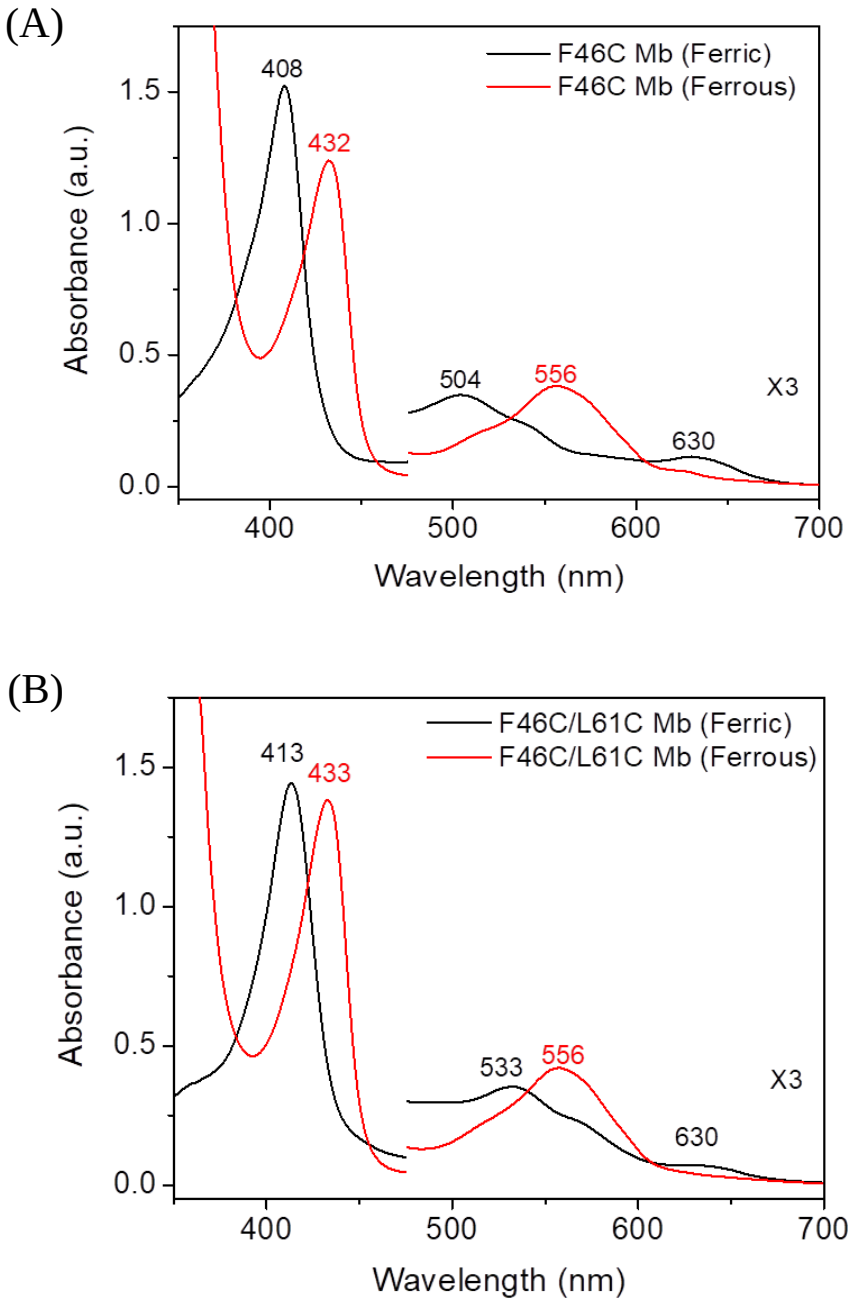


**Figure S1**. UV-vis spectra of (A) F46C Mb, (B) F46C/L61C Mb in ferric (black) and ferrous (red) states.

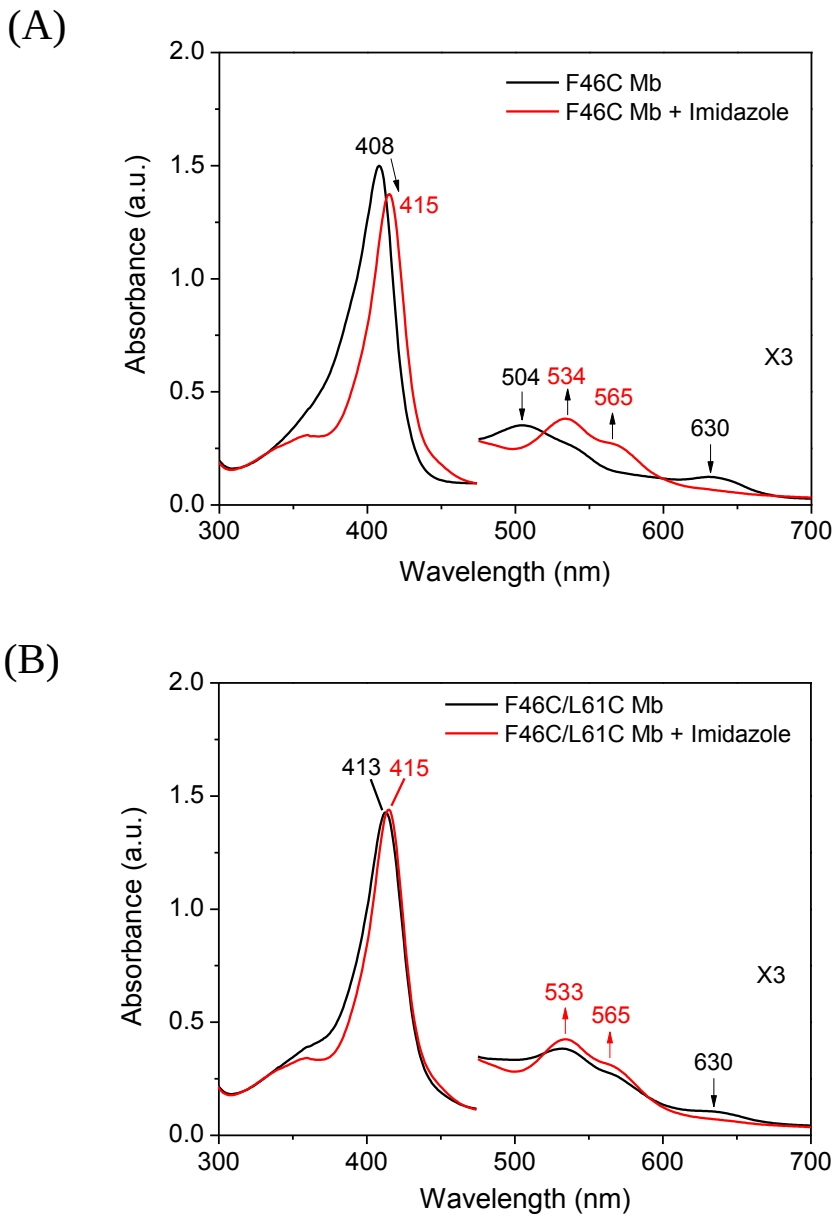


**Figure S2.** UV-vis spectra of 10 μM (A) F46C Mb and (B) F46C/L61C Mb in the absence and presence of 5 mM imidazole in 100 