Electronic Supplementary Material (ESI) for ChemComm.

Electronic Supplementary Information

A chemoselective reaction between protein *N*-homocysteinylation and azide catalyzed by Heme(Ⅱ)

Nan Chen,[‡a] Zeyu Qiao,[‡a] and Chu Wang*[a,b]

[a] Synthetic and Functional Biomolecules Center, Beijing National Laboratory for Molecular Sciences, Key Laboratory of Bioorganic Chemistry and Molecular Engineering of Ministry of Education, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China.

[b] Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China.

Email: chuwang@pku.edu.cn

Table of content

I Biological methods 2
II Supplementary schemes and figures 6
III Chemical Synthesis 10
IV Reference 13

I Biological methods

Reagents

Biotin-PEG4-azide (Product no. 271606) was purchased from ChemPep Inc. Rhodamine azide was synthesized by Beijing Okeanos Technology Co., Ltd. RNaseA, Pancreatic (Product no. 0675-250mg, Amresco), Myoglobin from equine skeletal muscle (Product no. M0630-250mg, Sigma), Lysozyme (Product no. 0665-5g, Amresco), Bovine Serum Albumin (BSA) (Fraction V) (Product no. BP1600-100, *L*-homocysteine thiolactone hydrochloride, HTL (H6503-100mg, Sigma), AC-Biotin-azide (Product Cat. 1330-25, Click Chemistry Tool), DTT, TCEP, chemical analogues of β-ME, amino acids and other chemical or biological reagents were obtained from commercial suppliers without any manipulation. B-hcy was synthesized according to the previous report[1].

Labelling of Mb via AAP

250 μM of Mb were prepared in PBS (0.2 mM EDTA). Either H_2O or HTL (0.2 mM) was added into 0.5 ml of protein solutions and the mixtures were incubated at 25 °C for 18 h. After induction, proteins were filtered with NAP-5 columns (GE Healthcare) to remove small molecule species. The modifications were characterized by LCQ fleet (Thermofisher Scientific) before.

Control or HTL-induced Mbs were diluted to 25 μM with citric acid solution (pH = 3, 50 mM). Then TCEP (final concentration 500 μM) and AAP (final concentration 0 or 250 μM) were added into solutions. The mixtures were incubated at 25°C for 3 h, followed by acetone precipitation. The protein pellets were washed twice with 70% acetone and resuspended in 100 μL of SDS/PBS (1%) with sonication. Samples were centrifuged (20,000 *g*, 3 min) and 50 μL of samples were aspirated for click chemistry (200 μM Rhodamine-azide, 100 μM Tris (benzyl triazolylmethyl)amine (TBTA), 1 mM $CuSO_4$ and 1 mM TCEP). Samples were allowed to react for 1 h at room temperature, quenched with 14 μL of 5*gel loading buffer and heated at 95°C for 10 min. The prepared samples were analyzed by SDS-PAGE (15% polyacrylamide; 10 μL of sample/lane) and visualized by in-gel fluorescence using Rhodamine channel (BioRad ChemiDoc MP).

Labelling of Mb via Rhodamine azide and Biotin azide

Control or HTL-induced Mbs were diluted to 25 μM with NH_4HCO_3 buffer (100 mM). Different concentration of Rhodamine azide (6.25, 12.5, 25 and 50 μM) or Biotin azide (50, 100, 200 and 400 μM) were added and the mixtures were incubated at 25°C for 3 h. Then 5*gel loading buffer were added directly and the samples were heated at 95°C for 10 min, followed by SDS-PAGE (15% polyacrylamide; 10 μL of sample/lane) analysis. The signals were

detected by in-gel fluorescence for Rhodamine azide labelling and western blotting for Biotin azide labelling, respectively.

Investigation on purified proteins

Control or HTL-induced Mbs were diluted to 25 μM with PBS (0.2 mM EDTA) or Tris-HCl (100 mM). The pH of protein solutions were adjusted to 6, 7, 8, 9 and 10. The reaction volume was 50 μL and Rhodamine azide (final concentration 400 μM) was added directly. Then 12.5 μL of 5*Gel loading buffer were added directly and the samples were heated at 95°C for 10 min, followed by in-gel fluorescence analysis.

Reaction solutions containing control or HTL-induced Mbs (25 μM) in PBS (0.2 mM EDTA) were adjusted to pH of 9. Volume of 50 μL were used for each reaction. All of samples were analyzed by in-gel fluorescence.

For experimental test of essential factors, Rhodamine azide (final concentration 400 μM) was added directly. Then 5*gel loading buffer were added without heating, without β-ME, without glycerol, without Tris, without SDS and without BPB. Other conditions were still kept the same.

