

Dissertation zur Erlangung des Doktorgrades  
der Fakultät für Chemie und Pharmazie  
der Ludwig-Maximilians-Universität München

# **Chemical proteomics profiling of AMPylation during neuronal differentiation**

Tobias Becker

aus

Traben-Trarbach, Deutschland

**2023**

## **Erklärung**

Diese Dissertation wurde im Sinne von § 7 der Promotionsordnung vom 28. November 2011 von *Herrn Dr. Pavel Kielkowski* betreut.

## **Eidesstattliche Versicherung**

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, 14.02.23

Tobias Becker

Dissertation eingereicht am: 16.01.2023

**1. Gutachter:** Dr. Pavel Kielkowski

**2. Gutachter:** Prof. Dr. Thomas Carell

Mündliche Prüfung am: 13.02.2023



*Ruhm liegt nicht darin, niemals zu fallen, sondern jedes Mal wieder aufzustehen, wenn wir gescheitert sind.*

Konfuzius

## Publication list

**Parts of this work have already been published or presented at conferences.**

Becker, T., Wiest, A., Telek, A., Bejko, D., Hoffmann-Röder, A., and Kielkowski, P. (2022). Transforming Chemical Proteomics Enrichment into a High-Throughput Method Using an SP2E Workflow. *JACS Au* 2, 1712–1723. 10.1021/jacsau.2c00284.

Makarov, D., Telek, A., Becker, T., von Wrisberg, M.K., Schneider, S., and Kielkowski, P. (2022). Clickable report tags for identification of modified peptides by mass spectrometry. *Journal of Mass Spectrometry* 57. 10.1002/jms.4812.

Becker, T., Cappel, C., di Matteo, F., Sonsalla, G., Kaminska, E., Spada, F., Cappello, S., Damme, M., and Kielkowski, P. (2021). AMPylation profiling during neuronal differentiation reveals extensive variation on lysosomal proteins. *iScience* 24. 10.1016/j.isci.2021.103521.

## Conference contributions

**Poster** “High-throughput chemical proteomics method to study protein post-translational modifications.” Next-Generation Protein Analysis and Detection (4th edition), Ghent, Belgium, 2022.

**Presentation** "AMPylation profiling during neuronal differentiation", FCI Fellows Meeting, Munich, Germany, 2022.

## Further publications

Kielkowski, P., Buchsbaum, I.Y., Becker, T., Bach, K., Cappello, S., and Sieber, S.A. (2020). A Pronucleotide Probe for Live-Cell Imaging of Protein AMPylation. *ChemBioChem* 21, 1285–1287. 10.1002/cbic.201900716.

Becker, T., and Kielkowski, P. (2020). Protein-AMPylierungs-Identifikation in lebenden Zellen. *BIOspektrum* 26, 743–746. 10.1007/s12268-020-1491-2.

## **Danksagung**

An erster Stelle möchte ich mich bei meinem Doktorvater *Dr. Pavel Kielkowski* bedanken. Danke Pavel für die Möglichkeit mit dir zusammen eine Forschungsgruppe an der LMU zu starten. Es war eine tolle Erfahrung, ein Labor mit dir von Grund auf neu aufzubauen. Danke dir für dein verständnisvolle Art und deine stets offenen und motivierenden Worte. Danke dir auch für dein Vertrauen in meine Ideen und dass ich stets die Möglichkeit hatte diese umzusetzen. Es war eine tolle und lehrreiche Zeit für mich.

Bei Professor *Dr. Thomas Carell* möchte ich mich herzlich für die Übernahme der Zweitkorrektur dieser Arbeit bedanken.

Ich danke dem Fonds der chemischen Industrie für die finanzielle und ideelle Unterstützung während meiner Promotion. Außerdem danke ich den anderen Stipendiaten des FCIs im Raum München für den interessanten Austausch während unserer Treffen.

*Dmytro Makarov* dir möchte ich für deine offenen und hilfreichen Ratschläge danken. Ich schätze dich als Person wirklich sehr und habe mich gefreut, dass du ein Teil dieses Labors bist. Deine Tischtenniskünste sind natürlich auch nicht zu verachten, ich freue mich schon auf eine neue Runde.

*Haoyu Chen* dir danke ich für die interessanten Einblicke in eine andere Kultur. Es hat Spaß gemacht Neues von dir zu lernen und unseren Oktoberfestbesuch werde ich auch in guter Erinnerung behalten.

*Andreas Wiest* ich habe mich sehr gefreut, dass du nach deinem Forschungspraktikum und deiner Masterarbeit nun auch als Doktorand in unserer Gruppe bist. Unsere fachlichen Diskussionen habe ich immer sehr genossen, denn diese haben mir immer wieder neue Blickwinkel auf Probleme aufgezeigt. Auch deine humorvolle und hilfsbereite Art weiß ich sehr zu schätzen. Und zu allerletzt, danke dir auch für das Korrekturlesen dieser Arbeit.

Ein riesiges Dankeschön geht an meine Familie, ohne euch wäre das alles nicht möglich gewesen. Ihr wart in jeder Situation für mich da und dafür bin ich unfassbar dankbar. Ich weiß, dass ich mich in der Zukunft immer auf euch verlassen kann und das macht mich glücklich.

Zu allerletzt möchte ich mich bei meiner Verlobten *Verena Hänle* bedanken. Danke für deinen riesigen Rückhalt während der gesamten Promotion. Danke dafür, dass du in jedem Moment weißt was mir gut tut und mich zum Lachen bringst. Danke, dass du mit deiner Art mein Leben bereicherst. Ich freue mich auf unsere gemeinsame Zukunft.



# Table of contents

Summary .....	I
1 Introduction .....	1
1.1 Protein post-translational modifications.....	1
1.2 Protein AMPylation.....	2
1.2.1 General features of protein AMPylation .....	2
1.2.2 AMPylation in eukaryotes.....	3
1.2.3 AMPylation in neurodevelopment and neurodegeneration.....	5
1.2.4 Methods to analyze AMPylated proteins .....	5
1.3 iNGNs model of neuronal differentiation .....	11
1.4 Neurodegenerative disorders.....	12
1.4.1 Characteristics of PLD3 .....	13
1.5 Mass spectrometry.....	15
1.5.1 Sample preparation in chemical proteomics .....	15
1.5.2 Orbitrap-based mass spectrometer .....	16
1.5.3 Data acquisition.....	18
2 Aim of this thesis.....	21
3 Published work.....	22
3.1 AMPylation profiling during neuronal differentiation reveals extensive variation on lysosomal proteins.....	22
3.2 Transforming Chemical Proteomics Enrichment into a High-Throughput Method Using an SP2E Workflow .....	45
3.3 Clickable report tags for identification of modified peptides by mass spectrometry	59
4 Unpublished work .....	67
4.1 Further optimization of the SP2E workflow .....	67
4.1.1 Comparison of DDA and DIA acquisition.....	67
4.1.2 Reduction of carboxylate-coated paramagnetic beads amount .....	70
4.1.3 Materials and methods .....	72



4.2	AMPylation profiling in various cell lines using different inhibitors .....	73
4.2.1	AMPylation profiling in SH-SY5Y neuroblastoma cells using chloroquine and GW4869 .....	73
4.2.2	AMPylation profiling in cerebral organoids using bafilomycin, monensin and chloroquine.....	75
4.2.3	Materials and methods .....	78
5	Bibliography.....	79
6	List of abbreviations.....	92
7	List of figures .....	94
8	Appendix .....	95
8.1	Supplementary information: AMPylation profiling during neuronal differentiation reveals extensive variation on lysosomal proteins .....	95
8.2	Supplementary information: Transforming Chemical Proteomics Enrichment into a High-Throughput Method Using an SP2E Workflow.....	104
8.3	Supplementary information: Clickable report tags for identification of modified peptides by mass spectrometry.....	117

## Summary

Post-translational modifications (PTMs) are important regulators of the activity, localization and interaction of cellular proteins. Therefore, the dysregulation of a specific PTM can have dramatic consequences on cellular homeostasis and can lead to a particular disease. One PTM that was recently connected to neurodegenerative diseases and neuronal development is AMPylation, which is defined as the covalent attachment of adenosine 5'-*O*-monophosphate (AMP) to serine, threonine or tyrosine residues. The link between AMPylation and neurodegeneration is based on several studies, such as those by *Truttmann et al.*, who showed that the overexpression of the highly active AMPylator FICD (E274G) increased the aggregation of  $\beta$ -amyloid and  $\alpha$ -synuclein.<sup>1</sup> Furthermore, *Kielkowski et al.* have screened AMPylation with chemical proteomics in various cell lines, such as induced pluripotent stem cells (iPSCs), neural progenitor cells (NPCs), neurons and cerebral organoids (COs) using the cell-permeable pro-N6pA AMPylation probe. Thereby, they identified a new set of AMPylated proteins, including cathepsins, which are known to be associated with neurodegenerative disorders. In addition, they observed a change in the AMPylation pattern of cytoskeletal proteins during the differentiation of iPSCs into functional neurons.<sup>2,3</sup> To get deeper insights in the AMPylation dynamics during neuronal differentiation, the aim of this project was to analyze the AMPylation changes with a higher time resolution. Therefore, iNGNs were used as a cellular model system, since they differentiate to neurons within 4 days.<sup>4,5</sup> The chemical proteomics analysis of the different time points during the iNGN differentiation revealed that several lysosomal proteins such as PLD3, ACP2 and ABHD6 changed their AMPylation levels. Thereby, we provide evidence that AMPylation can be a lysosomal PTM in addition to its well-known localization in the endoplasmic reticulum (ER) and mitochondria. To validate the identified AMPylation targets in a probe-independent manner, a gel-based approach was developed to separate the AMPylated protein from its unmodified counterpart using a Phos-tag ligand. The Phos-tag gel allows to estimate the extent of modification and additionally shows which form of the protein is modified. For the validation, we focused on PLD3 as several studies have associated PLD3 with Alzheimer's disease.<sup>6-10</sup> However, the exact molecular function of PLD3 in Alzheimer's disease is still unclear. The Phos-tag gel of PLD3 revealed that the amount of AMPylated PLD3 increases during the differentiation of iNGNs and in addition that mainly the soluble form of PLD3 is AMPylated. With regards to this, it was observed that the higher the AMPylation status of PLD3, the lower its activity in iNGNs. Furthermore, it was confirmed that the soluble form of PLD3 is fully AMPylated in physiological young and mature neurons.

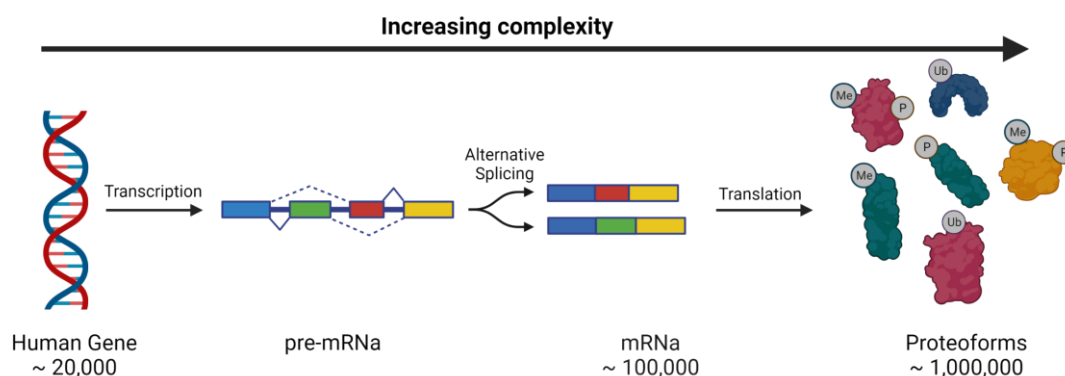
In the second part of this thesis, a systematic AMPylation screen was performed to elucidate metabolic pathways affected by AMPylation. Therefore, a chemical proteomics experiment was carried out in SH-SY5Y neuroblastoma cells using the pro-N6pA AMPylation probe and an inhibitor that blocks either cellular respiration, glycolysis, mTOR, autophagy, endolysosomal trafficking or protein secretion. By means of this, it was found that a certain group of proteins, including APP and NOTCH2, were only identified after treatment with bafilomycin, monensin or chloroquine. This indicates that AMPylation may have a function in endolysosomal trafficking and autophagy.

Taken together, we show that AMPylation is a PTM of luminal lysosomal proteins in neurons and that AMPylation may play a role in endolysosomal trafficking and autophagy.

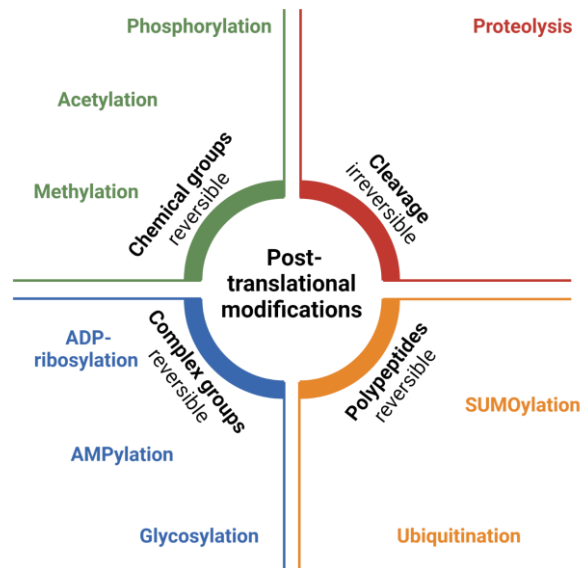
# 1 Introduction

## 1.1 Protein post-translational modifications

The human genome consists of roughly 20,000 protein coding genes, making it less complex than the human proteome which comprises of about 1,000,000 proteoforms.<sup>11</sup> One reason for this huge discrepancy in numbers is due to alternative splicing, a process in which the non-coding parts of the pre-mRNA (introns) are removed and the protein coding parts (exons) are assembled in various ways.<sup>12</sup> By means of this, the cell is able to generate more than one mRNA transcript per gene, finally resulting in approximately 100,000 different mRNAs.<sup>13</sup> However, post-translational modifications (PTMs) have the greatest impact on proteome complexity, as each protein can be post-translationally modified with one or more PTMs, resulting in about 1,000,000 proteoforms (Figure 1).<sup>14</sup> Generally, PTMs are reversible covalent modifications of protein side chains with various functional groups, which can be classified into chemical groups, complex groups and polypeptides (Figure 2). Chemical groups include modifications such as phosphorylation, methylation or acetylation, whereas complex groups contain ADP-ribosylation, AMPylation or glycosylation. Furthermore, polypeptides such as ubiquitin or SUMO can similarly be used to modify proteins. In contrast to the addition of functional groups, the proteolytic cleavage of the peptide backbone is referred as an irreversible PTM (Figure 2).<sup>15</sup> Dependent on the type and site of the PTM, the molecular properties of the modified protein are specifically changed. These include, for instance, alterations in the conformation, charge, hydrophobicity and hydrophilicity.<sup>16</sup> On this basis, PTMs are able to regulate the activity, localization and interaction of cellular enzymes.<sup>17</sup> Among the large number of PTMs, one PTM recently gained more attention, namely AMPylation.



**Figure 1: Schematic representation of the proteome complexity.** Around 20,000 human genes are transcribed into pre-mRNA. Due to alternative splicing, the number of mRNAs is increasing to around 100,000. After translation of mRNAs to proteins, each protein can be modified with one or more post-translational modifications, resulting in about 1,000,000 proteoforms.

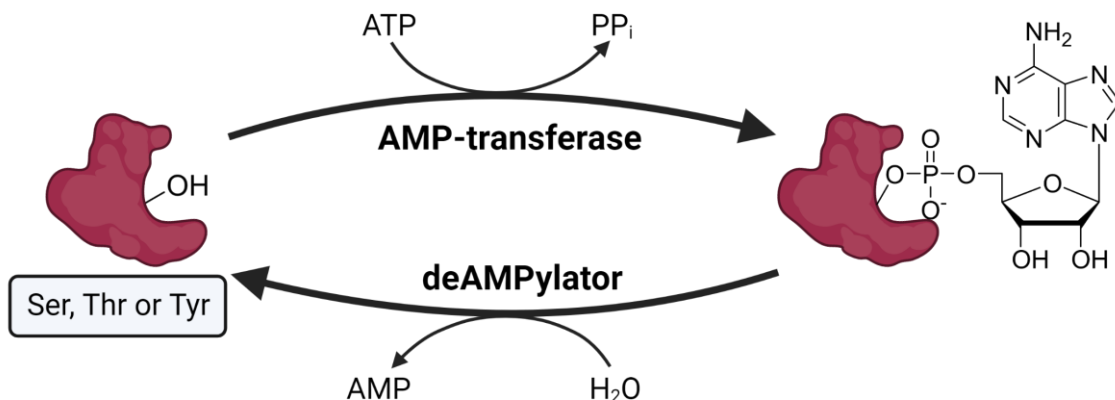


**Figure 2: Schematic representation of the diversity of PTMs.** PTMs can be classified in different groups based on their size, structure and reversibility.

## 1.2 Protein AMPylation

### 1.2.1 General features of protein AMPylation

Protein AMPylation, also known as adenylation, is a reversible PTM in eukaryotes and prokaryotes. It is defined as the covalent attachment of adenosine 5'-*O*-monophosphate (AMP) to serine, threonine or tyrosine under the formation of a phosphodiester bond. Like kinases, AMP-transferases use adenosine triphosphate (ATP) as a co-substrate to catalyze AMP attachment (Figure 3). Mechanistically, for AMPylation the  $\alpha$ -phosphate of ATP is attacked by a nucleophile, whereas for phosphorylation the  $\gamma$ -phosphate is targeted.<sup>18</sup> The catalytic activity of AMPylators is based on two distinct domains, the adenylyl transferase domain (ATase) domain with the signature motif G-X<sub>11</sub>-D-X-D in its active site and the filamentation induced cAMP (Fic) domain with the conserved HXFX(D/E)GNRXXR sequence motif.<sup>19–22</sup> The majority of AMPylators was identified in prokaryotes, such as VopS, DrrA and IbpA, where they function as bacterial virulence factors. During infection, they are secreted into mammalian host cells and alter the function of mammalian host cell proteins by AMPylation.<sup>23–26</sup>



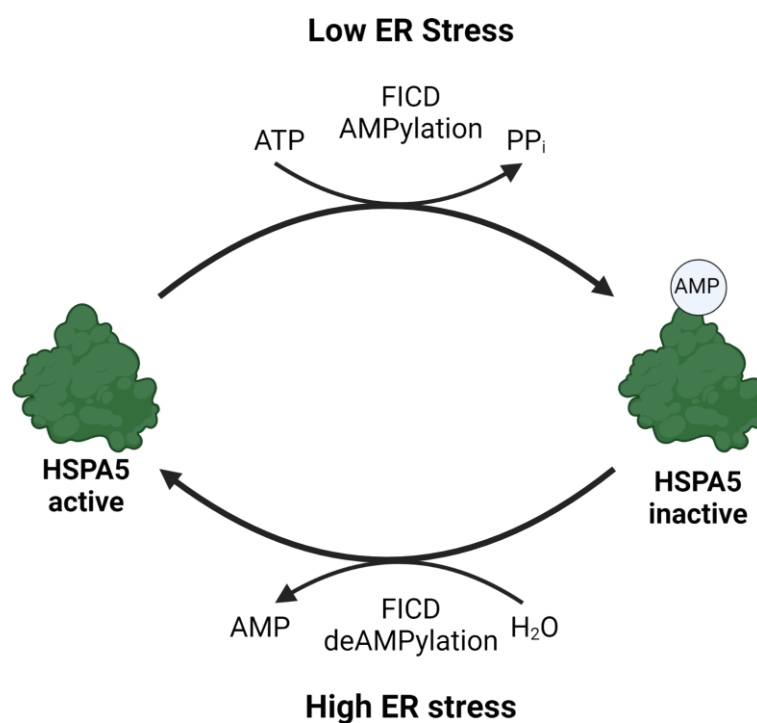
**Figure 3: Reaction scheme of protein AMPylation.** Amino acids that bear hydroxyl side group chains, such as serine (Ser), threonine (Thr) or tyrosine (Tyr), can be modified with an AMP catalyzed by an AMP-transferase that uses ATP as a co-substrate. Afterwards, AMP can be hydrolyzed by so-called deAMPylators.

### 1.2.2 AMPylation in eukaryotes

In contrast to prokaryotic cells, only two eukaryotic AMPylators were identified so far, namely the Fic domain containing protein (FICD) also known as huntingtin yeast-interacting protein E (HYPER) and the pseudokinase Selenoprotein-O (SELENOO).<sup>27,28</sup> FICD is localized in the endoplasmic reticulum (ER) and its reactivity is strictly regulated by an autoinhibitory  $\alpha$ -helix. To deactivate the inhibition loop, a crucial glutamate (E234) could be exchanged for glycine, resulting in constitutively activated FICD.<sup>21</sup> The best studied target of FICD is the molecular chaperone HSPA5, also known as BiP.<sup>29–32</sup> HSPA5 is one of the key enzymes in ER homeostasis as it triggers the unfolded protein response (UPR) and facilitates protein folding.<sup>33,34</sup> Interestingly, FICD was shown to not only AMPylate HSPA5 but also to be capable of deAMPylating HSPA5.<sup>35,36</sup> With regards to this, it was proposed that HSPA5 is AMPylated by FICD at low ER stress to deactivate its activity, while HSPA5 is deAMPylated by FICD at high ER stress to increase its activity (Figure 4).<sup>37,38</sup> The importance of the reversible AMPylation of HSPA5 was demonstrated in *Drosophila* that were lacking the Fic domain. It was observed that the *Drosophila* were blind, suggesting an important role of the Fic domain in visual neurotransmission.<sup>39</sup> In further *Drosophila* studies it was shown that the HSPA5 AMPylation mitigates photoreceptors degradation when exposed to constant light.<sup>40</sup> Recently, McCaul *et al.* prepared mFICD knockout mice and observed that the deficiency is well tolerated in unstressed mice. They proposed that mFICD could play a role in the adaptive immunity and neuronal plasticity.<sup>41</sup> Furthermore, a second independent laboratory prepared mFICD knockout

mice and investigated its role in the exocrine pancreas. Their findings suggest that mFICD acts as a molecular rheostat to regulate the UPR which is in line with its so far proposed role.<sup>42</sup>

In addition to FICD, a second human AMPylator, SELENOO, was identified by *Sreelatha et al.* in 2018 and is localized in the mitochondria. SELENOO is a pseudokinase and adopts a protein kinase fold. However, compared to conventional kinases, the ATP in the active site is flipped, which explains its AMPylation activity. It is suggested that SELENOO plays a role in the cellular response to oxidative stress, as its AMPylation targets are involved in redox homeostasis.<sup>28</sup> Recently, another group proposed that the *Salmonella* homolog of SELENOO, YdiU, has a UMPylation activity which is regulated by its self-AMPylation.<sup>43</sup>



**Figure 4: Regulation of the HSPA5 activity by AMPylation upon ER stress.** During high ER stress, FICD deAMPylates HSPA5 and activates it. In contrast, FICD AMPylates HSPA5 to inactivate it when the ER stress is low.

### 1.2.3 AMPylation in neurodevelopment and neurodegeneration

In recent years, *Kielkowski et al.* provided evidence that AMPylation plays a putative role in neurodevelopment and neurodegeneration. This observation was based on a chemical proteomics experiment which revealed that the AMPylation pattern during neurodevelopment was changed towards cytoskeletal proteins in neurons. In addition, several cathepsins, such as CTSB, CTSC and CTSL, were found to be AMPylated and are known to be associated with neurodegenerative diseases.<sup>2,3</sup> This hypothesis is also supported by a study in *C. elegans* in which the constitutively active homolog of FICD (E274G) was overexpressed and revealed to increase the aggregation of  $\beta$ -amyloid and  $\alpha$ -synuclein. This observation was explained by a potential loss of chaperone activity of the corresponding heat-shock proteins due to their AMPylation.<sup>1</sup> Further *in vitro* studies show that  $\alpha$ -synuclein is even directly AMPylated by FICD<sup>44</sup>. The connection of AMPylation and neurodegeneration was further corroborated by a patient study which shows that the Arg374His point mutation in FICD causes a motor neuron disease. The authors demonstrated that the point mutation in FICD increases its activity and thereby inactivates the molecular chaperone HSPA5 by AMPylating it.<sup>45</sup> However, the exact molecular mechanisms and defined targets to explain the connection between AMPylation and neurodevelopment as well as neurodegeneration remain unclear.

### 1.2.4 Methods to analyze AMPylated proteins

#### 1.2.4.1 *In vitro* methods

All approaches for the analysis of AMPylation presented in this section are performed *in vitro*, meaning in cell lysates. Most of them are based on methods that were already used to study phosphorylation as both use ATP as a co-substrate. One of the first approaches used to investigate AMPylation was radioactive labeling. In contrast to phosphorylation, which uses [ $\gamma$ -<sup>32</sup>P]-labeled ATP, AMPylation utilizes an [ $\alpha$ -<sup>32</sup>P]-labeled ATP because AMPylation involves the transfer of the  $\alpha$ -phosphate group with adenosine instead of the  $\gamma$ -phosphate group (Figure 5A). For the analysis, the radioactive [ $\alpha$ -<sup>32</sup>P]-labeled ATP is added to cell lysates, where it is accepted by the endogenous or recombinantly overexpressed AMPylator. Thereby, the AMPylated target proteins are radioactively labeled, enabling their subsequent separation by SDS-PAGE and visualization by autoradiography.<sup>27,28</sup>

Antibodies are another valuable tool for analyzing PTMs, since they allow the detection of the PTM of interest by Western blot and its enrichment by immunoprecipitation (IP). In contrast to other PTMs, there is not a wide range of antibodies commercially available to detect



AMPylation. Up to date, there are two commercially available antibodies, one against AMPylated threonine (Sigma-Aldrich 09-890) and one against AMPylated tyrosine (ABS184).<sup>46,47</sup> However, both have only been validated against small GTPases and are therefore not independent of the protein backbone. Furthermore, they cannot be used for IPs of native proteins as they only work against denatured targets. To overcome these drawbacks, *Höpfner et al.* have developed an antibody that recognizes protein AMPylation independent of the protein backbone. In addition, it detects AMPylation in native as well as in denatured proteins.<sup>48</sup>

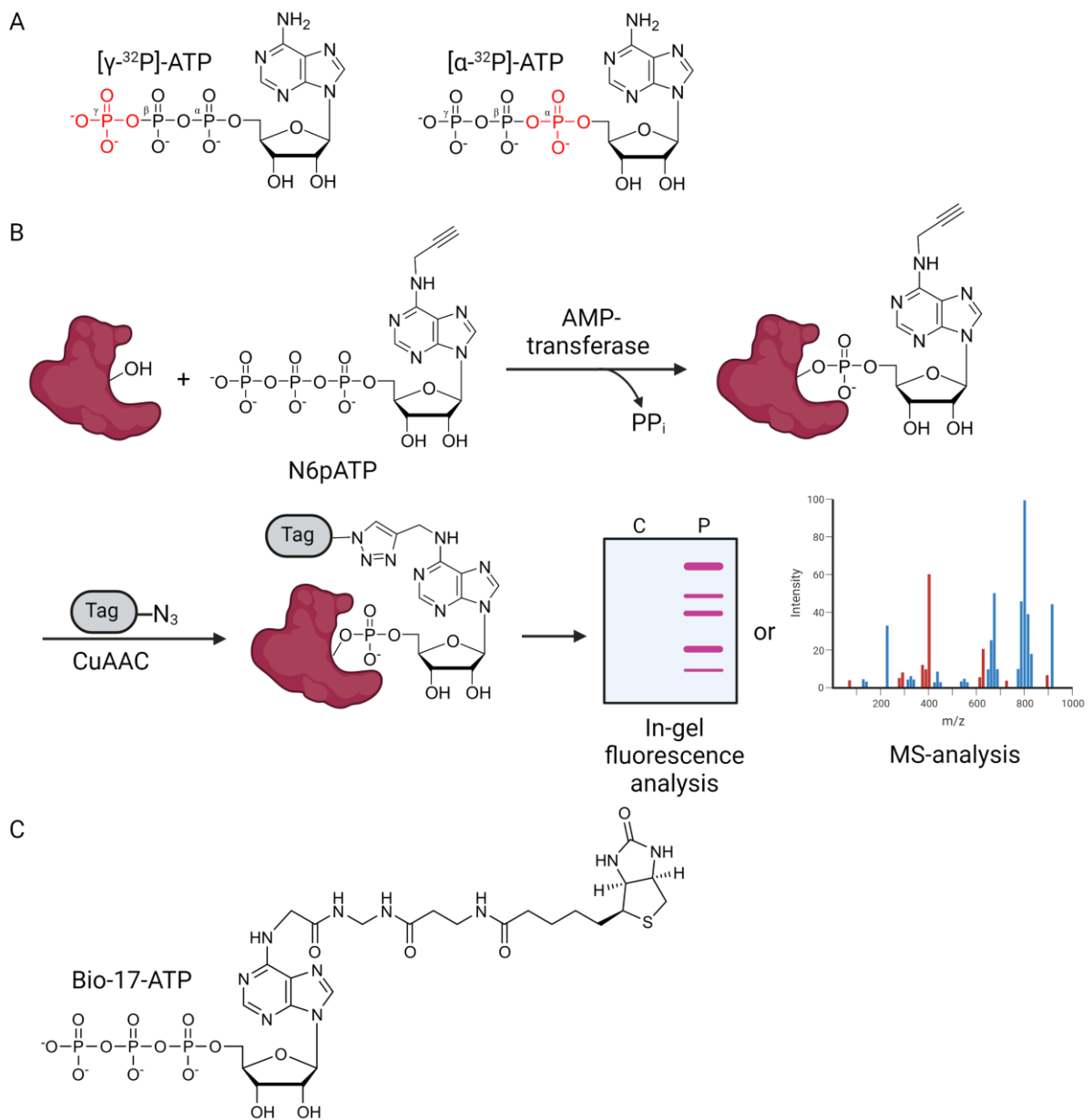
The following methods are based on bioorthogonal reactions to selectively modify the proteins of interest with fluorescent dyes or enrichment tags. Since the reaction is performed in lysates, it is crucial that the reactants do not cross-react with other endogenous functional groups. Furthermore, the reaction should tolerate low substrate concentrations, neutral pH and an aqueous buffer system.<sup>49</sup> These prerequisites are met by the well-known Cu(I)-catalyzed azide-alkyne cycloaddition (CuAAC), which is based on the reaction of a terminal alkyne with an azide under catalysis by Cu(I). The Cu(I) species is produced *in situ* by the reduction of Cu(II) with Tris(2-carboxyethyl)phosphine (TCEP). In addition to TCEP, ligands such as tris(benzyltriazolyl)methyl amine (TBTA) are added to accelerate the reaction and stabilize the catalytically active Cu(I) species in aqueous solution.<sup>50</sup>

One method that utilizes CuAAC to study AMPylation is based on the ATP-analogue N<sup>6</sup>-propargyl adenosine-5'-*O*-triphosphate (N6pATP) (Figure 5B). Similar to the radioactive labeled ATP, N6pATP is added to cell lysates and is accepted from the endogenous or recombinantly overexpressed AMPylators. By means of this, the target proteins are modified with an AMP that additionally bears an alkyne moiety. The alkyne moiety allows to click the AMPylated proteins to fluorescent dyes or affinity tags by CuAAC. Then, fluorescent dyes are used to visualize AMPylated proteins after SDS-PAGE by in-gel fluorescence. In contrast, affinity tags such as biotin can be used to enrich the AMPylated proteins and subsequently analyze them by LC-MS/MS.<sup>51,52</sup>

To identify the AMPylation targets of a specific AMPylator such as FICD, a high-throughput screen known as nucleic acid programmable protein array (NAPPA) was developed. This method uses a cDNA library to overexpress all possible target proteins with a specific tag by *in vitro* transcription and translation (IVTT). Subsequently, antibodies against the specific tag are utilized to attach the proteins to the array surface and the remaining cDNA is removed with DNase. In the next step, the captured proteins are treated with the N6pATP probe and the

recombinantly overexpressed AMPylator to label the target proteins. Followed by click chemistry with an azide-containing fluorophore, proteins are detected and identified on the array at a specific spot.<sup>53,54</sup>

In addition to radiolabeled ATP and N6pATP, a biotinylated ATP analogue (Bio-17-ATP) was used by *Sreelatha et al.* to identify the AMPylation targets (Figure 5C). Similar to the other methods, the ATP analogue is added to cell lysates where it is accepted by the endogenous or recombinantly overexpressed AMPylator. Since Bio-17-ATP already contains an affinity tag, the labeled proteins can be directly enriched on streptavidin-coated beads without any further click reaction. After the enrichment, the AMPylated proteins are analyzed with LC-MS/MS.<sup>28</sup> However, a disadvantage of this method is that the biotin group of Bio-17-ATP is bulky and therefore may decrease the acceptance by the AMPylator compared to N6pATP.



**Figure 5: *In vitro* methods to analyze AMPylation.** (A) Chemical structure of the radioactive ATP analogues  $[\gamma\text{-}^{32}\text{P}]\text{-ATP}$  and  $[\alpha\text{-}^{32}\text{P}]\text{-ATP}$ . (B) The ATP analogue N6pATP is added to cell lysates and modifies the AMPylation target proteins. The alkyne group of the modified proteins can either be coupled to fluorescent dyes or affinity tags using CuAAC. Fluorescently labeled proteins can be visualized by SDS-PAGE, while proteins clicked to an affinity tag are analyzed by MS after enrichment. (C) Chemical structure of Bio-17-ATP.

#### 1.2.4.2 *In vivo* methods based on the ProTide strategy

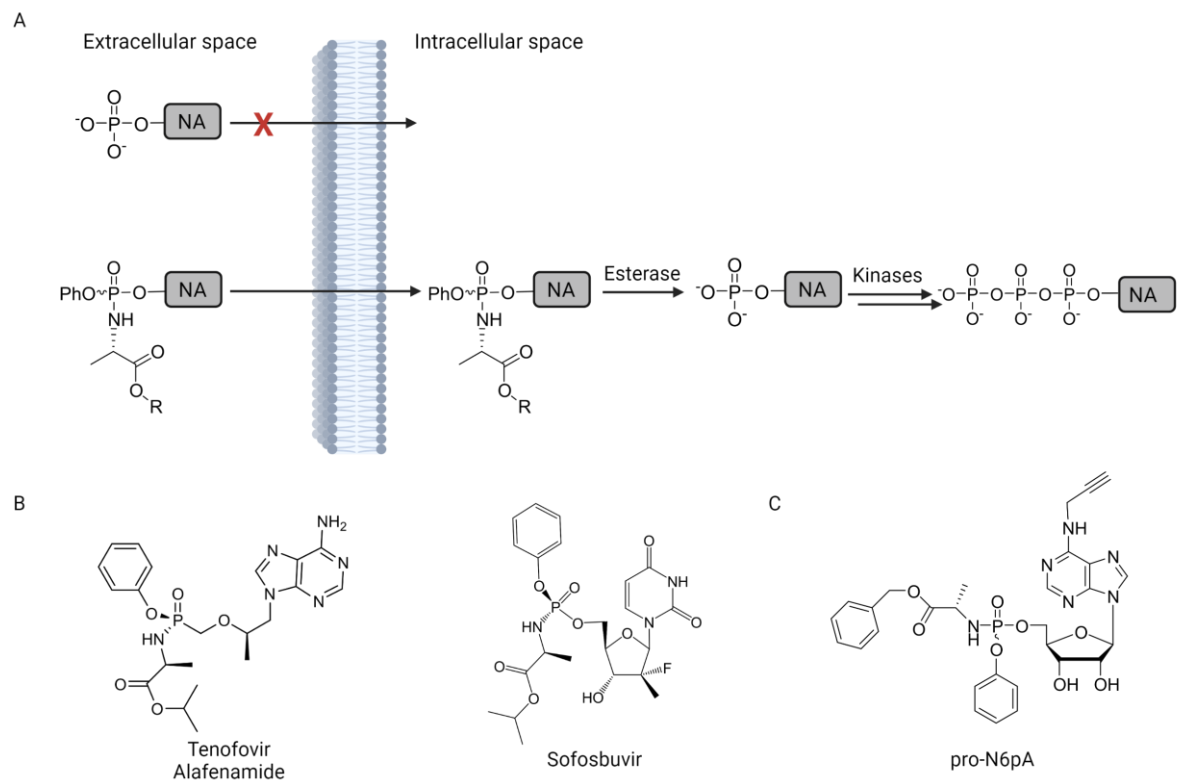
All ATP analogues mentioned in the section above are unable to efficiently cross the cell membrane due to their negative charges on the phosphate groups. Therefore, they are not suitable to study AMPylation under physiological conditions inside living cells. To enable cell permeability of nucleotide analogues, phosphoramidate pronucleotides, known as ProTides, have been developed.<sup>55,56</sup>

The ProTide strategy is a method to deliver nucleoside monophosphate analogues into living cells by masking the negative charges of the phosphate group. Another important feature of this approach is that the first rate-limiting phosphorylation step is bypassed as the nucleoside monophosphate is used instead of the corresponding nucleoside. Once the nucleoside monophosphate analogue is inside the cell, the masking groups are enzymatically cleaved, releasing the charged nucleoside monophosphate analogue. In the next step, intracellular kinases phosphorylate the nucleoside monophosphate analogue twice, finally producing the active nucleoside triphosphate analogue (Figure 6A).<sup>56</sup> This approach has already led to the FDA approval of several antiviral drugs such as sofosbuvir against hepatitis C or tenofovir alafenamide against hepatitis B (Figure 6B).<sup>57,58</sup>

In order to study AMPylation *in vivo*, Dr Kielkowski developed an AMPylation probe (pro-N6pA) based on the ProTide strategy suitable for chemical proteomics (Figure 6C). The masking groups of the phosphate group allow the pro-N6pA probe to cross the cellular membrane and be metabolically activated to its corresponding nucleoside triphosphate, N6pATP. In addition to its cell permeability, the pro-N6pA probe was equipped with a propargyl moiety on the N<sup>6</sup> amino group, which facilitates bioorthogonal coupling. In contrast to the *in vitro* methods, the pro-N6pA probe can be supplemented to the cell culture medium and is metabolically incorporated by AMP transferases. Followed by cell harvest and lysis, labeled proteins were coupled to a fluorophore for in-gel analysis or to biotin-azide for subsequent enrichment and LC-MS/MS analysis.<sup>2</sup>

One challenge in studying AMPylation under physiological conditions is the inherent competition of the metabolically activated N6pATP with its natural substrate ATP. This leads to a sub-stoichiometric amount of probe-modified proteins, as less N6pATP is utilized by the AMPylators. Another concern regarding the pro-N6pA probe was that ADP-ribosylation was detected instead of AMPylation. This assumption was based on the fact that ATP is a precursor in the biosynthesis of NAD<sup>+</sup>, which is the substrate for ADP-ribosylation. To rule out this hypothesis, HeLa cells were treated with olaparib, an inhibitor of the two main poly(ADP-

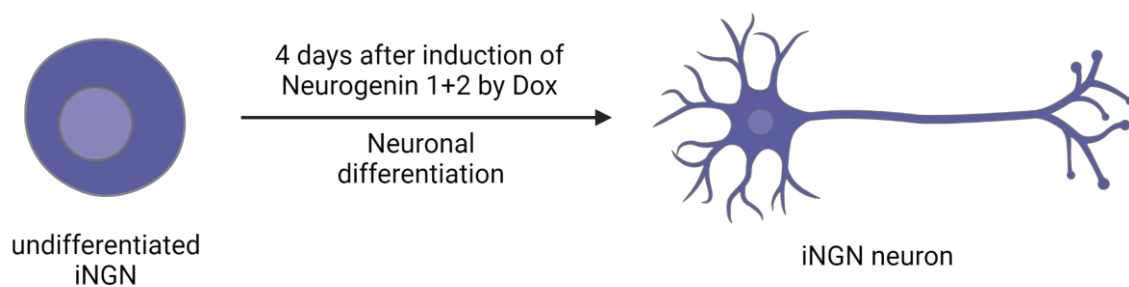
ribose)-polymerases 1 and 2 (PARPs), before the addition of pro-N6pA. The in-gel and MS-based chemical proteomics experiments revealed that the olaparib treatment had no effect on AMPylation, showing that the pro-N6pA probe does not label ADP-ribosylated proteins.<sup>2</sup>



**Figure 6: Schematic representation of the ProTide strategy.** (A) The negative charges of the phosphate groups are masked to allow the Nucleoside analogues (NA) to pass the cellular membrane. Inside the cell the masking groups are enzymatically cleaved by esterases and subsequently phosphorylated twice by intracellular kinases. Thereby, the biologically active triphosphate is released. (B) Chemical structure of the FDA approved drugs tenofovir alafenamide and sofosbuvir. (C) Chemical structure of the ProTide strategy based AMPylation probe pro-N6pA.

### 1.3 iNGNs model of neuronal differentiation

In order to develop novel therapeutics, the underlying molecular pathways of each disease have to be understood in detail. For this purpose, a suitable cellular system is of great importance. With regards to this, *Takahashi et al.* made great progress in 2006 by developing induced pluripotent stem cells (iPSCs), which allow to study diseases *in vitro* with cells derived directly from patients.<sup>59,60</sup> The generation of iPSCs is based on the four transcription factors Oct3/4, Sox2, c-Myc and Klf4, which reprogram adult fibroblasts to iPSCs that are similar to embryonic stem cells. Due to their embryonic stem cell like character, iPSCs can be differentiated to a variety of different cell-types, such as to neurons. However, the differentiation process from iPSCs to mature neurons is time-consuming and needs around 50 to 70 days.<sup>61</sup> To shorten the differentiation process, so-called iNGNs were developed that differentiate to functional neurons within 4 days.<sup>4,5</sup> The faster differentiation process is achieved by inducing the expression of both neurogenic transcription factors Neurogenin-1 and Neurogenin-2 in iPSCs with doxycycline (Figure 7). Because of the fast differentiation process from iPSCs to functional neurons, iNGNs represent a good model system to study changes in neuronal differentiation. In particular, iNGNs are suitable to study the AMPylation changes during neurogenesis in a reasonable time frame using the pro-N6pA probe in combination with chemical proteomics.



**Figure 7: Schematic representation of the iNGN differentiation process.** The expression of Neurogenin-1 and Neurogenin-2 is induced by the addition of doxycycline (Dox). 4 days after induction, iNGNs are differentiated to functional iNGN neurons.

## 1.4 Neurodegenerative disorders

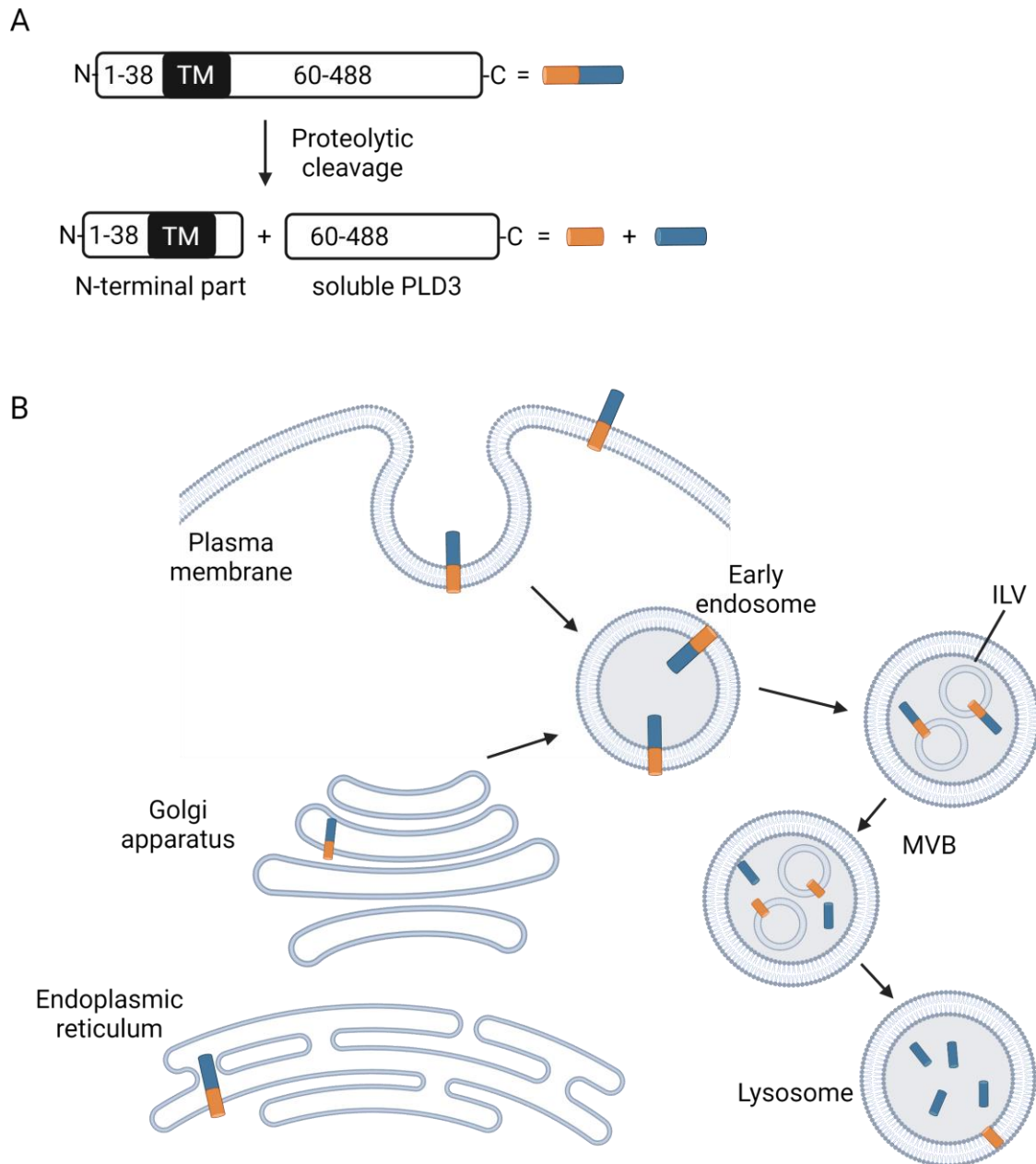
Life expectancy is increasing worldwide, which means that neurodegenerative disorders such as dementia become a growing social problem as it mainly affects people aged 60 years or older.<sup>62</sup> It is estimated that worldwide already in 2015 46.8 million people were living with dementia, but that number is expected to rise to about a 130 million people by 2050. This huge increase in dementia cases naturally translates into higher health care system costs, which were already \$818 billion in 2015 and are expected to rise to \$2 trillion by 2030.<sup>63</sup> However, not only the health system is burdened, also the patients, who suffer from symptoms such as cognitive deficits, mood disorders and disturbances in motor function.<sup>64</sup> For these reasons, great efforts have been made in the last decades to develop an effective treatment for neurodegenerative diseases. Just for clinical trials to treat Alzheimer's disease pharmaceutical companies have spent 42.5 billion dollars since 1995 without finding a breakthrough disease-modifying therapy.<sup>65</sup> In this regard, the majority of the therapeutic approaches aim to lower the levels of  $\beta$ -amyloid plaques or tau neurofibrillary tangles.<sup>66</sup> Surprisingly, last year the FDA approved aducanumab, an antibody that has been shown to reduce  $\beta$ -amyloid plaques. However, there has been a great deal of controversy whether the reduction in  $\beta$ -amyloid plaques observed in patients treated with aducanumab also significantly reduces their cognitive decline.<sup>67</sup> This raises the question whether  $\beta$ -amyloid is the right target for the treatment of Alzheimer's disease and if the real cause to tackle this disease is still unknown.<sup>68</sup> Therefore, it is of great importance to better understand the molecular mechanisms and identify novel pathways involved in Alzheimer's disease. One possible way to identify so far unknown risk factors for Alzheimer's disease is through genome-wide association studies (GWAS). In 2014, a large GWAS study of 11,000 cases and controls identified a potential novel risk factor, a rare variant of the protein phospholipase D3 (PLD3; V232M).<sup>10</sup> The connection between PLD3 and Alzheimer's disease was confirmed in several studies and will be discussed in more detail below.<sup>6-10</sup>

### 1.4.1 Characteristics of PLD3

PLD3 is a glycosylated type II transmembrane protein localized in the lysosomes.<sup>69,70</sup> The 38-amino-acid-long N-terminus resides in the cytosol whereby its 430-amino-acid-long C-terminus is on the luminal side of the membrane (Figure 8A). In contrast to other lysosomal proteins, PLD3 is transported there through an uncommon pathway that depends on the endosomal sorting complex required for transport (ESCRT). During its transport, PLD3 is proteolytically cleaved in intraluminal vesicles (ILV) of multivesicular bodies (MVB) resulting in a stable soluble form localized in the lysosomal lumen (Figure 8B).<sup>70</sup> As mentioned above, a rare coding variant of PLD3 (V232M) was identified to be a potential risk factor for Alzheimer. In regards to this, *Cruchaga et al.* have demonstrated that overexpression of PLD3 decreases the  $\beta$ -amyloid precursor protein (APP) and extracellular  $\beta$ -amyloid levels, whereas its knockout increases the amount of extracellular  $\beta$ -amyloid.<sup>10</sup> However, these results are still controversial, as different groups were not able to reproduce them at all or only partially.<sup>6-9</sup> In addition, the exact mechanism of how PLD3 might increase the risk of developing Alzheimer's disease is still unclear and needs further investigations in the future.

The catalytic activity of PLD3 is still under debate. While *Nackenhoff et al.* suggest that it possesses a lipase activity, *Gavin et al.* provide strong evidence that it has 5' exonuclease activity.<sup>71-73</sup> With respect to its 5' exonuclease activity, PLD3 was shown to digest ssDNA, a substrate of toll-like receptor 9 (TLR9) as well as ssRNA, a substrate of TLR7. By degrading these substrates, PLD3 is able to regulate the cytokine response. This was demonstrated in PLD3 knockout mice that developed spontaneous hemophagocytic lymphohistiocytosis (HLH), whereas PLD3 knockout mice lacking endosomal TLR signaling were healthy.<sup>71,73</sup>





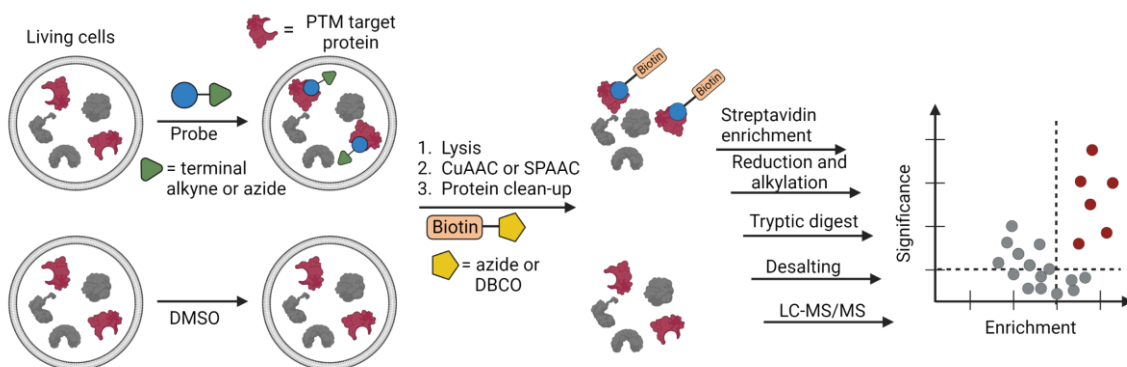
**Figure 8: Schematic representation of the transport mechanism and processing of PLD3.** (A) PLD3 can either exist in its full-length form or in its active soluble form after processing. (B) The transport of PLD3 is dependent on the ESCRT pathway. PLD3 is transported from the early endosomes to multivesicular bodies (MVB). Afterwards, PLD3 is proteolytically cleaved in intraluminal vesicles (ILV) of MVB resulting in its stable soluble form localized in the lysosomal lumen. Orange cylinder represents the N-terminal part of PLD3, while the blue cylinder visualizes the soluble form of PLD3.

## 1.5 Mass spectrometry

### 1.5.1 Sample preparation in chemical proteomics

Mass spectrometry based chemical proteomics is a powerful method to study PTMs under various conditions. The first challenge in chemical proteomics is to design a probe that mimics its natural counterpart, bears a bioorthogonal handle, is taken up by the cell and is accepted by the endogenous PTM writer enzyme. Once a suitable probe has been synthesized, such as the AMPylation probe pro-N6pA, the chemical proteomic workflow can be performed, which generally consists of the same key steps. The first step comprises metabolic labeling of the target proteins. Therefore, the probe is added to the cells where it is taken up and subsequently accepted by the PTM writer enzyme. In the next step, the cells are lysed and the probe-modified proteins are coupled to biotin using CuAAC or strain-promoted azide-alkyne cycloaddition (SPAAC).<sup>74,75</sup> Subsequently, the excess of click reagents and non-protein components of the cell lysates have to be removed, which was so far done with acetone precipitation or chloroform-methanol precipitation. The biotinylated proteins are then enriched on either agarose-based streptavidin-coated beads or magnetic streptavidin-coated beads. Before the enriched proteins are digested by trypsin or another protease, the disulfide bonds are reduced, and the free cysteines are alkylated. Finally, the peptides are desalted and measured with LC-MS/MS (Figure 9).<sup>76</sup>

Until now, there are many different variations of the general workflow with the aim to optimize the sample quality and accelerate the protocol. Since the overall workflow is quite time-consuming and tedious, its automation would be a tremendous improvement. However, some steps of the workflow are so far not suitable for automation, such as the precipitation that is required to remove excess of click reagents and remaining non-protein components. With regards to this, *Hughes et al.* developed a method based on carboxylate-coated paramagnetic beads that was used in whole proteome analysis to clean up the lysates from detergents and chaotropes.<sup>77,78</sup> This process is called Single-Pot Solid-Phase-enhanced Sample Preparation (SP3) and was already adapted by *Backus et al.* for their chemical proteomics workflow. In more detail, they used the SP3 protocol to remove contaminants of the click reaction, digest the proteins on the carboxylate-coated paramagnetic beads and finally transfer the peptides onto streptavidin-coated agarose beads.<sup>79</sup> Nevertheless, until now no chemical proteomics workflow is known to be fully automated.



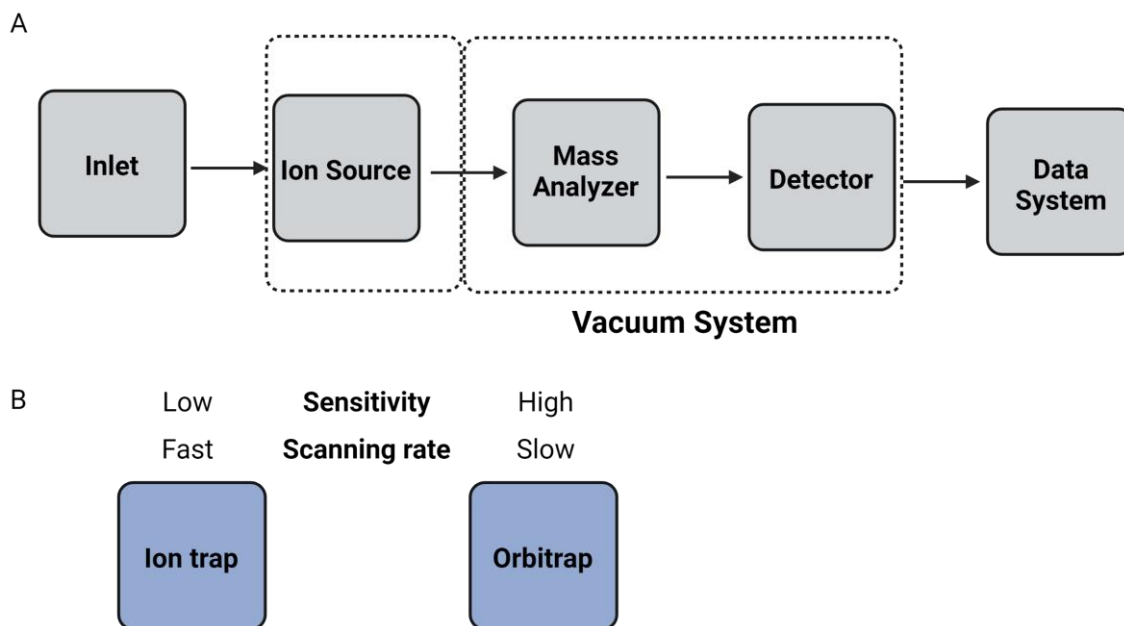
**Figure 9: Schematic representation of the general chemical proteomic sample preparation workflow.** First, living cells are treated with a probe that mimics the PTM of interest and either bears a terminal alkyne or an azide. During the incubation time, the probe is post-translationally attached to the target proteins by endogenous enzymes. Afterwards, the cells are lysed and coupled to biotin by CuAAC or SPAAC. Next, the excess of the click reagents is removed, and the biotinylated proteins are enriched with streptavidin-coated beads. Before the enriched proteins are digested with trypsin, they are reduced and alkylated. Finally, the peptides are desalted and analyzed by LC-MS/MS.

### 1.5.2 Orbitrap-based mass spectrometer

In general, the setup of a mass spectrometer to analyze proteomic samples always contains an inlet, an ion source, a mass analyzer, a detector and a data system (Figure 10A). The inlet is required to introduce the sample, which can be done either by an upstream high-performance liquid chromatography (HPLC) machine or by direct injection. Next, peptide ionization is achieved by the commonly used electrospray ionization (ESI) under atmospheric pressure, which is compatible with HPLC and direct injection.<sup>80</sup> Inside the mass spectrometer a high-vacuum is applied and the ions are separated based on their mass-to-charge ratio ( $m/z$ ) by a mass analyzer such as a quadrupole, ion trap, orbitrap or time-of-flight (TOF).<sup>81</sup> After passing the mass analyzer, the separated ions are detected by converting them into an electric signal. These signals can be used to prepare a MS spectrum by a specific data system. In contrast, the orbitrap detects the  $m/z$  value of an ion based on its frequency of oscillation using Fourier transformation.<sup>82</sup>

In addition to the advances made in sample preparation in recent years, great progress has also been made in the setup of mass spectrometers. Especially the sensitivity and the scanning rates have been improved enormously. Around 10 years ago, the Q Exactive orbitrap was the most advanced instrument with scanning rates of 12 Hz.<sup>83,84</sup> The Q Exactive is a hybrid instrument, meaning it has 2 mass analyzers, the quadrupole and the orbitrap. The quadrupole selects ions

with different  $m/z$  values for their further transport to the orbitrap where the ions are detected and analyzed. After the  $m/z$  values of the peptides are measured in a MS1 spectra, each peptide has to be fragmented, for example, in a higher energy collisional dissociation (HCD) cell to determine its identity in a MS2 scan by the orbitrap. Since the MS1 spectra as well as the MS2 spectra are both measured with the orbitrap, the scanning rate is limited to the speed of the orbitrap. Therefore, Thermo Fisher scientific has developed so-called tribrids such as the Orbitrap Eclipse, which contain a linear ion trap in addition to the quadrupole and orbitrap.<sup>85</sup> Although the resolution of the linear ion trap is much lower compared to the orbitrap, the resolving power is sufficient for the acquisition of MS2 spectra in a data-dependent proteomic experiment. In addition, the scanning speed of the linear ion trap is faster compared to the Orbitrap (Figure 10B). This allows parallelization of the measurement, where MS1 spectra can be recorded in the orbitrap and the MS2 spectra of the fragmented peptides are measured in the linear ion trap. Thereby, the scanning rate of a data-dependent proteomic experiment can reach around 40 Hz, which is a major advance.<sup>86</sup> In general, the performance of a mass spectrometer is always a compromise between the resolution of the MS spectra and the speed that they are acquired. On the one hand, it is desired to acquire MS spectra with the highest resolution as possible, but this leads to long scanning times. On the other hand, one would like to record as many MS spectra as possible, which is only feasible if they are acquired in low resolution. Thus, future advances in the setup of mass spectrometers will always aim to parallelize the measurement with fast and highly resolving mass analyzers.



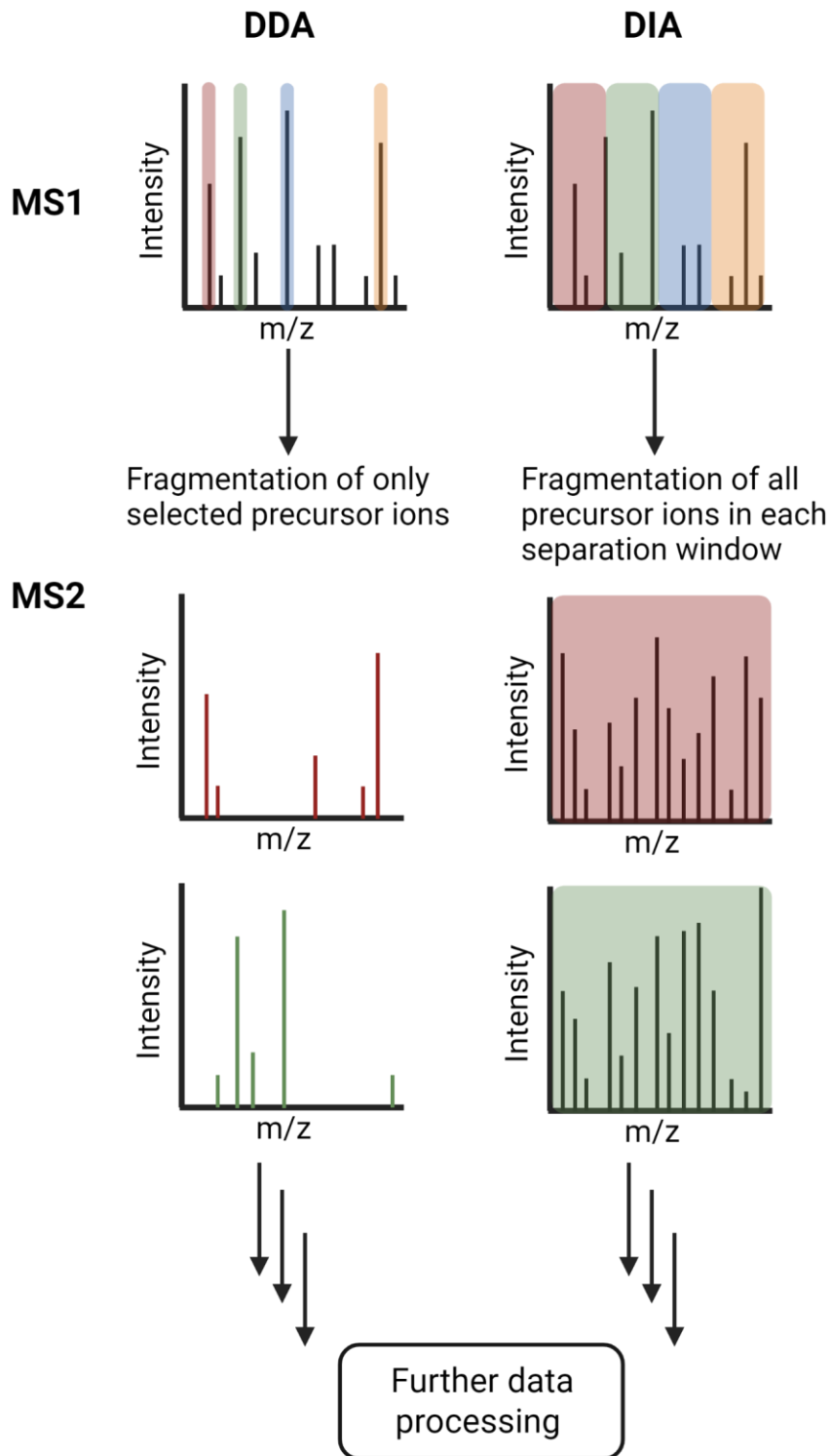
**Figure 10: General setup of a mass spectrometer.** (A) Schematic representation of a mass spectrometer setup. (B) Comparison of the sensitivity and scanning rate between the ion trap and the orbitrap.

### 1.5.3 Data acquisition

In general, there are two different types of data acquisition modes in MS/MS-experiments, the conventional data-dependent acquisition (DDA) and the recently growing data-independent acquisition (DIA).

In DDA, all precursor peptide ions are first measured in a high-resolution MS1 scan. Depending on their intensity, a pre-defined number is selected for subsequent fragmentation and acquisition of a MS2 spectrum. Commonly, the 10 or 20 most intense peaks are selected for a MS2 scan, which is referred to as a Top10 or Top20 method (Figure 11). Since the acquisition speed of novel mass spectrometers increases, it is also feasible to define a time window in which as many MS2 spectra as possible are recorded. The selection of only the most intense precursor ions has an obvious disadvantage, namely that low abundant peptides are underrepresented. Furthermore, due to the stochastic selection of precursor ions for fragmentation, the ability to identify the same group of proteins in technical replicates is limited.<sup>87</sup> In a dataset with multiple samples, this leads to a high number of missing values, which causes an error as the missing values have to be imputed for data analysis. In contrast, the number of missing values in a DIA measurement is much lower, because all precursor ions that were identified in a MS1 scan are

fragmented.<sup>88</sup> This is achieved by pre-defined isolation windows that cover a specific  $m/z$  range of the MS1. Since more than one precursor ion can be fragmented within a certain isolation window, the MS2 spectra of DIA are much more complex than those of DDA (Figure 11). Therefore, MS2 spectra in DIA have to be acquired with high resolution and in addition a spectral library is needed to enable the subsequent deconvolution. The spectral library contains information about the precursor and fragment  $m/z$  value, the elution time of each peptide and their charges.<sup>89,90</sup> The most common way for creating a spectral library is to measure the sample with DDA in advance.<sup>91</sup> However, this method is time-consuming and only allows the identification of peptides in DIA that were previously found in the DDA library run. Therefore, the DIA measurement depends entirely on the quality and depth of the DDA data set. To overcome this limitation, *Tsou et al.* have developed DIA-Umpire, a software capable of quantifying DIA results without the need of a DDA based spectral library. DIA-Umpire uses DIA MS1 and MS2 spectra to detect features that are utilized to construct pseudo-MS/MS spectra. Subsequently, the pseudo-MS/MS spectra can be used for conventional database-searching.<sup>92</sup> The most recent approach takes advantage of machine learning and creates spectral libraries *in silico*. This is possible as deep-learning algorithms are able to predict MS/MS spectra as well as their retention times.<sup>93-95</sup> This feature is already implemented in the freely available software DIA-NN and represents a major advance in DIA data analysis.<sup>96</sup>



**Figure 11: Comparison of DDA and DIA data acquisition.** In DDA, a pre-defined number of the most intense precursor ions of the MS1 spectrum are selected to be fragmented and analyzed in a MS2 scan. Thereby, each MS2 spectrum corresponds to exactly one fragmented peptide. In DIA, pre-defined isolation windows of the MS1 spectrum that cover a specific  $m/z$  range are selected for subsequent fragmentation. Each MS2 scan obtains all fragmented peptides within the specific  $m/z$  range. Therefore, the acquired MS2 spectra using DIA are more complex than the MS2 spectra of DDA.

