Development and Comparative Evaluation of Endolysosomal Proximity Labeling-based Proteomic Methods in Human iPSC-derived Neurons.

Ashley M. Frankenfield[1], Michael S Fernandopulle[2], Saadia Hasan[2], Michael Ward[2], Ling Hao[1*]

[1]Department of Chemistry, The George Washington University, Washington, DC, USA.;

[2]National Institute of Neurological Disorders and Stroke, NIH, Bethesda, MD, USA.

*Corresponding author

Tel.: +1 (202) 994-4492

E-mail: linghao@gwu.edu

ABSTRACT

Proximity-based *in situ* labeling techniques offer a unique way to capture both stable and transient protein-protein and protein-organelle interactions. Combining this technology with mass spectrometry (MS)-based proteomics allows us to obtain snapshots of molecular microenvironments with nanometer resolution, facilitating the discovery of complex and dynamic interaction networks. However, a number of technical challenges still exist, such as interferences from endogenously biotinylated proteins and other highly abundant bystanders, the challenge of selecting proper controls to minimize false discoveries, and experimental variations among biological/technical replicates. Here, we developed a new method to capture the proteomic microenvironment of the neuronal endolysosomal network, by knocking in (KI) an engineered ascorbate peroxidase (APEX2) gene to the endogenous locus of lysosome-associated membrane protein 1 (LAMP1). We systematically investigated and optimized the key parameters involved in proximity labeling MS to address the major challenges in the field. We also conducted comparative evaluation between this KI-LAMP1-APEX2 method and our overexpressed LAMP1-APEX2 probes, achieving complementary identification of both known and novel lysosomal membrane and membrane-interacting proteins in human iPSC-derived neurons. To summarize, this study demonstrated new analytical tools to characterize lysosomal functions and interactions in human neurons and filled critical gaps in the field for designing and optimizing proximity labeling technologies.

KEYWORDS: Proximity labeling mass spectrometry, Lysosome, Endolysosome, LAMP1, APEX, iPSC-derived neuron, Biotinylated proteins, Streptavidin

INTRODUCTION

Protein-protein interactions are essential for almost all cellular processes and provide key insights into disease processes.[1,2] Traditionally, protein-protein interactions are characterized by yeast two-hybrid assays or affinity-purification followed by mass spectrometry (AP-MS) or western blotting.[3–5] However, AP-MS often fails to capture transient and weak protein interactions, which are crucial for signal transduction and cellular transport processes. In recent years, proximity labeling (PL)-based techniques have emerged as a new class of tools for identifying both stable and transient protein interactions and molecular microenvironment in living cells and organisms.[6,7] PL enzymes can be genetically fused to a target protein (i.e. "bait") to biotinylate neighboring protein interactors (i.e. "prey") upon activation. Biotinylated proteins can then be selectively purified by streptavidin-coated beads, and subsequently analyzed by MS-based proteomics.

Biotin ligase and engineered ascorbate peroxidase represent the two most commonly used classes of PL enzymes, both developed in 2012.[8,9] Bifunctional ligase/repressor (BioID) is a promiscuous mutant of *Escherichia coli* biotin protein ligase, which can be genetically tagged onto a bait protein to biotinylate neighboring proteins upon the addition of biotin substrate.[8,10] However, BioID method often requires a 12-24 hour labeling period, causing the diffusion of the reactive biotin cloud and non-specific labeling.[11] Recently developed biotin ligase, TurboID, mitigated this issue and shortened the labeling time to 10 min.[12] The other type of PL enzyme is the engineered soybean ascorbate peroxidase (APEX/APEX2), originally developed by the Ting group to map protein interactions within the mitochondria lumen for electron microscopy.[9] Nowadays, APEX2-based proximity labeling has become a powerful tool for identifying protein-protein and protein-RNA interactions in different compartments in cells and whole organisms. [13–17] The major advantages of APEX over biotin ligases are the rapid and highly efficient labeling reaction (≤ 1 min) and smaller size (27 kDa) compared to BioID/TurboID (35 kDa), minimizing the potential impact on the function of the fusion protein. Upon addition of the biotin-phenol substrate, the APEX enzyme can be activated by rapid hydrogen peroxide treatment to generate highly reactive phenoxyl radicals that form a covalent bond with electron-rich aromatic amino acid residues (e.g. tyrosine) from proteins within a 20 nm

60 labeling radius. Biotinylated proteins can then be enriched by streptavidin-coated beads followed by MS-
61 based proteomic strategies.

62 Despite the advances of proximity labeling methods, a number of technical challenges still exist
63 for PL-MS. For example, the addition of biotin substrate in PL experiments enriches the endogenously
64 biotinylated proteins within the mitochondria, which can occupy streptavidin binding sites and compete
65 with other proteins/peptides in LC-MS/MS analysis.[18] The use of hydrogen peroxide for APEX activation
66 raised concerns about oxidative stress.[19,20] Experimental variations between different batches of cell culture
67 during PL activation hinder the accurate and precise quantification of protein-protein interactions. The
68 selection of proper controls is also critical to minimize both false-positive and false-negative protein
69 interactions. These challenges have limited the applicability of PL-MS and generate technical gaps between
70 method development and successful biological applications.

71 In this study, we aimed to address these major challenges by developing a new endogenous
72 APEX2-based PL-MS method and conducting a thorough investigation and optimization of key factors
73 involved in the PL-based proteomic workflow. We integrated the APEX2 gene onto the endogenous locus
74 of lysosomal-associated membrane protein 1 (LAMP1), localizing to the cytosolic surface of
75 endolysosomes in human induced pluripotent stem cells (hiPSCs)-derived neurons. Lysosomes are acidic
76 vacuoles that digest and recycle macromolecules through a cellular mechanism known as autophagy.[21]
77 Lysosomes are particularly important for neurons, because neurons are highly polarized, postmitotic, and
78 rely on autophagy to maintain cellular homeostasis.[22] LAMP1 is an abundant transmembrane protein on the
79 endolysosomal membrane and has served as a classical lysosomal marker in biochemistry assays.[23] The
80 LAMP1-APEX proteomic workflow was systematically optimized to reduce interferences and variations,
81 improve reproducibility and specificity, and address the potential effect of hydrogen peroxide on cellular
82 processes. A comparative evaluation was also conducted between this endogenous knock-in (KI)-LAMP1-
83 APEX method and our two other overexpressed LAMP1-APEX probes to discover known and novel
84 lysosome membrane and lysosome-interacting proteins in hiPSC-derived neurons.

85 **EXPERIMENTAL SECTION**

86 ***Human iPSC Culture and Development of LAMP1-APEX Lines***

87 Wild type hiPSCs from a WTC11 control male were obtained from Coriell and routinely
88 maintained in Matrigel-coated plates in Essential 8 Flex media as described previously.[24] To generate an
89 endogenously-expressed APEX2 probe targeted to the endolysosomal membrane, hiPSCs were transfected
90 with a custom DNA plasmid containing a LAMP1-APEX2 homologous recombination construct. Three
91 stable LAMP1-APEX2 iPSC lines were generated in this study for comparative analysis, endogenous
92 knock-in (KI) LAMP1-APEX, KuB-LAMP1-APEX2 (high overexpression) and KuD-LAMP1-APEX2
93 (modest overexpression). The expression levels of LAMP1-APEX2 in iPSCs are KuB> KuD> KI. The
94 successful integration of APEX2 into the LAMP1 locus was confirmed by PCR for each transgenic iPSC
95 line. Nuclear exporting signal (NES)-APEX2 (overexpression) was also generated as a cytosolic localized
96 APEX2 line. The detailed steps to generate stable transgenic iPSCs lines are described in Supporting
97 Information and our previous study.[25]

98 ***Differentiation of hiPSCs into i^{3}Neurons***

99 We implemented the advanced i^{3}Neuron platform developed in the Ward Lab as our primary
100 cellular platform in this study.[24,26] The hiPSCs used in this study contain a stably integrated doxycycline-
101 inducible neurogenin-2 cassette, which promotes the differentiation of the stem cells into functionally
102 mature glutamatergic cortical neurons in two weeks.[26] LAMP1-APEX hiPSCs were differentiated into
103 cortical neurons as described previously.[24] Briefly, transgenic hiPSCs were seeded into a 15cm dish with
104 Neuronal Induction Medium[24], which promotes terminal differentiation within three days. Then, the cells
105 were dissociated with Accutase (StemPro) and seeded onto polyornithine-coated plates. These cultures were
106 maintained in Cortical Neuron Culture Medium and underwent half-medium changes every other day until
107 neuronal maturation for APEX2 activation experiment and harvest.

