Electronic Supplementary Material (ESI) for Chemical Communications.

Photochemical bromination and iodination of peptides and proteins by photoexcitation of aqueous halides

Pan Luo[1,5], Zheyi Liu[1], Tingting Zhang[1,5], Xiaolei Wang[1,2], Jing Liu[3], Yiqiang Liu[2,5], Xiaohu Zhou[2,5], Yang Chen[2,5], Guangjin Hou[4], Wenrui Dong[2], Chunlei Xiao[2], Yan Jin[1], Xueming Yang[2], Fangjun Wang[1,5*]

[1] CAS Key Laboratory of Separation Sciences for Analytical Chemistry, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, China.

[2] State Key Laboratory of Molecular Reaction Dynamics, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian, 116023, China.

[3] College of Pharmacy, Dalian Medical University, Dalian, 116044, China.

[4] State Key Laboratory of Catalysis, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, China.

[5] University of Chinese Academy of Sciences, Beijing 100049, China.

Corresponding Author

Fangjun Wang − CAS Key Laboratory of Separation Sciences for Analytical Chemistry, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, China. Email: wangfj@dicp.ac.cn

Experimental

Materials

The deionized water used in the following experiments was purified by Mill-Q pure water system (Millipore Inc., Milford, MA). Fused silica capillaries with 75, 100, and 150 μm i.d. (#TSP075375; #TSP100375; #TSP150375) were purchased from Polymicro Technologies (Phoenix, USA). C_{18} AQ beads (3 μm, 120 Å; #I60828) were purchased from Sunchrom (Friedrichsdorf, Germany). C_{18} SPE column (#186000383) was purchased from Oasis®HLB (Waters, Milford, USA). HPLC grade methanol (#1.06007.4008) and acetonitrile (ACN; #1.00030.4000) were purchased from Merck (Darmstadt, Germany). Bovine serum albumin (BSA; #B2064), human serum albumin (HSA; #A1887), trypsin (#T1426), cocktail (#P8340), warfarin (#BP672), hydrogen peroxide (H_2O_2; #18304), 2,5-dihydroxybenzoic acid (DHB; #149357), DL-Dithiothreitol (DTT; #D0632), Tris (2-carboxyethyl) phosphine (TCEP; #C4507), iodoacetamide (IAA; # I1149), formic acid (FA; #695076), trifluoroacetic acid (TFA; #T6508), urea (#U6504), DMSO (#D8418), NH_4HCO_3 (#09830), Na_2HPO4 (#S5136), and NaH_2PO_4 (#S5011) were purchased from Sigma Aldrich (St. Louis, MO, USA). DL-methionine (Met; #M104886), NaBr (#S112315), and NaI (#BND0360) were purchased from Aladdin (Shanghai, China). Tris (#A600194-0500) was purchased from Sangon Biotech (Shanghai, China). SDS (#ST626) was purchased from Beyotime Biotech (Shanghai, China). $H_2{}^{18}O$ (#T3930) was purchased from Herunsheng Isotope (Henan, China). The mouse liver tissues were supported by Dalian Medical University. Standard peptides (GGGGYG, GGGGHG, GGGGWG, GGGGFG) were synthesized by ChinaPeptides Co., Ltd (Shanghai, China).

Preparation of mouse liver digest

The mouse liver tissues were obtained from Dalian Medical University in strict accordance with the recommendations in the EU animal management practices (1986). The mouse liver was sheared into small pieces, washed three times with cold-PBS solution, suspended in lysis buffer (8 M urea, 50 mM Tris-HCl, 2% SDS, and 1% cocktail pH 8.0), homogenized, and ultrasonicated (200 W, 5 s × 5 s × 30 times) in ice bath. The debris were removed by ultracentrifugation at 20,000 g and 4 ℃ for 30 min. Finally, the extracted protein samples were participated by adding 5 volumes of participation solution (acetone/ethanol/acetic acid,

50/50/0.1, v/v/v) and kept at -20 °C overnight. The participated proteins were collected by ultra-centrifugation at 20,000 g and 4 °C for 30 min and washed with ice-cold ethanol. The mouse liver protein samples were re-dissolved in the digestion buffer (8 M urea, 100 mM NH_4HCO_3, pH 8.0) and the concentration of protein samples were determined by BCA assay. 10 mg protein samples were reduced and alkylated with TCEP and IAA individually. Then, the protein solution was diluted by adding 4 volumes of 100 mM NH_4HCO_3. The digestion was performed by adding trypsin with a ratio of 1:50 (enzyme: protein, w/w) and kept at 37 °C overnight. Finally, the digested peptides were desalted by C_{18} SPE column, lyophilized, and stored at -80 °C until usage.

Sample preparation for 193-nm UV photochemical halogenation

In the 193-nm UV photochemical halogenation of peptide samples, standard peptides including GGGGYG, GGGGHG, GGGGWG, and GGGGFG, as well as mouse liver protein digest were all dissolved in 10 mM reaction solution (10 mM Na_2HPO_4, 10 mM NaH_2PO_4, 150 mM NaBr/ NaI, pH 7.4) with a concentration of 1 mg/mL. In the 193-nm UV photochemical halogenation of native proteins, 10 μM myoglobin and BSA samples were prepared in 10 mM reaction solution (10 mM Na_2HPO_4, 10 mM NaH_2PO_4, 150 mM NaBr/ NaI, pH 7.4). To elucidate the HSA-warfarin interaction details by photochemical halogenation analyses, 1 mL 10 μM HSA (in 10 mM reaction solution) was mixed with 1 μL 50 mM warfarin (in DMSO), while the control group was mixed with 1 μL DMSO.

Devices for 193-nm UV photochemical halogenation

For the direct irradiation device, protein samples were added into a 250 μL centrifugal tube. The 193-nm ArF excimer laser (10 mJ/pulse) was focused through a 250 mm convex lens (VM-TIM GmbH, Jena, German) and reflected directly into the sample tube with a diameter about 0.5 cm×1 cm by a reflector (VM-TIM GmbH, Jena, German) located at about 150 mm from the lens. Laser frequency was set at 5 Hz for standard peptides and 1 Hz for mouse liver digest, and 0.5 Hz for native proteins. To reduce the heat effect, protein irradiation experiments were all performed in ice bath.

The single-pulsed irradiation capillary reactor is established to further reduce the irradiation impact on protein structure. First of all, a section of fused silica capillary with 100 μm inner

diameter (i.d.) was related to a 500 μL syringe (Hamilton, USA; #1750) driven by a syringe pump (Longer, UK; #LSP01-2A). To form a light transparent window, about 5-mm wide capillary polyimide coating was removed by using butane flame. Then, we focused the 193-nm ArF excimer laser beam (Coherent LaserSystems Gmbh & Co.KG, Göttingen, German; #1124409) to about 2 mm width through a 250 mm convex lens and the focused laser beam was irradiated onto the laser transparent window which located at about 190 mm from the lens. The laser frequency was set at 5 Hz (5 mJ/pulse). Reaction sample solution was injected into the capillary reactor with a pumping speed of 5 μL/min (~1.07×10^{-2}m/s), followed by irradiation of focused laser beam when passing through the laser transparent window. The laser frequency and the syringe pumping speed were matched to realize irradiation each bolus of sample with just one pulse of laser shot (10 ns). Finally, the modified sample solution was collected by a tube containing equal volume of scavenger solution (0.5 mg/ml tyrosine) to quench the remaining radicals in the native protein halogenation experiments. After photochemical halogenation, the intact protein samples were digested by trypsin, desalted by C_{18} SPE column, lyophilized, and followed by LC-MS/MS analyses. Pictures of our setup were listed below.

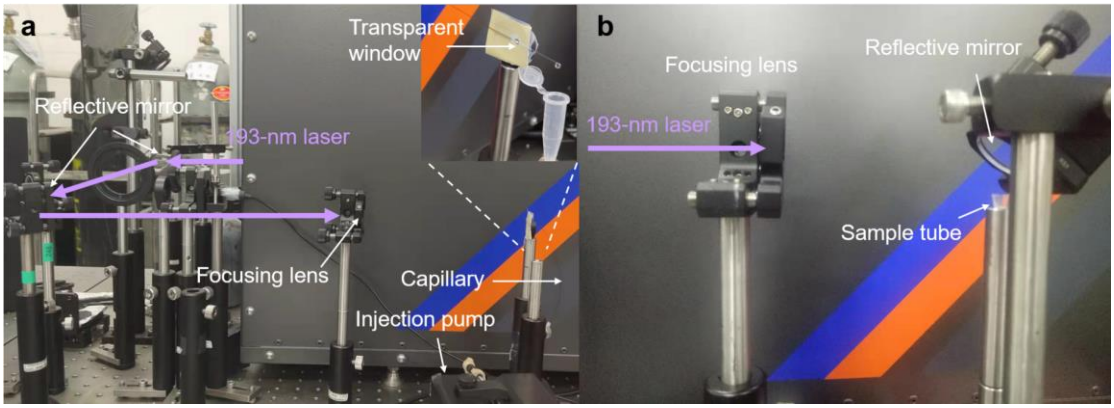


Figure S1. Setup pictures of the single-pulsed irradiation capillary reactor (a) and the direct irradiation device (b).

MALDI-TOF-MS analysis

The photochemically modified standard peptides were desalted by C_{18} SPE column, re-dissolved in 0.1% formic acid (FA) aqueous solution, and analyzed by MALDI-TOF-MS (AB, SCIEX, USA) in linear positive ion mode. 0.5 μL peptide samples and 0.5 μL matrix solution were successively dropped onto the sample plate for the follow-up MALDI-TOF-MS analysis.

The matrix solution was prepared by dissolving DHB into methanol/H_2O/TFA (50:50:0.2, v/v/v) with a concentration of 25 mg/mL.

LC-MS/MS analysis

After the photochemical halogenation, the tryptic peptide sequences and modification sites were determined by LC-MS/MS analyses. Vanquish Flex HPLC system coupled with Orbitrap Fusion Lumos MS (Thermo) was utilized for the LC-MS/MS analyses of standard protein and complex mouse liver protein digests. The LC separation mobile phase A was 0.1% FA aqueous solution and mobile phase B was 0.1% FA acetonitrile solution. The tryptic peptide samples were re-dissolved into 0.1% FA (v/v) solution to 0.1 μg/μL, loaded onto a 3 cm x 150 μm C_{18} (3 μm, 120 Å) trap column, and separated by a 15 cm x 75 μm i.d. C_{18} (3 μm, 120 Å) capillary column with flow rate at about 200 nL/min. The reversed-phase separation gradients for standard protein and complex mouse liver digests were set as 5-35% mobile phase B in 35 and 70 min, respectively. For Orbitrap Fusion Lumos MS analysis, the full MS spectra were collected with the orbitrap with a resolution of 120,000 and the MS/MS spectra were collected in a 'top-speed' manner (3 seconds) with a resolution of 15,000 in Orbitrap. The precursor ions with a charge state of 2 to 5 were isolated with a m/z window of 1.4 m/z and subjected to high-energy collisional dissociation (HCD) with 28.0% NCE.

Besides, surveyor LC system coupled with LTQ-Orbitrap Velos MS (Thermo) was utilized for the LC-MS analyses of standard peptides. The LC separation gradient was identical to standard protein. The LTQ-Orbitrap Velos MS parameters were set as ion transfer capillary 275 °C, electrospray voltage 1.8 kV, full MS scan from 350 to 1850 m/z with a resolution of 60,000 in Orbitrap.

Intact protein MS analysis

To explore the influences of photochemical bromination and iodination on protein structure, the halogenated intact protein samples were substituted with 50 mM NH_4OAc for native MS analysis. Intact protein MS analyses were performed on the LTQ-Orbitrap XL MS system (Thermo).

^{1}H NMR spectroscopy

500 μL of the photochemically modified standard peptide samples (1 mg/ml) were mixed with 25 μL D_2O for ^1H NMR analysis. NMR datasets were collected at 25°C by using a AVANCE III HD 700 MHz Bruker NMR spectrometer (Ettlingen, Germany) with a TCI cryoprobe.

Data processing and drawing software for MALDI-MS and the LC-MS analysis

The acquired data were opened with Data Explorer (TM) and Xcalibur Browser software in MALDI-MS and the LC-MS analysis, respectively. The spectrum was then exported to a excel table and opened in the Origin software for first drawing. Next, the origin vector graphs were copied to Adobe Illustrator for further assembly and drawing. In Adobe Illustrator, the noticeable peaks were draw in different colors.

MS dataset analysis

The intact MS datasets were deconvoluted by using Proteome Discoverer 1.4. The acquired LC-MS/MS datasets were searched against the database of corresponding protein sequences downloaded from Uniprot using Maxquant (version 1.6.0.1) to determine the peptide sequences and photochemical modification sites. Carbamidomethylation (C) was set as fixed modification in dataset searching. Additional variable modifications for photochemical bromination experiments: Br1 (YH) +77.9 Da, Br2 (YH) +155.8 Da, Br1O1 (W) +93.9 Da, Br2O1 (W) +171.8 Da, Br1O2 (W) +109.9 Da, and Oxi (MW) +15.99 Da were included. Additional variable modifications for photochemical iodination experiments: I1 (YH) +125.9 Da, I2 (YH) +251.8 Da, and Oxi (M) +15.99 Da were included. The other parameters were the default values of MaxQuant. Finally, the matched MS/MS spectra of modified peptides were manually validated and filtered to further improve the identification reliability.

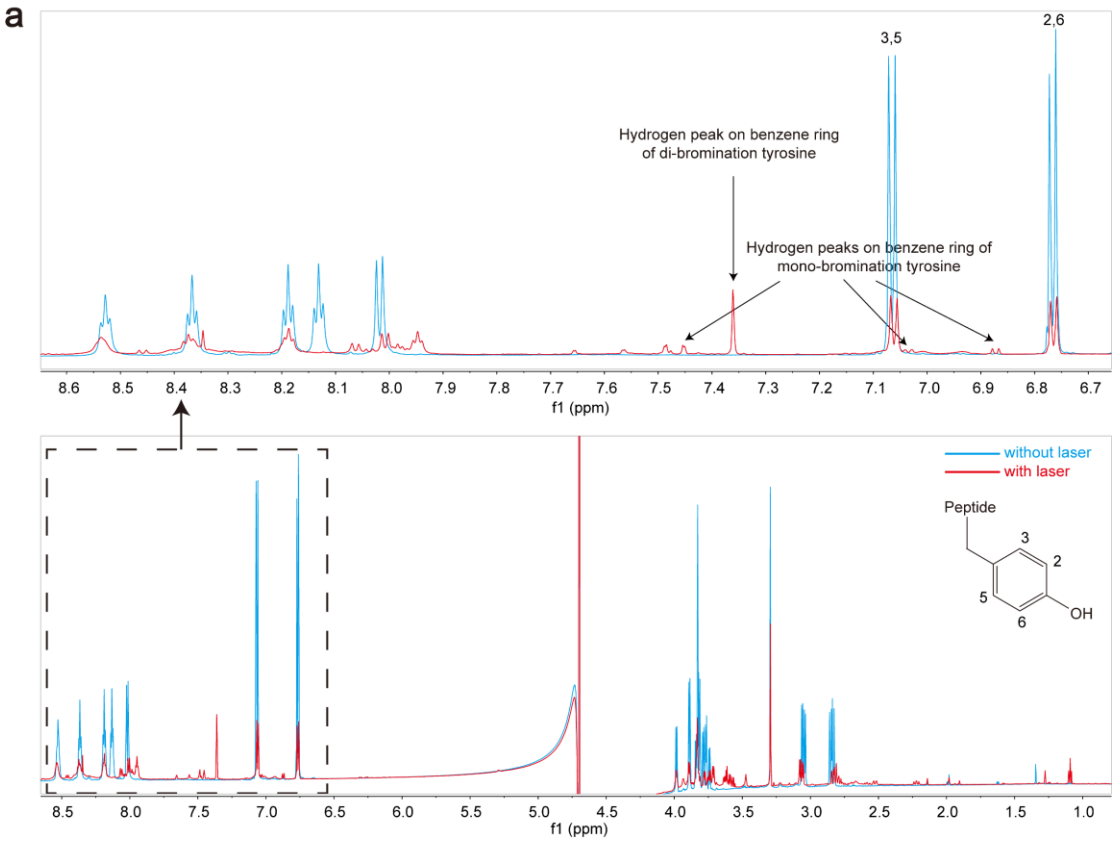
The peptide photochemical modification rate (%) was calculated by the sum intensity of the modified peptides divided by the sum of the parent peptides and modified peptides as the formula:

$$\text{Modification rate (\%)} = 100 * \frac{\sum I_{modified}}{I_{umodified} + \sum I_{modified}}.$$

^{1}H NMR analyses of bromination peptides

The photochemical bromination products were analyzed by ^1H NMR. The original and brominated product peaks were mainly assigned in Figure S1. It was difficult to match the

complex peaks of bromination Trp peptide (GGGGWG) due to the mix of multiple bromination and oxidation products, but we could clearly see the Trp oxidation leading to the great decrease of chemical shifts. For iodination experiments, abundant I_2 could be generated during the photochemical iodination process and the peptide ^1H NMR signals were mainly decreased or disappeared. Thus, the iodinated peptides and proteins were mainly characterized by LC-MS/MS analyses.



