Electronic Supplementary Material (ESI) for Chemical Science.

Supporting Information

Probing the Functional Hotspots Inside Protein Hydrophobic Pockets by In-Situ Photochemical Trifluoromethylation and Mass Spectrometry

Can Lai[1,5], Zhiyao Tang[2], Zheyi Liu[1], Pan Luo[1,3,4], Wenxiang Zhang[1], Tingting Zhang[1,5], Wenhao Zhang[1,5], Zhe Dong[2*], Xinyuan Liu[2], Xueming Yang[2,3,4], Fangjun Wang[1,3,5*]

* Fangjun Wang and Zhe Dong
Email: wangfj@dicp.ac.cn (F.W.); dongz@sustech.edu.cn (Z.D.)

This PDF file includes:

Supplemental experiments

Materials

Sodium trifluoromethanesulfinate ($NaSO_2CF_3$) was purchased from Innochem (Beijing, China), zinc trifluoromethanesulfonate ($Zn(SO_2CF_3)_2$) was purchased from JSENB (Hong Kong, China). Octreotide was purchased from Yuanye (Shanghai, China). Standard peptides were synthesized by ChinaPeptides (Shanghai, China). HPLC-grade of acetonitrile and methanol were purchased from Merck (Darmstadt, Germany). Fused silica capillaries (75, 150, 200 μm i.d.) were purchased from Polymicro Technologies (Phoenix, USA). PS-DVB (30 μm, 300Å) beads were bought from H&E (Beijing, China). SPE columns (Oasis HLB) were purchased from Waters (Shanghai, China). C18 AQ beads (2.4, 3 and 5 μm, 120Å) were purchased from Sunchrom (Friedchsdorf, Germany). Micro Bio-Spin Chromatography columns (P-6 Gel, Biorad) were purchased from LABLEAD (Beijing, China). Jupiter C5 column (5 μm, 4.6 x 250 mm) was bought from Phenomenex (California, USA). Centrifugal filter tubes (3K, Millipore) were purchased from HuiBai (Shenyang, China). Other regents were purchased from Sigma Aldrich (St. Louis, MO, USA). Deionized water was obtained with a Milli-Q system from Millipore (Milford, MA). 365-nm UV lamps (Actinic BL, 10W) were purchased from Philips (Shanghai, China).

Native MS analysis

For native MS analysis, protein samples were directly injected into Exactive plus EMR MS (Thermo Fisher, USA) equipped with a nano-ESI source at a flow rate of 500 nL/min. Full MS spectra were collected at positive MS mode in the range of 700 to 4000 m/z, and mass resolution was 70,000 and 17500 for HLDs and HSA. The AGC target, maximum inject time, capillary temperature, and S-lens level were 1×10^6, 120 ms, 275 ℃, and 120.0%, respectively.

LC-MS analysis

An Exactive plus EMR MS coupled to an Agilent HPLC system (1100 Series, Agilent, USA) was applied for LC/MS analysis, standard synthetic peptides (0.1 μg) were loaded onto a trap column (3 cm × 150 μm i.d.) packed with C18 AQ beads (5 μm, 120 Å) at a flow rate of 5 μL/min of 100% mobile phase A (0.1% formic acid (FA) in H_2O) for 5 min, then, peptide was eluted from the separation column (15 cm × 75 μm i.d.) packed with C18 AQ beads (3 μm, 120 Å) by using 35% mobile phase B (0.1% FA in acetonitrile (ACN)) at 300 nL/min. MS data was collected at positive mode with MS range of 430 to 1000 m/z and mass resolution was 70,000. The AGC target, maximum inject time, capillary temperature, and S-lens level were 1×10^6, 50 ms, 275 ℃, and 60.0%, respectively. Data was recorded at the start of sample loading step.

LC-UV analysis

The tryptic digested BSA samples were analyzed with Jupiter C5 column at flow rate of 500 μL/min using Agilent HPLC system. Mobile phase C (trifluoroacetic acid (TFA)/H_2O, 0.5%/99.5%, v/v) and mobile phase D (TFA/ ACN, 0.5%/99.5%, v/v) were utilized with a binary separation gradient of mobile phase D from 20% to 60% for 30 min applied. The eluted peptides were detected with Elite iChrom D5101 UV-Vis Detector (Dalian, China) at 214 nm.

UV-vis Spectroscopic analysis

UV-vis absorption spectrum was implemented with a 5-mm thick cuvette using UV-1900

spectrophotometer (Shimadzu, Japan), peptides and NaTFMS were dissolved in water (0.05 mg/mL for hexapeptides and 0.05 mg/mL for octreotide). The detection wavelength range was 200 to 500 nm for peptides and 200 to 800 nm for NaTFMS. The reaction dynamics analysis was performed by measuring the absorption changes of octreotide-NaTFMS aqueous mixture at 365 nm in a 2-mm thick cuvette for 30 min.

Circular dichroism (CD) spectroscopy measurement

CD analysis was performed on MOS-450 spectrometry (Bio-logic, France), holo-myoglobin was diluted to 5 μM by 10 x PB buffer (pH = 7.4) before detection. A 2-mm thick cuvette was used, acquisition wavelength was recorded from 190 nm to 260 nm with 1 nm per step, each scan was repeated 3 times with acquisition period of 1 s. Buffer spectra of 10 x PB was recorded firstly and automatically subtracted from the measured sample during analysis.

