Supporting Information for

Quantitative profiling of PTM stoichiometry by resolvable mass tags

Ying Chen,[#acde] Baiyi Quan,[#acde] Yuanpei Li,[#acde] Yuan Liu,[acde] Wei Qin [bcde] and Chu Wang *[abcde]

[a.]College of Chemistry and Molecular Engineering, Peking University, Beijing, China.

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

[c.]Beijing National Laboratory for Molecular Sciences, Peking University, Beijing, China.

[d.]Synthetic and Functional Biomolecules Center, Peking University, Beijing, China.

[e.]Key Laboratory of Bioorganic Chemistry and Molecular Engineering of Ministry of Education, Peking University, Beijing, China

#These authors contributed equally to this work

*chuwang@pku.edu.cn

This file contains Supplementary Table, Supplementary Dataset, Supplementary Figure 1-7, Supplementary Methods and Supplementary References.

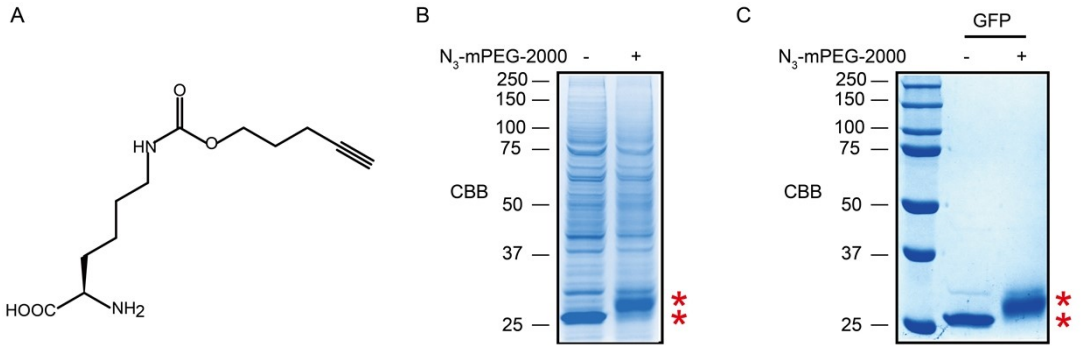
Supplementary Table (Excel):

Stoichiometry of HNEyne modification determined by STO-MS strategy

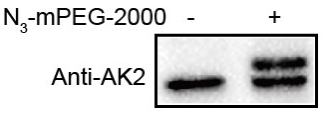
Supplementary Dataset (PDF):

Raw maps for determining stoichiometry of HNEyne modification by STO-MS strategy

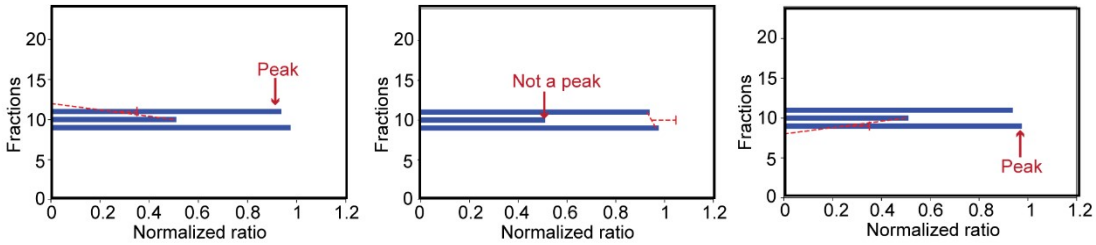
Supplementary Figures:



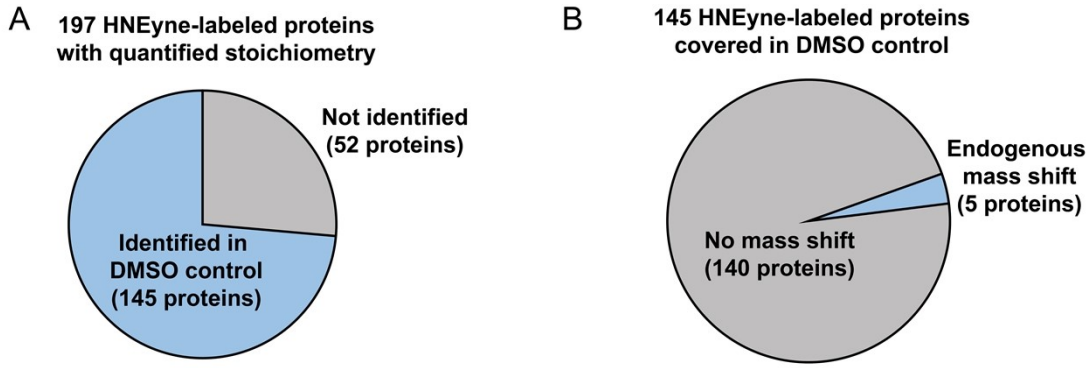
SI figure 1. A. Chemical structure of the unnatural amino acid used in this work; B. 100% of mass shift was observed when the purified GFP was spiked in HEK293T cell lysates; C. Full gel for figure 2G.



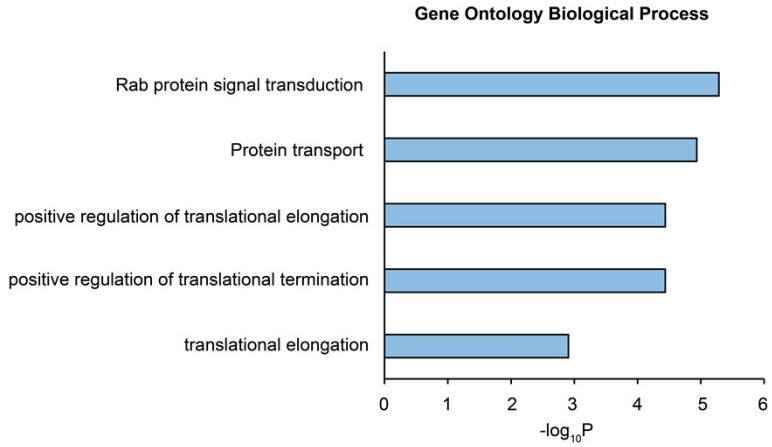
SI figure 2. HNE stoichiometry on endogenous AK2 was determined to be 49.5% at the concentration of 100 μM HNEyne.



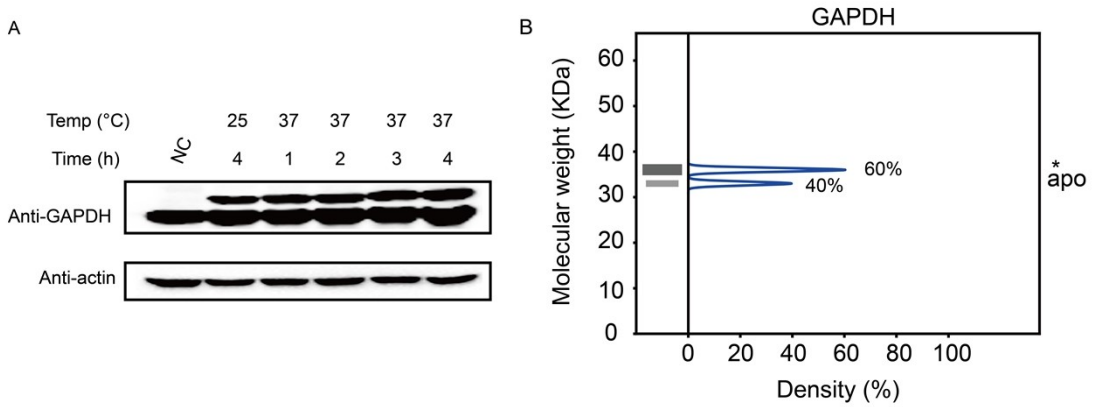
SI figure 3. Peak assignment by the peak-searching algorithm.