## 2 Aim of this thesis

In the last years, several studies provided evidence that AMPylation could play a role in neurodegeneration and neurodevelopment. The study of *Truttmann et al.* have shown that aggregation of  $\beta$ -amyloid and  $\alpha$ -synuclein increases upon overexpression of the highly active FICD E274G homolog.<sup>1</sup> In addition, *in vitro* studies of *Sanyal et al.* demonstrated direct AMPylation of  $\alpha$ -synuclein by FICD.<sup>44</sup> Furthermore, *Kielkowski et al.* observed that the protein AMPylation levels change during the differentiation of iPSCs to physiological neurons using mass spectrometry-based chemical proteomics.<sup>2,3</sup> However, physiological neurons have the disadvantage that they are difficult to obtain in larger quantities and take about 50 to 70 days to differentiate from iPSCs, which restricts swift screening of PTM changes. Therefore, the first aim of this thesis is to find a suitable model cell line that allows to obtain a homogenous neuronal culture in sufficient amounts and importantly in a shorter time frame, which would enable the optimization of the chemical proteomics workflow and protein AMPylation profiling. In addition, such neuronal cell line would make it possible to study the AMPylation changes during neuronal differentiation with a higher time resolution. Although chemical proteomics is able to identify novel AMPylation targets, the previous attempts to find the sites of modification in a high-throughput manner remained elusive. Furthermore, due to relatively low peptide coverage, the mass spectrometry-based chemical proteomics fails to distinguish different protein forms, for example after a protein scission. Moreover, a probe-independent validation is critical to exclude unspecific targets. Hence, the second aim of the thesis is to develop methods that allow to distinguish AMPylated proteins from their unmodified counterparts and identify the sites of modification in a high-throughput manner.

So far, the function of AMPylation has been studied mainly in detail for the UPR, including the well-known AMPylation of HSPA5 by FICD.<sup>29-32</sup> However, previous chemical proteomics profiling experiments by *Kielkowski et al.* revealed several AMPylation targets unrelated to the UPR, such as cytoskeletal proteins or cathepsins.<sup>2</sup> Therefore, the third aim of this thesis is to link AMPylation to corresponding metabolic pathways unrelated to the UPR. For large systematic screening experiments, the chemical proteomics workflow is time-consuming and have to be performed by specialized laboratory personnel. Furthermore, high protein quantities are required, which are hard to obtain for difficult-to-culture cell lines such as neurons. Hence, the final aim of this thesis is to develop a chemical proteomics workflow which is suitable for robotic automation with a lower protein input.



## 3 Published work

### 3.1 AMPylation profiling during neuronal differentiation reveals extensive variation on lysosomal proteins

Becker, T., Cappel, C., di Matteo, F., Sonsalla, G., Kaminska, E., Spada, F., Cappello, S., Damme, M., and Kielkowski, P. (2021). AMPylation profiling during neuronal differentiation reveals extensive variation on lysosomal proteins. *iScience* 24. 10.1016/j.isci.2021.103521.

#### Prologue

In a previous study *Kielkowski et al.* developed a cell-permeable AMPylation probe (pro-N6pA) which allows to study AMPylation in living cells. By means of this, they were able to screen AMPylation in several cell lines, such as human induced pluripotent stem cells (iPSCs), neural progenitor cells (NPCs), neurons and cerebral organoids (COs). Thereby, they observed changes in the AMPylation pattern between the cell lines, suggesting a possible role of AMPylation during neurodevelopment and neurodegeneration.<sup>2,3</sup> Based on this, the aim of the study was to gain deeper insights into the dynamics of protein AMPylation during neuronal differentiation. Therefore, AMPylation was analyzed at different time-points in iNGNs that differentiate to mature neurons within 4 days.<sup>4,5,97</sup> This analysis revealed several lysosomal proteins that change their AMPylation levels during neuronal differentiation, such as PLD3, ACP2 and ABHD6. The large number of identified lysosomal proteins points towards that AMPylation may be a lysosomal PTM in addition to its well-known localization in the ER and mitochondria. This finding was further corroborated by fluorescence imaging of AMPylation in SH-SY5Y neuroblastoma cells, which showed that the probe-labeled proteins clearly colocalized with the lysosomal marker LAMP2. To validate our results in a probe-independent manner, we developed a gel-based approach using a Phos-tag ligand to separate AMPylated proteins from their unmodified form. In addition to chemical proteomics, this approach even provides the information about which form of the protein of interest is modified. Since PLD3 was associated to Alzheimer's disease in several studies, we decided to focus on the validation of this protein and analyzed its AMPylation levels using a Phos-tag gel.<sup>6-10</sup> This approach revealed that the soluble form of PLD3 is AMPylated and that the amount of AMPylated PLD3 increases from 4 to 10 days differentiated iNGNs. In physiological young and mature neurons, it was even observed that PLD3 is fully AMPylated without any unmodified form left. Finally, the activity of PLD3 was investigated, and it was found that the higher the AMPylation status, the lower its activity in iNGNs. Taken together, in this study we provide evidence that AMPylation is a lysosomal PTM and show its connection to neuronal differentiation.

### **Author contribution**

In collaboration with *Pavel Kielkowski* and *Markus Damme*, I conceived this study. Furthermore, I cultured and differentiated the iNGN cells. I performed all proteomic experiments and analyzed the data. In addition, I developed the Phos-tag gel approach and carried out all gel-based analysis. Moreover, I wrote the manuscript together with *Pavel Kielkowski*. *Cedric Cappel* conducted the fluorescence imaging in SH-SY5Y cells and the PLD3 activity assay. *Francesco di Matteo* cultured physiological hiPSCs and differentiated them into NPCs and neurons. *Giovanna Sonsalla* cultured physiological hiPSCs and differentiated them into GPCs and astrocytes. *Ewelina Kaminska* and *Fabio Spada* assisted in setting up the iNGN cell culture.

### **License**

Copy of the open access publication based on the Creative Commons Attribution-NonCommercial License (CC BY-NC-ND 4.0), which allows the non-commercial reprint of the article.

### **Article available at:**

[https://www.cell.com/iscience/fulltext/S2589-0042\(21\)01492-9](https://www.cell.com/iscience/fulltext/S2589-0042(21)01492-9)

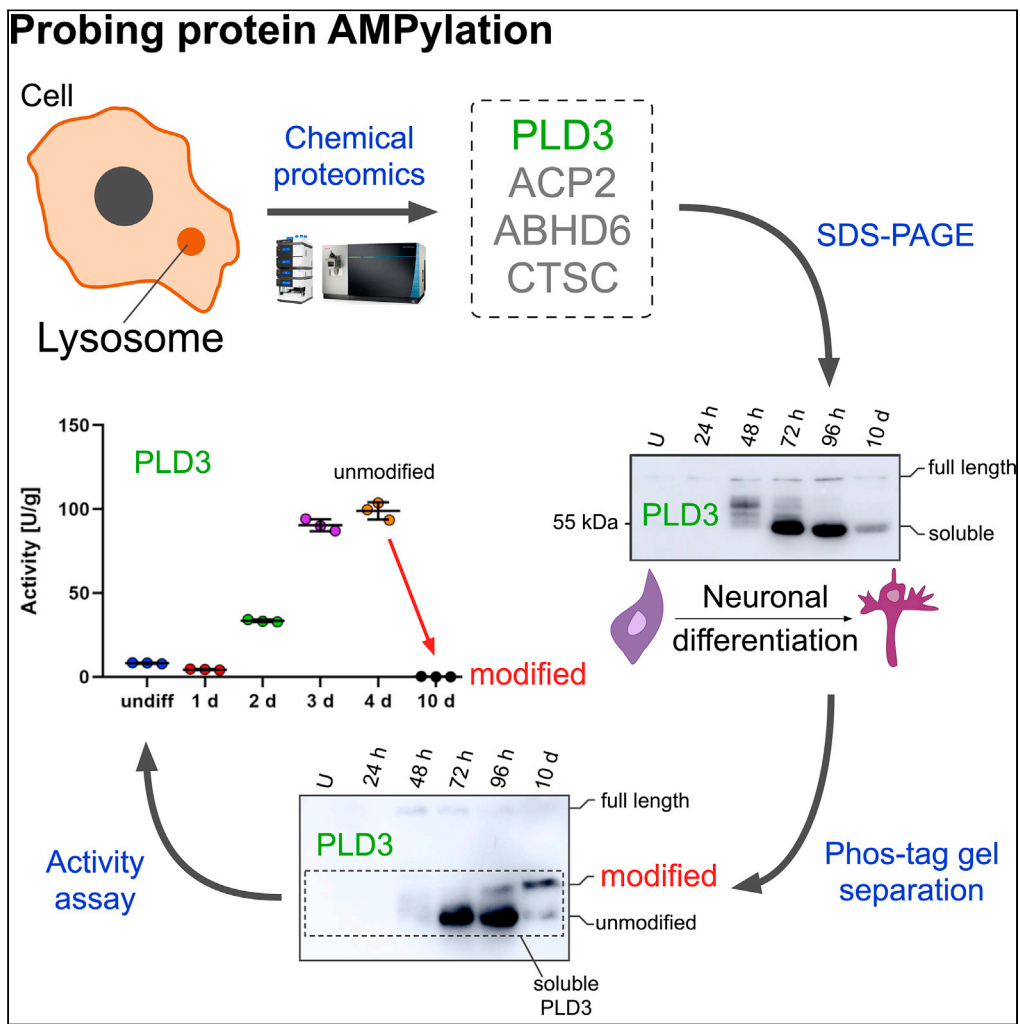
### **License available at:**

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Supporting information are available in the appendix (chapter 8.1)

Article

# AMPylation profiling during neuronal differentiation reveals extensive variation on lysosomal proteins



Tobias Becker, Cedric Cappel, Francesco Di Matteo, ..., Silvia Cappello, Markus Damme, Pavel Kielkowski

pavel.kielkowski@cup.lmu.de

**Highlights**

Profiling of AMPylation during neuronal differentiation

AMPylation is a potential PTM of luminal lysosomal proteins

Phos-tag gel enables the separation of non-AMPylated and AMPylated proteins

The modified lysosomal soluble form of PLD3 increases during neuronal maturation



## Article

## AMPylation profiling during neuronal differentiation reveals extensive variation on lysosomal proteins

Tobias Becker,<sup>1</sup> Cedric Cappel,<sup>2</sup> Francesco Di Matteo,<sup>3,4</sup> Giovanna Sonsalla,<sup>5,6,7</sup> Ewelina Kaminska,<sup>1</sup> Fabio Spada,<sup>1</sup> Silvia Cappello,<sup>3</sup> Markus Damme,<sup>2</sup> and Pavel Kielkowski<sup>1,8,\*</sup>

## SUMMARY

**Protein AMPylation is a posttranslational modification with an emerging role in neurodevelopment. In metazoans two highly conserved protein AMP-transferases together with a diverse group of AMPylated proteins have been identified using chemical proteomics and biochemical techniques. However, the function of AMPylation remains largely unknown. Particularly problematic is the localization of thus far identified AMPylated proteins and putative AMP-transferases. We show that protein AMPylation is likely a posttranslational modification of luminal lysosomal proteins characteristic in differentiating neurons. Through a combination of chemical proteomics, gel-based separation of modified and unmodified proteins, and an activity assay, we determine that the modified, lysosomal soluble form of exonuclease PLD3 increases dramatically during neuronal maturation and that AMPylation correlates with its catalytic activity. Together, our findings indicate that AMPylation is a so far unknown lysosomal posttranslational modification connected to neuronal differentiation and it may provide a molecular rationale behind lysosomal storage diseases and neurodegeneration.**

## INTRODUCTION

Protein posttranslational modifications (PTMs) provide the cell with mechanisms to swiftly react on internal and external clues to maintain the cellular homeostasis. Disruption of the protein homeostasis (proteostasis) is a hallmark of many neurodegenerative disorders (Hipp et al., 2019). Regulation of protein function by PTMs includes modulation of protein's catalytic activity, localization or protein-protein interactions (Abersold et al., 2018). Protein AMPylation comprises the attachment of adenosine 5'-O-monophosphate (AMP) onto serine, threonine, and tyrosine amino acid side chains (Figure 1A) (Sieber et al., 2020). So far two AMP transferases are known in metazoans to catalyze protein AMPylation, protein adenylyltransferase FICD (FICD), and SelO (SELENOO) (Casey and Orth, 2017; Sreelatha et al., 2018). FICD is characterized by its endoplasmic reticulum (ER) localization and dual catalytic activity of AMPylation and deAMPylation (Preissler et al., 2017; Sengupta et al., 2019). FICD catalyzes an AMP transfer from its substrate ATP and reverses the modification by the hydrolysis of the AMP-protein ester. FICD's catalytic activity is regulated by its  $\alpha$ -helix inhibition loop through the interaction of Glu234 positioned in the inhibition loop and Arg374, which is necessary for the complexation of ATP in the active site (Engel et al., 2012). Initial biochemical studies identified the ER localized heat shock protein HSPA5 (also called BiP or GRP78) as a cognate substrate of FICD (Ham et al., 2014; Sanyal et al., 2015). AMPylation of HSPA5 inhibits its chaperon activity and the downstream unfolded protein response cascade (Preissler et al., 2015). Furthermore, FICD's activity was recently shown to accelerate the neuronal differentiation of progenitor cells in human cerebral organoids, a tissue model of the human cerebral cortex (Kielkowski et al., 2020a). Surprisingly, apart from the increased number of neurons in tissue overexpressing FICD, some neurons displayed migratory defects. The neuronal role of the FIC domain was as well demonstrated in *Drosophila*, in which it is required for the visual neurotransmission in glial capitate projections (Rahman et al., 2012). In addition, FICD was shown to AMPylate  $\alpha$ -synuclein *in vitro* (Sanyal et al., 2019). In *Caenorhabditis elegans* FICD has been reported to regulate the aggregation of amyloid- $\beta$  and  $\alpha$ -synuclein (Truttmann et al., 2018). In contrast, there are only scarce data about the function of SelO and its protein substrates (Sreelatha et al., 2018).

Several complementary strategies were introduced to analyze protein AMPylation, including isotopically labeled adenosine probes (Pieles et al., 2014), radioactively labeled adenosine nucleotides, antibodies (Kingdon et al., 1967; Yarbrough et al., 2009), microarrays (Yu and LaBaer, 2015), N<sup>6</sup>-biotin-modified ATP

<sup>1</sup>LMU Munich, Department of Chemistry, Butenandtstr. 5-13, 81377 Munich, Germany

<sup>2</sup>University of Kiel, Institute of Biochemistry, Olshausenstr. 40, 24098 Kiel, Germany

<sup>3</sup>Max Planck Institute of Psychiatry, Kraepelinstraße 2, 80804 Munich, Germany

<sup>4</sup>International Max Planck Research School for Translational Psychiatry (IMPRS-TP), Kraepelinstraße 2-10, 80804 Munich, Germany

<sup>5</sup>LMU Munich, Department of Physiological Genomics, Biomedical Center (BMC), Großhadernerstr. 9, 82152 Planegg, Germany

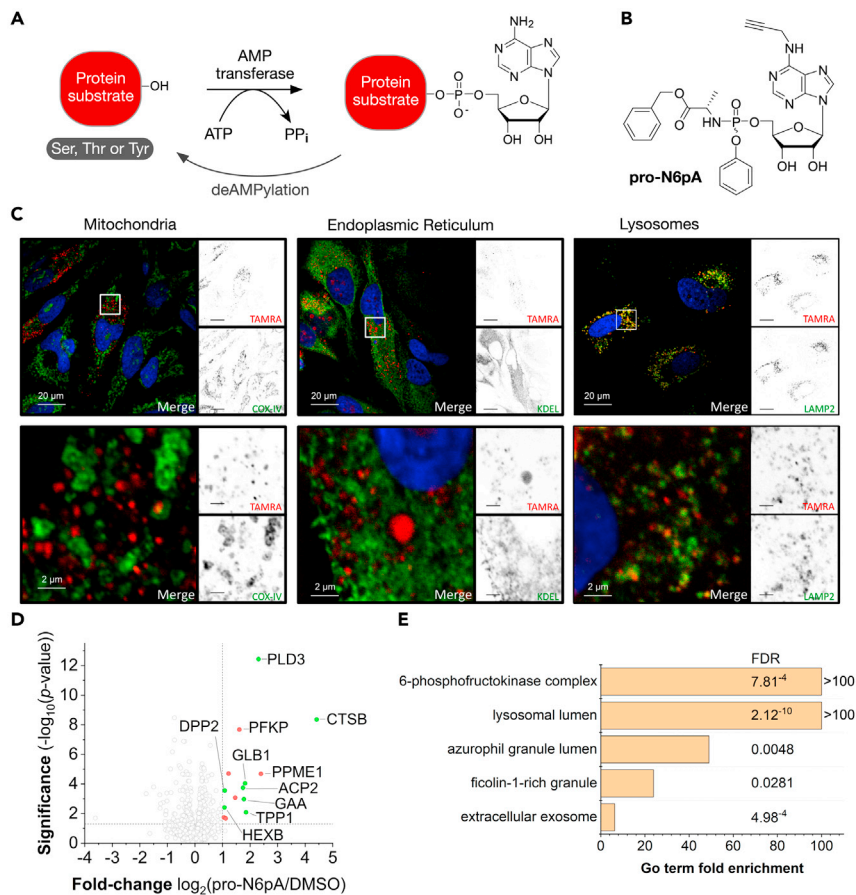
<sup>6</sup>Helmholtz Zentrum München, Institute for Stem Cell Research, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany

<sup>7</sup>Graduate School of Systemic Neurosciences (GSN), Großhadernerstr. 2, 82152 Planegg, Germany

<sup>8</sup>Lead contact

\*Correspondence: pavel.kielkowski@cup.lmu.de  
<https://doi.org/10.1016/j.isci.2021.103521>





**Figure 1. Human neuroblastoma cells (SH-SY5Y) display an enrichment of pro-N6pA probe-labeled proteins in lysosomes**

(A) Schematic representation of protein AMPylation.

(B) Structure of the pro-N6pA probe for *in situ* labeling of potentially AMPylated proteins.

(C) Click chemistry staining of pro-N6pA with TAMRA-N<sub>3</sub> (red) and immunocytochemical staining in SH-SY5Y and colocalizations with markers for lysosome (LAMP2), ER (KDEL), and mitochondria (COX-IV).

(D) Volcano plot showing the significantly enriched AMPylated proteins from SH-SY5Y cells. Proteins localized to lysosome are depicted in green. Significantly enriched protein hits with other subcellular localization are in red (cutoff lines at 2-fold enrichment and *p* value of 0.05).

(E) GO terms analysis of significantly enriched proteins in SH-SY5Y cells.

(Sreelatha et al., 2018), and N<sup>6</sup>-propargyl adenosine 5'-O-triphosphate (N6pATP) or phosphoramidate (Broncel et al., 2012; Grammel et al., 2011; Kielkowski et al., 2020a, 2020b). We have recently developed a chemical proteomic approach that allows high-throughput screening of protein AMPylation and comparison of the AMPylation levels between different conditions. This strategy utilizes the N<sup>6</sup>-propargyl or N<sup>6</sup>-ethylazide adenosine phosphoramidate (pro-N6pA and pro-N6azA, respectively) probes, which are metabolically activated to the corresponding N<sup>6</sup>-modified adenosine triphosphate upon uptake into cells and used by endogenous AMP-transferases or by bacterial effectors in infection studies for protein AMPylation (Figure 1B) (Kielkowski et al., 2020a, 2020b; Rauh et al., 2020). Of note, the resulting N<sup>6</sup>-modified ATP is inherently in competition with endogenous ATP and thus cannot report on the exact stoichiometry of protein AMPylation. The application of this chemical proteomic strategy in various cell types has revealed that protein AMPylation is more prevalent than previously assumed. Interestingly, it has led to the identification of a large group of protein substrates from different subcellular compartments that are not restricted to the ER, including cytosolic (PFKP, SQSTM1), nucleolar (PPME1), cytoskeletal (TUBB, MAP2), and lysosomal (CTSB, PLD3, and ACP2) proteins (Broncel et al., 2016; Kielkowski et al., 2020a). In particular, in the lysosome, there is only limited evidence on PTMs of luminal lysosomal proteins other than N-glycosylation amid numerous diseases associated with lysosomal proteins such as the lysosomal acid phosphatase

ACP2 in lysosomal storage diseases (LSDs) or the 5'-3' exonuclease PLD3 in Alzheimer's disease and auto-inflammatory diseases (Cappel et al., 2021; Cruchaga et al., 2014; Gavin et al., 2021; Gonzalez et al., 2018; Marques and Saftig, 2019; Schultz et al., 2011; Stadlmann et al., 2017; Stoka et al., 2016). Therefore, the discovery of lysosomal protein PTMs might shed new light on the regulation of their function, localization, and protein-protein interactions and hence putatively uncover unknown pathophysiologic mechanisms.

Here, we characterize protein AMPylation during neuronal differentiation and maturation by a combination of chemical proteomics and a gel-based approach. Thereby, we provide supporting evidence that the two lysosomal proteins PLD3 and ACP2 are increasingly AMPylated during neural differentiation and that the AMPylation of PLD3 correlates with the inhibition of its catalytic activity.

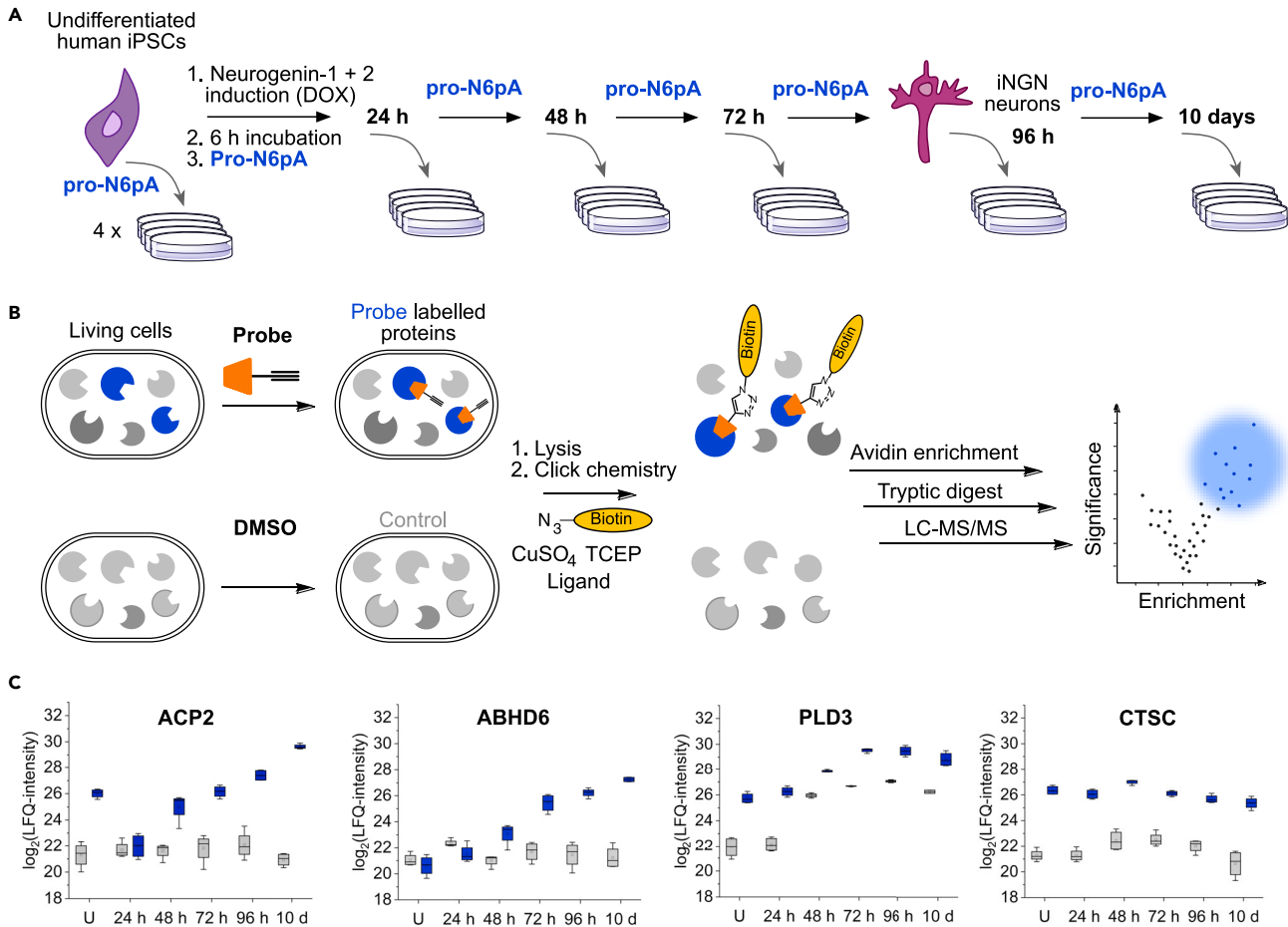
## RESULTS

### AMPylated proteins localize to lysosomes in SH-SY5Y neuroblastoma cells

We have developed a chemical proteomic strategy in previous studies to analyze the protein AMPylation in living cells using the cell-permeable pro-N6pA probe (Kielkowski et al., 2020a, 2020b). In this study, we have started with the completion of this dataset using fluorescence imaging of the SH-SY5Y cells treated with the pro-N6pA probe to visualize the subcellular localization of potentially AMPylated proteins. Interestingly, probe-labeled proteins exhibit a vesicular distribution pattern overlapping with lysosomes as confirmed by co-staining with the lysosomal-specific marker LAMP2 (Figure 1C). Colocalization of the majority of pro-N6pA probe signal with the ER and mitochondria was excluded by co-staining of the pro-N6pA probe with antibodies against KDEL and COX-IV, respectively. These findings point to an enrichment of AMPylated proteins in lysosomes and raise questions about the localization and origin of their AMPylation. GO analysis of chemical proteomics data from SH-SY5Y cells indicated an overrepresentation of lysosomal proteins among those enriched for AMPylation by 54% of all significantly enriched proteins. These include CTSB, PLD3, GAA, GLB1, TPP1, HEXB, DPP2, and GUSB, with PLD3 and CTSB showing the strongest enrichment (Figures 1D and 1E). Next, we asked whether AMPylation of lysosomal proteins changes during neuronal differentiation and maturation and how the AMPylation status correlates with non-neuronal cell types. Although the chemical proteomic analysis of protein AMPylation variation during neuronal differentiation was previously performed, it has focused solely on two differentiation stages, neuronal progenitors (NPCs) and mature neurons. To investigate the AMPylation of lysosomal proteins in differentiating neurons with higher temporal resolution, we searched for a suitable cellular system that would allow collecting sufficient amounts of total proteins for chemical proteomic analysis in shorter periods of time and thus overcoming the bottlenecks of the standard human induced pluripotent stem cells (iPSCs) differentiation protocols (Boyer et al., 2012).

### Chemical proteomic analysis of protein AMPylation in differentiating iNGN cells

To map the changes in protein AMPylation during the neuronal differentiation in more detail, we took advantage of the human induced pluripotent stem cells with inducible overexpression of a pair of transcription factors, Neurogenin-1 and Neurogenin-2, leading to their rapid differentiation into a homogeneous population of functional bipolar neurons within 4 days (iNGN) (Busskamp et al., 2014). The changes in AMPylation were followed at six time points during the iNGN differentiation and maturation using the previously described chemical proteomics strategy utilizing a metabolically activated pro-N6pA probe (Figure 2A). In brief, cells were treated with the pro-N6pA probe 16 h prior to their harvest to allow metabolic activation to the corresponding N<sup>6</sup>-propargyl ATP, which is used as a substrate by endogenous AMP-transferases to label proteins. Subsequently, the cells were lysed and the alkyne modified proteins were further coupled with biotin-PEG-azide by Cu(I) catalyzed click chemistry. The pro-N6pA labeled and biotinylated proteins were then enriched on avidin-coated agarose beads and on beads trypsinized to yield complex peptide mixtures, which were resolved by LC-MS/MS measurements. The resulting label-free quantification (LFQ) of proteins from four replicates and their comparison with DMSO-treated cells prepared in the same manner provided the quantitative differences in protein AMPylations between undifferentiated and differentiated iNGN cells (Figures 2B and S1, Tables S1, S2, S3, S4, S5, S6, and S8). Of note, owing to the metabolic activation of the pro-N6pA probe, a minor incorporation, for example, as ADP-ribosylation, cannot be completely excluded. The background from unspecific protein binding to the avidin-agarose beads in both vehicle control (DMSO) and probe-treated cells partially reflects the total protein expression level. To correct for this contribution, we have carried out the whole proteome analysis of iNGNs during the course of differentiation (Figure S2). In addition, the whole proteome analysis confirmed the identity of the cells and the progress of neuronal differentiation and maturation (Figure S3 and Tables



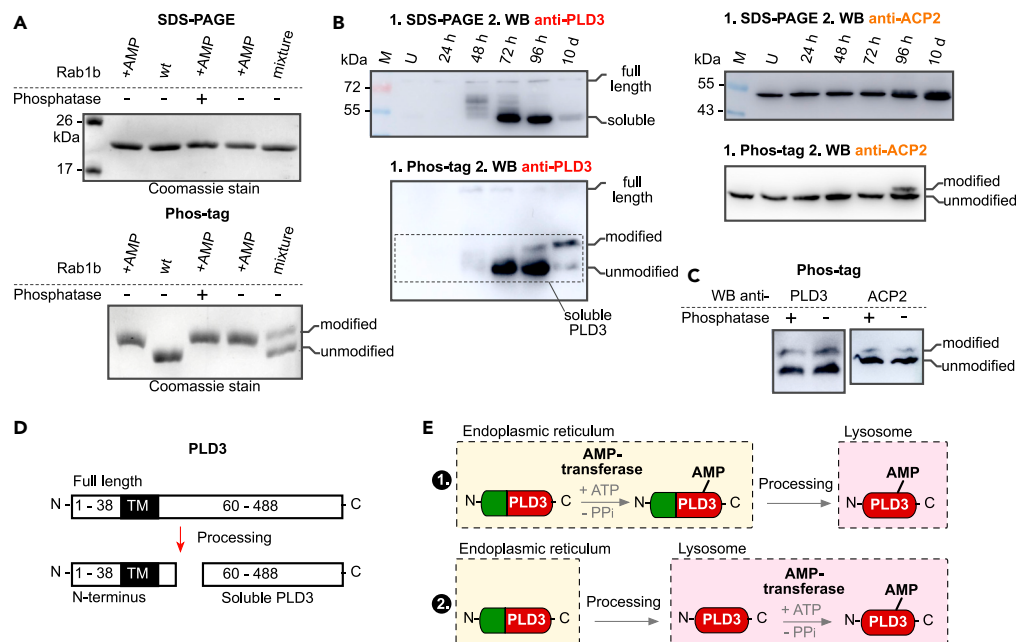
**Figure 2. Chemical proteomics of neuronal differentiation and maturation shows specific patterns of protein AMPylation**

(A) Schematic iNGNs differentiation procedure and pro-N6pA probe treatment.

(B) Chemical proteomic protocol comparing the probe and DMSO control-treated cells.

(C) Profile plots of significantly enriched proteins ACP2, ABHD6, PLD3, and CTSC from pro-N6pA probe-treated iNGNs. Blue boxes represent the LFQ-intensities from pro-N6pA probe-treated cells after enrichment. Gray boxes represent LFQ-intensities from control (DMSO-treated) cells, showing unspecific binding to avidin-coated agarose beads ( $n = 4$ , the box is defined by 25<sup>th</sup> and 75<sup>th</sup> percentile, whiskers show outliers, line is a median, and circle is a mean).

S7 and S8). The examination of the profile plots of the enriched proteins shows a distinct pattern of AMPylation dynamics during neuronal maturation (Figure 2C). The most distinct composition was observed for the two lysosomal proteins ACP2 and ABHD6, with a linear increase of probe incorporation during the course of differentiation, while the total expression level of both proteins remains stable during the process (Figures 2C and S2). On the other hand, other lysosomal significantly enriched proteins show a stable AMPylation level, for example, CTSC. A similar increase during iNGNs differentiation was observed for the cytosolic protein ATP-dependent 6-phosphofructokinase PFKP, a gatekeeper of glycolysis (Figure S2). Next, we focused on the lysosomal protein PLD3, which was recently associated with Alzheimer's disease, but its physiological function in neurons and regulation have been so far controversial (Arranz et al., 2017; Cruchaga et al., 2014). As shown previously, whole proteome analysis of PLD3 exhibits an increase of PLD3 in neurons compared with undifferentiated iNGNs. PLD3 is known to be transported from the ER and Golgi via endosomes to lysosomes, where the cytosolic N-terminal membrane-bound domain of full-length PLD3 is proteolytically cleaved. The resulting soluble luminal PLD3 containing the putative active site is delivered to the lysosomes (Gonzalez et al., 2018). Both chemical proteomics and whole proteome analysis cannot provide the information on which form of PLD3 is likely to be modified by AMPylation. Thus, we focused on the development of a gel-based methodology that would allow the separation of different protein forms as well as the quantification of the protein AMPylation levels. Previously, isoelectric focusing gels have been



**Figure 3. Phos-tag ligand SDS-PAGE distinguishes unmodified and AMPylated proteins**

(A) Coomassie-stained SDS-PAGE and Phos-tag ligand containing SDS-PAGE gel separation of unmodified (wt) and AMPylated (+AMP) recombinant Rab1b. With and without treatment with the shrimp alkaline phosphatase ( $\pm$ ). (B) PLD3 and ACP2 Phos-tag ligand SDS-PAGE separation and western blotting during iNGNs differentiation and maturation. Visualization using the anti-PLD3 and anti-ACP2 antibodies. (C) Phos-tag SDS-PAGE analysis of the PLD3 and ACP2 PTM status after phosphatase treatment. (D) Processing of the full-length PLD3 into soluble active PLD3. (E) Scheme showing the two different scenarios of putative intracellular trafficking and AMPylation of the PLD3. M stands for protein marker.

used for the separation of AMPylated HSPA5 from its unmodified form, but the separation of the two species appears to be rather limited (Preissler et al., 2015). Therefore, we explored the possibility to use another gel-based method, which also takes advantage of the presence of a phosphate moiety, but utilize the Phos-tag ligand (Kinoshita et al., 2004, 2006).

### Phos-tag ligand-based polyacrylamide gel electrophoresis separates the modified and unmodified proteins

In order to validate our results from chemical proteomics experiments by a complementary method, which would not rely on the pro-N6pA probe and mass spectrometry, we have concentrated on the development of a gel-based approach. Therefore, we have tested the possibility to capture the monophosphate moiety of AMPylated proteins using an alkoxide-bridge  $Mn^{2+}$  metal complex bound in the gel by Phos-tag ligand. We have used a recombinant Rab1b protein, which was AMPylated *in vitro* by the recombinant bacterial effector DrrA (Du et al., 2021). The Rab1b identities were confirmed by top-down mass spectrometry (Figure S4). To prepare a Phos-Tag gel, conventional sodium dodecyl sulfate (SDS)-polyacrylamide gels (SDS-PAGEs) were supplemented with the commercially available  $Mn^{2+}$  ion-binding Phos-tag ligand. Direct comparison of non-AMPylated, AMPylated, and 1:1 mixture of both modified and unmodified Rab1b showed that indeed the Phos-tag ligand added to standard SDS-PAGE resolves the two species and yields two clearly separated bands as visualized by Coomassie staining (Figure 3A). A control experiment using SDS-PAGE without addition of the Phos-tag ligand showed no separation of the two species (Figure 3A). Because the Phos-tag ligand was initially developed for separation of the phosphorylated proteins, we have treated the *in vitro* AMPylated Rab1b with shrimp alkaline phosphatase to exclude the potential separation due to protein phosphorylation and to show the resistance of the AMP moiety against the phosphatase cleavage (Figure 3A). The activity of the phosphatase was verified by hydrolysis of the phosphorylated ovalbumin (Figure S5). Next, we examined whether it is possible to separate AMPylated HSPA5 from its



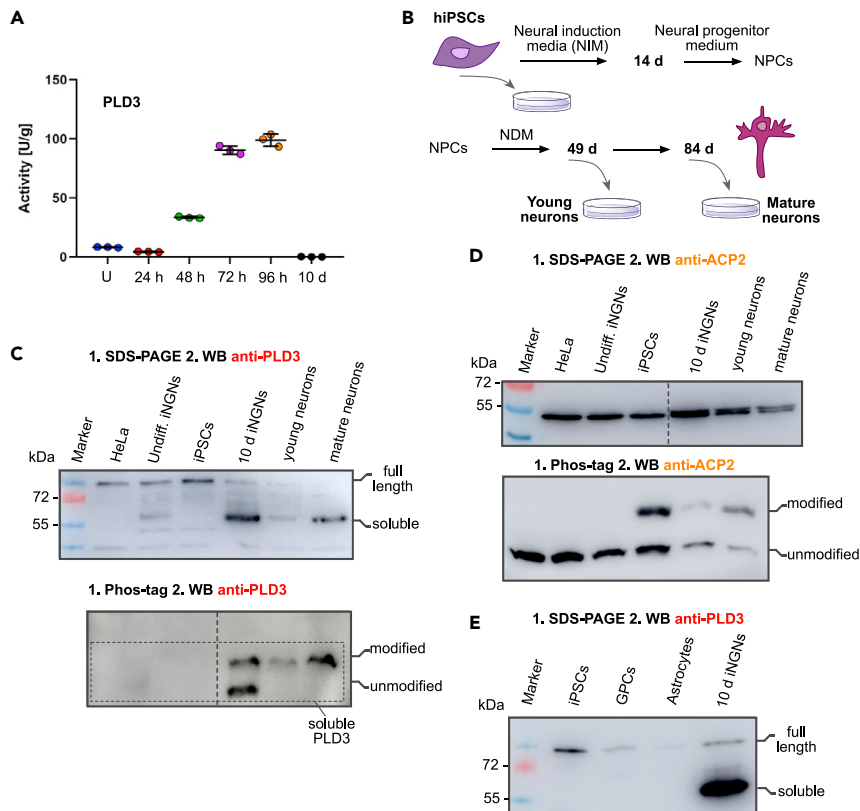
unmodified form in HeLa cell lysates. Therefore, after separation the Phos-Tag gel was blotted onto a PVDF membrane and visualized by staining with an anti-HSPA5 antibody. We observed a clear separation of the two forms, confirming our hypothesis that the Phos-tag ligand can be used for the separation of the AMPylated proteins in lysates (Figure S6). Hence, we have established the Phos-tag ligand-functionalized SDS-PAGE separation as a useful method for analysis of AMPylated proteins, which overcomes the necessity for treatment with the probe, and thus it can be used for analysis of protein AMPylation from a wider range of sources, for example, from animal tissues. In our chemical proteomics experiment, we have identified several AMPylated lysosomal proteins, including PLD3 and ACP2. PLD3 can exist either as the full-length protein or as the active soluble form localized in lysosomes, which is obtained by cleavage of the N terminus containing the transmembrane domain (Figure 3D) (Gonzalez et al., 2018). To obtain better insight into the changes of full-length, soluble, and modified PLD3 during neuronal differentiation, we performed the Phos-tag gel-based separation from undifferentiated iNGNs, and iNGNs differentiated between 1 and 10 days. To our surprise, we observed the modified soluble form of PLD3 only after 4 days of differentiation, with a majority of modified PLD3 in 10-day differentiated neurons, which is in line with our observation of increasing pro-N6pA labeling of PLD3 from the chemical proteomics study (Figures 3B and S7). Moreover, analysis by standard SDS-PAGE showed that the soluble lysosomal PLD3 form increases dramatically with increasing time of iNGN differentiation (up to 72 h), whereas it drops substantially upon maturation of iNGN neurons between 4 and 10 days post induction (Figure 3B). In addition to PLD3, a similar AMPylation trend was observed for another lysosomal protein, ACP2, further corroborating the chemical proteomic results and pointing to a specific function of probable protein AMPylation in neuronal maturation (Figure 3B) (Makrypidi et al., 2012). To exclude the separation due to the protein phosphorylation, lysates from the 10-day differentiated iNGNs were treated with the shrimp alkaline phosphatase (Figure 3C). Taken together, the combination of standard SDS-PAGE and Phos-tag-based gel electrophoresis enabled a detailed characterization of PLD3 post-translational processing and AMPylation dynamics during neuronal differentiation and maturation. More specifically, the amount of active soluble PLD3 increases during the differentiation process (up to 72), whereas its AMPylation raises during maturation of iNGN neurons (96 h–10 d). However, the identity and localization of the AMP transferase catalyzing the PLD3 AMPylation remains unknown, thus leaving two plausible scenarios of its AMPylation processing and trafficking into lysosomes (Figure 3E). Even though both methods correlate and point that the observed modification is indeed AMPylation, the exact identity of the PTM has not been confirmed by a direct mass spectrometry experiment. We have attempted to identify the AMPylation site of PLD3. For this, we enriched PLD3 by immunoprecipitation before digestion with trypsin and LC-MS/MS analysis. However, perhaps due to the low sequence coverage (53%) it did not lead to identification of the modification site. In order to specifically enrich only the modified peptides from lysates, we performed enrichment using the desthiobiotin linker, which allows one to selectively elute modified peptides after the enrichment and tryptic digestion (Zanon et al., 2020). Unfortunately, spectra analysis did not lead to identification of either PLD3 or any other site including T518 on the HSPA5. This is likely caused by the low ionization efficiency, fragmentation of the adenosine, and the low abundance of AMPylated peptides.

### PTM status correlates with the PLD3 activity in iNGN neurons

To correlate the catalytic activity of PLD3 with its PTM status during the course of iNGNs differentiation, a PLD3-specific acid 5' exonuclease activity assay was carried out (Cappel et al., 2021). Therefore, whole-cell lysates of iNGN cells were incubated with a fluorophore- and quencher-coupled oligodesoxynucleotide. A 5' exonucleolytic digest thereby led to a proportional increase in fluorescence signal, measured kinetically over 12 h. The overall activity of PLD3 per total cellular protein increases during the differentiation, which can be accounted for by the increasing levels of its soluble form in lysosomes. In contrast, the PLD3 activity decreases significantly from the 4th day to the 10th day after induction of the differentiation, coinciding with increasing AMPylation of the soluble form during neuronal maturation (Figure 4A). This observation suggests that AMPylation might inhibit PLD3's catalytic activity, which is in line with previous reports on inhibition of the chaperon activity of HSPA5 and the peptidase activity of CTSB.

### Directed differentiation of physiological human iPSCs shows the modified PLD3 as the only soluble form in young and mature neurons

In order to compare the AMPylation pattern obtained from the highly homogeneous and robust differentiation and maturation of iNGN cells we have collected lysates from physiological iPSCs, young and mature neurons differentiated for 5 and 10 weeks, respectively (Figures 4B and S8) (Gunhanlar et al., 2018). The modified, likely AMPylated soluble form was the only PLD3 species detectable by Phos-tag SDS-PAGE



**Figure 4. Modification of PLD3 correlates with its activity in neurons**

(A) PLD3 5' exonuclease activity assay correlates the amount of the soluble PLD3 form with its activity and shows the modification to inhibit the catalytic activity ( $n = 3$ , line indicates the mean, error bars show the standard deviation). (B) Directed differentiation protocol of human iPSCs into dopaminergic neurons. NDM, neural differentiation medium. (C) SDS-PAGE and Phos-tag ligand SDS-PAGE separations followed by western blot using the anti-PLD3 antibody show pronounced and nearly quantitative modification, likely AMPylation of PLD3 in dopaminergic neurons compared with the iNGN forward reprogramming. (D) ACP2 SDS-PAGE and Phos-tag ligand SDS-PAGE separations followed by western blot using anti-ACP2 antibody. (E) SDS-PAGE followed by western blot using the anti-PLD3 antibody detects solely the full-length PLD3 in iPSCs, GPCs, and astrocytes.

in young as well as in mature neurons, whereas in undifferentiated iPSCs the anti-PLD3 antibody detected no soluble form, whether unmodified or modified (Figures 4C and S9). Comparison of the PLD3 pattern in 10-day differentiated iNGNs, young and mature physiological neurons, with that in lysates from HeLa cells and undifferentiated iPSCs and iNGNs shows that the soluble forms, whether unmodified or AMPylated, are specific to neurons. Moreover, the analysis of PLD3 processing and AMPylation during differentiation of iPSCs into glial progenitor cells (GPCs) and astrocytes confirmed the specificity of this process for the neuronal cell lineage (Figure 4E, see Figure S9 for Phos-tag SDS-PAGE analysis). A similar degree of ACP2 modification was seen in 10-day differentiated iNGN and physiologically mature neural networks (Figure 4D). Taken together, these findings further corroborate the hypothesis that AMPylation plays a specific role in maturation of human neurons.

## DISCUSSION

Lysosomal dysfunction is linked to several human pathologies such as LSDs, cancer, neurodegeneration, and aging (Marques and Saftig, 2019). Neuronal cells are particularly sensitive to impaired lysosomal function owing to their tightly controlled differentiation process and postmitotic character. Even though the genetic basis and the biochemistry underlying these diseases are known, the cellular and molecular mechanisms leading to disruption of neuronal viability remain to be understood. Our study describes lysosomal

proteins to be significantly enriched among the AMPylated proteins with changing stoichiometry at various stages of neuronal differentiation, pointing to a specific function of protein AMPylation during differentiation and maturation. The fine-tuning of the lysosomal activity was recently reported to be critical for maintenance of the quiescent neural stem cells fitness and their activation responsiveness (Leeman et al., 2018). Thus, the precise orchestration of these processes might be achieved by adding an extra regulatory layer of protein PTMs, including AMPylation. Furthermore, comparison of the expression levels of the identified AMPylated proteins during the development of mouse and human embryos as well as human organoids shows clear differences between the species and model systems pointing to a regulation of the proteins at different levels and stages of gestation (Figure S10) (Klingler et al., 2021). The gel-based analyses of PLD3 showed a striking increase of the modified soluble form in differentiated iNGN neurons. Comparing the iNGN forward reprogramming model of neuronal differentiation with differentiation of physiological human iPSCs into young and mature neurons showed an even stronger trend with fully modified PLD3 in mature physiological neurons. Prolonging the neurogenic period gives rise to more neurons and expansion of the upper layer of neocortex (Stepien et al., 2020). Thus, lengthening the maturation period of the physiological neurons in comparison with the fast differentiation of iNGN cells may provide more time to establish a more strictly defined modification status of PLD3. In particular, the increased expression of PLD3 was shown to coincide with late neuronal development in the hippocampus and the primary somatosensory cortex (Pedersen et al., 1998). In consequence, the PLD3 AMPylation may modulate PLD3 activity or stability to ensure the proper lysosomal function during the migration and maturation of the basal progenitors from the cortical subventricular zone. The activity assay of PLD3 revealed that the probable AMPylation status correlates with its activity in iNGN neurons. Even though we have characterized the progress of the PLD3 AMPylation in detail, it remains to be elucidated where the AMPylation of PLD3 occurs and which AMP-transferase catalyzes the attachment of this modification (Figure 3E) or if it can be reversed by active deAMPylation. The heterogeneity of the cell types in the human cortex is enormous, including excitatory and inhibitory neurons, astrocytes, microglia, and oligodendrocytes. Analysis of the expression levels of the lysosomal proteins ACP2, ABHD6, PLD3, and CTSC in these cell types exhibits substantial differences (Figure S11) (Kanton et al., 2019). Together with considerable variance in the PTM status, as we have demonstrated for PLD3 in astrocytes and neuronal cell lineage, it highlights the diversity and specificity of the proteoforms in the central nervous system. Therefore, in future studies, it will be important to compare the AMPylation status of these proteins in different cell types and to elucidate the functional consequences. Our study provides evidence that AMPylation is likely a rare modification of the lysosomal proteins and thus gives insight into a putative mechanism modulating the protein activity in this critical cellular department. This finding may open up new possibilities for designing therapeutic strategies against the damage linked to lysosomal dysfunction during neurodevelopment, cancer, and aging.

### Limitations of the study

The combination of the two methods, MS-based chemical proteomics and Phos-tag ligand gel analysis, strongly supports the AMPylation of several lysosomal proteins; however, a direct proof for AMPylation is missing in this study. The main bottleneck in using mass spectrometry for validation is the inherent instability of the adenosine monophosphate moiety and the low abundance of the modified peptides. In cells, the pro-N6pA probe is converted to the corresponding N<sup>6</sup>-propargyl ATP, which might be further metabolized to N<sup>6</sup>-propargyl NAD<sup>+</sup> and incorporated onto proteins as ADP-ribosylation. Although we have shown in our previous study (Kielkowski et al., 2020a) that the amounts are negligible, it cannot be excluded under different experimental conditions. Together, the protein PTM status of lysosomal proteins observed in the chemical proteomic experiments correlates well with the Phos-tag ligand SDS-PAGE analysis. Nevertheless, the limitation of the Phos-tag gel separation is that it cannot distinguish between different NMPylations such as UMPylation or GMPylation. Further development of mass spectrometry experiments should determine the identity of the PTM.

### STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- [KEY RESOURCES TABLE](#)
- [RESOURCE AVAILABILITY](#)
  - Lead contact
  - Materials availability
  - Data and code availability

**● EXPERIMENTAL MODEL AND SUBJECT DETAILS**

- iNGN cells
- Generation of NPCs
- Generation of neurons
- iPSC culture for astrocytes differentiation
- Generation of GPCs and astrocytes
- Culturing of HeLa and SH-SY5Y cells

**● METHOD DETAILS**

- Fluorescence imaging
- Immunohistochemistry – physiological neurons
- PLD3 activity assay
- Chemical proteomics

**● QUANTIFICATION AND STATISTICAL ANALYSIS**

- Calculation chemical proteomics
- Statistical analysis chemical proteomics
- Profile plots

**SUPPLEMENTAL INFORMATION**

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2021.103521>.

**ACKNOWLEDGMENTS**

This research was supported by the Liebig fellowship from Fund of the association of the chemical industry (VCI) to P.K. and T.B., LMUexcellent Junior Fund to P.K., and Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – SFB 1309. We are grateful for the kind gift of recombinant Rab1b wt and AMPylated form from the group of Dr. Sabine Schneider by Dr. Marie-Kristin von Wisberg.

**AUTHOR CONTRIBUTIONS**

P.K., M.D., and T.B. conceived the study. T.B. cultured and differentiated the iNGN cells. T.B. carried out all proteomic experiments and gel-based analyses. C.C. performed the fluorescence imaging and PLD3 activity assay. F.D.M. cultured the physiological hiPSCs and differentiated them into neurons. G.S. cultured the hiPSCs and differentiated them into GPCs and astrocytes. E.K. and F.S. helped to establish the iNGN cell culture and reviewed the manuscript. P.K. and T.B. wrote the manuscript. All authors revised the manuscript.

**DECLARATION OF INTERESTS**

The authors declare no competing interests.

Received: March 10, 2021

Revised: July 20, 2021

Accepted: November 23, 2021

Published: December 17, 2021

**REFERENCES**

- Aebersold, R., Agar, J.N., Amster, I.J., Baker, M.S., Bertozzi, C.R., Boja, E.S., Costello, C.E., Cravatt, B.F., Fenselau, C., Garcia, B.A., et al. (2018). How many human proteoforms are there? *Nat. Chem. Biol.* 14, 206–214.
- Arranz, A.M., Strooper, B., Frigerio, C., Horre, K., Fazzari, P., Saido, T.C., and Saito, T. (2017). PLD3 gene and processing of APP. *Nature* 541, E1.
- Ayo-Martin, A.C., Kyrousi, C., Giaimo, R.D., and Cappello, S. (2020). GNG5 controls the number of apical and basal progenitors and alters neuronal migration during cortical development. *Front. Mol. Biosci.* 7, 578137.
- Biedler, J.L., Helson, L., and Spengler, B.A. (1973). Morphology and growth, tumorigenicity, and cytogenetics of human neuroblastoma cells in continuous culture. *Cancer Res.* 33, 2643–2652.
- Boyer, L.F., Campbell, B., Larkin, S., Mu, Y., and Gage, F.H. (2012). *Current Protocols in Stem Cell Biology* (Wiley), pp. 1H.6.1–1H.6.11.
- Broncel, M., Serwa, R.A., and Tate, E.W. (2012). A new chemical handle for protein AMPylation at the host–pathogen interface. *ChemBioChem* 13, 183–185.
- Broncel, M., Serwa, R.A., Bunney, T.D., Katan, M., and Tate, E.W. (2016). Global profiling of huntingtin-associated protein E (HYPE)-Mediated AMPylation through a chemical proteomic approach. *Mol. Cell Proteomics* 15, 715–725.
- Busskamp, V., Lewis, N.E., Guye, P., Ng, A.H., Shipman, S.L., Byrne, S.M., Sanjana, N.E., Murn, J., Li, Y., Li, S., et al. (2014). Rapid neurogenesis through transcriptional activation in human stem cells. *Mol. Syst. Biol.* 10, 760.
- Cappel, C., Gonzalez, A.C., and Damme, M. (2021). Quantification and characterization of the 5' exonuclease activity of the lysosomal nuclease PLD3 by a novel cell-based assay. *J. Biol. Chem.* 296, 100152.

- Casey, A.K., and Orth, K. (2017). Enzymes involved in AMPylation and deAMPylation. *Chem. Rev.* 118, 1199–1215.
- Cox, J., Hein, M.Y., Lubner, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell Proteomics* 13, 2513–2526.
- Cruchaga, C., Karch, C.M., Jin, S., Benitez, B.A., Cai, Y., Guerreiro, R., Harari, O., Norton, J., Budde, J., Bertelsen, S., et al. (2014). Rare coding variants in the phospholipase D3 gene confer risk for Alzheimer's disease. *Nature* 505, 550.
- Du, J., Wrisberg, M.-K., Gulen, B., Stahl, M., Pett, C., Hedberg, C., Lang, K., Schneider, S., and Itzen, A. (2021). Rab1-AMPylation by Legionella DrrA is allosterically activated by Rab1. *Nat. Commun.* 12, 460.
- Engel, P., Goepfert, A., Stanger, F.V., Harms, A., Schmidt, A., Schirmer, T., and Dehio, C. (2012). Adenylylation control by intra- or intermolecular active-site obstruction in Fic proteins. *Nature* 482, 107–110.
- Gavin, A.L., Huang, D., Blane, T.R., Thinnis, T.C., Murakami, Y., Fukui, R., Miyake, K., and Nemazee, D. (2021). Cleavage of DNA and RNA by PLD3 and PLD4 limits autoinflammatory triggering by multiple sensors. *Nat. Commun.* 12, 5874.
- Gonzalez, A.C., Schweizer, M., Jagdmann, S., Bernreuther, C., Reinheckel, T., Saftig, P., and Damme, M. (2018). Unconventional trafficking of mammalian phospholipase D3 to lysosomes. *Cell Rep.* 22, 1040–1053.
- Grammel, M., Luong, P., Orth, K., and Hang, H.C. (2011). A chemical reporter for protein AMPylation. *J. Am. Chem. Soc.* 133, 17103–17105.
- Gunhanlar, N., Shpak, G., van der Kroeg, M., Gouty-Colomer, L.A., Munshi, S.T., Lendemeijer, B., Ghazvini, M., Dupont, C., Hoogendijk, W.J.G., Gribnau, J., et al. (2018). A simplified protocol for differentiation of electrophysiologically mature neuronal networks from human induced pluripotent stem cells. *Mol. Psychiatry* 23, 1336–1344.
- Ham, H., Woolery, A.R., Tracy, C., Stenesen, D., Krämer, H., and Orth, K. (2014). Unfolded protein response-regulated Drosophila Fic (dFic) protein reversibly AMPylates BiP chaperone during endoplasmic reticulum homeostasis. *J. Biol. Chem.* 289, 36059–36069.
- Hipp, M.S., Kasturi, P., and Hartl, F.U. (2019). The proteostasis network and its decline in ageing. *Nat. Rev. Mol. Cell Biol.* 20, 421–435.
- Kanton, S., Boyle, M.J., He, Z., Santel, M., Weigert, A., Sanchis-Calleja, F., Guijarro, P., Sidow, L., Fleck, J.S., Han, D., et al. (2019). Organoid single-cell genomic atlas uncovers human-specific features of brain development. *Nature* 574, 418–422.
- Kielkowski, P., Buchsbaum, I.Y., Kirsch, V.C., Bach, N.C., Drukker, M., Cappello, S., and Sieber, S.A. (2020a). FICD activity and AMPylation remodel the endoplasmic reticulum to modulate human neurogenesis. *Nat. Commun.* 11, 517.
- Kielkowski, P., Buchsbaum, I.Y., Becker, T., Bach, K., Cappello, S., and Sieber, S.A. (2020b). A pronucleotide probe for live-cell imaging of protein AMPylation. *ChemBiochem* 21, 1285–1287.
- Kingdon, H., Shapiro, B., and Stadtman, E. (1967). Regulation of glutamine synthetase. 8. ATP: glutamine synthetase adenylyltransferase, an enzyme that catalyzes alterations in the regulatory properties of glutamine synthetase. *Proc. Natl. Acad. Sci. U S A* 58, 1703–1710.
- Kinoshita, E., Takahashi, M., Takeda, H., Shiro, M., and Koike, T. (2004). Recognition of phosphate monoester dianion by an alkoxide-bridged dinuclear zinc(II) complex. *Dalton Trans* 21, 1189–1193.
- Kinoshita, E., Kinoshita-Kikuta, E., Takiyama, K., and Koike, T. (2006). Phosphate-binding tag, a new tool to visualize phosphorylated proteins. *Mol. Cell Proteomics* 5, 749–757.
- Klaus, J., Kanton, S., Kyrousi, C., Ayo-Martin, A.C., Giaimo, R.D., Riesenberger, S., O'Neill, A.C., Camp, J.G., Tocco, C., Santel, M., et al. (2019). Altered neuronal migratory trajectories in human cerebral organoids derived from individuals with neuronal heterotopia. *Nat. Med.* 25, 561–568.
- Klingler, E., Francis, F., Jabaudon, D., and Cappello, S. (2021). Mapping the molecular and cellular complexity of cortical malformations. *Science* 371, eaba4517.
- Leeman, D.S., Hebestreit, K., Ruetz, T., Webb, A.E., McKay, A., Pollina, E.A., Dulken, B.W., Zhao, X., Yeo, R.W., Ho, T.T., et al. (2018). Lysosome activation clears aggregates and enhances quiescent neural stem cell activation during aging. *Science* 359, 1277–1283.
- Makrypidi, G., Damme, M., Müller-Loennies, S., Trusch, M., Schmidt, B., Schlüter, H., Heeren, J., Lübke, T., Saftig, P., and Braulke, T. (2012). Mannose 6 dephosphorylation of lysosomal proteins mediated by acid phosphatases Acp2 and Acp5. *Mol. Cell Biol.* 32, 774–782.
- Marques, A.R.A., and Saftig, P. (2019). Lysosomal storage disorders – challenges, concepts and avenues for therapy: beyond rare diseases. *J. Cell Sci.* 132, jcs221739.
- Pedersen, K.M., Finsen, B., Celis, J.E., and Jensen, N.A. (1998). Expression of a novel murine phospholipase D homolog coincides with late neuronal development in the forebrain. *J. Biol. Chem.* 273, 31494–31504.
- Pieles, K., Glatter, T., Harms, A., Schmidt, A., and Dehio, C. (2014). An experimental strategy for the identification of AMPylation targets from complex protein samples. *Proteomics* 14, 1048–1052.
- Preissler, S., Rato, C., Chen, R., Antrobus, R., Ding, S., Fearnley, I.M., and Ron, D. (2015). AMPylation matches BiP activity to client protein load in the endoplasmic reticulum. *ELife* 4, e12621.
- Preissler, S., Rato, C., Perera, L., Saudek, V., and Ron, D. (2017). FICD acts bifunctionally to AMPylate and de-AMPylate the endoplasmic reticulum chaperone BiP. *Nat. Struct. Mol. Biol.* 24, 23–29.
- Puck, T.T., and Marcus, P.I. (1955). A rapid method for viable cell titration and clone production with HeLa cells in tissue culture: the use of X-irradiated cells to supply conditioning factors. *Proc. Natl. Acad. Sci. U S A* 41, 432–437.
- Rahman, M., Ham, H., Liu, X., Sugiura, Y., Orth, K., and Krämer, H. (2012). Visual neurotransmission in Drosophila requires expression of Fic in glial capitate projections. *Nat. Neurosci.* 15, 871–875.
- Rauh, T., Brameyer, S., Kielkowski, P., Jung, K., and Sieber, S.A. (2020). MS-based in situ proteomics reveals AMPylation of host proteins during bacterial infection. *ACS Infect Dis.* 6, 3277–3289.
- Santos, R., Vadodaria, K.C., Jaeger, B.N., Mei, A., Lefcochilos-Fogelquist, S., Mendes, A., Erikson, G., Shokhirev, M., Randolph-Moore, L., Fredlender, C., et al. (2017). Differentiation of inflammation-responsive astrocytes from glial progenitors generated from human induced pluripotent stem cells. *Stem Cell Rep.* 8, 1757–1769.
- Sanyal, A., Chen, A.J., Nakayasu, E.S., Lazar, C.S., Zbornik, E.A., Worby, C.A., Koller, A., and Mattoo, S. (2015). A novel link between fic (filamentation induced by cAMP)-mediated adenylylation/AMPylation and the unfolded protein response. *J. Biol. Chem.* 290, 8482–8499.
- Sanyal, A., Dutta, S., Camara, A., Chandran, A., Koller, A., Watson, B.G., Sengupta, R., Ysselstein, D., Montenegro, P., Cannon, J., et al. (2019). Alpha-synuclein is a target of fic-mediated adenylylation/AMPylation: possible implications for Parkinson's disease. *J. Mol. Biol.* 431, 2266–2282.
- Schultz, M.L., Tecedor, L., Chang, M., and Davidson, B.L. (2011). Clarifying lysosomal storage diseases. *Trends Neurosci.* 34, 401–410.
- Sengupta, R., Poderycki, M.J., and Mattoo, S. (2019). CryoAPEX - an electron tomography tool for subcellular localization of membrane proteins. *J. Cell Sci.* 132, jcs222315.
- Sieber, S.A., Cappello, S., and Kielkowski, P. (2020). From young to old: AMPylation hits the brain. *Cell Chem. Biol.* 27, 773–779.
- Sreelatha, A., Yee, S.S., Lopez, V.A., Park, B.C., Kinch, L.N., Pilch, S., Servage, K.A., Zhang, J., Jiou, J., Karasiewicz-Urbanska, M., et al. (2018). Protein AMPylation by an evolutionarily conserved pseudokinase. *Cell* 175, 809–821.e19.
- Stadtmann, J., Taubenschmid, J., Wenzel, D., Gatterer, A., Dürnberger, G., Dusberger, F., Elling, U., Mach, L., Mechtler, K., and Penninger, J.M. (2017). Comparative glycoproteomics of stem cells identifies new players in ricin toxicity. *Nature* 549, 538–542.
- Stepien, B.K., Naumann, R., Holtz, A., Helppi, J., Huttner, W.B., and Vaid, S. (2020). Lengthening neurogenic period during neocortical development causes a hallmark of neocortex expansion. *Curr Biol* 30, 4227–4237.e5. <https://doi.org/10.1016/j.cub.2020.08.046>.
- Stoka, V., Turk, V., and Turk, B. (2016). Lysosomal cathepsins and their regulation in aging and neurodegeneration. *Ageing Res. Rev.* 32, 22–37.
- Truttmann, M.C., Pincus, D., and Ploegh, H.L. (2018). Chaperone AMPylation modulates aggregation and toxicity of neurodegenerative

disease-associated polypeptides. *Proc. Natl. Acad. Sci. U S A* 115, E5008–E5017.

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* 13, 731–740.

Yarbrough, M.L., Li, Y., Kinch, L.N., Grishin, N.V., Ball, H.L., and Orth, K. (2009). AMPylation of rho GTPases by *Vibrio* VopS disrupts effector binding and downstream signaling. *Science* 323, 269–272.

Yu, X., and LaBaer, J. (2015). High-throughput identification of proteins with AMPylation using

self-assembled human protein (NAPPA) microarrays. *Nat. Protoc.* 10, 756–767.

Zanon, P.R.A., Lewald, L., and Hacker, S.M. (2020). Isotopically labeled desthiobiotin azide (isoDTB) tags enable global profiling of the bacterial cysteinome. *Angew. Chem. Int. Ed.* 59, 2829–2836.

## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Antibodies</b>		
Goat polyclonal anti-mouse IgG, AF488-linked	Thermo Fisher Scientific	Cat# A-11001; RRID:AB_2534069
Goat polyclonal anti-rabbit IgG, AF488-linked	Thermo Fisher Scientific	Cat# A-11008; RRID: AB_143165
Goat polyclonal anti-rabbit IgG, HRP-linked	Sigma-Aldrich	Cat# A6667; RRID: AB_258307
Rabbit polyclonal anti-ABHD6	Thermo Fisher Scientific	Cat# PA5-38999; RRID AB_2555591
Rabbit polyclonal anti-ACP2	Thermo Fisher Scientific	Cat# PA5-29961; RRID: AB_2547435
Rabbit polyclonal anti-COX IV	Thermo Fisher Scientific	Cat# ab16056; RRID: AB_443304
Rabbit polyclonal anti-HSPA5	Thermo Fisher Scientific	Cat# PA5-34941; RRID: AB_2552290
Mouse monoclonal anti-KDEL (10C3)	Millipore	Cat# 420400; RRID: AB_212090
Mouse monoclonal anti-LAMP2 (H4B4)	DSHB	Cat# N/A; RRID: <a href="#">AB_528129</a>
Rabbit polyclonal anti-PLD3	Sigma-Aldrich	Cat# HPA012800; RRID: AB_1855330
Mouse monoclonal anti-TUBB3	Sigma-Aldrich	Cat# T8660 RRID: AB_477590
Guinea pig polyclonal anti-DCX	Millipore	Cat# AB2253 RRID: AB_1586992
Rabbit polyclonal anti-TBR1	Millipore	Cat# AB31940 RRID: AB_2200219
Mouse polyclonal anti-MAP2	Millipore	Cat# AB5392 RRID: AB_2138153
Mouse monoclonal anti-Neun	Millipore	Cat# MAB377 RRID: AB_2298772
Goat polyclonal anti-mouse IgG2b, AF647-linked	Thermo Fisher Scientific	Cat# A-21242 RRID: AB_2535811
Goat polyclonal anti-rabbit IgG, AF546-linked	Thermo Fisher Scientific	Cat# A-11010 RRID: AB_2534077
Goat polyclonal anti-Guinea pig IgG, AF647-linked	Thermo Fisher Scientific	Cat# A-21450 RRID: AB_2735091
Goat polyclonal anti-mouse IgG, AF546-linked	Thermo Fisher Scientific	Cat# A-21123 RRID: AB_2535765
Mouse monoclonal anti-S100beta	Sigma-Aldrich	Cat# S2532 RRID: AB_477499
Rabbit polyclonal anti-FGFR3	Santa Cruz	Cat# sc-123 RRID: AB_631511
Goat anti-mouse IgG1, Alexa fluor 488	Thermo Fisher Scientific	Cat# A-21121 RRID: AB_2535764
Rabbit polyclonal anti-Ovalbumin	Thermo Fisher Scientific	Cat# PA5-97525 RRID: AB_2812141
<b>Chemicals, peptides, and recombinant proteins</b>		
Acetone (HPLC grade)	VWR chemicals	Cat# 200067.320; CAS: 67-64-1
Acetonitrile (LC-MS grade)	Fisher Scientific	Cat# A955-212; CAS: 75-05-8
Alanyl-Glutamine	Sigma-Aldrich	Cat# G8541, CAS: 39537-23-0
Ammoniumperoxodisulfat (APS)	Sigma-Aldrich	Cat# 09913; CAS: 7727-54-0
β-Mercaptoethanol	Sigma-Aldrich	Cat# M3148; CAS: 60-24-2

(Continued on next page)

*Continued*

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biotin-PEG <sub>3</sub> -N <sub>3</sub>	Carbosynth	Cat# FA34890 CAS: 875770-34-6
Bromphenol blue	Fluka	Cat# 32712 CAS: 115-39-9
BSA	AppliChem	Cat# A6588; CAS: 9048-46-8
Coomassie Blue R-250	Fluka	Cat# 27816; CAS: 6104-59-2
CuSO <sub>4</sub> × 5 H <sub>2</sub> O	Acros	Cat# 10162, 7758-99-8
ddH <sub>2</sub> O (LC-MS grade)	Honeywell	Cat# 15350 CAS: 732-18-5
DMSO	Sigma-Aldrich	Cat# D4540, CAS: 67-68-5
Dithiothreitol	VWR (AppliChem)	Cat# A2948, CAS: 3483-12-3
Formic Acid (LC-MS grade)	Fisher Scientific	Cat# A117; CAS: 64-18-6
HEPES	Carl Roth	Cat# HN77.5; CAS: 7365-45-9
Iodacetamide	Sigma-Aldrich	Cat# I6125; CAS: 144-48-9
KCl	AppliChem	Cat# A2939; CAS: 7447-40-7
KH <sub>2</sub> PO <sub>4</sub>	Sigma-Aldrich	Cat# P9791; CAS: 7778-77-0
NA <sub>2</sub> HPO <sub>4</sub>	Sigma-Aldrich	Cat# T876; CAS: 7558-79-4
NA <sub>2</sub> SeO <sub>3</sub>	Sigma-Aldrich	Cat# S5261; CAS: 10102-18-8
NaCl	Bernd Kraft GmbH	Cat# KRAF04160; CAS: 7647-14-5
NP40	Sigma-Aldrich	Cat# 74385; CAS: 9016-45-9
Methanol (LC-MS grade)	Fisher Scientific	Cat# A456; CAS: 67-56-1
Powdered milk	AppliChem	Cat# A0830; CAS: 999999-99-4
SDS	AppliChem	Cat# A2572; CAS: 151-21-3
Sodium deoxycholate	Sigma-Aldrich	Cat# 30970; CAS: 302-95-4
TAMRA-N <sub>3</sub>	Baseclick	CAT# BCFA-008-1
TBTA	TCI	Cat# T2993; CAS: 510758-28-8
TCEP	Carbosynth	Cat# FT01756; CAS: 51805-45-9
TEAB (1 M)	Sigma-Aldrich	Cat# T7408, CAS: <a href="#">15715-58-9</a>
TEMED	Sigma-Aldrich	Cat# T9281; CAS: 110-18-9
Thiourea	Merck	Cat# 107979; CAS: 62-56-6
Tris-base	Fisher Scientific	Cat# 10344; CAS: 77-86-1
Trypan Blue	Fisher Scientific	Cat# 11886
Tween® 20	VWR (AppliChem)	Cat# A4974; CAS: 9005-64-5
Urea	AppliChem	Cat# A1049, CAS: 57-13-6
DPBS (1 ×)	Sigma-Aldrich	Cat# D8357
DMEM (1 ×)	Sigma-Aldrich	Cat# D6546
Ham's F-12 w/o L-Glu	Sigma-Aldrich	Cat# N4888
cOmplete® protease inhibitor	Sigma-Aldrich	Cat# 05001
Normal goat serum	Biozol	Cat# VEC-S-1000
FCS	Thermo Fisher Scientific	Cat# <a href="#">A3840001</a>
hHolo-transferrin	Sigma-Aldrich	Cat# 616424; CAS: 11096-37-0
hFGF-2	MACS Miltenyi Biotec	Cat# 130-104-921
hInsulin	BioXtra	Cat# I9278, CAS: 11061-68-0
hTGF- β1	MACS Miltenyi Biotec	Cat# 130-095-067
Geltrex	Thermo Fisher Scientific	Cat# A1413201
Immobilon® Western HRP substrate	Merck Millipore	Cat# WBKLS0500
Color prestained protein standard, Broad range (10-250 kDa)	New England BioLabs	Cat# P7719S
Blue prestained protein standard, Broad range (11-250 kDa)	New England BioLabs	Cat# P7718L

(Continued on next page)



**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Pen-strep	Sigma-Aldrich	Cat# P0781
Rotiphorese.Gel 30 (37, 5:1)	Carl Roth	Cat# 3029.1
TrypLE express	Thermo Fisher Scientific	Cat# 12013
Trypsin	Promega	Cat# V5113
Thiazovivin	Merck Millipore	Cat# 420220; CAS: 1056-71-8
NeuroBrew-21	Miltenyi Biotech	Cat# 130-093-566
4,6-Diamidino-2-phenylindole	Sigma-Aldrich	Cat# D9542
Accutase	Stem Cell Technologies	Cat# 07920
Ascorbic acid	Sigma-Aldrich	Cat# A92902; CAS: 50-81-7
B27-supplement (minus vitamin A)	Thermo Fisher Scientific	Cat# 12010
BDNF	Peprtech	Cat# 450-02
dcAMP	Sigma-Aldrich	Cat# D1256; CAS: 93839-95-3
GDNF	Peprtech	Cat# 450-10
Laminin	Sigma-Aldrich	Cat# L2020; CAS: 114,956-81-9
Matrigel® Basement membrane matrix, LDEV-free	Corning®	Cat# 354234
Minimum essential medium/non-essential amino acid	Thermo Fisher Scientific	Cat# 11050
mTESR1 medium	Stem Cell Technologies	Cat# 85850
N2-supplement	Thermo Fisher Scientific	Cat# 17048
Neurobasal medium	Thermo Fisher Scientific	Cat# 21103049
Polyornithine	Sigma-Aldrich	Cat# P4957
Rock inhibitor Y-27632(2HCl)	Stem Cell Technologies	Cat# 72304
Astrocyte media	ScienCell	Cat# 1801
Recombinant human LIF	Alomone Labs	Cat# L-200
Recombinant human FGF-basic	Peprtech	Cat# AF-100-18 B
Recombinant human EGF	Peprtech	Cat# AF-100-15
Recombinant human noggin	Peprtech	Cat# 120-10C
Recombinant human PDGF-AA	R&D Systems	Cat# 221-AA
Accutase	Thermo Fisher Scientific	Cat# A1110501
Collagenase	Stem Cell Technologies	Cat# 07909
<b>Critical commercial assays</b>		
Pierce® BCA protein Assay kit	Thermo Fisher Scientific	Cat# 23225
<b>Deposited data</b>		
MS raw data and calculation results	ProteomeXchange	PXD023873
<b>Experimental models: Cell lines</b>		
Human: HeLa	<a href="#">Puck and Marcus, 1955</a>	RRID: CVCL_0030
Human: SH-SY5Y	<a href="#">Biedler et al., 1973</a>	RRID: CVCL_0019
Human: iNGNs	Volker Busskamp, CRTD Dresden	
Human: iPSCs	Dr.Micha Drukker HHZ Munich	HMGU No 1 RRID: CVCL_YT30
<b>Software and algorithms</b>		
MaxQuant	<a href="#">Cox et al., 2014</a>	<a href="https://www.maxquant.org/download_asset/maxquant/latest">https://www.maxquant.org/download_asset/maxquant/latest</a>
Perseus	<a href="#">Tyanova et al., 2016</a>	<a href="https://maxquant.net/download_asset/perseus/latest">https://maxquant.net/download_asset/perseus/latest</a>
Origin	NA	<a href="https://www.originlab.com/">https://www.originlab.com/</a>

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Dr. Pavel Kielkowski ([pavel.kielkowski@cup.lmu.de](mailto:pavel.kielkowski@cup.lmu.de)).

### Materials availability

This study did not generate new unique reagents.

### Data and code availability

- Mass spectrometry-based proteomics data have been deposited at ProteomeXchange and are publicly available as of the date of publication. The accession number is listed in the [key resources table](#).
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

### iNGN cells

**Coating of petri dishes for iNGN culturing.** In order to culture iNGNs, petri dishes have to be coated. For this, 15 mg/mL Geltrex LDEV-free was diluted 1:1000 in cold coating media (49.5% DMEM (1×), 49.5% F-12, 1% Pen-Strep 100×) and was immediately added onto the petri dish. The coating volumes were as follows: p60, 3.5 mL; p100, 10 mL and p150, 25 mL. After addition of the coating media to the petri dishes, they were incubated at least one hour in the incubator at 37°C and were then ready for use.

**Passaging and culturing of iNGN cells in p100 dish.** First, media was removed and cells were washed with 4 mL PBS. Then, 1.5 mL TrypLE™ Express was added and the mixture was incubated 7 min at 37°C. The detached cells were resuspended with 2 mL pre-warmed E7 media (49% (v/v) DMEM, 49% (v/v) F-12, 1% (v/v) Alanine-Glutamine 100×, 77.6 nM Na<sub>2</sub>SeO<sub>3</sub>, 11.2 mM NaCl, 10 µg/mL hHolo-Transferrin, 10 µg/mL hInsulin) supplemented with 2 µM thiazovivin (E7+TZ) and the cell suspension was transferred to a 15 mL falcon. Before the cell suspension was spun down 5 min at 600 rpm, the concentration of the cell suspension was determined by mixing 10 µL cell suspension with 10 µL trypan blue. Afterwards, the supernatant was removed and the cell pellet was resuspended in the calculated amount of media to obtain a concentration of 1.5 million cells in 1 mL (For seeding iNGNs for differentiation the concentration was adjusted to 2.5 million cells in 1 mL). Next, the media of the previously coated p100 dish was removed and 9 mL E7+TZ media was added and topped with 1 mL of the cell suspension. Finally, the E7+TZ media was supplemented with 10 µL TGF-beta (2.0 µg/mL) and 10 µL FGF-1S (20 µg/mL) in order to obtain E9+TZ medium. The iNGN cells were cultured overnight at 37°C and 5% CO<sub>2</sub>, before the media was exchanged to E9 on the next day.

**Differentiation of iNGNs in p100 dish.** In order to differentiate iNGNs to neurons during 4 days, the expression of the two transcription factors Neurogenin-1 and Neurogenin-2 have to be induced by the addition of doxycycline (Dox). For this, 2.5 million cells in 10 mL E7 media containing 0.5 µg/mL Dox and 2 µM TZ (E7+Dox + TZ) were seeded in a p100 dish. After overnight incubation at 37°C in the incubator, the media was exchanged to 10 mL E7+Dox without TZ. On the next day, 10 mL fresh E7+Dox media was added onto the dish. On day 4, half of the E7+dox media was removed and 5 mL Neurobasal A media supplemented with 4% NS-21 was added. On day 5, the complete media was removed and 10 mL Neurobasal A media supplemented with 2% NS-21 and 2 mM L-alanyl-glutamine was added. On day 7, 2 mL of the media was removed and 2 mL Neurobasal A media supplemented with 2% NS-21 and 2 mM L-alanyl-glutamine was added. Until day 10, the iNGNs were incubated at 37°C in the incubator without any further media exchange.

**iPSC culture for physiological neurons differentiation.** iPSCs (RRID: CVCL\_YT30) were cultured as previously described in ([Ayo-Martin et al., 2020](#)). They were cultured in Matrigel Basement Membrane Matrix, LDEV-free coated plates in mTESR1 medium supplemented with 1× mTESR1 supplement. Media was changed every day. For passaging, the cells were dissociated using Accutase and the collected colonies

were resuspended in mTESR1 with 1 × mTESR1 supplement and 10 μM Rock inhibitor Y-27632(2HCl) and diluted in the desired density.

### Generation of NPCs

Neural progenitors were generated as previously described with modifications (Boyer et al., 2012; Klaus et al., 2019). In short, embryoid bodies were generated from iPSCs by plating colonies in suspension in neural induction medium consisting of DMEM F12 with N2 and B27 supplements (minus vitamin A). After 7 days in suspension, Embryoid bodies were plated on polyornithine and laminin coated dishes and cultured for 7 days in neural induction medium. Neural rosettes were manually picked, dissociated and plated in a new polyornithine/laminin-coated plate in neural progenitor medium (neural induction medium supplemented with bFGF at 20 ng/mL). For passaging, the cells were dissociated using Accutase and split at a maximum ratio of 1:4.

### Generation of neurons

Neurons were generated following the Gunhanlar protocol (Gunhanlar et al., 2018). Briefly, NPCs were plated on poly-L-ornithine and laminin coated dishes in neural differentiation medium consisting of Neurobasal with N2, B27 supplements (minus vitamin A), minimum essential medium/non-essential amino acid and laminin, supplemented with BDNF, GDNF, ascorbic acid and dcAMP. Media was changed every 2–3 days. Young and mature neurons were collected after 5 weeks and 10 weeks in culture respectively.

### iPSC culture for astrocytes differentiation

iPSCs were cultured on Geltrex LDEV-Free, Reduced Growth Factor Basement Membrane Matrix coated 6-well plates in mTESR1 medium supplemented with 1 × mTESR1 supplement. Media was changed every day. For passaging, the cells were incubated with Collagenase Type IV for 5–7 minutes at 37°C. The collagenase was aspirated and fresh mTESR1 with 1 × mTESR1 supplement was added to each well. A cell scraper was used to collect the cells, and they were subsequently plated on a fresh 6-well plate at the desired dilution.

### Generation of GPCs and astrocytes

Glial progenitor cells and astrocytes were generated as previously described with modifications (Santos et al., 2017). Briefly, confluent iPSC cultures were dissociated with collagenase, collected with a cell scraper and then cultured in suspension to form embryoid bodies. The first 24hrs the cells were cultured in mTESR1 with 1 × mTESR1 supplement and 10 μM Rock inhibitor Y-27632. For the next two weeks the cells were cultured in Astrocyte medium (AM) supplemented with 20ng/mL Noggin and 10 ng/mL PDGFAA, and an additional week with only PDGFAA. The embryoid bodies were then manually dissociated by pipetting and the resulting GPCs were plated on poly-L-ornithine and laminin-coated dishes in AM supplemented with 10ng/mL bFGF and 10ng/mL EGF. Astrocytes were differentiated from GPCs in AM supplemented with 10ng/mL LIF. Media was changed every other day. The GPCs and astrocytes were collected after 5 weeks and 9 weeks of differentiation respectively. The images were obtained on a LSM710 laser-scanning confocal (Carl Zeiss microscope, ZEN software) at ×40 magnification.

### Culturing of HeLa and SH-SY5Y cells

HeLa (RRID: CVCL\_0030) and SH-SY5Y (RRID: CVCL\_0019) cells were cultured in Dulbeccos Modified Eagles Medium – high glucose (DMEM) supplemented with 10% fetal calf serum (FCS) and 2 mM L-alanyl-glutamine at 37°C and 5% CO<sub>2</sub> atmosphere.

## METHOD DETAILS

### Fluorescence imaging

SH-SY5Y neuroblastoma cells were seeded on glass coverslips and grown in DMEM supplemented with 10% FCS. For Click staining of AMPylated proteins, cells were treated with 100 μM pro-N6pA for 24 h. After washing twice in cold PBS, the cells were fixed with 4% paraformaldehyde (PFA) in PBS for 20 min, washed two more times in PBS and permeabilized by incubation with 0.1% Triton X-100 in PBS for 5 min. The metabolically labelled proteins were coupled to TAMRA via CuAAC using 1 mM CuSO<sub>4</sub>, 5 mM THPTA, 10 μM TAMRA-N<sub>3</sub> and 100 mM sodium ascorbate in PBS for 1 h. Following washing twice, a second fixation in 4% PFA for 10 min and two more washing steps. For immunostaining, the cells were permeabilized in 0.2% saponin in PBS for 5 min, followed by quenching with 0.12% glycine and 0.2% saponin for 10 min

and incubation in blocking solution for 1 h (10% FCS and 0.2% saponin in PBS). Primary antibodies were diluted in blocking solution and incubated with the coverslips overnight at 4°C. After four washing steps in 0.2% saponin, the cells were incubated with AlexaFluor488 coupled secondary antibodies for 1 h at room temperature, washed four times in 0.2% saponin and twice in H<sub>2</sub>O. The coverslips were finally placed in 15 µL mounting medium (167 mg/mL Mowiol 4–88, 3% glycerol, 20 mg/mL DABCO, 1 µg/mL DAPI in PBS). Microscopic images were recorded at an Olympus FV1000 confocal laser scanning microscope using a U Plan S-Apo 100× oil objective (1.40 NA).

### Immunocytochemistry – physiological neurons

10 weeks old mature neurons were fixed using 4% PFA for 10 min and permeabilized with 0.3% Triton for 5 min. After fixation and permeabilization, cells were blocked with 0.1% Tween, 10% Normal Goat Serum. Primary and secondary antibodies were diluted in blocking solution. Nuclei were visualized using 0.5 mg/mL 4,6-diamidino-2-phenylindole (DAPI). Stained cells were analyzed using a Leica laser-scanning microscope.

### PLD3 activity assay

For quantitative determination of PLD3 5' exonuclease activity, lysates were prepared in TRIS-lysis buffer (TBS with 1% (v/v) Triton X-100 and 1 tablet cOmplete EDTA-free protease inhibitor cocktail). After collecting the whole cell lysates as described above, lysates were diluted in a final volume of 100 µL MES reaction buffer (50 mM MES, 200 mM NaCl) to a final concentration of 50 ng/µL in a lumox multiwell 96 plate (Sarstedt). The reaction was started by addition of 100 pmol quenched FAM-ssDNA substrate (6-FAM-AC CATGACGTTCC\*CT\*G\*-BMN-Q535 ([Biomers.net](http://Biomers.net)) with \* indicating a phosphothioate bond). After a pre-incubation period of 30 min fluorescence emission at 528 nm (following excitation at 485 nm) was measured in a microwell plate reader (SynergyHT from BioTek) from below the wells over a period of 12 h every 5 min while incubation at 37°C. For evaluation, a substrate control without lysate and a lysate control without substrate were measured together with the samples. The 5' exonuclease activity was calculated as the slope of the measured fluorescence in the samples minus both controls.

### Chemical proteomics

**Seeding of iNGNs for differentiation.** The identification of the AMPylation targets during neuronal differentiation was performed with four controls and four probe treated samples for each time point. In each 10 cm dish 2.5 million iNGNs in 10 mL E7+Dox+TZ media were seeded for differentiation. The media composition that is used for each day of differentiation is described in the section above, differentiation of iNGNs. One day before the cells were harvested, the samples were treated with 5 µL of 100 mM AMPylation probe (pro-N6pA) and the controls with 5 µL DMSO.

**Harvesting and cell lysis.** Cells were washed twice with 2 mL PBS. Then, 500 µL lysis buffer (PBS with 1% (v/v) NP40 and 1% (w/v) sodium deoxycholate and 1 tablet protease inhibitors per 10 mL buffer) was added and cells were scrapped into an Eppendorf tube. The cell suspension was incubated 15 min at 4°C with agitation, before cells were spun down 10 min at 12,000 rpm and 4°C. Subsequently, the cytosolic fraction of the lysate was transferred into a new 2 mL Eppendorf tube.

**Measurement of protein concentrations.** In order to measure the protein concentrations of the lysates bicinchoninic acid assay was performed. First, bovine serum albumin (BSA) standards with concentrations of 12.5, 25, 50, 100, 200 and 400 µg/mL were prepared and samples as well as controls were diluted 40 times to a total volume of 200 µL. To measure standards, samples and controls in triplicates, 50 µL of each was added to three wells of a transparent 96-well plate with flat bottom. Afterwards, 100 µL working reagent (2 µL R2 and 98 µL R1) was added to each well by a multistepper and the plate was incubated 15 min at 60°C. Then, the absorbance at 620 nm was measured by Tecan and the protein concentrations were calculated. For each replicate, except for the 24 h time point (250 µg), 400 µg of protein was used and the volumes were adjusted to a total volume of 970 µL with 0.2% SDS in PBS.

**Coupling of AMPylated proteins.** In order to couple the alkyne residue of AMPylated target proteins with Biotin-PEG<sub>3</sub>-N<sub>3</sub>, CuAAC was used. For this reaction, 10 µL 10 mM Biotin-PEG<sub>3</sub>-N<sub>3</sub>, 10 µL 100 mM TCEP and 1.2 µL 83.5 mM TBTA were added to 970 µL of each lysate. The mixture was vortexed and spun down before 20 µL 50 mM CuSO<sub>4</sub> was added to initiate the reaction. Finally, the reaction mixture was incubated 1.5 h shaking at 600 rpm and 25°C in the dark.

**Protein precipitation.** First, each reaction mixture of the previously performed click reaction was transferred to a 15 mL falcon. Then, 4 mL acetone was added to each falcon in order to precipitate the proteins. After 1 h of incubation at  $-20^{\circ}\text{C}$ , proteins were spinned down 15 min at 11,000 rpm and  $4^{\circ}\text{C}$ . Supernatant was discarded and each pellet was resuspended in 1 mL methanol by sonicating 3 times for 5 s at 20% intensity. Subsequently, each suspension was transferred to a 1.5 mL Eppendorf tube and was centrifuged 10 min at 11,000 rpm and  $4^{\circ}\text{C}$ . Each pellet was washed again with 1 mL methanol before it was dissolved in 1 mL 0.2% SDS in PBS by sonicating 3 times for 5 s at 20% intensity.

**Avidin beads enrichment.** In order to enrich the biotinylated proteins, avidin agarose beads (Sigma-Aldrich) were used. First, avidin beads were thawed on ice before 50  $\mu\text{L}$  avidin bead suspension for each sample or control was washed 3 times with 1 mL 0.2% SDS in PBS to equilibrate the beads. After each addition of washing solution, the Eppendorf tube was carefully inverted 10 times, the suspension was centrifuged 2 min at 2000 rpm at room temperature and the supernatant was discarded. Subsequently, each dissolved protein pellet was spinned down at maximum speed for 2 min and room temperature before the supernatant was added to the equilibrated beads. After each avidin bead suspension was incubated 1 h under continuous mixing at room temperature, the beads of each sample or control were washed 3 times with 1 mL 0.2% SDS in PBS, 2 times with 1 mL 6 M urea in  $\text{H}_2\text{O}$  and 3 times with 1 mL PBS.

**On beads digest of enriched proteins.** In order to prepare defined peptide fragments for the following MS-measurements, the enriched proteins were digested with trypsin. First, the washed avidin beads were resuspended in 200  $\mu\text{L}$  Xbuffer (7 M urea, 2 M thiourea in 20 mM HEPES pH 7.5). Then, 0.2  $\mu\text{L}$  1 M DTT was added to reduce the disulfide bonds. Afterwards, each mixture was vortexed and incubated 45 min at room temperature shaking at 600 rpm. Next, 2  $\mu\text{L}$  550 mM IAA was added to alkylate the cysteine residues. After each mixture was vortexed and incubated 30 min at room temperature shaking at 600 rpm in the dark, 0.8  $\mu\text{L}$  1 M DTT was added to quench the alkylation. Subsequently, each mixture was vortexed and incubated for another 30 min at room temperature shaking at 600 rpm in the dark before 600  $\mu\text{L}$  50 mM TEAB was added to increase the pH value to 8. Afterward, 1.5  $\mu\text{L}$  trypsin (0.5  $\mu\text{g}/\mu\text{L}$  in 50 mM acetic acid) was added to digest the enriched proteins. Finally, each mixture was vortexed and incubated overnight at  $37^{\circ}\text{C}$  shaking at 600 rpm.

On the next day, 4  $\mu\text{L}$  FA was added to stop the digest. Subsequently, each mixture was vortexed and centrifuged 1 min at 2000 rpm before the supernatant was transferred into a new Eppendorf tube. After 50  $\mu\text{L}$  0.1% FA was added to the avidin beads, each mixture was vortexed, the centrifugation step was repeated and the supernatant was again transferred to the new Eppendorf tube. Then, 50  $\mu\text{L}$  0.1% FA was added once more to the avidin beads and each mixture was vortexed. Finally, each mixture was centrifuged 3 min at 13,000 rpm and the supernatant was transferred to the new Eppendorf tube as before. For the following desalting step of the digested proteins, the pH was checked to be below 3.

**Desalting of peptides mixture.** In order to remove all disturbing salt from the digested proteins, Sep Pak C18 cartridges (50 mg columns, waters) were used. First, cartridges were washed with 1 mL ACN and 1 mL 80% ACN with 0.5% FA. Then, cartridges were equilibrated with 3 mL 0.5% FA before the acidified samples were loaded slowly. Afterwards, cartridges were washed with 3 mL 0.5% FA. The desalted peptides were eluted 2 times with 250  $\mu\text{L}$  80% ACN with 0.5% FA into a LoBind Eppendorf tube. Finally, the eluates were lyophilized.

**Final peptide preparation.** First, 30  $\mu\text{L}$  1% FA was added to the lyophilized peptides. Then, the mixture was vortexed, spinned down and sonicated 15 min to dissolve the complete peptides. Afterwards, the dissolved peptides were spinned down and added onto centrifugal filter units. Finally, the solution was centrifuged 1 min at 13,000 rpm and the filtrate was transferred into plastic MS vials.

**MS-measurement.** MS-measurements were performed on a Q Exactive HF mass spectrometer (Thermo Fisher Scientific) coupled to an UltiMate™ 3000 Nano HPLC (Thermo Fisher Scientific) via an EASY-Spray source (Thermo Fisher Scientific). First, peptides were loaded on a Acclaim PepMap 100  $\mu\text{-precolumn}$  cartridge (5  $\mu\text{m}$ , 100  $\text{\AA}$ , 300  $\mu\text{m}$  ID  $\times$  5 mm, Thermo Fisher Scientific). Then, peptides were separated at  $40^{\circ}\text{C}$  on a PicoTip emitter (noncoated, 15 cm, 75  $\mu\text{m}$  ID, 8  $\mu\text{m}$  tip, New Objective) that was in house packed with Reprosil-Pur 120 C18-AQ material (1.9  $\mu\text{m}$ , 120  $\text{\AA}$ , Dr. A. Maisch GmbH). The gradient was run from 1-36% ACN supplemented with 0.1% FA during a 120 min method (0–5 min 1%; 5–8 min to 6%; 8–98 min to 36%; 98–100 min to 85%; 100–105 min wash with 85%; 105–110 min to 1%, 110–120 min with 1%) at a flow rate of

200 nL/min. For measurements of chemical-proteomic samples on Q Exactive HF mass spectrometer, the following settings were used: The Q Exactive HF was operated in dd-MS<sup>2</sup> mode with the following settings: Polarity: positive; MS<sup>1</sup> resolution: 120 k; MS<sup>1</sup> AGC target:  $3 \times 10^6$  charges; MS<sup>1</sup> maximum IT: 20 ms; MS<sup>1</sup> scan range: m/z 300–1750; MS<sup>2</sup> resolution: 15 k; MS<sup>2</sup> AGC target:  $2 \times 10^5$  charges; MS<sup>2</sup> maximum IT: 100 ms; Top N: 20; isolation window: m/z 1.6; isolation offset: m/z 0.2; HCD stepped normalised collision energy: 28%; intensity threshold:  $5 \times 10^4$  counts; charge exclusion: unassigned, 1, 7, 8, >8; peptide match: off; exclude isotopes: on; dynamic exclusion: 90 s.

**Full proteome analysis.** In order to quantify the proteins levels during neuronal differentiation, a full proteome analysis was performed. For this approach, the same lysates (control samples) that were used for the chemical proteomic experiments were used. First, 100 µg protein of each lysate was diluted to a total volume of 200 µL with 0.2% SDS in PBS. Then, 800 µL acetone was added and incubated 1 h at  $-20^\circ\text{C}$  to precipitate proteins. Next, proteins were spun down 15 min at 11,000 rpm and  $4^\circ\text{C}$  and the supernatant was discarded. Each pellet was resuspended in 1 mL methanol by sonicating 3 times for 5 s at 20% intensity and the mixture was spun down once again as before. The supernatant was discarded and the washed protein pellets were dissolved in 200 µL Xbuffer. The following digest and desalting were performed as described in the sections above. Finally, the lyophilized peptides were dissolved in 200 µL 1% FA and filtered as described in the section, final peptide preparation.

**Phosphatase treatment.** In order to ensure that we only observed the separation of AMPylated and not phosphorylated proteins using a Phos-tag gel, the lysates were treated with a phosphatase prior to analysis. For the reaction, 10 µL 10× buffer (0.5 M Tris, 0.1 M MgCl<sub>2</sub>, pH 9.0), 1 µL shrimp alkaline phosphatase (1000u/mL) and 100 µg lysate were mixed and filled up with H<sub>2</sub>O to a total volume of 100 µL. The reaction mixture was incubated overnight at  $37^\circ\text{C}$ . As a positive control, the phosphorylated protein ovalbumin was used.

**Western blot analysis.** For each Western blot analysis, 20 µg cell lysate was used. In order to denature proteins, 4 µL 5× Laemmli buffer (10% (w/v) SDS, 50% (v/v) glycerol, 25% (v/v) β-mercaptoethanol, 0.5% (w/v) bromophenol blue, 315 mM Tris/HCl, pH 6.8) was added to 16 µL lysate solution and the samples were boiled 5 min at  $95^\circ\text{C}$ . Afterwards, 20 µL of each sample was loaded onto a 7.5, 10 or 12.5% SDS gel and proteins were separated according to their size by SDS-PAGE. Then, the separated proteins were transferred onto a membrane using a blotting sandwich moistened by blotting buffer (48 mM Tris, 39 mM glycine, 0.0375% (m/v) SDS, 20% (v/v) methanol), which was composed of one extra thick blot paper, the PVDF transfer membrane, the SDS-PAGE gel and again one extra thick blot paper. Before the protein transfer was carried out 45 min at 25 V using a Semi Dry Blotter (Bio-Rad), the transfer membrane was pre-incubated 5 min in methanol. In order to block non-specific binding sites, the membrane was incubated 60 min in blocking solution (0.5 g milk powder in 10 mL PBST (PBS +0.5% Tween)). Subsequently, 10 µL primary antibody with specificity for the protein of interest was added and the mixture was incubated 1 h at room temperature. The membrane was washed 3 times for 10 min with PBST before 1 µL of the secondary HRP antibody in 10 mL blocking solution was added. After 1 h of incubation at room temperature the membrane was washed again 3 times for 10 min with PBST. Then, 400 µL ECL Substrate and 400 µL peroxide solution were mixed and added to the membrane to stain the Western blot. Finally, images of the Western blot were taken by developing machine Amersham Imager 680 (GE Healthcare).

**Phos-tag gel.** In order to separate the unmodified from the post-translationally modified protein form, a Phos-tag gel was used. Compared to conventional SDS-gels, Phos-tag gels contain additionally Phos-tag reagent and MnCl<sub>2</sub>. To pore a 1 mm thick 7.5% Phos-tag separating gel, 1.25 mL 30% acrylamide solution, 1.25 mL 1.5 M Tris pH 8.8 solution, 50 µL 5 mM Phos-tag reagent, 50 µL 10 mM MnCl<sub>2</sub>, 50 µL 10% SDS, 2.33 mL H<sub>2</sub>O, 5 µL TEMED and 25 µL 10% APS were mixed. The stacking gel was prepared as for conventional SDS-gels by mixing 0.6 mL 30% acrylamide solution, 1 mL 0.5 M Tris pH 6.8 solution, 40 µL 10% SDS, 2.34 mL H<sub>2</sub>O, 4 µL TEMED and 20 µL 10% APS. After the Phos-tag gel was polymerized, 20 µg of the lysates were loaded and dependent on the size of the protein of interest, the gel was run for 2 to 3 h at 150 V. Following the separation, the gel was washed 2 times 15 min with blotting buffer (48 mM Tris, 39 mM glycine, 0.0375% (m/v) SDS, 20% (v/v) methanol) supplemented with 10 mM EDTA to remove the manganese-ions. Subsequently, the gel was washed 15 min in blotting buffer before it was blotted as described in the Western blot section.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Calculation chemical proteomics

MS Raw files were analysed using MaxQuant software 1.6.12.0 with the Andromeda search engine. Searches were performed against the Uniprot database for Homo sapiens (taxon identifier: 9606, March 2020). At least two unique peptides were required for protein identification. False discovery rate determination was carried out using a decoy database and thresholds were set to 1% FDR both at peptide-spectrum match and at protein levels. LFQ quantification was used as described for each sample.

### Statistical analysis chemical proteomics

Statistical analysis of the MaxQuant result table proteinGroups.txt was done with Perseus 1.6.10.43.

First, LFQ intensities were  $\log_2$ -transformed. Afterwards, potential contaminants as well as reverse peptides were removed. Then, the rows were divided into two groups - DMSO (control) and probe treated sample (sample). Subsequently, the groups were filtered for at least three valid values out of four rows in at least one group and the missing values were replaced from normal distribution. The  $-\log_{10}(p \text{ values})$  were obtained by a two-sided one sample Student's t-test over replicates with the initial significance level of  $p = 0.05$  adjustment by the multiple testing correction method of Benjamini and Hochberg ( $FDR = 0.05$ ) using the volcano plot function.

### Profile plots

In order to visualize the changes of the LFQ-intensities during the iNGN differentiation, profile plots were used. For this, the mean (circle), the median (line) and the whiskers for outliers from  $n = 4$  of each time point were prepared in Origin with box defined by 25<sup>th</sup> and 75<sup>th</sup> percentile.

## 3.2 Transforming Chemical Proteomics Enrichment into a High-Throughput Method Using an SP2E Workflow

Becker, T., Wiest, A., Telek, A., Bejko, D., Hoffmann-Röder, A., and Kielkowski, P. (2022). Transforming Chemical Proteomics Enrichment into a High-Throughput Method Using an SP2E Workflow. *JACS Au* 2, 1712–1723. 10.1021/jacsau.2c00284.

### Prologue

Chemical proteomics is the state-of-the-art technique to identify PTM proteins and is thus of great value for deciphering PTM networks and their association with diseases. For the sample preparation in chemical proteomics so far many different enrichment protocols have been developed. However, these workflows are time-consuming and must be performed by specialized laboratory personnel. In addition, high protein amounts are required, which are challenging to obtain for cell lines that are difficult to culture. Therefore, the aim of this study was to develop a protocol that is suitable for robotic automation and furthermore needs a significantly lower amount of protein. To enable automation, the workflow is based on carboxylate-coated and streptavidin-coated paramagnetic beads, which, unlike agarose beads, do not need an additional centrifugation step to separate the liquid and solid phases. Carboxylate-coated paramagnetic beads have already been used in the whole proteome sample preparation to remove detergents and remaining non-protein components after cell lysis.<sup>77,78</sup> We adapted the so-called SP3 protocol by *Hughes et al.* to clean-up the proteins after the click reaction by carboxylate-coated paramagnetic beads instead of protein precipitation, which could not be automated. Now that all prerequisites for automation were fulfilled, the AMPylation probe pro-N6pA was used to systematically optimize each step of the protocol.<sup>2</sup> First, a suitable lysis buffer for an efficient click reaction was determined. Next, it was found that the addition of urea after the click reaction increased the protein binding to the carboxylate-coated paramagnetic beads. Furthermore, to reduce background binding, the purified proteins were eluted from the carboxylate-coated paramagnetic beads before enrichment on streptavidin-coated paramagnetic beads. To demonstrate the robustness of the SP2E workflow, AMPylation was screened in SH-SY5Y neuroblastoma cells with different inhibitors including bafilomycin, monensin, rapamycin, thenoyltrifluoroacetone (TTFA) and 2-deoxy-D-glucose. In addition to AMPylation, another PTM was used to validate our approach, namely the *O*-linked  $\beta$ -*N*-acetyl-glucosamine glycosylation (*O*-GlcNAcylation).<sup>98</sup> After the SP2E workflow was shown to work reliably with 400  $\mu$ g of starting protein, the protocol was successfully scaled down into 96-well plate format using only 100  $\mu$ g of protein input. This allows to screen protein



PTMs with high-throughput and also enables the investigation of PTMs in difficult-to-culture cell lines. Taken together, we developed a fast, robust and sensitive chemical proteomics platform that can be automated by a robot.

### **Author contribution**

In this study, I carried out the AMPylation profiling in SH-SY5Y cells using different inhibitors. In collaboration with *Pavel Kielkowski*, I analyzed the data of all chemical proteomic experiments and monitored the entire study. Furthermore, I wrote the manuscript together with *Pavel Kielkowski*. In addition, *Pavel Kielkowski* designed the method. *Andreas Wiest*, *András Telek* and *Daniel Bejko* performed the experiments to optimize the method. *Anja Hoffmann-Röder* provided the Ac<sub>3</sub>4dGlcNAz probe for the *O*-GlcNAcylation analysis.

### **License**

Copy of the open access publication based on the Creative Commons Attribution-NonCommercial License (CC BY-NC-ND 4.0), which allows the non-commercial reprint of the article.

### **Article available at:**

<https://pubs.acs.org/doi/10.1021/jacsau.2c00284>

### **License available at:**

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Supporting information are available in the appendix (chapter 8.2)

# Transforming Chemical Proteomics Enrichment into a High-Throughput Method Using an SP2E Workflow

Tobias Becker, Andreas Wiest, András Telek, Daniel Bejko, Anja Hoffmann-Röder, and Pavel Kielkowski\*



Cite This: *JACS Au* 2022, 2, 1712–1723



Read Online

ACCESS |

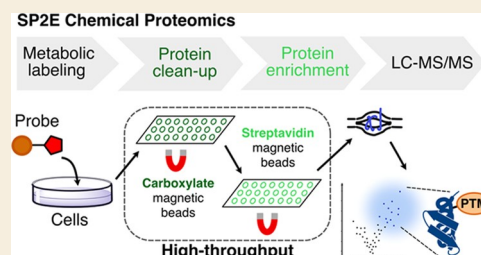
Metrics & More

Article Recommendations

Supporting Information

**ABSTRACT:** Protein post-translational modifications (PTMs) play a critical role in the regulation of protein catalytic activity, localization, and protein–protein interactions. Attachment of PTMs onto proteins significantly diversifies their structure and function, resulting in proteoforms. However, the sole identification of post-translationally modified proteins, which are often cell type and disease-specific, is still a highly challenging task. Substoichiometric amounts and modifications of low abundant proteins necessitate the purification or enrichment of the modified proteins. Although the introduction of mass spectrometry-based chemical proteomic strategies has enabled the screening of protein PTMs with increased throughput, sample preparation remains highly time-consuming and tedious. Here, we report an optimized workflow for the enrichment of PTM proteins in a 96-well plate format, which could be extended to robotic automation. This platform allows us to significantly lower the input of total protein, which opens up the opportunity to screen specialized and difficult-to-culture cell lines in a high-throughput manner. The presented SP2E protocol is robust and time- and cost-effective, as well as suitable for large-scale screening of proteoforms. The application of the SP2E protocol will thus enable the characterization of proteoforms in various processes such as neurodevelopment, neurodegeneration, and cancer. This may contribute to an overall acceleration of the recently launched Human Proteoform Project.

**KEYWORDS:** chemical proteomics, SP3, protein post-translational modifications, AMPylation, glycosylation

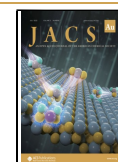


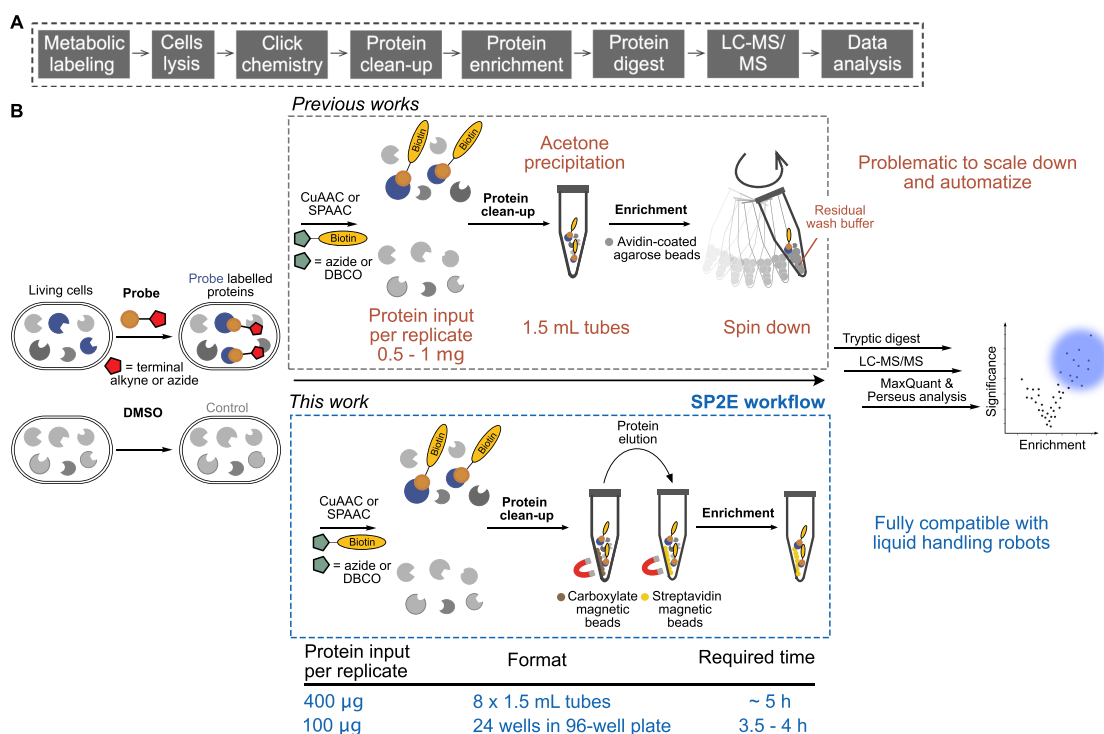
## INTRODUCTION

Protein post-translational modifications (PTMs) are crucial for the regulation and fine-tuning of many important biological processes such as neurodevelopment,<sup>1–4</sup> circadian clocks,<sup>5</sup> aging,<sup>6</sup> and impairment in numerous diseases.<sup>7–9</sup> The incredible diversity of genetic polymorphism, RNA splice variants, and PTMs results in many proteoforms,<sup>4,10,11</sup> which exceed the ~20,000 human genes by approximately 50 times. This biological network orchestrates the most complex processes including brain development and ensures a dynamic response of the cells to an external stimulus. However, the extent of protein PTMs in laboratory-cultured cells can differ significantly depending on cell types, diseases, and culture conditions. Mass spectrometry (MS)-based chemical proteomics has allowed to reliably map protein PTMs across various experimental conditions.<sup>12–17</sup> A widespread application of the chemical proteomic strategy was enabled by parallel improvements of liquid chromatography technologies, gains in speed, and sensitivity of mass spectrometers and bioinformatic pipelines for protein identification and quantification.<sup>18–21</sup> Nowadays, chemical proteomics is used to uncover the scope of protein PTMs in different cell types by the development of small-molecule probes, which mimic their natural counterparts. The utilization of these probes has provided valuable insights into protein acetylation, palmitoylation, myristoylation, pre-

nylation, glycosylation, ADP-ribosylation, and AMPylation.<sup>1,12–14,22–25</sup> In general, different chemical proteomic workflows follow the same sequence of key steps (Figure 1A). First, cultured cells are treated with the probe, which infiltrates the cellular system and competes with the endogenous substrate for the active site of the PTM writer enzymes. Next, the chemical proteomic probes contain an alkyne or azide handle to facilitate a bioorthogonal coupling to suitable biotin linkers with either Cu-catalyzed alkyne–azide cycloaddition (CuAAC) or copper-free strain-promoted azide–alkyne cycloaddition (SPAAC), respectively.<sup>15,26</sup> Following the click chemistry, proteins are precipitated to remove the excess biotin reagents and nonprotein components of the cell lysate.<sup>1,13</sup> In the next step, biotin-labeled proteins are enriched using avidin-coated beads. The critical part of this step is to maximize the efficiency of the washing to remove nonspecifically bound proteins and thus to reduce the complexity of the final MS sample while improving ratios

**Received:** May 9, 2022  
**Revised:** June 3, 2022  
**Accepted:** June 3, 2022  
**Published:** June 30, 2022



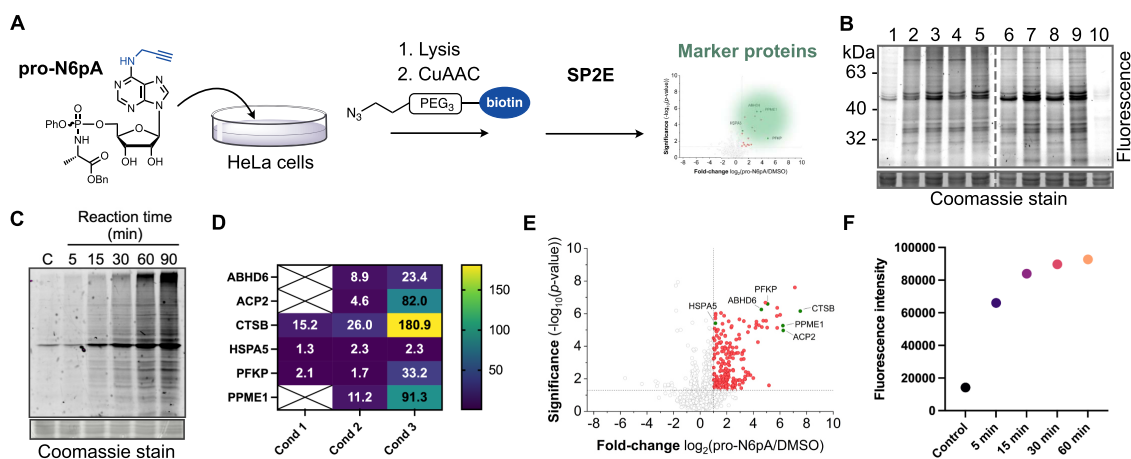


**Figure 1.** Schematic overview of the chemical proteomic workflow. (A) Key steps of the standard chemical proteomic workflow and (B) schematic characterization of the SP2E workflow and basic parameters of the procedure in comparison to the previously used workflow with avidin-coated agarose beads. In the table below the SP2E workflow, the typical times required to proceed with 8 samples (large scale) or 24 samples (small scale) are shown. For comparison, the previously used workflow would typically take about 8 h for eight samples.

between control- and probe-treated samples. After reduction and alkylation of the enriched modified proteins, they are digested by trypsin or another protease, desalted and concentrated for MS measurement.<sup>27</sup> Subsequently, each sample is measured separately for the label-free quantification (LFQ), providing the possibility to add more samples into the data set later on.<sup>28</sup> MS data are typically acquired on orbitrap or timsTOF-based liquid chromatography-tandem mass spectrometry (LC-MS/MS) instruments. Finally, peptide and protein identifications and quantifications are carried out using well-established commercial or free-of-charge pipelines such as MaxQuant or MSFragger.<sup>21,28</sup> A comparison of the probe-treated and control cells allows us to distinguish between unspecific protein binders and probe-modified proteins. Despite the success of chemical proteomic technology, the community of scientists combining organic synthesis, mass spectrometry, and biology is still rather small. With numerous validated commercial PTM probes and the widespread availability of mass spectrometers either used in individual groups or as a core service, the bottlenecks of the chemical proteomic approach still remain in the insufficient consistency, time inefficiency, and laboriousness of the enrichment techniques. Furthermore, the emerging field of chemical proteomic studies focused on neurodevelopment and the complex environment of the central nervous system composed of many different cell types is not compatible with the high amounts of total proteins that are required for the analysis so far. In a perfect scenario, the workflow would be efficient and reproducible even with a low protein input and the enrichment would require a minimum hands-on time or full automation. In comparison to standard MS whole proteome sample preparation methods, which include filter-assisted sample preparation (FASP), in-StageTip digestion (iST), and single-

pot solid-phase-enhanced sample preparation (SP3), the chemical proteomic method not only requires highly efficient protein and peptide purification but also needs to be combined with suitable bioorthogonal reaction conditions and much higher starting protein amounts (Figure 1B).<sup>29–32</sup> Thus far, the most common chemical proteomic methods for protein PTM enrichment utilize acetone or chloroform–methanol precipitation to remove the excess of click chemistry reagents.<sup>13</sup> Further, enrichment with avidin-coated agarose beads requires centrifugation or filtration to separate them from the wash buffer, which hinders the scale down and robotic automation of the approach. Of note, the Tate group combined avidin-coated magnetic beads and a trifunctional linker with azide, biotin, and rhodamine to visualize the enriched proteins by in-gel analysis,<sup>13,33</sup> and recently, the Backus group has implemented the SP3 peptide clean up into their chemical proteomic workflow before transferring the peptides including the biotin-modified peptides on avidin-coated agarose beads.<sup>34</sup> Although similar affinity enrichment techniques combined with MS have been reported, to our best knowledge, a procedure feasibly integrating all aspects of small-scale chemical proteomics is not available.<sup>35,36</sup>

Here, we report the development and optimization of a chemical proteomic method, which uses carboxylate-modified magnetic beads to clean up the proteins after CuAAC and streptavidin-coated magnetic beads for the enrichment of the labeled proteins (Figure 1B). The new method termed SP2E was further scaled down to a 96-well plate format, starting with 100  $\mu$ g total protein. The SP2E method has been successfully used for profiling protein glycosylation and the low abundant protein PTM called AMPylation. Together, the SP2E method provides a time-effective and robust platform for the routine and high-throughput profiling of protein PTMs, representing



**Figure 2.** Development and optimization of the SP2E workflow using the AMPylation probe. (A) Pro-N6pA probe structure and the workflow used for the optimization of the SP2E method. (B) Optimization of the lysis buffer based on the efficiency of the CuAAC click chemistry. Lysis buffer compositions: line 1 (control cells treated with plain dimethyl sulfoxide (DMSO) and lysed in 1% NP-40, 0.2% SDS in 20 mM HEPES), line 2 (1% NP-40 in PBS), line 3 (1% NP-40, 0.2% SDS in PBS), line 4 (0.5% Triton in PBS), line 5 (0.5% Triton, 0.2% SDS in PBS), line 6 (1% NP-40 in 20 mM HEPES), line 7 (1% NP-40, 0.2% SDS in 20 mM HEPES), line 8 (0.5% Triton in 20 mM Hepes), line 9 (0.5% Triton, 0.2% SDS in 20 mM Hepes), and line 10 (8 M urea in 0.1 M Tris/HCl). (C) In-gel fluorescence showing the click reaction time optimization. In the control C, cells were treated with plain DMSO and the lysate was incubated with the click reaction mixture for 90 min. (D) Heatmap visualizing the SP2E workflow optimization based on fold enrichment of six marker proteins. Condition 1 (without added urea to the click reaction mixture before protein loading onto carboxylate magnetic beads), condition 2 (with added urea to the click reaction before protein loading onto carboxylate beads and one pot clean up and enrichment of modified proteins), and condition 3 (with added urea into the click reaction, but the spatial separation of the protein clean up and enrichment). The numbers in boxes represent fold enrichments. (E) Volcano plot showing significantly enriched proteins (red dots) using the pro-N6pA AMPylation probe with highlighted marker proteins (green dots) using the optimized SP2E workflow (condition 3 from Figure 2D);  $n = 4$ , cutoff lines at  $p$ -value  $> 0.05$  and 2-fold enrichment. (F) Plot displaying total fluorescence intensity from the in-gel analysis of the time optimization of biotin–streptavidin complex formation.

an important step toward the robotic automation of the approach.

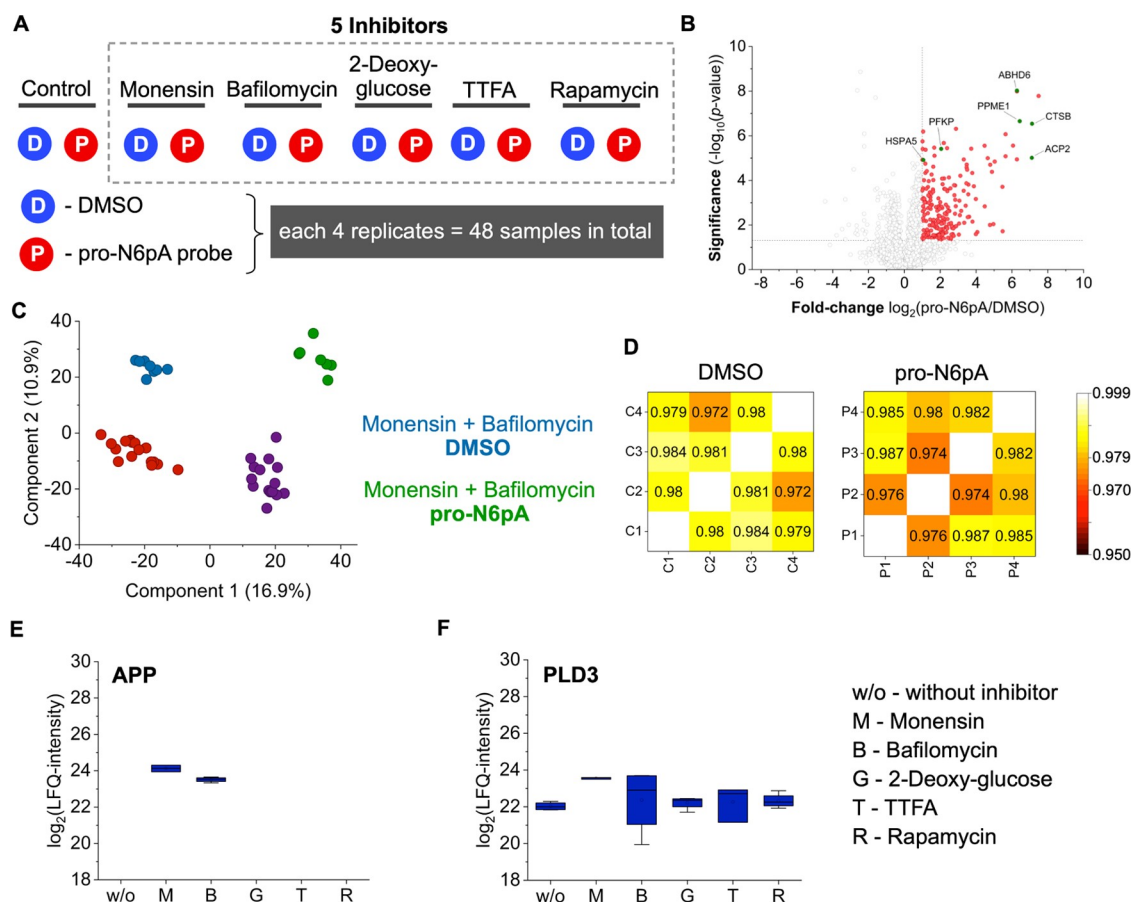
## RESULTS

### Development of the SP2E Workflow for Chemical Proteomics

To implement both carboxylate- and streptavidin-coated magnetic beads for protein clean up after click chemistry instead of precipitation and replacement of avidin agarose beads, every step of the chemical proteomic workflow was optimized. Here, we describe the most critical steps leading to an efficient enrichment of PTM proteins, which include the click chemistry conditions, protein loading on carboxylate magnetic beads, and spatial separation of protein clean up and enrichment. First, we set out to optimize the lysis buffer composition to maximize the efficiency of the click chemistry. To evaluate this, HeLa cells were treated with the cell-permeable pro-N6pA probe infiltrating protein AMPylation (Figure 2A).<sup>1,17</sup> After overnight incubation, the cells were harvested and lysed in nine different lysis buffers (Figure 2B). The CuAAC click chemistry was performed with 200  $\mu$ g of total protein per sample in 100  $\mu$ L of the respective lysis buffer and azide–TAMRA to visualize the conversion efficiency by in-gel fluorescence scanning after sodium dodecyl sulfate–polyacrylamide gel (SDS–PAGE) electrophoresis. The overall brightest fluorescence and highest amount of fluorescent protein bands was observed in the lysis buffer containing 1% NP-40, 0.2% SDS in 20 mM *N*-(2-hydroxyethyl)piperazine-*N'*-ethane sulfonic acid (HEPES) pH 7.5 (Figures 2B and S1), which was further used for all following experiments. To complete the click chemistry optimization, a time-dependent experiment was carried out to show that 1.5 h is necessary to

maximize the yield of the reaction (Figure 2C). For the following optimization steps of the MS workflow, we used a group of six known AMPylated maker proteins (HSPA5, CTSB, PFKP, PPME1, ACP2, ABHD6) to assess the PTM protein enrichment efficiency. The first optimization step was to examine the influence of the buffer composition on the efficiency of the protein binding to the carboxylate magnetic beads. Therefore, 400  $\mu$ g of total protein in 200  $\mu$ L of lysis buffer was used for the click reaction with azide–PEG<sub>3</sub>–biotin. After the CuAAC incubation time, the resulting click reaction mixture was directly transferred onto carboxylate-coated magnetic beads and followed by the addition of absolute EtOH to a final concentration of 60% to induce the protein complexation. After washing the beads three times with 80% EtOH, the streptavidin-coated magnetic beads were added directly to the carboxylate-coated magnetic beads and incubated for 1 h in 0.2% SDS in phosphate-buffered saline (PBS) to form the biotin–streptavidin complex. To remove the unmodified proteins, the bead mixture was washed thrice with 0.1% NP-40 in PBS, twice with 6 M urea, and thrice with water. The enriched proteins were subsequently reduced, alkylated, and trypsin digested in ammonium bicarbonate (ABC) buffer. The resulting peptides were eluted from the beads by two washes with ABC buffer before desalting on off-line Sep-Pak C18 columns and separation on a UHPLC using a 150 min gradient with the high-fidelity asymmetric waveform ion mobility spectrometry (FAIMS) device attached on to the Orbitrap Eclipse Tribrid mass spectrometer. The MS data were analyzed by MaxQuant and evaluated in Perseus.<sup>20,28</sup>

We have observed that protein loading onto carboxylate magnetic beads directly in lysis buffer with the click reagents did not give satisfactory results with poor enrichment of the marker proteins (Figures 2D and S2 and Table S1). Therefore,



**Figure 3.** Analysis of protein AMPylation under different stress conditions using the SP2E workflow. (A) Design of the experiment to test the impact of various inhibitors on protein AMPylation. (B) Volcano plot showing the enrichment of AMPylated proteins (pro-N6pA vs DMSO) from SH-SY5Y cells using the SP2E protocol;  $n = 4$ , cutoff lines at  $p$ -value  $> 0.05$  and 2-fold enrichment. (C) PCA of the inhibitor-treated cells and controls displaying separation of the monensin and bafilomycin as well as pro-N6pA-treated cells. Samples that were treated with DMSO and either rapamycin, 2-deoxy-D-glucose, TTFA, or without any inhibitor are depicted in red. Samples that were treated with pro-N6pA and either rapamycin, 2-deoxy-D-glucose, TTFA, or without any inhibitor are depicted in purple. (D) Representative heatmaps visualizing the Pearson correlation coefficients of LFQ intensities of DMSO and pro-N6pA replicates. (E) Profile plot displays the APP LFQ intensities under various conditions. The APP was not found in any other conditions, for example, in cells only treated with DMSO or some other inhibitors. (F) Profile plot displays the PLD3 LFQ intensities under various conditions.

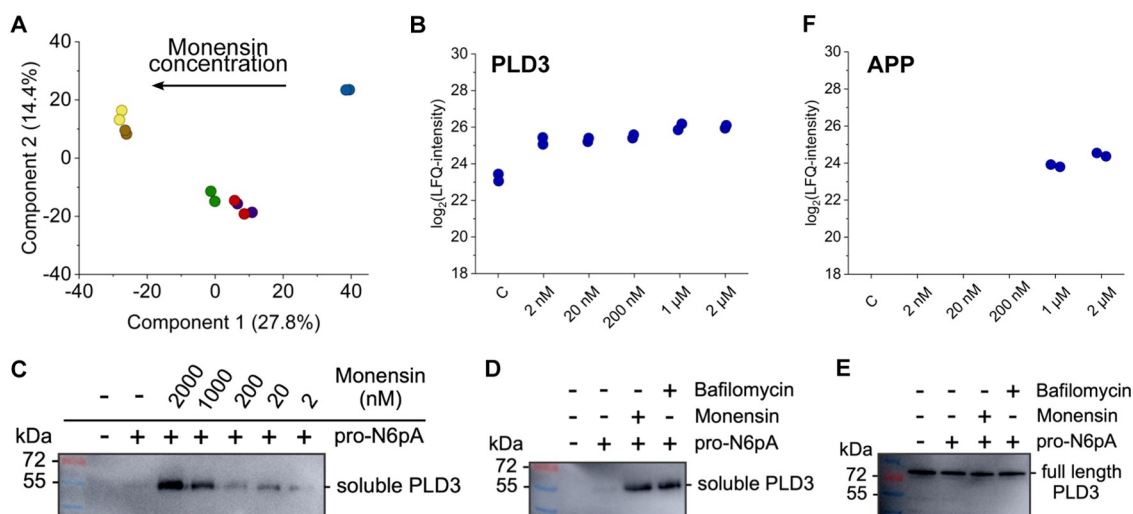
based on our previous experiments, we have tested whether the addition of concentrated urea to the click reaction mixture in advance may improve the protein loading onto the carboxylate magnetic beads. This could be explained by the interference of copper with the protein complexation to carboxylate magnetic beads. Thus, the addition of urea may lead to the neutralization of this effect.<sup>37,38</sup> Indeed, the dilution of the lysis buffer to a final concentration of 4 M urea has significantly improved the overall enrichment ratio (Figures 2D and S3 and Table S2).

Finally, we explored whether it is possible to reduce the overall protein background binding by spatial separation of the protein clean up and enrichment by eluting the proteins from the carboxylate magnetic beads before adding them to the streptavidin-coated magnetic beads in a new tube and thus improve the enrichment ratio. Hence, after the protein clean up on carboxylate magnetic beads, the proteins were eluted twice with 0.2% SDS in PBS and transferred onto streptavidin magnetic beads. The resulting volcano plot confirmed that spatial separation of protein clean up and enrichment is beneficial for PTM protein enrichment and hence outcompetes the advantage of performing both steps in one pot (Figure 2D,E and Table S3). In particular, the enrichment of ACP2

and PFKP has increased by more than 17-fold during the method development. Overall, when compared to the previous benchmark enrichment based on the avidin agarose beads, the optimized SP2E workflow yielded 10-fold more significantly enriched proteins.<sup>17</sup>

Additionally, the protein digest on beads and subsequent peptide elution conditions have been optimized.<sup>31,39</sup> However, standard reduction and alkylation with tris(2-carboxyethyl)-phosphine (TCEP) and chloroacetamide (CAA) at 95 °C for 5 min after the protein enrichment gave satisfying results (Figure 2E). To further shorten the time of the workflow, we analyzed the time dependency of the biotin–streptavidin complex formation to find out that 15 min is sufficient (Figures 2F and S4).

Taken together, we have established the feasible SP2E workflow for the MS-based analysis of PTM proteins. First, we showed that the lysis buffer containing 1% NP-40 and 0.2% SDS in 20 mM HEPES pH 7.5 efficiently lyse the cells and improves the yield of the CuAAC. Furthermore, the addition of urea into the lysis buffer after click chemistry enhances protein loading on the carboxylate magnetic beads. Next, we demonstrated that it is possible to reduce the background by



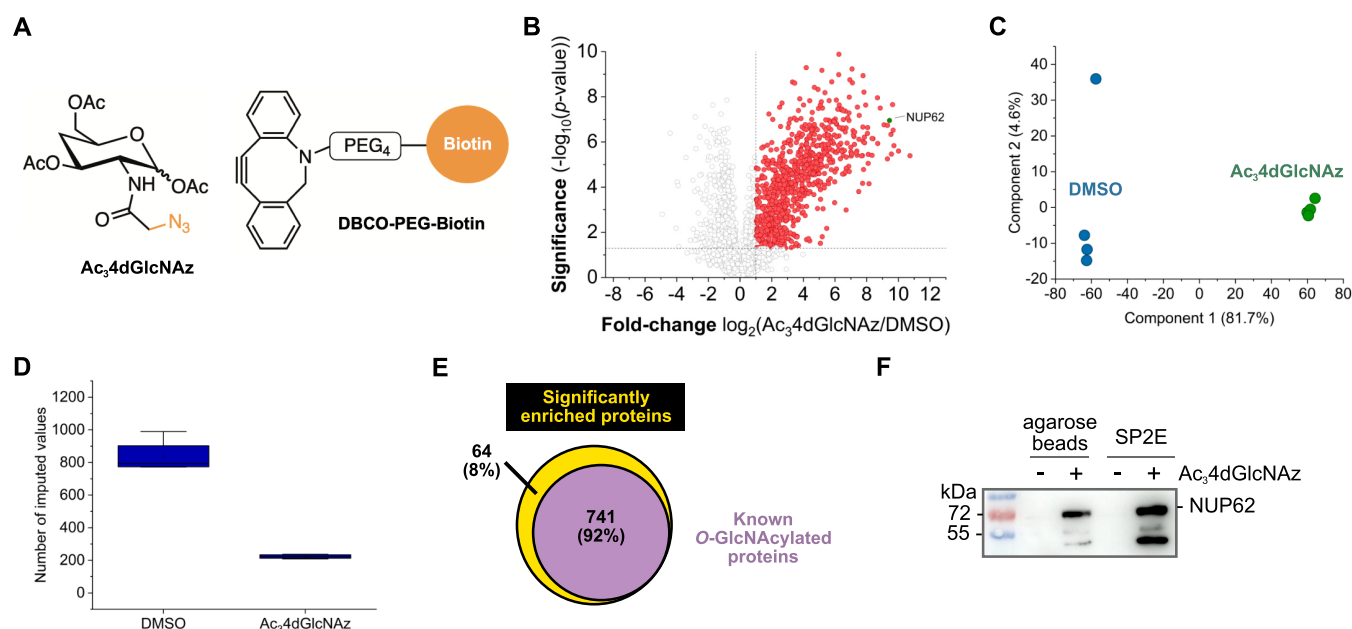
**Figure 4.** Monensin concentration-dependent increase in APP and PLD3 modification. (A) PCA displays distinct changes in the enriched proteins with increasing monensin concentration. (B) Profile plot of the PLD3 LFQ intensities shows a rapid increase in the PLD3 modification with a 2 nM monensin concentration. (C) Monensin concentration-dependent enrichment of the modified PLD3. For the enrichment, the SP2E protocol was used but the proteins were released from the streptavidin beads by the loading buffer, separated by SDS-PAGE, and analyzed via western blotting with the anti-PLD3 antibody. (D) Enrichment of the modified PLD3 after the treatment with bafilomycin (100 nM) and monensin (2 μM). For the enrichment, the SP2E protocol was used but the proteins were released from the streptavidin beads by loading buffer, separated by SDS-PAGE, and analyzed via western blotting with the anti-PLD3 antibody. (E) Western blotting of the whole proteome from cells treated with bafilomycin (100 nM) and monensin (2 μM) stained with the anti-PLD3 antibody. (F) In contrast to PLD3, the profile plot of the APP LFQ intensity reveals that APP is only enriched with the pro-N6pA probe with 1 and 2 μM monensin in cell culture media.

spatial separation of the protein clean up and enrichment by eluting the proteins from the carboxylate magnetic beads before adding them to the streptavidin magnetic beads. The SP2E workflow utility was demonstrated on productive enrichment of the low abundant AMPylated proteins.