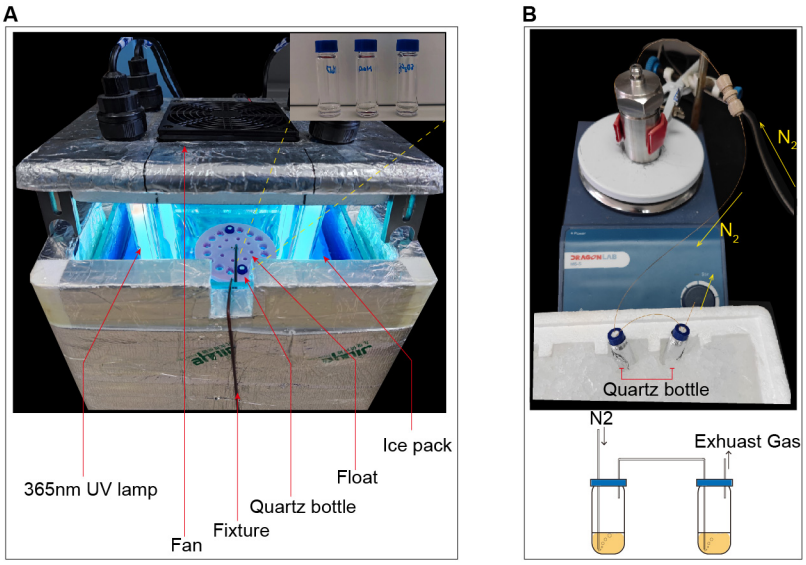

Fig. S1. Equipment for photochemical trifluoromethylation reaction (A) and reducing water dissolved oxygen (B).

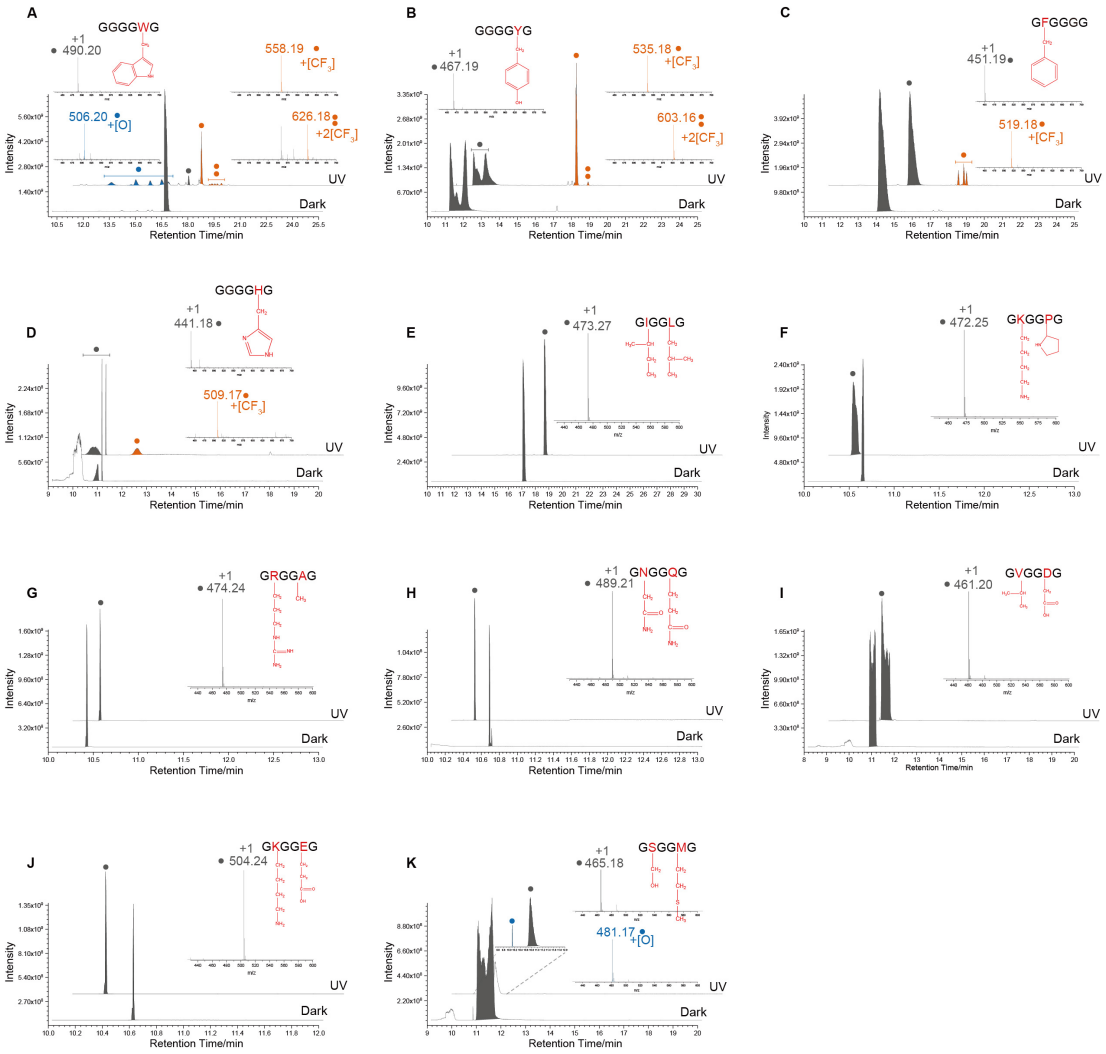

Fig. S2. Products base peak chromatograms of the standard synthetic peptide groups mixing with NaTFMS under 365-nm UV irradiation (120 min) and no UV conditions. (A)-(K) were peptides GGGGWG, GGGGYG, GFGGGG, GGGGHG, GIGGLG, GKGGPG, GRGGAG, GNGGQG, GVGGDG, GKGGEG, and GSGGMG. Insert was the peptide sequence and MS spectrum of single positive charged peptides of corresponding chromatogram peaks. The final concentration of peptide and NaTFMS were 1 mM and 100 mM.