For concentration-dependent labelling experiments, different concentration of Rhodamine azide (final concentration 0, 10, 50, 100, 200 and 400 μM) was added and the samples were heated at 95°C for 10 min after addition of 12.5 μL of 5*gel loading buffer.

For time-dependent labelling experiments, Rhodamine azide (final concentration 200 μM) and 12.5 μL of 5*Gel loading buffer were added, followed by heating at 95°C for 1, 5, 10, 20 and 30 min.

For temperature-dependent labelling experiment, Rhodamine azide (final concentration 200 μM) and 12.5 μL of 5*Gel loading buffer were added, followed by heating at 25, 45, 55, 65, 75, 85 and 95°C for 10 min.

For experimental test of β-ME, Rhodamine azide (final concentration 200 μM) and 12.5 μL of 5*gel loading buffer containing different concentrations of β-ME (final concentration 0, 10, 50, 100, 200 and 500 mM) were added, followed by heating at 95°C for 10 min.

After optimization, Cy5 azide (final concentration 200 μM) were used. 100 mM of β-ME and SDS (final concentration 0.4%) were added. The mixtures were heated at 75°C for 10 min.

For reductant-dependent labelling experiment, Rhodamine azide (final concentration 200 μM), 100 mM of β-ME or other reductants (DTT, TECP, chemical analogues) and SDS (final concentration 0.4%) were added. The mixtures were heated at 75°C for 10 min.

For experimental test of bioorthogonality, 1 mM of GSH, amino acids (Lysine, Cysteine, Histidine, Leucine, Serine, Arginine) and B-hcy were added into the protein solutions. Then

Rhodamine azide (final concentration 200 μM), β-ME (final concentration 100 mM) and SDS (final concentration 0.4%) were added. The mixtures were heated at 75°C for 10 min.

For selective labelling of other purified proteins, 50 μL of reaction solutions containing control or HTL-induced BSA, Lysozyme and RNaseA (25 μM) in PBS (0.2 mM EDTA) (pH 9) were used. Then freshly made Mb, Hb, Hematin, Hemin and ZnPP (final concentration 25 μM) were added, followed by addition of Rhodamine azide (final concentration 200 μM), β-ME (final concentration 100 mM) and SDS (final concentration 0.4%). The mixtures were heated at 75°C for 10 min.

For concentration-dependent test of Hemin, 50 μL of reaction solutions containing control or HTL-induced BSA (25 μM) in PBS (0.2 mM EDTA) (pH 9) were used. Then freshly made Hemin (final concentration 0, 1, 10, 25, 50, 100 μM) were added, followed by addition of Rhodamine azide (final concentration 200 μM), β-ME (final concentration 100 mM) and SDS (final concentration 0.4%). The mixtures were heated at 75°C for 10 min.

Small molecule experiments

Identification of product: Benzaldehyde or benzyl azide (final concentration 5 mM) were added to 100 μL of PBS (0.2 mM EDTA) (pH 9) containing B-hcy (final concentration 2.5 mM), followed by addition of β-ME (final concentration 100 mM) and Hemin (50 μM). Mixtures were heated at 75°C for 10 min. Then the solutions were diluted with 900 μL of water containing internal standard (ethyl benzoyltyrosinate, 11.1 μM). The samples were analyzed by LCQ fleet (ThermoFisher Scientific) and product A was characterized by Q Exactive (ThermoFisher Scientific).

Comparison of the reaction rates: Benzyl azide (final concentration 5 mM) were added to PBS (0.2 mM EDTA) (pH 9) containing B-hcy (final concentration 1 mM), followed by addition of β-ME (final concentration 100 mM) and Hemin (50 μM).The mixtures were heated at 75°C for 2, 4, 6 and 8 min. Benzaldehyde (final concentration 5 mM) were added to citric acid buffer (50 mM, 2 mM TCEP, pH 3) containing B-hcy (final concentration 1 mM). The mixtures reacted at 25°C for 2, 4, 6 and 8 min. At each point, 100 μL of reaction mixture were aspirated and added to 900 μL containing internal standard (ethyl benzoyltyrosinate, 11.1 μM) and 11.1% HCOOH for reaction quenching. 50 μL of final solutions were detected by LCQ fleet.

Calibration curve of Product A: different amounts of product A were added into 100 μL of PBS (0.2 mM EDTA) to give final concentration of 0, 1, 10, 50 and 100 μM. Then the solutions were diluted with 900 μL of water containing internal standard (ethyl benzoyltyrosinate, 11.1 μM). 50 μL of final solutions were detected by LCQ fleet.