mM potassium phosphate buffer (pH 7.0), 25 $^{o}$C.

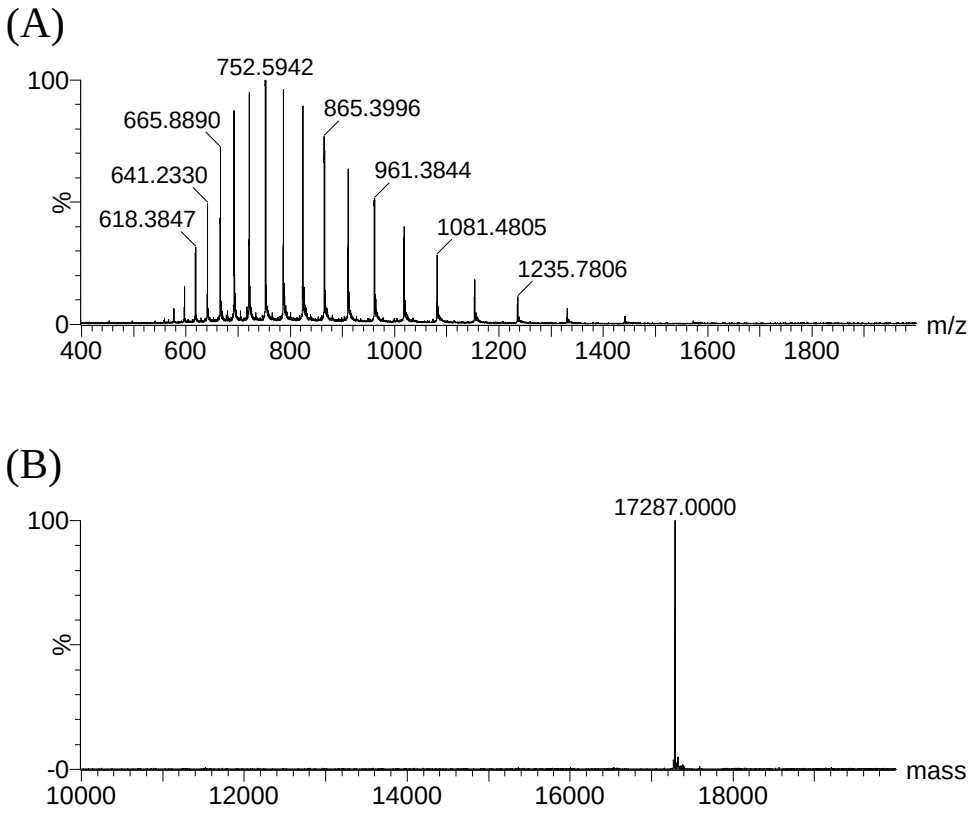


**Figure S3**. ESI-MS spectrum of F46C Mb: (A) Original multiply-charged series. (B) The MaxEnt survey spectrum showing the major component. Calculated molecular weight of apo-protein, 17287 Da; Observed, 17287.0 ± 0.5 Da.