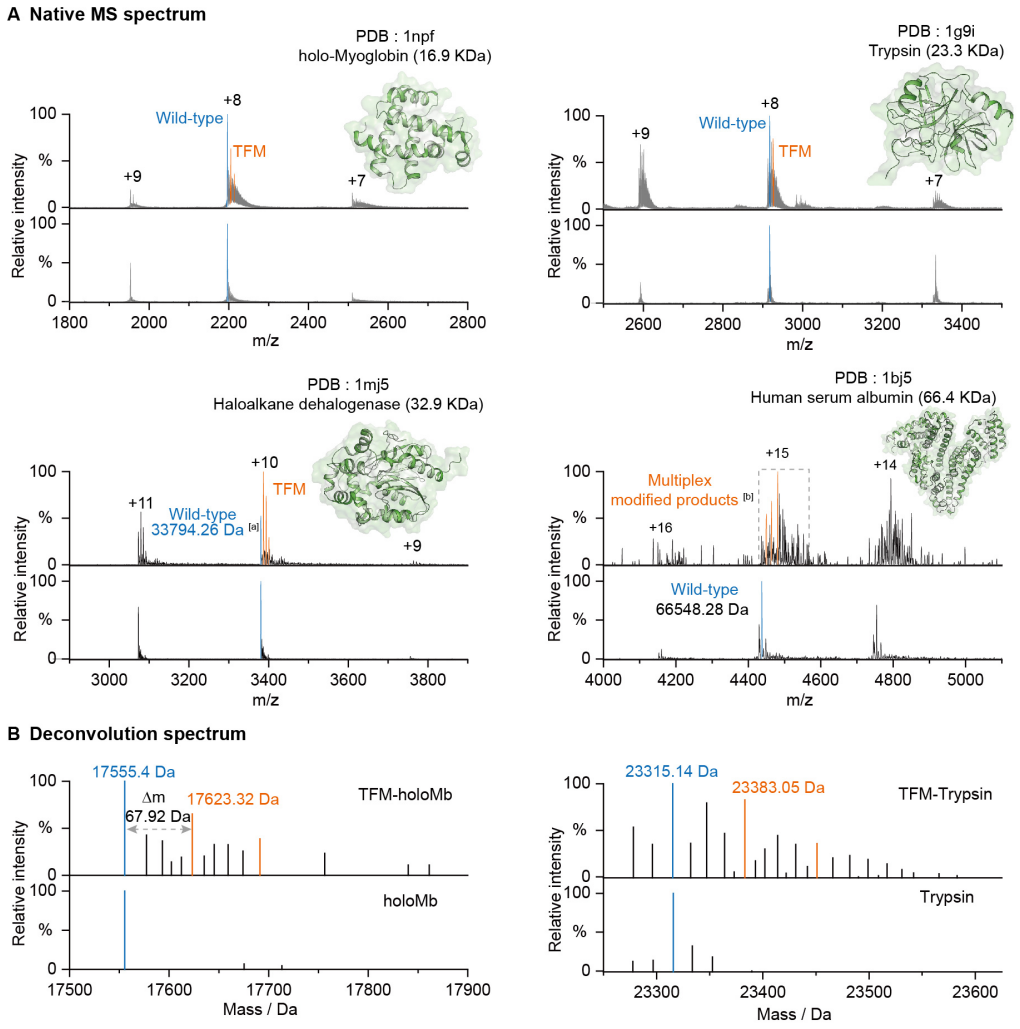

Fig. S3. Photochemical trifluoromethylation of native proteins. (A) Native MS spectrum of wild-type and TFM-Proteins. (B) Deconvolution spectrum of corresponding unmodified and TFM holo-Mb and Trypsin. Final concentration of proteins and reagent applied for the trifluoromethylation were holo-Mb (100 μM), Trypsin (100 μM), HLDs (15 μM), HSA (50 μM), and NaTFMS (50 mM) respectively. [a]Mass was calculated by the m/z of peak and corresponding charge state. [b]Due to the large mass of HSA and complex distribution of small mass TFM modifications, precise MS peaks was unable to be assigned by native MS detection, but further digestion and tandem MS detection can give the TFM ratio at each modified site.

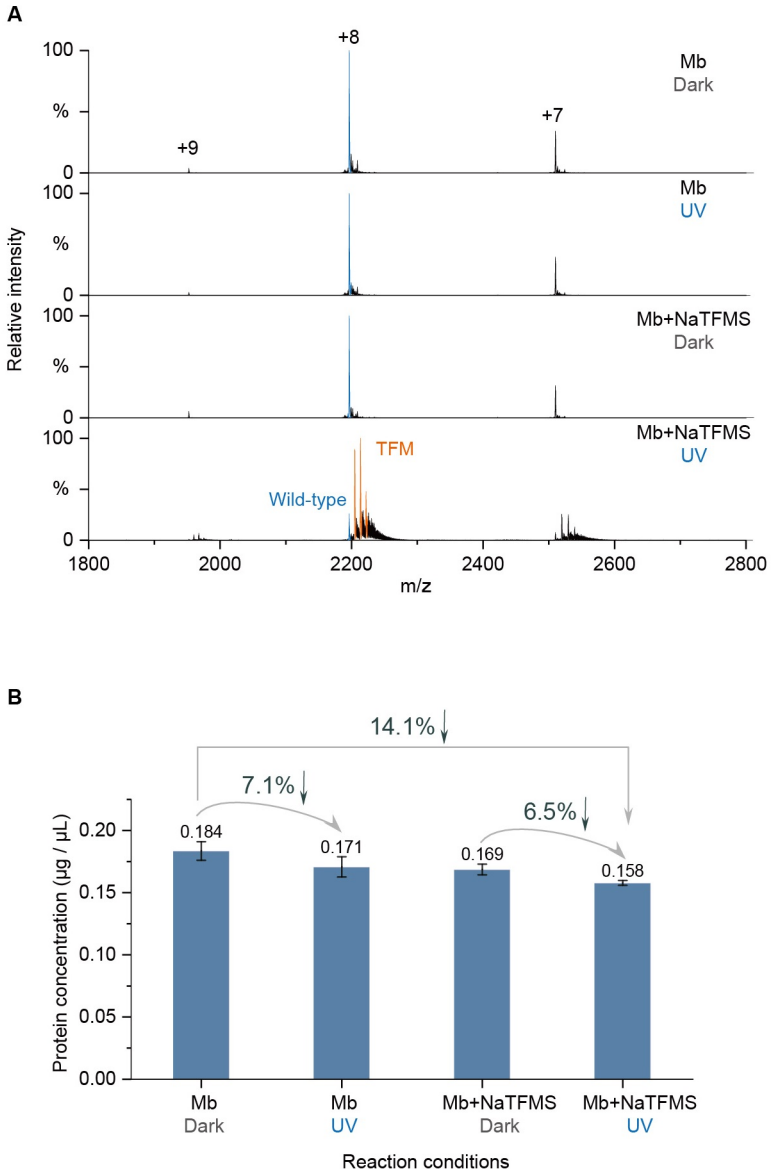

Fig. S4. Photochemical trifluoromethylation of native holo-Mb. (A) Native MS spectrum of wild-type holo-Mb, holo-Mb with UV irradiation, holo-Mb mixing with NaTFMS without UV irradiation and holo-Mb mixing with NaTFMS with UV irradiation. (B) comparison of protein concentration determined by BCA assay (to avoid the interference from reagent absorption, the background absorption of NaTFMS was determined and substrated from groups containing NaTFMS). Final concentration of holo-Mb and reagent applied were holo-Mb (10 μM) and NaTFMS (50 mM) respectively.