108 ***In vivo Proximity Labeling in hiPSC-derived Neurons***

109 Prior to APEX2 activation, neurons were incubated with 500µM biotin-phenol at 37 °C for 30 min.
110 Then, 1 mM of H_2O_2 was added to the cells for exactly 1 min at 37 °C to activate APEX2 enzyme. The PL
111 reaction was immediately terminated by aspirating the growth medium and rapid washing and incubating
112 with ice-cold quench buffer (10mM sodium azide, 10mM sodium ascorbate, 5mM TROLOX in PBS)
113 before the addition of the ice-cold lysis buffer (50 mM Tris-HCl, 500 mM NaCl, 0.2 % SDS, 1mM DTT,
114 10mM sodium azide, 10 mM sodium ascorbate, 5 mM TROLOX, cOmplete mini protease inhibitor).
115 Neurons were then rocked at 4°C for 10 min and scraped into QSonica (Q800R) sonication tubes and
116 sonicated for 15 min at 2 °C with alternating 40 seconds on, 20 seconds off cycles. The lysate was
117 centrifuged at 4 °C at 16,500 g to collect supernatants and stored at -80 °C.

118 ***Beads Titration Assay to Determine Optimal Streptavidin Beads/Proteins Ratio***

119 Biotinylated proteins can be captured by streptavidin-coated beads because of the strong non-
120 covalent interactions between biotin and streptavidin (Kd ~10^{-14}). The optimal beads/protein ratio can be
121 determined by the bead titration assay. Protein concentrations of neuron lysates were first determined by
122 DC (detergent-compatible) Colorimetric Protein Assay (Bio-Rad). Twenty micrograms of the protein lysate
123 were added to each 0.5 mL tube containing a series volume of streptavidin magnetic beads (GE) (0, 0.5, 1,
124 2, 5, 10, 20, 40 µL), followed by overnight rotation at 4°C. The next morning, the tubes were placed on a
125 magnetic rack, and 2 µl of the supernatant from each tube was spotted on a dry nitrocellulose membrane.
126 Once completely dried, the membrane was incubated in Odyssey Blocking Buffer for 1 hour, and then
127 Streptavidin Alexa Fluor 680 conjugate (1:1000 in blocking buffer) for 1hr and washed 5 times with TBST
128 buffer. The fluorescent signal of each dot on the membrane was measured under 700 nm wavelength to
129 calculate the optimal beads amount needed for each µg of input protein sample.

130 ***Biotinylated Protein Pull Down and On-Bead Digestion***

131 Protein lysate samples were incubated with streptavidin magnetic beads (with optimal
132 protein/beads ratio) overnight at 4°C. Supernatants were then removed and beads were washed twice with
133 each of the four sequential buffers (Buffer A: 2% SDS; Buffer B: 50mM Tris-HCl, 500mM NaCl, 0.1%

134 deoxycolic acid, 1% Triton-X, 1 mM EDTA; Buffer C: 50mM Tris-HCl, 250mM NaCl, 0.5% deoxycolic
135 acid, 0.5% NP-40, 1 mM EDTA; Buffer D: 50mM Tris-HCl, 2 M Urea). Magnetic beads were washed two
136 more times with Buffer D to remove all residue detergents. For on-beads digestion, 100 µL of Buffer D
137 was added to each tube to resuspend the magnetic beads. The beads were incubated with 5 mM of Tris(2-
138 carboxyethyl)phosphine (TCEP) for 30 min, 15 mM of iodoacetamide for 30 min, and 5 mM of TCEP for
139 10 min in a ThermoMixer shaking at 1200 rpm at 37 °C. Trypsin/Lys-C Mix enzyme (Promega) was then
140 added to the beads for a 16-hour digestion at 37 °C in a ThermoMixer shaking at 1200 rpm. Then a half
141 amount of Trypsin/LysC mix was added for an additional 3 hours of digestion. The sample tubes were then
142 briefly centrifuged and put on a magnetic rack to collect supernatants. The magnetic beads were washed
143 with 50 µL of 50 mM Tris buffer, and supernatants were combined. To quench the digestion, 10%
144 trifluoroacetic acid was added to the supernatant until pH <3. Peptide desalting was achieved by Waters
145 Oasis HLB 96-well extraction plate based on manufacture protocol. Peptide samples were dried under
146 SpeedVac and stored at -80°C.

147 ***LC-MS Analysis***

148 Dried peptide samples were resuspended in 10 µL of 2% acetonitrile (ACN), 0.1% formic acid
149 (FA) in LC-MS grade water and centrifuged to collect supernatant. LC-MS analyses were performed on a
150 Dionex UltiMate 3000 RSLCnano system coupled with a Thermo Scientific Q-Exactive HF mass
151 spectrometer. The mobile phase buffer A was 0.1% FA, 5% DSMO in H_2O, and buffer B was 0.1% FA,
152 5% DSMO in ACN. Two microliters were injected for each sample onto an Easy-spray PepMap C18
153 column (2 µM, 100 Å, 75 µM x 75 cm) with a 2-hour LC gradient and 60 °C column temperature. The LC
154 flow rate is 0.2 µl/min. MS was scanned from *m/z* 360 to1500 at 120K resolution in top 15 data dependent
155 acquisition. Parent masses were isolated (m/z 1.4 window) and fragmented with higher-energy collisional
156 dissociation (HCD) with a normalized collision energy of 27% and dynamic exclusion time of 22.5s.
157 Maximum injection times were 30 ms for MS and 35 ms for MS/MS. Automatic gain control (AGC) targets
158 were $1e^6$ for MS and $2e^5$ for MS/MS.

159 ***Proteomics Data Analysis***

160 LC-MS/MS raw files were analyzed by MaxQuant software for peptide/protein identification and
161 quantification.[27] Uniprot *Homo sapiens* proteome database was used for protein identification (1% false
162 discovery rate cutoff) with a fixed modification of cysteine carbamidomethylation and variable
163 modifications of oxidation of methionine, acetylation of protein N-terminus, and biotin-phenol
164 modification of tyrosine. PEAKS Studio software was used to conduct a semi-open post-translational
165 modification (PTM) search. The PEAKS PTM algorithm was used to search for 309 potential pre-set
166 modifications. Additional modifications for oxidation at other amino acid residues were imported from the
167 Unimod database.[28] Maxquant and PEAKS output files were analyzed in Excel and R for statistical analysis.
168 For data normalization to the most abundant biotinylated protein (PCCA), raw protein intensities from
169 Maxquant output were used to normalize to the intensity of PCCA followed by log-transformation before
170 statistical analysis.

171 ***Fluorescence Imaging of Neurons***

172 Human i^3Neurons expressing LAMP1-APEX2 were incubated in 500μM phenol-biotin at 37 °C
173 for 30 min, stimulated with 1mM H_2O_2 for one second, and immediately fixed in 4% paraformaldehyde.
174 Neurons were washed three times in PBS, then blocked and permeabilized (3% donkey serum and 0.1%
175 saponin in PBS) on a nutator (medium speed) for 30 min at room temperature (RT). Samples were incubated
176 with primary antibody against LAMP1 (mouse monoclonal H3A4, DHSB, 1:1000) overnight on a nutator
177 (medium speed) at 4°C. The following day, samples were washed three times with PBS, then incubated
178 with anti-mouse AF561 secondary antibody (Jackson ImmunoResearch) and Streptavidin-680 (1:1000) for
179 1 hour on a nutator (medium speed) at RT. After two washes with PBS, samples were incubated with
180 Hoechst (nuclear marker) for 10 min at RT. Samples were washed twice with PBS and visualized with a
181 Nikon Eclipse Ti spinning disk confocal microscope.

182 **RESULTS AND DISCUSSIONS**

183 ***Development and Validation of Endogenous LAMP1-APEX Probe***

184 Lysosomes frequently make contacts with other organelles for autophagy and other cellular
185 processes, but these transient and dynamic interactions have historically been restricted to individual
186 observations by live cell microscopy.[22] In order to study both stable and transient lysosomal interactions in
187 a high-throughput fashion, we developed a novel endogenous LAMP1-APEX2 probe in human iPSC-
188 neurons. The overall workflow of LAMP1-APEX proteomics is illustrated in **Figure 1A**. Human iPSCs
189 expressing the LAMP1-APEX2 fusion protein were differentiated into cortical neurons based on our
190 i^3Neuron protocol.[24] After neuronal maturation in 2 weeks, we incubated the neurons with biotin-phenol
191 substrate for 30 min and H_2O_2 for 1 min. Upon rapid activation by H_2O_2, the APEX2 enzyme transforms
192 biotin-phenol molecules into highly reactive phenoxyl radicals that form covalent bonds with electron-rich
193 aromatic amino acid residues (e.g. tyrosine) from proteins within a 20 nm labeling radius (**Figure 1B**).
194 After quenching the reaction, the neurons were harvested, and proteins were extracted from neurons.
195 Streptavidin magnetic beads were used to enrich biotinylated proteins followed by on-beads digestion,
196 peptide desalting, and LC-MS/MS analysis.