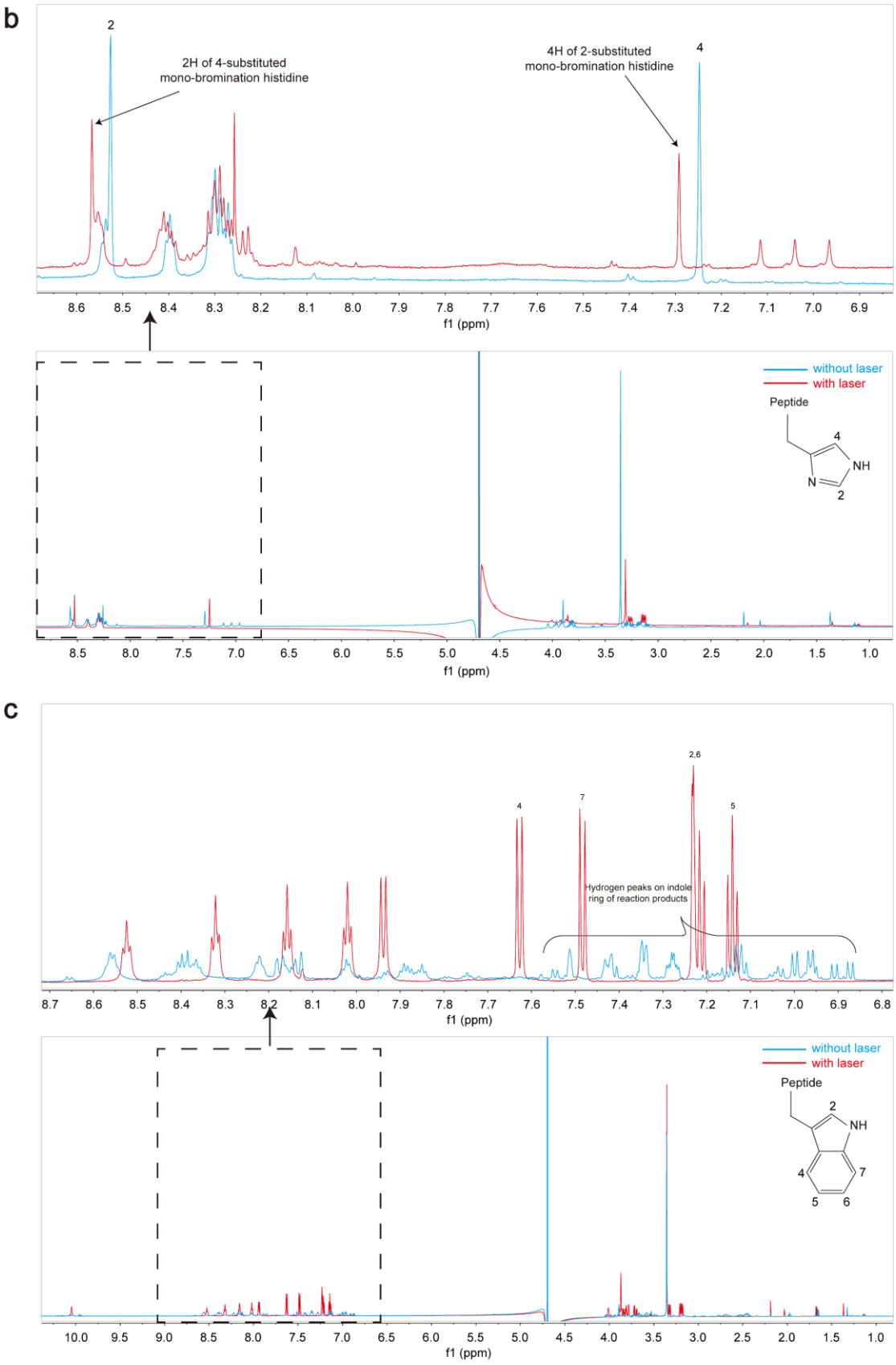


Figure S2. ^{1}H NMR characterization of brominated peptides with zoomed version on the top and, (a) GGGGYG with 60 s, (b) GGGGHG with 40 s, and (c) GGGGWG with 60 s irradiation

of 193-nm UV laser. The results of control experiments without laser irradiation were given in blue.

Photochemical peptide bromination

The results of phenylalanine peptide (GGGGFG) with different 193-nm UV laser irradiation time were shown in Figure S3. No obvious bromination Phe peptides could be observed.

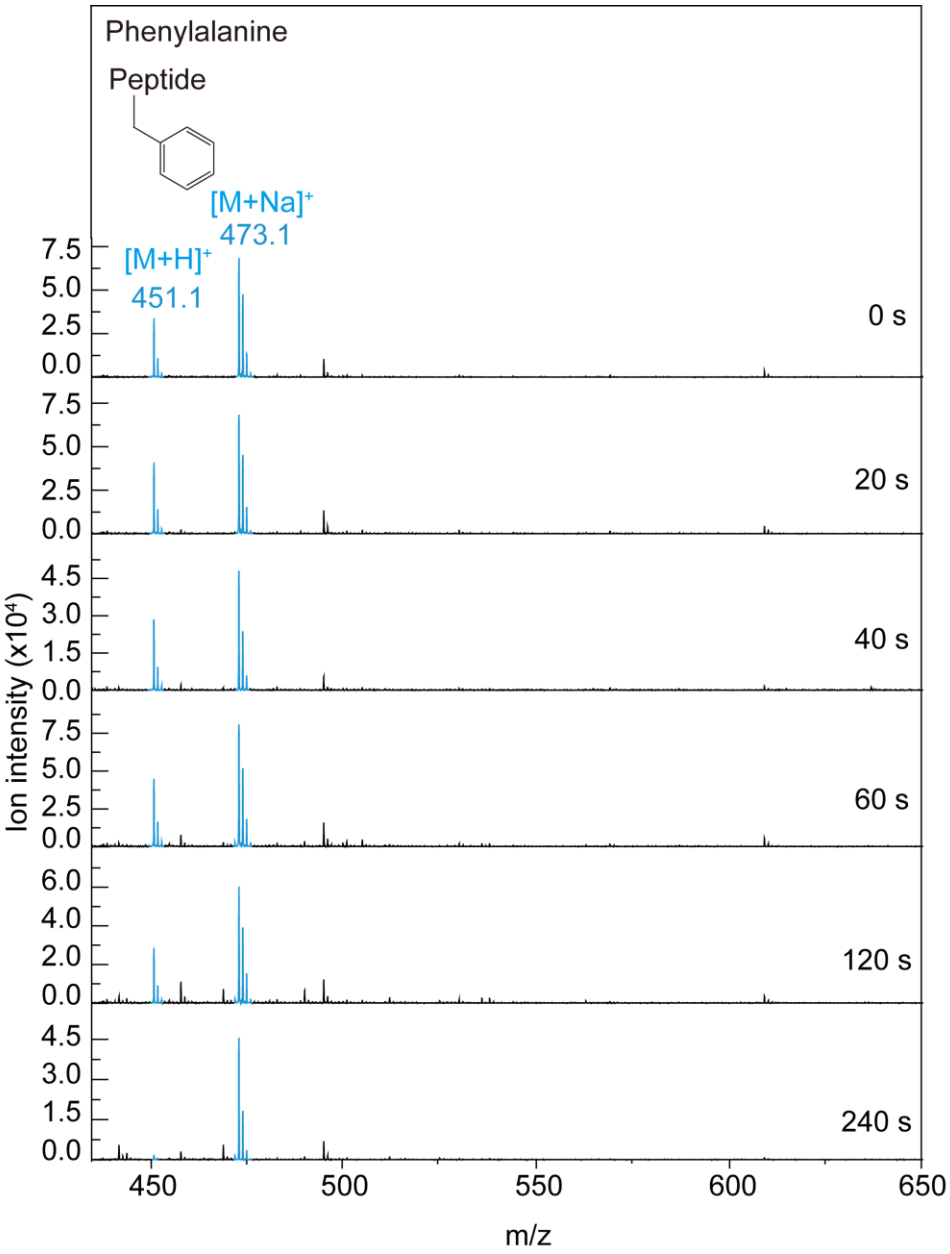


Figure S3. MALDI-TOF-MS characterization of peptide GGGGFG with 193-nm UV laser irradiation times of 0, 20, 40, 60, 120 and 240 s. The laser was operated at 5 Hz, 5 mJ/pulse. The peptides were prepared in reaction solution with 150 mM Br^- (10 mM Na_2HPO_4, 10 mM

NaH_2PO_4, 150 mM NaBr, pH 7.4).

Obviously, the tyrosine peptide (GGGGYG) could be brominated with significant bromine isotope peaks and the bromination efficiency increased along with the laser irradiation time (Figure S4). Products with 60 s irradiation were further subjected to LC-MS analyses and the mono- and di-bromination products were observed with characteristic bromine isotope peaks and molecular weight matching (Figure S5). The effect of NaBr concentration and solution volume on bromination efficiency were explored under 60 s irradiation (Figure S6). With the increase of NaBr concentration, di-bromination product was dominant and continued to increase. When the NaBr concentration increased to 300 mM, the bromination efficiency was hard to calculate due to the high salt concentration greatly decreased the S/N ratio MS detection. On the other hand, with the increasement of solution volume, the rate of di-bromination product greatly decreased, mainly due to relatively large solution volume might decrease the concentration of reactive ·Br radicals generated at identical laser irradiation.

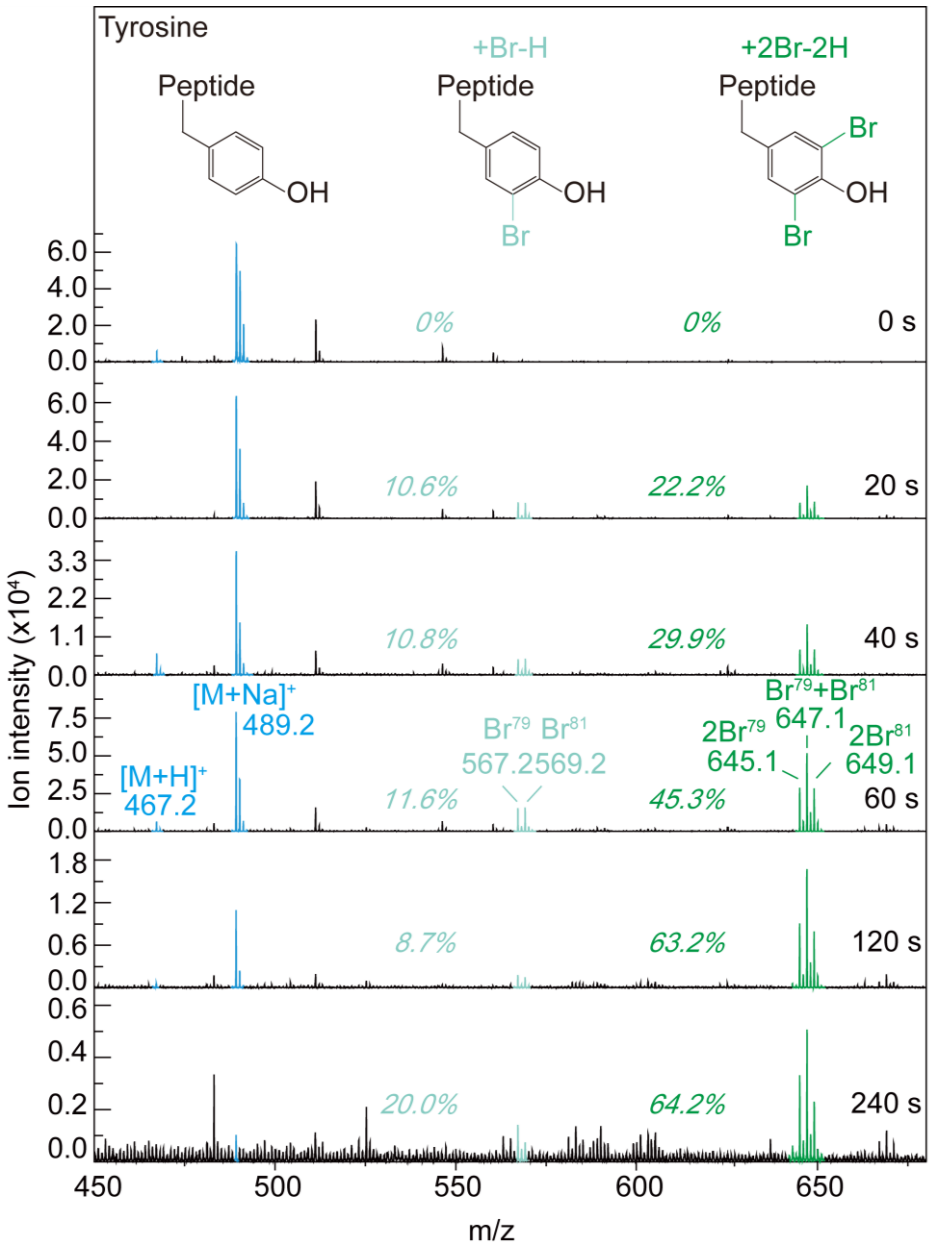


Figure S4. MALDI-TOF-MS characterization of peptide GGGGYG with different 193-nm UV laser irradiation times of 0, 20, 40, 60, 120, and 240 s. The 193-nm UV laser was operated at 5 Hz, 5mJ/pulse. The peptide samples were prepared in reaction solution with 150 mM Br^- (10 mM Na_2HPO_4, 10 mM NaH_2PO_4, 150 mM NaBr, pH 7.4).

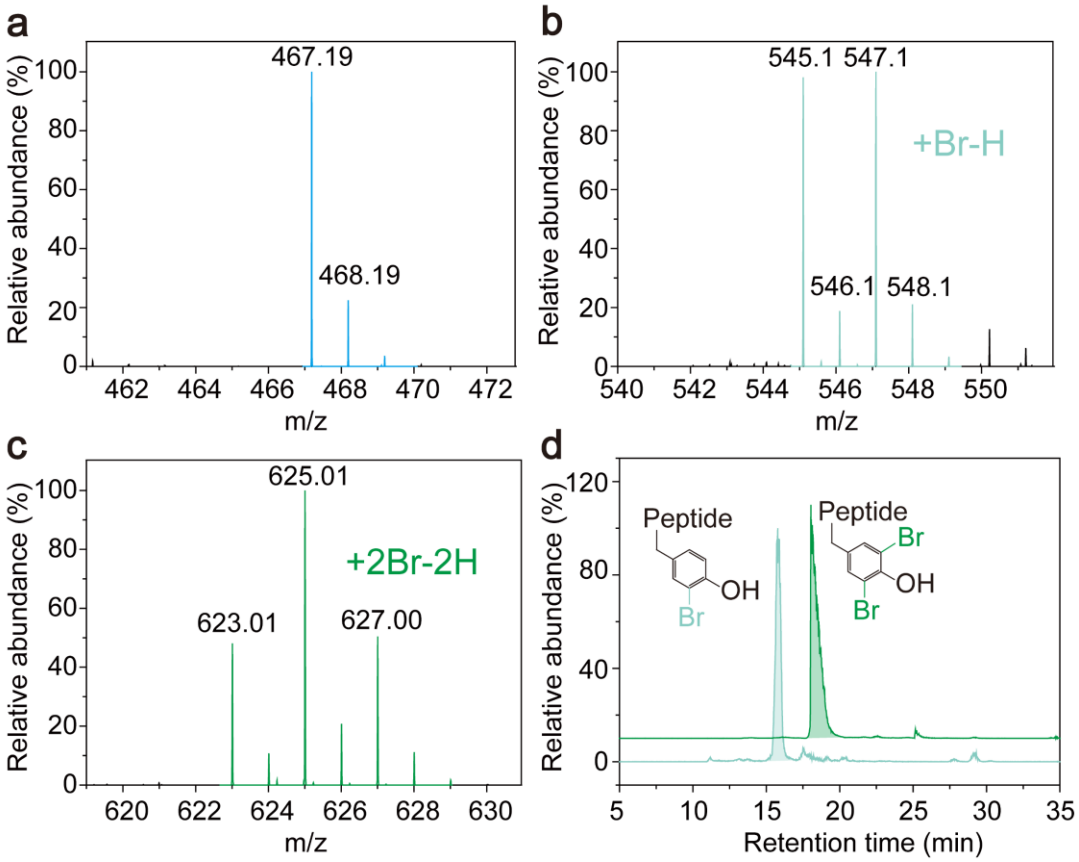


Figure S5. (a-c) The MS spectra in LC-MS analyses of peptide GGGGYG with 60 s irradiation, (a) the original form without laser irradiation, (b) the mono-bromination product, and (c) the di-bromination product. (d) The extracted product chromatographic peaks for the main products.