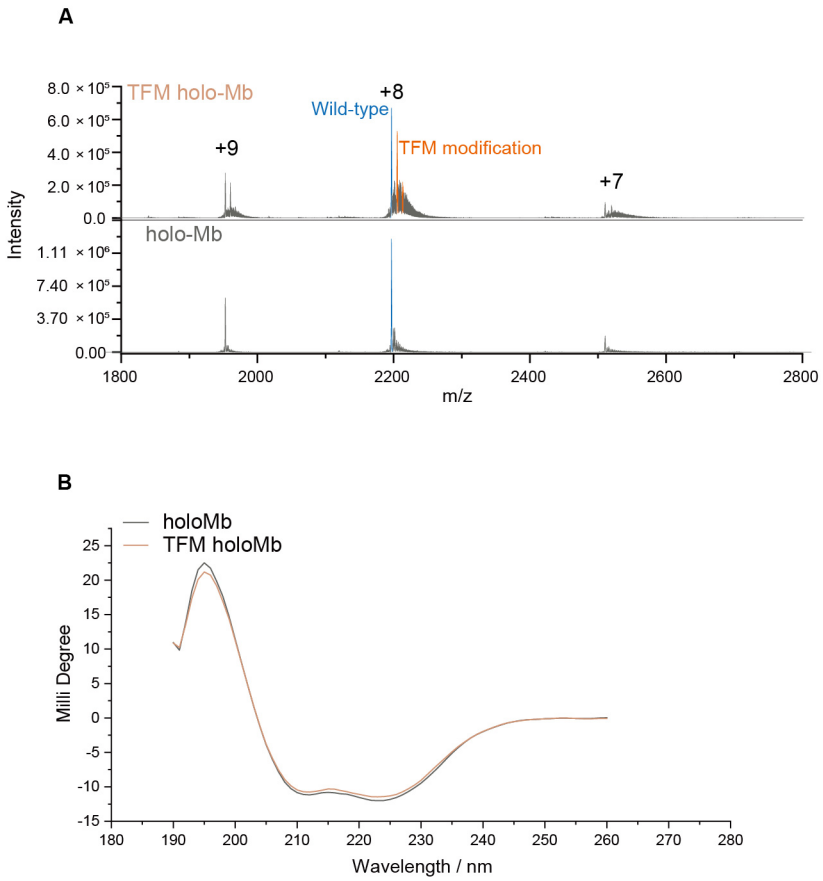

Fig. S5. CD measurement of wild-type holo-Mb and trifluoromethylated holo-Mb. (A) Corresponding native MS spectrum shows the extent of modification. (B) CD compare of wild-type holo-Mb and modified holo-Mb. Protein (100 μM) with NaTFMS (25 mM) in 10 x PB buffer (pH = 7.4) under UV irradiation time 15 min was applied for the holo-Mb trifluoromethylation. The modified holo-Mb was buffer exchanged to 10 x PB buffer to remove the unreacted NaTFMS before CD analysis.