Quantification of Product A: 100 μL of PBS (0.2 mM EDTA) were adjusted to pH of 6, 7, 8, 9 and10. Then B-hcy (1000 μM), β-ME (100 mM) and Hemin (50 μM) were added. 1 equivalent of benzyl azide (1000 μM) was added to PBS (0.2 mM EDTA) with pH of 6, 7, 8, 10 and 0.1, 0.5, 1, 2, 5 equivalent were added to PBS (0.2 mM EDTA) with pH of 9. All of the reaction were heated at 75°C for 10 min. Then the solutions were diluted with 900 μL of water containing internal standard (ethyl benzoyltyrosinate, 11.1 μM) for LCQ fleet analysis.

Protein labelling and sites identification

10 ml of reaction solutions containing HTL-induced BSA (167 μM) in PBS (0.2 mM EDTA) (pH 9) were labelled by AC-Biotin-azide (200 μM). Freshly made Hemin (final concentration 50 μM), β-ME (final concentration 100 mM) and SDS (final concentration 0.4%) were added together. The reaction mixtures were heated at 75°C for 10 min, followed by acetone precipitation. The protein pellets were washed twice with 70% acetone and resuspended in 4 ml of SDS/PBS (1.2%) with sonication. Samples were centrifuged (20,000 *g*, 3 min) and diluted with PBS to 0.2% SDS for enrichment by streptavidin beads (200 μL). The mixtures were incubated at room temperature for 3 h with rotation. After being washed with 3*5 ml of PBS and water sequentially, the beads were denatured in 6 M urea/PBS, reduced with 10 mM dithiothreitol (DTT) at 37°C for 30 min and alkylated with 20 mM iodoacetamide (IAA) at 35°C for 30 min in dark. Then the beads were incubated with 200 μL of 2M urea/PBS buffer, 2 μL of 100 mM $CaCl_2$ and 4 μL of trypsin (100 μg reconstituted in 200 μL of resuspend buffer) at 37 °C with agitation overnight. After trypsin digestion, the beads were washed with 3*PBS and 3*water. The adducted peptides were released from the beads by water containing 2% HCOOH (200 μL, two times) at room temperature. Then the beads were washed with water (50% acetonitrile, 1% HCOOH). All of the elution buffers were combined and dried in vacuum. Finally, the peptides samples were reconstituted in 10 μL of 0.1% (v/v) formic acid in water for MS analysis.

pFind3 software[2] was used for peptide identification with the newly released Bovine Proteome database by Uniprot. Cysteine residues were searched with a fixed modification for carboxyamidomethylation (+ 57.02146 Da). Oxidation (Met) (+ 15.9949 Da), Acetyl (Any N-term) (+ 42.010565 Da) and Lys *N*-homocysteinylation with AC-cleavable tags (HTL_AC_azide, + 215.09799 Da) were set as variable modifications. The search results were filtered with a defined peptide false positive rate of 1%.

II Supplementary schemes and figures

Table S1. List of gel loading buffer components used in the experiments

Reagent	Concentration
β-ME	1%
Tris-HCl (pH 6.8, 1 M)	12 mM
glycerol	5%
SDS	0.4%
BPB	0.02%

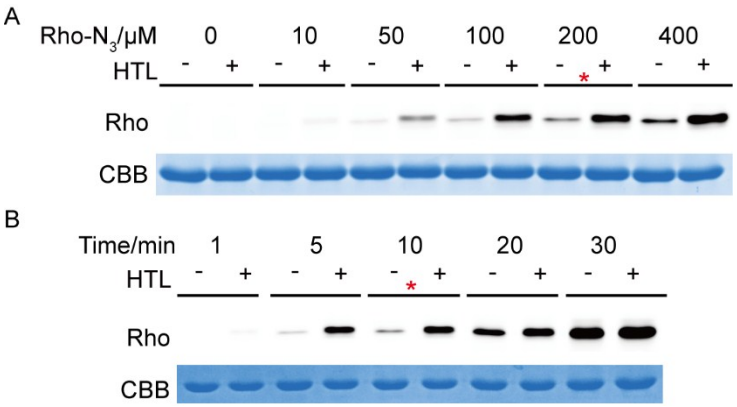


Fig. S1 Test of reaction conditions on Mb. (A) and (B) Concentration and time-dependent selective labelling of *N*-homocysteinylated Mb via rhodamine azide (Rho-N_3) (* denotes the optimal conditions used in the following experiments).