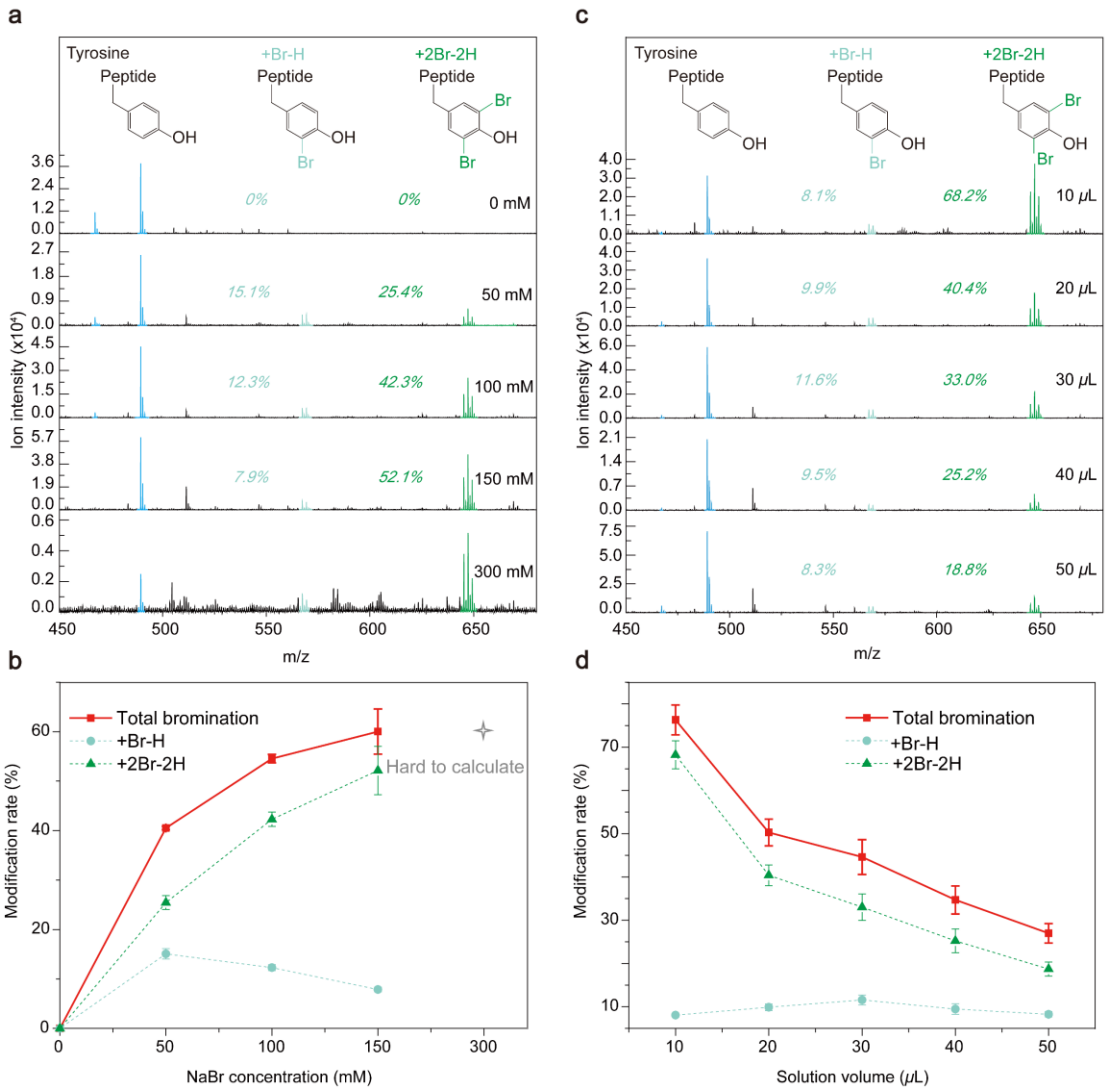


Figure S6. (a, c) MALDI-TOF-MS characterization of peptide GGGGYG, (a) with different NaBr concentration of 50, 100, 150, 300 mM, and (c) with different reaction solution volume of 10, 20, 30, 40 L, 50 μL. (b, d) The modification rate curve of total bromination product (red), mono-bromination product (light green), di-bromination product (dark green), (b) with different NaBr concentration, and (d) with different reaction solution volume. The 193-nm UV laser irradiation was operated at 5 Hz, 60 s, and 5 mJ/pulse. The peptide samples were prepared in reaction solution with 150 mM Br^- (10 mM Na_2HPO_4, 10 mM NaH_2PO_4, 150 mM NaBr, pH 7.4).

The histidine peptide (GGGGHG) could also be brominated with significant bromine isotope peaks and the bromination efficiency increased along with the laser irradiation time (Figure S7). Products with 40 s irradiation were further subjected to LC-MS analyses and the

mono- and di-bromination products were observed with characteristic bromine isotope peaks and molecular weight matching (Figure S8).

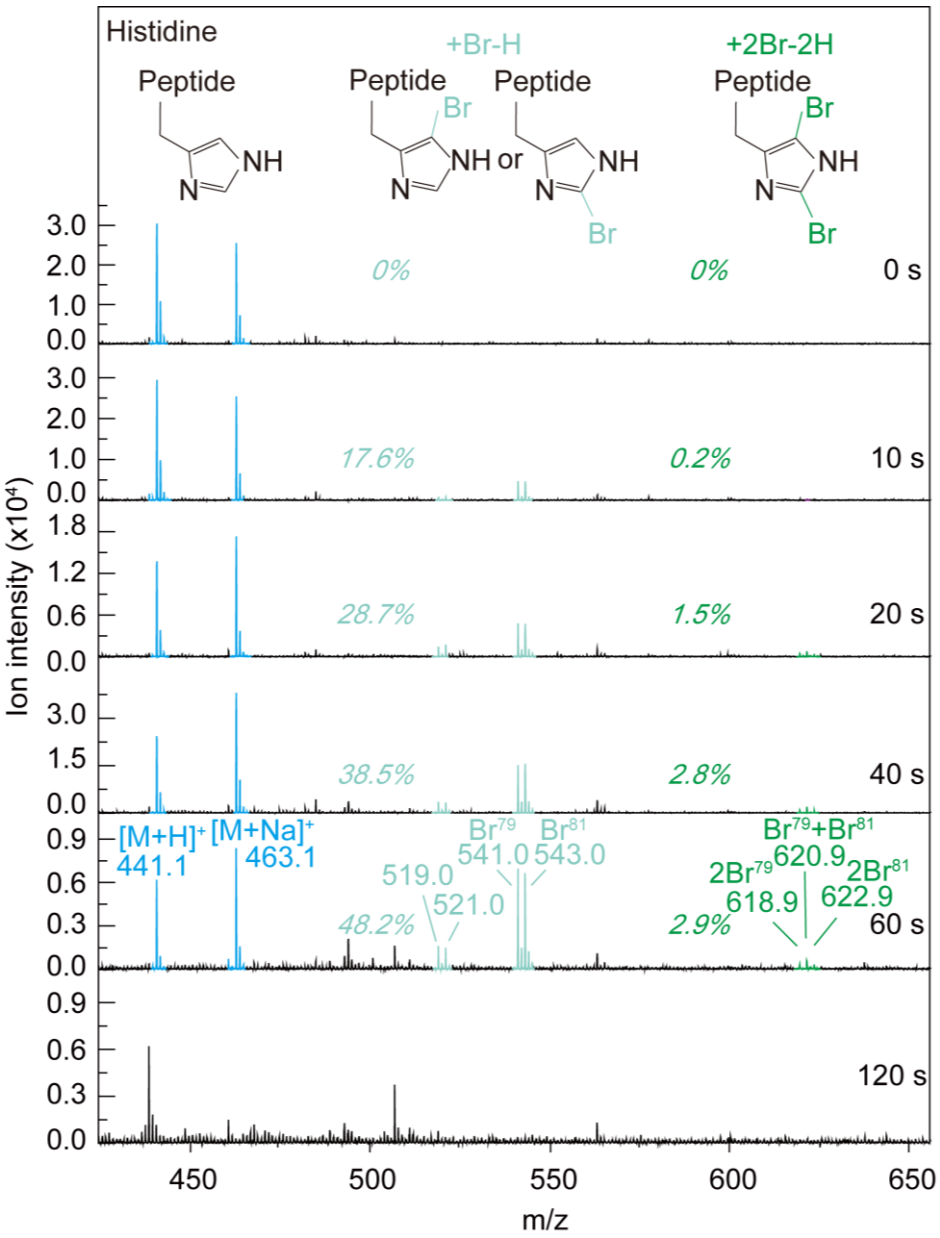


Figure S7. MALDI-TOF-MS characterization of peptide GGGGHG with different 193-nm UV laser irradiation times of 0, 10, 20, 40, 60, and 120 s. The 193-nm UV laser was operated at 5 Hz, 5 mJ/pulse. The peptide samples were prepared in reaction solution with 150 mM Br^- (10 mM Na_2HPO_4, 10 mM NaH_2PO_4, 150 mM NaBr, pH 7.4).

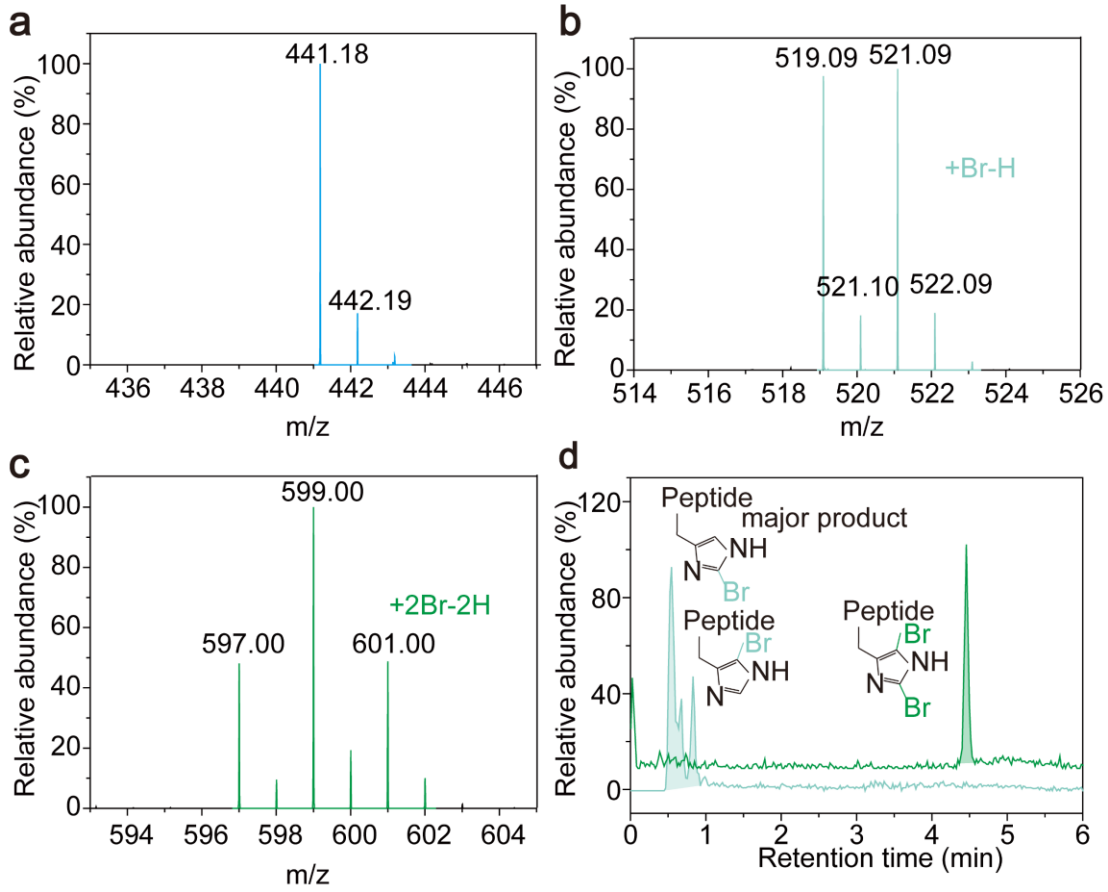


Figure S8. (a-c) The MS spectra from LC-MS analysis of peptide GGGGHG with 40 s irradiation, (a) the original form without laser irradiation, (b) the mono-bromination product, and (c) the di-bromination product. (d) The extracted product chromatographic peaks for the main products.

Moreover, the tryptophan peptide (GGGGWG) could also be brominated with significant bromine isotope peaks and the bromination efficiency changed along with the laser irradiation time (Figure S9). Products with 60 s irradiation were further subjected to LC-MS analyses and the corresponding products were observed with characteristic bromine isotope peaks and molecular weight matching (Figure S10).

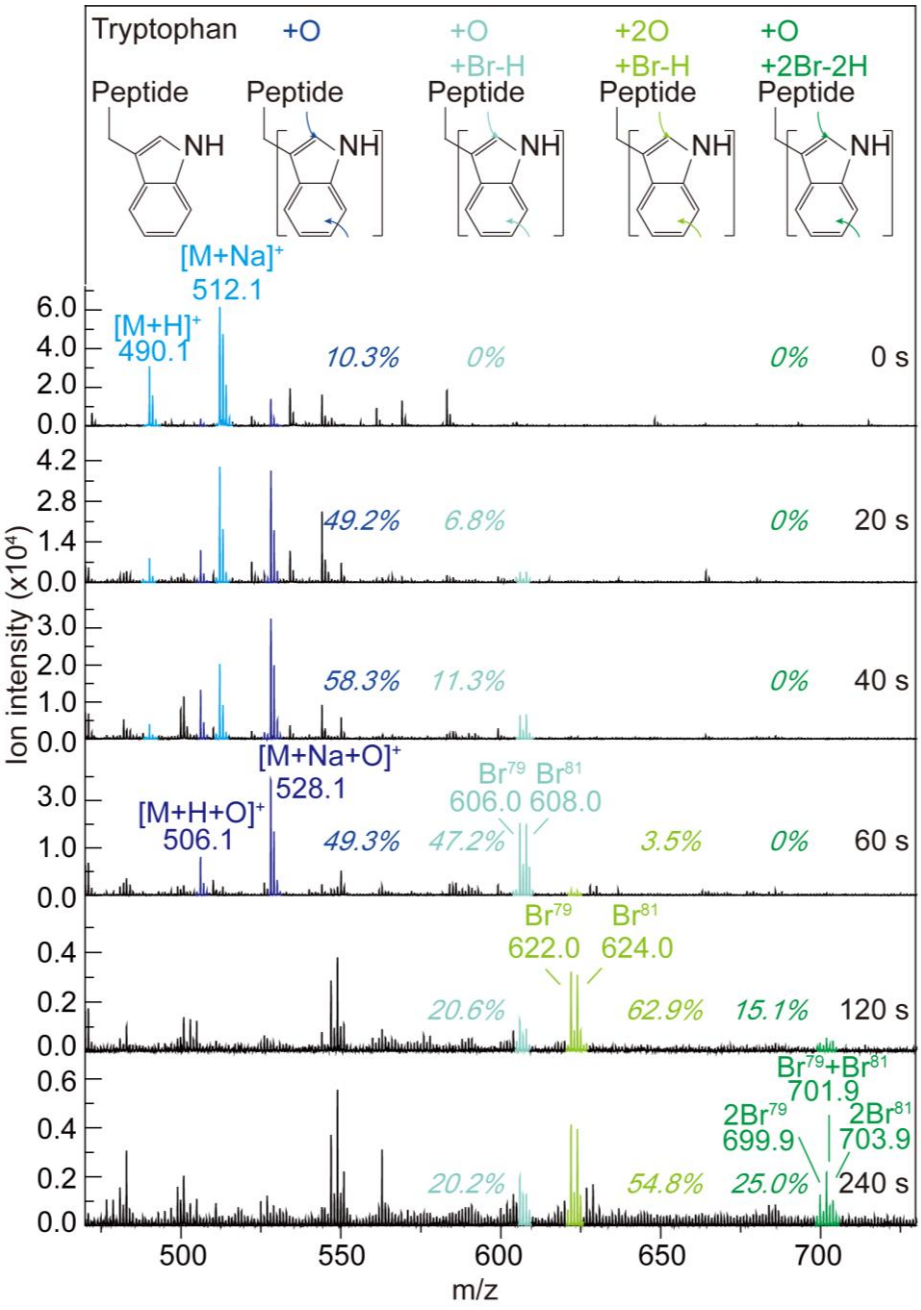


Figure S9. (a) MALDI-TOF-MS characterization of peptide GGGGWG with different 193-nm UV laser irradiation times of 0, 20, 40, 60, 120, and 240 s. (b-e) The 193-nm UV laser was operated at 5 Hz, 5 mJ/pulse. The peptide samples were prepared in reaction solution with 150 mM Br^- (10 mM Na_2HPO_4, 10 mM NaH_2PO_4, 150 mM NaBr, pH 7.4).