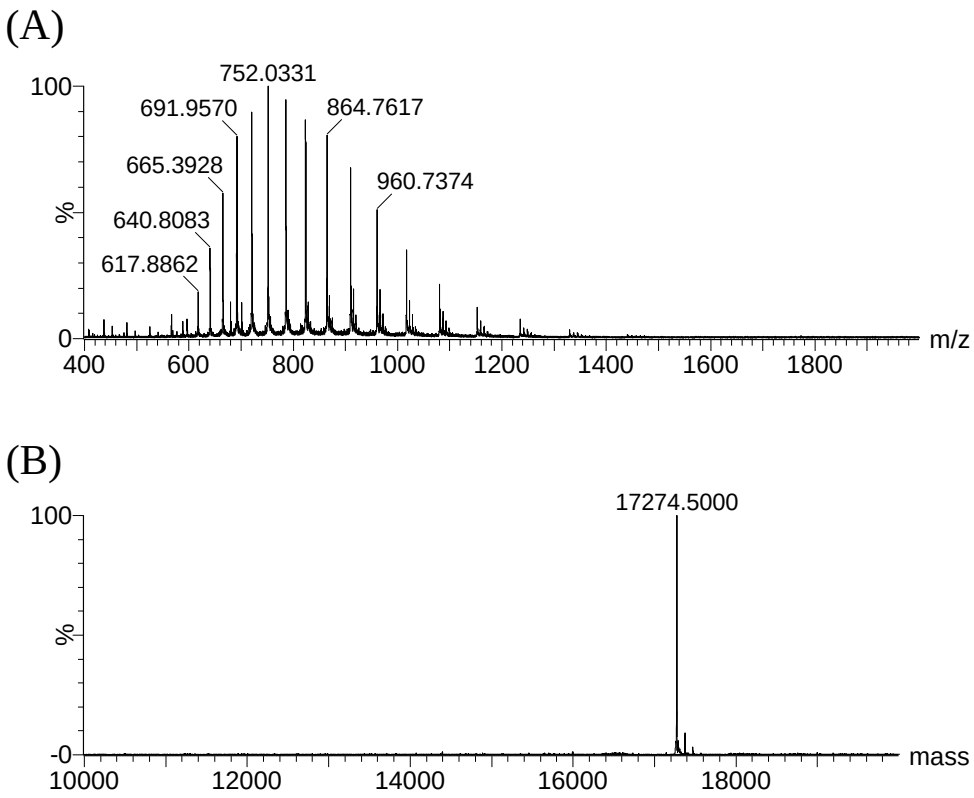


**Figure S4**. ESI-MS spectrum of F46C/L61C Mb: (A) Original multiply-charged series. (B) The MaxEnt survey spectrum showing the major component. Calculated molecular weight of apo-protein with an intramolecular disulfide bond, 17275 Da; Observed, 17274.5 ± 0.5 Da.

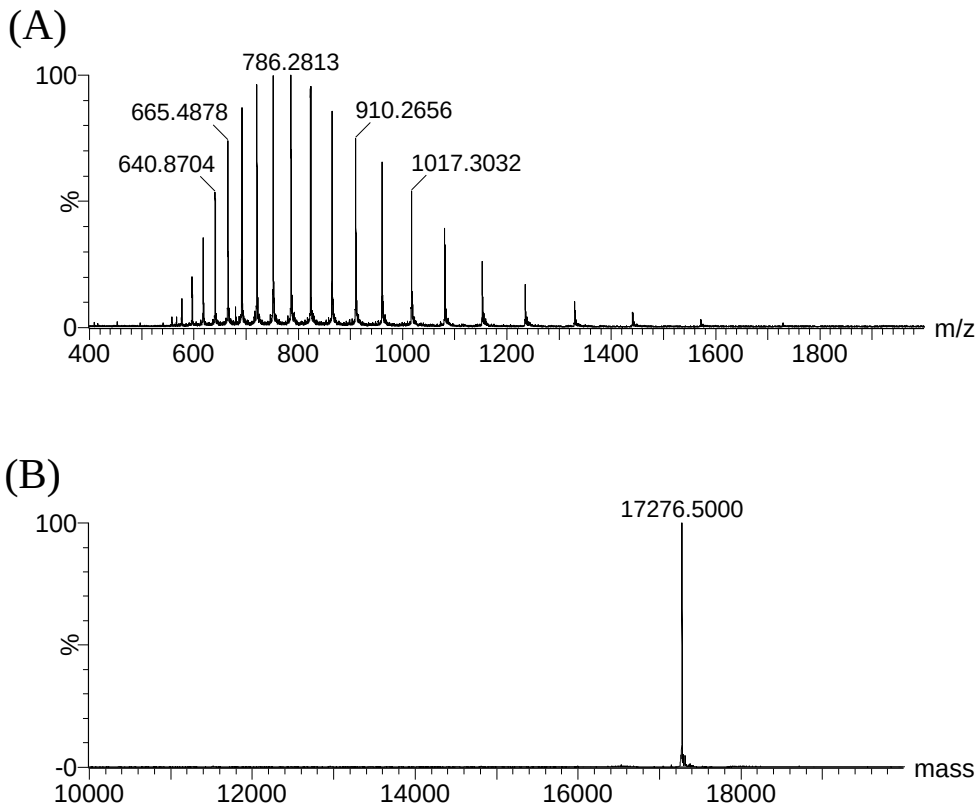


**Figure S5**. ESI-MS spectrum of F46C/L61C Mb after treated with TCEP: (A) Original multiply-charged series. (B) The MaxEnt survey spectrum showing the major component. Calculated molecular weight of apo-protein without an intramolecular disulfide bond, 17277 Da; Observed, 17276.5 ± 0.5 Da.