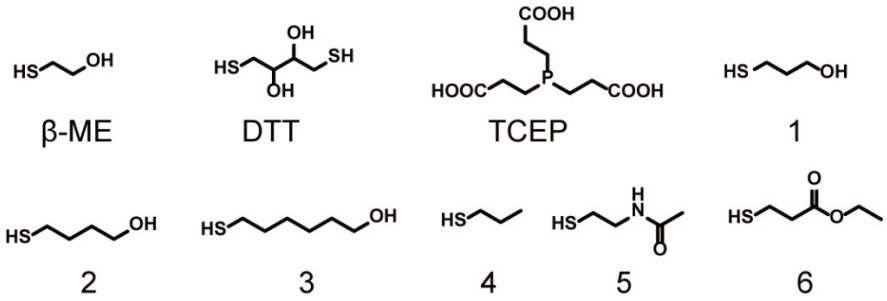


Fig. S2 Chemical structures of reductants and analogues used in the experiments.

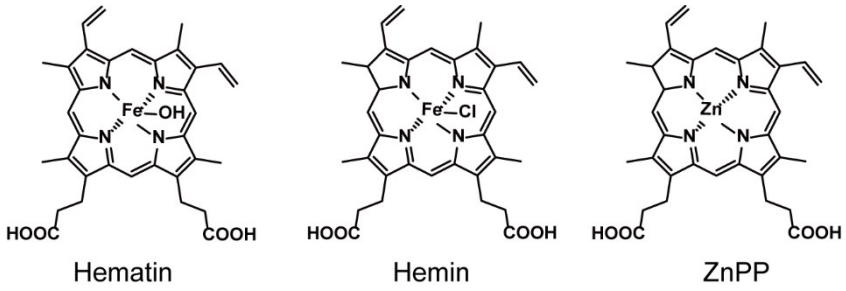


Fig. S3 Chemical structures of Hematin, Hemin and ZnPP.

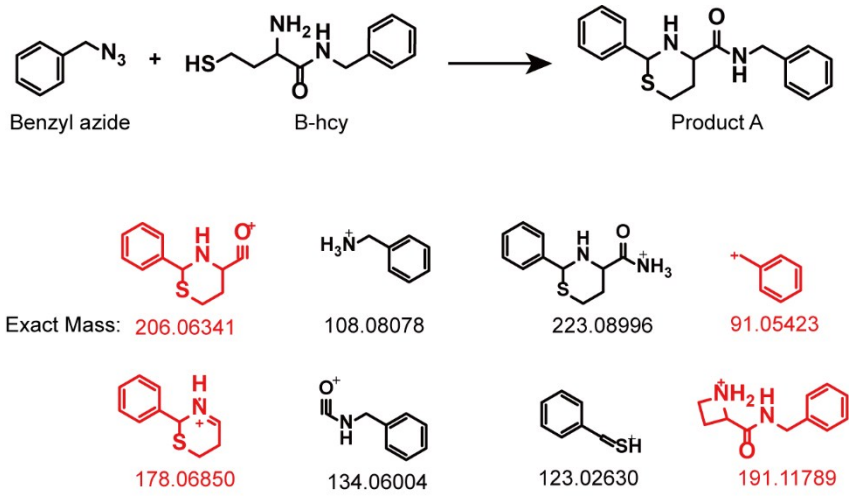


Fig. S4 Chemical reaction between benzyl azide and B-hcy. The structure of product A (BPTC) is determined by high-resolution LC-MS. The possible fragment ions upon HCD (Higher-energy C-trap dissociation) via Q Exactive are listed here. The identified ions (shown in Fig. 4B) in MS/MS spectrum are indicated in red.

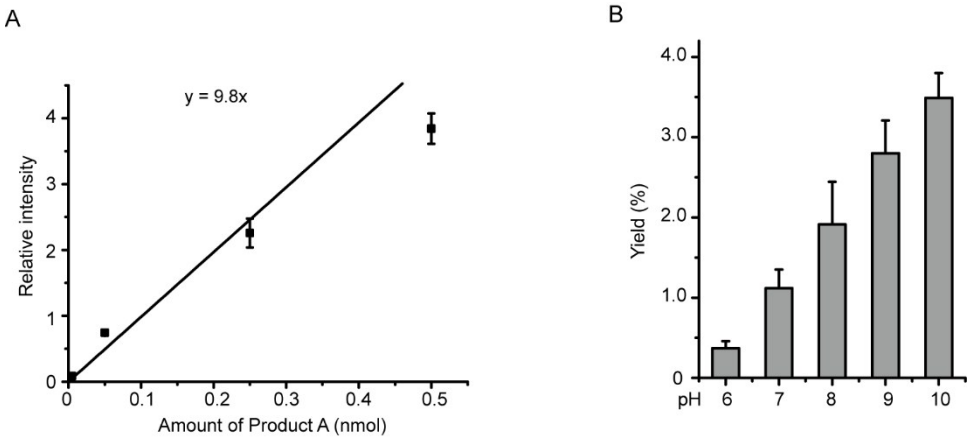


Fig. S5 Yield quantification of product A (BPTC). (A) Calibration curve of product A. The linear relation of relative intensity (divided by the intensity of internal standard) to the amount of product A is Y = 9.8X. (B) The yield of product A generated by the conjugation between

benzyl azide and B-hcy in PBS (0.2 mM EDTA) with different pHs (6.0, 7.0, 8.0, 9.0, 10.0). (data represent mean value ± standard deviation; n = 3 per group)

