Optimized Workflow for Enrichment and Identification of Biotinylated Peptides using Tamavidin 2-REV for BioID and Cell Surface Proteomics

Kohei Nishino^{1,2†}, Harunori Yoshikawa^{1†}, Kou Motani¹, and Hidetaka Kosako^{1*}

¹Division of Cell Signaling, Fujii Memorial Institute of Medical Sciences, Institute of Advanced
Medical Sciences, Tokushima University, Tokushima 770-8503, Japan

²Kuramoto Division, Technical Support Department, Tokushima University, Tokushima 770-8503,
Japan

*Corresponding Author

Hidetaka Kosako

Phone: +81-88-634-6413

Email: kosako@tokushima-u.ac.jp

[†]K.N. and H.Y. contributed equally to this work.

KEYWORDS: Tamavidin 2-REV, biotinylation, affinity enrichment, proximity-dependent
biotinylation, BioID, TurboID, cell surface labeling, protein–protein interactions

ABSTRACT: Chemical or enzymatic biotinylation of proteins is widely used in various studies, and proximity-dependent biotinylation coupled to mass spectrometry is a powerful approach for analyzing protein–protein interactions in living cells. We recently developed a simple method to enrich biotinylated peptides using Tamavidin 2-REV, an engineered avidin-like protein with reversible biotin-binding capability. However, the low abundance of protein biotinylation in cells required large amounts of cellular proteins to detect enough biotinylated peptides. Here we optimized the workflow for efficient enrichment and identification of biotinylated peptides. The most efficient recovery was achieved by heat inactivation of trypsin, prewashing Tamavidin 2-REV beads, clean-up of biotin solution, mock elution, and the optimal temperature and salt concentration for elution. Using the optimized workflow, over 2-fold more biotinylated peptides were identified with higher purity from RAW264.7 macrophages expressing TurboID-fused STING. In addition, sequential digestion with Glu-C and trypsin led to the identification of biotinylation sites that were not identified by trypsin digestion alone. Furthermore, the combination of this workflow with TMT labeling enabled large-scale quantification of cell surface proteome changes upon EGF stimulation. This workflow would be useful not only for BioID and cell surface proteomics but also for various other applications based on protein biotinylation.

38 INTRODUCTION

39 Chemical or enzymatic biotinylation of proteins is widely used in various studies including affinity
40 purification, enzyme-linked immunosorbent assay, Western blotting, cell staining, cell surface labeling,
41 and identification of post-translational modifications (1, 2). Recently, proximity-dependent
42 biotinylation approaches such as BioID and APEX have emerged as powerful techniques to identify
43 protein–protein interactions and subcellular proteomes in living cells and organisms (3-5). Biotinylated
44 proteins can be captured by streptavidin beads and identified by mass spectrometry after on-bead
45 digestion or digestion of eluted proteins (6-10). However, the extremely high affinity of streptavidin
46 for biotin makes it difficult to efficiently and specifically elute biotinylated proteins and remove
47 contaminants of non-biotinylated proteins. While the direct identification of biotinylated peptides can
48 eliminate these contaminants and provide further information about the topology of membrane proteins
49 and the sites of post-translational modifications (11-14), biotinylated peptides are rarely detected after
50 pull-down with streptavidin beads.

51 To efficiently enrich and identify biotinylated peptides, two groups have developed a revolutionary
52 method using anti-biotin antibody beads, from which biotinylated peptides can be efficiently but non-
53 specifically eluted with strong acid and high concentration of organic solvent (15, 16). Very recently,
54 we developed a simple method to specifically enrich biotinylated peptides using Tamavidin 2-REV
55 beads (17). Tamavidin 2-REV is an engineered version of Tamavidin 2 that acquired reversible biotin-
56 binding capability by mutating S36 that forms a hydrogen bond with biotin (17, 18). Because
57 biotinylated peptides can be mildly and specifically eluted from Tamavidin 2-REV beads by adding
58 excess free biotin, the enrichment efficiency was higher than that of anti-biotin antibody beads.
59 However, interference ions derived from our enrichment process prevented sensitive detection of
60 biotinylated peptides. Furthermore, the low abundance of protein biotinylation in BioID-expressing
61 cells required large amounts of cellular proteins to identify enough biotinylated peptides by mass

spectrometry (17, 19). Therefore, it is important to optimize the enrichment method for biotinylated peptides, as various enrichment methods for phosphorylated peptides have been extensively optimized so far.

Here, we optimized each step of the workflow for efficient enrichment and identification of biotinylated peptides using Tamavidin 2-REV. Consequently, approximately 1,400 biotinylated peptides were identified from 300 µg proteins of RAW264.7 macrophages expressing TurboID-fused STING (stimulator of interferon genes). This number of identified biotinylated peptides was more than twice as many as before optimization, and approximately 95% of all identified peptides were biotinylated. We also found that sequential digestion with Glu-C and trypsin led to the identification of biotinylation sites that were not identified by trypsin digestion alone. Furthermore, the combination of this workflow with TMT labeling enabled large-scale quantification of biotinylated peptides from EGF-stimulated and then surface-biotinylated HeLa cells. This optimized workflow using Tamavidin 2-REV would be useful for BioID and cell surface proteome profiling as well as for various other applications based on protein biotinylation.

77 MATERIALS AND METHODS

78 Reagents

79 MagCapture HP Tamavidin 2-REV magnetic beads and lysyl endopeptidase (Lys-C) were purchased
80 from FUJIFILM Wako Pure Chemical (Osaka, Japan). DMEM high glucose, penicillin-streptomycin
81 solution, and D-biotin were from Nacalai Tesque (Kyoto, Japan). Fetal bovine serum (FBS) and
82 Pefabloc SC were from Sigma-Aldrich (St. Louis, MO). EGF was from PeproTech (Rocky Hill, NJ).
83 EZ-Link Sulfo-NHS-Biotin, BCA protein assay kit, trypsin (MS grade), and endoproteinase Glu-C
84 (MS grade) were from Thermo Fisher Scientific (Waltham, MA). RapiGest SF was from Waters
85 (Milford, MA). All other standard laboratory reagents were obtained from Nacalai Tesque, Sigma-
86 Aldrich, or FUJIFILM Wako Pure Chemical.

