

Description of Additional Supplementary Files

File Name: Supplementary Data 1

Description: N-terminome profiling of HeLa cell line with MOCK, MG132, and MGTG treatment using tandem search strategy; PSMs of Nt-arginylated peptides for two biological replicates. MS2 similarity, RT evaluation, and MET test result are described.

File Name: Supplementary Data 2

Description: PSMs for Arg-starting peptide were identified using N-terminome profiling of HeLa cell line with MOCK, MG132, and MGTG treatment for two biological replicates.

File Name: Supplementary Data 3

Description: Similarity score output of MS2 prediction models with varied training method; containing both MS2 prediction result of trypsin and chymotrypsin

File Name: Supplementary Data 4

Description: Source data for supplementary figure 5. The data contains false discovery rates using Nt-arginylation specific target-decoy search strategy.

File Name: Supplementary Data 5

Description: Nt-arginylation sites after ML-based filtering; containing protein based information and LFQ data including localization, biological function, DeepLoc, and protease site prediction by SignalP, Target P, and Procleave.

File Name: Supplementary Data 6

Description: Structural and functional analysis of Nt-arginylation sites; Alphafold, Reactome, Gene Ontology, and GSEA

File Name: Supplementary Data 7

Description: Label free quantitation result of Nt-arginylation sites and paired non-Nt-arginylation sites; Sequence based LFQ intensities across biological duplicates.

File Name: Supplementary Data 8

Description: PRM-MS result of selected Nt-arginylation sites and paired non-Nt-arginylation sites; including intensities without CRMN normalization.

File Name: Supplementary Data 9

Description: PRM-MS quantitation result of selected Nt-arginylation sites and paired non-Nt-arginylation sites.

File Name: Supplementary Data 10

Description: Previously reported arginylated proteins and sites including corresponding references.