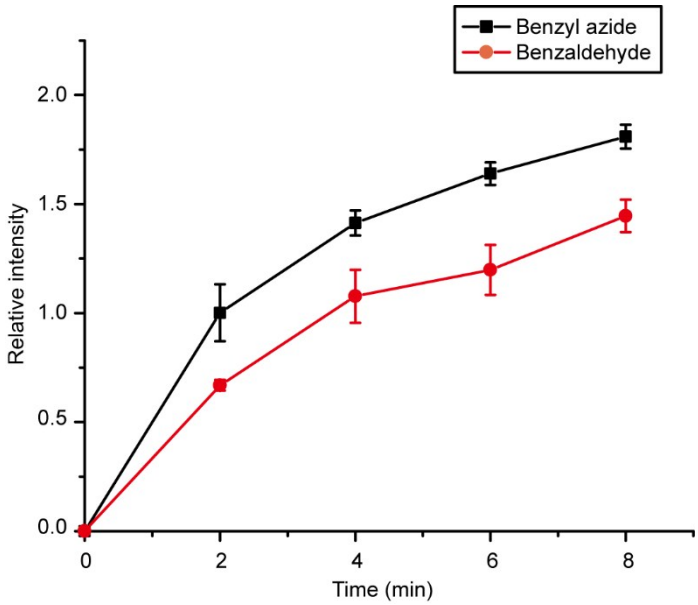
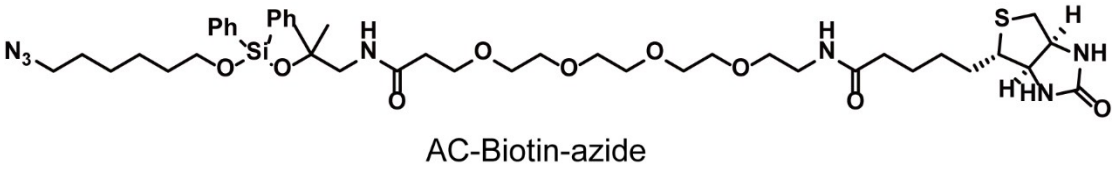


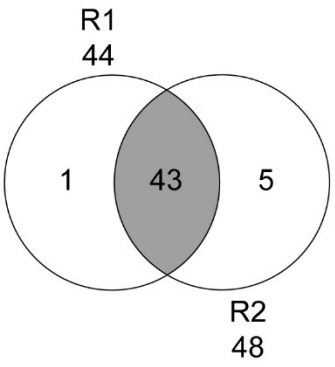
Fig. S6 The product generated by reacting benzyl azide or benzaldehyde with B-hcy were detected by LC-MS. The relative intensity of A was shown as the ratio of intensity of A *versus* that of an internal standard (ethyl benzoyltyrosinate). Data represent mean values ± standard deviation; n = 3 per group.



AC-Biotin-azide

Fig. S7 Chemical structure of acid cleavable biotin azide (AC-Biotin-azide).

A

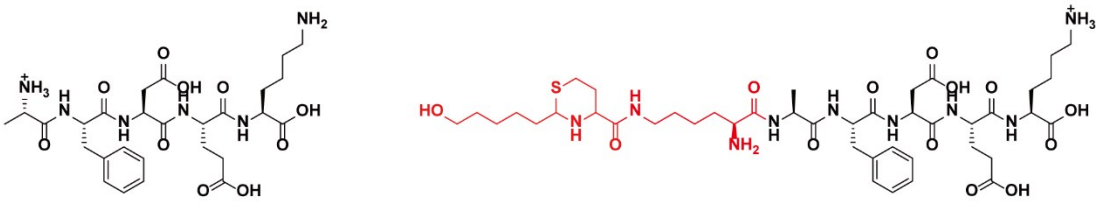


B

-K2----------K28-------K36-------
K44----------K65---------K75----------
--K88----------K100----------K117-------
-----K130-------K138-K140----------K151---
K155K156---K160----------K183--
----------K197------K204------K211----------
----K228------K235---------K245--K248-------
K256------K263--K266----------K285--
---------K297-K299----K304----K309-------K318
----------K336---K340-----K346----------
----------K374----------K386----------K399-
K401----------K412-------K420----------K437-
----------K455-------K463----------
K489-----K495--K498----------K523-----
K528----------K544--K547K548--------K557-
K559-K561------K568----------K580------K587----
------K597---------- C-Terminal

