Profiling Genome-wide Chromatin Methylation with Engineered Posttranslation Apparatus within Living Cells

Rui Wang[1,2,§], Kabirul Islam[1, §], Ying Liu[3], Weihong Zheng[1], Haiping Tang[4], Nathalie Lailler[5], Gil Blum[1,6], Haiteng Deng[4], Minkui Luo[1,*]

[1]Molecular Pharmacology and Chemistry Program, Memorial Sloan-Kettering Cancer Center, New York, NY 10065; [2]Program of Pharmacology, Weill Graduate School of Medical Science, Cornell University, New York, NY 10021; [3]Program of Biochemistry & Structural Biology, Cell & Developmental Biology, and Molecular Biology, Weill Graduate School of Medical Science, Cornell University, New York, NY 10021; [4]School of Life Sciences, Tsinghua University, Beijing, 100084 China; [5]Genomics Core Laboratory, Memorial Sloan-Kettering Cancer Center, New York, NY 10065; [6]Tri-Institutional Training Program in Chemical Biology, Memorial Sloan-Kettering Cancer Center, New York, NY 10065.

Supporting Information

Supplementary Methods

1. General materials and methods. Unless noted otherwise, all chemicals were purchased from Aldrich Chemical and used without further purification. HPLC grade reagents were obtained from Fisher Scientific; all solvents were pre-filtered and degassed prior to use. Analytical HPLC was carried out on a Waters 600 Controller HPLC/2998 diode array detector using XBridge™ Prep C18 5μm 4.6×150mm reverse phase column. Preparative HPLC purification was carried out on DELTA PAK C18 15μm, 300A 300×3.9 mm column. Concentration and lyophilization of aqueous solution of SAM analogues were performed with Savant Sc210A SpeedVac Concentrator (Thermo), followed by Flexi-Dry™ μP Freeze-Dryer (FTS™ System). Histone H3 peptide (amino acids 1-21, ARTKQTARKSTGGKAPRKQLA) and RGG peptide (GGRGGFGGRGGFGGRGGFG-GGK-biotin) were synthesized with standard Fmoc-protected solid-phase peptide synthesis strategy and purified by HPLC with 0 ~ 40% acetonitrile gradient in 0.1% trifluoroacetic acid/H_2O (the Proteomics Resource Center of the Rockefeller University). The integrity of the purified peptides was confirmed by mass spectrometry.

Nuclear magnetic resonance spectra (NMR) were recorded on Bruker AVIII 500MHz or 600MHz spectrometer. NMR Chemical shifts are reported in ppm; multiplicity are indicated by s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet), dd (doublet of doublet), etc.; coupling constants (*J*) are expressed in Hz; integration is provided. Formic acid-d_2 of 5μL was added into 600μL D_2O as the NMR solvent to increase the solubility of the compounds containing α-amino acid moiety. $^1H/^{13}C$-NMR signals were referenced against solvent peaks (^1H: $CHCl_3$ = 7.26 ppm, residual D_2O = 4.79 ppm; ^{13}C: formic acid-d_2 =166.2 ppm). Mass spectra were collected at the MSKCC Analytical Core Facility on a PE SCIEX API 100 or Waters Acuity SQD LC-MS with electron spray ionization (ESI) probe.

2. Synthesis and characterization of the precursors of methionine analogues.

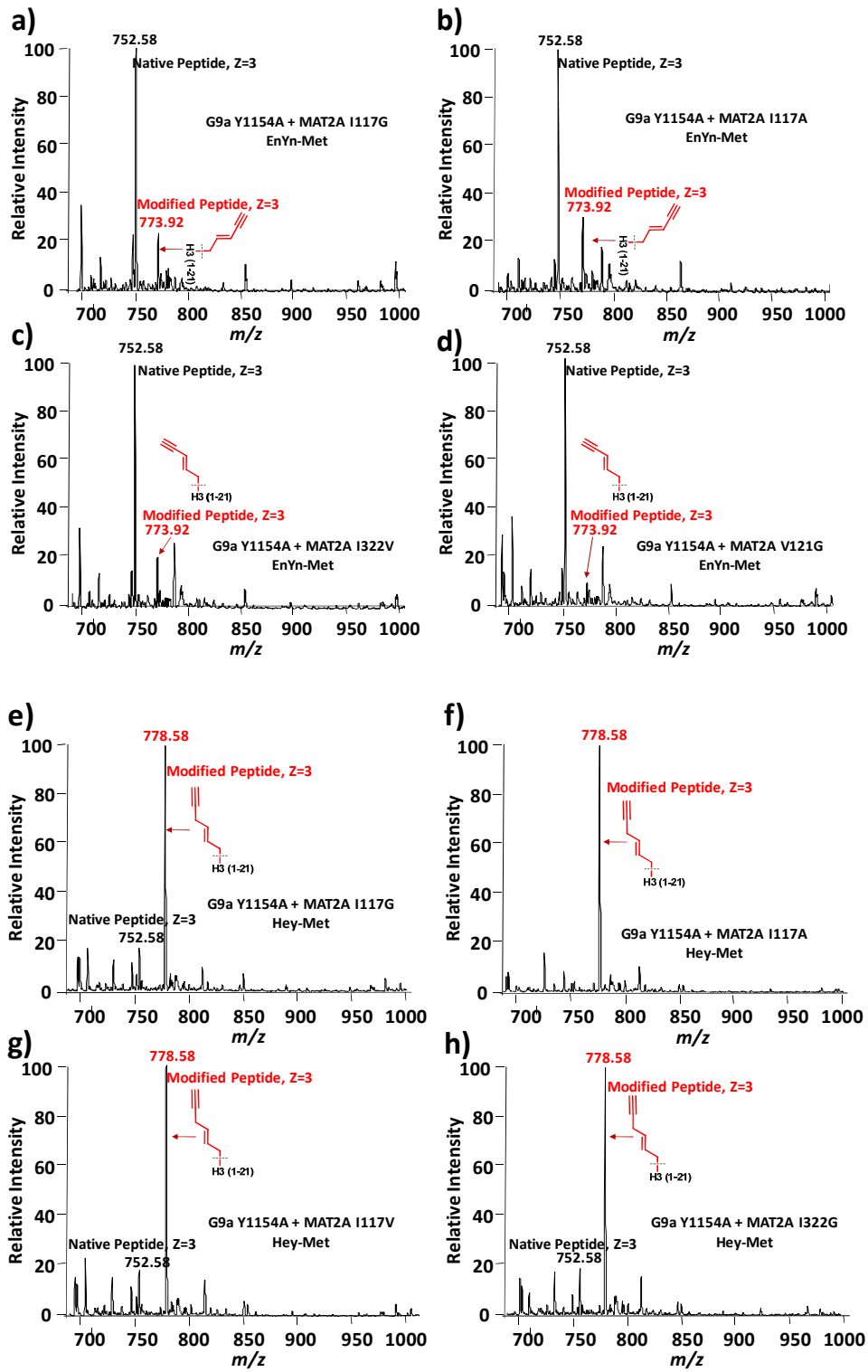
Starting materials: *S*-benzyl-*L*-homocysteine was prepared from commercially available *L*-methionine according to the reported protocol.[1] (*E*)-Pent-2-en-4-yn-1-ol was purchased from Alfa Aesar.

