

Supporting Information

Light-induced efficient and residue-selective bioconjugation of native proteins via indazolone formation

An-Di Guo^{a,c}, Ke-Huan Wu^b and Xiao-Hua Chen^{*a,b,c}

^a Chinese Academy of Sciences Key Laboratory of Receptor Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China. E-mail: xhchen@simm.ac.cn

^b School of Chinese Materia Medica, Nanjing University of Chinese Medicine, Nanjing, 210023, China.

^c University of Chinese Academy of Sciences, Beijing 100049, China.

General Experimental

Unless otherwise noted, all the materials were obtained from commercial suppliers and were used without further purification. Analytical TLC was performed using pre-coated plates (HSGF254) and visualized with UV light or an I₂ chamber. Silica chromatography was performed using the indicated solvent system on Sinopharm Chemical Reagent silica gel (200–300 mesh). ¹H NMR spectra and ¹³C NMR spectra were obtained on a Bruker AVANCE III 400 (400 MHz), Bruker AVANCE III 500 (500 MHz) and Bruker AVANCE III 600 (600 MHz) NMR spectrometer. Chemical shifts were reported in parts per million (ppm) on the δ scale from an internal standard (NMR descriptions: s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; br, broad). Coupling constants, *J*, are reported in Hertz. For HRMS analysis, samples were analyzed by flow-injection analysis into a Agilent 1290-6545 UHPLC-QTOF. LCMS analysis was performed on Waters UPLC-MS (ESI) with ACQUITY UPLC BEH C18 1.7 μ m column (UPLC: Waters HPLC H-CLASS, MS: Waters SQ Detector 2), using 5-95% acetonitrile gradient for 10 min. Protein molecular weight was determined by Waters XEVO G2-XS. Protein deconvolution was performed using Waters UNIFI. Peptide LC-MS/MS analysis was performed by Orbitrap fusion mass spectrometer. The fluorescent gel was visualized by GE Typhoon FLA 9500 and CLiNX ChemiScope 6300.

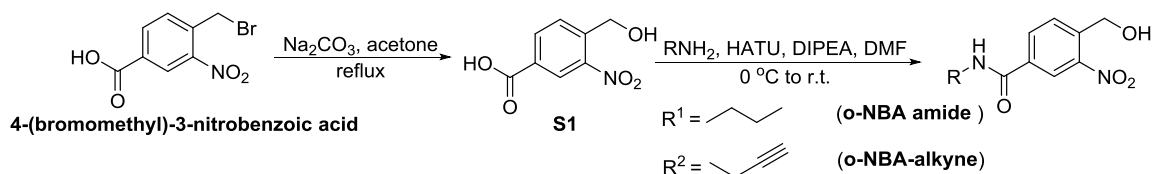
Photochemical reactions were carried out in Shanghai Heqi glassware B-002601 40×25 mm flat bottom flask and Titan 7.5×12.5×45 mm Quartz cuvette. Light sources used were a ZF-7A 16 W 365 nm UV source.

Peptide was purchased from Genescript.

Table of Contents

1. Synthesis of o-NBA compound.....	4
2. The reaction of o-NBA amide with peptide.....	6
3. Supplementary figures for protein labeling.....	9
4. Experimental procedures.....	12
5. 5. The selectivity of the o-NBA amide with NH ₂ -Lys (Cbz)-OMe (α -NH ₂) and Cbz-Lys-OMe (ϵ -NH ₂).....	15
6. Reference.....	16

1. Synthesis of *o*-NBA compound

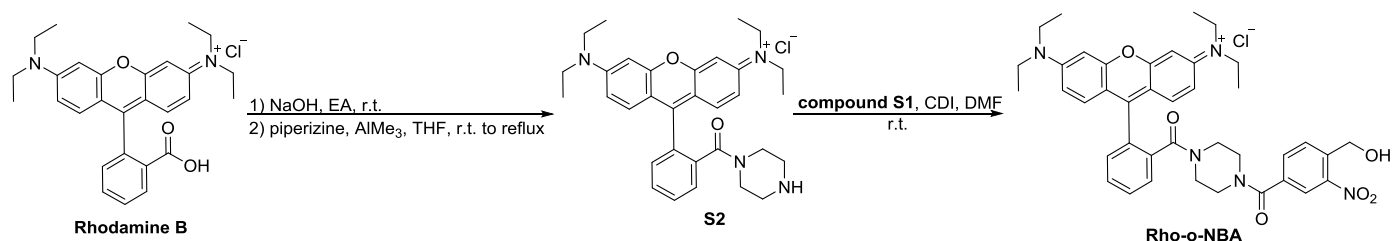


4-(bromomethyl)-3-nitrobenzoic acid (5.00 g, 19.2 mM) and Na_2CO_3 (7.13 g, 67.3 mM) in acetone/ H_2O (1:1, 150 mL) were refluxed for 3 h. Acetone was removed in vacuo and the aqueous phase was extracted with Et_2O twice. The aqueous phase was added concentrated HCl until pH = 1 or below and extracted with ethyl acetate 3 times. The organic layer was washed with H_2O and brine, dried over MgSO_4 and concentrated in vacuo without further purification to afford **compound S1** as a brown solid (3.78 g, 98%).⁵ ^1H NMR (500 MHz, MeOD) δ 8.60 (s, 1H), 8.29 (d, $J = 8.0$ Hz, 1H), 7.97 (d, $J = 8.1$ Hz, 1H), 5.00 (s, 2H).

Compound S1 (1.0 equiv), HATU (1.2 equiv) and propylamine (3.0 equiv) or propargylamine (3.0 equiv) or 3-azidopropan-1-amine (1.0 equiv) in anhydrous DMF were added DIPEA (3.0 equiv) dropwise at 0 $^\circ\text{C}$, and then the mixture was stirred at room temperature overnight. The mixture was added H_2O and extracted with ethyl acetate 3 times. The organic layer was washed with saturated NaHCO_3 , 0.1 M HCl, H_2O , brine, dried over Na_2SO_4 and concentrated in vacuo. The residue was purified by silica chromatography to afford product.