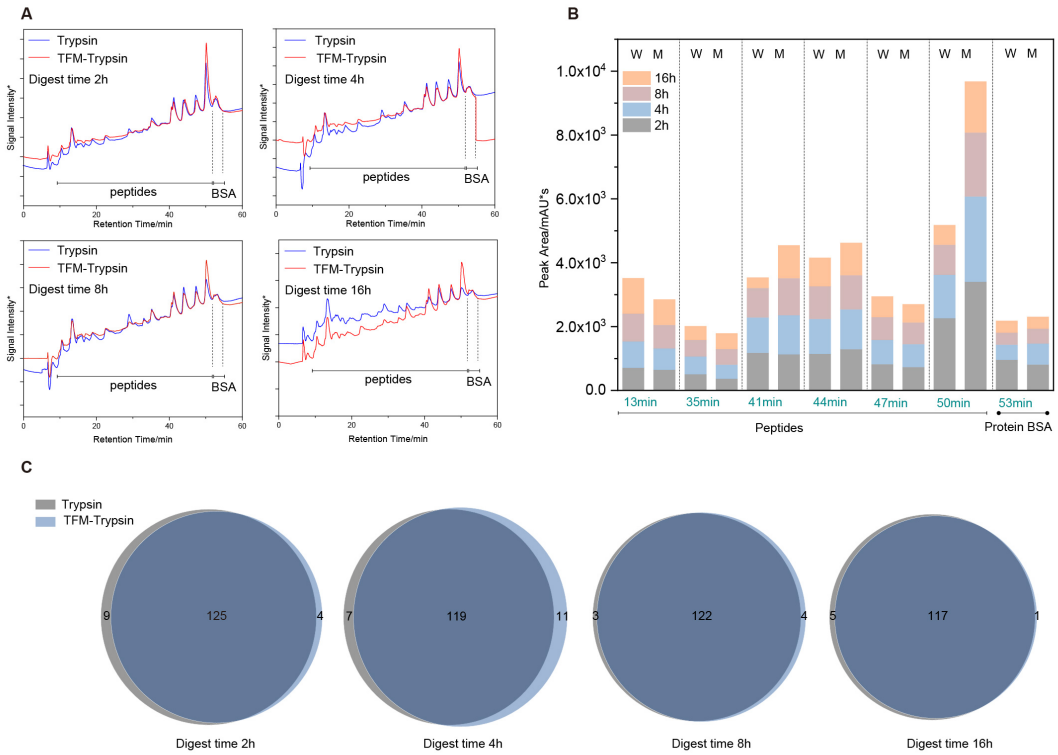

Fig. S6. Comparison of trypsin and TFM-trypsin for the digestion of native BSA. Digestion was performed in 100 mM NH_4HCO_3 (pH 8.0) buffer with an enzyme/substrate ratio of 1:50 (w/w) and incubated at 37 ℃ for 2, 4, 8, and 16 hours, individually. (A) LC-UV absorption spectrum of TFM-trypsin and trypsin digested native BSA with incubation time of 2, 4, 8 and 16 hours. *Baseline signal intensity of protein BSA were adjusted to the same level. (B) Peak area of TFM-trypsin and trypsin digested native BSA with incubation time of 2, 4, 8 and 16 hours at retention time 13, 35, 41, 47, 50 and 53 min (W denoted as wild-type trypsin and M denoted as TFM-trypsin). (C) Venn diagram of unique peptides identified from trypsin and TFM-trypsin digested BSA (Peptides were alkalized by TCEP/IAA prior to LC-MS/MS analysis).

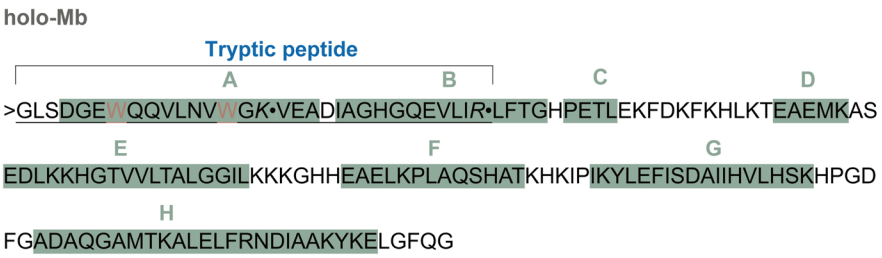

Fig. S7. Sequence of myoglobin with highlighted helix sequence region and labeled tryptic peptide sequence containing Trp residues.

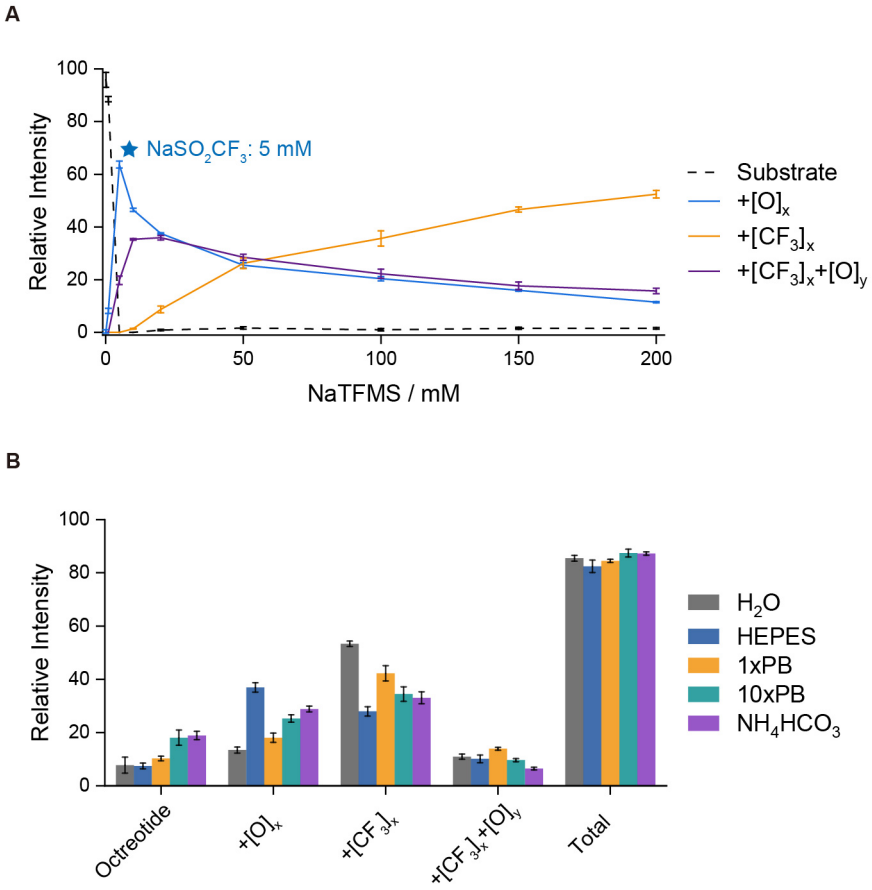

Fig. S8. Exploration of reaction conditions for the photochemical trifluoromethylation of octreotide. Relative MS intensity of products with increased $NaSO_2CF_3$ concentration (A), and reaction solvents (B). Octreotide was analyzed by directly injecting into Exactive plus EMR MS. Data was collected at positive mode with MS range of 400 to 1400 m/z, and mass resolution was 140,000. The AGC target, maximum inject time, capillary temperature, and S-lens level were 1×10^6, 50 ms, 320 ℃, and 60.0%, respectively. Total stands for the intensity sum of octreotide substrate and the known modified products.

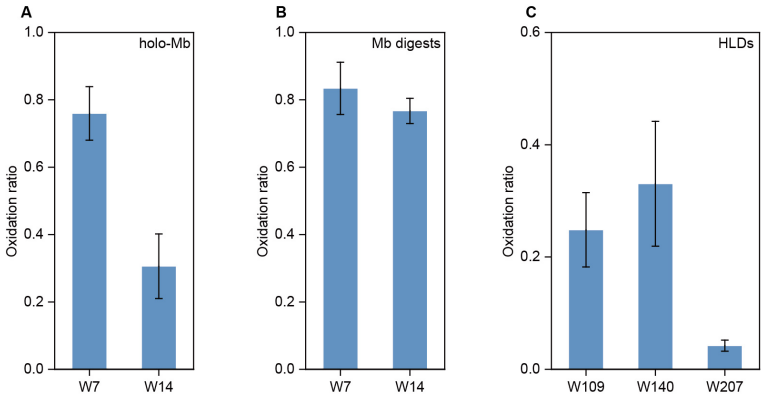

Fig. S9. Oxidation ratio of Trp residues in holo-Mb (A), Mb digests (B) and HLDs (C) after photochemical trifluoromethylation reaction.