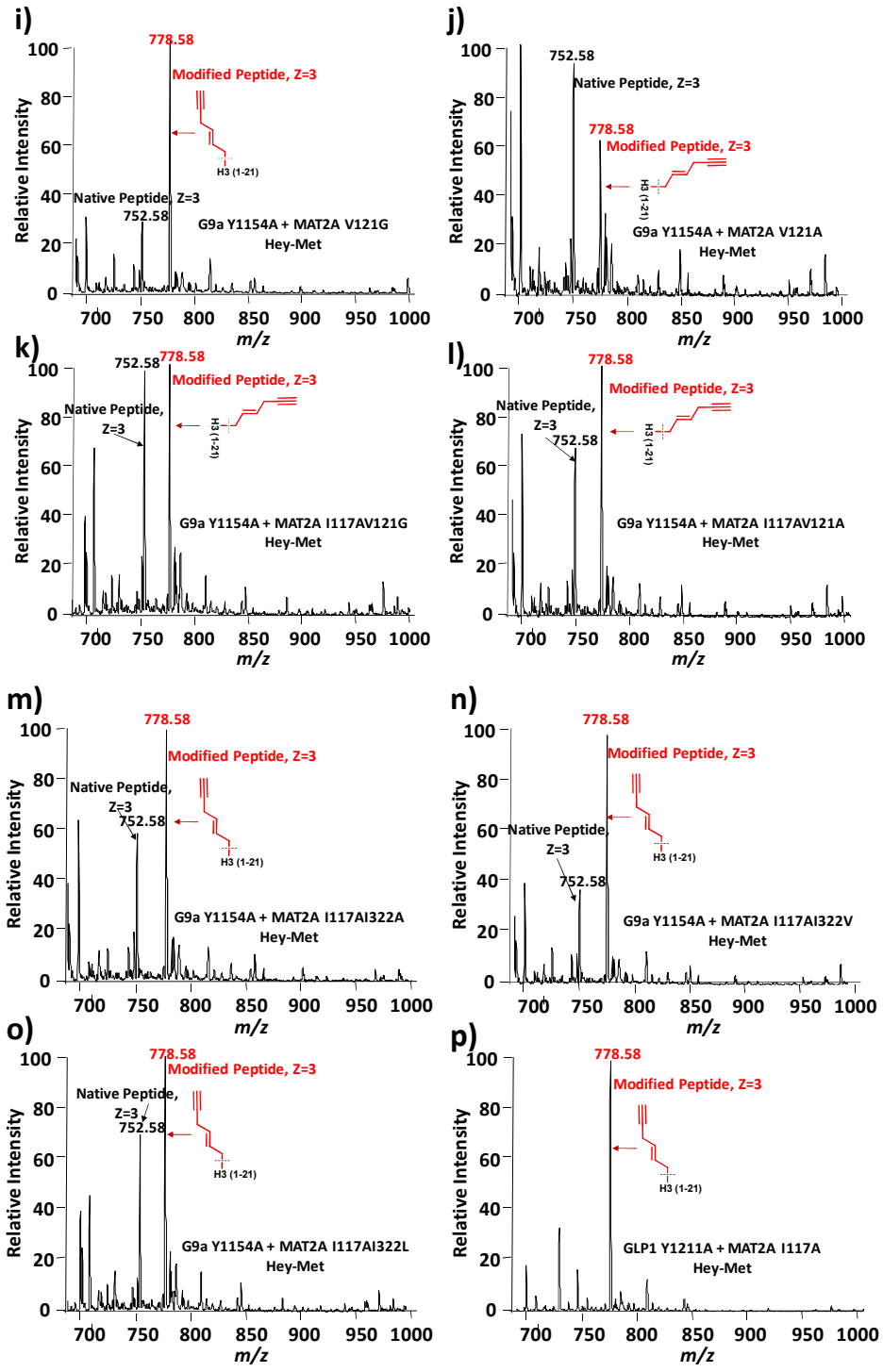
Synthesis of* (*E*)-6-bromohex-4-en-1-yne*: (*E*)-6-bromohex-4-en-1-yne was prepared as described previously.[2]

Synthesis of (E)-pent-2-en-4-ynyl tosylate: Potassium hydroxide (560 mg, 10 mmol) was added to the solution of (*E*)-pent-2-en-4-yn-1-ol (165 mg, 2 mmol) and *p*-toluenesulfonyl chloride (480 mg, 2.4 mmol) in diethyl ether (10 mL) at −5 ºC. The resulting mixture was stirred at this temperature for 1 h, and then stirred at ambient temperature (22 ºC) overnight. The slurry reaction mixture was poured into cold water (20 mL) and separated. The aqueous phase was extracted with diethyl ether (3×10 mL). The combined organic phase was washed with distilled deionized (DD) water, 1 N HCl, brine and then dried with anhydrous $MgSO_4$. Light brown solid (410 mg, 1.7 mmol, 85% yield) was obtained after removing the volatile solvents and used without further purification. ^{1}H-NMR (500MHz, $CDCl_3$): δ 2.45(s, 3H), 2.94(d, 1H, J = 1.9 Hz), 4.57(d, 2H, J = 5.9 Hz), 5.70(dd, 1H, J =15.9, 1.7 Hz) 6.11(d×t, 1H, J =15.9, 6.0 Hz), 7.35(d, 2H, J = 8.3 Hz), 7.79(d, 2H, J = 8.3Hz).

Synthesis of (E)-1-bromo-4-(prop-2-ynyloxy)but-2-ene: Aqueous sodium hydroxide solution (8 N, 6 mL) and tetrabutylammonium bromide (40 mg) were added into the solution of propargyl alcohol (0.59 mL, 10 mmol) and *trans*-1,4-dibromo-2-butene (2.57 g, 12 mmol) in dichloromethane (10 mL). The resulting biphasic reaction was stirred at ambient temperature (22 ºC) for 16 h. The reaction mixture was further diluted with water (30 mL) and dichloromethane (30 ml), and then separated. The aqueous layer was extracted with dichloromethane (2×30 mL). The combined organic phase was washed with water, brine and dried with anhydrous Na_2SO_4. Light brown liquid (610 mg, 3.2 mmol, 32% yield) was obtained after concentration and chromatography (15% dichloromethane/hexane). ^{1}H-NMR (500MHz, $CDCl_3$): δ 2.46(t, 1H, J = 2.4Hz), 4.02(d, 2H, J = 8.4 Hz), 4.16(d, 2H, J = 2.4 Hz), 4.20(dd, 2H, J = 6.6Hz, 1.2 Hz), 5.67-5.71(m, 1H), 5.89-5.94(m, 1H).

B. Supplementary Figures




i)
778.58
Modified Peptide, Z=3
H3 (1-21)
Native Peptide, Z=3
752.58
G9a Y1154A + MAT2A V121G
Hey-Met
j)
752.58
Native Peptide, Z=3
778.58
Modified Peptide, Z=3
H3 (1-21)
G9a Y1154A + MAT2A V121A
Hey-Met
k)
752.58
778.58
Modified Peptide, Z=3
Native Peptide, Z=3
H3 (1-21)
G9a Y1154A + MAT2A I117AV121G
Hey-Met
l)
778.58
Modified Peptide, Z=3
Native Peptide, Z=3
752.58
H3 (1-21)
G9a Y1154A + MAT2A I117AV121A
Hey-Met
m)
778.58
Modified Peptide, Z=3
Native Peptide, Z=3
752.58
H3 (1-21)
G9a Y1154A + MAT2A I117AI322A
Hey-Met
n)
778.58
Modified Peptide, Z=3
Native Peptide, Z=3
752.58
H3 (1-21)
G9a Y1154A + MAT2A I117AI322V
Hey-Met
o)
778.58
Modified Peptide, Z=3
Native Peptide, Z=3
752.58
H3 (1-21)
G9a Y1154A + MAT2A I117AI322L
Hey-Met
p)
778.58
Modified Peptide, Z=3
H3 (1-21)
GLP1 Y1211A + MAT2A I117A
Hey-Met
Relative Intensity
m/z

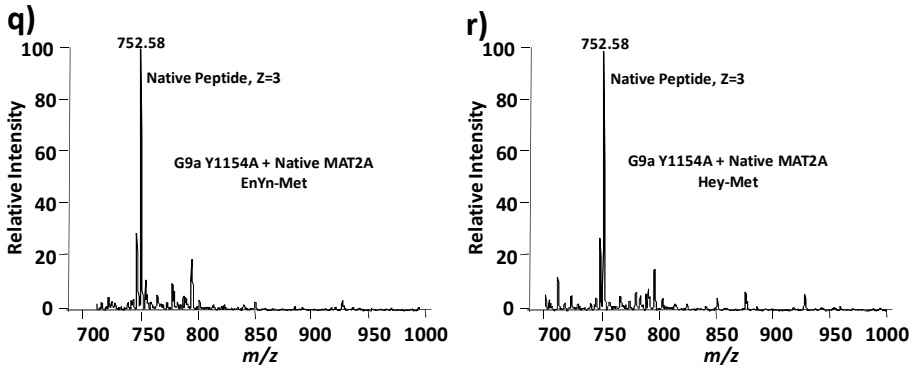


Figure S1. Modification of histone H3K9 peptide in the present of ATP, a methionine analogue (**1b** or **1c**), a MAT2A mutant and an engineered PMT (G9a Y1154A or GLP1 Y1211A). The products from the coupling reactomes were analyszed with LC- ESI-MS as described above. **(a-d)** Product profiles of EnYn-Met **1b** and MAT2A mutants (I17G, I117A, V121G and I322V) coupled with G9a Y1154A mutant. **(e-o)** Product profiles of Hey-Met **1c** and MAT2A mutants (I117G, I117A, I117V, V121G, V121A, I322G, I117AV121G, I117AV121A, I117AI322A, I117AI322V and I117AI322L) coupled with G9a Y1154A mutant. **(p)** Product profiles of Hey-Met **1c** and MAT2A I117A mutant coupled with GLP1 Y1211A mutant. **(q-r)** Product profiles of native MAT2A and EnYn-Met **1b** or Hey-Met **1c** coupled with G9a Y1154A mutant. Under these conditions, no alkylated product was detected, indicating that native MAT2A, unlike the MAT2A I117A mutant, cannot efficiently process the methionine analogues. The same results were obtained with other inactive MAT2A mutants (*e.g.* I117L, V121L, I322A, I322L, I117AI322G) (data not shown).