87

88 Cell Culture

89 STING-knockout RAW264.7 murine macrophages stably expressing TurboID-STING fusion protein
90 (17, 20) and HeLa cells (21) were maintained in DMEM supplemented with 10% FBS, penicillin (100
91 U/mL), and streptomycin (0.1 mg/mL) at 37 °C in 5% CO₂.

92

93 EGF Stimulation and Cell Surface Biotin Labeling

94 Cell surface biotin labeling was performed according to previous reports with slight modifications (22,
95 23). HeLa cells in 6-cm dishes were treated with 100 ng/mL EGF for 0, 5, 10, or 30 min after starvation
96 for 18 h in serum-free DMEM, washed twice with ice-cold PBS, and then incubated with EZ-Link
97 Sulfo-NHS-Biotin (0.25 mg/mL in PBS) for 30 min at 4 °C with gentle agitation. The cells were then
98 incubated with 0.1 M glycine in PBS for 10 min at 4 °C with gentle agitation. After washing with ice-
99 cold TBS1 (20 mM Tris-HCl, pH 7.5, and 150 mM NaCl), the cells were scraped and collected by
100 centrifugation at 4 °C.

Protein Extraction and Enzymatic Digestion of BioID Samples

STING-knockout RAW264.7 cells stably expressing TurboID-STING were seeded in 6-cm dishes and cultured overnight. The cells at 80% confluence were incubated with biotin (500 μ M final concentration) for 10 min, washed with ice-cold PBS, scraped on ice, collected by centrifugation at 4 °C, and lysed in 200 μ L of guanidine buffer (6 M guanidine-HCl and 100 mM HEPES-NaOH, pH 7.5) containing 10 mM Tris (2-carboxyethyl) phosphine (TCEP) and 40 mM chloroacetamide (CAA). The lysates were heated at 95°C for 10 min and then sonicated at 4°C for 150 s (5 $\times$ 30 s pulses with 30 s intervals) at high power using a Bioruptor II (Cosmo Bio, Tokyo, Japan). Protein concentrations were quantified using a BCA assay kit. Then, 300 μ g of proteins were purified by methanol–chloroform precipitation and resuspended in 200 μ L of 0.1% RapiGest SF in 50 mM triethylammonium bicarbonate or 100 μ L of PTS buffer [100 mM Tris-HCl, pH 8.0, 12 mM SDC, and 12 mM SLS (24–26)], followed by sonication and heating at 95°C for 10 min. The protein solution in PTS buffer was then diluted 5-fold with 100 mM Tris-HCl, pH8.0. The proteins were digested with trypsin (1:100) at 37 °C overnight, or with Lys-C (1:100) or Glu-C (1:100) at 37 °C for 4 h followed by digestion with trypsin (1:100) at 37 °C overnight.

Original Workflow for Enrichment of Biotinylated Peptides

A 15- μ L slurry of MagCapture HP Tamavidin 2-REV magnetic beads per sample was washed three times with TBS2 (50 mM Tris-HCl, pH 7.5, and 150 mM NaCl). The digested peptides containing 0.1% RapiGest SF were diluted 5-fold with TBS2 and incubated with the Tamavidin 2-REV beads in the presence of 1 mg/mL Pefabloc SC for 3 h at 4 °C. After washing five times with TBS2, biotinylated peptides were eluted with 100 μ L of 1 mM biotin in TBS2 for 15 min at 95 °C twice. The combined

eluates were desalted using GL-Tip SDB (GL Sciences, Tokyo, Japan), evaporated in a SpeedVac concentrator, and redissolved in 0.1% TFA and 3% acetonitrile (ACN).

Optimized Workflow for Enrichment of Biotinylated Peptides

A 15-μL slurry of the Tamavidin 2-REV beads per sample was washed twice with 10% ACN and then three times with TBS2. The digested peptides in diluted PTS buffer were heated at 95 °C for 10 min to inactivate trypsin, diluted 2-fold with TBS2, and incubated with the ACN-prewashed Tamavidin 2-REV beads for 3 h at 4 °C. During the incubation period, 1 mM biotin solution in TBS3 (50 mM Tris-HCl, pH 7.5, and 500 mM NaCl) was passed through GL-Tip SDB. After washing five times with TBS2, the beads were incubated with 200 μL of TBS3 for 15 min at 37 °C as a mock elution. After removing TBS3, biotinylated peptides were eluted for 15 min at 37 °C twice with 100 μL of the biotin solution that had been passed through GL-Tip SDB. The combined eluates were desalted using GL-Tip SDB, evaporated in a SpeedVac concentrator, and redissolved in 0.1% TFA and 3% ACN.

LC-MS/MS of BioID Samples

Peptides were analyzed by an EASY-nLC 1200 UHPLC-coupled Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher Scientific). Aliquots (10 μL) of each sample were loaded by 20 μL of 0.1% formic acid in water at a pressure of 200 bar onto a trap column (100 μm × 2 cm, PepMap nanoViper C18 column, 5 μm, 100 Å; Thermo Fisher Scientific). The following mobile phases were used: solvent A (0.1% formic acid in water) and solvent B (80% ACN and 0.1% formic acid in water). The trapped peptides were eluted from an analytical column (75 μm × 15 cm, nano HPLC capillary column, 3 μm; Nikkoy Technos, Tokyo, Japan) at a constant flow rate of 300 nL/min by changing the gradient: 5% B at 0 min, 40% B at 60 min, 100% B at 70 min, and 100% B at 80 min. The Orbitrap Fusion was operated in positive ion mode with nanospray voltage set at 2.0 kV, source temperature at

275 °C, and S-lens RF level at 60. A scan cycle comprised MS1 scan (m/z range from 375 to 1,500) with default charge set to 2, maximum ion injection time of 50 ms, a resolution of 120,000, and automatic gain control (AGC) value of 4e5. The MS1 scan cycle was followed by MS2 scans (ion trap) of the most intense ions fulfilling predefined selection criteria (AGC: 1e5; maximum ion injection time: 200 ms; isolation window: 1.6 m/z ; spectrum data type: centroid; exclusion of unassigned, singly, and >5 charged precursors; peptide match preferred; exclude isotopes on; dynamic exclusion time: 10 s) for 3 sec. The HCD collision energy was set to 30% normalized collision energy.