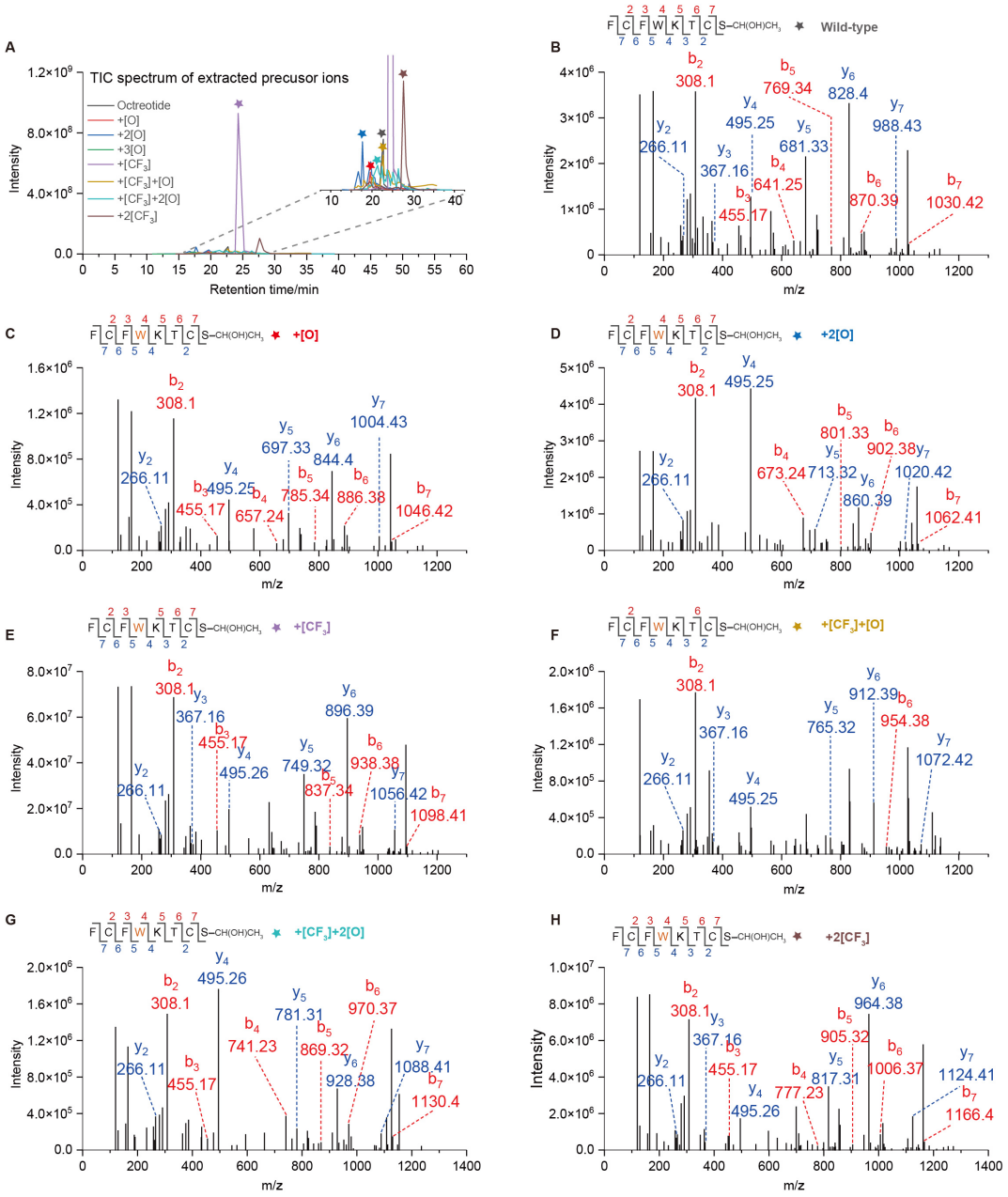

Fig. S10. LC-MS/MS analysis of the octreotide modification sites (Cys residues were reduced and alkylated by TCEP and IAA). (A) TIC spectrum of extracted precursor ions. (B)-(H) are corresponding deconvoluted MS^2 spectra of octreotide and octreotide with mono-oxidation, di-oxidation, mono-trifluoromethylation, mono-trifluoromethylation-mono-oxidation, mono-trifluoromethylation-di-oxidation and di-trifluoromethylation modifications. The modification groups assigned to Trp residue of octreotide can be inferred from mass difference of y5 and b4 fragment ions compared to none-modified octreotide.