***o*-NBA amide** Yield: 64%. ^1H NMR (400 MHz, MeOD) δ 8.50 (d, $J = 1.7$ Hz, 1H), 8.14 (dd, $J = 8.1, 1.7$ Hz, 1H), 7.96 (d, $J = 8.2$ Hz, 1H), 4.99 (s, 2H), 3.39 – 3.33 (m, 2H), 1.66 (dq, $J = 14.7, 7.4$ Hz, 2H), 1.02 – 0.94 (m, 3H). ^{13}C NMR (101 MHz, MeOD) δ 167.5, 148.4, 142.5, 135.6, 132.9, 129.7, 124.5, 61.8, 42.9, 23.6, 11.8. HRMS (ESI-Q-TOF): m/z $[\text{M}+\text{H}]^+$ Calcd for $\text{C}_{11}\text{H}_{15}\text{N}_2\text{O}_4^+$: 239.1026; found: 239.1040.

***o*-NBA alkyne** Yield: 58%. ^1H NMR (400 MHz, MeOD) δ 8.51 (d, $J = 1.7$ Hz, 1H), 8.15 (dd, $J = 8.2, 1.8$ Hz, 1H), 7.97 (d, $J = 8.2$ Hz, 1H), 4.99 (s, 2H), 4.18 (d, $J = 2.5$ Hz, 2H), 2.64 – 2.62 (m, 1H). ^{13}C NMR (101 MHz, MeOD) δ 167.1, 148.3, 142.8, 134.9, 133.0, 129.8, 124.6, 80.4, 72.3, 61.8, 30.1. HRMS (ESI-Q-TOF): m/z $[\text{M}-\text{H}]^-$ Calcd for $\text{C}_{11}\text{H}_9\text{N}_2\text{O}_4^-$: 233.0568; found: 233.0566.



Rhodamine B (5.00 g, 10.4 mmol) was dispersed in 1 M NaOH (100 mL) and ethyl acetate (100 mL), and then the aqueous layer was extracted with ethyl acetate twice. The organic layer was washed with brine, dried over Na_2SO_4 and concentrated in vacuo. The product Rhodamine B base was obtained without further purification as a red oil (4.89 g, 86%). Piperazine (1.89 g, 21.92 mmol) in anhydrous DCM was added AlMe_3 in toluene (2.0 M, 5.48 mL, 10.96 mmol) at room temperature. After stirring for 1 h, white precipitate appeared and Rhodamine B base (3.00 g, 5.48 mmol) in anhydrous DCM was added dropwise at room temperature. The reaction mixture was refluxed overnight. After cooled down to room temperature, the mixture was added 0.1 M HCl until no more bubbles were observed. The mixture was filtered and washed with DCM and DCM/MeOH (4:1). The filtrate was concentrated and dissolved in DCM, filtered to remove undissolved salt

and concentrated again. The residue was dispersed in NaHCO₃ (0.02%) and ethyl acetate, and the aqueous layer was washed with ethyl acetate 5 times to remove materials. The aqueous layer was saturated with NaCl, acidified with 1 M HCl and extracted with ⁱPrOH/DCM (4:1) 3 times. The organic layer was dried over Na₂SO₄ and concentrated. The residue was purified by silica chromatography to afford **compound S2** as a pink oil (2.11 g, 53%). ¹H NMR (400 MHz, MeOD) δ 7.83 – 7.74 (m, 3H), 7.53 (dd, *J* = 5.8, 2.9 Hz, 1H), 7.27 (d, *J* = 9.5 Hz, 2H), 7.09 (dd, *J* = 9.5, 2.4 Hz, 2H), 6.98 (d, *J* = 2.4 Hz, 2H), 3.74 – 3.66 (m, 12H), 3.13 (d, *J* = 4.5 Hz, 4H), 1.36 – 1.28 (m, 12H).

Compound S2 (132 mg, 0.241 mmol) in anhydrous DMF (2.5 mL) was added CDI (41 mg, 0.253 mmol) and stirred for 1 h. **Compound S3** (47 mg, 0.241 mmol) was added and stirred overnight. The solvent was removed in vacuo and the residue was purified by silica chromatography to afford **Rho-o-NBA** as a pink oil (59 mg, 34%). ¹H NMR (400 MHz, MeOD) δ 8.07 (d, *J* = 1.5 Hz, 1H), 7.96 (d, *J* = 8.0 Hz, 1H), 7.77 (s, 2H), 7.73 (dd, *J* = 8.0, 1.6 Hz, 2H), 7.52 (d, *J* = 4.9 Hz, 1H), 7.28 (d, *J* = 8.7 Hz, 2H), 7.07 (d, *J* = 8.8 Hz, 2H), 6.98 (d, *J* = 2.4 Hz, 2H), 4.96 (s, 2H), 3.70 (q, *J* = 7.1 Hz, 8H), 3.49 (s, 8H), 1.31 (t, *J* = 7.1 Hz, 12H). ¹³C NMR (126 MHz, MeOD) δ 170.0, 169.7, 159.3, 157.3, 156.9, 148.4, 141.4, 136.5, 136.0, 133.2, 133.1, 132.2, 131.8, 131.4, 131.3, 130.0, 128.9, 124.5, 115.4, 114.9, 97.4, 61.7, 46.9, 12.8. HRMS (ESI-Q-TOF): *m/z* [M-Cl]⁺ Calcd for C₄₀H₄₄N₅O₆⁺: 690.3286; found: 690.3274.

2. The reaction of o-NBA amide with peptide

Ac-RYKSHWN-OH/ Ac-SRKYDH-OH

Stock solution: o-NBA amide (5 mM) in MeOH, and peptide (10 mM) in H₂O. o-NBA amide (1.25 mM) and peptide (0.5 mM) in 100 mM PBS/MeOH (9:1, pH = 7.4) were treated with 365 nm UV light for 7 min and shaken at 25 °C for 30 min. The samples were collected, diluted with MeOH/H₂O and analyzed by UPLC-MS.