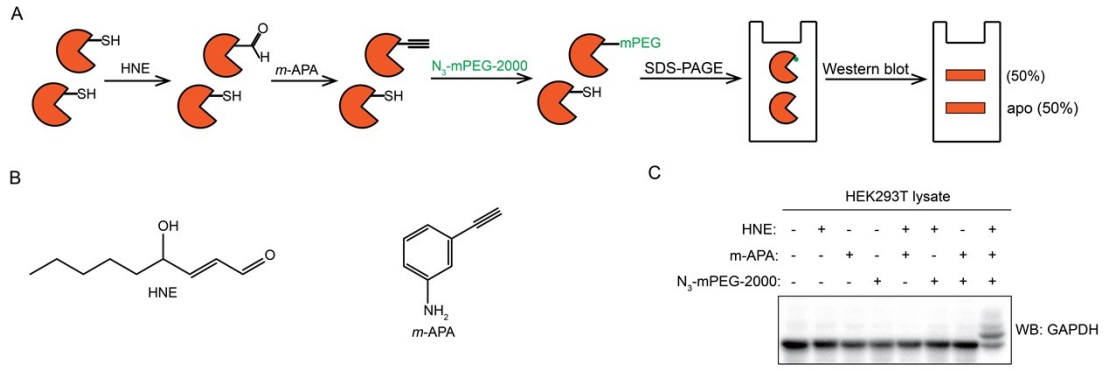
SI figure 4. A. 145 HNEyne-labeled proteins with quantified stoichiometry are identified in the HNEyne negative control sample; B. The majority of HNEyne-labeled proteins have no mass shift in negative control sample.



SI figure 5. Gene Ontology analysis of heavily HNEyne-modified proteins on biological processes.



SI figure 6. A. The stoichiometry of HNEyne modifications on endogenous GAPDH can be quantified by mass resolvable tag strategy; B. GAPDH was determined to mono-modified with stoichiometry of 60%.



SI figure 7. A. The scheme of *m*-APA-based method to measure stoichiometry. HNE modified proteins were react with *m*-APA probe, clicked with azide-mPEG-2000 and analyzed by SDS-PAGE. After transfer, the stoichiometry of individual protein was detected by its antibody. B. Chemical structure for HNE and *m*-APA probe. C. Validation of the reaction between *m*-APA labeling and azide-mPEG-2000 to quantify the stoichiometry of natural HNE modification on GAPDH.

Supplementary Methods

Cell culture

Cells were cultured at 37℃ under 5% CO_2 atmosphere in DMEM culture medium supplemented with 10% FBS and 1% P/S. For SILAC based LC-MS/MS experiment, HEK293T cells were passaged for seven times in DMEM for SILAC supplemented with 10% FBS, 1% P/S, 100 μg/mL L-arginine-HCl and L-lysine-HCl or [^{13}C6, ^{15}N4] L-arginine-HCl and [^{13}C6, ^{15}N2] L-lysine-HCl. The resulting "light" and "heavy" cells were then used for LC-MS/MS experiments.

Optimizing click reaction using BSA and HNEyne probe

4 mg/mL of BSA in PBS was incubated with 10 mM TCEP for 30 min at RT and then free TCEP was removed using NP-50 columns by following manufacturer's instructions. BSA was incubated with 100 μM HNEyne for 1 h at RT and then reacted with 1 mM Cu-BTTAA complex (Cu:BTTAA=1:2), 10 mM N_3-mPEG-2000 and 3 mM NaVc by conditions shown in Figure 2D. Afterwards, the samples were added with 5x sample buffer, resolved by 10% SDS-PAGE and stained by Coomassie blue.