MS Data Analysis of BioID Samples

Raw data were directly analyzed against the Swiss-Prot database, version 2017-10-25, restricted to *Mus musculus* (25,097 sequences) using Proteome Discoverer version 2.4 (Thermo Fisher Scientific) with the Sequest HT search engine. The search parameters were as follows: (a) trypsin as an enzyme with up to two missed cleavages; (b) precursor mass tolerance of 10 ppm; (c) fragment mass tolerance of 0.6 Da; (d) carbamidomethylation of cysteine as a fixed modification; (e) acetylation of protein N terminus, oxidation of methionine, and biotinylation of lysine as variable modifications. Peptides were filtered at a false discovery rate (FDR) of 1% using the Percolator node of Proteome Discoverer.

TMT-Based Quantitative Proteomics of Surface-Biotinylated Cells

EGF-stimulated and surface-biotinylated HeLa cells were lysed in guanidine buffer as described above. After heating and sonication, proteins (100 µg each) were purified by methanol–chloroform precipitation and resuspended in 20 µL of 0.1% RapiGest SF in 50 mM triethylammonium bicarbonate. After sonication and heating at 95°C for 10 min, the proteins were digested with 1 µg trypsin/Lys-C mix (Promega) at 37 °C overnight. After heating at 95°C for 10 min, the digested peptides from each sample were labeled with 0.1 mg of TMT 11-plex reagents (Thermo Fisher Scientific) and pooled. The

combined sample was then diluted 6-fold with TBS2 and incubated with the ACN-prewashed
 Tamavidin 2-REV beads for 3 h at 4 °C. Biotinylated peptides were enriched by the optimized
 workflow as described above. LC-MS/MS analysis of the resultant peptides was performed on an
 EASY-nLC 1200 UHPLC connected to a Q Exactive Plus mass spectrometer through a
 nanoelectrospray ion source (Thermo Fisher Scientific). The peptides were separated on the analytical
 column (75 $\mu\text{m} \times 15 \text{ cm}$, 3 μm ; Nikkyo Technos) with a linear gradient of 4–20% for 0–115 min and
 20–32% for 115–160 min, followed by an increase to 80% ACN for 10 min and finally hold at 80%
 ACN for 10 min. The mass spectrometer was operated in data-dependent acquisition mode with a top
 10 MS/MS method. MS1 spectra were measured with a resolution of 70,000, an AGC target of 3e6,
 and a mass range from 375 to 1,400 m/z . MS/MS spectra were triggered at a resolution of 35,000, an
 AGC target of 1e5, an isolation window of 0.7 m/z , a maximum injection time of 150 ms, and a
 normalized collision energy of 34. Dynamic exclusion was set to 20 s. Raw data were directly analyzed
 against the SwissProt database restricted to *Homo sapiens* using Proteome Discoverer version 2.4 with
 the Sequest HT search engine for identification and TMT quantification. The search parameters were
 as follows: (a) trypsin as an enzyme with up to two missed cleavages; (b) precursor mass tolerance of
 10 ppm; (c) fragment mass tolerance of 0.02 Da; (d) TMT of peptide N-terminus and
 carbamidomethylation of cysteine as fixed modifications; (e) TMT of lysine, biotinylation of lysine,
 and oxidation of methionine as variable modifications. Peptides were filtered at an FDR of 1% using
 the Percolator node. TMT quantification was performed using the Reporter Ions Quantifier node.
 Normalization was performed such that the total sum of the abundance values for each TMT channel
 over all peptides was the same.

Data Visualization

Results were analyzed using R version 4.1.1 and plotted using RStudio version 1.4.1717. The peptide coverage of Reticulon-4 protein in Figure 4C was illustrated using Protter (27). Proteins with the GO annotation terms “cell surface” and/or “plasma membrane” and/or “extracellular” listed in UniProtKB were classified as cell surface proteins as reported previously (28, 29). The schematic diagram of EGFR in Figure 5E was illustrated using Pfam (30).

Western Blotting

Cell lysates prepared in guanidine buffer as described above were subjected to methanol–chloroform precipitation and resuspended in SDS-sample buffer. The proteins (10 µg each) were separated by electrophoresis on a 9% polyacrylamide gel and transferred onto a PVDF membrane. After blocking with Blocking One (Nacalai Tesque) for 1 h at room temperature, the membrane was incubated with primary antibodies in TBS-T (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, and 0.05% Tween-20) containing 5% Blocking One overnight at room temperature. After washing twice with TBS-T for 10 min, the membrane was incubated with HRP-linked secondary antibodies for 1 h and then washed twice with TBS-T for 10 min. To detect biotinylated proteins, the transferred membrane was blocked with Blocking One and incubated with HRP-linked streptavidin (BioLegend, #405210, 1:10,000) in TBS-T containing 5% Blocking One for 1 h. Protein bands on the membrane were detected using the ChemiDoc Touch Imaging System (Bio-Rad) after incubation with Clarity Western ECL substrate (Bio-Rad). Antibodies and dilutions used in this study were as follows: anti-STING (D2P2F) rabbit mAb (Cell Signaling Technology, #13647, 1:10,000); anti-EGF receptor (D38B1) XP rabbit mAb (Cell Signaling Technology, #4267, 1:1,000); anti-phospho-ERK1/2 (Thr202/Tyr204) (D13.14.4E) XP rabbit mAb (Cell Signaling Technology, #4370, 1:2,500); anti- α -tubulin (DM1a) mouse mAb (Sigma-Aldrich, T6199, 1:4,000); anti-rabbit IgG, HRP-linked antibody (Cell Signaling Technology,

218 #7074, 1:10,000); anti-mouse IgG, HRP-linked antibody (Cell Signaling Technology, #7076,
219 1:10,000).
220

221 RESULTS AND DISCUSSION

222 Reduction of Interference Ions Generated in the Enrichment Process of Biotinylated Peptides

223 We have previously established STING-knockout (SKO) RAW264.7 macrophages stably expressing
224 TurboID-fused STING (17). Western blot analysis confirmed the expression of TurboID-STING in
225 these cells and its ability to biotinylate cellular proteins by adding 500 μ M biotin to the culture medium
226 for 10 min (Figure S1). This biotinylated sample was used throughout the experiments to optimize the
227 enrichment workflow for biotinylated peptides.