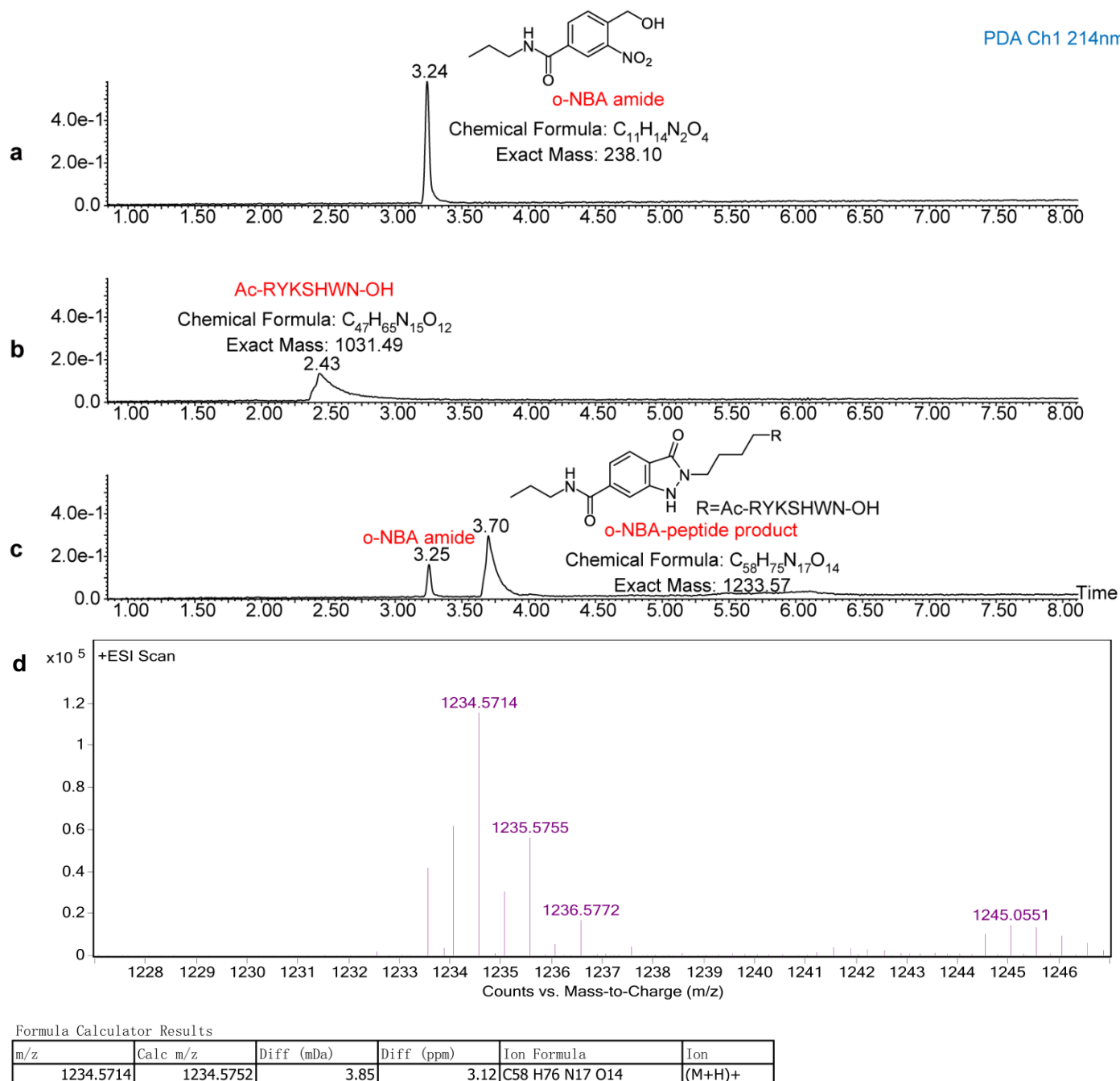


Figure S1. a) o-NBA amide (2 mM); b) Ac-RYKSHWN-OH (0.5 mM); c) o-NBA amide (1.25 mM) and Ac-RYKSHWN-OH (0.5mM) were treated with 365nm UV light; d) HRMS (ESI-Q-TOF) of o-NBA-peptide product.