197 In our recent study about lysosomal trafficking in neurons, we overexpressed KuB-APEX2 tagged
198 LAMP1 protein in hiPSC-derived neurons and identified Annexin A11 as a molecular tether that links RNA
199 granules to lysosomes during long-distance RNA transport in neuronal axons.[25] Although overexpressing
200 proteins is a common practice in cell biology, it might generate overexpression artifacts that can be picked
201 out by the sensitive LC-MS platform. For instance, overexpressing LAMP1 might expand a small
202 percentage of its localization to the plasma membrane, increasing non-specific labeling background.[29] To
203 address this issue, we developed a new LAMP1-APEX2 probe by knocking in (KI) APEX2 to the
204 endogenous LAMP1 locus, namely KI-LAMP1-APEX2 proteomics. The correct localization of LAMP1-
205 APEX2 probe was confirmed by fluorescence microscopy. As shown in **Figure 1C and Figure S1**,
206 biotinylated protein signals stained by streptavidin (SA) exhibit excellent colocalization with LAMP1-
207 staining in neurons. In the negative control without H_2O_2 activation, no biotinylation signal was observed.

208 The ability of APEX2 to biotinylate neighboring proteins was also validated in a dot-blot assay.
209 The neuronal protein lysate (50 μg) was spotted onto a nitrocellulose membrane and stained against
210 streptavidin. The APEX2 enzyme was only activated when exposed to both biotin-phenol and H_2O_2 (**Figure**
211 **2A**). In control neurons with no APEX expression, no clear signals were observed. Therefore, our KI-
212 LAMP1-APEX2 probe demonstrated specific endolysosomal localization and enzymatic activity to
213 biotinylate proteins on and surrounding the endolysosome membrane in hiPSC-derived neurons.

214 ***Optimization of the Proximity Labeling Experiment***

215 Despite the various applications of PL-proteomics in biological studies, technical challenges still
216 exist that compromise the specificity and accuracy in identifying and quantifying protein interactions. Here,
217 we conducted a systematic optimization of the key factors involved in PL-proteomics to achieve the optimal
218 beads-to-protein ratio, improved reproducibility, reduced interferences, and optimal digestion efficiency. [30]

219 In PL experiments, streptavidin-coated beads are often added in excess to ensure the complete and
220 reproducible capture of all biotinylated proteins from the samples. However, excess beads can cause
221 increased non-specific binding and overwhelming signals of streptavidin peptides from on-beads digestion.
222 Therefore, a precise control of beads-to-protein ratio is critical for PL-MS. We conducted a beads titration
223 assay to determine the optimal beads amount per μg of input protein. The dot-plots signals from supernatant
224 of beads-protein mix are directly proportional to the amount of biotinylated proteins that are not captured
225 by the streptavidin beads. As shown in **Figure 2B**, fluorescence signals of the dot-blots declined with an
226 increase of beads amount until reaching a plateau at 10 μL of beads, indicating the complete capture of
227 biotinylated proteins from the sample (50 μg of protein lysate). Therefore, the optimal beads-to-protein
228 ratio for KI-LAMP1-APEX proteomics was 0.2 μL of beads per μg of input proteins. Since different PL
229 probe captures different amount of prey proteins, beads titration assay should always be conducted for each
230 PL probe targeting on different proteins of interest or has different expression levels.

231 To evaluate the interference background signals from on-beads digestion, we added 2 μg of
232 Trypsin/LysC enzyme to 200 μL of pure streptavidin magnetic beads for overnight digestion without any
233 input protein. As shown in **Figure 3A**, LC-MS chromatography was overwhelmed with streptavidin
234 peptides from the magnetic beads. A total of 35 proteins and 384 peptides were identified as interference
235 signals resulted from streptavidin, trypsin, keratins, and others. When compared to an actual APEX2
236 proteomic experiment, these interference signals accounted for 2% of the identified protein, 3% of peptides,
237 and 8% of MS/MS scans. Based on these interference peptides, we generated an LC-MS/MS exclusion list
238 for all PL experiments. The starting *m/z* for full MS scan was also adjusted to *m/z* 360 to exclude several
239 abundant interference peaks (eg. Streptavidin peptide with *m/z* 354.7068). An alternative to on-beads
240 digestion is to use biotin antibody to enrich biotinylated proteins, which can preserve the information about
241 biotinylation sites, but has non-specific binding issues from antibodies.[31,32] Excess biotin in harsh detergent
242 buffer with heating can also be used to elute biotinylated proteins from streptavidin beads, but detergent
243 must be removed before LC-MS analysis.[33]

244 Although protein concentration was initially determined with a colorimetric protein assay, the
245 concentration of enriched proteins by streptavidin beads is unknown. To achieve the optimal on-beads
246 digestion efficiency and minimum interferences from trypsin, we optimized the amount of protease
247 concentration for on-beads digestion. Compared to trypsin alone, Trypsin/Lys-C mix resulted in 15% more
248 protein and peptide IDs with much less miscleavages (data not shown), in agreement with previous
249 findings.[34] With an increased amount of Trypsin/Lys-C mix, The number of identified proteins, peptides,
250 and PSMs (peptide spectral matches) increased and peptide miscleavages decreased until reaching a plateau
251 at ~ 1-1.5 μg of enzyme for 1mg of input protein from KuB-LAMP1-APEX2 iPSC-neurons (**Figure 3B**).
252 Most identified peptides had +2 or +3 charges, and increased amount of protease shifted the peptides
253 towards lower charges and smaller precursor masses (**Figure S2**).

254 ***Controlling the false discoveries in proximity labeling proteomics***

255 An inherent challenge in identifying protein interactions is the false discoveries, both false positives
256 and false negatives.[35] For PL-proteomics, false positive interactions include proteins that bind non-
257 specifically to beads, endogenously biotinylated proteins, and other highly abundant bystanders. False
258 negative interactions include low abundance and transient/weak protein interactions which can be ruled out
259 if using overly strict cut-off thresholds. Here, we demonstrated that rational experimental design with data
260 normalization and proper selection of controls can reduce these false positive and false negative
261 discoveries.[30,36]

262 The presence of highly abundant endogenously biotinylated proteins has been a major concern in
263 PL-proteomics.[20] Within the mitochondria, pyruvate carboxylase (PC), 3-methylcrontonyl-coA carboxylase
264 (MCC), propionyl-CoA carboxylase (PCC), and acetyl-CoA carboxylase 1 (ACACA) are known to be
265 endogenously biotinylated.[37] Upon the addition of the biotin substrate, these endogenously biotinylated
266 proteins as well as their interacting proteins in the mitochondria are enriched by streptavidin beads,
267 accounting for a major source of false positives in PL experiments. In our dataset, 6 endogenously
268 biotinylated proteins were identified, PCCA, PCCB, PC, MCCC1, MCCC2, and ACACA (heatmap
269 intensities shown in **Figure 4A**). These abundant proteins cannot be removed from the samples, but we
270 found that normalizing the proteomic dataset to the most abundant endogenously biotinylated protein
271 (PCCA) can greatly reduce variations for both technical and biological replicates. As shown in the scatter
272 plots of two biological replicates (**Figure 4B**), excellent reproducibility was achieved with both slope and
273 R^2 close to 1 after normalization to PCCA.

274 For the study of the lysosomal microenvironment in neurons, proper controls are important to
275 minimize false discoveries. We established two control lines for our LAMP1-APEX2 neurons: cytosol-
276 localized NES-APEX2 line, and negative control neurons that has no APEX2 expression. As illustrated in
277 **Figure 5**, all quantified proteins are ranked based on their abundance ratio between LAMP1-APEX2 vs.
278 control neurons. Using a known set of lysosomal transmembrane proteins (ratios should be positive) and
279 endogenously biotinylated proteins (ratios should be ~1), we can evaluate the false positives and false

280 negatives in these comparison groups: For KI-LAMP1-APEX2 neurons, the parental line lacking APEX2
281 is a better control (**Figure 5A**). Using NES-APEX2 control for KI neurons caused massive false negatives,
282 which could be due to the disparity in APEX expression levels (**Figure 5B)**. Whereas, for high expression
283 KuD-LAMP1-APEX or KuB neurons, the line lacking APEX2 created substantial false positives (**Figure**
284 **5C**). Thus, NES-APEX2 line is a better control for the overexpressed LAMP1-APEX2 constructs (**Figure**
285 **5D**). Therefore, designing the appropriate control group for each PL study is critical in controlling false
286 discoveries.