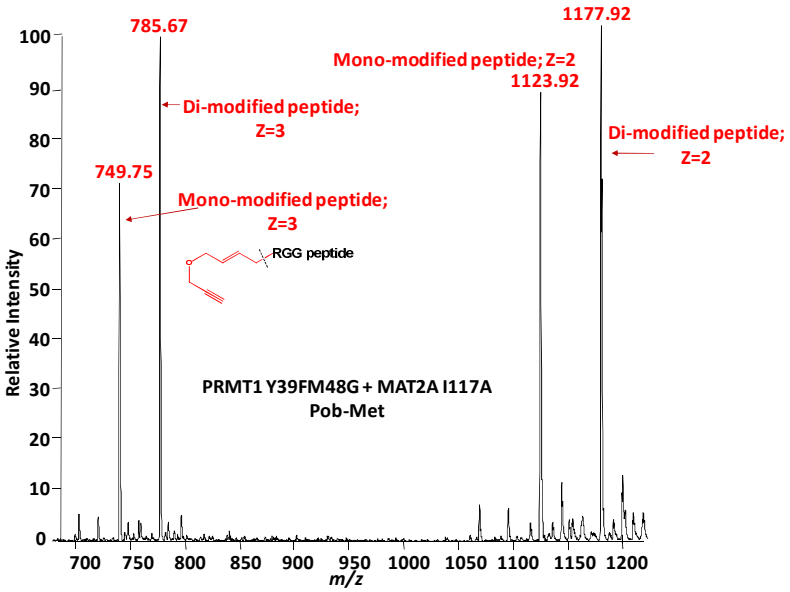


Figure S2. Representative LC-MS trace of RGG peptide modification with the reactome containing ATP, Pob-methionine analogue, the MAT2A I117A mutant and the PRMT1 Y39FM48G mutant.

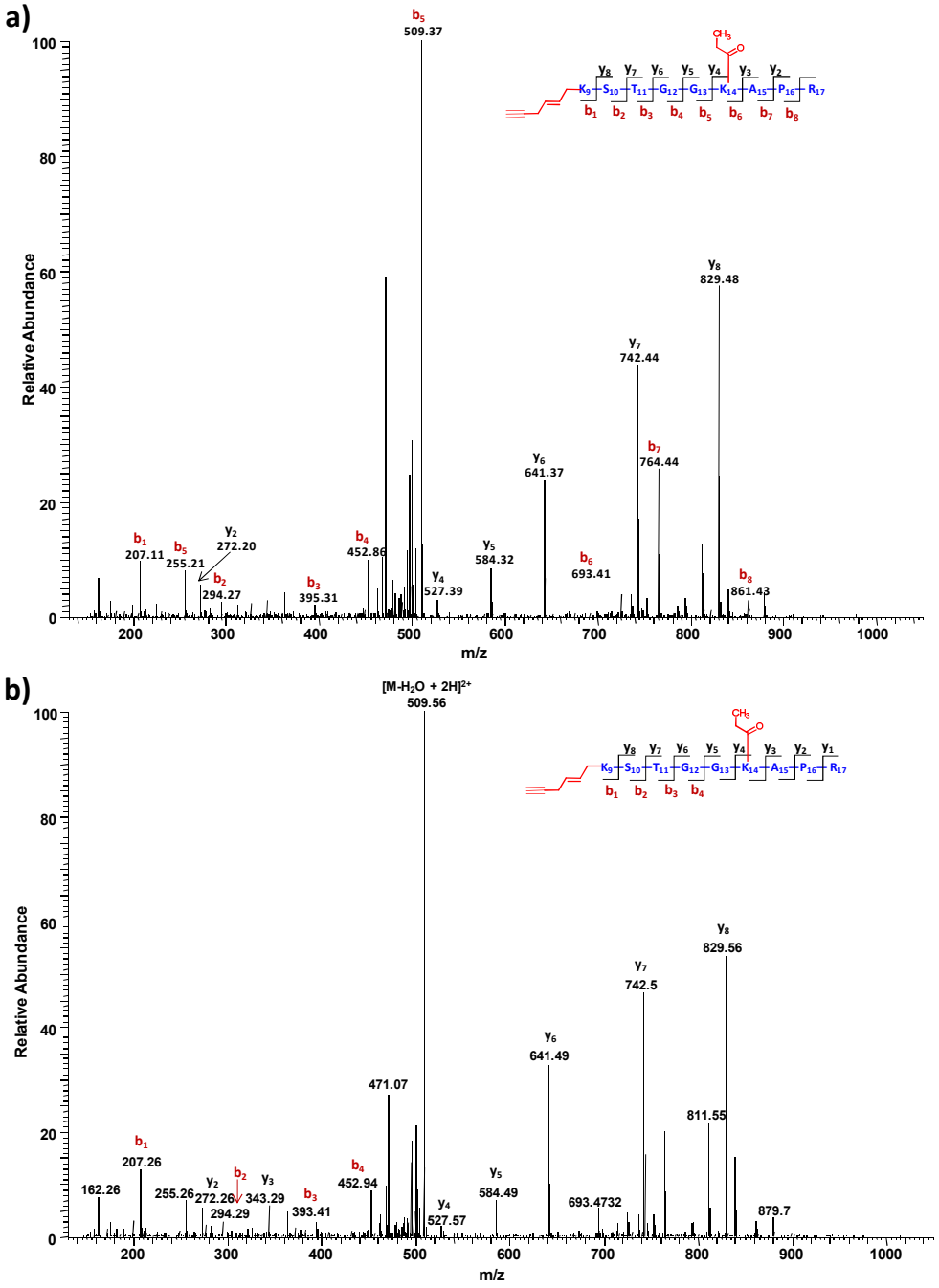


Figure S3. Tandem mass spectra of the modification on recombinant histone H3. MAT2A I117A can exploit Hey-Met **1c** and generate Hey-SAM **2c**. Hey-SAM can be further processed by **(a)** G9a's Y1154A mutant or **(b)** GLP1's Y1211A mutant *in vitro* and yield the corresponding H3K9-(*E*)-hex-2-en-5-ynylated histones as confirmed by the expected MS/MS peaks.

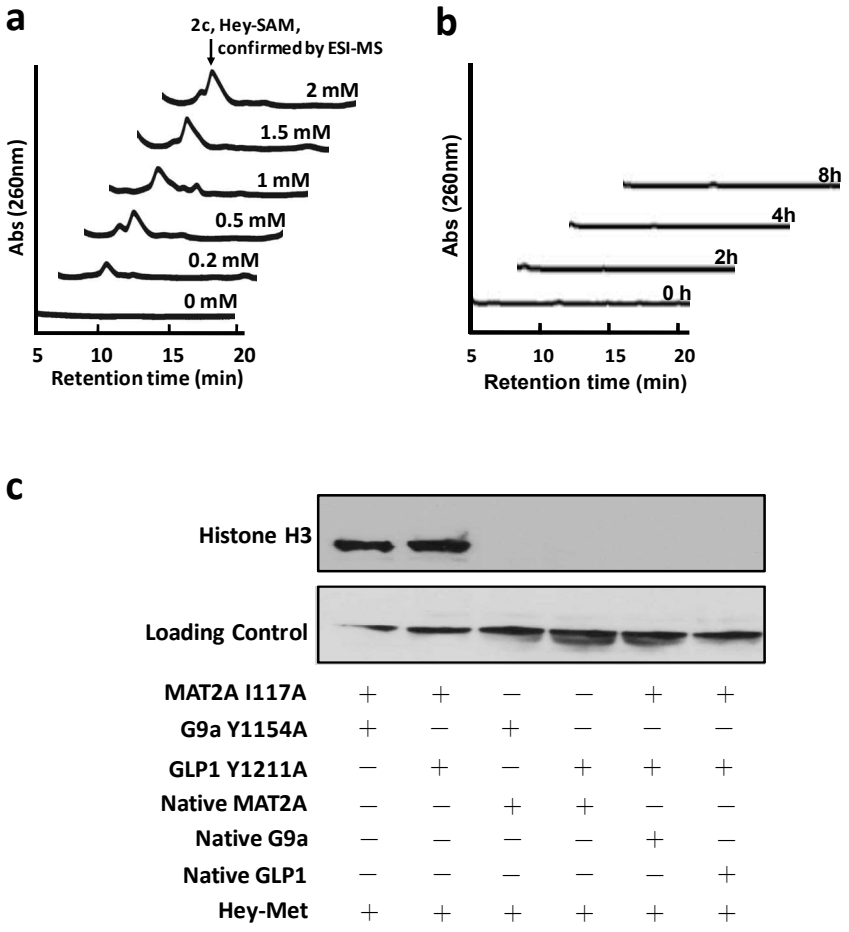


Figure S4. Synthesis of Hey-SAM **2c** and Hey-SAM-mediated histone alkylation within living cells. **(a)** Production of Hey-SAM **2c** within living cells. HEK293T cells were transfected with the MAT2A I117A and supplied with various concentrations of Hey-Met **1c**. The resultant production of Hey-SAM **2c** was detected by HPLC analysis as shown in Figure 3a. **(b)** No production of Hey-SAM **2c** in HEK293T cells that were transfected with native MAT2A, followed by the treatment of 2 mM Hey-Met **1c** for varied periods of time. **(c)** Enrichment of (*E*)-hex-2-en-5-ynylated histone H3. HEK293T cells were co-transfected with the native G9a/GLP1 or the G9a/GLP1 Y1154A/Y1211A mutants and native MAT2A or the MAT2A I117A mutant. After incubating HEK293T cells with 1 mM Hey-Met **1c** under a methionine-depleted condition, the resultant lysate proteins were conjugated with the azido-diazo-biotin probe, enriched with streptavidin beads, blotted with anti-H3 antibody blotting (the upper panel). The lower panel was shown as the loading control.