228 In the original workflow, cellular proteins were extracted with guanidine hydrochloride together
229 with one-step reduction and alkylation, and then purified by methanol–chloroform precipitation. After
230 digestion with trypsin in RapiGest surfactant, Pefabloc SC was added to inhibit trypsin activity.
231 Biotinylated peptides were then captured with Tamavidin 2-REV magnetic beads and eluted
232 competitively with excess biotin (17). When the peptide samples obtained by this workflow were
233 subjected to LC-MS/MS, high intensity peaks of interference and/or contaminant ions were observed
234 at the MS1 level, particularly when small amounts of biotinylated peptides were analyzed (Figure 1A,B,
235 Condition 1). These peaks could cause ion suppression and inefficient identification of biotinylated
236 peptides. Therefore, we attempted to optimize each step of the enrichment process to reduce these
237 interference ions and efficiently elute biotinylated peptides from Tamavidin 2-REV.

238 First, to examine the effect of surfactants that solubilize precipitated proteins and promote enzymatic
239 digestion (31-33), we compared RapiGest (34) with a phase-transfer surfactant (PTS) composed of
240 sodium deoxycholate (SDC) and sodium lauroylsarcosinate (SLS) (24-26). As a result, there were
241 almost no differences in MS1 chromatograms between RapiGest and PTS (Figure 1A,B, Condition 1
242 vs. Condition 2); however, we started using PTS because it is more cost-effective than RapiGest.

243 Second, we focused on the serine protease inhibitor Pefabloc SC, 4-(2-aminoethyl)-benzenesulfonyl
244 fluoride (35) that was used to inactivate trypsin and prevent digestion of Tamavidin 2-REV when

245 capturing biotinylated peptides. Because trypsin is inactivated by heat (36, 37), we compared the
246 effects of Pefabloc SC and heat treatment on MS1 chromatograms and found that the interference ions
247 observed around 50 min (Figure 1B, (i)) were dramatically reduced in the heat-treated sample (Figure
248 1A,B, Condition 2 vs. Condition 3). These results indicate that heat treatment is superior to the addition
249 of Pefabloc SC.

250 Third, we examined a series of interference ions detected between 35 min and 45 min on MS1
251 chromatograms (Figure 1B, (ii)). The repeating and characteristic m/z patterns suggested
252 contamination by polymer-like materials from the Tamavidin 2-REV magnetic beads. Therefore, when
253 the beads were pre-washed with 10% acetonitrile before capturing biotinylated peptides, these
254 interference ions were clearly reduced (Figure 1A,B, Condition 2 vs. Condition 4).

255 Fourth, to investigate contamination in the biotin solution used to elute biotinylated peptides from
256 the beads, we compared eluates with the biotin solutions with or without passing through the GL-Tip
257 SDB desalting tip. The results showed that some contaminant ions (Figure 1B, (iii)) were removed
258 when the biotin solution had been passed through GL-Tip SDB (Figure 1A,B, Condition 2 vs.
259 Condition 5). To confirm that these contaminant ions were not derived from digested peptides but from
260 the biotin solution, Tamavidin 2-REV beads that had not been incubated with the digested peptides
261 were eluted with the biotin solution or biotin-free solution, and the eluates were analyzed by LC-
262 MS/MS. While five intense peaks were detected in the eluate with the biotin solution (Figure 1C,
263 arrows), these peaks were strongly reduced in the eluate with the biotin-free solution. These
264 contaminants had hydrophobic properties and could not be removed by desalting the eluates using GL-
265 Tip SDB. Thus, we eluted biotinylated peptides with the biotin solution that had been passed through
266 GL-Tip SDB.

Figure 1. Interference ions generated under various conditions to enrich biotinylated peptides. (A) Table showing the conditions tested to reduce interference ions. Changes from the original workflow are highlighted in red. (B) MS1 chromatograms of peptide samples obtained under each condition. Interference ions indicated by (i), (ii), and (iii) are derived from Pefabloc SC, Tamavidin 2-REV beads, and biotin solution, respectively. Relative abundance is on the same scale across conditions. (C) MS1 chromatograms of the eluate with biotin solution (TBS2 containing 1 mM biotin, top) or with biotin-free solution (TBS2, bottom) from Tamavidin 2-REV beads.

Optimization of Elution Conditions for Biotinylated Peptides from Tamavidin 2-REV

Next, we attempted to optimize the conditions for competitive elution of biotinylated peptides from the Tamavidin 2-REV beads. First, we focused on the temperature at which biotinylated peptides were eluted from the beads. In the original workflow, biotinylated peptides were eluted with the biotin solution at 95 °C. We therefore compared four temperatures: 4, 37, 56, and 95 °C. Among them, the highest number of biotinylated peptides was identified at 37 °C (Figure 2A). The interference ions were increased with the increase in the elution temperature, suggesting that these interference ions were non-specifically eluted from the beads by heating (Figure S2).

We next investigated whether the contamination by non-biotinylated peptides in the eluates could be reduced by adding a step of ‘mock’ elution with biotin-free buffer immediately before elution with the biotin solution. After the beads were incubated with biotin-free buffer for 15 min at 37 °C, the buffer was removed and biotinylated peptides were eluted from the beads with the biotin solution at 37 °C. As a result, non-biotinylated peptides were efficiently removed from the eluate, which increased the ratio of biotinylated to non-biotinylated peptides (Figure 2B).

The previous crystal structure analysis had revealed that Tamavidin 2 forms multiple hydrogen bonds with biotin (38), and therefore we considered the possibility that biotinylated peptides could be efficiently eluted by increasing the salt concentration to weaken the hydrogen bonds. After the mock elution, biotinylated peptides were eluted from the beads at 37 °C with the biotin solutions containing various concentrations of NaCl. As a result, the highest number of biotinylated peptides was identified around 500 mM NaCl (Figure 2C). On the basis of these results of optimizing the elution conditions, we eluted biotinylated peptides from Tamavidin 2-REV beads with the biotin solution containing 500 mM NaCl at 37 °C after the mock elution.

Figure 2. Optimal elution conditions for biotinylated peptides. (A) Comparison of temperatures of biotin solution to elute biotinylated peptides from Tamavidin 2-REV beads. (B) Effect of ‘mock’ elution on the contamination by non-biotinylated peptides in the eluates. (C) Comparison of NaCl concentrations of biotin solution to elute biotinylated peptides.