287 APEX2 activation requires brief H_2O_2 (1 mM for 1 min at 37 °C) and biotin-phenol (500 μM for
288 30 min at 37 °C) treatment for cells, which raised the concerns in the field for possible oxidative stress and
289 perturbation of cellular processes. To address this question directly, we conducted whole cell lysate
290 proteomics from parental iPSC-neurons under different treatments (H_2O_2 treated vs. biotin-phenol treated
291 vs. control). Label-free proteomics quantification (N=3) identified 4339 proteins from neurons, but none of
292 them reached statistical significance (**Figure 6A**). Interestingly, glutathione peroxidase 4 (GPX4) and
293 mitogen-activated protein kinase 3 (MAPK3), which function to protect against oxidative stress, were
294 slightly upregulated in H_2O_2 vs. control group without reaching statistical significance.[38,39] To evaluate
295 possible PTMs caused by H_2O_2 and biotin-phenol, we conducted a semi-open PTM search for the
296 proteomics dataset with more than 300 PTMs from PEAKS software and Unimod database.[28] As expected,
297 methionine was found to be the most oxidized amino acid residue in peptides/proteins (**Figure 6B**). No
298 obvious increase of PTMs were observed in H_2O_2 and biotin-phenol groups vs. control, except that H_2O_2
299 treated group did present more oxidized peptide numbers compared to controls. Therefore, we proved that
300 the activation steps for APEX2 should not influence cellular processes other than enriching the
301 endogenously biotinylated proteins in mitochondria and slightly higher number of oxidized peptides.
302 However, a strict control of the concentration and time for H_2O_2 treatment is critical to minimize potential
303 oxidative stress.

304 ***Comparison of Different Lysosome Proximity Labeling Probes***

305 PL probes have been developed for various cell biology applications, but their proteomics
306 performance has rarely been compared. Here, we comparatively evaluated our three PL methods, KI-
307 LAMP1-APEX (endogenous expression), KuB-LAMP1-APEX (overexpression), and KuD-LAMP1-
308 APEX (modest overexpression). The expression levels of LAMP1-APEX2 in iPSC-neurons are KuB>
309 KuD> KI. The proteomic profiles are compared in **Figure 7A**. Overexpressing probes presented higher
310 numbers of identified proteins compared to the endogenous KI methods. The KuB method had slightly
311 more IDs compared to KuD. These results were validated by the beads titration assay, where the optimal
312 amount of beads per μg of input proteins was determined to be 0.2, 0.25, and 0.25 for KI, KuB, and KuD
313 neurons, respectively (**Figure 2 and S3**). Because the KI probe is expressed at physiological level, the
314 proteomics result is less prone to false positives and mislocalization of LAMP1. However, more starting
315 material is needed for the KI method compared to KuB/KuD. Therefore, scalability of the cell lines and
316 overexpression artifacts must be considered when determining the appropriate probe for PL experiments.
317 The complete protein list and comparisons from APEX probes are provided in **Table S1** and **Figure S4**.

318 After comparing to proper controls (KI vs. No-APEX2 control; KuB/KuD vs. NES-APEX2
319 control), all three of our LAMP1-APEX2 methods successfully identified and quantified specific
320 endolysosomal membrane and membrane interacting proteins with complementary coverage (**Figure 7A)**.
321 Protein network analysis was conducted for the overlapping proteins from at least two LAMP1-APEX2
322 methods, shown in **Figure 7B**. These proteins are lysosomal membrane proteins (LAMP1, LAMP2,
323 LAMTORs, PIP4P1, PIP4P2 et al.) and proteins related to lysosomal acidification, transport, degradation,
324 mobility and regulatory pathways. For instance, the identified Rab GTPases (Rab7 and Rab39) and SNARE
325 protein family (STX7, STX12, VTI1B) play key roles in autophagy, endolysosomal trafficking, and
326 lysosome biogensis.[40,41] The homotypic fusion and vacuole protein sorting (HOPS) complex is composed
327 of six vacuolar protein sorting (Vps) subunits that act as a tether between the autophagosome and lysosome
328 for autophagy and interact with the Rab GTPases.[40,41] Additionally, multiple components of the vacuolar-
329 type ATPase (V-ATPase), which maintains the acidic environment of lysosomes, were reproducibly

330 identified.[42] Besides controlling the lysosomal pH, the V-ATPases are also important for nutrient sensing
331 and amino acid efflux through the mTORC1 pathway.[42] Key proteins in the mTORC1 pathway
332 (LAMTOR1,2,3, RRAGC, RRAGB, RPTOR, MTOR) were also identified across our LAMP1-APEX2
333 proteomics experiments.

334

335

336 **CONCLUSIONS**

337 We developed a new proximity labeling proteomic method, KI-LAMP1-APEX2, to study the
338 endolysosomal membrane and membrane interacting proteins in human iPSC-derived neurons. Excellent
339 reproducibility, specificity, and accuracy were achieved after systematic investigation of key experimental
340 parameters in proximity labeling proteomics. We recommend conducting beads titration assay to achieve
341 the optimal protein/beads ratio, normalizing proteomics data to the endogenously biotinylated proteins, and
342 selecting proper controls for all PL experiment to reduce variation and false positive/negative interactions.
343 These findings not only provided helpful guides to design proximity labeling experiments, but also can be
344 applied to general affinity purification experiments. Additionally, hydrogen peroxide treatment during
345 APEX2 activation did not cause significant proteome perturbation but the time and concentration of H_2O_2
346 treatment need to be strictly controlled, addressing a major concern in the field. Furthermore, comparative
347 evaluations of different LAMP1-APEX2 probes (KI, KuD, and KuB) revealed complementary proteome
348 coverage and provided key references for the rational design of proximity labeling experiments. To
349 summarize, our study established lysosomal proximity labeling methods that can serve as promising tools
350 to study lysosomal functions and filled critical gaps in the field of proximity labeling and affinity
351 purification mass spectrometry.

352

353 **SUPPORTING INFORMATION**

Supplemental methods for transgenic LAMP1-APEX iPSC lines; Fluorescence imaging of APEX2 activity in overexpressing KuD-LAMP1-APEX2 and cytosolic NES-APEX2 iPSC-derived neurons; Distribution of peptide charges and precursor masses with different amount of proteases for on-beads digestion; Beads titration assay for overexpressed APEX probes; Venn diagram of all identified proteins from all APEX proteomics datasets; Table of identified proteins from all APEX-proteomics datasets.

ACKNOWLEDGEMENTS

The authors also want to thank Maia Parsadanian for technical assistance in cell culture. This project is financially supported by the GW Faculty Startup Fund. A.M.F acknowledges the Bourdon F. Scribner Fund Fellowship from GW Chemistry Department. M.S.F. was supported by the National Institutes of Health (F30AG060722). Raw proteomics data deposited to the PRIDE database will be available to the public upon acceptance.

The authors declare no competing financial interest.

REFERENCES

(1) Vidal, M.; Cusick, M. E.; Barabási, A.-L. Interactome Networks and Human Disease. *Cell* **2011**, *144* (6), 986–998. https://doi.org/10.1016/j.cell.2011.02.016.

(2) Huttlin, E. L.; Bruckner, R. J.; Paulo, J. A.; Cannon, J. R.; Ting, L.; Baltier, K.; Colby, G.; Gebreab, F.; Gygi, M. P.; Parzen, H.; Szpyt, J.; Tam, S.; Zarraga, G.; Pontano-Vaites, L.; Swarup, S.; White, A. E.; Schweppe, D. K.; Rad, R.; Erickson, B. K.; Obar, R. A.; Guruharsha, K. G.; Li, K.; Artavanis-Tsakonas, S.; Gygi, S. P.; Wade Harper, J. Architecture of the Human Interactome Defines Protein Communities and Disease Networks. *Nature* **2017**. https://doi.org/10.1038/nature22366.