P02769 (ALBU_BOVIN)

C



y5⁺ (Exact mass: 609.28787) y6⁺ (Exact mass: 952.48083)

Fig. S8 Identification of sites of *N*-homocysteinylation from HTL-induced BSA model by LC-MS/MS. (A) Venn diagram showing the number of sites of *N*-homocysteinylation identified across two biological replicates. (B) The total Lys residues and *N*-homocysteinylated Lys residues denoted in red. (C) Chemical structures and exact mass of y5+ and y6+ ions. The modification adduct is denoted in red (cal. m/z 343.19296; obs. m/z 343.1773).

Table S2 List of identified sites of *N*-homocysteinylation from HTL-induced BSA.

III Chemical Synthesis

General Methods

Unless otherwise noted, all chemicals, including anhydrous solvents, are purchased from Sigma-Aldrich or J&K. NMR spectra are recorded on Bruker UltraShield 400 MHz spectrometer. Chemical shifts are recorded in ppm relative to tetramethylsilane (TMS, ^{1}H, 0 ppm) or solvent signals: $CDCl_3$ (^{1}H, 7.26; ^{13}C, 77.16 ppm). Peaks are reported as follows: chemical shift, multiplicity (s = singlet, bs = broad singlet, d = doublet, t = triplet, q =quartet, m = multiplet), coupling constant(s) (Hz). High-resolution tandem mass spectra were obtained on Q Exactive (ThermoFisher Scientific).

Alkynyl aldehyde probe (AAP)

AAP was synthesized according to the reference[3]. In brief, pyridinium chlorochromate (PCC) (9 mmol, 1.9 g) and sodium acetate (0.9 mmol, 72.5 mg) were added to a round- bottom flask containing CH_2Cl_2 (10 ml). The reaction mixture was cooled to 0℃. Then Hept-6-yn-1-ol (6 mmol, 673 mg) was added and the mixtures were allowed to stir for 5 h. Then the mixtures were diluted with diethyl ether (15 ml) and the solvent was removed at ~10℃ by rotary evaporation under vacuum. After purification by silica gel chromatography, the product AAP was obtained as colorless oil (143 mg, yield 22%). The ^{1}H NMR data were similar to the reference. The proton of aldehyde (δ 9.79) and residual solvent peaks (ethyl acetate) were detected. ^{1}H NMR (400 MHz, $CDCl_3$) δ 2.63-2.38 (m, 2H), 2.30-2.17 (m, 2H), 1.98-1.93 (m, 1H), 1.87-1.79 (m, 2H), 1.77-1.53 (m, 2H).