#### Application of the SP2E Workflow for Analysis of Protein AMPylation

To validate our approach using a heterogeneous set of samples, we went on to investigate metabolic pathways which putatively regulate protein AMPylation. In our previous studies, we showed that primary lysosomal proteins in neuroblastoma SH-SY5Y cells are metabolically labeled by our probe and thus likely AMPylated. Moreover, the profiling of protein AMPylation during neuronal differentiation of the human-induced pluripotent stem cells (hiPSCs) revealed an accumulation of the lysosomal AMPylated proteins in mature neurons, including PLD3, ACP2, and ABHD6.<sup>1,17</sup> In the case of PLD3, the increased AMPylation correlates with a decrease in its exonuclease catalytic activity. However, in parallel, another group of proteins localized to cytosol and mitochondria has been consistently enriched, including PFKP, PPME1, SLC25A3, and cytoskeletal proteins. There are two known AMPylators thus far, FICD and SELENOO, which are localized in the endoplasmic reticulum (ER) and mitochondria, respectively. This raises the question about the localization and mechanism of metabolic or signaling pathways regulated by protein AMPylation apart from the previously described FICD-HSPAS axis involved in the unfolded protein response (UPR).<sup>6,40,41</sup> With our optimized SP2E workflow, we decided to address this question using a set of five inhibitors blocking mTOR, autophagy lysosomal pathway, glycolysis, and cellular respiration. Therefore, the neuroblastoma SH-SY5Y cells were treated with rapamycin, bafilomycin, 2-deoxy-D-glucose, thenoyltrifluoroacetone (TTFA), and monensin.<sup>42–46</sup> SH-SY5Y cells were treated either with the inhibitor or with

the inhibitor together with the pro-N6pA probe to avoid any influence of protein expression changes on protein enrichment triggered by the inhibitor (Figure 3A). Additionally, two more controls were included. Cells were treated with plain DMSO or the pro-N6pA probe to check the efficiency of the enrichment in neuroblastoma cells (Figure 3B and Table S4). For each condition, four replicates have been prepared and each enrichment was performed with 400 μg of total protein (Figure 3A). This screening resulted in overall 762 significantly enriched proteins. The control experiment without any inhibitor alone gave 236 significantly enriched proteins. In comparison, the previous agarose beads' enrichment workflow yielded only 14 significantly enriched proteins, marking a large improvement in the enrichment efficiency using the SP2E workflow.<sup>1,17</sup> Together, there is an overlap of 35 proteins, which were significantly enriched under all conditions. The principal component analysis (PCA) showed a clear difference between the DMSO control and probe-treated samples (Figures 3C and S5) for all experiments. Interestingly, the samples that were treated with bafilomycin and monensin clustered together, suggesting a common signaling pathway involving protein AMPylation. This demonstrates the robustness of the SP2E workflow with Pearson correlation coefficients of the LFQ intensities between the replicates over 95% (Figure 3D) and its feasibility to discover new pathways regulating protein AMPylation. The analysis of the enriched proteins revealed the amyloid-β precursor protein (APP) to be one of the top hits in cells treated with bafilomycin and monensin. Further examination of the profile plots showed that APP is only enriched in the cells that were treated with the pro-N6pA probe and either bafilomycin or monensin (Figure 3E). The subsequent search for similar enrichment profiles has uncovered a group of another 12 proteins including GPR56, FAT1, LAMA4, TGOLN2, RNF149, CRIM1, ITM2B, L1CAM, TMEM59, MCAM, LRP1, and CLU that were specifically enriched under these



**Figure 5.** Analysis of *O*-GlcNAcylation by the  $\text{Ac}_3\text{4dGlcNAz}$  probe, SPAAC, and SP2E workflow. (A) Chemical structure of the  $\text{Ac}_3\text{4dGlcNAz}$  probe for metabolic labeling of *O*-GlcNAcylated proteins and the DBCO–biotin reagent to functionalize the probe-modified proteins by SPAAC. (B) Volcano plot visualizing the enrichment of the *O*-GlcNAcylated proteins;  $n = 4$ , cutoff lines at  $p$ -value  $>0.05$  and 2-fold enrichment. Red dots are significantly enriched proteins. (C) PCA graph points to a clear separation of the control and probe-treated samples. Of note, component 1 possesses a high value of 81.7%. (D) Box plot shows the total number of imputed values across the replicates in DMSO and  $\text{Ac}_3\text{4dGlcNAz}$ -treated cells. The number of imputed values indicates how many proteins were not identified in the sample but were found in at least one other sample/replicate. The increased number of imputed values in the DMSO controls demonstrates the efficiency of washing steps removing the nonspecific binding proteins. (E) Diagram showing the overlap between all significantly enriched proteins using the  $\text{Ac}_3\text{4dGlcNAz}$  probe and previously described *O*-GlcNAcylated proteins. (F) Enrichment of the modified NUP62 with the agarose-based and the SP2E protocol. In addition, the SP2E enrichment was performed with 400  $\mu\text{g}$  of the protein input. Enriched proteins were released from the streptavidin beads by loading buffer, separated by SDS-PAGE, and analyzed via western blotting with the anti-NUP62 antibody.

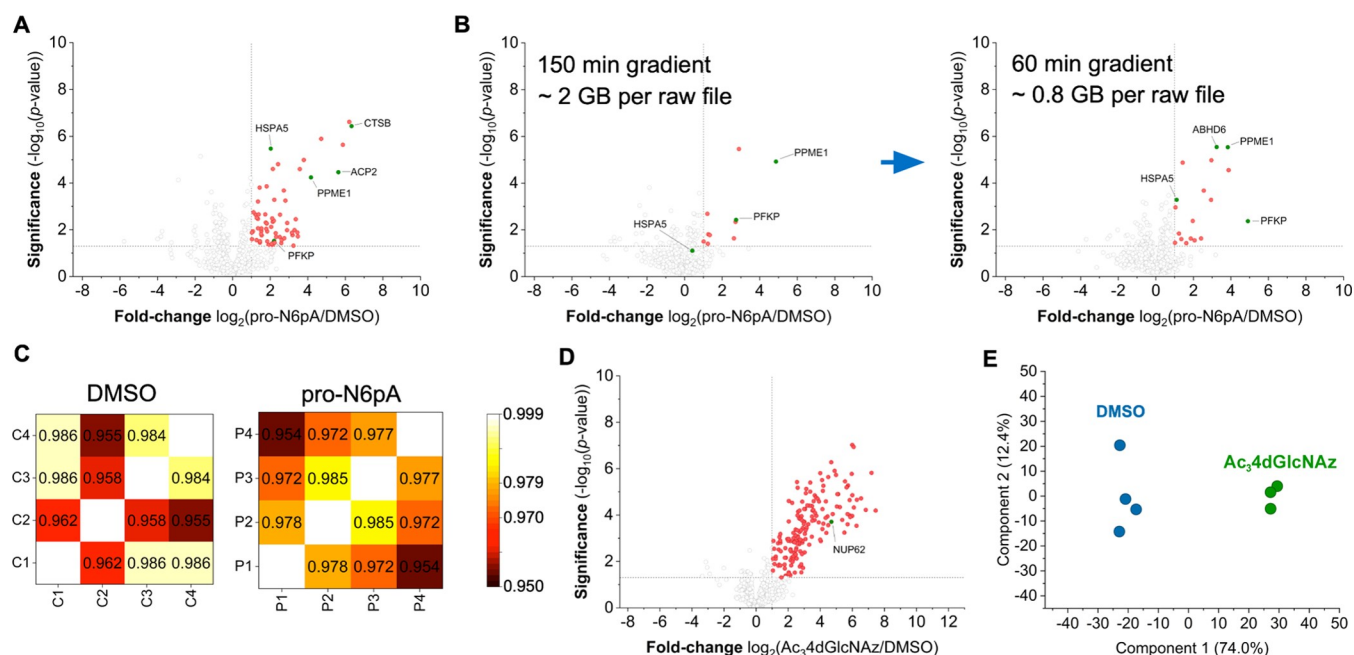
two conditions (Figure S5). Interestingly, the lysosomal exonuclease PLD3 has shown a strong response to monensin (Figure 3F). This would point toward the link between AMPylation and trafficking pathways from ER to lysosomes and autophagy.

To investigate the relationship between the monensin concentration and the protein AMPylation stoichiometry in more detail, SH-SY5Y cells were treated with an increasing concentration of monensin in the cell culture media ranging from 2 nM to 2  $\mu\text{M}$  (Figure 4A).<sup>46–48</sup> The concentration of the pro-N6pA probe was kept constant and each condition was performed in duplicate. Surprisingly, the modification of PLD3 increased by 4.0-folds with 2 nM monensin and further increased with a higher inhibitor concentration (Figure 4B and Table S5). The extent of the PLD3 modification was then confirmed in the following in-gel analysis (Figure 4C). Therefore, the SP2E-based enrichment was used, but instead of the protein digest on the beads, they were released from the streptavidin beads, separated on SDS-PAGE, and analyzed by western blotting. For the first time, we uncovered that only the soluble form of PLD3 is modified by the pro-N6pA probe and that there is a strong increase in the modified form upon the addition of monensin and bafilomycin (Figure 4D). The high efficiency of the SP2E workflow is demonstrated by a clear band of soluble PLD3 after enrichment, although this band is not even visible in the western blotting of the whole cell lysate (Figure 4E). In contrast, the modified APP was only found in cells treated with 1 and 2  $\mu\text{M}$  monensin (Figure 4F).<sup>48</sup> Indeed, PTMs of the amyloid- $\beta$  precursor protein are deemed to play an important role in the development of Alzheimer's disease

pathophysiology.<sup>49</sup> By application of our SP2E workflow, we showed that AMPylation might be an additional PTM involved in the regulation of the APP physiological function. Together, the screening of the AMPylation changes triggered by five different active compounds manifests the utility of the SP2E workflow, which is characterized by minimal background binding and robustness.

#### Application of the SP2E Workflow for Analysis of Protein *O*-GlcNAcylation

To assess the utility of our optimized workflow for other PTMs, we used the previously described  $\text{Ac}_3\text{4dGlcNAz}$  probe for the pull-down of *O*-linked  $\beta$ -*N*-acetyl-glucosamine glycosylated (*O*-GlcNAcylation) proteins (Figure 5A).<sup>12,50–52</sup> Numerous metabolic labels have been developed for the characterization of *O*-GlcNAcylated proteins.<sup>51</sup> However, they often suffer from low substrate specificity and labeling efficiency. Here, we used the 2,4-dideoxy-D-glucopyranose derivative, which shows improved specificity for cytosolic proteins.<sup>51</sup> In contrast to previous experiments with the pro-N6pA probe, the  $\text{Ac}_3\text{4dGlcNAz}$  probe contains an azide functional group for bioorthogonal protein labeling using SPAAC (Figure 5A).<sup>26</sup> The HeLa cells were treated with a 200  $\mu\text{M}$  final concentration for overnight and lysed. To avoid unspecific reactivity of free thiols with the DBCO–biotin reagent utilized for the SPAAC, they were capped with iodoacetamide.<sup>53</sup> Afterward, the lysate (400  $\mu\text{g}$  total protein) containing *O*-4dGlcNAz proteins was reacted with the DBCO–biotin reagent and enriched using the SP2E method in the same fashion as described above for AMPylation (Figure



**Figure 6.** Scale-down of the SP2E workflow into a 96-well plate format. (A) SP2E protocol with 100  $\mu\text{g}$  of the input protein performed in 1.5 mL tubes visualized in the volcano plot. (B) Optimization of the LC-MS/MS measurement with 100  $\mu\text{g}$  protein input using the 96-well plate format SP2E protocol. (C) Heatmaps representing the Pearson correlation coefficients between the replicates. (D) Volcano plot showing the enrichment of *O*-GlcNAcylated proteins starting from 100  $\mu\text{g}$  of the input protein in a 96-well plate format. (E) PCA of the small-scale  $\text{Ac}_3\text{4dGlcNAz}$  enrichment shows very good separation of controls from probe-treated samples, with component 1 value corresponding to 74%. All volcano plots,  $n = 4$ , cutoff lines at  $p$ -value  $>0.05$  and 2-fold enrichment. Red dots are significantly enriched proteins.

5B and Table S6). Interestingly, the high number of significantly enriched proteins leads to a clear separation of the probe-treated and control samples in the PCA plot, with one component over 81% (Figure 5C). Furthermore, the SP2E enrichment of the *O*-GlcNAcylated proteins had a strong impact on the number of imputed values in DMSO-treated control cells ( $\sim 800$ ) compared to around  $\sim 200$  values in  $\text{Ac}_3\text{4dGlcNAz}$  probe-treated cells. Thus, a high number of proteins was consistently identified only in the probe-treated samples pointing toward a low unspecific protein binding to the magnetic beads (Figure 5D). To our contentment, 92% of the 805 significantly enriched proteins were previously described as *O*-GlcNAcylated ([www.oglcnac.mcw.edu](http://www.oglcnac.mcw.edu)), with one of the most significant hits being the well-studied cotranslationally *O*-GlcNAcylated protein NUP62 (Figure 5B–E).<sup>54</sup> To benchmark the enrichment efficiency of the SP2E workflow, we have compared it to the previously used standard workflow, which is based on acetone precipitation and avidin agarose beads. For this purpose, we performed the avidin agarose-based workflow with the same  $\text{Ac}_3\text{4dGlcNAz}$  probe-treated lysates and observed that a substantially lower number of proteins was significantly enriched (123, Figure S6). However, 107 (87%) of the protein hits were significantly enriched by both methods, showing a very good fidelity. In addition to the MS-based experiment, we performed a gel-based experiment to benchmark our workflow. Therefore, both methods were performed with 400  $\mu\text{g}$  of starting protein, but instead of digesting the *O*-GlcNAcylated proteins, they were eluted and separated on SDS-PAGE. The enrichment of NUP62 was visualized by an anti-NUP62 antibody on a western blot to demonstrate the comparability of both methods (Figures S6 and S7). Together, these experiments demonstrate the utility of the SP2E protocol for the

enrichment of metabolically labeled proteins and the use of different bioorthogonal reactions such as SPAAC.

#### Scale-Down of the SP2E Workflow into a 96-Well Plate Format

Although the combination of the carboxylate and streptavidin magnetic beads using SP2E streamlined the PTM protein enrichment, it remained to be demonstrated whether this approach is efficient with a lower protein input. Therefore, we performed the enrichment starting with 100  $\mu\text{g}$  of total protein using lysates from pro-N6pA-treated HeLa cells in a standard 1.5 mL eppendorf tube. In this first attempt, it was already possible to significantly enrich five out of six AMPylation marker proteins (Figure 6A and Table S7). Encouraged by the general feasibility of the SP2E workflow with a lower protein input, we have moved on to scale-down the protocol into a 96-well plate format. The dynamic range of the SP2E enrichment efficiency could be shown in the enrichment of rather low abundant AMPylated proteins from pro-N6pA-treated HeLa cells in a 96-well plate format (Figure 6B). The initial testing with a simple decrease of the click reaction volume to 20  $\mu\text{L}$  and wash steps to 150  $\mu\text{L}$  showed only poor enrichment results (Figure S8). Thus, we have adjusted the protocol with the following steps. The clean up of the proteins after the click reaction was extended with an additional acetonitrile washing step, as used for the automated whole proteome sample preparation by Müller et al.<sup>30</sup> In addition, the reduction and alkylation steps were omitted,<sup>55</sup> proteins were digested in 50  $\mu\text{L}$  of triethylammonium bicarbonate (TEAB), and peptides were eluted from the streptavidin magnetic beads with 20  $\mu\text{L}$  of TEAB and 20  $\mu\text{L}$  of 0.5% formic acid (FA) buffer with incubation at 40  $^\circ\text{C}$  for 5 min (Figure S8). The resulting MS samples have been acidified by the addition of FA and the peptide mixture was resolved using a 60 min LC-MS/MS



measurement (Figure 6B and Table S8). In particular, the use of the shorter gradient is beneficial for two practical reasons. First, more samples can be measured in a shorter time, and second, the MS spectra files are smaller with overall less MS data to the process, leading to faster identification and quantification by search engines. Furthermore, since the total amount of enriched peptides is estimated to be very low, the shorter gradient is likely to result in more intense MS spectra and thus in more identified peptides (Figure 6B). Importantly, the Pearson correlation coefficients of the protein intensities remained over 95% (Figure 6C).

The efficiency of the optimized protocol in a 96-well plate was further tested with the Ac<sub>3</sub>4dGlcNAz probe (Figure 6D). Indeed, it was possible to significantly enrich 168 proteins including NUP62 among the most significantly enriched ones (Figure 6D and Table S9). In total, 511 proteins were identified. Moreover, both TEAB and ABC buffers used for the digest provided comparable results. To elute the peptides efficiently from the streptavidin magnetic beads after digestion, it is necessary to repeat the elution twice, although it results in peptide dilution in the final MS sample (Figure S8). Similar to the large-scale experiment, the SP2E procedure in the 96-well plate format displays an excellent separation in the PCA of the control and probe-treated samples after enrichment (Figure 6E). To compare the enrichment of NUP62 with the large-scale SP2E workflow, we performed the western blotting analysis of the enriched proteins showing the successful detection of NUP62 using the antibody (Figure S7). In summary, our protocol promises to provide a fast, robust, and high-throughput chemical proteomic platform, which may be used by biologists to assess the PTM status of a wide variety of cells. This is an important prerequisite to unraveling the complex PTM networks and elucidating the underlying functional consequences of protein PTMs.

## DISCUSSION

Chemical proteomics has enabled the characterization of many protein PTMs, which are otherwise inaccessible using the whole proteome analysis. Several enrichment workflows have been developed to make the procedure universal and feasible. However, the protocols often require to be carried out by specialized laboratory personnel, they are tedious, and time-consuming. Additionally, with the increasing demand to screen protein PTMs in specialized cell types that are difficult to culture or not accessible in larger amounts typically required by the chemical proteomic protocols, it is of paramount importance to streamline the enrichment workflow and to provide a platform that could be automated. This would parallel the development of the high-throughput automatized whole proteome MS analysis. The transformation of the MS field has been underlined by the rapid improvement in the speed and sensitivity of modern mass spectrometers. The progress of the MS instrumentation has been complemented by software tools allowing for fast and reliable protein identification and quantification. Together, these developments have created a suitable environment for the transition of the chemical proteomic analysis of protein PTMs from a specialized field to a more widely applied analytical tool.

Here, we describe the development and application of the SP2E workflow, which enables the chemical proteomic characterization of protein PTMs in a small-scale, robust, and time-effective manner. The SP2E workflow combines the previously reported protein clean up (SP3) and enrichment

protocols.<sup>13,30,33–36</sup> However, each step of the workflow was optimized to enable a smooth transition between the steps and to achieve high identification rates of the enriched proteins. The main difference in the previous MS-based chemical proteomic protocols is the utilization of the paramagnetic beads for both protein clean up and enrichment. It leads to a better separation of the solid and liquid phase and thus improves the removal of nonspecifically binding proteins during the enrichment steps. Furthermore, it allows a better separation in smaller volumes and it can be readily automated. The initial substitution of the standardly used avidin-coated agarose beads with streptavidin-coated magnetic beads resulted in only moderate enrichment of the AMPylated marker proteins. Therefore, we have systematically evaluated each step of the enrichment protocol with a focus on a possible scale-down of the whole procedure. We started our investigation by improving the lysis buffer composition, which is critical to ensure efficient cell lysis already in small volumes to provide protein concentrations of up to 10  $\mu\text{g}/\mu\text{L}$  and to facilitate efficient click chemistry (Figure 2A). Next, enrichment efficiency was improved by spatial separation of protein clean up and enrichment, which allows for more efficient washing and hence less unspecific protein binding. The optimized protocol has been tested in large-scale experiments starting with 400  $\mu\text{g}$  of total protein to explore metabolic and signaling pathways in which protein AMPylation plays a role. In total, 48 samples have been prepared using five different inhibitors and four replicates per condition. We used the same pro-N6pA probe for the optimization of the workflow but SH-SY5Y neuroblastoma cells instead of HeLa cells. The Pearson correlation coefficient of protein intensities among all samples showed a high correlation ( $>0.95$ ), demonstrating the robustness of the workflow. Interestingly, the PCA revealed a difference between the control and probe-treated cells and importantly displays a clear change in enriched proteins from monensin- and bafilomycin-treated cells, suggesting the specific role of AMPylation in cell stress response to these inhibitors. Moreover, this indicates the high efficiency of the wash steps during the enrichment and reproducibility of the SP2E workflow. Concentration-dependent analysis of monensin on pro-N6pA labeling showed that two replicates of each condition provide sufficient information due to the high reproducibility of the procedure and revealed that the PLD3 modification rapidly increases with a 2 nM monensin concentration.

To show the versatility of the procedure, we have performed the enrichment of O-GlcNAcylated proteins using the azido Ac<sub>3</sub>4dGlcNAz probe. The SPAAC click reaction followed by the SP2E workflow provided excellent enrichment of the well-described glycosylated protein NUP62 with a nominal value of 694-fold enrichment. Furthermore, another 740 known glycosylated proteins were significantly enriched. The outstanding enrichment efficiency is visible by the number of missing values in the control samples (Figure 5D). Since protein glycosylation plays an important role in numerous metabolic processes and often serves as a disease marker, our SP2E protocol offers the possibility to screen for the O-GlcNAcylation in a high-throughput manner. This might not only accelerate the progress in the field but also help to decipher the complex glycan patterns by application of different glycosylation labels.

Finally, the current availability of chemical proteomic data shared in public repositories together with the feasibility of

sophisticated data analyses necessitates the generation of high-quality data in a high-throughput manner. This can only be achieved by automation of the procedures, for example, with the autoSP3 protocol and other proteomic approaches. Here, we have scaled down the SP2E workflow into the 96-well plate format, which retains the principal operations paralleled in autoSP3. Importantly, we showed by the analysis of protein AMPylation and *O*-GlcNAcylation that 100  $\mu\text{g}$  of the input protein is sufficient for successful PTM protein profiling. This results in the same high correlations between the samples (in average >0.96) and achieves high fold enrichments (NUP62 26-fold and HSPA5 2-fold). Although there is a drop in the overall number of enriched proteins, this is outweighed by cost and time efficiency due to the smaller scale of cell culture, washing steps, usage of the multichannel pipettes, shorter measurement times, and data processing. Of note, in the case of AMPylation, 66% of marker proteins were successfully enriched and for *O*-GlcNAcylation, 167 proteins out of 168 significantly enriched proteins were previously described to be *O*-GlcNAcylated. The main advantage of the SP2E workflow and its application in the 96-well plate format is the possibility to scale down the starting biological material and screen beyond a few dozen samples, which was technically not possible using the previous protocol based on the agarose beads. Overall, the manual SP2E workflow with 24 samples in a 96-well format can be carried out in 3.5–4 h starting from the preparation of the click reaction to the addition of trypsin. After overnight trypsin digest, the peptides are eluted and transferred into MS vials within an additional 45 min.

The limitation of the SP2E workflow for chemical proteomic characterization of protein PTMs is the inherent necessity to treat the cells with small-molecule probes, which are not always commercially available and thus need to be synthesized. Furthermore, the labeling ratio is often determined by the metabolic activation of probes and the substrate selectivity of PTM writers. Thus, it might be difficult to estimate the exact stoichiometry of the protein modification. Similar to other proteomic approaches, the statistical evaluation includes the imputation of missing values, and therefore it might be challenging to identify the hit proteins with low protein intensities as significantly enriched. However, some of these bottlenecks can be overcome by increasing the number of replicates.

The small-scale SP2E allows us to screen many biologically relevant conditions simultaneously. By means of this, metabolic switches which are regulated by PTMs can be identified and subjected to deeper and focused analysis in larger scale. The SP2E workflow is therefore an important step toward robotic automation of chemical proteomics and its broad application to explore complex protein PTM networks to solve biological problems.

## METHODS

### Culturing of HeLa and SH-SY5Y Cells

HeLa (RRID: CVCL\_0030) and SH-SY5Y (RRID: CVCL\_0019) cells were cultured in Dulbecco's modified Eagle's medium (DMEM)—high glucose supplemented with 10% fetal calf serum (FCS) and 2 mM *l*-alanyl-*l*-glutamine at 37 °C and a 5% CO<sub>2</sub> atmosphere.

### Probe Treatments

In each 10 cm dish, 2.5 million HeLa or SH-SY5Y cells were seeded in 10 mL of media. Cells were either treated with 10  $\mu\text{L}$  of a probe (100

mM stock pro-N6pA or 200 mM stock Ac<sub>3</sub>4dGlcNAz) or with 10  $\mu\text{L}$  of DMSO as a control. After the addition of the probe and the inhibitor, the cells were incubated for 16 h at 37 °C before harvesting. For harvesting, the cells were washed twice with 2 mL of Dulbecco's PBS (DPBS), scraped into 1 mL of DPBS, and pelleted at 1000 rpm, 4 °C.

### Cell Lysis

Cells were lysed with 500  $\mu\text{L}$  of lysis buffer (20 mM HEPES, pH 7.5, 1% (v/v) NP-40, 0.2% (w/v) SDS) and sonicated for 10 s at a 20% intensity with a rod sonicator. The lysate was clarified at 11,000 rpm at 4 °C for 10 min and the protein concentration was determined by bicinchoninic acid (BCA) assay.

### Measurement of Protein Concentrations

To measure the protein concentrations of the lysates, a bicinchoninic acid assay (Pierce BCA Protein Assay Kit, Thermo Fisher Scientific) was performed. First, bovine serum albumin (BSA) standards with concentrations of 12.5, 25, 50, 100, 200, and 400  $\mu\text{g}/\text{mL}$  were prepared and samples as well as controls were diluted 40 times to a total volume of 200  $\mu\text{L}$ . To measure standards, samples, and controls in triplicates, 50  $\mu\text{L}$  of each was added to three wells of a transparent 96-well plate with a flat bottom. Afterward, 100  $\mu\text{L}$  of a working reagent (2  $\mu\text{L}$  of R2 and 98  $\mu\text{L}$  of R1) was added to each well by a multistepper and the plate was incubated for 15 min at 60 °C. Then, the absorbance at 562 nm was measured by Tecan and the protein concentrations were calculated.

### Large-Scale SP2E Workflow

A total of 400  $\mu\text{g}$  of protein of probe-treated and control lysates was diluted with lysis buffer (20 mM HEPES, pH 7.5, 1% (v/v) NP-40, 0.2% (w/v) SDS) to a 200  $\mu\text{L}$  reaction volume. To each replicate, 2  $\mu\text{L}$  of biotin-PEG-N<sub>3</sub> (10 mM in DMSO), 2  $\mu\text{L}$  of TCEP (100 mM in water), and 0.25  $\mu\text{L}$  of tris((1-benzyl-4-triazolyl)methyl)amine (TBTA) (83.5 mM in DMSO) were added. Samples were gently vortexed, and the click reaction was initiated by the addition of 4  $\mu\text{L}$  of CuSO<sub>4</sub> (50 mM in water) and incubated for 1.5 h (room temperature (rt), 600 rpm). Subsequently, 200  $\mu\text{L}$  of 8 M urea was added to each replicate. A total of 100  $\mu\text{L}$  of mixed hydrophobic and hydrophilic carboxylate-coated magnetic beads (1:1) was washed thrice with 500  $\mu\text{L}$  of water. The click reaction mixture was directly transferred onto the equilibrated carboxylate-coated magnetic beads, resuspended, and 600  $\mu\text{L}$  of ethanol was added. After resuspending the beads via vortexing, the suspension was incubated for 5 min at rt and 950 rpm. The beads were washed thrice with 500  $\mu\text{L}$  of 80% ethanol in water using a magnetic rack and the proteins were separately eluted by the addition of 0.5 mL of 0.2% SDS in PBS. For this, the beads were resuspended, incubated for 5 min at 950 rpm, rt, and the supernatant was directly transferred onto 50  $\mu\text{L}$  of equilibrated streptavidin-coated magnetic beads (three times prewashed with 500  $\mu\text{L}$  of 0.2% SDS in PBS). The procedure was repeated once and the supernatants were combined and incubated for 1 h, rt and 950 rpm for biotin/streptavidin binding. The streptavidin-coated magnetic bead mixture was washed thrice with 500  $\mu\text{L}$  of 0.1% NP-40 in PBS, twice with 500  $\mu\text{L}$  of 6 M urea, and twice with 500  $\mu\text{L}$  of water. Washed bead mixtures were resuspended in 80  $\mu\text{L}$  of 125 mM ABC buffer, and proteins were reduced and alkylated by the addition of 10  $\mu\text{L}$  of 100 mM TCEP and 10  $\mu\text{L}$  of 400 mM chloracetamide, followed by 5 min incubation at 95 °C. Proteins were digested overnight at 37 °C with 1.5  $\mu\text{L}$  of sequencing grade trypsin (0.5 mg/mL). The following day, the beads were washed thrice with 100  $\mu\text{L}$  of 100 mM ABC buffer and the supernatants were combined and acidified with 2  $\mu\text{L}$  of formic acid. Peptides were desalted using 50 mg of Sep-Pak C18 cartridges on a vacuum manifold. The columns were equilibrated with 1 mL of acetonitrile, 1 mL of elution buffer (80% acetonitrile with 0.5% formic acid in water), and 3 mL of wash buffer (0.5% formic acid in water). Subsequently, the samples were loaded on the cartridges and washed with 3 mL of wash buffer. The peptides were eluted two times with 250  $\mu\text{L}$  of elution buffer and vacuum-dried with a SpeedVac at 35 °C. Finally, dried peptides were

reconstituted in 30  $\mu\text{L}$  of 1% formic acid in water by vortexing and sonication (15 min) and transferred to an MS vial.

### Small-Scale SP2E Workflow

A total of 100  $\mu\text{g}$  of protein of probe-treated and control lysates was diluted with lysis buffer (20 mM HEPES, pH 7.5, 1% (v/v) NP-40, 0.2% (w/v) SDS) to a 19  $\mu\text{L}$  reaction volume in a 96-well plate. The master mix containing 0.2  $\mu\text{L}$  of biotin-PEG- $\text{N}_3$  (10 mM in DMSO), 0.2  $\mu\text{L}$  of TCEP (100 mM in water), and 0.125  $\mu\text{L}$  of TBTA (16.7 mM in DMSO) was added to each sample. Samples were gently vortexed and the click reaction was initiated by the addition of 0.4  $\mu\text{L}$  of  $\text{CuSO}_4$  (50 mM in water) and incubated for 1.5 h (rt, 600 rpm). Subsequently, 60  $\mu\text{L}$  of 8 M urea was added to each replicate. A total of 100  $\mu\text{L}$  of mixed hydrophobic and hydrophilic carboxylate-coated magnetic beads (1:1) were washed thrice with 100  $\mu\text{L}$  of water. The click reaction mixture was directly transferred onto the equilibrated carboxylate-coated magnetic beads, resuspended, and 100  $\mu\text{L}$  of absolute ethanol was added. After resuspending the beads via vortexing, the suspension was incubated for 5 min at rt and 950 rpm. The beads were washed thrice with 150  $\mu\text{L}$  of 80% ethanol in water using a magnetic rack and once with 150  $\mu\text{L}$  of acetonitrile (LC-MS). Proteins were separately eluted by the addition of 60  $\mu\text{L}$  of 0.2% SDS in PBS. For this, the beads were resuspended, incubated for 5 min at 40  $^\circ\text{C}$  and 950 rpm, and the supernatant was directly transferred onto 50  $\mu\text{L}$  of equilibrated streptavidin-coated magnetic beads (three times prewashed with 100  $\mu\text{L}$  of 0.2% SDS in PBS). The procedure was repeated twice, and the supernatants were combined and incubated for 1 h at rt and 800 rpm for biotin/streptavidin binding. The streptavidin-coated magnetic bead mixture was washed thrice with 150  $\mu\text{L}$  of 0.1% NP-40 in PBS, twice with 150  $\mu\text{L}$  of 6 M urea, and twice with 150  $\mu\text{L}$  of water. For each washing step, the beads were incubated 1 min at rt and 800 rpm. Washed bead mixtures were resuspended in 50  $\mu\text{L}$  of 50 mM TEAB and proteins were digested overnight at 37  $^\circ\text{C}$  by addition of 1.5  $\mu\text{L}$  of sequencing grade trypsin (0.5 mg/mL). The following day, the beads were washed twice with 20  $\mu\text{L}$  of 50 mM TEAB buffer and twice with 20  $\mu\text{L}$  of 0.5% FA and the wash fractions were collected and combined. For each washing step, the beads were incubated for 5 min at 40  $^\circ\text{C}$  and 600 rpm. The combined washed fractions were acidified by the addition of 0.9  $\mu\text{L}$  of formic acid (FA) and transferred directly without desalting to an MS vial.

### Large-Scale SPAAC Protocol

A total of 400  $\mu\text{g}$  of protein of probe-treated and control lysates was diluted with lysis buffer (20 mM HEPES, pH 7.5, 1% (v/v) NP-40, 0.2% (w/v) SDS) to a 200  $\mu\text{L}$  reaction volume. To each replicate, 3  $\mu\text{L}$  of 1 M IAA in water was added and incubated for 30 min at 750 rpm, 25  $^\circ\text{C}$ . Next, 2  $\mu\text{L}$  of a 2 mM DBCO-PEG- $\text{N}_3$  reagent was added to initiate the SPAAC reaction. The reaction mixtures were incubated at 25  $^\circ\text{C}$ , 750 rpm for 30 min. The samples proceeded further in the same way as for CuAAC.

### Small-Scale SPAAC Protocol

A total of 100  $\mu\text{g}$  of protein of probe-treated and control lysates was diluted with lysis buffer (20 mM HEPES, pH 7.5, 1% (v/v) NP-40, 0.2% (w/v) SDS) to a 19  $\mu\text{L}$  reaction volume in a 96-well plate. To each replicate, 0.3  $\mu\text{L}$  of 1 M IAA in water was added and incubated for 30 min at 750 rpm, 25  $^\circ\text{C}$ . Next, 0.2  $\mu\text{L}$  of the 2 mM DBCO-PEG- $\text{N}_3$  reagent was added to initiate the SPAAC reaction. The reaction mixtures were incubated at 25  $^\circ\text{C}$ , 750 rpm for 30 min. The samples proceeded further in the same way as for CuAAC.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/jacsau.2c00284>.

Supporting figures, cell culture conditions, SP2E workflow, LC-MS/MS acquisition conditions, data analysis, and list of reagents (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

Pavel Kielkowski – Institute for Chemical Epigenetics Munich, LMU Munich, 81375 Munich, Germany; [orcid.org/0000-0003-4910-6263](https://orcid.org/0000-0003-4910-6263); Email: [pavel.kielkowski@cup.lmu.de](mailto:pavel.kielkowski@cup.lmu.de)

### Authors

Tobias Becker – Institute for Chemical Epigenetics Munich, LMU Munich, 81375 Munich, Germany

Andreas Wiest – Institute for Chemical Epigenetics Munich, LMU Munich, 81375 Munich, Germany; [orcid.org/0000-0003-2525-6304](https://orcid.org/0000-0003-2525-6304)

András Telek – Institute for Chemical Epigenetics Munich, LMU Munich, 81375 Munich, Germany

Daniel Bejko – Institute for Chemical Epigenetics Munich, LMU Munich, 81375 Munich, Germany

Anja Hoffmann-Röder – Department of Chemistry, LMU Munich, 81377 Munich, Germany; [orcid.org/0000-0002-8352-8323](https://orcid.org/0000-0002-8352-8323)

Complete contact information is available at:

<https://pubs.acs.org/doi/10.1021/jacsau.2c00284>

### Author Contributions

P.K. designed the method; A.W., A.T., and D.B. performed the experiments and contributed to the optimization of the method; T.B. carried out the AMPylation study; A.H.-R. provided the  $\text{Ac}_3\text{4dGlcNAz}$  probe and contributed to  $\text{O-GlcNAcylation}$ ; T.B. and P.K. overviewed the study and analyzed the data; and T.B. and P.K. wrote the manuscript with input from all authors.

### Notes

The authors declare no competing financial interest. Proteomics data are freely available at ProteomeXchange Consortium via PRIDE partner repository with the data set identifier PXD033644. Tables S1–S14 are available via figshare data repository DOI: 10.6084/m9.figshare.19969907.

## ■ ACKNOWLEDGMENTS

This research project was supported by Liebig fellowship from VCI to P.K. and T.B., LMUexcellence Junior Fund to P.K., and generous support from SFB1309 by DFG. The authors thank Stefan Marchner (LMU Munich) for the preparation of the  $\text{Ac}_3\text{4dGlcNAz}$  probe and Dietrich Mostert (Technical University of Munich) for proofreading.

## ■ REFERENCES

- (1) Becker, T.; Cappel, C.; Matteo, F. D.; Sonsalla, G.; Kaminska, E.; Spada, F.; Cappello, S.; Damme, M.; Kielkowski, P. AMPylation Profiling during Neuronal Differentiation Reveals Extensive Variation on Lysosomal Proteins. *iScience* **2021**, *24*, No. 103521.
- (2) Mansfield, S. G.; Gordon-Weeks, P. R. Dynamic Post-Translational Modification of Tubulin in Rat Cerebral Cortical Neurons Extending Neurites in Culture: Effects of Taxol. *J. Neurocytol.* **1991**, *20*, 654–666.
- (3) Zheng, P.; Obara, C. J.; Szczesna, E.; Nixon-Abell, J.; Mahalingam, K. K.; Roll-Mecak, A.; Lippincott-Schwartz, J.; Blackstone, C. ER Proteins Decipher the Tubulin Code to Regulate Organelle Distribution. *Nature* **2022**, 132–138.
- (4) Aebersold, R.; Agar, J. N.; Amster, I. J.; Baker, M. S.; Bertozzi, C. R.; Boja, E. S.; Costello, C. E.; Cravatt, B. F.; Fenselau, C.; Garcia, B. A.; Ge, Y.; Gunawardena, J.; Hendrickson, R. C.; Hergenrother, P. J.;

- Huber, C. G.; Ivanov, A. R.; Jensen, O. N.; Jewett, M. C.; Kelleher, N. L.; Kiessling, L. L.; Krogan, N. J.; Larsen, M. R.; Loo, J. A.; Loo, R. R. O.; Lundberg, E.; MacCoss, M. J.; Mallick, P.; Mootha, V. K.; Mrksich, M.; Muir, T. W.; Patrie, S. M.; Pesavento, J. J.; Pitteri, S. J.; Rodriguez, H.; Saghatelian, A.; Sandoval, W.; Schlüter, H.; Sechi, S.; Slavoff, S. A.; Smith, L. M.; Snyder, M. P.; Thomas, P. M.; Uhlén, M.; Eyk, J. E. V.; Vidal, M.; Walt, D. R.; White, F. M.; Williams, E. R.; Wohlschlagler, T.; Wysocki, V. H.; Yates, N. A.; Young, N. L.; Zhang, B. How Many Human Proteoforms Are There? *Nat. Chem. Biol.* **2018**, *14*, 206–214.
- (5) Brüning, F.; Noya, S. B.; Bange, T.; Koutsouli, S.; Rudolph, J. D.; Tyagarajan, S. K.; Cox, J.; Mann, M.; Brown, S. A.; Robles, M. S. Sleep-Wake Cycles Drive Daily Dynamics of Synaptic Phosphorylation. *Science* **2019**, *366*, No. eaav3617.
- (6) Truttmann, M. C.; Pincus, D.; Ploegh, H. L. Chaperone AMPylation Modulates Aggregation and Toxicity of Neurodegenerative Disease-Associated Polypeptides. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, No. 201801989.
- (7) Rogowski, K.; van Dijk, J.; Magiera, M. M.; Bosc, C.; Deloulme, J.-C.; Bosson, A.; Peris, L.; Gold, N. D.; Lacroix, B.; Grau, M. B.; Bec, N.; Larroque, C.; Desagher, S.; Holzer, M.; Andrieux, A.; Moutin, M.-J.; Janke, C. A Family of Protein-Deglutamylating Enzymes Associated with Neurodegeneration. *Cell* **2010**, *143*, 564–578.
- (8) Hoch, N. C.; Polo, L. M. ADP-Ribosylation: From Molecular Mechanisms to Human Disease. *Genet. Mol. Biol.* **2020**, *43*, No. e20190075.
- (9) Kam, T.-I.; Mao, X.; Park, H.; Chou, S.-C.; Karuppagounder, S. S.; Umanah, G.; Yun, S.; Brahmachari, S.; Panicker, N.; Chen, R.; Andrabi, S. A.; Qi, C.; Poirier, G. G.; Pletnikova, O.; Troncoso, J. C.; Bekris, L. M.; Leverenz, J. B.; Pantelyat, A.; Ko, H.; Rosenthal, L. S.; Dawson, T. M.; Dawson, V. L. Poly(ADP-Ribose) Drives Pathologic  $\alpha$ -Synuclein Neurodegeneration in Parkinson's Disease. *Science* **2018**, *362*, No. eaat8407.
- (10) Smith, L. M.; Thomas, P. M.; Shortreed, M. R.; Schaffer, L. V.; Fellner, R. T.; LeDuc, R. D.; Tucholski, T.; Ge, Y.; Agar, J. N.; Anderson, L. C.; Chamot-Rooke, J.; Gault, J.; Loo, J. A.; Paša-Tolić, L.; Robinson, C. V.; Schlüter, H.; Tsybin, Y. O.; Vilaseca, M.; Vizcaíno, J. A.; Danis, P. O.; Kelleher, N. L. A Five-Level Classification System for Proteoform Identifications. *Nat. Method* **2019**, *16*, 939–940.
- (11) Smith, L. M.; Agar, J. N.; Chamot-Rooke, J.; Danis, P. O.; Ge, Y.; Loo, J. A.; Paša-Tolić, L.; Tsybin, Y. O.; Kelleher, N. L.; The Consortium for Top-Down Proteomics. The Human Proteoform Project: Defining the Human Proteome. *Sci. Adv.* **2021**, *7*, No. eabk0734.
- (12) Laughlin, S. T.; Bertozzi, C. R. Metabolic Labeling of Glycans with Azido Sugars and Subsequent Glycan-Profiling and Visualization via Staudinger Ligation. *Nat. Protoc.* **2007**, *2*, 2930–2944.
- (13) Kallemeijn, W. W.; Lanyon-Hogg, T.; Panyain, N.; Grocin, A. G.; Ciepla, P.; Morales-Sanfrutos, J.; Tate, E. W. Proteome-Wide Analysis of Protein Lipidation Using Chemical Probes: In-Gel Fluorescence Visualization, Identification and Quantification of N-Myristoylation, N- and S-Acylation, O-Cholesterylation, S-Farnesylation and S-Geranylgeranylation. *Nat. Protoc.* **2021**, *16*, 5083–5122.
- (14) Martin, B. R.; Wang, C.; Adibekian, A.; Tully, S. E.; Cravatt, B. F. Global Profiling of Dynamic Protein Palmitoylation. *Nat. Method* **2012**, *9*, 84–89.
- (15) Parker, C. G.; Pratt, M. R. Click Chemistry in Proteomic Investigations. *Cell* **2020**, *180*, 605–632.
- (16) Kielkowski, P.; Buchsbaum, I. Y.; Becker, T.; Bach, K.; Cappello, S.; Sieber, S. A. A Pronucleotide Probe for Live-Cell Imaging of Protein AMPylation. *ChemBioChem* **2020**, *21*, 1285–1287.
- (17) Kielkowski, P.; Buchsbaum, I. Y.; Kirsch, V. C.; Bach, N. C.; Drukker, M.; Cappello, S.; Sieber, S. A. FICD Activity and AMPylation Remodelling Modulate Human Neurogenesis. *Nat. Commun.* **2020**, *11*, No. 517.
- (18) Sinha, A.; Mann, M. A Beginner's Guide to Mass Spectrometry-Based Proteomics. *Biochemist* **2020**, *42*, 64–69.
- (19) Cox, J.; Mann, M. MaxQuant Enables High Peptide Identification Rates, Individualized p.p.b.-Range Mass Accuracies and Proteome-Wide Protein Quantification. *Nat. Biotechnol.* **2008**, *26*, 1367–1372.
- (20) Tyanova, S.; Temu, T.; Sinitcyn, P.; Carlson, A.; Hein, M. Y.; Geiger, T.; Mann, M.; Cox, J. The Perseus Computational Platform for Comprehensive Analysis of (Prote)Omics Data. *Nat. Method* **2016**, *13*, 731–740.
- (21) Yu, F.; Teo, G. C.; Kong, A. T.; Haynes, S. E.; Avtonomov, D. M.; Geiszler, D. J.; Nesvizhskii, A. I. Identification of Modified Peptides Using Localization-Aware Open Search. *Nat. Commun.* **2020**, *11*, No. 4065.
- (22) Grammel, M.; Luong, P.; Orth, K.; Hang, H. C. A Chemical Reporter for Protein AMPylation. *J. Am. Chem. Soc.* **2011**, *133*, 17103–17105.
- (23) Kliza, K. W.; Liu, Q.; Roosenboom, L. W. M.; Jansen, P. W. T. C.; Filippov, D. V.; Vermeulen, M. Reading ADP-Ribosylation Signaling Using Chemical Biology and Interaction Proteomics. *Mol. Cell* **2021**, *81*, 4552.e8–4567.e8.
- (24) Yang, Y.-Y.; Ascano, J. M.; Hang, H. C. Bioorthogonal Chemical Reporters for Monitoring Protein Acetylation. *J. Am. Chem. Soc.* **2010**, *132*, 3640–3641.
- (25) Sieber, S. A.; Cappello, S.; Kielkowski, P. From Young to Old: AMPylation Hits the Brain. *Cell. Chem. Biol.* **2020**, *27*, 773–779.
- (26) Agard, N. J.; Prescher, J. A.; Bertozzi, C. R. A Strain-Promoted [3 + 2] Azide–Alkyne Cycloaddition for Covalent Modification of Biomolecules in Living Systems. *J. Am. Chem. Soc.* **2004**, *126*, 15046–15047.
- (27) Zecha, J.; Satpathy, S.; Kanashova, T.; Avanesian, S. C.; Kane, M. H.; Clauser, K. R.; Mertins, P.; Carr, S. A.; Kuster, B. TMT Labeling for the Masses: A Robust and Cost-Efficient, In-Solution Labeling Approach\* [S]. *Mol. Cell. Proteomics* **2019**, *18*, 1468–1478.
- (28) Cox, J.; Hein, M. Y.; Lubner, C. A.; Paron, I.; Nagaraj, N.; Mann, M. Accurate Proteome-Wide Label-Free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. *Mol. Cell. Proteomics* **2014**, *13*, 2513–2526.
- (29) Wiśniewski, J. R.; Zougman, A.; Nagaraj, N.; Mann, M. Universal Sample Preparation Method for Proteome Analysis. *Nat. Method* **2009**, *6*, 359–362.
- (30) Müller, T.; Kalxdorf, M.; Longuespée, R.; Kazdal, D. N.; Stenzinger, A.; Krijgsveld, J. Automated Sample Preparation with SP3 for Low-input Clinical Proteomics. *Mol. Syst. Biol.* **2020**, *16*, No. e9111.
- (31) Hughes, C. S.; Moggridge, S.; Müller, T.; Sorensen, P. H.; Morin, G. B.; Krijgsveld, J. Single-Pot, Solid-Phase-Enhanced Sample Preparation for Proteomics Experiments. *Nat. Protoc.* **2019**, *14*, 68–85.
- (32) Sielaff, M.; Kuharev, J.; Bohn, T.; Hahlbrock, J.; Bopp, T.; Tenzer, S.; Distler, U. Evaluation of FASP, SP3, and IST Protocols for Proteomic Sample Preparation in the Low Microgram Range. *J. Proteome Res.* **2017**, *16*, 4060–4072.
- (33) Wright, M. H.; Clough, B.; Rackham, M. D.; Rangachari, K.; Brannigan, J. A.; Grainger, M.; Moss, D. K.; Bottrill, A. R.; Heal, W. P.; Broncel, M.; Serwa, R. A.; Brady, D.; Mann, D. J.; Leatherbarrow, R. J.; Tewari, R.; Wilkinson, A. J.; Holder, A. A.; Tate, E. W. Validation of N-Myristoyltransferase as an Antimalarial Drug Target Using an Integrated Chemical Biology Approach. *Nat. Chem.* **2014**, *6*, 112–121.
- (34) Yan, T.; Desai, H. S.; Boatner, L. M.; Yen, S. L.; Cao, J.; Palafox, M. F.; Jami-Alahmadi, Y.; Backus, K. SP3-FAIMS Chemo-proteomics for High Coverage Profiling of the Human Cysteineome. *ChemBioChem* **2021**, *22*, 1841–1851.
- (35) Klont, F.; Kwiatkowski, M.; Faiz, A.; van den Bosch, T.; Pouwels, S. D.; Dekker, F. J.; ten Hacken, N. H. T.; Horvatovich, P.; Bischoff, R. Adsorptive Microtiter Plates As Solid Supports in Affinity Purification Workflows. *J. Proteome Res.* **2021**, *20*, 5218–5221.
- (36) Makowski, M. M.; Gräwe, C.; Foster, B. M.; Nguyen, N. V.; Bartke, T.; Vermeulen, M. Global Profiling of Protein–DNA and

Protein–Nucleosome Binding Affinities Using Quantitative Mass Spectrometry. *Nat. Commun.* **2018**, *9*, No. 1653.

(37) Tian, Y.-P.; Zhang, X.-J.; Wu, J.-Y.; Fun, H.-K.; Jiang, M.-H.; Xu, Z.-Q.; Usman, A.; Chantrapromma, S.; Thompson, L. K. Structural Diversity and Properties of a Series of Dinuclear and Mononuclear Copper(II) and Copper(I) Carboxylate Complexes. *New J. Chem.* **2002**, *26*, 1468–1473.

(38) Stojceva Radovanovic, B. C.; Premovic, P. I. Thermal Behaviour of Cu(II)-Urea Complex. *J. Therm. Anal.* **1992**, *38*, 715–719.

(39) Hughes, C. S.; Foehr, S.; Garfield, D. A.; Furlong, E. E.; Steinmetz, L. M.; Krijgsveld, J. Ultrasensitive Proteome Analysis Using Paramagnetic Bead Technology. *Mol. Syst. Biol.* **2014**, *10*, 757.

(40) Truttmann, M. C.; Zheng, X.; Hanke, L.; Damon, J. R.; Grootveld, M.; Krakowiak, J.; Pincus, D.; Ploegh, H. L. Unrestrained AMPylation Targets Cytosolic Chaperones and Activates the Heat Shock Response. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, E152–E160.

(41) Sanyal, A.; Dutta, S.; Camara, A.; Chandran, A.; Koller, A.; Watson, B. G.; Sengupta, R.; Ysselstein, D.; Montenegro, P.; Cannon, J.; Rochet, J.-C.; Mattoo, S. Alpha-Synuclein Is a Target of Fic-Mediated Adenylation/AMPylation: Possible Implications for Parkinson's Disease. *J. Mol. Biol.* **2019**, *431*, 2266–2282.

(42) Raught, B.; Gingras, A.-C.; Sonenberg, N. The Target of Rapamycin (TOR) Proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 7037–7044.

(43) Leeman, D. S.; Hebestreit, K.; Ruetz, T.; Webb, A. E.; McKay, A.; Pollina, E. A.; Dulken, B. W.; Zhao, X.; Yeo, R. W.; Ho, T. T.; Mahmoudi, S.; Devarajan, K.; Passequé, E.; Rando, T. A.; Frydman, J.; Brunet, A. Lysosome Activation Clears Aggregates and Enhances Quiescent Neural Stem Cell Activation during Aging. *Science* **2018**, *359*, 1277–1283.

(44) Yoshimori, T.; Yamamoto, A.; Moriyama, Y.; Futai, M.; Tashiro, Y. Bafilomycin A1, a Specific Inhibitor of Vacuolar-Type H(+)-ATPase, Inhibits Acidification and Protein Degradation in Lysosomes of Cultured Cells. *J. Biol. Chem.* **1991**, *266*, 17707–17712.

(45) Zhang, J.-G.; Fariss, M. W. Thenoyltrifluoroacetone, a Potent Inhibitor of Carboxylesterase Activity. *Biochem. Pharmacol.* **2002**, *63*, 751–754.

(46) Mollenhauer, H. H.; Morré, D. J.; Rowe, L. D. Alteration of Intracellular Traffic by Monensin; Mechanism, Specificity and Relationship to Toxicity. *Biochim. Biophys. Acta, Rev. Biomembr.* **1990**, *1031*, 225–246.

(47) Pohlmann, R.; Krüger, S.; Hasilik, A.; von Figura, K. Effect of Monensin on Intracellular Transport and Receptor-Mediated Endocytosis of Lysosomal Enzymes. *Biochem. J.* **1984**, *217*, 649–658.

(48) Wang, X.; Wu, X.; Zhang, Z.; Ma, C.; Wu, T.; Tang, S.; Zeng, Z.; Huang, S.; Gong, C.; Yuan, C.; Zhang, L.; Feng, Y.; Huang, B.; Liu, W.; Zhang, B.; Shen, Y.; Luo, W.; Wang, X.; Liu, B.; Lei, Y.; Ye, Z.; Zhao, L.; Cao, D.; Yang, L.; Chen, X.; Haydon, R. C.; Luu, H. H.; Peng, B.; Liu, X.; He, T.-C. Monensin Inhibits Cell Proliferation and Tumor Growth of Chemo-Resistant Pancreatic Cancer Cells by Targeting the EGFR Signaling Pathway. *Sci. Rep.* **2018**, *8*, No. 17914.

(49) Long, J. M.; Holtzman, D. M. Alzheimer Disease: An Update on Pathobiology and Treatment Strategies. *Cell* **2019**, *179*, 312–339.

(50) Yang, X.; Qian, K. Protein O-GlcNAcylation: Emerging Mechanisms and Functions. *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 452–465.

(51) Li, J.; Wang, J.; Wen, L.; Zhu, H.; Li, S.; Huang, K.; Jiang, K.; Li, X.; Ma, C.; Qu, J.; Parameswaran, A.; Song, J.; Zhao, W.; Wang, P. G. An OGA-Resistant Probe Allows Specific Visualization and Accurate Identification of O-GlcNAc-Modified Proteins in Cells. *ACS Chem. Biol.* **2016**, *11*, 3002–3006.

(52) Pedowitz, N. J.; Pratt, M. R. Design and Synthesis of Metabolic Chemical Reporters for the Visualization and Identification of Glycoproteins. *RSC Chem. Biol.* **2021**, *2*, 306–321.

(53) Zhang, C.; Zhang, C.; Dai, P.; Vinogradov, A.; Gates, Z. Site-Selective Cysteine-Cyclooctyne Conjugation. *Angew. Chem., Int. Ed.* **2018**, *57*, 6459–6463.

(54) Wulff-Fuentes, E.; Berendt, R. R.; Massman, L.; Danner, L.; Malard, F.; Vora, J.; Kahsay, R.; Stichelen, S. O.-V. The Human O-GlcNAcome Database and Meta-Analysis. *Sci. Data* **2021**, *8*, No. 25.

(55) Petelski, A. A.; Emmott, E.; Leduc, A.; Huffman, R. G.; Specht, H.; Perlman, D. H.; Slavov, N. Multiplexed Single-Cell Proteomics Using SCoPE2. *Nat. Protoc.* **2021**, *16*, 5398–5425.

## Recommended by ACS

### Generic Plug-and-Play Strategy for High-Throughput Analysis of PTM-Mediated Protein Complexes

Yunqiu Qin, Ruijun Tian, *et al.*

APRIL 26, 2022  
ANALYTICAL CHEMISTRY

READ 

### Novel Strategies to Address the Challenges in Top-Down Proteomics

Jake A. Melby, Ying Ge, *et al.*

MAY 13, 2021  
JOURNAL OF THE AMERICAN SOCIETY FOR MASS SPECTROMETRY

READ 

### Mass Spectral Signatures of Complex Post-Translational Modifications in Proteins: A Proof-of-Principle Based on X-ray Irradiated Vancomycin

Marwa Abdelmouleh, Jean-Christophe Pouilly, *et al.*

JUNE 18, 2020  
JOURNAL OF THE AMERICAN SOCIETY FOR MASS SPECTROMETRY

READ 

### Prosit-TMT: Deep Learning Boosts Identification of TMT-Labeled Peptides

Wassim Gabriel, Mathias Wilhelm, *et al.*

MAY 12, 2022  
ANALYTICAL CHEMISTRY

READ 

Get More Suggestions >

### 3.3 Clickable report tags for identification of modified peptides by mass spectrometry

Makarov, D., Telek, A., Becker, T., von Wrisberg, M.K., Schneider, S., and Kielkowski, P. (2022). Clickable report tags for identification of modified peptides by mass spectrometry. *Journal of Mass Spectrometry* 57. 10.1002/jms.4812.

#### Prologue

PTMs play an important role in regulating the activity, localization and interaction of cellular enzymes. In order to analyze a specific PTM in detail, the site of the PTM has to be identified in advance to test protein point mutants for functional studies. However, the site identification of low abundant PTMs, such as AMPylation, remains challenging due to the lack of satisfactory techniques. Hence, the aim of this study was to develop novel MS-tags that facilitate the site identification in bottom-up proteomics. The idea of the approach was to combine isobaric labeling with PTM analogs used in chemical proteomics. More specifically, the PTM analogs bear an alkyne residue which can be coupled by CuAAC to MS-tags containing both an isobaric label and an azide residue. Thereby, the peptide of interest is modified with the isobaric label that generates a reporter ion upon MS/MS fragmentation, improving the identification and quantification of modified peptides. The structure of one of the novel MS-tags synthesized in this study is based on the 2,6-dimethylpiperidine group (DMP) used in commercially available TMT-tags.<sup>99</sup> In contrast, the second MS-tag is based on the recently reported sulfoxide-containing (SOX) tag.<sup>100</sup> First, the fragmentation properties of both were investigated using recombinant BSA labeled with iodoacetamide alkyne (IAA). The MS<sup>2</sup> spectra of both MS-tags showed the expected reporter ions after HCD fragmentation and the DMP-tag also for electron-dissociation (ETD) fragmentation. Subsequently, the ability of the MS-tags to identify reactive cysteines in peptides from HeLa whole proteome samples was analyzed. MaxQuant and MSFragger analysis revealed that more modified peptides were identified with the DMP-tag than with the SOX-tag, but for both, 99% of all MS<sup>2</sup> spectra contained the corresponding reporter ion. In the next step, the performance of the DMP-tag to study AMPylation was analyzed using *in vitro* AMPylated Rab1b. The DMP-AMP peptide showed the same fragmentation properties as the endogenously AMPylated peptide, but with an additional reporter ion. Finally, an anti-DMP antibody was tested to enrich DMP-AMP modified peptides from pro-N6pA treated cell lysates. However, although AMPylated peptides were identified there was no overlap with previously found AMPylated proteins using chemical proteomics.<sup>2</sup>

Taken together, we developed two novel MS-tags with the potential to facilitate the identification of PTM peptides.

### **Author contribution**

In this study, I performed the affinity enrichment of AMP-DMP modified peptides with an anti-DMP antibody. *Pavel Kielkowski* conceived the study, assisted with intact MS-measurements and wrote the manuscript. *Dmytro Makarov* synthesized the DMP-tag as well as the SOX-tag and participated in MS-measurements. *András Telek* carried out the MS-sample preparation and MS-measurements to validate the MS-tags. *Marie-Kristin von Wrisberg* provided the *in vitro* AMPylated Rab1b.

### **License**

Copy of the open access publication based on the Creative Commons Attribution-NonCommercial License (CC BY-NC 4.0), which allows the non-commercial reprint of the article.

### **Article available at:**





<https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/epdf/10.1002/jms.4812>

### **License available at:**

<https://creativecommons.org/licenses/by-nc/4.0/>

Supporting information are available in the appendix (chapter 8.3)

# Clickable report tags for identification of modified peptides by mass spectrometry

Dmytro Makarov<sup>1</sup>  | András Telek<sup>1,2</sup> | Tobias Becker<sup>1</sup> |  
Marie-Kristin von Wrisberg<sup>1</sup>  | Sabine Schneider<sup>1</sup>  | Pavel Kielkowski<sup>1</sup> 

<sup>1</sup>Department of Chemistry, LMU Munich, Munich, Germany

<sup>2</sup>Department of Applied Biotechnology, Budapest University of Technology and Economics, Budapest, Hungary

## Correspondence

Pavel Kielkowski, Department of Chemistry, LMU Munich, Butenandtstr. 5-13, 81377 Munich, Germany.  
Email: pavel.kielkowski@cup.lmu.de

## Funding information

Deutsche Forschungsgemeinschaft, Grant/Award Number: SFB1309; Fonds der Chemischen Industrie

## Abstract

The identification and quantification of modified peptides are critical for the functional characterization of post-translational protein modifications (PTMs) to elucidate their biological function. Nowadays, quantitative mass spectrometry coupled with various bioinformatic pipelines has been successfully used for the determination of a wide range of PTMs. However, direct characterization of low abundant protein PTMs in bottom-up proteomic workflow remains challenging. Here, we present the synthesis and evaluation of tandem mass spectrometry tags (TMT) which are introduced via click-chemistry into peptides bearing alkyne handles. The fragmentation properties of the two mass tags were validated and used for screening in a model system and analysis of AMPylated proteins. The presented tags provide a valuable tool for diagnostic peak generation to increase confidence in the identification of modified peptides and potentially for direct peptide-PTM quantification from various experimental conditions.

## KEYWORDS

AMPylation, chemical proteomics, MS-tags, protein post-translational modifications, reporter ions

## 1 | INTRODUCTION

Post-translational protein modifications play a critical role in many cellular functions.<sup>1</sup> This creates numerous PTM proteins or so-called proteoforms, which largely exceed the number of encoded genes and generates an extraordinary diversity of protein properties.<sup>2</sup> However, techniques to confidently quantify and identify the site of modification are missing. This is, in particular, a challenging issue for low abundant and unstable PTMs such as AMPylation.<sup>3-8</sup> Although the number of available linkers for enrichment complemented by various chemical proteomic approaches is quite large, there is a vacancy of the linker

that would yield a reporter ion upon MS/MS fragmentation and thus improve the site identification rates.<sup>9-11</sup> So far, isobaric labeling has been mainly used for protein quantification in bottom-up proteomics.<sup>12</sup> The large-scale employment of mass spectrometry-based proteomics has taken off. The isobaric strategies have allowed for the multiplication of the sample's measurement to minimize the measurement time while providing precise quantification of proteins prepared under different conditions. In parallel, chemical proteomic strategies have been utilized for the identification of PTM proteins using small compound PTM analogs containing an alkyne handle that allows the downstream enrichment of the modified proteins.<sup>13-18</sup> It remained challenging to quantify and compare PTM stoichiometry between conditions because modified and unmodified peptides displayed

Dmytro Makarov and András Telek contributed equally.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

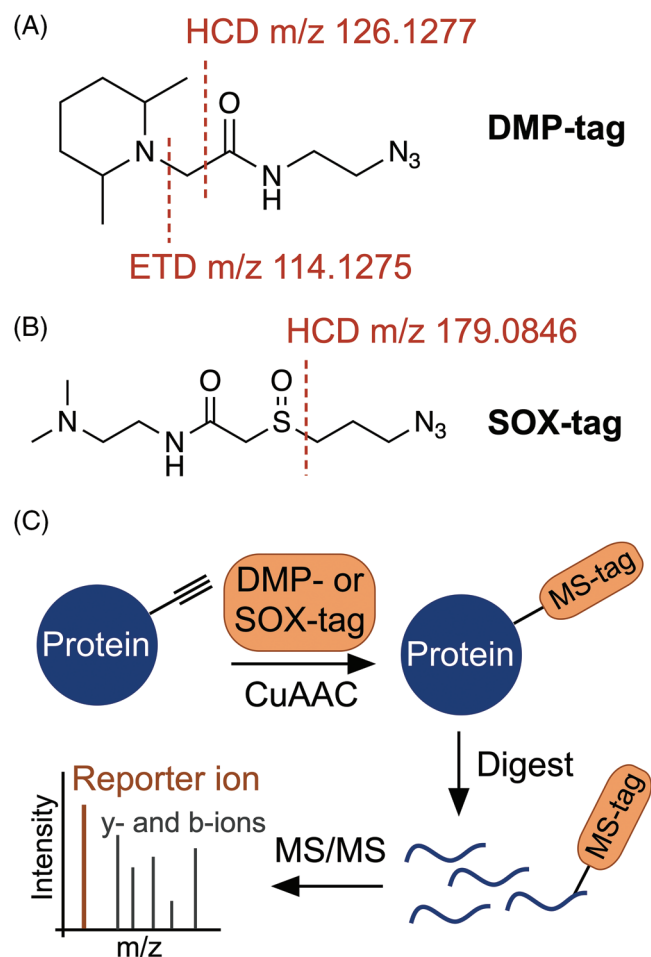
© 2022 The Authors. *Journal of Mass Spectrometry* published by John Wiley & Sons Ltd.



different ionization properties. Furthermore, different total protein amounts, as well as shifted retention times during LC separation, contribute to the abovementioned problems.<sup>19,20</sup> Here, we report the synthesis of two MS-tags conjugated to alkyne-modified proteins via click chemistry, which produce a reporter ion upon fragmentation in proof-of-concept experiments. The reporter ions are then used to improve the identification of the modified peptides by the search algorithms.

## 2 | RESULTS

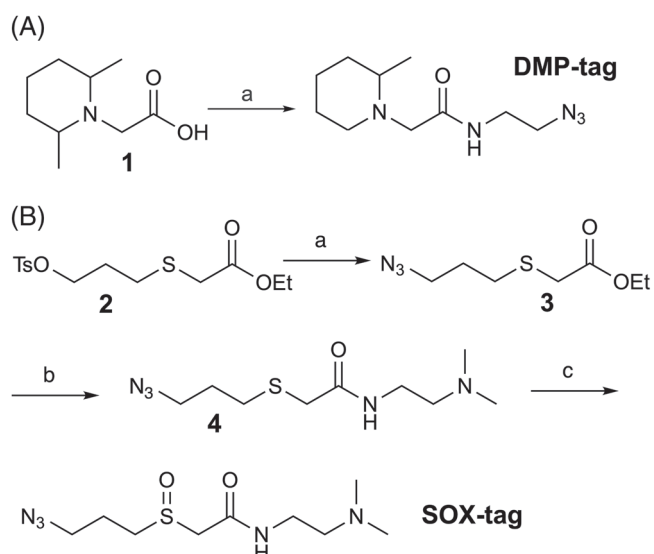
Based on the structure of commercially available TMT-tag and the recently reported sulfoxide-containing MS-tag, we have designed and synthesized two novel MS-tags.<sup>21,22</sup> The presented 2,6-dimethylpiperidine-based (DMP) and sulfoxide-containing (SOX) tags contain azido group for bioorthogonal Cu(I)-catalyzed azide-alkyne cycloaddition (CuAAC) with alkyne-modified peptides or proteins. Thus, the two new MS-tags enable selective labeling of modified peptides for MS analysis, in contrast to the original TMT-tag reagent bearing an *N*-hydroxysuccinimide ester (NHS) group to react with all available primary amines within the protein sample (Figure 1).<sup>18</sup>



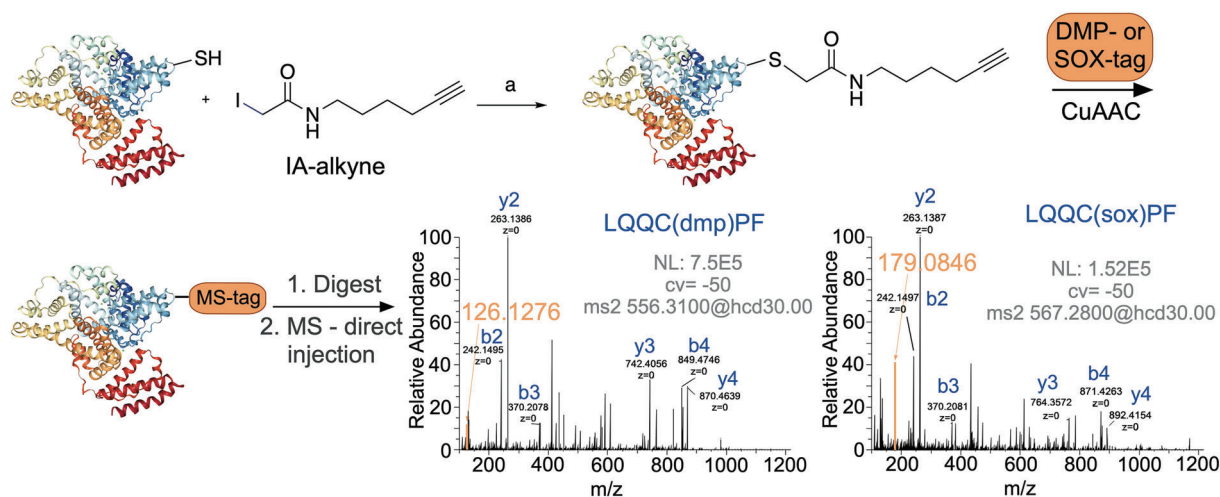
**FIGURE 1** (A,B) Structure of the DMP- and SOX-tag with characteristic reporter ion masses. (C) Overall strategy to identify modified peptides in complex samples using the DMP- and SOX-tag

The study has been initiated by the synthesis of two new MS-tags (Figure 2). First, the DMP-tag containing 2,6-dimethylpiperidine was prepared from the carboxylic acid derivative **1** by HATU catalyzed amide coupling with 2-azidoethylamine, yielding after 3 days the desired DMP-tag at a 63% yield. Next, the synthesis of the sulfoxide-based tag was carried out in a total of five steps. In brief, the synthesis starts with nucleophilic substitution of ethyl bromoacetate with 3-mercaptopropanol, followed by activation of the hydroxyl group by tosylation (**2**) and subsequent conversion to azide **3** to equip the linker of the MS-tag with moiety suitable for click chemistry. Even though hydrolysis of the ester was a side reaction during nucleophilic substitution, obtained acid **3** was used in 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) and hydroxybenzotriazole (HOBt) catalyzed amide coupling giving compound **4**, which was oxidized by mCPBA to the final SOX-tag. Although all the steps provided moderate to good yields, the final oxidation showed a somewhat lower yield of 17% caused by side reactions and instability during purification. Both reagents DMP- and SOX-tags show good stability when stored as ready-to-use solutions in DMSO at  $-20^{\circ}\text{C}$ .