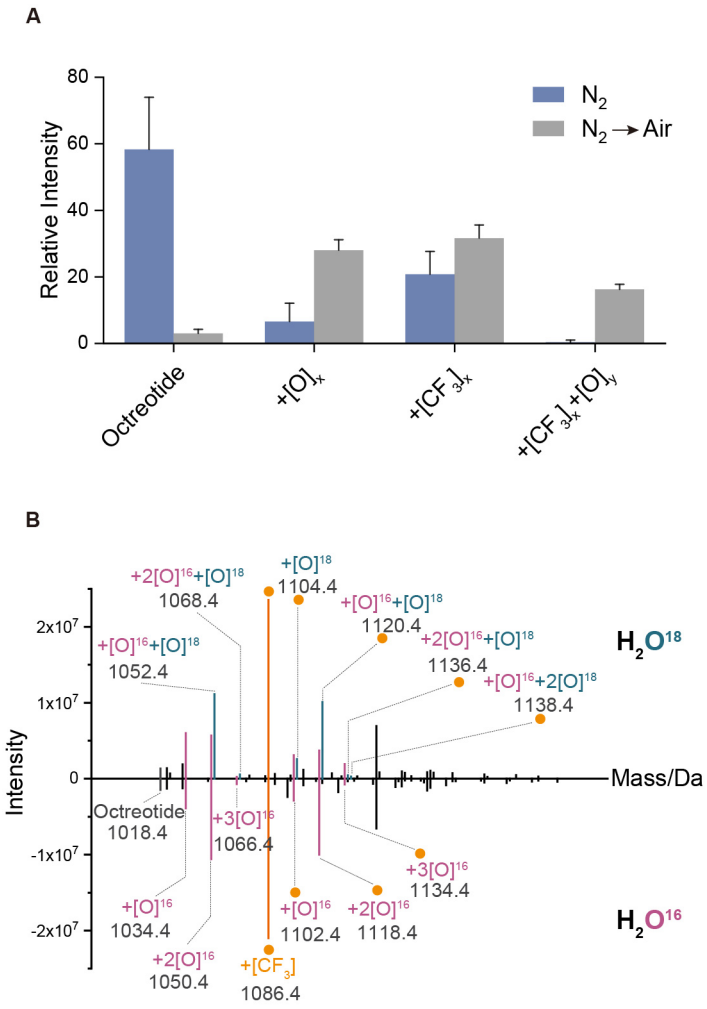

Fig. S11. Study of the oxygen role during photochemical trifluoromethylation process. (A) Relative MS intensity of product components changes after lowering water dissolved oxygen (NaTFMS: 100 mM). (B) Deconvolution spectrum of octreotide trifluoromethylation products in solvent H_2O^{16} and H_2O^{18}.

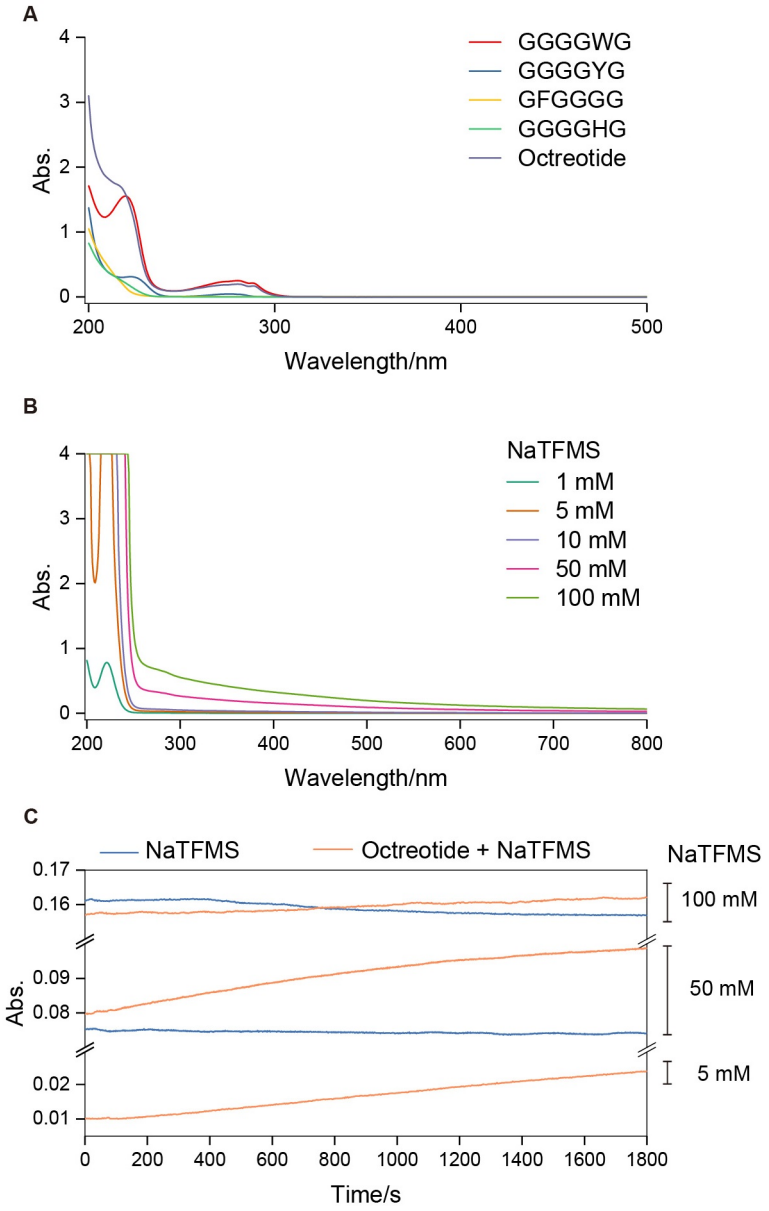

Fig. S12. UV-vis spectra of peptides (A), and different concentration of NaTFMS aqueous solution (B). (C) UV absorbance changes of NaTFMS and octreotide-NaTFMS mixture at 365-nm in 30 min measurement time.

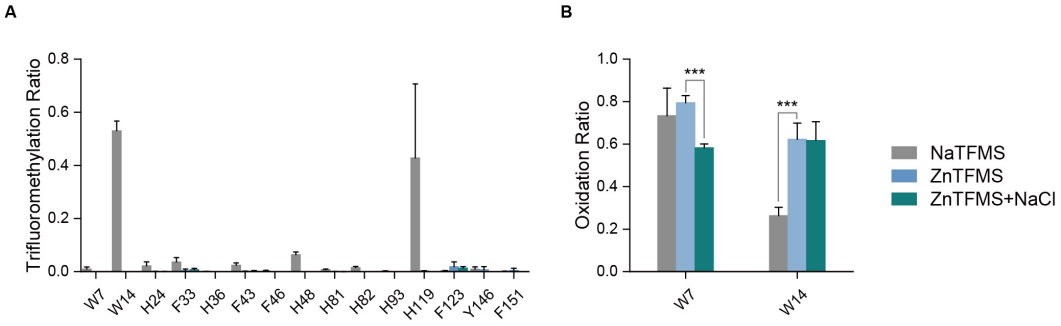

Fig. S13. Photochemically modification of holo-myoglobin by using Langlois's reagent NaTFMS (50 mM) and Baran's reagent ZnTFMS (25 mM) with and without adding of extra NaCl (50 mM) in 50 mM Tris-HCl (pH 7.4) buffer under 365-nm UV irradiation for 30min. (A) TFM ratio of modified sites (n = 3). (B) Oxidation ratio of residue W7 and W14. ***, $p<0.005$.

