Electronic Supplementary Material (ESI) for Analyst.

SUPPLEMENTARY INFORMATION

An SDS-PAGE based proteomic approach for N-terminome profiling

Prashant Kaushal, [a,b] Yumi Kwon, [a,c] Shinyeong Ju [a,c] and Cheolju Lee [*a,b,d]

a. Center for Theragnosis, Korea Institute of Science and Technology, Seoul 02792, Korea.

b. Division of Bio-Medical Science & Technology, KIST School, Korea University of Science and Technology, Seoul 02792, Korea.

c. Department of Life Science and Research Institute for Natural Sciences, Hanyang University, Seoul 04763, Korea.

d. KHU-KIST Department of Converging Science and Technology, Kyung Hee University, 26 Kyunghee-daero, Dongdaemun-gu, Seoul 02447, Korea.

* Corresponding author. E-mail: clee270@kist.re.kr; Tel: +82-2-958-6788

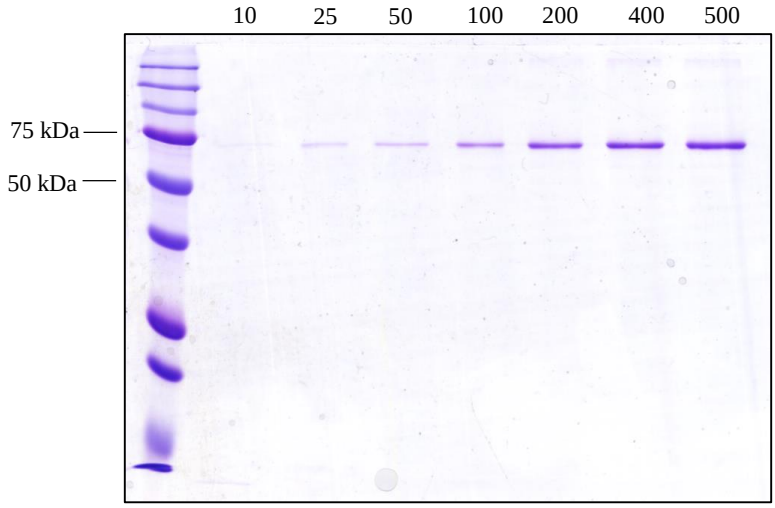
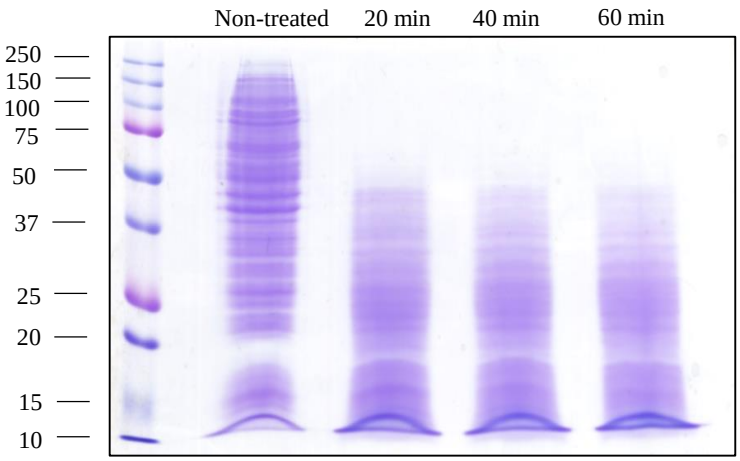
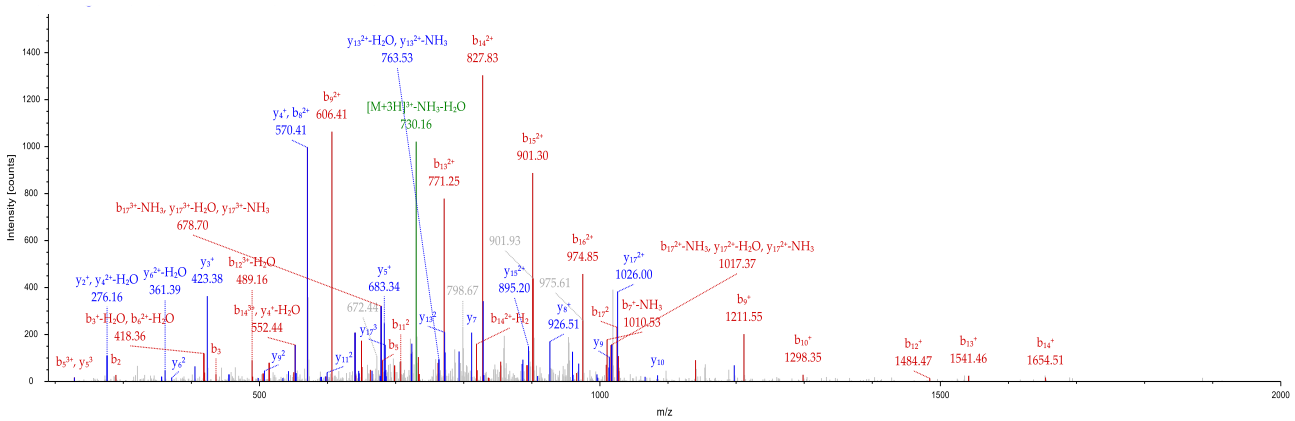