To explore in detail the fragmentation properties of the DMP- and SOX-tags, we have established a model protein-PTM system in which the free thiol of the cysteine residue C58 of bovine serum albumin (BSA) was modified with the cysteine reactive probe IA-alkyne, which contains a terminal alkyne (Figure 3). The alkyne decorated BSA resembles the protein PTM isolated from cells treated with an alkyne-containing probe, which is common in chemical proteomic workflows that aim to map protein PTMs.<sup>15,18,23–25</sup> The alkyne-modified BSA was further decorated with either DMP- or SOX-tag using the CuAAC. Next, the BSA was proteolytically cleaved by chymotrypsin, and the resulting peptide mixtures were desalted on



**FIGURE 2** (A) Synthetic approach to DMP-tag (a) 2-azidoethan-1-amine, EDC, HOBt, DIPEA, DMF, r.t., 72 h, 63%. (B) Synthesis of the SOX-tag. (a)  $\text{NaN}_3$ , EtOH,  $95^{\circ}\text{C}$ , 18 h, 60%. (b)  $N^1,N^1$ -dimethylethane-1,2-diamine, EDC, HOBt, DIPEA, DCM, r.t., 18 h, 55%. (c) mCPBA,  $\text{H}_2\text{O}$ , r.t., 1.5 h, 17%

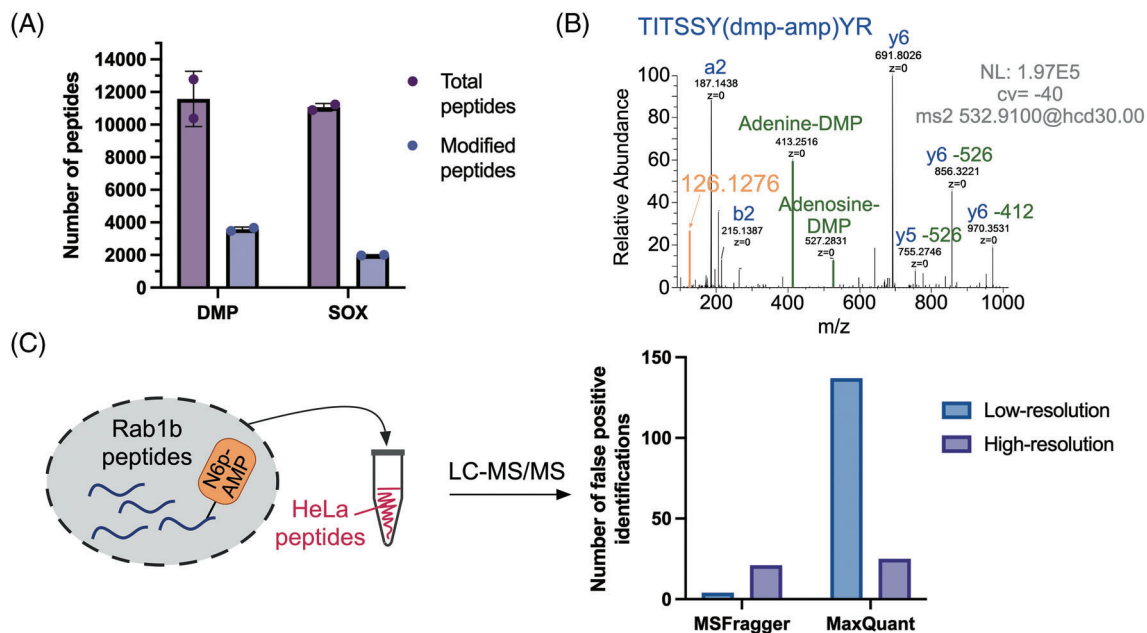


**FIGURE 3** Proof-of-concept experiments with the BSA model system and the DMP- and SOX-tags

C18 columns and analyzed by direct injection into the Orbitrap Eclipse Tribrid mass spectrometer with high-field asymmetric waveform ion mobility spectrometry (FAIMS) introduced between the ion source and the Eclipse mass spectrometer.<sup>26</sup> We have started with an analysis of the DMP-tag modified BSA and the resulting DMP-tagged peptide LQQC (dmp)PF and its missed cleavage peptide LQQC (dmp) DEHVKLVNELTEF where dmp refers to the attached DMP-tag. Step-wise optimization of the FAIMs compensation voltage (CV) provided suitable conditions at  $-50$  CV to acquire the MS1 spectra and select the target ions for the MS<sup>2</sup> experiment. To identify the suitable condition for the generation of the MS<sup>2</sup> spectra, with particular focus on the intensity of the reporter ion at 126.1277 m/z, resulting from the fragmentation of the DMP-tag, the higher-energy C-trap dissociation (HCD) energy has been gradually increased. Optimization has revealed that the most effective cleavage occurs at 30 V. A complementary set of experiments has been performed with the electron-transfer dissociation (ETD) fragmentation technique, showing a somewhat lower intensity of the corresponding reporter ion at 114.1275 m/z (Figure S1). In parallel, the fragmentation properties of the SOX-tag were assessed to show the anticipated reporter ion at 179.0846 m/z using HCD fragmentation. However, it surprisingly produced a complementary reporter ion at 131.1178 m/z as a major fragment upon ETD fragmentation (Figure S2). For both the DMP-tag and SOX-tag, better fragmentation was observed for species with higher charge peptide precursor ions. Measurement of the negative control, the BSA peptides, which were not modified with the DMP-tag or SOX-tag, but only with IA-alkyne, confirmed the specificity of all reporter ions (Figure S3).

Having characterized the fragmentation properties of the DMP- and SOX-tag in our model system, we have continued to test the possibility of using these MS-tags for the identification of modified peptides on the whole proteome level. For this, HeLa cells lysate was treated with IA-alkyne and further reacted with the DMP- or SOX-tag using click chemistry. Subsequently, the labeled proteome was acetone precipitated, trypsin digested and measured by LC-MS/MS using

the 2 h gradient with alternating FAIMS CV voltages between  $-50$  and  $-70$  V. The MS<sup>2</sup> has been acquired in the orbitrap, m/z range was adjusted to span from 110 to 1100 m/z and HCD fragmentation set to 30 V.<sup>27</sup> Next, MaxQuant searched the resulting spectra for the peptides modified with cysteine reactive probe and labeled with the DMP- or SOX-tag. The reporter ions were set up as diagnostic peaks. From the single run, MaxQuant identified an overall 10 802 peptides in the DMP-tag sample (Table S1). Among the total number of peptides, 3578 were modified with more than 99% of all MS<sup>2</sup> spectra containing the corresponding diagnostic peak of the DMP-tag (Figure 4A). In comparison, MSFragger search has found on average 18 673 peptides and 4783 modified peptides. SOX-label showed somewhat lower numbers with MaxQuant finding in total 14 601 peptides and 1992 modified peptides again with more than 99% containing the SOX-tag reporter ions, and again higher numbers resulted from the MSFragger search—17 380 peptides and 2916 modified peptides showing the efficiency of the offset search (Table S2). The average site identification probability for both tags with high-resolution MS<sup>2</sup> is over 99%. For comparison, the samples have been measured using the low-resolution ion trap MS<sup>2</sup> acquisition as well (Figure S4). This led to a significant increase in the total number of identified peptides but lower number of DMP-modified peptides (2325) and SOX-modified peptides (1550; both calculated with MaxQuant). This observation is in line with previously reported improvement of modified peptide identification rates by the high-resolution MS<sup>2</sup> spectra and demonstrates the feasibility of our approach compatible with a wide range of MS measurement setups.<sup>28</sup> Moreover, the identified DMP-modified peptides using the low-resolution MS<sup>2</sup> might contain interfering reporter ions at 126.0913 m/z from acetylated lysine.<sup>29</sup> Together, the application of the DMP- and SOX-tag with IA-alkyne in the proof-of-principle experiments show the high efficiency of the reporter ion release, which opens a way for relative quantification of modified peptides between various experimental conditions when used with isotopically labeled tags. However, the absolute quantification of modified peptides would need to be determined for each



**FIGURE 4** Analysis of cysteine DMP- and SOX-labeling on whole proteome level and Rab1b AMPylation with DMP-tag. (A) Total peptides and modified peptides found by MaxQuant in the DMP- and SOX-tag labeled cysteines using high-resolution MS<sup>2</sup> acquired in orbitrap. (B) Fragmentation properties of DMP-labeled AMPylated peptide from Rab1b. (C) Comparison of the MaxQuant and MSFragger search fidelity using the spiked in unnatural N<sup>6</sup>-propargyl AMPylated peptide from Rab1b

application because of its dependence on the used chemical proteomic probe, its metabolic incorporation rate or reactivity and CuAAC efficiency.

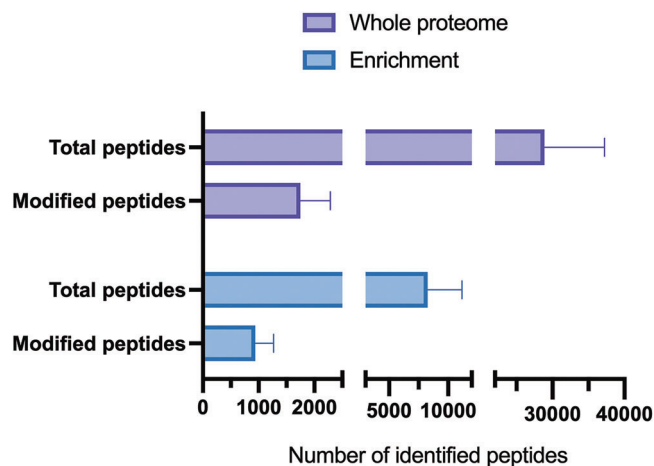
Encouraged by the results, we focused on our better performing DMP-tag to study protein AMPylation.<sup>30</sup> First, in an in vitro reaction of the well-described pair of AMP-transferase DrrA and its substrate Rab1b with ATP or N<sup>6</sup>-propargyl ATP the AMPylated Rab1b was prepared and characterized by intact protein MS (Figure S5 and S6).<sup>31,32</sup> Both, wt and N<sup>6</sup>-propargyl modified proteins were then coupled with DMP-tag, reduced, alkylated and trypsinized. The resulting peptide mixture was analyzed by direct injection into the mass spectrometer. Of note, Rab1b is modified on Y77 with the tryptic peptide TITSSYYR, which makes the site identification in particular challenging because of six possible modification sites. The FAIMS compensation voltage optimization allowed us to select and enhance the intensity of the desired peptides for MS<sup>2</sup> experiments. The unmodified Rab1b peptide served as a control. Interestingly, the unmodified peptides were found only as double-charged peptides, whereas the AMPylated peptide was predominantly triple charged. The following fragmentation experiments corroborate previous reports and add additional insight on fragmentation properties.<sup>3,4,33</sup> The fragmentation of the control AMPylated TITSSYYR peptide by HCD provided all characteristic ions and neutral losses. These were also paralleled in the analysis of the N<sup>6</sup>-propargyl AMP modified peptide. The measurement was repeated with the attached DMP-tag to explore the possibility to modulate fragmentation properties and improve the site identification rate of this unstable PTM. However, the DMP-AMP-peptide exhibits the same fragmentation properties, but as expected, it has yielded an

additional reporter ion at 126.1276 m/z with HCD and a low intensity 114.1275 m/z reporter ion when ETD was used (Figures 4B and S7). The artificial Rab1b DMP-AMP-peptides were spiked in the HeLa whole proteome tryptic digest and analyzed via LC-MS/MS. Indeed, it was possible to identify the desired DMP-AMP-peptide from the Rab1b by MaxQuant and MSFragger. Of note, the score was improved when the neutral losses were defined, but it led to the incorrect localization of the modification on the peptide (Table S3). This could be due to the fact that the modified Rab1b peptide contains six potentially modified sites out of eight amino acids in total. The MS<sup>2</sup> acquired in the ion trap resulted in false-positive identifications in both MaxQuant and MSFragger. The MaxQuant search was set up to search for unnatural modified peptides (with N<sup>6</sup>-propargyl AMP) but identified 137 modified peptides instead of one from Rab1b (Figure 4C). In comparison, the MSFragger search showed greater stringency by finding only four modified peptides, which were inherently incorrect but did not find the Rab1b peptide (Figure 4C). The high-resolution MS<sup>2</sup> acquired in the orbitrap has led to improvement of the MaxQuant search and in MSFragger to correct assignment of the modified Rab1b peptide (Figure 4C).

Our and others' previous attempts to search for AMPylated peptides in whole proteome tryptic digest proved to be challenging. In particular, Pieles et al have synthesized two adenosine analogs containing <sup>15</sup>N and <sup>13</sup>C stable isotopes, which were used for metabolic labeling of AMPylated proteins. Although it was possible to identify reporter ion clusters of labeled adenosines in in vitro activity assays, the search of the labeled peptides on a whole proteome has shown a rather low efficiency.<sup>4</sup> Therefore, in our study, we have decided to

use a commercial DMP-specific antibody to enrich the DMP-modified peptides.<sup>34</sup> In principle, the main advantage compared to other approaches is the possibility of enriching and selectively eluting only the modified peptides without the necessity to use additional chemical or enzymatic cleavage of the linker used for the enrichment. The possibility of enriching DMP-modified peptides was first tested on a model system with an IA-alkyne probe coupled with the DMP-tag. The anti-DMP antibody was applied on the peptide level and resulted in a two-fold increase of the DMP-modified peptides in the sample comparison to DMP-modified peptides without enrichment (Figure 5 and Table S4). The same approach was then applied to pro-N6pA treated cell lysates, which resulted in the labeling of DMP-AMP-modified peptides. Even though we have identified numerous AMPylated peptides, there was no overlap with previously found AMPylated proteins using complementary methods.<sup>15,33</sup> We hypothesize that this is mainly caused by challenging bioinformatic analysis, which has to deal with a complex mixture of ions after fragmentations. Although harnessing the potential of reporter ions and neutral losses presence might be a great advantage in future.

In summary, we have designed and synthesized two clickable MS-tags based on DMP- and SOX-moieties, which were evaluated using the single digested modified BSA protein and on whole proteome level with cysteine reactive IA-alkyne. Further on, we have applied the DMP-tag in the analysis of protein AMPylation and attempted enrichment of the AMPylated peptides using the DMP-specific antibody. This study extends the repertoire of available MS-linkers, opens the possibility to further develop isotopically labeled derivatives of DMP- and SOX-tags for quantification of PTM-peptides obtained from different cell types or stress conditions. Moreover, we have generated a high-quality MS spectra resource for optimization of the PTM search algorithms, which is freely available to the community.



**FIGURE 5** DMP-IA-alkyne-modified peptides identification rate in the whole proteome and after the enrichment using the DMP-specific antibody

## ACKNOWLEDGMENTS

This research was supported by the Liebig fellowship from Fonds der Chemischen Industrie to P.K. and T.B., LMUexcellent Junior Fund to P.K. and Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – SFB 1309.

## AUTHOR CONTRIBUTION

P.K. conceived the study, assisted with MS measurements, and wrote the manuscript. D.M. carried out the synthesis of the tags and participated in MS measurements. A.T. performed MS samples preparation and MS measurements. T.B. carried out the affinity enrichment of AMP-DMP modified peptides. M.-K.W. overexpressed the DrrA and Rab1b and performed in vitro AMPylation. All authors have revised the manuscript.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## ORCID

Dmytro Makarov <https://orcid.org/0000-0001-9971-8570>

Marie-Kristin von Wrisberg <https://orcid.org/0000-0002-7100-4565>

Sabine Schneider <https://orcid.org/0000-0003-1054-8689>

Pavel Kielkowski <https://orcid.org/0000-0003-4910-6263>

## REFERENCES

- Lin H, Caroll KS. Introduction: posttranslational protein modification. *Chem Rev.* 2018;118:887-888. doi:10.1021/acs.chemrev.7b00756
- Aebersold R, Agar JN, Amster IJ, et al. How many human proteoforms are there? *Nat Chem Biol.* 2018;14:206-214. doi:10.1038/nchembio.2576
- Li Y, Al-Eryani R, Yarbrough ML, Orth K, Ball HL. Characterization of AMPylation on threonine, serine, and tyrosine using mass spectrometry. *J Am Soc Mass Spectr.* 2011;22:752-761. doi:10.1007/s13361-011-0084-1
- Pieles K, Glatter T, Harms A, Schmidt A, Dehio C. An experimental strategy for the identification of AMPylation targets from complex protein samples. *Proteomics.* 2014;14:1048-1052. doi:10.1002/pmic.201300470
- Sanyal A, Dutta S, Camara A, et al. Alpha-synuclein is a target of Fic-mediated adenylation/AMPylation: possible implications for Parkinson's disease. *J Mol Biol.* 2019;431:2266-2282. doi:10.1016/j.jmb.2019.04.026
- Sanyal A, Chen AJ, Nakayasu ES, et al. A novel link between Fic (filamentation induced by cAMP)-mediated adenylation/AMPylation and the unfolded protein response. *J Biol Chem.* 2015;290:8482-8499. doi:10.1074/jbc.M114.618348
- Yu X, Woolery AR, Luong P, et al. Copper-catalyzed azide-alkyne cycloaddition (click chemistry)-based detection of global pathogen-host AMPylation on self-assembled human protein microarrays. *Mol Cell Proteomics.* 2014;13:3164-3176. doi:10.1074/mcp.M114.041103
- Preissler S, Rato C, Chen R, et al. AMPylation matches BiP activity to client protein load in the endoplasmic reticulum. *eLife.* 2015;4:e12621. doi:10.7554/elife.12621

9. Zanon PRA, Lewald L, Hacker SM. Isotopically labeled desthiobiotin azide (isoDTB) tags enable global profiling of the bacterial cysteinome. *Angew Chem Int Ed*. 2020;132:2851-2858. doi:10.1002/anie.201912075
10. Speers AE, Cravatt BF. Activity based protein profiling (ABPP) and click chemistry (CC)-ABPP by MudPIT mass spectrometry. *Current Protocols in Chemical Biology*. 2009;1:29-41. doi:10.1002/9780470559277.ch090138
11. Wright MH, Clough B, Rackham MD, et al. Validation of N-myristoyltransferase as an antimalarial drug target using an integrated chemical biology approach. *Nat Chem*. 2014;6:112-121. doi:10.1038/nchem.1830
12. Rauniyar N, Yates JR III. Isobaric labeling-based relative quantification in shotgun proteomics. *J Proteome Res*. 2014;13:5293-5309. doi:10.1021/pr500880b
13. Martin BR, Wang C, Adibekian A, Tully SE, Cravatt BF. Global profiling of dynamic protein palmitoylation. *Nat Methods*. 2011;9:84-89. doi:10.1038/nmeth.1769
14. Storck EM, Morales-Sanfrutos J, Serwa RA, et al. Dual chemical probes enable quantitative system-wide analysis of protein prenylation and prenylation dynamics. *Nat Chem*. 2019;11:552-561. doi:10.1038/s41557-019-0237-6
15. Kielkowski P, Buchsbaum IY, Kirsch VC, et al. FICD activity and AMPylation remodeling modulate human neurogenesis. *Nat Commun*. 2020;11:517. doi:10.1038/s41467-019-14235-6
16. Kielkowski P, Buchsbaum IY, Becker T, Bach K, Cappello S, Sieber SA. A pronucleotide probe for live-cell imaging of protein AMPylation. *Chembiochem*. 2020;21:1285-1287. doi:10.1002/cbic.201900716
17. Chuh KN, Pratt MR. Chemical methods for the proteome-wide identification of posttranslationally modified proteins. *Curr Opin Chem Biol*. 2015;24:27-37. doi:10.1016/j.cbpa.2014.10.020
18. Parker CG, Pratt MR. Click chemistry in proteomic investigations. *Cell*. 2020;180:605-632. doi:10.1016/j.cell.2020.01.025
19. Olsen JV, Mann M. Status of large-scale analysis of post-translational modifications by mass spectrometry. *Mol Cell Proteomics*. 2013;12:3444-3452. doi:10.1074/mcp.o113.034181
20. Sharma K, D'Souza RCJ, Tyanova S, et al. Ultradeep human phosphoproteome reveals a distinct regulatory nature of Tyr and Ser/Thr-based signaling. *Cell Rep*. 2014;8:1583-1594. doi:10.1016/j.celrep.2014.07.036
21. Stadlmeier M, Bogena J, Wallner M, Wühr M, Carell T. A sulfoxide-based isobaric labelling reagent for accurate quantitative mass spectrometry. *Angew Chem Int Ed*. 2018;57:2958-2962. doi:10.1002/anie.201708867
22. Li J, Vranken JGV, Vaites LP, et al. TMTpro reagents: a set of isobaric labeling mass tags enables simultaneous proteome-wide measurements across 16 samples. *Nat Methods*. 2020;17(4):399-404. doi:10.1038/s41592-020-0781-4
23. Grammel M, Hang HC. Chemical reporters for biological discovery. *Nat Chem Biol*. 2013;9:475-484. doi:10.1038/nchembio.1296
24. Backus KM, Correia BE, Lum KM, et al. Proteome-wide covalent ligand discovery in native biological systems. *Nature*. 2016;534(7608):570-574. doi:10.1038/nature18002
25. Heal WP, Tate EW. Getting a chemical handle on protein post-translational modification. *Org Biomol Chem*. 2010;8:731-738. doi:10.1039/b917894e
26. Bekker-Jensen DB, Martínez-Val A, Steigerwald S, et al. A compact quadrupole-orbitrap mass Spectrometer with FAIMS interface improves proteome coverage in short LC gradients\*. *Mol Cell Proteomics*. 2020;19:716-729. doi:10.1074/mcp.tir119.001906
27. Wiśniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. *Nat Methods*. 2009;6:359-362. doi:10.1038/nmeth.1322
28. Bakalarski CE, Haas W, Dephoure NE, Gygi SP. The effects of mass accuracy, data acquisition speed, and search algorithm choice on peptide identification rates in phosphoproteomics. *Anal Bioanal Chem*. 2007;389:1409-1419. doi:10.1007/s00216-007-1563-x
29. Nakayasu ES, Wu S, Sydor MA, et al. A method to determine lysine acetylation stoichiometries. *Int J Proteom*. 2014;2014:730725. doi:10.1155/2014/730725
30. Sieber SA, Cappello S, Kielkowski P. From young to old: AMPylation hits the brain. *Cell Chem Biol*. 2020;27(7):773-779. doi:10.1016/j.chembiol.2020.05.009
31. Du J, von Wrisberg M-K, Gulen B, et al. Rab1-AMPylation by legionella DrrA is allosterically activated by Rab1. *Nat Commun*. 2021;12:460. doi:10.1038/s41467-020-20702-2
32. Müller MP, Peters H, Blümer J, Blankenfeldt W, Goody RS, Itzen A. The Legionella effector protein DrrA AMPylates the membrane traffic regulator Rab1b. *Science*. 2010;329:946-949. doi:10.1126/science.1192276
33. Broncel M, Serwa RA, Bunney TD, Katan M, Tate EW. Global profiling of huntingtin-associated protein E (HYPE)-mediated AMPylation through a chemical proteomic approach. *Mol Cell Proteomics*. 2016;15:715-725. doi:10.1074/mcp.O115.054429
34. Qu Z, Meng F, Bomgarden RD, et al. Proteomic quantification and site-mapping of S-nitrosylated proteins using isobaric iodoTMT reagents. *J Proteome Res*. 2014;13:3200-3211. doi:10.1021/pr401179v

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Makarov D, Telek A, Becker T, von Wrisberg M-K, Schneider S, Kielkowski P. Clickable report tags for identification of modified peptides by mass spectrometry. *J Mass Spectrom*. 2022;57(3):e4812. doi:10.1002/jms.4812

## 4 Unpublished work

### 4.1 Further optimization of the SP2E workflow

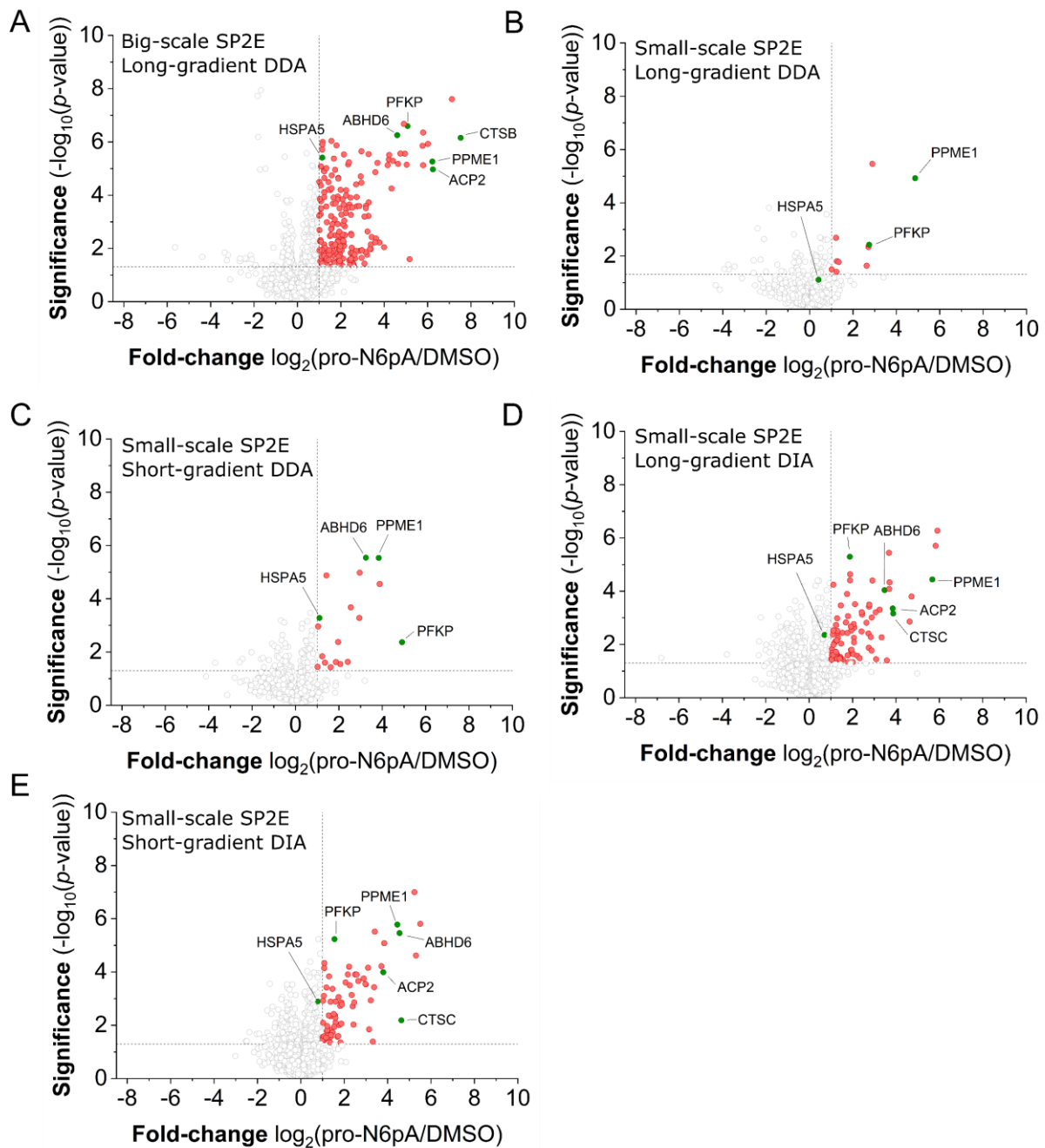
#### 4.1.1 Comparison of DDA and DIA acquisition

The above described SP2E workflow is a fast and robust chemical proteomic platform ready for automation.<sup>101</sup> The protocol uses paramagnetic beads that enable to separate the liquid and solid phases without the need of a centrifugation step. During the optimization process of this protocol, the starting protein amount was reduced from 400 µg to 100 µg to allow the analysis of different-to-culture cell lines. In addition to the lower protein amount, working volumes were decreased to transfer the protocol from 1.5 mL Eppendorf tubes to the 96-well plate format. The 96-well plate format has the advantage of increasing the throughput of the workflow, as multichannel pipettes can be used for the washing steps or even the entire protocol can be automated by a robot. However, apart from the benefits of performing the protocol with less protein in a shorter time, the number of significantly enriched proteins decreased compared to the big-scale SP2E protocol when both were measured with DDA in a 150 minutes gradient (Figure 12A,B). Therefore, the gradient was shortened to 60 minutes to concentrate the small amounts of enriched peptides in a shorter time window and thus obtain MS spectra with higher intensity. Although the number of significantly enriched proteins increased compared to the long gradient, the overall number is still lower than in the big-scale protocol (Figure 12A,C).<sup>101</sup> Thus, we had the idea to change the data acquisition mode in the MS-measurement from DDA to DIA, since DIA fragments all precursor ions regardless of their MS1 intensity. This leads to a higher reproducibility between the replicates and hence fewer missing values. In this way, the lower number of significantly enriched proteins in the small scale SP2E could be overcome.

To investigate whether DIA could improve the results of the small-scale SP2E workflow, we remeasured pro-N6pA treated HeLa cells with DIA and compared the results with the previously obtained DDA dataset. For the evaluation, we focused on the total number of significantly enriched proteins in the volcano plot and furthermore on the enrichment of the six AMPylation marker proteins HSPA5, CTSC, PFKP, PPME1, ACP2 and ABHD6.

In comparison to the small-scale DDA measurements, the volcano plots of the small-scale DIA measurements show more significantly enriched proteins (Figure 12D,E). More specifically, 4 times more proteins are significantly enriched in the short gradient DIA measurement than in the short gradient DDA measurement (Figure 12C,E). The ability to analyze chemical proteomics with DIA is further underlined by the fact that 5 out of 6 AMPylation marker

proteins are significantly enriched. The only marker protein that was not significantly enriched is HSPA5, but this was only slightly below the fold-enrichment cut-off. In contrast to the DDA measurement, the fold-enrichment of the marker proteins and the number of significantly enriched proteins are similar in the short and long gradient of the DIA measurements. Therefore, the short gradient DIA measurement is preferable compared to the long gradient since the measurement time is shorter. Taken together, the short-gradient DIA measurement is suitable to measure small-scale SP2E samples and outperforms the DDA measurement.



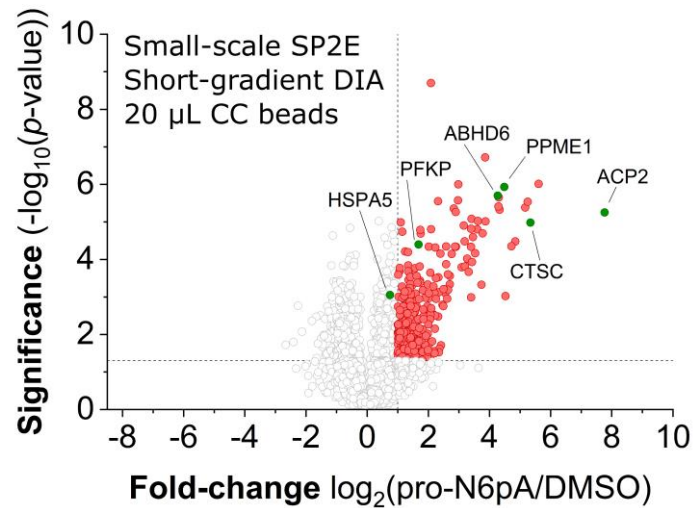
**Figure 12 Comparison of DDA and DIA measurement.** For the comparison, pro-N6pA treated HeLa cells were prepared using either the big-scale or small-scale SP2E protocol. Volcano plot visualization showing fold-enrichment versus significance using 100  $\mu\text{M}$  pro-N6pA probe compared to DMSO;  $n = 4$ , cut-off lines  $p\text{-value} > 0.05$  and 2-fold enrichment. (A) Big-scale SP2E measured with long-gradient DDA. (B) Small-scale SP2E measured with long-gradient DDA. (C) Small-scale SP2E measured with short-gradient DDA. (D) Small-scale SP2E measured with long-gradient DIA. (E) Small-scale SP2E measured with short-gradient DIA. Red dots are significantly enriched proteins. Green dots represent the AMPylation marker proteins.



#### **4.1.2 Reduction of carboxylate-coated paramagnetic beads amount**

In addition to the improvement of the MS-measurement using DIA, we wanted to further optimize the SP2E workflow. Since both the big-scale and small-scale SP2E protocol used the same amount of carboxylate-coated paramagnetic beads so-far, we aimed to reduce the amount of the beads. The reduction of the beads would be beneficial for two reasons, first the amounts of consumables are decreased and second the handling of the workflow is simplified as less beads are in each small well of the 96-well plate. To analyze whether the beads reduction shows no deterioration compared to the standard protocol, pro-N6pA treated HeLa cells were prepared using the small-scale SP2E workflow, but with 20  $\mu$ L carboxylate-coated paramagnetic beads instead of 100  $\mu$ L. The MS-measurement was already performed with the DIA method instead of the previously used DDA method. As described in the above section, the quality of the enrichment was evaluated with the six AMPylation marker proteins HSPA5, CTSC, PFKP, PPME1, ACP2 and ABHD6. The volcano plot of the reduced beads protocol revealed that 5 out of 6 AMPylation marker proteins were significantly enriched, only HSPA5 was slightly below the fold enrichment cut-off (Figure 13). This is in-line with the volcano plot of the AMPylation samples prepared with the standard SP2E workflow (Figure 12E). In addition to the marker proteins, the volcano plot of the reduced beads protocol showed a high number of significantly enriched proteins, demonstrating its feasibility (Figure 13). Based on these results, we concluded that the amount of carboxylate-coated paramagnetic beads in the SP2E protocol can be reduced from 100  $\mu$ L to 20  $\mu$ L. With only one-fifth of the carboxylate-coated paramagnetic beads required for the SP2E workflow, huge financial savings can be achieved as this method is designed for high-throughput use. Furthermore, the lower bead amount reduces the risk to unintentionally pipette up beads from the small wells of the 96-well plate. Taken together, we demonstrate that the standard SP2E workflow can be performed with 20  $\mu$ L carboxylate-coated paramagnetic beads.

Another parameter in the small-scale SP2E protocol that could be optimized in the future is the reduction of the streptavidin-coated paramagnetic beads. So far, the same amount of streptavidin-coated paramagnetic beads was used in the small-scale as well as in the big-scale protocol. Since it was possible to perform the small-scale SP2E with one-fifth of the carboxylate-coated paramagnetic beads, it may be feasible to reduce the streptavidin-coated paramagnetic beads from 50  $\mu$ L to 10  $\mu$ L. Thereby, the costs of the protocol are lowered and the handling is simplified.



**Figure 13 Reduction of carboxylate-coated paramagnetic beads in the small-scale SP2E workflow.** HeLa cells treated with pro-N6pA probe were prepared using the small-scale SP2E workflow with 20  $\mu$ L carboxylate-coated paramagnetic (CC) beads instead of 100  $\mu$ L. Volcano plot visualization showing fold-enrichment versus significance using 100  $\mu$ M pro-N6pA probe compared to DMSO;  $n = 4$ , cut-off lines  $p$ -value  $> 0.05$  and 2-fold enrichment. Red dots are significantly enriched proteins. Green dots represent the AMPylation marker proteins.

### **4.1.3 Materials and methods**

#### **4.1.3.1 Culturing of HeLa cells and probe treatment**

HeLa cells were cultured as described in chapter 3.2. The pro-N6pA treatment was performed as described in chapter 8.2.

#### **4.1.3.2 Cell harvest and lysis**

Cell were harvested and lysed as described in chapter 3.2.

#### **4.1.3.3 SP2E workflow**

The standard small-scale and big-scale SP2E workflow was performed as described in chapter 3.2. For the reduction of the carboxylate-coated paramagnetic beads the small-scale SP2E protocol was modified and carried out with 20  $\mu$ L carboxylate-coated paramagnetic beads.

#### **4.1.3.4 MS measurement**

The DDA measurement was performed as described in chapter 8.2. The DIA measurement was carried out with the same HPLC setup as described in chapter 8.2, but with different parameters of the Orbitrap Eclipse. The MS1 spectra were acquired in the orbitrap with the following settings: Polarity: positive; Resolution: 60k; AGC target: standard; maximum injection time: 50 ms; MS scan range: 200-1800  $m/z$ . The MS2 spectra were acquired in the orbitrap in a range of 500-740  $m/z$  with 60 4 Da wide windows, which overlap by 2  $m/z$ . The other settings were: AGC target: 200%; maximum injection time: Auto; HCD collision energy: 35%; resolution: 30k. FAIMS was used with a single CV of -45 V and the inner electrode temperature was set to 100 °C whereas the outer electrode was set to 90 °C. The FAIMS flow rate was adjusted to 3.5 L/min.

#### **4.1.3.5 Data analysis**

DDA samples were analyzed as described in chapter 8.2 using MaxQuant 2.0.1.0 and Perseus 1.6.10.43. Before calculating the DIA samples, they were converted into mzML files with "peakPicking" and "demultiplex" enabled using MSConvertGUI. The calculation was done with DIA-NN 1.8.1, searching against the uniprot database for Homo sapiens (taxon identifier: 9606, March 2020). DIA-NN was used with default settings except some changes. The FASTA digest for library-free search/library generation and Deep-learning based spectra, RTs and IMs prediction were both enabled. In addition, the precursor  $m/z$  range was set to 500-740  $m/z$ .

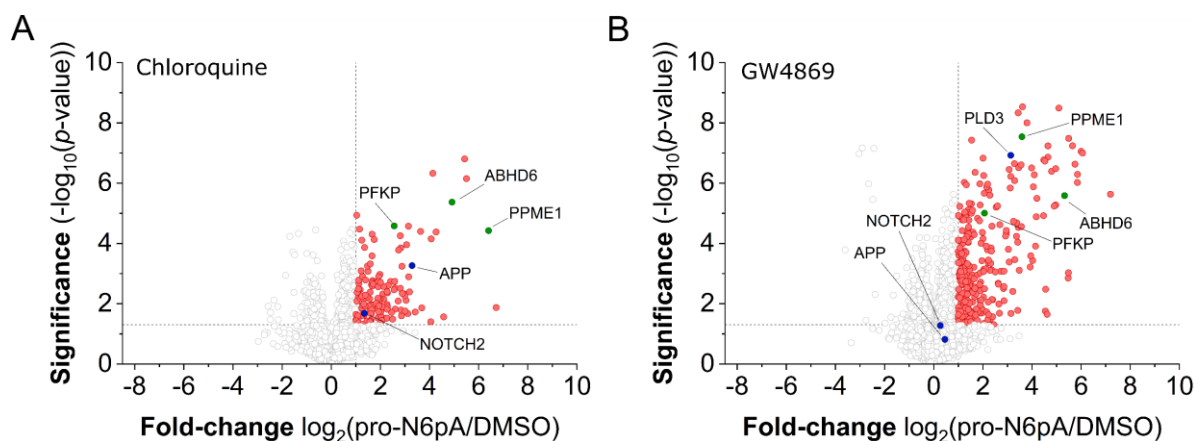
## 4.2 AMPylation profiling in various cell lines using different inhibitors

### 4.2.1 AMPylation profiling in SH-SY5Y neuroblastoma cells using chloroquine and GW4869

In order to gain insights into the metabolic pathways in which AMPylation is involved in addition to its known role in the UPR, AMPylation was analyzed in SH-SY5Y neuroblastoma cells using five different inhibitors: bafilomycin, monensin, rapamycin, thenoyltrifluoroacetone (TTFA) and 2-deoxy-D-glucose.<sup>31,32</sup> In this study, it was found that a certain group of proteins, such as APP, which is strongly connected to Alzheimer's disease or Neurogenic locus notch homolog protein 2 (NOTCH2), which is important for the cell fate determination, could only be identified after treatment with bafilomycin and monensin.<sup>101-104</sup> In addition to that, the AMPylation of PLD3 was shown to be upregulated upon bafilomycin and monensin treatment.<sup>101</sup> Since bafilomycin inhibits autophagy and monensin the endolysosomal trafficking, it was hypothesized that there is a link between AMPylation and these pathways.<sup>105,106</sup> To further support this assumption, we decided to expand this study by chloroquine, an inhibitor of the autophagic flux, similar to bafilomycin.<sup>107</sup> Therefore, SH-SY5Y neuroblastoma cells were either treated with chloroquine together with the pro-N6pA probe or just with chloroquine. The inhibitor treatment of the control as well as of the probe-treated samples ensures that the protein expression is equal in both. After the cells were incubated for 16 hours with the inhibitor and probe, the cells were harvested and the MS-samples were prepared with the big-scale SP2E protocol. The subsequent MS-measurement was performed with the long-gradient DDA method. The volcano plot showed that the pro-N6pA treatment and the sample preparation worked as the 3 AMPylation marker proteins PPME1, PFKP and ABHD6 were significantly enriched (Figure 14A). In addition, PLD3 which showed higher AMPylation levels upon bafilomycin and monensin treatment, was not found in the volcano plot of the chloroquine treated samples. However, PLD3 was identified in 2 out of 4 probe-treated replicates and not in the controls. This can be due to the reason that the MS-measurement was performed using the long-gradient DDA method, which is not as sensitive as the DIA method. Furthermore, just as observed with bafilomycin and monensin treatment, APP and NOTCH2 were both significantly enriched when treated with chloroquine. This underlines the hypothesis that AMPylation could play a role in endolysosomal trafficking and autophagy. To narrow down the pathway that causes the increase in AMPylation for a specific group of proteins, we decided to use an inhibitor that blocks the exosome production to secrete proteins. More specifically, we used the sphingomyelinases inhibitor GW4869 which blocks the endosomal sorting complexes for transport machinery (ESCRT)-independent pathway.<sup>108</sup> The

GW4869 treated SH-SY5Y neuroblastoma cells were prepared with the big-scale SP2E protocol and measured with the short-gradient DIA method. The volcano plot revealed that the pro-N6pA treatment and the workflow worked, as the 3 AMPylation marker proteins PPME1, PFKP and ABHD6 were significantly enriched (Figure 14B). In contrast to the chloroquine treatment, PLD3 was significantly enriched and identified in 4 out of 4 probe-treated samples. However, it is important to note that the measurement was performed with the more sensitive DIA method. Furthermore, NOTCH2 and APP were both not significantly enriched when treated with GW4869. This shows that the AMPylation of NOTCH2 and APP is independent of the exosome-mediated protein secretion. Therefore, the observed AMPylation of NOTCH2 and APP after monensin, bafilomycin or chloroquine treatment is more related to autophagy and endolysosomal trafficking rather than to protein secretion. Nevertheless, the exact molecular rationale for the change in AMPylation upon monensin, bafilomycin or chloroquine treatment remains unknown. Taken together, we analyzed AMPylation with 2 novel inhibitors, chloroquine and GW4869, and thereby provide evidence that AMPylation may play a role in endolysosomal trafficking and autophagy.

To better understand the function of the NOTCH2, APP and PLD3 AMPylation in endolysosomal trafficking and autophagy, it will be essential in the future to identify the site of modification. By means of this, point mutants of the protein of interest can be prepared to directly compare the effect of the modified and unmodified forms upon treatment with the inhibitor. This allows to narrow down the actual molecular function of the AMPylation of NOTCH2, APP and PLD3.



**Figure 14 AMPylation profiling in SH-SY5Y cells using chloroquine and GW4869.** SH-SY5Y cells treated with 100  $\mu\text{M}$  pro-N6pA probe and either 100  $\mu\text{M}$  chloroquine or 1  $\mu\text{M}$  GW4869 were prepared using the big-scale SP2E protocol. Volcano plot visualization showing fold-enrichment versus significance using 100  $\mu\text{M}$  pro-N6pA probe compared to DMSO;  $n = 4$ , cut-off lines  $p$ -value  $> 0.05$  and 2-fold enrichment. (A) Volcano plot of chloroquine treated SH-SY5Y cells measured with long-gradient DDA method. (B) Volcano plot of GW4869 treated SH-SY5Y cells measured with short-gradient DIA method. Red dots are significantly enriched proteins. Green dots represent 3 AMPylation marker proteins. Blue dots are proteins that were found only significantly enriched after monensin and bafilomycin treatment in our previous study.<sup>101</sup>

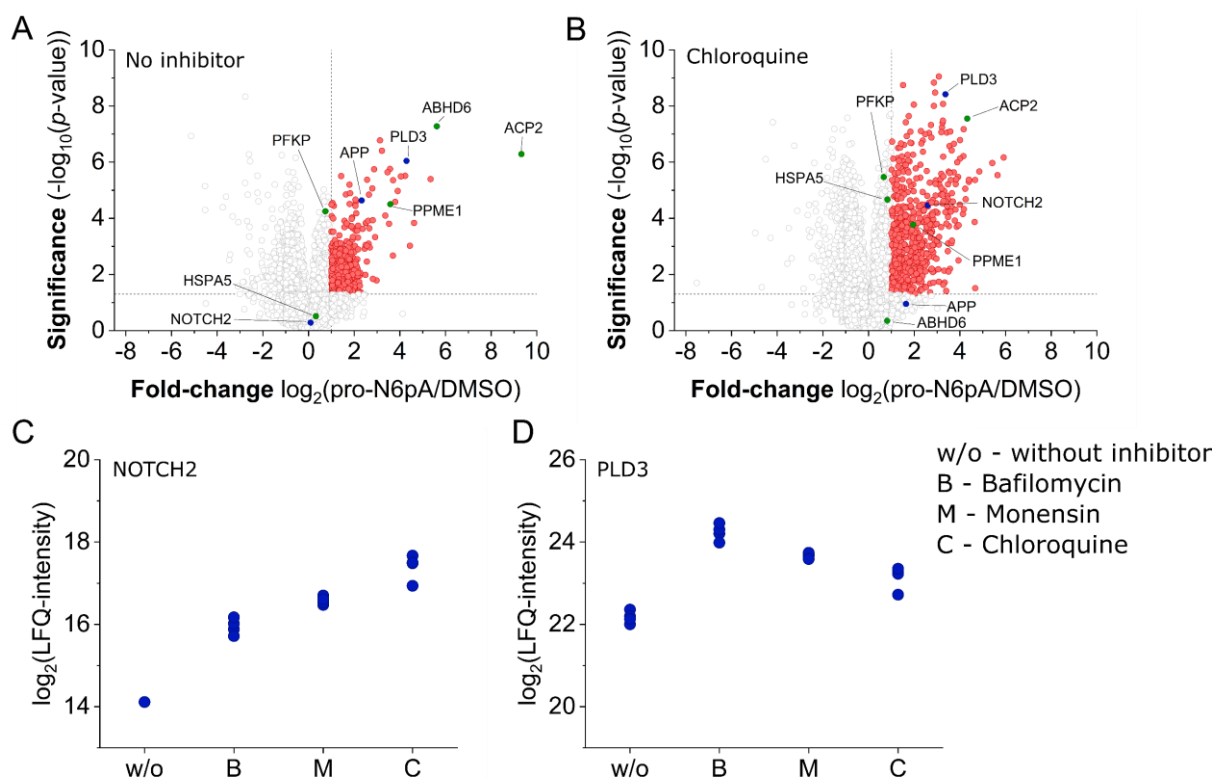
#### 4.2.2 AMPylation profiling in cerebral organoids using bafilomycin, monensin and chloroquine

In previous work, AMPylation was analyzed in SH-SY5Y neuroblastoma cells using a set of different inhibitors.<sup>101</sup> Thereby, a specific group of proteins such as NOTCH2 and APP were significantly enriched upon monensin, bafilomycin and chloroquine treatment. Based on these results it was concluded, that AMPylation could play a role in autophagy and endolysosomal trafficking. Since AMPylation has been shown to be important during neuronal development, we decided to investigate the effect of monensin, bafilomycin and chloroquine on AMPylation in a model system for neuronal development.<sup>2,3,109</sup> Therefore, cerebral organoids (COs) were used which represent a 3-dimensional *in vitro* model of human brain development derived from human pluripotent stem cells.<sup>110</sup> Kielkowski *et al.* have already used COs for their initial studies on AMPylation and have shown in fluorescent labeling experiments that the majority of AMPylated proteins are localized in the neuronal layer. Furthermore, they also identified AMPylation targets in COs using the pro-N6pA probe for chemical proteomics experiments.<sup>2</sup> However, they prepared the samples with the old method based on acetone precipitation and

streptavidin-agarose beads. Additionally, they measured the samples using the less sensitive DDA method. In contrast, we prepared the CO samples with the small-scale SP2E protocol and measured them with the short-gradient DIA method. The COs were a kind gift of *Dr. Pravata* (Max Planck Institute of Psychiatry) and were treated with the pro-N6pA probe as well as with monensin, bafilomycin and chloroquine in the same way as the SH-SY5Y cells. The volcano plot of COs treated with pro-N6pA without any inhibitor showed that the 3 AMPylation marker proteins PPME1, ABHD6 and ACP2 were significantly enriched (Figure 15A). These were not identified in the study of Kielkowski et al., demonstrating the sensitivity of the small-scale SP2E workflow in combination with the short-gradient DIA measurement. Although PFKP was significantly enriched in the previous study, it was slightly below the cut-off in this measurement.<sup>2</sup> In the next step, we compared whether APP, NOTCH2 and PLD3 changed their enrichment based on the treatment with monensin, bafilomycin and chloroquine. For the first evaluation, we compared the volcano plots of untreated and chloroquine-treated samples, whereas for the second evaluation we used profile plots that better visualize the changes of a protein under different conditions. APP was found significantly enriched in the volcano plot without inhibitor treatment, while it was not enriched upon chloroquine treatment (Figure 15A,B). Interestingly, this is the exact opposite of what was observed in SH-SY5Y cells for APP. In contrast, NOTCH2 behaved in the same way as in SH-SY5Y cells and was enriched in chloroquine-treated but not untreated COs (Figure 15A,B). The profile plot of NOTCH2 revealed that the LFQ-intensities of COs treated bafilomycin, monensin and chloroquine were higher than those of untreated COs (Figure 15C). More specifically, NOTCH2 was identified in every inhibitor-treated sample, whereas NOTCH2 was only found once in untreated COs. PLD3 was found significantly enriched in both untreated and chloroquine treated COs. To analyze whether the PLD3 LFQ-intensities in inhibitor treated COs were higher than those of untreated cells, a profile plot of PLD3 was prepared. The profile plot showed that the LFQ-intensities of PLD3 were higher in inhibitor treated COs than in untreated COs (Figure 15D). This was in line with the results that were observed for PLD3 in SH-SY5Y cells. The overall AMPylation profiling in COs using bafilomycin, monensin and chloroquine underline the suggested role of AMPylation in autophagy and endolysosomal trafficking, since PLD3 and NOTCH2 showed increased AMPylation levels upon inhibitor treatment. However, APP did not show the expected increase in AMPylation after inhibitor treatment which could be due to a different function of APP in COs or the cell type heterogeneity of COs. Taken together, we analyzed the AMPylation in COs with different inhibitors using the recently developed SP2E

protocol in combination with a DIA measurement and thereby provide new insights into potential pathways that AMPylation may be involved in a complex neuronal network.

Due to the inherent heterogeneity of COs, the experiment should be repeated in the future with a different batch of COs to show its reproducibility. Since SH-SY5Y cells were additionally treated with the inhibitor GW4869, it would be interesting to investigate the differences in AMPylation upon GW4869 treatment in COs. This will allow conclusions about whether the AMPylation in COs plays a role in the ESCRT-independent pathway of protein secretion.



**Figure 15 AMPylation profiling in cerebral organoids using bafilomycin, monensin and chloroquine.** Cerebral organoids treated with 100  $\mu\text{M}$  pro-N6pA probe and either with 100 nM bafilomycin, 2  $\mu\text{M}$  monensin or 100  $\mu\text{M}$  chloroquine were prepared using the small-scale SP2E protocol. The samples were measured with the short-gradient DIA method. Volcano plot visualization showing fold-enrichment versus significance using 100  $\mu\text{M}$  pro-N6pA probe compared to DMSO;  $n = 4$ , cut-off lines  $p\text{-value} > 0.05$  and 2-fold enrichment. (A) Volcano plot without inhibitor. (B) Volcano plot of chloroquine treated cerebral organoids. Red dots are significantly enriched proteins. Green dots represent the AMPylation marker proteins. Blue dots are proteins that were found only significantly enriched after monensin and bafilomycin treatment in our previous study.<sup>101</sup> (C) Profile plot displays the NOTCH2 LFQ-intensity under various conditions. (D) Profile plot displays the PLD3 LFQ-intensity under various conditions.



### **4.2.3 Materials and methods**

#### **4.2.3.1 Culturing of SH-SY5Y neuroblastoma cells**

SH-SY5Y neuroblastoma cells were cultured as described in chapter 3.2.

#### **4.2.3.2 Probe and inhibitor treatment of SH-SY5Y cells**

The pro-N6pA probe treatment was performed as described in chapter 3.2. For the inhibitor treatment 100  $\mu$ M chloroquine and 1  $\mu$ M GW4869 were used.

#### **4.2.3.3 Cell harvest and lysis**

Cell were harvested and lysed as described in chapter 3.2

#### **4.2.3.4 SP2E workflow**

SH-SY5Y cells were prepared using the big-scale SP2E workflow as described in chapter 3.2. Cerebral organoids were prepared using the small-scale SP2E method as described in chapter 3.2, but with 20  $\mu$ L carboxylate-coated paramagnetic beads.

#### **4.2.3.5 MS measurement**

Chloroquine treated SH-SY5Y cells were measured with the long-gradient DDA method as described in chapter 8.2. GW4869 treated SH-SY5Y cells and all cerebral organoid samples were measured with the short-gradient DIA method as described in chapter 4.1.3.4.

#### **4.2.3.6 Data analysis**

DDA samples were analyzed as described in chapter 8.2 using MaxQuant 2.0.1.0 and Perseus 1.6.10.43. DIA samples were analyzed as described in chapter 4.1.3.5 using DIA-NN 1.8.1.

## 5 Bibliography

1. Truttmann, M.C., Pincus, D., and Ploegh, H.L. (2018). Chaperone AMPylation modulates aggregation and toxicity of neurodegenerative disease-associated polypeptides. *Proc Natl Acad Sci U S A* *115*, E5008–E5017. 10.1073/pnas.1801989115.
2. Kielkowski, P., Buchsbaum, I.Y., Kirsch, V.C., Bach, N.C., Drukker, M., Cappello, S., and Sieber, S.A. (2020). FICD activity and AMPylation remodelling modulate human neurogenesis. *Nat Commun* *11*, 517. 10.1038/s41467-019-14235-6.
3. Sieber, S.A., Cappello, S., and Kielkowski, P. (2020). From Young to Old: AMPylation Hits the Brain. *Cell Chem Biol* *27*, 773–779. <https://doi.org/10.1016/j.chembiol.2020.05.009>.
4. Lam, R.S., Töpfer, F.M., Wood, P.G., Busskamp, V., and Bamberg, E. (2017). Functional maturation of human stem cell-derived neurons in long-term cultures. *PLoS One* *12*, 1–26. 10.1371/journal.pone.0169506.
5. Busskamp, V., Lewis, N.E., Guye, P., Ng, A.H., Shipman, S.L., Byrne, S.M., Sanjana, N.E., Murn, J., Li, Y., Li, S., et al. (2014). Rapid neurogenesis through transcriptional activation in human stem cells. *Mol Syst Biol* *10*, 760. 10.15252/msb.20145508.
6. Lambert, J.-C., Grenier-Boley, B., Bellenguez, C., Pasquier, F., Campion, D., Dartigues, J.-F., Berr, C., Tzourio, C., and Amouyel, P. (2015). PLD3 and sporadic Alzheimer’s disease risk. *Nature* *520*, E1–E1. 10.1038/nature14036.
7. Hooli, B. v, Lill, C.M., Mullin, K., Qiao, D., Lange, C., Bertram, L., and Tanzi, R.E. (2015). PLD3 gene variants and Alzheimer’s disease. *Nature* *520*, E7–E8. 10.1038/nature14040.
8. van der Lee, S.J., Holstege, H., Wong, T.H., Jakobsdottir, J., Bis, J.C., Chouraki, V., van Rooij, J.G.J., Grove, M.L., Smith, A. v, Amin, N., et al. (2015). PLD3 variants in population studies. *Nature* *520*, E2–E3. 10.1038/nature14038.
9. Heilmann, S., Drichel, D., Clarimon, J., Fernández, V., Lacour, A., Wagner, H., Thelen, M., Hernández, I., Fortea, J., Alegret, M., et al. (2015). PLD3 in non-familial Alzheimer’s disease. *Nature* *520*, E3–E5. 10.1038/nature14039.
10. Cruchaga, C., Karch, C.M., Jin, S.C., Benitez, B.A., Cai, Y., Guerreiro, R., Harari, O., Norton, J., Budde, J., Bertelsen, S., et al. (2014). Rare coding variants in the

- phospholipase D3 gene confer risk for Alzheimer's disease. *Nature* 505, 550–554. 10.1038/nature12825.
11. Aebersold, R., Agar, J.N., Amster, I.J., Baker, M.S., Bertozzi, C.R., Boja, E.S., Costello, C.E., Cravatt, B.F., Fenselau, C., Garcia, B.A., et al. (2018). How many human proteoforms are there? *Nat Chem Biol* 14, 206–214. 10.1038/nchembio.2576.
  12. Black, D.L. (2003). Mechanisms of Alternative Pre-Messenger RNA Splicing. *Annu Rev Biochem* 72, 291–336. 10.1146/annurev.biochem.72.121801.161720.
  13. Tress, M.L., Abascal, F., and Valencia, A. (2017). Alternative Splicing May Not Be the Key to Proteome Complexity. *Trends Biochem Sci* 42, 98–110. 10.1016/j.tibs.2016.08.008.
  14. Mann, M., and Jensen, O.N. (2003). Proteomic analysis of post-translational modifications. *Nat Biotechnol* 21, 255–261. 10.1038/nbt0303-255.
  15. Walsh, C.T., Garneau-Tsodikova, S., and Gatto, G.J. (2005). Protein posttranslational modifications: The chemistry of proteome diversifications. *Angewandte Chemie - International Edition* 44, 7342–7372. 10.1002/anie.200501023.
  16. Wang, Y.C., Peterson, S.E., and Loring, J.F. (2014). Protein post-translational modifications and regulation of pluripotency in human stem cells. *Cell Res* 24, 143–160. 10.1038/cr.2013.151.
  17. Beltrao, P., Bork, P., Krogan, N.J., and van Noort, V. (2013). Evolution and functional cross-talk of protein post-translational modifications. *Mol Syst Biol* 9, 714. <https://doi.org/10.1002/msb.201304521>.
  18. Yarbrough, M.L., and Orth, K. (2009). AMPylation is a new post-translational modification. *Nat Chem Biol* 5, 378–379. 10.1038/nchembio0609-378.
  19. Jiang, P., Mayo, A.E., and Ninfa, A.J. (2007). Escherichia coli Glutamine Synthetase Adenylyltransferase (ATase, EC 2.7.7.49): Kinetic Characterization of Regulation by PII, PII-UMP, Glutamine, and  $\alpha$ -Ketoglutarate. *Biochemistry* 46, 4133–4146. 10.1021/bi0620510.
  20. Hedberg, C., and Itzen, A. (2015). Molecular perspectives on protein adenylylation. *ACS Chem Biol* 10, 12–21. 10.1021/cb500854e.

21. Engel, P., Goepfert, A., Stanger, F. v., Harms, A., Schmidt, A., Schirmer, T., and Dehio, C. (2012). Adenylylation control by intra-or intermolecular active-site obstruction in Fic proteins. *Nature* 482, 107–110. 10.1038/nature10729.
22. Woolery, A., Luong, P., Broberg, C., and Orth, K. (2010). AMPylation: Something Old is New Again. *Front Microbiol* 1. 10.3389/fmicb.2010.00113.
23. Zekarias, B., Mattoo, S., Worby, C., Lehmann, J., Rosenbusch, R.F., and Corbeil, L.B. (2010). Histophilus somni IbpA DR2/Fic in virulence and immunoprotection at the natural host alveolar epithelial barrier. *Infect Immun* 78, 1850–1858. 10.1128/IAI.01277-09.
24. Rauh, T., Brameyer, S., Kielkowski, P., Jung, K., and Sieber, S.A. (2020). MS-Based in Situ Proteomics Reveals AMPylation of Host Proteins during Bacterial Infection. *ACS Infect Dis* 6, 3277–3289. 10.1021/acsinfecdis.0c00740.
25. Müller, M.P., Peters, H., Blümer, J., Blankenfeldt, W., Goody, R.S., and Itzen, A. (2010). The Legionella effector protein DrrA AMPylates the membrane traffic regulator Rab1b. *Science* (1979) 329, 946–949. 10.1126/science.1192276.
26. Yarbrough, M.L., Li, Y., Kinch, L.N., Grishin, N. v., Ball, H.L., and Orth, K. (2009). AMPylation of Rho GTPases by Vibrio VopS disrupts effector binding and downstream signaling. *Science* (1979) 323, 269–272. 10.1126/science.1166382.
27. Worby, C.A., Mattoo, S., Kruger, R.P., Corbeil, L.B., Koller, A., Mendez, J.C., Zekarias, B., Lazar, C., and Dixon, J.E. (2009). The Fic Domain: Regulation of Cell Signaling by Adenylylation. *Mol Cell* 34, 93–103. 10.1016/j.molcel.2009.03.008.
28. Sreelatha, A., Yee, S.S., Lopez, V.A., Park, B.C., Kinch, L.N., Pilch, S., Servage, K.A., Zhang, J., Jiou, J., Karasiewicz-Urbańska, M., et al. (2018). Protein AMPylation by an Evolutionarily Conserved Pseudokinase. *Cell* 175, 809--821.e19. 10.1016/j.cell.2018.08.046.
29. Ham, H., Woolery, A.R., Tracy, C., Stenesen, D., Krämer, H., and Orth, K. (2014). Unfolded protein response-regulated Drosophila Fic (dFic) protein reversibly AMPylates BiP chaperone during endoplasmic reticulum homeostasis. *Journal of Biological Chemistry* 289, 36059–36069. 10.1074/jbc.M114.612515.

30. Preissler, S., Rato, C., Chen, R., Antrobus, R., Ding, S., Fearnley, I.M., and Ron, D. (2015). AMPylation matches BiP activity to client protein load in the endoplasmic reticulum. *Elife* 4. 10.7554/elife.12621.
31. Sanyal, A., Chen, A.J., Nakayasu, E.S., Lazar, C.S., Zbornik, E.A., Worby, C.A., Koller, A., and Mattoo, S. (2015). A novel link between fic (filamentation induced by cAMP)-mediated adenylation/AMPylation and the unfolded protein response. *Journal of Biological Chemistry* 290, 8482–8499. 10.1074/jbc.M114.618348.
32. Truttmann, M.C., Zheng, X., Hanke, L., Damon, J.R., Grootveld, M., Krakowiak, J., Pincus, D., and Ploegh, H.L. (2017). Unrestrained AMPylation targets cytosolic chaperones and activates the heat shock response. *Proceedings of the National Academy of Sciences* 114, E152–E160. 10.1073/pnas.1619234114.
33. Munro, S., and Pelham, H.R.B. (1986). An hsp70-like protein in the ER: Identity with the 78 kd glucose-regulated protein and immunoglobulin heavy chain binding protein. *Cell* 46, 291–300. 10.1016/0092-8674(86)90746-4.
34. Gething, M.J. (1999). Role and regulation of the ER chaperone BiP. *Semin Cell Dev Biol* 10, 465–472. 10.1006/scdb.1999.0318.
35. Casey, X.A.K., Moehlman, A.T., Zhang, J., Servage, K.A., Krämer, X.H., and Orth, X.K. (2017). Fic-mediated deAMPylation is not dependent on homodimerization and rescues toxic AMPylation in flies. *Journal of Biological Chemistry* 292, 21193–21204. 10.1074/jbc.M117.799296.
36. Preissler, S., Rato, C., Perera, L.A., Saudek, V., and Ron, D. (2017). FICD acts bifunctionally to AMPylate and de-AMPylate the endoplasmic reticulum chaperone BiP. *Nat Struct Mol Biol* 24, 23–29. 10.1038/nsmb.3337.
37. Preissler, S., Rato, C., Chen, R., Antrobus, R., Ding, S., Fearnley, I.M., and Ron, D. (2015). AMPylation matches BiP activity to client protein load in the endoplasmic reticulum. *Elife* 4, e12621. 10.7554/eLife.12621.
38. Preissler, S., and Ron, D. (2019). Early events in the endoplasmic reticulum unfolded protein response. *Cold Spring Harb Perspect Biol* 11. 10.1101/cshperspect.a033894.
39. Rahman, M., Ham, H., Liu, X., Sugiura, Y., Orth, K., and Krämer, H. (2012). Visual neurotransmission in *Drosophila* requires expression of Fic in glial capitate projections. *Nat Neurosci* 15, 871–875. 10.1038/nn.3102.