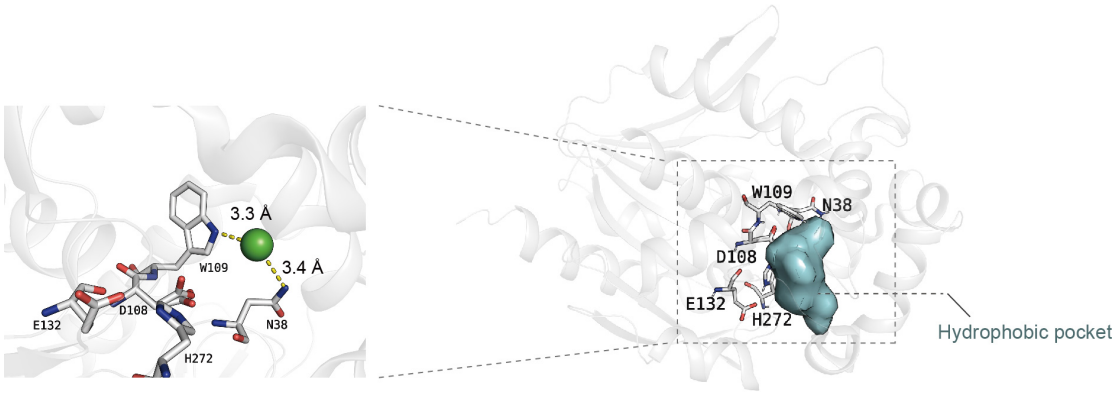

Fig. S14. Crystal structure of HLDs with catalytic associated residues shown in sticks and polar contact of W109 and N38 with chloride anion (green sphere) shown in yellow dash (PDB: 1mj5).

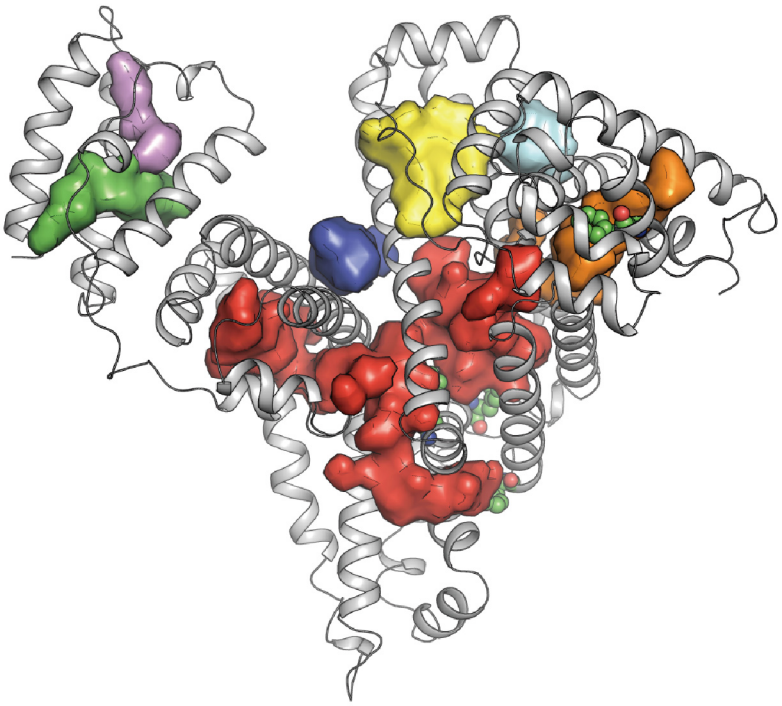

Fig. S15. Crystal structure of human serum albumin with binding pockets shown in solids (PDB: 1bj5).

Table S1. Trifluoromethylation (TFM) ratio (n = 3), surrounding hydrophobicity indices (SrHp), and solvent accessibility surface areas (SASAs) of modified Trp residues in holo-Mb and HLDs. SrHp was calculated by constructed protein crystal structure in Pymol (PDB of Mb is 1npf and HLDs is 1mj5), then residues within 8 Å radius around targeted residues were found and their hydrophobic indices were summed.

Protein	Residue Site	TFM ratio	Residues within 8 Å Radius	SrHp Index	SASAs
Mb	W7	0.007 ±0.004	GLSDGE**W**QQVLNVW---------- ---------- ILKKKGHH---------- ----------AM-KAL-LF---------- - ------WQQVLNV**W**GKVEADI--H--- V----------GT-	3.06	19.41
	W14	0.478 ±0.086	VLTALGGILK---------- -----I---L---H---F-------M---L---------- ------	14.53	8.85
HLDs	W109	0.513 ±0.049	---------- QHGNPT---LW----------L--M-- ----------AYA-HR--L---W---------- VHD**W**GSALGFD----------MEAI- -P-----F-------F---R----------F------- ---------- PTLSWPRQIPI--TP-DVV-IA--Y----- ----------L----------HF-- ---------- ---------- ---------- ---------- IE**W**ADFPE-DRDLFQAFRS-------- ---------- PR-IPIAGTP---------- -L---R---------- ------- ---------- QHGNPTS----------LIGM-- ----------RYAYAEHRDYL---------- ----DWG-ALG----------I------- ---------F--FR-------V-----F---------- ----------F------RRPTLS**W**PRQIPI-- -P--V--I---------- ---------- ----------	-2.73 3.21	52.92 0

Table S2. Initial pH of different solvents and pH after dissolving $NaSO_2CF_3$ of 150 mM. (Mixed pH was measured using precision pH test paper)

Solvent	Initial pH	Mixed pH
H_2O	neutral	<3.8
50mM HEPES	5.32	<3.8
1 x PB	7.64	3.8~4.1
10 x PB	7.39	7.0~7.5
100mM NH_4HCO_3	7.92	7.5~8.0