Quantifying the efficiency of click reaction using competitive assay

1 mg/mL of HEK293T cell lysates were incubated with 100 μM HNEyne for 4 h at RT. Afterwards, free HNEyne was discarded using methanol/chloroform precipitation and the proteomes were re-suspended in 2% SDS/PBS. For the first run of click reaction, proteomes were reacted with 1 mM Cu-BTTAA complex (Cu:BTTAA=1:2), 1, 2, 4, 8 or 10 mM N_3-mPEG-2000 and 3 mM NaVc for 4 h at 37℃. Afterwards, the samples were precipitated again using methanol/chloroform mixture and re-suspended in 0.4% SDS/PBS. For the second run of click reaction, proteomes were reacted with 1 mM Cu-BTTAA complex (Cu:BTTAA=1:2), 400 μM TAMRA-azide and 3 mM NaVc for 1 h at RT. Finally, all samples were added with 5x sample buffer and resolved by 10% SDS-PAGE.

Purification of unnatural amino acid (UAA) inserted GFP

E.coli cells expressing UAA inserted GFP were sonicated in 10% glycerol/PBS for 30 min and the lysates were centrifuged at 8000 g for 15 min under 4℃. Ni-NTA column was washed with 8 column volumes (CV) of PBS and then loaded with the lysates. Afterwards, Ni-NTA column was sequentially washed with 6 CV of 40 mM imidazole/PBS, 2 CV of 75 mM imidazole/PBS, 3 CV of 250 mM imidazole/PBS, 6 CV of 500 mM imidazole/PBS and finally 2 CV of water. The protein purity was determined by SDS-PAGE and then high-quality samples were collected. Finally, imidazole in the proteins was removed through ultrafiltration and proteins were aliquoted and stored at -80℃.

Quantifying the efficiency of click reaction using UAA inserted GFP

Purified GFP was diluted to final concentration of 0.5 mg/mL using PBS or HNEyne labeled HEK293T cell lysates. Then samples were reacted with 1 mM Cu-BTTAA complex (Cu:BTTAA=1:2), 10 mM N_3-mPEG-2000 and 3 mM NaVc for 4 h at 37℃. Afterwards, samples

were added with 5x sample buffer, resolved by 10% SDS-PAGE and stained by Coomassie blue.

Determining the stoichiometry of HNE modification on ZAK

HEK293T cells which stably overexpressing 6xhis-ZAK^{WT} and 6xhis-ZAK^{C22A} were lysed in 0.1% triton/PBS via sonication. Cell lysates were centrifuged at 20,000 g for 30 min under 4℃. Afterwards, the proteome concentration was determined by BCA assay and normalized to 3 mg/mL. The proteomes were incubated with 0, 5, 20, 100 μM HNEyne for 1 h at RT. The proteomes were precipitated using cold methanol and re-suspended in 2% SDS/PBS. Afterwards, click reaction was performed as above and then samples were added with 5x sample buffer and resolved by 10% SDS-PAGE. Finally, the gel was transferred and the membrane was incubated with anti-his antibody.

Determining the stoichiometry of HNE modification on AK2

3 mg/mL HEK293T cell lysates were incubated with 100 μM HNEyne for 1 h at RT. The lysates were precipitated using cold methanol and re-suspended in 2% SDS/PBS. Afterwards, click reaction was performed as above and then samples were added with 5x sample buffer and resolved by 10% SDS-PAGE. Afterwards, the gel was transferred and the membrane was incubated with anti-AK2 antibody.

Determining the stoichiometry of HNE modification on GAPDH by *m*-APA labeling

3 mg/mL of HEK293T cell lysates were incubated with 100 μM HNE for 1 h at RT. Then proteomes were labeled by 0.5 mM *m*-APA in the presence of 60 mM $NaBH_3CN$ at pH 5 for 1 h. Free *m*-APA was discarded using methanol/chloroform precipitation and the proteomes were washed using cold methanol and resuspended in 2% SDS/PBS. Proteomes were reacted with 1 mM Cu-BTTAA complex (Cu:BTTAA=1:2), 10 mM N_3-mPEG-2000 and 3 mM NaVc for 4 h at 37℃. Afterwards, the samples were precipitated again by methanol/chloroform mixture and re-suspended in 0.4% SDS/PBS. Finally, samples were added with 5x sample buffer and resolved by 10% SDS-PAGE. Afterwards, the gel was transferred and the membrane was incubated with anti-GAPDH antibody.