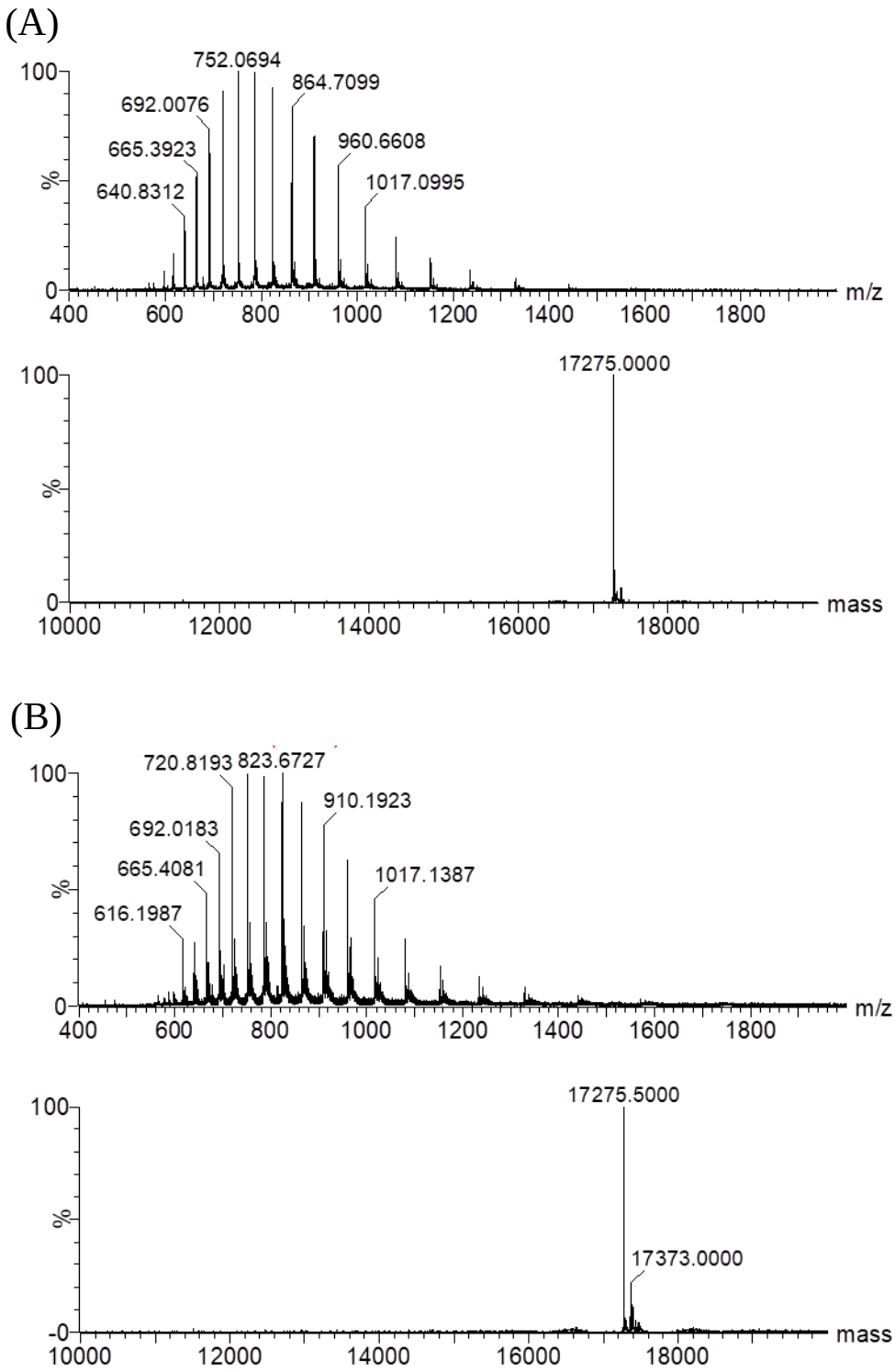


**Figure S6**. ESI-MS spectra of F46C/L61C Mb before (A) and after (B) reduction with dithionite. Calculated molecular weight of the apo-protein with an intramolecular disulfide bond, 17275 Da; Observed, 17275.5 ± 0.5 Da. The peak of 17373.0 Da could be due to an association of the apo-protein with an $SO_4^{2-}$ ion (96 Da).