290 Comparison of Biotinylated Peptides Enriched by the Original and Optimized Workflows

291 To evaluate the optimized workflow, biotinylated peptides enriched by the original and optimized
 292 workflows from 300 µg proteins of RAW264.7 cells expressing TurboID-STING were compared by
 293 LC-MS/MS analysis. MS1 chromatograms showed that many peaks of interference ions in the original
 294 workflow were greatly reduced in the optimized workflow (Figure 3A). Furthermore, the number of
 295 identified biotinylated peptides was increased by more than two-fold after the optimization, and the
 296 contamination by non-biotinylated peptides was reduced to almost one-third, resulting in
 297 approximately 95% of all identified peptides being biotinylated (Figure 3B). Thus, we have established
 298 an optimized workflow to efficiently and specifically enrich biotinylated peptides using Tamavidin 2-
 299 REV.

Figure 3. Comparison of biotinylated peptides enriched by the original and optimized workflows. (A) MS1 chromatograms of peptide samples obtained by the original workflow (top) and the optimized workflow (bottom). Relative abundance is on the same scale between the two workflows. (B) The numbers of biotinylated peptides and non-biotinylated peptides that were identified by the original workflow or the optimized workflow.

301 Sequential Digestion with Different Proteases

302 BioID-based proximity biotinylation of proteins occurs on lysine residues, which can inhibit efficient
 303 digestion with trypsin (39, 40). Previous studies of cross-linking MS using lysine-reactive cross-linkers
 304 (*e.g.*, DSSO) have shown that sequential digestion with Lys-C or other non-lysine-specific proteases
 305 (*e.g.*, Glu-C) followed by trypsin increases the coverage of identified cross-linked peptides (41, 42).
 306 We therefore examined three combinations of different proteases to digest biotinylated proteins from
 307 TurboID-STING-expressing cells: trypsin alone, Lys-C followed by trypsin, and Glu-C followed by
 308 trypsin. First, we examined whether PTS buffer used to solubilize precipitated proteins was compatible
 309 with digestion with these proteases, and confirmed that each protease cleaved at specific residues; that
 310 is, > 97% of all cleavages by trypsin occurred at lysine or arginine residues, >96% of all cleavages by
 311 Lys-C occurred at lysine residues, and 97% of all cleavages by Glu-C occurred at aspartate or
 312 glutamate residues (Figure S3) (43). Next, when biotinylated peptides were enriched after digestion
 313 with the three combinations of different proteases, digestion with trypsin alone identified the highest
 314 number of biotinylation sites (Figure 4A) and the highest number of unique biotinylation sites (Figure
 315 4B). Sequential digestion with Lys-C and trypsin yielded similar results to digestion with trypsin alone.
 316 For example, many of identified biotinylation sites were shared by the two digestions (Figure 4A,B).
 317 On the other hand, sequential digestion with Glu-C and trypsin yielded many unique biotinylation sites
 318 that were not identified by trypsin alone or by Lys-C and trypsin, although the total number of
 319 identified biotinylation sites was less (Figure 4A,B).

320 Consistent with the endoplasmic reticulum (ER) localization of STING at the steady state, many
 321 biotinylation sites identified by the three digestions were located on ER proteins (Table S1). In the
 322 case of an ER protein Reticulon-4/Nogo (44, 45), while four biotinylation sites were identified by
 323 digestion with trypsin alone, six other biotinylation sites were identified by sequential digestion with
 324 Glu-C and trypsin (Figure 4C). All 10 biotinylation sites were located in the N-terminal cytoplasmic

region of Reticulon-4, consistent with the fusion of TurboID to the cytoplasmic side of STING. Collectively, these results indicate that digestion with trypsin alone is most effective in identifying peptides with biotinylated lysine residues. In addition, sequential digestion with Glu-C and trypsin can identify further biotinylation sites and provide more detailed topological information of membrane proteins.

Figure 4. Biotinylation sites identified by digestion with combinations of different proteases. (A) The numbers of biotinylation sites identified by digestion with trypsin alone, Lys-C followed by trypsin, and Glu-C followed by trypsin. (B) Venn diagram showing the distribution of biotinylation sites identified by digestion with the three combinations of different proteases. (C) Schematic of biotinylated peptides and their sites identified by digestion with trypsin alone or Glu-C followed by trypsin on an ER protein Reticulon-4/Nogo.

Multiplexed Quantitative Cell Surface Proteomics by the Combination of TMT Labeling and Biotinylated Peptide Enrichment

Finally, quantitative cell surface proteome profiling was performed by applying this optimized workflow. Selective biotinylation of surface-exposed primary amines on cell surface proteins is one of the most frequent approaches, and the membrane-impermeable reagent Sulfo-NHS-biotin is typically used (40). We examined changes in the surface proteome of HeLa cells stimulated with epidermal growth factor (EGF), which is well known to induce the internalization and degradation of EGF

338 receptor (EGFR) (46). First, we confirmed that EGF stimulation of HeLa cells induced rapid and
339 transient phosphorylation of ERK1/2 and gradual degradation of EGFR, as previously reported (Figure
340 S4A) (47, 48). Next, surface biotinylation was performed at four time points after EGF stimulation in
341 multiple biological replicates, followed by digestion of cellular proteins with trypsin. After TMT
342 labeling and mixing of digested peptides, biotinylated peptides were enriched using Tamavidin 2-REV
343 (Figure 5A). In total, 7,996 biotinylated peptides were identified and quantified by LC-MS/MS
344 analysis (Table S2) and classified as peptides of ‘Surface proteins’ or ‘Non-surface proteins’ according
345 to the GO cellular component annotation used in previous reports (28, 29). The number of quantified
346 biotinylated peptides classified as ‘Surface proteins’ accounted for 66% of the total (Figure S4B),
347 indicating that biotinylated peptides derived from cell surface proteins were efficiently enriched. We
348 further examined changes in relative abundances of the biotinylated peptides classified as ‘Surface
349 proteins’ during the time course of EGF stimulation (Figure 5B). As a result, 11 biotinylated peptides
350 derived from the extracellular region of EGFR were detected (Figure 5C,D), and all 11 biotinylated
351 peptides were specifically decreased by EGF stimulation (Figure 5B,D), confirming the specific
352 internalization of EGFR by EGF stimulation. These results clearly demonstrated that the combination
353 of TMT-based multiplexing and this optimized workflow enables large-scale quantification of the cell
354 surface proteome.