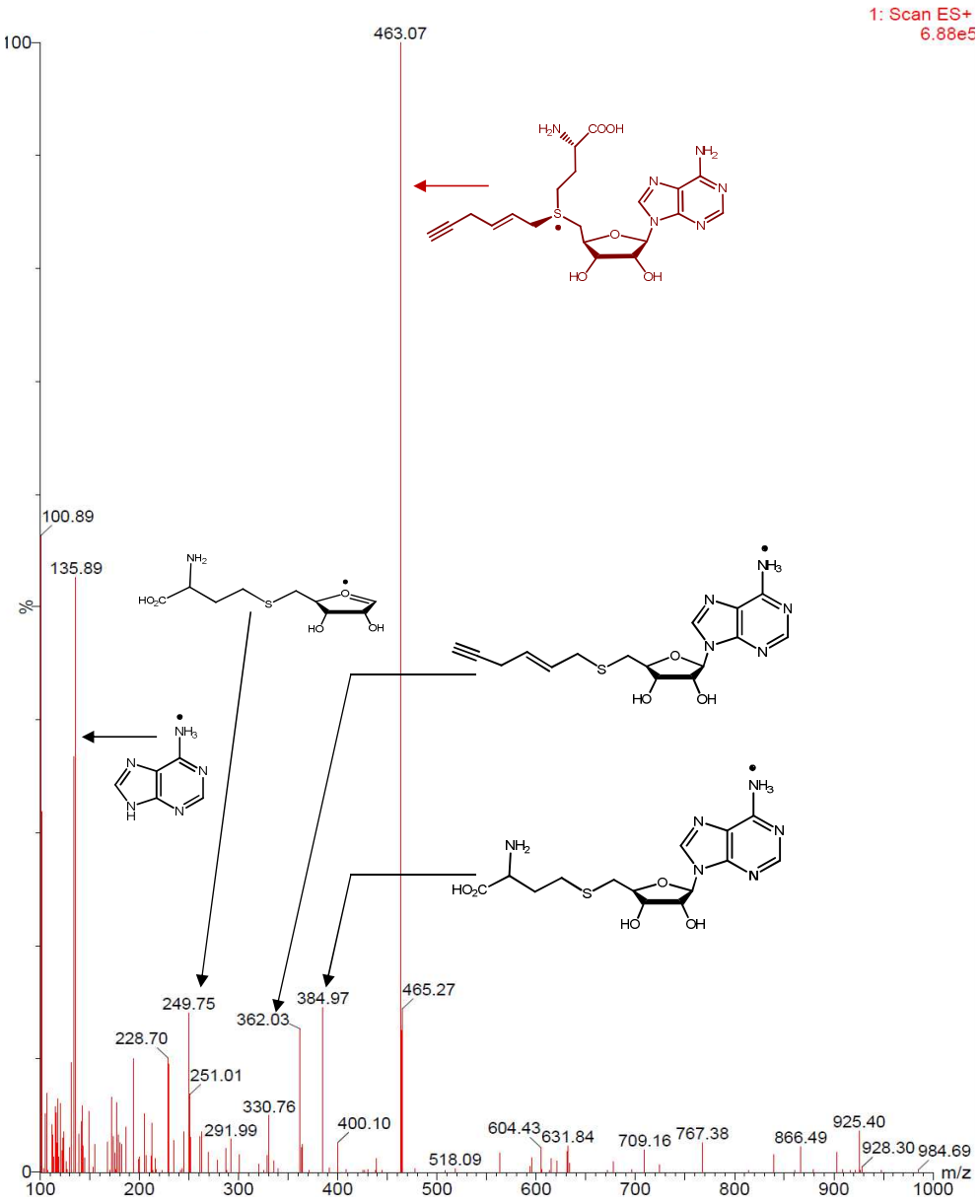


Figure S5. ESI-MS spectrum of in-cell-generated (*E*)-hex-2-en-5-ynyl SAM (Hey-SAM, **2c**). The sample was collected from the peak designated in Figure 3a and injected into MS for analysis. The characteristic peaks: 463.07 $[M]^+$, 384.97 [*S*-adenosyl-*L*-homocysteine+H]$^+$, 362.03 [5-(hex-2-ene-5-ynyl)thio-5-deoxyadenosine+H]$^+$, 249.75 [*S*-adenosyl-*L*-homocysteine+H]$^+$, 135.89 [adenine+H]$^+$.

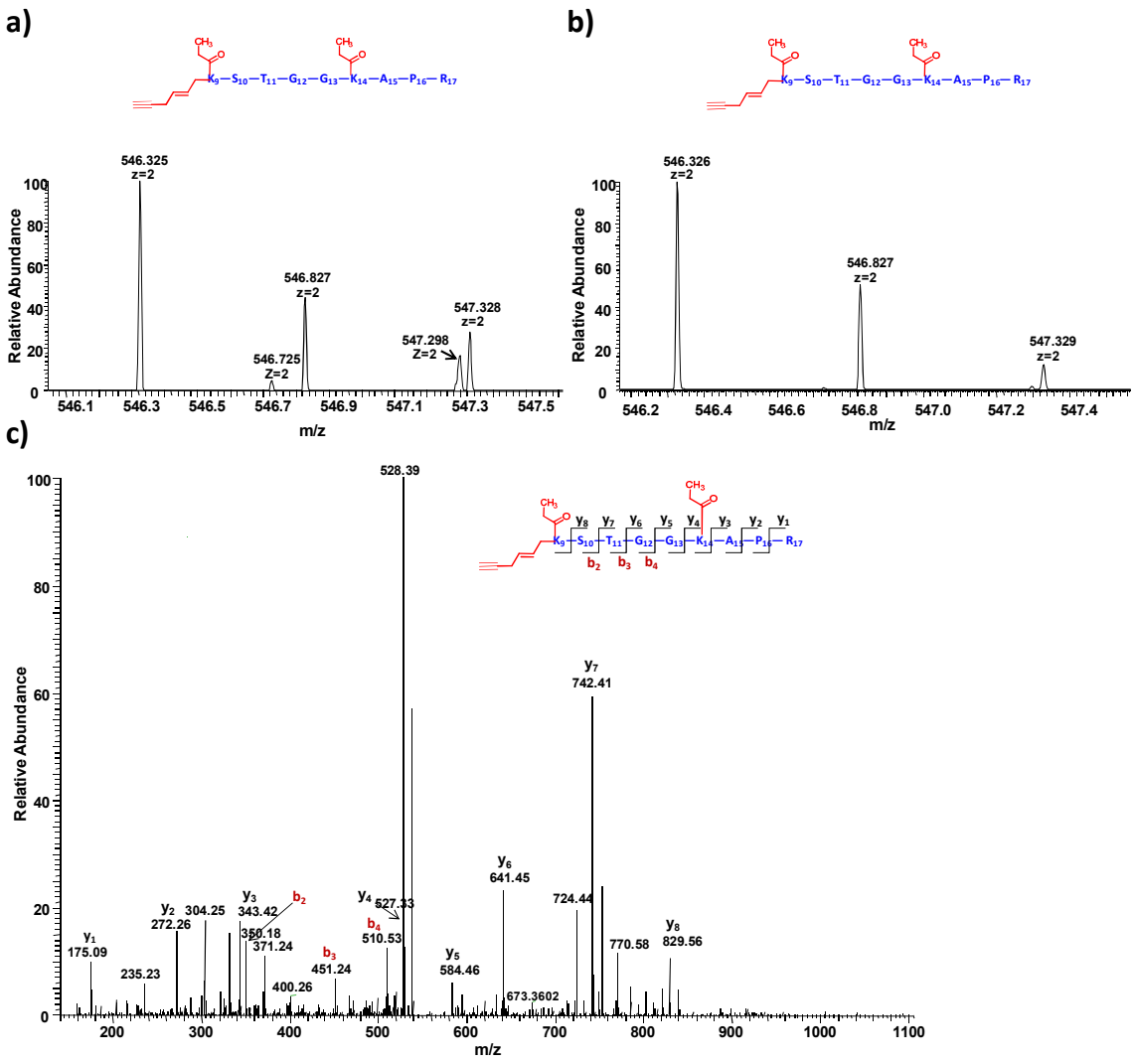


Figure S6. Mass spectroscopic characterization of in-cell (*E*)-hex-2-en-5-ynylation on H3K9. Total histones were acid-precipitated from the cells transfected with the engineered MAT2A and G9a, followed by the treatment of 1 mM Hey-Met **1c**. The precipitated histones were resolved in SDS-PAGE and the corresponding gel bands were propionylated, trypsinized and analyzed by LC/MS/MS as described previously.[3,4] **(a)** LC-MS trace of (*E*)-hex-2-en-5-ynylated histone H3K9 peptide from the cells co-transfected with the MAT2A I117A and G9a Y1154A variants. **(b)** LC-MS trace of (*E*)-hex-2-en-5-ynylated histone H3K9 peptide from the cells co-transfected with the MAT2A I117A and GLP1 Y1152A variants. **(c)** Representative tandem MS spectrum of the histone H3K9 fragment containing the characteristic (*E*)-hex-2-en-5-ynylation modification for the data in Figure S6b.