Figure S1: Representative SDS-PAGE gel image showing different amounts of bovine serum albumin (BSA), ranging from 10 ng to 500 ng, used for N-termini enrichment.

A



B



C

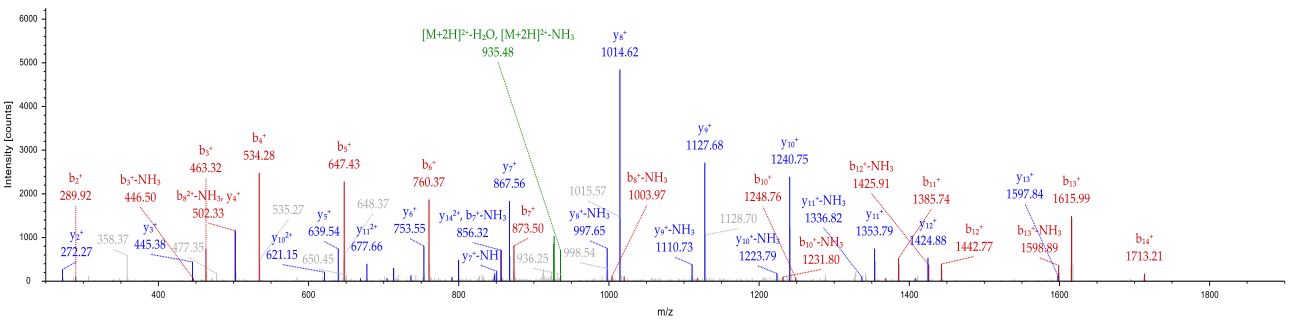


Figure S2: (A) SDS-PAGE gel image showing Glu-C treated HCT8 cell lysate for different time periods. 100 µg of sample was incubated with Glu-C in 1:100 ratio and loaded on to the gel for N-termini enrichment. MS/MS spectra of an (B) endogenously acetylated protein N-terminal peptide (Ac)MDFNM(d_3Ac)K(d_3Ac)KLASDAGIFFTR of Q9NR46 protein from position 1-18, detected with charge: +3, monoisotopic m/z: 742.03920 Da (-0.58 mmu/-0.77 ppm) and of an (C) endogenously free (i.e. d_3 acetylated) protein N-terminal peptide (d_3Ac)MI(d_3Ac)KAILIFNNHG(d_3Ac)KPR of Q9N2572 protein from position 1-15, detected with charge: +2, monoisotopic m/z: 944.05243 Da (-0.37 mmu/-0.4 ppm) in Glu-C untreated cell lysate sample.