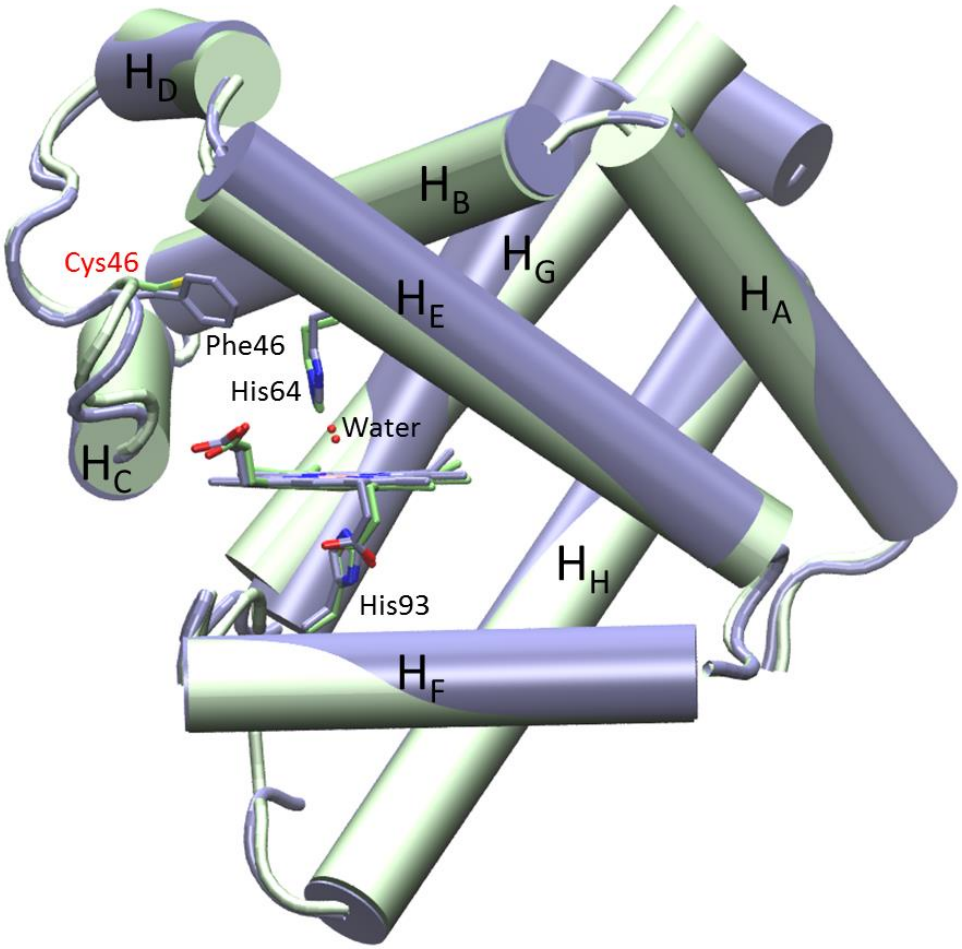


**Figure S7.** Overlay of the X-ray crystal structure of F46C Mb (green) with that of WT Mb (blue). The heme active site including the axial water molecules, and residues Cys46, Phe46, His64 and His93, are highlighted, and helices A-E are shown in Cartoon for clarification.