Figure 5. TMT-based multiplexed quantification of biotinylated peptides from EGF-stimulated and surface-biotinylated HeLa cells. (A) Schematic workflow of quantitative cell surface proteomics using surface biotinylation, TMT labeling, and Tamavidin 2-REV enrichment. (B) Changes in relative abundances of all 5,272 biotinylated peptides classified as 'Surface proteins' after EGF stimulation. The thick orange line and the surrounding blue ribbon indicate the mean abundance and the area between maximum and minimum abundance of 11 biotinylated peptides derived from the extracellular region of EGFR, respectively. Other biotinylated peptides are shown in gray. (C) Schematic diagram of EGFR with identified biotinylation sites in the extracellular region. (D) EGF-induced decrease of the 11 biotinylated peptides derived from the extracellular region of EGFR. Lines and surrounding ribbons indicate changes in mean abundances and standard deviations between biological replicates, respectively.

356 CONCLUSIONS

357 Global identification of protein–protein interactions and subcellular proteomes in living cells is
 358 essential to elucidate regulatory mechanisms of various biological processes. Proximity-dependent
 359 biotinylation has emerged as a promising approach that overcomes the limitations of classical
 360 biochemical fractionation and affinity purification. Various techniques for proximity-dependent
 361 biotinylation followed by MS analysis have been rapidly developed and improved. In particular, BioID
 362 enzymes have been subjected to intensive engineering and optimization. Although specific enrichment
 363 of biotinylated proteins/peptides from complex biological samples is a critical step for proximity-
 364 dependent identification, the enrichment methods have not been fully developed. In this study, we have
 365 established an optimized workflow for enrichment and identification of biotinylated peptides using
 366 Tamavidin 2-REV. In contrast to the conventional enrichment and identification of biotinylated
 367 proteins using streptavidin, identification of biotinylated peptides provides information about the
 368 biotinylation sites on proteins. In addition to trypsin digestion alone, sequential digestion with Glu-C
 369 and trypsin can provide more comprehensive biotinylation sites and more detailed topological
 370 information of membrane proteins. Because advances in proximity-dependent biotinylation and
 371 chemical biology such as acyl-biotin exchange are increasing the importance of MS-based
 372 identification of biotinylated proteins and their sites, this optimized workflow may be useful for a wide
 373 range of applications.

374

375 **Author Contributions**

376 K.N., H.Y., and H.K. conceived the project and analyzed the data. K.N., H.Y., and K.M. performed
377 the experiments. H.Y. and H.K. wrote the manuscript. H.K. supervised the project. All authors read
378 and approved the final manuscript.

379

380 **Notes**

381 The authors declare no competing financial interest. The mass spectrometry raw files and search results
382 have been deposited in the ProteomeXchange Consortium via the jPOST partner repository (49) with
383 the dataset identifiers PXD032059, PXD032060, PXD032061, and PXD032062.

384

385 **ACKNOWLEDGMENTS**

386 This work was supported by the JSPS KAKENHI (17K08661 and 20K07340 to K.M., 21K06138 to
387 H.Y., and 19H04966 and 20K06628 to H.K.) and the Takeda Science Foundation (to K.M. and H.Y.).
388 We thank Megumi Kawano and Mayumi Kajimoto for technical assistance and Mayumi Iwata for
389 secretarial assistance. We also thank Mitchell Arico from Edanz (<https://jp.edanz.com/ac>) for editing
390 a draft of this manuscript.

391

392 REFERENCES

- 393 1. Dundas, C. M.; Demonte, D.; Park, S., Streptavidin-biotin technology: improvements and
394 innovations in chemical and biological applications. *Appl Microbiol Biotechnol* **2013**, 97, (21),
395 9343-53.
- 396 2. Fairhead, M.; Howarth, M., Site-specific biotinylation of purified proteins using BirA. *Methods*
397 *Mol Biol* **2015**, 1266, 171-84.
- 398 3. Kim, D. I.; Roux, K. J., Filling the Void: Proximity-Based Labeling of Proteins in Living Cells.
399 *Trends Cell Biol* **2016**, 26, (11), 804-817.
- 400 4. Qin, W.; Cho, K. F.; Cavanagh, P. E.; Ting, A. Y., Deciphering molecular interactions by proximity
401 labeling. *Nat Methods* **2021**, 18, (2), 133-143.
- 402 5. Samavarchi-Tehrani, P.; Samson, R.; Gingras, A. C., Proximity Dependent Biotinylation: Key
403 Enzymes and Adaptation to Proteomics Approaches. *Mol Cell Proteomics* **2020**, 19, (5), 757-773.
- 404 6. Roux, K. J.; Kim, D. I.; Raida, M.; Burke, B., A promiscuous biotin ligase fusion protein identifies
405 proximal and interacting proteins in mammalian cells. *J Cell Biol* **2012**, 196, (6), 801-10.
- 406 7. Rhee, H. W.; Zou, P.; Udeshi, N. D.; Martell, J. D.; Mootha, V. K.; Carr, S. A.; Ting, A. Y.,
407 Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging.
408 *Science* **2013**, 339, (6125), 1328-1331.
- 409 8. Liu, G.; Papa, A.; Katchman, A. N.; Zakharov, S. I.; Roybal, D.; Hennessey, J. A.; Kushner, J.;
410 Yang, L.; Chen, B. X.; Kushnir, A.; Dangas, K.; Gygi, S. P.; Pitt, G. S.; Colecraft, H. M.; Ben-
411 Johny, M.; Kalocsay, M.; Marx, S. O., Mechanism of adrenergic CaV1.2 stimulation revealed by
412 proximity proteomics. *Nature* **2020**, 577, (7792), 695-700.
- 413 9. St-Germain, J. R.; Samavarchi Tehrani, P.; Wong, C.; Larsen, B.; Gingras, A. C.; Raught, B.,
414 Variability in Streptavidin-Sepharose Matrix Quality Can Significantly Affect Proximity-
415 Dependent Biotinylation (BioID) Data. *J Proteome Res* **2020**, 19, (8), 3554-3561.
- 416 10. Go, C. D.; Knight, J. D. R.; Rajasekharan, A.; Rathod, B.; Hesketh, G. G.; Abe, K. T.; Youn, J. Y.;
417 Samavarchi-Tehrani, P.; Zhang, H.; Zhu, L. Y.; Popiel, E.; Lambert, J. P.; Coyaud, E.; Cheung, S.
418 W. T.; Rajendran, D.; Wong, C. J.; Antonicka, H.; Pelletier, L.; Palazzo, A. F.; Shoubbridge, E. A.;
419 Raught, B.; Gingras, A. C., A proximity-dependent biotinylation map of a human cell. *Nature*
420 **2021**, 595, (7865), 120-124.