(3) Huttlin, E. L.; Ting, L.; Bruckner, R. J.; Gebreab, F.; Gygi, M. P.; Szpyt, J.; Tam, S.; Zarraga, G.; Colby, G.; Baltier, K.; Dong, R.; Guarani, V.; Vaites, L. P.; Ordureau, A.; Rad, R.; Erickson, B. K.; Wühr, M.; Chick, J.; Zhai, B.; Kolippakkam, D.; Mintseris, J.; Obar, R. A.; Harris, T.; Artavanis-Tsakonas, S.; Sowa, M. E.; De Camilli, P.; Paulo, J. A.; Harper, J. W.; Gygi, S. P. The BioPlex Network: A Systematic Exploration of the Human Interactome. *Cell* **2015**. https://doi.org/10.1016/j.cell.2015.06.043.

(4) Wang, X.; Huang, L. Identifying Dynamic Interactors of Protein Complexes by Quantitative Mass Spectrometry. *Mol. Cell. Proteomics* **2008**. https://doi.org/10.1074/mcp.M700261-MCP200.

386 (5) Kaake, R. M.; Wang, X.; Huang, L. Profiling of Protein Interaction Networks of Protein
387 Complexes Using Affinity Purification and Quantitative Mass Spectrometry. *Molecular and*
388 *Cellular Proteomics*. 2010. https://doi.org/10.1074/mcp.R110.000265.

389 (6) Kim, D. I.; Roux, K. J. Filling the Void: Proximity-Based Labeling of Proteins in Living Cells.
390 *Trends Cell Biol.* **2016**, *26* (11), 804–817. https://doi.org/10.1016/j.tcb.2016.09.004.

391 (7) Samavarchi-Tehrani, P.; Samson, R.; Gingras, A. C. Proximity Dependent Biotinylation: Key
392 Enzymes and Adaptation to Proteomics Approaches. *Mol. Cell. Proteomics* **2020**, *19* (5), 757–773.
393 https://doi.org/10.1074/mcp.R120.001941.

394 (8) Roux, K. J.; Kim, D. I.; Raida, M.; Burke, B. A Promiscuous Biotin Ligase Fusion Protein
395 Identifies Proximal and Interacting Proteins in Mammalian Cells. *J. Cell Biol.* **2012**, *196* (6), 801–
396 810. https://doi.org/10.1083/jcb.201112098.

397 (9) Martell, J. D.; Deerinck, T. J.; Sancak, Y.; Poulos, T. L.; Mootha, V. K.; Sosinsky, G. E.;
398 Ellisman, M. H.; Ting, A. Y. Engineered Ascorbate Peroxidase as a Genetically Encoded Reporter
399 for Electron Microscopy. *Nat Biotechnol* **2012**, *30* (11), 1143–1148.
400 https://doi.org/10.1038/nbt.2375.

401 (10) John E. Cronan, J. Biotination of Proteins in Vivo: A Post-Translational Modification to Label,
402 Purify, and Study Proteins. *J. Biol. Chem.* **1190**, *265* (18), 10327–10333.

403 (11) Béganton, B.; Solassol, I.; Mangé, A.; Solassol, J. Protein Interactions Study through Proximity-
404 Labeling. *Expert Rev. Proteomics* **2019**, *16* (8), 717–726.
405 https://doi.org/10.1080/14789450.2019.1638769.

406 (12) Branon, T. C.; Bosch, J. A.; Sanchez, A. D.; Udeshi, N. D.; Svinkina, T.; Carr, S. A.; Feldman, J.
407 L.; Perrimon, N.; Ting, A. Y.; Author, N. B. Efficient Proximity Labeling in Living Cells and
408 Organisms with TurboID HHS Public Access Author Manuscript. *Nat Biotechnol* **2018**, *36* (9),
409 880–887. https://doi.org/10.1038/nbt.4201.

410 (13) Rhee, H. W.; Zou, P.; Udeshi, N. D.; Martell, J. D.; Mootha, V. K.; Carr, S. A.; Ting, A. Y.
411 Proteomic Mapping of Mitochondria in Living Cells via Spatially Restricted Enzymatic Tagging.
412 *Science (80-.).* **2013**, *339* (6125), 1328–1331. https://doi.org/10.1126/science.1230593.

413 (14) Lam, S. S.; Martell, J. D.; Kamer, K. J.; Deerinck, T. J.; Ellisman, M. H.; Mootha, V. K.; Ting, A.
414 Y. Directed Evolution of APEX2 for Electron Microscopy and Proximity Labeling. *Nat. Methods*
415 **2014**. https://doi.org/10.1038/nmeth.3179.

416 (15) Lobingier, B. T.; Hüttenhain, R.; Eichel, K.; Miller, K. B.; Ting, A. Y.; von Zastrow, M.; Krogan,
417 N. J. An Approach to Spatiotemporally Resolve Protein Interaction Networks in Living Cells. *Cell*
418 **2017**, *169* (2), 350-360.e12. https://doi.org/10.1016/j.cell.2017.03.022.

419 (16) Paek, J.; Kalocsay, M.; Staus, D. P.; Wingler, L.; Pascolutti, R.; Paulo, J. A.; Gygi, S. P.; Kruse,
420 A. C. Multidimensional Tracking of GPCR Signaling via Peroxidase-Catalyzed Proximity
421 Labeling. *Cell* **2017**, *169* (2), 338-349.e11. https://doi.org/10.1016/j.cell.2017.03.028.

422 (17) Fazal, F. M.; Han, S.; Parker, K. R.; Kaewsapsak, P.; Xu, J.; Boettiger, A. N.; Chang, H. Y.; Ting,
423 A. Y. Atlas of Subcellular RNA Localization Revealed by APEX-Seq. *Cell* **2019**, *178* (2), 473-
424 490.e26. https://doi.org/10.1016/j.cell.2019.05.027.

425 (18) Tytgat, H. L. P.; Schoofs, G.; Driesen, M.; Proost, P.; Van Damme, E. J. M.; Vanderleyden, J.;
426 Lebeer, S. Endogenous Biotin-Binding Proteins: An Overlooked Factor Causing False Positives in
427 Streptavidin-Based Protein Detection. *Microb. Biotechnol.* **2015**, *8* (1), 164–168.
428 https://doi.org/10.1111/1751-7915.12150.

429 (19) Verrastro, I.; Tveen-Jensen, K.; Woscholski, R.; Spickett, C. M.; Pitt, A. R. Reversible Oxidation
430 of Phosphatase and Tensin Homolog (PTEN) Alters Its Interactions with Signaling and Regulatory
431 Proteins. *Free Radic. Biol. Med.* **2016**, *90*, 24–34.
432 https://doi.org/10.1016/j.freeradbiomed.2015.11.004.

433 (20) Han, S.; Li, J.; Ting, A. Y. Proximity Labeling: Spatially Resolved Proteomic Mapping for
434 Neurobiology. *Curr Opin Neurobiol* **2018**, *50*, 17–23. https://doi.org/10.1016/j.conb.2017.10.015.

435 (21) Lawrence, R. E.; Zoncu, R. The Lysosome as a Cellular Centre for Signalling, Metabolism and
436 Quality Control. *Nat. Cell Biol.* **2019**, *21* (2), 133–142. https://doi.org/10.1038/s41556-018-0244-
437 7.

438 (22) Ferguson, S. M. Neuronal Lysosomes. *Neurosci. Lett.* **2019**, *697* (March 2018), 1–9.
439 https://doi.org/10.1016/j.neulet.2018.04.005.

440 (23) Li, J.; Pfeffer, S. R. Lysosomal Membrane Glycoproteins Bind Cholesterol and Contribute to
441 Lysosomal Cholesterol Export. *Elife* **2016**. https://doi.org/10.7554/eLife.21635.

442 (24) Fernandopulle, M. S.; Prestil, R.; Grunseich, C.; Wang, C.; Gan, L.; Ward, M. E. Transcription
443 Factor–Mediated Differentiation of Human IPSCs into Neurons. *Curr. Protoc. Cell Biol.* **2018**, *79*
444 (1), 1–48. https://doi.org/10.1002/cpcb.51.

445 (25) Liao, Y. C.; Fernandopulle, M. S.; Wang, G.; Choi, H.; Hao, L.; Drerup, C. M.; Patel, R.; Qamar,
446 S.; Nixon-Abell, J.; Shen, Y.; Meadows, W.; Vendruscolo, M.; Knowles, T. P. J.; Nelson, M.;
447 Czekalska, M. A.; Musteikyte, G.; Gachechiladze, M. A.; Stephens, C. A.; Pasolli, H. A.; Forrest,
448 L. R.; St George-Hyslop, P.; Lippincott-Schwartz, J.; Ward, M. E. RNA Granules Hitchhike on
449 Lysosomes for Long-Distance Transport, Using Annexin A11 as a Molecular Tether. *Cell* **2019**,
450 *179* (1), 147-164.e20. https://doi.org/10.1016/j.cell.2019.08.050.