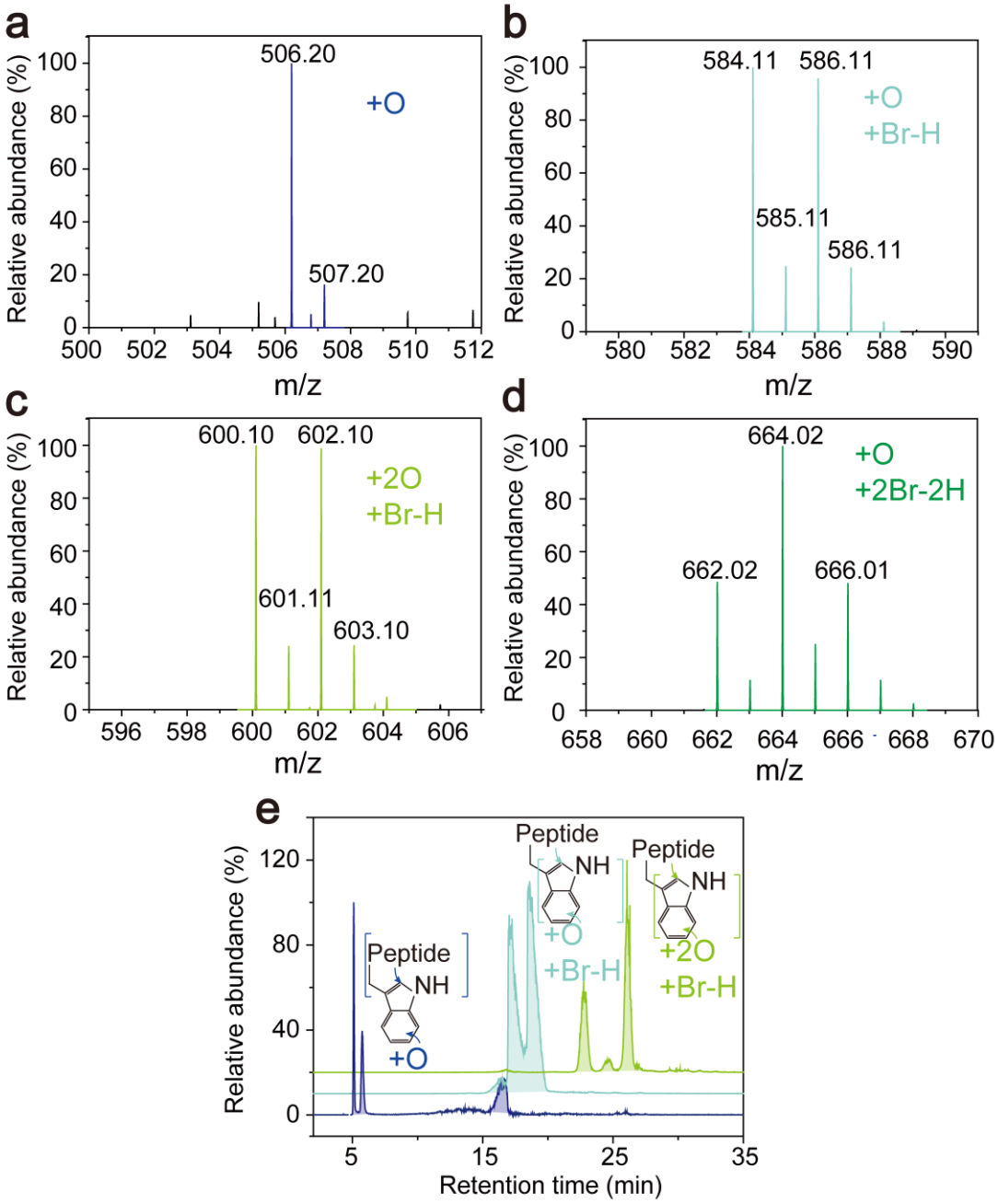


Figure S10. The MS spectra from LC-MS analysis of peptide GGGGWG with 60 s irradiation, (a) the mono-oxidation product, (b) the mono-oxidation and -bromination product, (c) the di-oxidation and mono-bromination product, and (d) the mono-oxidation and di-bromination product. (e) The extracted product chromatographic peaks for the main products.

Photochemical peptide iodination

The results of tryptophan (GGGGWG) and phenylalanine (GGGGFG) peptide with different 193-nm UV laser irradiation time were shown in Figure S11. No obvious iodination Trp and Phe peptides could be observed.

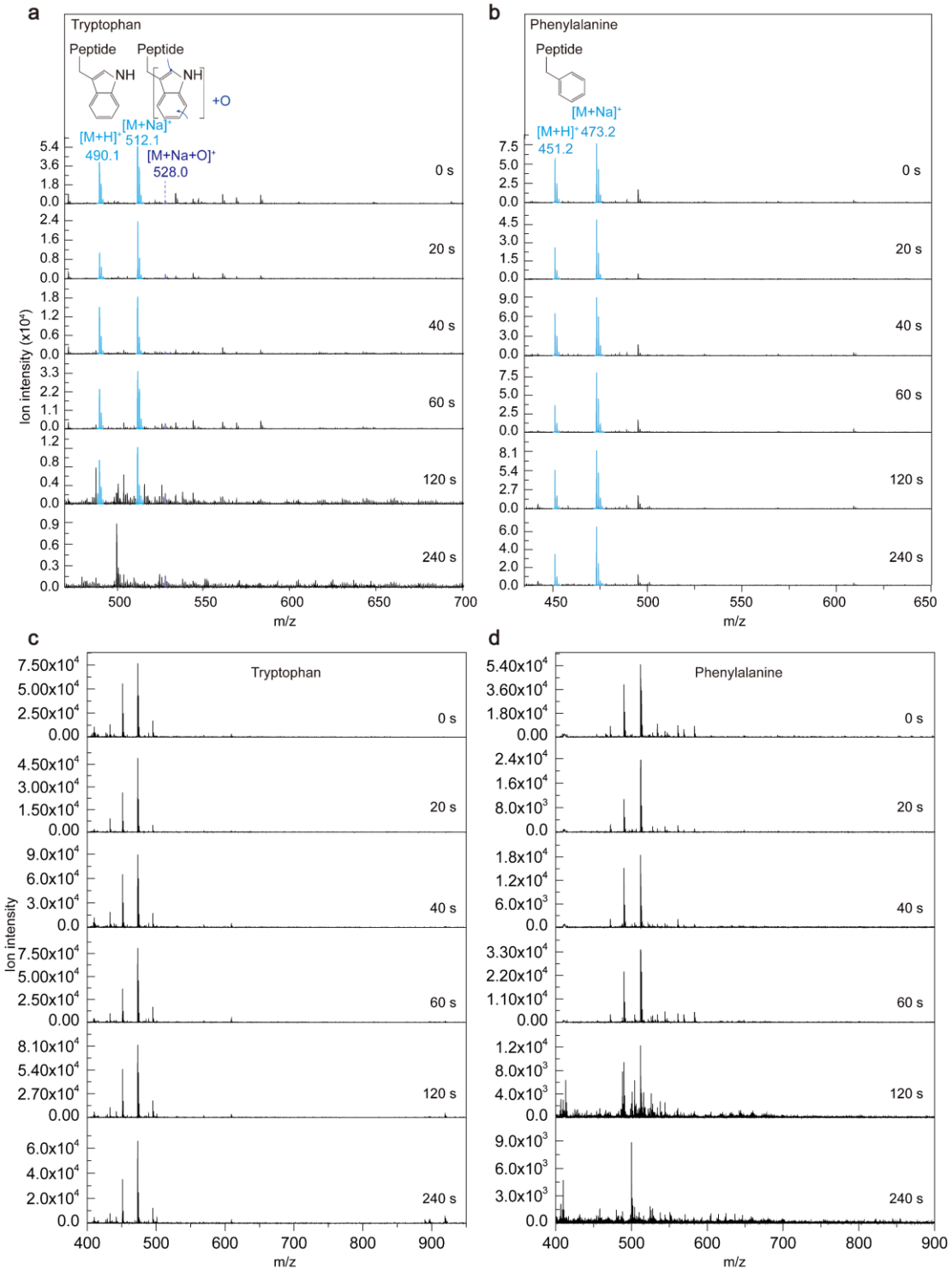
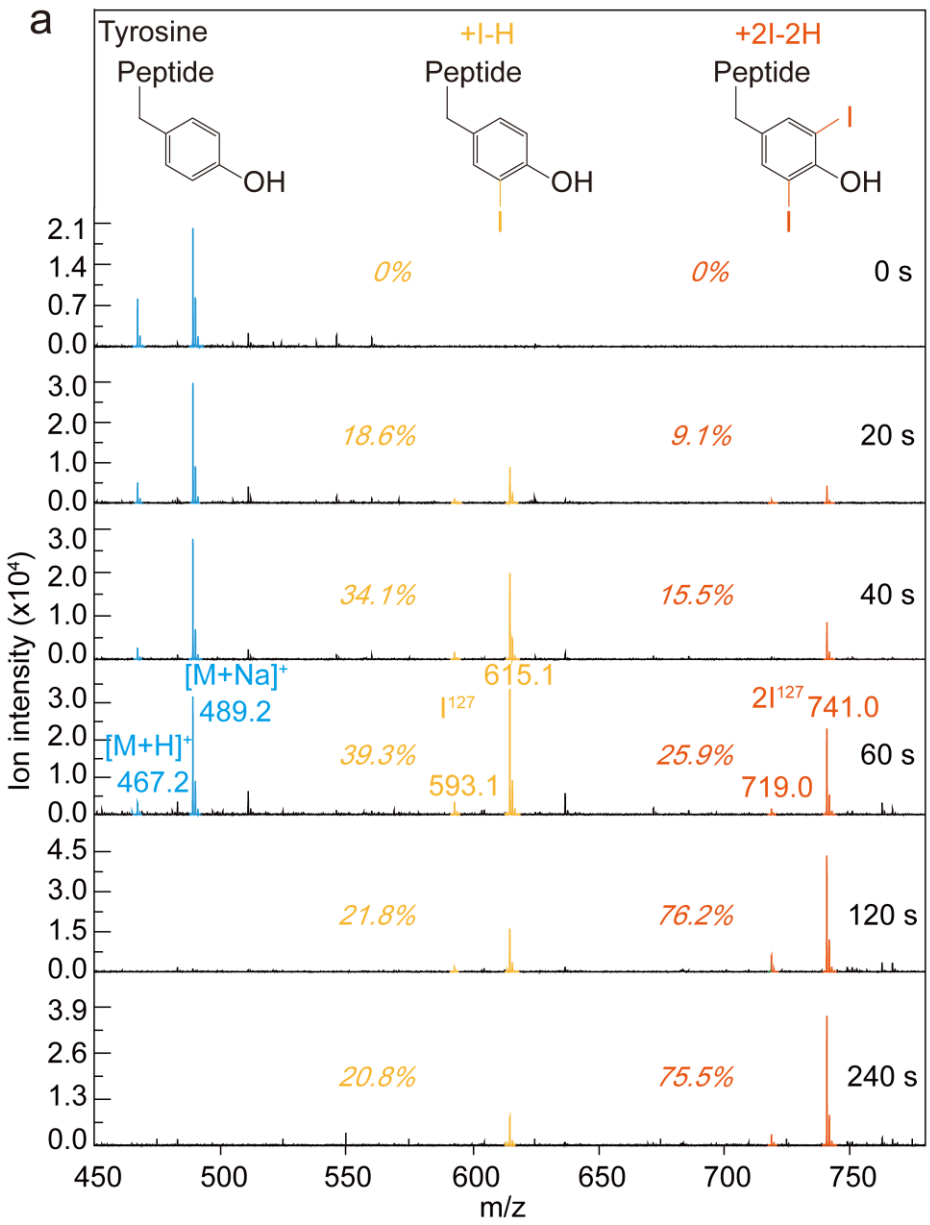


Figure S11. MALDI-TOF-MS characterization of peptide (a) GGGGWG and (b) GGGGFG and their original unprocessed spectrum (c, d) with different laser irradiation times of 0, 20, 40, 60, 120, and 240 s. The 193-nm UV laser was operated at 5 Hz, 5 mJ/pulse. The peptide samples were prepared in reaction solution with 150 mM I^- (10 mM Na_2HPO_4, 10 mM NaH_2PO_4, 150 mM NaI, pH 7.4).

Obviously, the iodination products of tyrosine peptide (GGGGYG) could be seen by molecular weight matching and the iodination efficiency increased along with the laser irradiation time (Figure S12). Products with 60 s irradiation were further subjected to LC-MS analyses and the mono- and di-iodination products were observed by molecular weight matching (Figure S13).



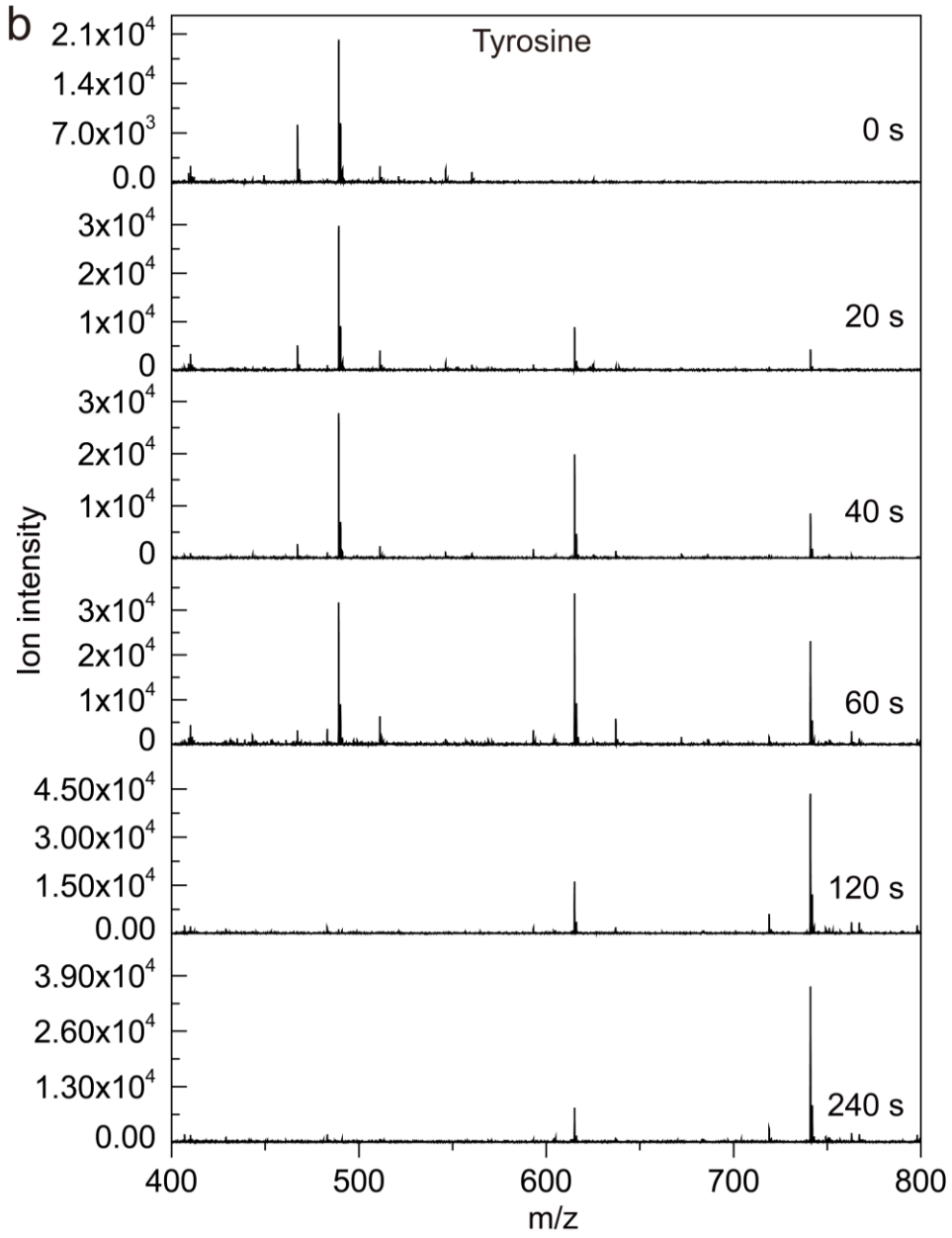


Figure S12. MALDI-TOF-MS characterization of peptide GGGGYG (a) and the original unprocessed spectrum (b) with different laser irradiation times of 0, 20, 40, 60, 120, and 240 s. The 193-nm UV laser was operated at 5 Hz, 5 mJ/pulse. The peptide samples were prepared in reaction solution with 150 mM I^- (10 mM Na_2HPO_4, 10 mM NaH_2PO_4, 150 mM NaI, pH 7.4).