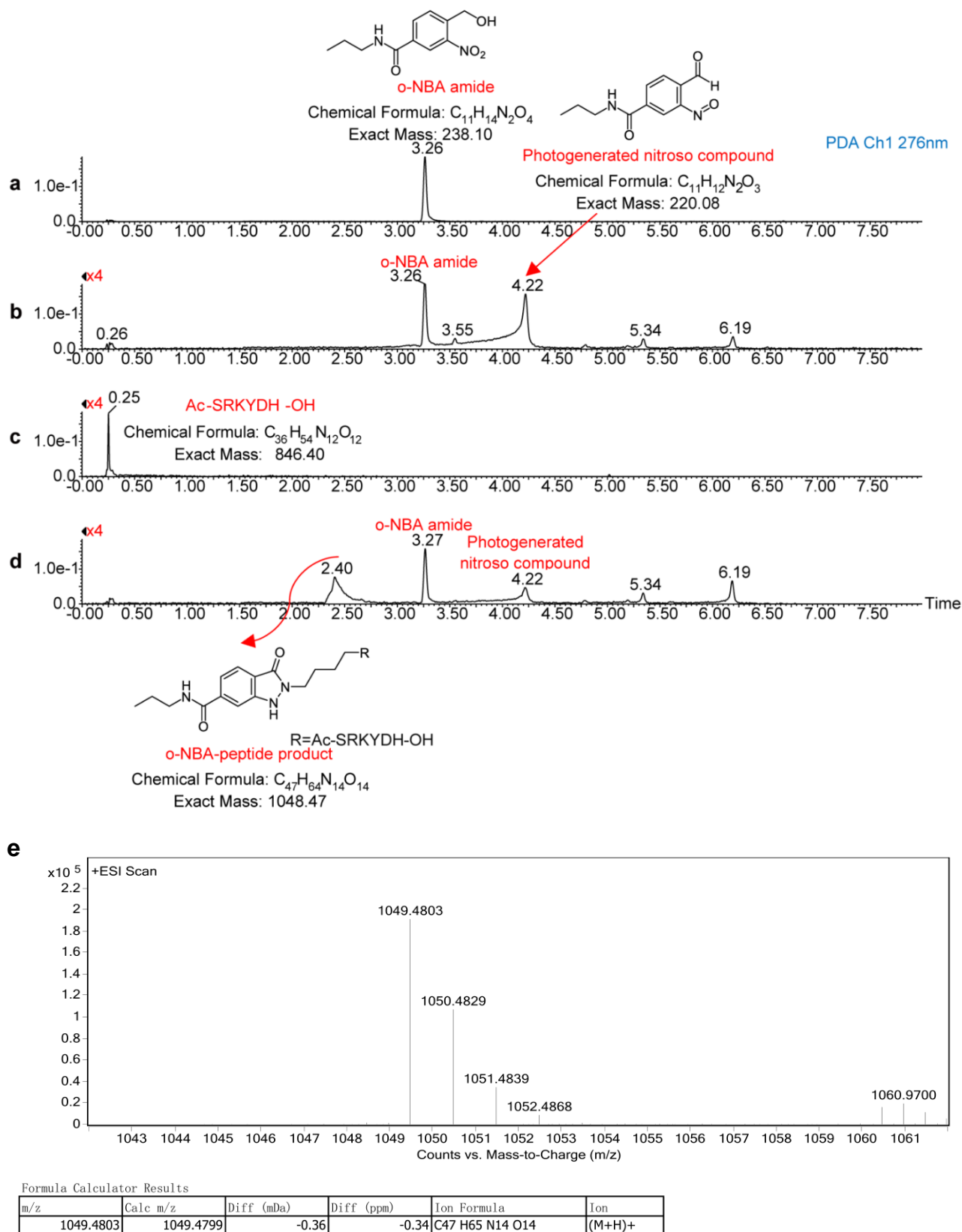


Figure S2. a) o-NBA amide (1.25 mM); b) Ac-SRKYDH-OH (0.5 mM); c) o-NBA amide (1.25 mM) and Ac-SRKYDH-OH (0.5 mM) were treated with 365 nm UV light; d) HRMS (ESI-Q-TOF) of o-NBA-peptide product. e) HRMS (ESI-Q-TOF) of o-NBA-peptide product.

Ac-RCYMNK-OH

Stock solution: o-NBA amide (10 mM) in MeOH, peptide (10 mM) in H₂O and DTT (160 mM) in H₂O. o-NBA amide (2 mM) was treated with 365 nm UV light for 7 min and then was added to peptide (0.5 mM) in 100 mM PBS/MeOH (9:1, pH = 7.4) 5 aliquots with a 6 min-interval. Then, DTT (8 mM) was added and shaken for 1 minute. The samples were collected, diluted with MeOH/H₂O and analyzed by UPLC-MS.

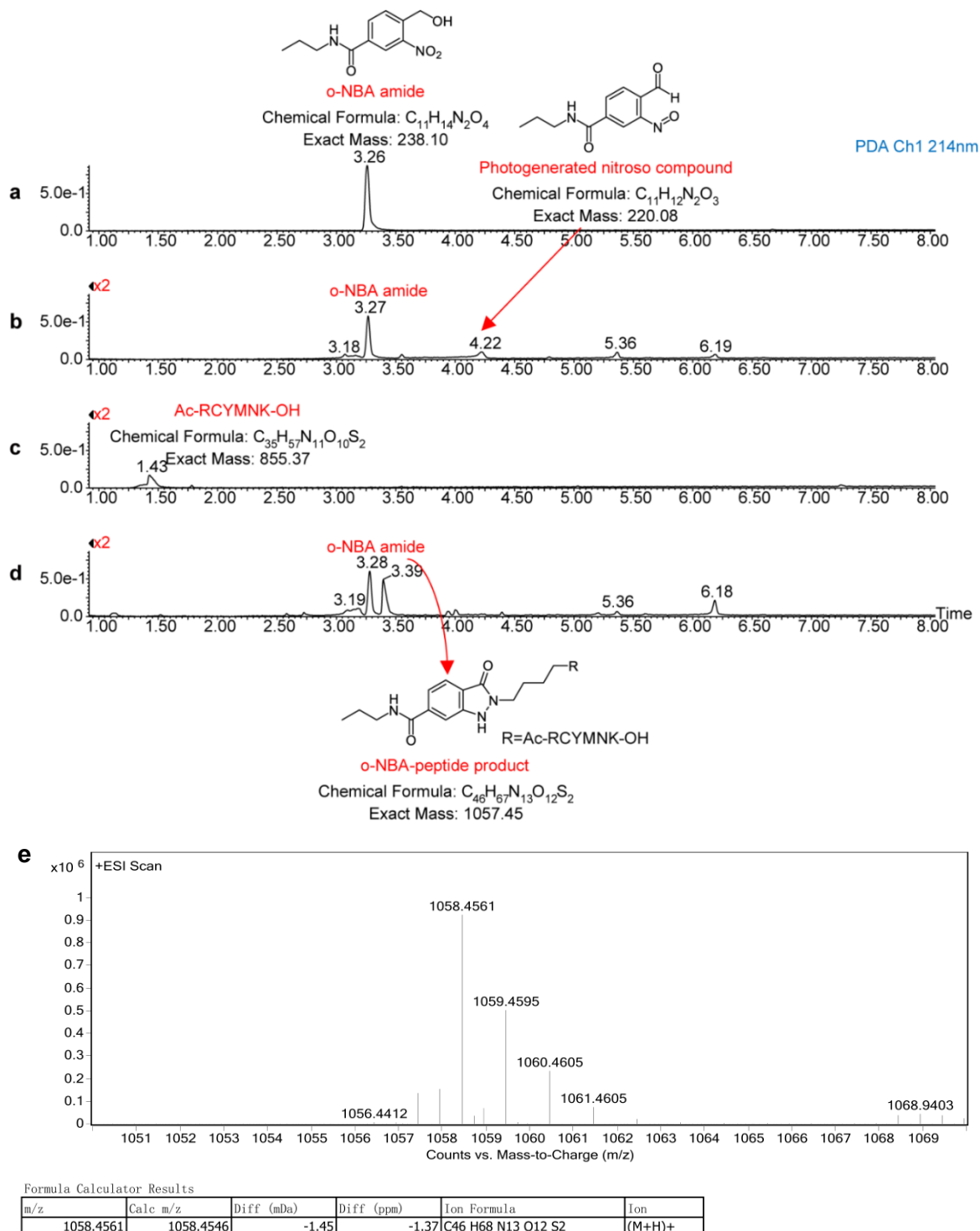


Figure S3. a) o-NBA amide (2 mM); b) o-NBA amide (2 mM) was irradiated with 365 nm UV light; c) Ac-RCYMNK-OH (0.5 mM); d) o-NBA amide (2 mM) was irradiated with 365 nm UV light for 7 min and then was added to Ac-RCYMNK-OH (0.5 mM) 5 aliquots and added DTT (8 mM); e) HRMS (ESI-Q-TOF) of o-NBA-peptide product.