40. Moehlman, A.T., Casey, A.K., Servage, K., Orth, K., and Krämer, H. (2018). Adaptation to constant light requires Fic-mediated AMPylation of BiP to protect against reversible photoreceptor degeneration. *Elife* 7, e38752. 10.7554/eLife.38752.
41. McCaul, N., Porter, C.M., Becker, A., Tang, C.-H.A., Wijne, C., Chatterjee, B., Bousbaine, D., Bilate, A., Hu, C.-C.A., Ploegh, H., et al. (2021). Deletion of mFICD AMPylase alters cytokine secretion and affects visual short-term learning in vivo. *Journal of Biological Chemistry* 297. 10.1016/j.jbc.2021.100991.
42. Casey, A.K., Gray, H.F., Chimalapati, S., Hernandez, G., Moehlman, A.T., Stewart, N., Fields, H.A., Gulen, B., Servage, K.A., Stefanius, K., et al. (2022). Fic-mediated AMPylation tempers the unfolded protein response during physiological stress. *Proceedings of the National Academy of Sciences* 119, e2208317119. 10.1073/pnas.2208317119.
43. Yang, Y., Yue, Y., Song, N., Li, C., Yuan, Z., Wang, Y., Ma, Y., Li, H., Zhang, F., Wang, W., et al. (2020). The YdiU Domain Modulates Bacterial Stress Signaling through Mn<sup>2+</sup>-Dependent UMPylation. *Cell Rep* 32. 10.1016/j.celrep.2020.108161.
44. Sanyal, A., Dutta, S., Camara, A., Chandran, A., Koller, A., Watson, B.G., Sengupta, R., Ysselstein, D., Montenegro, P., Cannon, J., et al. (2019). Alpha-Synuclein Is a Target of Fic-Mediated Adenylation/AMPylation: Possible Implications for Parkinson's Disease. *J Mol Biol* 431, 2266–2282. 10.1016/j.jmb.2019.04.026.
45. Rebelo, A.P., Ruiz, A., Dohrn, M.F., Wayand, M., Farooq, A., Danzi, M.C., Beijer, D., Aaron, B., Vandrovцова, J., Houlden, H., et al. (2022). BiP inactivation due to loss of the deAMPylation function of FICD causes a motor neuron disease. *Genetics in Medicine*. 10.1016/j.gim.2022.08.019.
46. Hao, Y.H., Chuang, T., Ball, H.L., Luong, P., Li, Y., Flores-Saaib, R.D., and Orth, K. (2011). Characterization of a rabbit polyclonal antibody against threonine-AMPylation. *J Biotechnol* 151, 251–254. 10.1016/j.jbiotec.2010.12.013.
47. Smit, C., Blümer, J., Eerland, M.F., Albers, M.F., Müller, M.P., Goody, R.S., Itzen, A., and Hedberg, C. (2011). Efficient synthesis and applications of peptides containing adenylylated tyrosine residues. *Angewandte Chemie - International Edition* 50, 9200–9204. 10.1002/anie.201103203.

48. Höpfner, D., Fauser, J., Kaspers, M.S., Pett, C., Hedberg, C., and Itzen, A. (2020). Monoclonal Anti-AMP Antibodies Are Sensitive and Valuable Tools for Detecting Patterns of AMPylation. *iScience* 23, 101800. <https://doi.org/10.1016/j.isci.2020.101800>.
49. Best, M.D. (2009). Click chemistry and bioorthogonal reactions: Unprecedented selectivity in the labeling of biological molecules. *Biochemistry* 48, 6571–6584. 10.1021/bi9007726.
50. Hein, J.E., and Fokin, V. V (2010). Copper-catalyzed azide-alkyne cycloaddition (CuAAC) and beyond: new reactivity of copper(I) acetylides. *Chem Soc Rev* 39, 1302–1315. 10.1039/b904091a.
51. Casey, A.K., and Orth, K. (2018). Enzymes Involved in AMPylation and deAMPylation. *Chem Rev* 118, 1199–1215. 10.1021/acs.chemrev.7b00145.
52. Grammel, M., Luong, P., Orth, K., and Hang, H.C. (2011). A chemical reporter for protein AMPylation. *J Am Chem Soc* 133, 17103–17105. 10.1021/ja205137d.
53. Yu, X., and Labaer, J. (2015). High-throughput identification of proteins with AMPylation using self-assembled human protein (NAPPA) microarrays. *Nat Protoc* 10, 756–767. 10.1038/nprot.2015.044.
54. Yu, X., Woolery, A.R., Luong, P., Hao, Y.H., Grammel, M., Westcott, N., Park, J., Wang, J., Bian, X., Demirkan, G., et al. (2014). Copper-catalyzed azide-alkyne cycloaddition (click chemistry)-based detection of global pathogen-host AMPylation on self-assembled human protein microarrays. *Mol Cell Proteomics* 13, 3164–3176. 10.1074/mcp.M114.041103.
55. Mehellou, Y., Balzarini, J., and McGuigan, C. (2009). Aryloxy phosphoramidate triesters: A technology for delivering monophosphorylated nucleosides and sugars into cells. *ChemMedChem* 4, 1779–1791. 10.1002/cmdc.200900289.
56. Mehellou, Y., Rattan, H.S., and Balzarini, J. (2018). The ProTide Prodrug Technology: From the Concept to the Clinic. *J Med Chem* 61, 2211–2226. 10.1021/acs.jmedchem.7b00734.
57. Chapman, H., Kernan, M., Prisbe, E., Rohloff, J., Sparacino, M., Terhorst, T., and Yu, R. (2001). PRACTICAL SYNTHESIS, SEPARATION, AND STEREOCHEMICAL

ASSIGNMENT OF THE PMPA PRO-DRUG GS-7340. *Nucleosides Nucleotides Nucleic Acids* 20, 621–628. 10.1081/NCN-100002338.

58. Sofia, M.J., Bao, D., Chang, W., Du, J., Nagarathnam, D., Rachakonda, S., Reddy, P.G., Ross, B.S., Wang, P., Zhang, H.-R., et al. (2010). Discovery of a  $\beta$ -d-2'-Deoxy-2'- $\alpha$ -fluoro-2'- $\beta$ -C-methyluridine Nucleotide Prodrug (PSI-7977) for the Treatment of Hepatitis C Virus. *J Med Chem* 53, 7202–7218. 10.1021/jm100863x.
59. Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell* 131, 861–872. <https://doi.org/10.1016/j.cell.2007.11.019>.
60. Takahashi, K., and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* 126, 663–676. 10.1016/j.cell.2006.07.024.
61. Gunhanlar, N., Shpak, G., van der Kroeg, M., Gouty-Colomer, L.A., Munshi, S.T., Lendemeijer, B., Ghazvini, M., Dupont, C., Hoogendijk, W.J.G., Gribnau, J., et al. (2018). A simplified protocol for differentiation of electrophysiologically mature neuronal networks from human induced pluripotent stem cells. *Mol Psychiatry* 23, 1336–1344. 10.1038/mp.2017.56.
62. Subramaniam, M., Chong, S.A., Vaingankar, J.A., Abdin, E., Chua, B.Y., Chua, H.C., Eng, G.K., Heng, D., Hia, S.B., Huang, W., et al. (2015). Prevalence of Dementia in People Aged 60 Years and Above: Results from the WiSE Study. *Journal of Alzheimer's Disease* 45, 1127–1138. 10.3233/JAD-142769.
63. Prince, M.J., Wimo, A., Guerchet, M.M., Ali, G.C., Wu, Y.-T., and Prina, M. (2015). World Alzheimer Report 2015 - The Global Impact of Dementia (Alzheimer's Disease International).
64. Cerejeira, J., Lagarto, L., and Mukaetova-Ladinska, E.B. (2012). Behavioral and psychological symptoms of dementia. *Front Neurol* MAY. 10.3389/fneur.2012.00073.
65. Cummings, J.L., Goldman, D.P., Simmons-Stern, N.R., and Ponton, E. (2022). The costs of developing treatments for Alzheimer's disease: A retrospective exploration. *Alzheimer's and Dementia* 18, 469–477. 10.1002/alz.12450.



66. Long, J.M., and Holtzman, D.M. (2019). Alzheimer Disease: An Update on Pathobiology and Treatment Strategies. *Cell* 179, 312–339. 10.1016/j.cell.2019.09.001.
67. McCleery, J., and Quinn, T.J. (2021). Aducanumab and the certainty of evidence. *Age Ageing* 50, 1899–1900. 10.1093/ageing/afab167.
68. Thambisetty, M., Howard, R., Glymour, M.M., and Schneider, L.S. (2021). Alzheimer’s drugs: Does reducing amyloid work? *Science* (1979) 374, 544–545. 10.1126/science.abl8366.
69. Fazzari, P., Horre, K., Arranz, A.M., Frigerio, C.S., Saito, T., Saido, T.C., and de Strooper, B. (2017). PLD3 gene and processing of APP. *Nature* 541, E1–E2. 10.1038/nature21030.
70. Gonzalez, A.C., Schweizer, M., Jagdmann, S., Bernreuther, C., Reinheckel, T., Saftig, P., and Damme, M. (2018). Unconventional Trafficking of Mammalian Phospholipase D3 to Lysosomes. *Cell Rep* 22, 1040–1053. 10.1016/j.celrep.2017.12.100.
71. Gavin, A.L., Huang, D., Huber, C., Mårtensson, A., Tardif, V., Skog, P.D., Blane, T.R., Thinnes, T.C., Osborn, K., Chong, H.S., et al. (2018). PLD3 and PLD4 are single-stranded acid exonucleases that regulate endosomal nucleic-acid sensing. *Nat Immunol* 19, 942–953. 10.1038/s41590-018-0179-y.
72. Nackenoff, A.G., Hohman, T.J., Neuner, S.M., Akers, C.S., Weitzel, N.C., Shostak, A., Ferguson, S.M., Mobley, B., Bennett, D.A., Schneider, J.A., et al. (2021). PLD3 is a neuronal lysosomal phospholipase D associated with  $\beta$ -amyloid plaques and cognitive function in Alzheimer’s disease. *PLoS Genet* 17, e1009406-.
73. Gavin, A.L., Huang, D., Blane, T.R., Thinnes, T.C., Murakami, Y., Fukui, R., Miyake, K., and Nemazee, D. (2021). Cleavage of DNA and RNA by PLD3 and PLD4 limits autoinflammatory triggering by multiple sensors. *Nat Commun* 12, 5874. 10.1038/s41467-021-26150-w.
74. Agard, N.J., Prescher, J.A., and Bertozzi, C.R. (2004). A Strain-Promoted [3 + 2] Azide–Alkyne Cycloaddition for Covalent Modification of Biomolecules in Living Systems. *J Am Chem Soc* 126, 15046–15047. 10.1021/ja044996f.
75. Parker, C.G., and Pratt, M.R. (2020). Click Chemistry in Proteomic Investigations. *Cell* 180, 605–632. <https://doi.org/10.1016/j.cell.2020.01.025>.

76. Kallemeijn, W.W., Lanyon-Hogg, T., Panyain, N., Goya Grocin, A., Ciepla, P., Morales-Sanfrutos, J., and Tate, E.W. (2021). Proteome-wide analysis of protein lipidation using chemical probes: in-gel fluorescence visualization, identification and quantification of N-myristoylation, N- and S-acylation, O-cholesterylation, S-farnesylation and S-geranylgeranylation. *Nat Protoc* 16, 5083–5122. 10.1038/s41596-021-00601-6.
77. Hughes, C.S., Moggridge, S., Müller, T., Sorensen, P.H., Morin, G.B., and Krijgsveld, J. (2019). Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nat Protoc* 14, 68–85. 10.1038/s41596-018-0082-x.
78. Hughes, C.S., Foehr, S., Garfield, D.A., Furlong, E.E., Steinmetz, L.M., and Krijgsveld, J. (2014). Ultrasensitive proteome analysis using paramagnetic bead technology. *Mol Syst Biol* 10, 757. 10.15252/msb.20145625.
79. Yan, T., Desai, H.S., Boatner, L.M., Yen, S.L., Cao, J., Palafox, M.F., Jami-Alahmadi, Y., and Backus, K.M. (2021). SP3-FAIMS Chemoproteomics for High-Coverage Profiling of the Human Cysteinome\*\*. *ChemBioChem* 22, 1841–1851. <https://doi.org/10.1002/cbic.202000870>.
80. Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F., and Whitehouse, C.M. (1989). Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science* (1979) 246, 64–71. 10.1126/science.2675315.
81. El-Aneed, A., Cohen, A., and Banoub, J. (2009). Mass Spectrometry, Review of the Basics: Electrospray, MALDI, and Commonly Used Mass Analyzers. *Appl Spectrosc Rev* 44, 210–230. 10.1080/05704920902717872.
82. Scigelova, M., and Makarov, A. (2006). Orbitrap Mass Analyzer – Overview and Applications in Proteomics. *Proteomics* 6, 16–21. <https://doi.org/10.1002/pmic.200600528>.
83. Michalski, A., Damoc, E., Hauschild, J.-P., Lange, O., Wiegand, A., Makarov, A., Nagaraj, N., Cox, J., Mann, M., and Horning, S. (2011). Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer. *Mol Cell Proteomics* 10, M111.011015-M111.011015. 10.1074/mcp.M111.011015.

84. Eliuk, S., and Makarov, A. (2015). Evolution of Orbitrap Mass Spectrometry Instrumentation. *Annual Review of Analytical Chemistry* 8, 61–80. 10.1146/annurev-anchem-071114-040325.
85. Senko, M.W., Remes, P.M., Canterbury, J.D., Mathur, R., Song, Q., Eliuk, S.M., Mullen, C., Earley, L., Hardman, M., Blethrow, J.D., et al. (2013). Novel Parallelized Quadrupole/Linear Ion Trap/Orbitrap Tribid Mass Spectrometer Improving Proteome Coverage and Peptide Identification Rates. *Anal Chem* 85, 11710–11714. 10.1021/ac403115c.
86. Yan, X., Zhang, Y., Zhou, Y., Li, G., and Feng, X. (2022). Technical Overview of Orbitrap High Resolution Mass Spectrometry and Its Application to the Detection of Small Molecules in Food (Update Since 2012). *Crit Rev Anal Chem* 52, 593–626. 10.1080/10408347.2020.1815168.
87. Tabb, D.L., Vega-Montoto, L., Rudnick, P.A., Variyath, A.M., Ham, A.J.L., Bunk, D.M., Kilpatrick, L.E., Billheimer, D.D., Blackman, R.K., Cardasis, H.L., et al. (2010). Repeatability and reproducibility in proteomic identifications by liquid chromatography-tandem mass spectrometry. *J Proteome Res* 9, 761–776. 10.1021/pr9006365.
88. Bruderer, R., Bernhardt, O.M., Gandhi, T., Miladinović, S.M., Cheng, L.Y., Messner, S., Ehrenberger, T., Zanotelli, V., Butscheid, Y., Escher, C., et al. (2015). Extending the limits of quantitative proteome profiling with data-independent acquisition and application to acetaminophen-treated three-dimensional liver microtissues. *Molecular and Cellular Proteomics* 14, 1400–1410. 10.1074/mcp.M114.044305.
89. Schubert, O.T., Gillet, L.C., Collins, B.C., Navarro, P., Rosenberger, G., Wolski, W.E., Lam, H., Amodei, D., Mallick, P., MacLean, B., et al. (2015). Building high-quality assay libraries for targeted analysis of SWATH MS data. *Nat Protoc* 10, 426–441. 10.1038/nprot.2015.015.
90. Gillet, L.C., Navarro, P., Tate, S., Röst, H., Selevsek, N., Reiter, L., Bonner, R., and Aebersold, R. (2012). Targeted Data Extraction of the MS/MS Spectra Generated by Data-independent Acquisition: A New Concept for Consistent and Accurate Proteome Analysis\*. *Molecular & Cellular Proteomics* 11, O111.016717. <https://doi.org/10.1074/mcp.O111.016717>.

91. Ludwig, C., Gillet, L., Rosenberger, G., Amon, S., Collins, B.C., and Aebersold, R. (2018). Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial. *Mol Syst Biol* *14*, e8126. <https://doi.org/10.15252/msb.20178126>.
92. Tsou, C.-C., Avtonomov, D., Larsen, B., Tucholska, M., Choi, H., Gingras, A.-C., and Nesvizhskii, A.I. (2015). DIA-Umpire: comprehensive computational framework for data-independent acquisition proteomics. *Nat Methods* *12*, 258–264. [10.1038/nmeth.3255](https://doi.org/10.1038/nmeth.3255).
93. Tiwary, S., Levy, R., Gutenbrunner, P., Salinas Soto, F., Palaniappan, K.K., Deming, L., Berndl, M., Brant, A., Cimermancic, P., and Cox, J. (2019). High-quality MS/MS spectrum prediction for data-dependent and data-independent acquisition data analysis. *Nat Methods* *16*, 519–525. [10.1038/s41592-019-0427-6](https://doi.org/10.1038/s41592-019-0427-6).
94. Gessulat, S., Schmidt, T., Zolg, D.P., Samaras, P., Schnatbaum, K., Zerweck, J., Knaute, T., Rechenberger, J., Delanghe, B., Huhmer, A., et al. (2019). Prosit: proteome-wide prediction of peptide tandem mass spectra by deep learning. *Nat Methods* *16*, 509–518. [10.1038/s41592-019-0426-7](https://doi.org/10.1038/s41592-019-0426-7).
95. Yang, Y., Liu, X., Shen, C., Lin, Y., Yang, P., and Qiao, L. (2020). In silico spectral libraries by deep learning facilitate data-independent acquisition proteomics. *Nat Commun* *11*, 146. [10.1038/s41467-019-13866-z](https://doi.org/10.1038/s41467-019-13866-z).
96. Demichev, V., Messner, C.B., Vernardis, S.I., Lilley, K.S., and Ralser, M. (2020). DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat Methods* *17*, 41–44. [10.1038/s41592-019-0638-x](https://doi.org/10.1038/s41592-019-0638-x).
97. Beers, J., Gulbranson, D.R., George, N., Siniscalchi, L.I., Jones, J., Thomson, J.A., and Chen, G. (2012). Passaging and colony expansion of human pluripotent stem cells by enzyme-free dissociation in chemically defined culture conditions. *Nat Protoc* *7*, 2029–2040. [10.1038/nprot.2012.130](https://doi.org/10.1038/nprot.2012.130).
98. Yang, X., and Qian, K. (2017). Protein O-GlcNAcylation: emerging mechanisms and functions. *Nat Rev Mol Cell Biol* *18*, 452–465. [10.1038/nrm.2017.22](https://doi.org/10.1038/nrm.2017.22).
99. Li, J., van Vranken, J.G., Pontano Vaiteș, L., Schweppe, D.K., Huttlin, E.L., Etienne, C., Nandhikonda, P., Viner, R., Robitaille, A.M., Thompson, A.H., et al. (2020). TMTpro reagents: a set of isobaric labeling mass tags enables simultaneous proteome-wide

- measurements across 16 samples. *Nat Methods* 17, 399–404. 10.1038/s41592-020-0781-4.
100. Stadlmeier, M., Bogena, J., Wallner, M., Wühr, M., and Carell, T. (2018). A Sulfoxide-Based Isobaric Labelling Reagent for Accurate Quantitative Mass Spectrometry. *Angewandte Chemie International Edition* 57, 2958–2962. <https://doi.org/10.1002/anie.201708867>.
  101. Becker, T., Wiest, A., Telek, A., Bejko, D., Hoffmann-Röder, A., and Kielkowski, P. (2022). Transforming Chemical Proteomics Enrichment into a High-Throughput Method Using an SP2E Workflow. *JACS Au* 2, 1712–1723. 10.1021/jacsau.2c00284.
  102. Zhang, Y., Thompson, R., Zhang, H., and Xu, H. (2011). APP processing in Alzheimer's disease. *Mol Brain* 4, 3. 10.1186/1756-6606-4-3.
  103. Dontu, G., Jackson, K.W., McNicholas, E., Kawamura, M.J., Abdallah, W.M., and Wicha, M.S. (2004). Role of Notch signaling in cell-fate determination of human mammary stem/progenitor cells. *Breast Cancer Research* 6, R605. 10.1186/bcr920.
  104. Cheng, H.-T., Kim, M., Valerius, M.T., Surendran, K., Schuster-Gossler, K., Gossler, A., McMahon, A.P., and Kopan, R. (2007). Notch2, but not Notch1, is required for proximal fate acquisition in the mammalian nephron. *Development* 134, 801–811. 10.1242/dev.02773.
  105. Mollenhauer, H.H., Morré, D.J., and Rowe, L.D. (1990). Alteration of intracellular traffic by monensin; mechanism, specificity and relationship to toxicity. *Biochim Biophys Acta* 1031, 225–246.
  106. Yoshimori, T., Yamamoto, A., Moriyama, Y., Futai, M., and Tashiro, Y. (1991). Bafilomycin A1, a specific inhibitor of vacuolar-type H<sup>+</sup>-ATPase, inhibits acidification and protein degradation in lysosomes of cultured cells. *Journal of Biological Chemistry* 266, 17707–17712. 10.1016/s0021-9258(19)47429-2.
  107. Mauthe, M., Orhon, I., Rocchi, C., Zhou, X., Luhr, M., Hijlkema, K.-J., Coppes, R.P., Engedal, N., Mari, M., and Reggiori, F. (2018). Chloroquine inhibits autophagic flux by decreasing autophagosome-lysosome fusion. *Autophagy* 14, 1435–1455. 10.1080/15548627.2018.1474314.

108. Catalano, M., and O'Driscoll, L. (2020). Inhibiting extracellular vesicles formation and release: a review of EV inhibitors. *J Extracell Vesicles* 9. 10.1080/20013078.2019.1703244.
109. Becker, T., Cappel, C., di Matteo, F., Sonsalla, G., Kaminska, E., Spada, F., Cappello, S., Damme, M., and Kielkowski, P. (2021). AMPylation profiling during neuronal differentiation reveals extensive variation on lysosomal proteins. *iScience* 24. 10.1016/j.isci.2021.103521.
110. Lancaster, M.A., Renner, M., Martin, C.-A., Wenzel, D., Bicknell, L.S., Hurles, M.E., Homfray, T., Penninger, J.M., Jackson, A.P., and Knoblich, J.A. (2013). Cerebral organoids model human brain development and microcephaly. *Nature* 501, 373–379. 10.1038/nature12517.

## 6 List of abbreviations

<b>AMP</b>	Adenosine monophosphate
<b>APP</b>	$\beta$ -amyloid precursor protein
<b>ATase</b>	Adenylyl transferase domain
<b>ATP</b>	Adenosine triphosphate
<b>COs</b>	Cerebral organoids
<b>CuAAC</b>	Cu(I)-catalyzed azide-alkyne cycloaddition
<b>DDA</b>	Data-dependent acquisition
<b>DIA</b>	Data-independent acquisition
<b>Dox</b>	Doxycycline
<b>ER</b>	Endoplasmic reticulum
<b>ESCRT</b>	Endosomal sorting complex required for transport
<b>ESI</b>	Electrospray ionization
<b>ETD</b>	Electron-dissociation
<b>Fic</b>	Filamentation induced cAMP
<b>FICD</b>	Fic domain containing protein
<b>GWAS</b>	Genomic wide association study
<b>HCD</b>	Higher energy collisional dissociation
<b>HLH</b>	Hemophagocytic lymphohistiocytosis
<b>HPLC</b>	High-performance liquid chromatography
<b>HYPE</b>	Huntingtin yeast-interacting protein E
<b>ILV</b>	Intraluminal vesicles
<b>IP</b>	Immunoprecipitation
<b>iPSCs</b>	Induced pluripotent stem cells
<b>IVTT</b>	<i>In vitro</i> transcription and translation
<i>m/z</i>	Mass-to-charge ratio
<b>MVB</b>	Multivesicular bodies
<b>N6pATP</b>	N6-propargyl adenosine-5'-O-triphosphate
<b>NA</b>	Nucleoside analogues
<b>NAPPA</b>	Nucleic acid programmable protein array

<b>NOTCH2</b>	Neurogenic locus notch homolog protein 2
<b>NPCs</b>	Neural progenitor cells
<b>PARPs</b>	Poly(ADP-ribose)-polymerases
<b>PTMs</b>	Post-translational modifications
<b>SELENOO</b>	Selenoprotein-O
<b>SP3</b>	Single-Pot Solid-Phase-enhanced Sample Preparation
<b>SPAAC</b>	Strain-promoted azide-alkyne cycloaddition
<b>TBTA</b>	Tris(benzyltriazolyl)methyl amine
<b>TCEP</b>	Tris(2-carboxyethyl)phosphine
<b>TLR9</b>	Toll-like receptor 9
<b>TOF</b>	Time-of-flight
<b>UPR</b>	Unfolded protein response



## 7 List of figures

Figure 1: Schematic representation of the proteome complexity.....	1
Figure 2: Schematic representation of the diversity of PTMs.....	2
Figure 3: Reaction scheme of protein AMPylation. ....	3
Figure 4: Regulation of the HSPA5 activity by AMPylation upon ER stress. ....	4
Figure 5: <i>In vitro</i> methods to analyze AMPylation.....	8
Figure 6: Schematic representation of the ProTide strategy.....	10
Figure 7: Schematic representation of the iNGN differentiation process. ....	11
Figure 8: Schematic representation of the transport mechanism and processing of PLD3. ....	14
Figure 9: Schematic representation of the general chemical proteomic sample preparation workflow.....	16
Figure 10: General setup of a mass spectrometer. ....	18
Figure 11: Comparison of DDA and DIA data acquisition. ....	20
Figure 12 Comparison of DDA and DIA measurement.....	69
Figure 13 Reduction of carboxylate-coated paramagnetic beads in the small-scale SP2E workflow. ....	71
Figure 14 AMPylation profiling in SH-SY5Y cells using chloroquine and GW4869.....	75
Figure 15 AMPylation profiling in cerebral organoids using bafilomycin, monensin and chloroquine. ....	77

## 8 Appendix

### 8.1 Supplementary information: AMPylation profiling during neuronal differentiation reveals extensive variation on lysosomal proteins

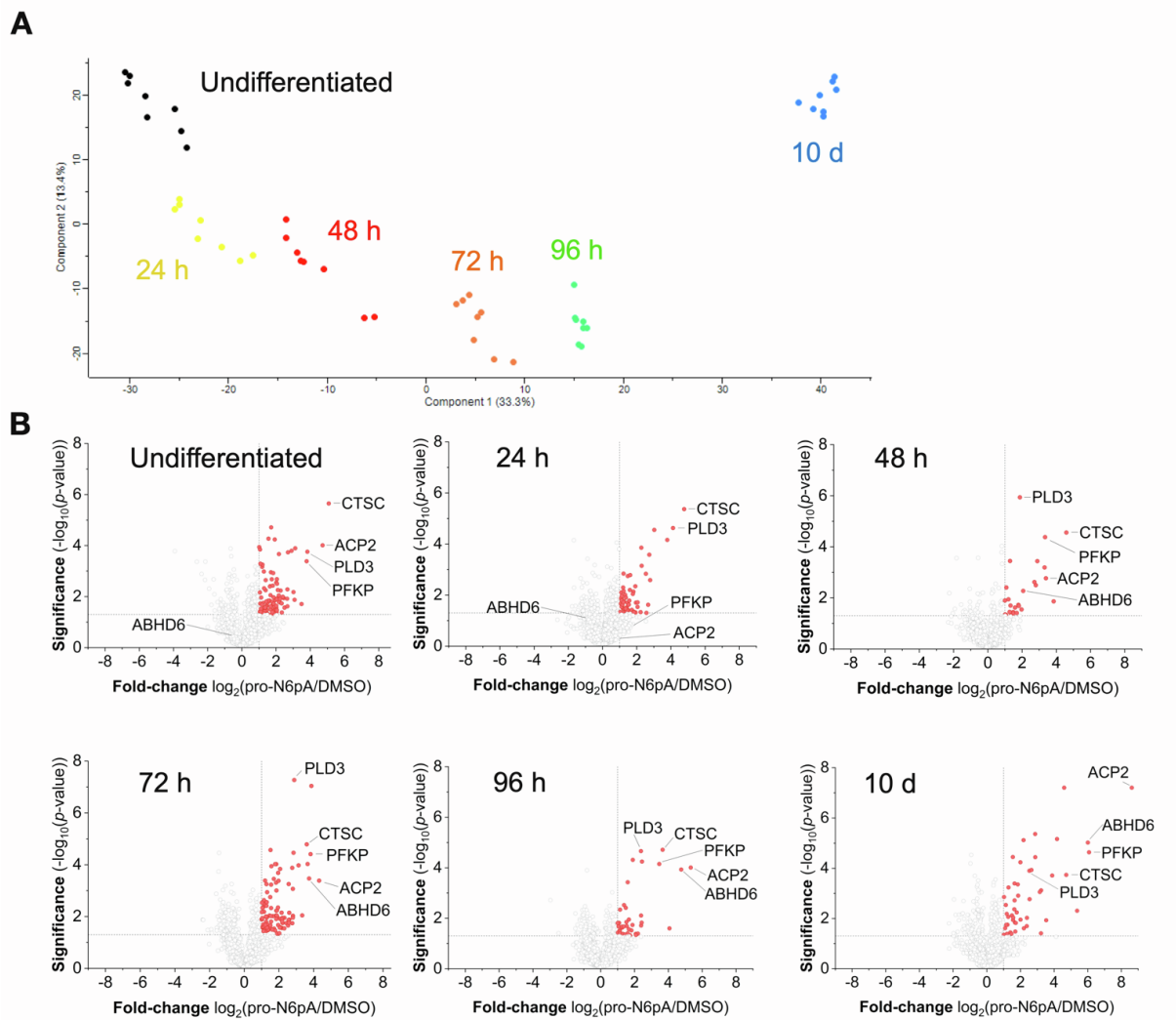
Becker, T., Cappel, C., di Matteo, F., Sonsalla, G., Kaminska, E., Spada, F., Cappello, S., Damme, M., and Kielkowski, P. (2021). AMPylation profiling during neuronal differentiation reveals extensive variation on lysosomal proteins. *iScience* 24. 10.1016/j.isci.2021.103521.

iScience, Volume 24

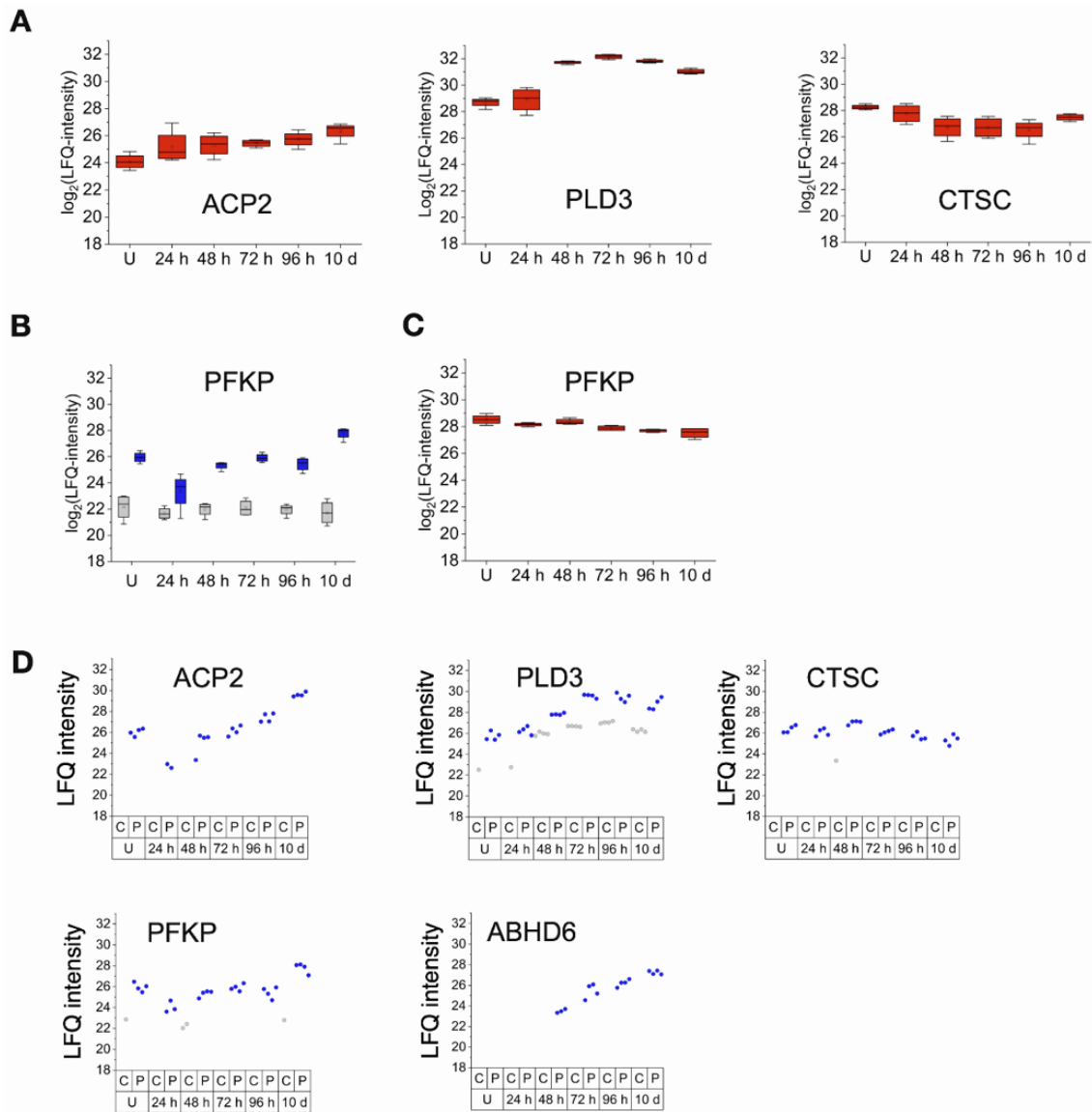
## **Supplemental information**

### **AMPylation profiling during neuronal differentiation reveals extensive variation on lysosomal proteins**

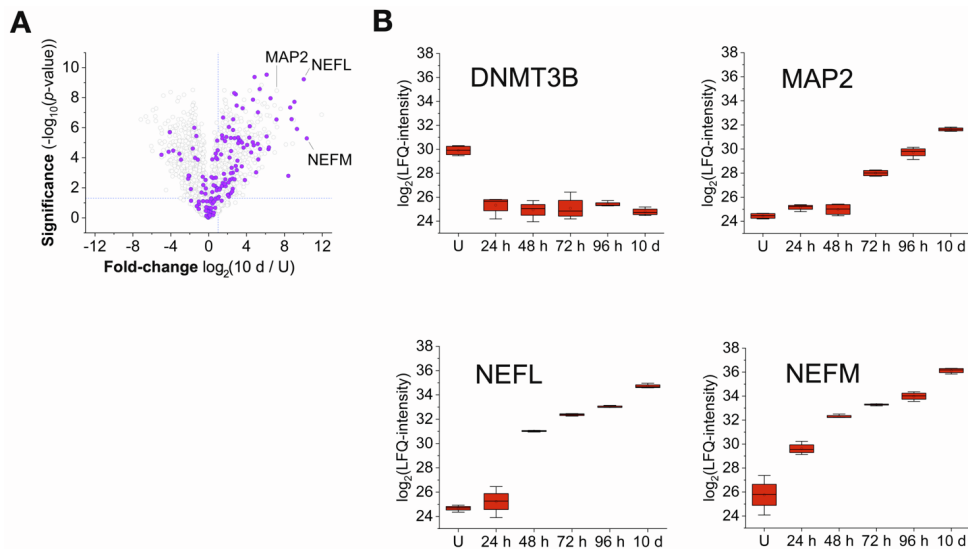
**Tobias Becker, Cedric Cappel, Francesco Di Matteo, Giovanna Sonsalla, Ewelina Kaminska, Fabio Spada, Silvia Cappello, Markus Damme, and Pavel Kielkowski**



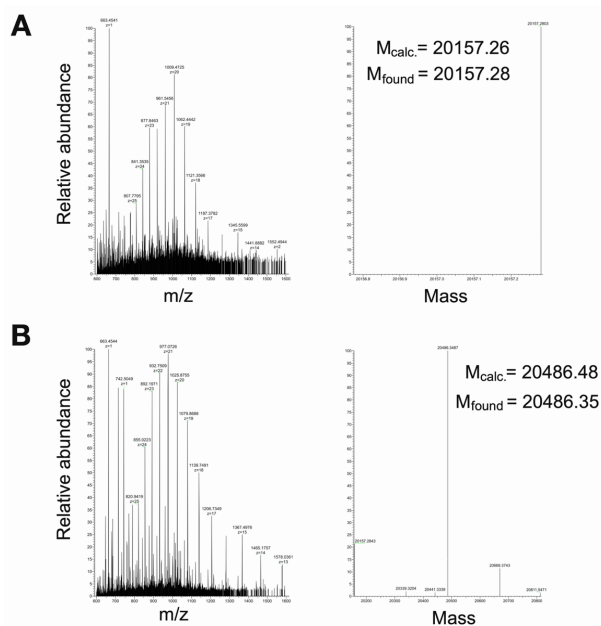
**Supplemental Figure 1.** Chemical proteomics of neuronal differentiation and maturation. Related to **Figure 2**. **A**) Principal component analysis (PCA) of the chemical proteomic LFQ MS data. **B**) Volcano plots showing the enrichment of AMPylated proteins at different times during differentiation. Red circles highlight the significantly enriched proteins. ( $n = 4$ , cut-off lines are at  $p$ -value  $>0.05$  and at least 2-fold enrichment).



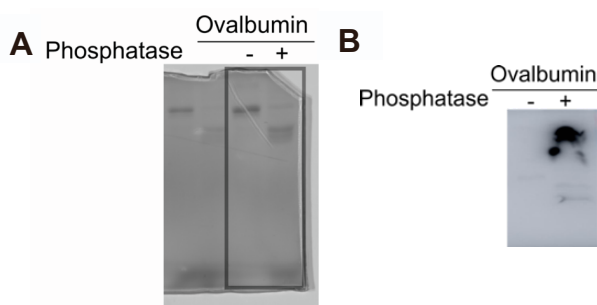
**Supplemental Figure 2.** Profile plots of chemical proteomics of neuronal differentiation and maturation. Related to **Figure 2**. **A)** Profile plots showing total protein expressions estimated from the whole proteome analysis during the iNGN differentiation and maturation. **B)** Profile plot visualizing the enrichment (blue boxes) of the PFKP in the chemical proteomic experiment. Grey boxes stand for PFKP background binding to the agarose-beads found in the DMSO treated control. **C)** Profile plot showing total expression of PFKP estimated from the whole proteome analysis during the course of differentiation. **D)** Profile plots of the selected hits before imputation of the missing values to visualize the efficiency of the enrichment process and utility of the chemical proteomic workflow. Blue dots represent the probe treated samples (P), gray dots represent the DMSO treated samples (C).



**Supplemental Figure 3.** iNGN differentiation and maturation. Related to **Figure 2**. **A)** Volcano plot representing the difference in protein expression between undifferentiated iNGNs and 10 d old neurons. Purple circles label protein annotated as neuronal processes by GO terms. **B)** Selected protein markers of neural maturation.

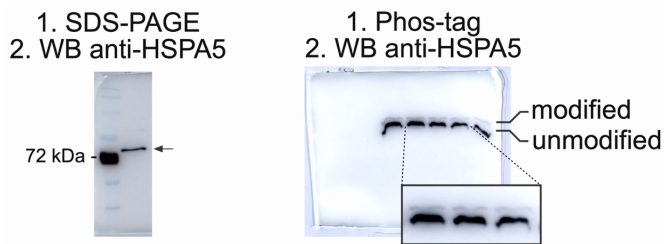


**Supplemental Figure 4.** Top-down mass spectrometry of Rab1b. Related to **Figure 3**. **A)** Unmodified Rab1b. **B)** AMPylated Rab1b.  $\Delta M$  is 329.07 Da corresponding to the AMP moiety.

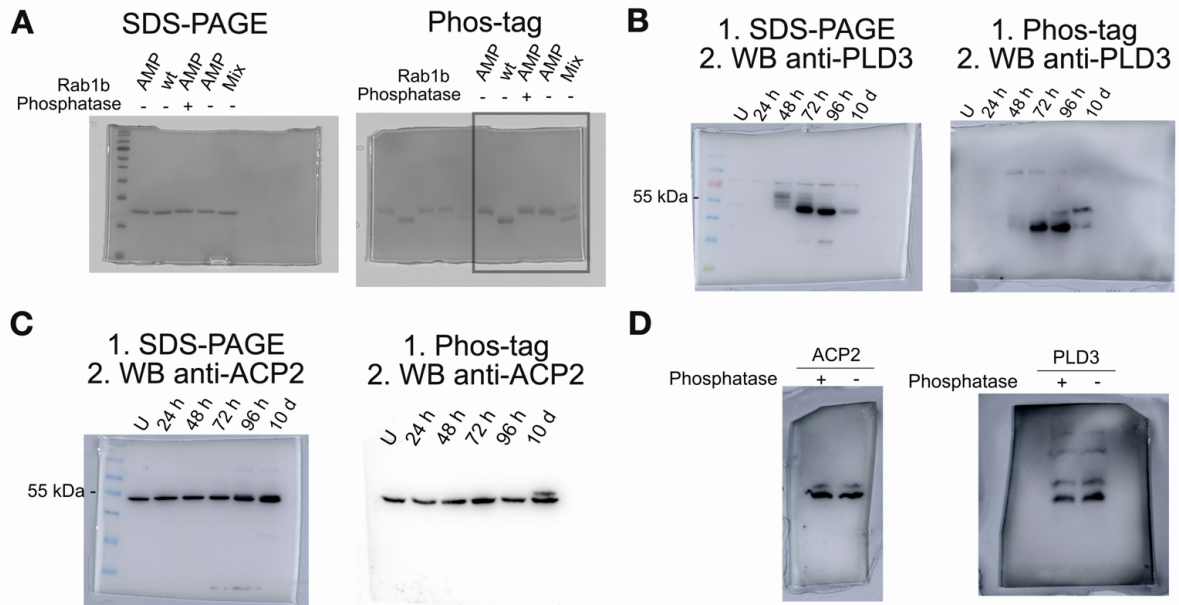


**Supplemental Figure 5.** Dephosphorylation of Ovalbumin. Related to **Figure 3**. **A)** Phos-tag gel separation of Ovalbumin with and without the phosphatase treatment, Coomassie stain. **B)** Phos-tag

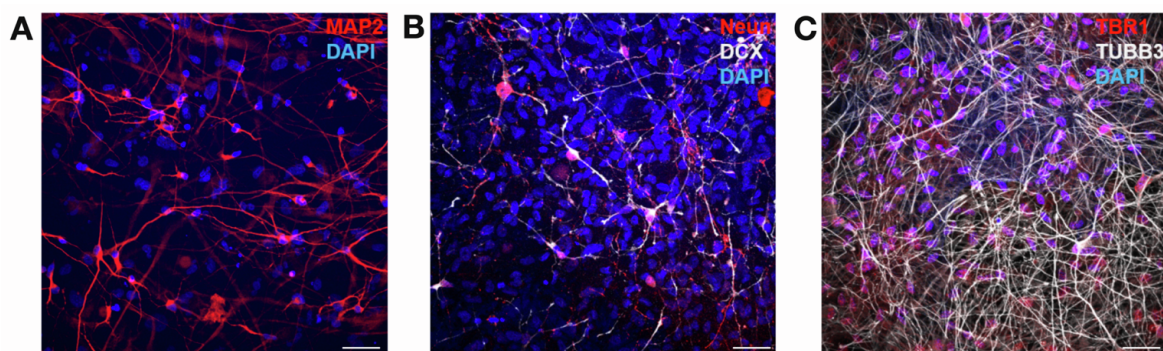
gel separation of Ovalbumin with and without the phosphatase treatment, 5 ng Ovalbumin added to the lysis buffer. Staining using anti-Ovalbumin antibody.



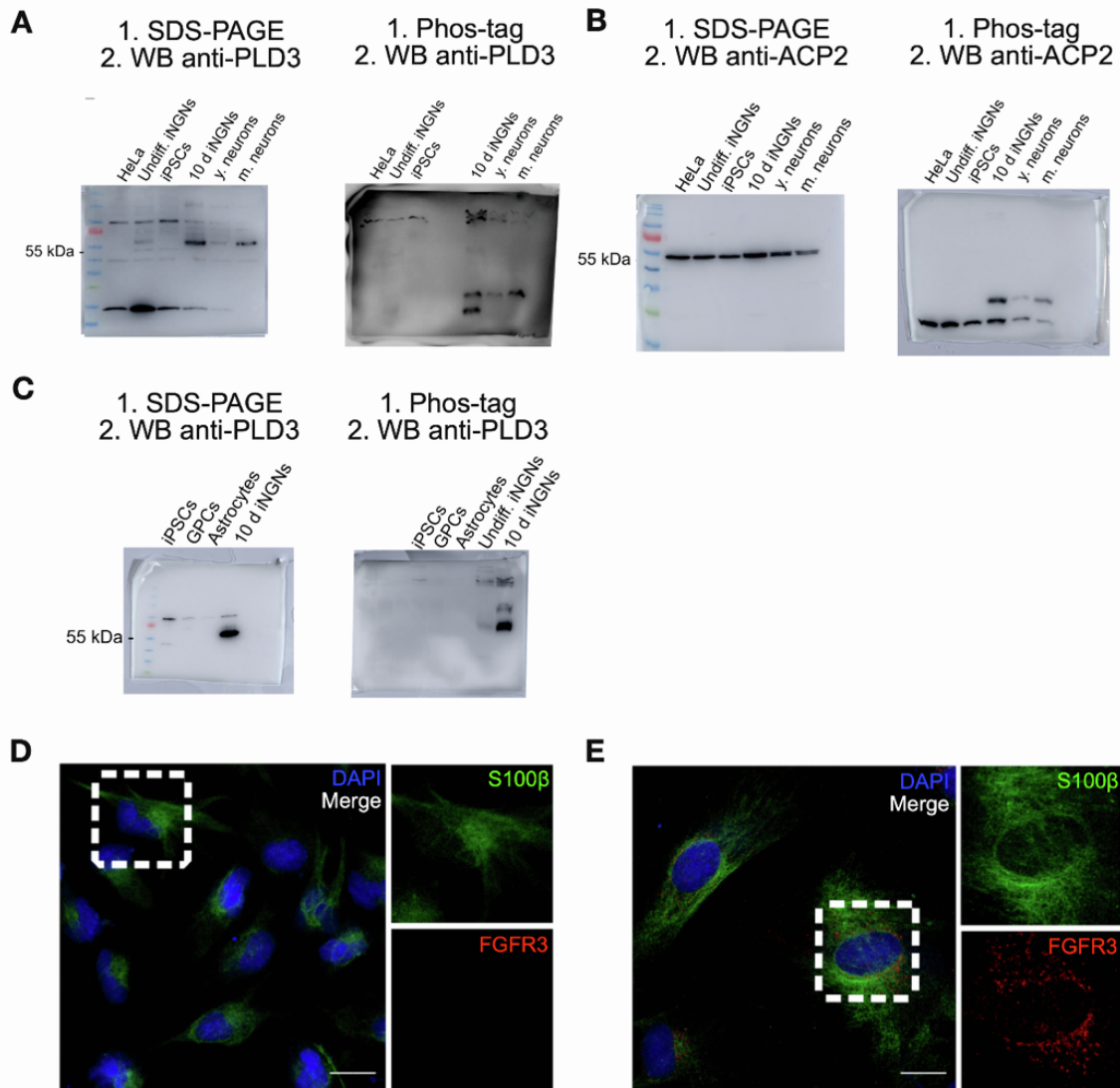
**Supplemental Figure 6.** HSPA5 separation. Related to **Figure 3**. All lines are from HeLa lysates.



**Supplemental Figure 7.** Uncut SDS-PAGE and Phos-tag gels used in **Figure 3**. **A)** Coomassie stained SDS-PAGE and Phos-tag gel separation of AMPylated and non-modified Rab1b. **B)** Western blots of the SDS-PAGE and Phos-tag gel separation of PLD3 during differentiation course of iNGN cells. **C)** Western blots of the SDS-PAGE and Phos-tag gel separation of ACP2 during differentiation course of iNGN cells. **D)** Western blots of the SDS-PAGE and Phos-tag gel separation of PLD3 and ACP2 in 10 days differentiated iNGN cells treated with shrimp alkaline phosphatase.

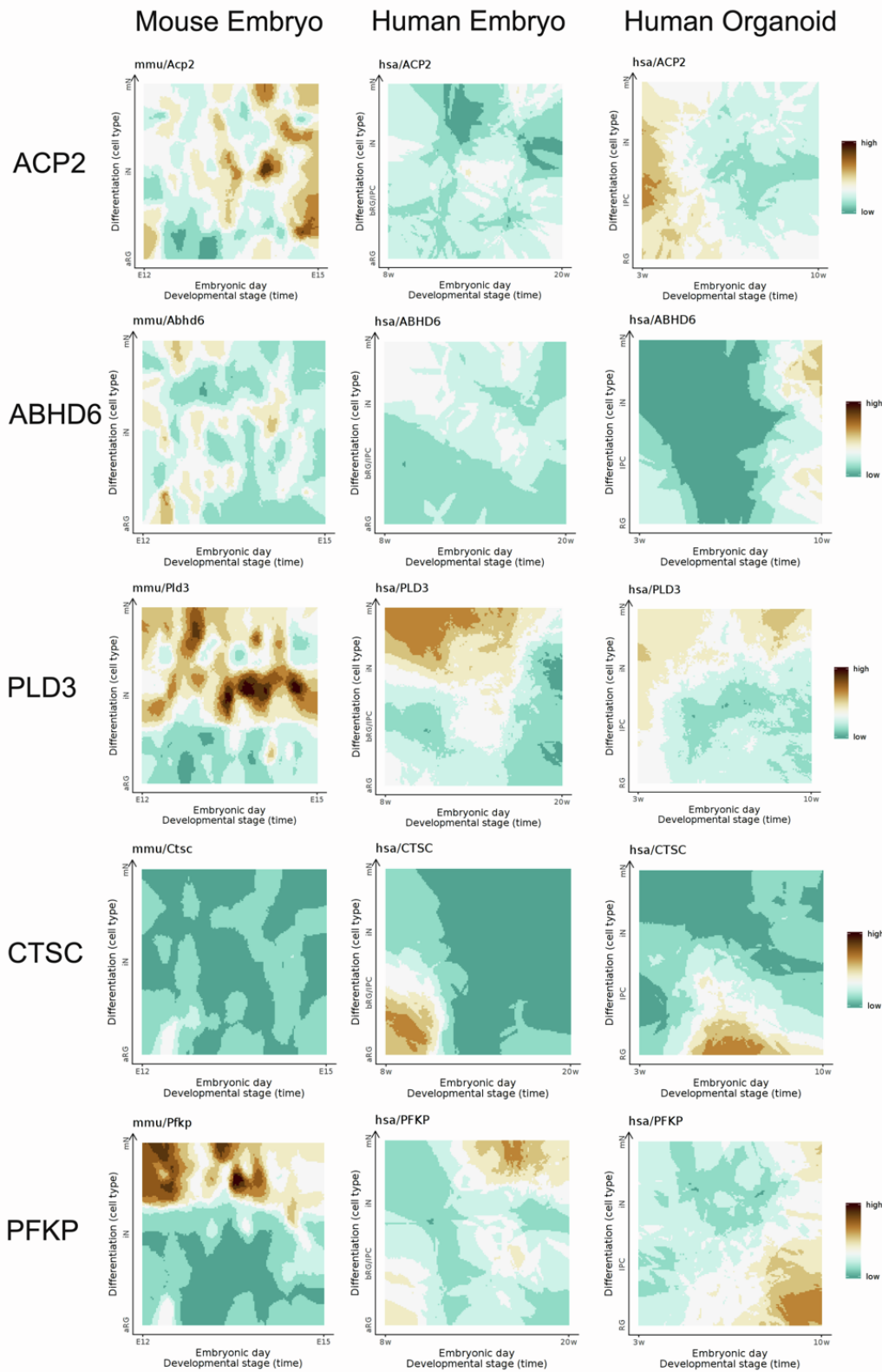


**Supplemental Figure 8.** Characterization of differentiation and maturation of physiological neurons. Related to **Figure 4**. **A - C)** Micrographs of 2D neuronal cultures after 10 weeks of differentiation. Cells were immunostained for **A)** MAP2, **B)** Neun, Doublecortin (DCX), **C)** TBR1 and TUBB3. Nuclei (blue) are stained with DAPI. Scale bars: 50  $\mu$ m.

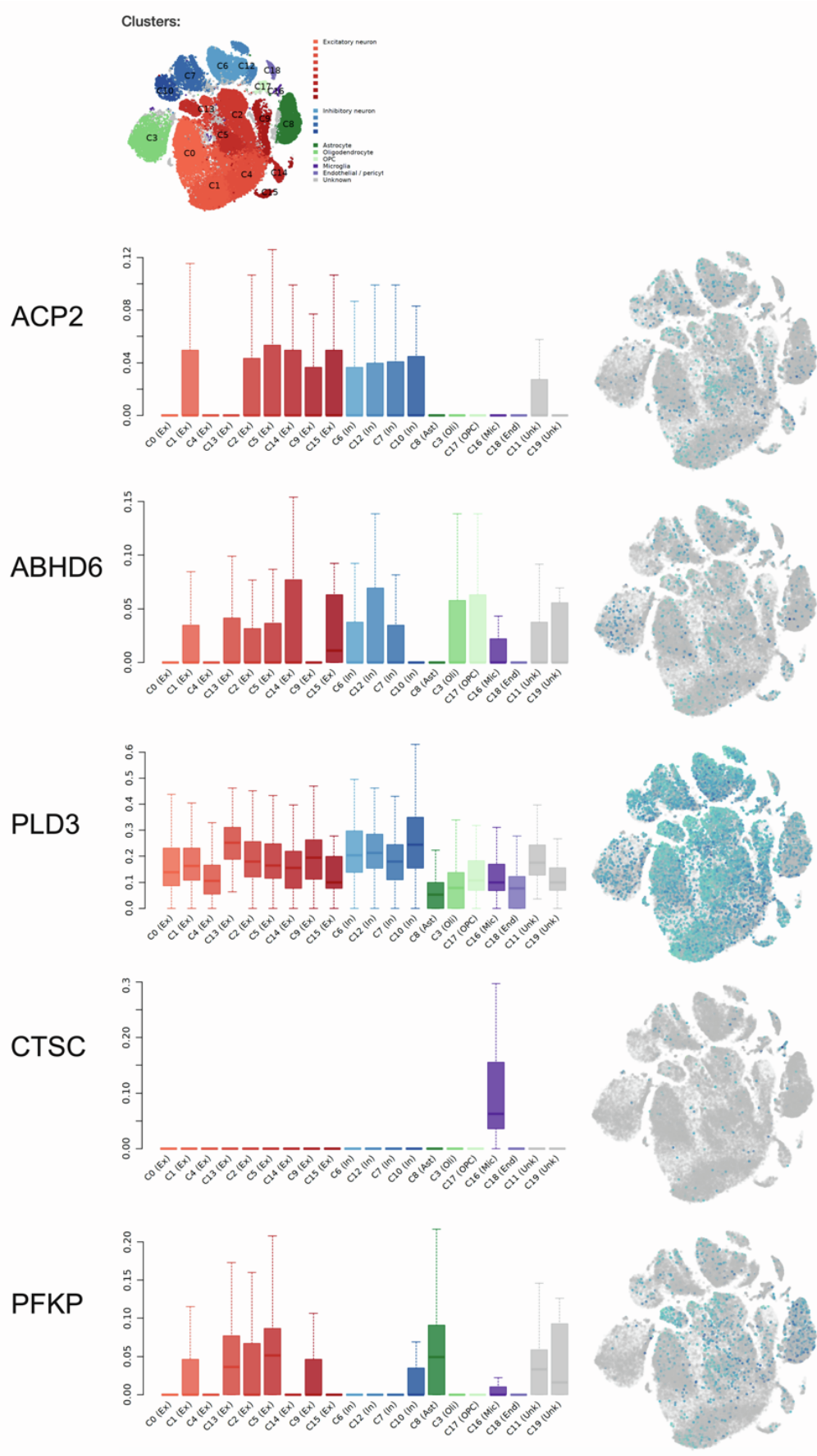


**Supplemental Figure 9.** Uncut SDS-PAGE and Phos-tag gels used in **Figure 4**. **A)** Western blots of the SDS-PAGE and Phos-tag gel separation of PLD3 comparing HeLa, iNGNs and dopaminergic neurons. **B)** Western blots of the SDS-PAGE and Phos-tag gel separation of ACP2 comparing HeLa, iNGNs and dopaminergic neurons. **D)** and **E)** Characterization of astrocytes differentiation. Nuclei (blue) are stained with DAPI. Scale bars: 20  $\mu$ m. **D)** GPCs. **E)** Astrocytes.





**Supplemental Figure 10.** Protein expression changes during embryo development. Related to discussion.



**Supplemental Figure 11.** Analysis of protein expression levels in various cell types. Related to discussion.

## **8.2 Supplementary information: Transforming Chemical Proteomics Enrichment into a High-Throughput Method Using an SP2E Workflow**

Becker, T., Wiest, A., Telek, A., Bejko, D., Hoffmann-Röder, A., and Kielkowski, P. (2022). Transforming Chemical Proteomics Enrichment into a High-Throughput Method Using an SP2E Workflow. *JACS Au* 2, 1712–1723. [10.1021/jacsau.2c00284](https://doi.org/10.1021/jacsau.2c00284).

## Transforming Chemical Proteomics Enrichment into a High-Throughput Method Using an SP2E workflow.

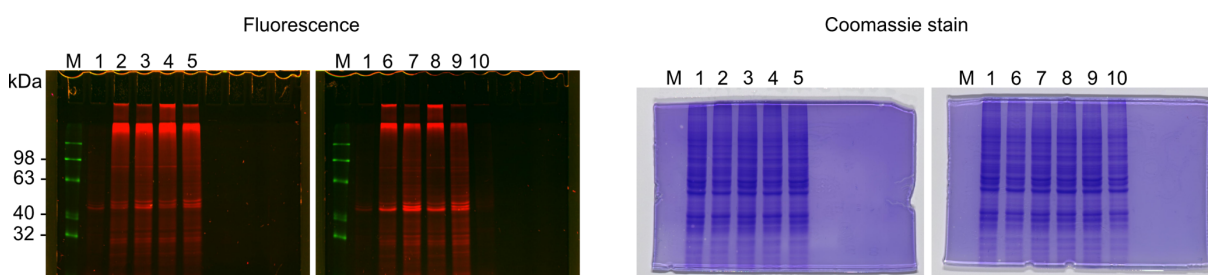
Tobias Becker<sup>1</sup>, Andreas Wiest<sup>1</sup>, András Telek<sup>1</sup>, Daniel Bejko<sup>1</sup>, Anja Hoffmann-Röder<sup>2</sup>, Pavel Kielkowski<sup>1\*</sup>

<sup>1</sup> LMU Munich, Institute for Chemical Epigenetics Munich, 81375 Munich, Germany.

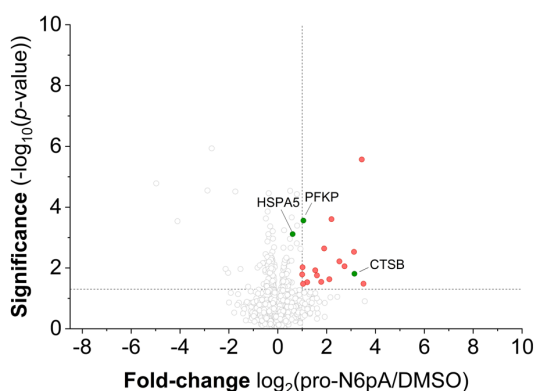
<sup>2</sup> LMU Munich, Department of Chemistry, 81377 Munich, Germany.

\* correspondence to [pavel.kielkowski@cup.lmu.de](mailto:pavel.kielkowski@cup.lmu.de)

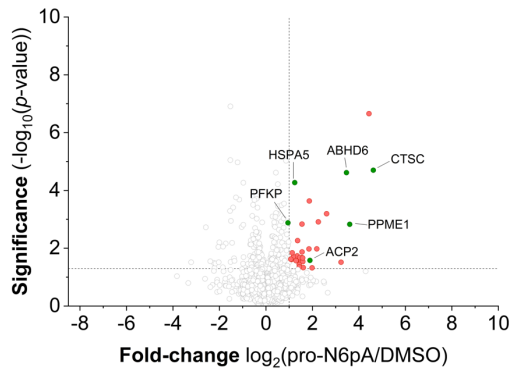
### Supporting Figures



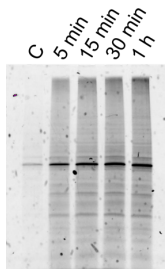
**Figure S1.** Full SDS-PAGE fluorescence analysis and loading controls. M (fluorescence protein molecular weight marker), line 1 (control cells not treated with the pro-N6pA probe and lysed in 1% NP-40, 0.2% SDS in 20mM Hepes), line 2 (1% NP-40 in PBS), line 3 (1% NP-40, 0.2% SDS in PBS), line 4 (0.5% Triton in PBS), line 5 (0.5% Triton, 0.2% SDS in PBS), line 6 (1% NP-40 in 20mM Hepes), line 7 (1% NP-40, 0.2% SDS in 20mM Hepes), line 8 (0.5% Triton in 20mM Hepes), line 9 (0.5% Triton, 0.2% SDS in 20mM Hepes), line 10 (8M urea in 0.1M Tris/HCl).



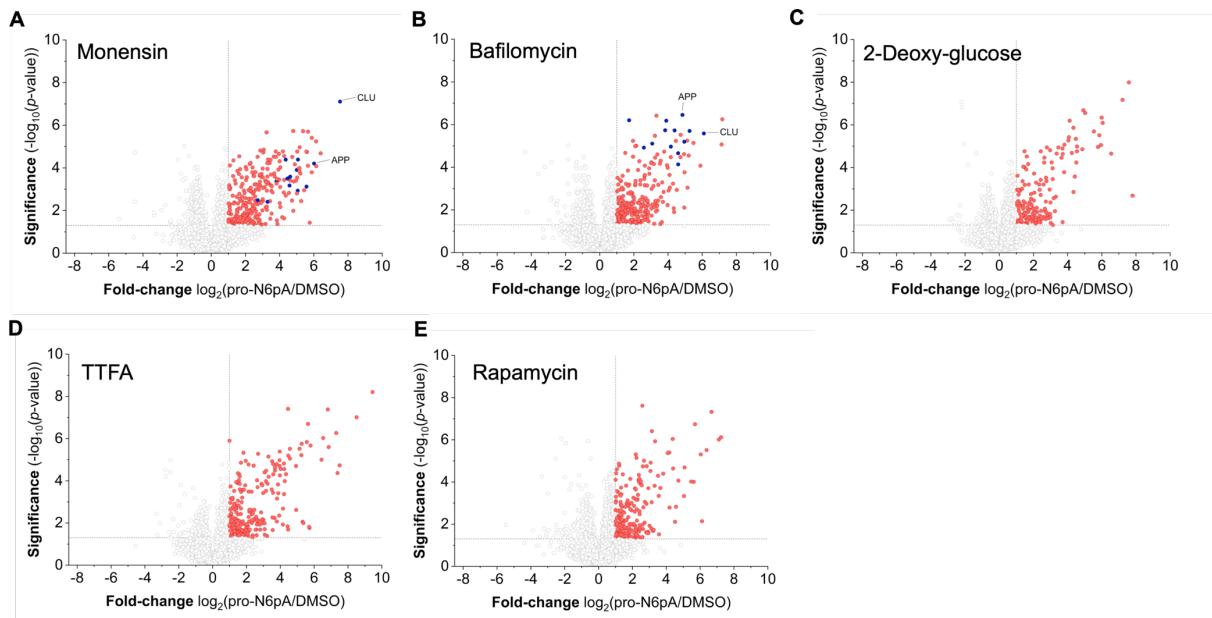
**Figure S2.** Volcano plot showing the enrichment of pro-N6pA modified proteins from HeLa cells using 400  $\mu$ g total protein. Development of the SP2E workflow. The proteins after the click reaction were directly loaded onto carboxylate magnetic beads.  $n = 4$ , cut-off lines at  $p$ -value  $> 0.05$  and 2-fold enrichment. Significantly enriched proteins are in red. The AMPylation marker proteins are labelled in green. Detailed values can be found in Supporting Table 1.



**Figure S3.** Volcano plot showing the enrichment of pro-N6pA modified proteins from HeLa cells using 400  $\mu$ g total protein. Development of the SP2E workflow. After the click reaction urea was added to the reaction mixture to improve the protein magnetic beads complexation.  $n = 4$ , cut-off lines at  $p$ -value  $>0.05$  and 2-fold enrichment. Significantly enriched proteins are in red. The AMPylation marker proteins are labelled in green. Detailed values can be found in Supporting Table 2.

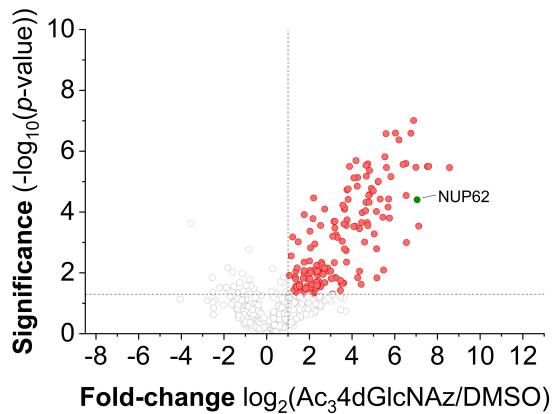


**Figure S4.** Biotin-streptavidin complex formation – incubation time optimization.

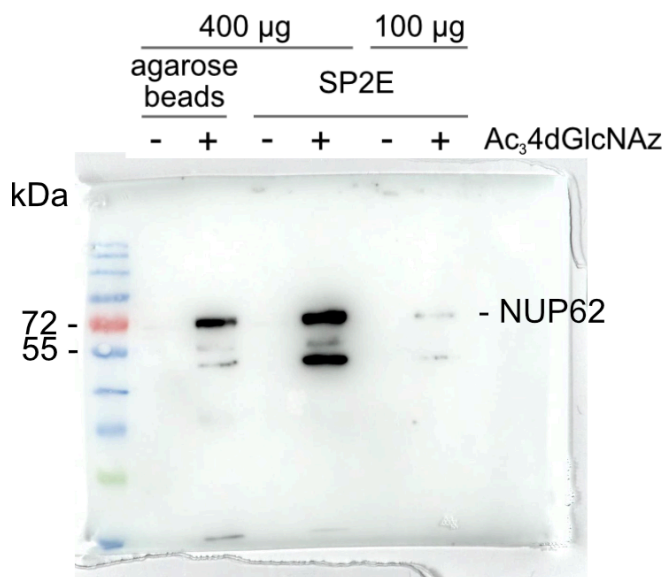


**Figure S5. A-E** Volcano plots showing the enrichment of pro-N6pA modified proteins in cell cultures treated with different inhibitors.  $n = 4$ , cut-off lines at  $p$ -value  $>0.05$  and 2-fold enrichment. Red dots are significantly enriched proteins. Blue dots represent proteins with a similar profile plot to APP, which were found significantly enriched only after monensin and bafilomycin treatment. This group includes following proteins: GPR56, FAT1, LAMA4,

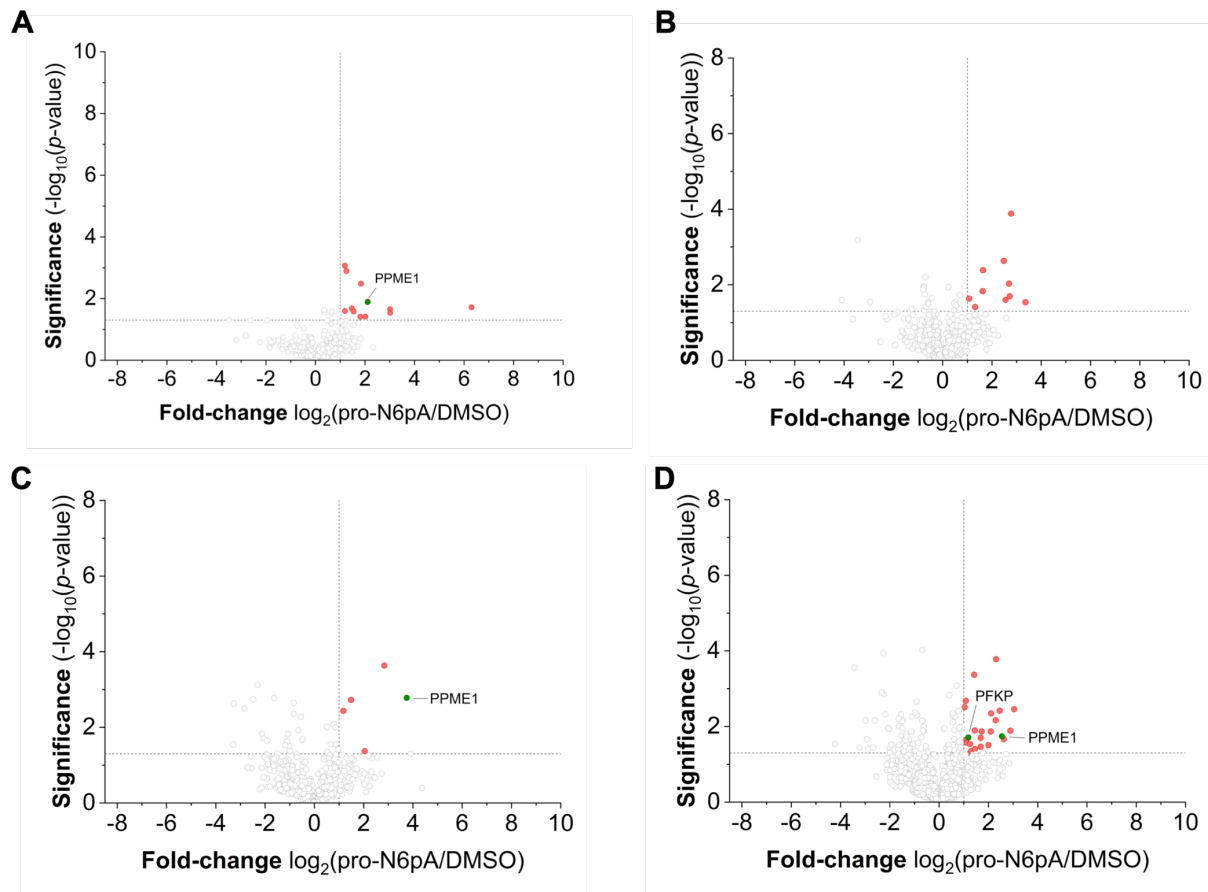
TGOLN2, RNF149, CRIM1, ITM2B, L1CAM, TMEM59, MCAM, LRP1 and CLU. Detailed values can be found in Supporting Table 4.