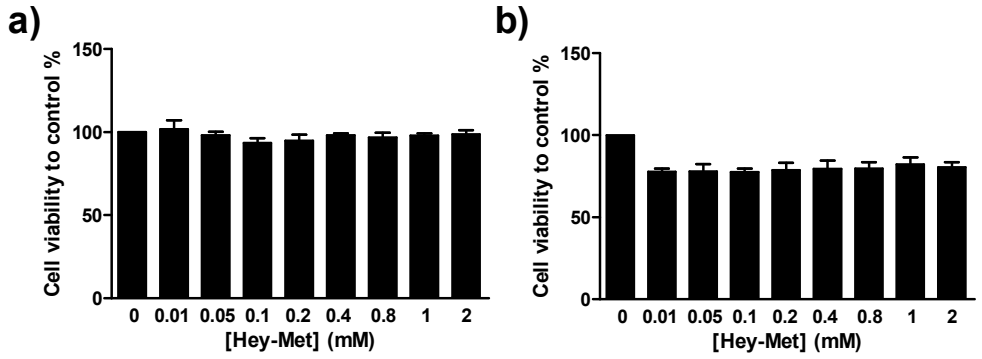


Figure S7. Toxicity evaluation of Hey-Met **1c** and Hey-SAM **2c** with alamar blue assay. **(a)** HEK293T cells were treated with varied concentrations (0-2mM) of Hey-Met **1c** for 8 h. The cell viability was determined by alamar blue assay and presented as the percentage of the control (without Hey-Met treatment). **(b)** HEK293T cells were transfected with the MAT2A I117A mutant prior to the treatment of Hey-Met **1c** (0–2 mM), which is expected to be processed into Hey-SAM **2c** intracellularly by the MAT2A I117A mutant. Cell viability was determined with alamar blue assay after 8 h incubation and was presented as the percentage of the control (without Hey-Met **1c**). Error bars represent means ± S.E.M (n=3). No obvious toxicity of Hey-Met **1c** and Hey-SAM **2c** was observed under the examined conditions.

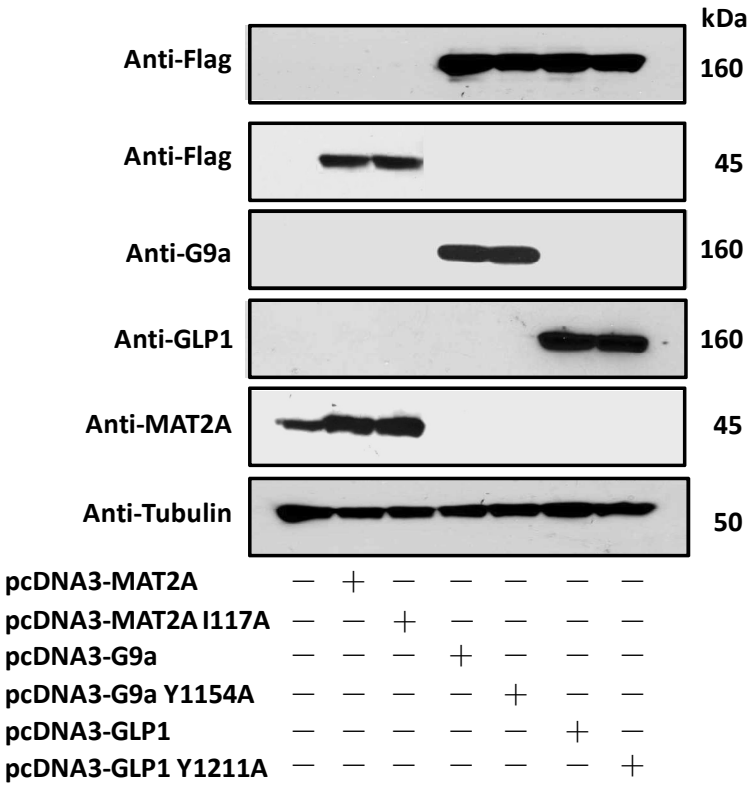


Figure S8. Western blot analysis of transfection efficiency of native MAT2A, G9a, GLP1 and their variants in HEK293T cells. Around 50 μg extracted protein was loaded per well and the equivalence of the loading was confirmed by tubulin.

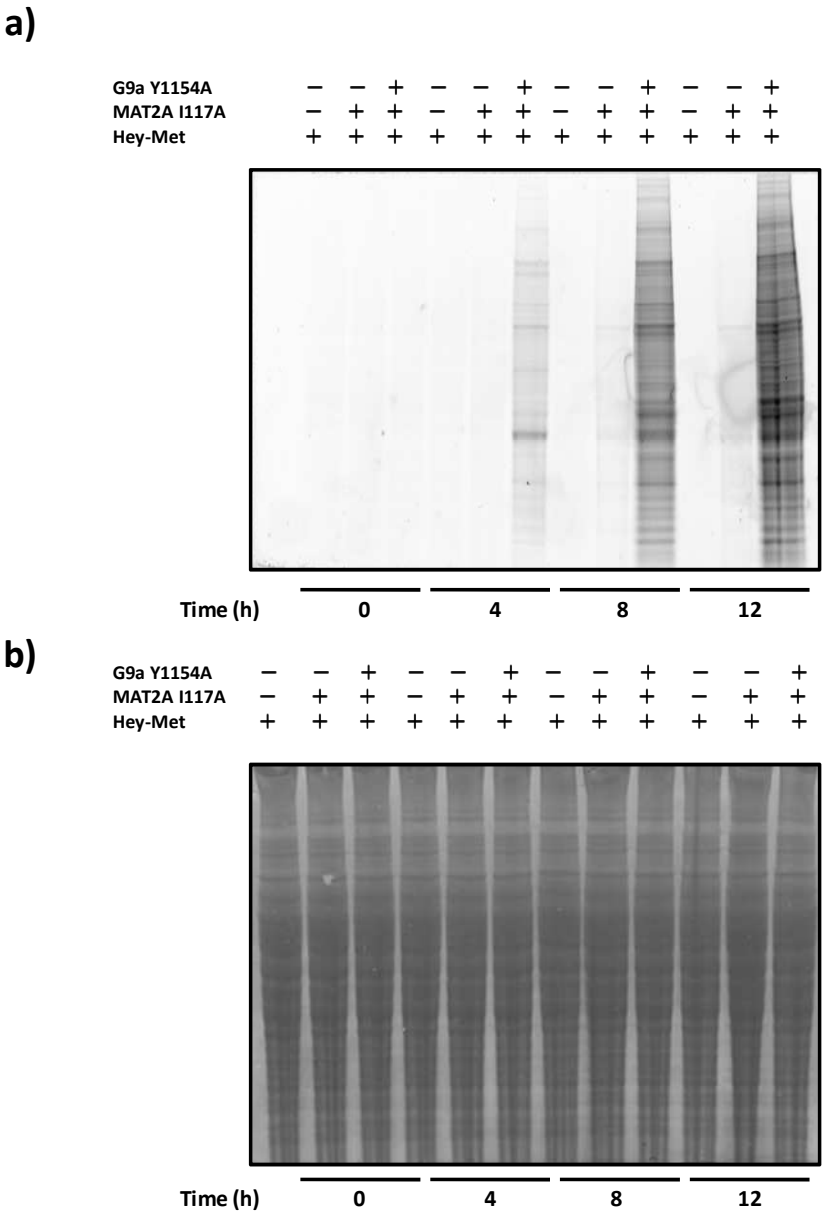


Figure S9. Hey-Met-mediated, time-dependent labeling of G9a targets within living cells. **(a)** HEK293T cells were transfected with MAT2A I117A and/or G9a Y1154A and treated with Hey-Met **1c** (1mM) for varied periods of time. Cell lysates were reacted with an azido fluorescent dye, and the labeled proteins were resolved by SDS-PAGE followed by in-gel fluorescence. **(b)** Coomassie blue staining of the SDS-PAGE as corresponding loading controls.