- 421 11. Forrester, M. T.; Foster, M. W.; Stamler, J. S., Assessment and application of the biotin switch
422 technique for examining protein S-nitrosylation under conditions of pharmacologically induced
423 oxidative stress. *J Biol Chem* **2007**, 282, (19), 13977-83.
- 424 12. Wan, J.; Roth, A. F.; Bailey, A. O.; Davis, N. G., Palmitoylated proteins: purification and
425 identification. *Nat Protoc* **2007**, 2, (7), 1573-84.
- 426 13. Rexach, J. E.; Clark, P. M.; Hsieh-Wilson, L. C., Chemical approaches to understanding O-
427 GlcNAc glycosylation in the brain. *Nat Chem Biol* **2008**, 4, (2), 97-106.
- 428 14. Yang, J.; Gupta, V.; Tallman, K. A.; Porter, N. A.; Carroll, K. S.; Liebler, D. C., Global, in situ,
429 site-specific analysis of protein S-sulfenylation. *Nat Protoc* **2015**, 10, (7), 1022-37.
- 430 15. Udeshi, N. D.; Pedram, K.; Svinkina, T.; Fereshetian, S.; Myers, S. A.; Aygun, O.; Krug, K.;
431 Clauser, K.; Ryan, D.; Ast, T.; Mootha, V. K.; Ting, A. Y.; Carr, S. A., Antibodies to biotin enable
432 large-scale detection of biotinylation sites on proteins. *Nat Methods* **2017**, 14, (12), 1167-1170.
- 433 16. Kim, D. I.; Cutler, J. A.; Na, C. H.; Reckel, S.; Renuse, S.; Madugundu, A. K.; Tahir, R.;
434 Goldschmidt, H. L.; Reddy, K. L.; Hukanir, R. L.; Wu, X.; Zachara, N. E.; Hantschel, O.; Pandey,
435 A., BioSITE: A Method for Direct Detection and Quantitation of Site-Specific Biotinylation. *J*
436 *Proteome Res* **2018**, 17, (2), 759-769.
- 437 17. Motani, K.; Kosako, H., BioID screening of biotinylation sites using the avidin-like protein
438 Tamavidin 2-REV identifies global interactors of stimulator of interferon genes (STING). *J Biol*
439 *Chem* **2020**, 295, (32), 11174-11183.
- 440 18. Takakura, Y.; Sofuku, K.; Tsunashima, M., Tamavidin 2-REV: an engineered tamavidin with
441 reversible biotin-binding capability. *J Biotechnol* **2013**, 164, (1), 19-25.
- 442 19. Yamanaka, S.; Horiuchi, Y.; Matsuoka, S.; Kido, K.; Nishino, K.; Maeno, M.; Shibata, N.; Kosako,
443 H.; Sawasaki, T., A proximity biotinylation-based approach to identify protein-E3 ligase
444 interactions induced by PROTACs and molecular glues. *Nat Commun* **2022**, 13, (1), 183.
- 445 20. Motani, K.; Kosako, H., Activation of stimulator of interferon genes (STING) induces ADAM17-
446 mediated shedding of the immune semaphorin SEMA4D. *J Biol Chem* **2018**, 293, (20), 7717-7726.
- 447 21. Ohsugi, M.; Adachi, K.; Horai, R.; Kakuta, S.; Sudo, K.; Kotaki, H.; Tokai-Nishizumi, N.; Sagara,
448 H.; Iwakura, Y.; Yamamoto, T., Kid-mediated chromosome compaction ensures proper nuclear
449 envelope formation. *Cell* **2008**, 132, (5), 771-82.

450 22. Kim, K. L.; Kim, D.; Lee, S.; Kim, S. J.; Noh, J. E.; Kim, J. H.; Chae, Y. C.; Lee, J. B.; Ryu, S.
451 H., Pairwise detection of site-specific receptor phosphorylations using single-molecule blotting.
452 *Nat Commun* **2016**, 7, 11107.

453 23. Li, Y.; Wang, Y.; Mao, J.; Yao, Y.; Wang, K.; Qiao, Q.; Fang, Z.; Ye, M., Sensitive profiling of
454 cell surface proteome by using an optimized biotinylation method. *J Proteomics* **2019**, 196, 33-41.

455 24. Masuda, T.; Tomita, M.; Ishihama, Y., Phase transfer surfactant-aided trypsin digestion for
456 membrane proteome analysis. *J Proteome Res* **2008**, 7, (2), 731-40.

457 25. Masuda, T.; Saito, N.; Tomita, M.; Ishihama, Y., Unbiased quantitation of Escherichia coli
458 membrane proteome using phase transfer surfactants. *Mol Cell Proteomics* **2009**, 8, (12), 2770-7.

459 26. Masuda, T.; Sugiyama, N.; Tomita, M.; Ohtsuki, S.; Ishihama, Y., Mass Spectrometry-Compatible
460 Subcellular Fractionation for Proteomics. *J Proteome Res* **2020**, 19, (1), 75-84.

461 27. Omasits, U.; Ahrens, C. H.; Muller, S.; Wollscheid, B., Protter: interactive protein feature
462 visualization and integration with experimental proteomic data. *Bioinformatics* **2014**, 30, (6), 884-
463 6.

464 28. Ozlu, N.; Qureshi, M. H.; Toyoda, Y.; Renard, B. Y.; Mollaoglu, G.; Ozkan, N. E.; Bulbul, S.;
465 Poser, I.; Timm, W.; Hyman, A. A.; Mitchison, T. J.; Steen, J. A., Quantitative comparison of a
466 human cancer cell surface proteome between interphase and mitosis. *EMBO J* **2015**, 34, (2), 251-
467 65.