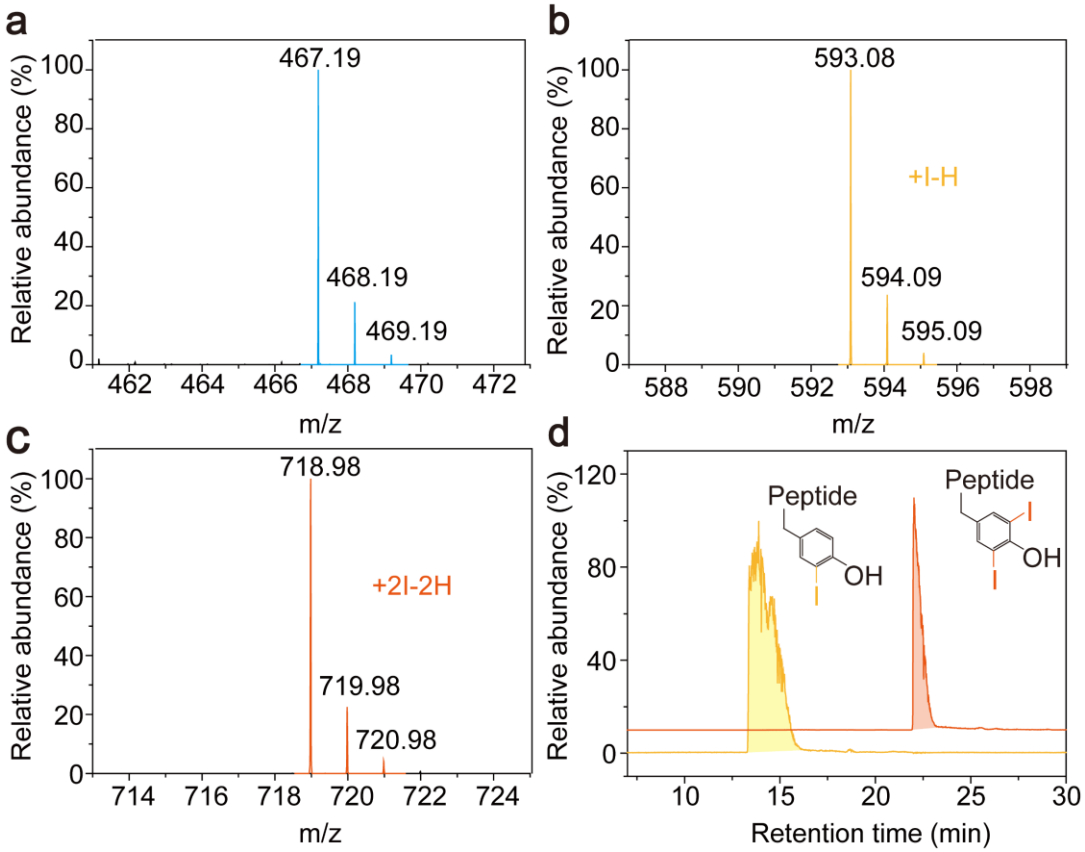


Figure S13. (a-c) The MS spectra from LC-MS analysis of peptide GGGGYG with 60 s irradiation, (a) the original form without laser irradiation, (b) the mono-iodination product, and (c) the di-iodination product. (d) The extracted product chromatographic peaks for the main products.

Besides, the iodination products of histidine peptide (GGGGHG) could be seen by molecular weight matching and the iodination efficiency increased along with the laser irradiation time (Figure S14). Products with 60 s irradiation were further subjected to LC-MS analyses and the mono- and di-iodination products were observed by molecular weight matching (Figure S15).

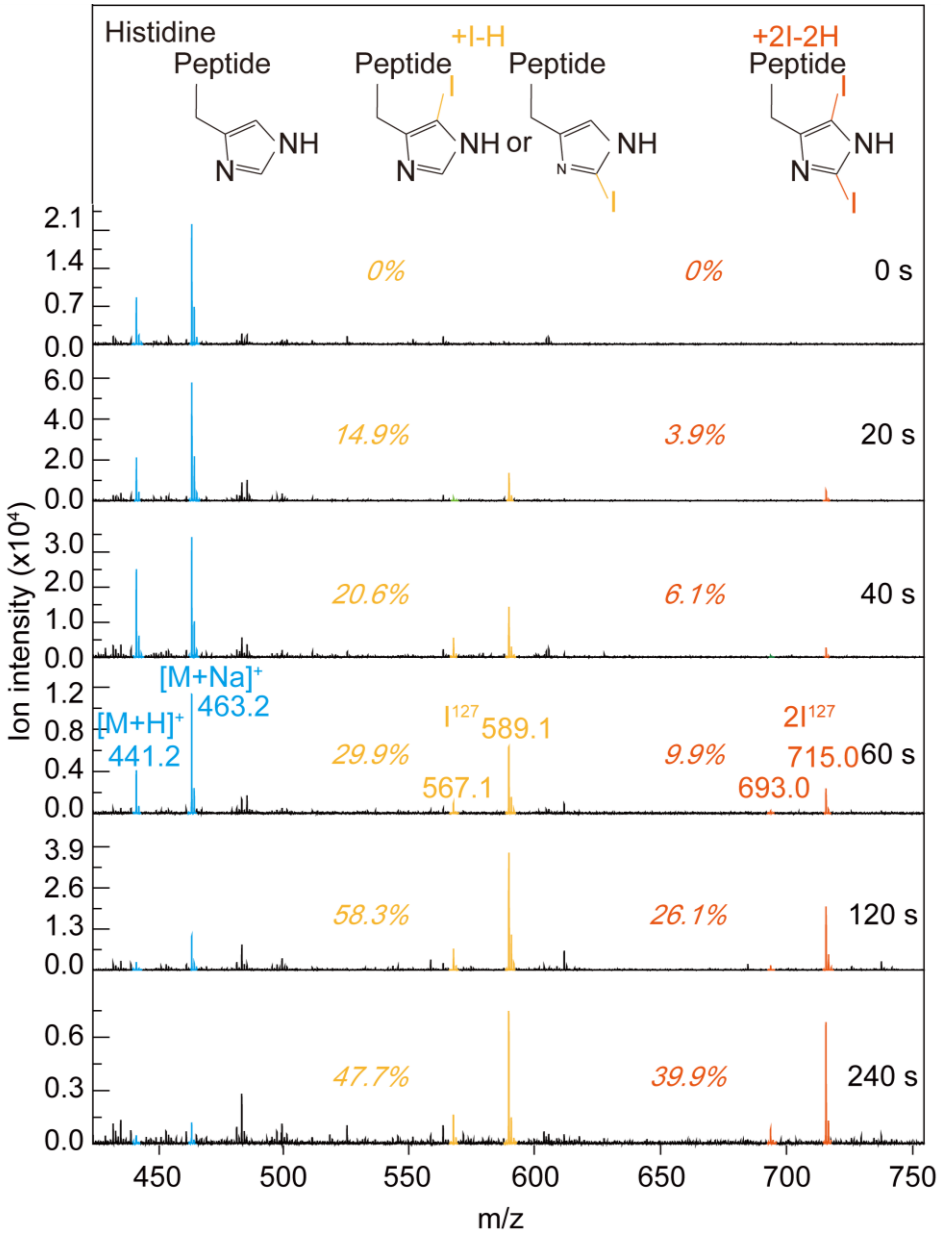


Figure S14. (a) MALDI-TOF-MS characterization of peptide GGGGHG with different laser irradiation times of 0, 20, 40, 60, 120, and 240 s. The 193-nm UV laser was operated at 5 Hz, 5 mJ/pulse. The peptide samples were prepared in reaction solution with 150 mM I^- (10 mM Na_2HPO_4, 10 mM NaH_2PO_4, 150 mM NaI, pH 7.4).

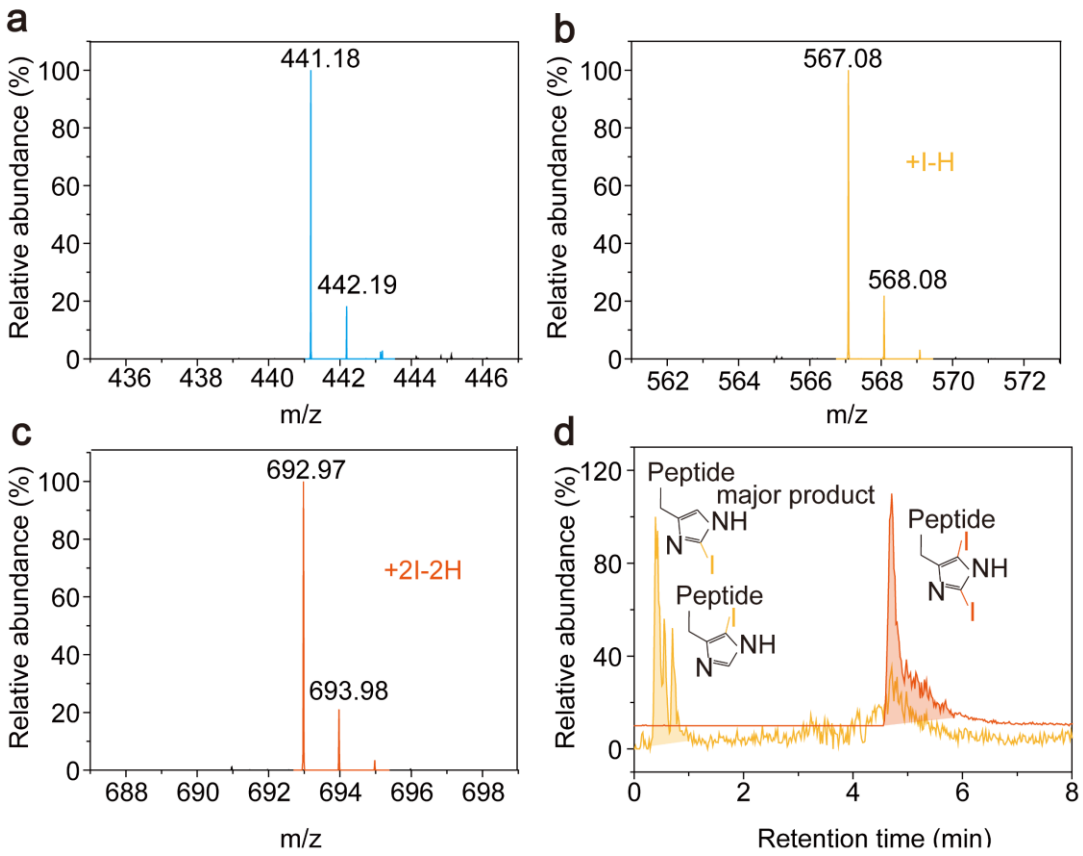


Figure S15. (b-d) The MS spectra from LC-MS analysis of peptide GGGGHG with 60 s irradiation, (b) the original form without laser irradiation, (c) the mono-iodination product, and (d) the di-iodination product. (d) The extracted product chromatographic peaks for the main products.

The proposed reaction pathways for photochemical halogenations

Monatomic halogen radicals (•Cl, •Br, •I) can be formed by photo-excited electron transfer from halide X^- to water[1]. Compared with F^- and Cl^-, Br^- and I^- have much stronger 193-nm photo-adsorption to generate abundant •Br and •I radicals with higher reactivity[2, 3]. The reduction potentials between halogen radicals and halide X^- further supported that direct photochemical bromination and iodination of aromatic species could be achieved and the iodination was more prone to occur[1]. The reaction solutions with F^- and Cl^- ions were also utilized in the 193-nm photochemical modification, yet no protein fluorination or chlorination was observed due to the difficulty in •F radical generation from photo-excitation of F^- without catalyst[4] and the low reactivity of the •Cl radicals[5]. Based on our experiment results, we speculated the major photochemical halogenation pathway as shown in Figure S16.

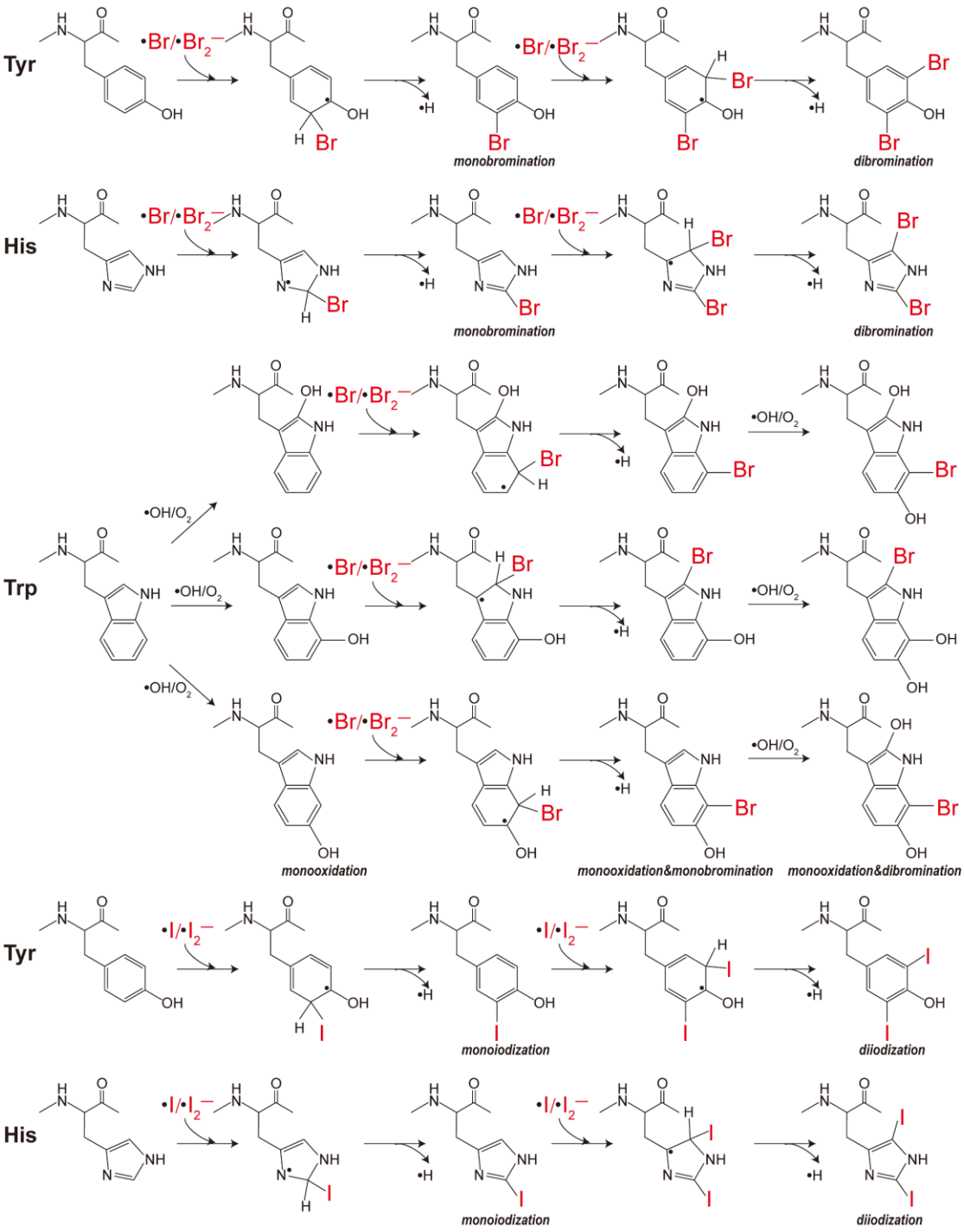


Figure S16. Potential photochemical bromination pathway for Tyr, His, Trp and iodination pathway for Tyr, His. Bromine and iodine atoms were highlighted in red.

Table S1. A summary of synthetic peptide bromination and iodination.