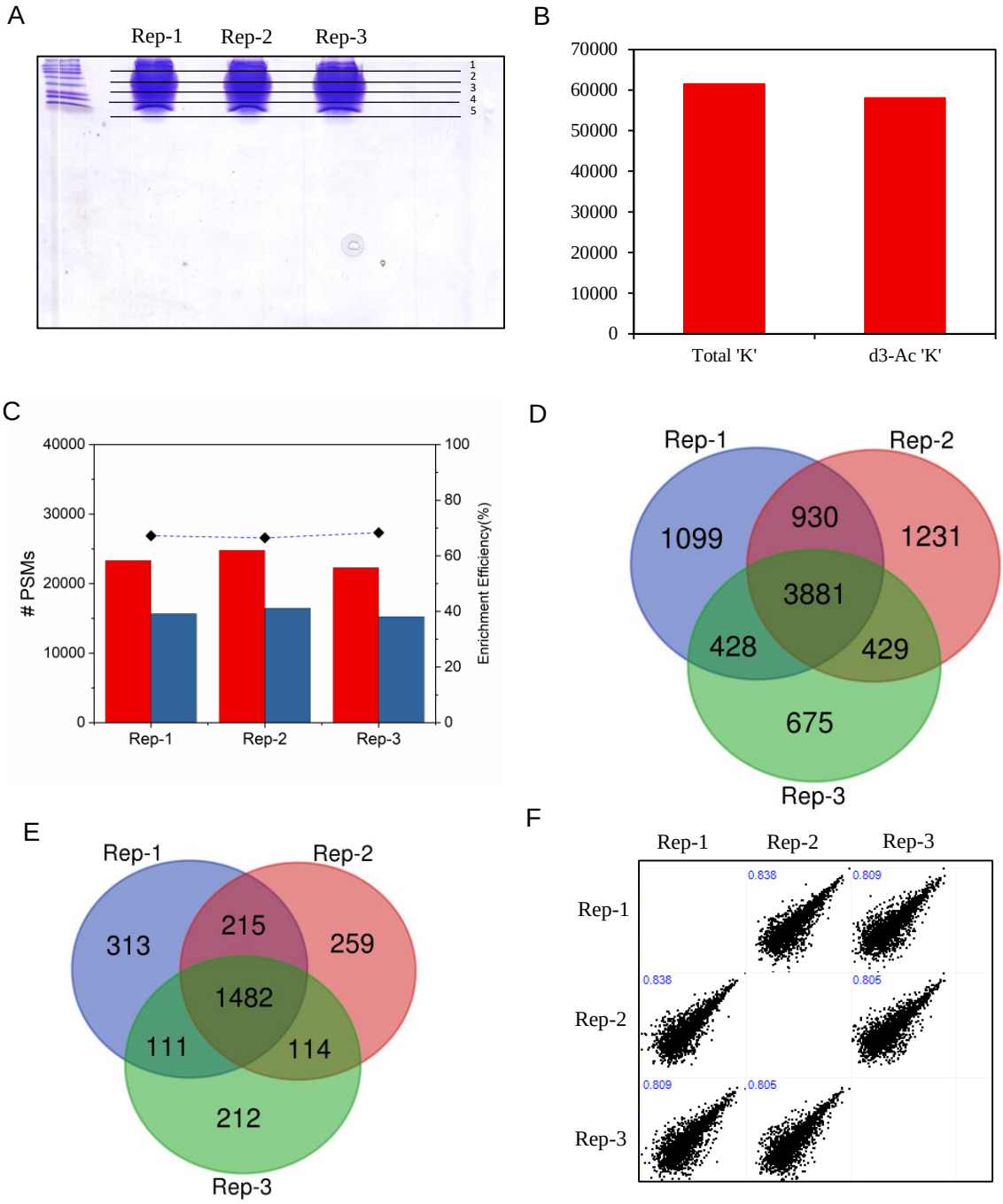


Figure S3: (A) Gel image showing short run SDS-PAGE containing 100 µg of U87MG cell lysate in each lane, which is further divided into five fractions for N-terminome profiling. (B) Bar graph showing (B) total and labeled lysine residues identified in all the PSMs using optimized d_6-actic anhydride labeling protocol and (C) N-termini enrichment efficiency in all three technical replicates of cell lysate. (D) Venn diagram showing number of identified N-terminal peptides and (E) their corresponding proteins in the triplicate