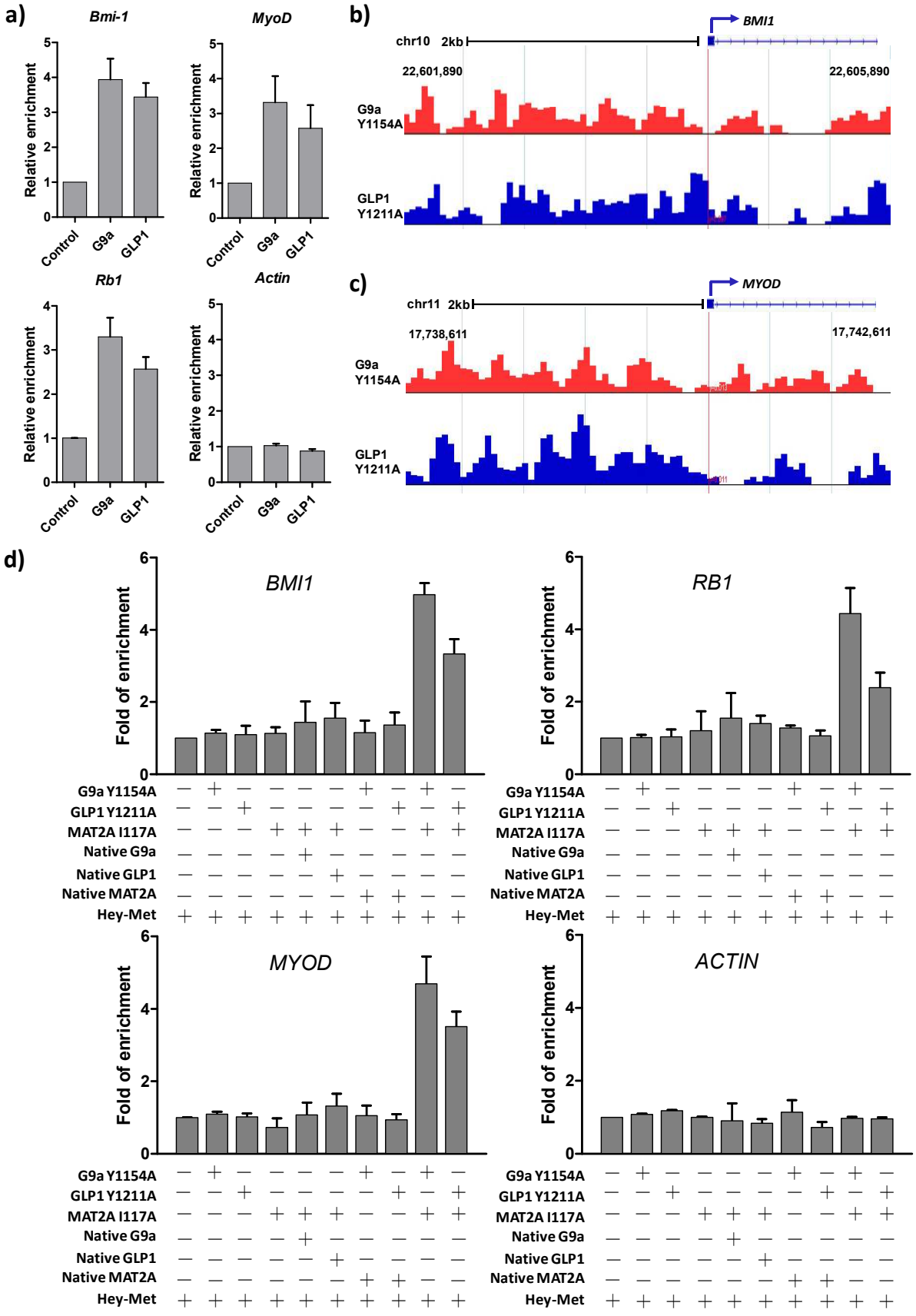

a)
Bmi-1
MyoD
Rb1
Actin
Relative enrichment
Control
G9a
GLP1
b)
BMI1
chr10 2kb
22,601,890
22,605,890
G9a Y1154A
GLP1 Y1211A
c)
MYOD
chr11 2kb
17,738,611
17,742,611
d)
BMI1
RB1
MYOD
ACTIN
Fold of enrichment
G9a Y1154A
GLP1 Y1211A
MAT2A I117A
Native G9a
Native GLP1
Native MAT2A
Hey-Met

Figure S10. Characterization of H3K9 modifications with ChIP and CliEn. **(a)** ChIP analysis of the promoters of *Rb-1*, *MyoD* and *Bmi-1* genes with *anti*-H3K9me2 antibody. ChIP-qPCR data showed that the overexpression of native G9a or GLP1 enhances H3K9me2 marker at the promoter regions of these previously-validated G9a/GLP1 target genes. Methylation level at the promoter of *Actin*, which is not affected by G9a/GLP1, was used as the control. ChIP data are presented as the fold changes versus the control groups (without the overexpression of native G9a or GLP1). **(b-c)** Clickable Chromatin Enrichment (CliEn)-seq showed the levels of the (*E*)-hex-2-en-5-ynylation at the *MyoD* and *Bmi-1* genes in the HEK293T cells co-transfected with the MAT2A I117A and G9a/GLP1 Y1154A/Y1211A mutants. The CliEn-seq files were normalized by the total reads and processed in Genplay tool. Red, the cells co-transfected with the MAT2A I117A and G9a Y1154A mutants, followed by the treatment of Hey-Met. Blue, the cells co-transfected with the MAT2A I117A and GLP1 Y1211A mutants, followed by the treatment of Hey-Met. **(d)** CliEn-qPCR analysis of the promoters of *Rb-1*, *MyoD* and *Bmi-1* genes. CliEn-qPCR data showed that co-transfection of the MAT2A I117A with G9a Y1154A or GLP1 Y1211A variant rather than their native forms enhances the H3K9 (*E*)-hex-2-en-5-ynylation at the promoters of *Rb-1*, *MyoD* and *Bmi-1* genes. CliEn-qPCR data were presented as the fold of enrichment versus the control group (without the overexpression of the native or engineered MAT2A, G9a and GLP1). Modification level at the promoter of *Actin*, which is not affected by the engineered MAT2A, G9a or GLP1 variants, was used as the control. Error bars represent means ± S.E.M (n=3).