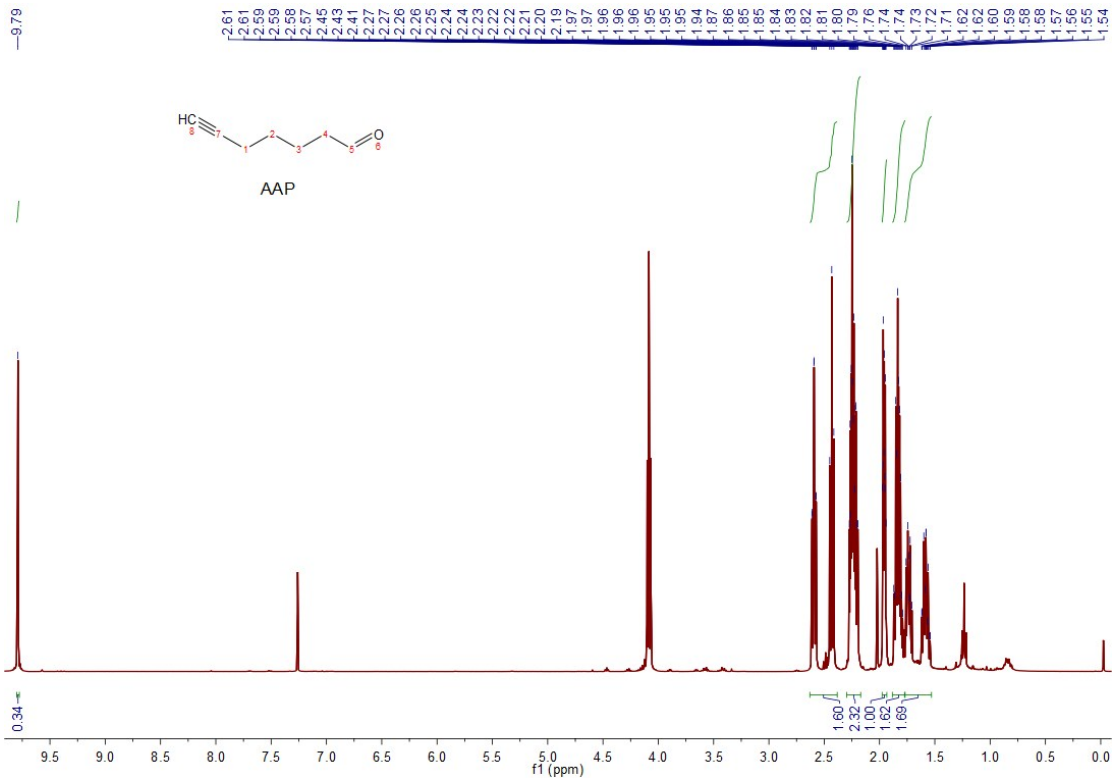


Fig. S9 ^{1}H NMR spectrum of AAP.

Product A (BPTC)

Scalable BPTC were prepared via the conjugation between B-hcy and benzaldehyde. B-hcy (6.25 mmol, 1.4 g) were dissolved into 55 ml PBS (pH 10, 9% methanol). Then 10 equivalent β-ME and 3 equivalent benzaldehyde were added into the solution. The mixtures were stirred at room temperature for 4 h. The reaction was monitored by LC-MS. After completion, the mixtures were filtered and washed with methanol. The crude products were vapored under vacuum and product BPTC were isolated as white solid (660 mg, 2.1 mmol, yield 34%) by preparative HPLC (Waters). ^{1}H NMR (400 MHz, DMSO) δ 8.32 (t, *J* = 5.6 Hz, 1H), 7.51-7.14 (m, 10H), 5.29 (d, *J* = 12.3 Hz, 1H), 4.29 (d, *J* = 4.0 Hz, 2H), 3.52 (t, *J* = 11.4 Hz, 1H), 3.28-2.86 (m, 2H), 2.52-2.46 (m, 2H), 2.08 (d, *J* = 12.7 Hz, 1H) ^{13}C NMR (100MHz, DMSO) δ 171.71, 141.04, 139.43, 128.31, 128.24, 127.72, 127.18, 126.72, 126.58, 64.64, 60.55, 41.89, 29.18, 27.55. HRMS *m/z* calc'd for $C_{18}H_{20}N_2OS$ $[M+H]^+$ 313.13746, found: 313.13674.