451 (26) Wang, C.; Ward, M. E.; Chen, R.; Liu, K.; Tracy, T. E.; Chen, X.; Xie, M.; Sohn, P. D.; Ludwig,
452 C.; Meyer-Franke, A.; Karch, C. M.; Ding, S.; Gan, L. Scalable Production of IPSC-Derived
453 Human Neurons to Identify Tau-Lowering Compounds by High-Content Screening. *Stem Cell*
454 *Reports* **2017**, *9* (4), 1221–1233. https://doi.org/10.1016/j.stemcr.2017.08.019.

455 (27) Tyanova, S.; Temu, T.; Cox, J. The MaxQuant Computational Platform for Mass Spectrometry-
456 Based Shotgun Proteomics. *Nat. Protoc.* **2016**, *11* (12), 2301–2319.
457 https://doi.org/10.1038/nprot.2016.136.

458 (28) Creasy, D.; Cottrell, J. UniMod: Protein Modifications for Mass Spectrometry. *Proteomics* **2004**,
459 *4*, 1534–1536. https://doi.org/10.1002/pmic.200300744.

460 (29) Cheng, X. T.; Xie, Y. X.; Zhou, B.; Huang, N.; Farfel-Becker, T.; Sheng, Z. H. Characterization of
461 LAMP1-Labeled Nondegradative Lysosomal and Endocytic Compartments in Neurons. *J. Cell*
462 *Biol.* **2018**, *217* (9), 3127–3139. https://doi.org/10.1083/jcb.201711083.

463 (30) Hecht, E. S.; Oberg, A. L.; Muddiman, D. C. Optimizing Mass Spectrometry Analyses: A Tailored
464 Review on the Utility of Design of Experiments. *Journal of the American Society for Mass*
465 *Spectrometry*. 2016. https://doi.org/10.1007/s13361-016-1344-x.

466 (31) Kim, D. I.; Cutler, J. A.; Na, C. H.; Reckel, S.; Renuse, S.; Madugundu, A. K.; Tahir, R.;
467 Goldschmidt, H. L.; Reddy, K. L.; Huganir, R. L.; Wu, X.; Zachara, N. E.; Hantschel, O.; Pandey,
468 A. BioSITe: A Method for Direct Detection and Quantitation of Site-Specific Biotinylation. *J.*
469 *Proteome Res.* **2018**. https://doi.org/10.1021/acs.jproteome.7b00775.

470 (32) Udeshi, N. D.; Pedram, K.; Svinkina, T.; Fereshetian, S.; Myers, S. A.; Aygun, O.; Krug, K.;
471 Clauser, K.; Ryan, D.; Ast, T.; Mootha, V. K.; Ting, A. Y.; Carr, S. A. Antibodies to Biotin Enable

472 Large-Scale Detection of Biotinylation Sites on Proteins. *Nat. Methods* **2017**, *14* (12), 1167–1170.
473 https://doi.org/10.1038/nmeth.4465.

474 (33) Cheah, J. S.; Yamada, S. A Simple Elution Strategy for Biotinylated Proteins Bound to
475 Streptavidin Conjugated Beads Using Excess Biotin and Heat. *Biochem. Biophys. Res. Commun.*
476 **2017**, *493* (4), 1522–1527. https://doi.org/10.1016/j.bbrc.2017.09.168.

477 (34) Glatter, T.; Ludwig, C.; Ahrné, E.; Aebersold, R.; Heck, A. J. R.; Schmidt, A. Large-Scale
478 Quantitative Assessment of Different in-Solution Protein Digestion Protocols Reveals Superior
479 Cleavage Efficiency of Tandem Lys-C/Trypsin Proteolysis over Trypsin Digestion. *J. Proteome*
480 *Res.* **2012**, *11* (11), 5145–5156. https://doi.org/10.1021/pr300273g.

481 (35) Walzthoeni, T.; Claassen, M.; Leitner, A.; Herzog, F.; Bohn, S.; Förster, F.; Beck, M.; Aebersold,
482 R. False Discovery Rate Estimation for Cross-Linked Peptides Identified by Mass Spectrometry.
483 *Nat. Methods* **2012**. https://doi.org/10.1038/nmeth.2103.

484 (36) Gokce, E.; Shuford, C. M.; Franck, W. L.; Dean, R. A.; Muddiman, D. C. Evaluation of
485 Normalization Methods on GeLC-MS/MS Label-Free Spectral Counting Data to Correct for
486 Variation during Proteomic Workflows. *J. Am. Soc. Mass Spectrom.* **2011**.
487 https://doi.org/10.1007/s13361-011-0237-2.

488 (37) Tong, L. Structure and Function of Biotin-Dependent Carboxylases. *Cell. Mol. Life Sci.* **2013**, *70*
489 (5), 863–891. https://doi.org/10.1007/s00018-012-1096-0.

490 (38) Franco, R.; Cidlowski, J. A. Glutathione Efflux and Cell Death. *Antioxidants Redox Signal.* **2012**,
491 *17* (12), 1694–1713. https://doi.org/10.1089/ars.2012.4553.

492 (39) Barata, A. G.; Dick, T. P. A Role for Peroxiredoxins in H2O2- and MEKK-Dependent Activation
493 of the P38 Signaling Pathway. *Redox Biol.* **2020**, *28* (October 2019), 101340.
494 https://doi.org/10.1016/j.redox.2019.101340.

495 (40) Balderhaar, H. J. klein.; Ungermann, C. CORVET and HOPS Tethering Complexes - Coordinators
496 of Endosome and Lysosome Fusion. *J. Cell Sci.* **2013**, *126* (6), 1307–1316.
497 https://doi.org/10.1242/jcs.107805.

498 (41) Zhang, M.; Chen, L.; Wang, S.; Wang, T. Rab7: Roles in Membrane Trafficking and Disease.
499 *Biosci. Rep.* **2009**, *29* (3), 193–209. https://doi.org/10.1042/BSR20090032.

500 (42) Kissing, S.; Rudnik, S.; Damme, M.; Lüllmann-Rauch, R.; Ichihara, A.; Kornak, U.; Eskelinen, E.
501 L.; Jabs, S.; Heeren, J.; De Brabander, J. K.; Haas, A.; Saftig, P. Disruption of the Vacuolar-Type
502 H+-ATPase Complex in Liver Causes MTORC1-Independent Accumulation of Autophagic
503 Vacuoles and Lysosomes. *Autophagy* **2017**, *13* (4), 670–685.
504 https://doi.org/10.1080/15548627.2017.1280216.