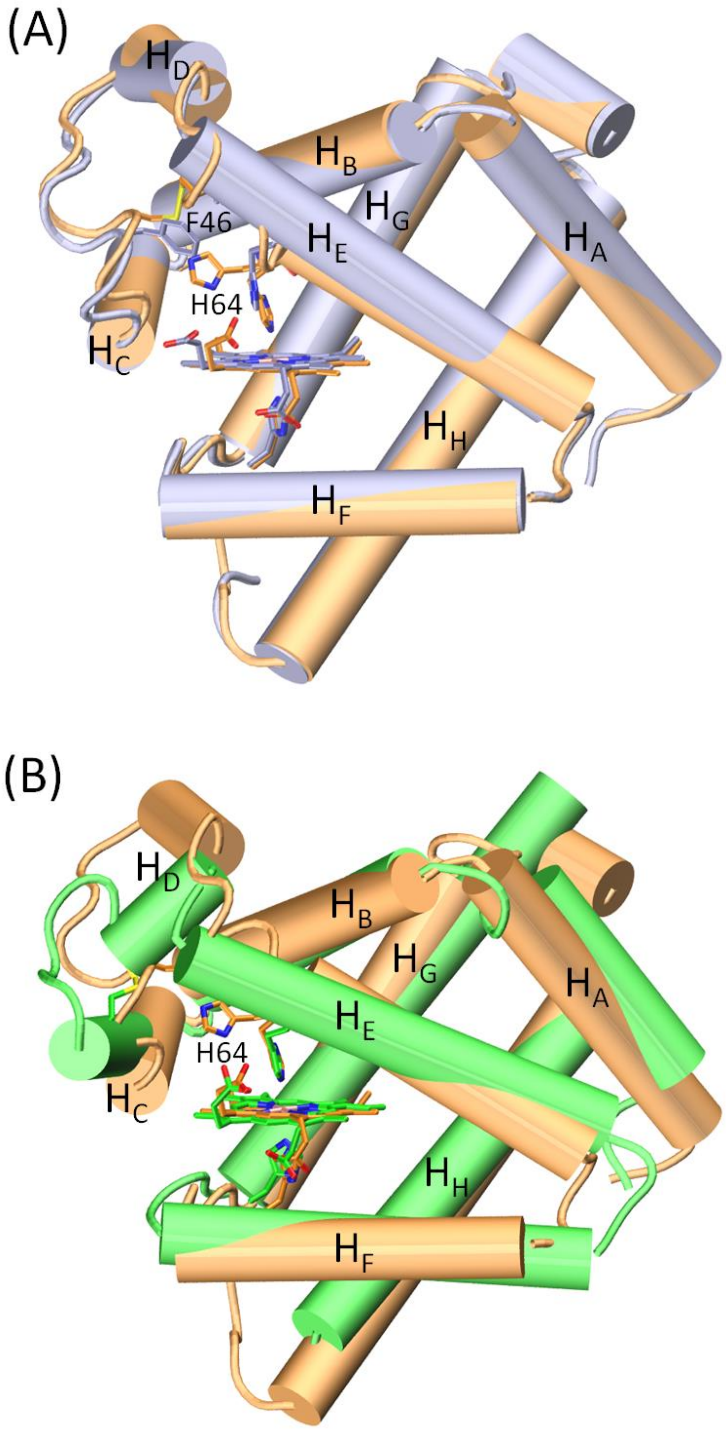


**Figure S8.** Overlay of the X-ray crystal structure of F46C/L61C Mb (orange) with that of WT Mb (A, blue) and human Ngb (B, green), with helices A-E shown in Cartoon for clarification.