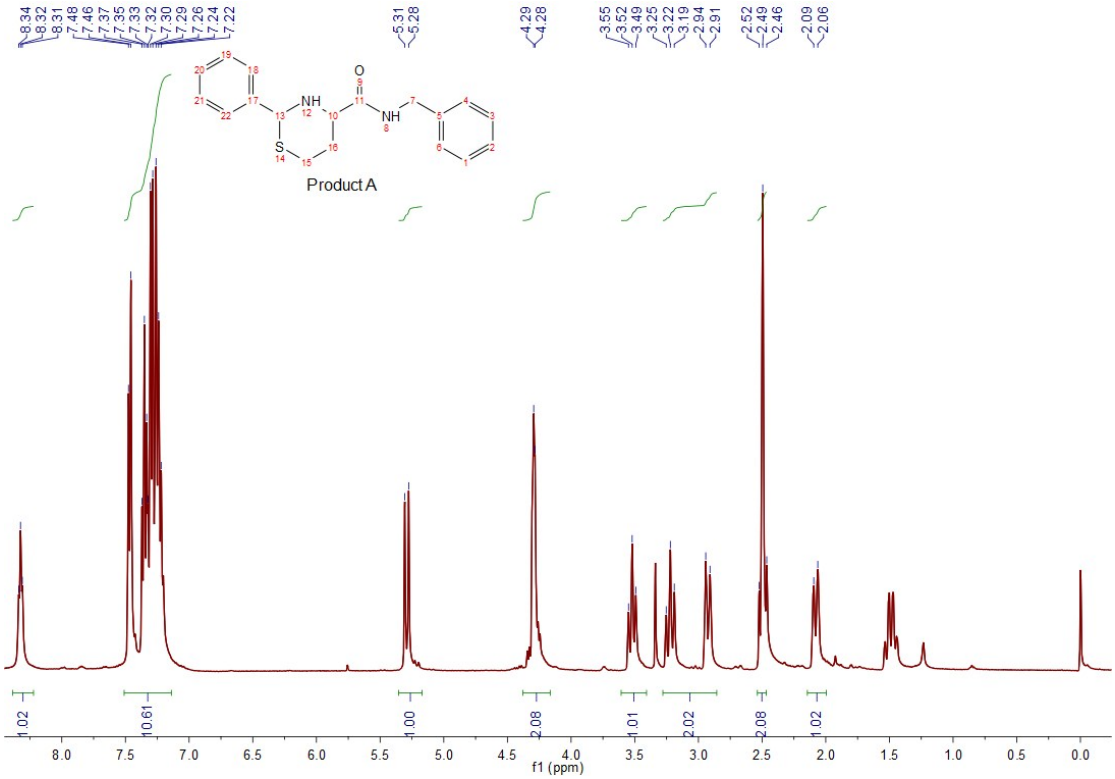


Fig. S10 ^{1}H NMR spectrum of Product A (BPTC).

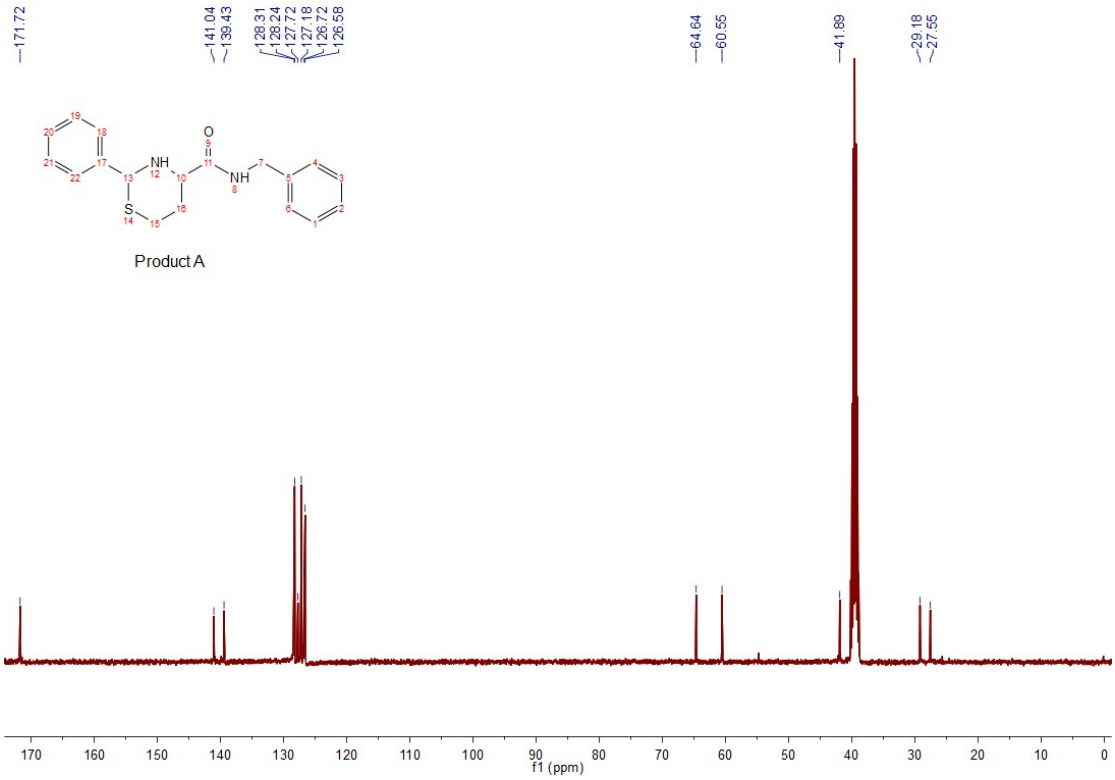


Fig. S11 ^{13}C NMR spectrum of Product A (BPTC).

IV Reference

1 N. Chen, J. Liu, Z. Qiao, Y. Liu, Y. Yang, C. Jiang, X. Wang and C. Wang, *Chem Sci.* 2018, **9**, 2826.

2 H. Chi, C. Liu, H. Yang, W. F. Zeng, L. Wu, W. J. Zhou, R. M. Wang, X. N. Niu, Y. H. Ding, Y. Zhang, Z. W. Wang, Z. L. Chen, R. X. Sun, T. Liu, G. M. Tan, M. Q. X. Dong, P., P. H. Zhang and S. M. He, *Nat Biotechnol.* 2018.

3 X. Cheng, L. Li, M. Uttamchandani and S. Q. Yao, *Org Lett.* 2014, **16**, 1414.