**Figure S6.** Volcano plot visualizing the enrichment of the O-GlcNAcylated proteins using the acetone precipitation for protein clean-up and avidin-coated agarose beads for enrichment;  $n = 4$ , cut-off lines at  $p$ -value  $>0.05$  and 2-fold enrichment. Red dots are significantly enriched proteins. Detailed values can be found in Supporting Table 10.



**Figure S7.** Enrichment of the modified NUP62 with the agarose-based and SP2E protocol. The SP2E enrichment was performed with 400  $\mu$ g and 100  $\mu$ g protein input. Enriched proteins were released from the streptavidin beads by loading buffer, separated by SDS-PAGE and analysed via Western-blot with the anti-NUP62 antibody.



**Figure S8.** Volcano plot showing the optimization of the 96-well plate format SP2E workflow starting with 100  $\mu$ g total protein. All samples have been analysed using the 150 min LC-MS/MS gradient. For comparison see Figure 6B. In the following is the description of the alternations to the protocol described in Methods and Materials. (A) The clean-up of the proteins was performed without ACN wash step and proteins were digested in ABC buffer. (B) The clean-up of the proteins was performed without ACN wash step. Protein digest was carried out in ABC buffer. Additional elution with 1% FA after proteins' digest was omitted to decrease the dilution of the peptides. (C) The amounts of carboxylate magnetic and streptavidin magnetic beads was halved. The clean-up of the proteins was performed without ACN wash step and proteins were digested in ABC buffer. (D) Additional elution with 1% FA after proteins' digest was omitted to decrease the dilution of the peptides. Detailed values can be found in Supporting Table 11 (Figure S9 A), Supporting Table 12 (Figure S9 B), Supporting Table 13 (Figure S9 C) and Supporting Table 14 (Figure S9 D).

## Materials and Methods

### Reagents and Tools Table

Reagent/Resource	Reference or Source	Identifier or Catalog Number
Experimental Models		
Human: SH-SY5Y	<a href="#">Biedler et al., 1973</a>	RRID: CVCL_0019

Human: HeLa	Puck and Marcus, 1955	RRID: CVCL_0030
<b>Recombinant DNA</b>		
/		
<b>Antibodies</b>		
Rabbit polyclonal anti-PLD3	Sigma-Aldrich	Cat# HPA012800; RRID: AB_1855330
Mouse monoclonal anti-NUP62	BD Biosciences	Cat# 610498; RRID: AB_397864
Goat polyclonal anti-rabbit IgG, HRP-linked	Sigma-Aldrich	Cat# A6667; RRID: AB_258307
Goat polyclonal anti-mouse IgG, HRP linked	Sigma-Aldrich	Cat# AP130P; RRID: AB_91266
<b>Oligonucleotides and other sequence-based reagents</b>		
/		
<b>Chemicals, Enzymes and other reagents</b>		
Acetonitrile (LC-MS grade)	Thermo Fisher Scientific	Cat# A955-212
Alanyl-Glutamine	Sigma-Aldrich	Cat# G8541
Ammoniumperoxodisulfat (APS)	Sigma-Aldrich	Cat# 09913
Biotin-PEG <sub>3</sub> -N <sub>3</sub>	Carbosynth	Cat# FA34890
BSA	AppliChem	Cat# A6588
Coomassie Blue R-250	Fluka	Cat# 27816
CuSO <sub>4</sub> x 5 H <sub>2</sub> O	Acros	Cat# 10627162
DBCO-PEG <sub>4</sub> -Biotin	Jena Bioscience	Cat# CLK-A105P4-10
ddH <sub>2</sub> O (LC-MS grade)	Honeywell	Cat# 15665350
DMSO	Sigma-Aldrich	Cat# D4540
Formic Acid (LC-MS grade)	Thermo Fisher Scientific	Cat# A117
Hepes	Carl Roth	Cat# HN77.5
NP40	Sigma-Aldrich	Cat# 74385
Methanol (LC-MS grade)	Thermo Fisher Scientific	Cat# A456
Powdered milk	AppliChem	Cat# A0830
SDS	AppliChem	Cat# A2572
TAMRA-N <sub>3</sub>	Baseclick	Cat# BCFA-008-1
TBTA	TCI	Cat# T2993
TCEP	Carbosynth	Cat# FT01756
TEAB (1 M)	Sigma-Aldrich	Cat# T7408
TEMED	Sigma-Aldrich	Cat# T9281
Tris-base	Thermo Fisher Scientific	Cat# 10724344
Trypan Blue	Thermo Fisher Scientific	Cat# 11538886
Urea	AppliChem	Cat# A1049
DPBS (1x)	Sigma-Aldrich	Cat# D8357



DMEM (1x)	Sigma-Aldrich	Cat# D6546
Immobilon® Western HRP Substrate	Merck Millipore	Cat# WBKLS0500
Color Prestained Protein Standard, Broad Range (10-250 kDa)	New England BioLabs	Cat# P7719S
Rotiphorese.Gel 30 (37, 5:1)	Carl Roth	Cat# 3029.1
TrypLE Express	Thermo Fisher Scientific	Cat# 12604013
Trypsin	Promega	Cat# V5113
Ammonium bicarbonate (ABC)	Fluka	Cat# 09830-100g
Streptavidin magnetic beads	New England BioLabs	Cat# S1420S
Carboxylate-coated magnetic beads (hydrophobic)	Cytiva	Cat# 65152105050250
Carboxylate-coated magnetic beads (hydrophilic)	Cytiva	Cat# 45152105050250
Monensin	ABCR	Cat# AB349208
Bafilomycin	ABCR	Cat# AB355587
Rapamycin	ABCR	Cat# AB352180
TTFA	ABCR	Cat# AB104115
2-Deoxy-D-glucose	ABCR	Cat# AB173119
Fetal bovine serum	Thermo Fisher Scientific	Cat# A3840001
Ethanol (EtOH)	Merck	Cat# 34852
Chloracetamide (CAA)	Sigma-Aldrich	Cat# C0267-100g
pro-N6pA probe	/	Ref. 17
GlcNAz probe	/	Ref. 47
Avidin-agarose beads	Sigma-Aldrich	Cat# A9207-5ML
5/6-TAMRA-Azide-Biotin	Jena Bioscience	Cat# CLK-1048-5
<b>Software</b>		
MaxQuant	Cox et al., 2014	<a href="https://www.maxquant.org/download_asset/maxquant/latest">https://www.maxquant.org/download_asset/maxquant/latest</a>
Perseus	<a href="#">Tyanova et al., 2016</a>	<a href="https://maxquant.net/download_asset/perseus/latest">https://maxquant.net/download_asset/perseus/latest</a>
Origin	N/A	<a href="https://www.originlab.com/">https://www.originlab.com/</a>
<b>Other</b>		
Pierce® BCA Protein Assay Kit	Thermo Fisher Scientific	Cat# 23225
Orbitrap Eclipse Tribrid Mass Spectrometer	Thermo Fisher Scientific	N/A
FAIMS Pro Duo Interface	Thermo Fisher Scientific	N/A

PicoTip™ Emitter, Silica Tip™	New Objectives	FS360-75-8-N-20-C15
ReproSil-Pur 120 C18-AQ	Dr. Maisch GmbH	Cat# R119.aq.0001
PEPMAP100 C18 5UM 0.3X5MM	Thermo Fisher Scientific	Cat# 160454

## Methods and Protocols

### *Inhibitor treatment*

In case the cells were treated with an inhibitor in addition to the probe, the final concentrations were as follows: 100 nM rapamycin, 100 nM bafilomycin, 2 µM monensin, 100 µM TTFa and 4 µM 2-Deoxy-D-glucose. For the monensin concentration dependency experiments the final monensin concentrations were 2 nM, 20 nM, 200 nM, 1 µM and 2 µM. After the simultaneous addition of probe and inhibitor, the cells were incubated 16 h at 37 °C before harvesting. For harvesting, the cells were washed twice with 2 mL DPBS, scraped into 1 mL DPBS and pelleted at 1000 rpm, 4 °C.

### *Cell lysis optimization*

For optimization of different lysis buffers, following compositions were used: Lysis buffer 1 (8 M urea in 100 mM Tris/HCl pH 8.5) Lysis buffer 2 (1% NP-40 in PBS), Lysis buffer 3 (1% NP-40, 0.2% SDS in PBS), Lysis buffer 4 (0.5% Triton in PBS), Lysis buffer 5 (0.5% Triton, 0.2% SDS in PBS), Lysis buffer 6 (1% NP-40 in 20 mM Hepes pH 7.5), Lysis buffer 7 (0.5% Triton, 0.2% SDS in 20 mM Hepes pH 7.5), Lysis buffer 8 (0.5% Triton in 20 mM Hepes pH 7.5).

### *In-gel fluorescence analysis*

The click reaction was performed with 200 µg protein in 100 µL of the respective lysis buffer. To each sample, 1 µL azide-TAMRA (10 mM in DMSO), 1 µL TCEP (100 mM in water) and 0.125 µL TBTA (83.5 mM in DMSO) were added. After gentle vortexing, the reaction was initiated by the addition of 2 µL CuSO<sub>4</sub> (50 mM in water) and incubated for 1.5 h (rt, 600 rpm). Subsequently, 400 µL of pre-cooled acetone (1:4 ratio) was added to each click reaction mixture and proteins were precipitated (1 h, -20 °C). Proteins were pelleted by centrifugation (15 min, 11,000 rpm, 4 °C) and reconstituted in 100 µL of 0.2% SDS in PBS by sonication (10 s, 20% intensity). 10 µL of each protein solution (20 µg) was boiled for 5 min at 95 °C with 2.4 µL 5x Laemmli buffer (10% (w/v) SDS, 50% (v/v) glycerol, 25% (v/v) β-mercaptoethanol, 0.5% (w/v) bromphenol blue, 315 mM Tris/HCl, pH 6.8), loaded and run on a 10% Tris/Glycine home-made SDS-PAGE gel (150 V, 1 h) with 1x running buffer (25 mM Tris, 0.192 M glycine, 0.1% (m/v) SDS). In-gel fluorescence was scanned with Amersham Imager (GE Healthcare) and further stained with the coomassie solution (0.25% Coomassie Blue R-250, 10% (v/v) acetic acid, 50% (v/v) methanol) as the loading control.

The time-dependent click reaction experiment was performed with 5, 15, 30, 60 and 90 min incubation time after CuSO<sub>4</sub> addition.

### ***SP2E workflow large scale optimization***

During the optimization process, we performed this workflow with the following changes. The first attempts were performed without a separate elution step of the proteins from the carboxylated beads (Figure 2C). In these attempts one was performed with the addition of 200  $\mu$ L 8 M urea after the click reaction (Figure 2D). In addition, TEAB was used instead of ABC buffer for the digest. Another attempt was to elute the proteins from the carboxylated beads before adding them to the streptavidin-coated magnetic beads (Figure 2E). Finally, the starting protein amount was reduced from 400  $\mu$ g to 100  $\mu$ g.

### ***SP2E workflow small scale optimization***

During the optimization process, we performed this workflow with the following changes. In the first attempt, the clean-up of the proteins was performed without ACN wash step and the proteins were digested in ABC buffer (Figure S9 A). In the second attempt, the clean-up of the proteins was performed without ACN wash step, the proteins were digested in ABC buffer and the additional elution step with 1% FA after the digest was omitted (Figure S9 B). In the third attempt, the amounts of carboxylate magnetic and streptavidin magnetic beads were halved, the clean-up of the was performed without ACN wash step and the proteins were digested in ABC buffer (Figure S9 C). In the fourth attempt, the additional elution step with 1% FA after the protein digest was omitted (Figure S9 D).

### ***Agarose-avidin beads enrichment workflow***

400  $\mu$ g protein of probe treated and control lysates were diluted with lysis buffer (20 mM Hepes, pH 7.5, 1% (v/v) NP40, 0.2% (w/v) SDS) to 200  $\mu$ L reaction volume. To each replicate, 3  $\mu$ L of 1M IAA in water was added and incubated for 30 min at 750 rpm, 25 °C. Next, 2  $\mu$ L of 2 mM DBCO-PEG-biotin reagent was added to initiate SPAAC reaction. The reaction mixtures were incubated at 25 °C, 750 rpm for 30 min. Then, 800  $\mu$ L acetone was added to each reaction mixture. After 1 h of incubation at -20°C, proteins were spun down 15 min at 11,000 rpm and 4°C. Supernatant was discarded and each pellet was resuspended in 1 mL methanol by sonicating 10 s at 20% intensity. Subsequently, each suspension was centrifuged 10 min at 11,000 rpm and 4°C. Each pellet was washed again with 1 mL methanol before it was dissolved in 1 mL 0.2% SDS in PBS by sonicating 10 s at 20% intensity. Next, avidin beads were thawed on ice before 50  $\mu$ L avidin bead suspension for each sample or control was washed 3 times with 1 mL 0.2% SDS in PBS to equilibrate the beads. After each addition of washing solution, the Eppendorf tube was carefully inverted 10 times, the suspension was

centrifuged 2 min at 2000 rpm at room temperature and the supernatant was discarded. Subsequently, each dissolved protein pellet was spun down at maximum speed for 2 min and room temperature before the supernatant was added to the equilibrated beads. After each avidin bead suspension was incubated 1 h under continuous mixing at room temperature, the beads of each sample or control were washed 3 times with 1 mL 0.2% SDS in PBS, 2 times with 1 mL 6 M urea in H<sub>2</sub>O and 3 times with 1 mL PBS. For in-gel analysis, the enriched proteins were eluted by the addition of 50  $\mu$ L SDS-loading buffer and incubation at 95 °C for 5 min. 20  $\mu$ L of the eluates were loaded on a SDS-gel and the western blot was performed as described in the above section. To perform the MS-measurement, the enriched proteins were digested with trypsin. For this, the washed avidin beads were resuspended in 200  $\mu$ L Xbuffer (7 M urea, 2 M thiourea in 20 mM HEPES pH 7.5). Then, 0.2  $\mu$ L 1 M DTT was added to reduce the disulfide bonds. Afterwards, each mixture was vortexed and incubated 45 min at room temperature shaking at 600 rpm. Next, 2  $\mu$ L 550 mM IAA was added to alkylate the cysteine residues. After each mixture was vortexed and incubated 30 min at room temperature shaking at 600 rpm in the dark, 0.8  $\mu$ L 1 M DTT was added to quench the alkylation. Subsequently, each mixture was vortexed and incubated for another 30 min at room temperature shaking at 600 rpm in the dark before 600  $\mu$ L 50 mM TEAB was added to increase the pH value to 8. Afterward, 1.5  $\mu$ L trypsin (0.5  $\mu$ g/ $\mu$ L in 50 mM acetic acid) was added to digest the enriched proteins. Finally, each mixture was vortexed and incubated overnight at 37°C shaking at 600 rpm. On the next day, 4  $\mu$ L FA was added to stop the digest. Subsequently, each mixture was vortexed and centrifuged 1 min at 2000 rpm before the supernatant was transferred into a new Eppendorf tube. After 50  $\mu$ L 0.1% FA was added to the avidin beads, each mixture was vortexed, the centrifugation step was repeated and the supernatant was again transferred to the new Eppendorf tube. Then, 50  $\mu$ L 0.1% FA was added once more to the avidin beads and each mixture was vortexed. Finally, each mixture was centrifuged 3 min at 13,000 rpm and the supernatant was transferred to the new Eppendorf tube as before. Desalting and the MS-measurement were performed as described in the large scale SP2E workflow.

### ***MS-measurement***

MS measurements were performed on an Orbitrap Eclipse Tribrid Mass Spectrometer (Thermo Fisher Scientific) coupled to an UltiMate 3000 Nano-HPLC (Thermo Fisher Scientific) via an EASY-Spray source (Thermo Fisher Scientific) and FAIMS interface (Thermo Fisher Scientific). First, peptides were loaded on an Acclaim PepMap 100  $\mu$ -precolumn cartridge (5  $\mu$ m, 100 Å, 300  $\mu$ m ID x 5 mm, Thermo Fisher Scientific). Then, peptides were separated at 40 °C on a PicoTip emitter (noncoated, 15 cm, 75  $\mu$ m ID, 8  $\mu$ m tip, New Objective) that was packed in house with Reprosil-Pur 120 C18-AQ material (1.9  $\mu$ m, 150 Å, Dr. A. Maisch GmbH). The long gradient was run from 4-35.2% acetonitrile supplemented with 0.1% formic acid

during a 150 min method (0-5 min 4%, 5-6 min to 7%, 7-105 min to 24.8%, 105-126 min to 35.2%, 126-140 min 80%, 140-150 min 4%) at a flow rate of 300 nL/min. The short gradient was run from 4-35.2% acetonitrile supplemented with 0.1% formic acid during a 60 min method (0-5 min 4%, 5-6 min to 7%, 7-36 min to 24.8%, 37-41 min to 35.2%, 42-46 min 80%, 47-60 min 4%) at a flow rate of 300 nL/min. FAIMS was performed with two alternating CVs including -50 V and -70 V. For measurements of chemical-proteomics samples, the Orbitrap Eclipse Tribrid Mass Spectrometer was operated in dd-MS<sup>2</sup> mode with following settings: Polarity: positive; MS<sup>1</sup> resolution: 240k; MS<sup>1</sup> AGC target: standard; MS<sup>1</sup> maximum injection time: 50 ms; MS<sup>1</sup> scan range: m/z 375-1500; MS<sup>2</sup> ion trap scan rate: rapid; MS<sup>2</sup> AGC target: standard; MS<sup>2</sup> maximum injection time: 35 ms; MS<sup>2</sup> cycle time: 1.7 s; MS<sup>2</sup> isolation window: m/z 1.2; HCD stepped normalised collision energy: 30%; intensity threshold: 1.0e4 counts; included charge states: 2-6; dynamic exclusion: 60 s.

### ***Quantification and statistical analysis***

MS raw files were analysed using MaxQuant software 2.0.1.0 with the Andromeda search engine. Searches were performed against the Uniprot database for Homo sapiens (taxon identifier: 9606, March 2020). At least two unique peptides were required for protein identification. False discovery rate determination was carried out using a decoy database and thresholds were set to 1% FDR both at peptide-spectrum match and at protein levels. LFQ quantification was used as described for each sample. For calculation of the large scale samples, carbamidomethylation was set as a fixed modification and methionine oxidation as well as N-terminal acetylation as a variable modification. In contrast, the small scale samples of the 96-well plate were calculated without carbamidomethylation as a fixed modification.

Statistical analysis of the MaxQuant result table proteinGroups.txt was done with Perseus 1.6.10.43. First, LFQ intensities were log<sub>2</sub>-transformed. Afterwards, potential contaminants as well as reverse peptides were removed. Then, the rows were divided into two groups - DMSO (control) and probe treated sample (sample). Subsequently, the groups were filtered for at least three valid values out of four rows in at least one group and the missing values were replaced from normal distribution. The -log<sub>10</sub>(p-values) were obtained by a two-sided one sample Student's t-test over replicates with the initial significance level of p = 0.05 adjustment by the multiple testing correction method of Benjamini and Hochberg (FDR = 0.05) using the volcano plot function.

### ***Western blot analysis***

For each Western blot analysis, 20 µg cell lysate was used. In order to denature proteins, 4 µL 5× Laemmli buffer (10% (w/v) SDS, 50% (v/v) glycerol, 25% (v/v) β-mercaptoethanol, 0.5% (w/v) bromphenol blue, 315 mM Tris/HCl, pH 6.8) was added to 16 µL lysate solution and the

samples were boiled 5 min at 95 °C. Afterwards, 20 µL of each sample was loaded onto a 7.5, 10 or 12.5% SDS gel and proteins were separated according to their size by SDS-PAGE. Then, the separated proteins were transferred onto a membrane using a blotting sandwich moistened by blotting buffer (48 mM Tris, 39 mM glycine, 0.0375% (m/v) SDS, 20% (v/v) methanol), which was composed of one extra thick blot paper, the PVDF transfer membrane, the SDS-PAGE gel and again one extra thick blot paper. Before the protein transfer was carried out 45 min at 25 V using a Semi Dry Blotter (Bio-Rad), the transfer membrane was pre-incubated 5 min in methanol. In order to block non-specific binding sites, the membrane was incubated 60 min in blocking solution (0.5 g milk powder in 10 mL PBST (PBS +0.5% Tween)). Subsequently, 10 µL primary antibody with specificity for the protein of interest was added and the mixture was incubated 1 h at room temperature. The membrane was washed 3 times for 10 min with PBST before 1 µL of the secondary HRP antibody in 10 mL blocking solution was added. After 1 h of incubation at room temperature, the membrane was washed again 3 times for 10 min with PBST. Then, 400 µL ECL Substrate and 400 µL peroxide solution were mixed and added to the membrane to stain the Western blot. Finally, images of the Western blot were taken by developing machine Amersham Imager 680 (GE Healthcare).

### ***SP2E in-gel analysis***

SP2E large scale protocol was performed as described above, but instead of digesting the proteins after the enrichment on streptavidin beads, they were released by addition of 50 µL 1x Laemmli buffer and 5 min incubation at 95 °C. Subsequently, 20 µL of the eluates were loaded on a SDS-gel and the western blot was performed as described in the above section.

### ***Time optimization of biotin-streptavidin complex formation***

The click reaction was performed with 200 µg protein in 100 µL lysis buffer (20 mM Hepes, pH 7.5, 1% (v/v) NP40, 0.2% (w/v) SDS). To each sample, 1 µL 5/6-TAMRA-Azide-Biotin (10 mM in DMSO), 1 µL TCEP (100 mM in water) and 0.125 µL TBTA (83.5 mM in DMSO) were added. After gentle vortexing, the reaction was initiated by the addition of 2 µL CuSO<sub>4</sub> (50 mM in water) and incubated for 1.5 h (rt, 600 rpm). Subsequently, 100 µL of 8 M urea was added to each replicate. 50 µL of mixed hydrophobic and hydrophilic carboxylate-coated magnetic beads (1:1) were washed thrice with 500 µL water. Click reaction mixture was directly transferred onto the equilibrated carboxylate-coated magnetic beads, resuspended and 300 µL ethanol was added. After resuspending the beads via vortexing, suspension was incubated for 5 min at rt and 950 rpm. The beads were washed thrice with 500 µL of 80% ethanol in water and the proteins were separately eluted by the addition of 0.5 mL of 0.2% SDS in PBS. The eluates were combined, added to 50 µL equilibrated streptavidin-coated magnetic beads (3 times pre-washed with 500 µL 0.2% SDS in PBS) and were incubated for

5, 15, 30 or 60 min. Afterwards the beads were washed thrice with 500  $\mu$ L 0.1% NP40 in PBS, twice with 500  $\mu$ L 6 M urea and twice with 500  $\mu$ L water. Then, the proteins were released from the beads by addition of 50  $\mu$ L 1x Laemmli buffer and 5 min incubation at 95 °C. Finally, 20  $\mu$ L of the eluates were loaded on a SDS-gel and in-gel fluorescence was scanned with Amersham Imager (GE Healthcare).

### **8.3 Supplementary information: Clickable report tags for identification of modified peptides by mass spectrometry**

Makarov, D., Telek, A., Becker, T., von Wrisberg, M.K., Schneider, S., and Kielkowski, P. (2022). Clickable report tags for identification of modified peptides by mass spectrometry. *Journal of Mass Spectrometry* 57. 10.1002/jms.4812.



## Supplementary Information.

### Clickable report tags for identification of modified peptides by mass spectrometry.

Dmytro Makarov<sup>1\*</sup>, András Telek<sup>1,2\*</sup>, Tobias Becker<sup>1</sup>, Marie-Kristin von Wrisberg<sup>1</sup>, Sabine Schneider<sup>1</sup>, Pavel Kielkowski<sup>1\*</sup>

<sup>1</sup> LMU Munich, Department of Chemistry, Butenandtstr. 5-13, 81377, Munich, Germany.

<sup>2</sup> current address: Budapest University of Technology and Economics, Department of Applied Biotechnology, Szent Gellért sq. 4, 1111, Budapest, Hungary.

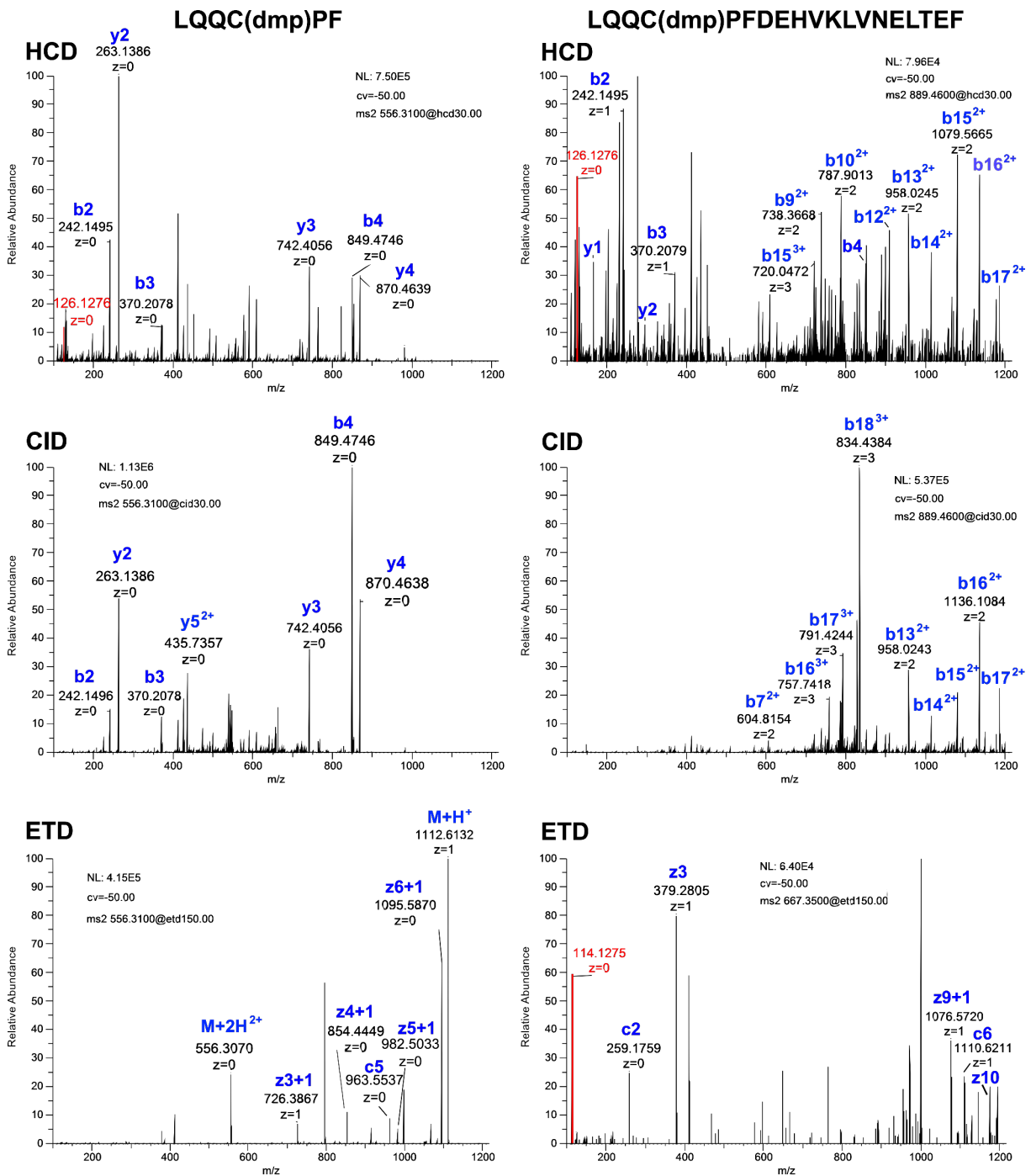
\* These authors contributed equally.

\* Correspondence to [pavel.kielkowski@cup.lmu.de](mailto:pavel.kielkowski@cup.lmu.de)

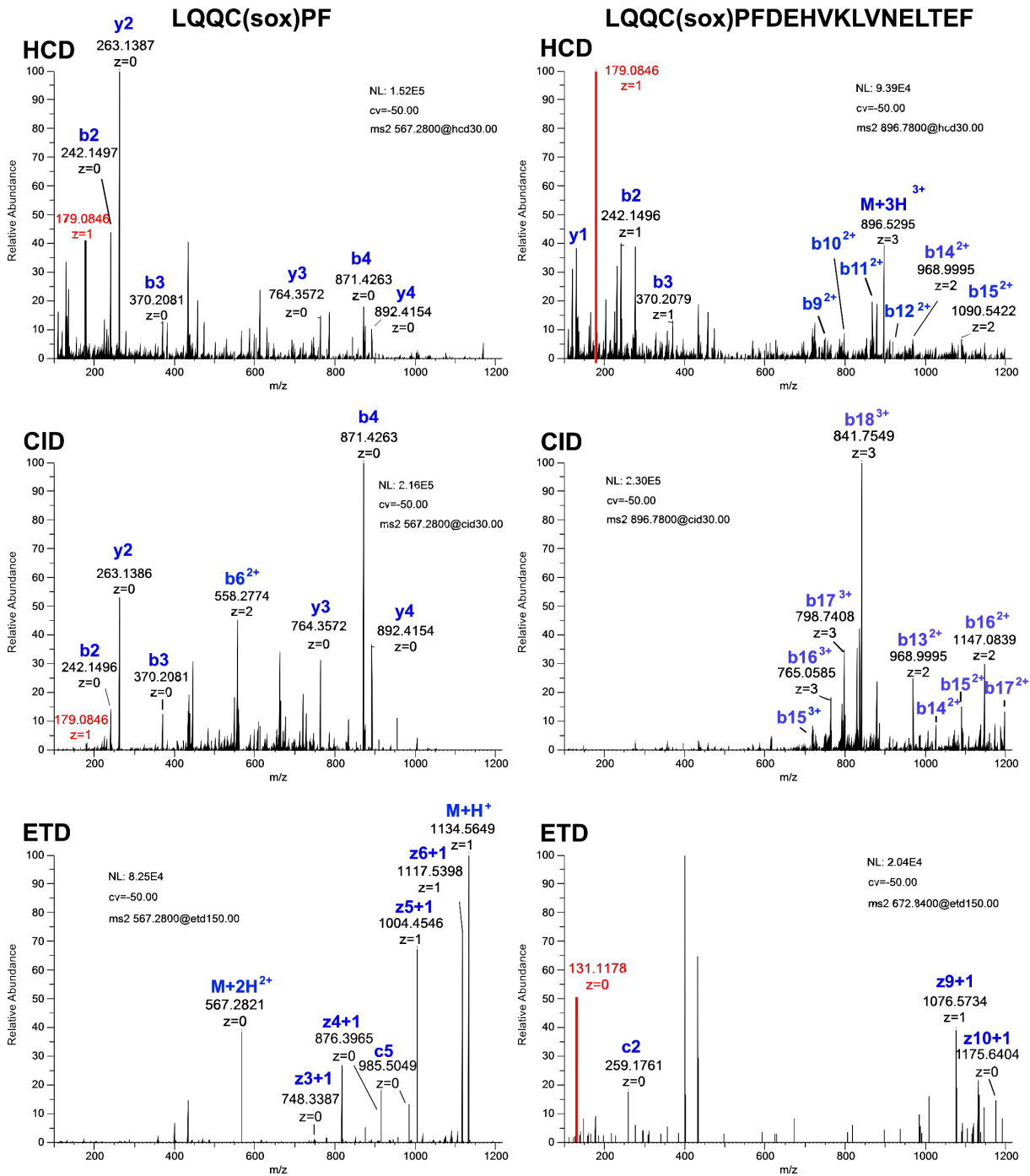
### Table of content:

<b>Supplementary Figures and Tables</b>	<b>2</b>
<b>Supplementary Methods</b>	<b>6</b>
Synthesis of small compounds	6
Biochemistry and Proteomics	11
<b>NMR and MS Spectra of Small Compounds</b>	<b>15</b>

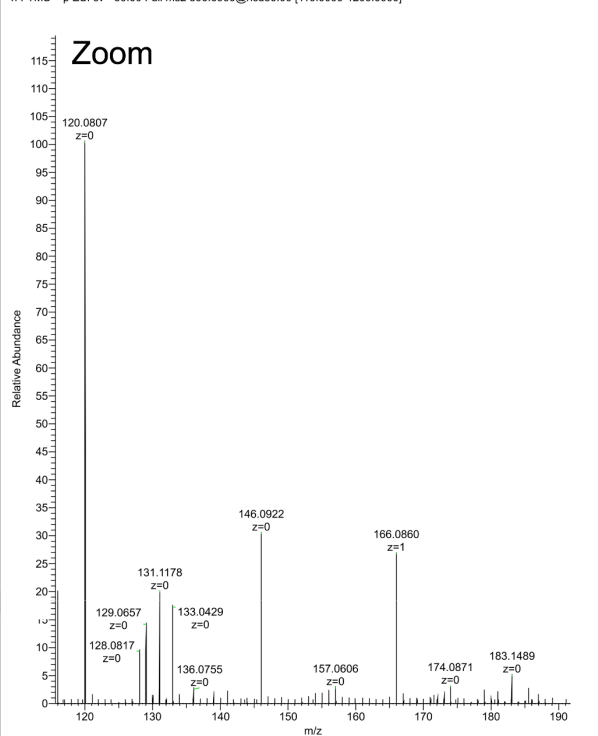
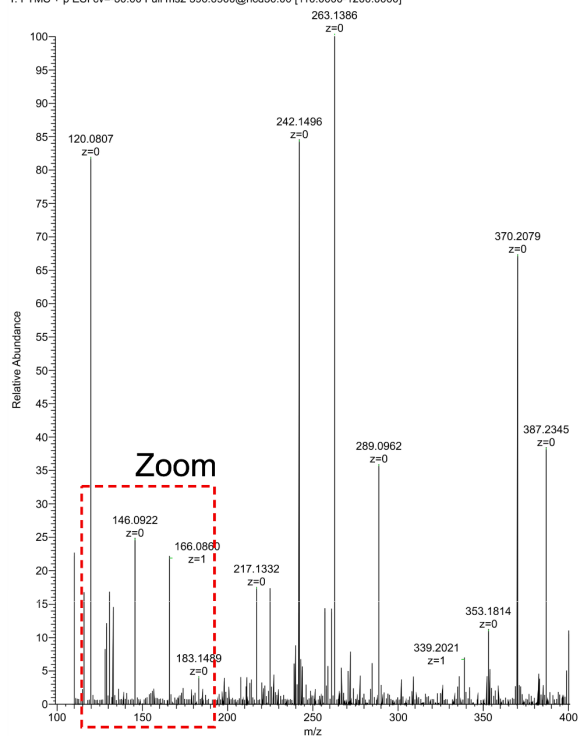
## Supplementary Figures



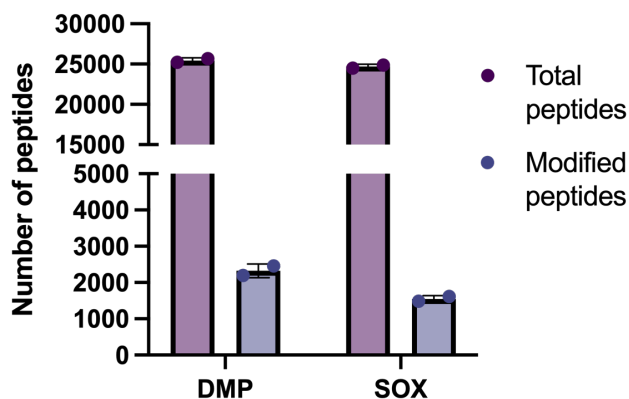
**Figure S1:** Fragmentation and reporter ion release from BSA peptide resulting from chymotrypsin digest labelled with DMP-tag.



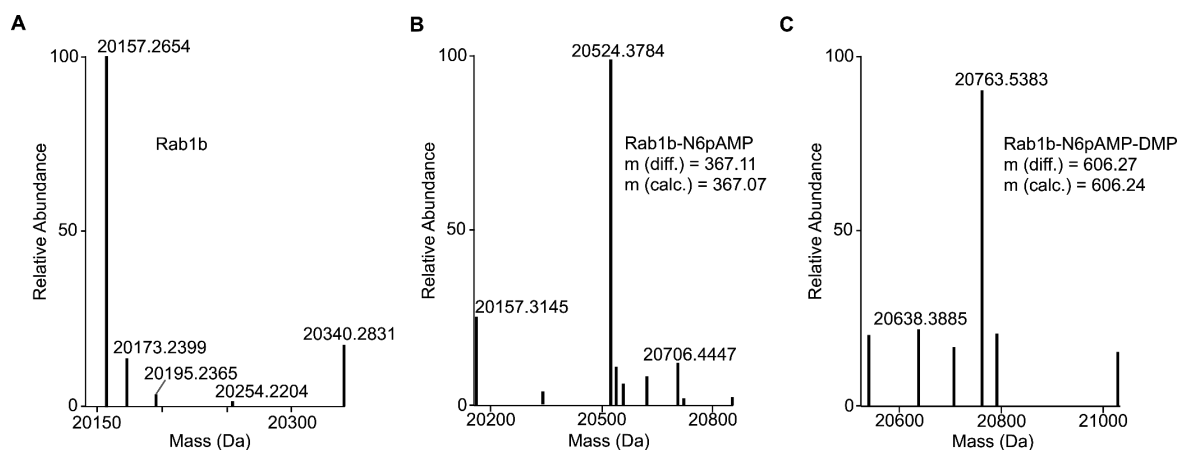
**Figure S2:** Fragmentation and reporter ion release from BSA peptide resulting from chymotrypsin digest labelled with SOX-tag.



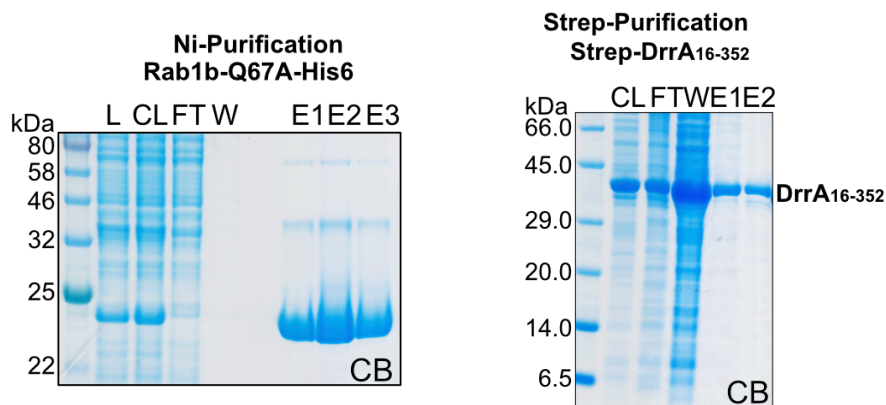
**Figure S3:** Control fragmentation experiment of BSA peptide showing absence of detectable peaks, which would overlap with DMP- or SOX-reporter ions.



**Figure S4:** Total peptides and modified peptides found by MaxQuant in the DMP- and SOX-tag labelled cysteines using low-resolution MS2 acquired in ion trap.

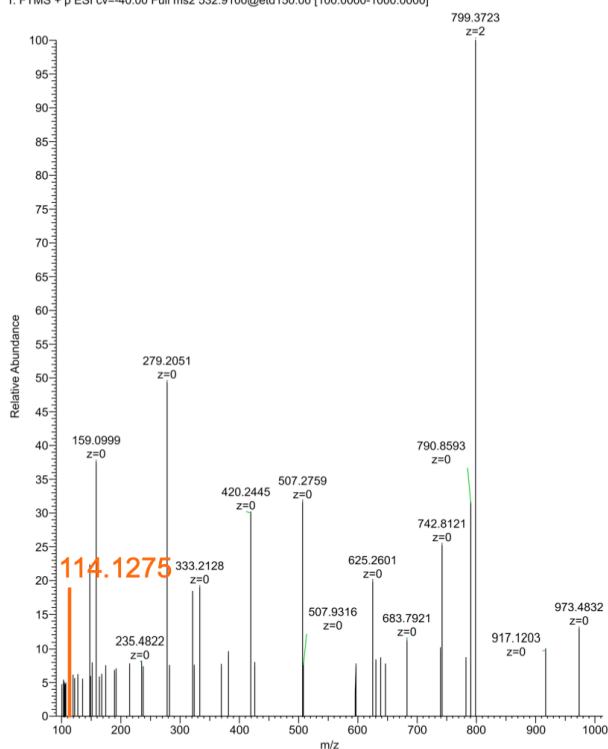


**Figure S5:** Deconvoluted intact protein MS spectra of **A)** Rab1b, **B)** N6pAMPylated Rab1b and **C)** N6pAMPylated Rab1b with DMP-tag.



**Figure S6:** Rab1b and DrrA purification.

20210225\_AT15\_Rab1b\_AMPTMT\_digested-40cv\_etd150 #1 RT: 0.00 AV: 1 NL: 7.14E3  
 T: FTMS + p ESI cv=-40.00 Full ms2 532.9100@etd150.00 [100.0000-1000.0000]



**Figure S7:** ETD fragmentation of AMPylated TITSSYYR peptide decorated with the DMP-tag yielding the 114.1275 reporter ion.

**Table S3:** Andromeda scores of modified Rab1b peptides with and without including neutral loss containing peptide fragments in the search. When these fragments were included, the modification was falsely assigned to the T3 of the peptide in case of the highest scoring PSMs.

Peptide	Score without neutral losses	Score with neutral losses
TITSSY(N6pAMP)YR	92.19	101.65
TITSSY(AMP-DMP)YR	75.18	103.55

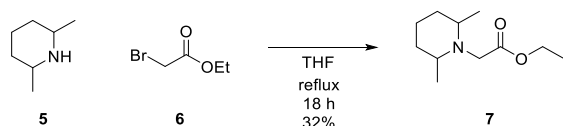
## Supplementary Methods

### Synthesis of small compounds

#### General Information

Unless noted otherwise, all reactions were performed using oven-dried glassware under an atmosphere of argon. Molsieve-dried solvents were used from *Sigma Aldrich* and chemicals were bought from *Sigma Aldrich*, *TCl*, *abcr*, *Roth* and *Carbosynth*. Reaction controls were performed using TLC-Plates from *Merck* (Merck 60 F<sup>254</sup>), flash column chromatography purifications were performed on *Merck* Geduran Si 60 (40–63 $\mu$ M). Visualization of the TLC plates was achieved through UV-absorption or through staining with *Potassium permanganate stain*. NMR spectra were recorded in deuterated solvents on *Varian VXR400S*, *Varian Inova 400*, *Bruker AMX 600*, *Bruker Ascend 400* or *Bruker Avance III HD*. HR-ESI-MS spectra were obtained from a *Thermo Finnigan LTQ FT-ICR*. For MS-experiments, only highest grade solvents and reagents were used. The mass spectrometry proteomics data are available upon request.

#### Synthesis of *N*-(2-azidoethyl)-2-(2,6-dimethylpiperidin-1-yl)-acetamide:



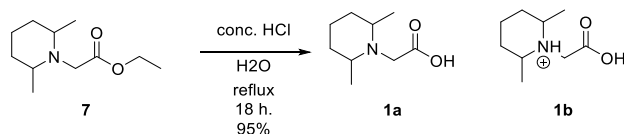
#### Ethyl 2-(2,6-dimethylpiperidin-1-yl)acetate (**7**) [33]

2,6-Dimethylpiperidine (**5**) (1.19 mL, 8.83 mmol) and ethyl 2-bromoacetate (**6**) (0.98 mL, 8.83 mmol) were dissolved in THF (15 mL) and the RM was stirred overnight under reflux. The solvent was evaporated and the residue was dissolved in H<sub>2</sub>O (20 mL). The mixture was adjusted to pH 12 with 1M NaOH solution so a suspension was formed. The suspension was then extracted with EtOAc (3  $\times$  15 mL). The combined organic phase was dried over anhydrous Mg<sub>2</sub>SO<sub>4</sub> and evaporated *in vacuo*. The residue was dissolved in hexane (15 mL) and solids were formed. The solids were filtered off and the solvent was evaporated. The residue was purified by column chromatography using 50% EtOAc in hexane. Product **7** was obtained as a yellowish liquid (0.56 g, 32%).

<sup>1</sup>H NMR (400 MHz, Chloroform-*d*)  $\delta$  4.14 (q,  $J$  = 7.1 Hz, 2H), 3.56 (s, 2H), 2.86 – 2.78 (m, 2H), 1.70 – 1.62 (m, 1H), 1.60 – 1.53 (m, 2H), 1.44 – 1.32 (m, 1H), 1.26 (t,  $J$  = 7.1 Hz, 5H), 1.10 (d,  $J$  = 6.3 Hz, 6H).  
<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>)  $\delta$  207.20, 68.12, 54.79, 31.09, 30.48, 25.74, 23.01, 19.42.

MS(ESI<sup>+</sup>):  $m/z$  (%): 200.25 (100) [M+H]<sup>+</sup>.

#### 2-(2,6-Dimethylpiperidin-1-yl)acetic acid (**1**)



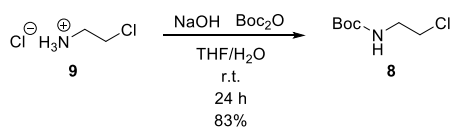
Ethyl 2-(2,6-dimethylpiperidin-1-yl)acetate **7** (0.55 g, 2.76 mmol) was dissolved in H<sub>2</sub>O (5.50 mL) and concentrated HCl (0.55 mL) was added to the RM. The mixture was heated overnight to reflux. After being cooled to ambient temperature, the RM was extracted with Et<sub>2</sub>O (3  $\times$  10 mL). The combined organic phase was dried over anhydrous Mg<sub>2</sub>SO<sub>4</sub> and concentrated *in vacuo*. Products **1a** and **1b** were obtained as a mixture of protonated and non-protonated molecules in a 1:2 ratio and as an off-white solid (0.45 g, 95%).

Compound **1a** and **1b**. <sup>1</sup>H NMR (400 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  4.22 (s, 1H, [-CH<sub>2</sub>-CO (**1a**)]), 4.00 (s, 2H, [-CH<sub>2</sub>-CO (**1b**)]), 3.63 – 3.52 (m, 3H, [-CH-CH<sub>3</sub> (**1a** + **1b**)]), 1.99 – 1.51 (m, 9H, [-CH<sub>2</sub>-CH<sub>2</sub>-CH<sub>2</sub>-(**1a** + **1b**)]), 1.40 (d,  $J$  = 6.4 Hz, 3H, [-CH<sub>3</sub>-CH (**1a**)]), 1.28 (d,  $J$  = 6.6 Hz, 6H, [-CH<sub>3</sub>-CH-(**1b**)]).

<sup>13</sup>C NMR (101 MHz, MeOD)  $\delta$  170.29, 63.77, 63.07, 42.72, 32.85, 26.52, 23.48, 23.19, 18.59, 17.17.

MS(ESI<sup>+</sup>):  $m/z$  (%): 172.22 (100) [M+H]<sup>+</sup>.

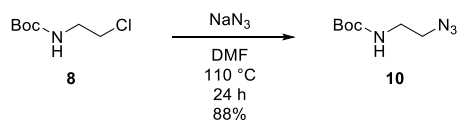
### **tert-Butyl (2-chloroethyl)carbamate (8)[34]**



2-Chloroethan-1-amine hydrochloride (**9**) (2.00 g, 17.24 mmol) was suspended in THF (25 mL) and a solution of NaOH (0.83 g, 20.69 mmol) was added to H<sub>2</sub>O (20 mL). The RM was cooled on an ice bath before a solution of Boc<sub>2</sub>O (4.52 g, 20.69 mmol) in THF (30 mL) was added dropwise over 40 min. The RM was stirred for 24 h at r.t. THF was evaporated *in vacuo*. The aqueous phase was diluted with H<sub>2</sub>O (30 mL) and extracted with EtOAc (3 × 25 mL). The combined organic layer was dried over anhydrous Mg<sub>2</sub>SO<sub>4</sub> and concentrated *in vacuo*. Purification of the product was carried out by column chromatography using a mixture of EtOAc in hexane (0 – 25%). Product **8** was obtained as a transparent viscous liquid (2.58 g, 83%). NMR data were in agreement with the literature.[34]

<sup>1</sup>H NMR (400 MHz, Chloroform-*d*) δ 4.99 (s, 1H), 3.58 (t, *J* = 5.7 Hz, 2H), 3.45 (q, *J* = 5.8 Hz, 2H), 1.43 (s, 9H).

### **tert-Butyl (2-azidoethyl)carbamate (10)**

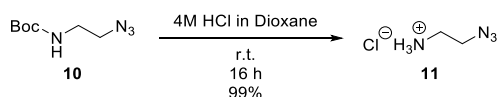


Carbamate **8** (2.58 g, 14.36 mmol) was dissolved in DMF (45 mL). NaN<sub>3</sub> (3.73 g, 57.43 mmol) was added to the mixture. The resulting suspension was stirred for 24 h at 115 °C. The solvent was removed *in vacuo* and the residue was dissolved in H<sub>2</sub>O (100 mL). The aqueous phase was extracted with EtOAc (3 × 50 mL). The combined organic phase was dried over anhydrous Mg<sub>2</sub>SO<sub>4</sub> and concentrated *in vacuo*. Product **10** was obtained as a yellow viscous liquid (2.35 g, 88%). NMR data were in agreement with the literature.[35]

<sup>1</sup>H NMR (400 MHz, Chloroform-*d*) δ 4.86 (s, 1H), 3.40 (t, *J* = 5.6 Hz, 2H), 3.29 (q, *J* = 5.8 Hz, 2H), 1.44 (s, 9H).

MS(ESI<sup>+</sup>): *m/z* (%): 187.22 (30) [M+H]<sup>+</sup>, 373.29 (25) [2M+H]<sup>+</sup>.

### **2-Azidoethan-1-amine hydrochloride (11)**

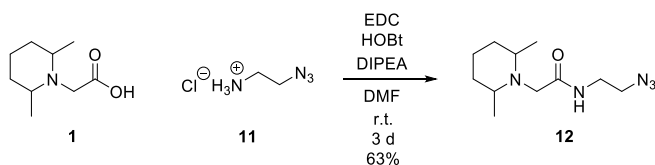


Carbamate **10** (1.15 g, 6.18 mmol) was added to the solution of 4M HCl in Dioxane (4.20 mL). The resulting mixture was stirred overnight at r.t. A suspension was formed and the solvent was removed *in vacuo*. Product **11** was obtained as a yellow solid after drying it under a high vacuum (0.75 g, 99%). NMR data were in agreement with the literature.[36]

<sup>1</sup>H NMR (400 MHz, Deuterium Oxide) δ 3.72 – 3.65 (m, 2H), 3.17 – 3.11 (m, 2H).

MS(ESI<sup>+</sup>): *m/z* (%): 128.24 (100) [M+ACN+H]<sup>+</sup>.

### **N-(2-azidoethyl)-2-(2,6-dimethylpiperidin-1-yl)acetamide (12)**



Acid **1** (1.13 g, 0.76 mmol), HOBt (0.18 g, 1.14 mmol), EDC.HCl (0.22 g, 1.14 mmol) and DIPEA (0.60 mL, 3.42 mmol) were dissolved in anhydrous DMF (15 mL) under inert atmosphere. The RM was stirred for 3 h at r.t. before the solution changed color. The mixture of azide **11** (0.14 g, 1.14 mmol) with

DIPEA (0.20 mL, 1.14 mmol) in anhydrous DMF (5 mL) was added to the RM. The RM was stirred under an inert atmosphere for 72 h at r.t. DMF was evaporated *in vacuo* and the residue was dissolved in DCM. The mixture was washed with H<sub>2</sub>O (2 × 15 mL) and brine (1 × 15 mL). The organic phase was dried over anhydrous Mg<sub>2</sub>SO<sub>4</sub> and the solvent was removed *in vacuo*. The product was purified by column chromatography using 5% MeOH in DCM as eluent. Product **12** was obtained as a yellow liquid (0.12 g, 63%).

<sup>1</sup>H NMR (400 MHz, Chloroform-*d*) δ 7.87 (s, 1H), 3.44 (s, 4H), 3.08 (s, 2H), 2.49 – 2.41 (m, 2H), 1.61 – 1.54 (m, 2H), 1.42 – 1.17 (m, 4H), 1.02 (d, *J* = 6.2 Hz, 6H).

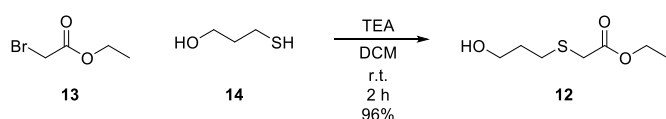
<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 174.02, 58.95, 51.38, 50.92, 38.35, 24.33, 21.50.

MS(ESI<sup>+</sup>): *m/z* (%): 240.25 (100) [M+H]<sup>+</sup>.

HR-MS(ESI<sup>+</sup>): calculated for C<sub>11</sub>H<sub>22</sub>N<sub>5</sub>O<sup>+</sup> 240.18244, found 240.18196.

### Synthesis of 2-((3-azidopropyl)sulfinyl)-*N*-(2-(dimethylamino)ethyl)acetamide:

#### Ethyl 2-((3-hydroxypropyl)thio)acetate (**12**) [37]



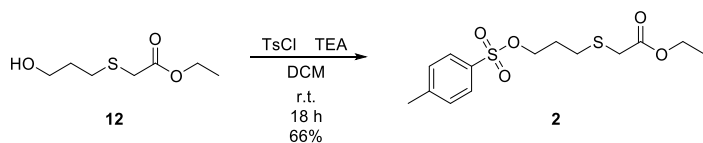
In a round bottom flask, ethyl 2-bromoacetate (**13**) (2.57 mL, 23.22 mmol) and triethylamine (4.85 mL, 34.83 mmol) were dissolved in DCM (20 mL). 3-Mercaptopropan-1-ol (**14**) (2.00 mL, 23.22 mmol) was added dropwise to the RM and the solution was stirred for 2 h at r.t. After completion of the reaction, the solution was washed with H<sub>2</sub>O (1 × 15 mL), 1M HCl (1 × 15 mL) and with a saturated solution of NaHCO<sub>3</sub> (1 × 15 mL). The organic phase was collected and dried over anhydrous Mg<sub>2</sub>SO<sub>4</sub>, the solvent was evaporated *in vacuo*. Product **12** was obtained without purification as a yellow viscous liquid (3.98 g, 96%).

<sup>1</sup>H NMR (400 MHz, Chloroform-*d*) δ 4.18 (q, *J* = 7.2 Hz, 2H), 3.78 – 3.71 (m, 2H), 3.22 (s, 2H), 2.75 (t, *J* = 7.1 Hz, 2H), 1.89 – 1.81 (m, 2H), 1.28 (t, *J* = 7.1 Hz, 3H).

<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 207.23, 73.18, 72.42, 26.51, 21.52, 20.92, 10.39.

MS (ESI<sup>+</sup>): *m/z* (%): 179.24 (100) [M+H]<sup>+</sup>.

#### Ethyl 2-((3-(tosyloxy)propyl)thio) acetate (**2**) [ TSAI, C.-S.; HUNG, T.-C.; CHUANG Hong-Yang. ANTIBODY-DRUG CONJUGATES. WO2018090045, 2018.]



Ethyl thioacetate **12** (3.95 g, 22.16 mmol), *p*-Toluene sulfonyl chloride (6.34 g, 33.24 mmol) and TEA (6.17 mL, 44.32 mmol) were mixed in DCM (20 mL). The resulting RM was stirred overnight at r.t. After that, the mixture was washed with concentrated NaHCO<sub>3</sub> (1 × 20 mL), H<sub>2</sub>O (2 × 15 mL) and brine (1 × 15 mL). The solvent was evaporated *in vacuo*. The product was purified by column chromatography using 50% EtOAc in hexane. Product **2** was obtained as a yellow liquid (4.88 g, 66%).

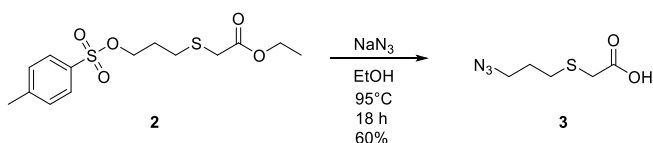
<sup>1</sup>H NMR (400 MHz, Chloroform-*d*) δ 7.79 (d, *J* = 8.3 Hz, 2H), 7.35 (d, *J* = 7.7 Hz, 2H), 4.20 – 4.10 (m, 4H), 3.14 (s, 2H), 2.66 (t, *J* = 7.1 Hz, 2H), 2.45 (s, 3H), 1.95 (p, *J* = 6.6 Hz, 2H), 1.27 (t, *J* = 7.2 Hz, 3H).

<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 175.41, 145.08, 132.80, 130.03, 127.98, 68.61, 61.58, 33.36, 28.56, 28.24, 21.74, 14.21.

MS (ESI<sup>+</sup>): *m/z* (%): 228.17 (100) [M+3ACN+2H]<sup>2+</sup>.

#### 2-((3-azidopropyl)thio)acetic acid (**3**)





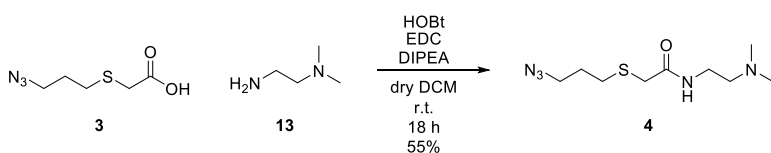
The suspension of tosylated ethyl thioacetate **2** (4.80 g, 14.44 mmol) and NaN<sub>3</sub> (2.82 g, 43.43 mmol) in EtOH (20 mL) was stirred overnight at 95°C. After the reaction was complete, the solids were filtered off and EtOH was evaporated *in vacuo*. The residue was diluted with H<sub>2</sub>O (10 mL) and the mixture was acidified with 1M HCl to pH 2. The product was extracted with EtOAc (3 × 15 mL). The organic phase was subsequently washed with H<sub>2</sub>O (3 × 10 mL), brine (1 × 10 mL) and dried over anhydrous Mg<sub>2</sub>SO<sub>4</sub>. The solvent was evaporated *in vacuo*. The product was purified by column chromatography using 10% EtOAc in hexane. Product **3** was obtained as a yellow liquid (1.51 g, 60%).

<sup>1</sup>H NMR (400 MHz, Chloroform-*d*) δ 3.43 (t, *J* = 6.6 Hz, 2H), 3.26 (s, 2H), 2.76 (t, *J* = 7.1 Hz, 2H), 1.89 (p, *J* = 6.8 Hz, 2H).

<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 175.16, 49.87, 33.34, 29.74, 28.11.

MS (ESI<sup>+</sup>): *m/z* (%): 176.24 (100) [M+H]<sup>+</sup>.

### 2-((3-Azidopropyl)thio)-*N*-(2-(dimethylamino)ethyl)acetamide (**4**)



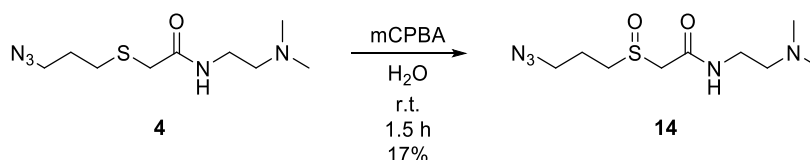
HOBT (1.74 g, 12.84 mmol) and EDC hydrochloride (2.46 g, 12.84 mmol) were dissolved in dry DCM (5 mL) under an inert atmosphere and the mixture was cooled down to 0°C on an ice bath. A mixture of azide **3** (1.50 g, 8.56 mmol) in dry DCM (7 mL) was added dropwise to the RM. After 15 min of stirring, a solution of *N*<sup>1</sup>,*N*<sup>1</sup>-dimethylethane-1,2-diamine (**13**) (1.49 mL, 13.70 mmol) and DIPEA (5.97 mL, 34.24 mmol) was added dropwise to the mixture. The RM was removed from the ice bath and stirred overnight at r.t. The RM was washed with conc. NaHCO<sub>3</sub> solution (1 × 10 mL) and H<sub>2</sub>O (2 × 10 mL). The organic phase was dried over anhydrous Mg<sub>2</sub>SO<sub>4</sub> and the solvent was evaporated *in vacuo*. The purification of the product was made by column chromatography using 70% EtOAc in hexane as a first eluent to elute polar impurities and using 10% MeOH in DCM with the addition of 0.5% of TEA as a second to elute the desired product. Product **4** was obtained as a transparent liquid (1.15 g, 55%).

<sup>1</sup>H NMR (400 MHz, Chloroform-*d*) δ 7.14 (s, 1H), 3.41 (t, *J* = 6.6 Hz, 2H), 3.38 – 3.33 (m, 2H), 3.21 (s, 2H), 2.64 (t, *J* = 7.1 Hz, 2H), 2.43 (t, *J* = 6.0 Hz, 2H), 2.24 (s, 6H), 1.87 (p, *J* = 6.8 Hz, 2H).

<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 168.82, 57.87, 50.04, 45.25, 37.16, 36.16, 31.08, 29.98, 28.41.

MS (ESI<sup>+</sup>): *m/z* (%): 246.21 (100) [M+H]<sup>+</sup>.

### 2-((3-Azidopropyl)sulfinyl)-*N*-(2-(dimethylamino)ethyl)acetamide (**14**) [21]



Acetamide **4** (0.116 g, 0.470 mmol) was dissolved in distilled H<sub>2</sub>O (3 mL) and acidified with 1M HCl to pH 2. mCPBA (0.080 g, 0.470 mmol) was added to the flask and the reaction mixture was stirred for 1.5 h at r.t. The completion of the reaction was monitored by LC-MS. The aqueous phase was washed with DCM (3 × 3 mL). The crude product was obtained by removing water *in vacuo*. The purification of the desired compound was made by semi-preparative HPLC. Product **14** was obtained as a yellowish viscous liquid (0.021 g, 17%).

<sup>1</sup>H NMR (400 MHz, Methanol-*d*<sub>4</sub>) δ 7.77 (dd, *J* = 39.1, 8.1 Hz, 1H), 7.49 – 7.38 (m, 1H), 3.79 (dd, *J* = 110.2, 13.4 Hz, 2H), 3.53 (t, *J* = 6.5 Hz, 2H), 3.21 (q, *J* = 7.3 Hz, 2H), 3.11 – 3.02 (m, 2H), 2.94 (s, 6H), 2.08 – 2.00 (m, 2H), 1.31 (t, *J* = 7.3 Hz, 2H).

$^{13}\text{C}$  NMR (101 MHz, MeOD)  $\delta$  164.84, 57.55, 49.82, 48.64, 44.03, 36.87, 22.10, 15.88, 15.69.  
MS (ESI<sup>+</sup>): m/z (%): 262.24 (100) [M+H]<sup>+</sup>.  
HR-MS(ESI<sup>+</sup>): calculated for C<sub>9</sub>H<sub>20</sub>N<sub>5</sub>S 262.13377, found 262.13327

## Biochemistry and proteomics.

### Expression and Purification of Rab1b.

Competent *E. coli* NEB 10-beta cells were transformed with pBAD\_Rab1b<sub>3-174</sub>-Q67A-His6 plasmid. The transformation was directly inoculated into 50 mL non-AI medium (Hammill *et al.*, 2007) containing ampicillin (100 µg/mL) and incubated overnight at 37 °C, 200 rpm. The pre-culture was diluted to an OD<sub>600</sub> between 0.04 – 0.07 in 500 mL AI medium (Hammill *et al.*, 2007) supplemented with ampicillin (100 µg/mL). The cells were cultivated at 37 °C, 200 rpm for 12 hours at 37 °C, 200 rpm.

The obtained cell pellets were resuspended in His-wash buffer containing 50 mM HEPES pH 8.0, 500 mM LiCl, 20 mM imidazole, 2 mM β-mercaptoethanol, 1 mM MgCl<sub>2</sub>, 0.01 mM GTP supplemented with 0.1 mg/mL DNase I (AppliChem) and one cOmplete™ ULTRA EDTA-free protease inhibitor tablet (Roche). High pressure cell lysis was performed using the Emulsiflex C5 homogenizer (Avestin). The cell lysate was centrifuged (18,000 rpm, 1 h, 4 °C) and the cleared lysate was incubated together with 1 mL Ni-NTA slurry (Jena Bioscience) for 1 h at 4 °C slightly shaking. The beads were transferred to a gravity flow column and washed with His-wash buffer before eluting the proteins in 1 mL fractions using His-wash buffer supplemented with 500 mM imidazole pH 8.0. The fractions containing the protein were pooled, concentrated and rebuffed in storage buffer (20 mM HEPES pH 8.0, 150 mM NaCl, 5 mM MgCl<sub>2</sub>, 2 mM DTT and 0.01 mM GTP) with Amicon® Ultra-4 10K NMWL centrifugal filter units (Millipore). Purified proteins were analyzed by 15 % SDS PAGE and stored at -80 °C. Protein concentration was determined using NanoPhotometer® N60 (Implen GmbH).

### Expression and Purification of DrrA.

Competent *E. coli* NEB 10-beta cells were transformed with pBAD-RSF1031K-StrepII-TEV-DrrA<sub>16-352</sub> and directly inoculated into 50 mL non-AI medium (Hammill *et al.*, 2007) containing kanamycin (50 µg/mL). Next day the pre-culture was diluted to an OD<sub>600</sub> between 0.04 – 0.07 in 500 mL AI medium (Hammill *et al.*, 2007) supplemented with kanamycin (50 µg/mL) and incubated for 16 hours at 37 °C, 200 rpm.