datasets. (F) Multiple scatter plot (drawn using Perseus (v. 1.6.7.0) showing correlation between triplicates datasets based on LFQ intensity of all the peptides identified by MaxQuant (ver. 1.5.8.3). Venn diagrams shown here are generated using webtool available at: http://bioinformatics.psb.ugent.be/webtools/Venn/

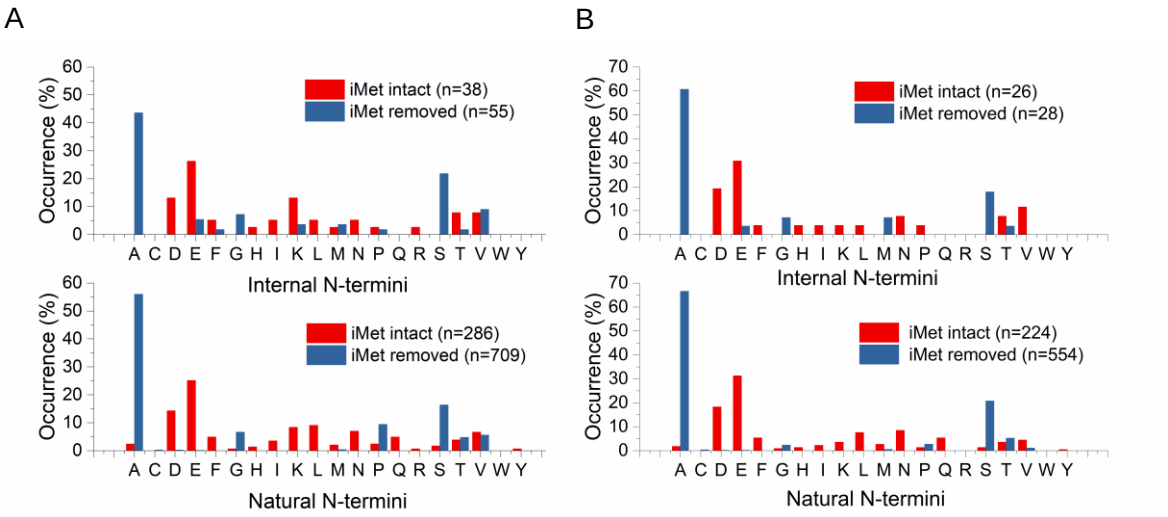


Figure S4: Frequency distribution plots of amino acids at P1' (when iMet is removed) and P2' position (when iMet is intact) in the natural and internal N-termini (A) irrespective of acetylation, or (B) with acetylation state.