	Peptide	Modification	Modification rate with different irradiation time (%±stdev)					
			10 s	20 s	40 s	60 s	120 s	240 s
Photochemical bromination	GGGGYG	Mono-bromination (+Br-H)		10.6±0.6	10.8±2.9	11.6±1.5	8.7±0.4	20.0±1.1
		Di-bromination (+2Br-2H)		22.2±2.3	29.9±3.7	45.3±1.9	63.2±1.3	64.2±3.0
		Total		32.8±1.7	40.7±6.6	56.9±0.7	71.9±1.8	84.2±3.3
	GGGGHG	Mono-bromination (+Br-H)	17.6±1.9	28.7±1.8	38.5±8.0	48.2±10.3		
		Di-bromination (+2Br-2H)	0.2±0.4	1.5±0.6	2.8 ±0.3	2.9±0.4		
		Total	17.8±2.3	30.2±1.3	41.3±8.4	51.1±10.3		
	GGGGWG	Mono-oxidation (+O)		49.2±2.1	58.3±2.6	49.3±6.0	1.4±1.4	
		Mono-oxidation&-bromination (+O+Br-H)		6.8±0.6	11.3±5.1	47.2±5.7	20.6±2.6	20.2±5.8
		Di-oxidation&mono-bromination (+2O+Br-H)				3.5±1.0	62.9±4.9	54.8±5.2
		Mono-oxidation&di-bromination (+O+2Br-2H)					15.1±1.7	25.0±0.7
		Total		56.0±2.7	69.6±5.5	100±0	100±0	100±0
	GGGGFG	Can not be brominated						
	Peptide	**Modification**	**Modification rate with different irradiation time (%±stdev)**					
			10 s	**20 s**	**40 s**	**60 s**	**120 s**	**240 s**
Photochemical iodination	GGGGYG	Mono-iodination (+I-H)		18.6±1.6	34.1±2.5	39.3±3.5	21.8±2.9	20.8±3.2
		Di-iodination (+2I-2H)		9.1±1.7	15.5±1.4	25.9±2.1	76.2±3.3	75.5±4.5
		Total		27.7±3.3	49.6±3.6	65.2±4.6	98.0±0.4	96.3±1.3
	GGGGHG	Mono-iodination (+I-H)		14.9±4.9	20.6±3.1	29.9±3.1	58.3±2.1	47.7±1.6
		Di-iodination (+2I-2H)		3.9±0.1	6.1±0.3	9.9±2.7	26.1±6.4	39.9±3.7
		Total		18.8±4.9	26.7±3.4	39.8±0.4	84.4±4.5	87.6±2.3
	GGGGWG	Can not be iodinated						
	GGGGFG							

Photochemical halogenations of mouse liver digest

The typical MS/MS spectra of the brominated and iodinated residues were shown in Figure S17 and Figure S19, respectively. The bromination and iodination statistical results of the complex mouse liver digest were shown in Figure S18 and Figure S20, respectively. Interestingly, the Tyr mono-iodination preferred single-pulse irradiation, while superior di-iodination results were achieved in direct irradiation device with multiple-pulse irradiation, providing more flexibility and adaptability. In addition, no sequence specificity was observed in both photochemical bromination and iodination through WebLogo motif analyses[6]. The sequence motif analyses of the bromination residues Y, H, W and iodination residues Y, H were shown in

Figure S21.

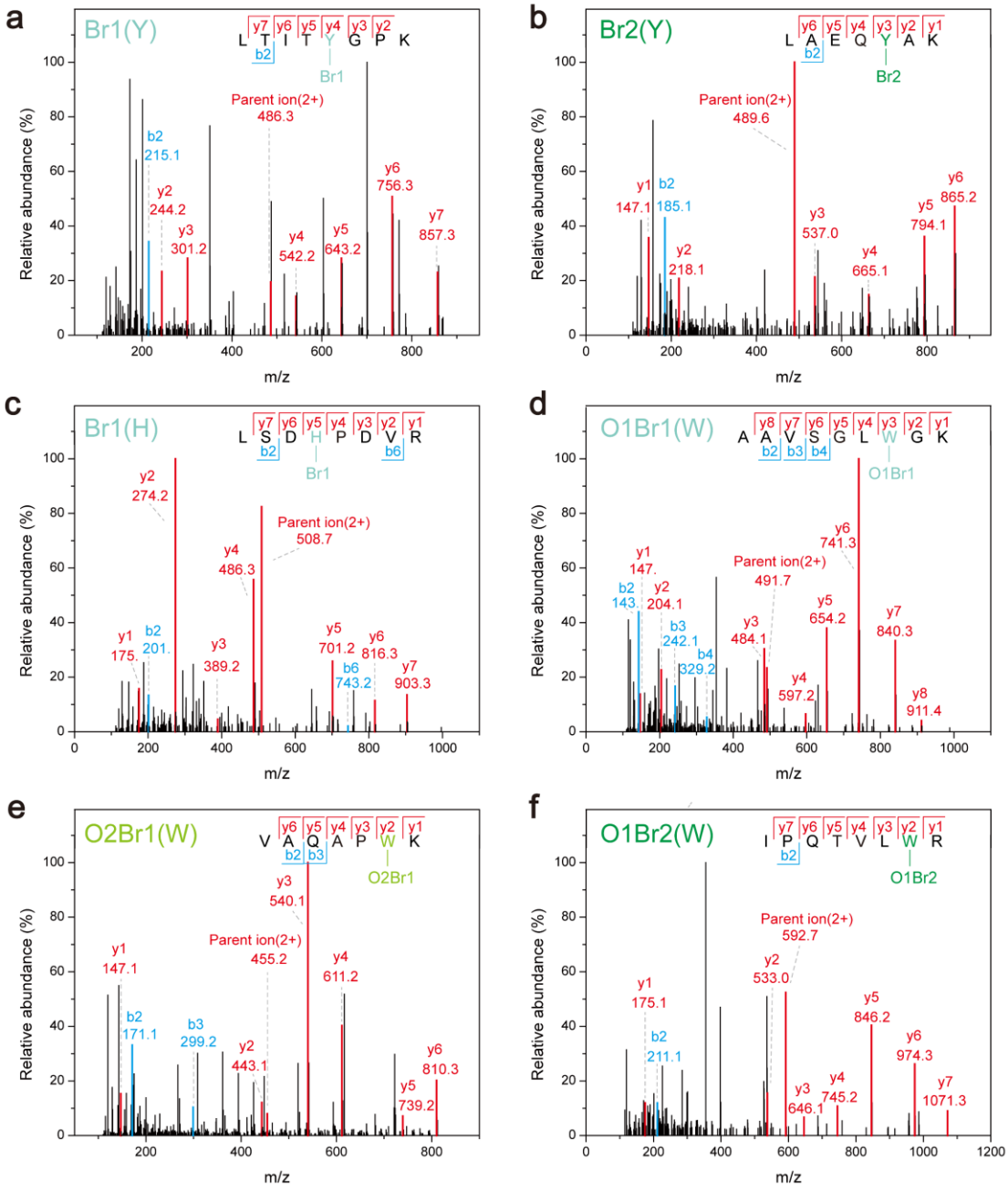


Figure S17. (a-f) Randomly selected examples of MS/MS spectra with confident peptide sequence and bromination sites determination.

a

Modifications	Number of brominated peptides		
	150 pulses in direct irradiation device	300 pulses in direct irradiation device	Single-pulsed irradiation capillary reactor
Br1(Y)	1		3
Br2(Y)	1		210
Br1(H)	1		585
O1Br1(W)	13	2	9
O2Br1(W)	129	49	19
O1Br2(W)			14
Total	145	51	770

b

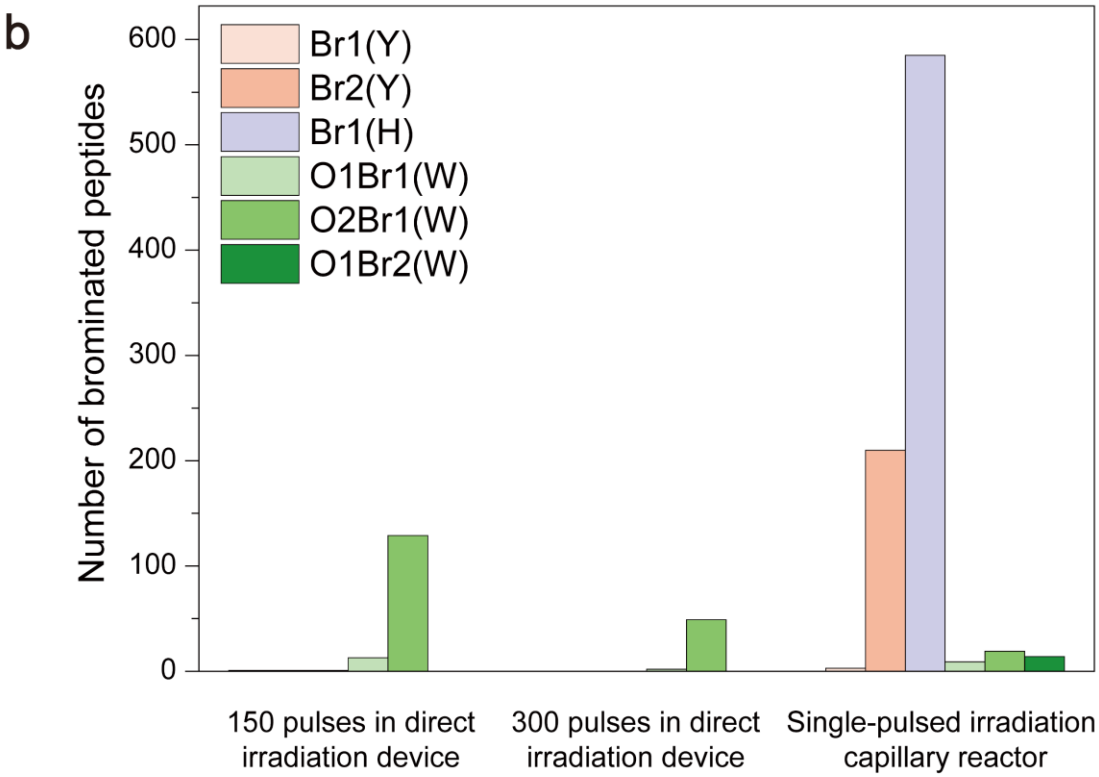


Figure S18. The statistical results of bromination residues with 150/300 pulses laser irradiation in the direct irradiation device or 1 pulse of laser irradiation in the capillary reactor.

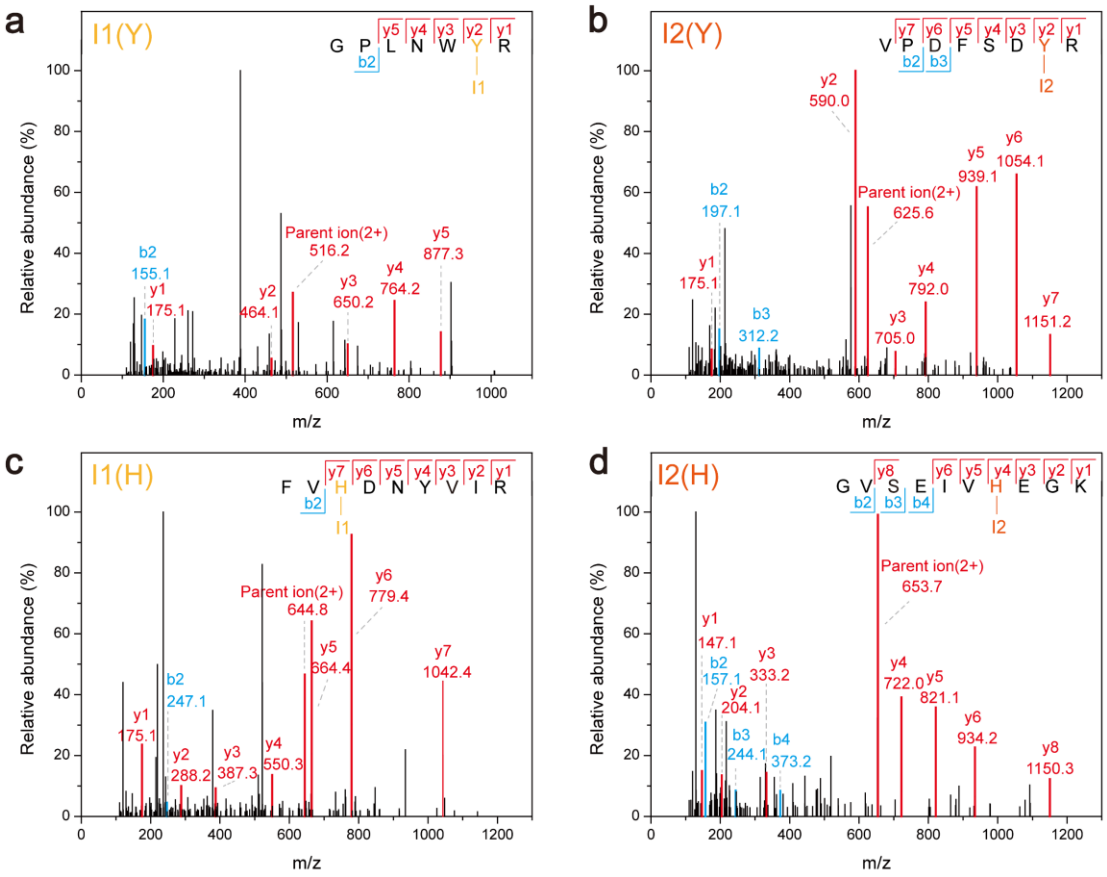


Figure S19. Randomly selected examples of MS/MS spectra with confident peptide sequence and iodination sites determination.

a

Modifications	Number of iodinated peptides		
	150 pulses in direct irradiation device	300 pulses in direct irradiation device	Single-pulsed irradiation capillary reactor
I1(Y)	17		1348
I2(Y)	950	920	326
I1(H)	462	2	174
I2(H)	863	691	25
Total	2194	1539	1812

b

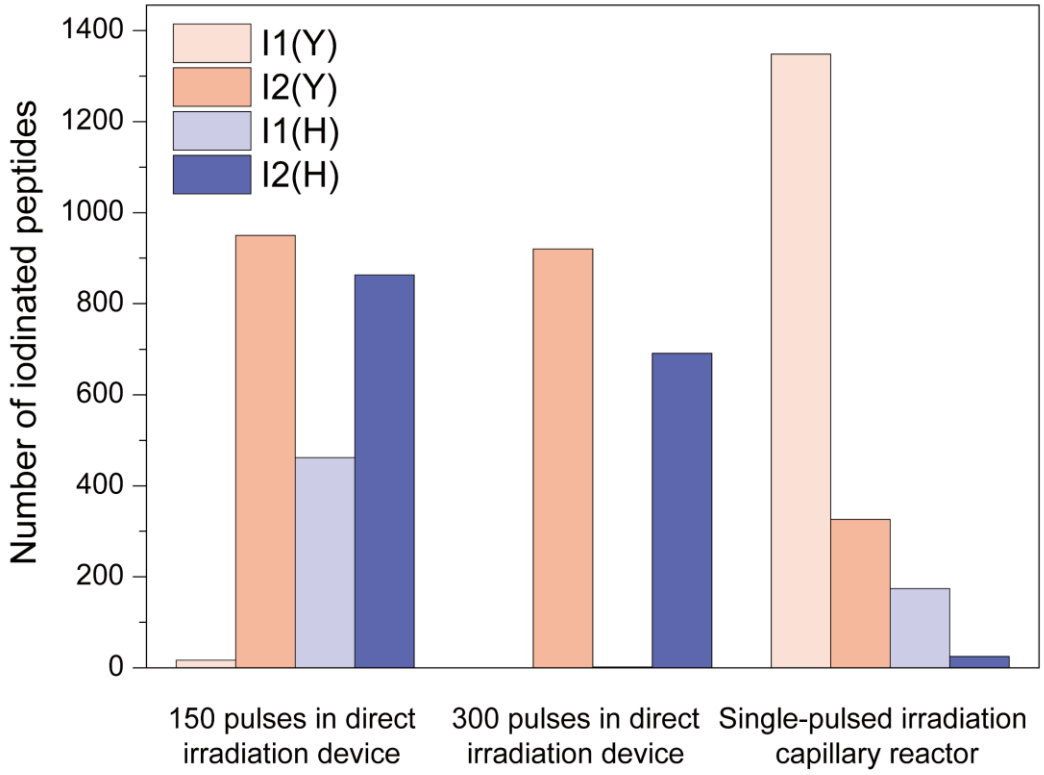


Figure S20. The statistical results of iodination residues with 150/300 pulses laser irradiation in the direct irradiation device or 1 pulse of laser irradiation in the capillary reactor.