3. Supplementary figures for protein labeling

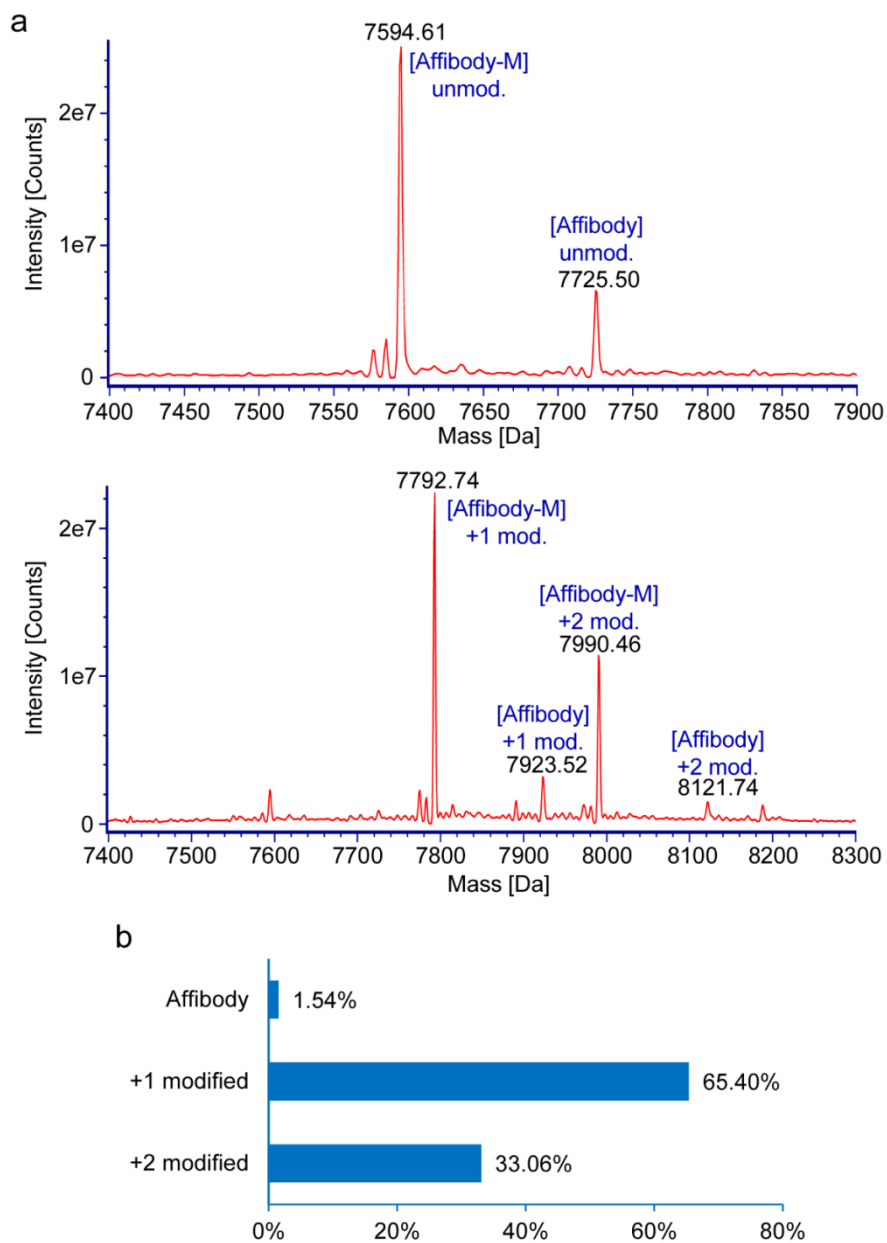


Figure S4 Light-induced Affibody labeling with *o*-NBA alkyne. a) ESI-TOF Spectrum of unmodified Affibody and *o*-NBA-alkyne (2.5 equivalent) modified Affibody. b) Quantitative analysis of *o*-NBA-alkyne (2.5 equivalent) modified Affibody by peak height of mass spectrometry.