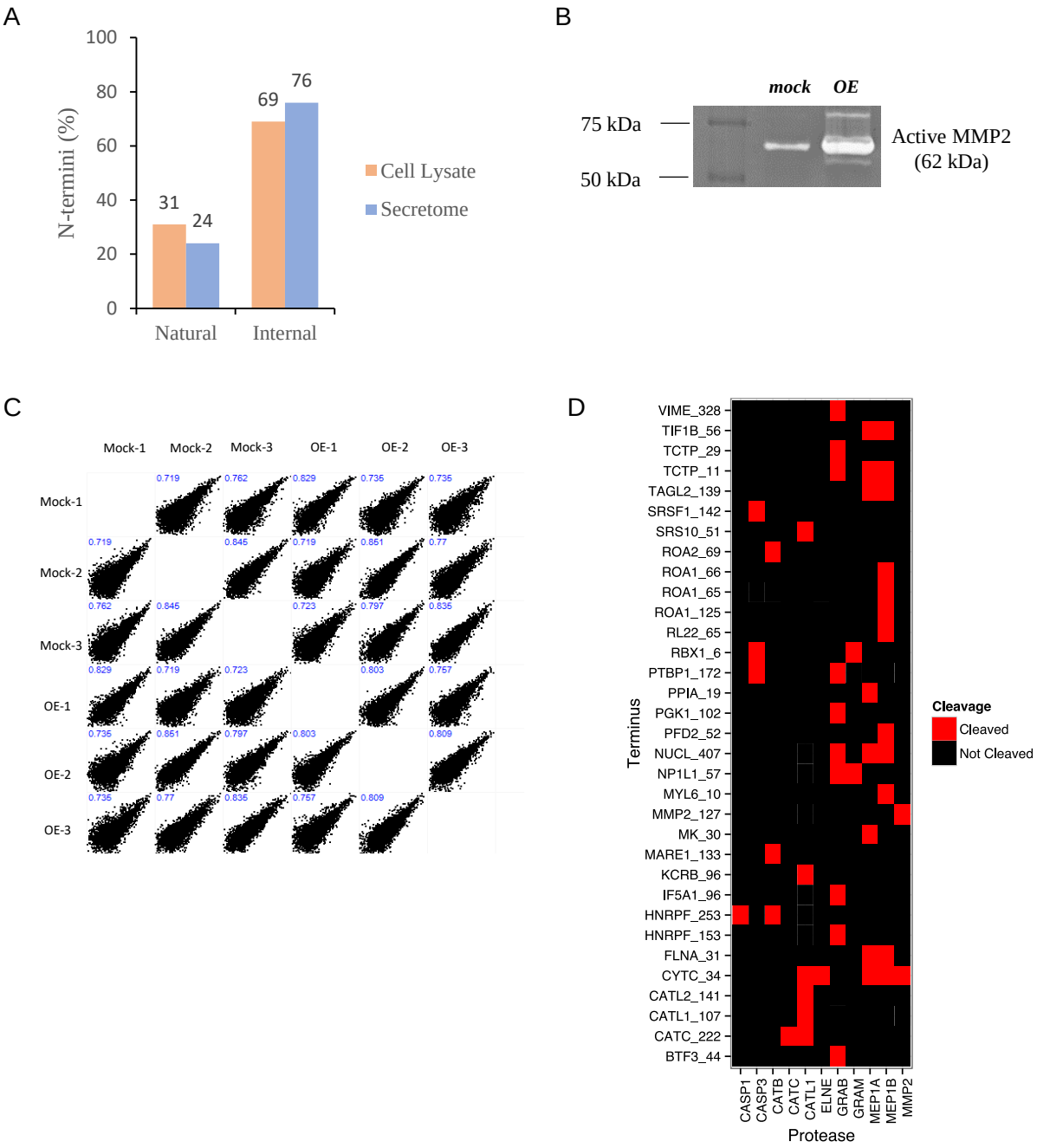


Figure S5: (A) Comparison of identified natural and internal N-termini in between cell lysate and HCT8 secretome datasets. (B) Gelatin zymography result of HCT8 *mock* and *OE* stable transfectants. 5 µg of conditioned media was loaded in each lane. (C) Multiple scatter plot (drawn using Perseus 1.6.7.0) showing correlation between biological replicates of *Mock* and *OE* datasets based on LFQ intensity of peptides identified in all datasets. (D) Matrix showing known evidences of substrates identified in LFQ analysis in between *Mock* and *OE* datasets, and their corresponding protease identified through TopFINDer webtool.