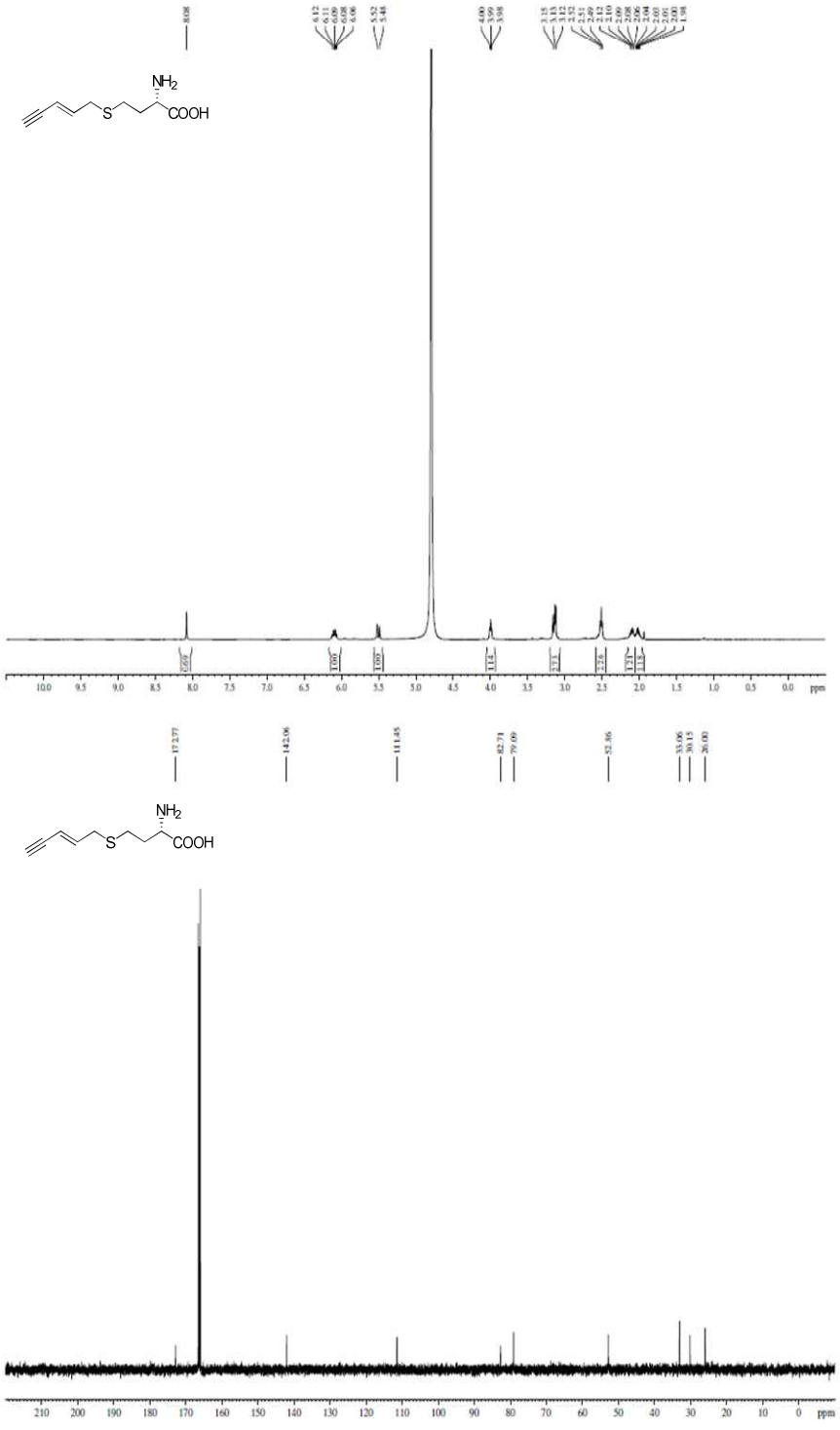


Figure S11. ^{1}H- and ^{13}C-NMR spectra of *S*-(pent-2-en-4-yny)-*L*-homocysteine (EnYn-Met, **1b**) in D_2O.

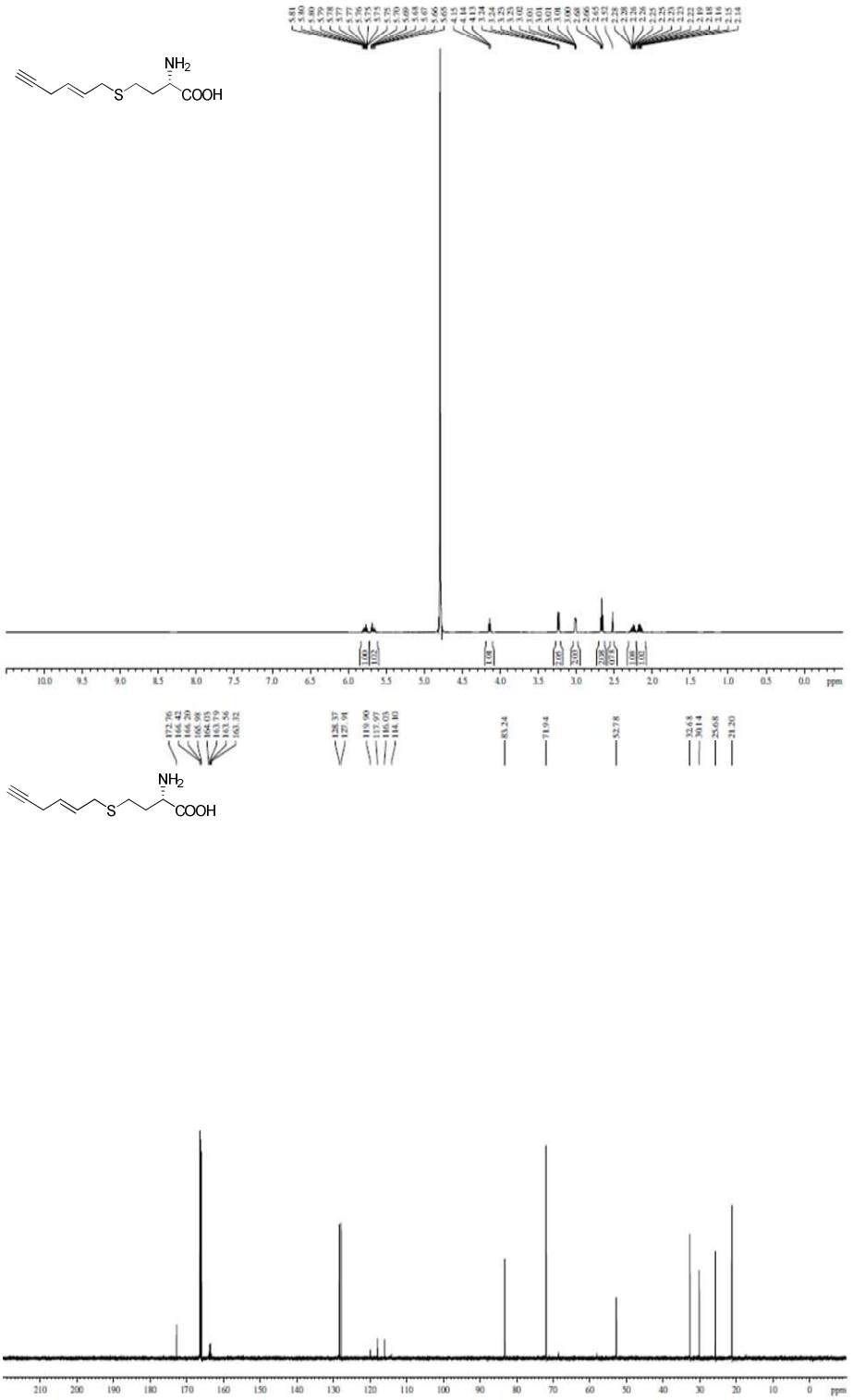


Figure S12. ^{1}H- and ^{13}C-NMR spectra of *S*-(hex-2-en-5-yny)-*L*-homocysteine trifluoroacetic salt (Hey-Met, **1c**) in D_2O.

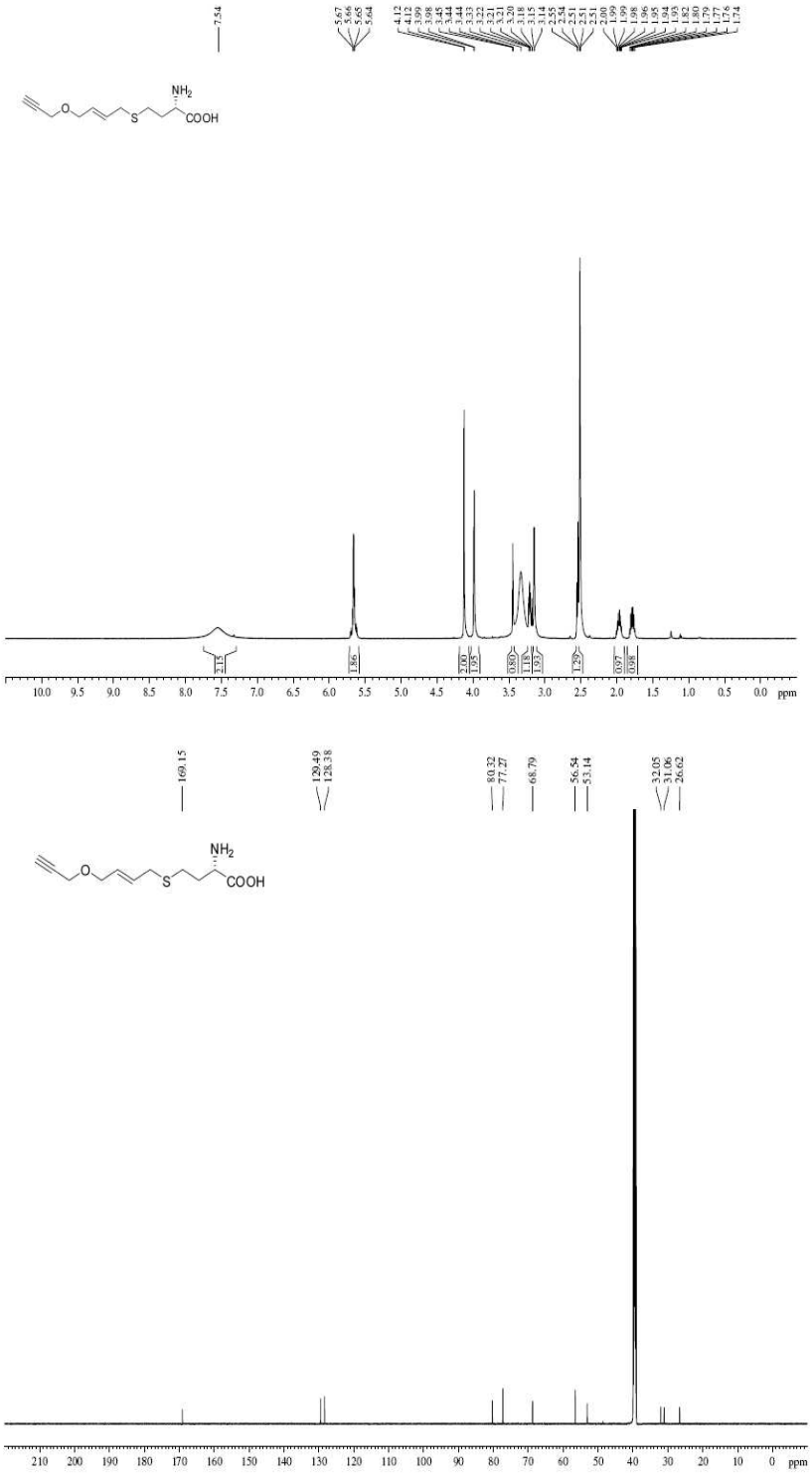


Figure S13. ^{1}H- and ^{13}C-NMR spectra of *S*-[4-(prop-2-ynyloxy)but-2-enylthio]-*L*-homocysteine (Pob-Met, **1d**) in D_2O.