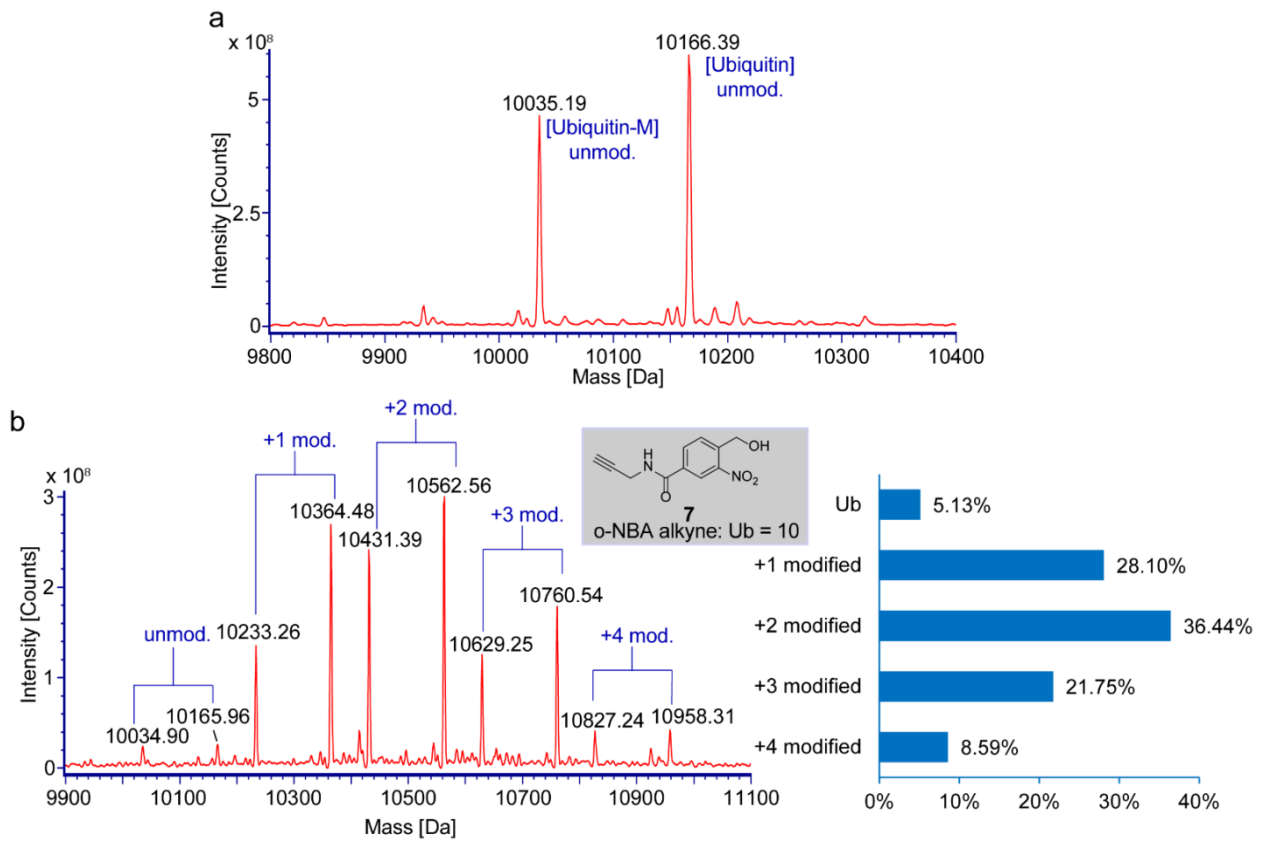


Figure S5 Light-induced ubiquitin labeling with o-NBA alkyne. a) ESI-TOF Spectrum of unmodified ubiquitin. b) Quantitative analysis of o-NBA-alkyne (10 equivalent) modified ubiquitin by peak height of mass spectrometry.

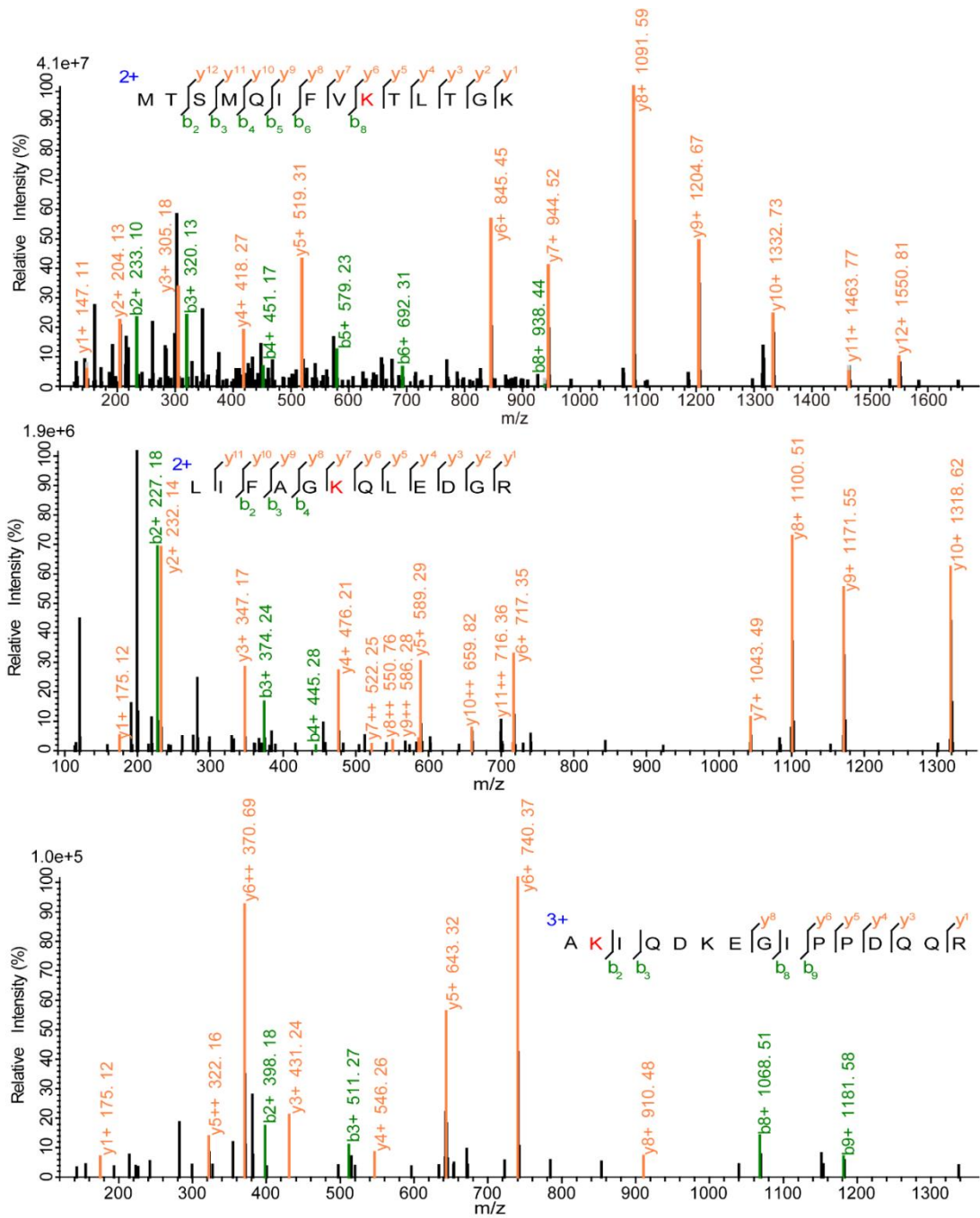


Figure S6 Supplementary tandem MS Analysis of the fragments of Ubiquitin-o-NBA-alkyne conjugate before CuAAC.

4. Experimental procedures

4.1 The optimization of o-NBA amide photo-reacting with Cbz-Lys-OMe

Stock solution: ONBA (10 mM/2 mM/1 mM) in MeOH and Cbz-Lys-OMe (10 mM/2 mM) in MeOH. ONBA (2 mM/0.4 mM/0.2 mM) and Cbz-Lys-OMe (0.5 mM/0.1 mM) in 100 mM PBS/MeOH = 2:1 (pH = 7.4/6.0/8.0) were treated with 365 nm UV light for corresponding time and shaken at 25 °C for corresponding time. The samples were collected, diluted with MeOH/H₂O and analyzed by UPLC-MS. Yields were determined by ratio of peak area value of experiment to that of internal standard product on reverse-phase HPLC, are reported as an average of three independent trials.