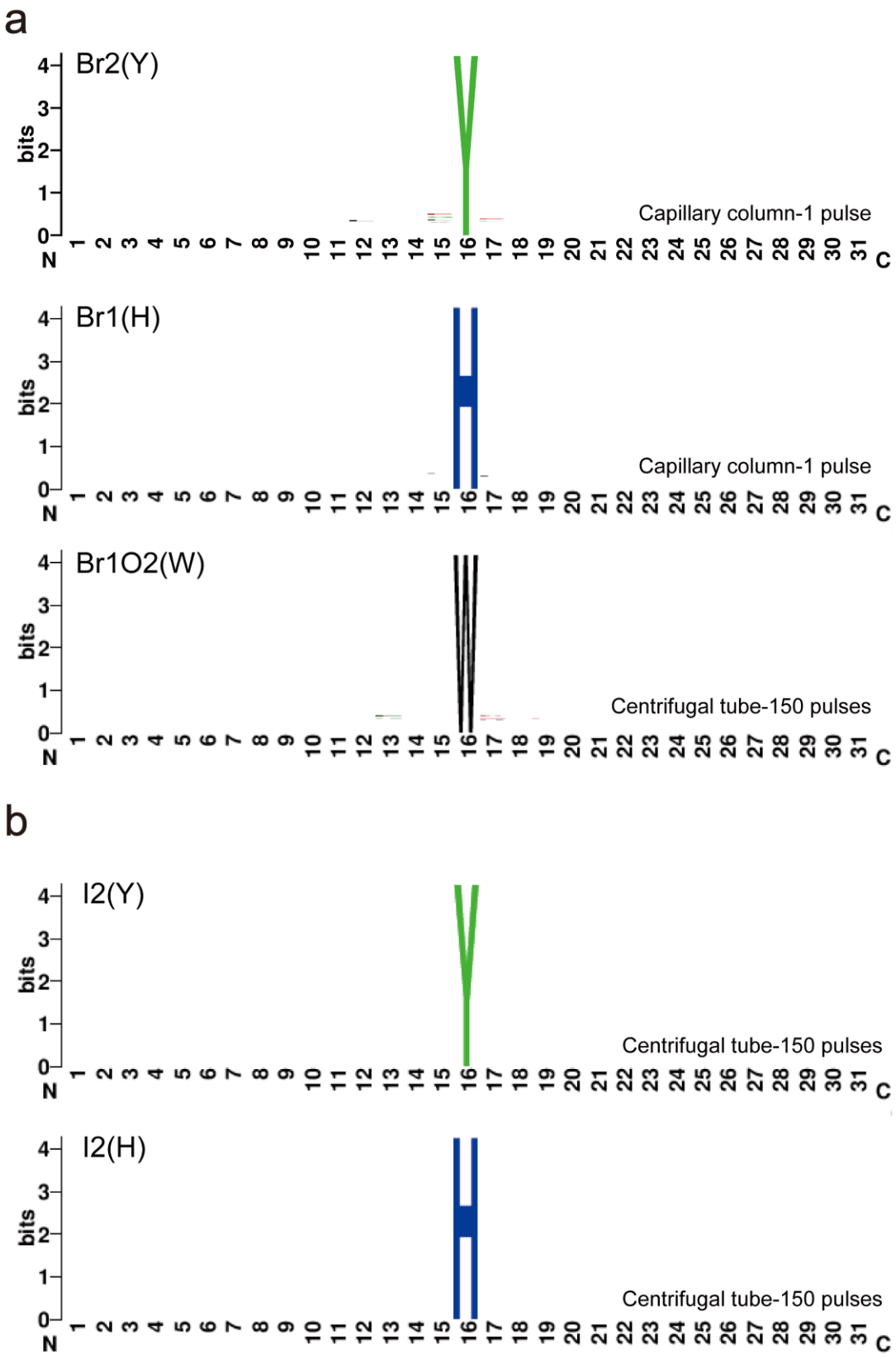


Figure S21. Sequence motif analyses of the (a) brominated and (b) iodinated residues.

The influences of photochemical halogenations on native protein structure

To explore the potential effects of photochemical halogenations on native protein structure, 0, 100, 150 pulses of 193-nm UV laser irradiation were applied to native myoglobin samples in ice bath. Almost coincident CD spectra were observed for myoglobin with and without bromination as shown in Figure S22, indicating the myoglobin native structure was largely remained. For iodination experiments, abundant I_2 could be generated during the photochemical iodination process and the myoglobin CD signals were mainly decreased or disappeared.

The charge state distributions of native and halogenated myoglobin samples were further characterized in native MS[7-9]. Similar charge state distributions were observed for native, brominated, and iodinated myoglobin, implying the compact protein structure could be largely retained in the photochemical halogenation without inducing structure relaxing and charge state increasing[10]. The MS deconvolution results demonstrated the success of multiple photochemical halogenations on the myoglobin samples as shown in Figure S23.

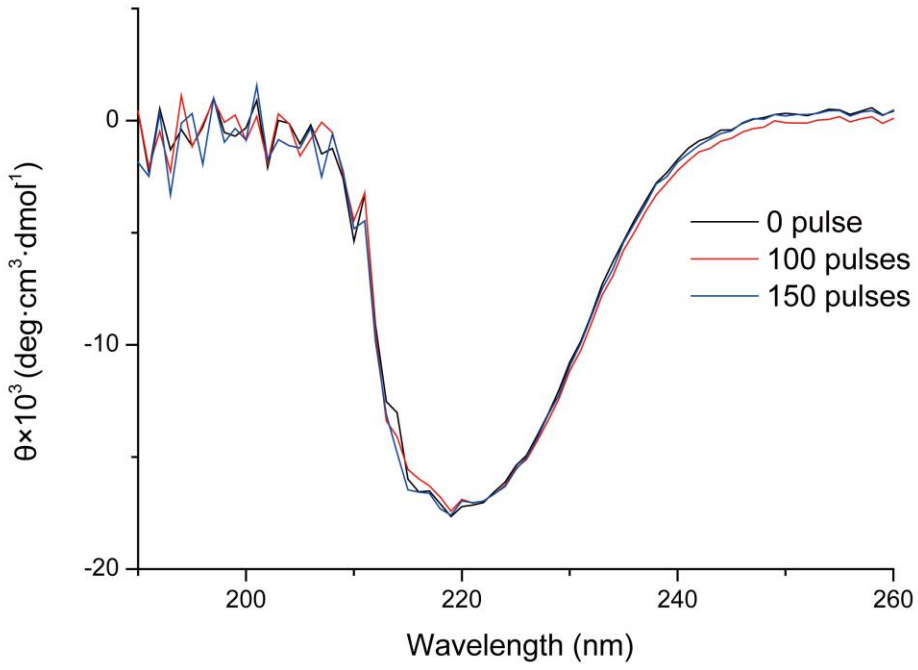


Figure S22. CD spectra of myoglobin with 0, 100, 150 pulses of laser irradiation for photochemical bromination. The 193-nm UV laser was operated at 0.5 Hz, 10 mJ/pulse and the protein samples were kept in ice bath.

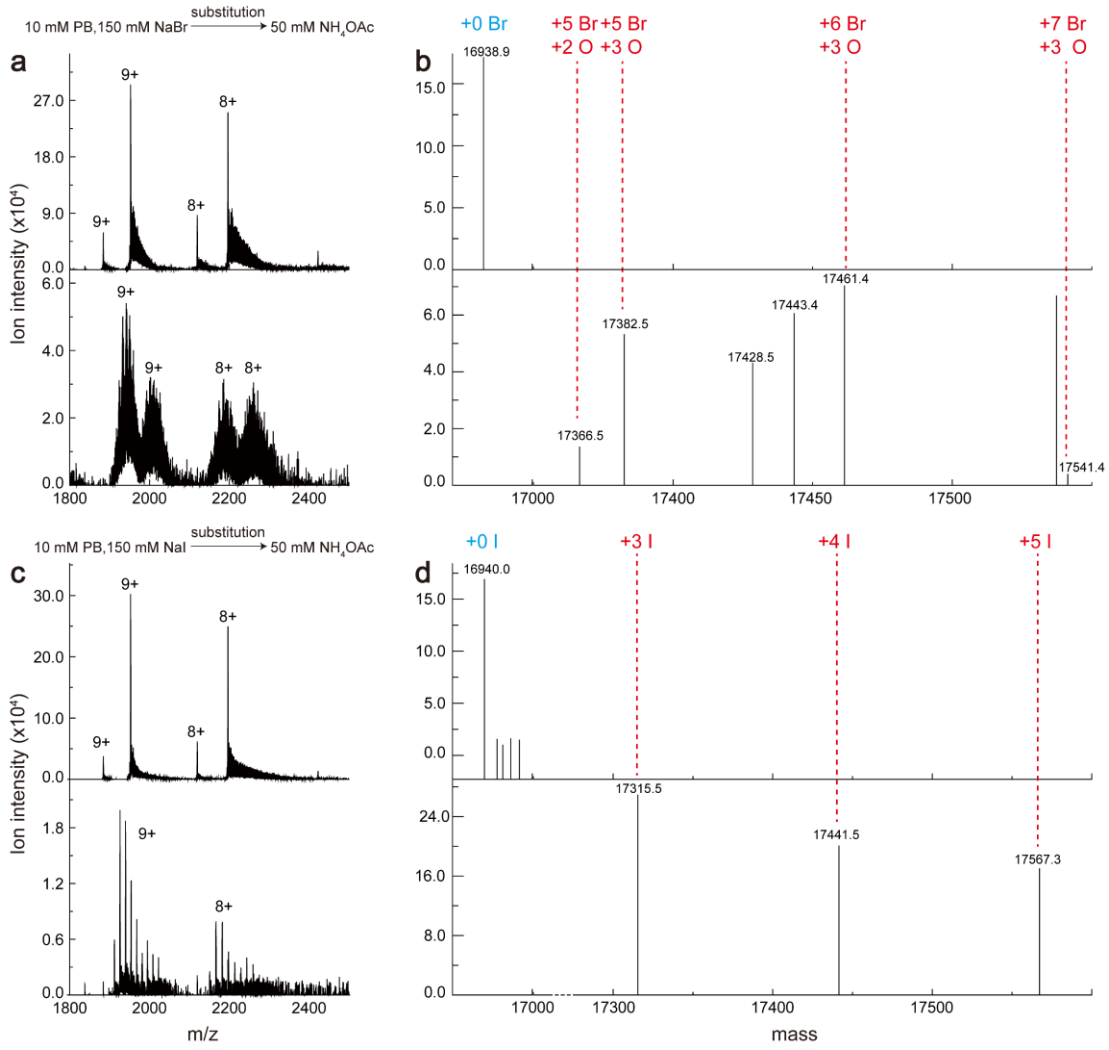


Figure S23. (a, c) Native MS characterization of photochemically (a) brominated and (c) iodinated myoglobin. Upper: native form, lower: halogenated form. The phosphate reaction solutions were changed to 50 mM NH_4OAc by ultrafiltration before the native MS characterization. (b, d) The deconvolution MS spectra of native, brominated, and iodinated forms of myoglobin. The decrease of MS intensity could be attributed the incomplete solvent replacement.

Photochemical halogenation of native BSA

After database searching by MaxQuant, all brominated and iodinated peptides in the 193-nm UV photochemical halogenation of native BSA were listed in Table S2 and Table S3, respectively. Besides, all brominated and iodinated residues were listed in Table S4.

Table S2. The photochemical bromination sequences and bromination rates of native BSA with 1 pulse of 193-nm UV laser irradiation. Experiments were performed in 3 replicates.

		BSA
Residue	**Sequence**	**MR%±stdev**
11-20	FKDLGEEHFK	18.6±32.2
52-76	TCVADESHAGCEKSLHTLFGDELCK	8.3±7.2
65-76	SLHTLFGDELCK	0.3±0
65-81	SLHTLFGDELCKVASLR	2.2±1.3
94-106	QEPERNECFLSHK	65±56.4
94-114	QEPERNECFLSHKDDSPDLPK	2.2±3.8
99-106	NECFLSHK	98.6±2
99-114	NECFLSHKDDSPDLPK	11.9±20.6
137-143	YLYEIAR	6.2±7.3
205-217	FGERALKAWSVAR	100±0
209-217	ALKAWSVAR	90.4±2
225-242	AEFVEVTKLVTDLTKVHK	53.6±46.4
262-273	YICDNQDTISSKLK	92.1±2.2
274-294	LKECCDKPLLEKSHCIAEVEK	100±0
286-316	SHCIAEVEKDAIPENLPPLTADFAEDKDVCK	23.3±40.4
336-347	RHPEYAVSVLLR	8.1±9.2
348-362	LAKEYEATLEECCAK	4.9±8.4
351-362	EYEATLEECCAK	13.4±23.2
363-377	DDPHACYSTVFDKLK	94.1±10.2
378-387	HLVDEPQNLIK	9.5±2.6
389-409	QNCDQFEKLGEYGFQNALIVR	0.3±0.6
397-409	LGEYGFQNALIVR	32.4±37.4
459-465	LCVLHEK	57.5±5.2
459-471	LCVLHEKTPVSEK	14.7±25.2
459-474	LCVLHEKTPVSEKVTK	1.4±2.3
475-504	CCTESLVNRRPCFSALTPDETYVPKAFDEK	4.4±4.2
484-499	RPCFSALTPDETYVPK	3.5±6
500-520	AFDEKLFTFHADICTLPDTEK	0.9±1.6
505-520	LFTFHADICTLPDTEK	33.3±57.7
505-523	LFTFHADICTLPDTEKQIK	1.5±2.6
525-537	KQTALVELLKHKPK	3.6±4.7
525-537	QTALVELLKHKPK	12.8±11.1

Table S3. The photochemical iodination sequences and iodination rates of native BSA with 1 pulse of 193-nm UV laser irradiation. Experiments were performed in 3 replicates.

		BSA
Residue	Sequence	MR%±stdev
1-10	DTHKSEIAHR	87.6±4.8
1-12	DTHKSEIAHRFK	98.7±0.8
5-12	SEIAHRFK	100±0
5-20	SEIAHRFKDLGEEHFK	100±0
11-20	FKDLGEEHFK	17.3±0.3
13-20	DLGEEHFK	27.2±3.2
42-64	LVNELTEFAKTCVADESHAGCEK	22.8±5.8
52-81	TCVADESHAGCEKSLHTLFGDELCKVASLR	47.2±2.8
65-76	SLHTLFGDELCK	47.3±9.6
65-81	SLHTLFGDELCKVASLR	41.7±17
94-106	QEPERNECFLSHK	14.4±0.3
94-114	QEPERNECFLSHKDDSPDLPK	40.7±2
99-106	NECFLSHK	14.5±5
99-114	NECFLSHKDDSPDLPK	24.4±2.5
137-143	YLYEIAR	20.9±26.8
225-242	AEFVEVTKLVTDLTKVHK	25.8±6
233-242	LVTDLTKVHK	22.9±8.3
233-256	LVTDLTKVHKECCHGDLLECADDR	73.7±3.4
240-261	VHKECCHGDLLECADDRADLAK	56.7±3.4
262-273	YICDNQDTISSK	97.5±0.6
262-275	YICDNQDTISSKLK	98.9±0.1
262-285	YICDNQDTISSKLKECCDKPLLEK	99.6±0.7
286-294	SHCIAEVEK	82.4±3.4
286-312	SHCIAEVEKDAIPENLPPLTADFAEDK	60.7±13.5
286-316	SHCIAEVEKDAIPENLPPLTADFAEDKDVCK	61.9±7.6
295-322	DAIPENLPPLTADFAEDKDVCKNYQEAK	55.8±4.6
317-335	NYQEAKDAFLGSFLYEYSR	9.3±2
336-347	RHPEYAVSVLLR	98.8±0.6
337-346	HPEYAVSVLLR	97.8±3.1
348-362	LAKEYEATLEECCAK	79.3±3.3
351-377	EYEATLEECCAKDDPHACYSTVFDKLK	99.3±0.8
363-377	DDPHACYSTVFDKLK	93.4±5.3
378-387	HLVDEPQNLIK	37.9±1.6
378-396	HLVDEPQNLIKQNCDQFEK	55.4±0.5
378-409	HLVDEPQNLIKQNCDQFEKLGEYGFQNALIVR	62.8±13.2
389-409	QNCDQFEKLGEYGFQNALIVR	83.7±2.2
397-409	LGEYGFQNALIVR	90.4±1.9
410-427	YTRKVPQVSTPTLVEVSR	100±0
436-458	CCTKPESERMPCTEDYLSLILNR	82.4±6.3
445-458	MPCTEDYLSLILNR	71.2±25
459-465	LCVLHEK	14.4±0.8
459-471	LCVLHEKTPVSEK	12.5±1.5
459-474	LCVLHEKTPVSEKVTK	13.7±2.8
475-499	CCTESLVNRRPCFSALTPDETYVPK	81.1±5.4
475-504	CCTESLVNRRPCFSALTPDETYVPKAFDEK	86.3±7.4
484-499	RPCFSALTPDETYVPK	82.5±8.5
484-504	RPCFSALTPDETYVPKAFDEK	85.5±2
500-520	AFDEKLFTFHADICTLPDTEK	85.3±4.4
500-523	AFDEKLFTFHADICTLPDTEKQIK	90.5±1.4
505-520	LFTFHADICTLPDTEK	86.4±2.3
505-523	LFTFHADICTLPDTEKQIK	81.6±4.6
505-524	LFTFHADICTLPDTEKQIKK	92.1±1.3
525-537	KQTALVELLKHKPK	57.6±7.1
525-537	QTALVELLKHKPK	35.1±3.3
525-544	QTALVELLKHKPKATEEQLK	16.4±3.2
534-544	HKPKATEEQLK	50.1±43.4

Table S4. The photochemical halogenation sites of native BSA characterized in LC-MS/MS analysis.