C. Supplementary Table

Table S1. Primers used for ChIP-qPCR and CliEn-qPCR.

Gene name	Forward primer (5'-3')	Reverse primer (5'-3')
Bmi-1	CACGTCTTT TAGGACTTGGTGGGG	TCCTCTGCACTCGGGCGGAT
Rb-1	CACAGCTCCCTCCCTTCCTTTC	CTACCCAGAACCACCCCCTCCA
MyoD	TGGCTATATTTATCTCTGGGTTCATG	GCTCCCGCAGCCCCT
Actin	ACTCTACCTCTCAAGCCCAGGTCTC	CTAAATGTGCTGGGTGGGTCACTA

Table S2. Selected genes for the analysis of the (*E*)-hex-2-en-5-ynylation enrichment. The loci were known to carry H3K9 methylation markers as reported previously.[5-28]

Gene Symbol				
BMI1	CHAT	MLL	CBX5	SPRY2
CIITA	GH1	CIDEA	ZBTB7C	LIPE
ZNF436	TFAP2C	TFF1	ALK	MYO1B
SCN2A	PPARG	TBC1D5	CXCL10	DHFR
FST	IL12B	CGA	MDC1	OPRM1
HOXA1	LAMB1	MET	GPT	GRIN1
IFNB1	NOTCH1	KDM5C	OCRL	POLA1

References

(1) Erchegyi, J.; Grace, C. R.; Samant, M.; Cescato, R.; Piccand, V.; Riek, R.; Reubi, J. C.; Rivier, J. E. *J. Med. Chem.* **2008**, *51*, 2668.

(2) Gaoni, Y.; Leznoff, C. C.; Sondheim.F *J. Am. Chem. Soc.* **1968**, *90*, 4940.

(3) Islam, K.; Bothwell, I.; Chen, Y.; Sengelaub, C.; Wang, R.; Deng, H.; Luo, M. *J. Am. Chem. Soc.* **2012**, *134*, 5909.

(4) Islam, K.; Zheng, W.; Yu, H.; Deng, H.; Luo, M. *ACS Chem. Biol.* **2011**, *6*, 679.

(5) Li, Y.; Tan, T.; Zong, L.; He, D.; Tao, W.; Liang, Q. *Chromosome Res.* **2012**, *20*, 769.

(6) Kubicek, S.; O'Sullivan, R. J.; August, E. M.; Hickey, E. R.; Zhang, Q.; Teodoro, M. L.; Rea, S.; Mechtler, K.; Kowalski, J. A.; Homon, C. A.; Kelly, T. A.; Jenuwein, T. *Mol. Cell* **2007**, *25*, 473.

(7) Yamane, K.; Toumazou, C.; Tsukada, Y.; Erdjument-Bromage, H.; Tempst, P.; Wong, J.; Zhang, Y. *Cell* **2006**, *125*, 483.

(8) Tsukada, Y.; Ishitani, T.; Nakayama, K. I. *Genes Dev.* **2010**, *24*, 432.

(9) Rigal, M.; Kevei, Z.; Pelissier, T.; Mathieu, O. *EMBO J.* **2012**, *31*, 2981.

(10) van Essen, D.; Zhu, Y.; Saccani, S. *Mol. Cell* **2010**, *39*, 750.

(11) Qi, H. H.; Sarkissian, M.; Hu, G. Q.; Wang, Z.; Bhattacharjee, A.; Gordon, D. B.; Gonzales, M.; Lan, F.; Ongusaha, P. P.; Huarte, M.; Yaghi, N. K.; Lim, H.; Garcia, B. A.; Brizuela, L.; Zhao, K.; Roberts, T. M.; Shi, Y. *Nature* **2010**, *466*, 503.

(12) Wang, J.; Scully, K.; Zhu, X.; Cai, L.; Zhang, J.; Prefontaine, G. G.; Krones, A.; Ohgi, K. A.; Zhu, P.; Garcia-Bassets, I.; Liu, F.; Taylor, H.; Lozach, J.; Jayes, F. L.; Korach, K. S.; Glass, C. K.; Fu, X. D.; Rosenfeld, M. G. *Nature* **2007**, *446*, 882.

(13) Shi, L.; Sun, L.; Li, Q.; Liang, J.; Yu, W.; Yi, X.; Yang, X.; Li, Y.; Han, X.; Zhang, Y.; Xuan, C.; Yao, Z.; Shang, Y. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 7541.

(14) Bannister, A. J.; Zegerman, P.; Partridge, J. F.; Miska, E. A.; Thomas, J. O.; Allshire, R. C.; Kouzarides, T. *Nature* **2001**, *410*, 120.

(15) Garcia-Bassets, I.; Kwon, Y. S.; Telese, F.; Prefontaine, G. G.; Hutt, K. R.; Cheng, C. S.; Ju, B. G.; Ohgi, K. A.; Wang, J.; Escoubet-Lozach, L.; Rose, D. W.; Glass, C. K.; Fu, X. D.; Rosenfeld, M. G. *Cell* **2007**, *128*, 505.

(16) Takada, I.; Mihara, M.; Suzawa, M.; Ohtake, F.; Kobayashi, S.; Igarashi, M.; Youn, M. Y.; Takeyama, K.; Nakamura, T.; Mezaki, Y.; Takezawa, S.; Yogiashi, Y.; Kitagawa, H.; Yamada, G.; Takada, S.; Minami, Y.; Shibuya, H.; Matsumoto, K.; Kato, S. *Nat. Cell Biol.* **2007**, *9*, 1273.

(17) Musri, M. M.; Carmona, M. C.; Hanzu, F. A.; Kaliman, P.; Gomis, R.; Parrizas, M. *J. Biol. Chem.* **2010**, *285*, 30034.

(18) Chen, H.; Yan, Y.; Davidson, T. L.; Shinkai, Y.; Costa, M. *Cancer Res.* **2006**, *66*, 9009.

(19) Chen, H.; Ke, Q.; Kluz, T.; Yan, Y.; Costa, M. *Mol. Cell Biol.* **2006**, *26*, 3728.

(20) Martins, G.; Calame, K. *Annu. Rev. Immunol.* **2008**, *26*, 133.

(21) Chen, H.; Kluz, T.; Zhang, R.; Costa, M. *Carcinogenesis* **2010**, *31*, 2136.

(22) Smith, M. A.; Wright, G.; Wu, J.; Tailor, P.; Ozato, K.; Chen, X.; Wei, S.; Piskurich, J. F.; Ting, J. P.; Wright, K. L. *J. Biol. Chem.* **2011**, *286*, 7893.

(23) Riddle, N. C.; Jung, Y. L.; Gu, T.; Alekseyenko, A. A.; Asker, D.; Gui, H.; Kharchenko, P. V.; Minoda, A.; Plachetka, A.; Schwartz, Y. B.; Tolstorukov, M. Y.; Kuroda, M. I.; Pirrotta, V.; Karpen, G. H.; Park, P. J.; Elgin, S. C. *PLoS Genet.* **2012**, *8*, e1002954.

(24) Jun, H. J.; Kim, J.; Hoang, M. H.; Lee, S. J. *PLoS One* **2012**, *7*, e44345.

(25) Wang, D.; Xia, X.; Weiss, R. E.; Refetoff, S.; Yen, P. M. *PLoS One* **2010**, *5*, e9853.

(26) Mysliwiec, M. R.; Carlson, C. D.; Tietjen, J.; Hung, H.; Ansari, A. Z.; Lee, Y. *J. Biol. Chem.* **2012**, *287*, 1235.

(27) Vucetic, Z.; Kimmel, J.; Reyes, T. M. *Neuropsychopharmacology* **2011**, *36*, 1199.

(28) Barski, A.; Cuddapah, S.; Cui, K.; Roh, T. Y.; Schones, D. E.; Wang, Z.; Wei, G.; Chepelev, I.; Zhao, K. *Cell* **2007**, *129*, 823.