4.2 Protein Sequences:

Affibody:

MTSVDNKFNKELSVAGREIVTLPNLNDPQKAFIFSLWDDPSQSANLLAEAKKLNDAAQAPKGSHHHHHH

Nanobody-EGFR:

QVKLEESGGGSVQTGGSLRLTCAASGRTRSRSYGMGWFRQAPGKEREFVSGISWRGDSTGYADSVKGRFTISRDN
KNTVDLQMNSLKPEDTAIYYCAAAGSAWYGTLYEYDYWGQGTQVTVSSALEHHHHHH

Ubiquitin:

MTSMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQORLIFAGKQLEDGRTLSDYNIQKESTLHLVLRRLRG
LEHHHHHHHH

Nanobody-HER2:

MQVOLQESGGGSVQAGGSLKLTCAASGYIFNSCGMGWYRQSPGRERELVSRISGDGDTWHKESVKGRFTISQDNV
KKTLYLQMNSLKPEDTAVYFCAVCYNLETYWGQGTQVTVSSGGHHHHHH

4.3 Expression and purification of Affibody/Ubiquitin/Nanobody-EGFR/Nanobody-HER2 proteins^[1-3]

DH10B cells were respectively transformed with expression plasmid PBAD-Affibody or PBAD-Ubiquitin or PET-26b-7D12 (Nanobody-EGFR) or PET-26b-2RS15d (Nanobody-HER2) using heat shock and recovered in 500 µL SOC media and incubated at 37 °C for 1 hour before plating to 2YT agar plate containing 100 µg/mL Ampicillin. A single colony from the plate was picked and used to inoculate 5-mL 2YT containing corresponding probiotic. A 1-mL aliquot of overnight culture was then used to inoculate 100 mL 2YT containing the same concentrations of antibiotics. The cells were grown at 37 °C until OD₆₀₀ reached ~0.5, and the protein expression was induced by adding 1 mM IPTG for 8 hours. The cells were pelletized in 50-mL conical tubes by centrifugation at 4200 g for 30 min at 4 °C and stored at -80 °C. Next day, the cell pellets were resuspended in 10mL lysis buffer. After incubation on ice for 30 min, the lysates were further treated with ultrasonication. Following centrifuge, the supernatants were incubation with 100 µL Ni-NTA agarose beads at 4 °C for 2 hours with gentle shaking. The resin was centrifuged briefly and washed three times with washing buffer. Finally, the protein was eluted with elution buffer. A 10-µL aliquot from the elution fraction was mixed with one fourth amount of 5X SDS loading buffer and heated at 98 °C for 8 minutes, before loading onto 12% SDS-PAGE gel. After exchanging to a phosphate buffer (pH 7.5), the protein was analyzed by LC/ESI-MS and the protein mass was obtained by deconvoluting the charge ladder.

4.4 The modification of proteins with o-NBA alkyne

o-NBA-alkyne was prepared as a 2.5 mM/5 mM/10 mM stock in MeOH solution. Affibody/Ubiquitin/Nanobody-HER2 was prepared as a 55 μ M stock in PBS solution. o-NBA-alkyne (2.5 mM) in MeOH was treated with 365 nm UV light for 7 min and then was added to Affibody/Ubiquitin/Nanobody-HER2 (50 μ M) in PBS one portion to a final concentration of 125 μ M/250 μ M/500 μ M and mixed. The mixture was shaken at 25 °C for 1 h. The samples were collected, diluted with H₂O and analyzed by ESI-TOF.

4.5 SDS-PAGE of labeled Ubiquitin after CuAAC

TAMRA-N₃ was prepared as a 1.25 mM stock in DMSO solution. CuSO₄ and THPTA were prepared as a 50 mM stock in H₂O solution and premixed with a volume ratio of 1:5. Sodium ascorbate was prepared as a 100 mM stock in H₂O solution. One hundred microliters of o-NBA-alkyne (2.5 equivalent) labeled Ubiquitin (50 μ M) were added TAMRA-N₃ (100 μ M), premixed CuSO₄ (100 μ M), THPTA (500 μ M) and sodium ascorbate (5 mM). The mixture was rotated at 25 °C for 1 h and analyzed by SDS-PAGE.

4.6 The modification of Nanobody-HER2 with Rho-o-NBA

Rho-o-NBA was prepared as a 2.5 mM stock in MeOH solution. Nanobody-HER2 was prepared as a 55 μ M stock in PBS solution. Rho-o-NBA (2.5 mM) in MeOH was treated with 365 nm UV light for 10 min and then was added to Nanobody-HER2 (50 μ M) in PBS one portion to a final concentration of 125 μ M and mixed. The mixture was shaken at 25 °C for 1 h. The samples were collected, diluted with H₂O and analyzed by ESI-TOF and SDS-PAGE (without light irradiation as control).

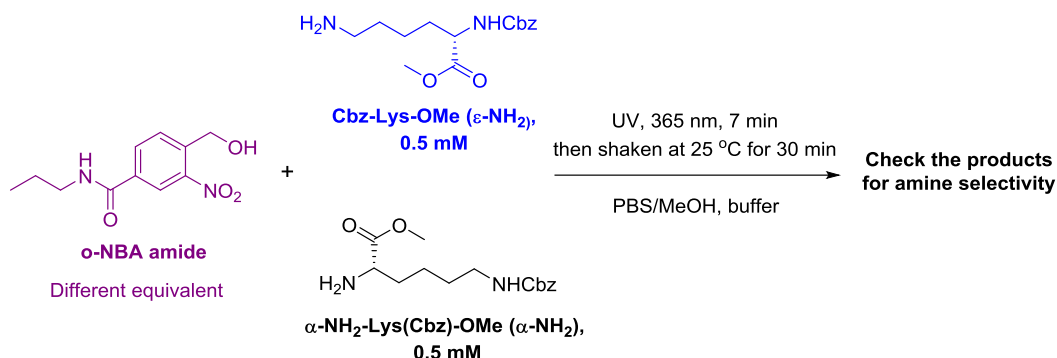
4.7 Real-time and in vivo labeling of expressed Ubiquitin protein versus time