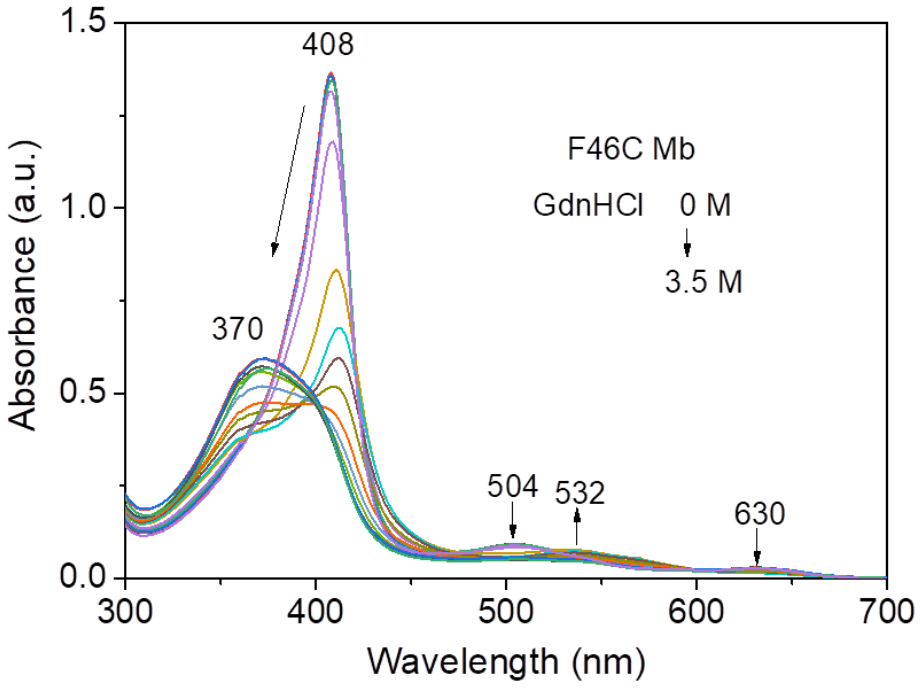


**Figure S9.** UV-vis spectra of Gdn·HCl-induced unfolding of F46C Mb.

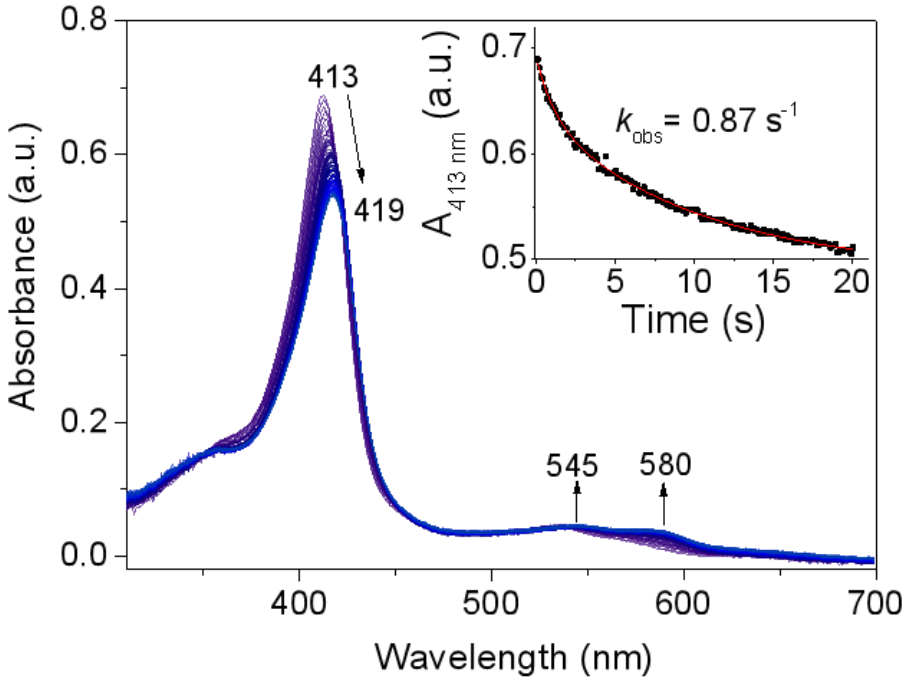


**Figure S10.** Stopped-flow spectra of 10 μM F46C/L61C Mb mixing with 1 mM $H_2O_2$ in 100 mM potassium phosphate buffer (pH 7.0), 25 $^{o}$C, for 20 sec. The absorbance of 413 nm versus time was shown as an inset.

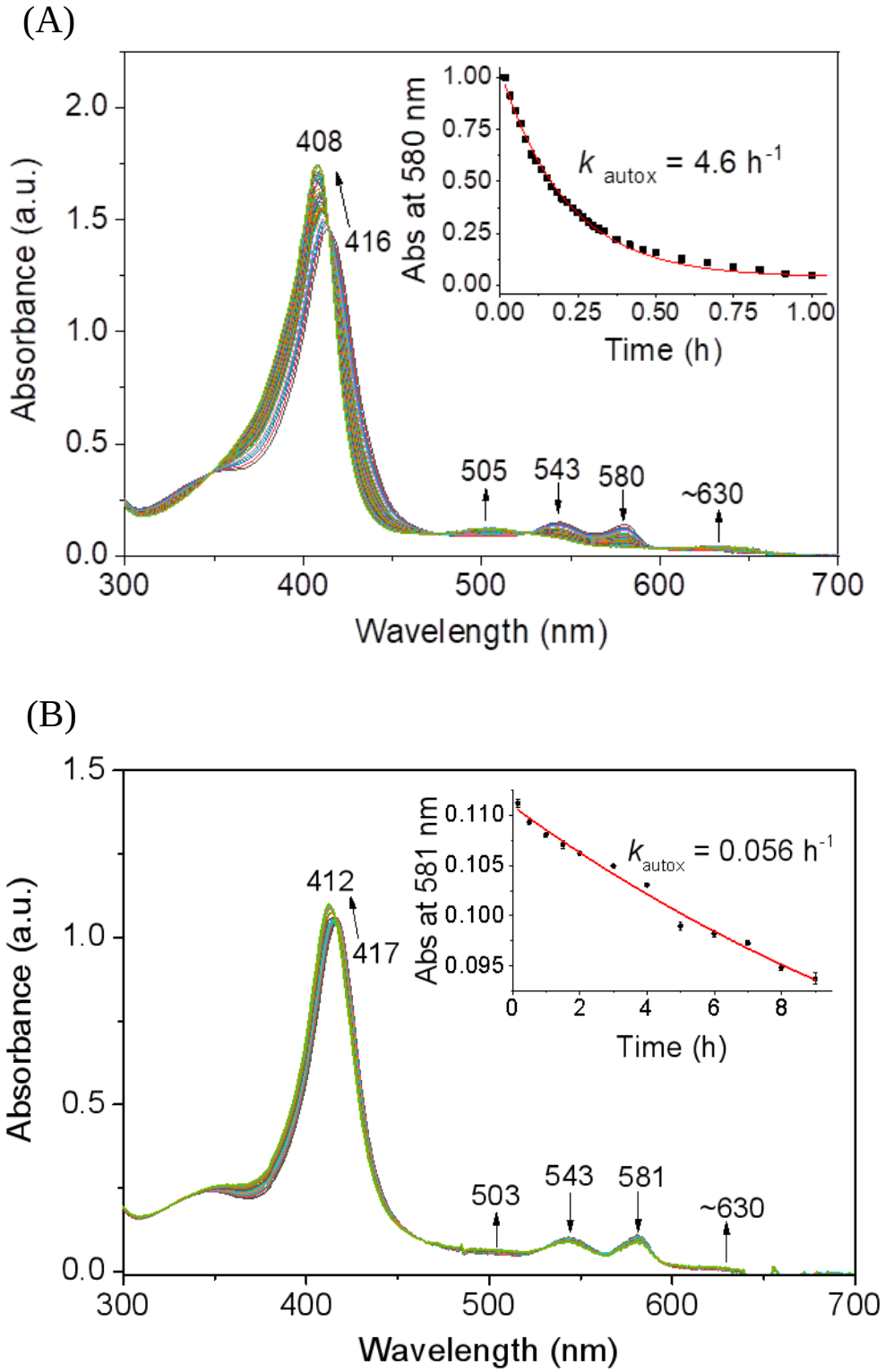


**Figure S11.** UV-vis spectra changes of (A) oxy-F46C Mb and (B) oxy-WT Mb upon autooxidation in 100 mM potassium phosphate buffer, pH 7.0, 25 $^{o}$C. The single-exponential fits of the decay of ~580 nm were shown as insets.