Determining the stoichiometry of endogenous carbonylation on ZAK by *m*-APA labeling

1×10^7 HEK293T cells which stably overexpressing 6xhis-ZAK^{WT} were seeded in 10 cm dish to grow overnight and then incubated with 1, 10, 50, 100, 200 μM arachidonic acid (AA) or H_2O_2 for 30 min in serum-free medium. Cells were collected and lysed in ice-cold 0.1% triton/PBS. Then the lysates were centrifuged at 20,000 g for 30 min under 4℃ and protein concentration was determined by BCA kit and normalized to 1.5 mg/mL.

Proteomes were labeled by 0.5 mM *m*-APA in the presence of 60 mM $NaBH_3CN$ at pH 5 for 1 h. Free *m*-APA was discarded using methanol/chloroform precipitation and the proteomes were washed using cold methanol and resuspended in 2% SDS/PBS. Proteomes were reacted with 1 mM Cu-BTTAA complex (Cu:BTTAA=1:2), 10 mM N_3-mPEG-2000 and 3 mM NaVc for 4 h at 37℃. Afterwards, the samples were precipitated again by methanol/chloroform mixture and re-suspended in 0.4% SDS/PBS. Samples were added with 5x sample buffer and resolved by 8% SDS-PAGE. Finally, the gel was transferred and the membrane was incubated with anti-his antibody.

Application of STO-MS to globally quantify the stoichiometry of HNE modification

HEK293T SILAC cells were lysed in cold PBS by sonication and the lysates were centrifuged at 20,000 g for 30 min under 4℃. Afterwards, proteome concentration was determined by BCA assay and normalized to 1 mg/mL. 1 mL of proteomes from light and heavy cells were incubated with 100 μM HNEyne for 4 h at RT. For negative control sample, proteomes from light and heavy cells were incubated with DMSO for 4 h at RT. Afterwards, proteomes were precipitated with methanol/chloroform mixture and re-suspended in 2% SDS/PBS. Proteomes were then reacted with 1 mM Cu-BTTAA complex (Cu:BTTAA=1:2), 10 mM N_3-mPEG-2000 and 3 mM NaVc for 4 h at 37℃. Proteomes were precipitated again with methanol/chloroform mixture and re-suspended in 0.4% SDS/PBS. Afterwards, heavy and light samples were resolved by 12% SDS-PAGE (17.5 x 11.5 cm) with the loading ratio of 1:3. Then the gel was cut into 23 slices and each slice was further cut into 1 mm x 1 mm cubes for following in-gel trypsin digestion. The gel cubes were incubated with 500 μL of 50 mM NH_4HCO_3 (containing 50% ACN) for 15 min and then with 500 μL ACN for 10 min. Afterwards, gel cubes were incubated with 50 μL of 10 mM DTT at 56℃ for 30 min and further treated with 50 μL of 55 mM iodoacetamide for 60 min in dark. Gel cubes were incubated with 500 μL ACN again for 30 min. Finally, gel cubes were treated with 100 μL of 10 ng/μL trypsin for 12 h at 37℃. Gel cubes were incubated with 2x100 μL of 50 mM NH_4HCO_3 (containing 50% ACN and 5% formic acid) for 30 min to extract digested peptides. All peptides from light sample were combined and equally distributed to each fraction of heavy sample, then the samples were dried by speed-vacuum and stored at -80℃ until LC-MS/MS analysis.

LC-MS/MS data analysis

The samples were reconstituted in 10 μL of 0.1% FA/H_2O and centrifuged at 20000 g for 30 min. Afterwards, 6.6 μL was analyzed by Q Exactive mass spectrometer. Full-scan mass spectra were acquired with m/z range of 350-1800 and resolution of 70000. TOP 20 most abundant ions were selected for MS2 spectra acquisition with a resolution of 17500 using collision mode of HCD. Other important parameters were set as following: isolation window, 2.0 m/z; normalized collision energy, 28%; dynamic exclusion, 20.0 s.