Site	SASA from PDB data ($Å^2$)	Modification	
		Bromination	Iodination
H3	100.26	Y	Y
H9	62.651	Y	Y
H18	34.797	Y	Y
Y30	5.890	N	N
H39	0.994	N	N
H59	82.645	Y	Y
H67	47.289	Y	Y
Y84	34.206	Y	Y
H105	70.321	Y	Y
Y137	27.583	Y	Y
Y139	9.047	Y	Y
W213	44.509	Y	
H241	7.032	Y	Y
H246	87.770	Y	Y
Y262	63.733	Y	Y
H287	25.041	Y	Y
Y318	8.118	N	Y
Y331	4.197	N	N
Y333	31.467	N	Y
H337	50.294	Y	Y
Y340	42.293	Y	Y
Y352	1.736	Y	Y
H366	36.514	Y	Y
Y369	6.929	Y	Y
H378	56.758	Y	Y
Y400	87.779	Y	Y
Y410	21.918	Y	Y
Y451	72.461	Y	Y
H463	11.242	Y	Y
Y496	47.028	Y	Y
H509	134.985	Y	Y
H534	61.480	Y	Y

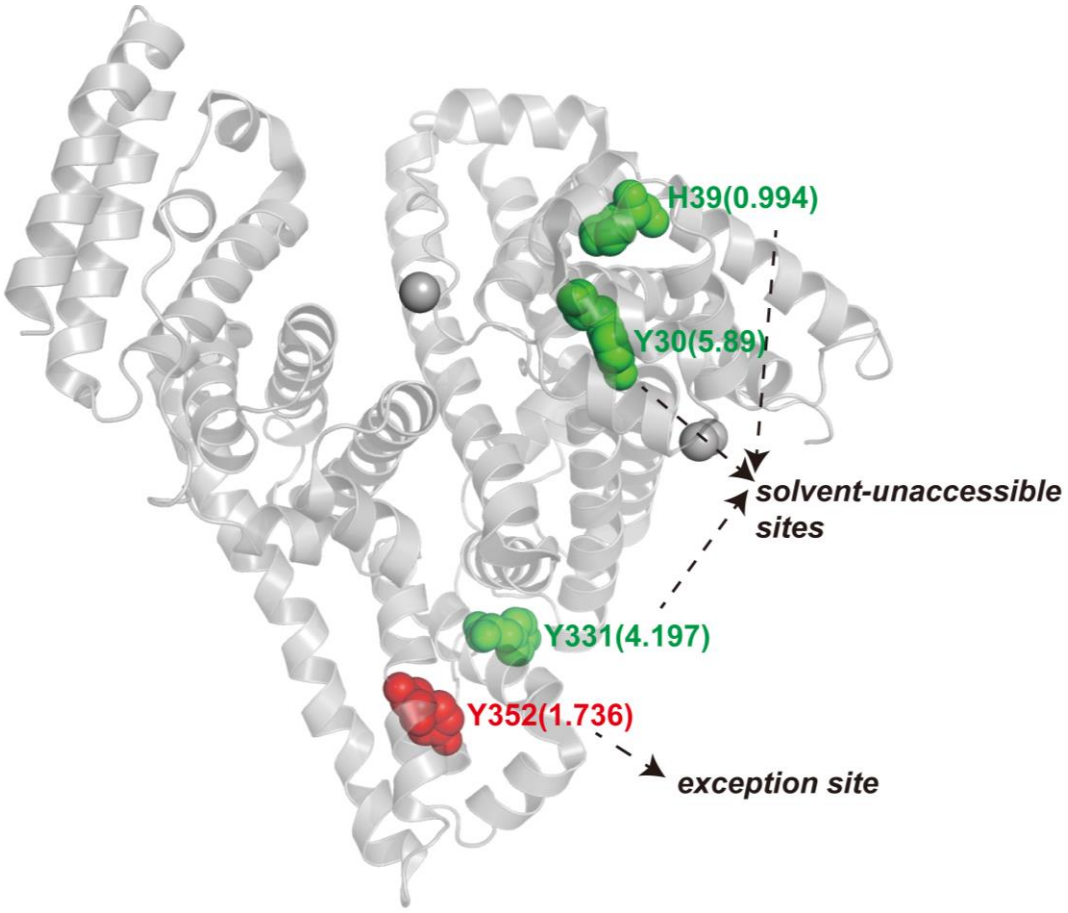


Figure S24. The locations of aromatic reisudes with low solvent-accessible surface areas (SASAs). Green, wihtout photochemical halogenation; Red, with photochemical halogenation.

Probing protein-drug interaction by 193-nm UV photochemical modifications

The human serum albumin (HSA) is the most abundant protein in human plasma which contains three homologous helical domains (I–III) and can bind a remarkable wide range of endogenous metabolites and small-molecule drugs. Elucidating the binding mechanisms of HSA with administered drugs has important physiological significance for solving the problems associated with the drug absorption, distribution, metabolism, and elimination (ADME)[11]. Previous X-ray diffraction (XRD) crystallographic analyses of HSA-drug complexes reveal that there consist of two primary drug-binding sites locating in subdomains IIA and IIIA[12]. After database searching by MaxQuant, all brominated and iodinated peptides in the 193-nm UV photochemical modifications of HSA with or without warfarin combination were listed in Table S5 and Table S6, respectively. Peptide regions of HSA with significant differences in photochemical halogenation upon warfarin combination were listed in Table S7. Besides, a

summary table of protein digest and native protein halogenation were listed in Table S8.

Table S5. The photochemical bromination sequences and bromination rates of HSA and HSA+warfarin with 1 pulse of 193-nm UV laser irradiation. All experiments were performed in 3 replicates.

		HSA	HSA+warfarin
Residue	Sequence	MR%±stdev	MR%±stdev
1-10	DAHKSEVAHR	90.4±4.6	85.3±5.3
1-12	DAHKSEVAHRFK	72.4±10.1	74.6±5.9
52-73	TCVADESAENCDKSLHTLFGDK	1.2±2	1±1.7
52-81	TCVADESAENCDKSLHTLFGDKLCTVATLR	7.2±2.7	4.9±2.3
65-73	SLHTLFGDK	0.9±1.6	1.8±3.2
65-81	SLHTLFGDKLCTVATLR	0.1±0.1	0.1±0.1
65-93	SLHTLFGDKLCTVATLRETYGEMADCCAK	0.6±0.5	0.3±0.3
94-106	QEPERNECFLQHK	12±9	3.4±4.6
99-114	NECFLQHKDDNPNLPR	10±14.2	16.3±18.2
115-136	LVRPEVDVMCTAFHDNEETFLK	0.2±0.3	0.1±0.3
115-137	LVRPEVDVMCTAFHDNEETFLKK	0.4±0.4	0.2±0.2
115-144	LVRPEVDVMCTAFHDNEETFLKKYLYEIAR	7.9±13.3	0.2±0.1
137-145	KYLYEIARR	1.3±1.3	0.8±1.3
138-144	YLYEIAR	0.1±0	0.1±0
138-145	YLYEIARR	2.5±4.4	0±0
145-159	RHPYFYAPELLFFAK	0.2±0.2	0.3±0.5
145-160	RHPYFYAPELLFFAKR	1.1±1	0.8±0.9
161-174	YKAAFTECCQAADK	0.1±0.1	0±0
206-218	FGERAFKAWAVAR	59.5±14.9	11.5±9.2
210-218	AFKAWAVAR	44.1±8.8	16.3±4
234-262	LVTDLTKVHTECCHGDLLECADDRADLAK	0.5±0.8	0±0
241-262	VHTECCHGDLLECADDRADLAK	0±0	0.2±0.2
258-274	ADLAKYICENQDSISSK	33.4±14	12.3±11.4
263-274	YICENQDSISSK	0±0	0.1±0.1
263-276	YICENQDSISSKLK	6.5±11.3	5.3±1.9
277-313	ECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESK	1.1±1.9	0±0
287-323	SHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAK	0.1±0.1	0±0
314-336	DVCKNYAEAKDVFLGMFLYEYAR	0±0.1	0±0
337-348	RHPDYSVVLLLR	0±0	1±0.9
338-348	HPDYSVVLLLR	1.2±0.5	0.2±0.3
360-389	CCAAADPHECYAKVFDEFKPLVEEPQNLIK	0±0	0±0
390-410	QNCELFEQLGEYKFQNALLVR	0.1±0	0±0
446-475	MPCAEDYLSVVLNQLCVLHEKTPVSDRVTK	0.6±1.1	0±0
525-538	KQTALVELVKHKPK	2.2±2.5	2±3.5
526-538	QTALVELVKHKPK	1±1.7	0±0
526-541	QTALVELVKHKPKATK	0.8±1.3	1±0.8

Table S6. The photochemical iodination sequences and iodination rates of HSA and HSA+warfarin with 1 pulse of 193-nm UV laser irradiation. All experiments were performed in 3 replicates.

		HSA	HSA+warfarin
Residue	Sequence	MR%±stdev	MR%±stdev
5-12	SEVAHRFK	87 ± 4.4	90.8 ± 5.4
52-81	TCVADESAENCDKSLHTLFGDKLCTVATLR	8.1 ± 2.2	4.8 ± 1.4
65-73	SLHTLFGDK	18.9 ± 1.6	13.2 ± 1.7
65-81	SLHTLFGDKLCTVATLR	3.5 ± 0.2	4 ± 1.2
94-106	QEPERNECFLQHK	5.1 ± 1	4 ± 0.7
94-114	QEPERNECFLQHKDDNPNLPR	12.6 ± 2.8	12.2 ± 2.3
99-106	NECFLQHK	5.2 ± 2.4	4.4 ± 0.9
99-114	NECFLQHKDDNPNLPR	8.4 ± 2.8	7.3 ± 1.4
115-136	LVRPEVDVMCTAFHDNEETFLK	0.3 ± 0.1	0.3 ± 0.1
115-137	LVRPEVDVMCTAFHDNEETFLKK	3.6 ± 1.6	2.4 ± 0.7
137-144	KYLYEIAR	84.3 ± 3	81.7 ± 2.2
138-144	YLYEIAR	69.7 ± 3.6	64.1 ± 5.8
138-145	YLYEIARR	84.9 ± 4.3	81.6 ± 1.9
146-160	HPYFYAPELLFFAKR	35.1 ± 15	23.8 ± 22.7
161-181	YKAAFTECCQAADKAACLLPK	37.2 ± 9.1	25.7 ± 5.7
234-262	LVTDLTKVHTECCHGDLLECADDRADLAK	21.6 ± 13.1	13.8 ± 11
263-274	YICENQDSISSK	81.7 ± 0.8	77.9 ± 2.3
263-276	YICENQDSISSKLK	73 ± 22	67.8 ± 10.4
263-286	YICENQDSISSKLKECCEKPLLEK	74.1 ± 23.1	73.5 ± 0.5
287-313	SHCIAEVENDEMPADLPSLAADFVESK	1.9 ± 0.5	1.4 ± 0.5
314-323	DVCKNYAEAK	79.7 ± 8.5	79.7 ± 5.4
337-348	RHPDYSVVLLLR	46.2 ± 22.6	67 ± 5.6
338-348	HPDYSVVLLLR	56.4 ± 17	68.7 ± 6.4
349-359	LAKTYETTLEK	40.9 ± 2.1	37.1 ± 6.6
352-359	TYETTLEK	66.6 ± 4.8	60.8 ± 3.7
352-372	TYETTLEKCCAAADPHECYAK	5.4 ± 7.6	5 ± 8.7
390-402	QNCELFEQLGEYK	58.4 ± 2.1	59.5 ± 6.7
390-410	QNCELFEQLGEYKFQNALLVR	28.9 ± 12.1	30.7 ± 10.8
411-428	YTKKVPQVSTPTLVEVSR	75.6 ± 3.8	63.2 ± 11.9
445-472	RMPCAEDYLSVVLNQLCVLHEKTPVSDR	N/A	50 ± 70.7
476-500	CCTESLVNRRPCFSALEVDETYVPK	36.7 ± 11.2	29.7 ± 5.9
485-500	RPCFSALEVDETYVPK	46.5 ± 5.1	49.4 ± 9.4
501-521	EFNAETFTFHADICTLSEKER	17.6 ± 4.8	17.4 ± 1
526-538	QTALVELVKHKPK	1.7 ± 3	0 ± 0

Table S7. Peptide regions of HSA with significant differences in photochemical halogenation upon warfarin combination.

Photochemical bromination				Photochemical iodization			
Site	Peptide sequence	Change of oxidation level	Modification Sites	Site	Peptide sequence	Change of oxidation level	Modification Sites
206-218	FGERAFKAWAVAR	↓	W214	146-160	HPYFYAPELLFFAKR	↓	H146 Y148 Y150
210-218	AFKAWAVAR	↓	W214	161-181	YKAAFTECCQAADKAACLLPK	↓	Y161
258-274	ADLAKYICENQDSISSK	↓	Y263	234-262	LVTDLTKVHTECCHGDLLECADDRADLAK	↓	H242 H247
				338-348	HPDYSVVLLLR	↑	H338 Y341
				411-428	YTKKVPQVSTPTLVEVSR	↓	Y411

Table S8. A summary of protein digest and native protein bromination and iodination.

Sample	Device	Experimental results	
		Photochemical bromination	Photochemical iodination
Mouse liver digest	150 pulses in direct irradiation device	145 brominated peptides	2194 iodinated peptides
	300 pulses in direct irradiation device	51 brominated peptides	1539 iodinated peptides
	Single-pulsed irradiation capillary reactor	770 brominated peptides	1812 iodinated peptides
		All the brominated and iodinated peptides and residues were packaged in the **MouseLiver_MsMs spectrum**	
BSA	single-pulsed irradiation capillary reactor	The brominated peptides were listed in **Table S2** and 27 bromination residues were observed in **Table S4**	The iodinated peptides were listed in **Table S3** and 28 iodination residues were observed in **Table S4**
HSA-warfarin interaction	single-pulsed irradiation capillary reactor	The brominated peptides were listed in **Table S5** and peptide regions of HSA with significant differences upon warfarin combination were listed in **Table S7**	The iodinated peptides were listed in **Table S6** and peptide regions of HSA with significant differences upon warfarin combination were listed in **Table S7**

Data availability

The LC-MS/MS datasets in this study are deposited at the ProteomeXchange Consortium via the PRIDE[13] partner repository with the data set identifier PXD024795.

References

1. K. Zhang and K. M. Parker, *Environ. Sci. Technol.*, 2018, **52**, 9579-9594.
2. M. F. Fox, *J. Chem. Soc., Faraday Trans. 1*, 1977, **73**, 872-882.
3. M. F. Fox and E. Hayon, *J. Chem. Soc., Faraday Trans. 1*, 1977, **73**, 1003-1016.
4. W. Chen, Z. Huang, N. E. S. Tay, B. Giglio, M. Wang, H. Wang, Z. Wu, D. A. Nicewicz and Z. Li, *Science*, 2019, **364**, 1170-1174.
5. M. F. Fox, B. E. Barker and E. Hayon, *J. Chem. Soc., Faraday Trans. 1*, 1978, **74**, 1776-1785.
6. G. E. Crooks, G. Hon, J. M. Chandonia and S. E. Brenner, *Genome Res.*, 2004, **14**, 1188-1190.
7. B. T. Ruotolo and C. V. Robinson, *Curr. Opin. Chem. Biol.*, 2006, **10**, 402-408.
8. A. C. Susa, Z. Xia, H. Y. H. Tang, J. A. Tainer and E. R. Williams, *J. Am. Soc. Mass Spectrom.*, 2017, **28**, 332-340.
9. A. C. Leney and A. J. R. Heck, *J. Am. Soc. Mass Spectrom.*, 2017, **28**, 5-13.
10. S. Mehmood, J. Marcoux, J. T. S. Hopper, T. M. Allison, I. Liko, A. J. Borysik and C. V. Robinson, *J. Am. Chem. Soc.*, 2014, **136**, 17010-17012.
11. J. Hodgson, *Nat. Biotechnol.*, 2001, **19**, 722-726.
12. J. Ghuman, P. A. Zunszain, I. Petitpas, A. A. Bhattacharya, M. Otagiri and S. Curry, *J. Mol. Biol.*, 2005, **353**, 38-52.
13. Y. Perez-Riverol, A. Csordas, J. Bai, M. Bernal-Llinares, S. Hewapathirana, D. J. Kundu, A. Inuganti, J. Griss, G. Mayer, M. Eisenacher, E. Perez, J. Uszkoreit, J. Pfeuffer, T. Sachsenberg, S. Yilmaz, S. Tiwary, J. Cox, E. Audain, M. Walzer, A. F. Jarnuczak, T. Ternent, A. Brazma and J. A. Vizcaino, *Nucleic Acids Res.*, 2019, **47**, D442-D450.