**Table S1.** Data collection and refinement statistics for F46C Mb and F46C/L61C Mb.

| | **F46C Mb** | **F46C/L61C Mb** |
|---|---|---|
| Wavelength | 0.9791 | 0.9791 |
| Resolution range | 24.63 - 1.9 (1.968 - 1.9) | 24.01 - 1.8 (1.864 - 1.8) |
| Space group | P 21 21 21 | P 21 21 21 |
| Unit cell (Å, °) | 39.7891 49.2505 77.5737 90 | 39.944 48.016 76.847 90 90 |
| Total reflections | 153101 (11302) | 180041 (17631) |
| Unique reflections | 12432 (1147) | 13634 (1299) |
| Multiplicity | 12.3 (9.9) | 13.2 (13.6) |
| Completeness (%) | 99.19 (94.33) | 95.63 (94.34) |
| Mean I/sigma(I) | 18.47 (6.95) | 25.39 (6.15) |
| Wilson B-factor | 15.10 | 22.61 |
| R-merge | 0.11 (0.3584) | 0.08057 (0.6268) |
| R-pim | 0.1147 (0.3782) | 0.08386 (0.6511) |
| CC1/2 | 0.03231 (0.117) | 0.02286 (0.1739) |
| R-work | 0.998 (0.946) | 0.999 (0.961) |
| R-free | 1 (0.986) | 1 (0.99) |
| Number of non-hydrogen | 12432 (1147) | 13634 (1299) |
| macromolecules | 1243 (115) | 1043 (99) |
| ligands | 0.1729 (0.1644) | 0.1794 (0.2335) |
| solvent | 0.2077 (0.2071) | 0.2253 (0.3198) |
| Protein residues | 0.955 (0.894) | 0.968 (0.877) |
| RMS (bonds) | 0.948 (0.875) | 0.963 (0.655) |
| RMS (angles) | 1402 | 1412 |
| Ramachandran favored (%) | 1215 | 1234 |
| Ramachandran allowed (%) | 43 | 43 |
| Ramachandran outliers (%) | 144 | 135 |
| Rotamer outliers (%) | 153 | 152 |
| Clashscore | 0.011 | 0.013 |
| Average B-factor | 1.39 | 1.56 |
| macromolecules | 96.69 | 97.33 |
| ligands | 3.31 | 2.67 |
| solvent | 0.00 | 0.00 |

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

## Reference:


1. Springer, B. A.; Sligar, S. G., High-level expression of sperm whale myoglobin in Escherichia coli. *Proc.Nat. Acad. Sci.* **1987,** *84* (24), 8961.

2. Zhang, P.; Yuan, H.; Xu, J.; Wang, X.-J.; Gao, S.-Q.; Tan, X.; Lin, Y.-W., A Catalytic Binding Site Together with a Distal Tyr in Myoglobin Affords Catalytic Efficiencies Similar to Natural Peroxidases. *ACS Catalysis* **2020,** *10* (1), 891-896.

3. Yan, D. J.; Li, W.; Xiang, Y.; Wen, G. B.; Lin, Y. W.; Tan, X., A novel tyrosine-heme C-O covalent linkage in F43Y myoglobin: a new post-translational modification of heme proteins. *Chembiochem* **2015,** *16* (1), 47-50.

4. Morrison, M.; Horie, S., Determination of heme a concentration in cytochrome preparations by hemochromogen method. *Anal. Biochem.* **1965,** *12* (1), 77-82.

5. Wu, L. B.; Yuan, H.; Zhou, H.; Gao, S. Q.; Nie, C. M.; Tan, X.; Wen, G. B.; Lin, Y. W., An intramolecular disulfide bond designed in myoglobin fine-tunes both protein structure and peroxidase activity. *Arch. Biochem. Biophys.* **2016,** *600*, 47-55.

6. Brantley, R. E.; Smerdon, S. J.; Wilkinson, A. J.; Singleton, E. W.; Olson, J. S., The mechanism of autooxidation of myoglobin. *J. Biol. Chem.* **1993,** *268* (10), 6995-7010.

7. Otwinowski, Z.; Minor, W., [20] Processing of X-ray diffraction data collected in oscillation mode. In *Methods in Enzymology*, Academic Press: 1997; Vol. 276, pp 307-326.

8. Urayama, P.; Phillips, G. N.; Gruner, S. M., Probing Substates in Sperm Whale Myoglobin Using High-Pressure Crystallography. *Structure* **2002,** *10* (1), 51-60.

9. Emsley, P.; Cowtan, K., Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **2004,** *60* (Pt 12 Pt 1), 2126-32.

10. Adams, P. D.; Grosse-Kunstleve, R. W.; Hung, L.-W.; Ioerger, T. R.; McCoy, A. J.; Moriarty, N. W.; Read, R. J.; Sacchettini, J. C.; Sauter, N. K.; Terwilliger, T. C., PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallographica Section D* **2002,** *58* (11), 1948-1954.

11. Murshudov, G. N.; Vagin, A. A.; Dodson, E. J., Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallographica Section D* **1997,** *53* (3), 240-255.

12. Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M., PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* **1993,** *26* (2), 283-291.