The obtained cell pellet was thoroughly resuspended in 30 mL of Strep-wash buffer containing 100 mM Tris pH 8.0, 150 mM NaCl, 1 mM EDTA and supplemented with 0.1 mg/mL DNase I (AppliChem) and one cOmplete™ ULTRA EDTA-free protease inhibitor tablets (Roche). The cells were lysed using an Emulsiflex C5 high pressure homogenizer (Avestin). Cell debris was removed by centrifugation (18,000 rpm, 45 min, 4 °C) and the cleared lysate was applied to an ÄKTA Pure FPLC system (Cytiva, former GE Healthcare) equipped with a 1 mL StrepTrap HP column (Cytiva). The column was washed with Strep-wash buffer and eluted in 1 mL fractions with Strep-wash buffer supplemented with 50 mM biotin. The fractions containing the protein were pooled together, concentrated and rebuffed (20 mM HEPES pH 8.0, 150 mM NaCl, 2 mM DTT, and 5 mM MgCl<sub>2</sub>) using Amicon® Ultra-4 10K NMWL centrifugal filter units (Millipore). Purified proteins were analyzed by 15 % SDS PAGE and stored at -80 °C. Protein concentration was determined using NanoPhotometer® N60 (Implen GmbH).

### *In vitro* AMPylation of Rab1b.

Purified Rab1b<sub>3-174</sub> was mixed with DrrA<sub>16-352</sub> in a 50:1 Rab1b:DrrA ratio in AMPylation buffer (20 mM HEPES pH 8.0, 150 mM NaCl, 1 mM DTT, 1 mM MgCl<sub>2</sub>, 0.01 mM GTP) supplemented with an 2.5-fold excess of ATP compared to the Rab1b. Samples were incubated at 25 °C for 3 hours at 200 rpm and AMPylated Rab1b was purified via size exclusion chromatography using a superdex 75 10/300 GL column and SEC buffer (20 mM HEPES pH 8.0, 150 mM NaCl, 1 mM DTT, 1 mM MgCl<sub>2</sub>, 0.01 mM GTP).

### Cysteine labelling on BSA and click chemistry.

To a 1.6 mg/mL BSA solution in 20 mM Hepes pH=7.5 or 50 mM ABC buffer was added IAA-alkyne to a final concentration of 10 mM. The mixture was incubated for 1h at 60°C. Excess IAA-alkyne was removed by buffer exchange on a 50k cutoff Amicon® Ultra Centrifugal Filter. The solution was transferred to a new reaction tube and the click reaction was performed with 2 mM DMP-N<sub>3</sub> or SOX-N<sub>3</sub>, 4 mM TCEP, 4 mM CuSO<sub>4</sub> and 0.4 mM TBTA.

### **Peptide preparation from purified proteins.**

If proteins contained disulphide bridges (e.g. BSA) those were reduced and the free cysteines alkylated by 5 mM TCEP and 20 mM CAA by incubation for 5 minutes at 95°C, 550 rpm. Trypsin digestion was performed in 20 mM Hepes pH=7.5 buffer with 1:200 trypsin to protein ratio by incubation for 16 h at 37°C, 550 rpm. Chymotrypsin digestion was carried out with 1:100 chymotrypsin to protein ratio in 50 mM ABC buffer containing 2mM CaCl<sub>2</sub> by incubation for 16 h at 25°C, 550 rpm. The samples were acidified by the addition of FA to 0.5%. The peptides were desalted on SepPak C18 50 mg cartridges (Waters) using vacuum manifold with final elution into 300 µL 0.5% FA, 80% ACN in water.

### **HeLa proteome preparation.**

HeLa cells were cultivated in high glucose Dulbeccos's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 2 mM L-glutamine. Cells were grown at 37 °C and 5% CO<sub>2</sub>. Cells were seeded into 15 cm diameter dishes and grown to 90-100% confluency. The medium was removed, the cells were washed with PBS and the scrapped into 1 mL PBS and centrifuged. The pellets were lysed in 600 µL 20 mM Hepes pH=7.5, 1% NP40, 0.2% SDS by sonication and the lysates were centrifuged. 2 µL samples were diluted 100x and used for determining protein concentration by BCA assay (carried out as recommended by the supplier). To the lysates 4x volume of acetone was added, and the proteins were let precipitate for 2 h or overnight at -20 °C. Precipitated proteins were pelletized by centrifuging for at 4 °C at 4000g. The pellet was washed with cold methanol and redissolved in X-buffer (3.9 M urea, 1.1 M thiourea in 20 mM Hepes pH=7.5). Proteins were reduced by incubating with 1 mM DTT at 37 °C for 45 min, then alkylated by IAA 5.5 mM IAA at 25 °C for 30 min in the dark. The excess IAA was quenched with 4 M DTT at 25 °C for 30 min. To this solution 3x volume of 50 mM TEAB buffer (pH=8) was added, and the proteins were digested by trypsin. The samples were acidified by the addition of FA to 0.5%. The peptides were desalted on SepPak C18 50 mg cartridges (Waters) using vacuum manifold with final elution into 750 µL 0.5% FA, 80% ACN in water and dried in SpeedVac (Thermo Scientific).

**Cysteine labelling, click chemistry and DMP-enrichment.** 1 mg HeLa cell lysate per sample was diluted to 2 mg/mL protein concentration with 0.2% SDS in PBS. IAA-alkyne (or DMSO for controls) was added to the samples to 1 mM final concentration and the mixture was incubated for 1 h at 25°C 550 rpm. Thereafter, DMP-N<sub>3</sub> / SOX-N<sub>3</sub> was added to 2 mM final concentration together with 2 mM TCEP, 2 mM CuSO<sub>4</sub> and 0.2 mM TBTA. The mixture was incubated for 1.5 h at 25°C at 550 rpm. The proteins were precipitated by addition of 4x volume of acetone and incubation for 2 h at -20°C. Peptides were prepared from the labelled proteome as described in HeLa proteome preparation above.

### **Spike-in sample preparation.**

Peptides prepared from purified proteins were dried in SpeedVac (Thermo Scientific) and reconstituted in 1% FA in water. 25 ng/µl digested HeLa proteome standard was prepared in 1% FA in water. To this solution the prepared peptide solution was mixed to yield a 20 ng/µl theoretical concentration of the spiked-in peptides. From these samples 8 µl was injected in every LC/MS/MS measurement.

### **Intact protein measurements.**

The 10 mM (10 pmol/µl) desalted protein solutions were injected into the Orbitrap Eclipse Tribrid mass spectrometer by the microinjection source. For injection 10 µl/min sample flow was applied, for data acquisition the flow was lowered to 2-4 µl/min. The measurements were carried out in Standard Pressure Mode (8 mtorr) with ion transfer tube temperature set to 320 °C. The full MS scans were acquired on the orbitrap with 120k FWHM (at 200 *m/z*) resolution applying 5 microscans. The AGC target was set to 2e5 allowing maximal ion injection time of 100 ms. FAIMS CV was scanned from -70 to 30 V in 2 V steps to find the optimal setting for each protein. The CV where the most intense protein peaks could be detected were selected for measurement.

### **Peptide measurements with direct injection.**

The desalted 10 mM peptide solutions (calculated from the initial protein concentration assuming 100% peptide recovery) were injected into the Orbitrap Eclipse Tribrid mass spectrometer by the microinjection

source. The measurements were carried out in Standard Pressure Mode (8 mtorr) with ion transfer tube temperature set to 320 °C. The full MS scans were acquired on the orbitrap with 60k FWHM (at 200  $m/z$ ) resolution applying 5 microscans. The AGC target was set to  $2e5$  allowing maximal ion injection time of 100 ms. FAIMS CV was scanned from -80 to -30 V in 2 V steps to find the optimal setting for the peptide of interest. The CV where the peak of the desired peptide was the most intense was selected for measurement. For MS2 fragmentation the peptide ion of interest was selected on the quadrupole (with a 1.5  $m/z$  window) in all available charge states. Fragmentation parameters were varied systematically, CID energy between 5-50%, HCD energy between 10-40%, ETD reaction time between 2-150 ms.

#### **Standard LC/MS/MS method.**

The default LC/MS/MS analysis was performed on the Orbitrap Eclipse Tribrid mass spectrometer (Thermo Fisher Scientific) coupled to an UltiMate3000 Nano-HPLC (Thermo Fisher Scientific) via a NanoFlex source (Thermo Fisher Scientific) and FAIMS interface (Thermo Fisher Scientific). First, peptides were loaded on an Acclaim PepMap 100  $\mu$ -precolumn cartridge (5  $\mu$ m, 100 Å, 300  $\mu$ m ID  $\times$  5 mm, Thermo Fisher Scientific). Then, peptides were separated at 40 °C on a PicoTip emitter (noncoated, 15 cm, 75  $\mu$ m ID, 8  $\mu$ m tip, New Objective) that was in house packed with Reprosil-Pur 120 C18-AQ material (1.9  $\mu$ m, 150 Å, Dr. Maisch GmbH). The gradient was run from 4-35.2% acetonitrile supplemented with 0.1% formic acid during a 150 min method (0-5 min at 4%, 5-6 min to 7%, 7-105 min to 24.8%, 105-126 min to 35.2%, 126-140 at 80%, 140-150 min at 4%) at a flow rate of 300 nL/min. Unless otherwise stated, this LC method was used in all LC/MS/MS analyses. FAIMS was performed with two alternating CVs including -50 V and -70 V with cycle times of 1.7 and 1.3 s respectively. The Orbitrap Eclipse Tribrid Mass Spectrometer was operated in data dependent MS2 mode with following settings: Polarity: positive; MS1 resolution: 240k; MS1 AGC target: standard; MS1 maximum injection time: 50 ms, MS1 scan range:  $m/z$  375-1500; MS2 ion trap scan rate: rapid; MS2 AGC target: standard; MS2 maximum injection time: 35 ms; MS2 isolation window:  $m/z$  1.2; HCD stepped normalized collision energy: 30%; intensity threshold:  $1.0e4$  counts; included charge states: 2-6; dynamic exclusion: 60 s; MS2 scan range:  $m/z$  120-1200. The high-resolution MS2 was acquired in orbitrap with 15k resolution.

#### **Data analysis.**

To analyse the acquired MS data with MaxQuant the .raw files were split by FAIMS-MzXML-generator (GitHub) to separate the information acquired using different FAIMS CVs. For most analyses the MaxQuant version 1.6.15.0 was used. The spectra were searched against the *in silico* digested Uniprot database for Homo sapiens (taxon identifier: 9606). False discovery rate (FDR) was determined by using a decoy database and set to 1% as thresholds for both peptide-spectrum match and protein levels. In these analyses the focus was on identifications, so no quantification was performed. Carbamidomethylation (+57.0215) was set up as fixed modification except for cysteine labelling experiments, where it was set up as variable modification. Protein N-terminal acetylation (+42.0106) and methionine oxidation (+15.9949) were always included as variable modifications. Other modifications were included in the search depending on the sample. The reverse and potential contaminant identifications as well as identifications from control samples were filtered out. For all the analyses MSFragger version 3.2 was used together with Philosopher 3.4.13 integrated into the graphical user interface FragPipe v15.0. The modification search was performed both in the conventional way (modification of the *in silico* digested peptides and matching to the altered fragmentation spectrum) and with mass offset search (allowing for certain precursor mass shifts for peptide spectrum matching). To achieve this, the mass of the modifications were supplied both as a variable modification on certain amino acids and as mass offsets restricted to the same amino acids. For the mass offset search, the labile modification mode was used, where the diagnostic fragment masses were defined. This search mode was only activated if the diagnostic ion intensity exceeded the relative intensity threshold of 10%. In AMPylation searches Y-ion mass of 79.9663 was also provided corresponding to the neutral loss off the adenosine fragment from the intact peptide. Validation of the identifications was performed by the TransProteomic Pipeline tools PeptideProphet and ProteinProphet.

**DMP-enrichment.**

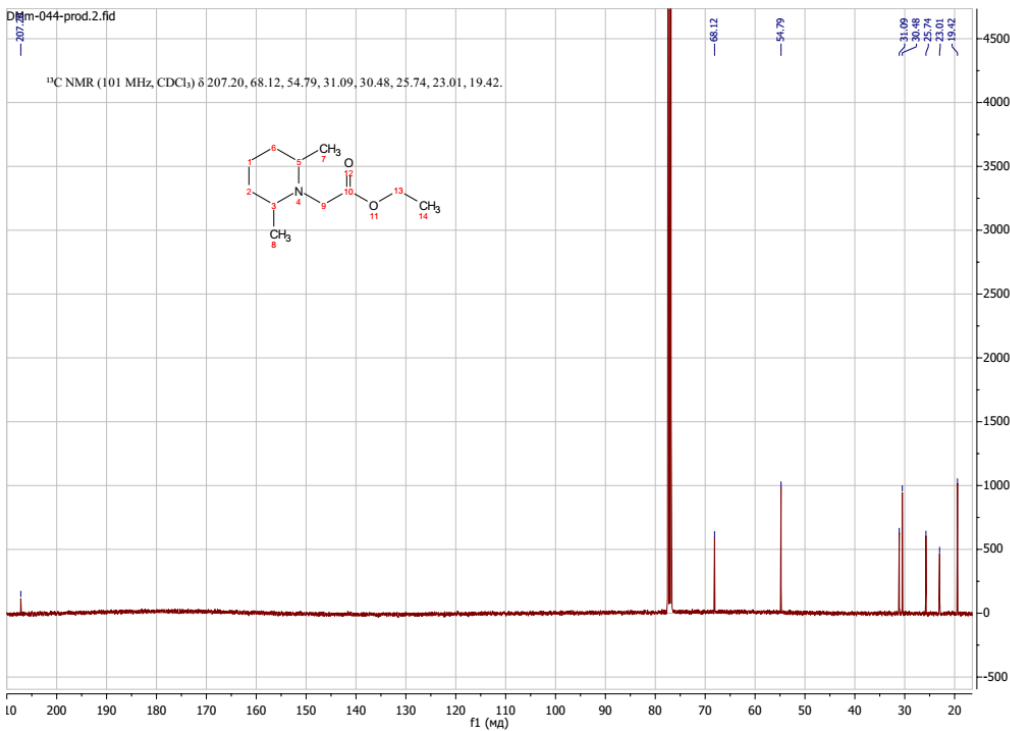
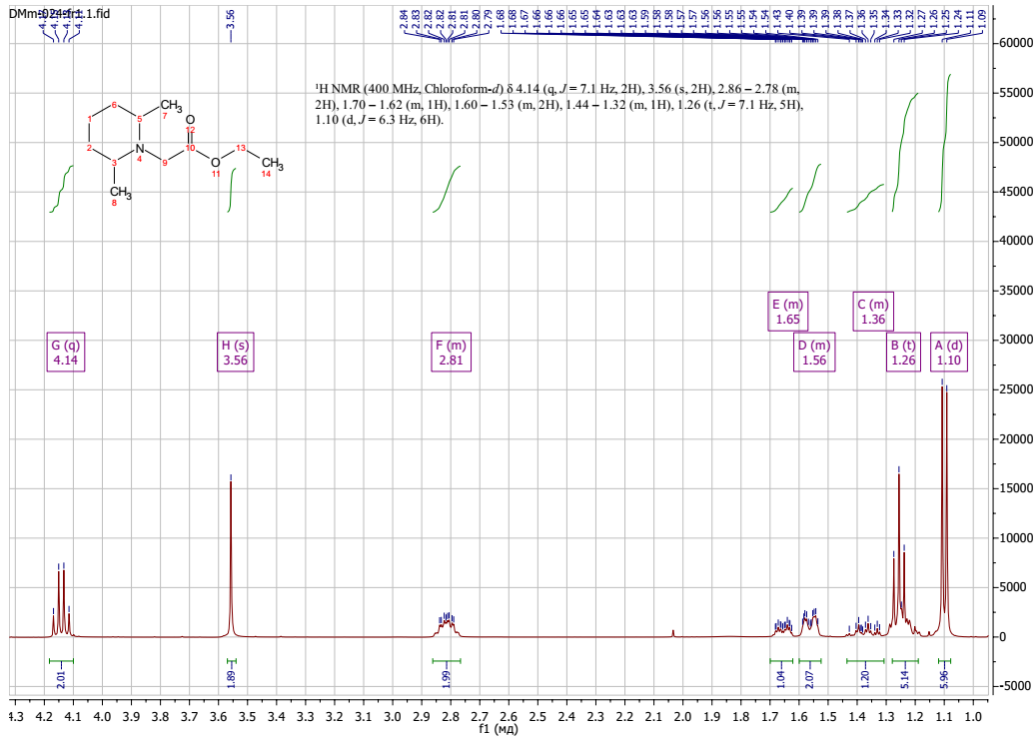
1 mg HeLa cell lysate per sample was diluted to 2 mg/mL protein concentration with 0.2% SDS in PBS. IAA-alkyne was added to the samples to 1 mM final concentration and the mixture was incubated for 1 h at 25°C. Thereafter, DMP-N3 was added to 2 mM final concentration together with 2 mM TCEP, 2 mM CuSO<sub>4</sub> and 0.2 mM TBTA. The mixture was incubated for 1.5 h at 25°C. The proteins were precipitated by the addition of 4× volume of acetone and incubation for 2 h or overnight at -20°C. Peptides were prepared from the labelled proteome as described in HeLa proteome preparation above. Dried peptides were redissolved in IP buffer (50 mM Tris HCl pH=8, 250 mM NaCl) to a final theoretical peptide concentration of 2 mg/mL. 2 µL of these solutions were taken for full proteome measurements. The rest was loaded on 100 µL equilibrated anti-TMT resin (Thermo Fisher Scientific) and incubated for 16 h at 4°C. The supernatant was removed, and the beads were washed with 2×500 µL 2M urea in IP buffer, 2×500 µL 0.05% SDS in IP buffer, 2×500 µL IP buffer and 2×500 µL water. The captured peptides were eluted by 3×500 µL TMT elution buffer (Thermo Fisher Scientific). The eluate was desalted on SepPak C18 50 mg cartridges (Waters) as described above, and dried in SpeedVac (Thermo Scientific). Peptides were reconstituted in 25 µL 1% FA in water for MS measurement.

**Data availability.**

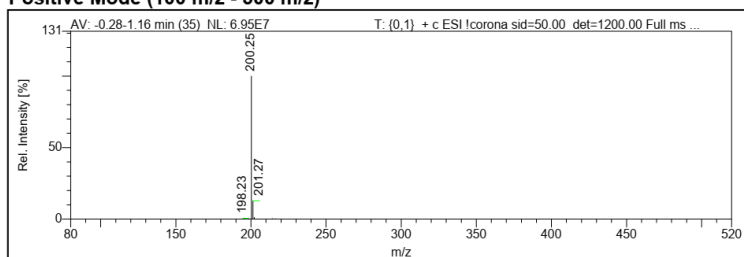
Mass spectrometry-based proteomics data have been deposited at ProteomeXchange and are publicly available as of the date of publication. The accession number is PXD030608.

# NMR and MS Spectra of Small Compounds.

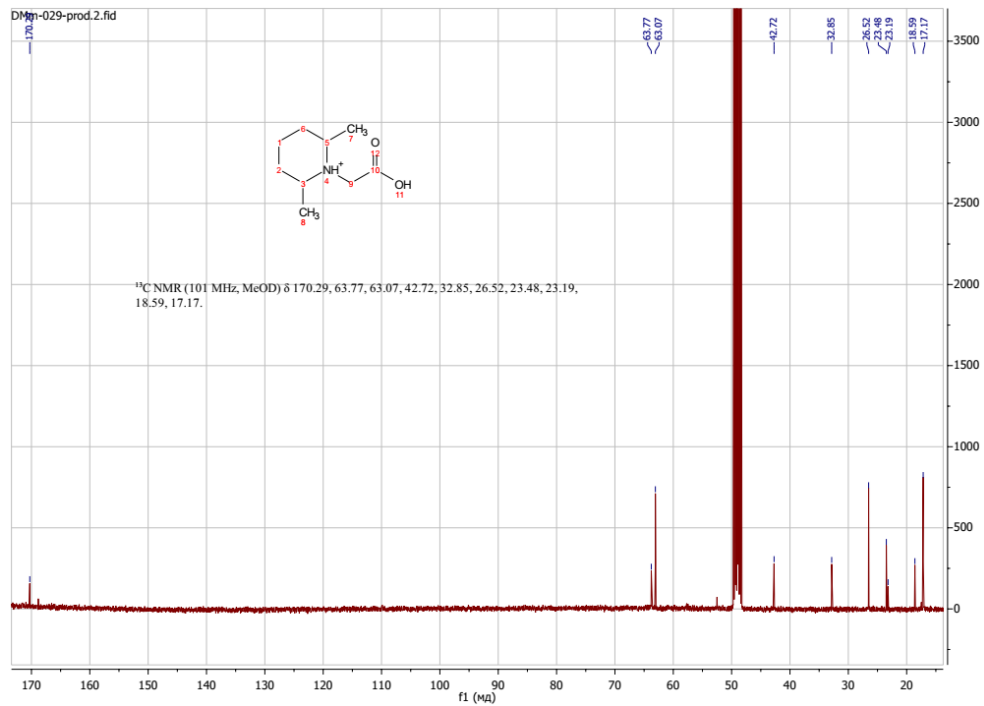
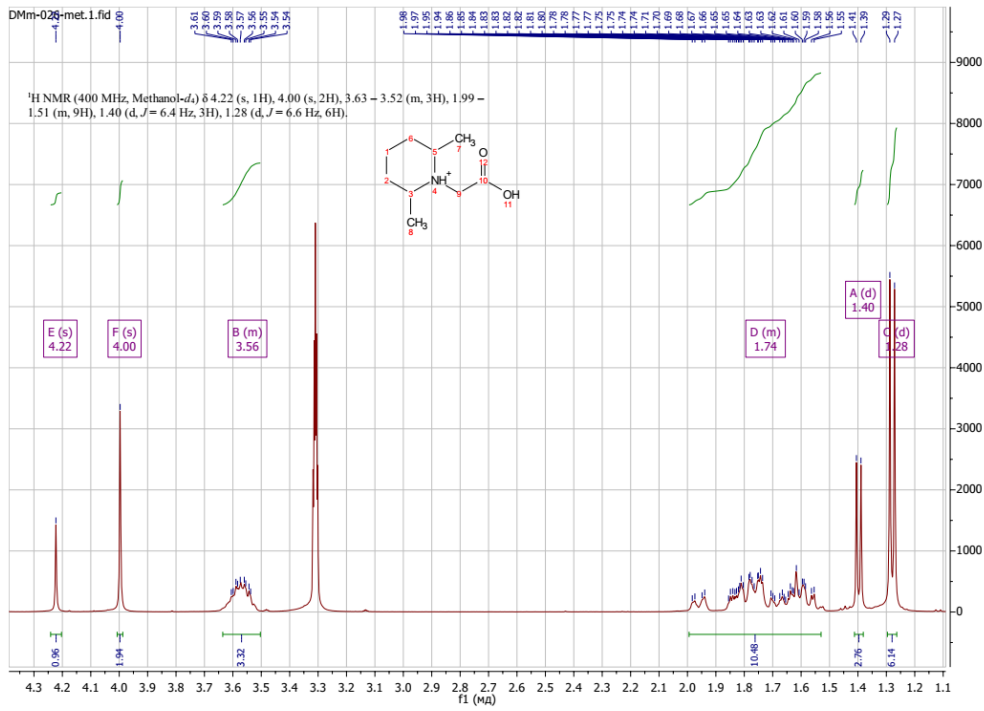
## Ethyl 2-(2,6-dimethylpiperidin-1-yl)acetate (7)



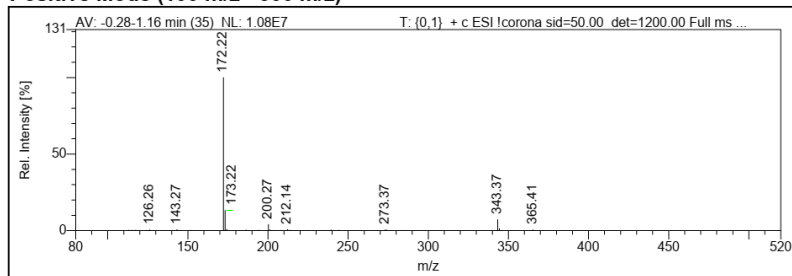
### Positive Mode (100 m/z - 500 m/z)



## 2-(2,6-Dimethylpiperidin-1-yl)acetic acid (1)

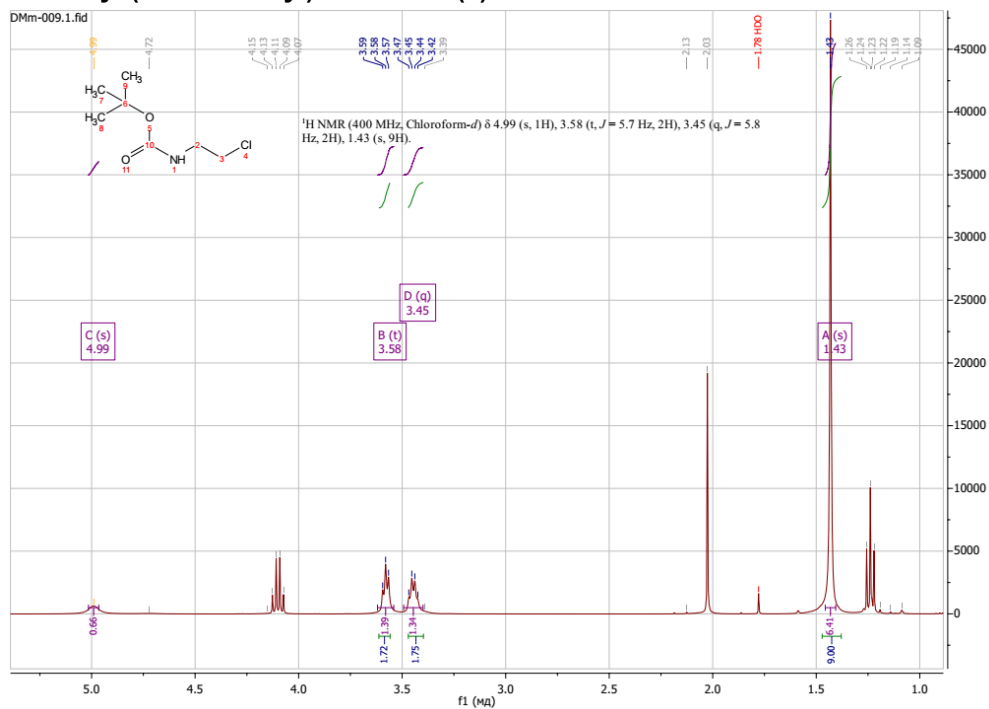


### Positive Mode (100 m/z - 500 m/z)

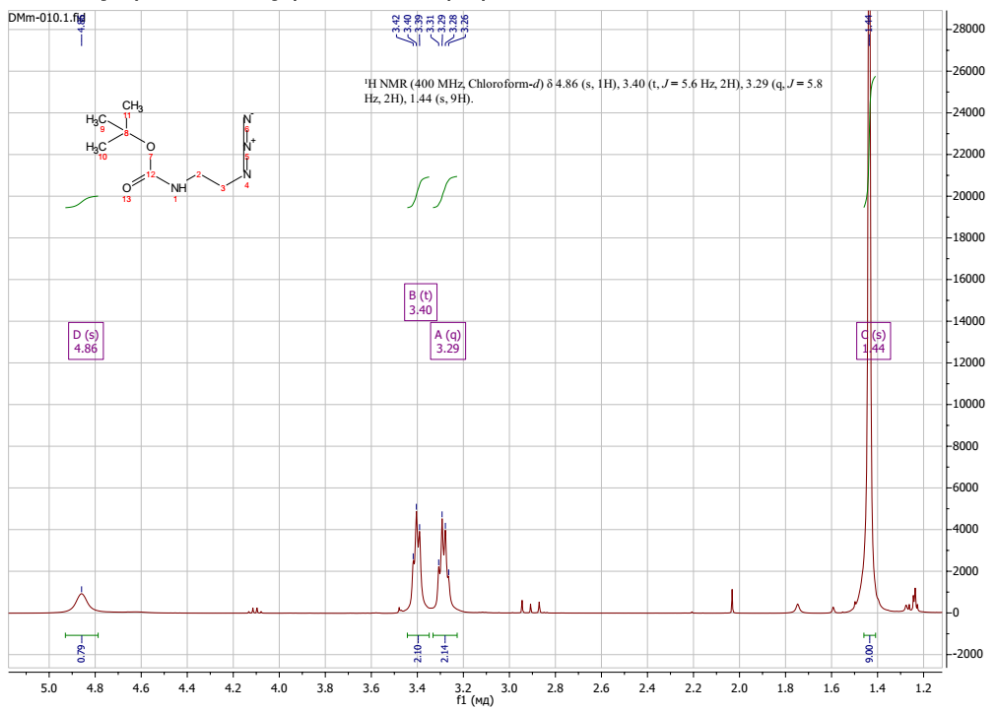




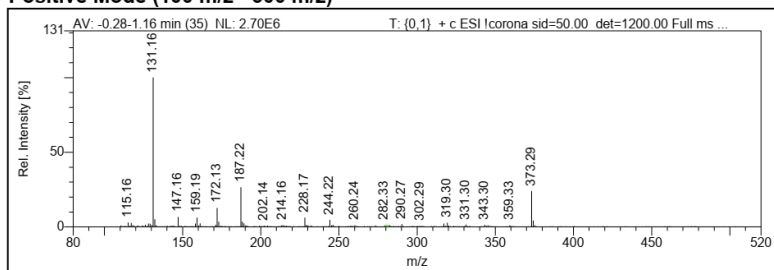
**tert-Butyl (2-chloroethyl)carbamate (8)**



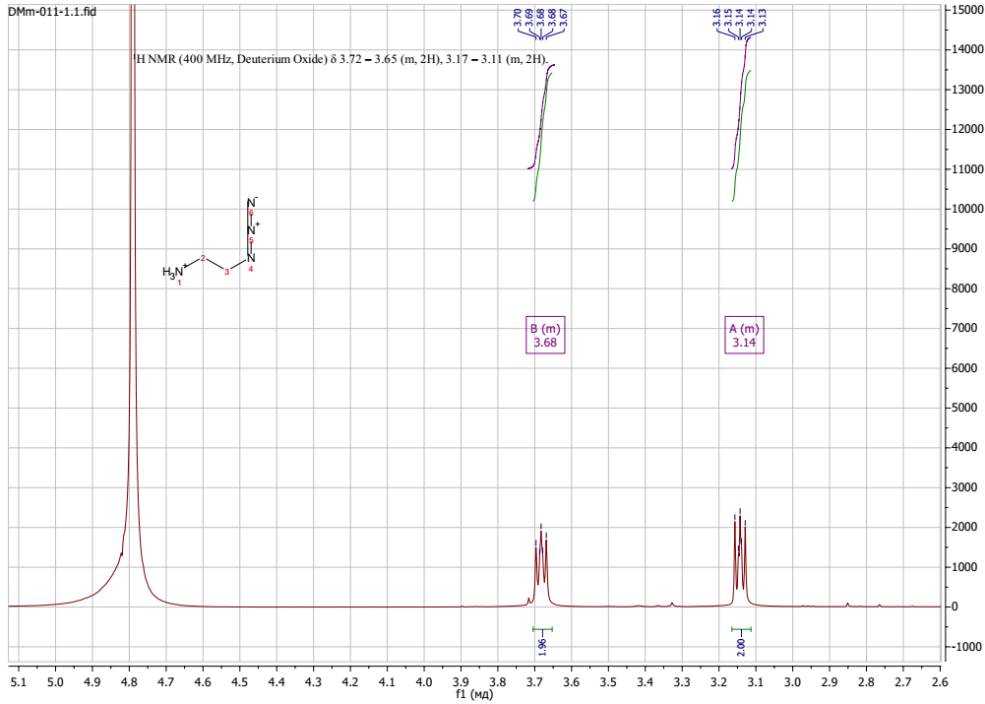
**tert-Butyl (2-azidoethyl)carbamate (10)**



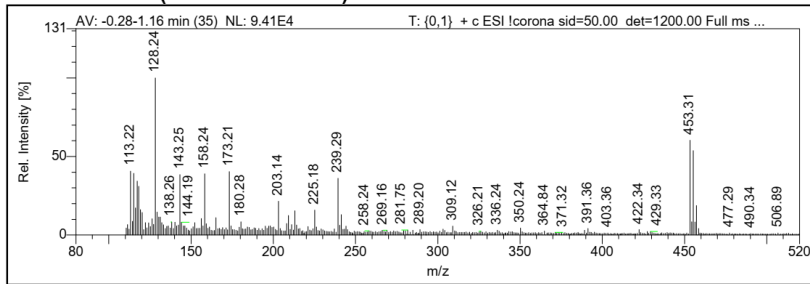
**Positive Mode (100 m/z - 500 m/z)**



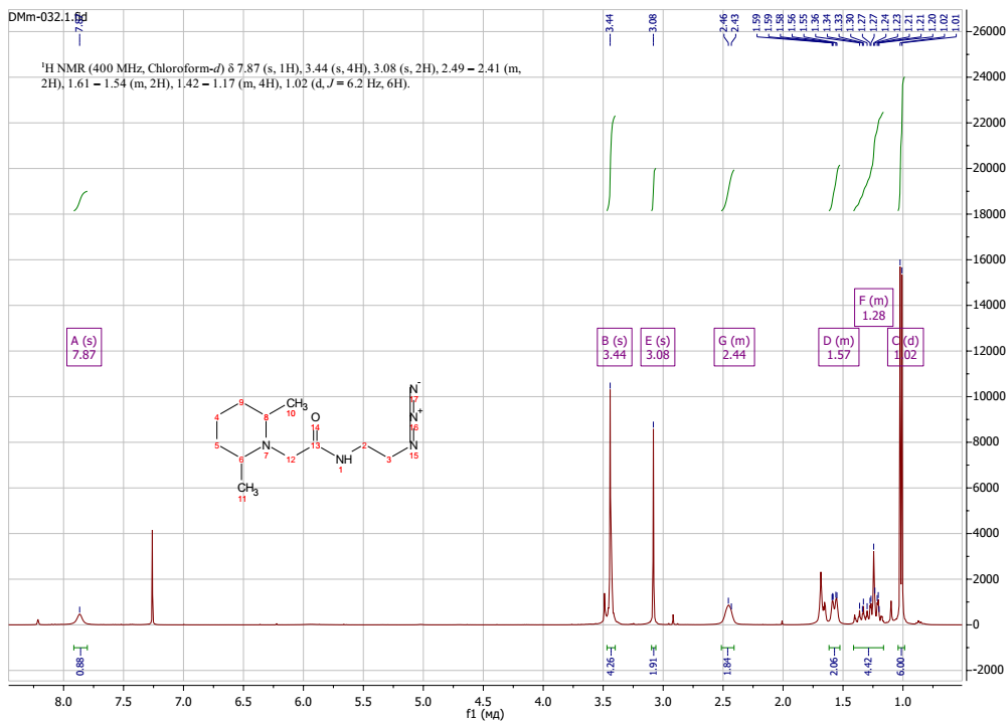
## 2-Azidoethan-1-amine hydrochloride (11)

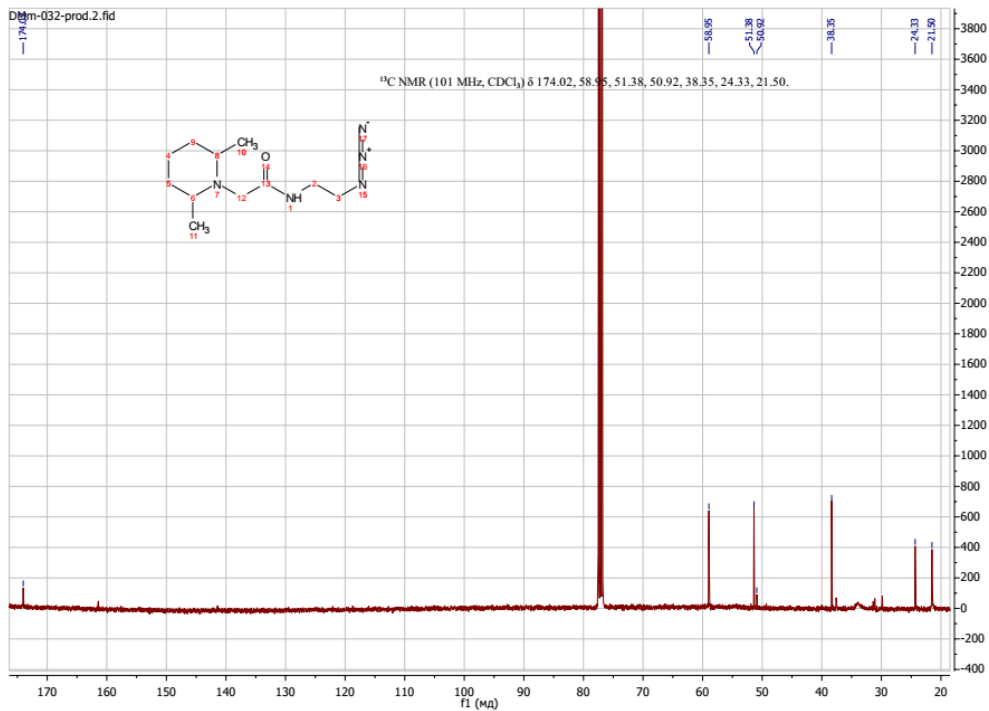


### Positive Mode (100 m/z - 500 m/z)

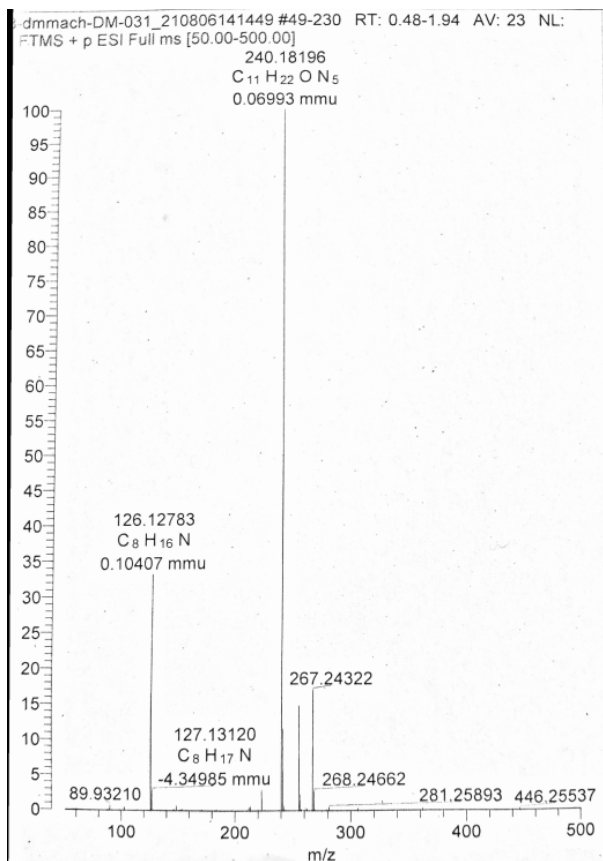
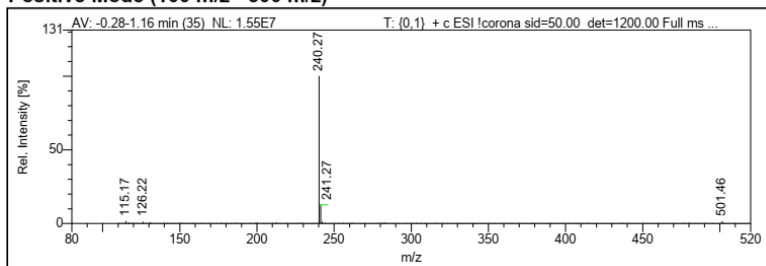


## N-(2-azidoethyl)-2-(2,6-dimethylpiperidin-1-yl)acetamide (12)

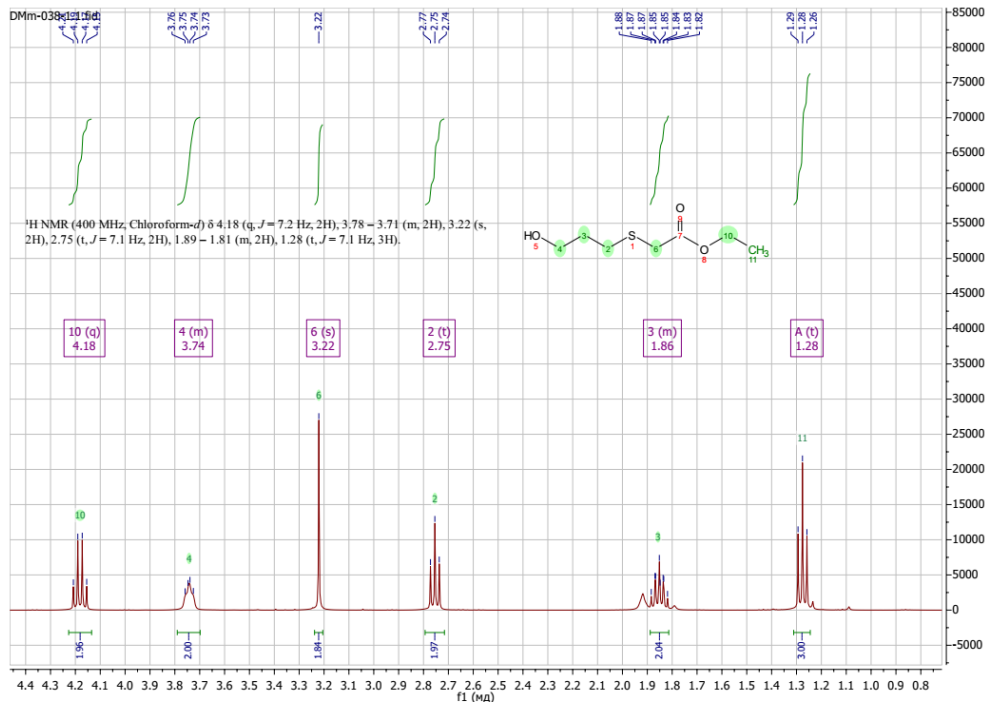


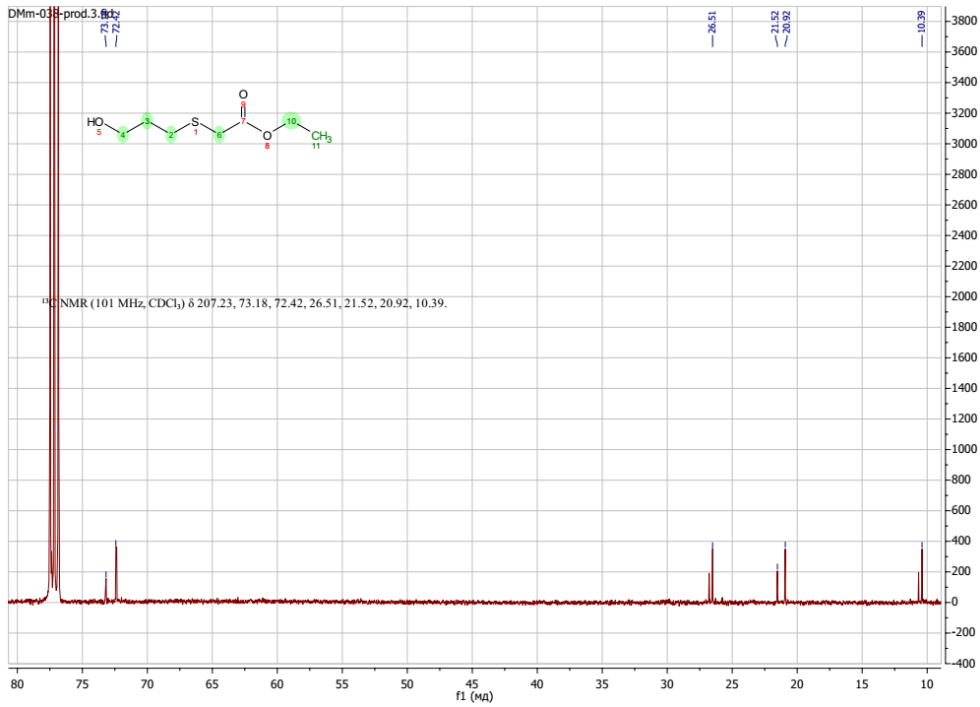


**Positive Mode (100 m/z - 500 m/z)**

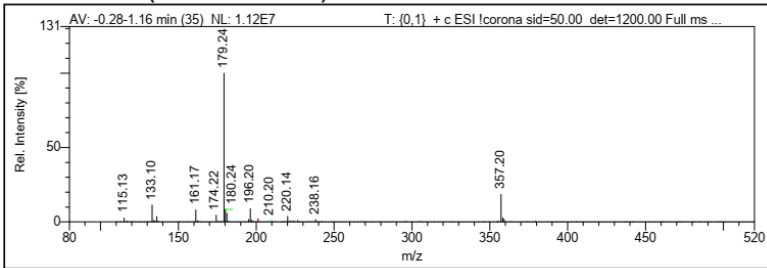


# Ethyl 2-((3-hydroxypropyl)thio)acetate (12)

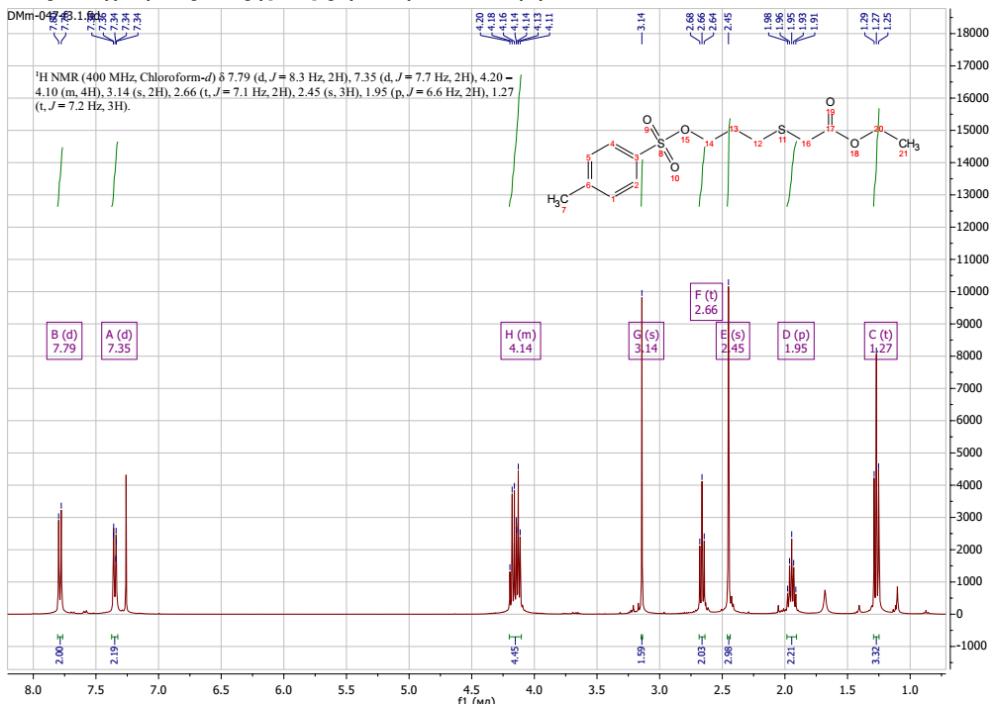


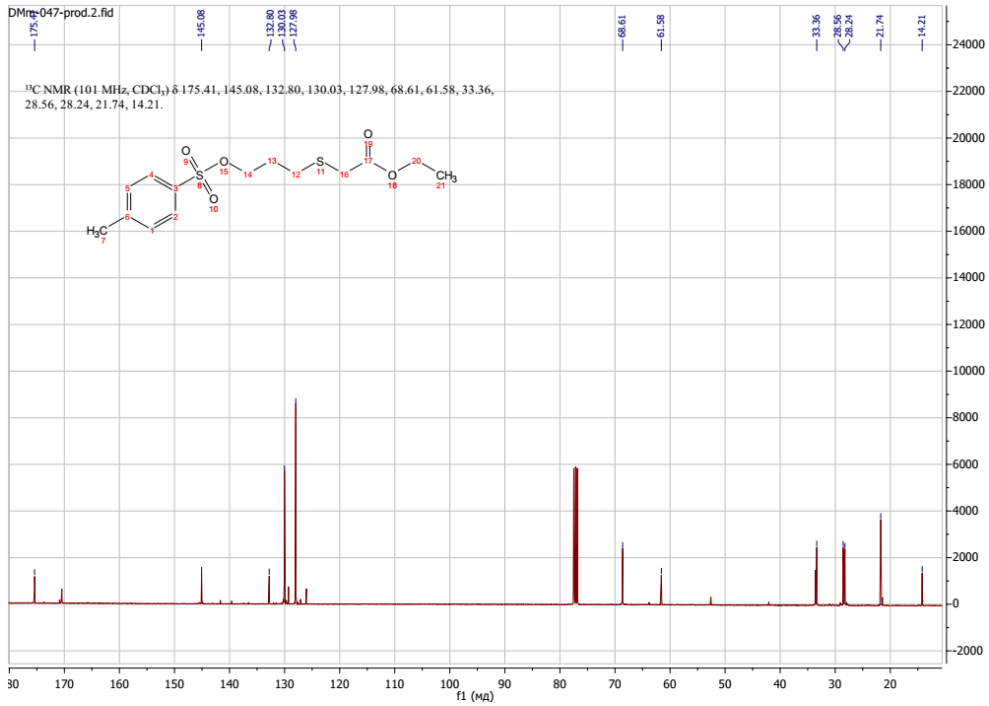


**Positive Mode (100 m/z - 500 m/z)**

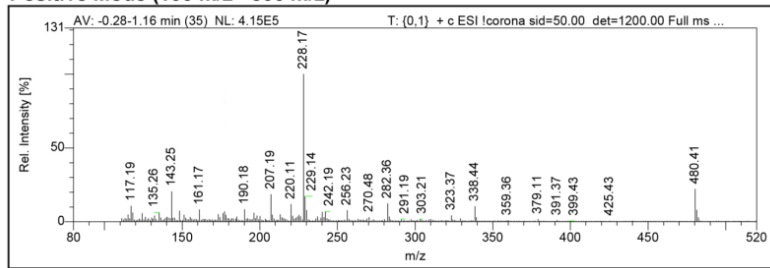


**Ethyl 2-(3-(tosyloxy)propyl)thioacetate (2)**

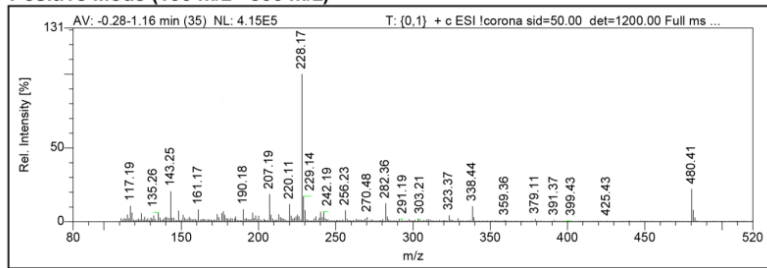




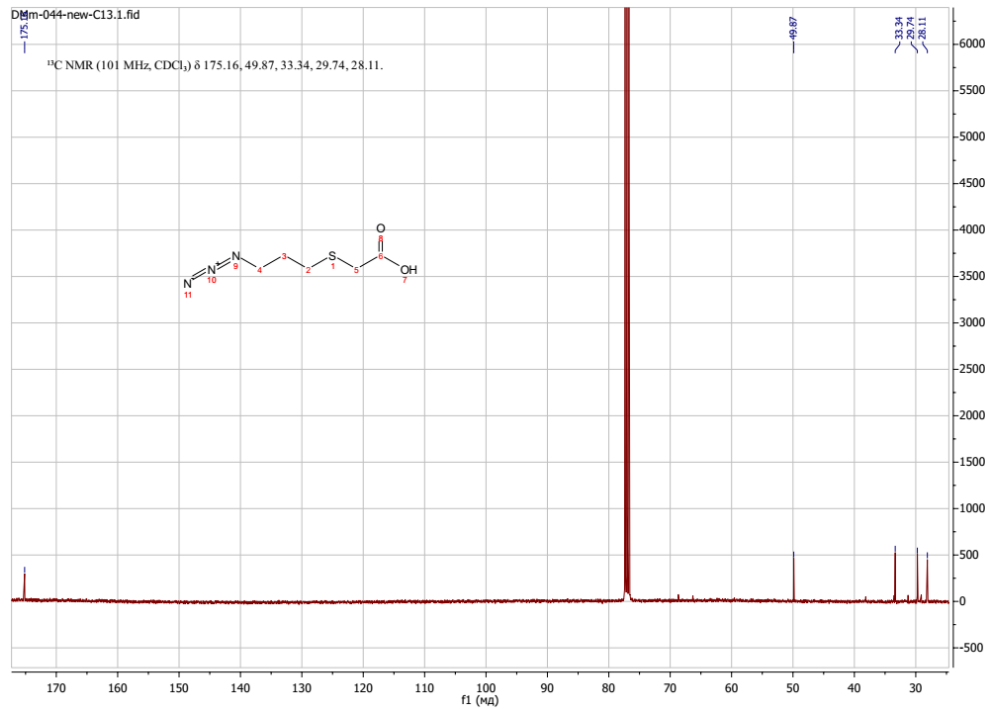
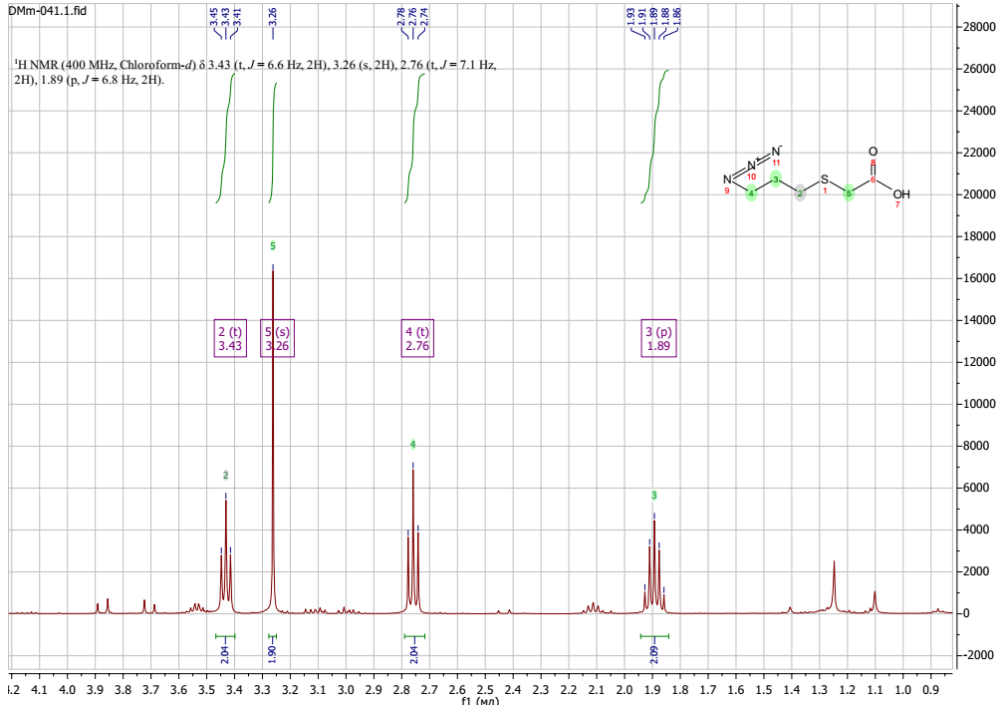
**Positive Mode (100 m/z - 500 m/z)**



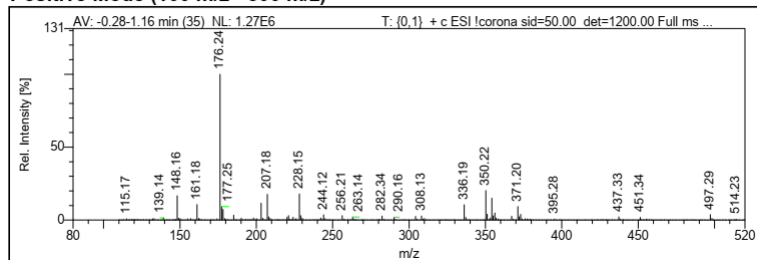
**Positive Mode (100 m/z - 500 m/z)**



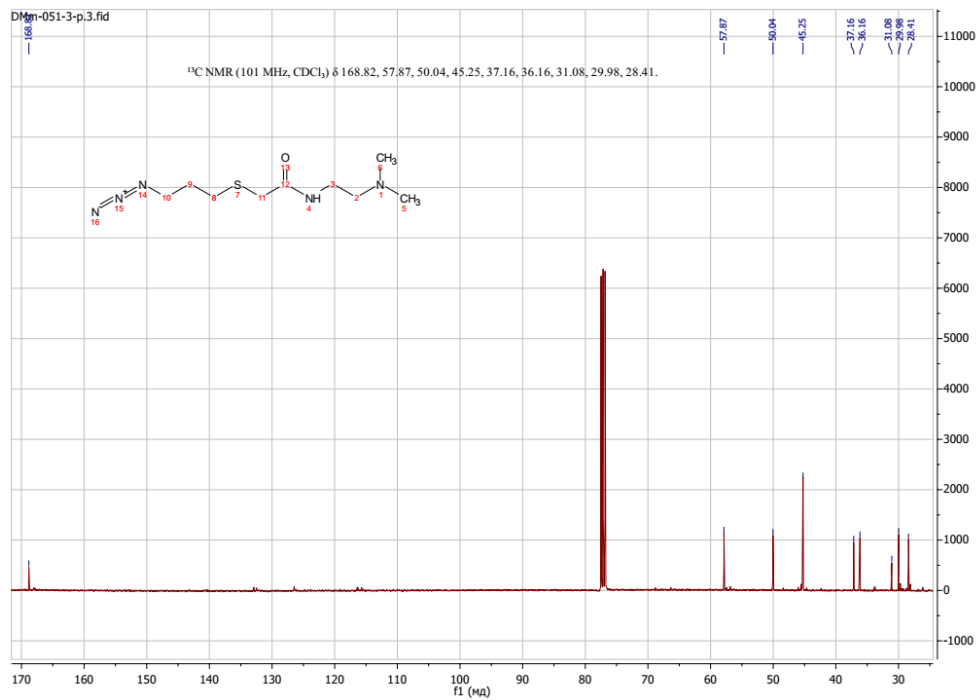
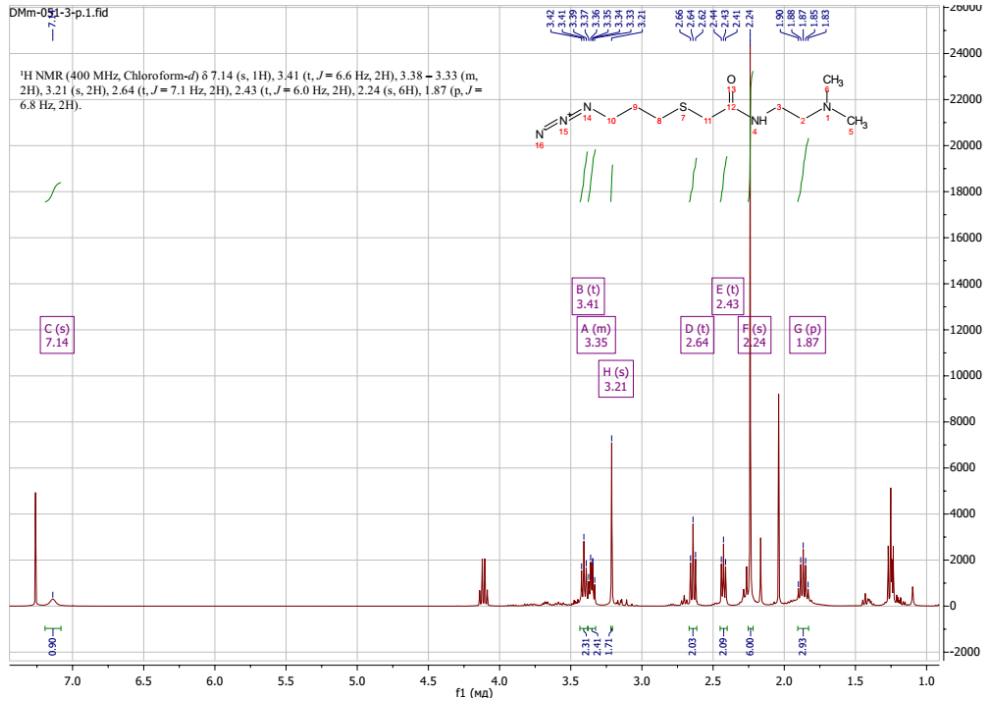
## 2-((3-azidopropyl)thio)acetic acid (3)



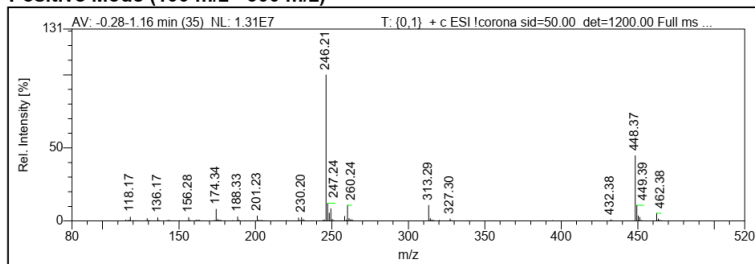
### Positive Mode (100 m/z - 500 m/z)



## 2-((3-Azidopropyl)thio)-N-(2-(dimethylamino)ethyl)acetamide (4)

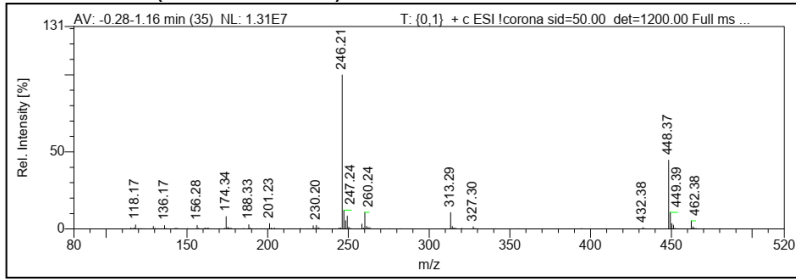


### Positive Mode (100 m/z - 500 m/z)

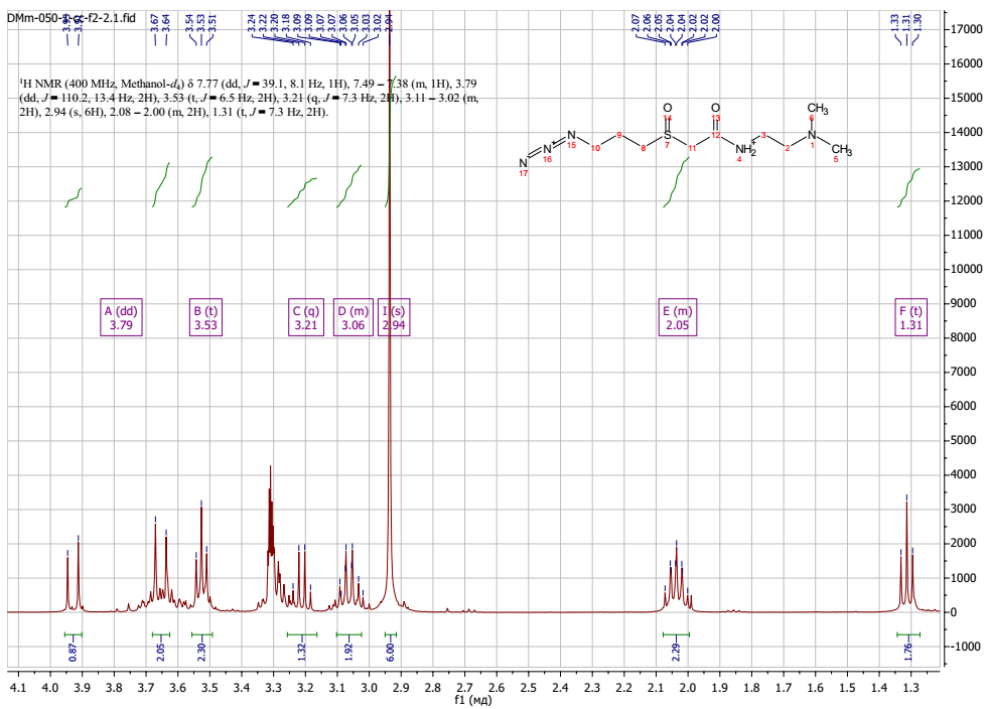
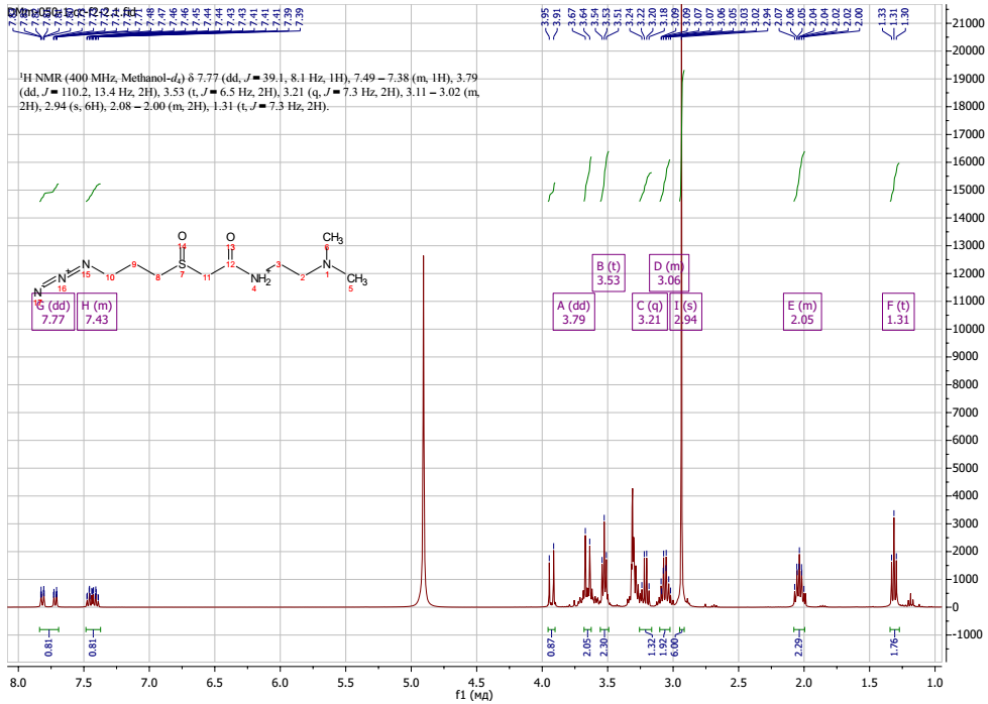




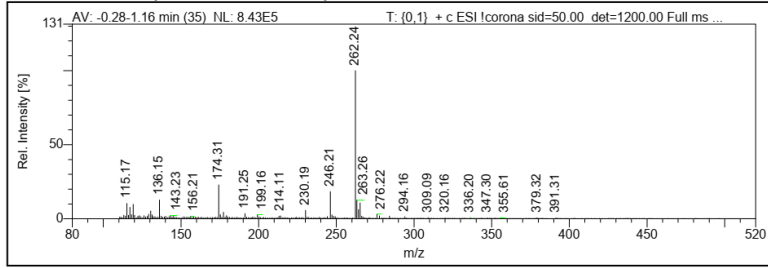
Positive Mode (100 m/z - 500 m/z)



2-((3-Azidopropyl)sulfinyl)-N-(2-(dimethylamino)ethyl)acetamide (14)



**Positive Mode (100 m/z - 500 m/z)**



09-dmmach-DM-067 #49-227 RT: 0.56-1.97 AV: 22 NL: 1.92E6  
 F: FTMS + p ESI Full ms [50.00-500.00]