505

506

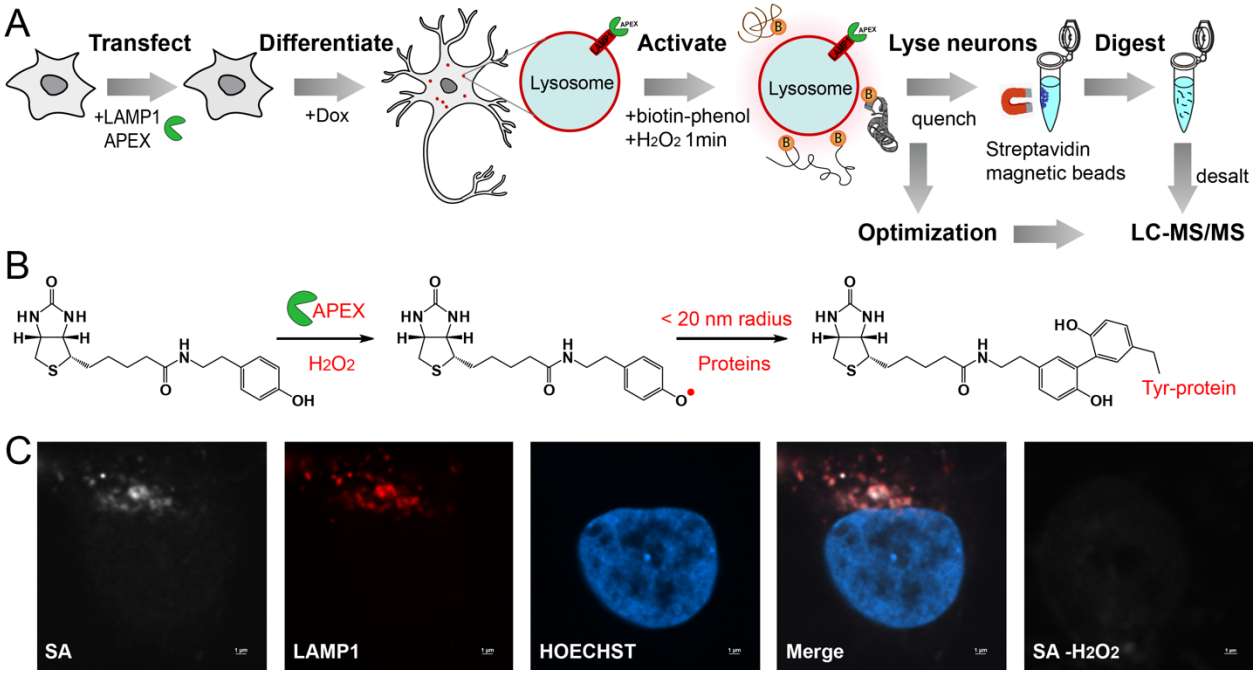
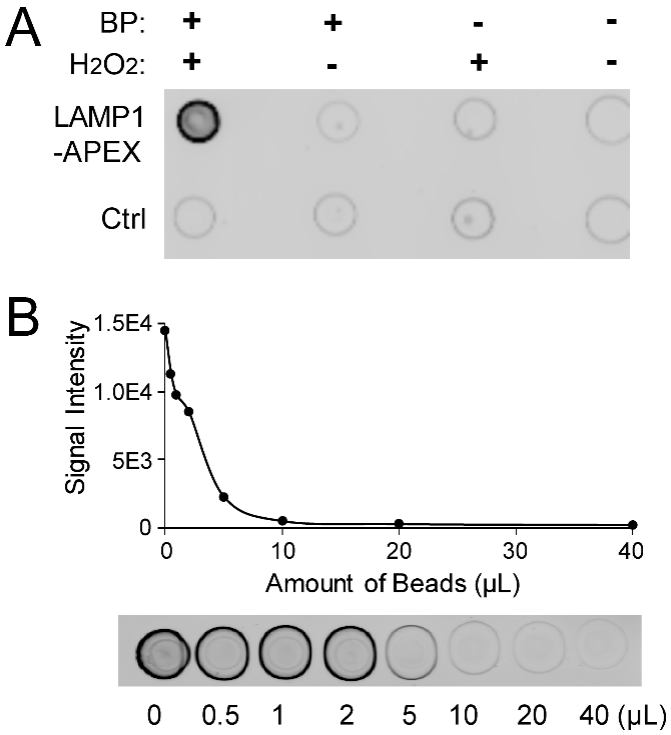


Figure 1. **Development of endogenous lysosomal proximity labeling proteomics.** (A) Illustration of LAMP1-APEX2 proteomic workflow in human iPSC-derived neurons. (B) Schematic of APEX enzymatic labeling reaction. (C) Fluorescence imaging of APEX2 activity in an iPSC-derived neuron with endogenously tagged KI-LAMP1-APEX2 probe. Biotinylation is visualized by staining against streptavidin (SA) Fluor 680 (far red). LAMP1 (red) is used as an endolysosome marker. Hoechst is a nuclear marker (blue). Control neurons without H_2O_2 treatment exhibit no SA signals.



516

517

518 **Figure 2. Validation and optimization of endogenous lysosomal proximity labeling proteomics (KI-**

519 **LAMP1-APEX) with streptavidin dot-blots.** (A) Streptavidin dot-blot validation of APEX2 activity. (B)

520 Determination of the optimal beads-to-protein ratio using dot blot intensities in the beads titration assay.

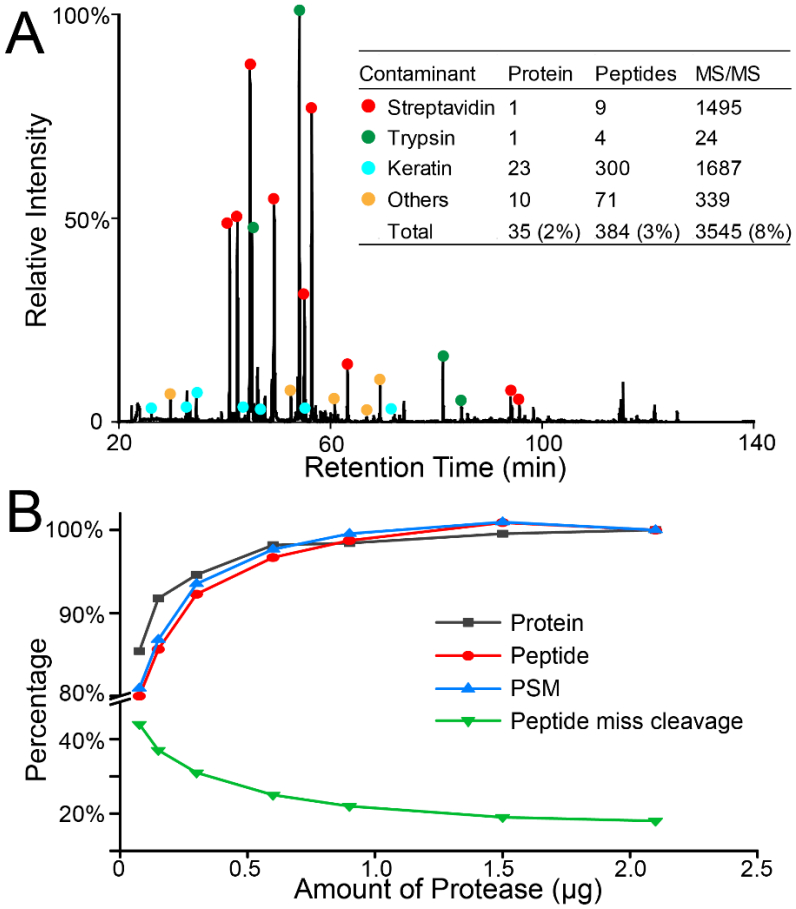
521 Increasing volume of bead slurry were incubated with 50 µg of input protein lysate in eight tubes overnight.

522 Supernatants were spotted on the nitrocellulose membrane and stained against streptavidin. Dot blot signal

523 intensities were measured and plotted against the amount of beads.

524

525



526

527

528 **Figure 3. Evaluation of interference signals and different amount of proteases for on-beads digestion.**

529 (A) Base peak chromatogram of digesting streptavidin magnetic beads with no input proteins. Contaminant

530 proteins/peptides in the table was used to generate an exclusion list for LC-MS/MS experiment. The

531 percentages were generated in comparison with an actual APEX proteomics experiment. (B) The

532 normalized number of identifications and missed cleavages with different amount of Trypsin/LysC mix for

533 on-beads digestion with 1 mg of input protein samples.

534

535

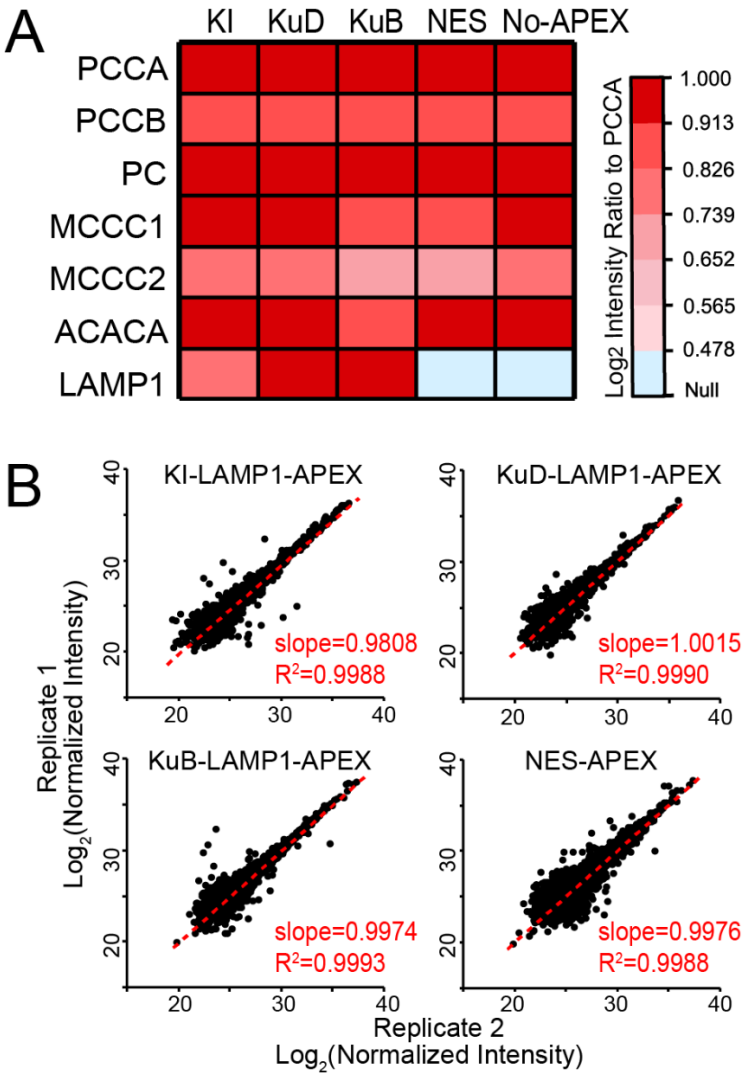


Figure 4. **Normalization of APEX2 proteomics data to endogenously biotinylated protein.** (A) Heatmap intensity of six endogenously biotinylated proteins and LAMP1 bait proteins across different APEX2 probes. LAMP1 is not detected in cytosolic NES-APEX2 and parental iPSC-derived neurons without APEX. (B) Scatter plots between biological replicates with excellent reproducibility for different APEX2 proteomic datasets in iPSC-derived neurons.