The LC-MS/MS data were searched by MaxQuant software[1]. All 23 fractions were analyzed together with “Experiments” set from 01 to 23 and “Fractions” all set as “1”. Under “Group-specific parameters”, the “Multiplicity” was set as 2, “Lys8” and “Arg10” were set as “Heavy labels”, “Label-free quantification” was set as “LFQ”, “LFQ min number of neighbors” was set as 2, “skip normalization” and “match between run” were checked. Finally, for each identified protein, we got quantified SILAC ratio as well as light and heavy LFQ intensity of all fractions. As a result, we normalized SILAC (L/H) ratios with assumption that light LFQ intensities are equal across all fractions since all light samples were combined as a standard and equally distributed to heavy fractions when we prepared the mass spec sample (detailed codes about normalization can be found on website https://github.com/wangchulab/STO-MS).

The normalized ratios were further analyzed by an in-home python script (detailed codes about the python script can be found on website https://github.com/wangchulab/STO-MS). In detail, a predicted fraction distribution curve was first generated for each protein according to its molecular weight. In principle, for the real sample, the fraction with the highest ratio (the most abundant fraction of the protein) should be close to the predicted fraction (within plus/minus 3 fractions). As a result, proteins which failed to meet the criterion were excluded. Afterwards, the ratios for each protein were normalized against the highest ratio across 23 fractions and proteoforms with different

numbers of HNE modifications were recognized by different "peaks" on the ratio/fraction graph. The peaks were recognized by finding the saddle point which was defined by its ratio should be smaller than the ratios of both former and latter fractions plus 0.2, considering the peaks may overlap. At the same time, only peaks with both area and height larger than 0.1 were considered. After peak recognition, the stoichiometry was quantified by the formula: $R_{modified}/ (R_{modified}+ R_{native})$. To improve the modification identification in the HNEyne sample, data from the DMSO sample is utilized as negative control. All the proteins that are only identified in only one fraction in the DMSO-treated sample are coerced into a "reference list". For any protein identified in the HNE-treated sample and in the reference list, the protein is considered as a modified protein if this protein showed up in more than one consecutive fractions. The stoichiometry of the modification is later calculated by a peak-searching algorithm similar as previous described.

Validating the stoichiometry of HNE modification on FBL

1 mg/mL of HEK293T cell lysates were incubated with 100 μM HNEyne for 4 h at RT. Afterwards, the proteomes were precipitated with methanol/chloroform mixture and re-suspended in 2% SDS/PBS. Afterwards, click reaction was performed as above and then the samples were precipitated again with methanol/chloroform mixture and re-suspended in 0.4% SDS/PBS. The samples were added with 5x sample buffer and resolved by 10% SDS-PAGE. Afterwards, the gel was transferred and the membrane was incubated with anti-FBL antibody.

Validating the stoichiometry of HNE modification on GAPDH

HepG2 cells which stably overexpressing 6xhis-$GAPDH^{WT}$, 6xhis-$GAPDH^{C152A}$ or 6xhis-$GAPDH^{C247A}$ were lysed in 0.1% triton/PBS via sonication. Then the lysates were centrifuged at 20,000 g for 30 min under 4℃. Afterwards, the concentration was determined by BCA kit and normalized to 3 mg/mL. The proteomes were incubated with 100 μM HNEyne for 1 h at RT. Then the samples were precipitated using cold methanol and re-suspended in 2% SDS/PBS. Afterwards, click reaction was performed as above and then the samples were added with 5x sample buffer and resolved by 10% SDS-PAGE. Finally, the gel was transferred and the membrane was incubated with anti-his antibody.

Supplementary References

1. J. Cox and M. Mann, *Nat. Biotechnol.*, 2008, **26**, 1367-1372.