468 29. Wojdyla, K.; Collier, A. J.; Fabian, C.; Nisi, P. S.; Biggins, L.; Oxley, D.; Rugg-Gunn, P. J., Cell-
469 Surface Proteomics Identifies Differences in Signaling and Adhesion Protein Expression between
470 Naive and Primed Human Pluripotent Stem Cells. *Stem Cell Reports* **2020**, 14, (5), 972-988.

471 30. Mistry, J.; Chuguransky, S.; Williams, L.; Qureshi, M.; Salazar, G. A.; Sonnhammer, E. L. L.;
472 Tosatto, S. C. E.; Paladin, L.; Raj, S.; Richardson, L. J.; Finn, R. D.; Bateman, A., Pfam: The
473 protein families database in 2021. *Nucleic Acids Res* **2021**, 49, (D1), D412-d419.

474 31. Leon, I. R.; Schwammle, V.; Jensen, O. N.; Sprenger, R. R., Quantitative assessment of in-solution
475 digestion efficiency identifies optimal protocols for unbiased protein analysis. *Mol Cell Proteomics*
476 **2013**, 12, (10), 2992-3005.

- 477 32. Kulak, N. A.; Pichler, G.; Paron, I.; Nagaraj, N.; Mann, M., Minimal, encapsulated proteomic-
478 sample processing applied to copy-number estimation in eukaryotic cells. *Nat Methods* **2014**, 11,
479 (3), 319-24.
- 480 33. Steger, M.; Demichev, V.; Backman, M.; Ohmayer, U.; Ihmor, P.; Muller, S.; Ralser, M.; Daub,
481 H., Time-resolved in vivo ubiquitinome profiling by DIA-MS reveals USP7 targets on a proteome-
482 wide scale. *Nat Commun* **2021**, 12, (1), 5399.
- 483 34. Chen, E. I.; Cociorva, D.; Norris, J. L.; Yates, J. R., 3rd, Optimization of mass spectrometry-
484 compatible surfactants for shotgun proteomics. *J Proteome Res* **2007**, 6, (7), 2529-38.
- 485 35. Powers, J. C.; Asgian, J. L.; Ekici, O. D.; James, K. E., Irreversible inhibitors of serine, cysteine,
486 and threonine proteases. *Chem Rev* **2002**, 102, (12), 4639-750.
- 487 36. Pace, J., The inactivation of trypsin by heat. *Biochem J* **1930**, 24, (3), 606-14.
- 488 37. Petritis, B. O.; Qian, W. J.; Camp, D. G., 2nd; Smith, R. D., A simple procedure for effective
489 quenching of trypsin activity and prevention of 18O-labeling back-exchange. *J Proteome Res*
490 **2009**, 8, (5), 2157-63.
- 491 38. Takakura, Y.; Tsunashima, M.; Suzuki, J.; Usami, S.; Kakuta, Y.; Okino, N.; Ito, M.; Yamamoto,
492 T., Tamavidins--novel avidin-like biotin-binding proteins from the Tamogitake mushroom. *FEBS*
493 *J* **2009**, 276, (5), 1383-97.
- 494 39. Herber, J.; Njavro, J.; Feederle, R.; Schepers, U.; Muller, U. C.; Brase, S.; Muller, S. A.;
495 Lichtenthaler, S. F., Click Chemistry-mediated Biotinylation Reveals a Function for the Protease
496 BACE1 in Modulating the Neuronal Surface Glycoproteome. *Mol Cell Proteomics* **2018**, 17, (8),
497 1487-1501.
- 498 40. Li, Y.; Qin, H.; Ye, M., An overview on enrichment methods for cell surface proteome profiling.
499 *J Sep Sci* **2020**, 43, (1), 292-312.
- 500 41. Ser, Z.; Cifani, P.; Kentsis, A., Optimized Cross-Linking Mass Spectrometry for in Situ Interaction
501 Proteomics. *J Proteome Res* **2019**, 18, (6), 2545-2558.
- 502 42. Stieger, C. E.; Doppler, P.; Mechtler, K., Optimized Fragmentation Improves the Identification of
503 Peptides Cross-Linked by MS-Cleavable Reagents. *J Proteome Res* **2019**, 18, (3), 1363-1370.

504 43. Nishida, H.; Ishihama, Y., One-Step Isolation of Protein C-Terminal Peptides from V8 Protease-
505 Digested Proteins by Metal Oxide-Based Ligand-Exchange Chromatography. *Anal Chem* **2022**,
506 94, (2), 944-951.

507 44. Jozsef, L.; Tashiro, K.; Kuo, A.; Park, E. J.; Skoura, A.; Albinsson, S.; Rivera-Molina, F.; Harrison,
508 K. D.; Iwakiri, Y.; Toomre, D.; Sessa, W. C., Reticulon 4 is necessary for endoplasmic reticulum
509 tubulation, STIM1-Orai1 coupling, and store-operated calcium entry. *J Biol Chem* **2014**, 289, (13),
510 9380-95.

511 45. Wang, S.; Tukachinsky, H.; Romano, F. B.; Rapoport, T. A., Cooperation of the ER-shaping
512 proteins atlastin, lunapark, and reticulons to generate a tubular membrane network. *Elife* **2016**, 5.

513 46. Gorden, P.; Carpentier, J. L.; Cohen, S.; Orci, L., Epidermal growth factor: morphological
514 demonstration of binding, internalization, and lysosomal association in human fibroblasts. *Proc*
515 *Natl Acad Sci U S A* **1978**, 75, (10), 5025-9.

516 47. Felder, S.; Miller, K.; Moehren, G.; Ullrich, A.; Schlessinger, J.; Hopkins, C. R., Kinase activity
517 controls the sorting of the epidermal growth factor receptor within the multivesicular body. *Cell*
518 **1990**, 61, (4), 623-34.

519 48. Dikic, I.; Giordano, S., Negative receptor signalling. *Curr Opin Cell Biol* **2003**, 15, (2), 128-35.

520 49. Okuda, S.; Watanabe, Y.; Moriya, Y.; Kawano, S.; Yamamoto, T.; Matsumoto, M.; Takami, T.;
521 Kobayashi, D.; Araki, N.; Yoshizawa, A. C.; Tabata, T.; Sugiyama, N.; Goto, S.; Ishihama, Y.,
522 jPOSTrepo: an international standard data repository for proteomes. *Nucleic Acids Res* **2017**, 45,
523 (D1), D1107-D1111.

524