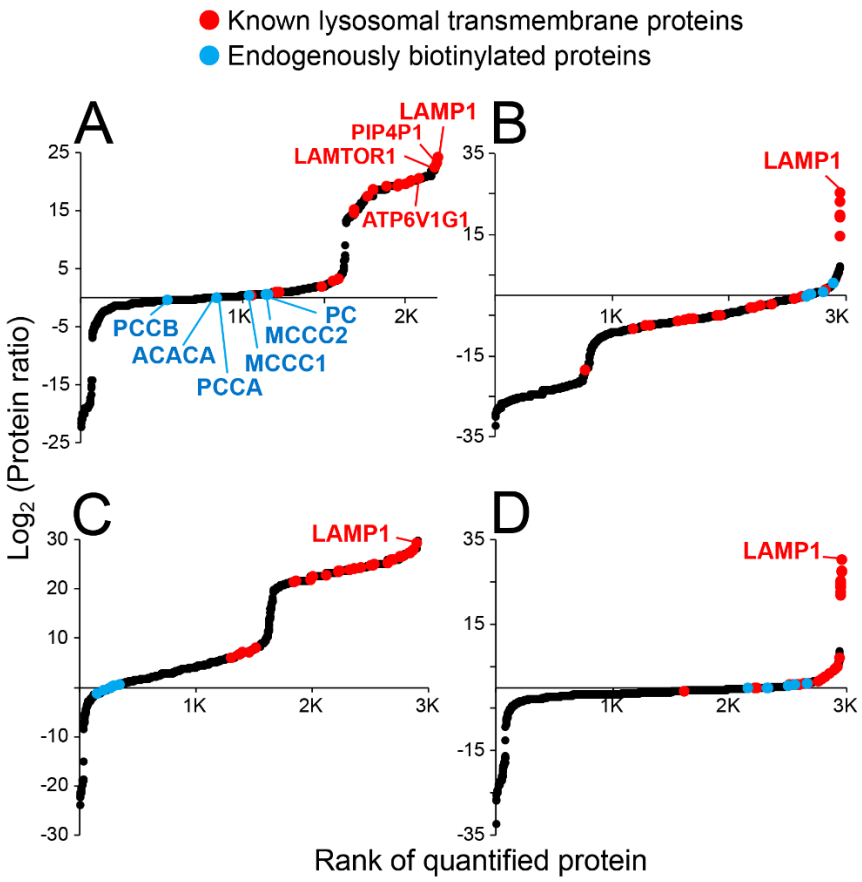


Figure 5. Evaluation of false discoveries in LAMP1-APEX proteomics using different control dataset. Scatter plots of the rank of protein ratios in (A) KI-LAMP1-APEX2 Proteomics with No-APEX2 parental line as control; (B) KI-LAMP1-APEX2 with NES-APEX2 as control; (C) KuD-LAMP1-APEX2 with No-APEX as control; (D) KuD-LAMP1-APEX2 with NES-APEX2 as control. Protein intensities were normalized to endogenously biotinylated protein PCCA.

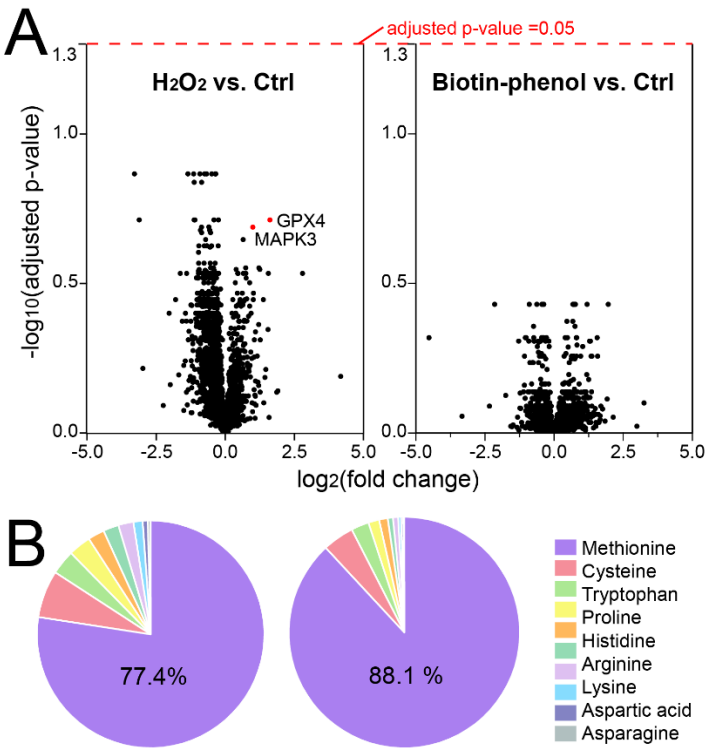


Figure 6. Evaluation of potential influence of H_2O_2 and PB treatment to the whole cell proteome. (A) Volcano plots of all the quantified proteins from H_2O_2 vs. Ctrl and PB vs. Ctrl comparison group in the whole cell protein extract. (B) Pie graph of oxidized amino acids from the identified proteins in iPSC-derived neurons.

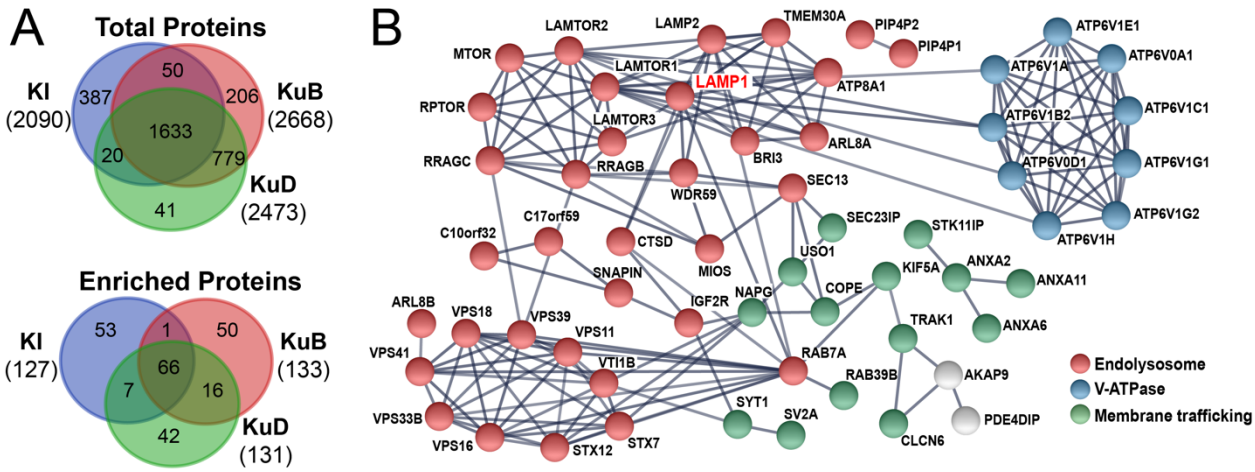


Figure 7. Comparative evaluation of different lysosome proximity labeling probes. (A) Venn diagram of all identified proteins and truly enriched proteins after ratiometric analysis vs. controls from KI, KuB, and KuD LAMP1-APEX iPSC-neurons. (B) STRING network of overlapped proteins from three LAMP1-APEX2 probes.