o-NBA-alkyne was prepared as a 100 mM stock in DMSO solution. TCEP•HCl and CuSO₄ were prepared as a 100 mM and 50 mM stock in H₂O solution respectively. TBTA was prepared as a 1.7 mM solution in DMSO/^tBuOH solution. TAMRA-N₃ was prepared as a 1.25 mM stock in DMSO solution. DH10B cells were transformed with expression plasmid PBAD-Ubiquitin using heat shock and recovered in 500 μ L SOC media and incubated at 37 °C for 1 h before plating to 2YT agar plate containing 100 μ g/mL Ampicillin. A single colony from the plate was picked and used to inoculate 5 mL 2YT containing corresponding probiotic. A 1 mL aliquot of overnight culture was then used to inoculate 100 mL 2YT containing the same concentrations of antibiotics. The cells were grown at 37 °C until OD₆₀₀ reached 0.4. DH10B was added o-NBA-alkyne (2 mM) after incubated at 37 °C with shaking, and then was incubated at 30 °C with shaking. After 15 min, ten milliliters of cells were collected by centrifugation, resuspended in o-NBA-alkyne (2 mM) in PBS, exposed to 365 nm UV light for 10 min and re-centrifuged as a negative control group without induced expression. Then ubiquitin expression was induced with IPTG (1 mM). Ten milliliters of the cell suspension were collected and centrifuged after induced 30 min/1 h/2 h/4 h/8 h/16 h. The cell pellets were resuspended in PBS, normalized in OD₆₀₀, exposed to 365 nm UV light for 10 min and collected by centrifugation. All collected cell pellets were resuspended in DPBS (containing 1 mM PMSF), lysed ultrasonically and centrifuged. The supernatant was denatured with 1% SDS and heating, normalized by BCA protein assay kit and diluted with DPBS to 2 mg/mL concentration for CuAAC. Samples were added TCEP•HCl (2 mM), CuSO₄ (1 mM), TBTA (100 μ M), TAMRA-N₃ (25 μ M) and rotated for 2 h. Then proteins were precipitated by acetone, washed with MeOH twice, re-dissolved in 1 \times loading buffer and analyzed by SDS-PAGE and western blot (His tag).

4.8 In gel digestion of o-NBA-alkyne Ubiquitin and LC-MS/MS analysis

Ubiquitin-o-NBA-alkyne conjugate obtained in the previous step was loaded to SDS-PAGE and Coomassie Blue staining. The bands of interest were excised from gels and cut into cubes (1 × 1 mm). The gel pieces were destained and dehydrated in acetonitrile and incubated in buffer I (20 mM DTT, 50 mM NH₄HCO₃) at 56 °C for 30 min followed by buffer II (100 mM iodoacetamide, 50 mM NH₄HCO₃) at ambient temperature for 20 min photophobia in dark room, before being dehydrated. The samples were in gel digested with trypsin at 37 °C for 12 hours. The resulting peptides were extracted thrice with extraction buffer (60% acrylonitrile, 0.1% trifluoroacetic acid) along with sonication, and then vacuum centrifuged to dryness. For LC-MS/MS analysis, the extracted peptides were reconstituted in 0.1% formic acid and analyzed by Orbitrap fusion mass spectrometer.

5. The selectivity of the *o*-NBA amide with NH₂-Lys (Cbz)-OMe (α -NH₂) and Cbz-Lys-OMe (ϵ -NH₂)



Stock solution: *o*-NBA amide (2.5 mM/5 mM/10 mM/17.5 mM/25 mM) in MeOH, Cbz-Lys-OMe (10 mM) in MeOH and NH₂-Lys (Cbz)-OMe (10 mM) in MeOH.

Reaction conditions: *o*-NBA amide (0.5 mM/1 mM/2 mM/3.5 mM/5 mM), Cbz-Lys-OMe (0.5 mM) and NH₂-Lys (Cbz)-OMe (0.5 mM) in 100 mM PBS/MeOH (2:1, pH = 5.0/6.0/6.5/7.0/7.4/8.0) were treated with 365 nm UV light for 7 min and shaken at 25 °C for 30 min. The samples were collected, diluted with MeOH/H₂O and analyzed by UPLC-MS. The yields of α -NH₂ and ϵ -NH₂ indazolone product Yields were determined by ratio of peak area value of experiment to that of internal standard product on reverse-phase HPLC and are reported as an average of three independent trials.

Table 1. Selectivity of *o*-NBA amide towards Cbz-Lys-OMe (ϵ -NH₂) than NH₂-Lys (Cbz)-OMe (α -NH₂) on different conditions.

Entry	Concentration <i>o</i> -NBA amide	The concentration of Cbz-Lys-OMe (ϵ -NH ₂) and NH ₂ -Lys (Cbz)-OMe (α -NH ₂)	pH	Yields of Cbz-Lys-OMe (ϵ -NH ₂)	Yields of NH ₂ -Lys (Cbz)-OMe (α -NH ₂)
1	2.0 mM	0.5 mM	5.0	39%	2%
2	2.0 mM	0.5 mM	6.0	48%	4%
3	2.0 mM	0.5 mM	6.5	65%	4%
4	2.0 mM	0.5 mM	7.0	72%	6%
5	2.0 mM	0.5 mM	8.0	97%	9%
6	0.5 mM	0.5 mM	7.4	55%	N.D.
7	1.0 mM	0.5 mM	7.4	80%	N.D.
8	2.0 mM	0.5 mM	7.4	95%	7%
9	3.5 mM	0.5 mM	7.4	95%	11%
10	5.0 mM	0.5 mM	7.4	95%	14%

N.D.: no detected on UPLC-MS.

6. References

- [1] X.- H. Chen, Z. Xiang, Y. S. Hu, V. K. Lacey, H. Cang, L. Wang, *ACS Chem. Biol.* **2014**, *9*, 1956-1961;
- [2] W. Hu, Y. Yuan, C. H. Wang, H. T. Tian, A. D. Guo, H. J. Nie, H. Hu, M. J. Tan, Z. Tang, X.- H. Chen, *Chem.* **2019**, *5*, 2955-2968.
- [3] A.-D. Guo, D. Wei, H.-J. Nie, H. Hu, C. Peng, S.-T. Li, K.-N. Yan, B.-S. Zhou, L. Feng, C. Fang, M. Tan, R. Huang X.-H. Chen ,*Nat. Commun.* **2020**, *11*, 5472.